Suche nach dem Higgs-Boson im Kanal $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ am TeVatron mit Hilfe Neuronaler Netze

D i p l o m a r b e i t

vorgelegt der
Fakultät für Physik
Ludwig-Maximilians-Universität München

von

D a n i e l a   G ö r i s c h

München, den 24. März 2006

Erstgutachter: Prof. Dr. Otmar Biebel

Zweitgutachter: Prof. Dr. Gerhard Buchalla

Captured effortlessly
that's the way it was.
Happened so naturally
I did not know it was love.
The next thing I felt was
you holding me close.
What was I gonna do?
I let myself go!

And now we're flyin' through the stars.
I hope this night will last forever.

I've been waitin' for you.
It's been so long.
I knew just what I would do,
when I heard your song.
Filled my heart with your bliss.
Gave me freedom.
You knew I could not resist.
I needed someone.

And now we're flyin' through the stars.
I hope this night will last forever.
Oh oh oh oh

Ain't nobody
Loves me better
Makes me happy
Makes me feel this way
Ain't nobody
Loves me better than you

...

Chaka Khan *Ain't nobody*

[*David "Hawk" Wolinski*]

dedicated to

...

A N D R E A S

...

Ich mag Dich nicht nur, weil Du bist wer Du bist,
sondern wer ich bin,
wenn ich mit Dir zusammen bin.

(GABRIEL GARCÍA MÁRQUEZ)

Zusammenfassung

Diese Arbeit beschreibt die Suche nach dem Standardmodell-Higgs-Boson im Zerfallskanal $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ mit dem DØ-Experiment am TeVatron. Hier kollidieren Protonen und Antiprotonen bei einer Schwerpunktsenergie von $\sqrt{s} = 1.96\ TeV$.
Ausgehend von Kollisionen dieser Art können Standardmodell-Higgs-Bosonen entstehen, welche ab einer Masse von 135 GeV vorwiegend in WW-Paare zerfallen. Der leptonische W-Zerfallskanal bietet eine klare Signatur bei der Signalselektion und Untergrundunterdrückung. Der Schwerpunkt der Analyse liegt bei einer Higgsmasse von 160 GeV. Das Prinzip wird darüberhinaus auf weitere Higgsmassen verallgemeinert. Die Vorselektion, welche den qualitativen Bereich eingrenzen soll, erfolgt anhand von Schnitten. Zur weiteren Optimierung des Signal-zu-Untergrund-Verhältnisses $\frac{S}{\sqrt{B}}$ werden Neuronale Netze untersucht, bearbeitet und trainiert. Ihre Generalisierungsfähigkeit wird in der Hauptselektion auf Signal wie auch auf Untergrund angewendet.
Die Ergebnisse dieser Analyse werden abschließend mit einer auf Schnitten basierenden Selektion verglichen. Dieser Vergleich erfolgt ausschließlich für eine Higgsmasse von 160 GeV. Eine Einschätzung der Sensitivität wird gegeben.

Inhaltsverzeichnis

Kapitel 1

Einleitung

"The standard model is working too well", das Standardmodell funktioniert einfach zu gut [1]. Diese Aussage von Richard P. Feynman steht als Symbol für die Stärke der bisherigen experimentell bestätigten Vorhersagen in und um dieses Modell. Der Großteil der Teilchenphysiker hält es für so überzeugend, daß ohne Zweifel die Annahme besteht, es weiter ergänzen und damit auch verbessern zu können.
Das Standardmodell regelt den Aufbau der Materie aus kleinsten elementaren Teilchen, sowie die zwischen ihnen wirkenden Kräfte. Bereits in der Antike stand die Frage nach dem Aufbau der Welt im Mittelpunkt. Gingen die Griechen noch davon aus, das Atom[1] sei der kleinste bekannte Baustein, so wissen die Wissenschaftler unserer Zeit, daß dieses Atom aus weiteren Subkomponenten aufgebaut ist. Als kleinste unteilbare Teilchen der heutigen Zeit sind Fermionen (die Leptonen und die Quarks) sowie Bosonen bekannt, welche durch das Standardmodell beschrieben werden. Wurden früher die Theorien den experimentellen Entdeckungen angepasst, so fordert heute die Theorie, sofern sie gültig sei, nach der Existenz neuer, weiterer elementarer Teilchen. So mußte beispielsweise erklärt werden, woher und vor allem wie die Teilchen ihre Massen erhalten. Dazu wurde das Standardmodell theoretisch mit der Einführung eines neuen Teilchens erweitert. Eine Lösung des Problems war gefunden. Dieses neue Teilchen, das Higgs-Boson, wurde nach Peter Higgs, dem Mann, der diese Theorie entwickelt hatte, benannt. Wird das Higgs-Boson gefunden, wäre das eine weitere Bestätigung für das Standardmodell.
Sein experimenteller Nachweis gestaltet sich dabei besonders schwierig. Anders als bei den Austauschbosonen der schwachen Wechselwirkung, deren Massen vor ihrer Entdeckung theoretisch so genau vorhergesagt werden konnten, daß sich mit zunehmend besseren technischen Geräten die experimentelle Bestätigung immer deutlicher abzeichnete, ist die Masse des Higgs-Bosons einer von vielen unbestimmten Parametern des theoretischen Standardmodells. Durch Eingrenzung des Massenbereiches kann eine Richtung für die Suche vorgegeben werden. So konnte bereits der Massenbereich unter 114.4 GeV durch die vier LEP-Experimente ausgeschlossen werden. Die obere Grenze von 290 GeV liefern Berechnungen, die Schleifenkorrekturen berücksichtigen.
Die vorliegende Analyse befaßt sich mit Suche nach dem Higgs-Boson. Ihr Hauptinteresse liegt dabei bei einer Higgsmasse von 160 GeV. Die Selektion soll mit Hilfe

[1] Atom: aus dem griechischen von ἄτομος - unteilbar

Neuronaler Netze erfolgen, deren Erfolg mit einer auf Schnitten basierenden Selektion verglichen wird. Kapitel 2 erläutert die theoretischen Grunglagen des Standardmodells und speziell die Details zu dieser Analyse. Das technische Hintergrundwissen zum Beschleuniger und zum Detektor liefert Kapitel 3. Wie die Rekonstruktion der für die Analyse wichtigsten Komponenten erfolgt, schildert Kapitel 4. Im fünften Kapitel wird die Vorselektion besprochen, danach in Kapitel 6 werden allgemein und spezifisch die Neuronalen Netze vorgestellt. Abschließend werden die Ergebnisse der Analyse für Higgsmassen von 120 GeV, 140 GeV, 160 GeV und 180 GeV präsentiert, speziell für die 160 GeV-Higgsmasse werden diese mit den Ergebnissen einer schnittbasierten Analyse verglichen.

Kapitel 2

Das Higgs-Boson im Standardmodell

Wie bereits in Kapitel 1 erwähnt, wird die Theorie der Teilchenphysik äußerst genau[1] im sogenannten Standardmodell der Elementarteilchen und den damit verbundenen Wechselwirkungen beschrieben. Einen Überblick über dieses Modell mit Bezug auf die für diese Analyse wichtigen Aspekte soll in diesem Kapitel gegeben werden.

2.1 Das Standardmodell der Elemantarteilchenphysik

Im Standardmodell werden zwei Arten von Elementarteilchen vereinigt. Zum einen sind dies die Materieteilchen selbst, zum anderen die für die Wechselwirkungen verantwortlichen Austauschteilchen. Beide sind elementare Bausteine unserer Welt, welche theoretisch als punktförmige Teilchen betrachtet werden. Experimentell bestimmen sich ihre radialen Ausmaße zu Größenordnungen $< 10^{-18} - 10^{-19}\ m$.
Die Gruppe der Materieteilchen besteht aus Fermionen. Dies sind Spin-1/2-Teilchen, die sogenannten Leptonen und Quarks. In ihrer Anordnung erkennbar sind drei Generationen (siehe dazu Tabelle 2.1). In der Gruppe der Leptonen sind das: das Elektron (e), das Myon (μ) und das Tau (τ) sowie deren dazugehörige Neutrinos. Analog werden die sechs verschiedenen *Quark-Flavours* Up (u), Down (d), Charm (c), Strange (s), Top (t) und Bottom (b) paarweise zu Generationen zusammengeschlossen. Neutrinos in diesem Fall wechselwirken ausschließlich schwach, Leptonen hingegen auch elektromagnetisch. Quarks nehmen an allen Wechselwirkungen teil, unterliegen somit zusätzlich der starken Wechselwirkung. Aufgrund der Annahme der Masselosigkeit der Neutrinos im minimalen Standardmodell werden nur linkshändige Eigenzustände des schwachen Isospins I_3 dargestellt. Generell bilden linkshändige Fermionen Dubletts, währenddessen die Rechtshändigen als Singuletts existieren. Die Händigkeit weist hierbei auf die Ausrichtung von Spin und Impuls der einzelnen Teilchen zueinander hin. Rechtshändigkeit signalisiert Parallelität, Linkshändigkeit Antiparallelität. Der experimentelle Befund, daß Massen- und Flavoureigenzustände unterschiedlich sind, findet per Konvention Ausdruck in einer Mischung der Quark-Eigenzustände mit $I_3 = -1/2$. Speziell die im Vergleich zu den Leptonen schwereren Quarks zeigen diese

[1] Die Genauigkeit der veranschlagten Tests liegt derzeitig bei einem Niveau $< 0.1\ \%$ [2]

Konvention in den apostrophierten Teilchen, Leptonen selbst hingegen nicht aufgrund der im (minimalen) Standardmodell verschwindenden Neutrinomasse.
Zu jedem dieser hier behandelten 12 Elementarteilchen existiert das dementsprechende Antiteilchen mit gleichen Eigenschaften jedoch entgegengesetzter elektrischer, schwacher und starker Ladung[2].
Austauschteilchen, Eichbosonen genannt, stellen den zweiten Typ an Elementarteilchen dar. Abgesehen von der Gravitation, erfolgen die fundamentalen Wechselwirkungen hier durch Austausch dieser Bosonen – sogenannte Spin-1-Teilchen (siehe Tabelle 2.2).

	1.Generation	2.Generation	3.Generation	Q	I_3
Lepton	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
	e_R^-	μ_R^-	τ_R^-	-1	0
Quark	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
	u_R	c_R	t_R	+2/3	0
	d_R	s_R	b_R	-1/3	0

Tabelle 2.1: Die Generationen im Standardmodell: Elektron (e), Myon (μ), Tau (τ) und die dazugehörigen Neutrinos (ν) sowie Up-, Down-, Charm-, Strange-, Top- und Bottom-Quark. Ebenfalls angegeben werden die elektrische Ladung Q in Einheiten der Elementarladung und die dritte Komponente des schwachen Isospins I_3. L steht hier für linkshändige, R für rechtshändige Teilchen und Apostrophe kennzeichnen Mischzustände.

Die elektromagnetische Wechselwirkung erfolgt durch das Photon, welches an alle elektrisch geladenen Teilchen koppelt. Die Theorie wird durch die *Quantenelektrodynamik* oder kurz QED beschrieben. Diese stellt die bis heute älteste Quantenfeldtheorie dar, in der elektromagnetische Prozesse störungstheoretisch in Abhängigkeit von der zugehörigen elektromagnetischen Kopplungskonstante α berechnet werden können. Diese Kopplungskonstante wird zu niedrigen Energien hin beschrieben durch

[2] Die Ladung ist die Eigenschaft eines Teilchens, mit einem Feld in Wechselwirkung zu treten. Die Art der Ladung gibt an, welcher Wechselwirkung die Teilchen unterliegen.

Wechselwirkung	elektromagnetisch	schwach		stark
Ursache	elektrische Ladung	schwache Ladung		Farbladung
Reichweite[m]	∞	10^{-18}		$< 10^{-15}$
Kopplungskonstante	1/137	$G_{\mathrm{F}} = 1.167 \cdot 10^{-5}\ GeV^{-2}$		≤ 1
Eichboson	Photon	$W^{\pm}$	Z^0	8 Gluonen
Masse [GeV]	0	80.2	91.2	0

Tabelle 2.2: Die fundamentalen Wechselwirkungen mit den dazugehörigen Eichbosonen im Standardmodell.

die Feinstrukturkonstante:

$$\alpha = \frac{\mathrm{e}^2}{4\,\pi\,\varepsilon_0\,\hbar\,c} \approx \frac{1}{137} \quad .$$

Sie gilt als sehr genau verstanden und überaus erfolgreich. Ihre wesentlichen Eigenschaften liegen wie bei allen Quantenfeldtheorien in der Eichinvarianz und der Renormierbarkeit. Diese ermöglichen es, die Phase des Fermionfeldes frei zu wählen und aufgrund von Selbst-Energie-Anteilen divergente Terme aufzuheben. Bedingt durch die Masselosigkeit des Photons erlangt die elektromagnetische Wechselwirkung eine unendliche Reichweite.

