

Alexander Wilk

Elektronen-Pionen-Separation
im ALICE TRD

— 2004 —

für meine Großeltern

Experimentelle Physik

Elektronen-Pionen-Separation
im ALICE TRD

Diplomarbeit

von

Alexander Wilk

Westfälische Wilhelms-Universität Münster

Institut für Kernphysik

— 2004 —

Inhaltsverzeichnis

1	Einleitung	3
2	Theoretisches Rüstzeug	5
2.1	Das Standardmodell der Teilchenphysik	5
2.1.1	Die Bausteine der Materie	5
2.1.2	Die fundamentalen Wechselwirkungen	7
2.2	Die starke Wechselwirkung	7
2.3	Das Quark-Gluon-Plasma	9
3	Experimenteller Rahmen	11
3.1	Der LHC	11
3.2	ALICE	13
3.3	Schwerionenreaktionen am LHC	15
4	Der Transition Radiation Detector (TRD)	17
4.1	Physikalische Motivation	17
4.2	Theorie für den TRD	18
4.2.1	Die Bethe-Bloch-Formel	19
4.2.2	Übergangsstrahlung	20
4.3	Aufbau eines TRD-Moduls	21
4.4	Funktionsweise des TRD	23
4.5	Die Teststrahlzeit 2002 am CERN	24
4.5.1	Der Versuchsaufbau	25
4.5.2	Das Triggersystem	25
5	Klassische Methoden zur e/π -Separation	27
5.1	N-Methoden und Q-Methoden	27

5.1.1	N-Methoden	27
5.1.2	Q-Methoden	28
5.2	Auswertung der Datensätze	32
5.2.1	Datenauslese	32
5.2.2	Simulation der einzelnen Kammern	33
5.2.3	Die Extrapolationsmethode	34
6	Künstliche neuronale Netzwerke	37
6.1	Einführung	37
6.1.1	Allgemeine Merkmale	38
6.1.2	Historisches	39
6.2	Vom natürlichen zum künstlichen Neuron	39
6.3	Feedforward Netzwerke	40
6.4	Der Backpropagation-Algorithmus	42
6.4.1	Herleitung	42
6.4.2	Arbeitsweise	43
6.4.3	Probleme mit dem Backpropagation-Algorithmus	44
6.5	Training und Validierung	45
7	e/π Separation mit neuronalen Netzwerken	49
7.1	Wahl des Netzes und der Parameter	49
7.2	Datenauswertung und Vergleich mit der LQX-Methode	51
8	Zusammenfassung und Ausblick	55
A	Variablen der Schwerionenphysik	57
	Literaturverzeichnis	59
	Danksagung	63

1. Einleitung

Materie, wie wir sie in unserem Alltag kennen, kann in verschiedenen Aggregatzuständen vorkommen. Jedem ist z.B. Wasser in seiner festen, flüssigen und gasförmigen Form bekannt. Inzwischen sind weitere Materiezustände entdeckt worden, etwa das Plasma oder das so genannte *Bose-Einstein-Kondensat*. Ein Ziel der Schwerionenphysik ist es Zustände der Kernmaterie zu untersuchen. Von der *Quantenchromodynamik* (QCD) wird unter bestimmten Bedingungen ein exotischer Zustand der Kernmaterie, das so genannte *Quark-Gluon-Plasma* (QGP), vorhergesagt. Zur Bildung eines QGP soll es bei extrem hohen Temperaturen (wie kurz nach dem Urknall) oder bei sehr hohen Dichten (z.B. in Neutronensternen) kommen.

Im Labor werden derartige Bedingungen durch Kollisionen schwerer Atomkerne (z.B. Gold oder Blei) in großen Teilchenbeschleunigern erzeugt. Dabei erreichen die Teilchen nahezu Lichtgeschwindigkeit. Aktuelle Experimente am RHIC (**R**elativistic **H**eavy **I**on **C**ollider) liefern einige Hinweise, dass es gelungen ist ein QGP künstlich herzustellen. Zur Bestätigung der Ergebnisse und weiteren Erforschung des QGP ist es nötig, weitere Experimente durchzuführen. Dies wird an modernen Beschleunigern geschehen, die in Planung oder bereits im Bau sind.

Ein Beschleuniger der nächsten Generation, der LHC (**L**arge **H**adron **C**ollider), wird derzeit am CERN (European Organisation for Nuclear Research¹) gebaut. Hier sollen Experimente stattfinden, von denen erwartet wird, dass man durch sie wichtige Erkenntnisse über den Aufbau der Materie erlangt. Unter anderem sind auch Strahlzeiten mit Bleikernen geplant. Hier wird es möglich sein, das QGP über einen längeren Zeitraum als am RHIC zu erhalten, und so genauer zu untersuchen. Das einzige Experiment, das darauf spezialisiert ist Schwerionenstöße zu untersuchen, ist das ALICE Experiment (**A** Large **I**on **C**ollider **E**xperiment). ALICE besteht aus mehreren Detektorsystemen, die verschiedene Aufgaben erfüllen sollen. Der Detektor, der in dieser Arbeit behandelt wird, ist der *Transition Radiation Detector* (TRD). Viele der Signaturen, über die man ein QGP erforschen kann, lassen sich durch hochenergetische Elektronen identifizieren. Aufgabe des TRD ist es Elektronen mit Impulsen größer als $1 \text{ GeV}/c$ zu erkennen. Dies gestaltet sich insofern schwierig, als dass der überwiegende Teil der erzeugten Teilchen Pionen sind². Mit dem TRD sollen 90% der Elektronen richtig identifiziert werden, bei einer Pionenkontamination von unter 1%.

¹früher: Conseil Européen de la Recherche Nucléaire

²Man erwartet bis zu 10^5 mal mehr Pionen als Elektronen.

Ziel dieser Arbeit ist es, verschiedene Verfahren zur e/π -Separation auf ihre Effizienz hin zu analysieren. Es werden klassische Methoden, die schon häufig eingesetzt worden sind, und eine Analysemethode mit neuronalen Netzwerken untersucht.

2. Theoretisches Rüstzeug

Für den weiteren Verlauf der vorliegenden Arbeit werden einige Grundlagen benötigt, über die hier ein kurzer Abriss gegeben werden soll.

2.1 Das Standardmodell der Teilchenphysik

2.1.1 Die Bausteine der Materie

Bei der Suche nach den Bausteinen der Materie hat man immer kleinere Teilchen entdeckt. Schon im alten Griechenland gab es philosophische Diskussionen darüber, ob ein Stein unendlich oft teilbar sei, oder ob man irgendwann zu elementaren Bausteinen (den so genannten Atomen)¹ käme, die nicht weiter teilbar wären.

Ende des 19. Jahrhunderts gelangte man schließlich zu der Überzeugung, dass auch diese „unteilbaren“ Atome² eine Substruktur haben müssen. Einen deutlichen Hinweis lieferten die fast 100 Elemente, deren Eigenschaften sich periodisch wiederholen. Durch Streuexperimente mit α -Teilchen an einer Goldfolie (Rutherford 1911) konnte gezeigt werden, dass Atome aus einem kleinen (im Verhältnis zum Atomdurchmesser) Kern bestehen und Elektronen vermutlich auf Bahnen um sie kreisen, was schließlich zum Bohrschen Atommodell führte. Bald fand man heraus, dass auch die Atomkerne wieder aus kleineren Teilchen bestehen, den Nukleonen (Protonen und Neutronen). Mit der Postulierung des Neutrinos (Pauli 1930) und der Entdeckung des Neutrons (Chadwick 1932) hatte man nun vier Teilchen (Proton, Neutron, Elektron und Neutrino), mit denen man alle damals bekannten Phänomene der Atom- und Kernphysik beschreiben konnte.

Mit Entdeckung des Myons 1946/47 tauchte ein neues Teilchen auf, das sich nicht in dieses einfache Schema einordnen ließ. Es hat ähnliche Eigenschaften wie das Elektron, ist jedoch um ein vielfaches schwerer. Im Laufe der 50er und 60er Jahre wurden zahlreiche weitere Teilchen gefunden, die man zwar einer von drei Gruppen (Leptonen, Mesonen und Baryonen) zuordnen konnte, es fehlte aber eine einfache Theorie um diese Teilchen in ein plausibles Schema zu bringen. In den 60er Jahren gelang es schließlich Gell-Mann und Zweig Ordnung in den „Teilchenzoo“ zu bringen. Hadronen, Teilchen die der starken Wechselwirkung unterliegen (also

¹ $\acute{\alpha}\tau\omicron\mu\omicron\varsigma$ (griech.) = unteilbar

²Hier ist die heutige Bedeutung der Atome gemeint, also die Teilchen aus denen die Elemente des Periodensystems bestehen. Die klassische griechische Bedeutung trifft heute eher auf die Quarks und die Leptonen zu.

Mesonen und Baryonen), bestehen aus elementarerer Teilchen, den Quarks. So setzen sich Mesonen aus einem Quark-Antiquark Paar zusammen, die Baryonen bestehen aus einem 3-Quark Bindungszustand. In den folgenden Jahren wurde diese Theorie immer weiter verfeinert und führte zum so genannten Standardmodell der Teilchenphysik.

Heute geht man davon aus, dass die Materie aus Leptonen und Quarks besteht, und die fundamentalen Wechselwirkungen durch Austauschteilchen beschrieben werden können (siehe Tab. 2.1 und Tab. 2.2). Leptonen und Quarks lassen sich in jeweils drei Familien (auch Generationen genannt) einordnen. Die Leptonen und ihre Antiteilchen lassen sich trennen in geladene (Elektron, Myon, Tauon) und ungeladene Teilchen (die zugehörigen Neutrinos). Die Quarks (und Antiquarks) unterscheiden sich durch ihren *Flavour* (Quarktyp). Jedes Quark besitzt einen zusätzlichen inneren Freiheitsgrad, die so gennante „Farbe“. Diese kommt in drei verschiedenen Varianten vor: rot, grün und blau. Hierbei ist zu beachten, dass die Farbkombination in den Hadronen additiv immer weiß ergeben muss, also für Mesonen Farbe und zugehörige Antifarbe, sowie für Baryonen rot, grün und blau. Außerdem sagt das Standardmodell noch ein weiteres Teilchen voraus, das so genannte Higgs-Boson, das schließlich für die Masse der Materieteilchen und der Austauschteilchen verantwortlich ist. Die Suche nach diesem Teilchen ist ein weiterer Schwerpunkt des LHC-Programms.

Fermionen	Familie			elektrische Ladung	Farbe	Spin
	1	2	3			
Leptonen	ν_e	ν_μ	ν_τ	0	-	$1/2$
	e^-	μ^-	τ^-	-1	-	$1/2$
Quarks	u	c	t	$+2/3$	r,b,g	$1/2$
	d	s	b	$-1/3$	r,b,g	$1/2$

Tabelle 2.1: Die Elementarteilchen... [Dem98].

Wechselwirkung	koppelt an	Austauschteilchen	Reichw. in m	rel. Stärke
stark	Farbladung	8 Gluonen	$\approx 10^{-15}$	1
elektromagnetisch	elektr. Ladung	1 Photon	∞	10^{-2}
schwach	schw. Ladung	3: W^+ , W^- , Z^0	$< 10^{-18}$	10^{-15}
Gravitation	Masse	Graviton?	∞	10^{-41}

Tabelle 2.2: ...und ihre Wechselwirkungen [Dem98],[Ger86].

2.1.2 Die fundamentalen Wechselwirkungen

Nach derzeitigem Kenntnisstand sind vier verschiedene Wechselwirkungen bekannt: die Gravitation, die elektromagnetische, die schwache und die starke Wechselwirkung.

Die schwächste Wechselwirkung ist die Gravitation. Sie ist z.B. verantwortlich für die Bildung der Himmelskörper und deren Bewegungen. Eine weitere Kraft, die uns aus dem Alltag bekannt ist, ist die elektromagnetische Wechselwirkung. Sie ist für die meisten alltäglichen Erscheinungen (Licht, chem. Bindungen, ...) zuständig. Die schwache Kraft, u.a. für Kernzerfälle verantwortlich, ist uns in der makroskopischen Welt ebenso wenig zugänglich wie auch die starke Wechselwirkung, die dafür Sorge trägt, dass Atomkerne und auch die Nukleonen zusammengehalten werden. Auf die starke Kraft wird im nächsten Unterkapitel eingegangen.

Die Wechselwirkungen werden nach heutigem Kenntnisstand über den Austausch von so genannten Vektorbosonen vermittelt [Pov93]. Das Vektorboson der elektromagnetischen Wechselwirkung ist das Photon. Es ist ungeladen und hat keine Masse. Die Austauschteilchen der schwachen Wechselwirkung sind das W^{+-} , das W^- und das Z^0 -Boson; sie besitzen alle eine Ruhemasse ($80\text{-}90 \text{ GeV}/c^2$), wobei aber nur die W -Teilchen eine elektrische Ladung besitzen. Die Gluonen sind die Austauschteilchen der starken Kraft, und tragen eine Farbladung (immer eine Farbe und eine Antifarbe). Es gibt 8 verschiedene Gluonen, die wie Photonen keine Ruhemasse besitzen und keine elektrische Ladung tragen.

Zum Austauschteilchen der Gravitation, dem Graviton, gibt es nur theoretische Überlegungen. Es sollte Spin 2 haben, masselos und sowohl elektrisch, wie auch farblich neutral sein. Man sollte hier ebenfalls noch erwähnen, dass die Gravitation nicht in das Standardmodell der Teilchenphysik eingebettet ist.

2.2 Die starke Wechselwirkung

Analog zur Quantenelektrodynamik (QED) für die elektromagnetische, hat man für die starke Wechselwirkung ebenfalls eine Eichtheorie, die Quantenchromodynamik, entwickelt. Während in der QED die Photonen als Austauschteilchen fungieren, sind dies in der QCD die Gluonen.

Die Gluonen werden als elektrisch neutrale Punktteilchen angesehen. Da sie masselos sind, müsste der Theorie zufolge die starke Wechselwirkung, wie auch die elektromagnetische und die gravitative Wechselwirkung, eine unendliche Reichweite haben³.

³Der Grund für die eingeschränkte Reichweite der schwachen Wechselwirkung sind eben ihre massiven Austauschteilchen.

Im Gegensatz zur QED, wo es nur eine Ladung gibt (die elektrische), gibt es in der QCD drei verschiedene Ladungen. Während die Photonen nicht geladen sind, tragen die Gluonen eine Farbladung, was bedeutet, dass sie sich wiederum gegenseitig beeinflussen. Dies wirkt sich auf größere Entfernungen abschirmend aus.

Die in der Natur vorkommenden Teilchen weisen alle eine nach außen neutrale Farbladung auf. So bestehen Mesonen aus zwei Quarks, die jeweils Farbe und Anti-Farbe tragen, Baryonen bestehen aus drei Quarks (rot, grün und blau). Es sind keine freien Quarks beobachtet worden (Confinement). Versucht man ein Quark aus seinem gebundenen Zustand zu befreien, so ist es irgendwann energetisch günstiger, ein neues Quark-Antiquark-Paar auszubilden, als ein Quark in einen freien Zustand zu bewegen. Dieses Verhalten spiegelt sich in dem Potenzial⁴ der starken Wechselwirkung wider:

$$V_s = -\frac{4}{3} \frac{\alpha_s}{r} + kr \quad [\text{Per90}]. \quad (2.1)$$

Hierbei ist α_s die Kopplungskonstante der starken Wechselwirkung, r der Abstand der beiden Quarks und k ein konstanter Faktor. Der erste Term spiegelt die Ähnlichkeit zum Coulomb-Potenzial bei kleinem r wider, er entspricht einem Ein-Gluon-Austausch [Per90]. Der zweite Term hingegen kommt von der Farb-Ladung der Gluonen und spielt bei größeren Abständen eine Rolle. Wird der Abstand genügend groß, erreicht die Energie in kr einen Wert, der ausreicht, um ein Quark-Antiquark-Paar zu bilden.

Eine weitere entscheidende Rolle spielt hier auch die Kopplungskonstante α_s . Bei ihr handelt es sich nicht um eine Konstante im eigentlichen Sinne, sie ist vielmehr abhängig vom Impulsübertrag, was der Entfernung zwischen den beteiligten Quarks entspricht. In erster Ordnung Störungsrechnung in der QCD ergibt sich:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f) \cdot \ln(Q^2/\Lambda^2)} \quad [\text{Pov93}]. \quad (2.2)$$

Wobei n_f die Zahl der beteiligten Quarktypen, Λ der Skalenfaktor der QCD und Q der Impulsübertrag bei der Streureaktion ist. Λ ist der einzige freie Parameter in der QCD und kann nur experimentell bestimmt werden. Sein Wert liegt bei ungefähr 250 MeV/c. Bei kleinen Quark-Quark-Abständen ist der Impulsübertrag groß und α_s wird klein. Die Kopplung verschwindet für sehr kleine Abstände nahezu. Man spricht auch von „asymptotischer Freiheit“. Falls die Abstände der beteiligten Quarks groß sind, wird α_s auch groß und das Problem lässt sich nicht mehr störungstheoretisch lösen. Statt dessen wendet man Gitterrechnungen an.

⁴Dieses Potenzial ist nur eine grobe Näherung und kann nur für schwere Quarks definiert werden.

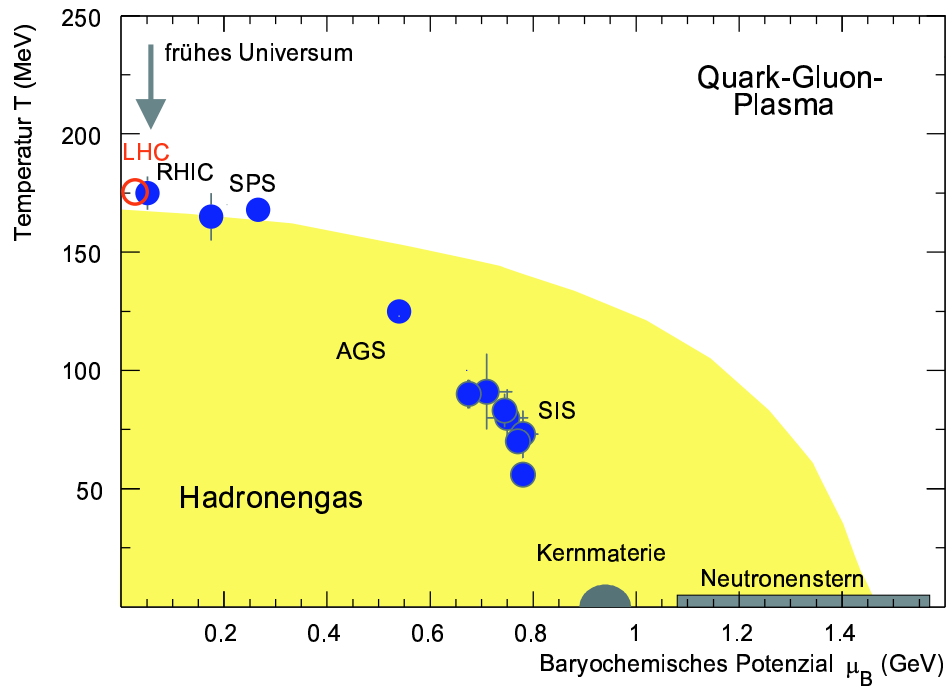


Abbildung 2.1: Das Phasendiagramm der Kernmaterie: Aufgetragen ist das baryochemische Potenzial μ_B gegen die Temperatur. Man erwartet sowohl für hohe Dichten (was einem großen μ_B entspricht), als auch für hohe Temperaturen einen Phasenübergang. Eingezeichnet sind die Messpunkte, die mit bisherigen Beschleunigern erreicht worden sind, sowie der Messpunkt, der mit dem LHC erreicht werden wird.

2.3 Das Quark-Gluon-Plasma

Gitterrechnungen sagen unter bestimmten Bedingungen einen Phasenübergang der Kernmaterie voraus. Bei sehr hohen Temperaturen (ca. 175 MeV [Kar02]) oder genügend hoher Netto-Baryondichte⁵ ($1 \text{ GeV}/\text{fm}^3$ [Lud03]) soll es zu einem *Deconfinement* kommen. Das bedeutet, die Quarks und Gluonen sind nicht mehr gebunden, sondern befinden sich in einem „quasi-freien“ Zustand, der als *Quark-Gluon-Plasma* (QGP) bezeichnet wird.

Geht man den Weg der höheren Dichte, so muss man einen kalten Kern ($T = 0 \text{ K}$) auf das ca. zehnfache der Dichte von Kernmaterie unter Normbedingungen bringen [Pov93]. Ist dies der Fall, so überlappen die Wellenfunktionen der Nukleonen, und die Quarks und Gluonen können nicht mehr einzelnen Nukleonen zugeordnet werden. Sie können sich im Kern frei bewegen. Man nimmt an, dass solche Bedingungen in Neutronensternen herrschen.

⁵Als Netto-Baryondichte bezeichnet man die Anzahl der Baryonen minus der Anzahl der Antibaryonen pro Volumen.

Bei ultrarelativistischen Schwerionenreaktionen geht man hingegen einen anderen Weg: Hier soll der Phasenübergang über die sehr hohe Temperatur erreicht werden. Ist bei so einer „heißen“ Kollision die Energiedichte ausreichend, so kann ein Quark-Gluon-Plasma entstehen, das aber nur kurze Zeit bestehen bleibt. Während das System sich ausdehnt, kühlt es sich ab, und die Quarks und Gluonen rekombinieren wieder zu Hadronen. Ist diese Rekombination vollständig, so liegt wieder ein Hadronengas vor. Das System kühlt sich noch weiter ab, die Hadronen können nicht mehr miteinander wechselwirken, es kommt zum *freeze out*. Solche Bedingungen haben vermutlich kurz nach dem Urknall geherrscht. In Abbildung 2.1 ist das Phasendiagramm der Kernmaterie dargestellt.

Das Quark-Gluon-Plasma kann nicht direkt nachgewiesen werden. Es gibt jedoch zahlreiche Signaturen, die auf eine Ausbildung eines Quark-Gluon-Plasmas hindeuten. Jede Signatur ist für sich alleine aber nicht hinreichend, erst die Messung und Kombination mehrerer Signaturen erlaubt es auf ein Quark-Gluon-Plasma zu schließen. Einige Signaturen sind elektromagnetischer Art. Diese sind ein gutes Mittel um Informationen aus dem Plasma zu erhalten, da Leptonen und Photonen nicht der starken Wechselwirkung unterliegen. Sind sie in der Reaktionszone entstanden, besteht eine große Wahrscheinlichkeit dafür, dass sie diese ungehindert verlassen und so Informationen direkt aus der heißen Anfangs-Phase der Reaktion liefern können. Zu den elektromagnetischen Observablen zählen Photonen, die thermodynamische Informationen aus allen Phasen der Reaktion liefern, und Dileptonenpaare, die durch Quark-Antiquark-Annihilation im QGP entstehen. Andere Signale sind hadronischer Natur. Sie geben Informationen über den Zustand nach dem thermischen Ausfrieren wieder. Die Produktion von Teilchen mit *strange*-Quarks gehört zu dieser Art von Signatur. Weitere Signale, die auf ein QGP hindeuten, sind die Unterdrückung von J/Ψ -Mesonen und Jet-Quenching.

Aufgabe des ALICE-TRD ist es einige dieser Signale zu identifizieren und zu untersuchen.

3. Experimenteller Rahmen

ALICE ist eines von fünf Experimenten, die am LHC am CERN in Genf aufgebaut werden. Während die anderen Experimente eher auf die Untersuchung von Proton-Proton-Stößen spezialisiert sind, wird ALICE Schwerionenstöße untersuchen, bei denen Schwerpunktsenergien von $\sqrt{s_{NN}} = 5,5$ TeV in *Pb-Pb*-Stößen erreicht werden sollen. In *p-p*-Kollisionen erreicht der LHC $\sqrt{s} = 14$ TeV. Der LHC soll im Jahr 2007 in Betrieb gehen.

3.1 Der LHC

Der LHC wird zur Zeit in dem 27 km langen ehemaligen LEP-Tunnel (Large Electron Positron Collider) gebaut. Es wird zwei Strahlröhren geben, in denen sowohl Proton-, als auch Schwerionen-Strahlen genutzt werden können. An vier Kollisionpunkten werden die verschiedenen Experimente aufgebaut. Der LHC wird nicht nur eine höhere Schwerpunktsenergie zur Verfügung stellen als RHIC¹, an dem die derzeit führenden Experimente der Schwerionenphysik stattfinden, sondern auch eine höhere Luminosität² aufweisen. Der LHC soll für *p-p*-Reaktionen eine Luminosität $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (RHIC: $\mathcal{L} = 1,4 \cdot 10^{31} \text{ cm}^{-2}\text{s}^{-1}$ [Har03]) und für *Pb-Pb*-Reaktionen einen Wert von $\mathcal{L} = 10^{27} \text{ cm}^{-2}\text{s}^{-1}$ (RHIC: $\mathcal{L} = 2 \cdot 10^{26} \text{ cm}^{-2}\text{s}^{-1}$ in *Au-Au*-Kollisionen [Har03]) erreichen.

Im Folgenden werden die Experimente am LHC, nämlich CMS, ATLAS, LHCb, TOTEM und ALICE kurz vorgestellt.

CMS

Eine der Hauptaufgaben des CMS-Experimentes (Compact Muon Solenoid) ist die Suche nach dem *Higgs-Boson*. Nach dem Standardmodell der Teilchenphysik sollten alle Elementarteilchen und Austauscheteilchen masselos sein. Masse erhalten sie erst durch Wechselwirkung mit einem Hintergrundfeld, das den ganzen Raum durchzieht. Das zugehörige Austauscheteilchen ist das Higgs-Boson. Durch seine

¹RHIC erreicht Schwerpunktsenergien bis zu $\sqrt{s_{NN}} = 200$ GeV bei *Au-Au*-Stößen und $\sqrt{s} = 500$ GeV für *p-p*-Kollisionen. [Hah03]

²Die Luminosität ist ein Maß für die Wechselwirkungsrate bei bekanntem Wirkungsquerschnitt.

