

**Studien zur Markierung von b-Quarks mit Hilfe  
niederenergetischer Elektronen und zur Suche nach  
unsichtbar zerfallenden Higgs-Bosonen im  
Phion-Modell**

**Diplomarbeit  
vorgelegt von  
Sascha Thoma**

**Betreuer: Prof. Dr. Karl Jakobs**



**Fakultät für Mathematik und Physik  
der Albert-Ludwigs-Universität Freiburg,  
im Dezember 2007**

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Das Standardmodell der Teilchenphysik</b>	<b>5</b>
2.1	Grundlagen . . . . .	5
2.2	Das Prinzip der Eichinvarianz . . . . .	7
2.3	Der Higgs-Mechanismus . . . . .	8
2.4	Die Theorie des Standardmodells . . . . .	10
2.4.1	Das Glashow-Weinberg-Salam-Modell . . . . .	10
2.4.2	Die Quantenchromodynamik . . . . .	13
2.5	Wichtige Higgs-Produktionskanäle und Higgs-Zerfälle . . . . .	15
2.6	Physik des b-Quark-Zerfalls . . . . .	16
<b>3</b>	<b>Eine Erweiterung des Standardmodells: Das Phion-Modell</b>	<b>21</b>
<b>4</b>	<b>Das ATLAS-Experiment</b>	<b>25</b>
4.1	Der ATLAS-Detektor . . . . .	27
4.1.1	Der Spurdetektor . . . . .	27
4.1.2	Das Kalorimetersystem . . . . .	29
4.1.3	Das Myonen-System . . . . .	31
4.1.4	Hermetizität des Detektors . . . . .	31
4.1.5	Das Triggersystem . . . . .	32
<b>5</b>	<b>Verwendete Programme</b>	<b>33</b>
5.1	ROOT . . . . .	33
5.2	Toolkit for Multivariate Analysis . . . . .	33
5.3	ATHENA . . . . .	34
5.4	PYTHIA . . . . .	34
5.5	GEANT . . . . .	35
5.6	ATLFAST . . . . .	35
5.7	Die Ereignisrekonstruktion . . . . .	36
5.7.1	Spurrekonstruktion . . . . .	36
5.7.2	Standardrekonstruktion der Elektronen und Photonen . . . . .	36
5.7.3	Myonen-Rekonstruktion . . . . .	37
5.7.4	Rekonstruktion der Tau-Leptonen . . . . .	38

5.7.5	Jet-Rekonstruktion . . . . .	38
5.7.6	Vertex-Rekonstruktion . . . . .	38
<b>6</b>	<b>Methoden zur Klassifikation von Ereignissen</b>	<b>41</b>
6.1	Die Likelihood-Methode . . . . .	42
6.2	Neuronale Netzwerke . . . . .	45
6.3	Entscheidungsbäume . . . . .	50
<b>7</b>	<b>Eine Studie zum Nachweis eines Higgs-Bosons im Phion-Modell</b>	<b>55</b>
7.1	Ereignisgenerierung und Detektor-Simulation . . . . .	56
7.2	Die assoziierte Higgs-Produktion mit einem Z-Boson . . . . .	58
7.3	Die Vektorboson-Fusion . . . . .	64
7.4	Die Signifikanz des Signals und das Entdeckungspotential . . . . .	68
7.5	Die Signifikanz und die Ausschlussgrenzen im Phion-Modell . . . . .	72
<b>8</b>	<b>Die Identifikation der niederenergetischer Elektronen</b>	<b>77</b>
8.1	Verwendete Datensätze . . . . .	77
8.2	Bisherige Identifikation niederenergetischer Elektronen . . . . .	78
8.3	Zusätzliche Variablen zur Identifikation der niederenergetischen Elektronen . . . . .	85
8.4	Training der neuen Methoden . . . . .	90
8.4.1	Konfiguration der Likelihood-Methode . . . . .	91
8.4.2	Konfiguration der Neuronalen Netzwerke . . . . .	93
8.4.3	Konfiguration der Entscheidungsbäume . . . . .	99
8.4.4	Vergleich der Methoden . . . . .	102
8.5	Effizienz der Methoden bei der Identifikation der niederenergetischen Elektronen in ATHENA . . . . .	102
<b>9</b>	<b>Anwendung der neuen Methoden auf die Markierung von b-Quarks</b>	<b>111</b>
9.1	Algorithmus zur Markierung von b-Quarks . . . . .	111
9.2	Anwendung der neuen Methoden . . . . .	112
9.2.1	Markierung von b-Quarks ohne Verwendung von $d0$ und $p_t^{rel}$ . . . . .	115
9.2.2	Markierung von b-Quarks mit der Information über $d0$ und $p_t^{rel}$ . . . . .	120
9.3	Abhängigkeit der Methoden vom transversalen Impuls der Jets . . . . .	124
9.4	Bestimmung der linearen Korrelation zwischen dem SoftElectronTag und dem Combined Tag . . . . .	127
<b>10</b>	<b>Zusammenfassung</b>	<b>129</b>
<b>A</b>	<b>Tabellen mit den Signifikanzen und Ausschlussgrenzen im Phion-Modell</b>	<b>133</b>
A.1	Die Tabellen für den assoziierten Higgs-Produktionskanal (ZH): . . . . .	133
A.2	Die Tabellen für die Vektorboson-Fusion (VBF): . . . . .	137
<b>B</b>	<b>Ausgabewerte der Methoden bei der Markierung von b-Quarks</b>	<b>141</b>

<b>C Bestimmung des linearen Korrelationskoeffizienten zweier Verteilungen</b>	<b>149</b>
<b>Tabellenverzeichnis</b>	<b>150</b>
<b>Abbildungsverzeichnis</b>	<b>154</b>
<b>Literaturverzeichnis</b>	<b>160</b>

# 1 Einleitung

Seit jeher versucht der Mensch, die Natur, die ihn umgibt, zu verstehen und dieses Verständnis für sich zu nutzen. Das Vorgehen ist dabei immer dasselbe: Durch genaue Beobachtung und Experimente versucht er, Theorien aufzustellen und vorhandene zu verifizieren oder falsifizieren. Dabei ist er immer wieder an Grenzen gestoßen, die zunächst einmal unüberwindlich schienen, sei es durch gesellschaftliche, technische oder verständnisbedingte Probleme. Dennoch konnten die meisten dieser Probleme überwunden und somit eine große Menge an Wissen angesammelt werden.

Zu diesem Wissen gehören inzwischen unzählige Theorien und Fakten über die mikroskopischen und makroskopischen Bestandteile unseres Universums. Eine dieser Theorien im Bereich der Teilchenphysik ist das so genannte Standardmodell, welches die mikroskopische Zusammensetzung der Materie und die Wechselwirkung ihrer Bestandteile untereinander nach bisherigem Verständnis beschreibt. In ihm sind alle bisher bekannten Elementarteilchen enthalten. Weiterhin wird durch das Standardmodell ein weiteres Teilchen vorhergesagt: Das Higgs-Boson. Ohne dieses Higgs-Boson wäre die theoretische Beschreibung massiver Teilchen nicht möglich. Warum dies so ist, wird in Kapitel 2 erklärt.

Doch auch das Standardmodell kann die Natur nicht exakt beschreiben. Zum Einen wäre es wünschenswert, alle bekannten Wechselwirkungen, und zwar die Gravitation, die starke, die schwache und die elektromagnetische Wechselwirkung, bei hohen Energien durch eine einzige Theorie, einer „Theory of Everything“ (TOE), zu beschreiben. Für niedrige Energien wäre das Standardmodell dann eine Näherung für diese alles beschreibende Theorie. Dass die Vereinheitlichung zunächst einmal sehr unterschiedlich scheinender Wechselwirkungen überhaupt möglich ist, wurde schon durch die gemeinsame Beschreibung der magnetischen und elektrischen Wechselwirkung durch J. Maxwell (1864) und später durch die Vereinheitlichung der schwachen und elektromagnetischen Wechselwirkung in der so genannten elektroschwachen Wechselwirkung durch S. L. Glashow, A. Salam und S. Weinberg (Nobel-Preis 1979) gezeigt. Die Vereinheitlichung dieser elektroschwachen Theorie und der starken Wechselwirkung zu einer „Großen Vereinheitlichten Theorie“ (GUT<sup>1</sup>) wäre der nächste Schritt zur TOE. Dieser wird jedoch mit dem Standardmodell allein nicht realisiert werden können.

Zum Anderen leidet das Standardmodell am so genannten Hierarchie-Problem. Eine extrem genaue Feinabstimmung der Parameter im Higgs-Sektor ist notwendig, damit das Standardmodell bis zur Energie der Planck-Masse von  $m_P \simeq 1.22 \cdot 10^{19} \text{ GeV}/c^2$  gültig sein kann. Ansonsten treten bei der Berechnung des Propagators des Higgs-Bosons mit Schleifenkorrekturen Divergenzen auf, die nicht kompensiert werden können.

---

<sup>1</sup>engl.: Grand Unified Theory

Ein weiteres Problem des Standardmodells ist, dass das Universum nicht nur aus der Materie bestehen kann, welche wir mit dem Standardmodell beschreiben können. Stabile Galaxien mit den Drehgeschwindigkeiten, die gemessen wurden, wären mit der Menge an Materie, wie sie das Standardmodell erlaubt, nicht möglich. Zusätzlich muss also noch weitere Materie existieren. Nach heutigem Kenntnisstand liegt der Beitrag der im Standardmodell beschriebenen Materie und Energie zur gesamten Energiedichte des Universums bei ca. 5%. Der Rest besteht zu ca. 23% aus dunkler Materie, während die dunkle Energie die übrigen 72% zur Energiedichte des Universums beiträgt. Aus was diese dunkle Materie und Energie bestehen, ist bis heute unklar.

Diese und andere Gründe veranlassen Wissenschaftler dazu, Erweiterungen des Standardmodells zu ersinnen. Viele dieser Erweiterungen sind in der Lage, einige Probleme des Standardmodells zu lösen, doch stößt man auch in diesen Theorien wieder an Grenzen.

Die wohl populärste dieser Erweiterungen ist die Gruppe der supersymmetrischen Theorien. Die Annahme der Supersymmetrie ist, dass zu jedem bosonischen Teilchen ein fermionisches Partnerteilchen existiert und umgekehrt. Dank dieser Annahme ist es möglich, mehrere Probleme des Standardmodells zu lösen. Zum Beispiel wird so das Hierarchie-Problem gelöst und eine Vereinheitlichung der starken und elektroschwachen Kräfte wäre möglich. Weiterhin wird durch die Einführung des supersymmetrischen Teilchens mit der geringsten Masse ein weiterer Beitrag zur dunklen Materie erklärt. Leider besitzen diese supersymmetrischen Theorien oft sehr viele freie Parameter.

Eine andere Theorie, die ein Kandidat für die große Vereinheitlichung ist, ist die Stringtheorie. Sie geht davon aus, dass die Welt aus mehr als den bisher bekannten vier Raum-Zeit-Dimensionen besteht. Jedoch werden die Mechanismen der Unbeobachtbarkeit der zusätzlichen Dimensionen ad hoc und ohne experimentelle Unterstützung eingeführt.

Auch das Axionen-Modell könnte einen Teil der dunklen Materie erklären. Das Axion ist ein massives Teilchen, welches nur an Photonen koppelt. Solche Axionen könnten daher in der Sonne produziert und durch Photonen nachgewiesen werden, welche durch den umgekehrten Produktionsmechanismus entstehen. Ein solcher Versuch, das Axion zu entdecken, wird im CAST<sup>2)</sup>-Experiment am CERN<sup>3)</sup> realisiert.

Um herauszufinden, welche der neuen Theorien die Natur am besten beschreibt, sind verschiedene Experimente mit sehr hohen Energien notwendig. Daher wird dieser Bereich der Physik auch als Hochenergiephysik bezeichnet. Nur durch aufwändige und präzise Messungen wird es möglich sein, die Theorien zu verifizieren oder zu falsifizieren.

Einige dieser Experimente werden in naher Zukunft am CERN bei Genf in der Schweiz stattfinden. Zu diesem Zweck wurde der Tunnel des ehemaligen LEP-Beschleunigerrings<sup>4)</sup> umgebaut und der Proton-Proton-Beschleunigerring LHC<sup>5)</sup> installiert. An dessen Wechselwirkungspunkten werden mehrere Experimente stattfinden. Das Hauptaugenmerk dieser Arbeit liegt dabei auf dem

---

<sup>2</sup>CERN Axion Solar Telescope

<sup>3</sup>Conseil Européenne pour la Recherche Nucléaire

<sup>4</sup>Large Electron-Positron Collider

<sup>5</sup>Large Hadron Collider

---

ATLAS-Experiment<sup>6)</sup>, welches in Kapitel 4 näher beschrieben wird. Ein Ziel des Experimentes ist, neue Teilchen zu entdecken. Zum Beispiel kann dort die Frage nach der Existenz von supersymmetrischen Teilchen mit Massen auf der TeV-Skala geklärt werden. Weiterhin, und dies ist eines der Hauptziele des Experimentes, kann die Existenz des im Standardmodell postulierten Higgs-Bosons nachgewiesen werden. Die Entdeckung wird es ermöglichen, einige der unzähligen neuen Theorien und Erweiterungen des Standardmodells in bestimmten Energiebereichen auszuschließen.

Ob eine Entdeckung des Higgs-Bosons tatsächlich möglich ist, hängt von dessen Eigenschaften ab. So sagen manche Theorien vorher, dass ein Nachweis am LHC vielleicht nicht möglich sein wird. Ein solches Modell ist das Phion-Modell [1], welches in Kapitel 3 vorgestellt wird. Hier werden zusätzlich zu den im Standardmodell beschriebenen Teilchen weitere skalare Teilchen eingeführt, die sogenannten Phionen. Diese Phionen koppeln ausschließlich an das Higgs-Boson und wechselwirken sonst nicht: Es handelt sich um eine Singlett-Erweiterung des Standardmodells. Diese Art der Erweiterung ist besonders einfach und besitzt nur wenige Parameter, die angepasst werden müssen, um die physikalischen Beobachtungen zu beschreiben. Die Motivation des Phionen-Modells ist es, dass damit ein möglicher Zerfall des Higgs-Bosons in nicht nachweisbare Teilchen, im Folgenden „unsichtbarer Zerfall“ genannt, erklärt werden kann. Die Phionen sind nämlich unsichtbar für jeden Detektor, da sie weder stark, noch elektroschwach wechselwirken. Ein weitere Konsequenz der Unsichtbarkeit der Phionen ist, dass durch sie das Vorhandensein der dunklen Materie erklärt werden kann, falls sie eine Masse besitzen. Wären die Phionen jedoch masselos, so könnten sie der Ursprung der dunklen Energie oder wenigstens ein Teil davon sein.

Einer der wenigen Parameter des Phion-Modells ist die Kopplungsstärke zwischen dem Higgs-Boson und den Phionen, welche nahezu beliebig groß werden kann. Dadurch wird es dem Higgs-Boson ermöglicht, mit nahezu beliebig großer Breite in Phionen zu zerfallen. Je breiter die Higgs-Resonanz ist, desto schwieriger wird es sein, diese aus dem Untergrund herauszufiltern und durch die Zerfallsprodukte des Higgs-Bosons zu vermessen. Hieraus ergibt sich die Frage, ob ein solches Higgs-Boson im Phion-Modell beim ATLAS-Experiment noch nachgewiesen werden kann. Ein weiteres Problem ergibt sich aus dem unsichtbaren Zerfall des Higgs-Bosons, wodurch dieses nicht direkt beobachtet und seine Resonanz nicht vermessen werden kann. Statt seines direkten Nachweises über die Zerfallsprodukte wird daher ein Zählexperiment durchgeführt: Jeder signifikante Überschuss an Ereignissen mit der Signatur des Signals über dem Standardmodell-Untergrund wird als Signalereignis gewertet. Durch diese Art des Nachweises ist die Abhängigkeit des Zählexperimentes von der Resonanzbreite auch nicht so groß, wie sie es bei einer direkten Messung wäre. Eine geringe Abhängigkeit von der Breite ergibt sich aus einer eventuellen Veränderung der Produktions-Wirkungsquerschnitte, die jedoch auch bei einem sichtbaren Zerfall auftreten würde, und der Tatsache, dass jeder Nachweis eines unsichtbar zerfallenden Higgs-Bosons auf dem Nachweis von fehlender transversaler Energie aufbaut.

Der Frage, ob ein Higgs-Boson im Phion-Modell bei ATLAS nachgewiesen werden kann, wird für den assoziierten Higgs-Produktionskanal, in dem das Higgs-Boson zusammen mit einem Z-

---

<sup>6)</sup>A Toroidal LHC ApparatuS

Boson erzeugt wird, und für die Vektorboson-Fusion in Kapitel 7 nachgegangen.

In vielen anderen Modellen zerfällt das Higgs-Boson hauptsächlich in nachweisbare Teilchen. Meist wird dann zum Nachweis die Tatsache genutzt, dass für Massen des Higgs-Bosons, welche kleiner als ca.  $135 \text{ GeV}/c^2$  sind, dieses hauptsächlich in b-Quarks zerfällt. Auch bei Prozessen, in denen t-Quarks auftreten, entstehen b-Quarks. In über 99% aller Fälle zerfällt ein solches t-Quark durch die schwache Wechselwirkung in ein b-Quark. Daher ist die Markierung von b-Quarks für das Experiment sehr wichtig. Diese Markierung wird dadurch begünstigt, dass wegen einigen Eigenschaften des b-Quarks die b-Jets<sup>7)</sup> im Detektor von den Jets leichter Quarks unterschieden werden können. So legen b-Mesonen oder b-Baryonen eine relativ weite Strecke im Detektor zurück, bevor sie durch die schwache Wechselwirkung hauptsächlich in c-Quarks zerfallen. Diese schwachen Zerfälle weisen eine hohe Anzahl an geladenen Spuren auf, wodurch der Zerfallsvertex sehr genau bestimmt werden kann. Mit Hilfe des Zerfallsvertex und des Primärvertex des Prozesses wird die Fluglänge des b-Hadrons berechnet, wodurch es als solches identifiziert werden kann. Wird zusätzlich zu diesem Zerfallsvertex noch der Zerfallsvertex des c-Quarks aus dem b-Quark-Zerfall rekonstruiert, so können die b-Quarks noch effektiver nachgewiesen werden.

Ein anderer Ansatz, einen b-Jet zu identifizieren, ist die Markierung mit Hilfe niederenergetischer Elektronen oder Myonen. In dieser Arbeit werden die niederenergetischen Elektronen untersucht. Diese entstehen durch den schwachen Zerfall des b-Quarks oder des dadurch entstandenen c-Quarks und besitzen im Mittel einen hohen transversalen Impuls relativ zur Jet-Achse und eine niedrige Energie. Wird ein solches Elektron innerhalb eines Jets nachgewiesen, so deutet dies auf einen b-Quark-Zerfall hin. Allerdings ist diese Methode schon aus physikalischen Gründen nicht sehr effektiv. Denn das Zerfallsverhältnis von b-Quarks in ein Elektron beträgt nur 10%. Auch der eventuelle Zerfall des aus dem b-Quark-Zerfall entstandenen c-Quarks endet nur in etwa 8% aller Fälle in einem Endzustand mit einem Elektron. Der Grund, weshalb die Markierung eines b-Jets mit Hilfe der niederenergetischen Elektronen dennoch wichtig ist, ist, dass dieser Nachweis keine Informationen über die Lebenszeit des b-Quarks verwendet. Daher sind die Markierung durch den Nachweis der Elektronen und die Markierung unter Verwendung der Lebenszeit sehr unabhängig voneinander, wodurch eine gegenseitige Kalibration und Überprüfung anhand von Daten, also ohne Informationen aus Monte-Carlo-Simulationen, ermöglicht wird. Weiterhin kann durch eine Kombination beider Methoden das Verhältnis von Nachweisefizienz und Reinheit weiter verbessert werden.

Eine Einführung in die Physik und die Identifikation von Elektronen aus dem b-Quark-Zerfall geben der Abschnitt 2.6 und das Kapitel 8. In Kapitel 8 wird weiterhin beschrieben, wie die bisherigen Verfahren durch neue Variablen und neue Methoden, welche in Abschnitt 6 beschrieben werden, verbessert wurden. In Kapitel 9 werden diese Ergebnisse zur Identifikation der Elektronen schließlich auf die Markierung von b-Quarks angewandt.

---

<sup>7)</sup>Ein Jet ist die Menge aller Endprodukte der Hadronisation eines Quarks oder Gluons nach einer Teilchenkollision. Diese Hadronisation findet statt, da in der Natur aufgrund der Eigenschaften der starken Wechselwirkung keine freien Quarks existieren (Diese Tatsache wird auch „confinement“ genannt).

## 2 Das Standardmodell der Teilchenphysik

Das Standardmodell der Teilchenphysik ist die Theorie, welche die beobachteten Teilchen, deren Eigenschaften und alle bekannten Wechselwirkungen bis auf die Gravitation bisher am besten beschreibt. Die im Standardmodell beschriebenen Teilchen sind in Tabelle 2.1 aufgelistet. Die Bausteine der uns bekannten Materie sind die Leptonen und Quarks. Gemeinsame Eigenschaft dieser Teilchen ist, dass sie allesamt einen halbzahligen Spin besitzen. Daraus folgt, dass die quantenmechanischen Wellenfunktionen bezüglich des Austauschs zweier Teilchen der Fermi-Dirac Statistik gehorchen: Die Gesamtwellenfunktion zweier Teilchen ändert bei Vertauschen dieser das Vorzeichen. Teilchen mit dieser Eigenschaft werden Fermionen genannt. Zwischen den Fermionen werden Kräfte ausgetauscht, welche durch weitere Teilchen vermittelt werden. Diese Teilchen besitzen ganzzahlige Spins, gehorchen daher der Bose-Einstein Statistik und werden Bosonen genannt.

Wie diese Kräfte zwischen den Fermionen zustande kommen und auf welche Weise der Austausch der Bosonen eingeführt wird, soll nun beschrieben werden. Weiterhin wird gezeigt, wie die Fermionen und Bosonen ihre Masse erhalten und welche Rolle das am LHC gesuchte Higgs-Boson dabei spielt.

### 2.1 Grundlagen

Die dem Standardmodell zu Grunde liegende Theorie ist die Quantenfeldtheorie, die eine Verbindung der Quantenmechanik und der speziellen Relativitätstheorie ist. In einer Quantenfeldtheorie werden Teilchen durch quantisierte Felder beschrieben. Hieraus folgt, dass Teilchen vernichtet und erzeugt werden können. Die Wechselwirkungen und Eigenschaften der Teilchen werden mit Hilfe des Lagrange-Formalismus beschrieben: Die Lagrange-Dichte  $L$  eines Feldes ist durch

$$L = T - V \tag{2.1}$$

gegeben, wobei  $T$  die kinetische Energie des Feldes und  $V$  seine potentielle Energie ist. Für ein skalares Feld  $\phi$  mit den Parametern  $x_\mu$  und der Masse  $m$  sieht die Lagrange-Dichte ohne Wechselwirkung folgendermaßen aus [2]:

$$L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi^* + \frac{1}{2} m \phi^* \phi. \tag{2.2}$$

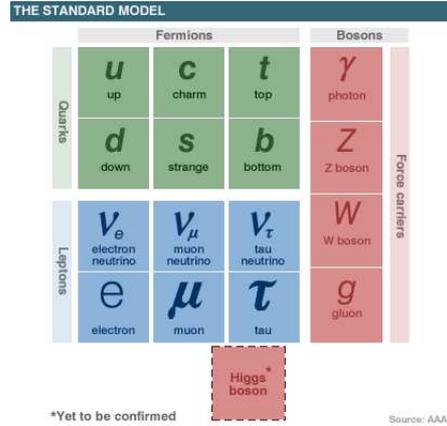


Abbildung 2.1: Die im Standardmodell bekannten Teilchen. Oben die drei Familien der Quarks, darunter die Leptonen, welche allesamt Fermionen sind. Rechts sind die Träger der Wechselwirkungskräfte abgebildet, die Eichbosonen. Unten ist das postulierte Higgs-Boson aufgelistet, nach dem immer noch gesucht wird. Diese Abbildung wurde der Referenz [3] entnommen.

Die Bewegungsgleichung dieses Teilchens kann über die Variation der Wirkung dieser Dichte berechnet werden:

$$\partial_\mu \left( \frac{\partial L}{\partial(\partial_\mu \phi)} \right) - \frac{\partial L}{\partial \phi} = 0. \quad (2.3)$$

Diese sind für ein freies Bosonenfeld  $\phi$  die Klein-Gordon-Gleichung

$$(\partial_\mu \partial^\mu + m^2)\phi = 0, \quad (2.4)$$

bzw. die Dirac-Gleichung

$$(\gamma_\mu \partial^\mu - m)\psi = 0 \quad (2.5)$$

für ein freies Fermionenfeld  $\psi$  mit der Lagrangedichte

$$L = \bar{\psi}(i\gamma_\mu \partial^\mu - m)\psi. \quad (2.6)$$

Die Matrizen  $\gamma_\mu$  sind hier die Dirac'schen Matrizen und  $\bar{\psi}$ , definiert als

$$\bar{\psi} = \psi^\dagger \gamma_0, \quad (2.7)$$

ist das zu  $\psi$  adjungierte Feld.

Ist die Lagrange-Dichte für Teilchen und deren Wechselwirkungen bekannt, so lassen sich auch die Wirkungsquerschnitte für alle möglichen Prozesse mit Hilfe der Störungsrechnung berechnen. Bei der Berechnung von Schleifenkorrekturen, d.h. Korrekturen höherer Ordnung, treten jedoch häufig Divergenzen auf, die zunächst einmal unphysikalisch erscheinen, da das gemessene Ergebnis endlich ist. Dieser scheinbare Widerspruch lässt sich lösen, indem man die

Massen und Kopplungskonstanten, die nur als Parameter in der Theorie auftreten, neu definiert. Dieser Prozess wird Renormierung genannt. Auf diese Weise werden die auftretenden Divergenzen in den Parametern absorbiert. Zwar sind die Parameter nun unendlich groß, jedoch wird im Experiment stets nur der Unterschied zwischen zwei Parameterwerten gemessen. Weiterhin stellt sich dadurch heraus, dass viele Parameter selbst von der Energie eines Prozesses abhängig sind. So sind die Kopplungskonstanten der schwachen, starken und elektromagnetischen Wechselwirkung Funktionen der Energie eines Prozesses. Dies ist auch der Grund dafür, dass ein Versuch, diese drei Wechselwirkungen ab einer gewissen Energie mit einer einzigen Theorie zu beschreiben, überhaupt möglich ist. Da die Parameter von der Energie abhängen, ist es notwendig, den Startpunkt für die Skala festzulegen. Der Startpunkt dieser Renormierungsskala wird Renormierungsmasse genannt.

Insgesamt kann festhalten werden, dass eine Theorie nur sinnvolle, also endliche Ergebnisse liefern kann, wenn sie renormierbar ist. M. Veltman und G. 't Hooft zeigten hierzu, dass eine Theorie immer renormierbar ist, wenn sie lokal eichinvariant ist. Für diese Arbeit [4] erhielten beide im Jahr 1999 den Nobelpreis. Das Prinzip der lokalen Eichinvarianz stellt ein zentrales Element der heutigen theoretischen Physik dar und soll im folgenden erklärt werden.

## 2.2 Das Prinzip der Eichinvarianz

Das Prinzip der Eichinvarianz besagt, dass sich die Physik unter bestimmten Transformationen nicht ändert. So darf das Ergebnis eines Experimentes zum Beispiel nicht vom Ort oder dem Zeitpunkt der Durchführung abhängen. Die zugehörigen Transformationen sind die Translation im Ort und die Zeit-Translation. Für die Beschreibung der Invarianz der Physik mit Hilfe des Lagrange-Formalismus bedeutet dies, dass die Lagrange-Dichte sich unter dieser Transformation nicht ändern darf. Mit dieser Invarianz der Lagrange-Dichte ist immer die Erhaltung eines Stroms verbunden. So führt die Invarianz unter Translation im Ort zum Erhalt des Impulses eines Systems, die Invarianz unter Zeit-Translation zur Energieerhaltung.

Ein weiteres einfaches Beispiel für eine solche Transformation ist die Phasentransformation eines skalaren Feldes:

$$\phi \rightarrow \phi e^{i\alpha}. \quad (2.8)$$

Der erhaltene Strom  $j^\mu$  ergibt sich dann aus der infinitesimalen Form dieser Transformation und der Forderung nach der Invarianz der Lagrange-Dichte 2.2 zu

$$\partial_\mu j^\mu = \partial_\mu \left( \frac{i\alpha}{2} \right) \left( \frac{\partial L}{\partial(\partial_\mu \phi)} \phi - \phi^* \frac{\partial L}{\partial(\partial_\mu \phi^*)} \right) = 0. \quad (2.9)$$

Eine *lokale* Eichtransformation ist eine solche Transformation, bei der nun der Parameter  $\alpha$  in Abhängigkeit des Raum-Zeitpunktes  $x$  transformiert wird:

$$\phi \rightarrow \phi e^{i\alpha(x)}. \quad (2.10)$$

Wird nun die Lagrange-Dichte 2.6 für ein Fermionfeld  $\psi$  betrachtet, so stellt man fest, dass die lokale Eichinvarianz nur dann gegeben ist, wenn die Ableitung mittransformiert wird. Denn

durch eine Transformation der Form 2.10 folgt für den kinetischen Term der Lagrangedichte 2.6

$$\partial_\mu \psi \rightarrow e^{i\alpha(x)} \partial_\mu \psi + ie^{i\alpha(x)} \psi \partial_\mu \alpha(x). \quad (2.11)$$

Dies bedeutet, dass der  $\partial_\mu \alpha(x)$ -Term die lokale Eichinvarianz bricht. Wird nun die Ableitung durch ihre kovariante Form gemäß

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - ieA_\mu \quad (2.12)$$

ersetzt und setzt man das Transformationsverhalten von  $A_\mu$  fest als

$$A_\mu \rightarrow A_\mu + \frac{1}{e} \partial_\mu \alpha, \quad (2.13)$$

so ergibt sich die lokale Eichinvarianz der so erhaltenen Lagrange-Dichte

$$L = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi + e\bar{\psi}\gamma^\mu \psi A_\mu. \quad (2.14)$$

Durch die Forderung nach lokaler Eichinvarianz der Lagrangedichte wird auf diese Weise eine Wechselwirkung auf natürlichem Wege eingeführt: Das Feld  $A_\mu$ , welches Eichfeld genannt wird, koppelt mit der Stärke  $e$  an das Fermionenfeld. Studiert man die Eichfelder  $A_\mu$  genauer, so stellt man fest, dass diese Bosonen sind, also die Eichbosonen, die die Kräfte zwischen den Fermionen austauschen.

Interessant ist nun, dass in der Lagrange-Dichte 2.14 ein Massenterm für das Eichfeld  $A_\mu$  die lokale Eichinvarianz verletzen würde. Mit dieser Theorie ist es also nicht möglich, ein massebehaftetes Eichboson zu realisieren, ohne die Renormierbarkeit zu verlieren. Dass die bekannten Eichbosonen teilweise eine Masse besitzen, wurde aber schon in verschiedenen Experimenten nachgewiesen. Auch die Tatsache, dass die Reichweite einer Wechselwirkung direkt mit der Masse der Eichbosonen zusammenhängt <sup>1)</sup>, lässt schon auf das Vorhandensein von Eichbosonen mit Masse schließen. Abhilfe schaffen nun eine Erweiterung um ein weiteres Feld und ein mathematischer Trick, welche 1964 von P. Higgs eingeführt und nach ihm benannt wurden [5].

## 2.3 Der Higgs-Mechanismus

Wie dieser sogenannte Higgs-Mechanismus funktioniert, soll am einfachen Beispiel eines komplexen skalaren Feldes  $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$  mit der Lagrange-Dichte

$$L = (\partial_\mu \phi)^* \partial^\mu \phi - \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2 \quad (2.15)$$

erläutert werden.  $\mu$  ist ein Parameter, der für  $\mu^2 > 0$  als Masse interpretiert werden kann,  $\lambda$  beschreibt die Stärke der Selbstkopplung des Feldes ( $\lambda > 0$ ). Diese Lagrange-Dichte ist invariant

---

<sup>1)</sup>Dieser Zusammenhang kann mit dem Yukawa-Potential  $e^{-mr}/r$  beschrieben werden.

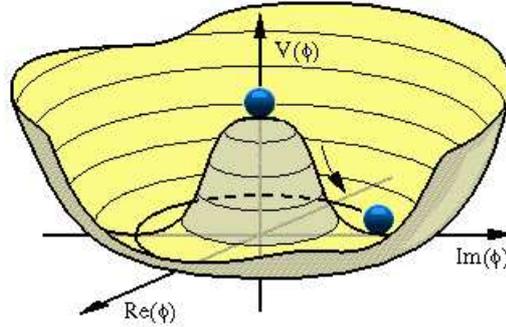


Abbildung 2.2: Das Potential  $V = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2$  eines komplexen skalaren Feldes, wie das des Higgs-Feldes, für  $\mu^2 < 0$  und  $\lambda > 0$ . Diese Abbildung wurde der Referenz [6] entnommen.

unter der  $U(1)$ -Transformation <sup>2)</sup> 2.8. Soll nun auch Invarianz unter der lokalen Transformation 2.10 gelten, so muss wiederum die partielle Ableitung durch die kovariante Ableitung 2.12 ersetzt und ein Eichfeld  $A_\mu$  mit 2.13 eingeführt werden. Für die Lagrange-Dichte ergibt dies

$$L = D_\mu \phi^* D^\mu \phi - \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2 - \text{kin. Term von } A_\mu. \quad (2.16)$$

Wird nun  $\mu^2 < 0$  gewählt, so kann dieser Parameter nicht mehr als Masse interpretiert werden. Stattdessen ergibt sich für das Potential

$$V = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2 \quad (2.17)$$

eine Form, wie sie in Abbildung 2.2 dargestellt ist.

Der Kreis der Minima dieses Potentials in der  $\phi_1 - \phi_2$ -Ebene wird durch  $\phi_1^2 + \phi_2^2 = v^2$ , mit  $v^2 = -\mu^2/\lambda$ , bestimmt. Wird das Feld  $\phi$  um ein solches Minimum  $v_0$  herum entwickelt, so ergibt sich eine neue Form

$$\phi(x) = \sqrt{\frac{1}{2}} v_0 + \eta(x) + i\theta(x) \quad (2.18)$$

für das entwickelte Feld. Für dieses neue Feld sieht das Potential nicht mehr symmetrisch aus. Man sagt daher, die Symmetrie sei durch die Wahl eines speziellen Grundzustandes spontan gebrochen worden.

Aus der neuen Form 2.18 des Feldes  $\phi$  folgt für die Lagrange-Dichte

$$L = \frac{1}{2} (\partial_\mu \eta)^2 + \frac{1}{2} (\partial_\mu \theta)^2 - v_0^2 \lambda \eta + \frac{1}{2} e^2 v_0^2 A_\mu A^\mu - e v_0 A_\mu \partial^\mu \theta - \dots \quad (2.19)$$

Betrachtet man diese Lagrange-Dichte genauer, so stellt man fest, dass durch die Wahl des neuen Grundzustandes des Feldes  $\phi$  das Eichboson  $A_\mu$  eine Masse  $m = e v_0$  erhalten hat. Dies ist allerdings so geschehen, dass die lokale Eichinvarianz der Lagrangedichte weiterhin gewährleistet ist.

<sup>2</sup>U steht für die Gruppe der Unitären Transformationen. Die Elemente dieser Gruppe zeichnen sich durch die Eigenschaft  $U^\dagger U = 1$ , mit  $U \in U(N)$ , aus.

Daher ist eine Theorie, bei der die massiven Teilchen ihre Masse durch den Higgs-Mechanismus erhalten, eine renormierbare Theorie mit vernünftigen Ergebnissen.

Zusätzlich zum massiven Eichboson erhält man ein massives skalares Feld  $\eta$  und ein masseloses skalares Feld  $\theta$ . Durch die spontane Brechung einer Symmetrie wird immer mindestens ein solches masseloses skalares Teilchen erzeugt<sup>3)</sup>, ein Goldstone-Boson. Dieses Goldstone-Boson kann jedoch mit einer Eichtransformation „wegtransformiert“ werden. Es wird dann als longitudinale Polarisation des jetzt massebehafteten Eichbosons interpretiert. Dabei werden also weder Freiheitsgrade vernichtet oder erzeugt.

## 2.4 Die Theorie des Standardmodells

In obigen Fall war die Eichtransformation eine Transformation der Gruppe  $U(1)$ . Insgesamt besteht die Gesamtsymmetrie des Standardmodells aus

$$U(1) \otimes SU(2) \otimes SU(3). \quad (2.20)$$

Dabei wird die elektroschwache Theorie aus den Gruppen  $U(1)$  der Hyperladung und  $SU(2)$ <sup>4)</sup> des schwachen Isospins gebildet. Die Gruppe  $SU(3)$  beschreibt die starke Wechselwirkung.

### 2.4.1 Das Glashow-Weinberg-Salam-Modell

Diese gemeinsame Beschreibung der elektroschwachen Theorie geht auf Glashow [7], Weinberg [8] und Salam [9] zurück, wofür sie 1979 den Nobel-Preis in Physik erhielten. Dazu wird zunächst ein Fermionenfeld mit Hilfe des Projektionsoperators der Chiralität

$$P = \frac{1}{2}(1 - \gamma^5), \quad (2.21)$$

mit  $\gamma^5 = \gamma_1 \gamma_2 \gamma_3 \gamma_4$ , in eine rechtshändige und eine linkshändige Komponente aufgespalten. Die Matrizen  $\gamma_\mu$  sind wieder die Dirac'schen Matrizen. In dieser Darstellung lässt sich das Fermionenfeld  $\psi$  zerlegen in

$$\psi = \frac{1}{2}(1 - \gamma^5)\psi + \frac{1}{2}(1 + \gamma^5)\psi = \psi_L + \psi_R. \quad (2.22)$$

Für die Beschreibung der paritätsverletzenden Natur der schwachen Wechselwirkung im Standardmodell gilt nun, dass nur linkshändige Felder an der schwachen Wechselwirkung teilnehmen. Dazu werden die linkshändigen Fermionen in Doublets des schwachen Isospins angeordnet. Die rechtshändigen Felder sind Singletts unter der  $SU(2)$ -Symmetrie, das heißt sie besitzen keinen schwachen Isospin. Weiterhin sind alle diese Felder zusätzlich noch Träger der schwachen Hyperladung  $Y$ . Alle Fermionen und ihre Quantenzahlen sind in Tabelle 2.1 aufgelistet.

<sup>3</sup>Dieser Sachverhalt wird Goldstone-Theorem genannt.

<sup>4</sup> $SU$  bezeichnet die Gruppe der speziellen unitären Gruppe. Die Elemente der Gruppe  $SU(N)$  besitzen folgende Eigenschaften:  $UU^\dagger = 1$  und  $\det(U) = 1$ .

Leptonen	$I_3$	$Q$	$Y$	Quarks	$I_3$	$Q$	$Y$
$\nu_L$	1/2	0	-1	$u_L$	1/2	2/3	1/3
$l_L^-$	-1/2	-1	-1	$d_L$	-1/2	-1/3	1/3
$\nu_R$	0	0	0	$u_R$	0	2/3	4/3
$l_R^-$	0	-1	-2	$d_R$	0	-1/3	-2/3

Tabelle 2.1: Die Fermionenfelder des Standardmodells und ihre Quantenzahlen. Nur die linkshändigen Dubletts sind Träger des schwachen Isospins  $I_3$  und nehmen so an der schwachen Wechselwirkung teil.  $Q$  ist die elektrische Ladung der Fermionen. Die Quantenzahl  $Y$  bezeichnet die Hyperladung der Felder. Die rechtshändigen Felder sind Singletts unter der  $SU(2)$ -Symmetrie des schwachen Isospins. Die rechtshändigen Neutrinos wechselwirken also weder durch die elektromagnetische und schwache, noch durch die starke Wechselwirkung. Ihr Vorhandensein wird aber durch die Masse der Neutrinos vorhergesagt: Durch einen geeigneten Lorentz-Boost kann ein linkshändiges Teilchen in ein rechtshändiges Teilchen transformiert werden, sofern es eine Masse besitzt.

Durch eine Eichtransformation der Gruppe  $SU(2)$  erhält man nun drei Eichfelder, die Felder  $W^i$  mit  $i = 1, 2, 3$ . Mit einer  $U(1)$ -Eichtransformation der Hyperladung wird ein weiteres Eichfeld  $B$  eingeführt. Dies geschieht, wie oben schon beschrieben, durch Ersetzen der partiellen Ableitung durch die kovariante Ableitung:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i\frac{g}{2}\tau_i W_\mu^i + i\frac{g'}{2}YB_\mu. \quad (2.23)$$

$g'$  und  $g$  beschreiben die Kopplungen der Bosonenfelder an die Fermionenfelder. Die  $\tau_i$  sind die Generatoren der Gruppe  $SU(2)$ , von denen eine mögliche Darstellung durch die Pauli-Matrizen realisiert ist,  $Y$  ist der Generator der Gruppe  $U(1)$ .

Wird nun ein Higgs-Dublett des schwachen Isospins

$$\phi = \begin{pmatrix} \phi_\alpha \\ \phi_\beta \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (2.24)$$

mit einem Potential der Form 2.17 eingeführt, so wird die  $U(1) \otimes SU(2)$ -Symmetrie durch Wahl eines Grundzustands für das Higgs-Dublett

$$\phi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (2.25)$$

mit Vakuumerwartungswert  $v$  spontan gebrochen. Nach der Diagonalisierung der dadurch entstandenen Massenmatrix für die  $W$ - und  $B$ -Felder ergeben sich die heute bekannten massiven

Eichbosonen  $W^+$ ,  $W^-$ ,  $Z$  und das masselose Photon <sup>5)</sup>. Die Stärke der Mischung der Felder  $B$  und  $W^i$ , die dadurch entsteht, kann durch den schwachen Mischungswinkel

$$\theta_W = \sin^{-1} \left( \frac{g'}{\sqrt{g^2 + g'^2}} \right), \quad (2.26)$$

auch Weinbergwinkel genannt, beschrieben werden [2]. Mit diesem ergeben sich für die Felder der Eichbosonen folgende Zusammenhänge:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \quad (2.27)$$

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & -\sin\theta_w \\ \sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}. \quad (2.28)$$

Für die Massen der massiven Eichbosonen folgt aus diesem Prozess (siehe Referenz [2])

$$M_{W^\pm} = \frac{1}{2}vg \quad \text{und} \quad M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}. \quad (2.29)$$

Durch diese Zusammenhänge konnte aus der Messung der Massen von W- und Z-Boson der Vakuumenergie des Higgs-Feldes bestimmt werden. Für diesen gilt  $v \approx 246 \text{ GeV}/c^2$ . Die Masse des Higgs-Bosons selbst ist dennoch unbekannt, da für die Selbstkopplung des Higgs-Bosons immer noch keine genaue Vorhersage oder Messung existiert.

Da auch die Fermionen Massen besitzen, muss das Higgs-Dublett  $\phi$  auch an diese koppeln. Dies kann durch die Yukawa-Kopplung realisiert werden:

$$L_{Yukawa} = -G_a \left[ (\bar{b}, \bar{a})_L \phi a_R + \bar{a}_R \phi^\dagger \begin{pmatrix} b \\ a \end{pmatrix}_L \right]. \quad (2.30)$$

Um die Massen der oberen Komponenten  $b$  der Dubletts des schwachen Isospins zu generieren, muss in dieser Gleichung  $\phi$  durch  $-i\tau_2\phi^*$  und das rechtshändige Feld entsprechend ersetzt werden. Für das Generieren der Massen der Teilchen mit Hilfe der Yukawa-Kopplung bedeutet dies also, dass für jedes Teilchen  $a$  eine separate Kopplung  $G_a$  an das Higgs-Boson existieren muss. Diese Kopplung wird künstlich eingeführt und die Wahl der Kopplungsstärke ist zunächst willkürlich, so dass eine Vorhersage der Fermionenmassen aus der Theorie heraus nicht möglich ist.

Um auch elektroschwache Zerfälle beschreiben zu können, bei denen ein Quark eines Dubletts in ein Quark eines anderen Dubletts übergeht, müssen die echten Quarks Mischungen der Quarkfelder der elektroschwachen Theorie sein. Aus diesem Grund wurde durch Cabibbo, Kobayashi

---

<sup>5)</sup>Die Quantenzahlen der Hyperladung  $Y$  und der dritten Komponente des schwachen Isospins  $T^3$  sind für das Higgs-Dublett so gewählt, dass die U(1)-Symmetrie der elektromagnetischen Wechselwirkung ungebrochen bleibt. Hierzu muss für die Quantenzahlen des Higgs-Bosons gelten:  $Q = T^3 - Y/2 = 0$ .

und Maskawa die nach ihnen benannte CKM-Mischungsmatrix eingeführt. Diese transformiert die Quarks gemäß

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = (CKM) * \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.31)$$

und gehört der Gruppe der SU(3) an. Dadurch wird ein Übergang innerhalb der Triplets (d s b) und (u c t) verhindert <sup>6)</sup>. Weiterhin besitzt diese Matrix eine komplexe Phase, durch die die CP-Verletzung, wie zum Beispiel im neutralen Kaonensystem, ermöglicht wird. Allerdings ist die Größe der CP-Verletzung nicht ausreichend, um das Ungleichgewicht von Materie zu Anti-Materie im Universum zu erklären.

## 2.4.2 Die Quantenchromodynamik

Die Theorie zur Beschreibung der starken Wechselwirkung wird Quantenchromodynamik genannt. Ihre Symmetrie ist eine Symmetrie der SU(3)-Gruppe der Farbladung. Somit gibt es drei Farben, bzw. deren Anti-Farben, deren Träger die Quarks sind. Die Forderung nach lokaler Eichinvarianz unter einer SU(3)-Symmetrie führt zu einer kovarianten Ableitung der Form

$$D_\mu = \partial_\mu + ig_s \lambda_a G_\mu^a. \quad (2.32)$$

Hierzu mussten passend zu den acht Generatoren  $\lambda_a$  der Gruppe SU(3) die acht Gluonenfelder  $G^a$  eingeführt werden. Da die Gruppe der SU(3) allerdings nicht abelsch ist, müssen die Gluonenfelder anders transformieren als zum Beispiel das Feld  $A_\mu$  in 2.13. Die Konsequenz ist, dass die Gluonen selbst zu Trägern der Farbladung werden: Sie tragen je eine Farbe und eine Anti-Farbe und ermöglichen so den Farbaustausch zwischen den Quarks <sup>7)</sup>. Hieraus folgt weiterhin, dass sie untereinander koppeln und dadurch die starke Kopplungskonstante  $\alpha_s$  nach der Renormierung bei großen Abständen sehr stark wird. Sie wird sogar so stark, dass die potentielle Energie ausreicht, um neue Quark-Anti-Quark-Paare aus dem Vakuum zu erzeugen. Aus diesem Grund treten Quarks nicht einzeln auf, sondern hadronisieren zu gebundenen Zuständen, den farbneutralen Hadronen. Ein Hadron wird immer durch Quarks gebildet, deren Farben sich zu weiß aufheben. Diese Tatsache ist unter dem Begriff *confinement* bekannt. Daraus folgt für Mesonen, dass sie aus einem Quark mit einer Farbe und einem mit der korrespondierenden Anti-Farbe bestehen. Bei Baryonen werden immer alle drei Farben oder Anti-Farben kombiniert. Weiterhin besteht ein Hadron nicht nur aus diesen Valenzquarks. Wäre dies so, so hätte zum Beispiel jedes Quark in einem Proton genau ein Drittel der Energie des Protons. Gemessen wurde bei Elektron-Proton-Streuungen aber eine andere Verteilung. Der Grund dafür ist, dass in Hadronen ständig Kräfte durch Abstrahlung und Einfangen von Gluonen ausgetauscht werden. Weiterhin werden ständig Quark-Anti-Quark-Paare, welche dann See-Quarks genannt werden, erzeugt und vernichtet. Dadurch werden Energie- und Impulsverteilung der Quarks in einem Hadron ausgeschmiert und an

<sup>6)</sup>Dieser Mechanismus wird nach Glashow, Iliopoulos und Maiani GIM-Mechanismus genannt.

<sup>7)</sup>Die Möglichkeit, drei Farben zu kombinieren lässt zunächst auf neun Gluonen schließen. Dieses neunte Gluon wäre aber ein farbloses Singlett der Gruppe SU(3) und könnte daher keine Farbladung übertragen.

einem Streuprozess von zum Beispiel Protonen können nun auch Anti-Quarks teilnehmen. Da für diese Prozesse innerhalb eines Hadrons störungstheoretische Berechnungen nicht möglich sind, weil die starke Kopplungskonstante bei diesen Prozessen zu groß wird, müssen die Impulsverteilungen der Quarks innerhalb eines Hadrons gemessen und parametrisiert werden. Diese Quark-Dichtefunktionen  $f_q(x)$  können dann als Wahrscheinlichkeit interpretiert werden, dass zum Beispiel in einem Proton  $p$  ein Quark  $q$  mit einem Anteil  $x = E_q/E_p$  der Protonenenergie existiert. Für die Berechnung des Wirkungsquerschnitts einer Elektron-Proton-Streuung folgt daraus

$$\sigma(pe^- \rightarrow X) = \sum_q \int_0^1 f_q(x) \sigma(qe^- \rightarrow X) dx. \quad (2.33)$$

Hier muss über alle Quarktypen summiert und über ihren Impulsanteil  $x$  integriert werden, da nicht bekannt ist, welches Quark mit welchem Impulsanteil an der Streuung teilgenommen hat. Die Messung der Quark-Dichtefunktionen geschieht am einfachsten durch die schon oben erwähnte Elektron-Proton- und Elektron-Neutron-Streuung. Da diese Messungen jedoch bei anderen Energien als bei denen am LHC stattfinden und die Quark-Dichtefunktionen energieabhängig sind, müssen diese Dichte-Funktionen auf die neuen Energien extrapoliert werden. Dies geschieht mit Hilfe der Entwicklungsgleichung von Altarelli und Parisi. Als einfaches Beispiel ändert sich die Quark-Dichte  $f$  eines Quarks mit der Energie  $Q$  wie

$$\frac{d}{d \log Q^2} f(x, Q^2) = \frac{\alpha_s}{2\pi} \int_x^1 q(y, Q^2) P_{qq} \left( \frac{x}{y} \right) \frac{dy}{y}. \quad (2.34)$$

Diese Gleichung beschreibt die Tatsache, dass ein Quark eines Hadrons mit dem Bruchteil  $x$  der Energie des Hadrons durch Gluonabstrahlung von einem Quark mit Bruchteil  $y$  der Hadronenergie entstanden ist. Die Wahrscheinlichkeit, dass genau dies geschieht, ist proportional zu  $\alpha_s P_{qq}(x/y)$ . Diese Gleichung muss entsprechend erweitert werden, berücksichtigt man die Gluon-Spaltung in Gluonen oder Quark-Anti-Quark-Paare.

Auch der Prozess der Hadronisation beinhaltet die Abstrahlung vieler Gluonen und die Produktion vieler Quark-Anti-Quark-Paare. Und auch dieser Prozess kann nicht mit der Störungstheorie beschrieben werden, da die starke Kopplungskonstante zu groß wird. Stattdessen werden so genannte Fragmentationsfunktionen der Quarks eingeführt, mit denen das Hadronisationsverhalten der Quarks parametrisiert werden kann. Diese Fragmentationsfunktionen,  $D_q^h(z)$ , können als Wahrscheinlichkeit interpretiert werden, dass ein bestimmtes Hadron  $h$  in einem Jet eines Quarks  $q$ , welches einen Anteil  $z = E_h/E_q$  der Energie des Quarks trägt, gefunden wird. Für die Berechnung von Wirkungsquerschnitten folgt also

$$\sigma(q_1 q_2 \rightarrow hX) = \sum_q \int_0^1 \sigma(q_1 q_2 \rightarrow q \bar{q}) [D_q^h(z) + D_{\bar{q}}^h(z)] dz, \quad (2.35)$$

wenn beachtet wird, dass nichts über die ursprüngliche Energie und Quantenzahlen des produzierten Quarks  $q$  bekannt ist. Die Fragmentationsfunktionen können zum Beispiel bei Elektron-Positron-Kollisionen vermessen werden, bei denen Jets entstehen.

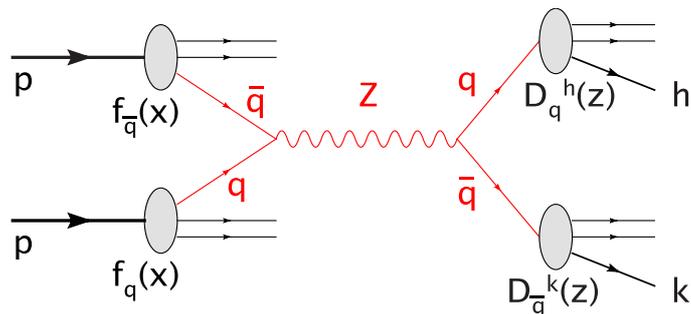


Abbildung 2.3: Ein Beispiel für einen Prozess, bei dem die Quark-Dichtefunktionen und die Fragmentationsfunktionen zur Berechnung des Wirkungsquerschnitts verwendet werden. In Rot ist der harte Streu-Prozess eingezeichnet, dessen Wirkungsquerschnitt mit Hilfe der Störungstheorie berechnet werden kann. Zusätzlich muss noch die Bremsstrahlung berücksichtigt werden, da die Quarks vor der Fusion oder Fragmentation Eichbosonen abstrahlen können. Dies ist in der Abbildung nicht dargestellt.

Die Abbildung 2.3 zeigt einen Feynman-Graphen, der einen Proton-Proton-Streuprozess mit Fragmentationsfunktionen und Quark-Dichtefunktionen darstellt.

## 2.5 Wichtige Higgs-Produktionskanäle und Higgs-Zerfälle

Um das im vorherigen Abschnitt eingeführte Higgs-Boson in einem Experiment entdecken zu können, müssen einige Eigenschaften des Higgs-Bosons bekannt sein und ausgenutzt werden. Anhand von Abbildung 2.4 (links) erkennt man, dass bei den Energien, bei denen der Large Hadron Collider operiert, einige Higgs-Produktionsprozesse besonders erfolgversprechend für eine Entdeckung sind. Besonders die Gluonenfusion besitzt über das ganze Massenspektrum hinweg einen relativ hohen Wirkungsquerschnitt. Dadurch kann es möglich sein, schon mit geringer Anzahl an Kollisionen die Existenz des Higgs-Bosons nachzuweisen. Ein weiterer potentieller Entdeckungskanal ist die assoziierte  $t\bar{t}$ -Higgs-Produktion. Diese beiden Produktionsmechanismen werden durch die Feynman-Graphen in Abbildung 2.5 dargestellt. Weitere wichtige Produktionsmechanismen sind vor allem die Vektorboson-Fusion und die assoziierte Higgs-Produktion. Denn auch die Vektorboson-Fusion besitzt einen sehr konstanten Wirkungsquerschnitt und die Prozesse der assoziierten Higgs-Produktion lassen sich dank der beteiligten Leptonen besonders leicht triggern und selektieren. Auf diese beiden Kanäle, deren Feynman-Graphen in Abbildung 2.6 gezeigt sind, wird später in den Unterkapiteln 7.2 und 7.3 noch genauer eingegangen. Da das Higgs-Boson an die Masse der Teilchen koppelt, sind nur direkte Zerfälle in massive Teilchen erlaubt und besonders schwere Teilchen werden bevorzugt. Hier allerdings sind auch

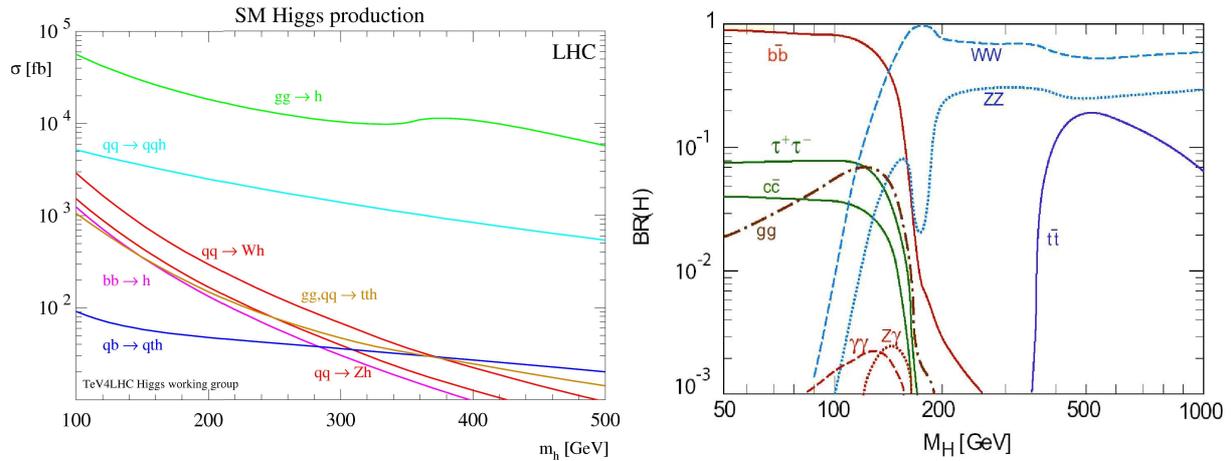


Abbildung 2.4: Verschiedene Higgs-Produktionsmechanismen und ihre Wirkungsquerschnitte (links), sowie die Verzweigungsverhältnisse der Higgs-Boson-Zerfälle als Funktionen der Higgs-Masse (rechts). Quellen der Abbildungen sind die Referenzen [10] und [11].

nur solche Prozesse erlaubt, die die Quantenzahlen der elektroschwachen und der starken Theorie erhalten. Das bedeutet, dass nur Paare aus Teilchen und Anti-Teilchen entstehen können. In Gebieten kleinerer Higgs-Masse sind dies bevorzugt  $b\bar{b}$ -Paare (siehe hierzu die rechte Seite der Abbildung 2.4), in den hohen Massenbereichen, bei denen  $m_{Higgs} \simeq 2 * m_W$  oder höher ist, dominiert der Zerfall in  $W^+W^-$ -Paare. Der Zerfall in Gluonen und Photonen wird über eine Quark-Schleife ermöglicht.