Dies würde man zunächst auch für die starke Wechselwirkung vermuten, welche durch die acht Gluonen, ebenfalls als masselos angesehen, getragen wird. Durch die zusätzliche Eigenschaft des *confinement* wird die Reichweite endlich mit einer Größenordnung $\sim 10^{-15}\ m$. Die Kopplung[3] hier erfolgt über die Farbladungen, welche mit rot, grün, blau und den Antifarben entsprechend mit antirot, antigrün und antiblau bezeichnet werden. Gluonen selbst tragen auch Farbladung und können demnach miteinander wechselwirken. Wie der Name der Wechselwirkung bereits vermuten läßt, muß eine vergleichsweise stärkere Kopplung existent sein, gegeben durch α_s, die bei kleinen Energien den Wert 1 annimmt und im Grenzwert zu hohen Energien verschwindet, so daß Quarks und auch Gluonen sich hier wie annähernd freie Teilchen verhalten können, wenn sie bei entsprechend geringen Distanzen (gleichbedeutend mit hohen Energien) beobachtet werden. Dieses Prinzip wird als asymptotische Freiheit bezeichnet. Vergrößert man nun die Betrachtungsskala, nimmt α_s im Vergleich so stark zu, daß Quarks lediglich in gebundenen Systemen – den Hadronen – auftreten (*confinement*). Hadronen nach außen sind farbneutrale Baryonen (Drei-Quark-Teilchen) sowie farbneutrale Mesonen (Zwei-Quark-Teilchen). Die gesamte Theorie spiegelt die *Quantenchromodynamik* oder QCD wider [3].

Als Träger der schwachen Wechselwirkung sind die Eichbosonen W^+, W^- und Z^0 bekannt. Die hohe Masse der Bosonen impliziert eine geringe Reichweite dieser Kraft von $\sim 10^{-18}\ m$. Im unteren Energiebereich wird die Stärke dieser Wechselwirkung beschrieben durch die Fermikonstante G_{F}. Diese bestimmt sich zu

$$G_{\mathrm{F}} = 1.167 \,\cdot\, 10^{-5}\, GeV^{-2} \quad .$$

[3]Kopplungen im physikalischen Sinne erfolgen durch Wechselwirkungen also durch Austausch von Energie oder Impuls der beteiligten Systeme.

Eine ausführlichere Behandlung der *Quantenflavourdynamik* oder QFD wird in [4] gegeben.
Bereits Glashow, Weinberg und Salam erlangten die Erkenntnis, daß sich die elektromagnetische und die schwache Wechselwirkung zu einer Theorie vereinheitlichen lassen, der elektroschwachen Wechselwirkung, beschrieben durch die *Glashow-Weinberg-Salam-Theorie*. In dieser sind Elektromagnetismus und schwache Wechselwirkung jeweils nur einer von zwei Teilaspekten, welche bei Energien $< 100\ GeV$ sehr gut getrennt diskutiert werden können.
Auf die Gravitation und das damit vermutete Graviton soll im Rahmen dieser Arbeit nicht näher eingegangen werden, aufgrund der zu vernachlässigenden Stärke der Kraft von $\approx 10^{-38}$ (bezogen auf die starke Wechselwirkung).
Komplette Darstellungen des Standardmodells finden sich in diversen Elementarteilchenbüchern wieder. Unter diesen befinden sich auch [3] und [4].

2.1.1 Symmetrien

Wie bereits erwähnt, bildet die theoretische Grundlage des Standardmodells die Quantenfeldtheorie, welche auf der $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ Eichsymmetrie basiert. Darin enthalten sind die Symmetriegruppe der starken Wechselwirkung $SU(3)_C$ und die der elektroschwachen Wechselwirkung $SU(2)_L \otimes U(1)_Y$. Spontane Symmetriebrechung (siehe Kapitel 2.1.3) führt dazu, daß von $SU(2)_L \otimes U(1)_Y$ nur noch die elektromagnetische Wechselwirkung $U(1)_{em}$ übrig bleibt. Daraus wird ersichtlich, wie schwache und elektromagnetische zur elektroschwachen Wechselwirkung vereint sind.
In der Teilchenphysik spielt die Existenz dieser Symmetrien eine bedeutende Rolle. Generell spricht man von einer Symmetrie, wenn das betrachtete physikalische System unter einer Transformation U invariant bleibt, also die Invarianz des Hamiltonoperators[4] vorliegt. Die algebraische Struktur einer Gruppe, die sogenannte Symmetriegruppe rührt von den unabhängigen Erzeugenden dieser Symmetrie her.
Im Standardmodell enthaltene Symmetriegruppen $SU(2)_L$, $SU(3)_C$ und $U(1)_Y$ gehören zu den kontinuierlichen Symmetrien, deren Parameter kontinuierliche Werte annehmen. Sie alle sind lokale Eichsymmetrien, deren Transformationsparameter explizit von den Raum-Zeit-Koordinaten abhängen. Dies bildet die Grundlage für das wichtige Eichprinzip lokaler (Eich-)Symmetrien:

Ein physikalisches System, genannt Ψ, dessen Dynamik durch die Lagrangedichte $\mathcal{L}$ beschrieben ist, soll invariant unter einer globalen Symmetrie G sein. Soll diese Symmetrie auch lokal gelten, so daß letztlich die Invarianz von $\mathcal{L}$ unter lokaler Transformation gewährleistet bleibt, dann müssen Eichfelder (z.B. Vektorbosonfelder) eingeführt werden. Diese wechselwirken mit dem Feld Ψ in eichinvarianter Weise. Dabei bestimmen die Eigenschaften der Symmetriegruppe G die Anzahl der Eichfelder wie auch die spezielle Form ihrer eichinvarianten Wechselwirkung. Demnach ist die Zahl assoziierter Eichbosonfelder der Zahl der Generatoren von G gleich.

[4] $UHU^+ = H$

2.1.2 Die Eichtheorie der elektroschwachen Wechselwirkung

In Kapitel 2.1.1 wurde bereits verdeutlicht, daß zur Beschreibung der elektroschwachen Wechselwirkung die Symmetriegruppe $SU(2)_L \otimes U(1)_Y$ benutzt wird, welche die lokale Symmetriegruppe der elektroschwachen Lagrangedichte sein muß. $SU(2)_L$ stellt die schwache Isospingruppe dar, deren Wirkung nur auf linkshändige Fermionen zielt, während $U(1)_Y$ für die schwache Hyperladungsgruppe steht. Insgesamt hat die elektroschwache Symmetriegruppe vier Generatoren, drei aus der $SU(2)_L$, die $I_i = \sigma_i/2$ (σ_i beschreiben die Pauli-Matrizen mit $i = 1,2,3$) und einer aus $U(1)_Y$, der $Y/2$. Fermionen werden demnach nicht nur durch ihre Ladung Q, sondern zusätzlich durch den schwachen Isospin I_3 sowie durch die schwache Hyperladung Y charakterisiert. Die genauen Werte der Quantenzahlen sind Tabelle 2.3 zu entnehmen. Wie sich die Verknüpfung dieser drei "Ladungen" darstellt, beschreibt die *Gell-Mann-Nishijima-Relation*:

$$Q = I_3 + \frac{Y}{2} \tag{2.1}$$

Lepton	I	I_3	Q	Y
$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} +1/2 \\ +1/2 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ -1 \end{pmatrix}$
e_R^-	0	0	-1	-2
Quark	I	I_3	Q	Y
$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} +1/2 \\ +1/2 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	$\begin{pmatrix} +1/3 \\ +1/3 \end{pmatrix}$
u_R	0	0	+2/3	+4/3
d_R	0	0	-1/3	-2/3

Tabelle 2.3: Die Fermion-Quantenzahlen der ersten Generation. Fermionen der 2. und 3. Generation weisen diesselben Quantenzahlen entsprechend der hier dargestellten Familie auf.

Den Generatoren zugehörige Eichbosonen sind die W_μ^i mit $i = 1,2,3$ aus der $SU(2)_L$ und B_μ aus der $U(1)_Y$. Als Besonderheit der schwachen Wechselwirkung zeigt sich, daß diese nur auf linkshändige Fermionen wirkt, die sich deshalb als Dubletts transformieren, während es bei den rechtshändigen Fermionen Singuletts sind.
Mit den vorangegangenen Bedingungen ergibt sich für die lokale Eichtransformation der Fermionfelder im Raum des schwachen Isospins und der Hyperladung folgender Zusammenhang:

$$\psi(x) \to \mathrm{e}^{i\frac{\sigma_i}{2}\Theta^i(x)}\mathrm{e}^{i\frac{Y}{2}\alpha(x)}\psi(x) \tag{2.2}$$

Die Funktionen $\Theta^i(x)$ und $\alpha(x)$ seien differenzierbar.
Zur Konstruktion der elektroschwachen Lagrangedichte wird zunächst die kovariante Ableitung eingeführt

$$D_\mu = \partial_\mu - ig\frac{\sigma_i}{2}W_\mu^i - ig'\frac{Y}{2}B_\mu \tag{2.3}$$

in welcher g und g' die beiden Kopplungskonstanten (g aus der $SU(2)_L$ und g' entsprechend der $U(1)_Y$) und W^i_μ sowie B_μ die den Symmetriegruppen zugehörige Eichbosonen darstellen. Um eine lokale Eichinvarianz zu gewährleisten, transformieren die Eichfelder selbst (siehe (2.4) und (2.5)). Dabei enthält die Transformation des W-Feldes (2.4) nicht verschwindende Strukturkonstanten und damit eine Selbstkopplung. Evidenzen hierfür sind wieder in der nicht-vertauschenden $SU(2)_L$ zu suchen.

$$W^i_\mu \to W^i_\mu - \frac{1}{g}\partial_\mu \Theta^i(x) - \varepsilon^{ijk}\Theta^j W^k_\mu \tag{2.4}$$

$$B_\mu \to B_\mu - \frac{1}{g'}\partial_\mu \alpha(x) \tag{2.5}$$

Dies führt zur Lagrangedichte der elektroschwachen Wechselwirkung, in der die Selbstwechselwirkungsvertizes in den kinetischen Termen der Eichfelder $W_{\mu\nu}W^{\mu\nu}$ enthalten sind:

$$\mathcal{L}_{elektroschwach} = \sum_{f=l,q} \bar{\psi}_f(x)\gamma^\mu D_\mu \psi_f(x) - \frac{1}{4}W^i_{\mu\nu}W^{\mu\nu}_i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.6}$$

$$W^i_{\mu\nu} \equiv \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\varepsilon^{ijk}W^j_\mu W^k_\nu \tag{2.7}$$

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.8}$$

Die Masseneigenzustände, also die physikalisch beobachtbaren Eichbosonen $W^\pm_\mu$, Z_μ und A_μ (Photon) ergeben sich aus den jeweiligen elektroschwachen Eigenzuständen durch Linearkombination oder Rotation um den schwachen Winkel Θ_W:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{2.9}$$

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\Theta_W & \sin\Theta_W \\ -\sin\Theta_W & \cos\Theta_W \end{pmatrix} \cdot \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix} \tag{2.10}$$

Über diesen Winkel sind auch die Kopplungen g und g' bestimmt zu

$$g = \frac{\mathrm{e}}{\sin\Theta_W} \tag{2.11}$$

$$g' = \frac{\mathrm{e}}{\cos\Theta_W} \tag{2.12}$$

Für umfassendere sowie detailliertere Einblicke in Eichtheorien sei [5] empfohlen.
In der bis hier behandelten Theorie sind sowohl die Fermionen als auch die Eichbosonen dieser Wechselwirkung entgegen unserer Beobachtung masselos. Die Lösung des Problems kann erst mit dem Higgs-Mechanismus geklärt werden.