Kopplung an die Teilchen erhalten diese ihre Massen. Aus experimentellen Ergebnissen weiß man, dass das Higgs-Boson eine Masse von mindestens $113,2 \text{ GeV}/c^2$ [Grz03], und aus theoretischen Überlegungen nicht mehr als $1000 \text{ GeV}/c^2$ haben sollte. Mit dem LHC ist es nun erstmals möglich diesen Energiebereich vollkommen zu überdecken, und so sollte dieses Teilchen auch nachgewiesen oder ausgeschlossen werden können.

Des Weiteren soll CMS wie auch LHCb die *CP-Verletzung* untersuchen und auch nach Signaturen *supersymmetrischer* Teilchen suchen.

ATLAS

Auch ATLAS (**A** Toroidal **LHC** **A**pparatus) hat ähnliche Aufgaben wie CMS. Die beiden Experimente sind so aufeinander abgestimmt, dass die von beiden produzierten Resultate miteinander verglichen werden können. ATLAS und CMS sind die größten Experimente, die am LHC geplant sind.

Wie bei ALICE soll auch beim ATLAS-Experiment ein Übergangsstrahlungsdetektor installiert werden. Bei diesem handelt es sich aber um einen *Transition Radiation Tracker*, der anders aufgebaut ist, und deshalb andere Methoden der e/π -Separation einsetzt als der ALICE-TRD. Der ATLAS-TRT besteht aus 4 mm dünnen *Straw Tubes*, die parallel zur Stahlachse ausgerichtet sind. Die Straw Tubes sind gasgefüllte Röhren, mit einem Anodendraht, der in der Mitte der Röhre die produzierten Cluster ausliest. Die Straws sind in Fasern eingebettet, in denen die Übergangsstrahlung erzeugt wird (siehe auch Kap. 5.1.1). Aufgrund der vollkommen anderen Konstruktion wird hier zur e/π -Separation die *Cluster-Counting*-Methode eingesetzt. Darauf wird später noch näher eingegangen.

LHCb

Hauptaufgabe des LHCb-Experiments (a Large **H**adron **C**ollider **B**eauty Experiment) ist die Untersuchung der CP-Verletzung in B-Mesonen Zerfällen. Die CP-Verletzung ist ein Phänomen, das nur in Prozessen auftritt, die durch die schwache Wechselwirkung³ bedingt werden. Einige Modelle, die über das Standardmodell

³Ein Beispiel für die P-Verletzung (P steht für Parität) ist der β -Zerfall. Es nehmen nur links-händige Teilchen, d.h. Teilchen deren Spin entgegengesetzt zu ihrem Impuls steht, am β -Zerfall teil. Man glaubte zunächst dies mit einer CP-Erhaltung (C steht hier für Ladungskonjugation) wieder beheben zu können, doch stellte sich heraus, dass dies in manchen Zerfällen von K-Mesonen nicht der Fall ist. B-Mesonen sind ähnlich aufgebaut wie K-Mesonen, nur enthalten sie statt eines *strange*-Quarks ein viel schwereres *beauty*-Quark.

hinausgehen, können mit Hilfe der CP-Verletzung das Übergewicht von Materie im Vergleich zu Antimaterie erklären. Außerdem soll LHCb einige Zerfallsmechanismen der B-Mesonen überprüfen, die das Standardmodell vorhersagt.

TOTEM

TOTEM (Total Cross Section, Elastic Scattering and Diffraction Dissociation at the LHC) ist das kleinste Experiment am LHC. Es wird in das CMS-Experiment integriert und soll den totalen Wirkungsquerschnitt der Kollisionen am LHC bestimmen. Hierzu wird der elastische und der inelastische Wirkungsquerschnitt gleichzeitig bei kleinem Impulsübertrag gemessen. Mit TOTEM kann die Luminosität des Colliders kalibriert werden.

3.2 ALICE

ALICE ist das einzige Experiment, das Fragestellungen der Schwerionenphysik in den Mittelpunkt stellt. Die Schwerpunktsenergie, die bei $Pb-Pb$ -Reaktionen erreicht wird, sollte ausreichen, um Bedingungen zu erzeugen, bei denen es zur Bildung eines Quark-Gluon-Plasmas kommt. Die einzelnen Detektoren sind optimiert, um verschiedene Signaturen eines möglichen Quark-Gluon-Plasmas nachzuweisen und zu untersuchen.

ALICE ist insgesamt 25 m lang und hat einen Durchmesser von ca. 16 m. Alle Detektorsysteme werden zusammen ein Gesamtgewicht von ca. 1000 t erreichen. ALICE lässt sich in zwei Bereiche gliedern, zum einen in den *Myonen-Arm* und zum anderen in den zentralen Detektorteil, der sich in dem *L3-Magneten* befindet (siehe Abb. 3.1). Im Myonen-Arm befinden sich Detektoren, die die bei Zerfällen von schweren Vektormesonen (z.B. J/Ψ , Υ) entstehenden Myonen nachweisen sollen. Der TRD befindet sich im zentralen Detektorteil, weshalb hier nur auf diese Detektoren näher eingegangen wird.

ITS

Das *Inner Tracking System* (ITS) ist der Detektor, der am nächsten zum Kollisionspunkt angeordnet ist. Das ITS besteht aus sechs zylinderförmigen, hochauflösenden Siliziumdetektoren. Seine primäre Aufgabe ist die Bestimmung des Vertex⁴ der $Pb-Pb$ -Kollisionen. Eine weitere Aufgabe ist die Bestimmung der Zerfallspunkte von

⁴Das ist der Punkt, an dem die Kollision stattfindet.

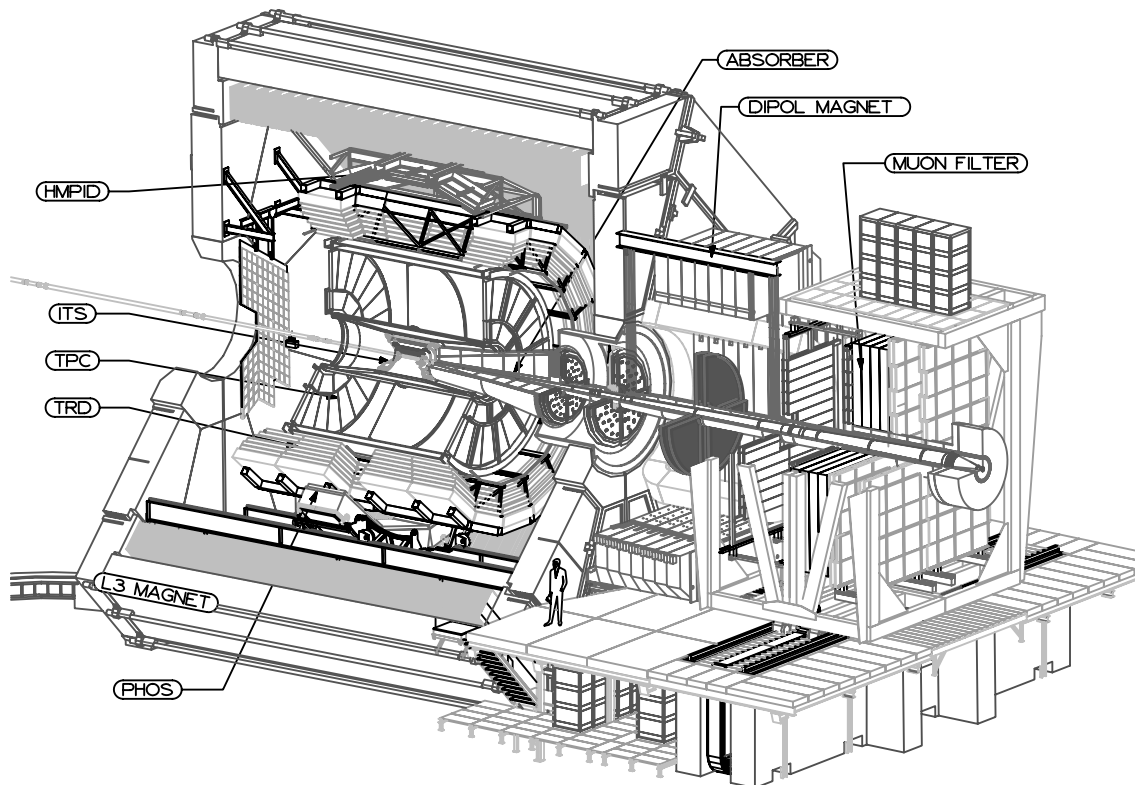


Abbildung 3.1: Das ALICE Experiment am LHC.

Hyperonen⁵ und Charm-Teilchen, die in den Kollisionen produziert werden. Das ITS kann außerdem Teilchen mit einem sehr niedrigen Impuls ($p < 100 \text{ MeV}/c$) identifizieren und erweitert damit das messbare Impulsspektrum hin zu niedrigen Werten.

TPC

Die *Time Projection Chamber* (TPC) ist der größte Detektor im ALICE-Experiment. Sie ist ein gasgefüllter Detektor, der sich um das ITS herum anordnet. Der innere Radius r_i beträgt 90 cm und der äußere $r_a = 250 \text{ cm}$. Hauptaufgabe der TPC ist die Spurrekonstruktion von geladenen Teilchen. Bei der Elektronenidentifizierung deckt sie den Impulsbereich $100 \text{ MeV}/c < p < 2,5 \text{ GeV}/c$ ab. Bei diesen Impulsen ist die Information über die deponierte Energie (dE/dx ; siehe Kapitel 4.2.1) ausreichend, um eine gute Rate bei der e/π -Separation zu erzielen.

⁵Baryonen mit *strange*-Quarks.

TOF und HMPID

Die *Time of Flight* Detektoren (TOF) und der *High Momentum Particle Identification Detector* (HMPID) sind beide Detektorsysteme, die ausschließlich der Teilchenidentifizierung dienen. Die TOF-Detektoren schließen sich an den TRD an, und überdecken den ganzen Raumwinkel, den auch der TRD überdeckt. Das TOF System soll Teilchen in einem Impulsbereich von $0,5 \text{ GeV}/c < p < 2,5 \text{ GeV}/c$ identifizieren. HMPID ist ein eher kleiner Detektor. Er ist weiter vom Vertex entfernt und hat eine viel kleinere Akzeptanz als die anderen bisher besprochenen Detektoren. Der HMPID wird aus RICH⁶-Detektoren bestehen, die mit flüssigem Freon gefüllt sind. HMPID soll Teilchen mit hohem Transversalimpuls⁷ bis $p_T < 5 \text{ GeV}/c$ identifizieren.

PHOS

Das *Photon Spectrometer* (PHOS) ist ein hochauflösendes elektromagnetisches Kalorimeter. Es wird im Abstand von 5 m zum Vertex aufgebaut und deckt eine Fläche von 18 m^2 ab. Mit PHOS sollen Photonen aus π^0 - und η -Zerfällen nachgewiesen werden. Außerdem sollen auch direkte Photonen gemessen werden.

TRD

Der letzte Detektor, der hier besprochen werden soll, ist der *Transition Radiation Detector* (TRD). Da sich die vorliegende Arbeit hauptsächlich mit diesem Detektor auseinandersetzt, ist ihm ein eigenes Kapitel (Kap. 4) gewidmet.

3.3 Schwerionenreaktionen am LHC

Bei zentralen Schwerionenkollisionen mit einer Schwerpunktsenergie von $\sqrt{s_{NN}} = 5,5 \text{ GeV}/c$, wie sie im LHC erreicht werden, werden vermutlich Energiedichten bis zu $\epsilon \simeq 500 \text{ GeV}/\text{fm}^3$ [Sch02] erreicht. Dies entspricht dem 3000fachen der Energiedichte eines Bleikerns im Grundzustand. Die Dichte, die erreicht werden müsste um ein QGP zu erzeugen, liegt bei etwa $\epsilon_c = 0,7 \text{ GeV}/c^2$ [Kar02]. Die Detektoren wurden so konzipiert, dass sie Multiplizitäten von bis zu 8000 geladenen Teilchen pro Event verarbeiten könnten. Dies entspricht bei zentralen Kollisionen einer Belegung von ca. 34% der Auslesepixel im ALICE-TRD [TDR01]. Die meisten dieser Teilchen sind Pionen.

⁶Ring Imaging Cherenkov Detektoren

⁷Zur Erläuterung siehe Anhang A.

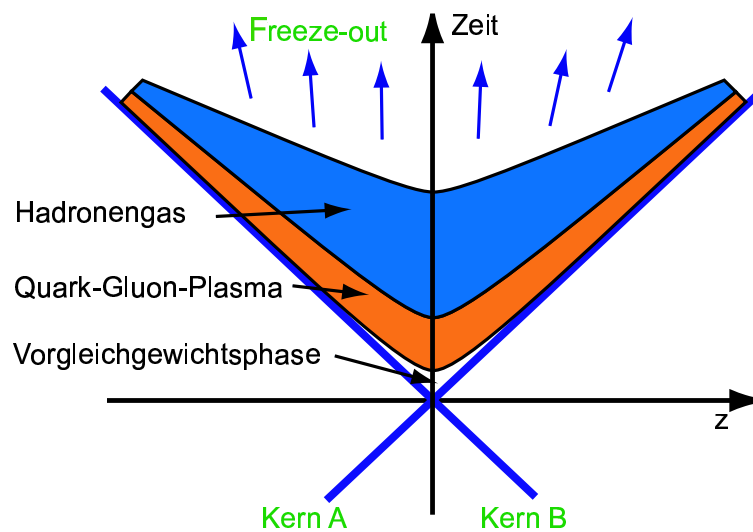


Abbildung 3.2: Verlauf einer Schwerionenreaktion.

In Abbildung 3.2 ist schematisch der zeitliche Verlauf einer Schwerionenreaktion dargestellt. Auf der y-Achse ist die Zeit aufgetragen. Von rechts und links nähern sich zwei Kerne (A und B) und kollidieren zum Zeitpunkt $\tau = 0$. Nach Durchlaufen der Vergleichgewichtsphase bildet sich ein QGP. Das System kühlt sich ab und es entsteht ein Hadronengas. Im weiteren Verlauf der Reaktion kommt es schließlich zum Freeze-out. Bei den Kollisionen, die am LHC stattfinden werden, beträgt die Lebensdauer der Plasmaphase bis zum Erreichen des vollständigen Freeze-outs etwa $\tau_f \simeq 30 - 40 \text{ fm}/c$ (ca. $1 - 1,3 \cdot 10^{-22} \text{ s}$) [Sch02]. Das Volumen, das der Feuerball beim Freeze-out einnimmt, liegt ungefähr bei $V \simeq 10^5 \text{ fm}$ [Sch02] (das 65-fache des Volumens eines Bleikerns).

4. Der Transition Radiation Detector (TRD)

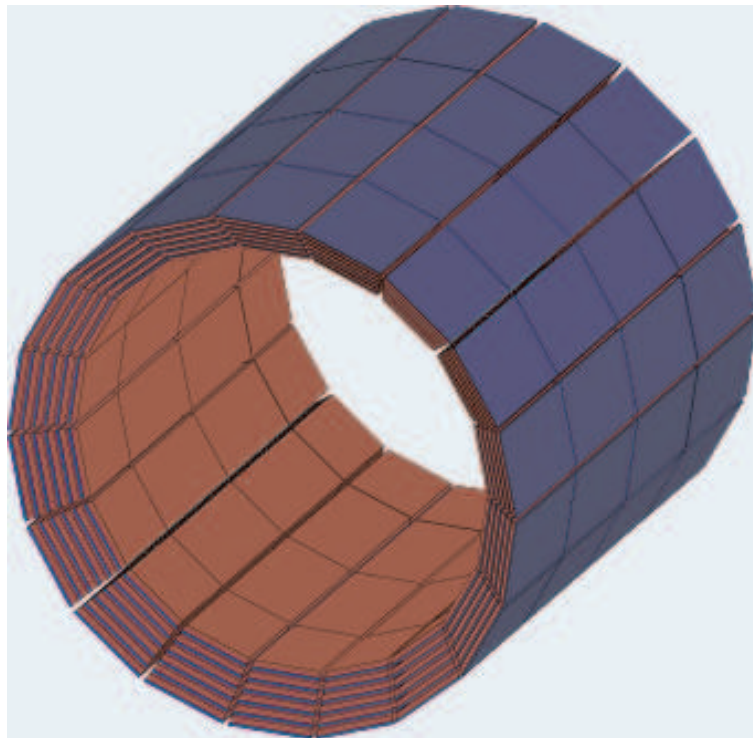


Abbildung 4.1: Schematische Darstellung des TRD. Man erkennt die tonnenförmige Struktur mit den 18 Supermodulen. Dies ist eine ältere Darstellung in der nur 4 Stacks pro Supermodul implementiert sind.

Der TRD schließt sich in radialer Richtung direkt an die TPC an und wird die Spurrekonstruktion und die Teilchenidentifizierung in ALICE vervollständigen. Zusammen mit dem ITS und der TPC werden somit Teilchenspuren vollständig bis zum Vertex rekonstruierbar sein. Außerdem können mit dem TRD Elektronen mit einem Impuls $p > 1$ GeV identifiziert werden, was die Elektronenidentifizierung im Experiment zu höheren Impulsen hin komplettiert.

Der TRD besteht aus insgesamt 540 einzelnen Modulen, die sich ringförmig zu 18 Supermodulen anordnen. Diese Supermodule bestehen in z -Richtung aus 5 Segmenten (*Stacks*) und jeweils 6 Ebenen (Abb. 4.1). Der innere Radius (r_i) beträgt 2,9 m, und der äußere Radius (r_a) 3,7 m. Der TRD wird über eine gesamte De-

tektorfläche von über 700 m² verfügen. Jedes TRD-Modul besteht aus Radiator, Driftkammer und Ausleseelektronik. Die Driftkammern werden mit insgesamt ca. 18 m³ eines Xe/CO_2 -Gemisches gefüllt. Die Auslese der Module wird über mehr als 1,2 Millionen Datenkanäle erfolgen [TDR01].

4.1 Physikalische Motivation

Die wichtigste Aufgabe des TRD (zusammen mit ITS und TPC) ist das Tracking von Teilchenspuren und die Teilchenidentifizierung. In Schwerionenreaktionen kommt es zu sehr hohen Multiplizitäten. Theoretische Berechnungen sagen für den LHC eine Teilchenmultiplizität von bis zu 8000 geladenen Teilchen pro *Rapiditätseinheit*¹ in zentralen Reaktionen voraus. Der überwiegende Teil dieser Teilchen sind geladene Pionen. Da Elektronen entscheidende Informationen aus und über die erfolgte Reaktion tragen, muss gewährleistet sein, dass sie auch als solche richtig identifiziert werden. Der TRD ist so konstruiert, dass mit ihm eine Spurerkennung der einzelnen Teilchen möglich wird. Der TRD soll außerdem als Elektronen-Trigger fungieren, und das ITS und die TPC bei der Teilchenerkennung unterstützen. Mit dem ITS und der TPC ist eine gute Teilchenidentifikation bis zu Impulsen von 2,5 GeV/*c* gewährleistet, der TRD erweitert dieses Spektrum hin zu höheren Impulsen. Im Folgenden wird ein kurzer Abriss über die QGP-Signaturen gegeben, die mit dem TRD nachgewiesen werden können.

J/Ψ- und Υ-Unterdrückung

Das J/Ψ und das Υ sind schwere Vektormesonen. Das J/Ψ besteht aus einem *charm*- und einem anti-*charm*- und das Υ aus einem *bottom*- und einem anti-*bottom*-Quark. Sollte es zur Ausbildung eines QGP kommen, sollte die Produktion von J/Ψ- und Υ-Mesonen, theoretischen Voraussagen zufolge, im Vergleich zu einem Szenario, in dem es zu keiner Ausbildung eines QGP kommt, unterdrückt sein.

Dileptonen Produktion

Während der Plasmaphase kommt es durch Quark-Antiquark-Annihilation zur Produktion von Dileptonenpaaren. Die gebildeten Leptonen unterliegen nicht der starken Wechselwirkung. Sie können daher die Reaktionszone weitestgehend ungehindert verlassen und führen so Informationen über die thermodynamischen

¹Zur Erläuterung siehe Anhang A.

Eigenschaften des Plasmas mit sich.

Verschiebung der Vektormesonmassen

Ein weiterer möglicher Phasenübergang, der aber nicht mit dem Übergang ins QGP übereinstimmen muss, ist die Wiederherstellung der *chiralen Symmetrie* [Pei97]. Die Konsequenz daraus wäre eine Verschiebung der leichten Vektormesonmassen zu kleineren Werten. Diese Massenverschiebung soll in ALICE über die Messung von e^+e^- -Paaren bestimmt werden, in die die Vektormesonen zerfallen können.

4.2 Theorie für den TRD

Ein geladenes Teilchen deponiert in den Driftkammern des TRD Energie. Dieser Energieverlust wird durch die *Bethe-Bloch-Formel* beschrieben. Ein anderer wichtiger Teil des TRD ist der Radiator, durch den unter bestimmten Bedingungen bei Teilchendurchgang *Übergangsstrahlung* produziert wird.

4.2.1 Die Bethe-Bloch-Formel

Durchquert ein geladenes Teilchen ein Medium, so gibt es durch Stöße mit den Atomen Energie ab. Dies führt zur Ionisation der Atome im Medium oder zu Anregungen der Hüllenelektronen. Durch die Bethe-Bloch-Formel wird dieser Energieverlust beschrieben:

$$-\frac{dE}{dx} = \frac{4\pi N_0 z^2 e^4}{mv^2} \frac{Z}{A} \left[\ln \left(\frac{2mv^2}{I(1-\beta^2)} \right) - \beta^2 \right] \quad [\text{Per90}]. \quad (4.1)$$

Dabei ist x die Dicke der durchquerten Schicht, N_0 die Loschmidt-Avogadro-Zahl, z ist die Ladung des Teilchens, v seine Geschwindigkeit, e die Elementarladung und m die Elektronenmasse. Z und A sind Ordnungszahl und Massezahl des durchquerten Mediums, I ist das mittlere Anregungspotenzial der Atome und $\beta = v/c$. In Abbildung 4.2 ist der schematische Verlauf des Energieverlusts gegen den *Lorentz-Faktor* ($\gamma = \frac{1}{\sqrt{1-\beta^2}}$) aufgetragen.

Am Verlauf der Funktion (Abb. 4.2) erkennt man zwei Bereiche. Bei sehr kleinem γ nimmt die deponierte Energie pro Wegstrecke zu größeren Werten von γ schnell ab. Dies lässt sich klassisch leicht verstehen, da die Teilchen auf ihrem Flug durch

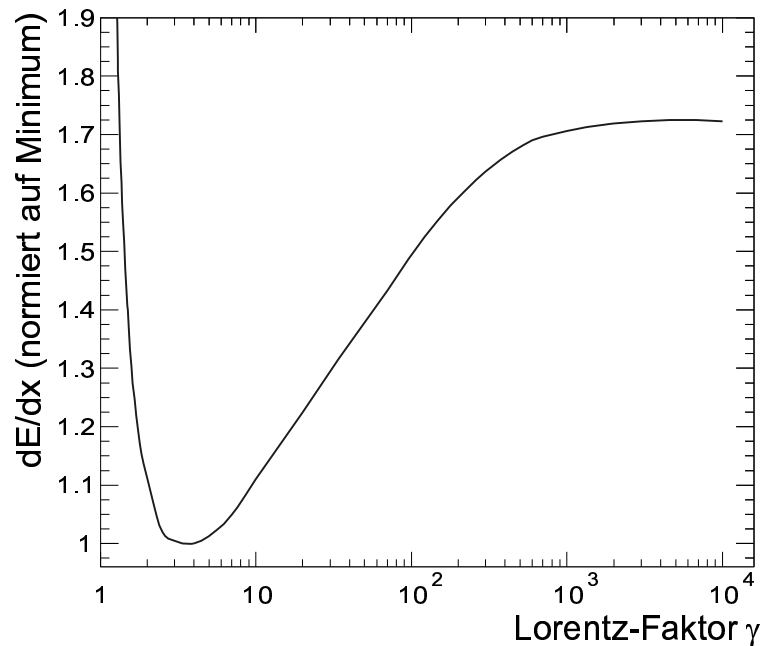


Abbildung 4.2: Schematische Darstellung des Energieverlusts eines geladenen Teilchens in einem Medium. Die Kurve ist auf ihr Minimum normiert.

das Medium weniger Zeit haben, mit den Atomen zu interagieren. Teilchen, deren γ genau im Minimum der Funktion liegt nennt man *minimal ionisierend*. Haben die Teilchen eine noch größere Geschwindigkeit, spielen relativistische Effekte eine Rolle. Das transversale elektrische Feld des Teilchens wächst an, so dass die Wechselwirkungsfläche auch Atome mit einschließt, die weiter entfernt von der Spur des Teilchens sind. Dies wird im logarithmischen Anstieg deutlich. Erreichen die Teilchen noch höhere Geschwindigkeiten, macht sich ein weiterer Effekt bemerkbar: Die Elektronen der Atome des Mediums schirmen das transversale elektrische Feld ab, es kann nicht beliebig groß werden. Die Kurve erreicht das so genannte *Fermi-Plateau*. Die Bethe-Bloch-Formel beschreibt nur den Energieverlust, der durch Stoßionisation und Anregung hervorgerufen wird. Es können aber auch andere Effekte eine Rolle spielen. Es kann zum Beispiel vorkommen, dass ein einzelnes Hüllenelektron relativ viel Energie absorbiert und seinerseits einen großen Ladungscluster erzeugt. Solche Elektronen nennt man δ -Elektronen.