## 2.6 Physik des b-Quark-Zerfalls

Wie gezeigt wurde, entstehen b-Quarks aus dem Higgs-Bosonenzerfall besonders oft für kleine Higgs-Massen. Doch nicht nur beim direkten Zerfall des Higgs-Bosons können b-Quarks entstehen. Jede t-Quarkproduktion endet in über 99% aller Fälle in einem Zerfall des t-Quarks in ein b-Quark <sup>8</sup>). Dies bedeutet zum Beispiel auch, dass in nahezu allen Ereignissen mit einer assoziierten  $t\bar{t}$ -Higgs-Produktion b-Jets auftreten. Weiterhin ist wichtig, dass die b-Quarks praktisch nicht in Hadronisationsprozessen entstehen. Kann in einem Ereignis also ein b-Jet identifiziert werden, so entstand dieser in fast allen Fällen direkt aus der Kollision der Protonen, einem t-Quark- oder einem Higgs-Boson-Zerfall.

Nach seiner Produktion hadronisiert das b-Quark und bildet ein b-Meson oder b-Baryon. Aufgrund der harten Fragmentationsfunktionen des b-Quarks tragen diese b-Mesonen oder b-Baryonen den größten Teil der Energie des b-Quarks mit sich und haben eine typische Lebens-

<sup>8</sup>Für den zugehörigen Eintrag  $V_{tb}$  in der CKM-Matrix gilt  $V_{tb} \simeq 0.999$  [12].

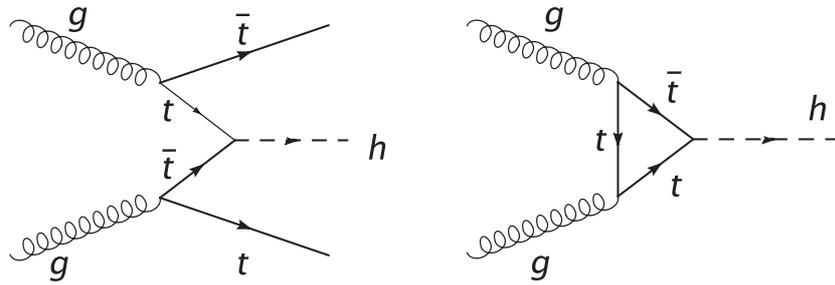


Abbildung 2.5: Feynman-Graphen zur assoziierten  $t\bar{t}$ -Higgs-Produktion (links) und zur Gluon-Fusion (rechts).

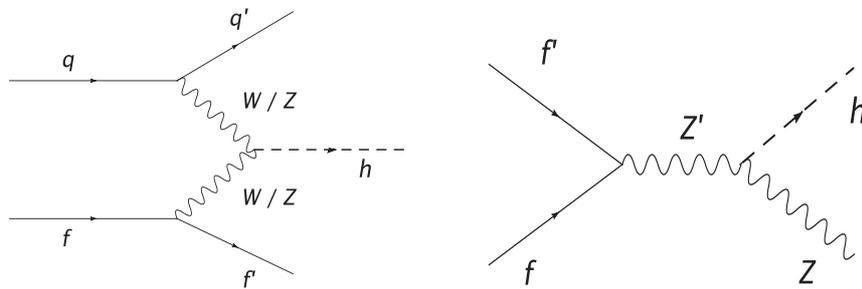


Abbildung 2.6: Feynman-Graphen zur Vektorboson-Fusion (links) und zur assoziierten Higgs-Produktion mit einem  $Z$ -Boson (rechts).

dauer von  $\tau \simeq 1.5$  ps. Dadurch legen sie eine Strecke von einigen Millimetern bis hin zu Zentimetern im Detektor zurück, bevor das am Hadron beteiligte b-Quark in ein c-Quark oder in seltenen Fällen in ein u-Quark zerfällt. Für die mittlere Flugstrecke eines Teilchens mit einem relativistischen Boost  $\beta\gamma$  gilt

$$\langle l \rangle = \beta\gamma c \tau \approx \mathcal{O}(\text{mm} - \text{cm}). \quad (2.36)$$

Der Grund für die hohe Lebensdauer ist, dass der Zerfall über die schwache Wechselwirkung stattfindet und die CKM-Matrixelemente für die Kopplung vom b-Quark an das c- und u-Quark sehr klein sind. Diese CKM-Matrixelemente sind  $V_{cb} \simeq 42.21 \times 10^{-3}$  und  $V_{ub} \simeq 3.96 \times 10^{-3}$ . Daher ist der Zerfall in ein c-Quark stark bevorzugt.

Das W-Boson, welches beim schwachen Zerfall erzeugt wird, zerfällt meist in zwei Quarks, die dann hadronisieren und Jets bilden. Die Zerfallskette eines b-Mesons ist graphisch in Abbildung 2.7 dargestellt.

Der Nachweis eines b-Jets gelingt also am besten direkt über die Lebensdauer des Zerfalls des b-Hadrons. Durch die hohe Anzahl an geladenen Spuren<sup>9)</sup> beim Zerfall eines b-Hadrons, welche durchschnittlich von der Größe fünf ist, gelingt es, den Zerfallsvertex des b-Hadrons zu

<sup>9)</sup>Diese Anzahl wird auch Multiplizität genannt.

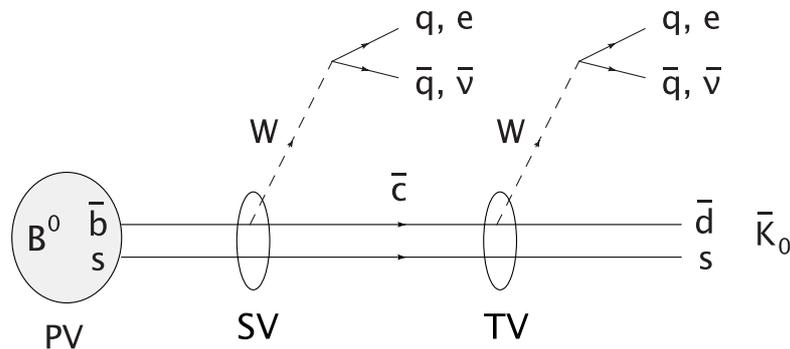


Abbildung 2.7: Gezeigt ist die Zerfallskette des b-Mesons  $B^0$ . Dieses wird am Primärvertex (PV) durch die Hadronisation eines  $\bar{b}$ -Quarks produziert. Nach einer Flugstrecke von einigen Millimetern bis hin zu Zentimetern zerfällt das im Meson enthaltene  $\bar{b}$ -Quark am Sekundärvertex (SV) durch einen schwachen Zerfall in ein  $\bar{c}$ -Quark. Dieses  $\bar{c}$ -Quark kann wiederum schwach in ein  $\bar{d}$ -Quark zerfallen. Der Zerfallsort des  $\bar{c}$ -Quarks wird Tertiärvertex (TV) genannt und liegt meist zu nahe am Sekundärvertex, um ihn getrennt rekonstruieren zu können.

bestimmen. Mit Hilfe dieses Vertex und dem Primärvertex des Prozesses kann die Flugstrecke des Mesons oder Baryons berechnet werden. Mit dieser Flugstrecke lassen sich die b-Mesonen oder b-Baryonen als solche identifizieren, denn die leichten Hadronen zerfallen entweder gar nicht, oder schon nach sehr kurzer Flugstrecke. Nur ein c-Hadron lebt lange genug, um eine ähnliche, doch meist geringere Flugstrecke zurückzulegen. Daher ist die Unterscheidung von b-Hadronen und c-Hadronen die schwierigste Aufgabe bei der Markierung der b-Quarks. Hierzu ist es hilfreich, wenn zusätzlich der Zerfallsvertex des aus dem Zerfall des b-Hadrons entstandenen c-Hadrons bestimmt werden kann. Dies gelingt in manchen Fällen, wodurch das Verhältnis aus Nachweiseffizienz und Reinheit für b-Jets noch einmal verbessert wird. Durch das oben beschriebene Verfahren in Verbindung mit der weiten Flugstrecke der b-Hadronen ist das b-Quark das am leichtesten zu identifizierende Quark. Daher werden für Analysen sehr oft Kanäle verwendet, in denen b-Quarks auftreten.

Doch dies ist nicht die einzige Methode, b-Quarks nachzuweisen. Auch durch den Zerfall des am Hadron beteiligten b-Quarks, bei dem das W-Boson in ein Elektron oder Myon zerfällt, kann ein b-Jet identifiziert werden. Da Elektronen oder Myonen nicht durch die Hadronisation gebildet werden können, deutet das Vorhandensein eines Elektrons oder Myons innerhalb eines Jets auf einen semileptonischen Zerfall eines Quarks hin. Im Folgenden wird nur der Zerfall des W-Bosons in ein Elektron und ein Elektron-Neutrino berücksichtigt.

Einen b-Jet mit Hilfe eines solchen Elektrons nachzuweisen, ist allerdings nicht sehr effizient. Nur in etwa 10% aller Fälle<sup>10)</sup> entsteht ein Elektron direkt aus dem Zerfall eines b-Hadrons. Zusätzlich zu diesen Elektronen kommen noch diejenigen, die aus dem Zerfall des eventuell

<sup>10)</sup>Zum Beispiel ist nach Referenz [12]  $\Gamma(B^0 \rightarrow l^+)/\Gamma_{B^0,gesamt} \simeq 10.4\%$  und  $\Gamma(B^\pm \rightarrow l^\pm)/\Gamma_{B^\pm,gesamt} \simeq 10.9\%$

entstandenen c-Hadrons entstehen. Die Wahrscheinlichkeit für diesen Zerfall liegt nach Referenz [12] bei ca.  $\Gamma(c \rightarrow l^+)/\Gamma_{c,gesamt} \simeq 9.6\%$ . Zusammen mit der Tatsache, dass pro Zerfall eines b-Hadrons im Durchschnitt  $n_c = 1.15$  c-Quarks entstehen [13], erhält man eine gesamte Wahrscheinlichkeit von

$$P = BR(b \rightarrow e + X) + n_c \cdot BR(c \rightarrow e + X) = 10.4\% + 1.15 \cdot 9.6\% \simeq 22\%, \quad (2.37)$$

ein Elektron in einem b-Jet zu finden. Gelänge es also, jeden b-Jet, in dem ein Elektron aus einer der oben genannten direkten oder indirekten Quellen vorkommt, in allen Fällen als solchen nachzuweisen, so könnten dennoch nur ca. 22% aller b-Jets gefunden werden. Schlimmer wird die Situation dadurch, dass der Nachweis solcher Elektronen recht schwierig und ineffizient ist. Vor allem die enorme Anzahl von bei der Hadronisierung erzeugten Pionen in Jets, die den Hauptuntergrund für die Elektronenidentifikation darstellen, erschwert diese.

Die Hauptmotivation für die Markierung von b-Jets mit Hilfe dieser Elektronen ist diejenige, dass zum Nachweis fast ausschließlich Informationen über die Schauer des Elektrons aus den Kalorimetern benötigt werden. Alle anderen Nachweismethoden basieren hauptsächlich auf den Informationen aus dem Spurdetektor. Somit sind diese Verfahren und die Markierung mit Hilfe der Elektronen sehr unabhängig voneinander. Dadurch ist es möglich, diese Methoden effektiv miteinander zu kombinieren und dadurch die Nachweiseffizienz für b-Jets zu erhöhen. Weiterhin können die auf der Lebensdauer basierenden Methoden mit Hilfe der Markierung von b-Quark durch die Elektronen anhand von echten Daten kalibriert werden. Durch den Vergleich dieser verschiedenen Algorithmen kann eine größere Unabhängigkeit von den Monte-Carlo-Simulationen und von theoretischen Modellen erreicht werden.

Die Identifikation von Elektronen aus dem Zerfall eines b-Hadrons oder des daraus entstandenen c-Hadrons erfolgt durch ihre spezifischen Eigenschaften.

- Die Elektronen aus dem Zerfall eines b-Hadrons oder des daraus entstandenen c-Hadrons sind nicht isoliert <sup>11)</sup>, wie zum Beispiel Elektronen aus einem W-Zerfall oder Z-Zerfall, sondern befinden sich nahe oder innerhalb eines Jets. Weiterhin besitzen sie wenig Energie im Vergleich zu solchen isolierten Elektronen. Daher werden diese Elektronen auch niederenergetische Elektronen genannt.
- Der transversale Impuls relativ zur Jet-Achse ist für die Elektronen aus dem Zerfall von b-Hadronen allerdings größer als der der Elektronen, die zum Beispiel aus einem Zerfall eines c-Hadrons stammen. Dies liegt an der großen Masse des zerfallenden b-Quarks, wodurch der transversale Impuls in Bezug auf die Jetachse bis zu  $p_T^{max.} \approx m_b/2$  betragen kann.
- Um die Elektronen von Pionen zu unterscheiden, werden die Schauerformen in den Kalorimetern betrachtet. Da die Pionen eine viel größere Masse besitzen, unterscheiden sich ihre Schauerformen von denen der Elektronen und reichen bis in das hadronische Kalorimeter.

Durch diese Eigenschaften gelingt es, die Elektronen aus dem Zerfall von b-Hadronen oder der daraus entstandenen c-Hadronen von anderen Teilchen oder Elektronen aus anderen Quellen

<sup>11</sup>Isolierte Elektronen sind Elektronen, in deren direkter Umgebung keine hadronische Aktivität zu finden ist.

zu unterscheiden. Wie die Identifikation dieser Elektronen beim ATLAS-Experiment momentan durchgeführt wird und welche Verbesserungen noch möglich sind, wird in Kapitel 8 beschrieben.

### 3 Eine Erweiterung des Standardmodells: Das Phion-Modell

Das Phion-Modell ([14], [1], [15]) ist eine Singlet-Erweiterung des Standardmodells um skalare Teilchen, welche Phionen genannt werden. Die Motivation, das Standardmodell auf diese Weise zu erweitern ist vielfältig. Zum Einen handelt es sich dabei um die einfachste mögliche Erweiterung des Standardmodells. Weiterhin sind diese skalaren Phionen ein guter Kandidat für die dunkle Materie, falls sie eine Masse besitzen. Falls sie keine Masse besitzen, könnte mit ihnen teilweise das Vorhandensein von dunkler Energie erklärt werden, die den Hauptteil der Energiedichte unseres Universums darstellt. Außerdem bewirkt die geringe Zahl der freien Parameter, dass dieses Modell ohne viele ad hoc eingeführte Annahmen auskommt und der Parameterraum, den es experimentell einzuschränken gilt, sehr überschaubar bleibt. Im Gegensatz zu anderen Theorien, wie der String-Theorie, werden hier auch keine neuen Dimensionen eingeführt.

Mathematisch können der hier relevante Teil der Lagrange-Dichte des Standardmodells und die Erweiterung durch das Phion-Modell in drei Teile aufgeteilt werden.

$$L_{Skalar} = L_{Higgs} + L_{Phion} + L_{Wechselwirkung}. \quad (3.1)$$

Diese Teile sind zum Einen der Higgs-Term des Standardmodells, welcher die normale Kinematik des Higgs-Feldes und dessen Selbstwechselwirkung beschreibt. Mit  $\phi$  als Higgs-Dublett, der Selbstwechselwirkungsstärke  $\lambda$  und dem Vakuumserwartungswert  $v$  des Higgs-Feldes lautet dieser explizit

$$L_{Higgs} = -\frac{1}{2}\partial_\mu\phi^\dagger\partial^\mu\phi - \frac{\lambda}{8}(\phi^\dagger\phi - v^2)^2. \quad (3.2)$$

Zusätzlich wird die Existenz weiterer skalarer Teilchen, der Phionen, angenommen, die ebenfalls mit sich selbst wechselwirken. Die Anzahl  $N$  der möglichen Phionenarten sowie die Stärke  $\lambda_S$  derer Selbstwechselwirkung sind freie Parameter dieser Theorie. Der kinematische und der Selbstwechselwirkungsbeitrag zur Lagrange-Dichte werden beschrieben durch

$$L_{Phion} = -\frac{1}{2}\partial_\mu\vec{S}\partial^\mu\vec{S} - \frac{1}{2}m_s^2\vec{S}^2 - \frac{\lambda_S}{8N}(\vec{S}^2)^2, \quad (3.3)$$

wobei  $\vec{S}$  der Vektor der  $N$  Phionen ist. Der wichtige Aspekt des Phion-Modells ist die Kopplung der Phionen an das Higgs-Boson mit einer Kopplungsstärke  $\omega$ :

$$L_{Wechselwirkung} = -\frac{\omega}{4\sqrt{N}}\vec{S}^2\phi^\dagger\phi. \quad (3.4)$$

Da die Phionen nur an sich selbst und das Higgs-Boson koppeln, ist es in diesem Modell möglich, dass das Higgs-Boson in die experimentell nicht nachweisbaren Phionen zerfällt. Man spricht daher von einem „unsichtbaren Zerfall“ des Higgs-Bosons. Die zugehörige Zerfallsbreite hängt dabei von der Kopplungsstärke  $\omega$  und der Anzahl  $N$  der Phionenarten ab. Da diese beiden Parameter nahezu frei wählbar sind, kann die Zerfallsbreite des Higgs-Bosons in unsichtbare Phionen und damit auch seine Gesamtbreite beliebig groß werden. Durch die Wahl eines Vakuumserwartungswertes einiger Phionen-Felder ungleich von Null, ist es in diesem Modell außerdem auch möglich, dass die Phionen-Felder und das Higgs-Feld mischen. Dies führt nach der Diagonalisierung der Massenmatrix zu mehreren realen Teilchen und somit mehreren und kleineren Maxima im Massenspektrum des Higgs-Bosons, beziehungsweise der Higgs-Bosonen. Das Mischen und der unsichtbare Zerfall können in diesem Modell auch gemeinsam auftreten. Dies kann zu einer beliebigen Form der Massenverteilung des Higgs-Bosons mit beliebigem Zerfallsverhältnis in unsichtbare Phionen führen, was die zukünftigen Experimente vor eine große Herausforderung stellen könnte.

Für die spätere Analyse wurde angenommen, dass das Zerfallsverhältnis des Higgs-Bosons in Phionen 100% beträgt, was für große Breiten des Higgs-Bosons eine gute Näherung ist. Weiterhin soll keine Mischung von Phionen und dem Higgs-Boson stattfinden. Die Masse der Phionen wurde zu Null gesetzt. Dies stellt allerdings keine Einschränkung des Modells dar, da für große  $N$  die Masse der Phionen vernachlässigt werden kann. Für diesen Fall mit  $N \rightarrow \infty$  lässt sich die unsichtbare Zerfallsbreite des Higgs-Bosons berechnen zu [14]:

$$\Gamma_{Higgs}^{unsichtbar} = \frac{\omega^2 v^2}{32\pi m_{Higgs}} = \frac{\omega^2 (\sin\theta_w \cos\theta_w M_Z)^2}{32\pi^2 \alpha_{em} M_{Higgs}}, \quad (3.5)$$

wobei  $v$  wieder der Vakuumserwartungswert des Higgs-Feldes und  $\theta_w$  der schwache Mischungswinkel ist. Abbildung 3.1 zeigt die Variation der Breite des Higgs-Bosons als Funktion der Masse des Higgs-Bosons für verschiedene Werte von  $\omega$ .

Wie man an dieser Abbildung leicht erkennen kann, sind besonders für leichte Higgs-Bosonen enorme Änderungen der Breite um mehrere Größenordnungen möglich. Daher stellt sich die Frage, ob ein solch breites und unsichtbar zerfallendes Higgs-Boson beim ATLAS-Experiment überhaupt beobachtet werden kann. Für zwei potentielle Entdeckungskanäle, die Vektorbosonen-Fusion und die assoziierte Produktion des Higgs-Bosons mit einem Z-Boson, wird diese Frage in Kapitel 7 untersucht werden. Eine ähnliche Studie zum Entdeckungspotential für den  $t\bar{t}$ -Produktionskanal existiert bereits [16].

Grenzen für die Gültigkeit des Phion-Modells wurden in den Referenzen [1, 17–19] aufgezeigt, in denen auch weitere Informationen über das Phion-Modell gefunden werden können.

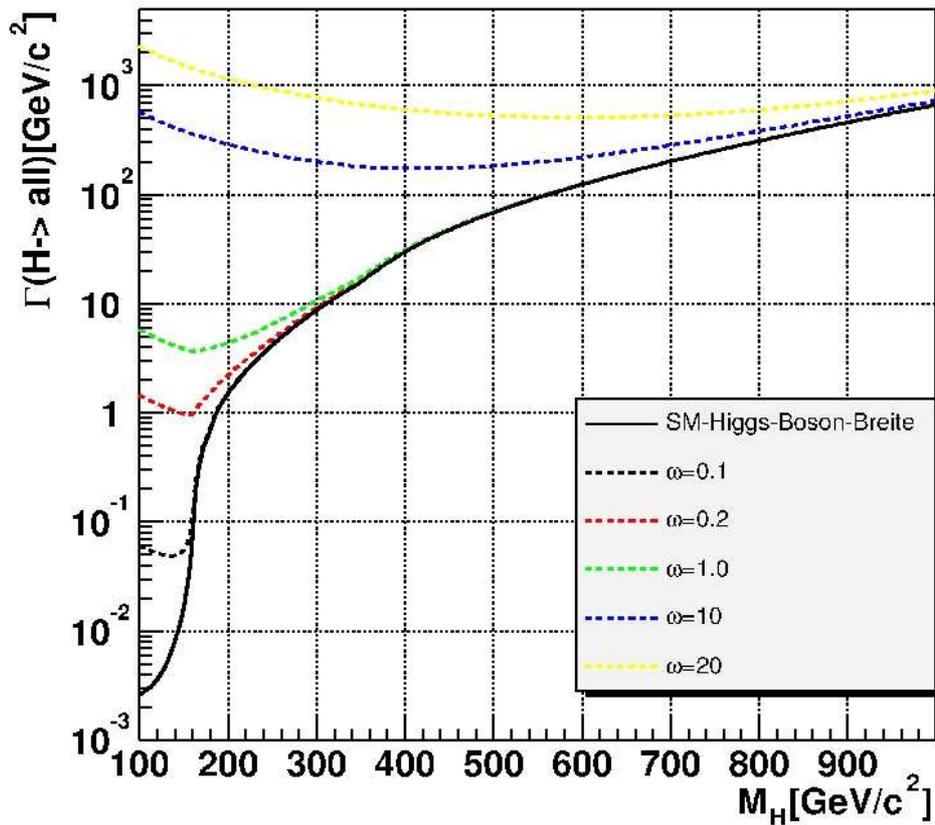


Abbildung 3.1: Zerfallsbreite des Higgs-Bosons als Funktion der Masse des Higgs-Bosons für verschiedene Werte der Kopplung  $\omega$  im Kontinuumfall ( $N \rightarrow \infty$ ) und für masselose Phionen. Eine mögliche Mischung von Phionen und Higgs-Boson wird hier nicht berücksichtigt. Diese Abbildung wurde der Referenz [16] entnommen.



## 4 Das ATLAS-Experiment

Um die Natur der kleinsten Bausteine der Materie zu entschlüsseln, betrachtet man Kollisionen von Teilchen. Aus den Endprodukten solcher Kollisionen lassen sich Rückschlüsse auf die abgelaufenen Prozesse und Wechselwirkungen ziehen. Um neue Physik zu erforschen, ist es notwendig, den Teilchen immer mehr Energie zur Verfügung zu stellen. Nur dadurch können immer schwerere Teilchen produziert und nachgewiesen werden. Eine Möglichkeit, dies zu bewerkstelligen, ist mit dem Large Hadron Collider (LHC), welcher in Abbildung 4.1 schematisch dargestellt ist, realisiert. Dieser Proton-Proton-Speicherring mit einem Umfang von rund 27 km liegt etwa 10 km von Genf entfernt. Mit einer Schwerpunktsenergie von 14 TeV und einer Luminosität von  $10^{33} \frac{1}{\text{cm}^2\text{s}}$  in den ersten drei Jahren und  $10^{34} \frac{1}{\text{cm}^2\text{s}}$  in der darauffolgenden Zeit wird er vermutlich eine genauere Einsicht in die Natur der Elementarteilchen bieten können. Die Luminosität ist definiert als

$$L = \frac{N_p^2 N_B f}{4\pi \sigma_x \sigma_y}. \quad (4.1)$$

$\sigma_x$  und  $\sigma_y$  sind hier die Strahlbreiten in x-, bzw. y-Richtung,  $N_B$  ist die Anzahl an Teilchenpaketen. Solche Teilchenpakete, bestehend aus  $N_p \simeq 10^{11}$  Protonen, werden sich alle  $f^{-1} = 25$  ns an den Wechselwirkungspunkten treffen. Die hohe Luminosität des LHC wird es ermöglichen, in den ersten Jahren Daten zu sammeln, die einer integrierten Luminosität von 30-60  $\text{fb}^{-1}$  entsprechen, und auf Dauer selbst Prozesse mit sehr kleinen Wirkungsquerschnitten genau untersuchen zu können. Der Nachteil dieser hohen Luminosität ist allerdings, dass Prozesse während eines Zusammenstoßes von den Teilchenpaketen nicht einzeln stattfinden, sondern im Mittel 23 inelastische Wechselwirkungen gleichzeitig. Dieser Effekt wird im Englischen als „pile up“ bezeichnet.

Der Grund, Protonen zu verwenden, ist, dass die große Masse der Protonen zu einer geringen Synchrotronstrahlung<sup>1)</sup> und somit zu einem geringeren Energieverlust führt, als es mit einer Beschleunigung von Elektronen mit deutlich weniger Energie möglich wäre. Dadurch ist die Maximalenergie der Protonen alleine durch die Stärke und Konstruktion der Magnete des LHC limitiert, welche die Protonen auf der Kreisbahn halten sollen. Geplant ist, dass die 1232 Dipolmagnete jeweils ein Magnetfeld der Stärke 8.3 T erzeugen sollen, wodurch die Schwerpunktsenergie von 7 TeV pro Strahl ermöglicht wird. Durch die Verwendung von Protonen entstehen allerdings auch Nachteile, denn ihre Kollisionen sind nicht rein wie die von Leptonen. Ein Proton selbst besteht nämlich aus drei Valenzquarks, vielen Seequarks und Gluonen. Daher ist bei einem Prozess nicht bekannt, welche der Konstituenten der Protonen bei einer Proton-Proton-Kollision wechselwirken und welche Energie diese besitzen. Weiterhin wird eine Analyse dadurch erschwert, dass, wie oben schon erwähnt, immer mehrere Kollisionen und Prozesse gleichzeitig stattfinden.

---

<sup>1)</sup>Die Synchrotronstrahlung eines Teilchens ist umgekehrt proportional zur Ruhemasse hoch vier.

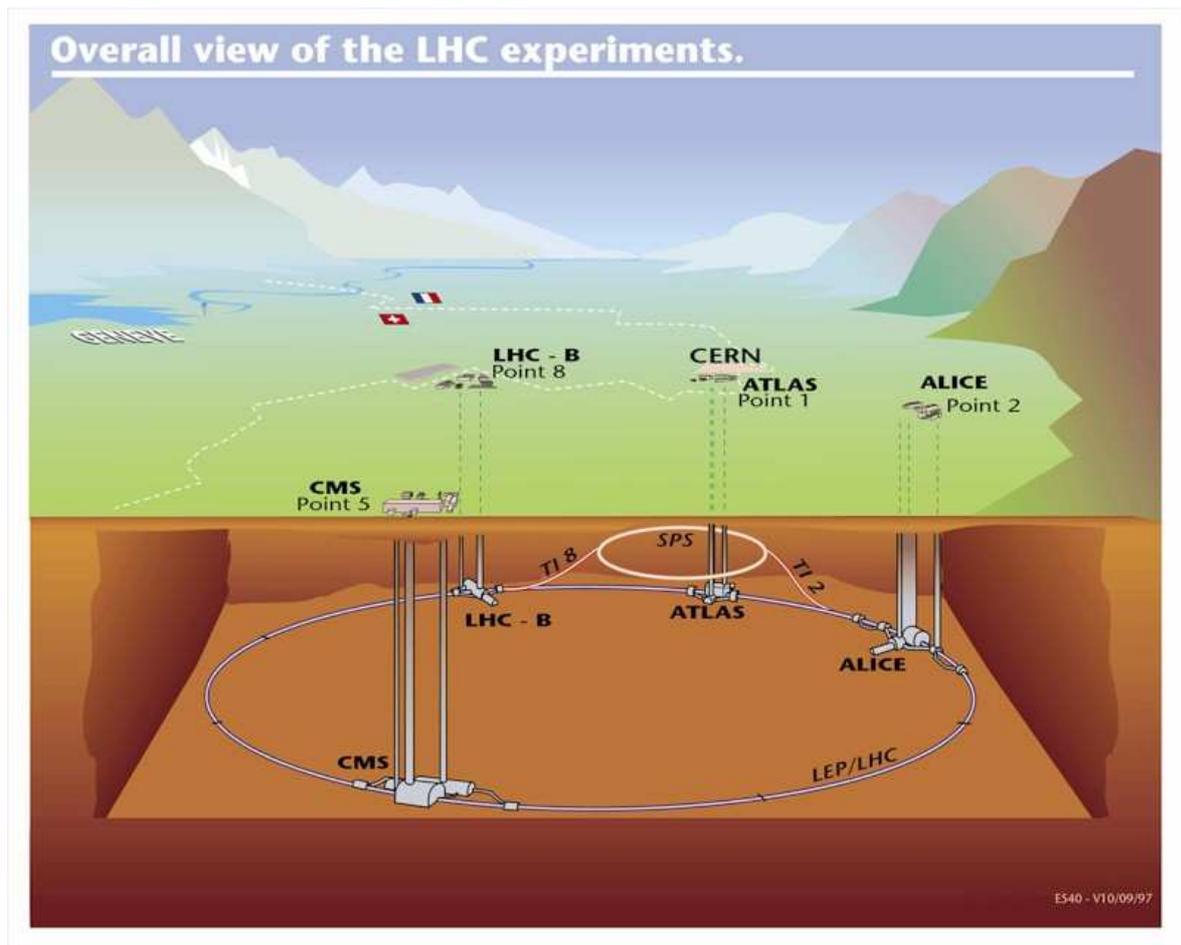


Abbildung 4.1: Schematische Abbildung des Large Hadron Collider in der Nähe von Genf. In einer Tiefe von knapp 100 m liegt dieser Beschleunigerring mit einem Umfang von ungefähr 27 km.

## 4.1 Der ATLAS-Detektor

Der ATLAS-Detektor<sup>2)</sup> ist als Allzweckdetektor konzipiert und wird an einem der vier Wechselwirkungspunkte des LHC installiert. Aufgebaut ist der Detektor aus zwei Teilen, dem zylindrischen Zentraldetektor, dem so genannten „Barrel“, und den Endkappen, die den Vorwärtsbereich abdecken. Das Gewicht des Detektors wird etwa 7000 Tonnen betragen, die auf eine Länge von 46 m und einen Durchmesser von 22 m verteilt sind. Im zwiebelartigen Aufbau des Barrels wechseln sich die Schichten aus verschiedenen Detektorkomponenten ab. Im Inneren, nahe des Wechselwirkungspunkts, liegen die präzisen Spurdetektoren, weiter außen die Kalorimeter. Diesen Aufbau und die Schichten illustriert Abbildung 4.2. Im Folgenden wird für die Beschreibung des Detektors auf folgendes Koordinatensystem zurückgegriffen:

- Der Wechselwirkungspunkt stellt den Ursprung dar.
- Die Strahlachse beschreibt die z-Achse des Systems.
- Über den zur Strahlachse gehörenden Polarwinkel  $\theta$  kann die Pseudorapidität  $\eta$  eingeführt werden:

$$\eta = -\ln(\tan\theta). \quad (4.2)$$

- Der Azimutalwinkel in der transversalen Ebene zur Strahlachse wird im Folgenden mit  $\phi$  bezeichnet.

### 4.1.1 Der Spurdetektor

Die innerste Schicht des Detektors bildet der Spurdetektor. Dieser ist in Abbildung 4.3 dargestellt. Zweck dieses Detektorteils ist die Messung des transversalen Impulses und der Richtung der geladenen Teilchen im Endzustand ab einem transversalen Impuls von  $p_t > 1 \text{ GeV}/c$ . Zu diesem Zweck wird ein Magnetfeld der Stärke 2 T erzeugt. Über die Krümmung der Spuren im Detektor kann der transversale Impuls der geladenen Teilchen bestimmt werden. Der Spurdetektor selbst beginnt nahe am Wechselwirkungspunkt mit dem Silizium-Pixeldetektor, um genaue Daten über die Spuren der geladenen Teilchen zu erhalten und den Primärvertex des Prozesses genau bestimmen zu können. Um den Silizium-Pixeldetektor herum ist der Silizium-Streifendetektor „SCT“<sup>3)</sup> angebracht, welcher vom Übergangsstrahlungsdetektor, dem „TRT“<sup>4)</sup>, umgeben ist. Dieser Übergangsstrahlungsdetektor besteht aus Gas-Zählrohren mit 4 mm Durchmesser. Mit Hilfe des TRT wird es möglich sein, Elektronen von Pionen zu unterscheiden. Um das starke Magnetfeld im Spurdetektor zu erzeugen, ist er von einem Solenoidmagnetsystem umgeben. Dieses erzeugte Magnetfeld muss genau bekannt sein, um eine gute Impulsauflösung zu gewährleisten.

<sup>2)</sup>Detaillierte Beschreibungen des Detektors findet man zum Beispiel in [20] oder [21]

<sup>3)</sup>SemiConductor Tracker

<sup>4)</sup>Transition Radiation Tracker

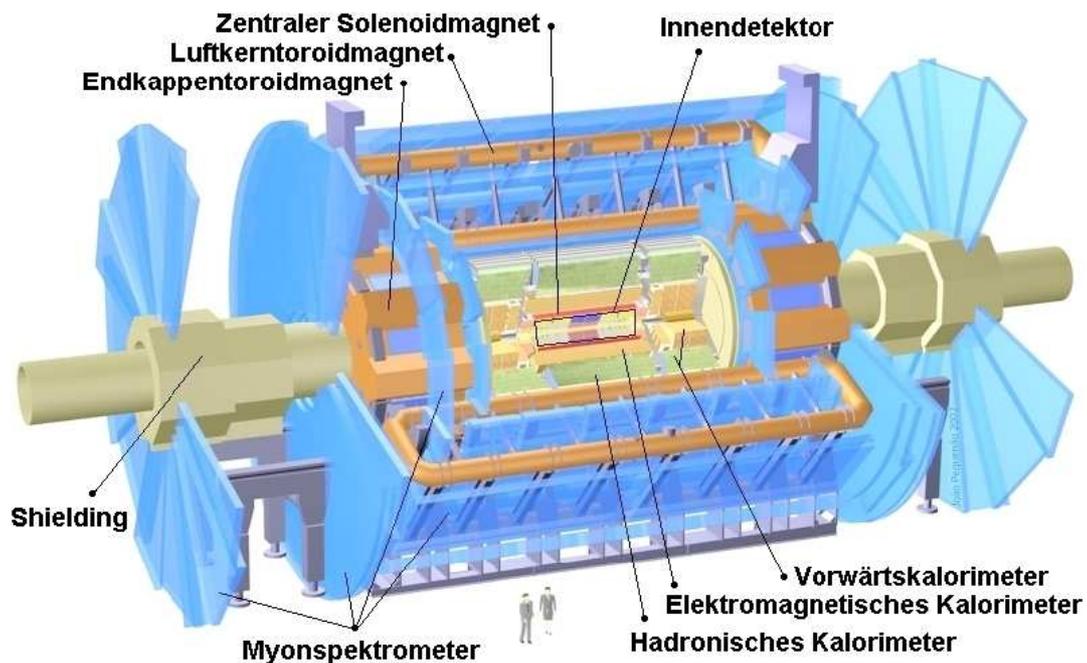


Abbildung 4.2: Schematischer Aufbau des ATLAS-Detektors. Um den Wechselwirkungspunkt sind die Komponenten von Innen nach Außen folgendermaßen angeordnet: Spurdetektor, welcher in Abbildung 4.3 genauer gezeigt wird, elektromagnetisches Kalorimeter, hadronisches Kalorimeter und als Abschluss der Myonendetektor mit seinen Magnetfeldspulen. An den Enden des Detektors werden die Endkappen angebracht, mit denen dann ein möglichst großer Raumwinkel bis zu  $\eta = 5.0$  abgedeckt wird.

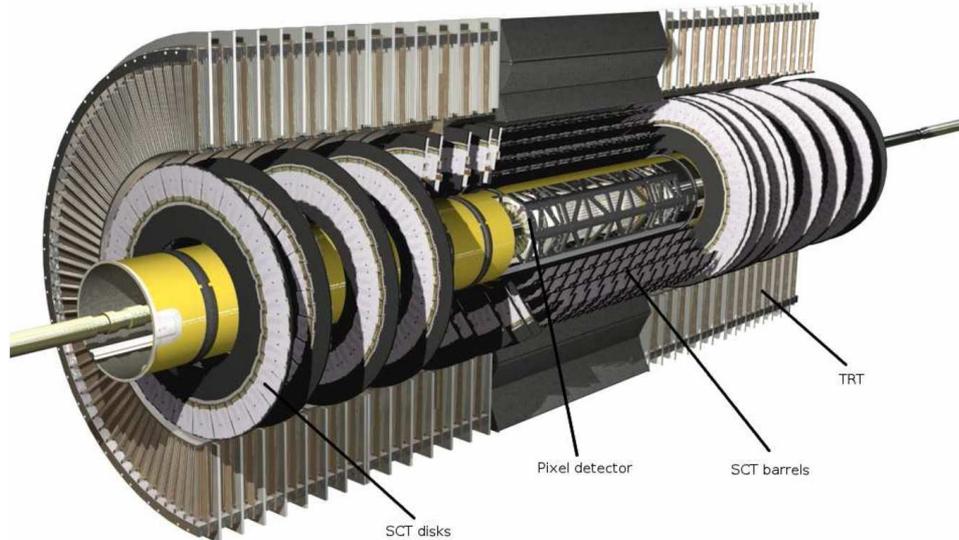


Abbildung 4.3: Schematischer Aufbau des Spurdetektors des ATLAS-Detektors. Von innen nach außen sind die Schichten der Pixeldetektor, der Silizium-Streifen Detektor (SCT), der Übergangstrahlungsdetektor (TRT) und das Magnet-system, welches hier nicht abgebildet ist.

### 4.1.2 Das Kalorimetersystem

Die nächsten Schichten des Detektors bilden die Kalorimeter. Die erste Lage bildet der so genannte „Presampler“, der die Unterscheidung von Elektronen und Jets und die Unterscheidung von Photonen und neutralen Pionen ermöglicht. Nach dem Presampler folgt das elektromagnetische Kalorimeter. Dieses muss verschiedensten Anforderungen genügen. So sollen Energien von einigen GeV bis hin zu TeV aufgelöst werden können. Auch für die Auflösung der Position eines Teilchens spielt das elektromagnetische Kalorimeter eine wichtige Rolle. Hierzu ist es sehr feinkörnig aufgebaut. Dies ermöglicht es, einzelne Teilchen in Jets voneinander zu unterscheiden und die Schauer dieser Teilchen genau zu vermessen. Auch der Rauschpegel der Elektronik muss so gering wie möglich gehalten werden, um eine gute Energieauflösung zu erhalten. Aufgebaut ist das elektromagnetische Kalorimeter aus drei Schichten von Flüssig-Argon „Sampling“-Kalorimetern. Ein Sampling besteht aus mehreren Schichten von Absorbermaterial, Bleiplatten, und aktivem Material, welches aus flüssigem Argon besteht. Diese Schichten sind in einer akkordeonartigen Struktur zusammengesetzt, um Kantenbereiche zu minimieren. Das erste Sampling besteht aus dünnen Streifen und ermöglicht eine präzise Ortsauflösung in  $\eta$ . Das zweite Sampling, auch „Main-Sampling“ genannt, ist transversal zur Strahlachse in die so genannten „Tower“ mit quadratischer Grundfläche aufgeteilt. Auch das letzte Sampling, die Rückseite des elektromagnetischen Kalorimeters, besteht aus Tower, die jedoch doppelt so breit sind, wie die des Main-Samplings. Abbildung 4.4 verdeutlicht diese Struktur und den Aufbau dieses Teils

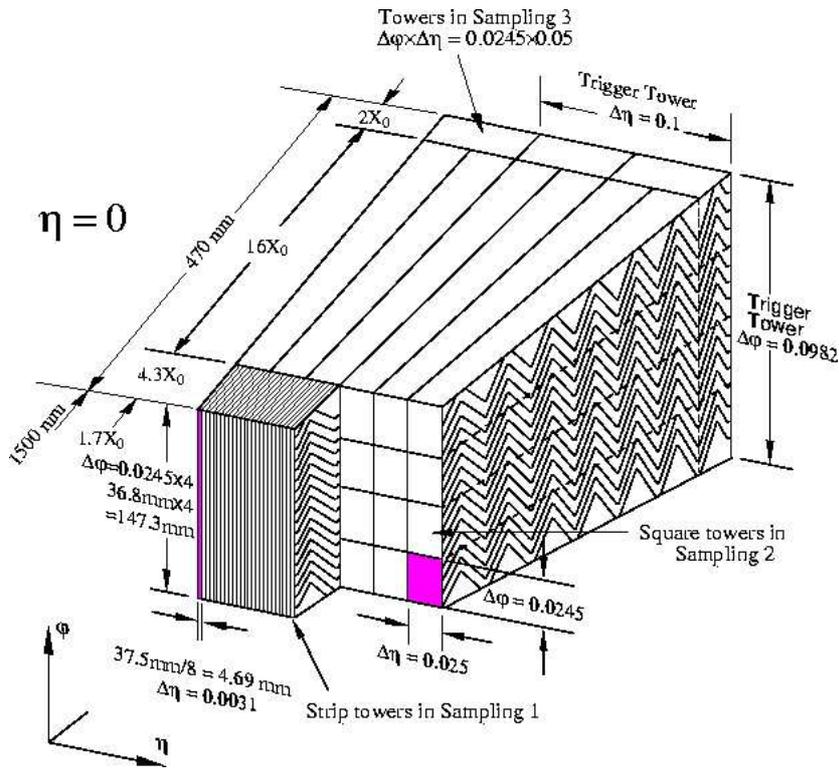


Abbildung 4.4: Schematische Abbildung des elektromagnetischen Kalorimeters des ATLAS-Detektors. Gezeigt sind die verschiedenen Schichten und die Akordeonartige Struktur des Aufbaus.

des Detektors. Die Granularität der einzelnen Komponenten ist in Tabelle 4.1 aufgelistet. Das elektromagnetische Kalorimeter soll folgende intrinsische Energieauflösung erreichen:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b \quad , \text{ mit } \begin{cases} a = 10\% , b = 0.2\% & \text{im Barrel} \\ a = 10 - 12\% , b = 0.35\% & \text{in den Endkappen} \end{cases} \quad (4.3)$$

Das hadronische Kalorimeter misst die Energie der Teilchen, welche das elektromagnetische Kalorimeter passieren. Dies sind hauptsächlich Hadronen, da diese ihre Energie durch die starke Wechselwirkung an den Detektor abgeben. Aufgebaut ist auch dieser Teil des Detektors aus mehreren Schichten. Dabei wechseln sich Absorber-Material und Szintillatoren ab. Einfliegende Teilchen wechselwirken zunächst mit dem Material, wobei auch Photonen erzeugt werden. Diese werden dann mit den Szintillatoren nachgewiesen und deren Energie gemessen. Als Absorber-Material wird dabei Eisen für das Barrel, Kupfer für das erste Forward-Sampling und die Endkappen und zusätzlich Wolfram für die weiteren Forward-Sampling-Komponenten verwendet. Als Szintillatormaterial wurde für das Forward-Sampling und das Barrel Plastikszintillatoren gewählt. In den Endkappen kommt flüssiges Argon zum Einsatz. Die unterschiedliche Wahl der Materialien ist hauptsächlich in der Strahlenbelastung der Komponenten begründet. Die Granularität der einzelnen Komponenten ist in Tabelle 4.1 aufgeführt. Die Energieauflösung des ha-

Komponente	Granularität [ $\Delta\eta \times \Delta\phi$ ]
ECAL Presampler	$0.025 \times 0.1$
ECAL Streifen	$0.003 \times 0.1$
ECAL Main-Sampling	$0.025 \times 0.025$
ECAL Rückseite	$0.05 \times 0.025$
HCAL Barrel	$0.1 - 0.2 \times 0.1$
HCAL Endkappen	$0.1 - 0.2 \times 0.1 - 0.2$
HCAL Vorwärtsrichtung	$0.2 \times 0.2$

Tabelle 4.1: Granularität des elektromagnetischen Kalorimeters (ECAL) und des hadronischen Kalorimeters (HCAL)

dronischen Kalorimeters wurde mit einem Teststrahl für einige Komponenten gemessen. Daraus ergaben sich für die Parameter  $a$  und  $b$ , welche in Formel 4.3 eingeführt wurden, folgende Werte: Im Barrel ist  $a = 45\%$  und  $b = 2.3\%$ , in den Endkappen konnten sie zu  $a = 85\%$  und  $b = 5.8\%$  bestimmt werden.

### 4.1.3 Das Myonen-System

Um Myonen, welche die Kalorimeter nahezu ungehindert durchqueren, nachweisen zu können, wurde um die Kalorimeter ein spezieller Myonen-Detektor gebaut. Dieser besteht im Zentralbereich aus Driftkammern, mit deren Hilfe die Spuren der Myonen rekonstruiert werden können. Für große Pseudorapiditäten  $\eta$  werden Kathodenstreifen-Kammern verwendet. Zusätzlich zu den Driftkammern kommen noch die „Resistive Plate Chambers“ im Zentralbereich hinzu, durch deren schnelle Antwortzeit und Zeitauflösung ein effektives „Triggern“ der Ereignisse möglich ist. In den Endkappen werden zu diesem Zweck „Thin Gap Chambers“ verwendet. Näheres zum Trigger folgt in den nächsten Abschnitten. Der Myonen-Detektor ist von einem Toroidmagneten umgeben, der durch die Erzeugung eines Magnetfeldes der Stärke 0.3 T die Spuren der Myonen krümmt. Durch diese Krümmung kann der transversale Impuls der Myonen berechnet werden.

### 4.1.4 Hermetizität des Detektors

Da bei vielen Prozessen auch Teilchen entstehen, welche vom Detektor nicht nachgewiesen werden können<sup>5)</sup>, gibt es viele Ereignisse, bei denen die fehlende transversale Energie des Ereignisses rekonstruiert werden muss. Dies geschieht, indem mit Hilfe aller sichtbaren Teilchen und der

<sup>5</sup>Außer den schon nachgewiesenen Neutrinos gibt es in vielen Erweiterungen des Standardmodells weitere nicht oder nur schwer nachweisbare Teilchen. Ein Beispiel hierfür hatten wir schon kennen gelernt: Die Phionen des Phion-Modells

Energie- und Impulserhaltung diese fehlende Energie berechnet wird. Hierfür ist es jedoch notwendig, den Detektor so hermetisch wie nur möglich zu konzipieren. Der ATLAS-Detektor deckt ein großes Gebiet in der Pseudorapidität  $\eta$  ab. Die Kalorimeter erreichen durch die Endkappen eine Abdeckung des Bereiches  $|\eta| < 5.0$ , beim Spurdetektor ist es der Bereich  $|\eta| < 2.5$ . Für die Messungen der Impulse der Myonen decken die Myonen-Kammern sogar einen Bereich von  $|\eta| < 2.7$  ab. Somit kann eine gute Rekonstruktion der fehlenden transversalen Energie erreicht werden. Die Rekonstruktion der totalen fehlenden Energie eines Prozesses ist nicht möglich, da die longitudinalen Impulse der am Prozess beteiligten Partonen nicht genau bekannt sind.

### 4.1.5 Das Triggersystem

Um mit Hilfe des Detektors die Teilchen im Endzustand eines Ereignisses nachzuweisen, müssen zunächst alle Daten aus dem Detektor ausgelesen werden. Die Menge der Daten für ein Ereignis beträgt ca. 1 bis 2 MB. Zusammen mit der Tatsache, dass alle 25 ns, also mit einer Rate von 40 MHz, etwa 23 dieser Ereignisse stattfinden, stellt dies für das Rechen- und Speichersystem des Experiments ein großes Problem dar. Außerdem sind nicht alle Prozesse physikalisch relevant, weshalb auch nicht alle von ihnen aufgezeichnet werden. Welche Ereignisse aber als interessant anzusehen sind, wird von einem komplizierten Triggersystem bestimmt. Dieses ist in 3 Stufen, den so genannten „Leveln“ 1 bis 3 angelegt. Der Level-1 Trigger verwendet nur Informationen aus den Kalorimetern und den Myonenkammern, wobei allerdings nicht die gesamte Granularität des Detektors verwendet wird. Dadurch gelingt es, uninteressante Ereignisse herauszufiltern und die Rate der Ereignisse auf 25 kHz zu reduzieren. Die den Level-1 Trigger passierenden Ereignisse werden dann vom Level-2 Trigger überprüft. Dieser arbeitet mit der vollen Kalorimeterinformation und reduziert die Ereignisrate auf 1 kHz. Erst der Level-3 Trigger rekonstruiert die übrig gebliebenen Ereignisse vollständig. Nach dieser letzten Trigger-Stufe beträgt die Ereignisrate noch ungefähr 100 Hz. Durch dieses Triggersystem also wird die Gesamtmenge der zu speichernden Daten eingeschränkt, in dem potentiell interessante Ereignisse selektiert werden.

# 5 Verwendete Programme

In diesem Kapitel werden die Programme vorgestellt, mit denen die Analysen dieser Arbeit erstellt wurden.

## 5.1 ROOT

ROOT [22] ist ein Objekt-Orientiertes Framework, mit dem enorme Mengen an Daten analysiert werden können. Dieses Framework wurde ursprünglich von R. Brun und F. Rademakers entwickelt und verwendet den C++-Interpreter *CINT* [23], um Kommandozeilen-Befehle und Skripte zu interpretieren. ROOT wurde und wird immer noch nach dem „Basar-Stil“ weiterentwickelt. Dies bedeutet, dass die Entwicklung frei ist und jeder dazu beitragen kann. Dadurch wurde das ROOT-Framework von den Physikern besonders für die Anforderungen der Forschung in der Hochenergie-Physik angepasst.

## 5.2 Toolkit for Multivariate Analysis

Das „Toolkit for Multivariate Analysis“, kurz TMVA, bietet eine in ROOT integrierte Umgebung, deren Implementierung in C++ umgesetzt ist [24]. Ziel dieser Analysen ist es, mehrere Klassen von Ereignissen, die sich in verschiedenen Variablen unterscheiden, mit Hilfe der integrierten Methoden zu diskriminieren. Hierbei können mehrere Methoden parallel verwendet werden. Der Anwender interagiert nur mit wenigen Klassenobjekten, welche die gewünschten Methoden verwalten und auf Wunsch das Training und/oder eine Auswertung vornehmen. Trotz der einfachen Bedienung ist eine gute Kenntnis der Methoden notwendig, um optimale Ergebnisse zu erzielen und Fehler zu vermeiden. Die im TMVA-Paket unterstützten Methoden zur Diskriminierung von Klassen sind

- Anlegen sequentieller Schnitte
- Projektive und multidimensionale Likelihood-Analysen
- Analysen mit linearen Diskriminanten (H-Matrix und Fisher-Methode)
- Drei verschiedene Implementierungen Neuronaler Netzwerke
- Boosted und Bagged Decision Trees
- RuleFit-Methoden

- Support Vector Maschinen

Von den oben genannten Methoden wurden in dieser Arbeit die projektive Likelihood-Methode, das MultiLayerPerceptron, welches eines der Neuronalen Netzwerke ist, und die Boosted Decision Trees verwendet. Diese Methoden werden später in Unterkapitel 6 noch genauer beschrieben.

### 5.3 ATHENA

ATHENA ist eine Sammlung von verschiedenen Programmen. Diese Programme sind das Gaudi-Framework und ATLAS-spezifische Programme, wie das Datenmodell, die Simulation des ATLAS-Detektors durch GEANT und Monte-Carlo-Generatoren für die Ereignissimulation. Außerdem sind Algorithmen implementiert, die es ermöglichen, die vom Detektor über ein Ereignis gelieferten Daten auszuwerten und als Teilchen zu interpretieren. Mit Hilfe von in der Programmiersprache PYTHON verfassten Skripten ist es möglich, nahezu jedes dieser Programme zu konfigurieren, hinzu- oder abzuschalten, ohne den Quellcode selbst ändern zu müssen. Dadurch gelingt es, für jede Analyse gesondert das Framework individuell anzupassen. In der hier vorgestellten Analyse zum Phion-Modell wurde ATHENA in der Version 10.0.4 verwendet. Bei der Identifikation der niederenergetischen Elektronen kam die ATHENA-Version 12.0.6 zum Einsatz. Eine Einführung in das Konzept und eine Hilfe für die Entwicklung von eigenem Code bietet der Developers Guide [25].

### 5.4 PYTHIA

PYTHIA [26] ist ein Monte-Carlo-Ereignisgenerator, der auch im ATHENA-Framework integriert ist. Das PYTHIA-Programm wird verwendet, um teilchenphysikalische Prozesse zu simulieren. Auf Basis der einkommenden Teilchen und deren Eigenschaften berechnet das Programm die Art des Prozesses und die Art der Teilchen und deren Eigenschaften im Endzustand. Alle Prozesse werden dabei mit Hilfe der Störungsrechnung in erster Ordnung berechnet. Der genaue Vorgang bei der Simulation von Ereignissen ist folgender:

1. Die zwei Strahl-Teilchen kommen aufeinander zu. Der genaue Elementarteilcheninhalt der Strahl-Teilchen wird durch die Strukturfunktionen CTEQ5L [27] parametrisiert. Diese basieren auf den Ergebnissen verschiedener Experimente. Die einfachste Möglichkeit, die Strukturfunktionen von Protonen oder Neutronen zu vermessen ist die Elektron-Proton- und die Elektron-Neutron-Streuung.
2. Von jedem der beiden Strahl-Teilchen aus beginnt ein Parton, einen Schauer zu bilden. Dieser Prozess wird „Initial State Radiation“ genannt. Diese Schauerbildung geschieht zum Beispiel durch Abstrahlen von Gluonen, welche wiederum in Quarks oder Gluonen zerfallen können. Von jedem dieser beiden Schauer wird jeweils ein Gluon oder Quark ausgewählt, welches am harten Streuprozess teilnehmen soll.

3. Im harten Streuprozess können die beiden ausgewählten Partonen miteinander wechselwirken und zum Beispiel zu einem Eichboson fusionieren, welches selbst wiederum in andere Elementarteilchen zerfällt. Falls dabei wieder Quarks oder Gluonen erzeugt werden, so können mit diesen wieder Schauer gebildet werden. In diesem Fall wird von „Final State Radiation“ gesprochen.
4. Zusätzlich können auch die anderen, nicht ausgewählten Quarks oder Gluonen des Strahlteilchens an semi-harten Prozessen teilnehmen.
5. Nach diesen Prozessen hadronisieren die Quarks und Gluonen des Endzustandes zu farbneutralen Hadronen. Falls ein so erzeugtes Hadron instabil ist, so zerfällt es weiter, bis am Ende nur stabile Hadronen übrig bleiben.

Dieser Algorithmus ist äquivalent zur Berechnung von Wirkungsquerschnitten, wie sie in Abbildung 2.3 schematisch dargestellt wurde.

## 5.5 GEANT

GEANT [28] ist ein Programm zur Simulation der Durchdringung von Teilchen durch Materie und auch ein Teil des ATHENA-Frameworks. Auf Basis dieses Programms wird der ATLAS-Detektor simuliert. So ist jede Komponente des ATLAS-Detektors in GEANT implementiert. GEANT berechnet auf dieser Basis die Signale, wie sie später im Experiment vom echten Detektor erwartet werden. Hierzu erhält GEANT die Informationen über den stattgefundenen Prozess von PYTHIA. Die erzeugten Teilchen werden dann durch die verschiedenen Teile des Detektors geführt. Bei diesem Prozess werden geometrische Volumenbegrenzungen, die Wechselwirkung der Teilchen mit dem Detektormaterial, wie zum Beispiel Photonenkonversion oder Vielfachstreuung, und die Wirkung des Magnetfeldes berücksichtigt. Die Spuren und Schauer der Teilchen im Detektor werden dann als simulierte Signale der aktiven Detektorkomponenten aufgezeichnet. Mit diesen Signalen kann dann die Rekonstruktion der Ereignisse erfolgen. Aufgrund dieser Vorgehensweise wird eine extrem hohe Genauigkeit bei der Simulation der Ereignisse erreicht. Der Nachteil, der dadurch entsteht, ist der enorme Zeitaufwand, der benötigt wird, um Ereignisse zu simulieren. Aus diesem Grund wurde die schnelle Simulation ATLFAST entwickelt, die im folgenden Abschnitt vorgestellt werden soll. Zusätzliche Informationen zu GEANT findet man zum Beispiel in Referenz [28].

## 5.6 ATLFAST

ATLFAST [29] ist eine Möglichkeit, den ATLAS-Detektor schnell zu simulieren. Statt jede Komponente des Detektors in GEANT einzeln nachzubilden werden die vom Monte-Carlo-Generator erzeugten Teilchen direkt wie rekonstruierte Teilchen behandelt. Somit werden die Simulation des Detektors und die Rekonstruktion der Teilchen vereinfacht und beschleunigt. Um Messungenauigkeiten der Energien und Impulse der rekonstruierten Teilchen zu berücksichtigen, werden

die mit GEANT gemessenen Auflösungen parametrisiert. Die Kalorimeter des ATLAS-Detektors werden in ATLFAST in einem einzigen Kalorimeter der Granularität  $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$  zusammengefasst, welches somit deutlich grobkörniger ist, als das Elektromagnetische Kalorimeter des ATLAS-Detektors. Jedes Teilchen, das mit diesem Kalorimeter wechselwirkt, hinterlässt dort seine mit einer Gauss-Verteilung verschmierte Energie in den verschiedenen Zellen. Auch das Magnetfeld des Solenoidmagneten wird nicht komplett simuliert, sondern ein konstantes Magnetfeld der Stärke 2 T angenommen, welches die Spuren der geladenen Teilchen krümmt. Durch diese Verfahrensweisen ist es möglich, eine für viele Zwecke ausreichende Detektor-Simulation mit möglichst geringem Rechen- und daher Zeitaufwand zu erhalten. Der Geschwindigkeitsvorteil von ATLFAST gegenüber einer kompletten Simulation beträgt üblicherweise vier bis fünf Größenordnungen. Jedoch müssen bei der Verwendung von ATLFAST bei einer Analyse die Rekonstruktionsineffizienzen von Teilchen und Jets, wie sie bei einer detaillierten Rekonstruktion auftreten, gesondert berücksichtigt werden. Eine etwas ältere Anleitung zu ATLFAST bietet Referenz [30], aktuellere Informationen findet man im Internet unter Referenz [29].