2.1.3 Higgs-Mechanismus und spontane Symmetriebrechung

Nach dem Einblick in die elektroschwache Theorie in Kapitel 2.1.2 soll nun die Lagrangedichte des Standardmodells angegeben werden. Nach dem jetzigen Standpunkt muß sie die Dynamik der Fermionen, also ihre kinetische Energie und Wechselwirkungen mit den Austauschbosonen, enthalten. Desweiteren ist ein zweiter Term notwendig, der die freie Propagation der Eichfelder und ihre Selbstwechselwirkung beschreibt. Ohne bisherige Betrachtung blieben die Massenterme für Fermionen und Bosonen. Diese sind nicht “von Hand” in die Theorie einführbar, da die Fermion-Massenterme in ihrer Form $m\bar{\psi}_R\psi_L$ kein $SU(2)_L$ Singulett darstellen, was die Invarianz von $\mathcal{L}$ verletzt. Bei Bosonen wiederum geht die Proportionalität der Massenterme mit $m^2 A_\mu A^\mu$, was keiner Invarianz der Eichtransformation des Feldes A entspricht. Die Lösung des Problems liegt in der Einführung zweier zusätzlicher Terme in der Lagrangedichte des Standardmodells (2.13), dem *Yukawa-Term* und dem *Higgs-Term*. Der Yukawa-Anteil ist verantwortlich für die Massenterme der Fermionen, erzeugt durch Kopplung des Higgs-Feldes an die Fermionen. Dem hingegen führt der Higgs-Mechanismus auf Grundlage der spontanen Symmetriebrechung zur Erzeugung der Bosonmassen. Diese zwei neuen Terme in der Lagrangedichte bilden ein zusätzliches $SU(2)_L$ Dublett, welches durch das Higgs-Feld mit Spin 0 beschrieben wird.

$$\mathcal{L} = \mathcal{L}_{Fermion} + \mathcal{L}_{Eich} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{2.13}$$

Aus der Quantenchromo- und Quantenflavourdynamik können

$$\mathcal{L}_{Fermion} = \sum_{\psi_L(x)} \bar{\psi}_L(x)\gamma^\mu D_\mu \psi_L(x) + \sum_{\psi_R(x)} \bar{\psi}_R(x)\gamma^\mu D_\mu \psi_R(x) \tag{2.14}$$

und

$$\mathcal{L}_{Eich} = -\frac{1}{4}W^i_{\mu\nu}W^{\mu\nu}_i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G^\alpha_{\mu\nu}G^{\mu\nu}_\alpha \tag{2.15}$$

übernommen werden. Dabei ist

$$G^\alpha_{\mu\nu} \equiv \partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu + g_s f^{\alpha\beta\gamma} A_{\mu\beta} A_{\nu\gamma} \tag{2.16}$$

der QCD-Feldstärketensor mit den Gluonfeldern A^α_μ. Das Konzept von spontaner Symmetriebrechung und darauf basierendem Higgs-Mechanismus soll im folgenden Abschnitt kurz erläutert werden:

Die Symmetrie eines physikalischen Systems ist genau dann spontan gebrochen, wenn die Lagrangedichte invariant unter der Symmetrietransformation dieses Systems ist, aber nicht der Grundzustand (Vakuumzustand) der Theorie.
Als Beispiel hierzu wird oft ein Ferromagnet unendlicher Ausdehnung mit einer Temperatur T nahe der Curie-Temperatur T_C verwendet. Unendlich viele elementare Spins mit rotationsinvarianten Wechselwirkungen liefern die Beschreibung solch eines Systems. Für den Grundzustand gelten dabei, abhängig von seiner Temperatur, zwei unterschiedliche Möglichkeiten: Getreu des Falles $T > T_C$ unterliegt die Spinausrichtung dem Zufall und setzt somit die mittlere Magnetisierung gleich Null. Der Grundzustand unterliegt hier der Rotationsinvarianz. Wird nun aber T auf $T < T_C$ abgesenkt, geht die freie Ausrichtung vieler Spins in eine bestimmte, aber willkürlich ausgerichtete über. Einen von unendlich vielen möglichen Grundzuständen wählt das System dann aus (spontane Magnetisierung). In der Regel ist es ein Zustand, dessen Rotationsinvarianz aufgrund der ausgezeichneten Richtung gebrochen ist. Man spricht von spontaner Symmetriebrechung. Abbildung 2.1 verdeutlicht dies anhand eines Potentialmodells. Bild (a) ist kennzeichnend für den Fall $T > T_C$ mit einer symmetrischen Struktur und einem einzigen Minimum, während in (b)

der Fall $T < T_C$ dargestellt wird mit unendlich vielen Minima (diese entsprechen den entarteten Grundzuständen). Ausschlaggebend ist die Wahl eines bestimmten Grundzustandes zur Erzeugung der Symmetriebrechung der Rotationssymmetrie.

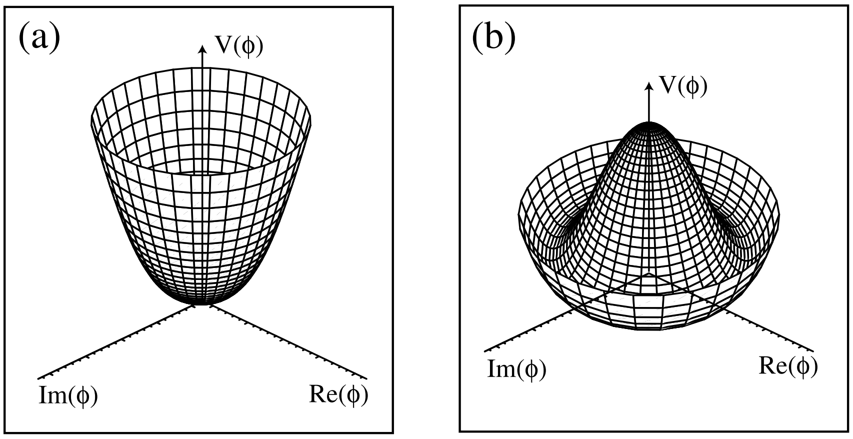

Abbildung 2.1: Potentialmodell zur Veranschaulichung der spontanen Symmetriebrechung. Die Situation $T > T_C$ ist in Bild (a) wiedergegeben, mit einem Potential rotationssymmetrischer Struktur und einem Minimum. Bild (b) verkörpert den Fall $T < T_C$ durch eine "mexican hat" Form und die unendlich vielen entarteten Minima.

Im folgenden Abschnitt soll die Symmetriebrechung innerhalb des Standardmodells erläutert werden, die dazu führt, daß von der elektroschwachen Wechselwirkung nur die elektromagnetische Wechselwirkung übrig bleibt:

$$SU(2)_L \otimes U(1)_Y \rightarrow U(1)_{em} \tag{2.17}$$

Die $U(1)_{em}$ muß erhalten bleiben, da sie eine Symmetrie des physikalischen Systems ist und somit zugleich eine Symmetrie des Vakuums bleiben muß. Es bleiben aber verschiedene Möglichkeiten offen, um Gleichung 2.17 dennoch zu erfüllen. Dazu wird der minimale Higgs-Sektor betrachtet, der ein Potential umfaßt, das analog zum Fall des Ferromagneten konstruiert ist. Die zu diesem Sektor gehörige Lagrangedichte lautet:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \tag{2.18}$$

Hierbei ist Φ ein komplexes, skalares, zum Isospin gehöriges Dublettfeld im Rahmen der $SU(2)_L \otimes U(1)_Y$

$$\Phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad , \tag{2.19}$$

welches wir später als Higgs-Feld indentifizieren können. Sein Potential wird durch nachfolgende Gleichung beschrieben:

$$V(\Phi) = -\frac{\mu^2}{2}\Phi^\dagger\Phi + \frac{\lambda}{4}(\Phi^\dagger\Phi)^2 \qquad \text{, wobei } \lambda > 0 \quad . \tag{2.20}$$

Dieses Potential wurde bereits in Abbildung 2.1 vorgestellt, in der Bild (a) der Form des Potentials mit Werten für $\lambda > 0$ und $\mu^2 > 0$ entspricht, während in Bild (b) die Form der Parameter $\lambda > 0$ und $\mu^2 < 0$ zugeordnet wird. Hier liegen alle Minima auf einem Kreis mit Radius $v = \sqrt{-\frac{\mu^2}{\lambda}} \simeq 246\ GeV$. Dieses v stellt den Vakuumerwartungswert des Higgs-Bosons dar [6], welcher auch über die Fermikonstante G_F bestimmt werden kann:

$$\frac{G_\mathrm{F}}{\sqrt{2}} = \frac{g^2}{2m_W^2} = \frac{1}{2v^2} \qquad \Longrightarrow \qquad v = (\sqrt{2}G_\mathrm{F})^{-\frac{1}{2}} = 246\ GeV$$

Das Minimum des Potentials muß so gewählt werden, daß das Photon masselos bleibt. Durch spontane Symmetriebrechung und Wahl der Parameter $\lambda > 0$ und $\mu^2 < 0$ besteht die Möglichkeit, dies zu erreichen. Weiter kann die experimentelle Vorgabe der Natur nur erfüllt werden, indem $\phi_1 = \phi_2 = \phi_4 = 0$ gesetzt wird, was eine 'gute' Wahl des Minimums zur Folge hat, so daß die Störungsrechnung um einen stabilen Wert um ϕ_3 herum betrieben werden kann. Für das Dublett, das Higgs-Feld, folgt demnach

$$\Phi(x) = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + \zeta(x)) \end{pmatrix} \qquad \text{mit } \ v = \sqrt{-\frac{\mu^2}{\lambda}} \tag{2.21}$$

mit dem Störparameter $\zeta(x)$, welcher das Higgs-Boson charakterisiert. Wie gefordert bleibt sein Vakuumwert entsprechend Gleichung (2.2) unter der Eichtransformation invariant.

Mit Einführung und Untersuchung der Massenterme liefert diese Theorie ein masseloses sowie drei massive Eichbosonen. Setzt man nun (2.21) in $\mathcal{L}$ (2.18) ein, werden im kinetischen Energieterm vier Felder ersichtlich. Für diese gilt:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \qquad \text{mit einer Masse } \ m_W = \frac{1}{2}v\,g \tag{2.22}$$

$$Z_\mu^0 = \frac{1}{\sqrt{g^2 + g'^2}}(gW_\mu^3 - g'B_\mu) \qquad \text{mit einer Masse } \ m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \tag{2.23}$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(g'W_\mu^3 + gB_\mu) \qquad \text{mit einer Masse } \ m_A = 0 \tag{2.24}$$

Jetzt ist es möglich, diese vier Felder mit den bereits besprochenen Eichbosonen zu assoziieren. So ergeben sich die massiven $W^\pm$ und Z^0 sowie das masselose Photon (γ). Für das Higgs-Boson selbst bleibt folgende Relation der Masse:

$$m_H = \sqrt{-2}\mu = \sqrt{2\lambda} \cdot v \tag{2.25}$$

Eine Zuordnung der Massen zu uns bekannten Teilchen durch die Postulierung eines Higgs-Feldes ist zwar erlaubt, jedoch bleibt in diesem Rahmen des Standardmodells das Higgs-Boson an sich (bis auf seinen Vakuumeigenwert) völlig unbestimmt. Seine Masse m_H sowie die Selbstkopplung λ sind über den Parameter v verknüpft.

Letztlich ist ersichtlich, daß alle Eichbosonmassen von lediglich diesem Massenparameter v und den Kopplungen g und g' abhängig sind. Mit (2.11) und (2.12) folgt eine Beziehung, die sich als überaus wichtig in der Benutzung zu Standardmodelltests erwiesen hat

$$\sin^2 \theta_W = 1 - \frac{m_W^2}{m_Z^2} \quad . \tag{2.26}$$

Einen weitaus komplexeren Überblick findet man in [7] und bei Weinberg [8].

Das Higgs-Dublett führt uns nicht nur zu den Massen der Eichbosonen, es kann genauso verwendet werden zur Erklärung der Fermionmassen [9]. Dafür muß das Higgs-Feld eichinvariant an die Fermionen ankoppeln. Zur Erklärung wird lediglich die erste Fermionfamilie betrachtet. Neutrinomassen bleiben weiterhin vernachlässigt. Als relevante Felder werden

$$\psi_L = \begin{pmatrix} \nu_f \\ f \end{pmatrix} \qquad \text{und} \qquad \psi_R = f_R \tag{2.27}$$

definiert. Der für Fermionen wichtige Term in der Lagrangedichte (2.13) ist der Yukawa-Term. Er wird durch

$$\mathcal{L}_{Yukawa} = -\tilde{g}_{\mathrm{f}} \left((\bar{\psi}_L \Phi)\, \psi_R + \bar{\psi}_R \, (\Phi^\dagger \psi_L) \right) \tag{2.28}$$

dargestellt. Die beschriebene Verbindung von Fermionen und skalaren Bosonen ist unter dem Begriff der *Yukawa-Kopplung* bekannt. Die Notation $\bar{\psi}_R$ steht in Gleichung (2.28) für $\psi_R^\dagger \gamma^0 \tilde{g}_{\mathrm{f}}$ und beschreibt einen Kopplungsparameter. Auch der Yukawa-Term erfüllt die Invarianz gegenüber Eichtransformationen vom Typ U(1) und SU(2) und verbindet links- mit rechtshändigen Fermionen.
Wird in Gleichung (2.28) noch das Higgs-Feld aus (2.21) eingesetzt, erhält man einen konstanten Term mit einer Higgs-Fermion-Wechselwirkung

$$\mathcal{L}_{Yukawa} \sim \tilde{g}_f \, \frac{v}{\sqrt{2}} \, (\bar{f}_R f_L + \bar{f}_L f_R) \quad . \tag{2.29}$$

Somit liest man aus Gleichung (2.29) aus dem ersten Summanden eine Fermionmasse von

$$m_f = \frac{\tilde{g}_f v}{\sqrt{2}} \tag{2.30}$$

ab. Durch sie ist auch $\tilde{g}_f$ festgelegt. Als Konsequenz wird ersichtlich, daß auch die im zweiten Summanden in (2.29) beschriebene Higgs-Fermion-Wechselwirkung – abgesehen vom Parameter v – bestimmt wird durch

$$\tilde{g}_f = \sqrt{2}\, \frac{m_f}{v} \quad . \tag{2.31}$$

Allgemein sollte zusätzlich erwähnt werden, daß die Stärke der Higgs-Kopplung direkt proportional zur Masse des Fermions ist und daß das Higgs-Boson in keiner Weise mit Neutrinos wechselwirkt. Zudem läßt sich phänomenologisch aus Gleichung (2.31) ableiten, daß der dominate Zerfall des Higgs-Teilchens den kinematisch erlaubten $f\bar{f}$-Endzustand mit der größten Fermionmasse wählt (siehe 2.5).