Betrachtet man leichte Teilchen, wie Elektronen oder Positronen, muss man zusätzlich auch die Bremsstrahlung beachten, über die auch ein Teil der Energie abgegeben wird. Ursache hierfür ist die Abbremsung und die Ablenkung der Teilchen durch die Coulomb-Potenziale der Atome im durchquerten Medium. Dies entspricht einer negativen Beschleunigung, was zur Emission elektromagnetischer Strahlung

führt (siehe auch [Dem98],[Ger86]). Der Anteil der Bremsstrahlung am Energieverlust kann den Anteil, der über Ionisation abgegeben wird, bei hohen Energien um ein Vielfaches überschreiten.

4.2.2 Übergangsstrahlung

Der Bremsstrahlung verwandt ist die *Übergangsstrahlung* (*Transition Radiation* (TR)). Sie entsteht, wenn ein Teilchen mit konstanter Geschwindigkeit die Grenzfläche zwischen zwei Medien mit von einander verschiedenen Dielektrizitätskonstanten durchquert. Diese Strahlung wurde zuerst im optischen Bereich von Ginzburg und Frank 1946 für nicht-relativistische Teilchen vorhergesagt [Dol92] (nach [Gin46]). Zunächst schien die Übergangsstrahlung wenig geeignet dafür zu sein, in Detektoren genutzt zu werden, da ihre Intensität im optischen Bereich sehr gering ist. 1957 wurde aber von Garibian gezeigt, dass ultrarelativistische Teilchen Übergangsstrahlung im Röntgenbereich bei gleichzeitiger Intensitätszunahme emittieren [Dol92] (nach [Gar57]).

Für den Fall ultrarelativistischer Teilchen ($\gamma \gg 1$) an einer Grenzschicht ergibt sich für die emittierte Strahlung, abhängig von ihrer Frequenz und dem Abstrahlungswinkel des TR-Photons, folgender Ausdruck:

$$\frac{d^2W}{d\omega d\vartheta} = \frac{2\alpha\hbar\vartheta^3}{\pi} \left(\frac{1}{1/\gamma^2 + \vartheta^2 + \omega_1^2/\omega^2} - \frac{1}{1/\gamma^2 + \vartheta^2 + \omega_2^2/\omega^2} \right)^2 \quad [\text{Ego00}] \quad (4.2)$$

Hier ist ω die Frequenz und ϑ der Abstrahlungswinkel des TR-Photons, α die *Feinstrukturkonstante*, $\omega_{1,2}$ sind die *Plasma-Frequenzen*² der durchquerten Medien und $\hbar (= h/2\pi)$ die *reduzierte Planck-Konstante*. Die Gleichung hat ein Maximum bei einem Winkel von $\vartheta = 1/\gamma$, Übergangsstrahlung wird also in einem engen Kegel um die Bewegungsrichtung des Teilchens emittiert. Für ultrarelativistische Teilchen liegt das Frequenzmaximum im Röntgenbereich. Integriert man nun über alle Winkel und alle Frequenzen, so erhält man die gesamte abgegebene Energie:

$$W = \frac{\alpha\hbar}{\pi} \frac{(\omega_1 - \omega_2)^2}{(\omega_1 + \omega_2)} \gamma \quad [\text{Ego00}]. \quad (4.3)$$

Die Intensität der Strahlung ist damit für eine bestimmte Grenzschicht nur vom Lorentz-Faktor γ abhängig. Elektronen haben eine viel geringere Masse als Hadronen. Die Abhängigkeit vom Lorentz-Faktor macht Übergangsstrahlung zu einem

²Unter der *Plasmafrequenz* versteht man eine Art Eigenschwingung eines insgesamt elektrisch neutralen Mediums, in dem aber die Raumladungen gegeneinander verschoben sind. Diese Raumladungen schwingen mit der Plasmafrequenz gegeneinander [Wah00].

geeigneten Hilfsmittel, um Teilchen mit gleichem Impuls, aber unterschiedlichen γ s (zum Beispiel Elektronen und Pionen), voneinander zu unterscheiden. Ein Nachteil ist allerdings die geringe Anzahl der emittierten Photonen. Sie liegt im Mittel pro Grenzschicht bei ungefähr $\alpha = \frac{1}{137}$. Deshalb benutzt man in Übergangsstrahlungsdetektoren entweder Folienstapel, Mikrofasern oder andere Materialien mit vielen Grenzschichten.

Die Anzahl der TR-Photonen lässt sich nicht beliebig steigern. Die Ausbeute an Übergangsstrahlung geht für viele Grenzschichten in eine Sättigung über. Bei vielen Grenzschichten spielen Interferenzeffekte eine große Rolle. Zusätzlich lässt sich die Absorption durch den Radiator nicht mehr vernachlässigen. Eine analytische Berechnung der TR lässt sich nur für einfache Anordnungen der Grenzschichten wie z.B. Folienstapeln durchführen. Hier befinden sich eine ganze Reihe von Folien in regelmäßigen Abständen zueinander, die Zwischenräume sind meist mit einem Gas gefüllt. Die anderen Radiatorarten (Schäume oder Fasermatten) weisen hingegen viele unregelmäßige Strukturen auf, was eine analytische Untersuchung unmöglich macht.

4.3 Aufbau eines TRD-Moduls

Jedes einzelne der 540 TRD-Module wird eine Dicke von etwas mehr als 10 cm haben. Dabei entfallen 4,8 cm auf den Radiator, 3,7 cm auf die Driftkammer und der Rest auf die Ausleseelektronik [TDR01].

Bei dem Radiator handelt es sich um einen *Sandwichradiator*. Das bedeutet, er ist aus mehreren Schichten verschiedener Materialien aufgebaut. Üblicherweise werden in Übergangsstrahlungsdetektoren Folienstapel verwendet. Um Folienstapel im TRD einzusetzen, müsste man stabile Metallrahmen einbauen, die die Folienstapel tragen und spannen. Dies hätte zur Folge, dass große Bereiche des Detektors aus inaktivem Material bestünden, was die geometrische Akzeptanz des Detektors stark verkleinerte. Außerdem könnten Detektoren, die weiter außen stehen (z.B. HMPID, TOF und PHOS), nur noch einen Bruchteil der entstandenen Teilchen in ihrem ursprünglichen Zustand detektieren. Ein weiterer Punkt, der eine wichtige Rolle spielt, sind Stabilitätskriterien. Der Radiator soll die Driftkammer nach vorne hin schließen, die Driftelektrode wird direkt auf den Radiator geklebt. In der Driftkammer herrscht ein kleiner Überdruck wegen des ständigen Gasaustausches, außerdem kann bei nicht-waagerechter Kammerorientierung die Schwere des Kammergases nicht vernachlässigt werden. Um zu verhindern, dass diese Faktoren zu einer nicht mehr akzeptablen Durchbiegung der Driftelektrode und damit zu

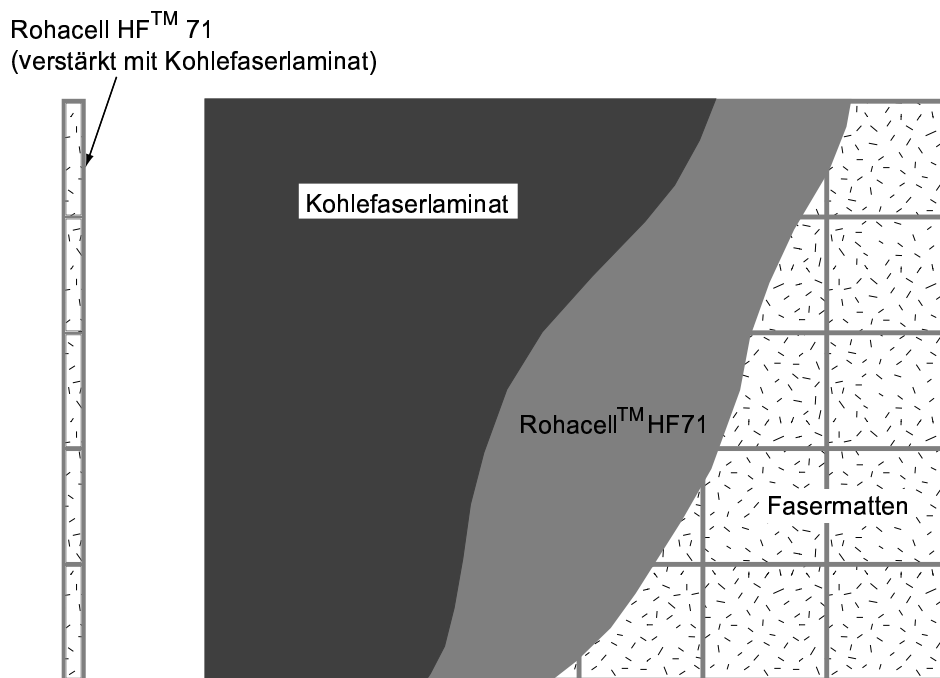


Abbildung 4.3: Schnitt durch einen Sandwichradiator. Die Fasermatten sind in ein Gerüst aus *Rohacell™ HF71* eingebettet. Zur zusätzlichen Verstärkung wird frontal und auf der Rückseite eine dünne Schicht aus Kohlefaserlaminat aufgetragen [TDR01].

verzerrten Driftfeldern führen, muss man darauf achten, dass der Radiator stabil genug ist. Letztendlich hat man sich für einen Sandwichradiator entschieden, weil er zum einen gute Übergangsstrahlungseigenschaften hat, zum anderen auch zur Stabilität der Kammer beiträgt.

Der Radiator besteht außen aus einer jeweils 8 mm dicken Schicht *Rohacell™ HF 71*, die durch ein dünnes Kohlefaserlaminat verstärkt ist. Im Inneren des Radiators befinden sich Fasermatten, die in ein *Rohacell™* Gitter einglassen werden (siehe auch Abb. 4.3). Die Fasermatten weisen ähnlich gute Übergangsstrahlungseigenschaften auf wie ein Folienstapel. Zusätzlich trägt auch *Rohacell™*, das stabilisierende Element, zur TR-Produktion bei.

An den Radiator schließt sich eine zeitauflösende Driftkammer an. Die Driftkammer unterteilt sich in eine Driftzone von 3 cm Tiefe und eine Verstärkungszone von 0,7 cm, diese sind durch geerdete Potenzialdrähte voneinander getrennt. Die Kammer ist mit einem Gasgemisch gefüllt, dass zu 85% aus Xenon und zu 15% aus Kohlendioxid besteht. Über einen geschlossenen Kreislauf wird das Gas ständig ausgetauscht und gereinigt. In der Kammer herrscht ein leichter Überdruck von max. 0,5 mbar. Je nach Position der Kammer kann die Schwere des Xenons diesen Druck auf bis zu 1 mbar erhöhen. Die Driftelektrode liegt auf einem Potenzial von

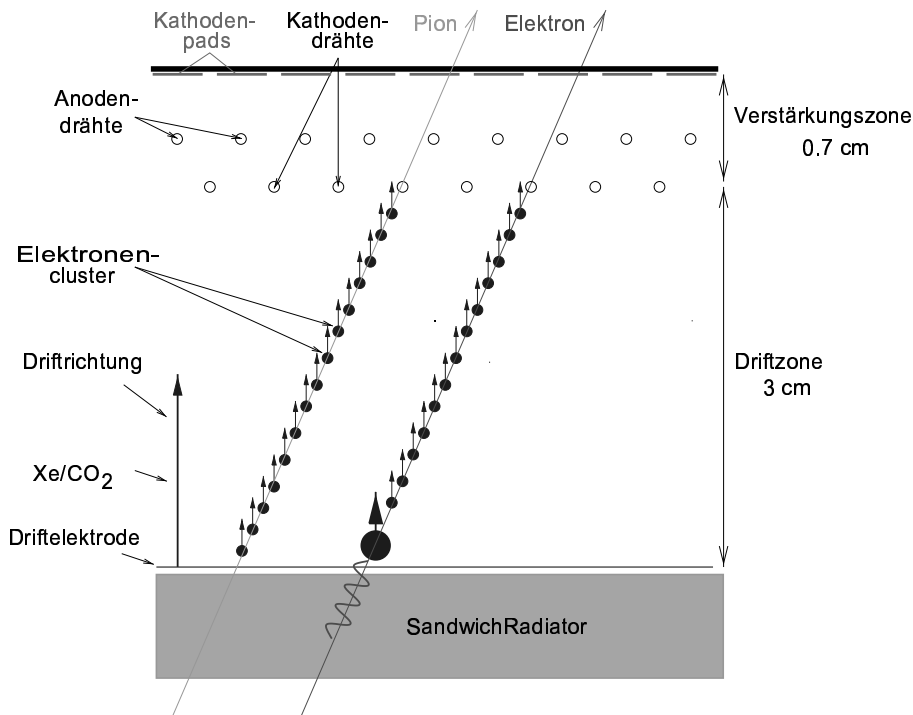


Abbildung 4.4: Funktionsweise eines TRD-Moduls. Ein Pion hinterlässt eine Ionisationsspur aus Elektronenclustern, während ein Elektron hier zusätzlich ein Übergangsstrahlungsphoton erzeugt [TDR01].

$U_d = -2,1$ kV. Sowohl Potenzialdrähte als auch Auslesepad liegen auf Erdpotential. In der Mitte der Verstärkungszone, zwischen Potenzialdrähten und Auslesepad, befinden sich die Anodendrähte. Auf die Anodendrähte wird ein positives Potenzial von $U_a = 1,6$ kV gelegt. Bei Teilchendurchflug werden in der Verstärkungszone Lawinen von Ladungsträgern produziert, die auf den Auslesepad weitere Ladungen influenzieren. Mit einem FADC³ werden diese mit einer Auflösung von 100 ns pro Zeitkanal ausgelesen.

4.4 Funktionsweise des TRD

Durchquert ein geladenes Teilchen den TRD, so passiert es zunächst den Radiator. Abhängig von seinem Lorentz-Faktor γ erzeugt es hierbei mit einer bestimmten Wahrscheinlichkeit ein Übergangsstrahlungsphoton. Dieses Photon wird praktisch in Flugrichtung des Teilchens emittiert (siehe Kap. 4.2). Beide Teilchen treten in

³Flash Analog Digital Converter.

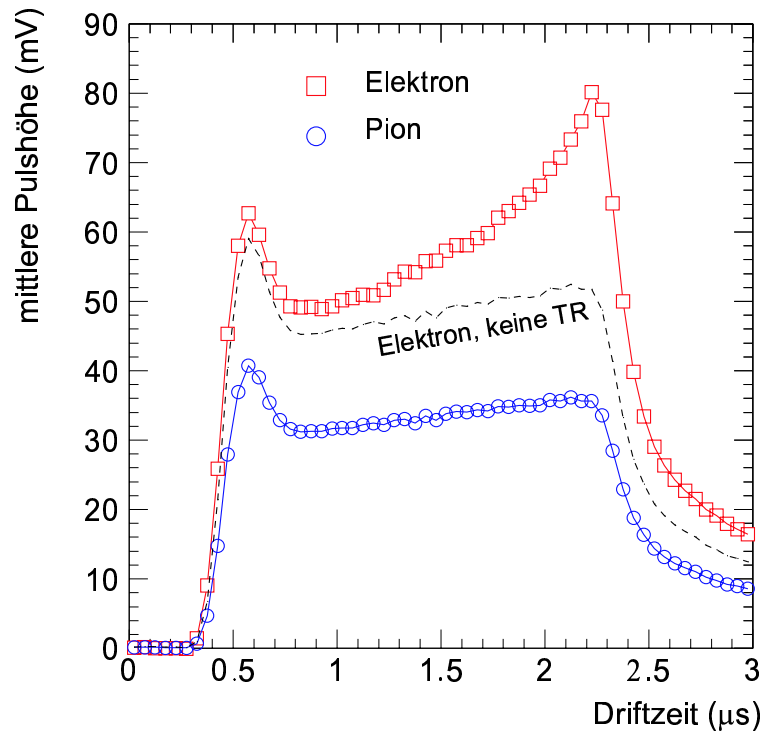


Abbildung 4.5: Gemitteltetes Driftspektrum für Elektronen und Pionen. Der Verstärkungspeak ist in beiden Verteilungen zu sehen, der TR-Peak nur für Elektronen. Die gestrichelte Linie gibt den mittleren Energieverlust für Elektronen ohne Übergangsstrahlung wieder [TDR01].

die Driftkammer ein, wo das Photon aufgrund der hohen Ordnungszahl des Kammergases Xenon schnell absorbiert wird und einen *Elektronencluster* erzeugt. Dieser Elektronencluster driftet in Richtung des Verstärkungsbereichs, wo er dann zu den Anodendrähten hin beschleunigt wird. Diese Beschleunigung ist so stark, dass sich durch Stöße mit den Gasatomen Lawinen ausbilden. Das Teilchen hinterlässt auf seinem Weg eine Ionisationsspur, die ebenfalls aus Elektronenclustern besteht. Auch diese driften in den Verstärkungsbereich und werden dort lawinenartig verstärkt.

Durch den Einsatz einer zeitauflösenden Driftkammer kann man nicht nur die gesamte erzeugte Ladung als Information nutzen. Zusätzlich erhält man auch Informationen, wo und in welcher Intensität Energie deponiert worden ist. In Abbildung 4.5 sind gemittelte Driftspektren von Elektronen und Pionen zu sehen. Man erkennt, dass das mittlere Signal der Elektronen deutlich höher ist, als das der Pionen (ungefähr ein Faktor von 1,4). Das liegt daran, dass Pionen mit einem Impuls von einigen GeV/c zu den minimalionisierenden Teilchen gehören, und weniger Energie pro Weglänge deponieren, als Elektronen, die bei gleichem Impuls schon auf dem Fermi-Plateau liegen. Bei kurzen Driftzeiten findet man den so genannten *Verstär-*

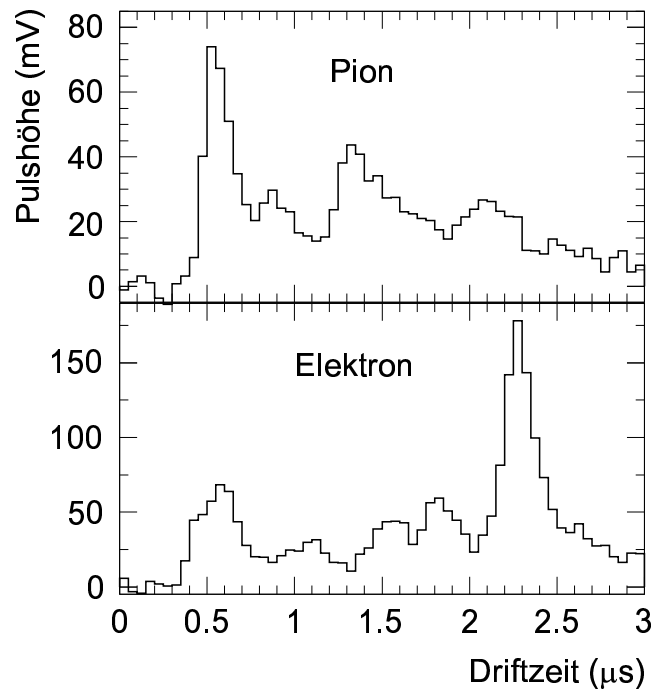


Abbildung 4.6: Typische Einzelspektren eines Elektrons und eines Pions. Die y-Achsen sind unterschiedlich skaliert [TDR01].

kungspeak. Er kommt dadurch zu Stande, dass im Verstärkungsbereich Ladungen, die vor und nach den Anodendrähten erzeugt werden, gleiche Driftzeiten haben. Die von den Übergangsstrahlungsphotonen produzierten Cluster müssen den ganzen Driftbereich durchqueren und werden kurz nach Eintritt in die Driftkammer absorbiert. Sie findet man in dem hohen Anstieg am Ende des Elektronenspektrums bei langen Driftzeiten wieder. Bei den Pionen fehlt dieser Anstieg. Die gestrichelte Linie gibt den Verlauf eines gemittelten Driftspektrums für Elektronen wieder, wenn kein Radiator verwendet wird, also ohne Übergangsstrahlung.

In Abbildung 4.6 sind Einzelspektren eines Elektrons und eines Pions dargestellt. Wieder erkennt man die im Mittel höhere deponierte Energie für das Elektron (man beachte die unterschiedliche Skalierung der y-Achsen). Ebenso sind in beiden Spektren für kleine Driftzeiten deutliche Peaks aus der Verstärkungsregion zu erkennen. Im Elektronenspektrum findet man einen sehr hohen Peak bei späten Driftzeiten, der typischerweise durch Übergangsstrahlung erzeugt worden ist. Ein anderer Vorgang, der zu hohen Peaks führt, ist die Produktion von δ -Elektronen. δ -Elektronen entstehen durch seltene harte Stöße des Primärteilchens mit den Gasatomen des Kammergases. δ -Elektronen können Energien von einigen keV besitzen, die sie wiederum

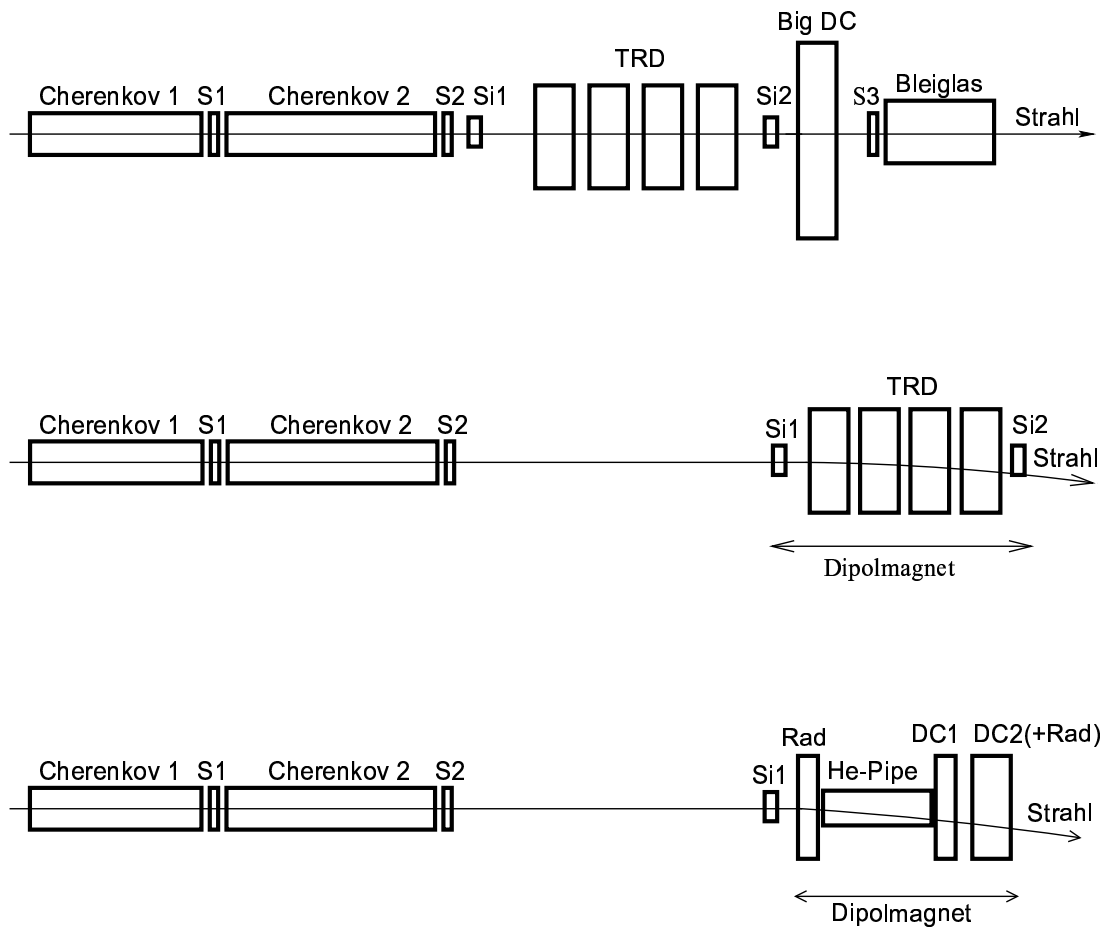


Abbildung 4.7: Versuchsaufbauten während der Strahlzeit 2002 am CERN PS. Oben ist die Versuchsanordnung für den Großteil der Messungen zu sehen, in der Mitte der Aufbau zu den Kammertests im Magnetfeld und unten das Setup zur Energiemessung der TR-Photonen.

durch Ionisationsprozesse abgeben. Dabei entstehen so genannte δ -Cluster, die als TR-Cluster missidentifiziert werden können. Ein Unterschied dabei ist aber, dass δ -Cluster zu beliebigen Zeiten auf der Spur der Primärteilchen entstehen können. Durch Ausnutzen dieser Information können von δ -Elektronen produzierte Cluster identifiziert werden.

4.5 Die Teststrahlzeit 2002 am CERN

Die Daten mit denen die Untersuchungen in dieser Arbeit gemacht worden sind, stammen aus einer Strahlzeit, die im Herbst 2002 am CERN PS⁴ stattgefunden hat.

⁴Proton Synchrotron.

Ziel dieser Strahlzeit war es das endgültige Radiatordesign inklusive der verwendeten Materialien festzulegen, sowie den ersten Prototypen einer großen Driftkammer⁵ und erstmals mehrere identische TRD-Module hintereinander geschaltet zu testen. Außerdem ist das Verhalten der Kammern in einem Magnetfeld und der Einfluss verschiedener Einstrahlwinkel getestet worden. Erstmals wurde bei den TRD-Modulen das Energiespektrum der Übergangsstrahlungsphotonen vermessen. Um den Einfluss von sekundär erzeugten Teilchen zu überprüfen, die eventuell durch die vorgelagerten Detektoren entstehen könnten, wurden Messungen mit einem *Dummy*-Material, das vor dem ersten Modul angebracht worden ist, gemacht.