## 5.7 Die Ereignisrekonstruktion

Um aus den von GEANT berechneten oder vom ATLAS-Detektor gelieferten Signalen für die Analyse verwendbare Informationen zu erhalten, sind in ATHENA mehrere Programme integriert. Das Ziel dabei ist, die Teilchen, die den Detektor durchquert haben, identifizieren zu können. Sehr wichtige Programme und Vorgehensweisen für die Rekonstruktion der Teilchen sollen im Folgenden kurz vorgestellt werden.

### 5.7.1 Spurrekonstruktion

Die Rekonstruktion von Spuren ([31, 32]) im Spurdetektor ist ein zentrales Element bei der Rekonstruktion der Ereignisse. Mit Mustererkennungsprogrammen werden die Treffer in der Pixel- und SCT-Lage des Spurdetektors zu lokalen Spursegmenten zusammengefasst. An mehrere Segmente werden dann komplette Spuren angepasst. Die verschiedenen Anpassungsalgorithmen, wie zum Beispiel der Standard-Algorithmus „TrkKalmanFitter“, berücksichtigen dabei unter anderem die Krümmung der Spuren durch das Magnetfeld und die Richtungsänderung der Spuren durch Vielfachstreuung. Ist eine Spur an mehrere Segmente angepasst worden, so wird diese bis in den Übergangstrahlungsdetektor extrapoliert und die dortigen Treffer-Informationen dazu verwendet, eine Anpassung einer neuen Spur vorzunehmen.

Die so erhaltenen Spuren sind die in den folgenden Rekonstruktionsschritten verwendeten Spuren der Teilchen.

### 5.7.2 Standardrekonstruktion der Elektronen und Photonen

Das Standardverfahren zur Rekonstruktion von Elektronen und Photonen beginnt mit dem so genannten „Sliding Window“-Algorithmus, der nach „Clustern“ im elektromagnetischen Kalorimeter sucht. Ein Cluster besteht aus mehreren benachbarte Zellen im Kalorimeter, in denen ein

Mindestmaß an Energie deponiert wurde. Ist ein Cluster gefunden, der gewisse Anforderungen erfüllt, so wird eine Spur gesucht, die, wird sie bis auf die Oberfläche des Kalorimeters extrapoliert, mit dem Cluster zusammenpasst.

Das genaue Vorgehen ist dabei das Folgende: Für jeden Cluster werden Variablen berechnet, mit denen die Elektronen und Photonen von anderen Teilchen, hauptsächlich Pionen, unterschieden werden können. Hauptsächlich handelt es sich bei diesen Variablen um Eigenschaften der Schauerform im Kalorimeter. Jeder Cluster muss bestimmte Anforderungen an diese Variablen erfüllen, damit es sich um ein Elektron oder Photon handeln kann.

Um zu einem Cluster, der die Anforderungen an die Variablen erfüllt, die passende Spur zu finden, werden zunächst alle Spuren betrachtet, die einen transversalen Impuls von mindestens  $5 \text{ GeV}/c$  aufweisen. Erfüllt eine Spur diese Anforderung, so werden  $\eta$  und  $\phi$  der Spur am Ursprung mit den Koordinaten  $\eta_C$  und  $\phi_C$  des Cluster-Zentrums am Eintrittspunkt verglichen. Gelten nun die Bedingungen

$$|\eta - \eta_C| < 0.05, |\phi - \phi_C| < 0.1 \text{ und } E/p < 4c, \quad (5.1)$$

so wird die Spur bis in jeden Bestandteil des Kalorimeters extrapoliert, wobei auch die Krümmung der Spur durch das Magnetfeld berücksichtigt wird.  $E$  ist hier die Energie, die vom Teilchen im Kalorimeter deponiert wurde,  $p$  ist der Betrag des Impulses, der im Spurdetektor bestimmt wurde. An den extrapolierten Punkten der Spur an der Oberfläche des Kalorimeters werden wieder die Unterschiede in  $\phi$  und  $\eta$  zum Cluster-Zentrum berechnet, für die folgendes gelten muss:

$$\Delta\eta < 0.025 \text{ und } \Delta\phi < 0.05. \quad (5.2)$$

Wird keine Spur gefunden, die die Bedingungen der Gleichung 5.2 erfüllt, so wird der Cluster als Photon gekennzeichnet.

Falls die zu einem Cluster passende Spur eine hohe Wahrscheinlichkeit aufweist, dass Photokonversion stattgefunden hat, so wird dieses Objekt als Photon gekennzeichnet, ansonsten handelt es sich bei dieser Kombination um ein Elektron. Die Wahrscheinlichkeit, ob eine solche Photokonversion stattgefunden hat, wird von einem speziellen Algorithmus bestimmt.

Zusätzlich zu diesem Algorithmus existiert ein Algorithmus zur Identifikation von Elektronen aus dem b-Quark-Zerfall. Dieser beginnt damit, die Spuren in das elektromagnetische Kalorimeter zu extrapolieren und dort nach dazugehörigen Clustern zu suchen, und wird in Kapitel 8.2 näher beschrieben.

### 5.7.3 Myonen-Rekonstruktion

Um Myonen zu rekonstruieren, sind die Algorithmen „Muonboy“ und „MOORE“ in ATHENA integriert. Im Folgenden soll jedoch nur der Algorithmus Muonboy kurz beschrieben werden. Muonboy ([33, 34]) beginnt damit, aktive Regionen in den Myonen-Kammern zu suchen. Die Treffer in den aktiven Regionen werden in lokalen Segmenten zusammengefasst, die einige Bedingungen erfüllen müssen. Zum Beispiel müssen die Segmente mit dem Primärvertex des Ereignisses kompatibel sein. Verschiedene Segmente werden danach miteinander verbunden, falls sie zu einer gemeinsamen Spur passen könnten. Dies wird mit mehreren Kombinationen der

Segmente versucht. Falls mehrere Segmente zusammenpassen, so ist ein Spur-Kandidat gefunden. Dieser wird mit den Spuren im Spurdetektor verglichen. Passt die rekonstruierte Spur in den Myonen-Kammern mit einer Spur im Spurdetektor zusammen, so wird diese Spur als Myon identifiziert.

### 5.7.4 Rekonstruktion der Tau-Leptonen

Die Rekonstruktion von Tau-Leptonen kann in ATHENA mit den Programmen „tau1p3p“ und „tauRec“ durchgeführt werden, von denen aber nur das Letztere, mit Hilfe der Referenz [35], beschrieben werden soll. Das Ziel dieser Algorithmen ist die Diskriminierung der Tau-Leptonen von Jets.

tauRec behandelt zunächst alle Kalorimeter-Cluster der so genannten „CombinedCluster“-Kollektion als Kandidaten für Tau-Leptonen<sup>1)</sup>. Für alle Kandidaten werden Spuren gesucht, die in einem Kegel mit einem Radius  $R = 0.3$  um das Zentrum des Clusters liegen und einen transversalen Impuls von mindestens  $2 \text{ GeV}/c$  besitzen. Diese Spuren werden dann bis zur Kalorimeteroberfläche extrapoliert. Mit Hilfe der Likelihood-Methode, die in Kapitel 6 beschrieben wird, wird dann mit mehreren Variablen der Spuren und des Clusters die Wahrscheinlichkeit berechnet, dass die Kombination von Cluster und Spuren zu einem Tau-Lepton gehört. Ein Beispiel für eine solche Variable ist der sogenannte EM-Radius, der dem kleineren transversalen Schauerprofil von  $\tau$ -Jets im Vergleich zum Profil der Quark-Jets Rechnung trägt.

### 5.7.5 Jet-Rekonstruktion

Auch für die Rekonstruktion von Jets sind mehrere Algorithmen entwickelt worden. Der am meisten verwendete Algorithmus ist der so genannte „Seeded Cone“-Algorithmus ([36,37]). Dieser bildet Jets, indem er mehrere Kalorimetercluster, die innerhalb eines Kegels mit einem bestimmten Radius  $R$ <sup>2)</sup> liegen, in einem Objekt zusammenfasst.

Um die wahre Achse des Jet-Kegels zu finden, wird zunächst eine zufällige Kegel-Achse gewählt. Mit allen Clustern, die innerhalb des Kegels um diese Achse mit Radius  $R$  liegen, wird eine neue Achse berechnet. Dies geschieht durch Berechnung des Mittels der Richtungen der Cluster, gewichtet mit ihrer Energie. Mit dieser neuen Achse wird das Verfahren wiederholt, bis eine stabile Lösung gefunden wird.

Die in dieser Arbeit verwendeten rekonstruierten Jets basieren auf diesem Algorithmus. Der Radius des Jet-Kegels betrug dabei  $R = 0.4$ .

### 5.7.6 Vertex-Rekonstruktion

Ziel der Primär-Vertex-Bestimmung ist es, die z-Komponente des Primär-Vertex zu bestimmen. Die x- und y-Koordinate sind durch die geringe Strahlausweitung gut bekannt. Weiterhin muss der richtige Vertex des Prozesses gefunden werden, da durch den Effekt des „Pile-Up“ mehrere

---

<sup>1</sup>In zukünftigen Versionen werden für diesen Zweck die sogenannten „TopoCluster“ verwendet.

<sup>2</sup>Die üblicherweise verwendeten Werte für den Kegelradius sind  $R = 0.4$  und  $R = 0.7$ .

Vertices existieren können.

Die Rekonstruktion des Primär-Vertex erfolgt durch Anpassung eines gemeinsamen Schnittpunktes an die Spuren. Hierzu können aber nicht alle Spuren verwendet werden, da wegen Pile-Up eventuell nicht alle Spuren von einem gemeinsamen Primär-Vertex stammen. Daher müssen die Spuren in verschiedenen Mengen, die die verschiedenen möglichen Vertices repräsentieren, gruppiert werden. An die Spuren der verschiedenen Gruppen werden dann Vertices angepasst. Hierbei werden Spuren, die die Anpassung des Vertex deutlich verschlechtern, wieder aus den Mengen entfernt. Der rekonstruierte Vertex, welcher die höchste Summe an transversalen Impulsen seiner Spuren aufweist, wird als Primär-Vertex verwendet.

Weiterhin können mit diesem Algorithmus sekundäre Zerfalls-Vertices, wie die von c- oder b-Hadronen, rekonstruiert werden. Hierzu werden Spuren betrachtet, die in Anbetracht der Strahl-Ausbreitung nicht mit einem Primär-Vertex vereinbar sind. An diese Spuren werden dann wieder Vertices angepasst, die den Zerfallsort der langlebigen Hadronen darstellen.



## 6 Methoden zur Klassifikation von Ereignissen

In diesem Kapitel werden Methoden zur Trennung von Ereignisklassen beschrieben. Als Werkzeug wurde die in ROOT integrierte Umgebung TMVA verwendet, welche in Unterkapitel 5.2 beschrieben wurde. Ziel dieser Methoden ist es, verschiedene Klassen von Ereignissen voneinander unterscheiden zu können. Diese Unterscheidung wird mit Hilfe mehrerer Variablen vollzogen, die die Ereignisse beschreiben. Eine totale Zugehörigkeit eines Ereignisses zu einer Klasse kann meist nicht entschieden werden, da die Variablen, die zur Trennung der verschiedenen Klassen verwendet werden, für die verschiedenen Klassen die gleiche Werte annehmen können. Daher versuchen alle Methoden, die a-posteriori Wahrscheinlichkeit  $P(K_i|V)$  zu bestimmen, mit der ein Ereignis mit dem Variablensatz  $V$  zu einer bestimmten Klasse  $K_i$  gehört. Für diese Wahrscheinlichkeit gilt mit dem Bayes-Theorem [38]:

$$P(K_i|V) = \frac{P(V|K_i) \cdot P(K_i)}{P(V)} \quad (6.1)$$

Hierbei ist  $P(K_i)$  die a-priori Wahrscheinlichkeit, dass ein Ereignis zur Klasse  $K_i$  gehört, und  $P(V)$  die a-priori Wahrscheinlichkeit, dass der Variablensatz  $V$  unabhängig von einer Klasse auftritt.  $P(V|K_i)$  beschreibt die Wahrscheinlichkeit für das Auftreten des Variablensatzes  $V$  in Klasse  $K_i$ .

Durch dieses Theorem kann ein Ereignis mit Variablensatz  $V$  der Klasse  $K_i$  zugeordnet werden, für die die a-posteriori Wahrscheinlichkeit  $P(K_i|V)$  maximal ist.

In den meisten Fällen sind die Wahrscheinlichkeiten  $P(V)$  und  $P(V|K_i)$  nicht analytisch bekannt. Existieren jedoch Beispielergebnisse, so können die Wahrscheinlichkeiten modelliert werden. Dies ist streng genommen nur dann möglich, falls die Anzahl der Beispielergebnisse unendlich groß ist. In realistischen Szenarien ist diese Anzahl stark begrenzt, wodurch die Annäherung an die analytischen Wahrscheinlichkeiten nicht exakt ist. Um dieses Problem zu verringern, werden für manche Methoden nicht die mehrdimensionalen Wahrscheinlichkeitsverteilungen modelliert, sondern für jede Variable eine eindimensionale Verteilung. Dies bedeutet aber, dass Informationen über Korrelationen zwischen den Variablen nicht mehr berücksichtigt werden.

Weiterhin stehen für die Modellierung der Wahrscheinlichkeitsdichten oft keine echten Daten zur Verfügung. Aus diesem Grund werden Monte-Carlo-Simulationen durchgeführt und die Ergebnisse zur Erstellung der Wahrscheinlichkeiten verwendet. Die durch die Simulation erhaltenen Ereignisse werden in zwei Untermengen aufgeteilt. Die eine Menge, im Folgenden als Trainingsereignisse bezeichnet, dient dazu, die Methoden zu trainieren und die Wahrscheinlichkeitsdichten zu modellieren. Die zweite Menge an Ereignissen, die sogenannten Testereignisse,

dienen dazu, die so erhaltenen Dichten zu verifizieren und die austrainierten Methoden auf ihre Leistungsfähigkeit hin zu überprüfen. Um dies zu gewährleisten, müssen die Testereignisse statistisch unabhängig von den Trainingsereignissen sein.

Im Folgenden werden nur zwei Klassen von Ereignissen berücksichtigt. Diese werden als „Signal“ und „Untergrund“ bezeichnet.

## 6.1 Die Likelihood-Methode

Die Likelihood-Methode verwendet die Verteilung im N-dimensionalen Variablenraum direkt als diskriminierende Variable. Denn die Likelihood-Funktion  $L_{Sig/Bg}$  eines Ereignisses mit den Variablenwerten  $x_1, \dots, x_n$  ist definiert als der Wert der Wahrscheinlichkeit für die Signal-/Untergrundzugehörigkeit im n-dimensionalen Variablenraum <sup>1)</sup>:

$$L_{Sig/Bg} = P_{Sig/Bg}(x_1, \dots, x_N). \quad (6.2)$$

Geht man von statistisch unabhängigen Variablen aus oder reicht die Zahl der Trainingsereignisse nicht aus, um gültige Wahrscheinlichkeitsverteilungen im n-dimensionalen Variablenraum zu modellieren, so lässt sich diese Darstellung zu

$$L_{Sig/Bg} = \prod_i P_{i,Sig/Bg}(x_i) \quad (6.3)$$

vereinfachen, wobei die  $P_{i,Sig/Bg}(x_i)$  nun getrennt die Wahrscheinlichkeitsverteilungen für Signal und Untergrund der einzelnen Variablen  $x_i$  sind. Für diese muss also gelten

$$\int_{-\infty}^{\infty} P_{i,Sig/Bg}(x_i) dx_i = 1 \quad , \forall i. \quad (6.4)$$

Die Leistung der Likelihood-Methode unter Verwendung der eindimensionalen Wahrscheinlichkeitsdichten ist dann nur optimal, falls die verwendeten Variablen nicht korreliert sind. Denn jegliche Korrelation zwischen den Variablen wird so nicht berücksichtigt. Andererseits ist die mathematische Beschreibung auf diese Weise deutlich einfacher und die Anzahl der benötigten Trainingsereignisse viel geringer.

Da oftmals auch für die Modellierung der eindimensionalen Wahrscheinlichkeitsdichten  $P_i(x_i)$  die Anzahl der Trainingsereignisse nicht ausreicht, um glatte Verteilungen zu erhalten, werden die Verteilungen in Histogramme eingetragen. Diese Histogramme werden durch bestimmte Verfahrensweisen geglättet und später durch Polynome angenähert.

Der Algorithmus, der im Programmpaket TMVA zum Glätten der Histogramme verwendet wird, wurde in Referenz [39] vorgestellt. Der Grad der Polynome kann je nach Verteilung der Variablen angepasst werden. Diese Polynome ergeben die Wahrscheinlichkeitsverteilungen der Variablen  $x_i$  für eine Analyse. Es ist aber auch möglich, statt eines Polynoms die geglätteten Verteilungen

---

<sup>1</sup>Die Likelihoodfunktion entspricht dem Quotienten  $P(V|K_i)/P(V)$  der in der Einleitung zu diesem Kapitel eingeführten a-priori Wahrscheinlichkeiten  $P(V|K_i)$  und  $P(V)$ .

direkt als Wahrscheinlichkeitsverteilungen zu verwenden. Obwohl die durch dieses Verfahren erhaltenen Polynome glatt sind, können die erhaltenen Wahrscheinlichkeitsdichten dennoch unerwünscht oszillieren. Weiterhin reagiert diese Methode sehr empfindlich auf die Anzahl der Kanäle der Histogramme. Wenn nur ein Kanal eines Histogramms etwas verschoben wird, so kann das resultierende Polynom schon ein anderes sein. Aus diesen Gründen wird gewöhnlicherweise der Grad  $g$  des Polynoms zu  $g = 2$  gewählt.

Eine weitere Möglichkeit, die Wahrscheinlichkeits-Dichtefunktionen aus den Variablenverteilungen zu extrahieren, ist die Parametrisierung mit Hilfe des Kern-Dichte-Abschätzers <sup>2)</sup>. Statt die Variablen, welche in Histogramme mit vielen Kanälen eingetragen werden, durch Polynome anzunähern, werden hier die Wahrscheinlichkeits-Dichtefunktionen gewonnen, indem die Variablenwerte der Trainingsereignisse ausgeschmiert werden. Statt Einträge in ein Histogramm vorzunehmen, werden für jeden Variablenwert Funktionen addiert. Für  $N$  Trainingsereignisse mit den Variablenwerten  $x_i$  folgt so für die Wahrscheinlichkeits-Dichtefunktionen  $p$  für eine Variable  $x$ :

$$p(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x-x_i}{h}\right). \quad (6.5)$$

Die Funktion  $K$  wird Kern-Funktion genannt, der Parameter  $h$  ist die Bandbreite der Kern-Funktion. Meist wird für die Kern-Funktion die Gauss-Funktion gewählt.

Für die Wahl der Bandbreite  $h$  existieren mehrere Varianten. Zum Einen kann die Bandbreite für alle Trainingsereignisse fest gewählt werden (Nicht-Adaptive Methode). Mit einer Gauss-Funktion als Kern-Funktion ergibt sich als sinnvoller Wert für die Bandbreite [40]:

$$h_{NA} = \left(\frac{4}{3}\right)^{1/5} \sigma_x \cdot N^{-1/5}. \quad (6.6)$$

Hier ist  $\sigma_x$  der quadratische Mittelwert der Verteilung der Variable  $x$ :

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}. \quad (6.7)$$

Zum Anderen kann die Bandbreite variabel sein (Adaptive Methode), wobei die lokale Ereignisdichte berücksichtigt wird:

$$h_A(x) = \frac{h_{NA}}{\sqrt{\tilde{p}(x)}}. \quad (6.8)$$

Die Wahrscheinlichkeitsdichte  $\tilde{p}(x)$  ergibt sich aus der Durchführung der Nicht-Adaptiven Methode.

Ein Problem der Methode des Kern-Dichte-Abschätzers ist das Grenzproblem. Denn scharfe Kanten werden durch diese Methode schlecht beschrieben. Und treten diese Kanten an den Histogrammgrenzen auf, so werden diese in ein Gebiet außerhalb der Histogrammgrenzen hinaus

<sup>2)</sup>Im Englischen „kernel density estimator“ genannt

verschmiert. Dadurch ändert sich aber die Fläche, und somit auch die Wahrscheinlichkeitsdichte, innerhalb der Grenzen. Weiterhin können so unphysikalische Werte für bestimmte Variablen auftreten, wie zum Beispiel eine negative Energie.

Um dieses Problem zu lösen, existieren zwei verschiedene Ansätze, die optional verwendet werden können:

1. Die Kern-Dichtefunktionen werden renormiert, sodass das Integral der Verteilung innerhalb der Histogrammgrenzen gleich 1 ist.
2. Die Dichte außerhalb der Histogrammgrenzen wird an der Grenze gespiegelt und somit in den relevanten Bereich verlagert.

Hat man die Wahrscheinlichkeitsdichten der Variablen durch eine dieser Methoden gefunden, so kann man die Likelihood-Funktionen wie oben beschrieben berechnen. Um nun noch zwischen Signal und Untergrund zu unterscheiden, wird das Likelihood-Verhältnis  $y_L$  eingeführt

$$y_L = \frac{L_{Sig}}{L_{Sig} + L_{Bg}}. \quad (6.9)$$

Häufig wird auch eine andere Definition des Likelihood-Verhältnisses verwendet

$$y_L = \log \left( \frac{L_{Sig}}{L_{Bg}} \right). \quad (6.10)$$

Indem man nun für Ereignisse einen Mindestwert für  $y_L$  fordert, erhält man eine Trennung von Signal und Untergrund. Die Vorteile der Likelihood-Methode sind die einfache Nachvollziehbarkeit der Ergebnisse und die Tatsache, dass schlecht diskriminierende Variablen die Endergebnisse nicht negativ beeinflussen.

Wie oben schon erwähnt wurde, werden durch die Likelihood-Methode bei Verwendung von eindimensionalen Wahrscheinlichkeitsdichten die Informationen über Korrelationen zwischen den Variablen nicht beachtet. Allerdings können durch Transformationen der Variablen zumindest die linearen Korrelationen berücksichtigt werden. Um eine geeignete Transformation zu erhalten, wird zunächst die Wurzel der linearen Kovarianzmatrix der Signalereignisse gesucht. Die Wurzel einer Matrix  $C$  ist definiert als die Matrix  $C'$ , für die  $C = (C')^2$  gilt. Um diese Matrix  $C'$  zu finden, wird die symmetrische Kovarianzmatrix diagonalisiert. Ist  $D$  die diagonale Form der Matrix  $C$ , d.h. es existiert eine Matrix  $S$  mit  $D = S^T C S$ , so folgt für die Wurzel von  $C$ :

$$D = S^T C S \implies C' = S \sqrt{D} S^T. \quad (6.11)$$

Die linear dekorrelierten Variablen erhält man nun durch multiplizieren des Vektors der Variablen mit der Transformations-Matrix  $(C')^{-1}$  für die Signalereignisse.

Dieses Verfahren funktioniert am besten für linear korrelierte oder gaussverteilte Variablen. Sind die Variablen jedoch stark nicht-linear voneinander abhängig, so kann die Anwendung dieses Verfahrens das Ergebnis der Likelihood-Methode verschlechtern.

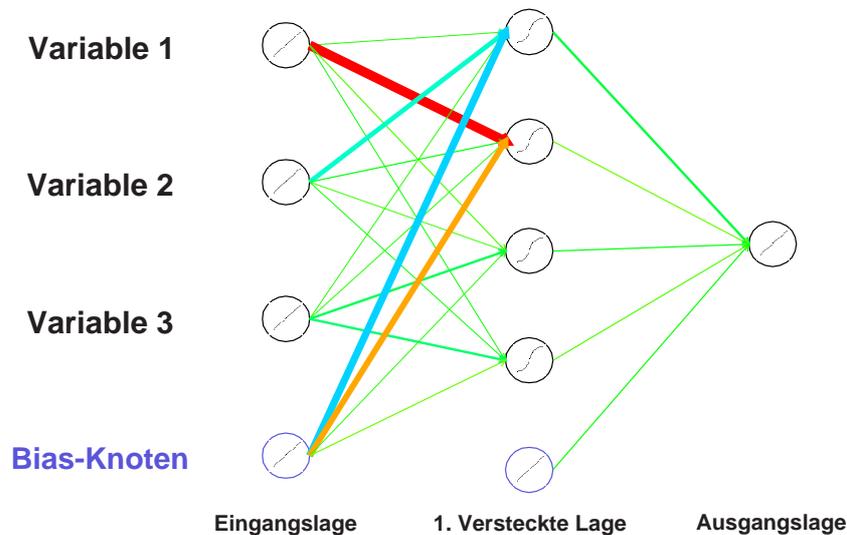


Abbildung 6.1: Beispiel für ein Perzeptron (Neuronales Netzwerk) mit drei Eingangsvariablen, einer versteckten Lage mit vier Neuronen und einem Ausgangsknoten. Die Schwellen-Knoten realisieren für jedes damit verbundene Neuron die Verschiebung der Antwortfunktion um eine Konstante  $b_i$  (siehe hierzu Gleichung 6.12). Die verschiedenen Farben symbolisieren die unterschiedlichen Gewichte  $w_i$  der Verbindungen zwischen den einzelnen Neuronen. Diese Abbildung wurde mit Hilfe des Programms TMVA angefertigt.

## 6.2 Neuronale Netzwerke

Ein künstliches Neuronales Netzwerk stellt den Versuch dar, ein natürliches Neuronales Netz, wie zum Beispiel das menschliche Gehirn, nachzuahmen. Dadurch soll es ermöglicht werden, komplizierte Klassifikationsprobleme zu lösen. Eine vereinfachte Form eines Neuronalen Netzes ist ein Perzeptron, welches nun mit Hilfe der Referenzen [24] und [41] beschrieben werden soll. Ein Beispiel für ein solches Perzeptron zeigt Abbildung 6.1.

Hier handelt es sich um ein Netzwerk mit drei Eingangsvariablen. Diese und ein Schwellen-Knoten sind mit allen Neuronen der nächsten Lage verbunden. Die Lagen zwischen Eingangs- und Ausgangslage werden versteckte Lagen genannt. Je nachdem, welche Eingangswerte nun auf ein Neuron wirken und wie groß die Gewichte der Verbindungen zu diesem Neuron sind, ändert sich der Ausgabewert dieses Neurons. Dieser Ausgabewert wirkt wiederum auf die Neuronen der nächsten Lage, bzw. am Ende auf den Ausgabeknoten. Da so die Ausgabe nur auf die darauffolgenden Lagen weitergegeben wird, spricht man auch von einem Vorwärtsvermittlungnetzwerk.

Die Ausgabe eines Neurons, also die Reaktion auf einen bestimmten Satz an Eingabewerten, wird mathematisch durch eine Antwortfunktion beschrieben. Sind  $x_1, \dots, x_N$  die Eingabewerte, die auf

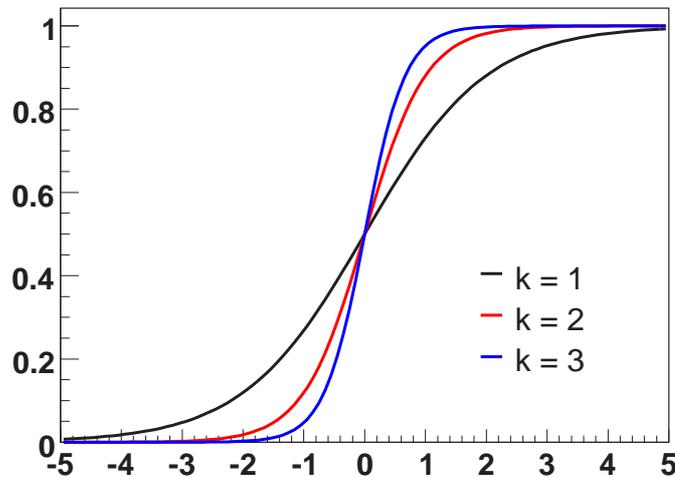


Abbildung 6.2: Die Sigmoid-Funktion für verschiedene Werte des Parameters  $k$ , dessen Kehrwert  $1/k$  auch Temperatur genannt wird.

das betroffene Neuron wirken, so ergibt sich mit den Gewichten  $w_1, \dots, w_N$  der Verbindungen zu diesem Neuron ein gesamter Eingabewert  $x$  aus der Synapsenfunktion

$$x = s(x_i) = \sum_i w_i * x_i + b_i \quad (6.12)$$

für das Neuron, wobei  $b_i$  das Gewicht der Verbindung zwischen Neuron und Schwellen-Knoten ist. Durch dieses Gewicht  $b_i$  wird eine Verschiebung der Antwortfunktion des Neurons realisiert. Der Ausgabewert des Neurons wird dann mit einer Antwortfunktion berechnet. Üblicherweise wird als Antwortfunktion die Sigmoidfunktion gewählt:

$$y = \frac{1}{1 + e^{-k \cdot (x_i)}} \quad , k = const. \quad (6.13)$$

Die Sigmoidfunktion ist für verschiedene Werte des Parameters  $k$  in Abbildung 6.2 dargestellt. Mathematisch bedeutet dies, dass der Variablenraum  $R^N$  der auf die Neuronen wirkenden Eingabewerte  $x_i$  mit Hilfe der Synapsenfunktion  $s$  und der Antwortfunktion  $g$  durch Anwendung von  $g \circ s(x_i)$  auf einen eindimensionalen Antwortraum abgebildet wird. Diese Abbildung ist nicht linear, da die Antwortfunktion nicht linear ist. Die Berechnung der Ausgabe eines Knotens ist graphisch in Abbildung 6.3 dargestellt.

Am Ende des Netzes liegt der Ausgabeknoten, mit dessen Ausgabewert die a-posteriori Wahrscheinlichkeit beschrieben werden soll, dass die Eingabewerte des Netzes zu einer bestimmten Klasse gehören.

Die Bestimmung der freien Parameter des Neuronalen Netzwerkes, der Gewichte  $w_i$ , geschieht

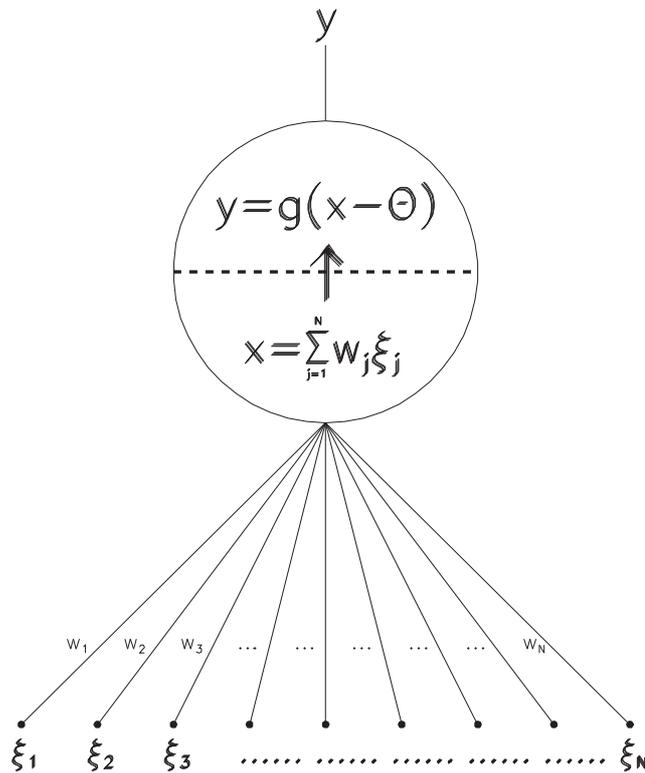


Abbildung 6.3: Graphische Darstellung der Berechnung der Ausgabe eines Knotens eines Neuronalen Netzwerkes. Mit den Eingangs-Variablen  $\xi_i$  und den Gewichten  $w_i$  wird der Wert der Synapsenfunktion, der gesamte Eingabewert  $x$  des Knotens, berechnet. Mit diesem Wert und der Antwortfunktion  $g(x - \theta)$  wird die Ausgabe  $y$  des Knotens bestimmt. Die Variable  $\theta$  realisiert eine Verschiebung der Antwortfunktion für diesen Knoten. Diese Abbildung wurde der Referenz [41] entnommen.

in der Trainingsphase. In dieser werden für Signal- und Untergrundereignisse Zielwerte für die Ausgabe des Netzwerkes definiert: Am Ausgangsknoten soll ein Signalereignis einen Ausgabewert von möglichst Eins erreichen, ein Untergrundereignis einen Wert möglichst nahe der Null. Die Gewichte der Verbindungen zwischen den Neuronen der verschiedenen Lagen müssen dann so gewählt werden, dass die Trennung der Verteilungen der Ausgabewerte von Signal und Untergrund möglichst gut wird.

Ein Algorithmus zum Trainieren eines Netzwerkes ist der „Backpropagation“-Algorithmus. Dieser überwacht die Vorwärtsvermittlung der Eingabewerte und minimiert durch Anpassung der Gewichte die Fehlklassifizierungen der Trainingsereignisse. Hierzu werden die Gewichte zum Ausgabeknoten hin als Erstes angepasst, die der Eingangslage zuletzt, wodurch der Name Backpropagation-Algorithmus entstand. Um die Fehlklassifizierungen von Trainingsereignissen feststellen zu können, muss diesem Algorithmus bekannt sein, ob ein bestimmtes Trainingsereignis zur Signal- oder Untergrundklasse gehört. Man spricht daher von einem beaufsichtigten Lernen, da die Zugehörigkeit der Trainingsereignisse bekannt sein muss. Die Gewichte des Neuronalen Netzes werden durch das Training derart verändert, dass folgende Fehlerfunktion, die die Fehlklassifizierung der Trainingsereignisse quantisiert, minimiert wird:

$$E(\vec{x}, \vec{w}) = \sum_{a=1}^N \frac{1}{2} (y_a - \hat{y}_a)^2. \quad (6.14)$$

In  $\vec{w}$  sind alle Gewichte des Neuronalen Netzwerkes enthalten. Der Vektor  $\vec{x}$  besteht aus den Eingabewerten aller verwendeten Trainingsereignisse.  $y_a$  sind die Ausgabewerte des Neuronalen Netzwerkes für die Menge der Trainingsereignisse,  $\hat{y}_a$  die gewünschten Ausgabewerte (1 für Signal- und 0 für Untergrundereignisse). Diese Funktion beschreibt somit die quadratische Abweichung der wirklich erzielten Ausgabewerte von den gewünschten Werten. Die Menge der  $w_i$ , die diese Fehlerfunktion minimieren und somit die beste Klassifikation der Ereignisse erreicht, wird mit Hilfe der Gradientenabstieg-Methode <sup>3)</sup>, gesucht. Hierzu werden, die Differenzierbarkeit der Antwortfunktion der Neuronen vorausgesetzt, die Gewichte des Netzes mit jedem Lernzyklus  $n$  folgendermaßen verändert <sup>4)</sup>:

$$w^{n+1} = w^n - \eta \nabla_{\vec{w}} E. \quad (6.15)$$

Die positive Zahl  $\eta$  heisst Lernrate,  $\nabla_{\vec{w}} E$  ist die Ableitung der Fehlerfunktion im Raum der Gewichte  $\vec{w}$ . Dies bedeutet, dass die Gewichte, beginnend von einem zufällig gewählten Anfangssatz, mit jedem Trainingszyklus in Richtung  $-\eta \nabla_{\vec{w}} E$  verschoben werden, in der die Fehlerfunktion  $E$  am schnellsten abfällt. Dafür muss die Fehlerfunktion streng monoton sein, um stets eine Änderung der Gewichte zu bewirken. Weiterhin gilt, dass eine kleine Lernrate  $\eta$  stets zu einem Minimum führt. Allerdings ist dann nicht gewährleistet, dass dieses Minimum ein globales Minimum ist. Auch dauert das Trainieren des Netzes mit einer kleinen Lernrate länger, will man das Minimum erreichen. Wählt man die Lernrate allerdings zu groß, so kann es sein, dass die Lösung um das Minimum herum oszilliert. Aus diesen Gründen, wird die Lernrate meist

---

<sup>3</sup>Im Englischen „gradient descent method“

<sup>4</sup>Zusätzlich kann noch ein sogenannter „Momentum“-Term berücksichtigt werden.

nicht konstant gewählt, sondern nach jedem Lernzyklus durch einen Parameter  $(1-\tau)$ , der Zerfallsrate des Lernparameters, verkleinert. Für die Lernrate  $\eta$  mit Anfangswert  $\eta_0$  gilt dann bei einem Trainingszyklus  $z$ <sup>5)</sup>:

$$\eta = (1 - \tau)^{z-1} \cdot \eta_0. \quad (6.16)$$

Allerdings muss bei dieser Methode darauf geachtet werden, dass die Zerfallsrate nicht zu klein gewählt wird, da sich sonst schon nach wenigen Trainingszyklen die Werte der Gewichte nicht mehr verändern und daher das Minimum der Fehlerfunktion nicht erreicht werden kann.

Eines der großen Probleme von Neuronalen Netzwerken ist die Undurchsichtigkeit der Ergebnisse und die Tatsache, dass nicht kontrolliert werden kann, ob ein globales oder nur ein lokales Minimum der Fehlerfunktion erreicht wurde. Weiterhin ist die optimale Wahl der Netztopologie für ein bestimmtes Problem im Voraus nicht bekannt.

Auch die optimale Anzahl der Knoten jeder Lage ist a priori nicht bekannt. Jedoch führt eine zu geringe Anzahl an Knoten zu einer geringeren Leistung des Netzes. Wird ihre Anzahl zu hoch angesetzt, so werden durch das Netz oft nur noch die ausgewählten Trainingsdaten beschrieben. Das Netz hat diese sozusagen auswendig gelernt. Dieser Effekt des Übertrainierens eines Netzwerkes kann erkannt werden, indem neben den Trainingsdaten ein weiterer statistisch unabhängiger Satz an Daten, die Testereignisse, während des Trainings durch das Netz propagiert wird. Ist die Klassifikation dieser Daten deutlich schlechter als die Klassifikation der Trainingsdaten, so hat mit hoher Wahrscheinlichkeit ein Übertrainieren des Netzes stattgefunden.

Die Vorteile eines Neuronalen Netzwerkes sind, dass die Ausgabe des Netzes am Ausgabeknoten  $i$  unter bestimmten Bedingungen mit der a-posteriori Wahrscheinlichkeit  $P(K_i|V)$  der Zugehörigkeit zu einer bestimmten Klasse  $K_i$  identifiziert werden kann. Dies wird zum Beispiel in Referenz [42] gezeigt. Die Bedingungen, die hierzu gelten müssen, sind:

1. Es müssen genügend Trainingsdaten vorhanden sein, damit die Wahrscheinlichkeiten  $P(V|K_i)$  ausreichend gut beschrieben werden.
2. Die einzelnen Klassen müssen entsprechend ihrer a-priori Wahrscheinlichkeit  $P(K_i)$  in der Menge der Trainingsereignisse vertreten sein.
3. Die verwendete Fehlerfunktion muss eine quadratische Form aufweisen.
4. Die Zielwerte für einen Ausgabeknoten  $i$  müssen zu Eins gewählt werden, falls ein Ereignis zur Klasse  $i$  gehört, Null sonst.
5. Das Neuronale Netzwerk muss ausreichend komplex sein. Nur dadurch kann das globale Minimum der Fehlerfunktion gefunden werden.

Weiterhin werden Korrelationen zwischen den verwendeten Variablen automatisch berücksichtigt. Auch konnte die prinzipielle Trennfähigkeit eines Neuronalen Netzes für mehrere Klassen

<sup>5)</sup>Für die letzten 5% der Trainingszyklen gilt für das Neuronale Netzwerk im TMVA-Paket:  $\eta = (1 - \tau^2)^{z-1} \cdot \eta_0$

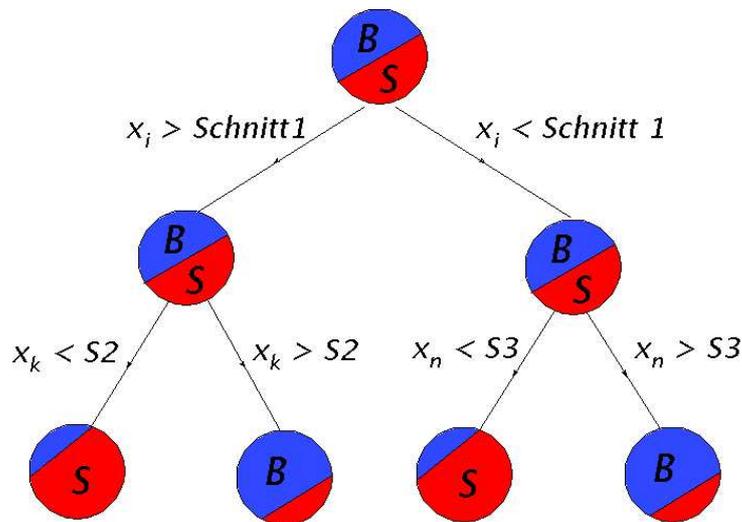


Abbildung 6.4: Beispiel für einen Entscheidungsbaum

von Problem mathematisch bewiesen werden. So können durch ein Neuronales Netzwerk im N-dimensionalen Variablenraum der Eingangsvariablen konvexe Gebiete <sup>6)</sup> selektiert werden [43]. Mit einem Neuronalem Netz ohne versteckte Lagen ist es möglich, im N-dimensionalen Variablenraum Ereignisse jenseits einer Halbebene zu selektieren. Denn die gewichtete Summe am Ausgabeknoten stellt eine N-1-dimensionale Hyperebene dar.

Mit zwei versteckten Lagen kann jeder Satz von Funktionen der Eingangsvariablen beschrieben werden [43]. Auch das XOR-Problem kann erst mit einem Netz mit zwei versteckten Lagen gelöst werden [44].

### 6.3 Entscheidungsbäume

Ein Entscheidungsbaum <sup>7)</sup> basiert auf einer binären Baumstruktur. Abbildung 6.4 zeigt einen solchen Entscheidungsbaum.

Anfangen von einem Startknoten, wird an jeder Verzweigung bestimmt, ob ein Ereignis den linken oder den rechten Ast herabwandert. Jedes Ereignis erreicht irgendwann einen Endknoten, der dem Ereignis eine gewisse Wahrscheinlichkeit zuweist, dass dieses zur Signal- oder Untergrundklasse gehört. Bei dieser Methode werden gewöhnlicherweise mehrere Bäume trainiert, mit denen eine kombinierte Wahrscheinlichkeit der Klassenzugehörigkeit berechnet wird.

Das Training eines solchen Entscheidungsbaums beginnt am Startknoten. Von diesem aus wird eine Variable gewählt und ein Schnitt auf diese, der die Menge der Trainingsereignisse in zwei Teile aufteilt. Dies wird für jeden darauffolgenden Knoten genau so fortgeführt, bis am Ende

<sup>6)</sup>Ein Gebiet ist genau dann konvex, falls zu je zwei beliebigen Elementen des Gebietes ein Verbindungsweg existiert, der noch ganz innerhalb des Gebietes liegt.

<sup>7)</sup>im Englischen „boosted decision tree“

Knoten entstehen, die eine bestimmte Anzahl an Ereignissen beinhalten oder eine bestimmte Größe des Baumes erreicht ist. Durch geschickte Wahl der Variablen und der Schnitte für jeden Knoten besitzen die Endknoten eine hohe Reinheit an Signal- oder Untergrundereignissen. Welche Variable jedoch für welche Verzweigung verwendet wird und welcher Schnitt darauf angewandt wird, wird im Training durch verschiedene Kriterien festgelegt. Der Trainingsalgorithmus wählt immer die Variable und den Schnitt aus, für die der Anstieg eines Trennungskriteriums für den Knoten selbst und die Summe der Trennungskriterien der beiden Töchterknoten, gewichtet mit der Reinheit des Signals in diesen Knoten, maximal ist. Ist  $p = N_S / (N_S + N_B)$  die Reinheit der resultierenden Trennung einer Variable-Schnitt-Kombination, so ergeben sich verschiedene Trennungskriterien zum Beispiel aus folgenden Funktionen:

- *Gini Index*:  $p * (1 - p)$
- *Entropie*:  $-p * \ln(p) - (1 - p) * \ln(1 - p)$
- *Statistische Signifikanz*:  $N_S / \sqrt{N_S + N_B}$
- *Fehlklassifizierung*:  $1 - \max(p, 1 - p)$

Durch dieses Verfahren entstehen im n-dimensionalen Variablenraum kleine Gebiete, welche als signal- oder untergrundartig eingestuft werden, je nachdem, wie viele Signal- oder Untergrundereignisse nach der Propagation der Ereignisse in dem entsprechenden Endknoten vorhanden sind. Theoretisch wäre es möglich, den Baum so lange weiterwachsen zu lassen, bis in den Endknoten nur noch Signal- oder Untergrundereignisse enthalten sind. Ein solcher Baum wäre jedoch hochgradig übertrainiert und würde nur noch speziell die verwendeten Trainingsereignisse beschreiben. Daher lässt man die Bäume nur bis zu einer bestimmten Größe wachsen oder bis die Knoten am Ende der Bäume eine bestimmte Reinheit aufweisen.

Um den Baum unabhängiger von statistischen Schwankungen in den Trainingsereignissen für Signal und Untergrund zu gestalten, existieren verschiedene Methoden.

Eine Methode ist das so genannte *Bagging*, bei dem das für einen Trainingsdurchlauf verwendete Ereignis zufällig aus der Menge aller Trainingsereignisse ausgewählt wird. Dies ist äquivalent dazu, dass man annimmt, dass die so ausgewählten Ereignisse, die zum Training eines bestimmten Baumes herangezogen wurden, die Wahrscheinlichkeits-Dichtefunktionen des gesamten Trainingsensembles repräsentieren. Denn durch diese Vorgehensweise werden Ereignisse aus einem stark bevölkerten Teil des Variablenraums öfter ausgewählt als andere. Da das selbe Ereignis mehrmals zufällig ausgewählt werden kann, wird also jeder Baum mit einem anderen Ensemble an Ereignissen trainiert. Dadurch wird eine größere statistische Unabhängigkeit der Gesamtmenge der Bäume erreicht. In der Implementation in TMVA werden die Trainingsereignisse für einen bestimmten Baum nicht zufällig ausgewählt, sondern jedes Ereignis mit einem zufälligen Gewicht versehen.

Eine andere Methode ist das *Boosting* mit dem Algorithmus „AdaBoost“. Bei diesem Algorithmus wird ein Entscheidungsbaum trainiert, mit dessen Ergebnisse durch Umgewichtung der Trainingsereignisse ein zusätzlicher neuer Baum trainiert wird. Landet zum Beispiel ein Signaleignis in einem Endknoten, der hauptsächlich Untergrundereignisse enthält, so wird dieses

Ereignis beim Trainieren des nächsten Baumes stärker gewichtet. Der Anfangswert der Gewichte  $w_i$  ist bei  $N$  Ereignissen  $1/N$ . Der Gewichtungsfaktor  $\alpha$  der Gewichte bei erneutem Training für ein falsch zugeordnetes Ereignis ergibt sich aus der Fehlerrate  $err$  des vorherigen Baums:

$$\alpha = \frac{1 - err}{err}. \quad (6.17)$$

Die Fehlerrate  $err$  für einen bestimmten Baum ist dabei definiert als

$$err = \frac{\sum_{\text{alle Ereignisse } n} w_n h_n(x_i)}{\sum_n w_n}. \quad (6.18)$$

Für die Funktion  $h$  soll dabei gelten

$$h_n(x_i) = \begin{cases} 0 & \text{falls ein Signalereignis } n \text{ mit dem Variablensatz } x_i \\ & \text{in einem Signalknoten endet,} \\ 1 & \text{falls ein Signalereignis } n \text{ mit dem Variablensatz } x_i \\ & \text{in einem Untergrundknoten endet.} \end{cases} \quad (6.19)$$

Für Untergrundereignisse gilt das Entsprechende. Durch diese Funktion werden in der Fehler-rate nur Ereignisse berücksichtigt, die den Knoten falsch zugeordnet wurden. Auch die Gewichte der Ereignisse werden nur korrigiert, falls das entsprechende Ereignis falsch zugeordnet wurde. Durch dieses Verfahren wird die Trennungsfähigkeit der darauffolgenden Bäume erhöht, da Ereignisse, die in den falschen Endknoten landen, besonders berücksichtigt werden. Nach dieser Berechnung der neuen Gewichte müssen diese wieder normiert werden:  $w_i \rightarrow w_i / \sum w_i$ .

Um nun nach diesem Trainieren der Bäume statistisch insignifikante Knoten zu entfernen und das Übertrainieren der Bäume weiter zu minimieren, können bestimmte Knoten wieder von den Bäumen abgetrennt werden. Dieser Prozess wird Beschneidung genannt (im Englischen „pruning“). Verschiedene Methoden des Beschneidens sind zum Beispiel:

- *Schneiden nach dem erwarteten Fehler* (im Englischen „expected error pruning“): Der statistische Fehlerabschätzer eines Knotens wird mit Hilfe des binomialen Fehlers

$$\sqrt{p + (1 - p)/N},$$

mit  $p$  als Reinheit (s.o.) und  $N$  als Anzahl der Trainingsereignisse in diesem Knoten, berechnet. Alle Knoten, welche eine kleinere statistische Fehlerabschätzung besitzen als die kombinierte Fehlerabschätzung der beiden Töchterknoten, werden rekursiv entfernt. Die Stärke der Beschneidung wird geregelt, indem dieser statistische Fehlerabschätzer mit einer Konstanten  $c$ , der Beschneidungsstärke, multipliziert wird. Konkret bedeutet dies, dass ein Knoten vom Baum entfernt wird, falls für seinen Fehlerabschätzer  $E$  gilt ([24, 45])

$$c \cdot E(\text{Knoten}) < P_1 \cdot E_1 + P_2 \cdot E_2,$$

wobei  $E_1$  und  $E_2$  die Fehlerabschätzer der beiden Tochterknoten sind.  $P_1$  und  $P_2$  sind die Wahrscheinlichkeiten, dass ein Ereignis an den jeweiligen Tochterknoten weitergereicht wird. Diese können mit Hilfe der Trainingsereignisse, die die Mutterknoten erreichen, und den Trainingsereignissen, die die jeweiligen Tochterknoten erreichen, abgeschätzt werden.

- *Cost Complexity Beschneidung*: Hier wird der Nutzen einer weiteren Aufspaltung betrachtet. Mit einem Kostenabschätzer  $R = 1 - \max(p, 1 - p)$  mit der Reinheit  $p$  ergibt sich die Cost Complexity zu:

$$\rho = \frac{R(\text{Knoten}) - R(\text{Unterbaum unterhalb des Knotens})}{(\text{Anzahl Knoten des Unterbaums}) - 1} \quad (6.20)$$

Die Knoten mit den kleinsten Werten  $\rho$  werden rekursiv entfernt, solange  $\rho < c = \text{const.}$  ist. Die Konstante  $c$  wird wieder Beschneidungsstärke genannt und bestimmt auch, wie stark ein Baum übertrainiert werden kann. Mit der Cost Complexity Beschneidung wird also versucht, den Nutzen gegen den Aufwand weiterer Spaltungen abzuwägen.

Sind alle Bäume fertig trainiert und beschnitten, so wird aus allen Bäumen eine entgeltige Klassifizierungsvariable  $y_{BDT}$  berechnet, wobei jedes Einzelergebnis der verschiedenen Bäume mit der Fehlerrate des entsprechenden Baumes gewichtet wird:

$$y_{BDT} = \sum_{\text{Alle Bäume } n} \alpha_n g_n(x_i). \quad (6.21)$$

$\alpha_n$  ist hier wiederum das Gewicht des jeweiligen Baums  $n$  (Siehe hierzu die Gleichung 6.17).  $g_n(x_i)$  ist 1, falls das Ereignis  $i$  in einem Signalknoten eingeordnet wird, sonst -1. Die Vorteile der Entscheidungsbäume sind, dass Korrelationen der Variablen die Leistung dieser Methode nicht negativ beeinflussen und dass sie nicht sehr anfällig auf Variablen mit geringer Trennungskraft sind. Denn solche Variablen werden bei einer Weggabelung an einem Knoten während des Trainings einfach nicht berücksichtigt. Andererseits sind Entscheidungsbäume relativ anfällig auf ein Übertrainieren und die Ergebnisse sind aufgrund der Größe der Bäume und ihrer Anzahl <sup>8)</sup> nicht durchsichtig.

<sup>8)</sup>Gewöhnlicherweise werden mehrere hundert Bäume trainiert. Die Größe der einzelnen Bäume hängt stark von der Beschneidungsstärke und der Anzahl der verwendeten Variablen ab.



## 7 Eine Studie zum Nachweis eines Higgs-Bosons im Phion-Modell

Im Folgenden wird eine Studie zum Zerfall eines Higgs-Bosons in nicht nachweisbare Phionen des Phion-Modells präsentiert. Es wird untersucht, wie groß das Potential des ATLAS-Experimentes ist, ein solches Higgs-Boson im assoziierten Higgs-Produktionskanal mit einem Z-Boson und im Vektorboson-Fusionskanal zu entdecken. Die Tatsache des unsichtbaren Zerfalls macht eine solche Analyse jedoch sehr schwierig. So ist es nicht möglich, Teilchen aus dem Zerfall und somit direkt ein Massenspektrum zu analysieren. Stattdessen müssen Zählexperimente durchgeführt werden. Das bedeutet, dass man Ereignisse einer bestimmten Signatur zählt, und einen signifikanten Überschuss an Ereignissen über dem gegebenen Untergrund aus Standardmodell-Prozessen als Signal interpretiert. Somit ist die genaue Kenntnis der Untergründe für eine solche Analyse extrem wichtig, jedoch auch sehr schwierig. Durch Simulationen erhofft man sich, das Verhalten der Signal- und Untergrundereignisse und deren Wirkungsquerschnitte so wirklichkeitsgetreu wie nur möglich nachzubilden. Mit dieser Kenntnis soll es im Experiment gelingen, das Signal in Form von zusätzlichen Ereignissen über dem Untergrund zu erkennen. Jedoch folgt hieraus auch, dass ein Erfolg auch davon abhängt, wie gut man die Physik der Natur kennt, wie genau man sie theoretisch beschreiben kann und wie gut die Monte-Carlo-Modelle sind, mit denen man die Prozesse simuliert. Um diese Abhängigkeit von den theoretischen Unsicherheiten und den Monte-Carlo-Simulationen zu verringern, werden die Standardmodell-Untergründe später im Experiment direkt aus Daten abgeschätzt. Dies erreicht man dadurch, dass man ähnliche und leicht zu rekonstruierende Prozesse misst und die Ergebnisse auf die relevanten Prozesse überträgt<sup>1)</sup>. Die Unsicherheiten, die man dadurch auf die Untergrundprozesse erhält, hängen dann nur noch von gut bekannten oder gut beschreibbaren theoretischen Größen ab und fallen daher wesentlich geringer aus, als die vorherigen Unsicherheiten aus den Simulationen.

Weiterhin ist man im Experiment darauf angewiesen, die fehlende transversale Energie genau rekonstruieren zu können. Diese fehlende transversale Energie ist zum Großteil die Energie der nicht nachweisbaren Teilchen und daher ein zentrales Element jeder Analyse eines unsichtbaren Zerfalls<sup>2)</sup>.

---

<sup>1</sup>Ein Beispiel hierfür ist die Abschätzung des Prozesses  $pp \rightarrow ZZ \rightarrow l^+l^- \nu\bar{\nu}$  durch Messung des Prozesses  $pp \rightarrow ZZ \rightarrow l^+l^-l^+l^-$ .

<sup>2</sup>Ein Teil zur fehlenden transversalen Energie tragen auch Ineffizienzen im Detektor und Ineffizienzen bei der Rekonstruktion der Teilchen und ihrer Energien bei. Weiterhin können Standardmodell-Zerfälle beitragen, bei denen Neutrinos auftreten. Beispiele hierfür sind semi-leptonische Zerfälle von Quarks oder  $Z \rightarrow \nu\bar{\nu}$ -Zerfälle.

## 7.1 Ereignisgenerierung und Detektor-Simulation

Die Ereignisse für diese Analyse wurden mit dem Monte-Carlo Generator PYTHIA (Version 6.226) erzeugt. Für jede Kombination der hier betrachteten Massen des Higgs-Bosons mit den hier verwendeten Kopplungen des Higgs-Bosons an die Phionen wurden 80000 Ereignisse simuliert. Diese große Anzahl stellt sicher, dass die Ergebnisse nicht sensitiv auf statistische Schwankungen in den generierten Ereignissen sind.

Der ATLAS-Detektor wurde mit dem Programm ATLFAST simuliert, welches im Unterkapitel 5.6 beschrieben wurde. Um Nachweiseffizienzen für die Leptonen zu berücksichtigen, wurde für jedes Lepton eine Nachweiswahrscheinlichkeit von 90% festgelegt. Um den Zerfall des Higgs-Bosons in nicht nachweisbare Teilchen zu simulieren, wurde dieses als unsichtbar für den ATLAS-Detektor deklariert und alle Zerfälle des Higgs-Bosons abgeschaltet.