2.2 Das Higgs-Boson innerhalb des Standardmodells

Wie sich nun das Higgs-Boson in die Theorie des Standardmodells eingliedert, soll in den nächsten Abschnitten verdeutlicht werden.

2.2.1 Das Higgs-Boson und seine Masse

Durch Störungsrechnung des Higgsfeldes mit Störparameter $\zeta(x)$ ergibt sich, wie in (2.25) bereits angegeben, die Higgsmasse zu

$$m_H = \sqrt{2\lambda} \cdot v$$

Innerhalb dieser Gleichung ist es lediglich möglich, eine Aussage zu dem Verhältnis $v = \sqrt{-\frac{\mu^2}{\lambda}}$ zu treffen, λ bleibt eine Unbekannte. Mit den theoretischen Grundlagen dieser Analyse ist eine Möglichkeit zur numerischen Berechnung des Massenwertes nicht gegeben.
Trotzdem sollte es möglich sein, wenigstens eine Einschränkung für m_H anzugeben. Dazu werden Streuungen $e^+e^- \longrightarrow W^+W^-$ betrachtet. Man stellt fest, daß aufgrund der Unitaritätsgrenze, also damit die Reaktionswahrscheinlichkeit den Wert 1 nicht übersteigt, als theoretische Obergrenze der Higgsmasse 1 TeV [10] anzusetzen ist:

$$m_H^2 \leq \frac{2\sqrt{2}\,\pi}{G_\mathrm{F}} \approx (1\ TeV)^2$$

Weitere Eingrenzungsmöglichkeiten bieten Korrekturterme höherer Ordnung, welche sogenannte Standardmodellparameter wie Top-Quark- und W-Boson-Masse, aber auch den schwachen Mischungswinkel θ_W beeinflussen. Präzise Messungen dieser Größen erlauben indirekt Rückschlüsse auf m_H.

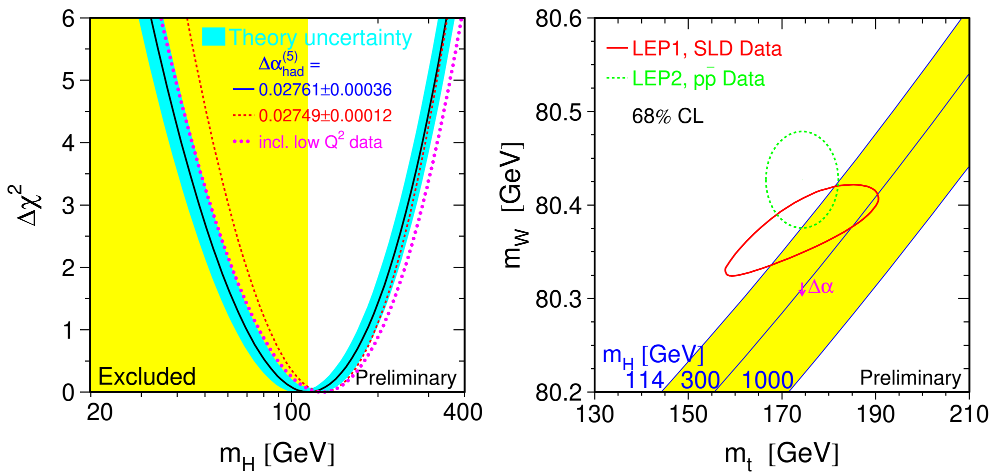

Abbildung 2.2: Die kombinierten Ergebnisse der vier LEP-Experimente (siehe auch 2.2.1). Linkerhand dargestellt ist der bereits experimentell ausgeschlossene Higgsmassenbereich (gelb) mit $m_H < 114.4$ GeV (95% C.L.) sowie ein theoretischer Fit an die Daten. Rechterhand der gelbe Streifen liefert die Beziehung von W- und Top-Masse (aus elektroschwachen Strahlungskorrekturen) in Abhängigkeit von der Higgsmasse [11].

Aber auch direkte Suchen nach dem Higgs-Boson helfen, noch unbekannte Parameter weiter einzugrenzen. Dies versuchten auch die damaligen LEP-Experimente. Die Ergebnisse der erzielten Erkenntnisse gibt Abbildung 2.2 wieder. Rechterhand ist der dato experimentell ausgeschlossene Bereich (gelb) von $m_H < 114.4$ GeV unter 95%igem C.L. aus der vergeblichen direkten Higgs-Suche dargestellt. Aus der $\Delta\chi^2$-Kurve des Fits des Standardmodells an 20 elektroschwachen Parametern ergibt sich ein Parabelminimum von $m_H = 91^{+58}_{-39}$ GeV. Damit gilt als obere Grenze der Masse des Higgs-Bosons ein Wert von etwa 290 GeV. Daneben gliedert sich ein Diagramm an, welches die W- und Top-Massen-Beziehung in Abhängigkeit von der Higgsmasse darstellt. Die Ovale und Konturen rechts basieren auf Analysen der Daten der vier LEP-Experimente, links werden Werte aus elektroschwachen Strahlungskorrekturen [11] von LEP-1 und SLD (durchgezogene Linie) sowie direkte Messungen von LEP-2 und TeVatron (gestrichelte Linie) kombiniert. Eine eventuelle Verschiebung der Relation beider Massen, sofern die Kopplungskonstante α um eine Standardabweichung variiert, bedeutet der mit $\Delta\alpha$ bezeichnete Pfeil.

2.2.2 Produktion und Zerfall von Higgs-Bosonen

Im Standardmodell ist das Higgs-Boson verantwortlich für die Massen der schweren Eichbosonen wie auch der Fermionen. Somit ist die Kopplung des Higgs-Bosons an diese Fermionen und Bosonen proportional zu deren Massen. Die dominanten Produktions- und Zerfallskanäle charakterisieren sich also duch eine Kopplung an $W^{\pm}$- und Z^0-Bosonen bzw. an Quarks und Leptonen der dritten Familie.

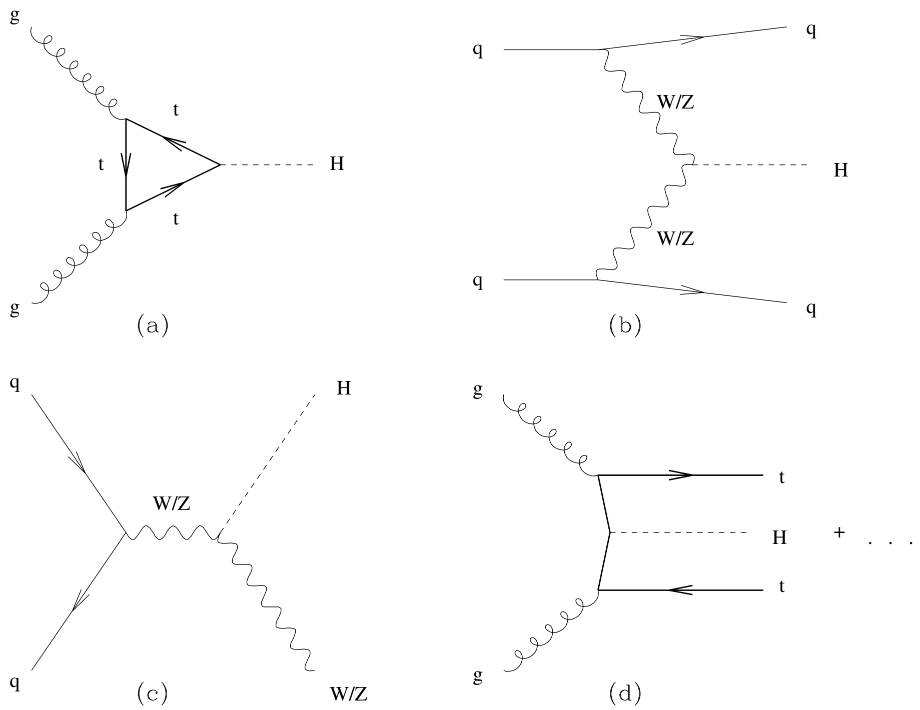

Abbildung 2.3: Die vorwiegenden Produktionsmechanismen für Higgs-Bosonen an Hadron-Collidern: (a) Gluon-Gluon-Fusion, (b) Vektorboson-Fusion, (c) assoziierte Produktion mit W- und Z-Bosonen oder (d) $t\bar{t}$-Paaren.

Die Hauptproduktionsmechanismen von Higgs-Bosonen an Hadron-Collidern sind in Abbildung 2.3 aufgezeigt. Ihre Bedeutung und auch Relevanz sind von Collider zu Collider verschieden, da je nach Schwerpunktsenergie und kollidierenden Teilchen verschiedene Ausgangssituationen geschaffen werden. Diese Analyse basiert auf Gluon-Gluon-Fusion. Hierbei wird ein Higgs erzeugt, welches in einem bestimmten Zerfallskanal (in dieser Analyse der W-Paarzerfall mit anschließendem Zerfall in Myonen sowie Neutrinos) untersucht werden kann. Auf der Suche nach leichteren H-Teilchen sei auf die Zerfallskanäle $WH \longrightarrow l\nu b\bar{b}$ und $ZH \longrightarrow \nu\bar{\nu}b\bar{b}$ verwiesen, in denen das Higgs jeweils in $b\bar{b}$-Paare zerfällt. Da die W/Z leptonisch zerfallen, kann dies zu einer sauberen Unterscheidung gegenüber der direkten $b\bar{b}$-Produktion genutzt werden. Diese beiden Mechanismen sind besonders für Analysen mit Daten des TeVatron-Beschleunigers zugrunde zu legen. Dahingegen spielen (c) und (d) aus Abb. 2.3 nur eine untergeordnete Rolle hinsichtlich des TeVatron, werden aber am künftigen LHC an Dominanz gewinnen.

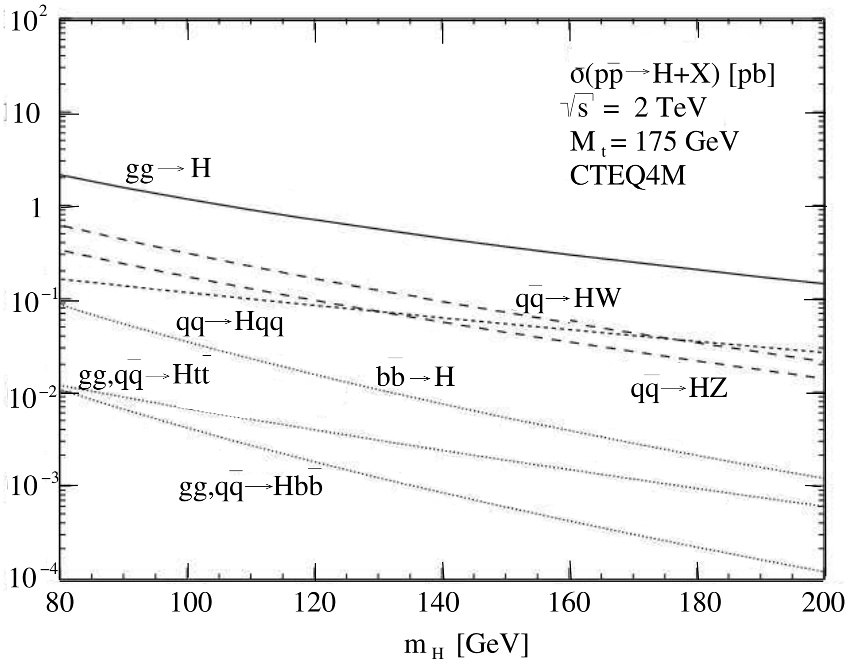

Abbildung 2.4: Darstellung der den diversen Produktionsmechanismen zugeordneten Wirkungsquerschnitte als Funktion von m_H (in Einheit pb) [12].