4.5.1 Der Versuchsaufbau

In Abbildung 4.7 sind die verschiedenen Versuchsaufbauten schematisch dargestellt. Der gemischte e/π -Strahl wurde durch einen 24 GeV/ c Protonenstrahl erzeugt, der auf ein Beryllium-Target traf [And04]. Die Messungen wurden bei Impulsen von 1 GeV/ c bis maximal 7 GeV/ c des sekundären e/π -Strahls durchgeführt. Der Elektronenanteil zeigte eine starke Impulsabhängigkeit. Bei maximalem Impuls lag er bei wenigen Prozent. Die TRD-Prototypen sind in der Abbildung mit TRD bezeichnet. Die große Kammer (Big DC) ist nur bei einigen Runs eingesetzt worden. In der Strahlachse befinden sich außerdem drei Szintillatoren (S1, S2 und S3), zwei Silizium-Streifenzähler (Si1 und Si2), zwei Cherenkov-Detektoren und ein Bleiglas-Kalorimeter. Die Szintillatoren definieren die Strahlachse durch die TRD-Module und sind Teil des Triggersystems. Die Ortsbestimmung des Strahls erfolgte durch die Silizium-Streifenzähler. Für die Messungen, die in einem Magnetfeld gemacht wurden, sind die kleinen TRD-Module in einen Dipolmagneten gestellt worden (siehe Abb. 4.7 Mitte).

Für die Energiemessung der Übergangsstrahlungsphotonen wurde zusätzlich noch ein mit Helium gefülltes Rohr (He-Pipe) dem Aufbau hinzugefügt (Abb. 4.7 unten). Die geladenen Teilchen durchqueren den Radiator, der in diesem Aufbau alleine ohne Driftkammer aufgestellt ist. Die Teilchen werden auf ihrem weiteren Weg durch das Magnetfeld abgelenkt, während die TR-Photonen von diesem unbeeinflusst durch die Helium-Pipe fliegen. Helium eignet sich gut, um die Absorption der γ s gering zu halten, da die Absorptionswahrscheinlichkeit für Photonen von der Ordnungszahl abhängt. Die TR-Photonen durchqueren die Röhre nahezu ungehindert. In der anschließenden Driftkammer können die Signale räumlich

⁵Dies war der erste Test mit einer Kammer, die ungefähr den Abmessungen entspricht, wie sie im TRD eingesetzt werden soll. Bei den bisherigen Strahlzeiten standen nur kleinere Prototypen zur Verfügung.

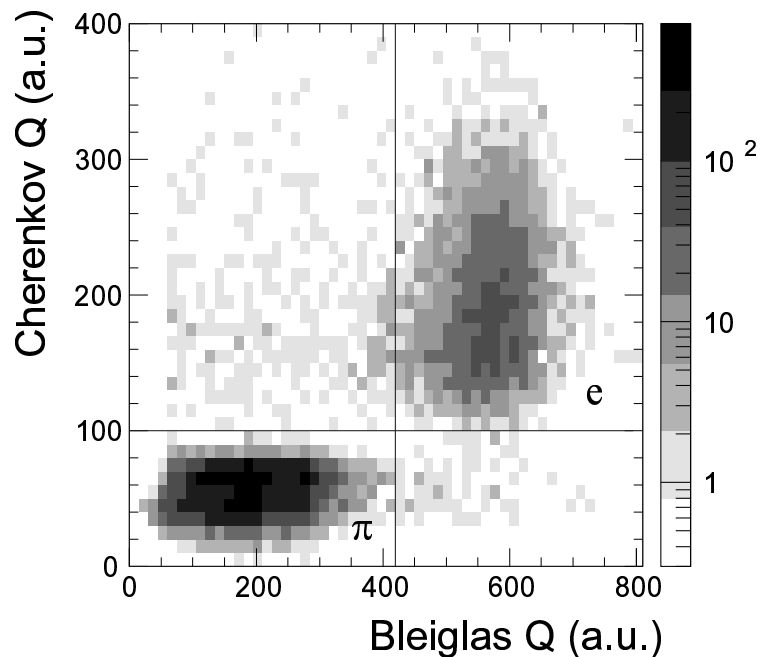


Abbildung 4.8: Separation von Elektronen und Pionen mit Hilfe von Cherenkov-Zähler und Bleiglas-Kalorimeter. Die gemessenen Werte für Elektronen (rechts oben) sind sowohl für Cherenkov- als auch für Bleiglas-Detektoren höher, als die für Pionen (links unten). Mit Schranken lassen sich die Teilchensorten trennen. Nur wenn die Einträge in allen drei Detektoren (hier nur zwei dargestellt) über diesen Schranken liegen, handelt es mit sehr großer Wahrscheinlichkeit um Elektronen, liegen die Werte in allen Detektoren unter den Cuts, hat ein Pion diesen Eintrag erzeugt[TDR01].

getrennt voneinander aufgenommen werden.

In der vorliegenden Arbeit wurden ausschließlich Daten verwendet, die mit dem ersten Versuchsaufbau aufgenommen worden sind.

4.5.2 Das Triggersystem

Der Trigger für das Experiment wurde durch eine Koinzidenzschaltung zwischen den Szintillatoren und den Cherenkov-Detektoren realisiert. Da Elektronen und Pionen im Experiment die gleichen Impulse, aber wegen ihrer Massen unterschiedliche Geschwindigkeiten hatten, konnten Elektronen durch die Cherenkov-Detektoren identifiziert werden. Die Pionengeschwindigkeiten lagen unter der *Cherenkov-Schwelle*⁶, so dass von ihnen keine *Cherenkov-Strahlung* emittiert wurde. Dies

⁶Haben Teilchen eine Geschwindigkeit, die die Geschwindigkeit des Lichtes in dem Medium übersteigt, wird Strahlung, die so genannte Cherenkov-Strahlung emittiert. Die Geschwindigkeit, ab der diese Strahlung ausgesandt wird, nennt man Cherenkov-Schwelle.

ermöglichte den Einbau eines *Scaledown-Faktors* in die Triggerlogik. Das bedeutet, dass man nur einen bestimmten Bruchteil von Pionen-Ereignissen triggert, aber jedes Elektronen-Ereignis, um einen ungefähr gleichen Anteil von Elektronen und Pionen zu erhalten.

Bei den späteren Offline-Analysen nutzt man zusätzlich zu den Signalen der Cherenkov-Detektoren noch das Signal des Bleiglas-Kalorimeters. Elektronen bilden im Bleiglas-Detektor elektromagnetische Schauer aus und verlieren nahezu ihre gesamte Energie. Die Pionen hingegen deponieren ihre Energie durch hadronische Streuprozesse. Die mittlere freie Weglänge ist bei diesen Vorgängen viel größer als bei elektromagnetischen Ereignissen, die Pionen deponieren nur einen kleinen Teil ihrer Energie. Setzt man nun Schwellen für die Signale aus Bleiglas- und Cherenkov-Detektoren, kann man die Teilchen voneinander separieren. In Abbildung 4.8 sind die erzeugten Signale und die Schwellen für zwei Detektoren eingezeichnet. Nur Teilchen, die in allen drei Detektoren Signale erzeugen, die über den jeweiligen Schwellen liegen, werden als Elektronen angesehen, während die Signale von Pionen in allen drei Detektoren unter diesen Schwellen liegen müssen.

5. Klassische Methoden zur e/π -Separation

5.1 N-Methoden und Q-Methoden

Um die Informationen der Übergangsstrahlung zur e/π -Separation zu nutzen gibt es zwei unterschiedliche Konzepte, die auf der Anzahl (N) oder auf der deponierten Ladung (Q) beruhen.

5.1.1 N-Methoden

Da man weiß, dass Elektronen mit hohem Impuls mit einer größeren Wahrscheinlichkeit Übergangsstrahlung produzieren als Pionen, kann man „feinkörnige“ (*fine grain*) Detektoren bauen, die aus der Anzahl der erzeugten TR-Photonen Rückschlüsse auf die Teilchenidentität erlauben. Ein Beispiel für einen solchen Detektor ist der ATLAS-TRT. In Abbildung 5.1 ist ein typisches ATLAS-TRT Modul dargestellt. Im Inneren des Moduls befinden sich Straw Tubes. Dabei handelt es sich um 4 mm dicke Proportional-Zählrohre. Je nach Position fasst ein Modul zwischen

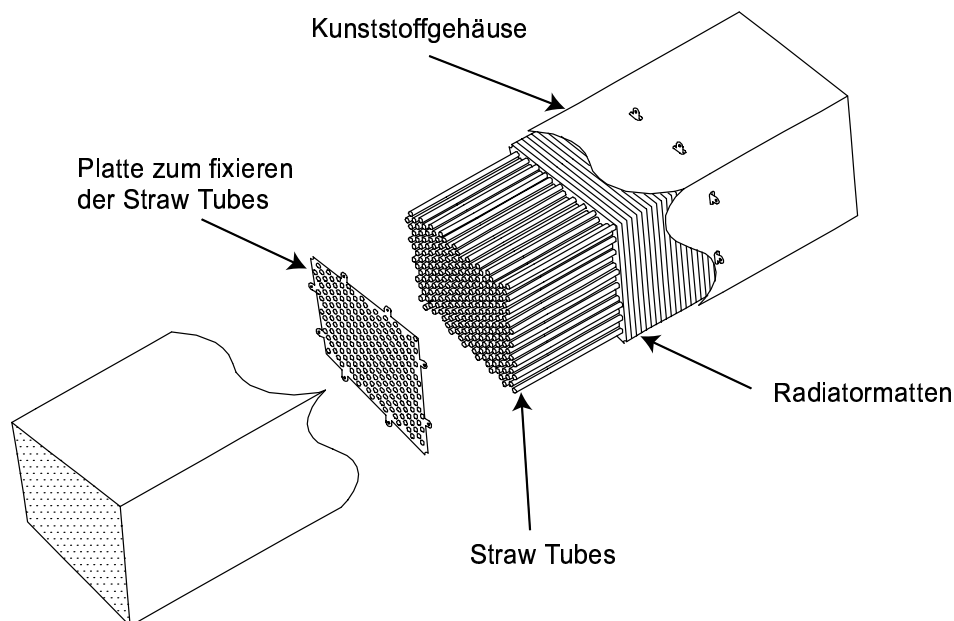


Abbildung 5.1: Ein ATLAS-TRT Modul [ATL97].

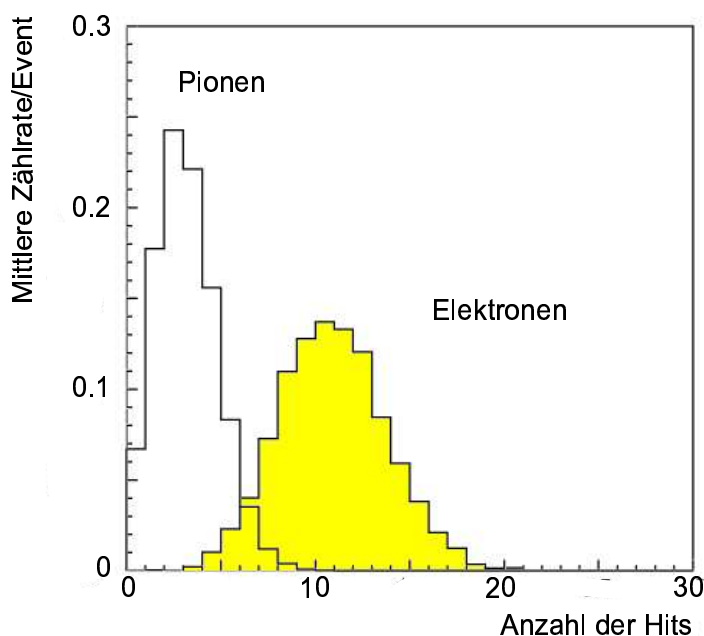


Abbildung 5.2: Die N-Methode: Im Mittel erzeugte Treffer über dem high threshold pro Track. Gezeigt ist eine typische Verteilung, wie sie von Transition Radiation Trackern produziert wird.

329 und 793 Straw Tubes. Die Tubes sind in Polypropylen/Polyethylen-Fasermatten eingebettet, die das Radiatormaterial des Detektors bilden. Außen sind die Module mit einem kohlefaserverstärkten Kunststoff umgeben. Im zentralen Teil¹ des TRT werden die Module in drei Schichten angeordnet. Weitere Informationen findet man in [ATL97].

Die N-Methoden zur e/π -Separation basieren auf dem so genannten *Cluster Counting*. Ein Teilchen durchquert auf seinem Weg durch den Detektor eine ganze Reihe von Tubes. Das Teilchen hinterlässt auf seinem Weg in den Tubes eine Ionisationsspur, und erzeugt eventuell Übergangsstrahlungsphotonen. Durch eine untere Schwelle (*low threshold*), um ein Verrauschen des Signals zu verhindern, kann man so angesprochene Tubes erkennen und die Teilchenspur rekonstruieren. Für die Trennung von Elektron und Pion setzt man eine weitere, höhere Schwelle ein (*high threshold*). Es werden nur die Straws gezählt, die einen besonders hohen Eintrag liefern. In Abbildung 5.2 ist die Anzahl der high threshold Hits gegen ihre Häufigkeit pro Event für Elektronen und Pionen beim Durchqueren des ATLAS-TRT dargestellt. Diese Methode lässt sich noch verbessern indem man zusätzlich noch die Zeit, die das Signal über dem high threshold (*time over threshold*) liegt, berücksichtigt. Da-

¹Der TRT besteht aus drei Teilen, dem zentralen *barrel module* und zwei *end-caps*, die eine andere geometrische Anordnung haben.

bei kommt zum tragen, dass das Signal eines größeren Peaks auch längere Zeit über dem high threshold liegt. Mit dieser Methode wird bei dem ATLAS-TRT eine Pionenunterdrückung von ungefähr 80 für Impulse $p = 20 \text{ GeV}/c$ erreicht [Mit03].

Im ALICE TRD hat man keine feinkörnige Struktur. Man könnte sich trotzdem vorstellen, eine ähnliche Methode einzusetzen (z.B. zählen der Zeitbins, die über einer „hohen“ Schwelle liegen). Es hat sich aber gezeigt, dass solche Methoden für den TRD nicht geeignet sind [ATP99].

5.1.2 Q-Methoden

Die Q-Methoden basieren auf der Ladung, die ein geladenes Teilchen beim Durchqueren des Detektors erzeugt. Sie eignen sich für Übergangsstrahlungsdetektoren bei denen sich an den Radiator eine Driftkammer anschließt. Zu dem Signal von eventuell produzierten Übergangsstrahlungsphotonen kommt hier noch das zeitaufgelöste dE/dx -Signal, das das Teilchen in der Kammer erzeugt.

Die LQ-Methode

Bei dieser Methode werden die Verteilungen, die für die deponierte Energie (Abb. 5.3) gemessen wurden, als Wahrscheinlichkeitsverteilungen aufgefasst. Elektron und Pion besitzen eine gewisse Wahrscheinlichkeit $P(E)$ in einer Kammer eine bestimmte Energie E zu deponieren. Während der Test-Strahlzeit am CERN, wurden vier identische Kammern getestet, die teilweise leicht unterschiedliche Konfigurationen² aufgewiesen haben. Nun gibt es verschiedene Herangehensweisen, um die erreichte Pioneneffizienz für sechs Kammern zu bestimmen.

Man kann jede Kammer einzeln „simulieren“ und dann die Pionenunterdrückung für sechs Module separat für jede Kammer bestimmen. Diese Methode wird im Weiteren als „*Simulationmethode*“ bezeichnet.

Bei dieser Methode werden zunächst, jeweils für Elektronen und Pionen getrennt, sechs Zufallszahlen E_n^e und E_n^π ($n = 1, \dots, 6$) entsprechend den zugehörigen Energieverteilungen erzeugt (vgl. [Win02]). $P^n(E_n | e)$ ist die Wahrscheinlichkeit, dass ein Elektron in Modul n ein Energiesignal der Größe E_n erzeugt. Die Gesamtwahrscheinlichkeit, dass ein Elektron einen bestimmten Satz von Messwerten produziert, ist:

$$P_e(E) = \prod_{n=1}^6 P^n(E_n | e) = \prod_{n=1}^6 P(E_n | e). \quad (5.1)$$

²Dazu zählen z.B. Drift- und Anodenspannung.

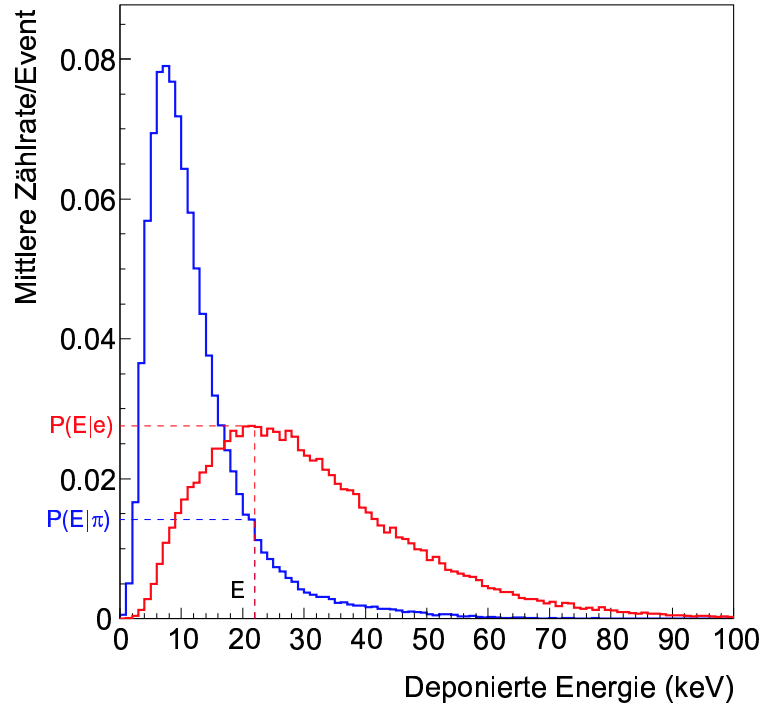


Abbildung 5.3: Relative Häufigkeit der deponierten Ladung in der Driftkammer. Elektronen (rot) deponieren im Mittel mehr Energie als Pionen (blau).

Hier gilt ($P = P^n$), da ja von einem Modul auf sechs Kammern geschlossen wird. Analog lässt sich die Gesamtwahrscheinlichkeit für Pionen formulieren:

$$P_\pi(E) = \prod_{n=1}^6 P^n(E_n | \pi) = \prod_{n=1}^6 P(E_n | \pi). \quad (5.2)$$

Mit den Werten für die Gesamtwahrscheinlichkeiten lässt sich eine Aussage über die Identifizierungswahrscheinlichkeit für Elektronen treffen. Anhand der für Elektronen erzeugten Zufallszahlen (E_n^e) ist dann die relative Wahrscheinlichkeit L_e (*Likelihood*) für ein Elektron:

$$L_e = \frac{P_e(E^e)}{P_e(E^e) + P_\pi(E^e)}, \quad \text{wobei } 0 \leq L_e \leq 1. \quad (5.3)$$

Für die Wahrscheinlichkeit, dass ein Pion fälschlicherweise als Elektron identifiziert wird, gilt mit den E_n^π :

$$L_e = 1 - L_\pi = \frac{P_e(E^\pi)}{P_e(E^\pi) + P_\pi(E^\pi)}, \quad 0 \leq L_e \leq 1. \quad (5.4)$$

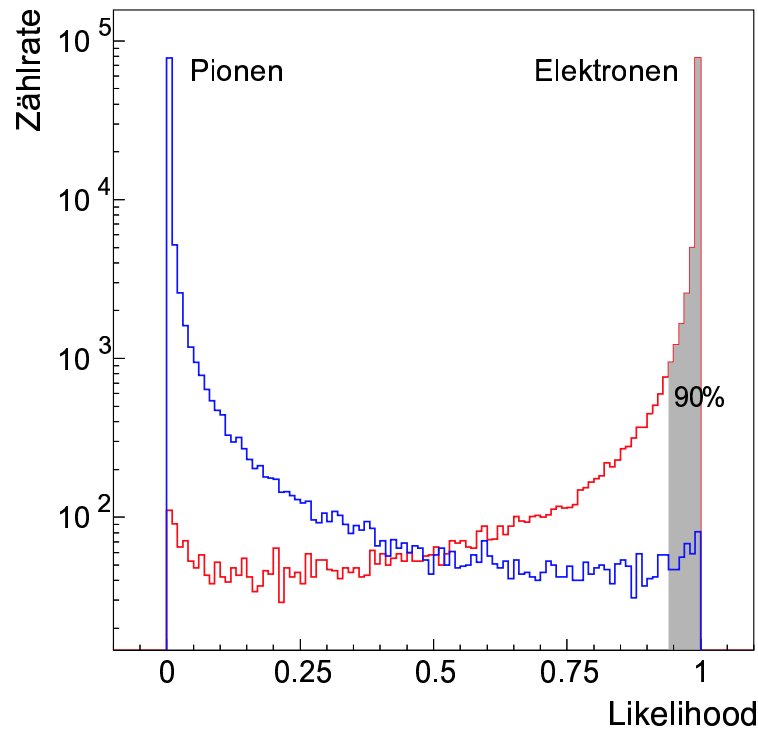


Abbildung 5.4: Likelihoodverteilung von Elektronen (rot) und Pionen (blau). Um die Pioneneffizienz zu bestimmen werden die Elektroneneinträge von rechts nach links integriert, bis 90% der Elektronen (grau unterlegt) erfasst sind. Der Anteil der Pionen, die noch innerhalb der Integrationsgrenzen liegen, ist die Pioneneffizienz.

Um die Pionenunterdrückung der sechs Module zu bestimmen, wird eine große³ Anzahl Elektronenereignisse und Pionenereignisse simuliert. Die so erhaltenen Likelihood-Werte werden in ein Histogramm gefüllt (siehe Abb. 5.4). Die Pionenunterdrückung ist der Kehrwert der Pioneneffizienz ε_π . Die Pioneneffizienz ist die relative Anzahl der missidentifizierten Pionen bei fester Elektroneneffizienz ε_e . Allgemein wird die Elektroneneffizienz auf 90% festgelegt. Das heißt, 90% der Elektronen werden richtig identifiziert. In Abbildung 5.4 wird die Vorgehensweise verdeutlicht. Es sind die Likelihoodverteilungen für Elektronen und Pionen dargestellt. Die Likelihoodverteilung der Elektronen wird integriert. Die untere Integrationsgrenze wird

³Bei den Analysen, die in dieser Arbeit gemacht worden sind, wurden jeweils 100 000 oder 1 000 000 Ereignisse simuliert.

so gewählt, dass 90% der Elektronen, und zwar die mit den größten Gesamtwahrscheinlichkeiten, berücksichtigt werden. Es gilt:

$$0,9 = \frac{\int_{L(\varepsilon_e=90\%)}^1 f^e(L_e)}{\int_0^1 f^e(L_e)}, \quad (5.5)$$

wobei f^e die Verteilung der Likelihood-Werte für Elektronen ist. Die Pioneneffizienz ist nun der Anteil der Pionen, der innerhalb der Integrationsgrenzen für Elektronen liegt:

$$\varepsilon_\pi = \frac{\int_{L(\varepsilon_e=90\%)}^1 f^\pi(L_e)}{\int_0^1 f^\pi(L_e)}. \quad (5.6)$$

Die Simulations-Methode ist geeignet, um verschiedene Kammerkonfigurationen miteinander zu vergleichen. Eine Aussage über die tatsächliche Pionenunterdrückung ist mit ihr nur schwer möglich.

Tatsächlich gibt es Korrelationen zwischen den Kammern, die die Pioneneffizienz positiv oder negativ beeinflussen. Will man diese Korrelationen in die Analysen einbeziehen, benutzt man die *Extrapolationsmethode*.

Die Vorgehensweise ähnelt der Simulationsmethode. Auch hier werden Wahrscheinlichkeitsverteilungen der deponierten Ladung für jede Kammer über in einem Messdurchgang gemessene Werte erstellt. Man kann jetzt für jede Kammer getrennt Zufallszahlen erzeugen oder einen zweiten Run⁴ benutzen, und für jedes Ereignis die zugehörige Wahrscheinlichkeit $P^n(E_n^e | e)$ bzw. $P^n(E_n^\pi | \pi)$ ablesen. In den Analysen dieser Arbeit wurden z.B. Runs mit Scaledown-Faktor benutzt, um die Wahrscheinlichkeitsverteilungen zu erzeugen. Die Daten zur Likelihood-Bestimmung stammen aus Runs ohne Scaledown-Faktor. Analog zu den Gleichungen 5.1 und 5.2 lassen sich Gesamtwahrscheinlichkeiten

$$P_{e,m}(E^e) = \prod_{n=1}^m P^n(E_n^e | e) \quad (5.7)$$

und

$$P_{\pi,m}(E^\pi) = \prod_{n=1}^m P^n(E_n^\pi | \pi) \quad (5.8)$$

⁴Dieser Run muss natürlich unter identischen oder zumindest ähnlichen Bedingungen (gleicher Impuls, gleicher Radiator, etc.) stattgefunden haben.