Um die Breite des Higgs-Bosons zu modifizieren und somit eine Kopplung an die Phionen zu simulieren, wurden die Variablen  $PARP(110)$  und  $MSTP(110)$  verändert. So war es möglich, beliebige Breiten des Higgs-Bosons einzustellen. Der Wirkungsquerschnitt der Prozesse ist jedoch sensitiv auf die Breite des Higgs-Bosons. Diese Abhängigkeit kann mit dem Programm PYTHIA nicht studiert werden. Der Grund hierfür ist, dass Wirkungsquerschnitte in PYTHIA unter Annahme der sogenannten „Narrow-Width-Approximation“ berechnet werden. Hier wird davon ausgegangen, dass die Breite eines ausgetauschten Teilchens vernachlässigbar klein ist. Weiterhin werden eventuelle Interferenzterme des Matrixelementes bei der Berechnung der Wirkungsquerschnitte nicht berücksichtigt. Mit dieser Näherung können die Wirkungsquerschnitte vieler bekannter Prozesse sehr gut beschrieben werden. Für Teilchen mit großer Breite ist diese Näherung jedoch nicht mehr gültig. Daher sind die von PYTHIA berechneten Wirkungsquerschnitte für große Breiten des Higgs-Bosons keine verlässlichen Angaben mehr [46].

Eine Studie über die Veränderungen der Wirkungsquerschnitte mit Hilfe eines angepassten Programms war im Rahmen dieser Arbeit nicht möglich. Daher wird im Folgenden für alle Prozesse stets der Standardmodell-Wirkungsquerschnitt für den jeweiligen Prozess angenommen.

Die mit PYTHIA berechneten Produktions-Wirkungsquerschnitte für die beiden Kanäle, die hier betrachtet werden, sind für die berücksichtigten Massen in Tabelle 7.1 zusammengefasst. Weiterhin existiert für die Vektorboson-Fusion eine Studie [47], welche die Rekonstruktionseffizienzen einer kompletten Simulation des ATLAS-Detektors (GEANT) für Jets im Vorwärtsbereich mit denen einer ATLFAST-Simulation vergleicht. Die Ergebnisse aus dieser Studie wurden hier verwendet, um die transversalen Impulse der Jets zu korrigieren.

Eine Schwierigkeit für diese Analyse wird durch die Abbildungen 7.1 und 7.2 aufgezeigt. Diese Abbildungen zeigen die Verteilung der von PYTHIA aktuell generierten Massen des Higgs-Bosons für beide Produktionsmechanismen, wobei eine Masse des Higgs-Bosons von  $m_{Higgs} = 130 \text{ GeV}/c^2$  und eine Breite von  $\Gamma_{Higgs} = 4.79 \text{ GeV}/c^2$  auf der jeweils linken, bzw.  $\Gamma_{Higgs} = 130 \text{ GeV}/c^2$  auf der jeweils rechten Seite der Abbildungen angenommen wurde. Wie an den Abbildungen zu erkennen ist, sind die Breiten der Verteilungen deutlich kleiner als die nominellen großen Breiten. Dieses Verhalten kann folgendermaßen erklärt werden: Für den assozi-

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$\sigma_{ZH} * BR(Z \rightarrow ll)$ [fb]	60.8	46.2	35.2	27.1	21.6	18.4	6.83	2.92	1.47	0.49
$\sigma_{VBF}$ [pb]	4.59	4.32	3.90	3.61	3.36	3.17	2.44	1.81	1.38	0.84

Tabelle 7.1: Wirkungsquerschnitte der Produktion eines Higgs-Bosons im Standardmodell in führender Ordnung der Störungsrechnung, wie sie von PYTHIA berechnet wurden. Die Abkürzung ZH steht für die assoziierte Higgs-Produktion mit einem Z-Boson, VBF steht für die Vektorboson-Fusion. Für die assoziierte Higgs-Produktion wurde weiterhin das Verzweigungsverhältnis von  $Z \rightarrow ll$  ( $l = e, \mu$ ) berücksichtigt.

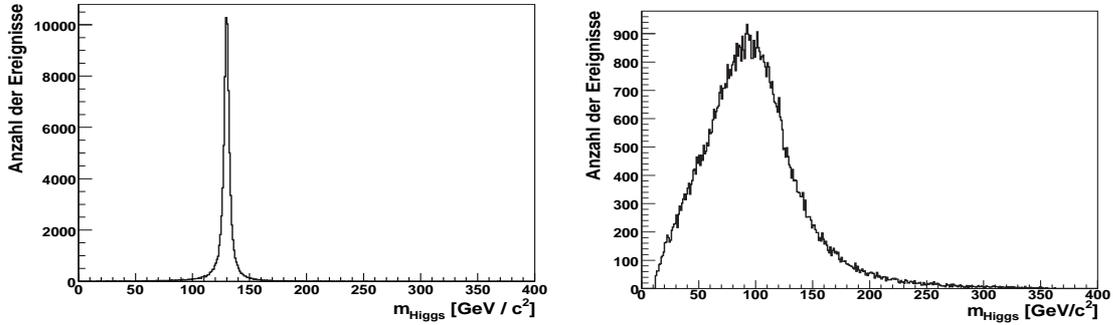


Abbildung 7.1: Generierte Massen des Higgs-Bosons für die assoziierte Higgs-Produktion für eine nominelle Masse des Higgs-Bosons von  $m_{Higgs} = 130 \text{ GeV}/c^2$  und einer Breite von  $\Gamma_{Higgs} = 4.79 \text{ GeV}/c^2$  (links) und  $\Gamma_{Higgs} = 130 \text{ GeV}/c^2$  (rechts).

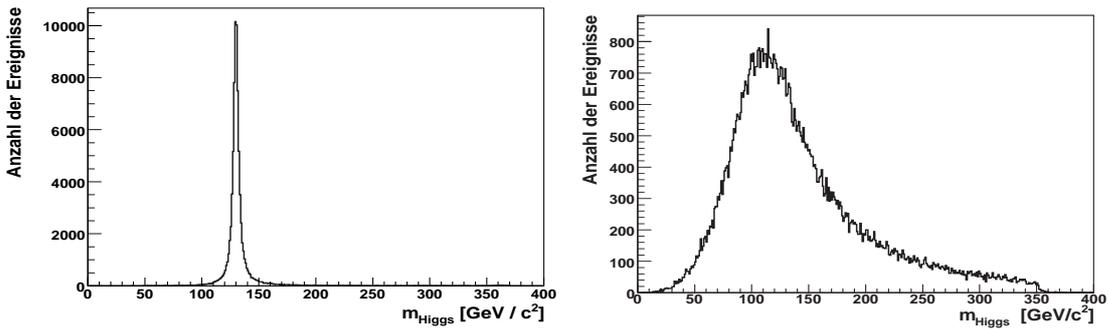


Abbildung 7.2: Generierte Massen des Higgs-Bosons für die Vektorboson-Fusion für eine nominelle Masse des Higgs-Bosons von  $m_{Higgs} = 130 \text{ GeV}/c^2$  und einer Breite von  $\Gamma_{Higgs} = 4.79 \text{ GeV}/c^2$  auf der linken Seite, bzw.  $\Gamma_{Higgs} = 130 \text{ GeV}/c^2$  auf der rechten Seite.

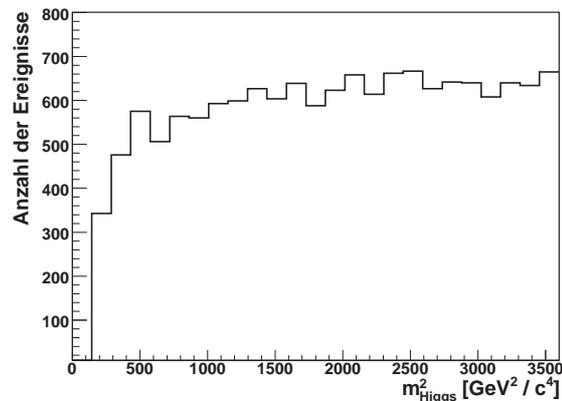


Abbildung 7.3: Die Anzahl der generierten Ereignisse über dem Quadrat der aktuell generierten Masse des Higgs-Bosons für den assoziierten Produktionskanal des Higgs-Bosons.

ierten Produktionskanal wird der Wirkungsquerschnitt von PYTHIA durch Faltung der Strukturfunktionen der beteiligten Partonen und den Breit-Wigner-Verteilungen für das Higgs-Boson und das Z-Boson berechnet. Die Formen der Strukturfunktionen für die See-Quarks und die Gluonen bewirken, dass der Wirkungsquerschnitt bei einem Higgs-Boson mit hoher Breite für kleinere Massen des Higgs-Bosons größer ist, wodurch diese öfter generiert werden. Gleiches gilt auch für den Vektorboson-Fusionskanal, bei dem die Strukturfunktionen der beteiligten Partonen, die Breit-Wigner-Verteilungen der Eichbosonen und die Breit-Wigner-Verteilung des Higgs-Bosons an der Berechnung des Wirkungsquerschnitts beteiligt sind.

Auch das starke Abfallen der Verteilung der aktuell generierten Massen im Bereich  $m < 30 \text{ GeV}/c^2$  im assoziierten Produktionskanal kann erklärt werden. Diese fällt ab, da der Wirkungsquerschnitt für diesen Kanal proportional zum Quadrat der Masse des Higgs-Bosons ist [48]. Dies erkennt man, indem die Anzahl der generierten Ereignisse über dem Quadrat der aktuell generierten Masse des Higgs-Bosons aufgetragen werden. Für eine Masse des Higgs-Bosons von  $m_{Higgs} = 130 \text{ GeV}/c^2$  und einer Breite von  $\Gamma_{Higgs} = 130 \text{ GeV}/c^2$  ist dies in Abbildung 7.3 gezeigt.

Zusätzlich existiert in PYTHIA ein Schnitt auf die invariante Masse von Teilchen. Diese invariante Masse muss mindestens  $12 \text{ GeV}/c^2$  betragen [46].

## 7.2 Die assoziierte Higgs-Produktion mit einem Z-Boson

Der erste Kanal der untersucht wurde ist die assoziierte Higgs-Produktion mit einem Z-Boson, welche folgende Signatur aufweist und durch den Feynman-Graphen in Abbildung 2.6 (rechts)

dargestellt wird:

$$q\bar{q} \longrightarrow ZH \longrightarrow ll + p_t^{miss} \quad , \text{ mit } l = e, \mu$$

Dies resultiert in zwei isolierten Leptonen mit hohem transversalen Impuls, welche durch den Zerfall des Z-Bosons entstehen, und einem großen fehlenden transversalen Impuls, welcher durch den Zerfall des Higgs-Bosons in nicht nachweisbare Phionen entsteht. Zusätzlich zur assoziierten Produktion mit einem Z-Boson kann auch die assoziierte Produktion mit einem W-Boson betrachtet werden. Doch wie eine bereits durchgeführte Studie [49] zeigte, ist es nicht möglich, das Signal vom irreduziblen Untergrund aus Drell-Yan-Prozessen zu extrahieren. Daher wird im Folgenden nur die assoziierte Produktion mit einem Z-Boson berücksichtigt.

Die relevanten Untergrundprozesse zu diesem Prozess sind:

- ZZ-Produktion, mit  $ZZ \rightarrow ll\nu\nu$ . Dieser Prozess liefert die gleiche Signatur wie das Signal und kann daher nur schwer vom Signal getrennt werden. Er stellt einen sogenannten irreduziblen Untergrund dar, da die Signatur exakt der des Signals entspricht.
- WZ-Produktion, mit  $Z \rightarrow ll$  und  $W \rightarrow l\nu$  oder  $W \rightarrow \tau\nu$  im Falle, dass das Lepton aus dem W-Zerfall nicht nachgewiesen werden kann. Weiterhin tragen Ereignisse, bei denen das  $\tau$ -Lepton hadronisch zerfällt und nicht von einem Jet aus einer Gluonabstrahlung unterschieden werden kann, zum Untergrund bei. Solche Jets aus einer Gluonabstrahlung im Anfangszustand (Initial State Radiation) treten auch beim Signal auf.
- Der Prozess  $WW \rightarrow l^+ \nu l^- \bar{\nu}$ . Da dieser Untergrund keine  $ll$ -Resonanz bei der Masse des Z-Bosons aufweist, lässt er sich gut vom Signal trennen.
- Der Prozess  $t\bar{t} \rightarrow WWb\bar{b} \rightarrow l\nu l\nu b\bar{b}$  trägt aufgrund des hohen Wirkungsquerschnitts und des leptonicen Zerfalls des t-Quarks einen großen Teil zum Untergrund bei. Doch auch dieser Untergrund lässt sich wegen der fehlenden Resonanz im Massenspektrum gut unterdrücken.
- Inklusive Produktion eines Z-Bosons:  $q\bar{q} \rightarrow \gamma/Z^* \rightarrow ll$ . Auch diese Ereignisse tragen wegen des hohen Wirkungsquerschnitts zum Untergrund bei. Hauptsächlich gilt es, die Ereignisse mit schweren Off-Shell <sup>3)</sup>  $Z \rightarrow \tau\tau$  Zerfällen, bei denen die  $\tau$ -Leptonen leptonic zerfallen, oder  $Z \rightarrow \mu\mu$  Ereignissen, falls die Energie der Myonen nicht korrekt rekonstruiert wird, zu unterdrücken.

Da bereits eine Analyse zum unsichtbaren Zerfall des Higgs-Bosons im assoziierten Higgs-Produktionskanal vorliegt [50], wurden die Untergrundprozesse nicht noch einmal generiert. Stattdessen wurden die Schnitte zur Separation von Signal und Untergrund und die Wirkungsquerschnitte der Untergründe aus Referenz [50] übernommen. Dieses Vorgehen ist gerechtfertigt, da im Rahmen dieser Arbeit die Signaleffizienzen aus Referenz [50] rekonstruiert werden konnten. Diese Schnitte sind:

<sup>3)</sup>Off-Shell bedeutet, dass für ein Teilchen mit Vierer-Impulsvektor  $p_\mu$  und Ruhemasse  $m$  gilt:  $p_\mu p^\mu \neq m^2 c^4$

- *Schnitt 1:* Ereignisse mit exakt zwei identifizierten und isolierten Leptonen, welche ein Paar aus Teilchen und Anti-Teilchen sind, wurden selektiert. Die Elektronen mussten einen transversalen Impuls von mindestens 15 GeV/c aufweisen, die Myonen einen von mindestens 10 GeV/c. Alle Leptonen mussten innerhalb der Pseudorapiditätsgrenze von  $|\eta| < 2.5$  liegen. Ereignisse mit einem dritten Lepton wurden zurückgewiesen. Die Nachweiseffizienz für jedes Lepton wurde zu 90% angenommen.
- *Schnitt 2:* Ein Schnitt auf die invariante Masse des Leptonenpaares wurde angewandt. Die Grenzen dieser Masse wurden auf  $m_Z \pm 10 \text{ GeV}/c^2$  gesetzt. Dieser Schnitt unterdrückt alle Untergründe.
- *Schnitt 3:* Ein Schnitt auf den fehlenden transversalen Impuls konnte einen großen Teil der Untergründereignisse unterdrücken:  $p_T^{miss} > 100 \text{ GeV}/c$ .
- *Schnitt 4:* Alle Ereignisse mit einem Jet mit den Eigenschaften  $p_T > 30 \text{ GeV}/c$  und  $|\eta| < 4.9$  wurden zurückgewiesen (Jet-Veto). Dieser Schnitt soll den  $t\bar{t}$ -Untergrund unterdrücken.
- *Schnitt 5:* Zusätzlich zum Jet-Veto wurde ein Veto auf b-Jets angewandt, welches die Unterdrückung des  $t\bar{t}$ -Untergrundes stark verbessert. Diese b-Jets mussten einen Impuls von mindestens  $p_T > 15 \text{ GeV}/c$  aufweisen. Die Nachweiseffizienzen für b-Jets wurden folgendermaßen angenommen: 60% für b-Jets, 10% für c-Jets und 1% für Jets, die von den leichteren Quarks u, d und s oder von Gluonen stammen.
- *Schnitt 6:* Ein Schnitt auf die transversale Masse des Systems, welches aus den zwei Leptonen und dem Vektor  $p_T^{miss}$  besteht, wurde angewandt:  $m_T > 200 \text{ GeV}/c^2$ . Die transversale Masse dieses Systems ist definiert als

$$m_T = \sqrt{2p_T^l p_T^{miss} * (1 - \cos(\Delta\phi))}$$

$\Delta\phi$  ist hier der azimutale Winkel zwischen dem fehlenden transversalen Impuls  $\vec{p}_T^{miss}$  und der Richtung des Di-Lepton-Systems.

Tabelle 7.2 vergleicht die Signaleffizienzen nach den jeweiligen Schnitten in dieser Analyse mit denen aus Referenz [50]. Grundlage hierfür ist ein Higgs-Boson im Standardmodell mit der Masse  $m_{Higgs} = 120 \text{ GeV}/c^2$ . Der Vergleich legt nahe, dass ein Übernehmen der Untergründereignisraten aus [50] gerechtfertigt ist. Der Unterschied in der Akzeptanz des Jet-Vetos (Schnitt 4) resultiert aus den unterschiedlichen PYTHIA-Versionen. Im Vergleich zu der in dieser Analyse verwendeten PYTHIA Version 6.226 wurde in der in Referenz [50] verwendeten PYTHIA Version 6.319 das Modell für die Mehrfach-Wechselwirkung (Multiple Interaction-Modell) grundlegend verändert. Eine Beschreibung dieses neuen Ansatzes für das Multiple Interaction-Modell findet man zum Beispiel in Referenz [51].

Die Wirkungsquerschnitte für die Untergründe vor und nach den Schnitten sind in Tabelle 7.3

	Schnitt 1	Schnitt 2	Schnitt 3	Schnitt 4	Schnitt 5	Schnitt 6
Rel. Effizienz dieser Analyse	0.445	0.892	0.304	0.744	0.984	0.957
Rel. Effizienz in Ref. [50]	0.448	0.896	0.313	0.652	0.988	0.952
Abs. Effizienz dieser Analyse	0.445	0.397	0.120	0.090	0.088	0.084
Abs. Effizienz in Ref. [50]	0.448	0.401	0.125	0.082	0.081	0.077

Tabelle 7.2: Effizienzen der Schnitte beim assoziierten Produktionskanal mit einer nominalen Masse des Higgs-Bosons von  $m_{Higgs} = 120 \text{ GeV}/c^2$  im Vergleich zu den in Referenz [50] erhaltenen. Die relativen Effizienzen beziehen sich jeweils auf den vorherigen Schnitt.

$\sigma$ [fb]	ZZ	WZ	WW	$t\bar{t}$	Z
Vor den Schnitten	$1.1 \cdot 10^4$	$2.7 \cdot 10^4$	$7.0 \cdot 10^4$	$4.9 \cdot 10^6$	$5.9 \cdot 10^8$
nach den Schnitten	10.7	3.97	0.77	0.31	$< 0.03$
Effizienz [%]	0.097	0.015	$1.1 \cdot 10^{-3}$	$6.3 \cdot 10^{-6}$	$\simeq 0$

Tabelle 7.3: Wirkungsquerschnitte der Standardmodelluntergründe für den assoziierten Higgs-Produktionskanal vor und nach den Schnitten. Diese wurden aus Referenz [50] übernommen.

aufgelistet.

Um die späteren Ergebnisse für das Entdeckungspotential bei großen Breiten des Higgs-Bosons interpretieren zu können, wurde die totale Schnittakzeptanz als Funktion der aktuell generierten Masse des Higgs-Bosons analysiert. Abbildung 7.4 zeigt, dass diese Akzeptanz für kleine generierte Massen des Higgs-Bosons sehr gering ist. Dieses Verhalten ist hauptsächlich durch zwei Schnitte geprägt:

- Der Schnitt auf den fehlenden transversalen Impuls: Abbildung 7.5 (links) zeigt, dass für kleine aktuell generierte Massen des Higgs-Bosons dieses hauptsächlich mit kleinem transversalen Impuls produziert wird. Dies führt dazu, dass diese Ereignisse vom Schnitt auf den fehlenden transversalen Impuls abgewiesen werden.
- Der Schnitt auf die invariante Masse des Di-Lepton-Systems, welche für das Signal in der Nähe der Masse des Z-Bosons erwartet wird: Für kleine Massen des Higgs-Bosons besteht die Möglichkeit, dass das Z-Boson On-Shell produziert wird und dann das Higgs-Boson abstrahlt. Dieser Sachverhalt wird durch Abbildung 7.5 (rechts) deutlich. Gleichzeitig bedeutet dies aber, dass das danach produzierte Di-Lepton-System eine invariante Masse besitzt, die sich stark von der On-Shell-Masse des Z-Bosons unterscheiden kann. Daher werden solche Ereignisse vom Schnitt auf die invariante Masse des Di-Lepton-Systems abgewiesen.

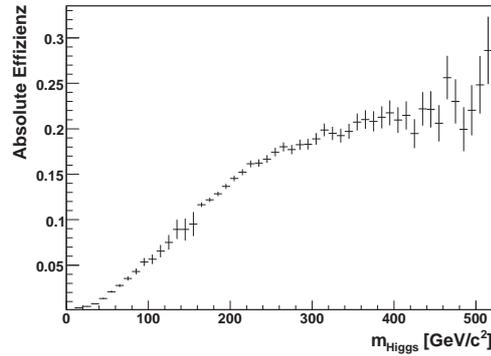


Abbildung 7.4: Totale Schnittakzeptanz bei der assoziierten Higgs-Produktion als Funktion der für jedes Ereignis aktuell generierten Masse des Higgs-Bosons. Diese entspricht nicht notwendigerweise der nominellen Masse. Die Fehlerbalken zeigen die statistischen Fehler.

Für größere Breiten des Higgs-Bosons, wodurch die Massenverteilung näher zu Null verschoben wird, wird also eine schlechtere Akzeptanz der Schnitte für Signalereignisse erwartet. Auch das Spektrum des fehlenden transversalen Impulses bestätigt diese Annahme. Dieses Spektrum für eine nominelle Masse des Higgs-Bosons von  $m_{Higgs} = 130 \text{ GeV}/c^2$  ist für eine Standardmodell-Breite von  $\Gamma_{Higgs} = 4.79 \text{ GeV}/c^2$  und eine Breite von  $\Gamma_{Higgs} = 43 \text{ GeV}/c^2$ , was einer Kopplung von  $\omega = 3$  entspricht, in Abbildung 7.6 normiert dargestellt. Sie zeigt, dass aus der zuvor diskutierten Verschiebung des Massenspektrums des Higgs-Bosons in Richtung kleinerer aktuell generierter Massen eine Verschiebung der Verteilung des fehlenden transversalen Impulses folgt.

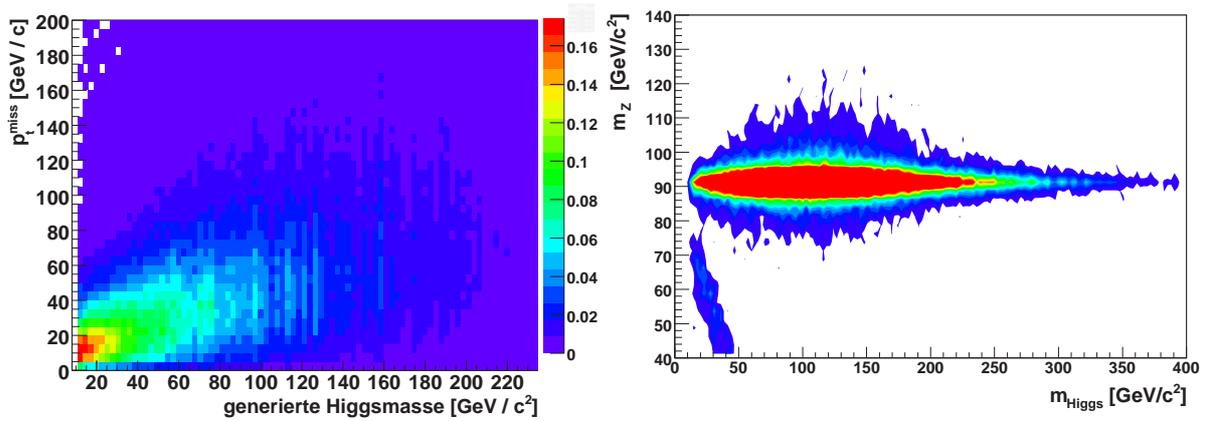


Abbildung 7.5: Links: Fehlender transversaler Impuls als Funktion der für jedes Ereignis aktuell generierten Masse des Higgs-Bosons. Jeder vertikale Streifen des Histogramms mit der Breite eines Kanals der Masse des Higgs-Bosons wurde auf die Anzahl der Ereignisse in diesem Streifen normiert. Rechts: Die aktuell generierte Masse des Z-Bosons als Funktion der aktuell generierten Masse des Higgs-Bosons nach der Abstrahlung des Higgs-Bosons für  $m_{Higgs} = 200 \text{ GeV}/c^2$  und  $\Gamma = 312 \text{ GeV}/c^2$ .

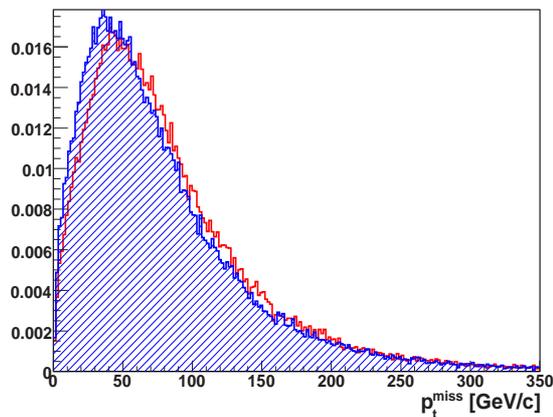


Abbildung 7.6: Normierte Verteilungen des fehlenden transversalen Impuls für ein Higgs-Boson der Masse  $m_{Higgs} = 130 \text{ GeV}/c^2$  für die assoziierte Higgs-Produktion. Die rote Verteilung zeigt Ereignisse mit einem Higgs-Boson mit Standardmodell-Breite  $\Gamma_{Higgs} = 4.79 \text{ GeV}/c^2$ . Die blaue Verteilung zeigt Ereignisse mit einem Higgs-Boson der Breite  $\Gamma_{Higgs} = 43 \text{ GeV}/c^2$ , was im Phion-Modell einer Kopplung von  $\omega = 3$  entspricht.

## 7.3 Die Vektorboson-Fusion

Der zweite Kanal, welcher hier diskutiert werden soll, ist die Vektorboson-Fusion. Dieser Prozess wird beschrieben durch:

$$qq \rightarrow qqVV \rightarrow qqH$$

Den dazugehörigen Feynman-Graph zeigt die Abbildung 2.6. Das Signal besteht aus zwei Jets, jeweils einer in Vorwärts- und Rückwärtsrichtung, und einem hohen fehlenden transversalen Impuls. Die relevanten Untergründe zu diesem Kanal sind:

- Die assoziierte Produktion eines Z-Bosons mit zwei Jets, bei der das Z-Boson in Neutrinos zerfällt.
- Die assoziierte Produktion eines W-Bosons mit zwei Jets. Dieser Kanal trägt zum Untergrund bei, wenn das W-Boson leptonic zerfällt und dieses Lepton nicht nachgewiesen werden kann.
- QCD Ereignisse mit mehreren Jets. Der fehlende transversale Impuls resultiert dabei aus nicht nachgewiesenen Jets, Jets mit falsch rekonstruierter Energie oder semileptonischen Zerfällen innerhalb der Jets.

Um diese Untergründe effektiv zu unterdrücken, werden Schnitte angewandt, welche wiederum aus einer bestehenden Analyse über unsichtbar zerfallende Higgs-Bosonen bei der Vektorbosonen-Fusion [52] übernommen wurden. Ereignisse werden als Signalereignisse angenommen, falls folgende Bedingungen erfüllt werden:

- *Schnitt 1:* Der fehlende transversale Impuls musste zur Vorselektion der Ereignisse größer als  $80 \text{ GeV}/c$  sein.
- *Schnitt 2:* Es mussten mindestens zwei Jets mit einem transversalen Impuls von  $p_T > 40 \text{ GeV}/c$  existieren. Die zwei Jets mit dem größten transversalen Impuls wurden ausgewählt und werden im Folgenden „Tag-Jets“  $j_1$  und  $j_2$  genannt. Sie mussten in verschiedenen Hemisphären liegen, d.h. für ihre Pseudorapiditäten musste gelten  $\eta_{j_1} * \eta_{j_2} < 0$ , und folgende Kriterien erfüllen:  
 $|\eta_{j_1}|, |\eta_{j_2}| < 5$  und  $|\eta_{j_1} - \eta_{j_2}| > 4.4$ .
- *Schnitt 3:* Die invariante Masse  $M_{jj}$  des Systems der beiden Tag-Jets musste größer als  $1200 \text{ GeV}/c^2$  sein.
- *Schnitt 4:* Aufgrund des Zerfalls in nicht nachweisbare Teilchen musste der fehlende transversale Impuls  $p_T^{miss} > 100 \text{ GeV}/c$  sein.
- *Schnitt 5:* Es durften keine weiteren Jets zwischen den Tag-Jets liegen, die einen transversalen Impuls von mehr als  $20 \text{ GeV}/c$  besitzen.

	Schnitt 1	Schnitt 2	Schnitt 3	Schnitt 4	Schnitt 5	Schnitt 6	Schnitt 7	Schnitt 8
Rel. Effizienzen in dieser Analyse	0.47	0.23	0.57	0.82	0.71	1.00	0.89	0.43
Rel. Effizienzen in Ref. [52]	0.47	0.23	0.56	0.83	0.80	1.00	0.88	0.47
Abs. Effizienzen in dieser Analyse	0.47	0.11	0.062	0.051	0.036	0.036	0.032	0.014
Abs. Effizienzen in Ref. [52]	0.47	0.11	0.061	0.050	0.040	0.040	0.035	0.017

Tabelle 7.4: Die in dieser Analyse bestimmten Signalakzeptanzen im Vergleich zu denen aus Referenz [52] für den Vektorboson-Fusionskanal

$\sigma$ [fb]	Wjj	Zjj	QCD
Vor den Schnitten	$6.7 \cdot 10^6$	$2.7 \cdot 10^6$	$\simeq 2 \cdot 10^{10}$
nach den Schnitten	121.5	123.0	$\simeq 0$
Effizienz [%]	$1.8 \cdot 10^{-3}$	$4.6 \cdot 10^{-3}$	$\simeq 0$

Tabelle 7.5: Wirkungsquerschnitte der wichtigsten Standardmodelluntergründe für den Vektorboson-Fusionskanal vor und nach Schnitten. Diese wurden aus Referenz [52] übernommen.

- *Schnitt 6*: Es wurde kein Ereignis zugelassen, bei dem ein Lepton mit den folgenden transversalen Impulsen nachgewiesen wurde:  $p_T > 5$  GeV/c für Elektronen,  $p_T > 6$  GeV/c für Myonen und  $p_T > 20$  GeV/c für  $\tau$ -Leptonen.
- *Schnitt 7*: Es wurde ein Schnitt auf die Isolationsvariable  $I$  der fehlenden Energie angewandt. Diese musste  $I > 1$  rad sein und ist folgendermaßen definiert:  

$$I = \min_{j=j_1, j_2} [\phi(P_T^{miss}) - \phi(j)].$$
Mit  $\phi$  ist der azimutale Winkel in der transversalen Ebene gemeint.
- *Schnitt 8*: Der azimutale Winkel zwischen den Tag-Jets muss folgende Bedingung erfüllen:  

$$\Delta\phi_{jj} = |\phi_{j_1} - \phi_{j_2}| < 1 \text{ rad.}$$

Auch die Wirkungsquerschnitte der Untergründe, welche in Tabelle 7.5 vor und nach den Schnitten aufgelistet sind, wurden aus Referenz [52] übernommen, nachdem die Signaleffizienzen reproduziert werden konnten. Ein Vergleich der in dieser Analyse bestimmten Signaleffizienzen mit denen aus Referenz [52] zeigt Tabelle 7.4.

Auch bei diesem Kanal sind kleinere Akzeptanzunterschiede in den Versionsunterschieden von PYTHIA begründet. Dies gilt ebenso für die deutlichen Unterschiede in den Schnittakzeptanzen des Schnittes 5 auf die Jets. In Referenz [52] wurde die PYTHIA Version 6.158 verwendet. Im Vergleich zu dieser Version wurden in der hier verwendeten PYTHIA Version 6.226 mehrere Veränderungen vorgenommen. So wurde unter anderem das Modell über die Abstrahlung von Gluonen im Anfangszustand (Initial State Radiation) verändert und die Massen von c- und b-Quark angepasst. Weiterhin wurden erst ab PYTHIA Version 6.2 die Strukturfunktionen CTEQ5L als Standard verwendet. Eine weitere große Veränderung in dieser Version war, dass viele Parameter des Modells über die Mehrfach-Wechselwirkung (Multiple Interaction) verändert wurden, um die Messungen von Minimum Bias-Ereignissen und Underlying Events

der Tevatron-Experimente *D0* und *CDF* zu beschreiben.

Wie zuvor schon erwähnt, wurde in dieser Analyse über den Vektorboson-Fusionskanal eine Energiekalibrierung der Jets aus Referenz [47] verwendet.

Um wiederum den Einfluss zu überprüfen, den die Vergrößerung der Breite des Higgs-Bosons auf die Akzeptanz der Schnitte haben wird, wurden auch hier die Schnittakzeptanzen als Funktion der für jedes Ereignis aktuell generierten Masse des Higgs-Bosons betrachtet. Das Ergebnis zeigt Abbildung 7.7 (links). Auch hier zeigt sich, dass die Schnittakzeptanzen in den Bereichen der kleineren Masse des Higgs-Bosons geringer ist. Der Grund hierfür ist wiederum der kleinere transversale Impuls des Higgs-Bosons, welcher gleichbedeutend ist mit dem kleineren fehlenden transversalen Impuls bei kleineren aktuell generierten Massen des Higgs-Bosons. Dieser Sachverhalt wird in Abbildung 7.7 (rechts) dargestellt. Auch bei diesem Kanal wird somit eine leichte Verringerung der Schnittakzeptanzen durch die Verschiebung des Massenspektrums bei einer Vergrößerung der Breite des Higgs-Bosons erwartet. Allerdings sollte diese deutlich geringer ausfallen, als es bei der assoziierten Produktion des Higgs-Bosons der Fall sein sollte, da sich die Massenverteilung des Higgs-Bosons für größere Breiten nur leicht in Richtung der Null verschiebt. Dieser Effekt der Verschiebung könnte allerdings dadurch ausgeglichen werden, da in der asymmetrischen Verteilung der Massen bei hohen Breiten auch die hohen Massen relativ stark vertreten sind. Für diese hohe Massen ist die Akzeptanz der Schnitte größer. Diese Vermutung legt auch die Verteilung des fehlenden transversalen Impulses nahe. In Abbildung 7.8 ist diese Verteilung für ein Higgs-Boson mit einer nominellen Masse von  $m_{Higgs} = 130 \text{ GeV}/c^2$  und einer Standardmodell-Breite von  $\Gamma_{Higgs} = 4.79 \text{ GeV}/c^2$ , beziehungsweise einer Breite von  $\Gamma_{Higgs} = 43 \text{ GeV}/c^2$ , was einer Kopplung von  $\omega = 3$  entspricht, normiert dargestellt. Die Verteilungen für diese beiden Breiten des Higgs-Bosons sind nahezu identisch, obwohl wegen der Verschiebung der Massenverteilung eine Verringerung erwartet wurde.

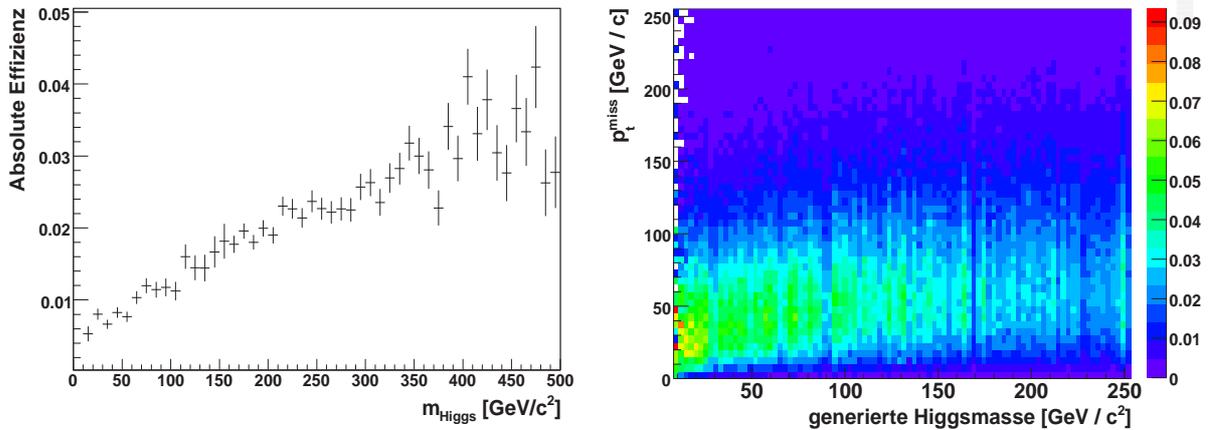


Abbildung 7.7: Links: Akzeptanz der Schnitte bei der Vektorboson-Fusion als Funktion der Masse des Higgs-Bosons, welche aktuell für die Ereignisse generiert wurde. Rechts: Der fehlende transversale Impuls als Funktion der aktuell generierten Masse des Higgs-Bosons. Jeder vertikale Streifen des Histogramms mit der Breite von einem Kanal der Masse des Higgs-Bosons wurde auf die Anzahl der Ereignisse in diesem Streifen normiert.

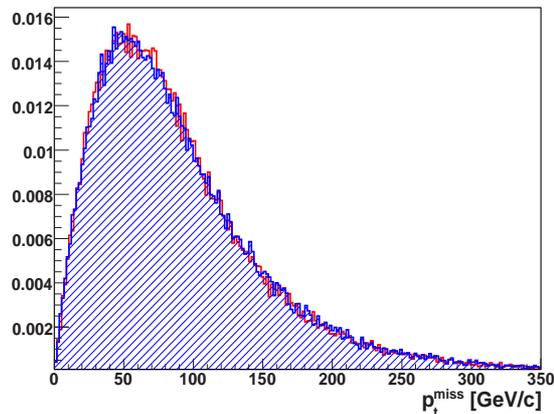


Abbildung 7.8: Normierte Verteilungen des fehlenden transversalen Impulses für ein Higgs-Boson der Masse  $m_{Higgs} = 130 GeV/c^2$  für die Vektorboson-Fusion. Die rote Verteilung zeigt Ereignissen mit einem Higgs-Boson mit Standardmodell-Breite  $\Gamma_{Higgs} = 4.79 GeV/c^2$ . Die blaue Verteilung zeigt Ereignisse mit einem Higgs-Boson der Breite  $\Gamma_{Higgs} = 43 GeV/c^2$ , was im Phion-Modell einer Kopplung von  $\omega = 3$  entspricht.

## 7.4 Die Signifikanz des Signals und das Entdeckungspotential

Die Signifikanz  $S$  des Signals kann aus der Anzahl der nachgewiesenen Signalereignisse  $N_S$  gegenüber der Anzahl  $N_B$  der erwarteten Standardmodell-Untergrundereignisse bestimmt werden. Hierzu kann zusätzlich ein systematischer Fehler  $\alpha$  berücksichtigt werden, der sich aus den Unsicherheiten über die Kenntnis der Untergründe ergibt. Dieser systematische Fehler kann bestimmt werden, indem die Auswirkungen der Schnitte einer Analyse auf Kanäle studiert werden, welche den Untergrundprozessen sehr ähnlich sind. Für den  $ZZ \rightarrow ll\nu\nu$ -Untergrund des assoziierten Produktionskanals des Higgs-Bosons bietet sich zum Beispiel eine Analyse des Kanals  $ZZ \rightarrow ll\mu\mu$  an. Diese Analyse kann später direkt mit Daten des Experiments durchgeführt werden, sodass nur noch eine geringe oder gar keine Abhängigkeit der Ergebnisse von theoretischen Größen und Monte-Carlo-Modellen mehr besteht. Die systematischen Fehler für die hier betrachteten Untergründe wurden schon in den Publikationen [50] und [52] studiert und zu einem gesamten Fehler kombiniert. Für Daten entsprechend einer integrierten Luminosität von  $30 \text{ fb}^{-1}$  ( $100 \text{ fb}^{-1}$ ) ergibt sich für die Untergründe des assoziierten Produktionskanals ein kombinierter Fehler von  $\alpha = 6.1\%$  ( $3.8\%$ ). Für die Vektorboson-Fusion und den Daten, die einer integrierten Luminosität von  $10 \text{ fb}^{-1}$  ( $30 \text{ fb}^{-1}$ ) entsprechen, wurde der kombinierte Fehler auf  $\alpha = 3.1\%$  ( $2.3\%$ ) abgeschätzt.

Für die Signifikanz  $S$  folgt unter Berücksichtigung des systematischen Fehlers auf die Untergründe:

$$S = \frac{N_S}{\sqrt{N_B + (\alpha N_B)^2}}, \quad (7.1)$$

statt der üblichen statistischen Signifikanz  $N_S/\sqrt{N_B}$ .

Da in verschiedenen Modellen auch verschiedene Produktionsmechanismen für das Higgs-Boson mit unterschiedlichen Wirkungsquerschnitten existieren, wurde eine modellunabhängige Variable eingeführt, die es erlaubt, das Entdeckungspotential des Higgs-Bosons beim ATLAS-Experiment abzuschätzen. Diese Variable wird definiert als

$$\xi^2 = BR(H \rightarrow inv.) * \frac{\sigma_{Prod. \text{ im neuen Modell}}}{\sigma_{Prod. \text{ im SM}}} \quad (7.2)$$

Für Modelle, in denen das Higgs-Boson gemäß dem Standardmodell produziert wird, bedeutet dies, dass  $\xi^2$  gleich dem Zerfallsverhältnis des Higgs-Bosons in nicht nachweisbare Teilchen ist. Um die Hypothese zu testen, ob eine Menge an Signalereignissen als statistische Fluktuation über dem erwarteten Untergrund interpretiert werden kann, wird das Vertrauensniveau, im englischen *confidence level*, definiert. Dieses gibt an, wie verträglich ein Messwert einer Variablen mit einer angenommenen Verteilung dieser Variablen ist. Soll ein Messwert mit einer Wahrscheinlichkeit von 95%, welche das Vertrauensniveau definiert, innerhalb des einseitigen Vertrauensintervalls liegen <sup>4)</sup>, so darf die Abweichung des Messwertes in eine Richtung nicht mehr als 1.65 Stan-

<sup>4)</sup>Es wird hier das einseitige Vertrauensintervall gewählt, da nur zusätzliche Ereignisse über dem Untergrund als Signal gewertet werden können.

Standardabweichungen vom Mittelwert der Verteilung betragen. Denn mit einer Normalverteilung  $G(x)$  der Variable  $x$  mit der Standardabweichung  $\sigma$  folgt aus

$$\int_{-\infty}^n G(x) dx = 0.95, \quad (7.3)$$

dass für  $n$  gelten muss:

$$n = 1.65 \sigma. \quad (7.4)$$

Genauer hierzu findet man zum Beispiel in [53].

Bei Kenntnis der Signifikanz einer Analyse kann die Ausschlussgrenze für  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% berechnet werden. Hierzu wird angenommen, dass die Anzahl der Untergrundereignisse bekannt und die Anzahl der Signalereignisse gaußförmig um den wahren Wert verteilt ist. Daraus folgt, dass die Signifikanz ebenfalls gaußförmig verteilt ist. Für die Ausschlussgrenzen von  $\xi^2$  ergibt sich dadurch [16]:

$$\xi^2 = \frac{S_{CL=95\%}}{S} = \frac{1.65}{S}. \quad (7.5)$$

In den Tabellen 7.6 und 7.7 sind Anzahl der erwarteten Signal- und Untergrundereignisse, sowie die Signifikanzen und die Ausschlussgrenzen auf die Variable  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% für ein Higgs-Boson im Standardmodell mit  $BR(H \rightarrow inv.) = 100\%$  für verschiedene Datenmengen und die in dieser Analyse betrachteten Kanäle und Massen des Higgs-Bosons aufgelistet.

In den Abbildungen 7.9 und 7.10 sind diese Ausschlussgrenzen für  $\xi^2$  für die verschiedenen Luminositäten und Kanäle als Funktion der Masse des Higgs-Bosons aufgetragen. Wird beim ATLAS-Experiment kein Überschuss an Ereignissen festgestellt, so kann ein Higgs-Boson, welches in der Natur mit einer Kombination aus  $BR(H \rightarrow inv.) * \sigma_{Prod. \text{ im neuen Modell}} / \sigma_{Prod. \text{ im SM}}$  realisiert ist, die über einer solchen Kurve liegt, zu 95% ausgeschlossen werden. Da im Phion-Modell das Higgs-Boson wie im Standardmodell produziert wird, können durch die Kurven direkt bestimmte Verzweigungsverhältnisse für verschiedene Massen ausgeschlossen werden.

Geht man andererseits von einem festen Verzweigungsverhältnis von  $BR(H \rightarrow inv.) = 100\%$  aus, so lassen sich die Kurven als Ausschlussgrenzen für mögliche Massen des Higgs-Bosons interpretieren. Alle Massen bis zu der Masse, bei der die Ausschlusskurve die Gerade  $f(m_{Higgs}) = 1$  schneidet, können so ausgeschlossen werden. Konkret bedeutet dies zum Beispiel für den assoziierten Higgs-Produktionskanal, dass nach Analyse der Daten, die einer integrierten Luminosität von  $30 \text{ fb}^{-1}$  entsprechen, eine Masse des Higgs-Bosons im Standardmodell bis zu  $m_{Higgs} \simeq 160 \text{ GeV}/c^2$  entsprechend eines Vertrauensniveaus von 95% ausgeschlossen werden kann.

## 7 Eine Studie zum Nachweis eines Higgs-Bosons im Phion-Modell

$m_{Higgs}$ [GeV/ $c^2$ ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	123.1	106.6	87.0	75.4	67.8	60.6	28.8	14.5	8.3	3.1
$N_{Bkg.}$	473.4									
Signifikanz $S$	3.41	2.95	2.41	2.08	1.87	1.68	0.80	0.40	0.23	0.09
$\xi^2_{95\% CL}$	0.48	0.56	0.69	0.79	0.88	0.98	2.07	4.11	7.17	19.21

$m_{Higgs}$ [GeV/ $c^2$ ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	410.5	355.5	289.9	251.3	225.8	202.1	95.8	48.4	27.7	10.4
$N_{Bkg.}$	1578.0									
Signifikanz $S$	5.71	4.94	4.03	3.49	3.14	2.81	1.33	0.67	0.39	0.14
$\xi^2_{95\% CL}$	0.29	0.33	0.41	0.47	0.53	0.59	1.24	2.45	4.28	11.46

Tabelle 7.6: Anzahl der erwarteten Signal- und Untergrundereignisse ( $N_{Sig.}$  und  $N_{Bkg.}$ ), die Signifikanzen und die Ausschlussgrenzen für  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% für die Entdeckung eines in nicht nachweisbare Teilchen zerfallenden Higgs-Bosons mit  $BR(H \rightarrow inv.) = 100\%$  im Standardmodell für die assoziierte Higgs-Produktion. Die Zahlen der oberen Tabelle entsprechen den erwarteten Datenmengen einer integrierten Luminosität von  $30 \text{ fb}^{-1}$ , die der unteren den Datenmengen einer integrierten Luminosität von  $100 \text{ fb}^{-1}$ . Die systematischen Fehler  $\alpha$  auf die Untergründe sind für diese Werte bereits berücksichtigt worden.

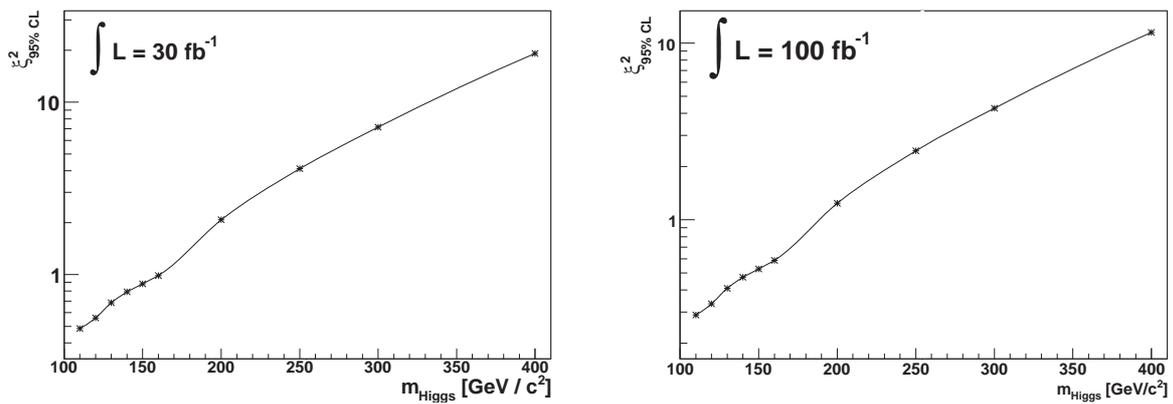


Abbildung 7.9: Ausschlussgrenzen im Standardmodell für den assoziierten Higgs-Produktionskanal auf 95% Vertrauensniveau für Daten, die einer integrierten Luminosität von  $30 \text{ fb}^{-1}$  (links) und  $100 \text{ fb}^{-1}$  (rechts) entsprechen.

$m_{Higgs}$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	617.1	536.2	551.3	573.3	547.1	537.2	499.0	440.4	367.2	271.8
$N_{Bkg.}$	2930.0									
Signifikanz $S$	5.84	5.07	5.21	5.42	5.17	5.08	4.72	4.17	3.47	2.57
$\xi^2_{95\% CL}$	0.28	0.33	0.32	0.30	0.32	0.32	0.35	0.40	0.48	0.64

$m_{Higgs}$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	1851.3	1608.5	1653.9	1719.8	1641.2	1611.4	1496.9	1321.3	1101.5	815.3
$N_{Bkg.}$	8790.0									
Signifikanz $S$	8.31	7.22	7.42	7.72	7.36	7.23	6.72	5.93	4.94	3.66
$\xi^2_{95\% CL}$	0.20	0.23	0.22	0.21	0.22	0.23	0.25	0.28	0.33	0.45

Tabelle 7.7: Anzahl der erwarteten Signal- und Untergrundereignisse ( $N_{Sig.}$  und  $N_{Bkg.}$ ), die Signifikanzen und die Ausschlussgrenzen für  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons mit  $BR(H \rightarrow inv.) = 100\%$  im Standardmodell für die Vektorboson-Fusion. Die Zahlen der oberen Tabelle entsprechen den erwarteten Datenmengen einer integrierten Luminosität von  $10 \text{ fb}^{-1}$ , die der unteren den Datenmengen einer integrierten Luminosität von  $30 \text{ fb}^{-1}$ . Die systematischen Fehler  $\alpha$  auf die Untergründe sind für diese Werte bereits berücksichtigt worden.

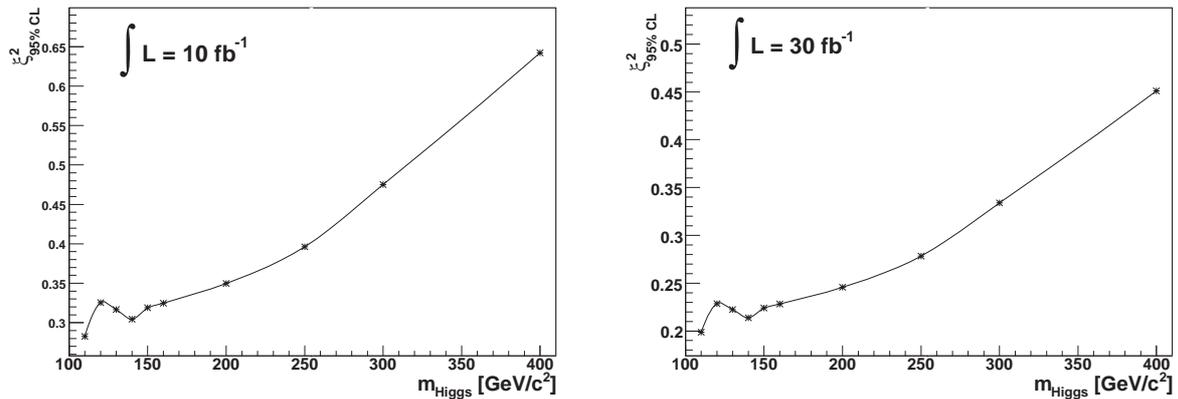


Abbildung 7.10: Ausschlussgrenzen im Standardmodell für den Vektorboson-Fusionskanal auf 95% Vertrauensniveau für Daten, die einer integrierten Luminosität von  $10 \text{ fb}^{-1}$  (links) und  $30 \text{ fb}^{-1}$  (rechts) entsprechen.

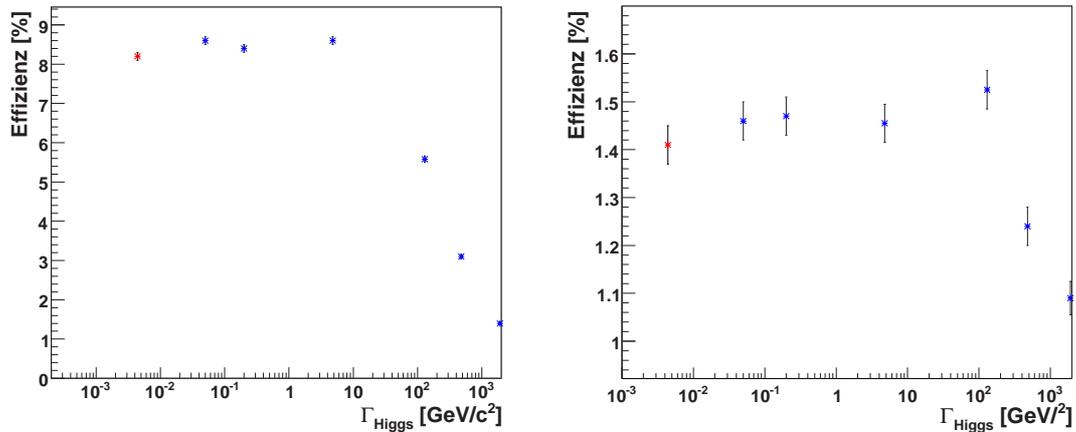


Abbildung 7.11: Schnittakzeptanz als Funktion der Breite des Higgs-Bosons mit  $m_{\text{Higgs}} = 130 \text{ GeV}/c^2$  für die assoziierte Higgs-Produktion (links) und die Vektorboson-Fusion (rechts). Die roten Punkte markieren die Akzeptanz für ein Standardmodell Higgs-Boson.

## 7.5 Die Signifikanz und die Ausschlussgrenzen im Phion-Modell

Durch die neuen physikalischen Gegebenheiten im Phion-Modell können sich diese Kurven jedoch verändern. Denn eine größere Breite des Higgs-Bosons kann Änderungen in den Wirkungsquerschnitten, welche in dieser Arbeit nicht berücksichtigt werden können, und der Kinematik der Prozesse bewirken. Um die Abhängigkeit der Ausschlussgrenzen von der Kopplung  $\omega$  abschätzen und erklären zu können, wurden daher die Schnittakzeptanzen als Funktion der Breite des Higgs-Bosons studiert. Diese Abhängigkeit ist allerdings nicht so groß, da in diesem Fall ein Zählexperiment durchgeführt wird. Daher ist diese Analyse nicht so sehr auf das Massenspektrum des Higgs-Bosons angewiesen, wie es der Fall wäre, wenn man dieses Spektrum direkt vermessen würde.

In Abbildung 7.11 zeigt sich, dass die Schnittakzeptanz im assoziierten Produktionskanal des Higgs-Bosons als Funktion der Breite des Higgs-Bosons zunächst konstant ist, für große Breiten des Higgs-Bosons jedoch abfällt. Dies ist der Fall, da für größere Breiten das Higgs-Boson öfter mit kleineren Massen generiert wird, was dazu führt, dass diese Ereignisse vom Schnitt auf die Di-Leptonmasse und den fehlenden transversalen Impuls abgewiesen werden. Dadurch fällt zum Beispiel die Schnittakzeptanz  $\varepsilon$  für ein Higgs-Boson mit einer Masse von  $130 \text{ GeV}/c^2$  von  $\varepsilon = 8.2\%$  bei Standardmodell-Breite auf  $\varepsilon = 7.4\%$  bei einer Breite von  $43 \text{ GeV}/c^2$ , was einer Kopplung von  $\omega = 3$  entspricht.

Für den Vektorboson-Fusionskanal konnte sich die Akzeptanz der Schnitte zunächst sogar erhöhen. Für eine Masse des Higgs-Bosons von  $130 \text{ GeV}/c^2$  verbessert sich die Akzeptanz von

## 7.5 Die Signifikanz und die Ausschlussgrenzen im Phion-Modell

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	105.47	91.82	78.63	67.16	59.72	56.36	26.67	14.82	8.33	3.07
$N_{Bkg.}$	473.4									
Signifikanz $S$	2.92	2.53	2.17	1.85	1.65	1.55	0.73	0.40	0.23	0.08
$S$ im <b>SM</b>	3.41	2.95	2.41	2.08	1.87	1.68	0.80	0.40	0.23	0.09
$\xi^2_{95\% CL}$	0.57	0.65	0.76	0.89	1.00	1.06	2.24	4.02	7.16	19.41
$\xi^2_{95\% CL}$ im <b>SM</b>	0.48	0.56	0.69	0.79	0.88	0.98	2.07	4.11	7.17	19.21

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	351.58	306.06	262.10	223.87	199.08	187.86	88.90	49.41	27.77	10.25
$N_{Bkg.}$	1578.0									
Signifikanz $S$	4.89	4.26	3.64	3.11	2.77	2.61	1.24	0.69	0.39	0.14
$S$ im <b>SM</b>	5.71	4.94	4.03	3.49	3.14	2.81	1.33	0.67	0.39	0.14
$\xi^2_{95\% CL}$	0.34	0.39	0.45	0.53	0.60	0.63	1.33	2.40	4.27	11.58
$\xi^2_{95\% CL}$ im <b>SM</b>	0.29	0.33	0.41	0.47	0.53	0.59	1.24	2.45	4.28	11.46

Tabelle 7.8: Aufgelistet sind die Anzahl der erwarteten Signal- und Untergrundereignisse ( $N_{Sig.}$  und  $N_{Bkg.}$ ), die Signifikanzen und die Ausschlussgrenzen für die Variable  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons mit  $BR(H \rightarrow inv.) = 100\%$  im Phion-Modell mit  $\omega = 3$  für die assoziierte Produktion eines Higgs-Bosons mit einem Z-Boson. Die Zahlen der oberen Tabelle entsprechen den erwarteten Datenmengen einer integrierten Luminosität von  $30 \text{ fb}^{-1}$ , die der unteren den Datenmengen einer integrierten Luminosität von  $100 \text{ fb}^{-1}$ . Die systematischen Fehler  $\alpha$  auf die Untergründe sind für diese Werte bereits berücksichtigt worden.