Die zugehörigen Wirkungsquerschnitte werden in Abbildung 2.4 veranschaulicht. Dargestellt ist ihr Verlauf in Abhängigkeit der Higgsmasse m_H. Die Berechnung dieser Graphen erfolgte unter der Annahme einer Schwerpunktsenergie von $\sqrt{s} = 2\ TeV$, wie sie in etwa dem TeVatron entspricht.
Für Gluon-Gluon-Fusion ergibt sich im vorhandenen Massenbereich der Wirkungsquerschnitt zu $\sigma = (1.5 - 0.1)\ pb$. Quark-Antiquark-Annihilationen der assoziierten Produktion erzielen eine Größenordnung kleinere Werte ($\sigma = (0.2 - 0.05)\ pb$).

Abbildung 2.5 zeigt die Verzweigungsverhältnisse der dominanten Higgs-Zerfälle. Ihre Darstellung erfolgt als Funktion dessen Masse. Deutlich erkennbar sind die beiden oberen Zerfallskurven, welche sich bei rund 135 *GeV* schneiden. Veranschaulichend heißt das: Ist ein Higgs-Boson leichter als dieser Grenzwert, zerfällt es vorwiegend in ein $b\bar{b}$-Paar, schwerere Teilchen zerfallen in $WW^{(*)}$-Paare. Eines dieser W-Bosonen kann dabei virtuell sein (mit * gekennzeichnet), d.h. es muß nicht zwingend auf seiner Massenschale liegen.

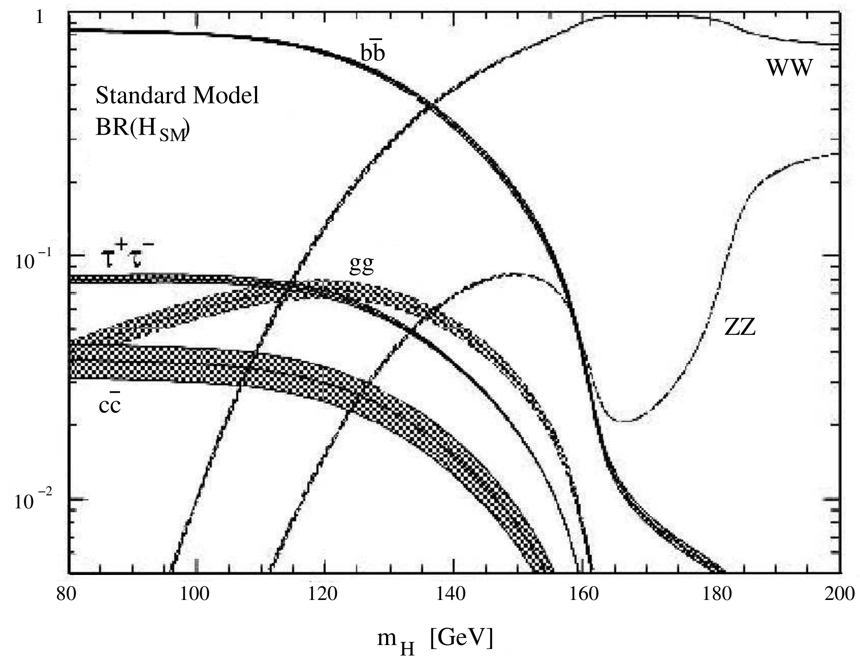

Abbildung 2.5: Die Abbildung stellt die unterschiedlichen Verzweigungsverhältnisse der diversen Zerfälle dar. Sie sind wiederum als Funktion der Higgsmasse angegeben [13].

Der $ZZ^{(*)}$-Kanal soll trotz seiner Unrelevanz für das TeVatron am FERMILAB aufgrund zu geringer integrierter Luminosität und zu niedrigen Wirkungsquerschnitten erwähnt werden, da seine Bedeutung in künftigen Analysen speziell der am LHC des CERN steigen wird.

Die vorliegende Analyse bearbeitet Higgs-Bosonen, deren Zerfall durch zwei W-Bosonen gekennzeichnet ist. Diese sollen myonisch zerfallen, d.h. das W^+-Boson geht über in ein Myon μ^+ und das zugehörige Myon-Neutrino ν_μ. Analog zerfällt W^- in ein μ^- und $\bar{\nu}_\mu$. Damit ist dieser Kanal gekennzeichnet durch den leptonischen Zerfall[5]. Die Signatur dieses Prozesses liegt in den hohen Transversalimpulsen der Myonen und damit einhergehend in der von den nichtdetektierten Neutrinos stammenden fehlenden transversalen Energie, welche aus den nicht gemessenen

[5] Weitere leptonische Kanäle sind Zerfälle mit Beteiligung zweier Elektronen e^+e^- sowie Kombinationen aus Elektronen $e^\pm$ und Myonen $\mu^\mp$ als auch Taus $\tau^+\tau^-$.

Transversalimpulsen beider Neutrinos entsteht. Der leptonische Zerfallskanal weist den Vorteil auf, weitaus weniger von QCD-Mehrjet-Untergrund betroffen zu sein.

2.2.3 Higgs-Bosonen und ihre Untergrundprozesse

Der in dieser Arbeit beschriebene Zerfallskanal $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ muß in Kombination mit weiteren Zerfallskanälen[6] betrachtet werden, welche kinematisch wie auch topologisch ähnliche Signaturen aufzeigen. Daten werden folglich mit dem Signal und weiteren Monte-Carlo-Simulationen verglichen.

Der wichtigste, wenn auch nicht ereignismäßig größte Untergrund stellt die Reaktion $p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ dar. Sie ist aufgrund der überaus ähnlichen Ereignis-Topologie und Kinematik besonders zu erwähnen und wird sich als irreduzibel erweisen. Lediglich die Spinkorrelation gibt eine Möglichkeit zur Diskrimination des Signals vom WW-Untergrund. Dadurch, daß das Higgs-Boson ein Spin-0-Teilchen ist, müssen die bei diesem Zerfall entstehenden Spin-1-Bosonen als WW-Paar korreliert sein. Dem hingegen sind die WW-Bosonen des Untergrundes im Hinblick auf eine Korrelation als frei und unabhängig voneinander zu betrachten. Kinematisch äußert sich dies in der Verteilung der Öffnungswinkel der Myonen der WW-Paare, welche in den Signal-Monte-Carlo-Simulationen zu kleineren Winkel tendiert, während der WW-Untergrund große Winkelwerte erwarten läßt.

Den Ereigniszahlen nach größte Bedeutung ist dem Drell-Yan-Prozess $q\bar{q} \longrightarrow Z/\gamma^* \longrightarrow \ell^+\ell^-$ zuzuordnen, der durch seine leptonischen Endzustände auch auf ähnliche Ereignissignaturen hindeutet. Darüberhinaus wird diese Analyse bestimmt durch die Top-Paar-Produktion $p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$, welche aufgrund der beteiligten Jets bis auf ein Minimum sehr gut separierbar sein sollten. Weiter müssen Ereignisse wie $p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ und $p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ betrachtet werden, sowie Higgs-Prozesse in denen via assozierter Produktion ein weiteres W-Boson neben dem WW-Paar auftritt, welches dann hadronisch zerfällt und damit charakteristischerweise Jets ausstrahlt. Auch der Υ-Zerfall und QCD-Mehrjet-Untergrund sind hier nebensächlich zu erwähnen.

[6] Die über das Signal heraus betrachteten Zerfallskanäle werden als Untergrund bezeichnet.

Kapitel 3

Das TeVatron und der DØ-Detektor

Das TeVatron mit dem DØ-Detektor [14] befindet sich am *Fermi National Accelarator Laboratory* oder kurz FERMILAB. Dies ist ein Forschungszentrum für Teilchenphysik, welches sich etwa 50 *km* westlich von Chicago, U.S.A., befindet und nach dem italienischen Physiker Enrico Fermi benannt wurde.

3.1 Das TeVatron

Mit rund 6 *km* Umfang ist das TeVatron der ringförmige Hauptbeschleuniger am FERMILAB, in dem an zwei Wechselwirkungspunkten Protonen p und Antiprotonen $\bar{p}$ bei einer Schwerpunktsenergie von $\sqrt{s} = 1.96\ TeV$ zur Kollision gebracht werden. Dieses $p\bar{p}$-Synchrotron wird bis zur Fertigstellung des LHC (*Large Hadron Collider*) am CERN der weltweit energiereichste Beschleuniger bleiben, der seine Tätigkeit mit bereits gemachten neuen Erkenntnissen und Entdeckungen untermauern kann. Zu den wichtigsten dieser Art zählt die Entdeckung des top-Quarks (1995) [15, 16]. Das gesamte Forschungsgelände ist in Abbildung 3.1 zu sehen.
Bevor Protonen und Antiprotonen in das Ringsystem des TeVatron eingespeist werden, müssen sie eine Reihe von Vorbeschleunigern durchlaufen, um genügend Energie aufzunehmen. Protonen werden dabei in drei Schritten auf ein kinetisches Energieniveau von 8 *GeV* gebracht. Dies erfolgt am Anfang in einem *Cockcroft-Walton-Beschleuniger* (Abb. 3.2). Dazu müssen aus Wasserstoffatomen und Elektronen e^- negativ geladene Wasserstoff-Ionen H^- erzeugt werden, welche mit einer Hochspannung auf 750 *keV* beschleunigt und erstmalig *gebuncht*[1] an den *LINAC* (*LIN*ear *AC*celerator) (Abb. 3.2) weitergegeben werden. Im *LINAC* wird auf 400 *MeV* kinetische Energie erhöht und mittels einer dünnen Carbonfaserfolie die Abtrennung der beiden e^- vorgenommen, so daß die Teilchenbunche im nächsten Schritt in das *Booster*-Synchrotron (Abb. 3.2) injiziert werden können.

[1] *Bunching* stellt den Prozess der Bündelung eines Teilchenstrahls dar, d. h. dieser wird in einzelne Teilchenpakete (*Bunche*) bestimmter Länge sowie Dichte unterteilt.

Abbildung 3.1: Zu sehen sind der Hauptbeschleuniger *TeVatron* (große Kreis) sowie der *Main Injector* (kleiner Kreis) zur Einspeisung der Teilchenbunche.

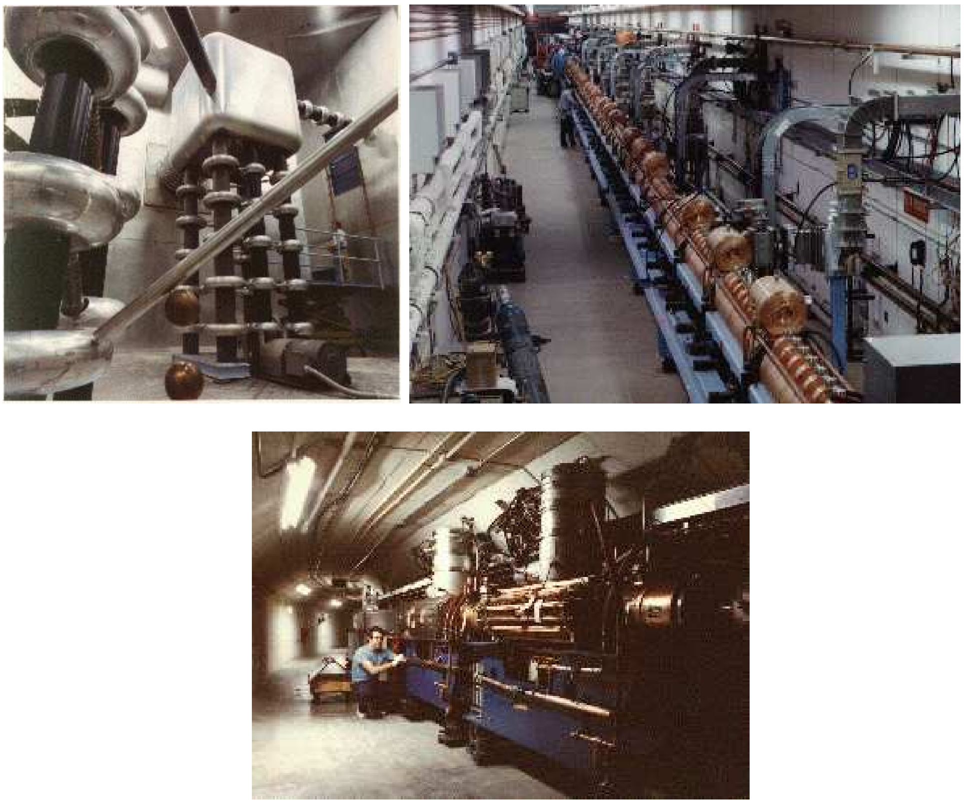

Abbildung 3.2: oben: links: *Cockcroft-Walton-Beschleuniger*, rechts: *LINAC*, unten: *Booster*