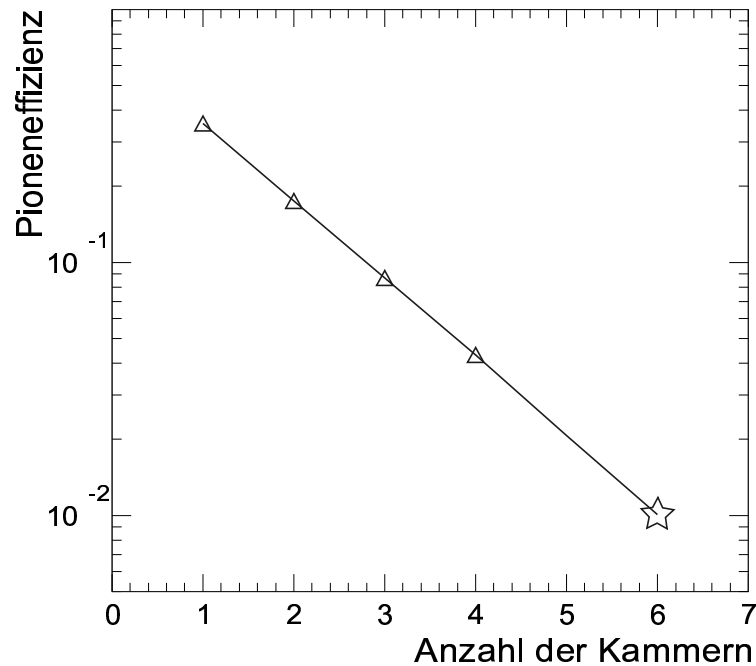


Abbildung 5.5: Bei der Extrapolationsmethode bestimmt man die Pioneneffizienz für eine aufsteigende Anzahl von Kammern (Dreiecke) mit gemessenen Werten. In einem Diagramm können sie durch eine Exponentialfunktion angefitet werden, um so die Pioneneffizienz für sechs Kammern (Stern) zu bestimmen [And02].

formulieren, wobei ($m=1, \dots, 4$) hier die Zahl der betrachteten Kammern ist. Die Gesamtwahrscheinlichkeiten werden zunächst nur für die erste Kammer berechnet, dann für die ersten beiden, bis man schließlich die Gesamtwahrscheinlichkeiten für alle vier TRD-Module erhält. Mit diesen Werten kann man wie in den Gleichungen (5.3) und (5.4) Likelihoodwerte $L_{e,m}$ berechnen. Die Pioneneffizienzen $\varepsilon_{\pi,m}$ für die unterschiedliche Anzahl der Kammern ergeben sich aus (5.5) und (5.6). Die errechneten Pioneneffizienzen lassen sich gegen die Anzahl der Kammern auftragen. Daraus kann nun die Pioneneffizienz für sechs Kammern extrapoliert werden (vgl. Abb. 5.5).

Die LQX-Methode

Die im Radiator erzeugten TR-Photonen werden kurz nach ihrem Eintritt in die Kammer absorbiert. Die durch TR produzierten Ladungs-Cluster sind im Verhältnis zu den durch Energieverlust erzeugten Clustern besonders groß. Die LQX-Methode versucht diesen Sachverhalt auszunutzen, indem die Position des größten Clusters im Driftbereich zusätzlich zur deponierten Ladung zur Likelihood-Bestimmung benutzt

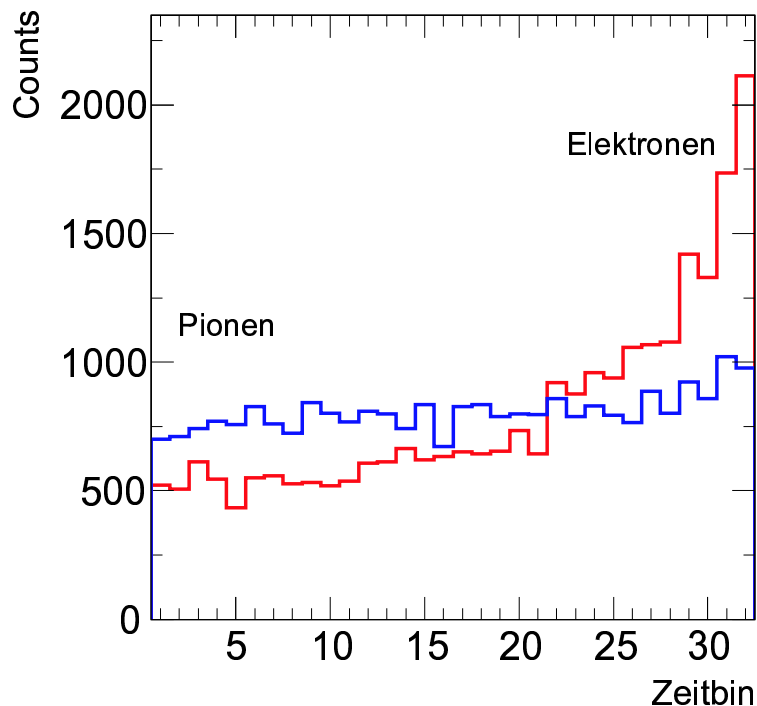


Abbildung 5.6: Verteilung der größten Cluster im Driftbereich. Bei Elektronen (rot) ist die Wahrscheinlichkeit den größten Cluster bei hohen Driftzeiten zu erzeugen deutlich größer als bei kleinen. Die Verteilung für die Pionen (blau) ist flach [And01].

wird.

In Abbildung 5.6 ist die mittlere Verteilung des größten Clusters⁵ im Driftbereich für Elektronen und Pionen dargestellt. Während die Verteilung für die Pionen flach ist, erkennt man für Elektronen einen Anstieg zu höheren Driftzeiten. Wie auch in der LQ-Methode kann man sowohl die Simulations-, als auch die Extrapolationsmethode anwenden. Der einzige Unterschied besteht darin, dass man statt über sechs nun insgesamt über zwölf Messwerte verfügt. Es gilt für die Simulationsmethode:

$$P_e(E) = \prod_{n=1}^6 P^n(E_n | e) \cdot \prod_{n=1}^6 P^n(X_n | e) \quad (5.9)$$

und

$$P_\pi(E) = \prod_{n=1}^6 P^n(E_n | \pi) \cdot \prod_{n=1}^6 P^n(X_n | \pi), \quad (5.10)$$

wobei X die Position des Clusters wiedergibt.

Die weitere Vorgehensweise ist identisch zu der in der LQ-Methode verwendeten.

⁵Der Einfachheit halber wird der höchste Energieeintrag benutzt. In der Abbildung sind jeweils die Zeitbins mit der höchsten deponierten Energie eingetragen.

Mit der LQX-Methode erzielt man eine Erhöhung der Pionenunterdrückung von 15 bis 30% gegenüber der LQ-Methode.

5.2 Auswertung der Datensätze

Ziel dieser Arbeit ist es, verschiedene Methoden zur Bestimmung der Pioneneffizienz zu analysieren. Bei der Untersuchung der klassischen Methoden sind die Analysen ausschließlich mit der LQ- und der LQX-Methode durchgeführt worden. Es wurde sowohl die Simulations- als auch die Extrapolationsmethode angewandt.

Die Daten stammen aus der Strahlzeit 2002 am CERN (siehe Kap. 4.5). Es wurden Daten mit unterschiedlichen Impulsen und verschiedenen Radiatoren (INV6 und INV8)⁶ ausgewertet. Zur Verarbeitung und Auswertung der Rohdaten wurden Macros benutzt, die mit dem Analyse-Program *ROOT* verwendet wurden. *ROOT* basiert auf der Programmiersprache *C++* und ist eine objekt-orientierte Umgebung zur Datenanalyse mit graphischer Oberfläche [ROO01].

5.2.1 Datenauslese

Die Prototypen, die in der Strahlzeit benutzt worden sind, verfügten über acht parallel zueinander verlaufende Auslese pads. Die FADCs hatten eine Auflösung von 50 ns pro Zeitkanal. Da die deponierte Ladung bei einem Event nicht alle Pads erreicht, musste evaluiert werden, wie viele Pads in die Analyse mit einbezogen werden sollten. Dazu wurden bei verschiedenen Runs jeweils ein, drei⁷ bzw. alle acht Pads ausgelesen.

In Abbildung 5.7 sind die verschiedenen Verteilungen für einen Run dargestellt. Links sind die verschiedenen Distributionen für den größten Cluster im Driftbereich und rechts die Energieverteilungen gezeigt. Die Anzahl der berücksichtigten Pads nimmt von oben nach unten zu. Man erkennt bei den Energieverteilungen keinen bedeutsamen Unterschied, nur die Anzahl der Treffer mit sehr hohen Energien nimmt zu. Die Positionsverteilungen der höchsten Einträge ist für drei und für acht ausgelesene Pads nahezu identisch. Die Verteilung für die Pionen verläuft flach, die der Elektronen steigt an. Liest man nur ein Pad aus, so erkennt man, dass die Verteilung für Pionen kontinuierlich abnimmt, während die der Elektronen flach verläuft. Dies

⁶Dabei handelt es sich um Sandwichradiatoren mit 6 mm (INV6) und 8 mm (INV8) dicken *RohacellTM*-Platten mit 8 bzw. 7 Fasermatten.

⁷Für die Analyse mit einem Pad wurde das Pad ausgewählt, das den höchsten integrierten Energieeintrag aufgewiesen hat. Für die Analysen mit drei Pads wurde das Pad mit dem höchsten Energieeintrag und zusätzlich seine beiden Nachbarpads miteinbezogen.

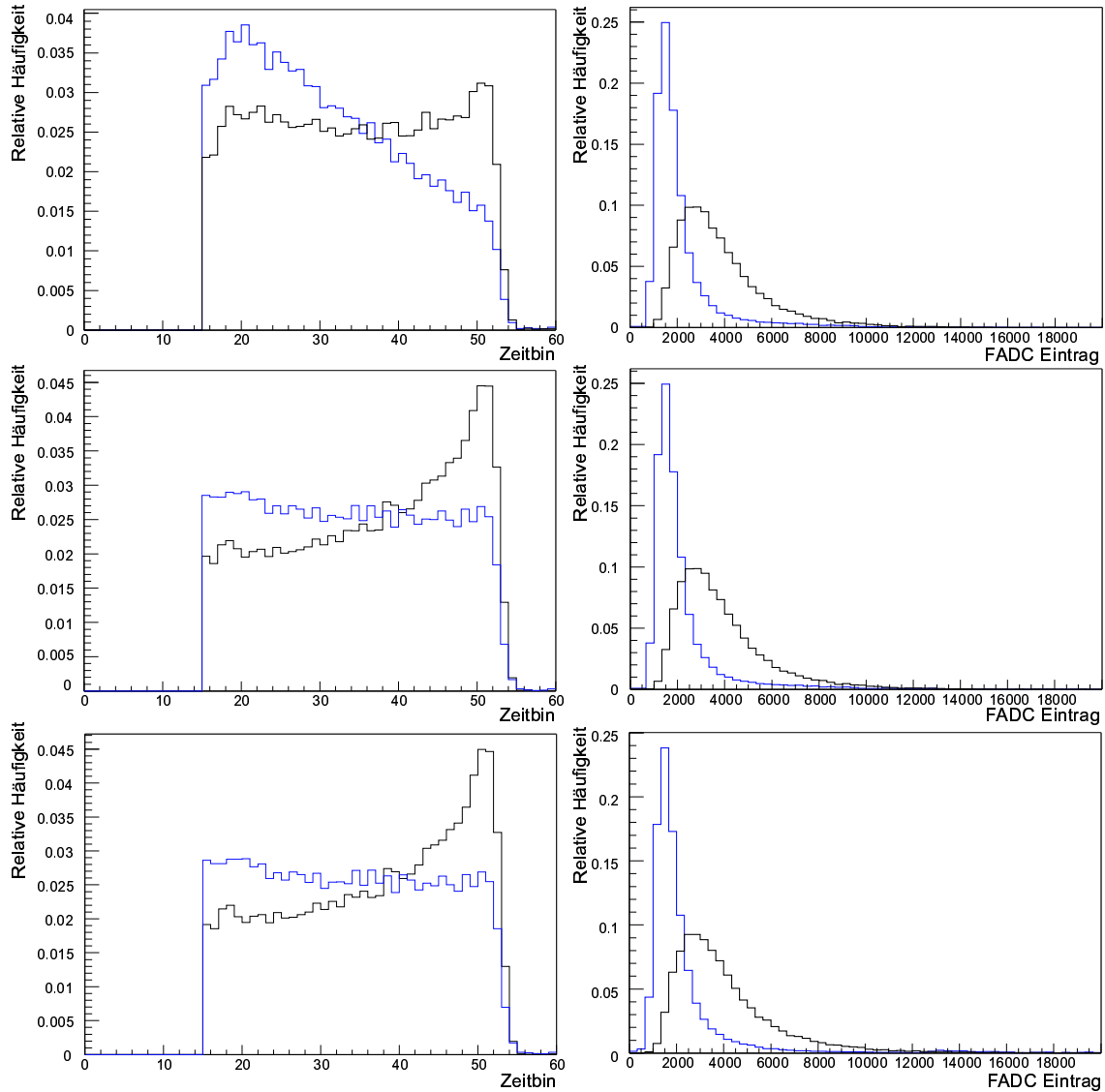


Abbildung 5.7: Datenauslese bei einer unterschiedlichen Anzahl von Pads. Links sind die Verteilungen der größten Cluster im Driftbereich, rechts die der deponierten Ladung. Oben wurde das Pad mit der höchsten deponierten Energie, integriert über alle Zeitbins, in der Mitte zusätzlich seine beiden Nachbarn und unten alle acht Pads ausgelesen. Dargestellt sind die Ergebnisse eines Runs mit einem Impuls von $3 \text{ GeV}/c$. Bei dem Radiator handelt es sich um den INV6 Radiator. Dieser Run wurde mit Pionen-Scaledown aufgenommen. Elektronen- sind schwarz, Pionenverteilungen blau dargestellt.

liegt daran, dass die Kammern in einem Winkel von 15° zur Strahlachse aufgestellt gewesen waren. Betrachtet man also nicht drei nebeneinanderliegende Pads, kann nicht die gesamte Spur erfasst werden. Besonders bei Pionen sind Peaks mit kurzen Driftzeiten bevorzugt, da ein großer Teil der Gesamtenergie in der Verstärkungszone deponiert wird und dann nur dieses Pad ausgelesen wird. Cluster die bei hohen Driftzeiten entstehen, werden wegen des schrägen Einfallswinkels nicht berücksichtigt. Es gibt auch Runs in denen das Verhalten der Kammern bei unterschiedlicher Ausrichtung der Kammern zur Strahlachse gemessen worden ist. Bei einem Winkel von 0° ergibt sich auch für ein Pad die typische Positionsverteilung, wie sie in Abbildung 5.6 zu sehen ist.

Bei der Auswertung der Daten wurden jeweils drei Pads zur Bestimmung des höchsten Peaks und zur Bestimmung der deponierten Energie verwendet.

5.2.2 Simulation der einzelnen Kammern

Mit der Simulationsmethode sind Runs mit INV6 und INV8 Radiatoren und Impulsen von $p = 1, 1,5, 2, 3, 4, 5$ und $6 \text{ GeV}/c$ untersucht worden. Alle Runs sind mit Pionen-Scaledown aufgenommen worden. Es wurden pro Kammer je 100 000 Elektronen- und Pionenereignisse simuliert.

In Abbildung 5.8 sind verschiedene Plots aus der Auswertung von Kammer 2 bei einem Impuls von $2 \text{ GeV}/c$ mit einem INV8 Radiator zu sehen. Oben (Abb. 5.8a und 5.8b) sind die Verteilungen für Position und deponierte Ladung, in der Mitte (Abb. 5.8c und 5.8d) die zugehörigen Likelihoodverteilungen abgebildet. Die Likelihoodverteilungen für die Position überschneiden sich stark. Bei der Energieverteilung hingegen sieht man bei den Pionen einen sehr hohen Eintrag bei kleiner Likelihood (also geringer Elektronen-Wahrscheinlichkeit), bei der Elektronenverteilung aber einen sehr hohen Eintrag bei $L_e = 1$ (große Elektronen-Wahrscheinlichkeit). Diese Likelihood-Verteilung wurde zur Bestimmung der Pioneneffizienz in der LQ-Methode eingesetzt. Links unten (Abb. 5.8e) ist die Likelihood-Verteilung für eine Kombination aus Position und deponierter Ladung abgebildet. Sie verläuft ähnlich wie die Likelihood-Verteilung der Energie. Auch hier hat man ausgeprägte Maxima für Pionen bei Null und für die Elektronen bei eins. Diese Verteilungen sind in der LQX-Methode benutzt worden. Abbildung 5.8f zeigt die Abhängigkeit der Pioneneffizienz bei unterschiedlichen Elektroneneffizienzen.

Die Entwicklung der Pioneneffizienz bei 90% Elektroneneffizienz für LQ- und LQX-Methode bei unterschiedlichen Impulsen ist in Abbildung 5.9 dargestellt. Abgebildet sind die Effizienzen für die einzelnen Kammern. Man erkennt eine Verschlechterung der Pionenunterdrückung bei höheren Impulsen. Die Pionen befinden sich bei diesen Impulsen schon ein wenig auf der aufsteigenden Flanke des

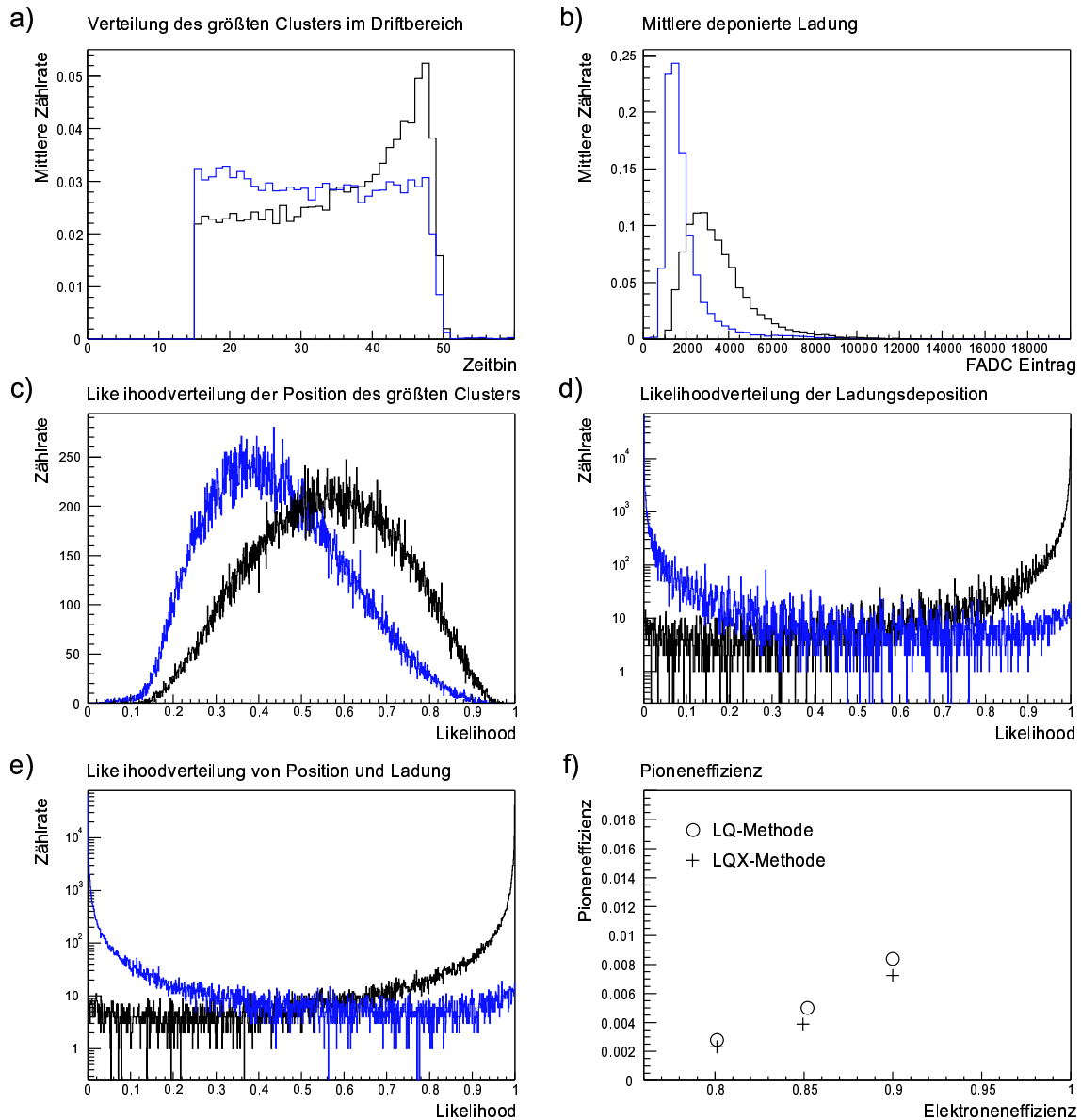


Abbildung 5.8: Typische Verteilungen wie sie in der LQX-Methode verwendet werden. Dargestellt sind Verteilungen aus einem Run mit einem Impuls von $2 \text{ GeV}/c$. **a)**: Die Position des größten Clusters, darunter **(c)** die zugehörige Likelihoodverteilung für sechs Kammern. **b)**: Die Verteilung der deponierten Energie mit Likelihoodverteilung **(d)**. **e)**: Die Kombination aus beiden Verteilungen und **f)** die Pioneneffizienz für verschiedene Elektroneneffizienzen. Elektronen- sind schwarz, Pionenverteilungen in blau dargestellt.

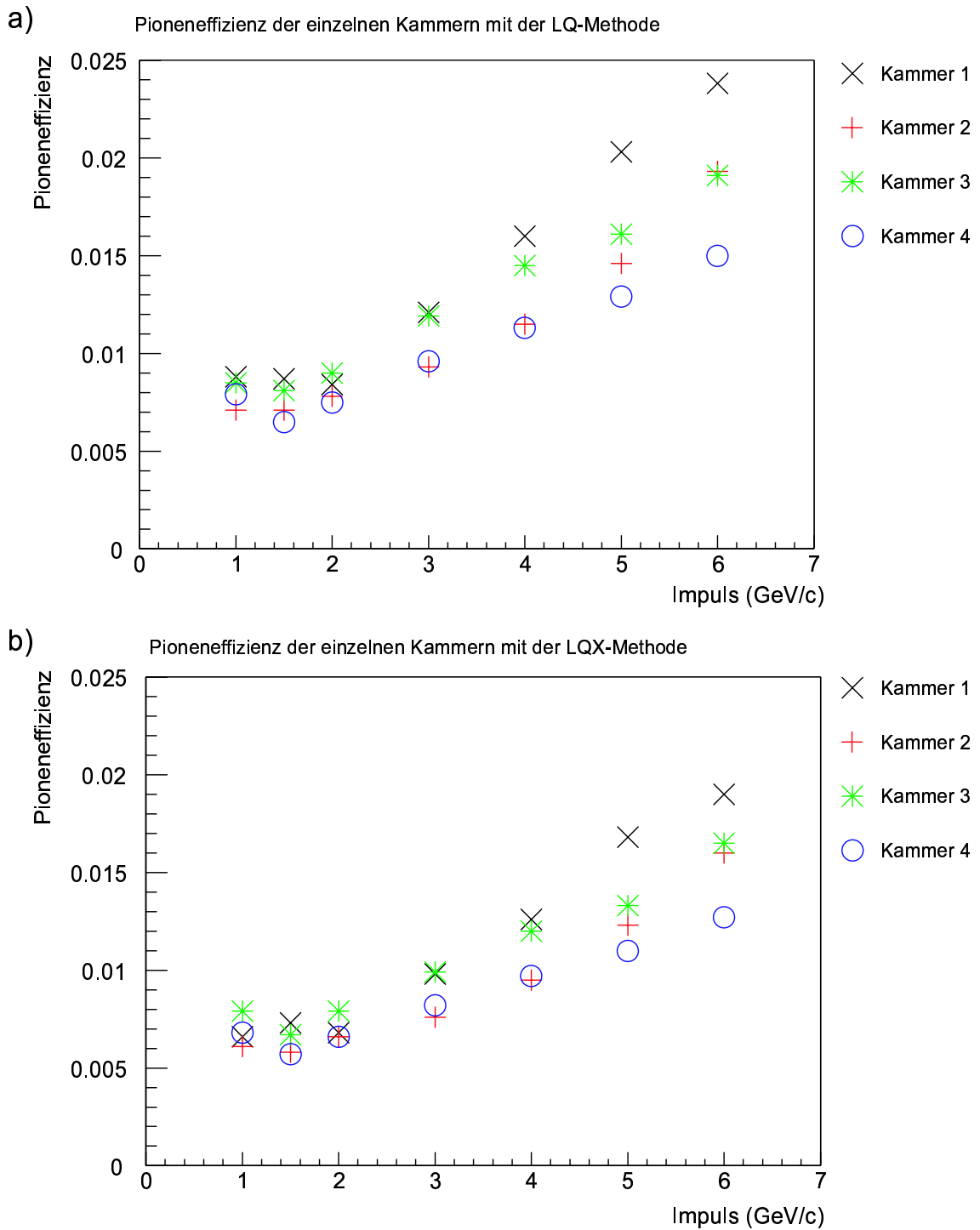


Abbildung 5.9: Pioneneffizienz der einzelnen Kammern bei verschiedenen Impulsen. Dargestellt sind die simulierten Pioneneffizienzen (bei 90% Elektroneneffizienz) aus Runs mit einem INV8 Radiator. Man erkennt die Impulsabhängigkeit der Pioneneffizienz. Ebenso sieht man, dass die Pionenunterdrückung zu hinteren Kammern hin ansteigt.

Energieverlustes (Abb. 4.2), sind also keine minimalionisierende Teilchen mehr. Das Verhältnis zwischen den deponierten Ladungen von Elektronen und Pionen wird kleiner. Außerdem erkennt man die Tendenz, dass hintere Kammern eine bessere e/π -Separation erreichen als die Kammern, die zuerst von den Teilchen durchquert werden. Eine Erklärung wäre, dass Elektronen auf ihrem Weg durch die vorderen Kammern zusätzlich noch Strahlung produzieren können (z.B. Bremsstrahlung), die dann in den hinteren Kammern zu einem erhöhten Signal führt.

5.2.3 Die Extrapolationsmethode

Bei der Extrapolationsmethode wurden Runs mit Pionen-Scaledown benutzt, um die Wahrscheinlichkeitsverteilungen zu erstellen, und Runs ohne Scaledown, um die Pionenunterdrückung zu bestimmen. Der Impulsbereich reichte von $2 \text{ GeV}/c$ bis $6 \text{ GeV}/c$. Bei dem Radiator handelte es sich um einen INV6 Sandwichradiator. Es wurde nur die LQX-Methode angewendet.