$\varepsilon = 1.41\%$  bei Standardmodell-Breite zu  $\varepsilon = 1.59\%$  bei einer Breite von  $43 \text{ GeV}/c^2$ . Dies liegt daran, dass die leichte Verschiebung des Maximums der Massenverteilung zwar eine geringere Schnittakzeptanz bewirkt, dieser Effekt aber durch die Asymmetrie dieser Verteilung aufgehoben wird. Denn in der Massenverteilung sind auch hohe Massen stark vertreten, welche im Bereich hoher Schnittakzeptanzen liegen. Erst bei sehr großen Breiten dominiert die Verschiebung der Massenverteilung über deren Asymmetrie. Dadurch verringert sich die Akzeptanz der Schnitte leicht im Vergleich zum Standardmodell. Diese Verschlechterung wird durch den Schnitt auf den fehlenden transversalen Impuls verursacht und wurde schon zuvor bei der Einführung dieses Kanals diskutiert und erwartet. Die Abbildung 7.11 zeigt dieses Verhalten.

Diese Auswirkungen der Higgs-Breite auf die Schnittakzeptanzen der beiden Kanäle lassen sich direkt auf das Entdeckungspotential übertragen.

Beispielhaft für eine Kopplung der Größe  $\omega = 3$  sind die Werte für die Anzahl der erwarteten Signal- und Untergrundereignisse, sowie die Werte für die Signifikanz des Signals und die Ausschlussgrenzen für  $\xi^2$  entsprechend einem Vertrauensniveau von 95% im Vergleich zu den Werten im Standardmodell in den Tabellen 7.8 und 7.9 aufgelistet. An den Werten zeigt sich deutlich, wie sich das Verhalten der Schnittakzeptanzen für die jeweiligen Kanäle auf die Ausschlussgrenzen auswirkt. So verbessert sich die Ausschlussgrenze beim Vektorboson-Fusionskanal für

## 7 Eine Studie zum Nachweis eines Higgs-Bosons im Phion-Modell

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	689.92	654.59	621.41	626.58	541.60	570.42	514.23	433.43	383.23	271.55
$N_{Bkg.}$	2930.0									
Signifikanz $S$	6.52	6.19	5.88	5.93	5.12	5.39	4.86	4.10	3.62	2.57
$S$ im SM	5.84	5.07	5.21	5.42	5.17	5.08	4.72	4.17	3.47	2.57
$\xi^2_{95\% CL}$	0.25	0.27	0.28	0.28	0.32	0.31	0.34	0.40	0.46	0.64
$\xi^2_{95\% CL}$ im SM	0.28	0.33	0.32	0.30	0.32	0.32	0.35	0.40	0.48	0.64

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	2069.76	1963.78	1864.24	1879.73	1624.81	1711.26	1542.69	1300.28	1149.68	814.66
$N_{Bkg.}$	8790.0									
Signifikanz $S$	9.29	8.81	8.37	8.43	7.29	7.68	6.92	5.83	5.16	3.66
$S$ im SM	8.31	7.22	7.42	7.72	7.36	7.23	6.72	5.93	4.94	3.66
$\xi^2_{95\% CL}$	0.18	0.19	0.20	0.20	0.23	0.21	0.24	0.28	0.32	0.45
$\xi^2_{95\% CL}$ im SM	0.20	0.23	0.22	0.21	0.22	0.23	0.25	0.28	0.33	0.45

Tabelle 7.9: Aufgelistet sind die Anzahl der erwarteten Signal- und Untergrundereignisse ( $N_{Sig.}$  und  $N_{Bkg.}$ ), die Signifikanzen und die Ausschlussgrenzen für die Variable  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons mit  $BR(H \rightarrow inv.) = 100\%$  im Phion-Modell mit  $\omega = 3$  für die Vektorboson-Fusion. Die Zahlen der oberen Tabelle entsprechen den erwarteten Datenmengen einer integrierten Luminosität von  $10 \text{ fb}^{-1}$ , die der unteren den Datenmengen einer integrierten Luminosität von  $30 \text{ fb}^{-1}$ . Die systematischen Fehler  $\alpha$  auf die Untergründe sind für diese Werte bereits berücksichtigt worden.

diesen Wert der Kopplung  $\omega$ , da sich die Schnittakzeptanzen leicht erhöhen. Ebenso wirkt sich die etwas schlechtere Schnittakzeptanz für diesen Wert der Kopplung beim assoziierten Produktionskanal negativ auf die Ausschlussgrenze aus.

In den Abbildungen 7.12 und 7.13 sind die Ausschlussgrenzen für  $\xi^2$  entsprechend einem Vertrauensniveau von 95% für die verschiedenen Luminositäten und für alle betrachteten Werte des Kopplungsparameters  $\omega$  für die beiden untersuchten Kanäle abgebildet. An diesen kann man jedoch sehen, dass sich die Ausschlussgrenzen für Breiten des Higgs-Bosons, welche kleiner sind als die Masse des Higgs-Bosons, für beide Kanäle kaum ändert. Übersetzt auf den Kopplungsparameter  $\omega$  bedeutet dies, dass bis zu einer Kopplung  $\omega = 1$  für kleine Massen und für größere Kopplungen bei höheren Massen des Higgs-Bosons für beide Kanäle nahezu keine Konsequenzen für die Ausschlussgrenzen zu befürchten sind. Für große Massen des Higgs-Bosons kann die Kopplung  $\omega$  sogar noch größer gewählt werden, bis sich die Ausschlussgrenzen verschieben.

Für sehr große Breiten verschlechtern sich die Ausschlussgrenzen im assoziierten Produktionskanal des Higgs-Bosons, da die Schnittakzeptanzen für diese Breiten abfallen. Die Vektorboson-Fusion hingegen profitiert von der Konstanz der Schnittakzeptanzen, wodurch sich die Ausschlussgrenzen auch für große Breiten nur sehr wenig verschlechtern.

Die genauen Werte für die Ausschlussgrenzen der beiden Kanäle für die hier betrachteten Massen sind für alle betrachteten Kopplungen in Anhang A tabellarisch aufgelistet.

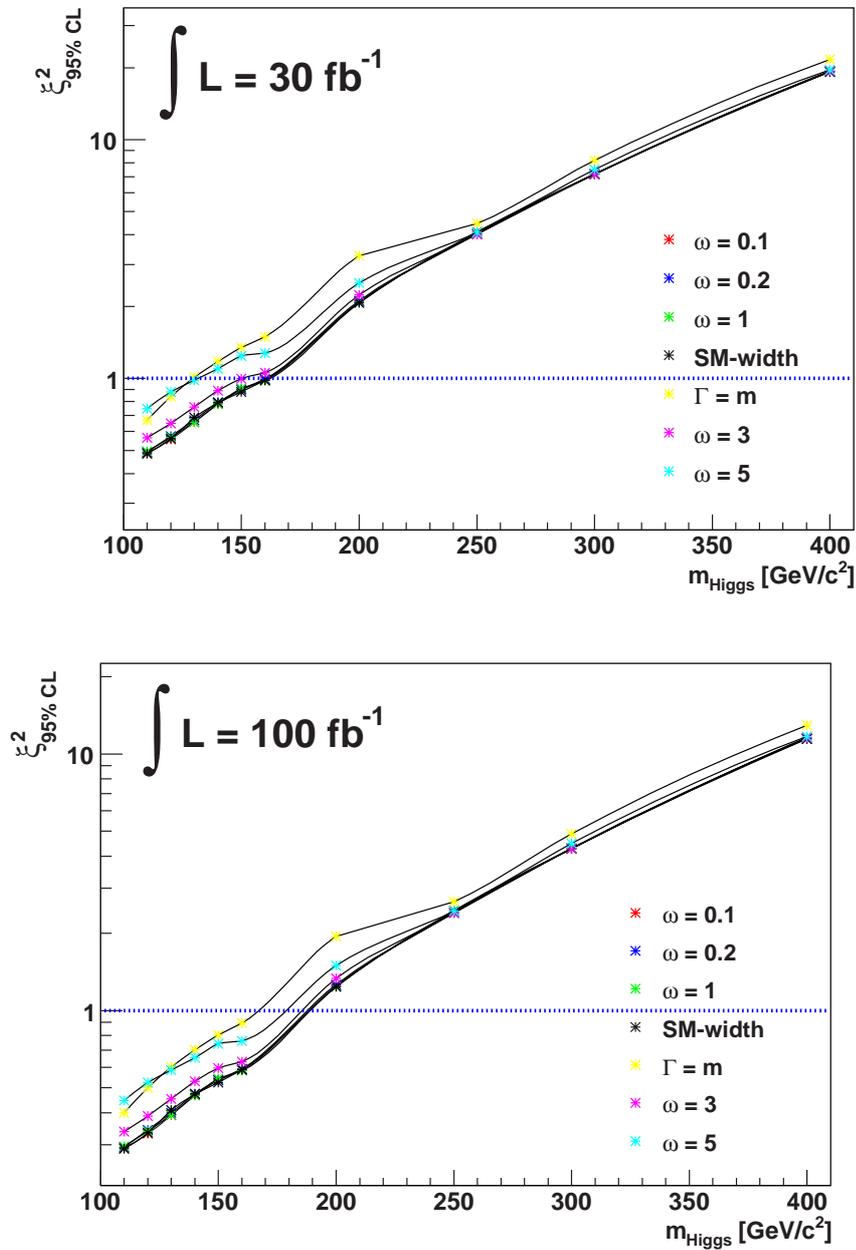


Abbildung 7.12: Ausschlussgrenzen für  $\xi^2$  im Phion-Modell für die assoziierte Higgs-Produktion mit einem Z-Boson auf 95% Vertrauensniveau für Daten, die einer integrierten Luminosität von  $30 \text{ fb}^{-1}$  (oben) und  $100 \text{ fb}^{-1}$  (unten) entsprechen. Die blauen, gestrichelten Linien entsprechen einem Higgs-Boson im Phion-Modell mit einem Zerfallsverhältnis in Phionen von  $BR(H \rightarrow \text{inv.}) = 100\%$ .

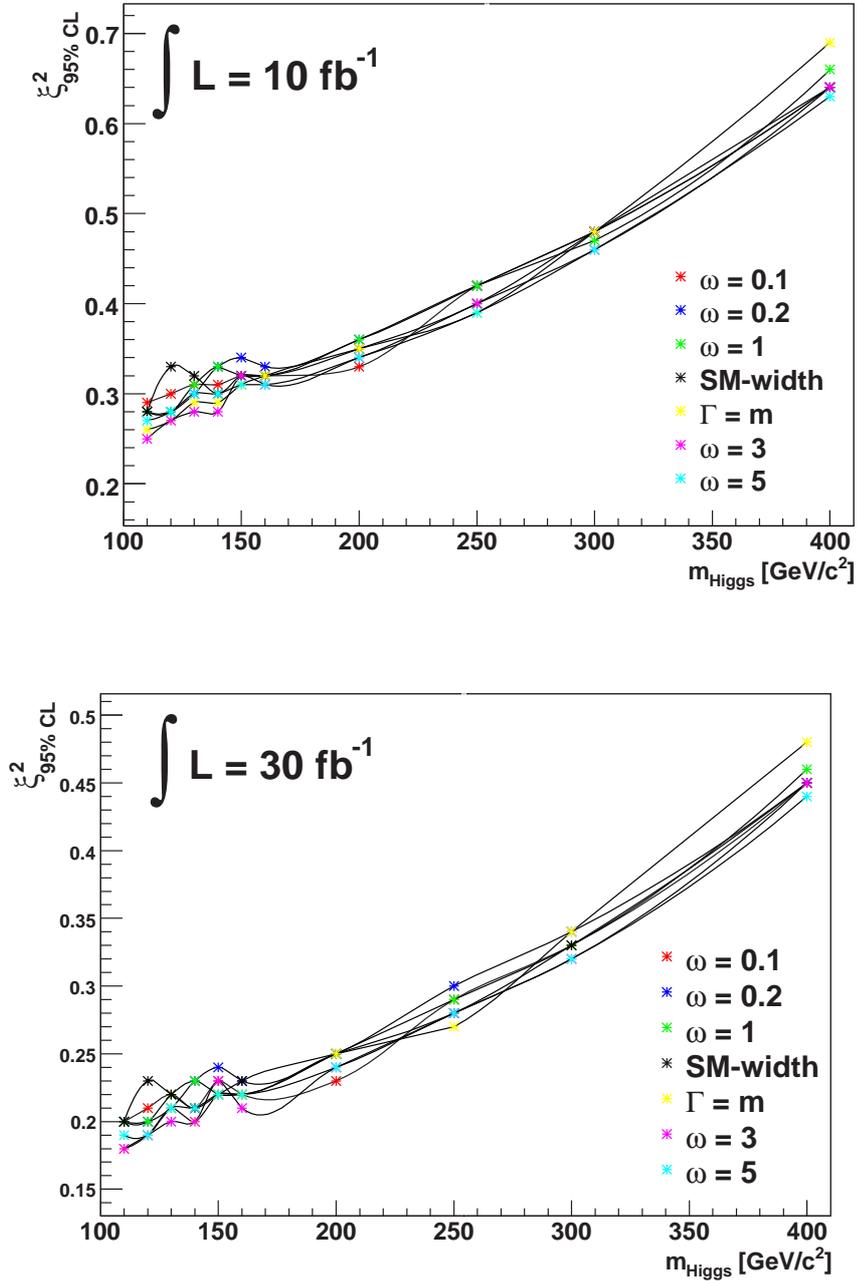


Abbildung 7.13: Ausschlussgrenzen für  $\xi^2$  im Phion-Modell für den Vektorboson-Fusionskanal auf 95% Vertrauensniveau für Daten, die einer integrierten Luminosität von  $10 \text{ fb}^{-1}$  (oben) und  $30 \text{ fb}^{-1}$  (unten) entsprechen.

# 8 Die Identifikation der niederenergetischer Elektronen

In diesem Kapitel wird zunächst beschrieben, wie die Identifikation von niederenergetischen Elektronen aus dem Zerfall von b-Hadronen oder der daraus entstandenen c-Hadronen zu diesem Zeitpunkt in ATHENA implementiert ist. Da diese Identifikation nur auf der in Kapitel 6 beschriebenen Likelihood-Methode basiert, die Korrelationen zwischen den Variablen nicht berücksichtigt<sup>1)</sup>, kann die Verwendung von anderen Methoden die Identifikationseffizienzen erhöhen. Außerdem können zusätzliche Variablen die Klassifikationsleistung weiter verbessern. Daher werden die in Kapitel 6 beschriebenen Methoden zur Klassifikation von Ereignissen auf die Identifikation von niederenergetischen Elektronen aus dem Zerfall von b-Hadronen oder den daraus entstandenen c-Hadronen unter Berücksichtigung von zusätzlichen Variablen angewandt und die Ergebnisse diskutiert. Die hierzu verwendeten Datensätze sollen zunächst im folgenden Unterkapitel vorgestellt werden.

## 8.1 Verwendete Datensätze

Da in den folgenden Abschnitten oftmals mehrere verschiedene Datensätze zu verschiedenen Zwecken verwendet werden, sollen diese Datensätze hier vorgestellt werden. Weiterhin werden für jeden Datensatz eindeutige Abkürzungen eingeführt, die die Lesbarkeit der folgenden Abschnitte verbessern sollen.

Alle Datensätze stammen aus der offiziellen Monte-Carlo-Produktion für das ATLAS-Experiment. Teilweise wurden sie jedoch mit verschiedenen Versionen von ATHENA erzeugt. Um eine eindeutige Zuordnung zur Version von ATHENA zu ermöglichen, wurden den Produktionen verschiedene Namen gegeben:

- Die „Computing system commissioning production“ wurde mit ATHENA Version 11 angefertigt. Diese wird im Folgenden mit *CSC11* bezeichnet.
- Die „Monte-Carlo validation production“ wurde mit ATHENA Version 12 angefertigt und wird im Folgenden mit *MC12* bezeichnet.

Die aus diesen Produktionen verwendeten Datensätze sind:

---

<sup>1</sup>Nur Korrelationen der Variablen mit der Pseudorapidität  $\eta$  werden teilweise berücksichtigt, indem die Likelihood-Verhältnisse für fünf  $\eta$ -Intervallen getrennt berechnet werden.

1. Ein Datensatz mit assoziierter Produktion des Higgs-Bosons mit einem W-Boson. Das Higgs-Boson zerfällt nach seiner Produktion in ein  $b\bar{b}$ -Quark-Paar, das W-Boson in ein Myon und ein Anti-Myon-Neutrino. Dieser Datensatz stammt aus der CSC11 Monte-Carlo-Produktion und wird im folgenden mit *CSC11WHbb* bezeichnet.
2. Ein Datensatz mit assoziierter Produktion des Higgs-Bosons mit einem W-Boson, bei dem das Higgs-Boson in ein  $b\bar{b}$ -Quark-Paar und das W-Boson in ein Myon und ein Anti-Myon-Neutrino zerfällt. Bei diesem Datensatz wurden die Ereignisse derart generiert, dass in jedem b-Jet des Ereignisses mindestens ein Elektron existiert. Dieser Datensatz stammt aus der MC12 Monte-Carlo-Produktion und wird im folgenden mit *MC12WHbbfiltered* bezeichnet.
3. Ein Datensatz, in dem das Higgs-Boson assoziiert mit einem W-Boson erzeugt wird. Das W-Boson zerfällt wiederum in ein Myon und ein Anti-Myon-Neutrino. In diesem Datensatz aus der CSC11-Produktion zerfällt das Higgs-Boson allerdings in ein  $u\bar{u}$ -Quark-Paar, weswegen er mit *CSC11WHuu* bezeichnet wird.
4. Ein Datensatz, in dem das Higgs-Boson assoziiert mit einem W-Boson erzeugt wird. Das W-Boson zerfällt in ein Myon und ein Anti-Myon-Neutrino. In diesem Datensatz, welcher aus der MC12-Produktion stammt, zerfällt das Higgs-Boson in ein  $u\bar{u}$ -Quark-Paar. Daher wird dieser Datensatz mit *MC12WHuu* bezeichnet.
5. Ein weiterer Datensatz mit einer mit einem W-Boson assoziierten Higgs-Produktion. Das Higgs-Boson zerfällt in diesem Datensatz aus der CSC11-Produktion in ein  $c\bar{c}$ -Quark-Paar. Im Folgenden wird dieser Datensatz *CSC11WHcc* genannt.
6. Der letzte Datensatz beinhaltet die Produktion von  $t\bar{t}$ -Quark-Paaren und stammt aus der MC12-Produktion. Er wird im Folgenden mit *MC12tt* bezeichnet.

## 8.2 Bisherige Identifikation niederenergetischer Elektronen

Der Algorithmus, welcher momentan zur Identifikation der Elektronen aus dem Zerfall von b-Hadronen oder den daraus entstandenen c-Hadronen in ATHENA implementiert ist, basiert auf der Likelihood-Methode<sup>2)</sup> und heißt „softeBuilder“. Zunächst werden von allen rekonstruierten Spuren diejenigen selektiert, welche folgende Kriterien erfüllen:

- Für den transversalen Impuls der Spur muss gelten:  $P_t \leq 1 \text{ GeV}/c$ . Spuren mit einem geringeren transversalen Impuls können nur sehr selten und mit großen Fehlern auf ihre Eigenschaften rekonstruiert werden.
- Der Betrag der Pseudorapidität  $|\eta|$  muss kleiner oder gleich 2.4 sein. Dies ist der Bereich, in dem mit dem Spurdetektor die Spuren rekonstruiert werden können.

---

<sup>2)</sup>Bei der in ATHENA implementierten Methode wird die Definition 6.10 des Likelihoodverhältnisses verwendet.

- Der Betrag des transversalen Stoßparameters  $d_0$  muss kleiner als 1 mm sein. Dieser wird für ein Teilchen bestimmt, indem dessen Spur auf die zur Strahlachse transversale Ebene durch den rekonstruierten Primärvertex projiziert und in Richtung des Primärvertex extrapoliert wird. Der transversale Stoßparameter ist dann der minimale Abstand des Primärvertex von dieser projizierten Spur eines Teilchens. Das Vorzeichen des transversalen Stoßparameters wird durch den azimuthalen Winkel  $\phi_0$  der Tangente an die Spur in dem Punkt, der am nächsten zum Primärvertex liegt, und dem Azimuthalwinkel  $\phi$  dieses Punktes der Spur berechnet. Gilt für diese Winkel, dass  $\phi - \phi_0 = (\pi/2) + n \cdot 2\pi$  ( $n \in \mathbb{Z}_0$ ), so wird das Vorzeichen positiv gewählt, ansonsten negativ. Diese Definition und Bestimmung des transversalen Stoßparameters wird Anhand der Abbildung 8.1 verdeutlicht. Durch den Schnitt auf diese Variable wird sichergestellt, dass Spuren eines Pile-Up-Ereignisses nicht berücksichtigt werden.
- In der innersten Lage des Pixeldetektors, der sogenannten b-Lage, muss mindestens ein Treffer registriert werden. Dies stellt sicher, dass die Spuren, beziehungsweise deren transversale Stoßparameter, mit guter Auflösung rekonstruiert werden können.
- Im Pixeldetektor müssen mindestens zwei Treffer registriert werden. Auch dieser Schnitt garantiert eine gute Auflösung bei der Rekonstruktion des transversalen Stoßparameters der Spuren.
- Die Summe der Treffer im Silizium-Streifendetektor (SCT) und der Treffer im Pixeldetektor muss mindestens neun betragen.
- Die Spur muss mindestens 20 Treffer im Übergangsstrahlungsdetektor aufweisen. Zusammen mit dem vorhergehenden Schnitt kann die Spur durch die große Menge an Information gut rekonstruiert werden. Dies erhöht die Auflösung bei der Bestimmung des transversalen Stoßparameters und der Auftreffpunkte der Spur auf die Kalorimeteroberflächen.
- Mindestens ein hochenergetischer Treffer der Spur im Übergangsstrahlungsdetektor muss registriert werden. Ein hochenergetischer Treffer ist ein Treffer, bei dem mindestens 5 keV an Energie gemessen wurde. Solche Treffer weisen hauptsächlich die Spuren von Elektronen auf. Die Anzahl solcher Treffer von Pionen ist im Mittel sehr gering. Ein schärferer Schnitt auf diese Variable wird nicht angewendet, da diese Variable als diskriminierende Variable bei der Identifikation der niederenergetischen Elektronen verwendet wird.

Die Menge dieser Schnitte wird im Folgenden mit „Qualitätsschnitte“ bezeichnet. Diese Kriterien werden von circa 15% der Spuren der Signal-Elektronen nicht erfüllt, die Anzahl der Pionen kann durch diese Schnitte um etwa 24% verringert werden [54].

Die so selektierten Spuren werden in das elektromagnetische Kalorimeter extrapoliert und die Energiedepositionen in den verschiedenen Samples des Kalorimeters in einem Kegelvolumen rekonstruiert. Hierzu werden die Spuren auch danach untersucht, ob Bremsstrahlung oder eine Photokonversion stattgefunden haben könnte. Die Menge der benachbarten Kalorimeterzellen, in denen die Energie deponiert wurde, bilden den Schauer eines Teilchens und werden „Cluster“ genannt. Aus den in den einzelnen Zellen deponierten Energien im Detektor werden

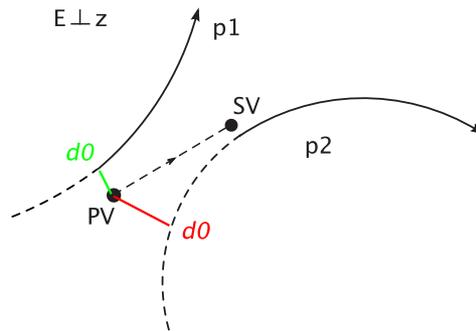


Abbildung 8.1: Graphische Darstellung der Bestimmung des transversalen Stoßparameters. Dieser wird bestimmt, indem die Spur des Teilchens auf die zur Strahlachse  $z$  transversale Ebene  $E$  durch den rekonstruierten Primärvertex  $PV$  projiziert und in Richtung des Primärvertex extrapoliert wird. Der transversale Stoßparameter ist dann der minimale Abstand des Primärvertex von dieser projizierten Spur. Ein Teilchen mit projizierter Spur  $p1$  wird am Primärvertex erzeugt. Für solche Teilchen ist der transversale Stoßparameter  $d0$  im Mittel sehr gering. Für ein Teilchen mit der projizierten Spur  $p2$ , welches am Sekundärvertex  $SV$  erzeugt wird, ist der Stoßparameter  $d0$  im Mittel größer.

dann mehrere Variablen berechnet. Für diese Berechnung wird auf folgendes Koordinatensystem zurückgegriffen. Die Kalorimeterzellen des elektromagnetischen Kalorimeters beschreiben dabei Einheiten in den Koordinaten  $\eta$  und  $\phi$ . Wird im Folgenden von einem Fenster von  $\Delta\eta \times \Delta\phi = (2 \cdot x + 1) \times (2 \cdot y + 1)$  gesprochen, so sind damit alle folgenden Kalorimeterzellen gemeint:

- Alle  $x$  benachbarten Kalorimeterzellen einer Referenzzelle in  $+\eta$ -Richtung und in  $-\eta$ -Richtung.
- Alle  $y$  benachbarten Kalorimeterzellen einer Referenzzelle in  $+\phi$ -Richtung und in  $-\phi$ -Richtung.

So bezeichnet zum Beispiel  $3 \times 3$  ein Fenster von 9 Zellen, dessen Zentrum die Referenzzelle ist. Mit der Referenzzelle ist immer die Kalorimeterzelle gemeint, welche mit dem Auftreffpunkt der extrapolierten Spur eines Teilchens auf das Kalorimeter zusammenfällt.

Weiterhin wird im Folgenden die Summe der Energien aller Zellen in einem Fenster der Größe  $\Delta\eta \times \Delta\phi = x \times y$  mit  $E_{x \times y}$  bezeichnet.

Die Variablen, die zur Identifikation der Elektronen berechnet werden, sind:

- $N_{tr}$ : Anzahl der hochenergetischen Treffer im Übergangsstrahlungsdetektor entlang der Spur eines Teilchens. Diese Anzahl ist für Elektronen im Mittel deutlich höher als zum Beispiel für Pionen.

- $E_t/P_t$ : Transversale Energie des Schauers, gemessen im Kalorimeter, geteilt durch den transversalen Impuls der Spur, welcher im Spurdetektor bestimmt wurde. Da Elektronen eine relativ geringe Masse besitzen und fast ihre gesamte Energie im elektromagnetischen Kalorimeter deponieren, ist die Verteilung dieser Variable für Elektronen um eins verteilt. Für Teilchen, die Ihre Energie auch im hadronischen Kalorimeter deponieren, kann dieser Wert deutlich unter eins liegen.
- $f1core$ : Die Energie in den drei Streifen im ersten Sampling des elektromagnetischen Kalorimeters, die um den Streifen mit maximaler Energie liegen, geteilt durch  $E_{core}$ , also  $E3 \times 1 / E_{core}$ .  
Die Energie  $E_{core}$  bezeichnet die Energie, die von einem Teilchen im elektromagnetischen Kalorimeter deponiert wurde, und wird aus der Summe der folgenden Energien berechnet:  $E3 \times 3$  im Presampler,  $E15 \times 3$  in den Streifen,  $E5 \times 5$  im Main-Sampling und  $E3 \times 5$  in der Rückseite des Kalorimeters. Da Elektronen im Mittel schon früher mit der Schauerbildung beginnen als Pionen, hat diese Variable im Durchschnitt einen höheren Wert für Elektronen.
- $f3core$ :  $E3 \times 3$  auf der Rückseite des elektromagnetischen Kalorimeters geteilt durch  $E_{core}$ . Diese Variable besitzt für Elektronen normalerweise einen geringeren Wert, da die Elektronen den Großteil ihrer Energie bereits früher im elektromagnetischen Kalorimeter deponieren.
- $ShProfileEta2$ :  $E3 \times 3$  geteilt durch  $E3 \times 7$  im Main-Sampling des elektromagnetischen Kalorimeters. Der Wert für diese Variable liegt für Elektronen nahe der Eins, da die Schauer der Elektronen im Durchschnitt sehr schmal sind. Für Pionen sind diese Schauer breiter, wodurch der Wert für die Variable  $ShProfileEta2$  für diese meist deutlich kleiner als eins ist.
- $Weta1$ : Korrigierte Breite des Schauers in einem  $3 \times 1$ -Fenster in den Streifen des ersten Samplings des elektromagnetischen Kalorimeters. Ist  $i_m$  die Zelle, auf die die extrapolierte Spur auftrifft, und  $E_i$  die in der  $i$ -ten Zelle deponierte Energie, so kann diese Variable folgendermaßen berechnet werden:

$$Weta1 = \sqrt{\sum_{i=i_m-1}^{i_m+1} E_i \cdot (i - i_m)^2 / \sum_{i=i_m-1}^{i_m+1} E_i}$$

Hier wird nur in  $\eta$ -Richtung über den Index  $i$  summiert. Die Schauer der Elektronen sind im Durchschnitt schmäler als die der Pionen. Daher wird zunächst erwartet, dass der Wert dieser Variable für Elektronen im Mittel geringer ist als der Wert für Pionen. Allerdings kommt es häufig vor, dass ein Pion erst spät im elektromagnetischen Kalorimeter damit beginnt, dort seine Energie zu deponieren, sodass der Schauer bei Erreichen des ersten Samplings noch nicht weit ausgebreitet ist. Insgesamt folgt aus diesen Überlegungen, dass die Verteilung dieser Variable für Pionen breiter ist als die Verteilung für Elektronen.

- $Weta2$ : Korrigierte Schauerbreite in einem  $3 \times 5$ -Fenster im Main-Sampling des elektromagnetischen Kalorimeters:

$$Weta2 = \sqrt{\sum_{i \in 3 \times 5 \text{-Fenster}} E_i \cdot (i - i_m)^2 / \sum_{i=i_m-1}^{i_m+1} E_i}$$

Diese Variable ist für Pionen im Durchschnitt größer als für Elektronen, da deren Schauer im Mittel breiter ist.

- *Wtots1*: Totale Breite des Schauers in einem Fenster 41x1 im Presampler des elektromagnetischen Kalorimeters:

$$Wtots1 = \sqrt{\sum_{i \in 41 \times 1 \text{-Fenster}} E_i \cdot (i - i_m)^2 / \sum_{i=i_m-1}^{i_m+1} E_i}$$

Auch diese Variable ist für Pionen aufgrund der größeren Breite ihrer Schauer größer als für Elektronen.

- *POS* (innerhalb des ATHENA-Frameworks auch *pos7* genannt): Differenz der Position des Schauers zur Position der extrapolierten Spur in einem  $15 \times 1$ -Fenster. Ist  $i_m$  die Zelle, auf die die extrapolierte Spur auftrifft, und  $E_i$  die Energie in der  $i$ -ten Zelle, so kann diese Variable folgendermaßen berechnet werden:

$$POS = \sum_{i=i_m-7}^{i_m+7} E_i \cdot (i - i_m) / \sum_{i=i_m-7}^{i_m+7} E_i$$

Über den Index  $i$  wird nur in  $\eta$ -Richtung summiert. Auch diese Variable ist für Pionen wegen der größeren Breite der Schauer größer als für Elektronen.

Die Verteilungen aller oben genannten Variablen für Signal-Elektronen und den Untergrund, bestehend aus allen übrigen Teilchen in den entsprechenden Datensätzen, sind nach den Qualitätsschnitten in Abbildung 8.2 dargestellt. Aus dieser Abbildung wird ersichtlich, dass einige dieser Variablen sehr gut zur Trennung von Signal-Elektronen und dem Untergrund geeignet sind.

Für diese Variablen werden separat für fünf Intervalle der Pseudorapidität  $\eta$  die Likelihoodwerte für Signal- und Untergrundzugehörigkeit berechnet. Diese Intervalle der Absolutbeträge der Pseudorapidität sind:

$$(0, 0.8), (0.8, 1.4), (1.4, 1.8), (1.8, 2.0), (2.0, 2.4).$$

Weiterhin besteht die Möglichkeit, den transversalen Stoßparameter der Spuren der Elektronenkandidaten zu verwenden. Für Teilchen, die am Primärvertex erzeugt werden, folgt aus der Definition, dass der transversale Stoßparameter wegen der Auflösung bei der Rekonstruktion der Spuren um Null verteilt ist. Falls ein Teilchen weit entfernt vom Primärvertex erzeugt wird, so wie es bei den Spuren der niederenergetischen Elektronen aus dem Zerfall der b-Hadronen oder der daraus entstandenen c-Hadronen der Fall ist, besitzt die in Richtung des Primärvertex extrapolierte Spur einen im Mittel größeren transversalen Stoßparameter. Somit ist dieser transversale Stoßparameter für Signal-Elektronen im Mittel größer als für die Pionen im Jet, da viele der Pionen am Primärvertex erzeugt werden. Aus diesem Grund kann die Verwendung des transversalen Stoßparameters die Identifikation der niederenergetischen Elektronen aus dem Zerfall von

b-Hadronen oder der daraus entstandenen c-Hadronen deutlich verbessern. Im Folgenden soll auf die Verwendung dieser Variable an dieser Stelle verzichtet werden. Dadurch kann zum Einen eine größere Unabhängigkeit von den auf der Lebensdauer der b-Hadronen basierenden Methoden zur Markierung der b-Quarks erreicht werden, da diese Methoden die Information über die transversalen Stoßparameter der Spuren in einem Jet bereits verwenden. Zum Anderen kann nur so ein fairer Vergleich der Likelihood-Methode mit den neuen Methoden, die zur Identifikation der Elektronen auf den transversalen Stoßparameter verzichten sollen, gewährleistet werden.

## 8 Die Identifikation der niederenergetischer Elektronen

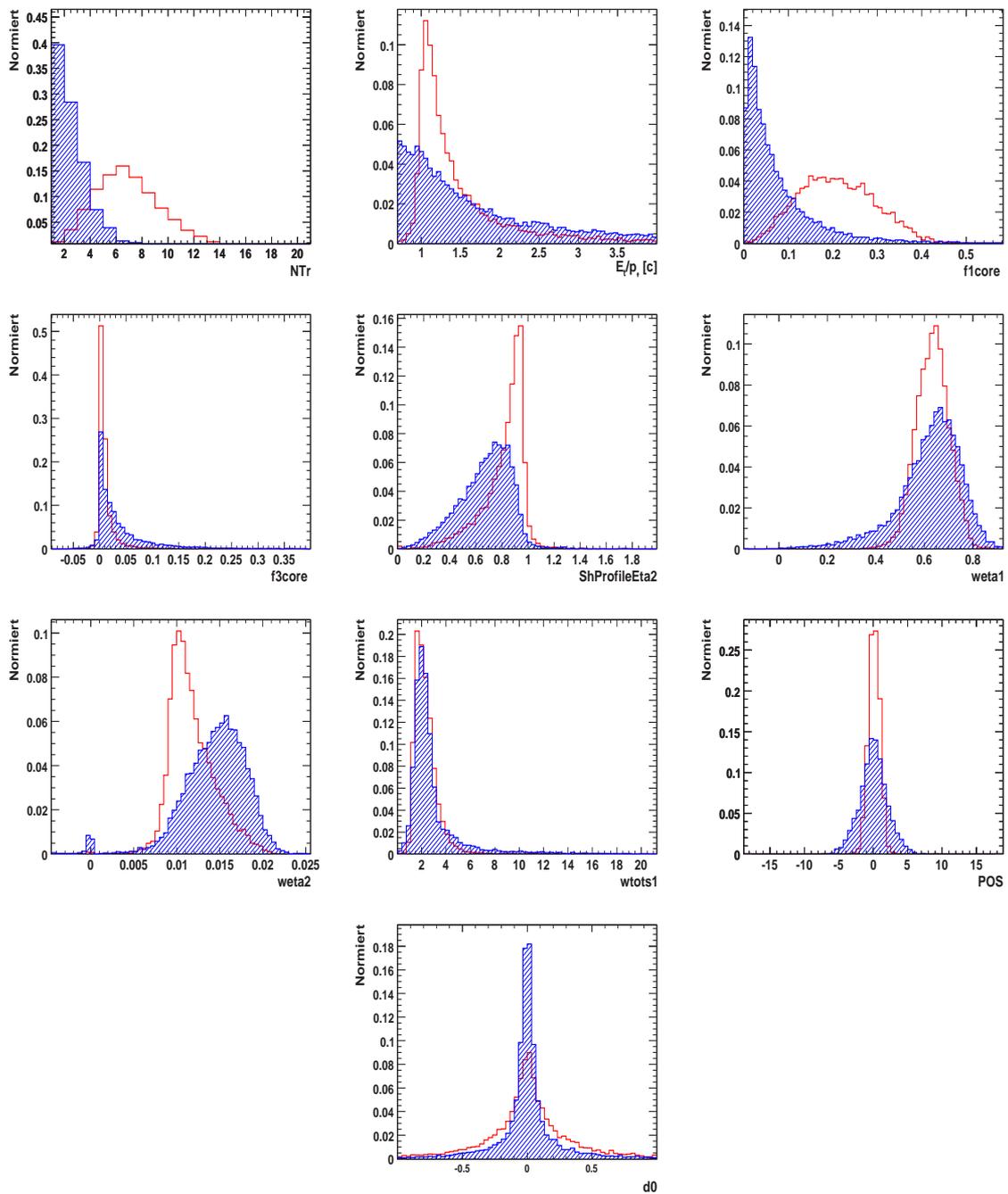


Abbildung 8.2: Gezeigt sind die verschiedenen Variablen der Elektronen-Kandidaten nach den Qualitätsschnitten auf die Spuren. In Rot sind jeweils die Verteilungen der Signal-Elektronen aus dem Zerfall der b-Hadronen oder dem Zerfall der daraus entstandenen c-Hadronen dargestellt. Die blauen schraffierten Flächen zeigen die Verteilungen der restlichen Spuren, welche hauptsächlich die Spuren von Pionen sind. Alle Verteilungen wurden mit dem Datensatz *MC12WHbbfiltered* angefertigt.

## 8.3 Zusätzliche Variablen zur Identifikation der niederenergetischen Elektronen

Um die Identifikationseffizienz von niederenergetischen Elektronen aus den Zerfällen von b-Hadronen oder den daraus entstandenen c-Hadronen weiter zu verbessern, wurden für die neuen Methoden weitere Variablen betrachtet.

- *frac1*: Diese Variable beschreibt die Energie des Schauers, außerhalb des Bereichs im Presampler, in dem  $E_{core}$  bestimmt wurde.

$$frac1 = \frac{E_{15 \times 15} - E_{7 \times 7}}{E_{15 \times 15}}.$$

Durch die geringere Ausbreitung der Schauer von Elektronen im Vergleich zu denen der Pionen hat diese Variable für Elektronen einen geringeren Wert.

- *ISO*: Diese Variable beschreibt die Isolierung des Schauers im elektromagnetischen Kalorimeter. Die zur Berechnung von *ISO* benötigten Variablen werden aus den Energien der folgenden Fenster im elektromagnetischen Kalorimeter berechnet:

$E_{3 \times 3} = E_1 \times 1$  im Presampler +  $E_3 \times 1$  in den Streifen +  $E_3 \times 3$  im Main-Sampling +  $E_3 \times 3$  im Rückteil.

$E_{3 \times 7} = E_3 \times 3$  im Presampler +  $E_{15} \times 3$  in den Streifen +  $E_3 \times 7$  im Main-Sampling +  $E_3 \times 7$  im Rückteil.

Mit diesen Variablen folgt für die Variable  $ISO = 1 - \frac{E_{3 \times 3}}{E_{3 \times 7}}$ . Für Elektronen, deren Schauer ausbreitung im Mittel geringer ist als die der Pionen, ist der Wert für diese Variable im Mittel geringer.

- *ethad*: Energie im ersten Teil des hadronischen Kalorimeters in einem Kegel mit Radius  $R = 0.45$ . Im Vergleich zu Pionen deponieren Elektronen nahezu ihre gesamte Energie im elektromagnetischen und fast keine im hadronischen Kalorimeter. Daher ist diese Variable für Elektronen im Mittel geringer als für Pionen.
- *emins1*: Energie des Streifens im Presampler, welcher zwischen den beiden Streifen maximaler Energie liegt und in dem die wenigste Energie deponiert wurde. Diese Energie ist aufgrund der geringen Schauerbreite der Elektronen meist ein wenig größer als für ein Pion, da diese Energie für ein Elektron auf einen kleineren Bereich verteilt wird und damit pro Zelle im Mittel mehr Energie deponiert wird.
- $\eta$ : Die Pseudorapidität der Spur wird verwendet, damit durch die neuen Methoden die Korrelation zwischen den anderen Variablen und der Pseudorapidität berücksichtigt wird. Dabei wurde sichergestellt, dass die Variable  $\eta$  selbst sich nicht diskriminierend auswirkt.

Um eine möglichst große Unabhängigkeit der neuen Methoden von den auf der Lebensdauer der b-Hadronen basierenden Methoden zu gewährleisten, wird auf die Verwendung des transversalen Stoßparameters der Spuren und des transversalen Impulses der Spuren relativ zur Jet-Achse an

dieser Stelle verzichtet. Diese Informationen werden bereits von den Methoden verwendet, die auf der Lebensdauer der b-Hadronen basieren. Die oben genannten Variablen sind nach den Qualitätsschnitten auf die Spuren in Abbildung 8.3 dargestellt. In dieser ist zu sehen, dass einige der neuen Variablen ein großes Potential zur Trennung von Signal und Untergrund besitzen.

Weiterhin wurden zusätzlich zu den oben genannte Variablen weitere Variablen der Elektronenkandidaten betrachtet. Diese sollen hier jedoch nicht vorgestellt werden, da das Trennungspotential dieser Variablen von Signal und Untergrund geringer war als das der hier vorgestellten Variablen. Um diese Trennungsfähigkeit der einzelnen Variablen von Signal und Untergrund quantisieren zu können, kann die Variable  $S$  betrachtet werden. Diese ist definiert als [24]

$$S_y = \int \frac{(P_{Sig}(y) - P_{Bg}(y))^2}{P_{Sig}(y) + P_{Bg}(y)} dy. \quad (8.1)$$

Die  $P(y)$  sind die Wahrscheinlichkeitsverteilungen der Variable  $y$  für Signal (*Sig*) und Untergrund (*BG*). Diese Variable quantisiert die Unterschiede in den Wahrscheinlichkeitsverteilungen von Signal und Untergrund. Sind die Verteilungen einer Variablen  $y$  für beide Klassen identisch, so folgt für die Trennungsfähigkeit  $S_y = 0$ . Gibt es zwischen Signal- und Untergrundverteilung einer Variable keinen Überlapp, so ist die Trennungsfähigkeit dieser Variable maximal:  $S_y = 1$ . Die Wahrscheinlichkeitsverteilungen für diese Analyse für das Signal entstanden aus den Verteilungen der Variablen von 10000 niederenergetischen Signal-Elektronen aus den Datensätzen *CSC11WHbb* und *MC12WHbbfiltered*, nachdem die Qualitätsschnitte auf die Spuren angewandt wurden. Die Untergrundverteilungen entstanden entsprechend aus den Verteilungen von 10000 Teilchen, die keine Signal-Elektronen waren, ebenfalls nach den Qualitätsschnitten auf die Spuren. Um die Verwendung zweier Datensätze aus Produktionen mit verschiedenen ATHENA Versionen zu rechtfertigen, wurden die Verteilungen der Variablen der Signal-Elektronen für beide Datensätze miteinander verglichen. In Abbildung 8.4 sind diese Vergleiche für einige Variablen dargestellt. Es zeigte sich, dass beide Datensätze nahezu identische Verteilungen der Variablen aufwiesen, sodass eine gemeinsame Nutzung gerechtfertigt ist.

Für die Variablen, welche die beste Trennungsfähigkeit aufwiesen, sind die Werte für  $S_y$  in Tabelle 8.1 aufgelistet. Aufgrund dieser guten Trennungsfähigkeit dieser Variablen werden diese im Folgenden von den vorgestellten neuen Methoden zur Identifikation der niederenergetischen Elektronen verwendet. Die linearen Korrelationen der Variablen, deren Berechnung in Anhang C erläutert wird, sind in Abbildung 8.5 in Matrixform dargestellt. Die hohen linearen Korrelationskoeffizienten einiger Variablen für das Signal lassen darauf schließen, dass mit einer anderen Methode als der bisher verwendeten Likelihood-Methode die Trennung von Signal und Untergrund noch verbessert werden kann. Zum Beispiel wurde die lineare Korrelation der Variablen *ISO* und *weta2* zu ca. 48% bestimmt. Dies war zu erwarten, denn eine geringe Breite eines Schauers, also ein geringer Wert für *weta2*, bedeutet, dass die meiste Energie dieses Clusters in den Zellen nahe des Zentrums deponiert wurde. Für die Berechnung der Variable *ISO*, welche die Energie in einem gewissen Bereich außerhalb des Zentrums quantisiert, folgt daraus, dass der Wert für *ISO* geringer ausfällt. Für den Untergrund, welcher hauptsächlich aus Pionen besteht, fallen die Korrelationen geringer aus. Dies liegt daran, dass die zweidimensionalen Verteilungen zweier Variablen für Pionen im Mittel deutlich breiter sind als die der Elektronen.

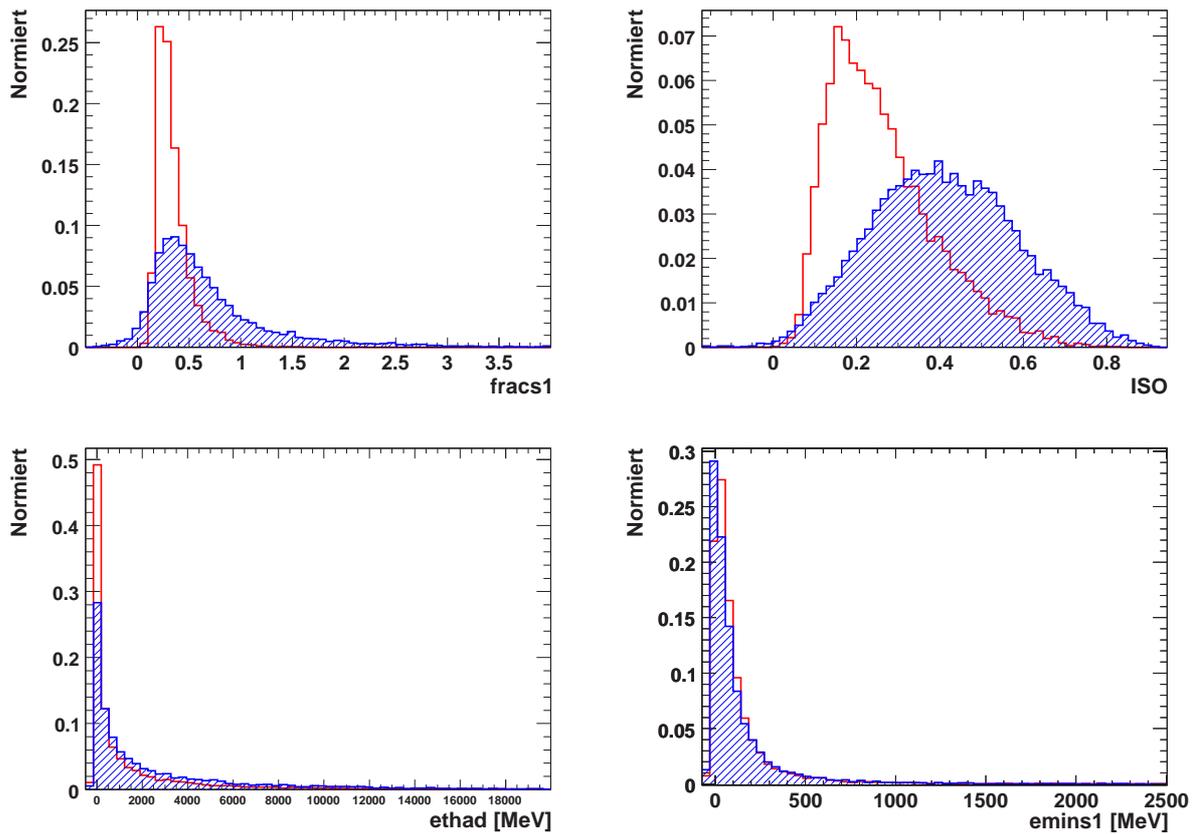


Abbildung 8.3: Gezeigt sind weitere Variablen der Elektronen-Kandidaten nach den Qualitätsschnitten auf die Spuren. In rot sind die Verteilungen der Signalelektronen aus dem Zerfall von b-Hadronen oder der daraus entstandenen c-Hadronen dargestellt. Die blaue schraffierte Fläche zeigt die Verteilung der restlichen Spuren, welche hauptsächlich Spuren von Pionen sind. Die Abbildungen wurden mit dem Datensatz *MC12WHbbfiltered* angefertigt.

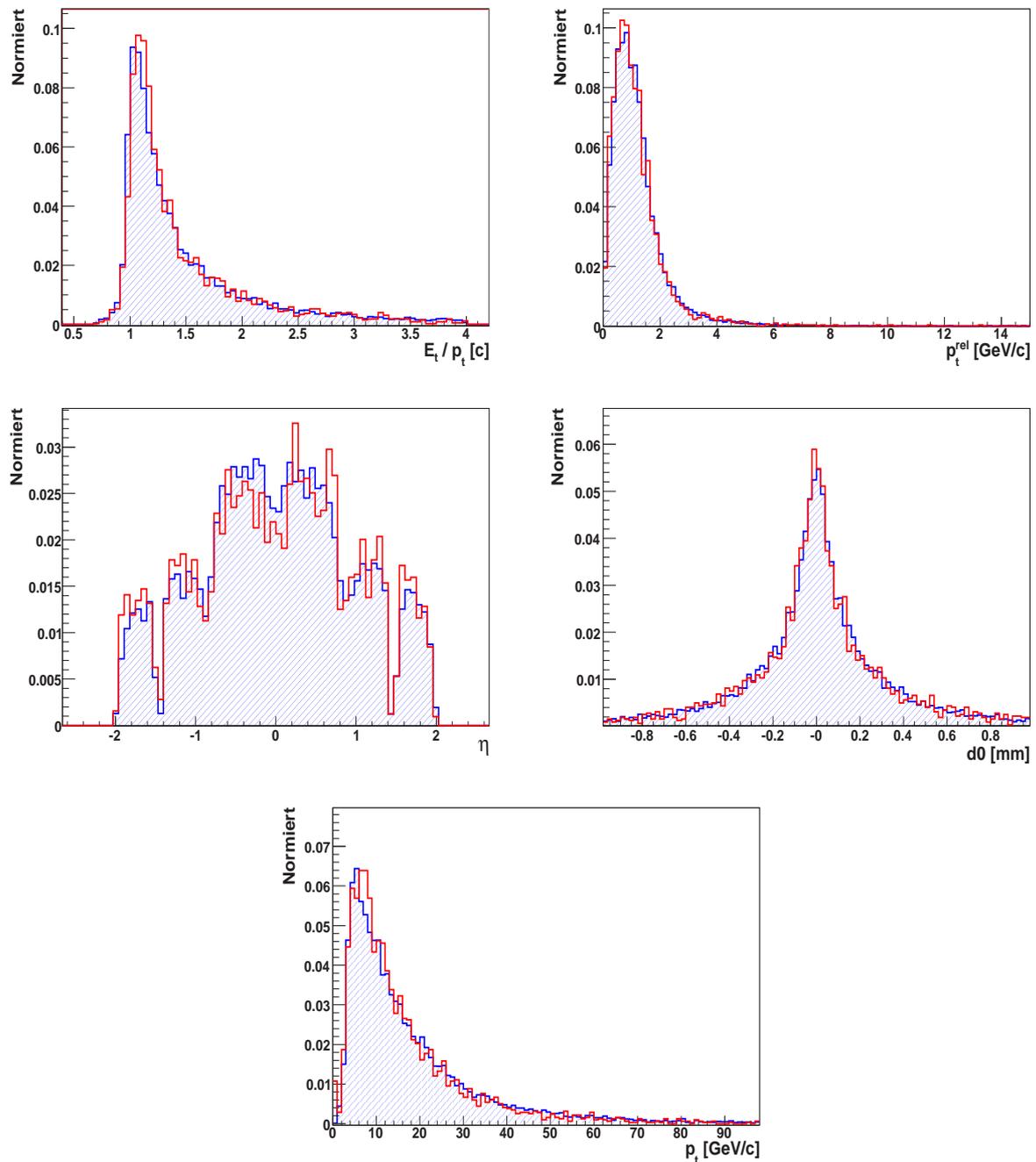


Abbildung 8.4: Gezeigt ist ein Vergleich der zwei für das Training der Methoden verwendeten Datensätze mit Hilfe mehrerer Variablen der Signal-Elektronen. Die Histogramme sind normiert, um einen direkten Vergleich liefern zu können. Die blauen Histogramme zeigen die Verteilungen des Datensatzes *CSC11WHbb*, welcher mit ATHENA Version 11 erzeugt wurde. Die roten Histogramme zeigen die Verteilungen des Datensatzes *MC12WHbbfiltered* (ATHENA Version 12).

Variable	Trennungsfähigkeit $S$
NTr	0.586
f1core	0.504
fracs1	0.260
weta2	0.236
ISO	0.219
POS	0.193
ShProfileEta2	0.189
f3core	0.174
$E_t/P_t$	0.153
weta1	0.111
ethad	0.059
emins1	0.027

Tabelle 8.1: Trennungsfähigkeit  $S$  der verschiedenen Variablen zur Identifikation der niederenergetischen Elektronen aus dem Datensatz *MC12WHbbfiltered*. Die Definition für  $S$  ist in Gleichung 8.1 gegeben. Diese Variablen werden später von den neuen Methoden verwendet.

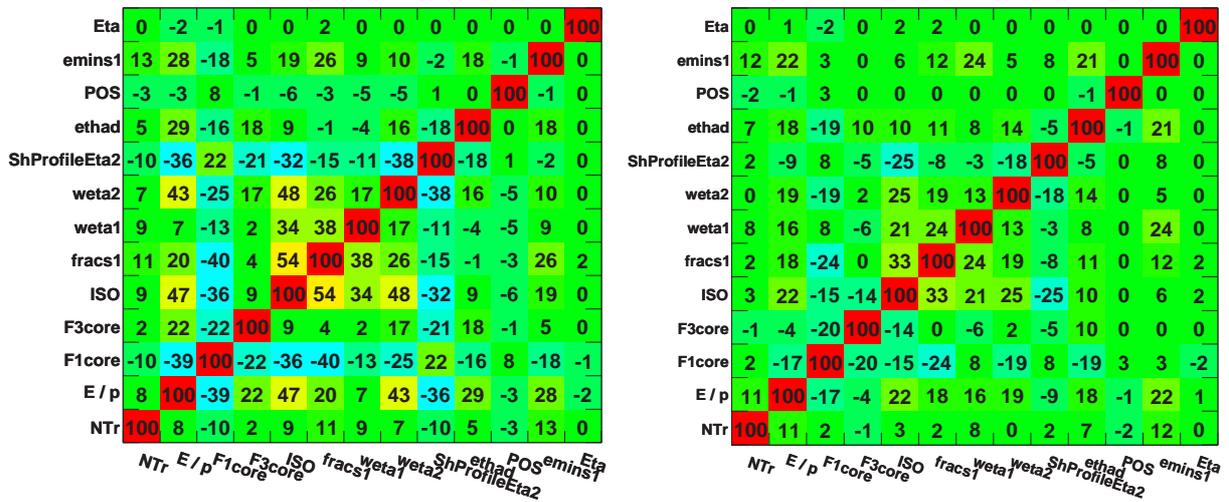


Abbildung 8.5: Lineare Korrelationsmatrizen für die Variablen, die von den neuen Methoden zur Identifikation der niederenergetischen Elektronen verwendet werden. Links sind die prozentualen Werte für Signal-Elektronen angegeben, rechts die der Untergrund-Teilchen. Diese Matrizen wurden mit Hilfe des Programms TMVA und dem Datensatz *MC12WHbbfiltered* erstellt.

## 8.4 Training der neuen Methoden

Um die neuen Methoden zur Identifikation der niederenergetischen Elektronen zu trainieren, zu testen und konfigurieren, wurden die hierzu benötigten Variablen für Signal-Elektronen und Untergrundteilchen separat in Dateien geschrieben. Als Quelle für die Variablen dienten die Datensätze *CSC11WHbb* und *MC12WHbbfiltered*.

Die Konsistenz dieser Datensätze wurde im vorherigen Abschnitt gezeigt. Insgesamt standen somit ca. 18000 niederenergetische Elektronen und mehrere hunderttausend Untergrundteilchen zur Verfügung, von denen jeweils 10000 zum Training der Methoden und 8000 Signal-Elektronen und 20000 Untergrund-Teilchen zur Auswertung der Methoden verwendet wurden. Diese Teilchen, welche zum Training verwendet wurden, und die Teilchen, die zur Auswertung der Methoden verwendet wurden, waren statistisch unabhängig voneinander. Dadurch konnte Anhand der Ereignisse zur Auswertung die Methoden auf ein eventuelles Übertrainieren getestet werden. Für das Training und die Auswertung der verschiedenen Methoden wurde das in ROOT integrierte TMVA-Programmpaket verwendet.