Der *Booster* ist mit 475 *m* Umfang das kleinste Synchrotron dieser Reihe und beschleunigt die Protonen in 0.033 Sekunden auf eine Energie von 8 *GeV* bevor sie dem *Main Injector* (Abbildung 3.1), einem weiteren 3 *km* langen Synchrotron zugeführt werden. Hier erfolgt die Energiesteigerung auf 150 *GeV* für Protonen wie auch Antiprotonen gleichermaßen. Darüberhinaus ist es die Aufgabe des *Main Injectors* den schon vorgebündelten Strahl dem TeVatron entsprechend weiter zu bunchen, d. h. er erzeugt alle 1.5 Sekunden ein Protonpaket mit je $5 \cdot 10^{12}$ Teilchen.
Nicht alle diesen Zyklus durchlaufenden Protonen werden ins TeVatron eingespeist, da parallel die Erzeugung des Antiprotonenstrahls erfolgt. Durch Trennung von Hauptbeschleuniger und *Main Injector* ist es möglich, zeitgleich zu den Kollisionen im Ring neue Antiprotonen zu erzeugen. Als Quelle dient hierbei ein Nickel-Kupfer-Target, bestehend aus Targetstation, Debuncher und Akkumulator, auf das alle 2.4 Sekunden 120 *GeV*-Protonen aus dem *Main Injector* aufschlagen. Die Rate der erzeugten Teilchen ist dabei verschwindend gering (nur etwa 20 Antiprotonen aus einer Million Protonen). Unter Zuhilfenahme einer Lithiumstromlinse werden die Antiprotonen zu einem Strahl fokussiert. Da bei diesem Targetbeschuß nicht nur Antiprotonen entstehen, sondern auch weitere unerwünschte Bestandteile, wird der Strahl durch einen gepulsten Dipolmagneten gelenkt, der diese Bestandteile mit einem nicht Antiprotonen entsprechendem Verhältnis von Ladung zu Masse in der Funktionsweise eines Massenspektrometers herausfiltert. Im Debuncher werden die Antiprotonen auf 8 *GeV* beschleunigt und zu einem kontinuirlichen Strahl fokussiert, bevor sie im Akkumulator gespeichert werden können. Durch stochastisches Kühlen im Akkumulator wird die mittlere Abweichung vieler Teilchen vom Sollorbit gemessen und (durch Kicker) korrigiert. Dies führt langfristig zu einer transversalen Kühlung des Strahls und zu einer Verringerung der longitudinalen Impulsstreuung, da Teilchen mit abweichenden Longitudinalimpuls ebenfalls vom Sollorbit abweichen ($r = p/(q \cdot B)$).

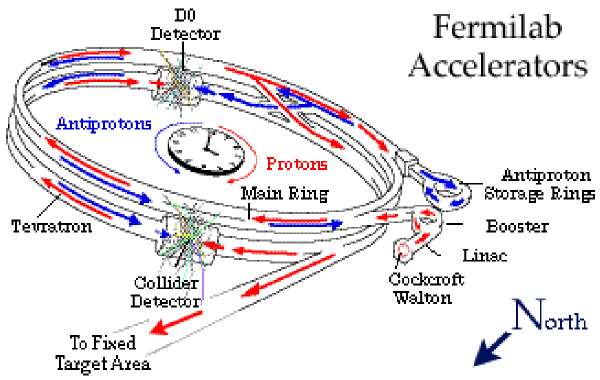

Abbildung 3.3: Das schematische Zusammenspiel aller Beschleunigerkomponenten, sowie die Standpunkte der Detektoren werden in dieser Abbildung gezeigt.

Nach Erzeugung der Antiprotonen, können diese ebenfalls dem *Main Injector* zugeführt und wie oben beschrieben beschleunigt werden. Die Injektion ins TeVatron erfolgt entgegen der Richtung der Protonen. Beide Teilchensorten erfahren hier eine

Beschleunigung auf etwa 1 TeV, ehe sie an den beiden Wechselwirkungspunkten[2] zur Kollision gebracht werden. Eine Gesamtübersicht des Beschleunigersystems gibt Abbildung 3.3. Für eine detailliertere Beschreibung des TeVatron wird auf [17] und [18] verwiesen.

3.2 Der DØ-Detektor

Der DØ-Detektor, im Querschnitt dargestellt in Abbildung 3.4, ist einer der beiden Teilchendetektoren am FERMILAB. In seinem Zentrum kommt es zu Proton-Antiproton-Kollisionen bei $\sqrt{s} = 1.96\ TeV$, aus denen diverse Reaktionsprodukte hervor gehen. Diese sowie deren Zerfallsprodukte zu detektieren und zu identifizieren, stellt die Aufgabe des Experiments dar.

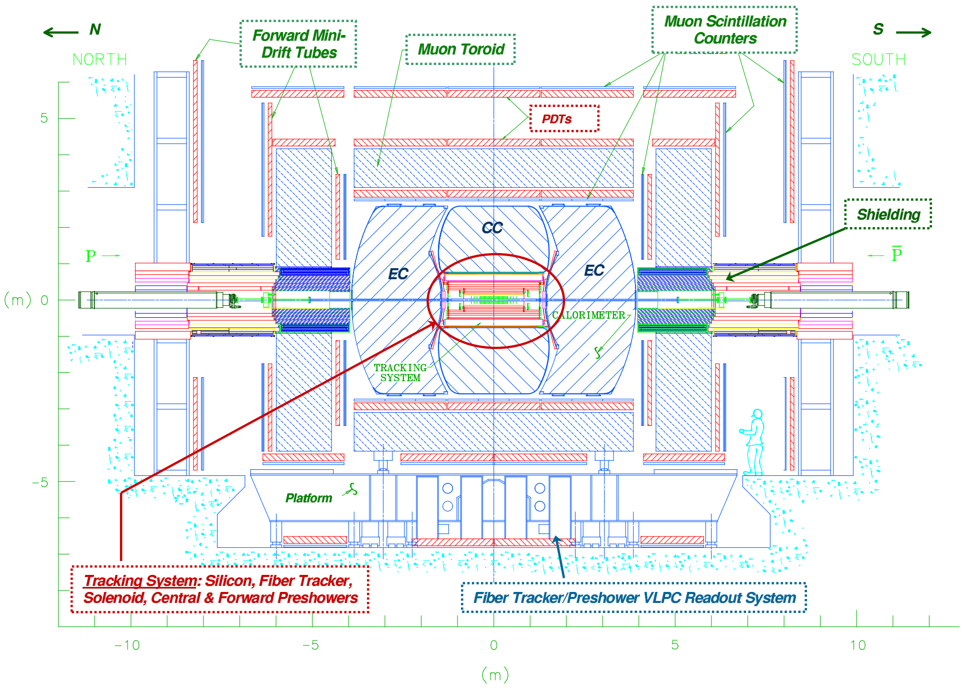

Abbildung 3.4: Querschnitt durch den DØ-Detektor.

Dazu trägt die weitreichende Raumwinkelabdeckung[3] von nahezu 4π bei, die nicht nur die Messung der Teilchen an sich, sondern auch die Rekonstruktion der Zerfallsprodukte zu einem Ursprung begünstigen.
Werden die Schwerpunktsenergien immer größer, so nimmt die Produktion der schweren Teilchen einen immer umfangreicheren Rahmen ein. Je nach Masse und Art der

[2]Wechselwirkungspunkte sind diejenigen Punkte im Beschleuniger, an denen beide Strahlen, sprich also Bunch für Bunch aufeinandertreffen. Diese Punkte sind i. a. umgeben von den Teilchendetektoren. Im Falle des TeVatron sind dies der CDF- und der DØ-Detektor.

[3]Der Raumwinkelabdeckung stellt hier die geometrische Akzeptanz des Detektors dar.

möglichen Zerfälle sind diese Teilchen kurz- oder langlebiger. Bei Kurzlebigkeit muß dieses Teilchen über Sekundärprodukte nachgewiesen werden. Aus diesem Grund ist es besonders wichtig, Eigenschaften und auch Spuren aller am Prozess beteiligten Reaktionspartner genaustens zu analysieren.
Dazu sind in Zylindersymmetrie um den Wechselwirkungspunkt herum die verschiedenen funktionsbasierten Subdetektoren angeordnet. Ein Blick in die räumliche Systemanordnung bietet Abbildung 3.5.

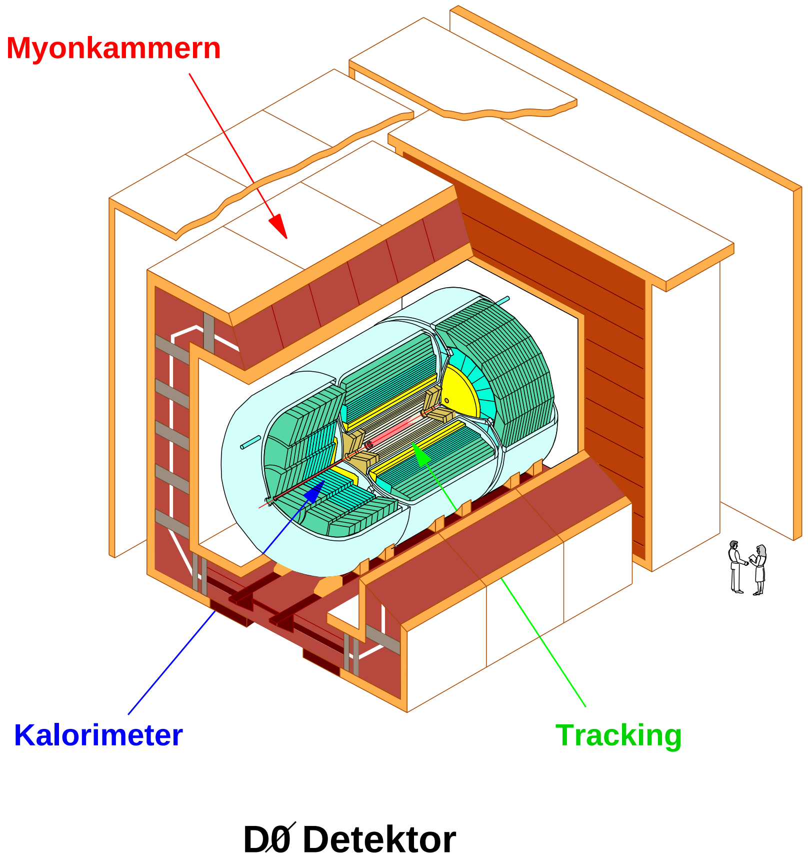

Abbildung 3.5: Ein Blick in den räumlichen DØ-Detektor.

Der DØ-Detektor, wie in den Abbildungen 3.4 und 3.5 ersichtlich, besteht aus dem *Silicon Microstrip Tracker* (SMT), dem *Central Fiber Tracker* (CFT), einem supraleitenden Magneten, dem Kalorimeter und in der Außenhülle aus dem Myonsystem. SMT und CFT gehören beide zum zentralen Spursystem. Darüberhinaus sind weitere Komponenten um die Strahlachse in Vorwärts- und Rückwärtsrichtung angeordnet. Dies soll den Studien der Phänomene kleiner Winkel dienen. Eine fundiertere Beschreibung des gesamten Detektors ist in [19] nachzulesen.
Konventionell wird der DØ-Detektor durch ein rechtshändiges kartesisches Koordinatensystem beschrieben. Der Ursprung dieses Systems liegt im nominellen Wechselwirkungspunkt, der zugleich den geometrischen Mittelpunkt des Detektors darstellt. Die x-Achse zeigt radial zum Mittelpunkt des TeVatron-Beschleunigers, die y-Achse

weist nach oben und die z-Achse fällt mit der Strahlachse des Protonenstrahls zusammen. Alle physikalischen Prozesse sind invariant gegenüber Rotationen um die Strahlachse, da Protonen sowie Antiprotonen unpolarisiert sind. Damit ist eine Anwendung von Kugelkoordinaten gerechtfertigt. Hierbei wird der Polarwinkel θ, der Winkel zur Strahlachse, zumeist durch die Pseudorapidität η charakterisiert, welche beschrieben wird als

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \quad .$$

Eine nähere Erläuterung der Pseudorapidität befindet sich im Anhang A. Natürliche Einheiten mit $\hbar = c = 1$ werden verwendet. Massen, Impulse und Energien werden alle in Energieeinheiten beschrieben.

3.2.1 Das zentrale Spursystem

Wie bereits erwähnt, besteht das zentrale Spursystem des DØ-Detektors im Inneren aus einem Siliziumdetektor, dem *Silicon Microstrip Tracker* (SMT) und einem mit szintillierenden Fasern ausgestatteten zweiten Detektor, dem *Central Fiber Tracker* (CFT). Beide sind von einem supraleitenden Magneten umgeben. Eine schematische Übersicht wird in Abbildung 3.6 geliefert.