In Abbildung 5.10 sind typische Histogramme zu sehen, wie sie in der Extrapolationsmethode verwendet oder erstellt werden. Oben (Abb. 5.10a und 5.10b) sind die Verteilungen für deponierte Energie und Position des größten Clusters für einen $2 \text{ GeV}/c$ Run mit Pionen-Scaledown dargestellt, darunter (Abb. 5.10c und 5.10d) für den entsprechenden Run ohne Scaledown. Man sieht, dass die Statistik für Elektronen bei den Runs ohne Scaledown sehr niedrig ist. Besonders die Positionsverteilungen sind sehr „zerfranst“. Dieser Run weist ein e/π -Verhältnis von etwa $1 : 2$ auf. Bei Runs mit höherem Impuls sinkt dieses Verhältnis auf ca. $1 : 20$ (vgl. Tab: 7.1). In der Abbildung 5.10e sind die Likelihoodverteilungen für vier nacheinandergeschaltete Kammern dargestellt. Aus den Verteilungen für unterschiedliche Kammerzahl werden die Effizienzen bestimmt. Abbildung 5.10f zeigt die Extrapolation auf den Wert für sechs Kammern. Trotz der geringen Statistik für Elektronen wird deutlich, dass alle Punkte in dem logarithmischen Plot auf einer Gerade liegen.

Bei der Bestimmung der Pioneneffizienz mit neuronalen Netzen wird ein ähnliches Verfahren verwendet. Die Ergebnisse der Extrapolationsmethode werden in Kapitel 7 mit den Ergebnissen, die durch die Verwendung neuronaler Netze erzielt werden, verglichen (Abb. 7.5).

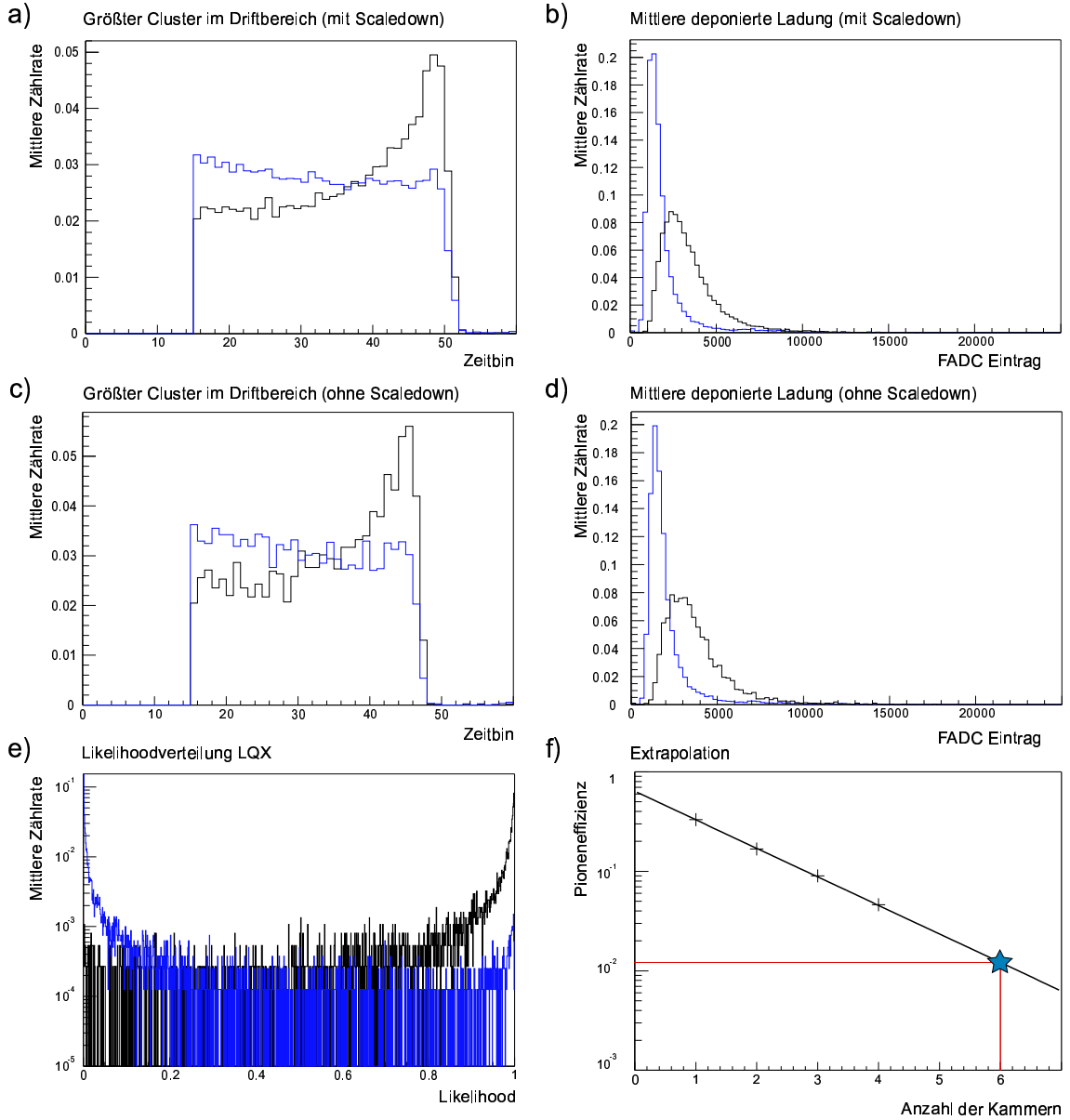


Abbildung 5.10: Die Extrapolationsmethode bei einem Impuls von $p = 2 \text{ GeV}/c$. **a)** und **b)**: Die Verteilungen für den größten Cluster im Driftbereich und der deponierten Ladung, jeweils mit Scaledown-Faktor. **c)** und **d)**: Die zugehörigen Verteilungen ohne Pionen-Scaledown. **e)**: Die Likelihoodverteilungen aus der Effizienzbestimmung für vier Kammern. **f)**: Aus den Effizienzen für eine bis vier Kammern wird die Pioneneffizienz für sechs Kammern extrapoliert. Die Elektronenverteilungen sind schwarz, die Pionenverteilungen blau dargestellt.

6. Künstliche neuronale Netzwerke

Die LQX-Methode liefert von den klassischen Methoden die besten Ergebnisse und liegt für die Pionenunterdrückung im Bereich des angestrebten TRD-Designzieles. Dennoch stellte sich die Frage, ob dieses Ergebnis nicht noch verbessert werden kann.

Die klassischen Methoden, die hier besprochen worden sind, nutzen nicht die gesamte Information aus, die der ALICE-TRD aufnehmen kann. Der TRD liefert ja Auskunft über die deponierte Ladung zeitaufgelöst über die gesamte Driftstrecke. Bei der LQ-Methode werden alle Zeitbins zusammengefasst und nur die über alle Zeitbins integrierte Energie ist für diese Methode relevant. Die LQX-Methode verwendet mehr Information, aber auch hier geht ein großer Teil verloren. Die N-Methoden verwenden zwar die Zeitauflösung der Kammern, aber die Information über die deponierte Energie wird nicht genutzt. Ziel ist einen Algorithmus zu finden, der zum einen die deponierte Energie, zum anderen aber auch die Position und Größe der erzeugten Cluster berücksichtigt. So ein Algorithmus ist nicht leicht zu finden, da es Korrelationen¹ gibt, die hierbei berücksichtigt werden müssen. Einen Ansatz diese Probleme zu umgehen, und trotzdem die maximal mögliche Information zu nutzen, liefern neuronale Netzwerke.

6.1 Einführung

Moderne Rechner sind heute in der Lage große Datenmengen in sehr kurzer Zeit zu verarbeiten und komplizierte Berechnungen durchzuführen. Sie scheitern aber an einigen Aufgaben, die jeder Mensch ohne Probleme bewältigen kann. Dazu zählen das Erkennen von Gesichtern, Spracherkennung, das Lesen von Handschriften und Verarbeitung von Texten in denen Syntaxfehler auftauchen². Die entscheidende Eigenschaft, die uns das ermöglicht, ist die Adaptionsfähigkeit. Durch Lernen sind wir befähigt bestimmte Dinge schneller und effizienter zu erledigen. Neuronale Netzwerke stellen den Versuch dar, die Leistungsfähigkeit von Computern mit der Anpassungsfähigkeit und Lernfähigkeit biologischer Systeme zu verknüpfen. Dabei bemüht man sich ein System aus Nervenzellen im Computer künstlich zu simulieren.

¹Ein Beispiel wäre Überlagerung oder Überschneidung von zwei Clustern.

²Kein C-Compiler verzeiht ein vergessenes Semikolon.

6.1.1 Allgemeine Merkmale

Heutzutage werden künstliche neuronale Netze in verschiedenen Bereichen, etwa der Kategorisierung, Funktionsapproximation, Optimierung, in der Robotik (z.B. zur autonomen Bewegung in unbekanntem Terrain) und zur Mustererkennung eingesetzt. Die Vorgehensweise bleibt dabei jedesmal gleich. Man kann ein neuronales Netz als Abbildungsvorschrift verstehen, die eine als Vektor kodierte Eingangsgröße in eine Ausgabe, dargestellt durch einen anderen Vektor, transformiert. Dies ist in Abbildung 6.1 schematisch dargestellt. Dabei können Ein- und Ausgabevektor durchaus unterschiedliche Dimensionen haben.

Neuronale Netze nutzen einfache Prozessoren (die *Neuronen*), die ihrerseits nur einfache Kalkulationen³ durchführen können. Die Verbindungen zwischen den Neuronen werden mit einem Gewicht bewertet. Durch Änderung der Gewichte kann eine Anpassung an die Zielvektoren erreicht werden. Zu den wichtigsten Eigenschaften neuronaler Netze zählen [Sch97]:

- **Lernfähigkeit:** Neuronale Netze können „lernen“ einen Ausgabevektor, der einem Eingabevektor zugeordnet ist, zu reproduzieren.
- **Fehlertoleranz:** Sie besitzen eine gewisse Fehlertoleranz, sollten Teile des Netzwerkes ausfallen.
- **Geschwindigkeit:** Das Training kann durchaus zeitaufwendig sein; ist ein Netz aber einmal trainiert, können Muster sehr schnell verarbeitet werden.
- **Robustheit:** Neuronale Netze können auch verrauschte Daten verarbeiten.
- **Generalisierungsfähigkeit:** Sie sind fähig auch unbekannte Muster richtig zuzuordnen.

Um ein neuronales Netz erfolgreich einzusetzen, benötigt man folgende Komponenten:

- ein geeignetes Netzwerk
- einen geeigneten Lernalgorithmus
- Trainingsdaten
- Validierungsdaten zur Überprüfung der Verallgemeinerungsfähigkeit.

³Z.B. Summation der gewichteten Verbindungen.

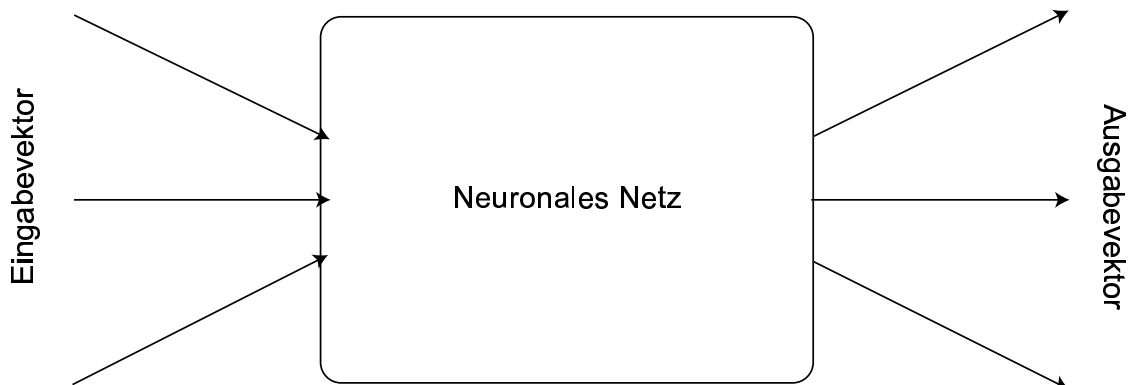


Abbildung 6.1: Neuronale Netze als Abbildungsvorschrift für eine Transformation eines Eingabevektors in einen Ausgabevektor.

6.1.2 Historisches

Dieser Abschnitt soll einen (sehr) kurzen historischen Überblick über die Entwicklung der neuronalen Netze geben.

Den ersten Entwurf eines künstlichen Neurons wurde 1943 von McCulloch und Pitts beschrieben. Dieses so gennante *McCulloch-Pitts-Neuron* wurde zum Ausgangspunkt verschiedener Neuronenmodelle, die auch heute noch gebräuchlich sind. Von Hebb wurde 1949 eine lokale Trainingsstrategie für Neuronen eingeführt, die Hebb'sche-Lernregel, die auf der Annahme beruht, dass Gewichtungen zwischen zwei Neuronen, sollten sie beide aktiv sein, verstärkt werden sollten. Auf diesen Grundlagen basieren die heutigen Lern-Algorithmen. 1962 wurden von Rosenblatt *Feedforward-Netzwerke* (auch *Perzeptronen* genannt) entwickelt. Auf diesen Netzwerktypus wird in Kapitel 6.3 näher eingegangen. Die letzte wichtige Entwicklung, die in dieser Arbeit benutzt wird, ist der so genannte *Backpropagation-Algorithmus*. Er wurde Mitte der achtziger Jahre von mehreren Personen unabhängig entdeckt. Die Arbeit, die hier am meisten Beachtung fand, ist die von Rumelhart, Hinton und Williams. Der Backpropagation-Algorithmus wird in Kapitel 6.4 besprochen. Weitergehende Informationen zur Historie neuronaler Netze kann man zum Beispiel in [Sch97] und [Per92] finden.

6.2 Vom natürlichen zum künstlichen Neuron

Das menschliche Gehirn besteht aus etwa $2,5 \cdot 10^{10}$ Nervenzellen, den so genannten Neuronen. Ein Neuron besteht aus einem Zellkörper mit Kern, den Dendriten, einem Axon und Synapsen (siehe Abb. 6.2). Die Dendriten sind die Eingangsfasern,

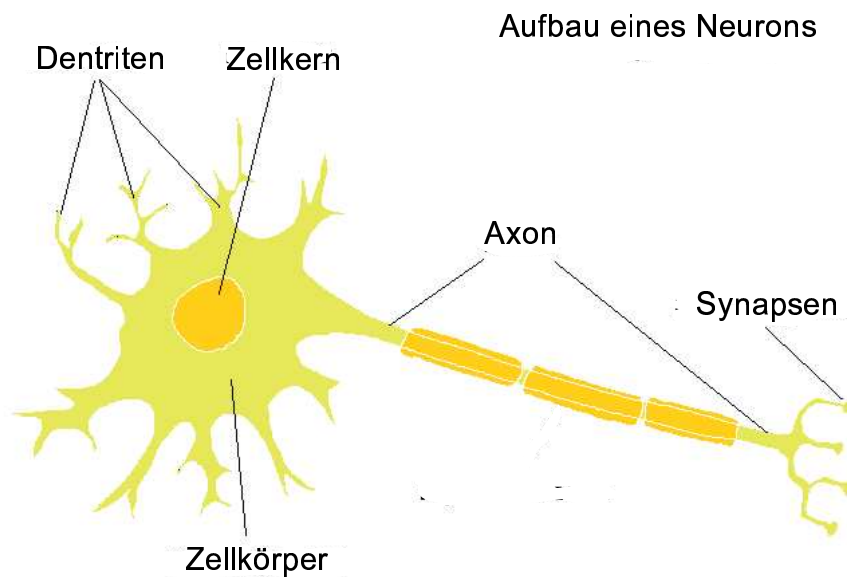


Abbildung 6.2: Schematische Darstellung eines typischen Neurons.

über die ein Signal aufgenommen wird. Das Neuron summiert diese Signale auf und leitet sie, wird ein bestimmter Schwellenwert überschritten, über das Axon weiter. Die Synapsen wandeln das elektrische Signal in ein chemisches um. Es wird über *Neurotransmitter* an andere Zellen weitergegeben. Die Synapsen haben aber nicht nur die Aufgabe das ankommende Signal weiterzuleiten, sie verstärken es auch (exhibitorische Synapse) oder schwächen es ab (inhibitorische Synapse).

Ein künstliches Neuron ist funktional ganz ähnlich aufgebaut (Abb. 6.3). Die Eingangssignale I_i der n Neuronen, mit denen das Neuron j verbunden ist, werden mit den Gewichten w_{ij} multipliziert, also verstärkt oder abgeschwächt, und hier nach im Neuron j mittels einer *Propagierungsfunktion* weiter verarbeitet. In der Regel werden die Eingaben entsprechend ihrer Gewichtung aufsummiert⁴. Mittels einer Aktivierungsfunktion a wird nun die Ausgabe berechnet. Typische Aktivierungsfunktionen sind die *Identität*, *Schwellenfunktionen* oder *sigmoide Funktionen*. Häufig wird zusätzlich noch ein neuronspezifischer Schwellenwert, das so genannte *Bias*, benutzt. Dies bewirkt eine Verschiebung der Aktivierungsfunktion auf der x-Achse. In der vorliegenden Arbeit ist eine sigmoide Aktivierungsfunktion mit Bias verwendet worden. Es ergibt sich als Ausgabe eines Neurons:

$$O_j(t+1) = \frac{1}{1 + e^{-(\sum_i w_{ij} O_i(t) - \theta_j)}} \quad (6.1)$$

⁴Es sind aber auch andere Propagierungsfunktionen gebräuchlich. Manchmal wird nur der Maximalwert oder der Minimalwert, bisweilen auch das Produkt der gewichteten Eingabe, verwendet.

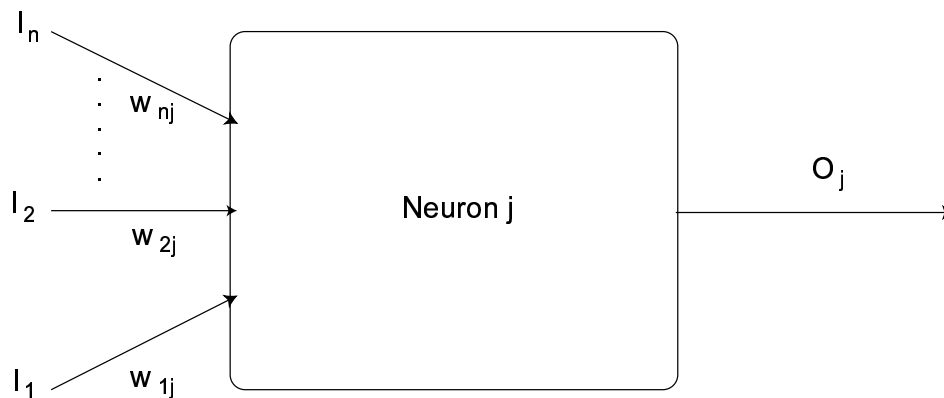


Abbildung 6.3: Das mathematische Modell eines Neurons.

Dabei ist $O_j(t+1)$ die Ausgabe des Neurons j zum Zeitpunkt⁵ $t+1$, w_{ij} die Gewichtung der Verbindung zwischen den Neuronen i und j , $O_i(t)$ die Ausgabefunktion des Neurons i zum Zeitpunkt t und θ_j der Schwellenwert von j .

6.3 Feedforward Netzwerke

Es gibt zahlreiche verschiedene Netzwerktypen. Im Allgemeinen unterscheidet man zwischen *Feedforward-Netzen* und *Feedback-Netzen*.

Neuronale Netzwerke besitzen eine Eingabeschicht (*Input Layer*), eine Ausgabeschicht (*Output Layer*) und eine beliebigen Anzahl von verdeckten Schichten (*Hidden Layer*). Die Anzahl der verdeckten Schichten kann auch Null sein (*Single-Layer-Perzeptron*), sie richtet sich im Allgemeinen nach der Komplexität des zu lösenden Problems.

Feedforward-Netze zeichnen sich dadurch aus, dass man ihre verschiedenen Schichten streng trennen kann und die Neuronen jeder Schicht nur über Verbindungen zu den Neuronen einer übergeordneten Schicht verfügen. Ein Eingabemuster (auch *Pattern* genannt) wird also schrittweise von der Eingabe zur Ausgabe verarbeitet. Man unterscheidet zusätzlich noch zwischen Feedforward-Netzen erster und zweiter Ordnung. Ein Netz gehört zur ersten Ordnung, wenn die Verbindungen immer nur zu der direkt darüber liegenden Schicht verlaufen. Existieren Verbindungen, die eine oder mehrere Schichten überspringen, spricht man von Feedforward-Netzen zweiter Ordnung.

In Abbildung 6.4 ist ein zweilagiges Perzeptron erster Ordnung dargestellt. Das Netz erhält über die Eingabeneuronen seinen Eingabevektor (*Input*) $\mathbf{I} = (I_1, I_2, I_3, I_4)$. Dieser wird mit den zugehörigen Gewichten $\mathbf{w}_{ij}^h = (w_{1j}, w_{2j}, w_{3j}, w_{4j})$ multipliziert.

⁵Mit Zeitpunkt ist ein bestimmter Schritt bei der Berechnung des Netzes gemeint.

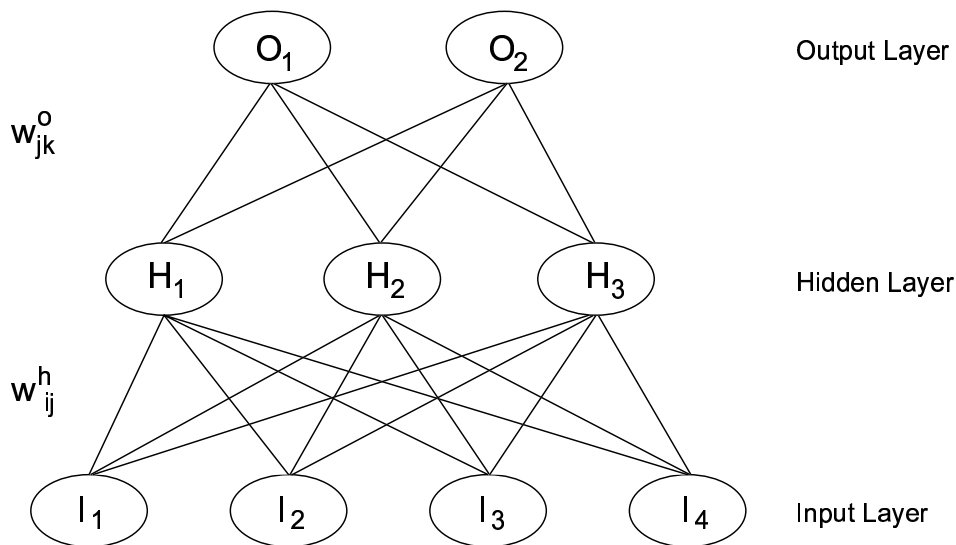


Abbildung 6.4: Ein zweistufiges Feedforward Netz 1. Ordnung.

Die Ausgaben des Input Layers werden von den Neuronen des Hidden Layers aufsummiert: $h_j = \sum_{i=1}^4 w_{ij}^h I_i$ und anschließend mit der Aktivierungsfunktion a_j neu berechnet, so dass sich als Aktivierung der Neuronen im Hidden Layer ergibt:

$$H_j(I) = a_j \left(\sum_{i=1}^4 w_{ij}^h I_i \right). \quad (6.2)$$

Diese Werte werden wiederum multipliziert mit den zu den Verbindungen gehörenden Gewichten $\mathbf{w}_{jk}^o = (w_{1k}, w_{2k}, w_{3k})$, die zum Output Layer führen. Die Output Neuronen summieren wieder auf: $o_k = \sum_{j=1}^3 w_{jk}^o H_j$, und als Ausgabe O_k erhält man über die Aktivierungsfunktion: $O_k = a_k(o_k)$. Die Ausgaben O_k des Netzes lassen sich in Abhängigkeit vom Input \mathbf{I} wie folgt darstellen:

$$O_k(\mathbf{I}) = a_k \left(\sum_{j=1}^3 w_{jk}^o a_j \left(\sum_{i=1}^4 w_{ij}^h I_i \right) \right). \quad (6.3)$$

Feedback-Netze verfügen im Gegensatz zu den Feedforward-Netzen auch über Verbindungen, die Neuronen der gleichen Schicht verbinden, die wieder zum gleichen Neuron zurückführen oder sogar Neuronen höherer Schichten mit Neuronen niedrigerer Layer verbinden. Oft ist eine eindeutige Zuordnung der Schichten nicht mehr möglich. Ein Beispiel für Feedback-Netze sind die so genannten *Hopfield*-Netze. Hier ist jedes Neuron mit jedem anderen Neuron verbunden, unabhängig welcher Schicht es angehört. Die Feedback-Netze sollen hier aber nicht

naher besprochen werden, da sie in der vorliegenden Arbeit keine Anwendung finden.

Das geeignete Netz fur ein gegebenes Problem zu finden, ist oft nicht einfach. Hier muss man sich auf Erfahrungswerte beziehen oder eine ganze Reihe von Tests durchfuhren, um das Optimum zu finden. Im Allgemeinen gilt die Regel, dass Feedforward-Netze hierarchisch von unten nach oben durchzustufen sind. Das heist, dass in einem nachfolgenden Layer weniger Neuronen sein sollten, als in der vorhergehenden Schicht. Dies ist aber nicht als festes Gesetz zu verstehen.

6.4 Der Backpropagation-Algorithmus

Mit dem Backpropagation-Algorithmus konnen Feedforward-Netze erster Ordnung trainiert werden, die eine stetige Funktion (oft eine sigmoide Funktion) als Aktivierungsfunktion nutzen. Backpropagation ist ein Gradientenabstiegsverfahren, bei dem das Netz von hinten nach vorne durchgearbeitet wird.