Das Herausschreiben der Variablen geschah nach der Spurselektion mit Hilfe der Qualitätsschnitte. Durch das Herausschreiben der Daten musste in der Testphase das ATHENA-Framework nicht verwendet werden, sondern nur das TMVA-Programmpaket. Allerdings konnten so während der Optimierungsphase der Methoden die Leistung der Likelihood-Methode des SoftElectronTag und der in TMVA implimentierten Methoden nicht direkt verglichen werden. Denn die Likelihood-Methode des SoftElectronTag verwendet im Gegensatz zur Likelihood-Methode des TMVA die Definition 6.10 für das Likelihood-Verhältnis und wird, wie zuvor schon beschrieben, für fünf Intervalle der Pseudorapidität  $\eta$  getrennt berechnet.

Um die Leistung von Methoden beurteilen und die Leistung mehrerer Methoden miteinander vergleichen zu können, müssen zwei Gesichtspunkte berücksichtigt werden:

1. Die Leistungsfähigkeit einer Methode wird durch die Effizienz bestimmt, mit der diese Methode Ereignisse der Signalklasse als solche erkennt.
2. Es ist die Effizienz zu berücksichtigen, mit der Ereignisse der Untergrundklasse fälschlicherweise als Signal klassifiziert werden.

Um diese Effizienzen für eine Methode zu bestimmen, werden von den Testereignissen die Ausgabewerte über die Klassenzugehörigkeit von der jeweiligen Methode berechnet. Die normierten Verteilungen dieser Ausgabewerte für Signal und Untergrund ergeben näherungsweise die Wahrscheinlichkeitsverteilungen der Ausgabewerte dieser Methode für Signal ( $P_{Sig}$ ) und für den Untergrund ( $P_{Bkg}$ ), im Folgenden kurz  $P_{Sig/Bkg}$  genannt, für die Klassifikation mit der jeweiligen Methode. Als nächster Schritt muss ein Arbeitspunkt gewählt werden. Dieser Arbeitspunkt  $x$  gibt den Mindestausgabewert an, die ein Ereignis haben muss, um als Signalereignis zu gelten. Dies bedeutet, dass für die von einer Methode berechneten Ausgabewerte  $p$  für ein Ereignis  $p > p_0$  gelten muss. Ist dies der Fall, so wird das Ereignis als Signalereignis akzeptiert, ansonsten gilt es als Untergrundereignis. Mit einem solchen Arbeitspunkt  $p_0$  und den Wahrscheinlichkeitsverteilungen  $P_{Sig/Bkg}$  für die Ausgabewerte kann die Effizienz einer Methode für die Signalklasse  $\epsilon_{Sig}$

oder die Effizienz für die Untergrundklasse  $\epsilon_{Bkg}$  wie folgt definiert werden:

$$\epsilon_{Sig/Bkg} = \frac{N_{Sig/Bkg}^+}{N_{Sig/Bkg}} = \frac{\int_{p_0}^{\infty} P_{Sig/Bkg}}{\int_{-\infty}^{\infty} P_{Sig/Bkg}}. \quad (8.2)$$

$N_{Sig/Bkg}^+$  ist hier die Anzahl der als Signal akzeptierten Signal- oder Untergrundereignisse im Testensemble,  $N_{Sig/Bkg}$  die gesamte Anzahl an Signal- oder Untergrundereignisse im Testensemble. Mit einem gewählten Arbeitspunkt erhält man somit stets eine bestimmte Effizienz, ein Signalereignis als solches zu identifizieren. Andererseits werden fast immer Untergrundereignisse fälschlicherweise als Signalereignisse akzeptiert. Eine gute Trennleistung für eine Methode bedeutet also eine möglichst große Signaleffizienz bei möglichst geringer Untergrundeffizienz. Um im Folgenden den direkten Vergleich mehrerer Methoden zu vereinfachen, wird daher in einer Graphik die inverse Untergrundeffizienz  $1/\epsilon_{bkg}$  gegen die Signaleffizienz  $\epsilon_{sig}$  aufgetragen. Liegt diese Kurve für eine Methode in einem Bereich der Signaleffizienz höher als die Kurven anderer Methoden, so ist die Trennleistung dieser Methode in diesem Bereich besser.

### 8.4.1 Konfiguration der Likelihood-Methode

Die Likelihood-Methode des TMVA-Programmpaketes wurde verwendet, um die Leistung der anderen Methoden abschätzen zu können, und nicht, um sie direkt in ATHENA einzubinden. Dennoch wurde eine Feinabstimmung der Parameter durchgeführt und getestet, welche der in Kapitel 6 beschriebenen Optionen der Likelihood-Methode die beste Trennung von Signal und Untergrund erreicht. Es zeigte sich, dass die Trennleistung aller Einstellungsmöglichkeiten sehr ähnlich war. So erreichen die Analysen mit Kern-Dichte-Abschätzer und verschiedenen Randbehandlungen nahezu die gleiche Leistung wie die Methoden, die die Wahrscheinlichkeits-Dichte mit Polynomen annähern. Dies wird in Abbildung 8.6 auf der linken Seite dargestellt und folgt aus der Tatsache, dass sich die modellierten Wahrscheinlichkeitsdichten der beiden Methoden nicht signifikant voneinander unterscheiden.

Für die anschließenden Vergleiche mit den anderen Klassifikations-Methoden wird eine Likelihood-Methode verwendet, die die Wahrscheinlichkeitsdichtefunktionen mit Hilfe von Polynomen zweiten Grades annähert.

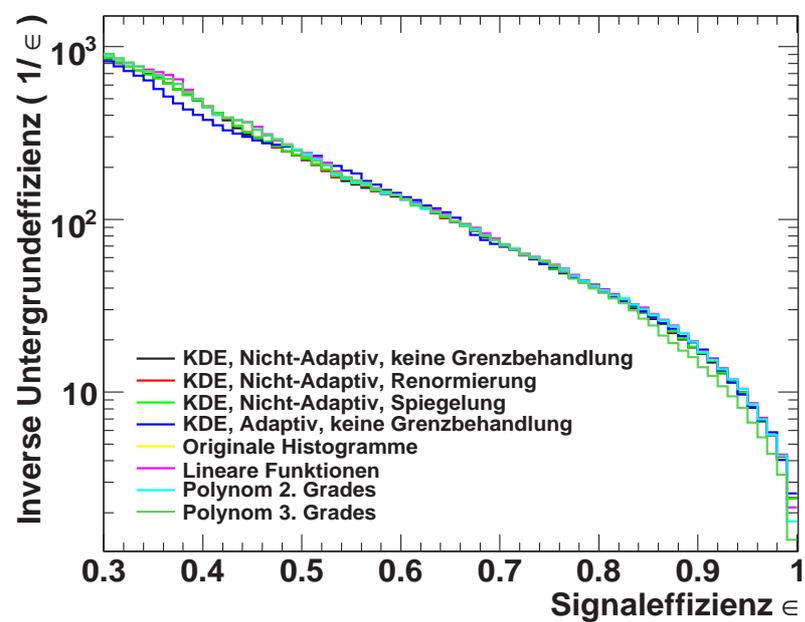


Abbildung 8.6: Vergleich der Trennungsfähigkeit zwischen den verschiedenen Einstellungen der Likelihood-Methoden.

### 8.4.2 Konfiguration der Neuronalen Netzwerke

Um eine möglichst gute Trennung von Signal und Untergrund zu erreichen, wurden verschiedene Konfigurationen für die Neuronalen Netzwerke getestet. Im Folgenden bezeichnet  $N$  die Anzahl der Variablen, die zur Trennung von Signal und Untergrund verwendet werden. In dieser Analyse ist  $N = 13$  zu setzen, da die Variablen aus Tabelle 8.1 zusammen mit der Pseudorapidität  $\eta$  verwendet werden.

#### Neuronale Netzwerke mit einer versteckten Lage

Bei den Netzwerken mit einer versteckten Lage zeigte sich, dass die Trennungsfähigkeit nicht sehr stark von der Zerfallsrate des Trainingsparameters abhängt. Dies zeigt zum Beispiel die Abbildung 8.7 (oben) für ein Netzwerk mit  $N + 12$  Knoten in der versteckten Lage. Schon mit einer kleinen Zerfallsrate (einem großen  $\tau$ ) ist die Konvergenz dieses Netzwerkes gewährleistet. Diese Konvergenz, also die Konvergenz der Gewichte gegen einen Satz, der die Fehlerfunktion minimiert, wird für dieses Netzwerk mit  $\tau = 0.005$  in der Abbildung 8.7 (unten) gezeigt. Hier sind die Werte der Fehlerfunktion über der Anzahl der Trainings-Zyklen aufgetragen. Nach genügend vielen Trainingszyklen verändern sich die Werte der Fehlerfunktion kaum noch. Dies gilt sowohl für die Werte der Funktion für die Trainingsereignisse, als auch für die Werte der Funktion für die statistisch unabhängigen Testereignisse. Eine größere Zerfallsrate (ein kleineres  $\tau$ ) bewirkt, dass die Anzahl der Trainingszyklen erhöht werden muss, um die Konvergenz gegen das Minimum der Fehlerfunktion zu erreichen.

Die Anzahl der Knoten des Netzwerkes ist für die Leistung eines Neuronalen Netzwerkes mit einer versteckten Lage viel entscheidender. Nur große Netzwerke sind in der Lage, die teilweise geringen Unterschiede zwischen signal- und untergrundartigen Ereignissen im  $N$ -dimensionalen Variablenraum zu beschreiben und ein globales Minimum der Fehlerfunktion zu erreichen. Ab etwa  $N + 8$  Knoten in der versteckten Lage zeigte sich keine signifikante Änderung im Trennungspotential der Netzwerke mehr. Dies zeigt Abbildung 8.8. Für die abgebildeten Netzwerke wurden die Anzahl der Zyklen und der Wert für  $\tau$  jeweils so angepasst, dass die Konvergenz der Netzwerke erreicht werden konnte.

Im Folgenden wurde als Konfiguration für ein Neuronales Netzwerk mit einer versteckten Lage das Netzwerk mit  $N + 12$  Knoten gewählt. Weiterhin wurden  $\tau = 0.005$  und  $N_{\text{Zyklen}} = 2000$  gewählt. Für dieses Netzwerk sind die Werte der Fehlerfunktion in Abbildung 8.7 (unten) dargestellt.

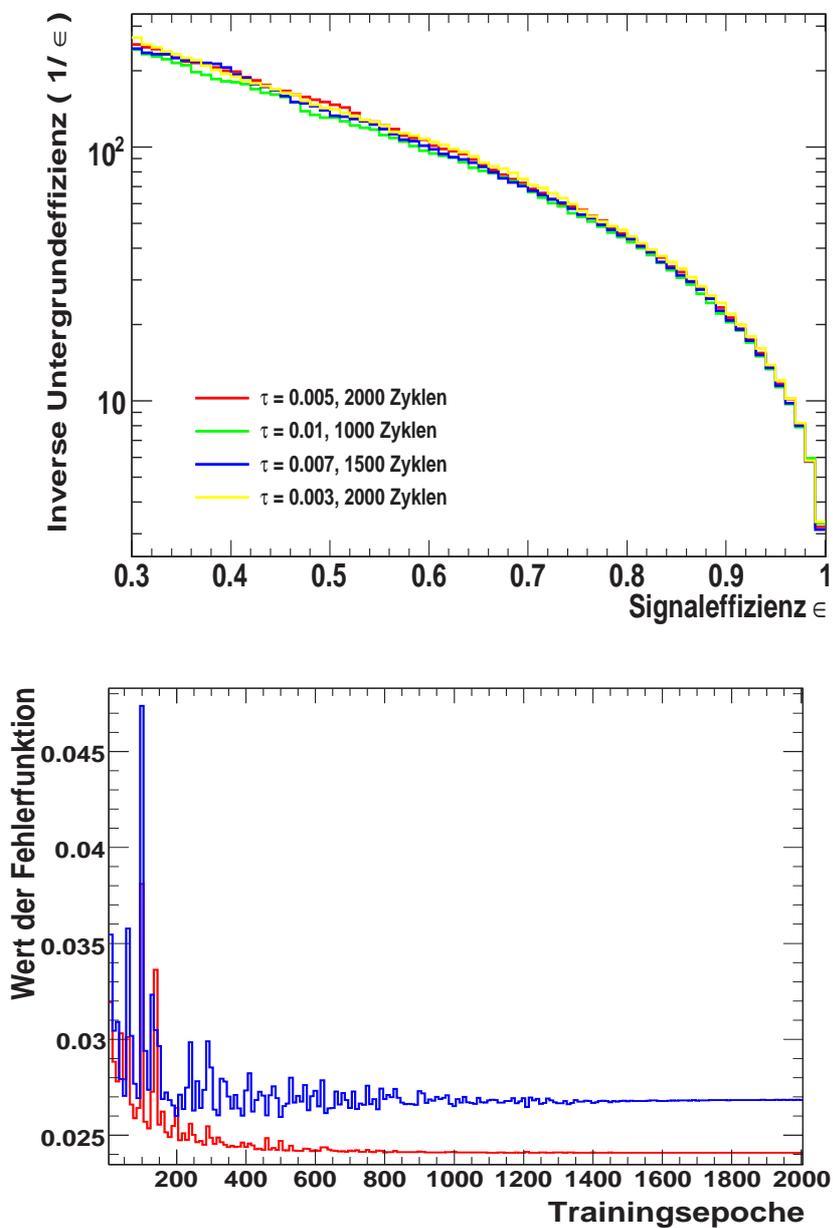


Abbildung 8.7: Oben ist die Trennleistung eines Neuronalen Netzwerkes mit  $N + 12$  Knoten in der versteckten Lage für verschiedene Werte von  $\tau$  gezeigt. Die untere Abbildung zeigt die Werte der Fehlerfunktion  $E$  für das Netzwerk für die Trainingsereignisse (in Rot) und die Testereignisse (in Blau) als Funktion des Trainingszyklus (der Epoche) mit  $\tau = 0.005$ .

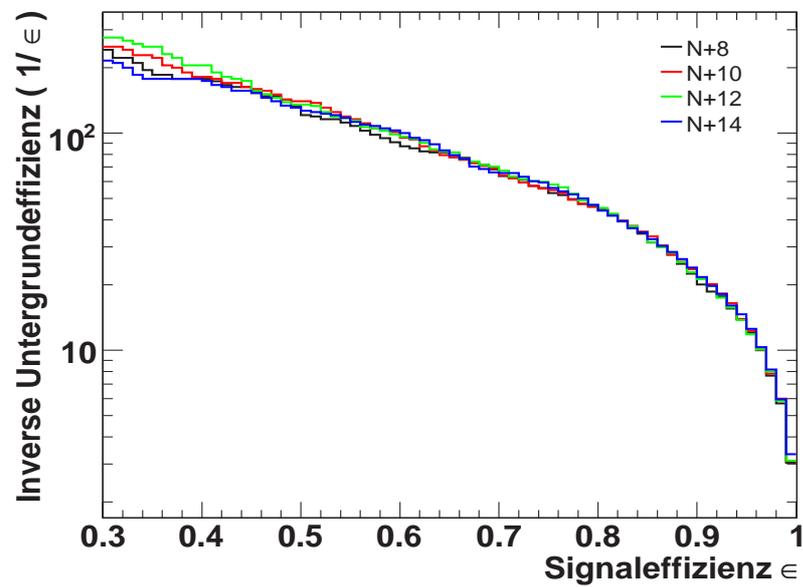


Abbildung 8.8: Die Auswirkung der Konfiguration eines Neuronalen Netzwerkes mit einer versteckten Lage auf die Trennfähigkeit. Die Zerfallsrate und die Anzahl der Trainingszyklen wurden für jedes Netzwerk angepasst.

### Neuronale Netzwerke mit zwei versteckten Lagen

Da sich, wie in Kapitel 6 beschrieben, nur einfache Probleme mit Netzwerken mit einer versteckten Lage lösen lassen, kann das Trennungspotential von Netzwerken mit zwei Lagen besser ausfallen. Für diese Netzwerke mit zwei versteckten Lagen ist die Einstellung der Zerfallsrate sehr wichtig. Abbildung 8.9 zeigt, dass eine zu geringe Zerfallsrate (ein zu großer Wert für  $\tau$ ) für das Training eines Netzwerkes der Konfiguration  $N, N - 1$  bewirkt, dass die Gewichte des Netzes zu früh nicht mehr angepasst werden. Dadurch kann sich nicht das bestmögliche Trennungspotential dieser Netzwerkkonfiguration entfalten. Dies wird besonders gut ersichtlich, wenn man die Werte der Fehlerfunktion  $E$  während des Trainings betrachtet. In Abbildung 8.10 werden für ein Neuronales Netzwerk mit dem Layout  $N, N - 1$  diese Werte während des Trainings für die Trainingsereignisse (rote Linie) und für die Testereignisse (blaue Linie) als Funktion der Trainingsepoche gezeigt. Es ist deutlich zu sehen, dass die Werte der Fehlerfunktion bei einem geringeren Wert für  $\tau$  kleiner werden können. Die Gewichte des Netzwerkes wurden also besser an das Minimum der Fehlerfunktion angepasst, wodurch ein besseres Trennungspotential erreicht wird. Die Abbildung 8.9 zeigt, dass eine weitere Verkleinerung von  $\tau < 0.001$  für ein Neuronales Netzwerk mit  $N, N - 1$  Knoten keine Leistungssteigerung mehr herbeiführt: Das Minimum der Fehlerfunktion scheint erreicht zu sein.

Die Anzahl der Knoten wirkte sich ebenfalls stark auf das Trennungspotential der Neuronalen Netzwerke mit zwei versteckten Lagen aus. Dies wird in Abbildung 8.11 gezeigt. Eine sehr gute Leistung erreichte hierbei ein Netzwerk der Konfiguration  $N, N - 1$ . Durch noch mehr Knoten konnte keine Verbesserung der Klassifikationsleistung mehr erzielt werden.

Im Folgenden wurde als Konfiguration für ein Neuronales Netzwerk mit zwei versteckten Lagen die Anzahl der Knoten in der ersten versteckten Lage zu  $N$ , die in der zweiten versteckten Lage zu  $N - 1$  gewählt. Die Zerfallsrate wurde mit  $\tau = 0.001$  festgelegt und das Netzwerk 3000 Zyklen lange trainiert. Die Werte der Fehlerfunktion für dieses Netzwerk sind in Abbildung 8.12 abgebildet.

### Neuronale Netzwerke mit drei versteckten Lagen

Auch Neuronale Netzwerke mit drei versteckten Lagen wurden hinsichtlich ihrer Fähigkeit zur Trennung von Signal-Elektronen und Untergrundteilchen untersucht. Die Leistung der Neuronalen Netzwerke mit zwei versteckten Lagen konnte durch eine zusätzliche Lage jedoch nicht mehr verbessert werden. Dies war zu erwarten, da selbst für die Lösung komplizierter Probleme zwei versteckte Lagen ausreichen (Siehe hierzu Kapitel 6).

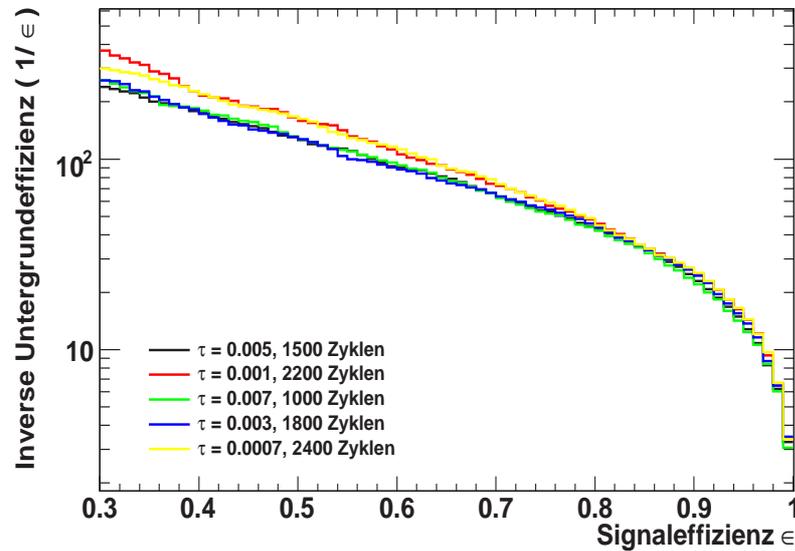


Abbildung 8.9: Das Trennpotential für ein Neuronales Netzwerk der Konfiguration  $N, N - 1$  für verschiedene Werte von  $\tau$ . Aus der Abbildung folgt, dass eine weitere Verkleinerung von  $\tau < 0.001$  keine Leistungssteigerung mehr mit sich bringt.

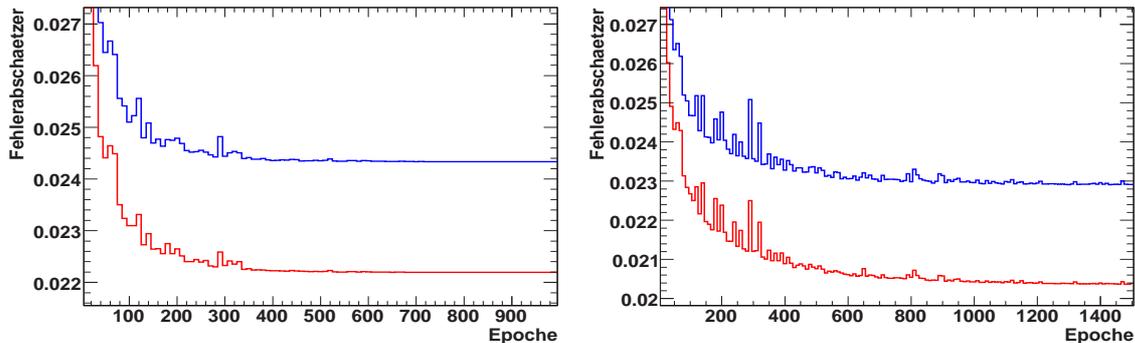


Abbildung 8.10: Links sind die Werte der Fehlerfunktion  $E$  (Fehlerabschätzer) für ein Neuronales Netzwerk der Konfiguration  $N, N - 1$  während des Trainings mit  $\tau = 0.01$  gezeigt. Die rote Linie entspricht den Werten für die Trainingsereignisse, die blauen denen der Testereignisse. Rechts wurde das gleiche Netzwerk trainiert, wobei  $\tau$  verkleinert wurde:  $\tau = 0.003$ . Dadurch konnte das Netzwerk einen geringeren Wert für die Fehlerfunktion erreichen.

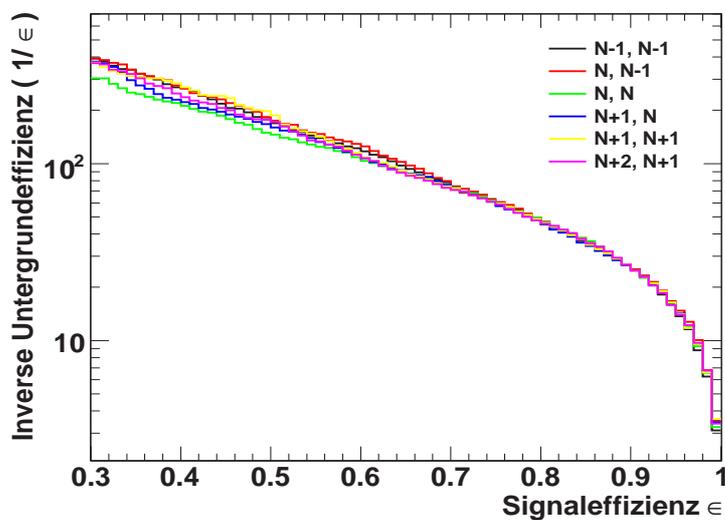


Abbildung 8.11: Die Auswirkung der Konfiguration eines Neuronalen Netzwerkes mit zwei versteckten Lagen auf die Trennfähigkeit. Die Zerfallsrate und die Anzahl der Trainingszyklen wurden für jedes der Netzwerke angepasst.

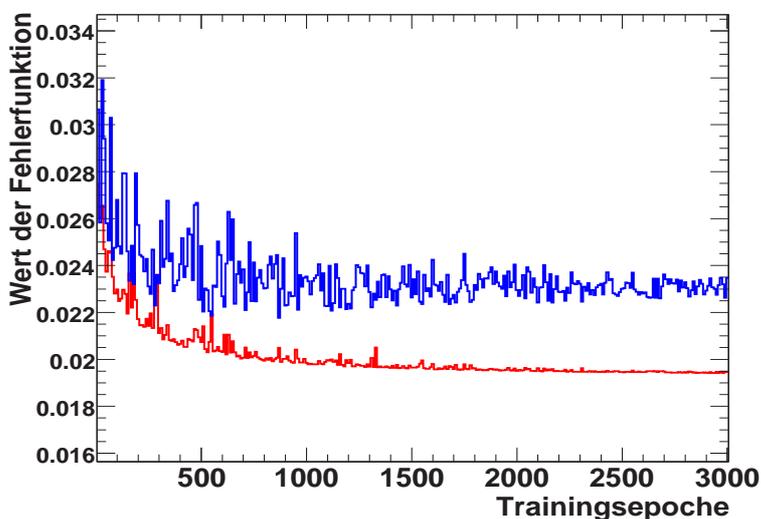


Abbildung 8.12: Werte der Fehlerfunktion des Neuronalen Netzwerkes mit der Konfiguration  $(N, N - 1)$  während des Trainings. Die rote Linie zeigt die Werte des Trainingsensembles, die blaue Linie die Werte des Testensembles. Der Wert für  $\tau$  betrug  $\tau = 0.001$  und das Netzwerk wurde 3000 Zyklen lange trainiert.

### 8.4.3 Konfiguration der Entscheidungsbäume

Für das Training der Entscheidungsbäume wurde im Rahmen dieser Arbeit nur die Boost-Methode *Boosting* betrachtet. Die Anzahl der Bäume hatte keinen großen Einfluss auf die Leistung der Methode. So zeigt sich auf der rechten Seite der Abbildung 8.13, dass bereits 400 Bäume ausreichen, um eine gute Leistung zu erhalten, und weitere Bäume keine Verbesserung herbeiführen. Auch die verschiedenen Trennungskriterien unterschieden sich hinsichtlich der Leistung der Bäume bei der Klassifikation der Ereignisse mit Ausnahme der Methode der Statistischen Signifikanz kaum voneinander. Dies wird in Abbildung 8.14 auf der rechten Seite gezeigt. Da das Trennungskriterium der Entropie im Bereich der mittleren und hohen Signaleffizienzen eine gute Leistung erbrachte, wurde die Entropie als Trennungskriterium für zukünftige Vergleiche gewählt.

Die Auswahl der Beschneidungsmethode ist für das gegebene Problem nicht sehr entscheidend, wie die linke Seite der Abbildung 8.14 zeigt. Die Beschneidung nach dem Erwarteten Fehler in Verbindung mit einer Beschneidungsstärke von 7.0 oder die CostComplexity-Beschneidung mit einer Stärke von 0.6 wiesen ein nahezu gleich großes Potential zur Trennung von Signal und Untergrund auf. Dies war allerdings nur nach geeigneter Wahl der Beschneidungsstärken der Fall. Abbildung 8.15 (unten) zeigt, dass für die Cost Complexity-Methode eine Variation um den Wert 0.6 eine kleine Verschlechterung der Trennleistung bewirkt. Eine größere Abweichung, wie zum Beispiel eine Beschneidungsstärke von 6.0, resultiert in einer deutlich schlechteren Leistung der Methode. Die Methode der Beschneidung nach dem Erwarteten Fehler lieferte ab einer Beschneidungsstärke von 6.0 bis hin zu einer getesteten Beschneidungsstärke von 11.0 konstante Ergebnisse, was in Abbildung 8.15 (oben) dargestellt wird.

Als Standardeinstellungen wurden für die weitere Analyse folgende Einstellungen für die Entscheidungsbäume gewählt:

- Methode: Boosting
- Anzahl der Bäume: 400
- Trennungskriterium: Entropie
- Beschneidungsmethode: Erwarteter Fehler
- Beschneidungsstärke: 7.0

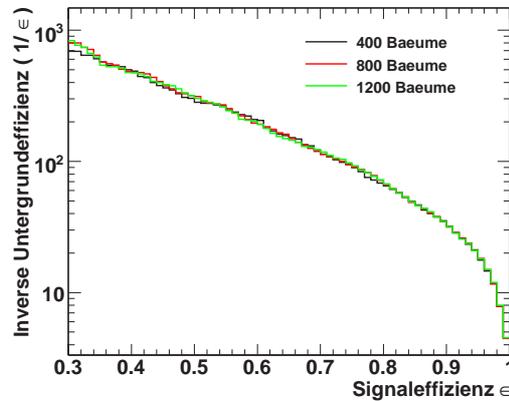


Abbildung 8.13: Gezeigt ist der Einfluss der Anzahl der trainierten Bäume auf die Leistung der Entscheidungsbäume. Für diese Abbildung wurden als Trennkriterium der Gini-Index, das Beschneiden nach dem Erwarteten Fehler und eine Beschneidungsstärke 7.0 gewählt.

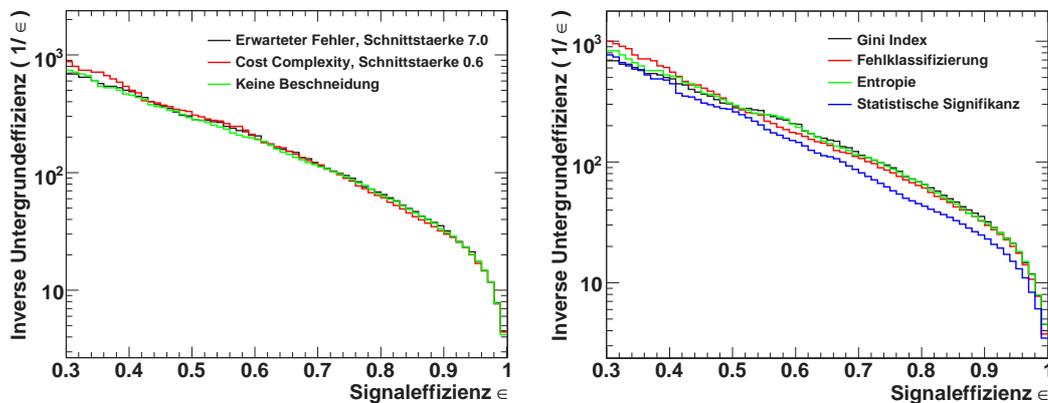


Abbildung 8.14: Gezeigt ist auf der linken Seite der Einfluss der Beschneidungsmethode auf die Leistung der Entscheidungsbäume, die mit der Boost-Methode trainiert wurden. Die Beschneidungsstärke wurde für das Schneiden nach dem Erwarteten Fehler zu 7.0 und für die Cost Complexity-Beschneidung zu 0.6 gewählt. Die Anzahl der trainierten Bäume betrug 400. Die rechte Seite zeigt den Einfluss des Trennungskriteriums auf das Trennungspotential der Entscheidungsbäume, die mit der Boost-Methode trainiert wurden. Beschnitten wurden die Bäume mit der Methode des Erwarteten Fehlers mit einer Beschneidungsstärke von 7.0 .

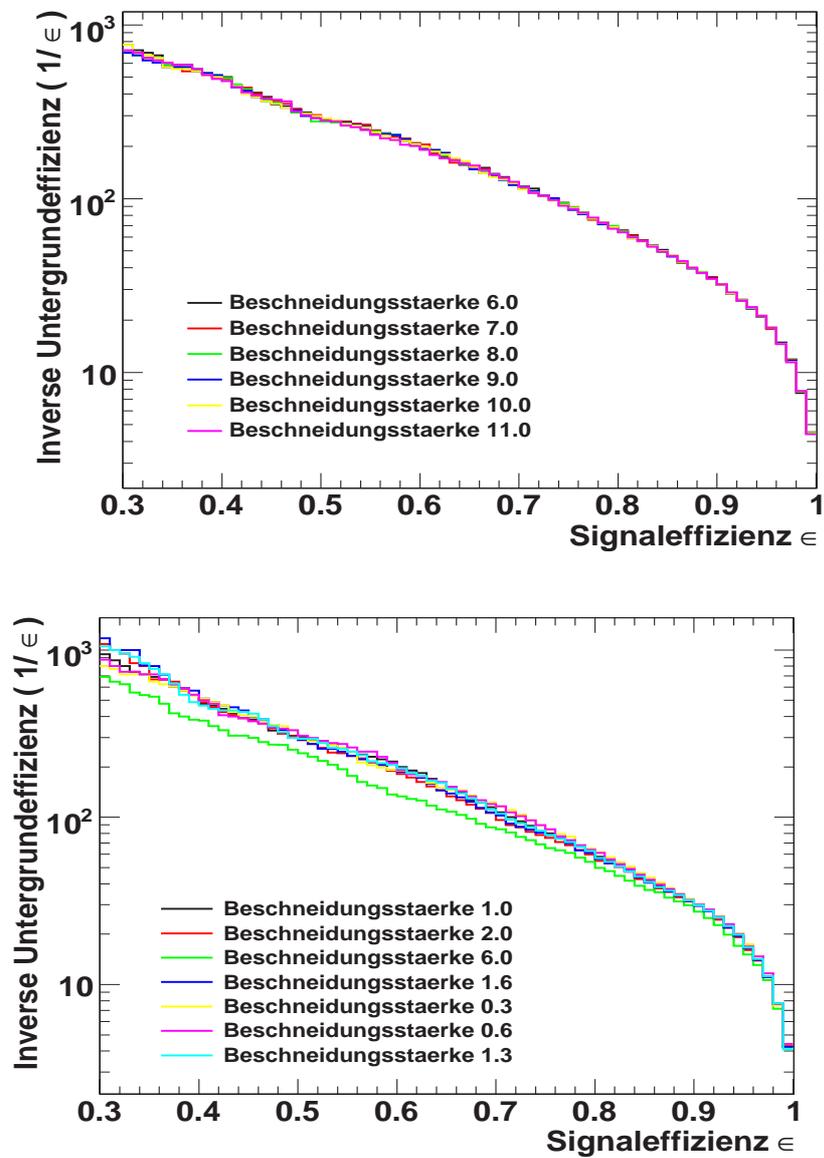


Abbildung 8.15: Gezeigt ist der Einfluss der Beschneidungsstärke auf die Leistungsfähigkeit der Entscheidungsbäume, die mit der Boost-Methode trainiert wurden und mit den Methoden des Erwarteten Fehlers (oben) und der Cost Complexity (unten) beschnitten wurden.

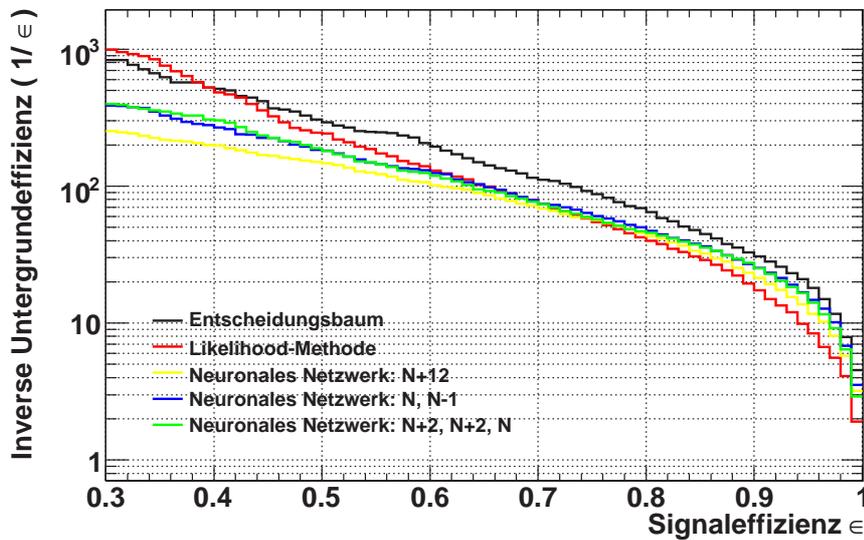


Abbildung 8.16: Gezeigt ist ein Vergleich der Klassifikationseffizienzen der verschiedenen Methoden.

#### 8.4.4 Vergleich der Methoden

Nach der Auswahl der jeweils besten Einstellungen der verschiedenen Methoden wurden die Klassifikationsleistungen für die Identifikation der niederenergetischen Elektronen miteinander verglichen. Dies wird in Abbildung 8.16 gezeigt. Aus der Abbildung geht hervor, dass in den Bereichen mit hohen Signaleffizienzen durch die Verwendung der Neuronalen Netzwerke und der Entscheidungsbäume eine Verbesserung gegenüber der Likelihood-Methode erreicht werden konnte. Besonders die Klassifikation mit Hilfe der Entscheidungsbäume konnte in den Bereichen der mittleren und hohen Signaleffizienzen ab etwa  $\epsilon > 0.4$  die Trennung von Signal und Untergrund stark verbessern. Die Neuronalen Netzwerke lieferten im Vergleich zur Likelihood-Methode erst in den Bereichen ab  $\epsilon > 0.7$  eine verbesserte Klassifikationsleistung. Wie im vorherigen Abschnitt schon erwähnt wurde, konnte die Leistung der Neuronalen Netzwerke mit zwei versteckten Lagen durch hinzunahme einer weiteren versteckten Lage nicht mehr verbessert werden.

### 8.5 Effizienz der Methoden bei der Identifikation der niederenergetischen Elektronen in ATHENA

Um die neuen Methoden der Elektronenidentifikation mit der in ATHENA implementierten Likelihood-Methode vergleichen zu können und eine zukünftige Verwendung dieser neuen Methoden zu ermöglichen, wurden diese in ATHENA integriert. Hierzu wurden mit dem TMVA-

Programmpaket die im Training ermittelten Gewichte und Einstellungen in Dateien gespeichert, welche über das TMVA-Programm wieder eingelesen und verwendet werden können.

Da in der in dieser Analyse verwendeten ATHENA-Version das TMVA-Programmpaket noch nicht integriert war <sup>3)</sup>, wurde die TMVA-Bibliothek der Version 3.8.7 mit den benötigten ATHENA-Bibliotheken verknüpft.

Im Folgenden wird die Klassifikationsleistung der neuen Methoden bei der Identifikation der niederenergetischen Elektronen aus dem Zerfall von b-Hadronen oder der daraus entstandenen c-Hadronen im Vergleich zur Leistung der Likelihood-Methode des *softeBuilder* für verschiedene Datensätze präsentiert und diskutiert. Dabei wurden alle Teilchen berücksichtigt, deren Spuren die Qualitätsschnitte erfolgreich passierten.

Die Klassifikation der Neuronalen Netzwerke mit einer Lage war jedoch, wie oben beschrieben, deutlich ineffizienter als mit den anderen Methoden. Weiterhin erreichten die Neuronalen Netzwerke mit zwei und drei versteckten Lagen eine nahezu identische Trennungsfähigkeit. Dies war zu erwarten, da die Hinzunahme einer zusätzlichen dritten Lage selbst für komplizierte Probleme nicht notwendig ist. Aufgrund dieser Fakten wurden für die folgenden Vergleiche von den zuvor beschriebenen Methoden nur das Neuronale Netzwerk mit zwei versteckten Lagen und die Entscheidungsbäume berücksichtigt.

Da ungefähr die Hälfte der Signal-Elektronen des Datensatzes *MC12WHbbfiltered* zum Training der neuen Methoden verwendet wurde, wird für diesen eine leichte Verbesserung der Klassifikation im Vergleich zu einem unabhängigen Datensatz erwartet, welcher nicht zum Training der Methoden verwendet wurde. Dieser Effekt sollte jedoch sehr gering sein, da ein Übertrainieren der Methoden nicht beobachtet werden konnte und somit auch mit einem unabhängigen Ensemble an Ereignissen nahezu die gleiche Trennleistung von Signal und Untergrund erreicht werden sollte, wie mit dem Trainingsensemble. Die mit den verschiedenen Methoden erreichten inversen Untergrundeffizienzen als Funktion der Signaleffizienz sind in Abbildung 8.17 graphisch dargestellt. Die Ausgabewerte der Methoden zeigen die Abbildungen 8.18, 8.19 und 8.20. Für einige ausgewählte Signaleffizienzen sind die genauen Werte für die inversen Untergrundeffizienzen in Tabelle 8.2 aufgelistet. Besonders die Entscheidungsbäume konnten in allen Bereichen der Signaleffizienz die Identifikation der niederenergetischen Elektronen um einen Faktor  $\geq 1.5$  verbessern. Die Leistung des Neuronalen Netzwerkes und der Likelihood-Methode sind im Bereich Signaleffizienzen von 0.4 bis 0.9 nahezu identisch. Für Signaleffizienzen kleiner 0.4 ist die Leistung der Likelihood-Methode besser als die des Neuronalen Netzwerkes.

Auch der Datensatz *CSC11WHbb* wurde zum Trainieren der Methoden verwendet. Daher wird auch für diesen im Vergleich zu einem Datensatz, welcher nicht zum Training der Methoden verwendet wurde, eine leicht verbesserte Trennleistung von Signal und Untergrund erwartet. Tabelle 8.3 listet die Werte der inversen Untergrundeffizienzen für einige Signaleffizienzen auf. Graphisch ist die inverse Untergrundeffizienz als Funktion der Signaleffizienz in Abbildung 8.21 dargestellt. Die mit diesem Datensatz erzielten Ausgabewerte der Methoden zeigen die Abbil-

---

<sup>3</sup>Eine ausreichend aktuelle ROOT-Version, in der das TMVA-Programmpaket integriert ist, befindet sich erst in ATHENA ab Version 13.

dungen 8.22, 8.23 und 8.24. Auch mit diesem Datensatz können die Entscheidungsbäume die Leistung der Likelihood-Methode weit übertreffen. Die Unterschiede zwischen diesen beiden Methoden betragen für diesen Datensatz für fast alle Signaleffizienzen einen Faktor von über zwei. Das Neuronale Netzwerk hingegen konnte nur im Bereich der Signaleffizienz zwischen 0.7 und 0.94 die Trennleistung der Likelihood-Methode erreichen oder verbessern.

Als weiterer Datensatz, mit dem die Identifikationseffizienzen der neuen Methoden getestet wurden, wurde der Datensatz *MC12t* gewählt. Dieser stellt einen wichtigen Untergrundkanal für viele Analysen dar. Durch den schwachen Zerfall der t-Quarks können ebenfalls Elektronen entstehen. Diese Elektronen werden in dieser Analyse dem Untergrund zugeordnet, da sie nicht aus dem Zerfall eines b-Hadrons oder dem Zerfall des daraus entstandenen c-Hadrons stammen. Daher wird mit diesem Datensatz eine deutlich schlechtere Leistung der Methoden bei der Identifikation der Signal-Elektronen erwartet, als es mit den anderen Datensätzen der Fall war. Abbildung 8.25 zeigt die mit diesem Datensatz erreichte Leistung aller Methoden. Das Neuronale Netzwerk kann die Identifikationsleistung für Signaleffizienzen kleiner 4.5 deutlich steigern. Die Entscheidungsbäume erreichten mit diesem Datensatz nur eine leicht bessere Klassifikationsleistung gegenüber der Likelihood-Methode. In Tabelle 8.4 sind für einige Signaleffizienzen die zugehörigen inversen Untergrundeigenschaften eingetragen. Diese Untergrundeigenschaften ergeben sich aus den erzielten Ausgabewerten der Methoden, welche in den Abbildungen 8.26, 8.27 und 8.28 gezeigt sind. An diesen Abbildungen lässt sich erkennen, dass die Elektronen aus dem Zerfall der t-Quarks eine signifikante Erhöhung der Untergrundverteilungen der Ausgabewerte für alle Methoden im Bereich der hohen Ausgabewerte bewirkt.

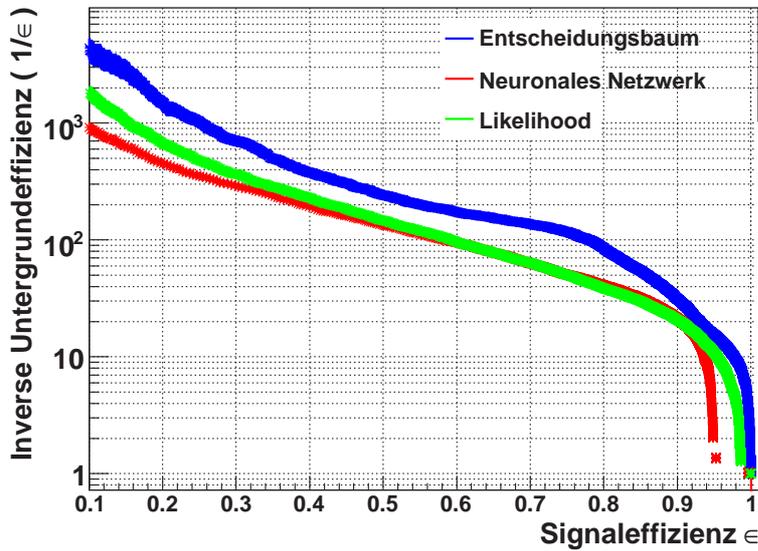


Abbildung 8.17: Gezeigt ist die Klassifikationsleistung bei der Identifikation der niederenergetischen Elektronen des Neuronales Netzwerkes mit zwei versteckten Lagen, der Entscheidungsbaume und der Likelihood-Methode des *softeBuilder* für den Datensatz *MC12WHbbfiltered* gezeigt.

$1/\epsilon_{Bkg}$	$\epsilon_{Sig} = 90\%$	$\epsilon_{Sig} = 80\%$	$\epsilon_{Sig} = 60\%$
Entscheidungsbäume	$31.30 \pm 0.54$	$85.11 \pm 2.46$	$173.17 \pm 7.18$
Neuronales Netzwerk	$21.28 \pm 0.30$	$40.79 \pm 0.81$	$95.01 \pm 2.91$
Likelihood-Methode	$21.55 \pm 0.31$	$38.63 \pm 0.74$	$95.94 \pm 2.94$

Tabelle 8.2: Inverse Untergrundeﬃzienzen  $1/\epsilon_{Bkg}$  der verschiedenen Methoden bei der Identifikation der niederenergetischen Elektronen für den Datensatz *MC12WHbbfiltered* und verschiedene Signaleﬃzienzen  $\epsilon_{Sig}$ .

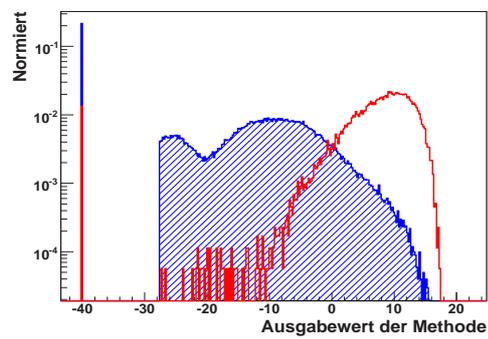


Abbildung 8.18: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

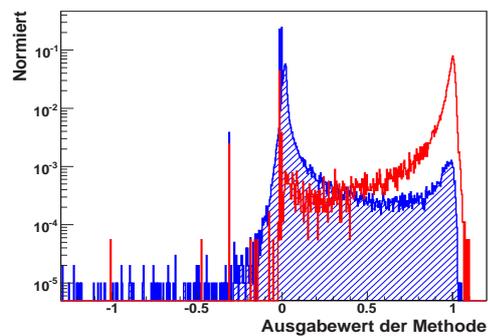


Abbildung 8.19: Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

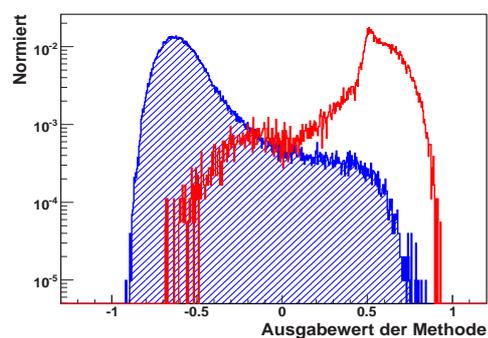


Abbildung 8.20: Ausgabewerte der Entscheidungsbäume für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

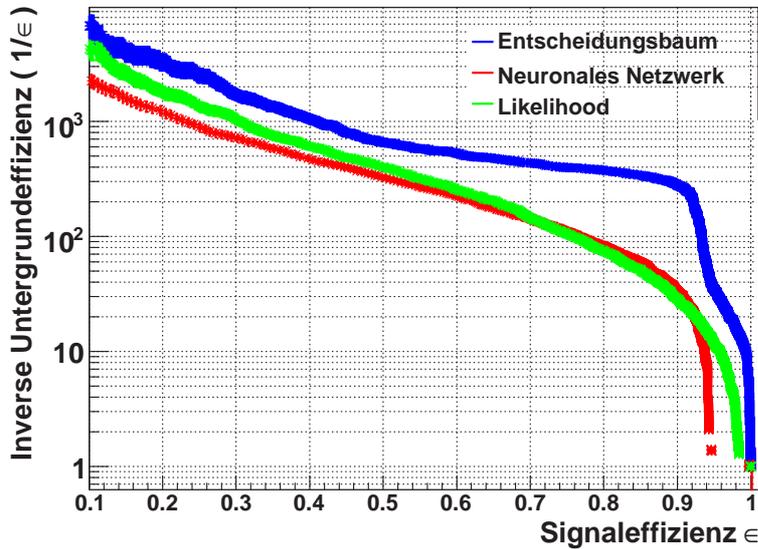


Abbildung 8.21: Gezeigt ist die Klassifikationsleistung des Neuronalen Netzwerkes mit zwei versteckten Lagen, der Entscheidungsbäume und der Likelihood-Methode des *SoftBuilder* bei der Identifikation der niederenergetischen Elektronen für den Datensatz *CSC11WHbb*.

$1/\epsilon_{Bkg}$	$\epsilon_{Sig} = 90\%$	$\epsilon_{Sig} = 80\%$	$\epsilon_{Sig} = 60\%$
Entscheidungsbäume	$280.00 \pm 13.03$	$374.42 \pm 20.16$	$530.04 \pm 33.97$
Neuronales Netzwerk	$33.15 \pm 0.52$	$82.05 \pm 2.06$	$223.56 \pm 9.31$
Likelihood-Methode	$28.50 \pm 0.41$	$74.42 \pm 1.77$	$253.94 \pm 11.25$

Tabelle 8.3: Inverse Untergrundeﬃzienzen  $1/\epsilon_{Bkg}$  der verschiedenen Methoden bei der Identifikation der niederenergetischen Elektronen für den Datensatz *CSC11WHbb* und verschiedene Signaleﬃzienzen  $\epsilon_{Sig}$  aufgelistet.

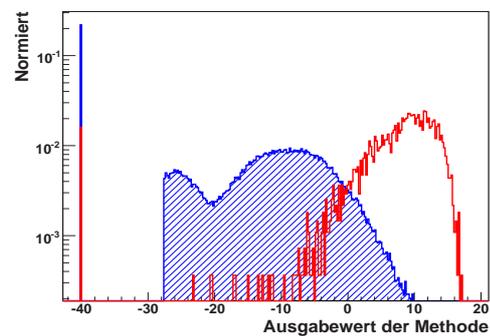


Abbildung 8.22: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

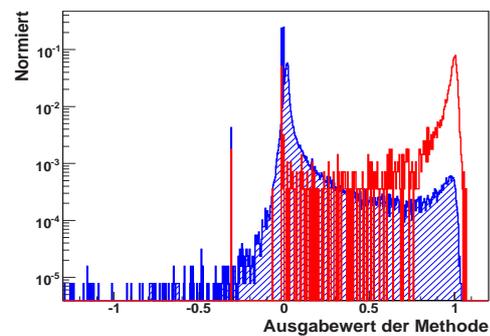


Abbildung 8.23: Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

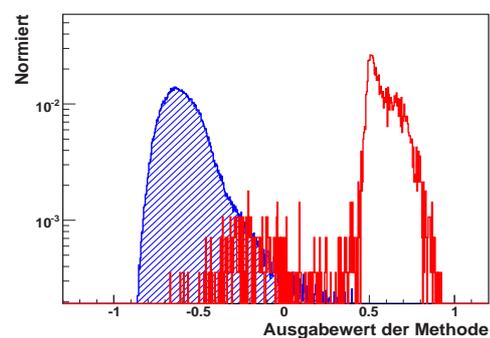


Abbildung 8.24: Ausgabewerte der Entscheidungsbäume für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

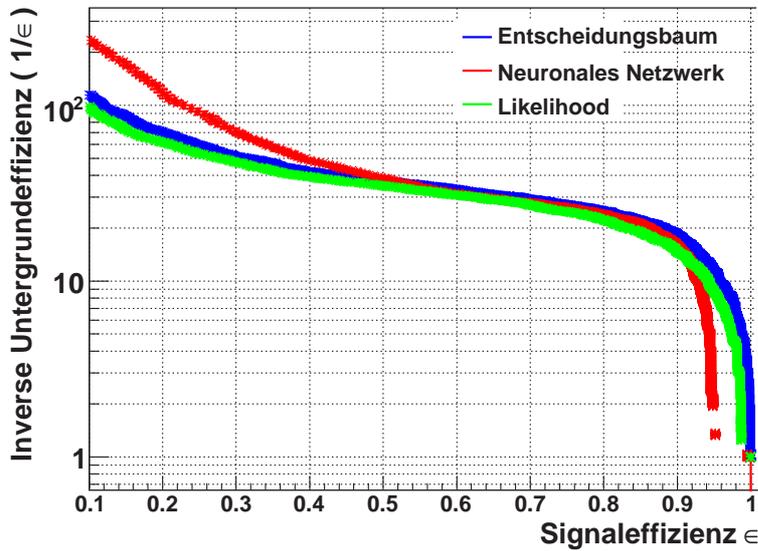


Abbildung 8.25: Klassifikationsleistung bei der Identifikation der niederenergetischen Elektronen gezeigt, die das Neuronale Netzwerk mit zwei versteckten Lagen, die Entscheidungsbäume und die Likelihood-Methode des *softeBuilder* mit dem Datensatz *MC12tt* erreichen.

$1/\epsilon_{Bkg}$	$\epsilon_{Sig} = 90\%$	$\epsilon_{Sig} = 80\%$	$\epsilon_{Sig} = 60\%$
Entscheidungsbäume	$18.27 \pm 0.23$	$24.90 \pm 0.37$	$33.12 \pm 0.58$
Neuronales Netzwerk	$15.86 \pm 0.18$	$23.79 \pm 0.35$	$31.50 \pm 0.53$
Likelihood-Methode	$14.70 \pm 0.16$	$22.41 \pm 0.31$	$30.83 \pm 0.52$

Tabelle 8.4: Inverse Untergrundeﬃzienzen  $1/\epsilon_{Bkg}$  der verschiedenen Methoden bei der Identifikation der niederenergetischen Elektronen für den Datensatz *MC12tt* und verschiedene Signaleﬃzienzen  $\epsilon_{Sig}$ .

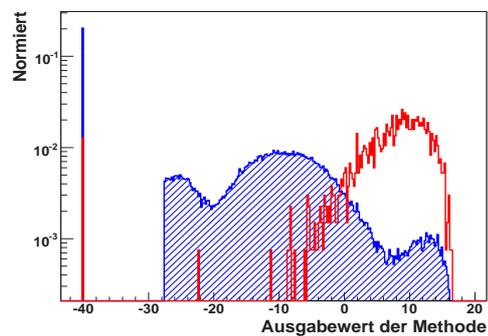


Abbildung 8.26: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

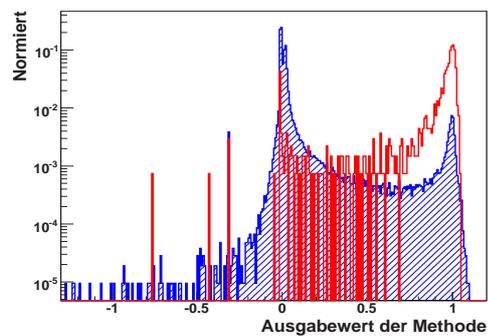


Abbildung 8.27: Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

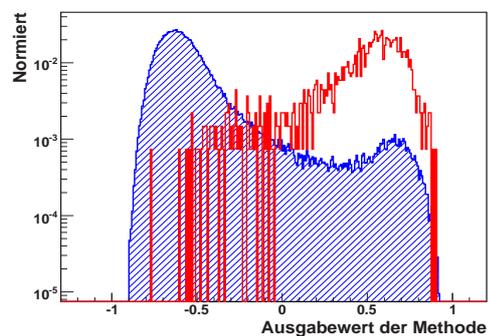


Abbildung 8.28: Ausgabewerte der Entscheidungsbäume für Signal (rot) und Untergrund (blau) bei der Identifikation der niederenergetischen Elektronen.

## 9 Anwendung der neuen Methoden auf die Markierung von b-Quarks

In diesem Kapitel wird zunächst der in ATHENA implementierte Algorithmus zur Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen vorgestellt. Mit allen im vorherigen Kapitel verwendeten Methoden zur Identifikation dieser Elektronen wird dann anhand von mehreren Datensätzen die Leistung dieser Methoden bei der Markierung der b-Quarks analysiert. Einer dieser Datensätze ist der Datensatz *MC12tt*. Wie im vorherigen Kapitel schon erwähnt wurde, ist mit diesem Datensatz die Identifikation der Elektronen aus dem Zerfall der b-Hadronen oder der daraus entstandenen c-Hadronen sehr schwierig, da auch beim Zerfall eines t-Quarks ein Elektron entstehen kann. Dieses Elektron kann mit den hier verwendeten Methoden kaum von den Signal-Elektronen unterschieden werden. Wie im vorherigen Kapitel schon gezeigt wurde, verschlechtert sich dadurch die Leistung der Methoden bei der Identifikation der niederenergetischen Elektronen aus dem Zerfall der b-Hadronen oder der daraus entstandenen c-Hadronen und damit auch die Leistung bei der Markierung von b-Quarks mit Hilfe dieser Elektronen. Zum Zeitpunkt dieser Analyse existierte noch keine Methode, um die Signal-Elektronen effektiv von den Elektronen aus dem Zerfall eines t-Quarks zu unterscheiden. Daher wurden für das Folgende alle Ereignisse, in denen aus dem Zerfall eines t-Quarks ein Elektron entsteht, in der Auswertung nicht berücksichtigt.