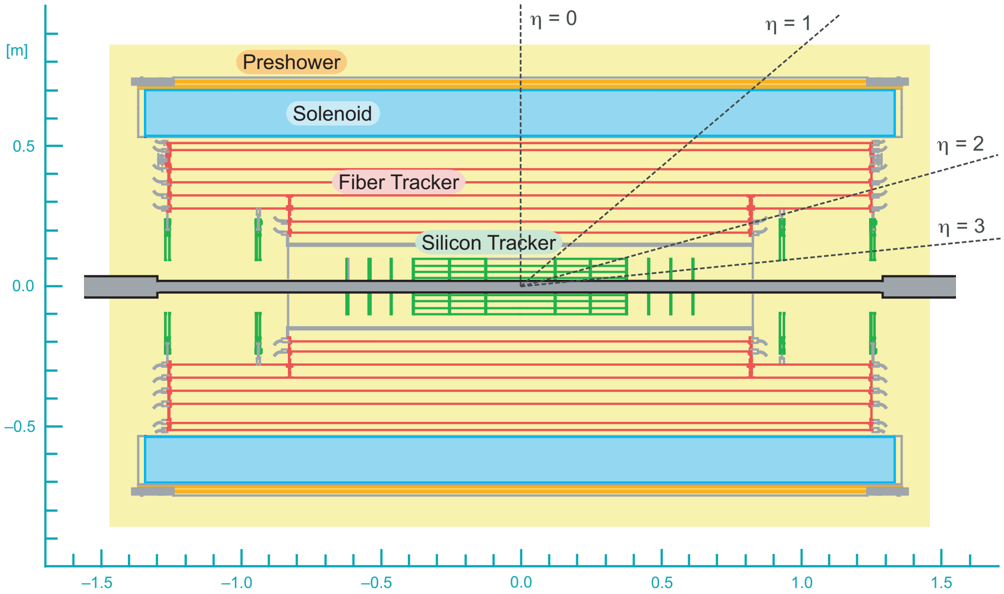

Abbildung 3.6: Die Darstellung zeigt den Querschnitt durch das Spursystem.

Silicon Microstrip Tracker

Der *Silicon Microstrip Tracker* (SMT) ist ein Silizium-Vertex-Detektor, der es ermöglicht, mit seinen sechs zylindrischen Segmenten, bestehend aus je vier doppellagigen Siliziumdetektoren, einen Winkelbereich von $|\eta| < 1.7$ abzudecken. Durch weitere 16 Silizium-Disks senkrecht zur Strahlachse kann der Detektierungsbereich auf $|\eta| < 3.0$ erweitert werden. Die auftreffenden Zerfallselemente ionisieren die Disks, so daß letztenendes über die ionisierte Ladung die Position des bzw. der Teilchen bestimmbar ist. Die Auslese zur Bestimmung der Primär- und Sekundärvertices erfolgt durch die 800 000 Auslesekanäle. Die Auflösung ergibt sich zu rund 50 μm.

Central Fiber Tracker

Den SMT umschließt in acht konzentrischen Zylinderlagen der *Central Fiber Tracker* (CFT). Dieser erstreckt sich innerhalb eines radialen Abstandes von 20 bis 51 *cm* und vereinigt rund 77 000 szintillierende Fasern von 0.8 *mm* Stärke. Jeder dieser acht Zylinder besteht aus einer Doppellage an Fasern, die Geradzahligen parallel und die Ungeradzahligen unter einem Winkel von 3° zur z-Achse angeordnet. Dies ermöglicht die Messung des Azimutwinkels ϕ bei bekanntem Radius und begünstigt einen räumlichen Informationsgewinn entlang der Strahlachse. Beide Detektoren zusammengenommen, erlauben eine Aussage über die z-Koordinate. Als Auslesetechnik für das Szintillationslicht werden Halbleiter-Photomultiplier (VLPC's) verwendet, welche das Signal über optische Fasern erhalten. Die geometrische Akzeptanz des CFT liegt bei $|\eta| \approx 1.6$, seine räumliche Auflösung ergibt sich zu 100 μm.

Werden hindurchtretende Teilchen sowohl im SMT als auch im CFT detektiert, ist von diesen eine höhere Impulsauflösung zu erwarten, als von Spuren die in jeweils nur einem der beiden wahrgenommen werden. Die Transversalimpulsauflösung dieser Teilchen ergibt sich für $|\eta| = 0$ nach [19] zu

$$\frac{\Delta p_T}{p_T} = 0.02 + 0.002 \cdot p_T/GeV \quad .$$

Supraleitender Solenoid

Silicon Microstrip Tracker und *Central Fiber Tracker* sind von einem *Solenoid*, einem supraleitendem Magneten von 2.8 *m* Länge und ca. 2 Tesla Stärke umgeben. Die radialen Ausmaße erstrecken sich über 60 *cm*. Durch Einsatz dieses Magneten werden die Bahnen geladener Teilchen mehr oder weniger gekrümmt, so daß anhand des Krümmungsradius und der Magnetfeldstärke eine Aussage über die Impulse der Teilchen getroffen werden kann. Damit ergibt sich

$$p = 0.3 \cdot q \cdot r \cdot B \quad ,$$

sofern die angegebenen Größen in GeV, m und T verwendet werden. Das Hauptwirken des zentralen Spursystems begrenzt sich auf einen Winkelbereich von $-3.0 < \eta < +3.0$. In diesem Bereich sollen Spuren geladener Teilchen so rekonstruiert werden, daß aus ihrer Bahnkrümmung der Impuls bestimmbar bleibt und anhand der Spuren ein Elektron von einem Photon unterschieden werden kann.

3.2.2 Das Kalorimeter

Kalorimeter werden in der experimentellen Teilchenphysik zur Bestimmung der Teilchenart und zur Bestimmung ihrer Energie eingesetzt, unabhängig von der etwaigen

Ladung. In ihrer Funktionsweise kommen die Prinzipien der starken sowie der elektromagnetischen Wechselwirkung zum Tragen. Deshalb ist beim DØ-Detektor das Kalorimeter abwechselnd aus absorbierenden und detektierenden Schichten aufgebaut. Eine stufenweise Energiemessung des sich ausbreitenden Schauers kann deshalb gewährleistet werden. Die dabei abgegebene Energie wird genutzt, um einen Rückschluß auf die Primärteilchenenergie zu erhalten.

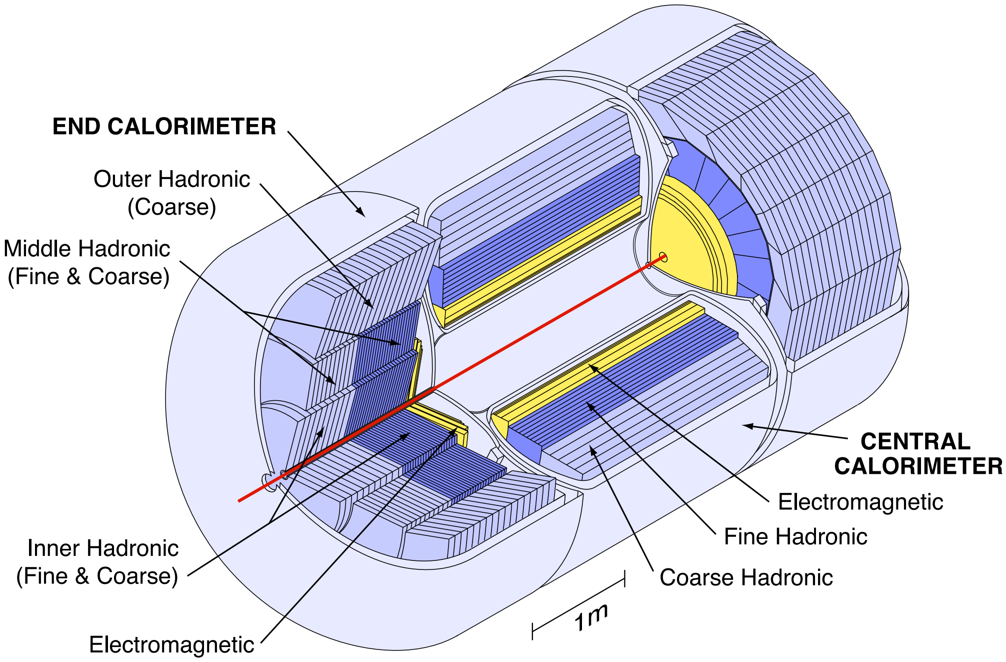

Abbildung 3.7: Eine 3D-Abbildung des DØ-Kalorimeters.

Das Kalorimeter des DØ-Detektors ist in Abbildung 3.7 dargestellt. Seine räumliche Ausdehnung erstreckt sich über 2.5 m Länge und 1.5 m im radialen Bereich. Wegen seiner hohen Granularität, seiner Uniformität und einem Pseudorapiditätsbereich von $|\eta| < 4.0$ zählt es zu den wichtigsten Komponenten des Detektors.
Gesamt setzt es sich aus drei separaten Kryostaten[4] zusammen, in denen das Zentralkalorimeter ($|\eta| < 1.0$) sowie die beiden Endkappenkalorimeter ($|\eta| < 4.0$) eingebettet sind. Desweiteren unterscheidet man noch zwischen elektromagnetischen und hadronischen Kalorimeteranteilen. Der Zentralbereich sowie die Endkappen sind daher aus vier elektromagnetischen, drei fein-hadronischen und einer grob-hadronischen Zylinderlage aufgebaut. Ersteres konzentriert sich durch eine feinere Segmentierung besonders auf die Energieermittlung von Elektronen und von Photonen. Darüberhinaus wird insbesondere für Photonen die Positionierung des Primärvertex verbessert, aufgrund der verfügbaren Richtungsinformationen durch die longitudinale Segmentierung. Erste Anzeichen hadronischer Schauer werden bereits hier erfaßt und im Hadron-Kalorimeter über die Wechselwirkung mit dem Material aufgenommen.

[4]Ein Kryostat (griech.kryos = kalt) ist ein Kühlgerät, in dem sehr tiefe Temperaturen möglichst konstant gehalten werden können.

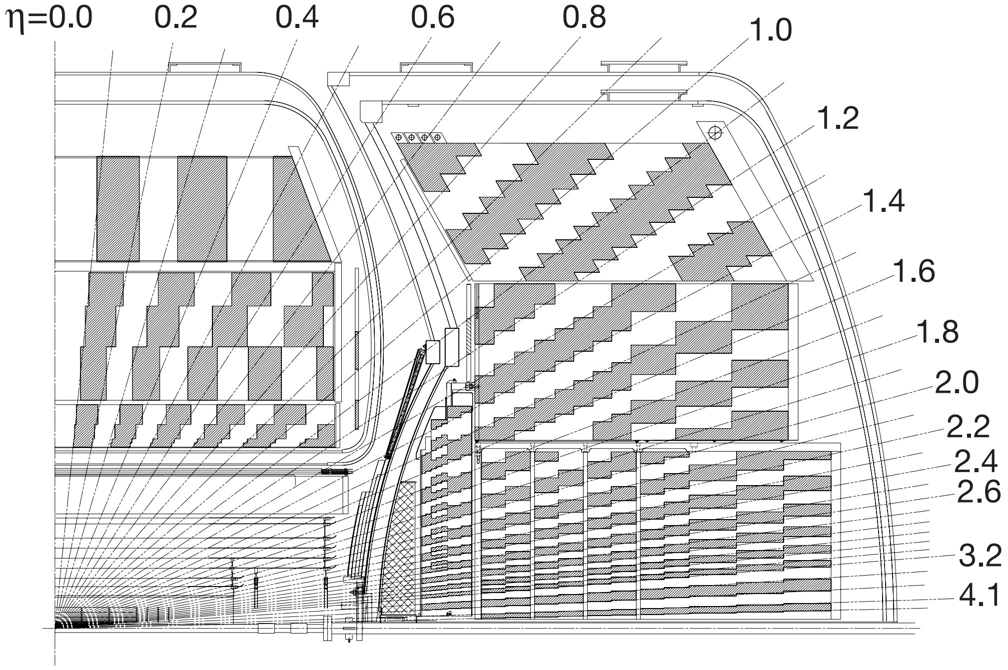

Abbildung 3.8: Querschnitt eines Quadranten des DØ-Kalorimeters und des Spursystems, projiziert in die x-z-Ebene. Die radialen Linien illustrieren die Detektor-Pseudorapidität und die pseudoprojektive Geometrie der Kalorimeter-Türme (tower). Die Türme haben im Durchschnitt die Größe $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ [20]. Die grauen und hellen Bereiche illustrieren die einzelnen Segmente des Kalorimeters.