6.4.1 Herleitung

Jedes Muster P hat einen so genannten Zielvektor T_p . Dies ist der Vektor, der nach Durchlaufen des Netzes als Ausgabevektor erreicht werden soll. Die Differenz

$$\delta_{p,k} = T_{p,k} - O_{p,k} \quad (6.4)$$

gibt an, welchen Abstand die gewunschte Ausgabe fur ein bestimmtes Ausgabe-Neuron n_k zur tatsachlichen Ausgabe $O_{p,k}$ des Neurons fur einen festen Eingabevektor $I_{p,i}$ hat. Um den Lernerfolg zu messen, wird noch eine Fehlerfunktion ϵ

$$\epsilon = \sum_p \epsilon_p \quad (6.5)$$

benotigt, mit:

$$\epsilon_p = \frac{1}{2} \sum_k (T_{p,k} - O_{p,k})^2. \quad (6.6)$$

Die Gewichte zu einem Zeitpunkt t seien durch w_{jk}^t gegeben. Sie werden zum Zeitpunkt $t + 1$ folgendermaen geandert:

$$w_{jk}^{t+1} = w_{jk}^t + \Delta w_{jk} \quad (6.7)$$

mit:

$$\Delta w_{jk} = -\eta \frac{\partial \epsilon_p}{\partial w_{jk}}. \quad (6.8)$$

Der Parameter η ist die Lernrate, die bestimmt mit welcher Geschwindigkeit gelernt wird. Mittels der Kettenregel kann man schreiben:

$$\frac{\partial \epsilon_p}{\partial w_{jk}} = \frac{\partial \epsilon_p}{\partial a(w_{jk}H_{p,j})} \frac{\partial a(w_{jk}H_{p,j})}{\partial w_{jk}}. \quad (6.9)$$

Die Aktivierungsfunktion $a(w_{jk}H_{p,j})$ ist abhängig von der Gewichtung w_{jk} und dem Output $H_{p,j}$, den das Neuron j liefert. Für das Ausgabe-Neuron k gilt:

$$\epsilon_p = \frac{1}{2} (T_{p,k} - O_{p,k})^2 \quad (6.10)$$

und mit

$$O_{p,k} = a(w_{jk}H_{p,j}) \quad (6.11)$$

folgt dann :

$$\frac{\partial \epsilon_p}{\partial w_{jk}} = \frac{1}{2} \frac{\partial}{\partial a(w_{jk}H_{p,j})} (T_{p,k} - a(w_{jk}H_{p,j}))^2 \frac{\partial a(w_{jk}H_{p,j})}{\partial w_{jk}} \quad (6.12)$$

$$= \frac{\partial \epsilon_p}{\partial a(w_{jk}H_{p,j})} a'(w_{jk}H_{p,j}) H_{p,j}. \quad (6.13)$$

Mit der Ersetzung

$$\delta_{p,k} = \frac{\partial \epsilon_p}{\partial a(w_{jk}H_{p,j})} a'(w_{jk}H_{p,j}) \quad (6.14)$$

gilt dann:

$$\Delta w_{jk} = \eta \delta_{p,k} H_{p,j}. \quad (6.15)$$

$\delta_{p,k}$ lässt sich einfach berechnen:

$$\delta_{p,k} = a'(w_{jk}H_{p,j}) \cdot (T_{p,k} - O_{p,k}). \quad (6.16)$$

Will man $\delta_{p,j}$ für die Neuronen n_j einer verdeckten Schicht berechnen, so ist:

$$\delta_{p,j} = a'(w_{ij}^h H_{p,i}^{h-1}) \cdot \sum_k \delta_{p,k} w_{jk}^{h+1}. \quad (6.17)$$

Wobei $H_{p,i}^{h-1}$ die Ausgaben der Neuronen m_i der darunter liegenden Schicht, w_{ij}^h die Gewichte der Verbindungen **zu** Neuron j und $w_{j,k}^{h+1}$ die Gewichte der Verbindungen **von** Neuron j zu all seinen Nachfolgeneuronen sind (nach [Sch97] und [Ott96]).

6.4.2 Arbeitsweise

Backpropagation ist ein iteratives Verfahren. Zunächst wird ein Eingabe-Muster eingelesen und dann im Netz nach vorne propagiert (Feedforward). Es wird also Schicht für Schicht durchgerechnet, bis man zur Ausgabe gelangt. Der Output wird nun mit der Zielvorgabe verglichen. Ausgehend von der Differenz zwischen Zielvektor und Ausgabevektor wird der Beitrag jedes Neurons an dem Gesamtfehler berechnet und seine Gewichtungen werden dementsprechend mit den Formeln aus Kapitel 6.4.1 angepasst. Die Gewichtsänderung verläuft in Richtung des steilsten Abstiegs der Fehlerfunktion. Angefangen mit der Ausgabeschicht werden die einzelnen δ 's berechnet. Sind diese durchlaufen, geht man zur davorliegenden Schicht. Das Netz wird also von hinten nach vorne angepasst (Backpropagation).

Hier wurde die Arbeitsweise im *Online*-Verfahren dargestellt. Das heißt, nach jedem einzelnen Muster erfolgt eine erneute Anpassung des Netzes. Damit wird dafür gesorgt, dass die Fehlerfunktion „verrauscht“. Dies ist oft erwünscht, da dadurch eventuell lokale Minima verlassen werden können (vgl. Kap. 6.4.3). Im *Offline*-Verfahren werden zunächst alle Muster durchlaufen und erst dann wird das Netz über den mittleren Fehler über alle Pattern angepasst. Bei der Datenauswertung dieser Arbeit wurde ausschließlich das Online-Verfahren angewendet.

6.4.3 Probleme mit dem Backpropagation-Algorithmus

Obwohl der Backpropagation-Algorithmus eine sehr erfolgreiche Strategie ist, um Feedforward Netze zu trainieren, gibt es doch einige Umstände, die bei seiner Anwendung beachtet werden müssen (vgl. [Per92] und [Sch97]):

Der Lernparameter η

Es ist von großer Wichtigkeit darauf zu achten, den Lernparameter η so zu wählen, dass er weder zu groß noch zu klein ist. Ist die Lernrate groß, so wird schneller gelernt, es besteht aber die Gefahr, dass das Minimum der Fehlerfunktion nicht erreicht wird. Ist die Lernrate klein, so konvergiert der Fehler erst nach langer Zeit, oder gar nicht.

Initialisierung der Gewichte

Bevor man beginnt Backpropagation auf ein Netz anzuwenden, müssen den einzelnen Gewichte der Verbindungen zwischen den Neuronen unterschiedliche Werte

zugewiesen werden. Dies geschieht meist dadurch, dass man die Gewichte zufällig initialisiert, ihnen wird also ein zufälliger Wert zugewiesen. Geschieht dies nicht, so werden die verschiedenen von einem Neuron ausgehenden Verbindungen immer den gleichen Wert haben, egal wie viele *Epochen*⁶ trainiert werden.

Lokale und globale Minima

Man kann sich die Fehlerfunktion als eine mehrdimensionale „Landschaft“⁷ vorstellen. Ziel des Backpropagation-Algorithmus ist es in dieser Landschaft ein Minimum zu finden. Es kann passieren, dass man mit dem Algorithmus statt des globalen nur ein lokales Minimum erreicht. Dies liegt an dem zufälligen Anfangswert der Gewichtungen w_{ij} . Befindet sich zwischen Startwert und optimalem Minimum ein „Gebirge“ auf der Fehleroberfläche, so kann dieses Minimum nie erreicht werden. Diesem Problem kann man mit dem Online-Verfahren begegnen. Da eine Neuberechnung der Gewichte nach jedem Pattern erfolgt, ändert sich auch die Gestalt der Fehleroberfläche, sie wird verwaschen, und ein Gebirge könnte so überwunden werden.

Plateaus

Die Größe der Gewichtsänderungen der w_{ij} hängt hauptsächlich von dem Betrag des Gradienten ab. Erreicht man ein Plateau mit geringer Steigung, kann das Lernen nur noch langsam vorstatten gehen, ist die Steigung Null, so kommt das Lernverfahren zu einem vollständigen Stillstand.

Schluchten

Unter bestimmten Umständen kann es in Schluchten zu Oszillationen kommen. Sind die Schluchtwände nahe genug beieinander, so bewirken die Gewichtsänderungen, dass das Verfahren immer von einer Wand zur anderen Wand springt und dabei nie das Tal erreicht. Unter ganz ungünstigen Bedingungen kann die Schlucht sogar verlassen werden und man gelangt in Bereiche, die nur lokale Minima aufweisen.

6.5 Training und Validierung

Um ein Netz zu trainieren und seine Eignung für ein bestimmtes Problem zu bestimmen, braucht man mindestens zwei Datensätze. Mit dem einen Datensatz, den Trainingsdaten, wird das Netz trainiert, und mit dem anderen, den Testdaten, die

⁶Als eine Epoche bezeichnet man den kompletten Durchlauf eines Trainingsdatensatzes.

⁷Die Dimension entspricht der Anzahl der Verbindungen.

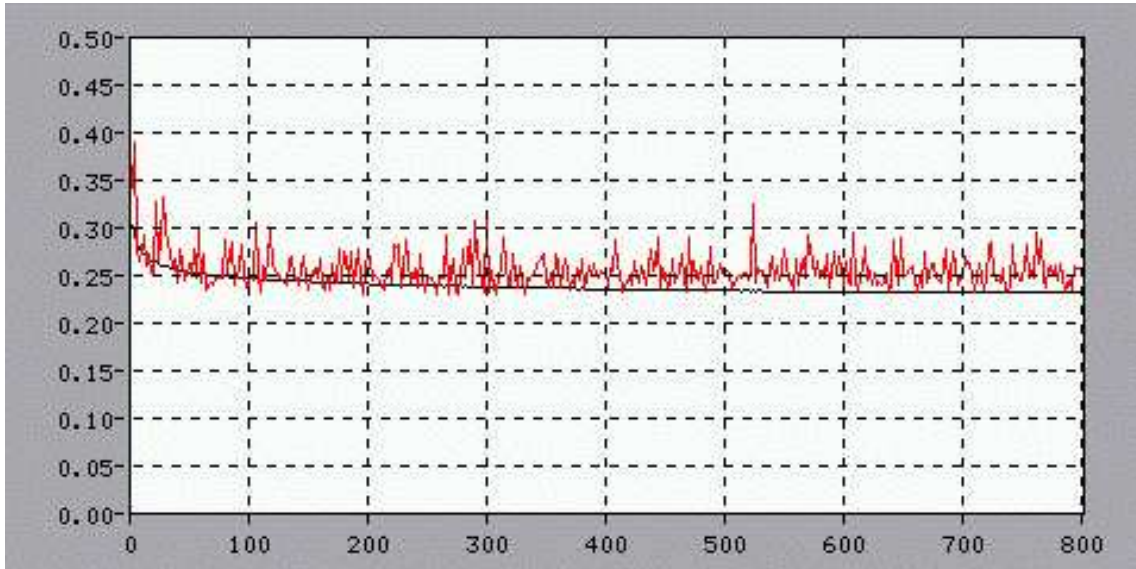


Abbildung 6.5: Entwicklung des *Mean Square Errors* (MSE) eines dreilagigen Feedforward-Netzwerkes, die x-Achse gibt die Anzahl der trainierten Epochen an. Die glatte Kurve ist der Verlauf für die Trainingsdaten, die andere der für die Testdaten. Hier wurde das Netz mit Daten aus einem Run mit $p = 2 \text{ GeV}/c$ (INV6-Radiator) mit Scaledown trainiert. Als Testdaten diente ein Run mit gleichem Radiator und gleichem Impuls, jedoch ohne Pionen-Scaledown. Der Lernparameter ist zu $\eta = 0.1$ gewählt worden.

Verallgemeinerungsfähigkeit des trainierten Netzes überprüft.

Die Anzahl der verschiedenen Trainingsmuster sollte groß sein, im Idealfall gegen unendlich gehen. Mit ihnen erfolgt die Anpassung des Netzes über den Trainingsalgorithmus. Die Testmuster haben keinen Einfluss auf die Änderung der Gewichte. Dies sind dem Netz unbekannte Daten. Mit ihnen kann überprüft werden, inwieweit das Netz unbekannte Muster richtig zuordnen kann. Eine ständige Überprüfung des Netzes ist erforderlich, um das Training bei maximaler Generalisierungsfähigkeit abbrechen. Die Generalisierungsfähigkeit kann zusätzlich noch über einen weiteren Datensatz überprüft werden. Hierzu wird oft ein dritter Datensatz verwendet, es kann aber auch der Datensatz verwendet werden, der für das Abbruchkriterium benutzt worden ist. Dann ist der Lernerfolg leicht positiv beeinflusst. Wird das Training nach einer fest vorgegebenen Anzahl von Epochen abgebrochen, hat die Verwendung des Testdatensatzes aber keinen Einfluss. Demgegenüber beeinflusst die Verwendung der Trainingsdaten als Testdaten für die Verallgemeinerungsfähigkeit das Ergebnis sehr stark in die positive Richtung. Man kann dann nicht mehr auf die Verallgemeinerungsfähigkeit des trainierten Netzes schließen, sondern nur noch darauf, wie gut das Netzwerk bekannte Muster richtig zuordnen kann. Für optimalen Lernerfolg sollten außerdem die Trainingsmuster von Epoche zu Epoche in zufälliger

Reihenfolge ins Netz eingespeist werden.

Ein weiterer Parameter, der für die Verallgemeinerungsfähigkeit eines Netzes entscheidend sein kann, ist das Abbruchkriterium. Es sind unterschiedliche Abbruchkriterien gebräuchlich. Zum einen kann der Lernerfolg *online* verfolgt werden und das Training abgebrochen werden, wenn es entweder zu keiner signifikanten Veränderung der Trainingskurve und der Testkurve kommt oder es zum *Overfitting* kommt. Overfitting bedeutet, dass sich das Netz zu sehr an die Trainingsdaten anpasst und die Generalisierungsfähigkeit nachlässt. Netze mit sehr vielen versteckten Neuronen neigen oft zu Overfitting. Durch ihre hohe Dimensionalität können sie sich an das Rauschen der Trainingsdaten anpassen. Eine andere Möglichkeit für den Abbruch ist eine zuvor festgelegte Zahl von Epochen zu trainieren. Diese Möglichkeit kann dazu benutzt werden, um die Leistungsfähigkeit verschiedener Netze mit den gleichen Daten zu untersuchen. Die dritte Möglichkeit ist das Training bei Erreichen eines gewissen Wertes von ϵ (Formel (6.5)) oder des *Mean Square Errors* (MSE) zu beenden.

In Abbildung 6.5 ist ein typischer zeitlicher Verlauf des MSE abgebildet. Man erkennt eine stetige Abnahme des MSE für die Trainingsdaten. Der Verlauf des MSE der Testdaten hingegen fluktuiert etwas oberhalb der Trainingsfunktion, fällt aber im Mittel dennoch weiter ab. Bei einem Overfitting des Netzes könnte man beobachten, dass die Testfunktion nicht mehr weiter abfällt, sondern im Mittel wieder ansteigen würde.

7. e/π Separation mit neuronalen Netzwerken

Bei den Analysen mit neuronalen Netzen ist das Programm *Stuttgart Neural Network Simulator* (SNNS) in der Version 4.2 benutzt worden. Dabei handelt es sich um ein Programm mit graphischer Oberfläche, bei dem eine ganze Reihe von gebräuchlichen Netzwerktypen, Lernalgorithmen, Analysetools und andere „Hilfswerkzeuge“ implementiert sind [SNNS4.2].

Die Theorie neuronaler Netze liefert keine Systematik, welche Netzwerktopologie verwendet, welcher Lernalgorithmus benutzt, oder in welcher Form Daten eingespeist werden sollten, um für ein gegebenes Problem die optimale Lösung zu finden. Pionen und Elektronen erzeugen im TRD verschiedene „Muster“. Bei dem vorliegenden Problem handelt es sich also um eine Variante der Mustererkennung. Feedforward-Netzwerke haben sich als besonders geeignet zur Mustererkennung herausgestellt. Das Standardlernverfahren für Feedforward-Netzwerke ist der Backpropagation-Algorithmus.

7.1 Wahl des Netzes und der Parameter

Elektronen deponieren im TRD im Mittel eine höhere Energie als Pionen (Q-Methoden). Dem Netz sollten also die absoluten Energiewerte zur Verfügung gestellt werden. Auf einer Elektronen-Spur findet man eine größere Zahl an Ladungsclustern als bei Pionen (N-Methoden). Der Input-Vektor sollte eine gewisse Auflösung in x-Richtung (was der Zeitachse entspricht) haben, um die Cluster voneinander trennen zu können. Da die Teilchen die Kammern in einem Winkel von 15° durchquert haben, sollten auch mehrere Pads berücksichtigt werden. In den vorliegenden Analysen ist eine 15×3 Eingabematrix verwendet worden. Das Signal der Kammer ist in 15 gleich große Zeitbins aufgeteilt worden, d.h. jeweils vier Zeitbins sind zu einem zusammengefasst worden. Es wurden, wie bei der LQ- und der LQX-Methode, drei nebeneinanderliegende Auslesepads berücksichtigt. Da sich herausgestellt hat, dass eine einfache Summation der vier FADC-Einträge zu einer Vollaktivierung der Eingabeneuronen führt, sind diese um einen Faktor 2000 herunter skaliert worden. Der Faktor 2000 ist gerade ausreichend dafür, dass die höchsten

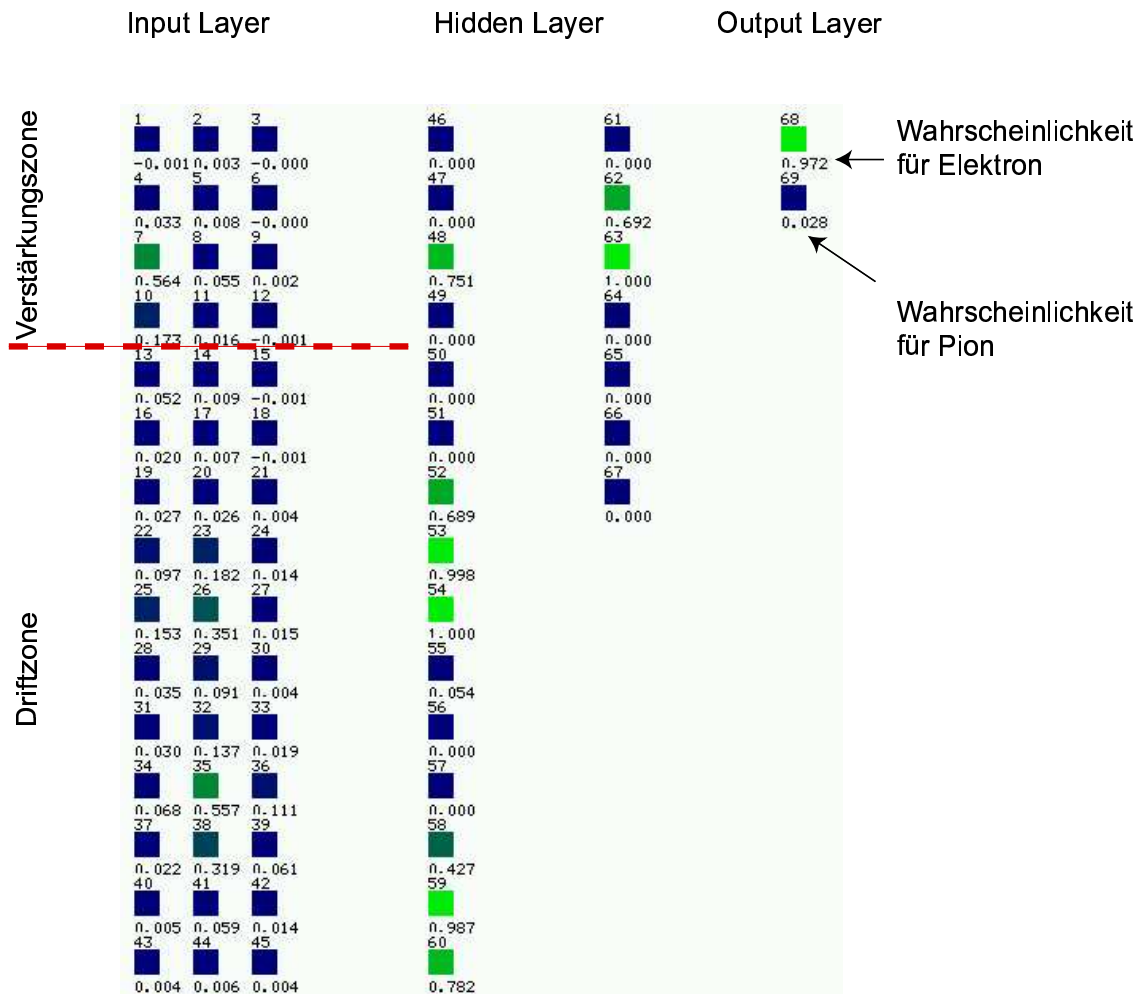


Abbildung 7.1: Das verwendete Netz für ein TRD-Modul. Links ist die Eingabeschicht, in der Mitte die versteckten Schichten und rechts die Ausgabeschicht. Im Input Layer erkennt man nebeneinander die drei ausgelesenen Pads und die Einteilung in 15 Zeitzeonen. Die Zahl unter den Neuronen ist die aktuelle Anregung. Sie ist auch farblich dargestellt: Grün bedeutet hohe, blau niedrige Anregung. Im Output Layer kann man die sich ergebenden Wahrscheinlichkeiten für ein Elektron- und für ein Pion-Ereignis ablesen. Dargestellt ist ein typisches Elektronereignis. Die Verbindungen der Neuronen untereinander sind aus Gründen der Übersicht nicht dargestellt.

Einträge nicht mehr zu einer Vollaktivierung der einzelner Eingabeneuronen führen¹. Oft werden zur größeren Verallgemeinerungsfähigkeit auch Drehungen und Spiegelungen der ursprünglichen Muster verwendet [Ott96]. Dies macht hier keinen Sinn, da keine Symmetrien vorhanden sind.

¹Eine Vollaktivierung der Eingabeneuronen hätte ähnliche Konsequenzen, wie die Überbelichtung eines Fotos. Bei „hellen“ Flächen, im Muster also Zonen mit hoher Aktivierung, lässt sich dann keine Substruktur mehr erkennen.

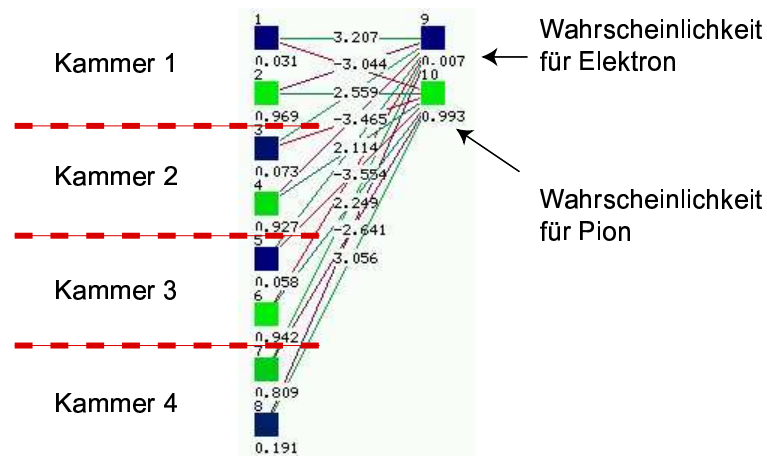


Abbildung 7.2: Die Zusammenführung der einzelnen Kammern. Dargestellt ist die Zusammenführung von vier Kammern. Als Input fungieren nun die Ausgaben, die die Netze der einzelnen Kammern liefern. Der Output ist die Gesamtwahrscheinlichkeit über alle vier Kammern. Zu sehen ist ein Pionereignis. Zusätzlich zu den Aktivierungen der Neuronen ist die Gewichtung einzelner Verbindungen zu erkennen.

Zunächst musste eine geeignete Netzwerktopologie gefunden werden. Dazu sind verschiedene Netzwerke (zunächst für eine einzelne Kammer) getestet, und ihre mittleren quadratischen Fehler (MSE) miteinander verglichen worden, die ein Maß für den Lernerfolg des jeweiligen Netzes sind. Es wurden Netze ohne, mit einem, zwei und drei Hidden Layern untersucht. Je mehr verdeckte Schichten vorhanden waren, desto kleiner war der MSE nach einer bestimmten Anzahl von Epochen. Das Netz mit drei Hidden Layern hatte zwar bei den Trainingsdaten den kleinsten MSE, doch konnte bei den Testdaten ein bestimmter Grenzwert nicht unterschritten werden. Netze mit zwei Hidden Layern zeigten vergleichbare Ergebnisse, bei ihnen ist die Gefahr der Überanpassung aber geringer. Von den Netzen mit zwei verdeckten Schichten war das Netz mit den meisten versteckten Neuronen am erfolgreichsten.

Als Lernalgorithmen sind Backpropagation-Algorithmen mit unterschiedlichen Lernparametern η (0,05, 0,1 und 0,2), sowie ein Backpropagation-Algorithmus mit zusätzlichem Impulsterm getestet worden. Es zeigte sich, dass die klassischen Backpropagation-Algorithmen für alle η s ein ähnliches Verhalten zeigten. Sie erreichten alle ähnliche MSE-Werte für die Trainingsfunktion und zeigten starke Schwankungen des MSE bei der Testfunktion (siehe auch Abb. 6.5). Hinzufügen eines Impulstermes brachte keine Steigerung der Lernrate.

In Abbildung 7.1 ist das letztlich verwendete Netz zu sehen. Als Lernparameter wurde $\eta = 0,1$ gewählt. Das Netzwerk besteht aus der Eingabeschicht (insgesamt 45

Impuls (GeV/c)	Trainingsdaten			Testdaten		
	Gesamt	Pionen	Elektronen	Gesamt	Pionen	Elektronen
2	58 032	31 677	26 355	11 691	7 965	3 726
3	58 706	32 453	26 253	17 829	14 595	3 234
4	49 326	26 434	22 892	25 834	22 256	2 578
5	37 450	17 707	19 743	25 245	23 477	1 768
6	44 144	16 393	27 751	28 404	26 986	1 418

Tabelle 7.1: Anzahl der Teilchen bei verschiedenen Impulsen. Die Trainingsdaten sind mit und die Testdaten ohne Pionen-Scaledown aufgenommen worden. Bei höheren Impulsen nimmt die Anzahl der Elektronen im Verhältnis zu den Pionen immer weiter ab.