### 9.1 Algorithmus zur Markierung von b-Quarks

Um für die Markierung von b-Quarks speziell die Elektronen aus einem b-Jet und dadurch den b-Jet als solchen zu identifizieren, ist ein separater Algorithmus in ATHENA implementiert. Dieser heißt „SoftElectronTag“ und verbessert die Identifikation der niederenergetischen Elektronen aus dem b-Quark-Zerfall weiter und wendet diese Identifikation auf die Markierung von b-Quarks an. Für jeden Jet, von dem mindestens ein Elektronenkandidat die Qualitätsschnitte erfüllt, sucht dieser Algorithmus den Kandidaten innerhalb des Jets, der das höchste vom *softeBuilder* berechnete Likelihood-Verhältnis aufweist. Das Likelihood-Verhältnis dieses besten Elektronenkandidaten kann dann um das Likelihood-Verhältnis von zwei weiteren Variablen, dem transversalen Stoßparameter  $d_0$  der Spur und dem transversalen Impuls der Spur relativ zur Jet-Achse, erweitert werden. Diese beiden Variablen werden verwendet, um die Elektronen der b-Jets von den Elektronenkandidaten aus den Jets der leichteren Quarks zu unterscheiden.

Da die b-Quarks eine im Mittel weite Flugstrecke von der Größenordnung  $O(\text{mm-cm})$  zurücklegen, bevor sie zerfallen, haben die niederenergetischen Elektronen aus dem b-Quark-Zerfall einen relativ großen transversalen Stoßparameter. Bei den Elektronenkandidaten aus den u-Jets

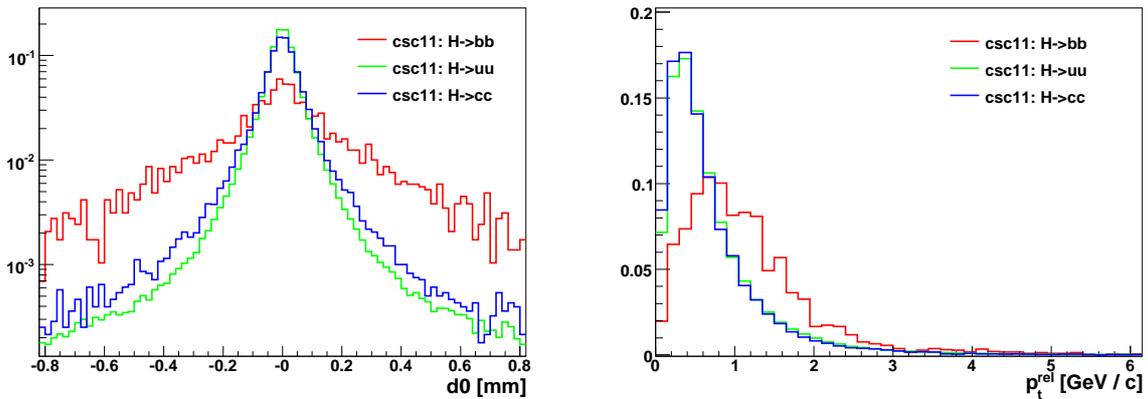


Abbildung 9.1: Die normierten Verteilungen des transversalen Stoßparameters sowie des transversalen Impulses relativ zur Jetachse für Signal-Elektronen aus dem Datensatz *CSC11WHbb* in rot und für alle Untergrund-Teilchen aus  $c$ -Jets (Datensatz *CSC11WHcc*) in blau, beziehungsweise  $u$ -Jets (Datensatz *CSC11WHuu*) in grün.

ist diese Variable sehr gering und aufgrund der Auflösung bei der Rekonstruktion der Spuren um null verteilt. Auch für die Kandidaten aus  $c$ -Jets ist der Wert dieser Variable im Mittel kleiner. Allerdings besitzen auch die  $c$ -Hadronen eine Lebensdauer und legen eine gewisse Flugstrecke zurück, bevor sie, eventuell sogar unter Auftreten eines Elektrons, zerfallen.

Der transversale Impuls relativ zur Jetachse ist für Signal-Elektronen aufgrund der großen Masse des  $b$ -Quarks im Mittel größer als für Teilchen, die von einem Jet eines leichteren Quarks stammen. Daher wird diese Variable ebenfalls verwendet, um die Signal-Elektronen der  $b$ -Jets von den Elektronenkandidaten der leichteren Jets zu unterscheiden.

Die Verteilungen dieser beiden Variablen werden für Signal-Elektronen aus dem Datensatz *CSC11WHbb* mit den Verteilungen der Elektronenkandidaten aus den Datensätzen *CSC11WHcc* und *CSC11WHuu* in Abbildung 9.1 verglichen.

Das aus allen Variablen berechnete Likelihood-Verhältnis des besten Elektronenkandidaten eines Jets bestimmt die Wahrscheinlichkeit, dass dieser Jet von einem  $b$ -Quark stammt. Auf dieses Likelihood-Verhältnis wird ein Schnitt angewandt und dadurch eine Trennung der  $b$ -Jets von anderen Jets erreicht.

## 9.2 Anwendung der neuen Methoden

Bei der Markierung von  $b$ -Quarks mit Hilfe der niederenergetischen Elektronen müssen zwei Szenarien berücksichtigt werden:

1. Der *SoftElectronTag* wird dazu verwendet, die auf der Lebensdauer der  $b$ -Hadronen ba-

sierenden Methoden zu kalibrieren oder mit diesen Methoden kombiniert zu werden. Dabei soll auf die Verwendung des transversalen Impulses relativ zur Jet-Achse  $p_t^{rel}$  und des transversalen Stoßparameters  $d0$  des Elektronenkandidaten eines Jets, der die höchste Wahrscheinlichkeit aufweist, ein Elektron aus einem b-Quark-Zerfall zu sein, verzichtet werden. Dadurch kann eine möglichst große Unabhängigkeit der beiden Ansätze zur Markierung von b-Quarks gewährleistet werden. Für diesen Fall entspricht die Wahrscheinlichkeit des Jets, ein b-Jet zu sein, der höchsten Wahrscheinlichkeit eines Elektronenkandidaten in diesem Jet, ein Elektron aus einem b-Quark-Zerfall zu sein.

2. Der *SoftElectronTag* soll selbstständig die Trennung der b-Jets von den Jets leichterer Quarks vornehmen und die bestmögliche Klassifikationsleistung erreichen. In diesem Fall können die Variablen  $p_t^{rel}$  und  $d0$  des Elektronenkandidaten eines Jets, der die höchste Wahrscheinlichkeit aufweist, ein Elektron aus einem b-Quark-Zerfall zu sein, verwendet werden.

Das erste Szenario wird durch die neuen Methoden im vorherigen Kapitel direkt abgedeckt.

Um das zweite Szenario abdecken zu können, wurde für jede der beiden neuen Methoden der höchste Ausgabewert für die Elektronenkandidaten für jeden Jet zusammen mit den Werten für  $p_t^{rel}$  und  $d0$  dieses Elektronenkandidaten mit dem höchsten Ausgabewert als Variablen für das Training weiterer Methoden verwendet. Der jeweils beste Kandidat für die Entscheidungsbäume und der des Neuronalen Netzwerkes mit zwei versteckten Lagen wurden getrennt betrachtet, da die Methoden für einen Jet nicht notwendigerweise denselben Kandidaten als den besten Elektronenkandidat auswählen. Die Signal-Elektronen, mit denen die Netzwerke trainiert wurden, stammen wieder aus den Datensätzen *MC12WHbbfiltered* und *CSC11WHbb*. Als Signal-Elektron wurde bei diesem Vorgang der beste Elektronenkandidat eines Jets angesehen, falls der Jet ein b-Jet war und das Elektron aus dem Zerfall des b-Hadrons oder des daraus entstandenen c-Hadrons stammt. Als Untergrund dienten alle besten Elektronenkandidaten aus den Jets von u-Quarks aus dem Datensatz *MC12WHuu*. Die besten Kandidaten von c-Quarks wurden nicht berücksichtigt. Dies wäre möglich gewesen, falls zum Beispiel ein Neuronales Netzwerk verwendet worden wäre, welches über die gleiche Menge an Ausgabeknoten verfügt, wie es zu trennenden Klassen gibt. In diesem Fall mit den b-Jets als Signal und den c- und u-Jets als Untergrund müsste ein solches Netzwerk drei Ausgabeknoten besitzen. Netzwerke mit mehreren Ausgabeknoten sind im TMVA-Programmpaket jedoch nicht enthalten. Eine weitere Möglichkeit wäre gewesen, für jede Untergrundklasse eine eigene Methode zu trainieren.

Die Methoden, welche die Variablen des besten Elektronenkandidaten für die Entscheidungsbäume verwendet und die beste Leistung bei der Trennung von b-Jets und u-Jets aufwies, waren die Neuronalen Netzwerke mit zwei versteckten Lagen ab einer Konfiguration der Anzahl an Knoten von  $N + 3, N + 1$ . Neuronale Netzwerke mit einer versteckten Lage und Entscheidungsbäume konnten diese Leistung nicht erreichen. Dies wird beispielhaft für jeweils einen Vertreter dieser beiden Methoden, welcher die beste Klassifikationsleistung für diese Methode erreichte, in Abbildung 9.2 auf der linken Seite gezeigt. Der optimale Wert für den Zerfallsparemeter für das Neuronale Netzwerk mit zwei versteckten Lagen konnte mit  $\tau = 0.00095$  be-

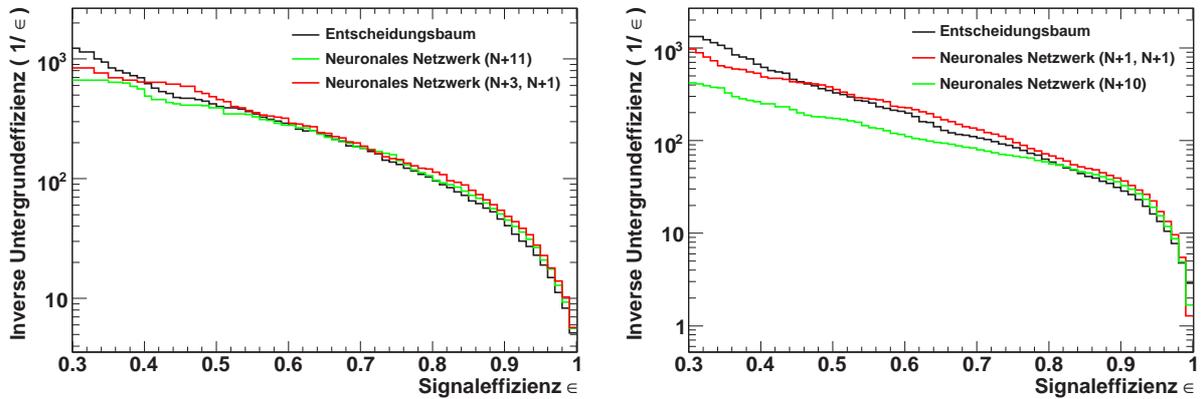


Abbildung 9.2: In dieser Abbildung ist die Leistung der Methoden zur Kombination der Ausgabewerte der besten Elektronenkandidaten eines Jets und dessen Werten für  $p_t^{rel}$  und  $d0$  gezeigt. Links ist diese Klassifikationsleistung für die besten Elektronenkandidaten der Entscheidungsbäume gezeigt, rechts die der besten Kandidaten des Neuronalen Netzwerkes mit zwei versteckten Lagen.

stimmt werden. Mit diesem Zerfallsparameter wurde das Netzwerk 2500 Zyklen lange trainiert.

Für die Verwendung des Ausgabewertes für die Elektronenkandidaten, welcher vom Neuronalen Netzwerk mit zwei versteckten Lagen bestimmt wurde, konnten ebenfalls die zweilagigen Neuronalen Netzwerke ab einer Konfiguration von  $N + 1, N + 1$  die beste Klassifikationsleistung erzielen. Die rechte Seite der Abbildung 9.2 vergleicht dieses Netzwerk mit einem Netzwerk mit einer versteckten Lage und den Entscheidungsbäumen. Mit den beiden letzteren Methoden konnte unabhängig von den Einstellungen die Leistung der zweilagigen Netzwerke nicht erreicht werden. Der Trainingsparameter des Neuronalen Netzwerkes mit zwei versteckten Lagen wurde mit  $\tau = 0.001$  eingestellt und das Netzwerk 2500 Zyklen lange trainiert.

Die neuen Netzwerke, die die Ausgabe für die Jets mit Hilfe der von den neuen Methoden berechneten Ausgabe für die Elektronen bestimmen, wurden ebenfalls in den *SoftElectronTag* integriert. Die Leistung bei der Markierung von b-Quarks, die die neuen Methoden und die Likelihood-Methode des *SoftElectronTag* erreichten, wurde mit dem Standardalgorithmus zur Analyse der Leistung von Algorithmen bei der Markierung von b-Quarks ausgewertet. Dieser Algorithmus trägt den Namen „BTagAna“. Die so erhaltenen Ergebnisse werden im Folgenden für beide der oben genannten Szenarien vorgestellt. Zum Zeitpunkt dieser Analyse lagen nicht genügend Datensätze vor, in denen ein Higgs-Boson in ein  $c\bar{c}$ -Quark-Paar zerfällt. Aus diesem Grund kann im Folgenden nur die Leistung der Methoden bei der Trennung der b-Jets von den Jets der leichten Quarks u, d und s studiert werden. Dies ist allerdings auch das hauptsächliche

Aufgabengebiet der Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen, da durch Zerfälle von c-Hadronen ebenfalls Elektronen entstehen, die schwer von den Elektronen aus b-Hadronen-Zerfällen getrennt werden können.

### 9.2.1 Markierung von b-Quarks ohne Verwendung von $d0$ und $p_t^{rel}$

In diesem Abschnitt werden die Ergebnisse für die Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen präsentiert, ohne den transversalen Stoßparameter der Spuren der Teilchen und den transversalen Impuls der Teilchen im Bezug zur Jet-Achse zu verwenden.

Die Abbildungen 9.3 und 9.4 zeigen die Leistung der neuen Methoden und der Likelihood-Methode bei der Markierung der b-Quarks mit den Datensätzen *CSC11WHbb* und *MC12WHbbfiltered*.

Als Quelle für u-Jets diente der Datensatz *MC12WHuu*. Da der Datensatz *MC12WHbbfiltered* bei der Generierung der Ereignisse künstlich mit b-Quark-Zerfällen angereichert wurde, bei denen niederenergetische Elektronen auftreten, ist die Signal-Effizienz dieses Datensatzes nicht mehr durch das Zerfallsverhältnis  $BR(b \rightarrow e\bar{\nu}c)$  beschränkt. Aus diesem Grund wird für diesen Datensatz stets ein größerer Bereich für die Signaleffizienz abgebildet. An der Abbildung lässt sich erkennen, dass sich die verbesserte Identifikation der niederenergetischen Elektronen direkt auf die Markierung der b-Quarks auswirkt. Dadurch erreichen die Entscheidungsbäume eine wesentlich effizientere Trennung von b-Jets und den leichteren Jets als die beiden anderen Methoden. Diese Verbesserung beträgt in manchen Bereichen der Signaleffizienz über 50%. Die Leistung des Neuronalen Netzwerkes mit zwei versteckten Lagen und die der Likelihood-Methode unterscheiden sich unter Berücksichtigung der Fehler ab einer Signaleffizienz von  $\varepsilon > 0.07$  nicht voneinander. Dies war zu erwarten, da auch die Leistung der beiden Methoden bei der Identifikation der niederenergetischen Elektronen nahezu identisch war.

Auch mit dem Datensatz *CSC11WHbb* zeigen sich große Unterschiede zwischen der Likelihood-Methode und den Entscheidungsbäumen. So kann die inverse Untergrundeffizienz auch mit diesem Datensatz um teilweise mehr als 50% vergrößert werden, was eine direkte Folge der verbesserten Identifikation der Elektronen ist. Deutlich zu erkennen ist auch die erhöhte inverse Untergrundeffizienz im Bereich der Signaleffizienzen  $0.4 < \varepsilon < 0.65$  für die Entscheidungsbäume, welche sich bereits bei der Identifikation der Elektronen im Signalbereich bei Effizienzen größer 0.6 angedeutet hatte. Das Neuronale Netzwerk mit zwei versteckten Lagen ist in kleinen Bereichen der Signaleffizienz ineffizienter als die Likelihood-Methode. Auch dies deutete sich schon in der Leistung bei der Identifikation der niederenergetischen Elektronen an.

Mit dem Datensatz *MC12tt* zeigen die Entscheidungsbäume bis zu einer Signaleffizienz von  $\varepsilon < 0.06$ , das Neuronale Netzwerk bis zu  $\varepsilon < 0.04$  eine verbesserte Leistung gegenüber der Likelihood-Methode bei der Markierung der b-Quarks. Dies ist in Abbildung 9.5 gezeigt. Einen direkten Zusammenhang zur Identifikation der Elektronen herzustellen ist in diesem Fall allerdings nicht ohne weiteres möglich, da für diesen Datensatz, wie zuvor schon erwähnt, die Ereignisse, in denen durch den Zerfall eines t-Quarks ein Elektron entstand, nicht berücksichtigt wurden.

Die genauen Ergebnisse für die Markierung der b-Quarks ohne die Verwendung des transversalen Stoßparameters und des Impulses der Spuren relativ zur Jet-Achse können für die verschiedenen Datensätze den Tabellen 9.1, 9.2 und 9.3 entnommen werden. Die von den Methoden

erzielten Ausgabewerte sind in Anhang B aufgeführt.

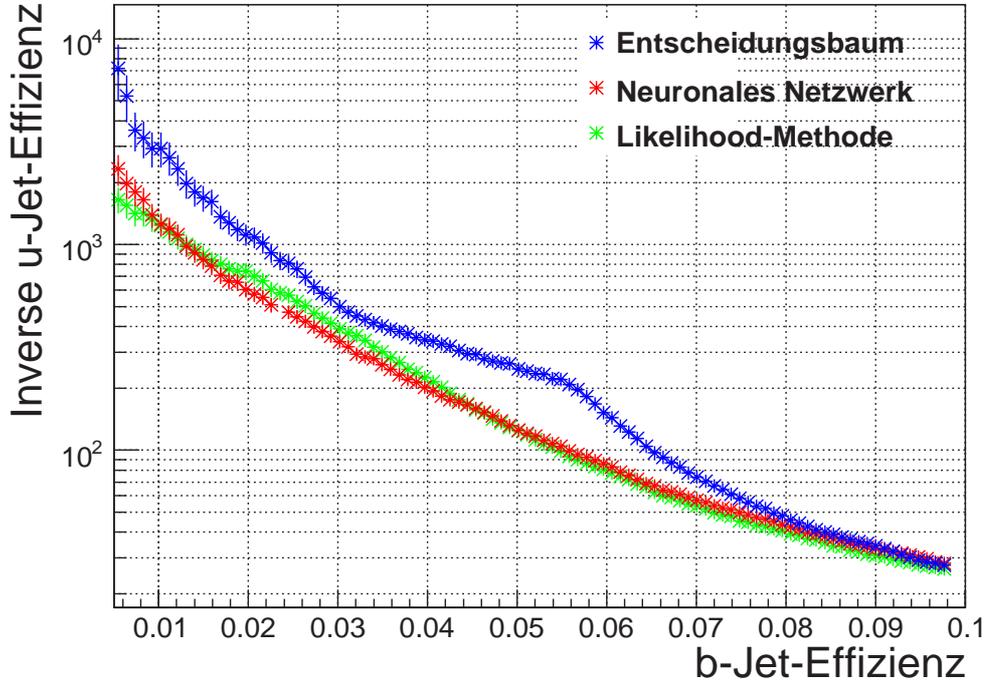


Abbildung 9.3: Gezeigt sind die inversen Untergrundeigenschaften für u-Jets als Funktionen der Identifikationseffizienzen für b-Jets für die Entscheidungsbäume, das Neuronale Netzwerk mit zwei versteckten Lagen und die Likelihood-Methode des *SoftElectronTag* mit den Datensätzen *CSC11WHbb* und *MC12WHuu*.

$1/\epsilon_{u-Jets}$	$\epsilon_{b-Jets} = 6\%$	$\epsilon_{b-Jets} = 4\%$	$\epsilon_{b-Jets} = 2\%$
Entscheidungsbäume	$152.43 \pm 6.66$	$344.64 \pm 22.69$	$1116.45 \pm 132.43$
Neuronales Netzwerk	$85.97 \pm 2.81$	$193.80 \pm 9.55$	$605.09 \pm 52.82$
Likelihood-Methode	$80.63 \pm 2.55$	$225.83 \pm 12.02$	$701.48 \pm 65.94$

Tabelle 9.1: Inverse Untergrundeigenschaften der verschiedenen Methoden für u-Jets bei der Markierung von b-Quarks für die Datensätze *CSC11WHbb* und *MC12WHuu* und verschiedene Signaleffizienzen  $\epsilon_{b-Jets}$ . Die Variablen  $p_t^{rel}$  und  $d_0$  wurden nicht verwendet.

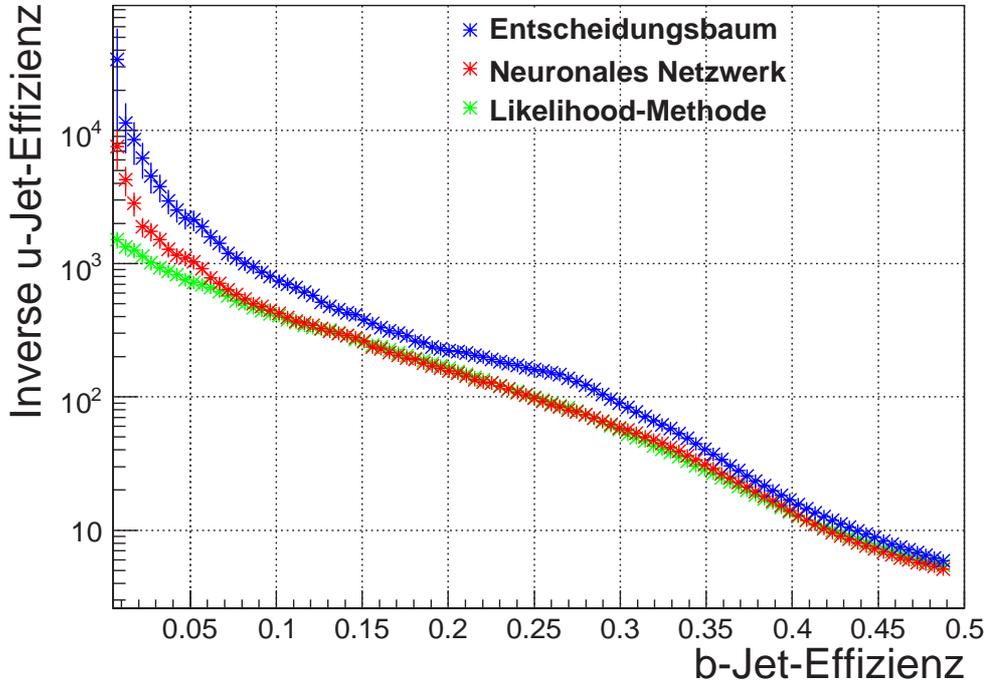


Abbildung 9.4: Hier sind die inversen Untergrundeigenschaften für u-Jets als Funktionen der Identifikationseffizienzen für b-Jets für die Entscheidungsbäume, das Neuronale Netzwerk mit zwei versteckten Schichten und die Likelihood-Methode des *SoftElectronTag* mit dem Datensatz *MC12WHbbfiltered* gezeigt. Als Quelle für u-Jets diente der Datensatz *MC12WHuu*.

$1/\epsilon_{u-Jets}$	$\epsilon_{b-Jets} = 30\%$	$\epsilon_{b-Jets} = 20\%$	$\epsilon_{b-Jets} = 10\%$
Entscheidungsbäume	$89.90 \pm 3.25$	$222.84 \pm 12.73$	$738.77 \pm 76.97$
Neuronales Netzwerk	$58.49 \pm 1.70$	$156.60 \pm 7.49$	$424.79 \pm 33.54$
Likelihood-Methode	$56.92 \pm 1.63$	$164.17 \pm 8.04$	$406.98 \pm 31.45$

Tabelle 9.2: Aufgelistet sind die inversen Untergrundeigenschaften der verschiedenen Methoden für u-Jets bei der Markierung von b-Quarks für den Datensatz *MC12WHbbfiltered* und verschiedene Signaleffizienzen  $\epsilon_{b-Jets}$ . Die Variablen  $p_t^{rel}$  und  $d0$  wurden nicht verwendet. Die u-Jets stammen aus dem Datensatz *MC12WHuu*.

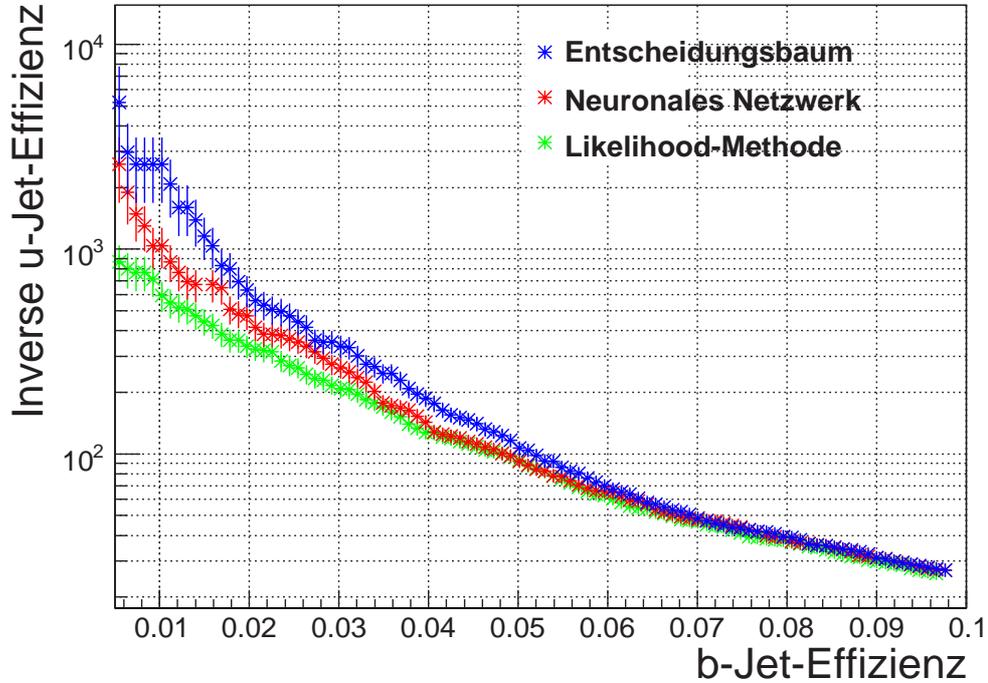


Abbildung 9.5: Gezeigt sind die inversen Untergrundeffizienzen für u-Jets als Funktionen der Identifikationseffizienzen für b-Jets für die Entscheidungsbäume, das Neuronale Netzwerk mit zwei versteckten Lagen und die Likelihood-Methode des *SoftElectronTag* mit dem Datensatz *MC12tt*.

$1/\epsilon_{u-Jets}$	$\epsilon_{b-Jets} = 6\%$	$\epsilon_{b-Jets} = 4\%$	$\epsilon_{b-Jets} = 2\%$
Entscheidungsbäume	$69.99 \pm 4.03$	$187.28 \pm 17.72$	$629.97 \pm 109.57$
Neuronales Netzwerk	$65.58 \pm 3.65$	$142.39 \pm 11.74$	$472.47 \pm 71.15$
Likelihood-Methode	$62.61 \pm 3.40$	$127.54 \pm 9.95$	$335.30 \pm 42.52$

Tabelle 9.3: Inverse Untergrundeffizienzen der verschiedenen Methoden für u-Jets bei der Markierung von b-Quarks für den Datensatz *MC12tt* und verschiedene Signaleffizienzen  $\epsilon_{b-Jets}$ . Die Variablen  $p_t^{rel}$  und  $d0$  wurden nicht verwendet.

### 9.2.2 Markierung von b-Quarks mit der Information über $d0$ und $p_t^{rel}$

Dieser Abschnitt enthält die Ergebnisse für die Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen, wobei die Informationen über den transversalen Stoßparameter der Spuren der Teilchen ( $d0$ ) und deren transversalen Impuls im Bezug zur Jet-Achse ( $p_t^{rel}$ ) verwendet wurden.

Für den Datensatz *CSC11WHbb* zeigt Abbildung 9.6, dass die neuen Methoden in allen für den SoftElectronTag relevanten Bereichen der Signal-Effizienz im Vergleich zur Likelihood-Methode eine deutlich verbesserte Leistung bei der Trennung der b-Jets von u-Jets erreichen. Diese Verbesserung beträgt über den gesamten Bereich der Signaleffizienz über 40%. Besonders das Neuronale Netzwerk, welches die Ausgabe des Neuronalen Netzwerkes zur Markierung der b-Quarks verwendet, kann stark von der Verwendung der Variablen  $d0$  und  $p_t^{rel}$  profitieren. Der Zuwachs an Leistung durch die zusätzlichen Variablen ist in allen Bereichen der Signaleffizienzen bei 50% und mehr. Für das Neuronale Netzwerk, welches die Ausgabe der Entscheidungsbäume verwendet, und für die Likelihood-Methode fällt dieser Leistungszuwachs geringer aus. Dieses Verhalten zeigt sich auch bei den anderen Datensätzen. Für den Datensatz *MC12WHbbfiltered* ist dies in Abbildung 9.7 gezeigt. Für diesen Datensatz erreichen die neuen Methoden ebenfalls eine gesteigerte Leistung bei der Trennung von b-Jets und u-Jets gegenüber der Likelihood-Methode. Ab einer Signaleffizienz von 20% beträgt der Vorsprung der neuen Methoden über 25%. Auch mit dem Datensatz *MC12tt*, für den die Klassifikationsleistungen der Methoden in Abbildung 9.8 gezeigt ist, können die neuen Methoden in allen relevanten Bereichen die Trennung der b-Jets von den Jets der leichteren Quarks deutlich effizienter durchführen als die Likelihood-Methode des SoftElectronTag. Die inversen Untergrundeffizienz für u-Jets vergrößert sich durch die neuen Methoden in manchen Bereichen sogar um einen Faktor  $>2$ . Der Leistungsgewinn durch die Verwendung der zusätzlichen Variablen  $d0$  und  $p_t^{rel}$  betrug für das Neuronale Netzwerk, welches die Ausgabe des Neuronalen Netzwerkes verwendete, teilweise deutlich mehr als 100%. Dadurch konnte sogar die Leistung des Netzwerkes übertroffen werden, welches die Ausgabe der Entscheidungsbäume verwendet. Dieses konnte die Leistung bei der Markierung von b-Quarks gegenüber den Entscheidungsbäumen selbst durch die zusätzlichen Variablen um 50% steigern. Dies macht deutlich, wie wichtig diese Variablen für die Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen ist, will man eine möglichst gute Trennung der b-Jets von den u-Jets erreichen.

Die genauen Werte für die inversen Untergrundeffizienzen für u-Jets, welche durch die Methoden erreicht wurden, sind für einige ausgewählte Signaleffizienzen und alle verwendeten Datensätze in den Tabellen 9.6, 9.7 und 9.8 aufgelistet. Die von den Methoden erzielten Ausgabe-werte können dem Anhang B entnommen werden.

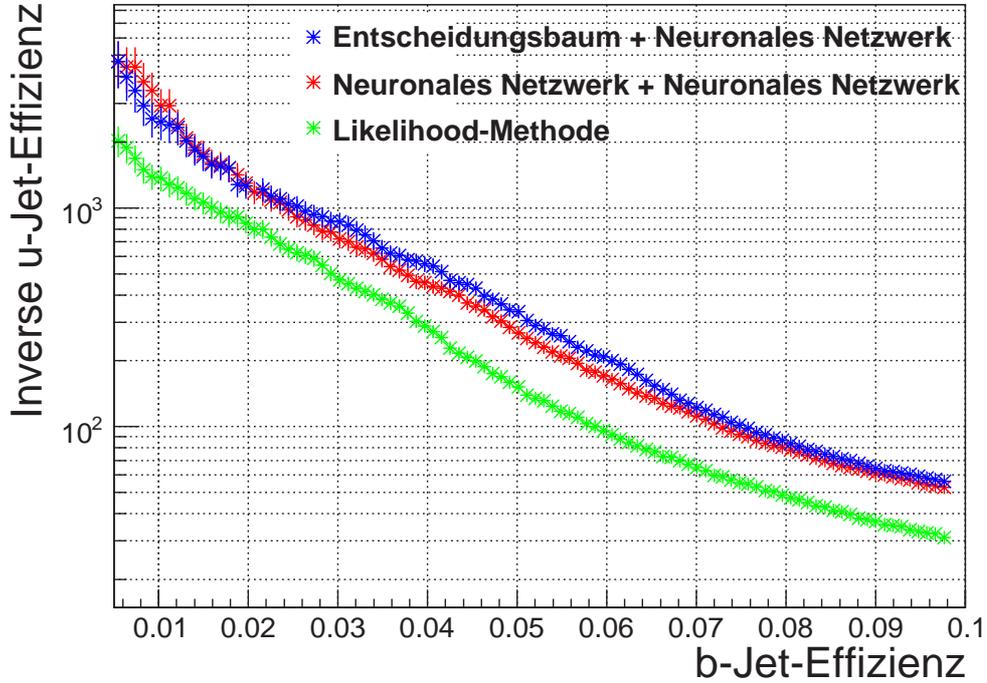


Abbildung 9.6: Diese Abbildung zeigt die inversen Untergrundeﬃzienzen für u-Jets als Funktionen der Identifikationseﬃzienzen für b-Jets für die Entscheidungs bäume, das Neuronale Netzwerk mit zwei versteckten Lagen und die Likelihood-Methode des *SoftElectronTag* mit dem Datensatz *CSC11WHbb*. Hier wurden die Variablen  $d0$  und  $p_t^{rel}$  verwendet. Die u-Jets stammen aus dem Datensatz *MC12WHuu*.

$1/\epsilon_{u-Jets}$	$\epsilon_{b-Jets} = 6\%$	$\epsilon_{b-Jets} = 4\%$	$\epsilon_{b-Jets} = 2\%$
Entscheidungsbäume	$208.05 \pm 10.63$	$554.32 \pm 46.31$	$1258.22 \pm 158.45$
Neuronales Netzwerk	$170.83 \pm 7.90$	$455.56 \pm 34.49$	$1299.48 \pm 166.31$
Likelihood-Methode	$95.61 \pm 3.30$	$288.24 \pm 17.35$	$852.34 \pm 88.33$

Tabelle 9.4: Inverse Untergrundeﬃzienzen der verschiedenen Methoden für u-Jets bei der Markierung von b-Quarks für den Datensatz *CSC11WHbb* und verschiedene Signaleﬃzienzen  $\epsilon_{b-Jets}$ . Als Quelle für u-Jets wurde der Datensatz *MC12WHuu* verwendet.

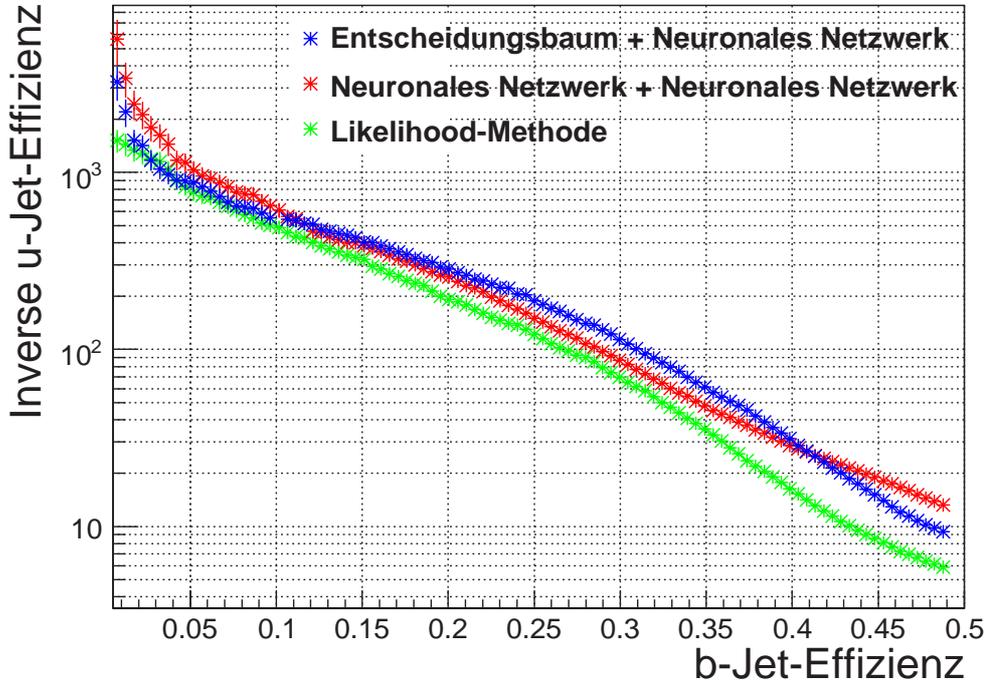


Abbildung 9.7: Gezeigt sind die inversen Untergrundeﬃzienzen für u-Jets als Funktionen der Identifikationseﬃzienzen für b-Jets für die Entscheidungsbäume, das Neuronale Netzwerk mit zwei versteckten Lagen und die Likelihood-Methode des *SoftElectronTag* mit dem Datensatz *MC12WHbbfiltered*. Die Variablen  $d_0$  und  $p_t^{rel}$  wurden hier verwendet. Die u-Jets stammen aus dem Datensatz *MC12WHuu*.

$1/\epsilon_{u-Jets}$	$\epsilon_{b-Jets} = 30\%$	$\epsilon_{b-Jets} = 20\%$	$\epsilon_{b-Jets} = 10\%$
Entscheidungsbäume	$113.84 \pm 4.63$	$284.38 \pm 18.36$	$543.73 \pm 48.58$
Neuronales Netzwerk	$86.91 \pm 3.09$	$256.47 \pm 15.72$	$612.31 \pm 58.07$
Likelihood-Methode	$69.35 \pm 2.19$	$191.99 \pm 10.17$	$488.97 \pm 41.43$

Tabelle 9.5: Tabelle der inversen Untergrundeﬃzienzen der verschiedenen Methoden für u-Jets bei der Markierung von b-Quarks für den Datensatz *MC12WHbbfiltered* und verschiedene Signaleﬃzienzen  $\epsilon_{b-Jets}$ . Die u-Jets stammen aus dem Datensatz *MC12WHuu*.

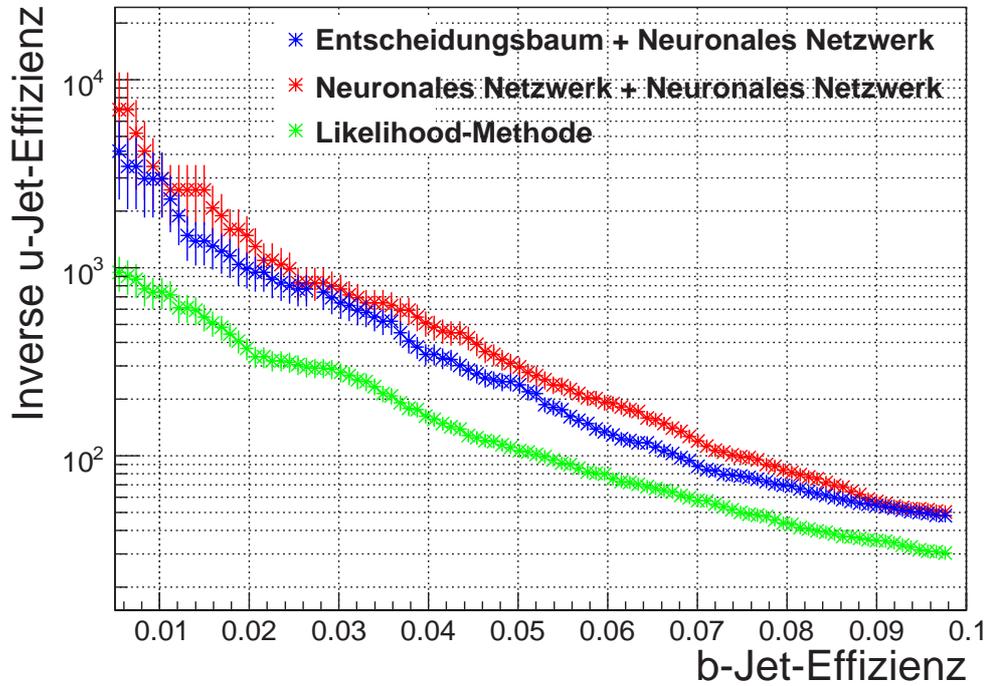


Abbildung 9.8: Gezeigt sind die inversen Untergrundeigenschaften für u-Jets als Funktionen der Identifikationseffizienzen für b-Jets für die Entscheidungsbäume, das Neuronale Netzwerk mit zwei versteckten Schichten und die Likelihood-Methode des *SoftElectronTag* mit dem Datensatz *MC12tt*. Hier wurden die Variablen  $d0$  und  $p_t^{rel}$  verwendet.

$1/\epsilon_{u-Jets}$	$\epsilon_{b-Jets} = 6\%$	$\epsilon_{b-Jets} = 4\%$	$\epsilon_{b-Jets} = 2\%$
Entscheidungsbäume	$134.12 \pm 10.73$	$346.48 \pm 44.66$	$989.95 \pm 215.91$
Neuronales Netzwerk	$192.35 \pm 18.46$	$506.68 \pm 79.05$	$1483.86 \pm 396.44$
Likelihood-Methode	$79.95 \pm 4.92$	$162.41 \pm 14.31$	$371.23 \pm 49.54$

Tabelle 9.6: Inverse Untergrundeigenschaften der verschiedenen Methoden für u-Jets bei der Markierung von b-Quarks für den Datensatz *MC12tt* und verschiedene Signaleffizienzen  $\epsilon_{b-Jets}$ .

### 9.3 Abhängigkeit der Methoden vom transversalen Impuls der Jets

Zusätzlich zur allgemeinen Leistung der Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen wurde auch die Abhängigkeit der Leistungsfähigkeit dieser Methoden von den transversalen Impulsen der Jets qualitativ betrachtet. Hierzu diente der Datensatz *CSC11WHbb* als Quelle für b-Jets, der Datensatz *MC12WHuu* als Quelle für u-Jets. Der Zweck dieser Betrachtungen war ein Vergleich der Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen und der Markierung durch die Methoden, die auf der Lebenszeit der b-Hadronen basieren. Als Vertreter dieser auf der Lebenszeit basierenden Methoden wurde der sogenannte „Combined Tag“ gewählt. Dieser Combined Tag kombiniert die für einen Jet berechneten b-Jet-Wahrscheinlichkeiten der Methoden „IP3D“ und „SV1“ und wird als Standardmethode zur Markierung von b-Quarks verwendet.

Die Methode IP3D berechnet den transversalen und longitudinalen Stoßparameter der Spuren eines Jets. Da diese direkt mit der Lebensdauer der b-Hadronen zusammenhängen, kann mit diesen die Wahrscheinlichkeit bestimmt werden, dass der Jet ein b-Quark-Jet ist. Der Algorithmus SV1 versucht, mit den im Spurdetektor rekonstruierten Spuren einen möglichen Zerfallsvertex des b-Quarks, den Sekundärvertex, zu rekonstruieren. Dies ist nur für Teilchen möglich, welche eine hohe Lebensdauer besitzen. Wird ein Sekundärvertex gefunden, so kann der dazugehörige Jet ein b-Jet sein und mit sehr hoher Wahrscheinlichkeit kein Jet der leichten Quarks.

Für die auf der Lebenszeit der b-Hadronen basierenden Methoden zeigt Abbildung 9.9, dass die Leistung bei niedrigen Jet-Impulsen sehr gering ist. Dies liegt daran, dass diese Methoden sehr stark auf die Auflösung bei der Rekonstruktion der Spuren angewiesen sind. Durch den im Mittel niedrigen transversalen Impuls der Spuren in Jets mit insgesamt niedrigen transversalen Impulsen, wird durch Vielfachstreuung diese Auflösung stark beeinträchtigt, wodurch auch die Leistung bei der Markierung von b-Quarks negativ betroffen ist. Für größere transversale Impulse der Jets steigt die Leistung an, bis bei einem transversalen Jet-Impuls von etwa 100 GeV/c ein Maximum erreicht wird. Danach wird die Leistung wieder schlechter, da durch den starken Boost der Jets diese stärker kollimiert sind und dadurch viele Spuren sehr nahe beieinander liegen. Die Folge ist, dass Treffer im Spurendetektor mehreren Spuren zugeordnet werden können („shared hits“) und somit die Auflösung bei der Rekonstruktion der Spuren schlechter wird.

Für die Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen kann diese Abhängigkeit von den transversalen Impulsen der Jets anders sein, da die Identifikation der Elektronen hauptsächlich auf der Information des elektromagnetischen Kalorimeters basiert. Für die Markierung der b-Quarks ohne Verwendung des transversalen Stoßparameters und des transversalen Impulses relativ zur Jet-Achse wird die Abhängigkeit der Leistung vom transversalen Impuls der Jets beispielhaft mit dem Neuronalen Netzwerk mit zwei versteckten Lagen präsentiert. Das Ergebnis zeigt Abbildung 9.10. Die Leistung bei der Trennung von b-Jets und u-Jets ist für kleine transversale Jet-Impuls am besten und fällt dann stetig ab. Der Grund hierfür ist, dass mit zunehmender Energie die Jets stärker kollimiert sind. Dies kann dazu führen, dass

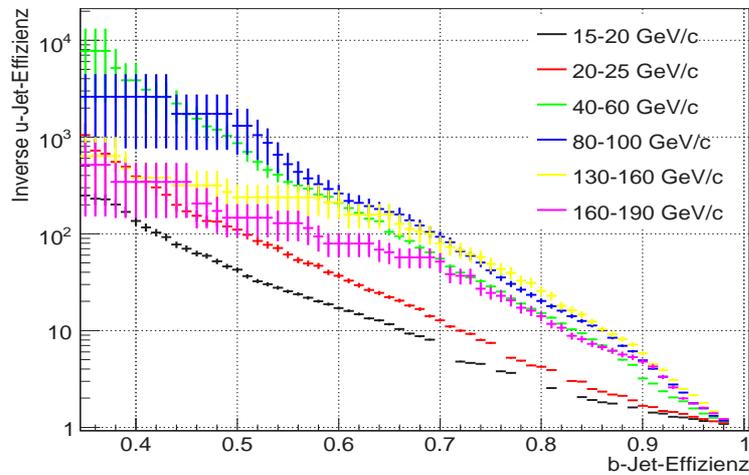


Abbildung 9.9: Gezeigt ist die Leistung der auf der Lebenszeit der b-Hadronen basierenden Methoden bei der Trennung von b-Jets und u-Jets in Abhängigkeit vom transversalen Impuls der Jets. Die b- und u-Jets stammen aus den Datensätzen *CSC11WHbb* und *MC12WHuu*.

sich die Schauer der Teilchen im Jet öfter überlappen, wodurch die Identifikation der niederenergetischen Elektronen, die ja auf der Information über diese Schauer beruht, ineffizienter wird. Durch die gute Leistung bei kleinen transversalen Jet-Impulsen kann eine Kombination mit den Methoden, die auf der Lebenszeit der b-Hadronen basieren, in diesem Bereich besonders effektiv sein.

Die Studie über die Abhängigkeit der Leistung bei der Markierung von b-Quarks mit den niederenergetischen Elektronen, wobei der transversale Stoßparameter und der transversale Impuls der Elektronenkandidaten zur Identifikation dieser verwendet wird, soll hier beispielhaft mit dem Neuronalen Netzwerk, welches die Ausgabe des Neuronalen Netzwerkes zur Identifikation der niederenergetischen Elektronen verwendet, präsentiert werden. In Abbildung 9.11 ist deutlich zu erkennen, dass die Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen nahezu unabhängig vom transversalen Impuls der Jets ist. Dies folgt aus der Kombination von zwei Effekten. Zum einen ist, wie oben beschrieben, die Auflösung des transversalen Stoßparameters für kleine transversale Impulse schlechter. In diesem Bereich ist die Identifikation der niederenergetischen Elektronen jedoch gut, da die Schauer der Teilchen eines Jets im Kalorimeter selten überlappen. Für größere transversale Impulse wird die Identifikation der niederenergetischen Elektronen aufgrund der oben genannten Gründe zwar schlechter, jedoch wird dies durch die in diesem Bereich bessere Auflösung des transversalen Stoßparameters wieder ausgeglichen.

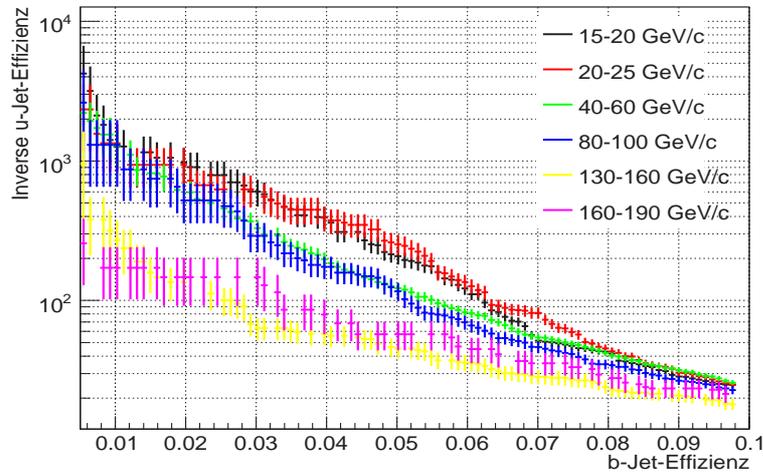


Abbildung 9.10: Gezeigt ist die Leistung des Neuronalen Netzwerkes mit zwei versteckten Lagen bei der Trennung von  $b$ -Jets und  $u$ -Jets in Abhängigkeit vom transversalen Impuls der Jets. Als Quelle für  $b$ -Jets diente der Datensatz *CSC11WHbb*, die  $u$ -Jets stammen aus dem Datensatz *MC12WHuu*.

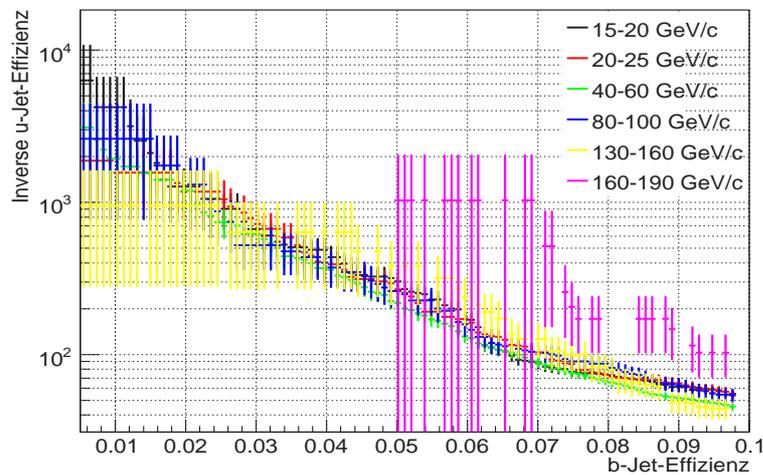


Abbildung 9.11: Gezeigt ist die Leistung des Neuronales Netzwerkes, welches die Ausgabe des Neuronales Netzwerkes und die Variablen  $d_0$  und  $p_t^{rel}$  zur Markierung von  $b$ -Quarks verwendet, bei der Trennung von  $b$ -Jets und  $u$ -Jets in Abhängigkeit vom transversalen Impuls der Jets. Verwendet wurden die Datensätze *CSC11WHbb* und *MC12WHuu*.

## 9.4 Bestimmung der linearen Korrelation zwischen dem SoftElectronTag und dem Combined Tag

Um die Korrelation zwischen der Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen und den Methoden zu bestimmen, die auf der Lebensdauer der b-Hadronen basieren, wurden die innerhalb des ATHENA-Frameworks bestimmten Ausgabewerte der verschiedenen Methoden für Jets mit Hilfe des Standard-Algorithmus zur Auswertung der Leistung von Methoden bei der Markierung von b-Quarks (BTagAna) gespeichert. Diese Methoden waren das Neuronale Netzwerk mit zwei versteckten Lagen, die Entscheidungsbäume, die Neuronale Netzwerke zur Kombination der Ausgabewerte dieser beiden Methoden mit den Variablen  $d0$  und  $p_t^{rel}$ , die Likelihood-Methode des SoftElectronTag und der Combined Tag, welcher im vorherigen Kapitel vorgestellt wurde.

Für die Korrelationsbestimmung zwischen der Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen und der Markierung durch den Combined Tag wurden die von den verschiedenen Methoden bestimmten Ausgabewerten von 50000 Jets aus dem Datensatz *CSC11WHbb* verwendet, welche mindestens einen Elektronenkandidaten beinhalteten, der die Qualitätsschnitte auf die Spuren passierte. Die Werte der linearen Korrelationskoeffizienten sind in den Tabellen 9.7 und 9.8 aufgelistet. Die Berechnung dieser Korrelationskoeffizienten wird in Anhang C beschrieben.

Anhand der Tabelle 9.7 ist ersichtlich, dass die Korrelation zwischen dem SoftElectronTag mit den verschiedenen Methoden und dem Combined Tag sehr gering ist, falls auf die Verwendung der Variablen  $p_t^{rel}$  und  $d0$  verzichtet wird. Diese geringen Korrelationen zeigen, dass eine gute Unabhängigkeit der Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen von der Markierung mit Hilfe der Information über die Lebensdauer der b-Hadronen gewährleistet ist. Dadurch kann im späteren Experiment eine gegenseitige Kalibration vorgenommen oder die Ergebnisse der komplementären Methoden kombiniert werden.

Für den Fall, dass die oben genannten Variablen verwendet werden, um die beste Trennung von b-Jets von den Jets der leichteren Quarks zu erhalten, vergrößern sich die linearen Korrelationskoeffizienten. Dies wird aus Tabelle 9.8 ersichtlich und war zu erwarten, da die auf der Lebensdauer der b-Hadronen basierenden Methoden die Informationen über den transversalen Stoßparameter ebenfalls verwenden.

Methode (ohne $d0$ und $p_t^{rel}$ )	$r$
Likelihood-Methode des SoftElectronTag	0.149
Neuronales Netzwerk mit zwei versteckten Lagen	0.097
Entscheidungsbäume	0.105

Tabelle 9.7: Liste der linearen Korrelationskoeffizienten  $r$  zwischen den Verteilungen der Ausgabewerte des Combined Tag und der Ausgabewerte der verschiedenen Methoden der Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen. Die Variablen  $p_t^{rel}$  und  $d0$  wurden hier nicht verwendet.

Methode (mit $d0$ und $p_t^{rel}$ )	$r$
Likelihood-Methode des SoftElectronTag	0.165
Neuronales Netzwerk mit der Information des Netzwerkes	0.268
Neuronales Netzwerk mit der Information der Entscheidungsbäume	0.192

Tabelle 9.8: Die linearen Korrelationskoeffizienten  $r$  zwischen den Verteilungen der Ausgabewerte des Combined Tag und der Ausgabewerte der verschiedenen Methoden der Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen. Für die Erstellung dieser Tabelle wurden die Informationen über  $p_t^{rel}$  und  $d0$  von den neuen Methoden verwendet.

# 10 Zusammenfassung

Ein Ziel des ATLAS-Experimentes ist es, die Existenz des im Standardmodell vorhergesagten Higgs-Bosons nachzuweisen. Dieser Nachweis kann schwierig sein, sollte das Higgs-Boson in nicht nachweisbare Teilchen zerfallen. Ein Modell, in dem ein solcher Zerfall erlaubt ist, ist das sogenannte Phion-Modell. Diese Arbeit befasste sich damit, welche Konsequenzen der Zerfall des Higgs-Bosons in nicht nachweisbare Phionen des Phion-Modells für das Entdeckungspotential des Higgs-Bosons hat.

Weiterhin befasste sich diese Arbeit mit der Identifikation von niederenergetischen Elektronen aus dem Zerfall von b-Hadronen beim ATLAS-Experiment. Es wurde untersucht, ob die aktuell verwendete Identifikation dieser Elektronen durch neue Methoden verbessert werden kann. Die Identifikation der Elektronen aus dem Zerfall von b-Hadronen wurde dann dazu verwendet, die Jets der b-Quarks von den Jets der leichteren Quarks zu unterscheiden.

Im Phion-Modell wird die Physik des Higgs-Bosons durch die zusätzlichen Zerfallsmöglichkeiten in nicht nachweisbare Phionen sehr stark beeinflusst. Insbesondere kann sich die Breite des Higgs-Bosons in diesem Modell stark vergrößern. In dieser Arbeit wurde untersucht, welche Auswirkung die vergrößerte Breite des Higgs-Bosons auf den Nachweis eines in nicht nachweisbare Teilchen zerfallenden Higgs-Bosons beim ATLAS-Experiment hat. Aufgrund dieses Zerfalls in nicht nachweisbare Teilchen, kann keine Massenresonanz gemessen werden, sondern es muss ein Zählexperiment durchgeführt werden: Es wird nach einem statistisch signifikanten Überschuss an Ereignissen über dem Untergrund des Standardmodells gesucht. Wird ein solcher gefunden, so wird dieser Überschuss als Signal interpretiert. Falls kein signifikanter Überschuss an Ereignissen gemessen werden kann, so können dennoch gewisse Bereiche der Masse des Higgs-Bosons oder seiner Zerfallsrate in nicht nachweisbare Phionen ausgeschlossen werden. Die Grenzen zu diesen Bereichen werden Ausschlussgrenzen genannt. Eine Verschiebung dieser Ausschlussgrenzen durch die zusätzlichen Zerfallsmöglichkeiten des Higgs-Bosons im Phion-Modell kann zum Einen durch eine Veränderung in den Wirkungsquerschnitten der Produktionsprozesse herbeigeführt werden. Zum Anderen kann sich die Kinematik eines Prozesses derart ändern, dass die Schnitte zur Selektion der Signalereignisse effizienter oder ineffizienter werden. Da keine Berechnung von Wirkungsquerschnitten im Phion-Modell mit zusätzlicher Berücksichtigung von Interferenzen mit Standardmodell-Untergründen vorlag, wurde im Rahmen dieser Arbeit nur der Einfluss der Kinematik auf die Schnittakzeptanz untersucht. Die Wirkungsquerschnitte der Prozesse als Funktion der Breite wurden konstant als die Wirkungsquerschnitte im Standardmodell gewählt.

Die Analyse zeigte, dass sich die Akzeptanzen und damit die Ausschlussgrenzen bei der mit einem Z-Boson assoziierten Higgs-Produktion erst für sehr große Breiten leicht verschlechtern, da dort die Schnittakzeptanzen mit großen Breiten des Higgs-Bosons kleiner werden. Für das hier

betrachtete Phion-Modell folgt daraus, dass sich die Ausschlussgrenzen im assoziierten Higgs-Produktionskanal für ein leichtes Higgs-Boson mit einer Masse von  $m_{Higgs} \leq 160 \text{ GeV}/c^2$  bis zu einer Kopplung des Higgs-Bosons an die nicht nachweisbaren Phionen von  $\omega = 1$  nicht signifikant ändert. Für Massen des Higgs-Bosons bis zu  $m_{Higgs} \leq 250 \text{ GeV}/c^2$  werden erst ab einer Kopplung von  $\omega = 3$  Änderungen in den Ausschlussgrenzen sichtbar, für größere Massen zeigten sich auch bei Kopplungen  $\omega > 3$  keine Veränderungen in den Ausschlussgrenzen mehr.

Die Ausschlussgrenzen für den Vektorboson-Fusionskanal verändern sich nur minimal und nur für sehr große Breiten des Higgs-Bosons, da die Schnittakzeptanzen auch für große Breiten nahezu konstant sind. Insgesamt konnte gezeigt werden, dass sich die Ausschlussgrenzen für moderate Änderungen der Breite des Higgs-Bosons kaum ändern. Moderat bedeutet hier, dass die Breite des Higgs-Bosons kleiner sein soll als seine Masse. Für das Higgs-Boson im Phion-Modell bedeutet dies, dass sich die Ausschlussgrenzen für alle betrachteten Massen erst für Kopplungen  $\omega > 5$  verschlechtern.

Weiterhin wurde in dieser Arbeit die Identifikation von niederenergetischen Elektronen aus den Zerfällen von b-Hadronen untersucht und diese auf die Markierung von b-Quark-Jets angewandt. Dabei wurden die Variablen der Elektronenkandidaten, welche in den Standardalgorithmen zur Identifikation der Elektronen verwendet werden, teilweise ersetzt und durch weitere Variablen ergänzt. Mit diesen Variablen wurden neue Methoden bezüglich ihrer Leistung bei der Identifikation der niederenergetischen Elektronen überprüft. Diese neuen Methoden waren die Neuronale Netzwerke mit einer, zwei und drei versteckten Lagen, sowie die Entscheidungsbäume. Besonders das Neuronale Netzwerk mit zwei versteckten Lagen und die Entscheidungsbäume erwiesen sich als sehr effizient bei der Identifikation der Signal-Elektronen. Daher wurden diese beiden Methoden in die offizielle ATLAS-Software ATHENA integriert und auf die Identifikation der niederenergetischen Elektronen aus dem b-Quark-Zerfall angewandt. Im Vergleich zur bisher verwendeten Likelihood-Methode konnten durch die neuen Methoden die Kombination aus Effizienz  $\varepsilon_{Sig}$  bei der Erkennung der niederenergetischen Elektronen aus dem b-Quark-Zerfall und Abweisung  $1/\varepsilon_{Bkg}$  von Teilchen, die keine Signal-Elektronen sind, verbessert werden. Je nach verwendetem Datensatz und Bereich der Signaleffizienz betrug der Vergrößerungsfaktor der inversen Untergrundeffizienz etwa zwei.

Diese beiden Methoden zur Identifikation der niederenergetischen Elektronen wurden weiterhin auf die Markierung von b-Quarks angewandt. Dies geschah für zwei Szenarien.

Im ersten Szenario wird die Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen dazu verwendet, die auf der Lebensdauer der b-Hadronen basierenden Methoden zu kalibrieren oder beide Methoden zu kombinieren. Grundlage hierfür ist eine geringe Korrelation zwischen den beiden Methoden. Im Rahmen dieser Arbeit wurde der lineare Korrelationskoeffizient zwischen den beiden Methoden bestimmt und es konnte gezeigt werden, dass dieser unter 15% liegt. Die Größe dieses Korrelationskoeffizienten rechtfertigt eine Kombination der beiden Ansätze oder eine Kalibration der auf der Lebensdauer der b-Hadronen basierenden Methoden während des laufenden Experimentes durch die Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen. Die Leistung der neuen Methoden bei der Markierung von b-Quarks konnte die bisher verwendete Likelihood-Methode für die hier verwendeten Datensätze um Faktoren von teilweise über zwei übertreffen oder war mindestens gleich gut.

---

Im zweiten Szenario soll durch die Markierung der b-Quarks mit Hilfe der niederenergetischen Elektronen selbst eine möglichst gute Trennung der b-Jets von den Jets der leichteren Quarks erreicht werden. Dies wird erreicht, indem weitere Variablen der niederenergetischen Elektronen zur Markierung der b-Quarks herangezogen werden. Um diese für die Markierung der b-Quarks spezifischen Variablen zusammen mit den Ausgabewerten der neuen Methoden zur Identifikation der niederenergetischen Elektronen auswerten zu können, wurde für jede der neuen Methoden ein weiteres Neuronales Netzwerk trainiert. Durch dieses Verfahren wurde die Kombination aus Effizienz  $\varepsilon_{Sig}$  bei der Markierung von b-Quarks und Abweisung  $1/\varepsilon_{Bkg}$  von Jets, die keine b-Jets waren, im Vergleich zur bisher verwendeten Likelihood-Methode des SoftElectronTag zum Teil um einen Faktor  $\simeq 2$  verbessert. Durch die Verwendung der zusätzlichen Variablen geht jedoch ein erheblicher Teil der Unabhängigkeit von den auf der Lebensdauer der b-Quarks basierenden Methoden verloren. So steigen die linearen Korrelationskoeffizienten zwischen dem SoftElectronTag und den neuen Methoden und dem Combined Tag auf bis zu etwa 27% an.

Insgesamt konnte sowohl die Identifikation der niederenergetischen Elektronen, als auch die Markierung von b-Quarks beim ATLAS-Experiment für viele der hier verwendeten Datensätze und große Bereiche der jeweiligen Signaleffizienzen verbessert werden. Die in dieser Arbeit verwendeten Methoden können in Zukunft dazu verwendet werden, in sowohl simulierten Daten, als auch in echten Daten des ATLAS-Experiments die niederenergetischen Elektronen aus dem Zerfall von b-Hadronen zu identifizieren und mit dieser Identifikation die Markierung von b-Quarks durchzuführen. Auch die oben genannte Kombination mit oder die Kalibration der Methoden zur Markierung von b-Quarks, welche auf der Lebensdauer des b-Quarks basieren, kann mit den hier verwendeten Methoden vor und während des Experimentes durchgeführt werden.



# A Tabellen mit den Signifikanzen und Ausschlussgrenzen im Phion-Modell

Im folgenden werden die erwartete Anzahl an Signalereignissen, die Signifikanzen und die Ausschlussgrenzen auf  $\xi^2$  entsprechend eines Vertrauensniveaus von 95% für die in dieser Analyse betrachteten Kanäle und ein Higgs-Boson im Phion-Modell mit verschiedenen Kopplungen  $\omega$  aufgelistet. Für alle Zahlen wurde für das Higgs-Boson eine Zerfallsrate in unsichtbare Phionen von  $BR(H \rightarrow inv.) = 100\%$  angenommen. Die Anzahl der Untergrundereignisse für beide Kanäle verändern sich durch das Phion-Modell nicht und können daher den entsprechenden Tabellen aus Kapitel 7 mit den Werten für das Standardmodell entnommen werden. Weiterhin wurde für alle Tabellen die systematischen Fehler  $\alpha$  auf die Untergründe berücksichtigt. Die Werte für  $\alpha$  können dem Text in Kapitel 7 entnommen werden.

## A.1 Die Tabellen für den assoziierten Higgs-Produktionskanal (ZH):

Tabelle A.1:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.1$  und  $\int L dt = 30 \text{ fb}^{-1}$  (ZH).

$m_H$ [ $\text{GeV}/c^2$ ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	122.12	107.46	91.02	76.12	66.37	60.72	28.48	14.82	8.32	3.09
Signifikanz $S$	3.37	2.97	2.51	2.10	1.83	1.67	0.78	0.40	0.23	0.08
$\xi^2_{95\% CL}$	0.48	0.55	0.65	0.78	0.89	0.98	2.09	4.02	7.16	19.29

Tabelle A.2:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.1$  und  $\int L dt = 100 \text{ fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	407.06	358.18	303.39	253.73	221.23	202.41	94.92	49.41	27.74	10.31
Signifikanz $S$	5.66	4.98	4.22	3.53	3.08	2.81	1.32	0.69	0.39	0.14
$\xi^2_{95\% CL}$	0.29	0.33	0.39	0.47	0.54	0.59	1.25	2.40	4.28	11.52

Tabelle A.3:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.2$  und  $\int L dt = 30 \text{ fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	121.27	104.04	89.05	75.10	66.42	59.98	28.19	14.82	8.30	3.11
Signifikanz $S$	3.35	2.87	2.46	2.07	1.83	1.65	0.77	0.40	0.22	0.08
$\xi^2_{95\% CL}$	0.49	0.57	0.67	0.79	0.90	0.99	2.12	4.03	7.19	19.21

Tabelle A.4:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.2$  und  $\int L dt = 100 \text{ fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	404.24	346.81	296.83	250.35	221.39	199.95	93.96	49.39	27.65	10.35
Signifikanz $S$	5.62	4.82	4.13	3.48	3.08	2.78	1.31	0.69	0.38	0.14
$\xi^2_{95\% CL}$	0.29	0.34	0.40	0.47	0.54	0.59	1.26	2.40	4.29	11.46

Tabelle A.5:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 1$  und  $\int L dt = 30 \text{ fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	120.32	105.13	90.76	76.07	65.57	61.04	28.71	14.67	8.34	3.07
Signifikanz $S$	3.32	2.90	2.51	2.10	1.81	1.68	0.79	0.40	0.23	0.08
$\xi^2_{95\% CL}$	0.50	0.57	0.66	0.78	0.91	0.98	2.08	4.07	7.15	19.41

Tabelle A.6:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 1$  und  $\int L dt = 100 \text{ fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	401.05	350.45	302.55	253.56	218.56	203.47	95.70	48.91	27.80	10.24
Signifikanz $S$	5.58	4.87	4.21	3.53	3.04	2.83	1.33	0.68	0.39	0.14
$\xi^2_{95\% CL}$	0.30	0.34	0.39	0.47	0.54	0.58	1.24	2.43	4.27	11.59

Tabelle A.7:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 3$  und  $\int L dt = 30\text{fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	105.47	91.82	78.63	67.16	59.72	56.36	26.67	14.82	8.33	3.07
Signifikanz $S$	2.92	2.53	2.17	1.85	1.65	1.55	0.73	0.40	0.23	0.08
$\xi^2_{95\%CL}$	0.57	0.65	0.76	0.89	1.00	1.06	2.24	4.02	7.16	19.41

Tabelle A.8:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 3$  und  $\int L dt = 100\text{fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	351.58	306.06	262.10	223.87	199.08	187.86	88.90	49.41	27.77	10.25
Signifikanz $S$	4.89	4.26	3.64	3.11	2.77	2.61	1.24	0.69	0.39	0.14
$\xi^2_{95\%CL}$	0.34	0.39	0.45	0.53	0.60	0.63	1.33	2.40	4.27	11.58

Tabelle A.9:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 5$  und  $\int L dt = 30\text{fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	79.78	67.99	60.69	54.34	48.03	46.79	23.75	14.57	7.96	3.04
Signifikanz $S$	2.20	1.88	1.67	1.50	1.32	1.29	0.65	0.40	0.22	0.08
$\xi^2_{95\%CL}$	0.75	0.88	0.98	1.10	1.24	1.27	2.51	4.10	7.50	19.62

Tabelle A.10:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 5$  und  $\int L dt = 100\text{fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	265.92	226.63	202.32	181.13	160.11	155.97	79.15	48.55	26.52	10.13
Signifikanz $S$	3.7	3.15	2.81	2.52	2.23	2.17	1.10	0.67	0.37	0.14
$\xi^2_{95\%CL}$	0.45	0.52	0.59	0.66	0.74	0.76	1.50	2.44	4.47	11.71

Tabelle A.11:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\Gamma_{Higgs} = m_{Higgs}$  und  $\int L dt = 30\text{fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	89.15	71.11	58.94	50.64	44.27	39.86	18.31	13.40	7.29	2.76
Signifikanz $S$	2.46	1.96	1.63	1.40	1.22	1.10	0.50	0.37	0.20	0.07
$\xi^2_{95\%CL}$	0.67	0.84	1.01	1.18	1.35	1.50	3.26	4.45	8.19	21.65

Tabelle A.12:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\Gamma_{Higgs} = m_{Higgs}$  und  $\int L dt = 100 \text{fb}^{-1}$  (ZH).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	297.16	237.02	196.47	168.79	147.55	132.85	61.02	44.67	24.30	9.19
Signifikanz $S$	4.13	3.30	2.73	2.35	2.05	1.85	0.85	0.62	0.34	0.13
$\xi^2_{95\%CL}$	0.40	0.50	0.60	0.70	0.80	0.89	1.95	2.66	4.88	12.92

## A.2 Die Tabellen für die Vektorboson-Fusion (VBF):

Tabelle A.13:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.1$  und  $\int L dt = 10 \text{ fb}^{-1}$  (VBF).

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	609.63	590.09	567.84	571.46	538.66	547.05	524.30	419.89	367.17	271.76
Signifikanz $S$	5.77	5.58	5.37	5.40	5.09	5.17	4.96	3.97	3.47	2.57
$\xi^2_{95\% CL}$	0.29	0.30	0.31	0.31	0.32	0.32	0.33	0.42	0.48	0.64

Tabelle A.14:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.1$  und  $\int L dt = 30 \text{ fb}^{-1}$  (VBF).

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	1828.89	1770.26	1703.53	1714.39	1615.98	1641.15	1572.89	1259.66	1101.52	815.28
Signifikanz $S$	8.21	7.94	7.64	7.69	7.25	7.36	7.06	5.65	4.94	3.66
$\xi^2_{95\% CL}$	0.20	0.21	0.22	0.21	0.23	0.22	0.23	0.29	0.33	0.45

Tabelle A.15:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.2$  und  $\int L dt = 10 \text{ fb}^{-1}$  (VBF).

$m_H$ [GeV/c <sup>2</sup> ]	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	614.79	619.17	574.66	534.42	513.01	527.24	484.34	410.86	364.24	271.76
Signifikanz $S$	5.81	5.86	5.43	5.05	4.85	4.99	4.58	3.89	3.44	2.57
$\xi^2_{95\% CL}$	0.28	0.28	0.30	0.33	0.34	0.33	0.36	0.42	0.48	0.64

Tabelle A.16:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 0.2$  und  $\int L dt = 30 \text{ fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	1844.38	1857.50	1723.98	1603.26	1539.03	1581.73	1453.02	1232.59	1092.72	815.28
Signifikanz $S$	8.28	8.34	7.74	7.19	6.91	7.10	6.52	5.53	4.90	3.66
$\xi_{95\% CL}^2$	0.20	0.20	0.21	0.23	0.24	0.23	0.25	0.30	0.34	0.45

Tabelle A.17:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 1$  und  $\int L dt = 10 \text{ fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	618.81	615.47	568.33	533.52	553.80	546.26	488.61	419.89	370.80	266.11
Signifikanz $S$	5.85	5.82	5.37	5.05	5.24	5.17	4.62	3.97	3.51	2.52
$\xi_{95\% CL}^2$	0.28	0.28	0.31	0.33	0.32	0.32	0.36	0.42	0.47	0.66

Tabelle A.18:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 1$  und  $\int L dt = 30 \text{ fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	1856.42	1846.40	1704.99	1600.55	1661.40	1638.77	1465.83	1259.66	1112.40	798.32
Signifikanz $S$	8.33	8.29	7.65	7.18	7.46	7.35	6.58	5.65	4.99	3.58
$\xi_{95\% CL}^2$	0.20	0.20	0.22	0.23	0.22	0.22	0.25	0.29	0.33	0.46

Tabelle A.19:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 3$  und  $\int L dt = 10 \text{ fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	689.92	654.59	621.41	626.58	541.60	570.42	514.23	433.43	383.23	271.55
Signifikanz $S$	6.52	6.19	5.88	5.93	5.12	5.39	4.86	4.10	3.62	2.57
$\xi_{95\% CL}^2$	0.25	0.27	0.28	0.28	0.32	0.31	0.34	0.40	0.46	0.64

Tabelle A.20:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 3$  und  $\int L dt = 30 \text{ fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	2069.76	1963.78	1864.24	1879.73	1624.81	1711.26	1542.69	1300.28	1149.68	814.66
Signifikanz $S$	9.29	8.81	8.37	8.43	7.29	7.68	6.92	5.83	5.16	3.66
$\xi_{95\% CL}^2$	0.18	0.19	0.20	0.20	0.23	0.21	0.24	0.28	0.32	0.45

Tabelle A.21:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 5$  und  $\int L dt = 10\text{fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	637.16	633.97	585.86	578.69	562.21	556.16	520.64	442.45	377.70	275.74
Signifikanz $S$	6.03	6.00	5.54	5.47	5.32	5.26	4.92	4.18	3.57	2.61
$\xi^2_{95\%CL}$	0.27	0.28	0.30	0.30	0.31	0.31	0.34	0.39	0.46	0.63

Tabelle A.22:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\omega = 5$  und  $\int L dt = 30\text{fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	1911.48	1901.91	1757.58	1736.08	1686.63	1668.48	1561.91	1327.35	1133.11	827.22
Signifikanz $S$	8.58	8.53	7.89	7.79	7.57	7.49	7.01	5.96	5.08	3.71
$\xi^2_{95\%CL}$	0.19	0.19	0.21	0.21	0.22	0.22	0.24	0.28	0.32	0.44

Tabelle A.23:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\Gamma_{Higgs} = m_{Higgs}$  und  $\int L dt = 10\text{fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	682.47	649.31	594.14	602.63	554.22	545.46	495.32	452.15	365.79	253.33
Signifikanz $S$	6.45	6.14	5.62	5.70	5.24	5.16	4.68	4.28	3.46	2.40
$\xi^2_{95\%CL}$	0.26	0.27	0.29	0.29	0.31	0.32	0.35	0.39	0.48	0.69

Tabelle A.24:  $N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von  $\xi^2$  zu 95% Vertrauensniveau für  $\Gamma_{Higgs} = m_{Higgs}$  und  $\int L dt = 30\text{fb}^{-1}$  (VBF).

$m_H [GeV/c^2]$	110	120	130	140	150	160	200	250	300	400
$N_{Sig.}$	2047.40	1947.92	1782.42	1807.90	1662.66	1636.39	1485.96	1356.46	1097.38	759.99
Signifikanz $S$	9.19	8.74	8.00	8.11	7.46	7.34	6.67	6.09	4.92	3.41
$\xi^2_{95\%CL}$	0.18	0.19	0.21	0.20	0.22	0.22	0.25	0.27	0.34	0.48



## B Ausgabewerte der Methoden bei der Markierung von b-Quarks

Im Folgenden werden die von den verschiedenen Methoden erzielten Ausgabewerte bei der Markierung von b-Quarks mit Hilfe der niederenergetischen Elektronen präsentiert.

**Ausgabewerte für den Datensatz *CSC11WHbb* ohne  $d_0$  und  $p_t^{rel}$**

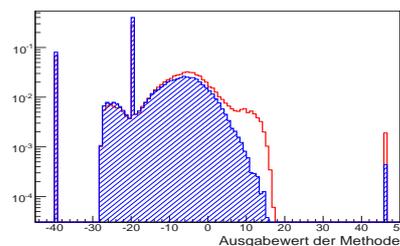


Abbildung B.1: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *CSC11WHbb*.

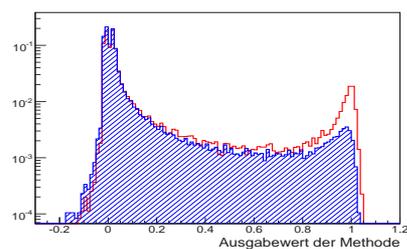


Abbildung B.2: Ausgabewerte des Neuronales Netzwerkes mit zwei versteckten Lagen bei der Markierung der b-Quarks für Signal (rot) und Untergrund (blau) für den Datensatz *CSC11WHbb*.

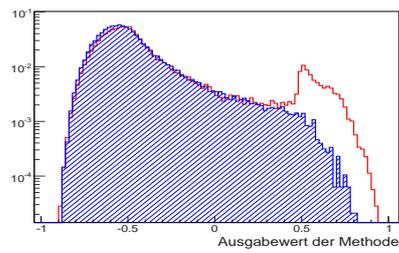


Abbildung B.3: Ausgabewerte der Entscheidungsbäume für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *CSC11WHbb*.

## Ausgabewerte für den Datensatz *MC12WHbbfiltered* ohne $d0$ und $p_t^{rel}$

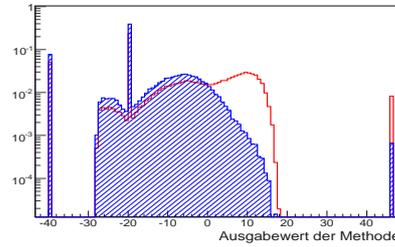


Abbildung B.4: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12WHbbfiltered*.

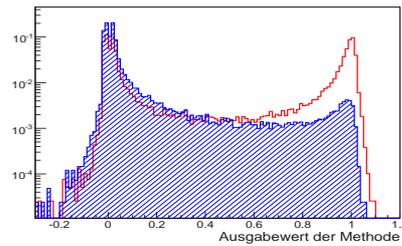


Abbildung B.5: Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen bei der Markierung der b-Quarks für Signal (rot) und Untergrund (blau) für den Datensatz *MC12WHbbfiltered*.

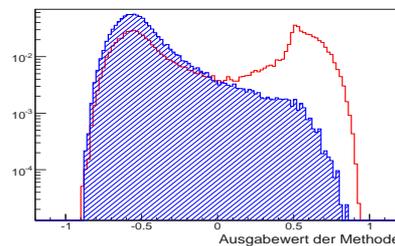


Abbildung B.6: Ausgabewerte der Entscheidungsbäume für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12WHbbfiltered*.

## Ausgabewerte für den Datensatz *MC12tt* ohne $d0$ und $p_t^{rel}$

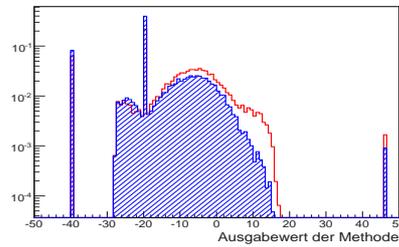


Abbildung B.7: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12tt*.

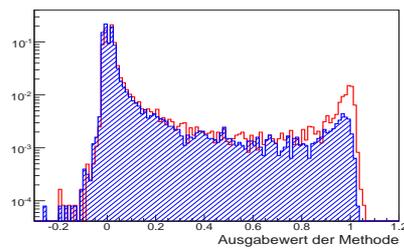


Abbildung B.8: Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen bei der Markierung der b-Quarks für Signal (rot) und Untergrund (blau) für den Datensatz *MC12tt*.

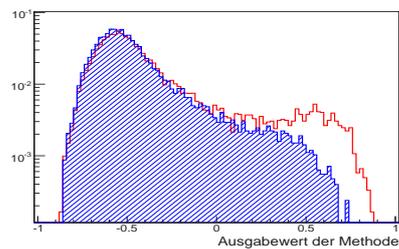


Abbildung B.9: Ausgabewerte der Entscheidungsbäume für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12tt*.

## Ausgabewerte für den Datensatz *CSC11WHbb* mit $d_0$ und $p_t^{rel}$

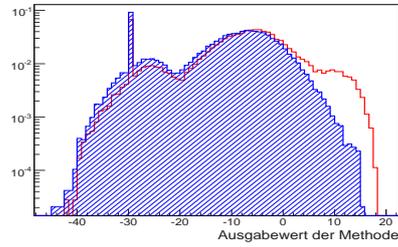


Abbildung B.10: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *CSC11WHbb*.

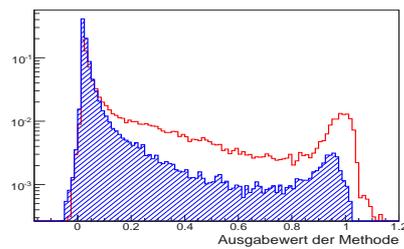


Abbildung B.11: Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Neuronalen Netzwerkes verwendet, für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *CSC11WHbb*.

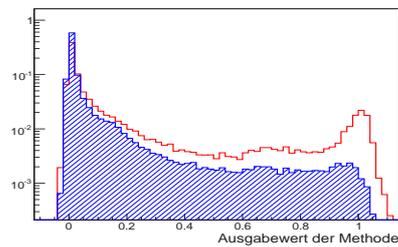


Abbildung B.12: Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Entscheidungsbäume verwendet, für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *CSC11WHbb*.

Ausgabewerte für den Datensatz *MC12filtered* mit  $d_0$  und  $p_t^{rel}$

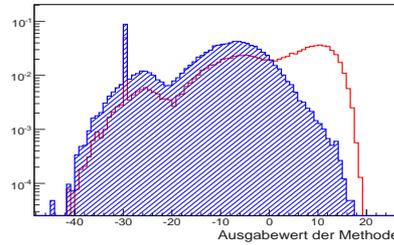


Abbildung B.13: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12filtered*.

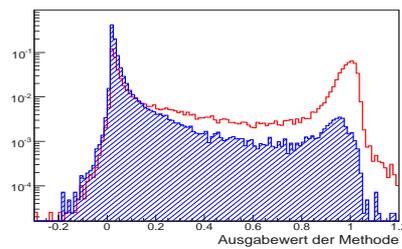


Abbildung B.14: Ausgabewerte des Neuronales Netzwerkes, welches die Ausgabe der Neuronales Netzwerkes verwendet, für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12filtered*.

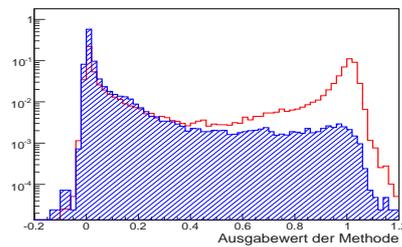


Abbildung B.15: Ausgabewerte des Neuronales Netzwerkes, welches die Ausgabe der Entscheidungsbäume verwendet, für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12filtered*.

## Ausgabewerte für den Datensatz *MC12tt* mit $d_0$ und $p_t^{rel}$

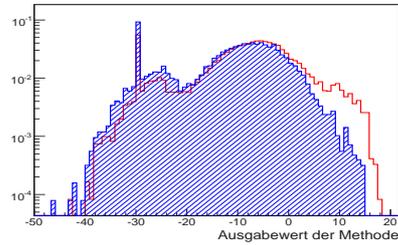


Abbildung B.16: Ausgabewerte der Likelihood-Methode für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12tt*.

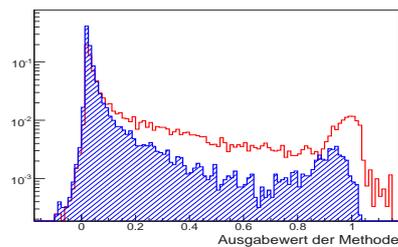


Abbildung B.17: Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Neuronalen Netzwerkes verwendet, für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12tt*.

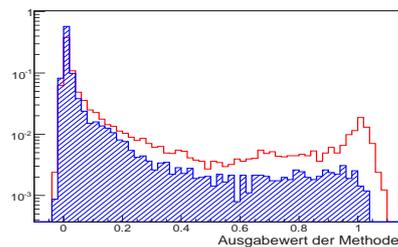


Abbildung B.18: Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Entscheidungsbäume verwendet, für Signal (rot) und Untergrund (blau) bei der Markierung der b-Quarks für den Datensatz *MC12tt*.



## C Bestimmung des linearen Korrelationskoeffizienten zweier Verteilungen

Im Folgenden soll kurz erläutert werden, wie die lineare Korrelation zweier Verteilungen bestimmt werden kann. Eine detailliertere Beschreibung findet man zum Beispiel in den Referenzen [55, 56].

Der lineare Korrelationskoeffizient  $\rho_{X,Y}$  zwischen zwei Verteilungen  $X$  und  $Y$  mit den Erwartungswerten  $\mu_X = \langle X \rangle$  und  $\mu_Y = \langle Y \rangle$  und den Standardabweichungen  $\sigma_X$  und  $\sigma_Y$  ist definiert als

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{\langle X - \mu_X \rangle \cdot \langle Y - \mu_Y \rangle}{\sigma_X \cdot \sigma_Y}. \quad (\text{C.1})$$

Bei einer Serie von  $n$  Messungen, das heisst einer Stichprobe vom Umfang  $n$ , von  $X$  und  $Y$  am selben System mit den Messwerten  $x_i$  und  $y_i$  ( $i=1,2,\dots,n$ ) lässt sich der empirische Korrelationskoeffizient  $r_{x,y}$  berechnen, mit dem die Korrelation der Verteilungen  $X$  und  $Y$  abgeschätzt werden kann:

$$r_{x,y} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (\text{C.2})$$

Hierbei sind die empirischen Erwartungswerte der Stichproben von  $X$  und  $Y$  definiert als

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{und} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (\text{C.3})$$

In Gleichung C.2 wurde berücksichtigt, dass die Standardabweichungen der Verteilungen durch die Varianzen der Stichproben ersetzt werden müssen. Für die Varianz  $s$  einer Stichprobe des Umfangs  $n$  gilt nämlich:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (\text{C.4})$$



# Tabellenverzeichnis

2.1	Die Fermionenfelder des Standardmodells und ihre Quantenzahlen. . . . .	11
4.1	Granularität des elektromagnetischen Kalorimeters (ECAL) und des hadronischen Kalorimeters (HCAL) . . . . .	31
7.1	Wirkungsquerschnitte der Higgs-Produktion im Standardmodell für die betrachteten Massen des Higgs-Bosons. . . . .	57
7.2	Schnittakzeptanzen für den ZH-Kanal. . . . .	61
7.3	Wirkungsquerschnitte der wichtigsten Untergründe für den ZH-Kanal . . . . .	61
7.4	Schnittakzeptanzen für den VBF-Kanal . . . . .	65
7.5	Wirkungsquerschnitte der wichtigsten Untergründe für den VBF-Kanal . . . . .	65
7.6	Anzahl der Signal- und Untergrundereignisse, die Signifikanzen und Ausschlussgrenzen für $\xi^2$ zu 95% Vertrauensniveau für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons im Standardmodell für den ZH-Kanal. . . . .	70
7.7	Anzahl der Signal- und Untergrundereignisse, die Signifikanzen und Ausschlussgrenzen für $\xi^2$ zu 95% Vertrauensniveau für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons im Standardmodell für den VBF-Kanal. . . . .	71
7.8	Anzahl der Signal- und Untergrundereignisse, die Signifikanzen und Ausschlussgrenzen für $\xi^2$ zu 95% Vertrauensniveau für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons im Phion-Modell mit $\omega = 3$ für den ZH-Kanal. . . . .	73
7.9	Anzahl der Signal- und Untergrundereignisse, die Signifikanzen und Ausschlussgrenzen für $\xi^2$ zu 95% Vertrauensniveau für die Entdeckung eines unsichtbar zerfallenden Higgs-Bosons im Phion-Modell mit $\omega = 3$ für den VBF-Kanal. . . . .	74
8.1	Trennungsfähigkeit $S$ der Variablen zur Identifikation der Elektronen . . . . .	89
8.2	Die inversen Untergrundeffizienzen bei der Identifikation der Elektronen für den Datensatz <i>MC12WHbbfiltered</i> und verschiedene Signaleffizienzen. . . . .	105
8.3	Die inversen Untergrundeffizienzen bei der Identifikation der Elektronen für den Datensatz <i>CSC11WHbb</i> und verschiedene Signaleffizienzen. . . . .	107
8.4	Die inversen Untergrundeffizienzen bei der Identifikation der Elektronen für den Datensatz <i>MC12tt</i> und verschiedene Signaleffizienzen. . . . .	109
9.1	Die inversen Untergrundeffizienzen für u-Jets bei der Markierung von b-Quarks für den Datensatz <i>CSC11WHbb</i> und verschiedene Signaleffizienzen (ohne $p_t^{rel}$ und $d0$ ). . . . .	117

9.2	Die inversen Untergrundeﬃzienzen fur u-Jets bei der Markierung von b-Quarks fur den Datensatz <i>MC12WHbbfiltered</i> und verschiedene Signaleﬃzienzen (ohne $p_t^{rel}$ und $d0$ ). . . . .	118
9.3	Tabelle der inversen Untergrundeﬃzienzen fur u-Jets bei der Markierung von b-Quarks fur den Datensatz <i>MC12tt</i> und verschiedene Signaleﬃzienzen (ohne $p_t^{rel}$ und $d0$ ). . . . .	119
9.4	Tabelle der inversen Untergrundeﬃzienzen fur u-Jets bei der Markierung von b-Quarks fur den Datensatz <i>CSC11WHbb</i> und verschiedene Signaleﬃzienzen (mit $p_t^{rel}$ und $d0$ ). . . . .	121
9.5	Tabelle der inversen Untergrundeﬃzienzen fur u-Jets bei der Markierung von b-Quarks fur den Datensatz <i>MC12WHbbfiltered</i> und verschiedene Signaleﬃzienzen (mit $p_t^{rel}$ und $d0$ ). . . . .	122
9.6	Tabelle der inversen Untergrundeﬃzienzen fur u-Jets bei der Markierung von b-Quarks fur den Datensatz <i>MC12tt</i> und verschiedene Signaleﬃzienzen. . . . .	123
9.7	Lineare Korrelationskoeffizienten der Verteilungen der Ausgabewerte des SoftElectronTags und des Combined Tag ohne Verwendung von $p_t^{rel}$ und $d0$ . . . . .	128
9.8	Lineare Korrelationskoeffizienten der Verteilungen der Ausgabewerte des SoftElectronTags bei Verwendung von $p_t^{rel}$ und $d0$ und des Combined Tag. . . . .	128
A.1	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 0.1$ und $\int L dt = 30 \text{ fb}^{-1}$ (ZH). . . . .	133
A.2	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 0.1$ und $\int L dt = 100 \text{ fb}^{-1}$ (ZH). . . . .	134
A.3	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 0.2$ und $\int L dt = 30 \text{ fb}^{-1}$ (ZH). . . . .	134
A.4	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 0.2$ und $\int L dt = 100 \text{ fb}^{-1}$ (ZH). . . . .	134
A.5	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 1$ und $\int L dt = 30 \text{ fb}^{-1}$ (ZH). . . . .	134
A.6	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 1$ und $\int L dt = 100 \text{ fb}^{-1}$ (ZH). . . . .	134
A.7	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 3$ und $\int L dt = 30 \text{ fb}^{-1}$ (ZH). . . . .	135
A.8	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 3$ und $\int L dt = 100 \text{ fb}^{-1}$ (ZH). . . . .	135
A.9	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 5$ und $\int L dt = 30 \text{ fb}^{-1}$ (ZH). . . . .	135
A.10	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\omega = 5$ und $\int L dt = 100 \text{ fb}^{-1}$ (ZH). . . . .	135
A.11	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\Gamma_{Higgs} = m_{Higgs}$ und $\int L dt = 30 \text{ fb}^{-1}$ (ZH). . . . .	135
A.12	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau fur $\Gamma_{Higgs} = m_{Higgs}$ und $\int L dt = 100 \text{ fb}^{-1}$ (ZH). . . . .	136

A.13	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 0.1$ und $\int L dt = 10 \text{fb}^{-1}$ (VBF).	137
A.14	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 0.1$ und $\int L dt = 30 \text{fb}^{-1}$ (VBF).	137
A.15	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 0.2$ und $\int L dt = 10 \text{fb}^{-1}$ (VBF).	137
A.16	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 0.2$ und $\int L dt = 30 \text{fb}^{-1}$ (VBF).	138
A.17	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 1$ und $\int L dt = 10 \text{fb}^{-1}$ (VBF).	138
A.18	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 1$ und $\int L dt = 30 \text{fb}^{-1}$ (VBF).	138
A.19	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 3$ und $\int L dt = 10 \text{fb}^{-1}$ (VBF).	138
A.20	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 3$ und $\int L dt = 30 \text{fb}^{-1}$ (VBF).	138
A.21	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 5$ und $\int L dt = 10 \text{fb}^{-1}$ (VBF).	139
A.22	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\omega = 5$ und $\int L dt = 30 \text{fb}^{-1}$ (VBF).	139
A.23	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\Gamma_{Higgs} = m_{Higgs}$ und $\int L dt = 10 \text{fb}^{-1}$ (VBF).	139
A.24	$N_{Sig.}$ , Signifikanzen und Ausschlussgrenzen von $\xi^2$ zu 95% Vertrauensniveau für $\Gamma_{Higgs} = m_{Higgs}$ und $\int L dt = 30 \text{fb}^{-1}$ (VBF).	139



# Abbildungsverzeichnis

2.1	Die Teilchen des Standardmodells . . . . .	6
2.2	Das Higgs-Potential . . . . .	9
2.3	Feynman-Graph für einen Prozess mit Quark-Dichtefunktionen und Fragmentationsfunktionen . . . . .	15
2.4	Produktion und Zerfall des Higgs-Bosons . . . . .	16
2.5	Feynman-Graphen zur assoziierten $t\bar{t}$ -Higgs-Produktion (links) und zur Gluon-Fusion (rechts). . . . .	17
2.6	Feynman-Graphen zur Vektorboson-Fusion (links) und zur assoziierten Higgs-Produktion mit einem Z-Boson (rechts). . . . .	17
2.7	Zerfallskette eines b-Mesons. . . . .	18
3.1	Breite des Higgs-Bosons im Phionmodell als Funktion der Kopplung $\omega$ . . . . .	23
4.1	Schematische Abbildung des LHC . . . . .	26
4.2	Der ATLAS-Detektor . . . . .	28
4.3	Der Spurdetektor von ATLAS . . . . .	29
4.4	Schematische Abbildung des elektromagnetischen Kalorimeters . . . . .	30
6.1	Ein Perzeptron . . . . .	45
6.2	Die Sigmoid-Funktion für verschiedene Werte des Parameters $k$ . . . . .	46
6.3	Graphische Darstellung der Berechnung der Ausgabe eines Knotens eines Neuronalen Netzwerkes. . . . .	47
6.4	Beispiel für einen Entscheidungsbaum . . . . .	50
7.1	Aktuell generierte Massen des Higgs-Bosons für den assoziierten Produktionskanal . . . . .	57
7.2	Aktuell generierte Massen des Higgs-Bosons für die Vektorboson-Fusion . . . . .	57
7.3	Die Anzahl der generierten Ereignisse über dem Quadrat der aktuell generierten Masse des Higgs-Bosons für den assoziierten Produktionskanal des Higgs-Bosons. . . . .	58
7.4	Totale Schnittakzeptanzen für den ZH-Kanal . . . . .	62
7.5	$p_{i}^{miss}$ und die für jedes Ereignis aktuell generierte Masse des Z-Bosons für den ZH-Kanal . . . . .	63
7.6	Normierte Verteilungen des fehlenden transversalen Impulses für ein Higgs-Boson der Masse $m_{Higgs} = 130 \text{ GeV}/c^2$ im Standardmodell und im Phion-Modell mit $\omega = 3$ für den assoziierten Produktionskanal. . . . .	63

7.7	Akzeptanz der Schnitte und die $p_t^{miss}$ -Verteilung als $F(m_{Higgs})$ für den VBF-Kanal	67
7.8	Normierte Verteilungen des fehlenden transversalen Impulses für ein Higgs-Boson der Masse $m_{Higgs} = 130 \text{ GeV}/c^2$ im Standardmodell und im Phion-Modell mit $\omega = 3$ für die Vektorboson-Fusion.	67
7.9	Ausschlussgrenzen für $\xi^2$ im Standardmodell entsprechend eines Vertrauensniveaus von 95% für den ZH-Kanal.	70
7.10	Ausschlussgrenzen für $\xi^2$ im Standardmodell entsprechend eines Vertrauensniveaus von 95% für den VBF-Kanal.	71
7.11	Schnittakzeptanz als $F(\Gamma_{Higgs})$ für den ZH-Kanal und den VBF-Kanal	72
7.12	Ausschlussgrenzen für $\xi^2$ im Phion-Modell zu 95% Vertrauensniveau bei verschiedenen Kollungen $\omega$ (ZH-Kanal)	75
7.13	Ausschlussgrenzen für $\xi^2$ im Phion-Modell zu 95% Vertrauensniveau bei verschiedenen Kollungen $\omega$ (VBF)	76
8.1	Graphische Darstellung der Bestimmung des transversalen Stoßparameters.	80
8.2	Die Verteilungen der verschiedenen Variablen der Elektronen-Kandidaten für Signal-Elektronen und Untergrundteilchen.	84
8.3	Die Verteilungen der zusätzlichen Variablen der Elektronen-Kandidaten.	87
8.4	Vergleich der zwei für das Training der Methoden verwendeten Datensätze mit Hilfe mehrerer Variablen.	88
8.5	Lineare Korrelationsmatrizen der Variablen für die Identifikation der Elektronen in Prozent	89
8.6	Vergleich zwischen verschiedenen Einstellungen der Likelihood-Methode.	92
8.7	Leistung eines Neuronalen Netzwerkes mit einer versteckten Lage mit verschiedenen Zerfallsraten und die Werte der Fehlerfunktion während eines Trainings für eines dieser Netzwerke.	94
8.8	Verschiedene Konfigurationen des Neuronalen Netzwerkes mit einer versteckten Lage und ihre Leistung.	95
8.9	Leistung eines Neuronalen Netzwerkes mit zwei versteckten Lagen mit verschiedenen Zerfallsraten.	97
8.10	Vergleich der Werte der Fehlerfunktion eines Neuronalen Netzes mit zwei versteckten Lagen mit unterschiedlichen Werten für $\tau$	97
8.11	Verschiedene Konfigurationen des Neuronalen Netzwerkes mit zwei versteckten Lagen und ihre Leistung.	98
8.12	Konvergenz des Neuronalen Netzwerkes ( $N, N - 1$ ).	98
8.13	Einfluß der Anzahl der Bäume auf die Leistung der Entscheidungsbäume	100
8.14	Einfluss der Beschneidungsmethode und des Trennungskriteriums auf die Leistung der Entscheidungsbäume, die mit der Boost-Methode trainiert wurden.	100
8.15	Einfluss der Beschneidungsstärke auf Beschneidungsmethoden der BDT-Methode.	101
8.16	Vergleich der verschiedenen Methoden außerhalb des ATHENA-Frameworks.	102
8.17	Vergleich der Identifikation der niederenergetischen Elektronen in ATHENA mit den verschiedenen Methoden für den Datensatz <i>MC12WHbbfiltered</i> .	105

8.18	Ausgabewerte der Likelihood-Methode bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>MC12WHbbfiltered</i> . . . . .	106
8.19	Ausgabewerte des Neuronalen bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>MC12WHbbfiltered</i> . . . . .	106
8.20	Ausgabewerte der Entscheidungsbäume bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>MC12WHbbfiltered</i> . . . . .	106
8.21	Vergleich der Identifikation der niederenergetischen Elektronen in ATHENA mit den verschiedenen Methoden für den Datensatz <i>CSC11WHbb</i> . . . . .	107
8.22	Ausgabewerte der Likelihood-Methode bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>CSC11WHbb</i> . . . . .	108
8.23	Ausgabewerte des Neuronalen Netzwerkes bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>CSC11WHbb</i> . . . . .	108
8.24	Ausgabewerte der Entscheidungsbäume bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>CSC11WHbb</i> . . . . .	108
8.25	Vergleich der Identifikation der niederenergetischen Elektronen in ATHENA mit den verschiedenen Methoden für den Datensatz <i>MC12tt</i> . . . . .	109
8.26	Ausgabewerte der Likelihood-Methode bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>MC12tt</i> . . . . .	110
8.27	Ausgabewerte des Neuronalen Netzwerkes bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>MC12tt</i> . . . . .	110
8.28	Ausgabewerte der Entscheidungsbäume bei der Identifikation der niederenergetischen Elektronen für den Datensatz <i>MC12tt</i> . . . . .	110
9.1	Der transversale Stoßparameter (links) und der transversale Impuls relativ zur Jetachse (rechts) für Signal- und Untergrund-Jets. . . . .	112
9.2	Leistung der Methoden zur Kombination der Ausgabewerte des besten Elektronenkandidaten eines Jets und dessen Werten für $p_t^{rel}$ und $d0$ . . . . .	114
9.3	Leistung der neuen Methoden bei der Markierung von b-Quarks im Vergleich zur Likelihood-Methode des <i>SoftElectronTag</i> mit dem Datensatz <i>CSC11WHbb</i> (ohne $p_t^{rel}$ und $d0$ ). . . . .	117
9.4	Leistung der neuen Methoden bei der Markierung von b-Quarks im Vergleich zur Likelihood-Methode des <i>SoftElectronTag</i> mit dem Datensatz <i>MC12WHbbfiltered</i> (ohne $p_t^{rel}$ und $d0$ ). . . . .	118
9.5	Leistung der neuen Methoden bei der Markierung von b-Quarks im Vergleich zur Likelihood-Methode des <i>SoftElectronTag</i> mit dem Datensatz <i>MC12tt</i> (ohne $p_t^{rel}$ und $d0$ ). . . . .	119
9.6	Leistung der Methoden bei der Markierung von b-Quarks mit dem Datensatz <i>CSC11WHbb</i> , wobei die Variablen $d0$ und $p_t^{rel}$ verwendet wurden. . . . .	121
9.7	Leistung der Methoden bei der Markierung von b-Quarks mit dem Datensatz <i>MC12WHbbfiltered</i> , wobei die Variablen $d0$ und $p_t^{rel}$ verwendet wurden. . . . .	122
9.8	Leistung der Methoden bei der Markierung von b-Quarks mit dem Datensatz <i>MC12tt</i> , wobei die Variablen $d0$ und $p_t^{rel}$ verwendet wurden. . . . .	123

9.9	Leistung der auf der Lebenszeit der b-Hadronen basierenden Methoden bei der Markierung von b-Quarks mit dem Datensatz <i>CSC11WHbb</i> in Abhängigkeit des transversalen Impulses der Jets. . . . .	125
9.10	Leistung des Neuronalen Netzwerkes bei der Markierung von b-Quarks mit dem Datensatz <i>CSC11WHbb</i> in Abhängigkeit des transversalen Impulses der Jets. . . . .	126
9.11	Leistung des Neuronalen Netzwerkes, welches die Ausgabe des Neuronalen Netzwerkes verwendet, bei der Markierung von b-Quarks mit dem Datensatz <i>CSC11WHbb</i> in Abhängigkeit des transversalen Impulses der Jets. . . . .	126
B.1	Ausgabewerte der Likelihood-Methode bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>CSC11WHbb</i> . . . . .	141
B.2	Ausgabewerte des Neuronalen Netzwerkes bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>CSC11WHbb</i> . . . . .	141
B.3	Ausgabewerte der Entscheidungsbäume bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>CSC11WHbb</i> . . . . .	142
B.4	Ausgabewerte der Likelihood-Methode bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12WHbbfiltered</i> . . . . .	143
B.5	Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12WHbbfiltered</i> . . . . .	143
B.6	Ausgabewerte der Entscheidungsbäume bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12WHbbfiltered</i> . . . . .	143
B.7	Ausgabewerte der Likelihood-Methode bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12tt</i> . . . . .	144
B.8	Ausgabewerte des Neuronalen Netzwerkes mit zwei versteckten Lagen bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12tt</i> . . . . .	144
B.9	Ausgabewerte der Entscheidungsbäume bei der Markierung der b-Quarks ohne $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12tt</i> . . . . .	144
B.10	Ausgabewerte der Likelihood-Methode bei der Markierung der b-Quarks mit $d0$ und $p_t^{rel}$ für den Datensatz <i>CSC11WHbb</i> . . . . .	145
B.11	Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Neuronalen Netzwerkes verwendet, bei der Markierung der b-Quarks für den Datensatz <i>CSC11WHbb</i> . . . . .	145
B.12	Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Entscheidungsbäume verwendet, bei der Markierung der b-Quarks für den Datensatz <i>CSC11WHbb</i> . . . . .	145
B.13	Ausgabewerte der Likelihood-Methode bei der Markierung der b-Quarks mit $d0$ und $p_t^{rel}$ für den Datensatz <i>MC12filtered</i> . . . . .	146
B.14	Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Neuronalen Netzwerkes verwendet, bei der Markierung der b-Quarks für den Datensatz <i>MC12filtered</i> . . . . .	146
B.15	Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Entscheidungsbäume verwendet, bei der Markierung der b-Quarks für den Datensatz <i>MC12filtered</i> . . . . .	146

B.16	Ausgabewerte der Likelihood-Methode bei der Markierung der b-Quarks mit $d_0$ und $p_i^{rel}$ für den Datensatz <i>MC12tt</i> . . . . .	147
B.17	Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Neuronalen Netzwerkes verwendet, bei der Markierung der b-Quarks für den Datensatz <i>MC12tt</i> . . . . .	147
B.18	Ausgabewerte des Neuronalen Netzwerkes, welches die Ausgabe der Entscheidungsbäume verwendet, bei der Markierung der b-Quarks für den Datensatz <i>MC12tt</i> . . . . .	147



# Literaturverzeichnis

- [1] T.Binoth, J.J.van der Bij, *Influence of strongly coupled, hidden scalars on Higgs signals* arXiv:hep-ph/9608245 v1 (1996).
- [2] F. Halzen, A. D. Martin, *Quarks & Leptons*, John Wiley & Sons, New York (1984).
- [3] American Association for the Advancement of Science, <http://www.aaas.org/>.
- [4] G. 't Hooft, M.J.G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B44 (1972), Nobelpreis 1999.
- [5] P. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys.Rev.Lett 12 (1964) 508.
- [6] Imperial College London, <http://www.hep.ph.ic.ac.uk/cms/physics/higgs.html>.
- [7] S. L. Glashow *Partial symmetries of weak interactions*, Nucl. Phys. 22 (1961), Nobelpreis 1979.
- [8] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. 19 (1967), Nobelpreis 1979.
- [9] A. Salam, *Weak and electromagnetic interactions*, Proc. of the 8th Nobel Symposium on 'Elementary particle theory, relativistic groups and analyticity', Stockholm, Sweden (1968), Nobelpreis 1979.
- [10] TeV4LHC Higgs working group, *Standard Model Higgs cross sections at hadron colliders*, <http://maltoni.home.cern.ch/maltoni/TeV4LHC/SM.html>.
- [11] Young-Kee Kim, Talk at the XXIII International Symposium on Lepton and Photon Interactions at High Energy, August 13-18, 2007, Korea.
- [12] Particle Data Group, *Journal of Physics G: Nuclear and Particle Physics*, (July 2006).
- [13] D.E. Groom et. al., *Review of Particle Physics*, Eur. Phys. J. C15, 1 (2000).
- [14] T. Binoth, J.J. van der Bij, *The stealthy Higgs model at future Linear Colliders*, arXiv:hep-ph/9908256 (1999).
- [15] J.J.van der Bij, *The minimal non-minimal standard model*, arXiv:hep-ph/0603082 v1 (2006).

- [16] Dietrich Schroff, *Studie zur elektronischen Kalibration der Flüssig-Argon-Kalorimeter und Entdeckung unsichtbar zerfallender Higgs-Bosonen im ATLAS-Experiment*, Doktorarbeit an der Albert-Ludwigs-Universität Freiburg (2004).
- [17] Delphi collaboration, Eur. Phys. J. C32 (2004).
- [18] L3 collaboration, Phys. Lett. B609 (2005).
- [19] OPAL Collaboration, Eur. Phys. J. C27 (2003).
- [20] ATLAS Collaboration, *Technical Design Report: Calorimeter Performance*, CERN/LHCC 96-40 (1997).
- [21] D. Froidevaux, P. Sphicas, *General-Purpose Detectors for the Large Hadron Collider*, Annu. Rev. Nucl. Part. Sci. 2006. 56: 375-440 .
- [22] R. Brun, F. Rademakers, et al. , *ROOT - An Object Oriented Data Analysis Framework*, <http://root.cern.ch/>.
- [23] M. Goto, <http://root.cern.ch/twiki/bin/view/ROOT/CINT> .
- [24] A. Höcker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, *Users Guide TMVA: Toolkit for Multivariate Analysis*, arXiv physics/0703039, <http://tmva.sf.net>.
- [25] *Athena: The ATLAS Common Framework*, Developer Guide, <https://twiki.cern.ch/twiki/bin/view/Atlas/AthenaFramework>.
- [26] T. Sjöstrand, L. Lönnblad, S. Mrenna, P. Skands, *PYTHIA 6.4: Physics and Manual*.
- [27] CTEQ: *The Coordinated Theoretical-Experimental Project on QCD*, <http://www.phys.psu.edu/cteq/> .
- [28] *GEANT - Detector Description and Simulation Tool*, Cern Program Library Long Writeup W5013, Geneva (1993).
- [29] ATLFAST Homepage, <http://www.hep.ucl.ac.uk/atlas/atlfast/>.
- [30] E. Richter-Was, D. Froidevaux, *ATLFAST 2.0: a fast simulation package for ATLAS*, ATLPHYS-98-131.
- [31] <https://twiki.cern.ch/twiki/bin/view/Atlas/CommonTrackingSoftware>.
- [32] W. Liebig, *Preparation of the ATLAS Inner Detector Track Reconstruction Software*, Vortrag beim Topical Seminar on Innovative Particle and Radiation Detectors, Siena (3. Oktober 2006).
- [33] <https://twiki.cern.ch/twiki/bin/view/Atlas/TrackingOverviewRecoMuonboy>.
- [34] <http://atlas-samusog.web.cern.ch/atlas-samusog/muonboy/Muonboy.htm>.

- [35] M. Heldmann, D. Cavalli, *An Improved  $\tau$ -Identification For The ATLAS Experiment*, ATLAS-PHYS-PUB-2006-008.
- [36] The QCD and Standard Model Working Group, *Summary Report from Les Houches*, arXiv:hep-ph/0005114v1 (1999).
- [37] <https://twiki.cern.ch/twiki/bin/view/Atlas/JetAlgorithms> .
- [38] J. Joyce, *Bayes Theorem*, Stanford Encyclopedia of Philosophy (2003).
- [39] J. Friedman in Proc. of the 1974 CERN School of Computing, *353QH twice smoothing algorithm*, Norway, Aug. 11-24, 1974.
- [40] D. W. Scott, *Multivariate Density Estimation, Theory, Practice and Visualization*, Wiley-Interscience, New York (1992).
- [41] C. Weiser, *Separation von Quark- und Gluonjets und Bestimmung des relativen  $K_s^0$ - und  $\Lambda^0$ -Produktionsverhältnisses mit dem DELPHI-Detektor bei LEP*, Diplomarbeit an der Universität Karlsruhe (1994).
- [42] M. D. Richard, R. P. Lippmann, *Neural Network Classifiers Estimate Bayesian a posteriori Probabilities*, Neural Computation 3 (1991).
- [43] K.P. Kratzer, *Neuronale Netze*, Hanser Verlag (1991).
- [44] H. Burkhardt, *Material zur Kursvorlesung Mustererkennung*, WS 06/07, Kap. 8a.
- [45] Y. Zhang, Z.-X. Chi, D.-G. Wang *Decision Tree's Pruning Algorithm Based on Deficient Data Sets*, Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'05) (2005).
- [46] T. Sjöstrand, Private Mitteilung.
- [47] V.Cavasinni, D.Costanzo, I.Vivarelli, *Forward Tagging and Jet Veto Studies for Higgs Events Produced via Vector Boson Fusion*, ATL-PHYS-2002-008 (2002).
- [48] J.J. van der Bij, Private Mitteilung.
- [49] P. Gagnon, *Invisible Higgs boson decays in the ZH and WH channels*, ATL-PHYS-PUB-2005-011 (2005).
- [50] F.Meisel, M.Dührssen, M.Heldmann and K.Jakobs, *Study of the discovery potential for an invisibly decaying Higgs boson via the associated ZH production in the ATLAS experiment*, ATL-PHYS-PUB-2006-009 (2006).
- [51] T. Sjöstrand, P. Skand, *Multiple interactions and the structure of beam remnants*, JHEP 03 (2004) 053.

- [52] L. Neukermans, B. Di Girolamo, *Observing an invisibly decaying Higgs boson in ATLAS via vector boson fusion*, ATL-PHYS-2003-006 (2003).
- [53] Glen Cowan, *Statistical Data Analysis*, (1998).
- [54] F. Derue, A. Kaczmarzka, *Soft electron identification and b-tagging with DC1 data*, ATL-PHYS-2004-026.
- [55] B. Elpelt, J. Hartung, K.-H. Klösener, *Statistik*, Oldenbourg Verlagsgruppe (1999).
- [56] S. Brandt, *Datenanalyse*, Spektrum Verlag (1999).
- [57] DELPHI Collaboration, *b-tagging in DELPHI at LEP*, CERN-EP/2002-088.

# Danksagungen

Ich danke Prof. Dr. K. Jakobs für die Möglichkeit, diese Arbeit verfassen zu können, und für seine Hilfe dabei.

Weiterhin danke ich Dr. C. Weiser, der mir während meiner Diplomarbeit ein stets hilfreicher und motivierender Ansprechpartner für alle erdenklichen Probleme war.

Ständige und geduldige Hilfen bei der Lösung zahlreicher Probleme waren auch Dr. R. Bernhard, M. Dührssen, Dr. M. Heldmann, I. Ludwig, J. Mechnich, F. Meisel, H. Nilsen, G. Piacquadio und viele mehr. Vielen Dank dafür!

Ein weiterer Dank geht an Prof. Dr. J. J. van der Bij, der mir in Fragen bezüglich des Phionen-Modells stets eine große Hilfe war.

Ein herzliches Dankeschön geht auch an meine Familie und an meine Freundin, die mich zu jeder Zeit unterstützt haben !

Freiburg, den 3. Dezember 2007

# Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe. Weiterhin wurden von mir nur die angegebenen Quellen und Hilfsmittel zur Anfertigung dieser Arbeit verwendet.

---

(Ort) (Datum)

---

Sascha Thoma