Neuronen), zwei verdeckten Schichten (15 und 7 Neuronen) und der Ausgabeschicht (2 Neuronen). In der Abbildung sind die Neuronen durchnummeriert (Zahl über dem Neuron). Die Zahl unter dem Neuron, sowie seine Einfärbung geben den Anregungszustand wieder. Die Eingabeschicht kann man als schematisches Abbild der Driftkammer verstehen. Man erkennt, dass drei nebeneinanderliegende Pads ausgelesen werden, sowie die Einteilung der Kammer in 15 Zeitzonen. Zu sehen ist ein typisches Elektronevent. In diesem Ereignis wurden drei große Cluster erzeugt (bei den Neuronen 7, 26 und 35), von denen derjenige bei hohen Driftzeiten (Neuron 35) eventuell durch Übergangsstrahlung entstanden ist. Die Ausgabeschicht gibt die einzelnen Wahrscheinlichkeiten wieder [Ott96] (nach [Ric91], [Den91]). In diesem Fall wird das Teilchen mit einer Wahrscheinlichkeit von 97.2% als Elektron identifiziert.

7.2 Datenauswertung und Vergleich mit der LQX-Methode

Während der Strahlzeit sind Datensätze mit und ohne Pionen-Scaledown unter identischen Bedingungen erfasst worden. In dieser Arbeit werden Daten ausgewertet, die mit einem INV6 Radiator und Impulsen $p = 2, 3, \dots, 6$ GeV/c aufgenommen worden sind. Als Trainingsdaten dienen die Runs mit Scaledown, als Test- und Validierungsdaten sind die Runs ohne Scaledown benutzt worden.

Die Trainingsdatensätze enthielten zwischen 40 000 und 60 000 Muster mit ungefähr gleich vielen Elektronen und Pionen, die Testdatensätze zwischen 10 000 und 30 000 Muster. Bei den Testdaten nimmt der Anteil, der durch Elektronen erzeugten Mu-

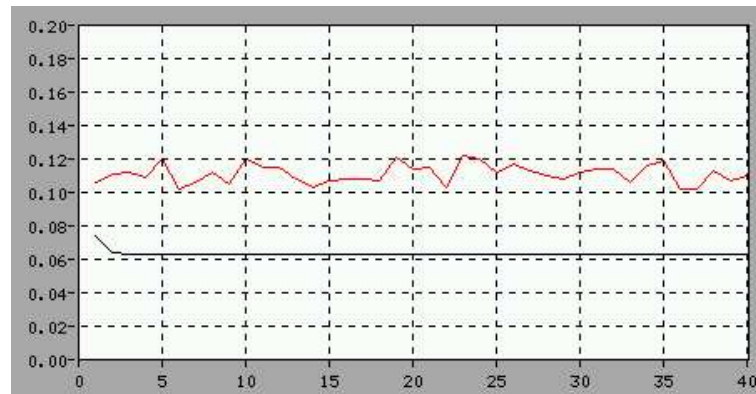


Abbildung 7.3: Fehlerentwicklung beim Training eines Netzes, das vier Kammern zusammenführt. Die Trainingskurve konvergiert schon nach sehr kurzer Zeit, während die Testkurve um einen bestimmten Wert fluktuiert. In x-Richtung ist die Anzahl der Epochen und in y-Richtung der MSE aufgetragen. Dieses Beispiel stammt aus dem Training für einen Impuls von 6 GeV/c.

ster, bei höheren Impulsen stark ab (siehe Tabelle 7.1). Die Netze wurden jeweils 1 000 Epochen trainiert. Weiteres Training machte keinen Sinn, da eine weitere Fehleroptimierung nicht abzusehen war.

Da es sehr aufwendig ist ein neuronales Netz zu finden, das geeignet ist alle Kammern gleichzeitig auszuwerten, wurden nur Netze für einzelne Kammern trainiert. Die Ergebnisse von Trainings- und Testdatensatz wurden in weitere neuronale Netzwerke eingespeist. Diese waren so konzipiert, dass die Ergebnisse von zwei, drei und vier Kammern gleichzeitig eingelesen werden konnten. Die Ergebnisse aus Runs mit Pionen-Scaledown dienten wieder als Trainingsdaten, die ohne als Test- und Validierungsdaten. In Abbildung 7.2 ist das Netz für vier Kammern mit einem typischen Pionereignis dargestellt. In der Eingabeschicht sind die jeweiligen Wahrscheinlichkeiten, die die Netze für die einzelnen Kammern produzieren, als Anregung der Neuronen wieder zu erkennen. In dieser Darstellung sind außerdem noch die Verbindungen zwischen den Neuronen mit ihrer jeweiligen Gewichtung dargestellt. Die Ausgabeschicht gibt schließlich die Gesamtwahrscheinlichkeit für das jeweilige Ereignis wieder. Einfügen von versteckten Schichten brachte bei diesen Netzen keine weitere Verbesserung des Ergebnisses, daher wurde darauf verzichtet.

Anzumerken ist noch, dass das Training dieser Netze sehr schnell abgeschlossen war (Abb.7.3). Die Trainingskurve konvergiert schon nach den ersten Epochen und ändert sich praktisch nicht mehr. Die Testkurve fluktuiert um einen bestimmten Wert. Bei der Auswertung sind die Netze dennoch 500 Epochen trainiert worden, um eine gute Teichenerkennung sicher zu stellen. Es war keine Tendenz zur Überanpassung zu erkennen. Der Lernparameter η war auf 0,05 festgelegt worden.

Die weitere Vorgehensweise ist identisch zu der Vorgehensweise, wie sie bei der Extrapolationsmethode angewendet wird. Die Ergebnisse der neuronalen Netze für zwei, drei und vier Kammern werden, ebenso wie die Ergebnisse, die aus dem Netz für Kammer 1 stammen, in Likelihoodverteilungen gefüllt. Aus diesen Likelihoodverteilungen lassen sich dann die jeweiligen Effizienzen mit den Formeln (5.5) und (5.6) bestimmen. Wie in Abbildung 5.5 können so die Pionen-Effizienzen für sechs Kammern extrapoliert werden (Abb. 7.4).

In Abbildung 7.5 sind die Pioneneffizienzen für die LQX-Methode und für neuronale Netze aufgeführt. Wie in Abbildung 5.9 erkennt man auch hier eine Impulsabhängigkeit. Die Pionenunterdrückung mit neuronalen Netzen ist signifikant (bis zu einem Faktor fünf) besser, als mit der LQX-Methode.

Die Genauigkeit der verwendeten Methoden, und damit der Fehler der Pionenunterdrückung, ist relativ schwer zu bestimmen. So ist z.B. in den „Extrapolations-Diagrammen“ häufig zu beobachten, dass die Effizienzen nicht ganz auf einer Gerade liegen. Es bleibt zu untersuchen, ob eine andere Wahl des Lernparameters η zu einer Glättung führt. Ein anderer kritischer Aspekt ist die niedrige Statistik der Elektronereignisse. Um dieses Problem zu umgehen, sind Netze mit Daten aus Runs mit Impulsen von 1 und 2 GeV/c (INV8, mit Pionen-Scaledown) trainiert und mit einem Impuls von 1,5 GeV/c getestet worden. Das Resultat reiht sich in die erzielten Ergebnisse ein und erreicht eine Pionenunterdrückung von etwa einem Faktor 1000.

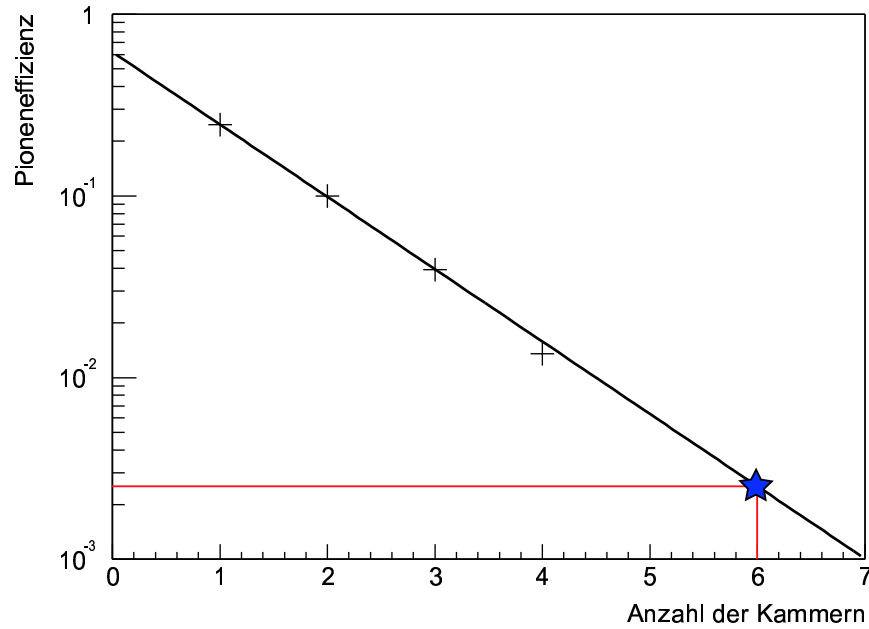


Abbildung 7.4: Extrapolationsmethode bei den neuronalen Netzen. Wie bei der LQX-Methode lässt sich aus den Effizienzen für eine unterschiedliche Anzahl von Kammern die Pioneneffizienz für sechs Kammern bestimmen. Hier ist die Extrapolation für einen Impuls von $2 \text{ GeV}/c$ dargestellt.

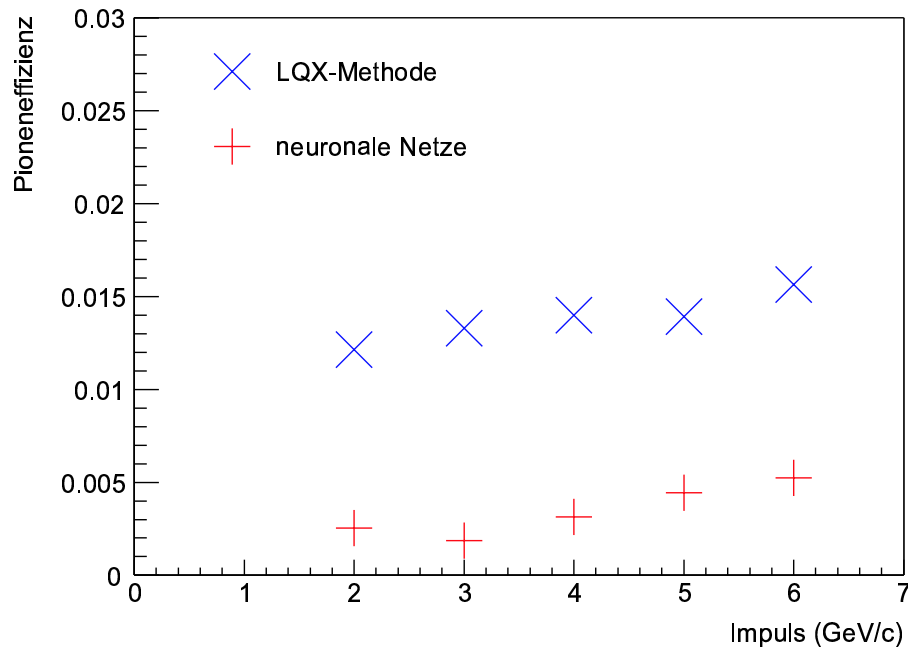


Abbildung 7.5: Vergleich der Pioneneffizienz mit der LQX-Methode und mit neuronalen Netzen.

8. Zusammenfassung und Ausblick

In dieser Arbeit sind klassische Methoden und Analysen mit neuronalen Netzen zur e/π -Separation untersucht worden. Die untersuchten Daten stammen aus der Strahlzeit, die 2002 am CERN PS durchgeführt worden ist. Dabei zeigte sich, dass mit neuronalen Netzen eine deutlich höhere Pionenunterdrückung erreicht werden kann als mit klassischen Methoden. Die Untersuchung mit neuronalen Netzen erfolgte in zwei Schritten. Zunächst wurden Netze mit den Daten der einzelnen Kammern trainiert. Die Ergebnisse dieser Netze wurden durch ein zweites neuronales Netz ausgewertet, um damit die Pioneneffizienzen mit der Extrapolationsmethode zu bestimmen.

Bei den Netzen zur Auswertung der Kammern handelte es sich um Feedforward-Netzwerke mit zwei verdeckten Schichten. Als Input diente eine 15×3 Matrix, deren Elemente der deponierten Ladung in einem bestimmten Bereich der Kammer entsprachen. Der Ausgabevektor spiegelt die jeweilige Wahrscheinlichkeit für ein Elektron- bzw. Pion-Ereignis wider. Als Lernalgorithmus wurde ein Backpropagation-Algorithmus mit einem Lernparameter $\eta = 0,1$ verwendet. Die Netze wurden jeweils 1000 Epochen lang trainiert. Die Netze zur Kombination der Kammer-Ergebnisse waren Einschicht-Perzeptronen. Die Dimension des Eingabevektors richtete sich dabei nach der Anzahl der betrachteten Kammern. Diese Netze wurden 500 Epochen mit $\eta = 0,05$ trainiert. Mit diesen Ergebnissen wurden Pioneneffizienzen bestimmt und daraus die Pioneneffizienz für sechs Kammern extrapoliert. Die Auswertung mit neuronalen Netzen zeigt bei den analysierten Daten eine signifikante Steigerung der Pionenunterdrückung im Vergleich zu der LQX-Methode. Ein großer Vorteil der neuronalen Netze ist ihre Schnelligkeit. Ist ein Netz einmal trainiert, so sind zur Teilchen-Klassifikation nur noch einige Additionen und Multiplikationen durchzuführen. So wäre es vorstellbar, dass ein solches Verfahren Online verwendet werden kann. So könnte ein neuronales Netz im Onlinetrigger von ALICE eingesetzt werden, um hochenergetische Elektronen zu erkennen.

Zusammenfassend lässt sich sagen, dass eine e/π -Separation im ALICE TRD mit neuronalen Netzen eine viel versprechende Alternative zu den klassischen Analyseverfahren darstellt. Die Untersuchungen in dieser Arbeit sind ein erster Versuch, ein solches Verfahren zu entwickeln. Weitere Analysen sind nötig, um es zu optimieren und seine Einsatzfähigkeit im realen Detektor zu zeigen.

A. Variablen der Schwerionenphysik

In der Regel werden zur Beschreibung ultrarelativistischer Schwerionenstöße spezielle kinematische Größen verwendet. Dazu zählen die Rapidität y und die Pseudorapidität η . Diese haben den Vorteil, dass sie unter Lorentztransformation ein einfaches Verhalten aufweisen. Die Energie E eines Teilchens ist¹:

$$E^2 = m_0^2 + \vec{p}^2 \quad (\text{A.1})$$

mit m_0 als Ruhemasse des Teilchens und dem Impuls \vec{p} . Der Impuls \vec{p} lässt sich bezüglich der Strahlachse in einen transversalen (p_T) und einen longitudinalen (p_L) Anteil zerlegen. Es gilt:

$$p = \sqrt{p_L^2 + p_T^2}. \quad (\text{A.2})$$

Der Transversalimpuls p_T ist invariant gegenüber Lorentztransformationen, der Longitudinalimpuls p_L jedoch nicht. Deshalb führt man die so genannte Rapidität ein:

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right). \quad (\text{A.3})$$

Diese ist zwar auch nicht lorentzinvariant, die Transformationseigenschaft ist jedoch einfach. Die Rapiditäten verhalten sich bei einer Lorentztransformation additiv, die Form der Rapiditätsverteilung ändert sich bei einem Wechsel des Bezugssystems nicht.

Im Experiment ist die Rapidität häufig schwer zugänglich, denn zur Rapiditätsbestimmung wird die Energie und der Longitudinalimpuls benötigt. Meistens ist jedoch nur der Emissionswinkel ϑ zur Strahlachse bekannt. Man führt eine weitere Größe ein, die so genannte Pseudorapidität:

$$\eta = \frac{1}{2} \ln \left(\frac{p + p_L}{p - p_L} \right) = -\ln \left(\tan \frac{\vartheta}{2} \right). \quad (\text{A.4})$$

Für ultrarelativistische Teilchen sind Rapidität und Pseudorapidität identisch.

¹Es wird hier und in den folgenden Gleichungen $c = 1$ gesetzt.

Literaturverzeichnis

- [Ake98] T. Akeson et al., *Electron identification with a prototype of the Transition Radiation Tracker for the ATLAS experiment*, Nucl. Instr. Meth. A412, 1998.
- [And01] A. Andronic et al., *Prototype tests for the ALICE TRD*, IEEE Trans. Nucl. Sc., vol. 48, 1259, 2001.
- [And02] A. Andronic et al., *Prototype tests for the ALICE TRD*, GSI Scientific Report 2002, GSI 2003-1, p.239, 2003.
- [And04] A. Andronic, private Mitteilung, 2004.
- [ATL97] ATLAS *Inner Detector Technical Design Report, Volume II*, Cern/LHC, 30 April 1997.
- [ATP99] ALICE *Addendum to Technical Proposal*, CERN/LHCC 99-13, 1999.
- [Dem98] W. Demtröder, *Experimentalphysik 4, Kern-, Teilchen- und Astrophysik*, Springer Verlag, 1998.
- [Den91] J. S. Denker, Y. leCun, *Transforming Neural-Net Output Levels to Probability Distributions*, Neural Information Processing Systems **3** 853-859, 1991.
- [Dol92] B. Dolgosheim, *Transition Radiation Detectors*, Nucl. Instr. Meth. A329, 1993.
- [Ego00] V. Egorytchev, V. Saveliev, S. J. Aplin, *Particle identification via transition radiation and detectors*, Nucl. Inst. Meth. A453, 2000.
- [Gar57] G. M. Garibian, Zh. Eksp. Theor. Fiz **33**, 1957.
- [Ger86] Gerthsen, Kneser, Vogel, *Physik*, 15. Auflage, Springer Verlag, 1986.
- [Gin46] V. L. Ginzburg, I. M. Frank, Zh. Eksp. Theor. Fiz **16**, 15, 1946.

- [Grz03] B. Grzadkowski, J. Pliszka, J. Wudka, *Triviality and stability limits on the Higgs boson mass in effective theories*, Acta Phys. Polon. B34, 2003.
- [Hah03] H. Hahn et al., *The RHIC design overview*, Nucl. Instr. Meth. A499, 2003.
- [Har03] M. Harrison et al., *RHIC project overview*, Nucl. Instr. Meth. A499, 2003.
- [Jac62] Jackson, *Classical Electrodynamics*, Wiley Sons, 1962.
- [Kar02] F. Karsch, *Lattice Results on QCD Thermodynamics*, Nucl. Phys. A 698, 199c-208c, 2002.
- [Kle92] K. Kleinknecht, *Detektoren für Teilchenstrahlung*, Teubner Verlag, 1992.
- [Lud03] T. Ludlum, L. McLerran, *What we have learned from the Relativistic Heavy Ion Collider*, Physics Today, October 2003.
- [May94] T. Mayer-Kuckuk, *Kernphysik*, Teubner Verlag, 1994.
- [Mit03] V. A. Mitsou for the ATLAS Collaboration, *The ATLAS Transition Radiation Tracker*, hep-ex/0311058, 2003.
- [Ott96] R. Ottenhus, *Teilchenklassifizierung im Bleiglaskalorimeter mit künstlichen neuronalen Netzen*, Diplomarbeit, Institut für Kernphysik, Münster, 1996.
- [PDG00] Particle Data Group *Review of Particle Physics*, Springer Verlag, 2000.
- [Pei97] T. Peitzmann, *Kernmaterie unter extremen Bedingungen - Die experimentelle Suche nach dem Quark-Gluon-Plasma*, Habilitationsschrift, Institut für Kernphysik, Münster, 1997.
- [Per90] D. H. Perkins, *Hochenergiephysik*, Addison-Wesley, 1990.
- [Per92] P. Peretto, *An Introduction to the Modeling of Neural Networks*, Cambridge University Press, 1992.
- [Pov93] Povh, Rith, Scholz, Zetsche, *Teilchen und Kerne*, Springer Verlag, 1993.
- [Ric91] M. D. Richards, R. P. Lippmann, *Neural Network Classifiers Estimate Bayesian a posteriori Probabilities*, Neural Computation **3** (4), 1991.
- [ROO01] *ROOT, An Object-Oriented Data Analysis Framework*, Users Guide 3.1c, <http://www.root.cern.ch/root/RootDoc.html>, 2001.

- [Sch97] A. Scherer *Neuronale Netze: Grundlagen und Anwendungen*, Vieweg, 1997.
- [Sch02] J. Schuhkraft, *Heavy Ions at the LHC: Physics Perspectives and Experimental Program*, Nucl. Phys. A698, 2002.
- [SNNS4.2] SNNS *Stuttgart Neural Network Simulator, UsersManual, Version 4.2*, Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, 1998, <http://www-ra.informatik.uni-tuebingen.de/downloads/SNNS/SNNSv4.2.Manual.pdf>.
- [TDR01] ALICE TRD, *Technical Design Report*, CERN/LHC, 2001-021, ALICE TDR 9, 3 October 2001.
- [Wah00] M. Wahn, *Teststrahlzeit und Simulationen mit einem Übergangsstrahlungsdetektor für das ALICE-Experiment*, Diplomarbeit, Institut für Kernphysik, Münster, 2000.
- [Win02] O. Winkelmann, *Teststrahlzeit mit einem Übergangsstrahlungsdetektor zur e/π -Separation*, Diplomarbeit, Institut für Kernphysik, Münster, 2002.
- [Zau03] O. Zaudtke, *Simulation einer TRD-Driftkammer des ALICE-Experiments mit GARFIELD*, Diplomarbeit, Institut für Kernphysik, Münster, 2003.

Danksagung

An dieser Stelle möchte ich mich bei meiner Familie, meinen Freunden und meinen Kollegen bedanken, die mich während meines Studiums und der Anfertigung dieser Arbeit unterstützt haben und zu ihrem Gelingen beigetragen haben.

Bei meiner Familie möchte ich mich dafür bedanken, dass sie mir mit moralischer und materieller Unterstützung dieses Studium ermöglicht hat und immer ein offenes Ohr für meine Probleme und Problemchen hatte.

Herrn Prof. Dr. R. Santo danke ich für die interessante Aufgabenstellung, die Möglichkeit diese Arbeit am Institut für Kernphysik anzufertigen, die lehrreichen Aufenthalte an der GSI und für die Möglichkeit an der DPG-Tagung in Tübingen sowie der ALICE-Week in Heidelberg teilzunehmen.

Bei Dr. Damian Bucher bedanke ich mich für die Betreuung sowie für zahlreiche Anregungen und Hilfestellungen ohne die diese Arbeit kaum so zu Stande gekommen wäre.

Großer Dank geht an Dipl.-Phys. Christian Klein-Bösing (für die große und leider viel zu oft benötigte Hilfe bei Software- und Computerproblemen), Dipl.-Phys. Oliver Zaudtke und Baldo Sahlmüller (für die nette Büroatmosphäre und ihre Hilfsbereitschaft) wie auch an Prof. Dr. Johannes P. Wessels (für die stete Diskussionsbereitschaft und seine zahlreichen Anregungen).

Dr. Anton Andronic danke ich für die geduldige Beantwortung verschiedener Fragen, die sich während dieser Arbeit gestellt haben.

Für die angenehme Zusammenarbeit in der Arbeitsgruppe bedanke ich mich bei: Dipl.-Phys. H.-H. Adam, Dr. S. Bathe, C. Baumann, Dr. H. Büsching, Dr. R. Glasow, H. Gottschlag, N. Heine, M. Hoppe, PD Dr. A. Khoukaz, Dipl.-Phys. T. Korfmeier, Dipl.-Phys. N. Lang, R. Menke, Dipl.-Phys. T. Mersmann, T. Rausmann, Dipl.-Phys. S. Steltenkamp, Dipl.-Phys. J. Stevering, Dipl.-Phys. A. Täschner, W. Verhoeven.

Ein dickes Dankeschön geht an Dr. Klaus Reygers, Dr. Damian Bucher und Christian Klein-Bösing für die kritische Durchsicht der Arbeit, sowie an Holger Schmalde, Björn Siemer und Dr. Andre Wenning für stilistische Ratschläge. Cornelia Röser danke ich für die professionelle Beratung in Rechtschreibangelegenheiten.

Des weiteren möchte ich mich bei Joey's Pizza Service für vorzügliche und vitaminreiche Nahrung während des Zusammenschreibens (eure Muffins sind echt lecker), bei der Krombacher Brauerei, die es mir ermöglichte einige Quadratmeter Regenwald zu retten, und bei Blizzard Entertainment für grandiose Unterhaltung während meiner freien Stunden bedanken.

Für zahlreiche Ablenkungen außerhalb der Physik danke ich Amin, Andreundanja, Beate (für das „Fenster zum Glück“), Boris, Christian K. (schöne Grüße nach Düsseldorf), Christian K.-B., Conny, Floh, Isa, Jovi (für zahllose ziellose Touren durch diverse Innenstädte), Magic (für zahlreiche, mehr oder weniger sinnvolle Diskussionen ;-)), Mela (der ich nicht nur Tine zu verdanken habe), Melly und Lucas, Nicola, Nicole (für zahlreiche Kinobesuche), Ricki, Tim (schöne Grüße an die Kellerbar), Tobi, Zoltan, meiner DOKO-Runde (nächstes Mal gewinn ich wieder), den Theater-Cafe Gängern, meinen Mitbewohnern Holger, Regina und Björn und nicht zuletzt Tine.

*Hiermit bestätige ich, daß ich diese Arbeit selbstständig
verfasst und keine anderen als die angegebenen
Quellen und Hilfsmittel verwendet habe.*

Münster, 3. März 2004

Alexander Wilk