Als Nachweismittel dient im Fall des DØ-Kalorimeters flüssiges Argon[5]. In den dazwischenliegenden Absorberschichten wurde Uran verwendet. Innerhalb der abwechselnden Schichten kommt es durch Energieverlust (Wechselwirken der Primärprodukte mit dem Uran) zur eigentlichen Ausbildung der Teilchenschauer (Sekundärprodukte), welche das Argon der darauffolgenden Schicht ionisieren. Die daraus vorhandene ionisierte Ladung, welche einem Bruchteil der Energie des Schauers entspricht, wird innerhalb der Zellen auf Kupferplatten gespeichert und ausgelesen. Die Summe der in einer Zelle vorhandenen Ladungen kann damit zur Gesamtenergie des Primärteilchens in Beziehung gesetzt werden. Aus den gemessenen Größen kann zusätzlich der fehlende Transversalimpuls bestimmt werden. Er wird den Neutrinos zugeschrieben. Diese wechselwirken nur äußerst schwach mit den Detektorkomponenten und entziehen sich somit meist jeglicher Registrierung. Aus der im Kalorimeter deponierten Energie sind desweiteren auch die hadronischen sowie elektromagnetischen Jets rekonstruierbar, deren Energieauflösung in etwa bei $\frac{80\ \%}{\sqrt{E\,[GeV]}}$ liegt.

Die Unterteilung des Kalorimeters in pseudoprojektive Türme (siehe Abbildung 3.8),

[5]Aufgrund des Argons befinden sich die Kalorimeterkomponenten innerhalb der Kryostate, da diese ein Verdampfen verhindern.

deren Zellstruktur eine Granularität von etwa 0.1×0.1 in η sowie in ϕ erreicht, liefert eine überdurchschnittliche Auflösung der Schauerposition.

Zusätzlich befinden sich zwischen den Kryostaten sogenannte *Interkryostat-Detektoren*, welche einen Informationsgewinn zum Energieverlust zwischen Zentral- und Endkappenkryostat liefern, der wiederum hinsichtlich der Messung und Erfassung von Jets innerhalb dieser Region von Bedeutung ist.
Zur fundierteren Kalorimeterbeschreibung sei auf [21] verwiesen. Abschließend sollen auch *Central* und *Forward Preshower Detector* (CPS und FPS) erwähnt werden. Der CPS, an der Außenseite des supraleitenden Magneten positioniert, überdeckt einen Pseudorapiditätsbereich von $|\eta| < 1.2$, während der FPS, vor den Endkappenkalorimetern, einen Arbeitsbereich zwischen $1.4 < |\eta| < 2.5$ bietet. Beide Detektoren bestehen aus Blei-Absorbern, gefolgt von mehreren Szintillatorlagen. Ihre Arbeitsbasis dient der besseren Identifizierung von Elektron und Photon. Dies wird erreicht durch die weiteren Messungen zur Energiedeposition pro Längeneinheit $\frac{dE}{dx}$, also der Ionisierung je Weglänge. Damit bleibt der Energieverlust des Teilchens vor Eintritt in das Kalorimeter abschätzbar.

3.2.3 Das Myonsystem

Als letzte Komponente des DØ-Detektors schließt sich das Myonsystem an. Seine Hauptbedeutung liegt in der Detektierung der Myonen.
Myonen $\mu^{\pm}$, die als Leptonen nicht der starken Wechselwirkung unterliegen und die aufgrund der im Vergleich zu Elektronen größeren Masse im DØ-Kalorimeter lediglich eine Energie von etwa 2.5 *GeV* verlieren, können somit eigenständig erfaßt werden. Dazu werden Treffer (Hits) aus den Myonkammern Spuren des zentralen Detektorsystems zugeordnet. Um eine Aussage zum Impuls treffen zu können, wird im Subdetektor ein Toroidmagnet eingesetzt.
Der zentrale Myondetektor sowie die Endkappen-Myondetektoren bilden das den restlichen Detektor umschließende Myonsystem, welches in Abbildung 3.9 dargestellt ist. Zentral wird ein Bereich $|\eta| < 1.0$ abgedeckt, die Endkappen erschließen zusätzlich $1.0 < |\eta| < 2.0$. Mit einer Impulsauflösung von etwa 0.1 für niederenergetische Myonen (Vielfachstreuung) und knapp 0.5 bei einem Transversalimpuls $p_T > 50\ GeV$ wird allerdings ein schlechteres Ergebnis als im zentralen Spursystem erzielt.

Der zentrale Myondetektor

Rechteckige Driftkammern, angeordnet innerhalb von drei Lagen (A-, B- und C-Lage) ergeben den *zentralen Myondetektor*, welcher nach der A-Lage durch einen Toroidmagneten von ca. 2 Tesla Stärke unterbrochen ist. Diese Kammerschicht besteht über und neben dem inneren Detektor aus vier Lagen von Driftröhren. Der Teil unter dem Detektor verfügt lediglich über drei Lagen. Insgesamt wurden hier 94 der Proportionaldriftrohre (PDT's) verarbeitet. Das Ziel liegt in der Bestimmung der Positionen der Myonen, welche aus der Driftzeit der Ladung des ionisierten Gases ermittelbar sind. Die Auflösung hinsichtlich Drift und Distanz der Rohre ergibt sich zu 1 *mm*.
Darüberhinaus wurden Szintillatoren um den zentralen Myondetektor angebracht.

Diese sogenannten *Cosmic Caps* bieten, über eine Zeitmessung, die Möglichkeit, kosmische Myonen von Zerfallsprodukten zu unterscheiden und damit letztenendes für ihre Unterdrückung zu sorgen. Dabei sind die zwischen Kalorimeter und A-Lage positionierten $A\phi$-Zähler sehr hilfreich, die ferner zu Triggerbelangen sowie zur Verbesserung der Bestimmung des Azimutwinkels ϕ beitragen. Aussagen über den Teilchenimpuls werden auch hier aus der Krümmung der Bahn im Magnetfeld gewonnen.

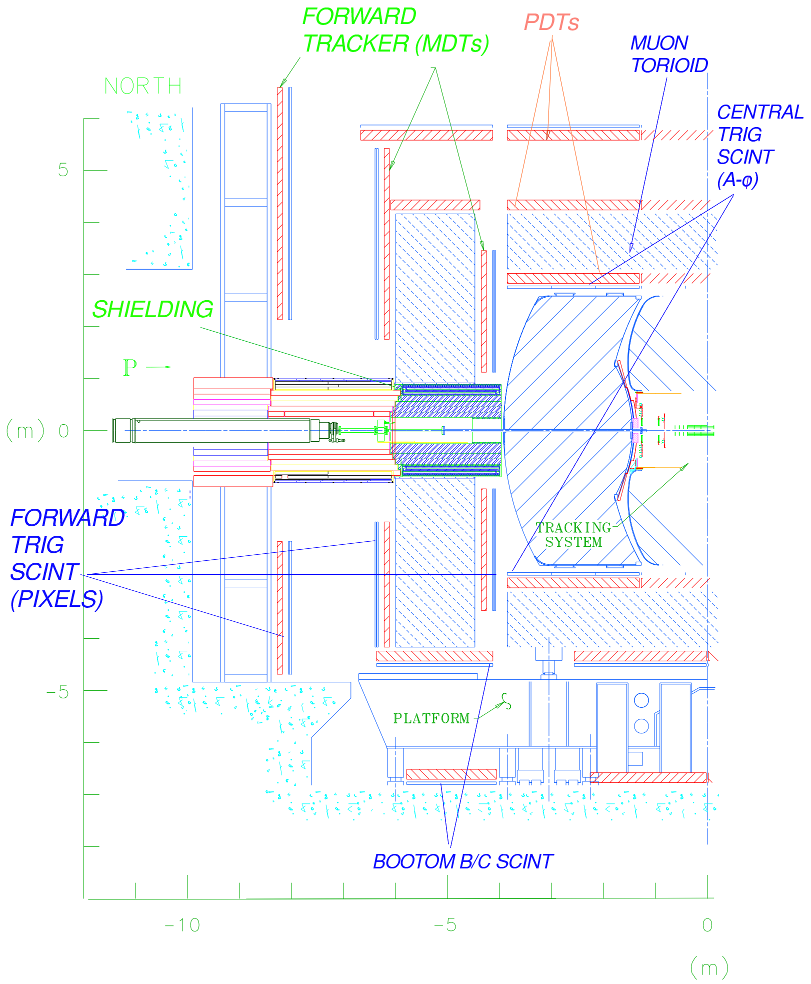

Abbildung 3.9: Ein zweidimensionaler Blick auf das Myonsystem.

Die Endkappen-Myondetektoren

Wie auch das zentrale Myonsystem müssen die Endkappen-Detektoren aus den drei Lagen mit dazwischenliegendem Toroiden aufgebaut sein. Demnach befinden sich in der A-Lage vier Schichten mit über 2400 Mini-Driftrohren (sogenannte MDT's mit einer Querschnittsfläche von $1\ cm \times 1\ cm$), sowie Pixelszintillatoren. B- und C-Lage erstrecken sich jeweils über drei Schichten. Sie sind in diesem Fall etwas größer und mit mehr Rohren ausgestattet. Aus oben genannten Gründen werden auch hier wieder Szintillationszähler in *Pixel Arrays* verwendet. Die erreichte Auflösung beläuft sich auf 0.7 *mm*.

In einem Bereich von $|\eta| < 1.0$ und $4.25 < |\phi| < 5.15$ unterhalb des Kalorimeters stellt sich durch zusätzliche Haltevorrichtungen des Detektors und der Ausleseelektronik das Triggern von Myonen und damit ihre Rekonstruktion äußerst schwierig dar. Dort ist zwar eine Szintillatorschicht vorhanden, aber auslesetechnisch sehr uneffektiv (siehe Abbildung 3.10), so daß dieser Bereich auszuschließen ist.

3.2.4 Die Luminositätsmonitore

Die Luminositätsmonitore, bestehend aus zwei Hodoskopen, messen die vom TeVatron gelieferte Luminosität. Zusammengesetzt sind die Hodoskope aus Szintillatoren, welche mit Hilfe von Photomultiplier ausgelesen werden. Sie sitzen zwischen den Kryostaten der Endkappenkalorimeter (Nord- und Süd-Detektor) und überdecken eine Pseudorapidität von $2.7 < |\eta| < 4.0$.
Die Luminosität ermittelt sich aus Strahlzusammenstößen in inelastischen $\mathrm{p\bar{p}}$-Wechselwirkungen. Via Zeitmessung zwischen nördlichem und südlichem Monitor ist eine Unterscheidung von Wechselwirkung zu Strahl-Halo[6] möglich, da die Restprodukte von Proton und Antiproton innerhalb einer Zeit von 9 *ns* den Detektor erreicht haben müssen. Für Teilchen, die dem nominellen Wechselwirkungspunkt entstammen, sollten diese Flugzeiten in Nord und Süd in etwa vergleichbar sein. Ein dabei auftretender Zeitversatz zweier Zerfallsprodukte unterstützt die exakte Bestimmung des Primärvertex. Als instantane Luminosität ergibt sich somit

$$\mathscr{L} = \frac{1}{\sigma_{eff}(\mathrm{p\bar{p}})} \frac{dN}{dt}$$

mit

$$\sigma_{eff}(\mathrm{p\bar{p}}) = \varepsilon_{LM} \cdot A \cdot \sigma_{\mathrm{p\bar{p}}} \ .$$

Die Effizienz des Luminositätsdetektors ε_{LM} wurde dabei gewonnen aus Studien der Kalorimeterzellenenergie, A steht im Bezug zu dessen Akzeptanz, welche aus Monte-Carlo-Simulationen berechnet wird. $\sigma_{\mathrm{p\bar{p}}}$ stellt den inelastischen $\mathrm{p\bar{p}}$-Wirkungsquerschnitt dar, der in anderen Experimenten zu 60.7 ± 2.4 *mb* [22] ermittelt wurde. Nach [23] beläuft sich der abgeschätzte Fehler der integrierten Luminosität auf derzeit 6.5% .

3.2.5 Das DØ-Triggersystem

Bei einer Rate von $2.5 \cdot 10^6$ Kollisionen pro Sekunde und einer einem Ereignis zugeordneten Datengröße von 250 *kB* liegt das Hauptproblem der Ereignisspeicherung im Datenvolumen, welches dementsprechend ein spezielles Auswahlverfahren verlangt. Ein Auswahlverfahren mittels eines Triggersystems ist von essentiellem Nutzen, da Ereignisse mit einer Kollisionsrate von 2.5 *MHz* durch eine Auswahl auf 50 *Hz*

[6] Der Strahl ist umgeben von einem Halo (griech. für Hof), in dem sich weitere Teilchen befinden können, welche aufgrund des abweichenden Sollorbits dem Strahl nicht zugerechnet werden dürfen.

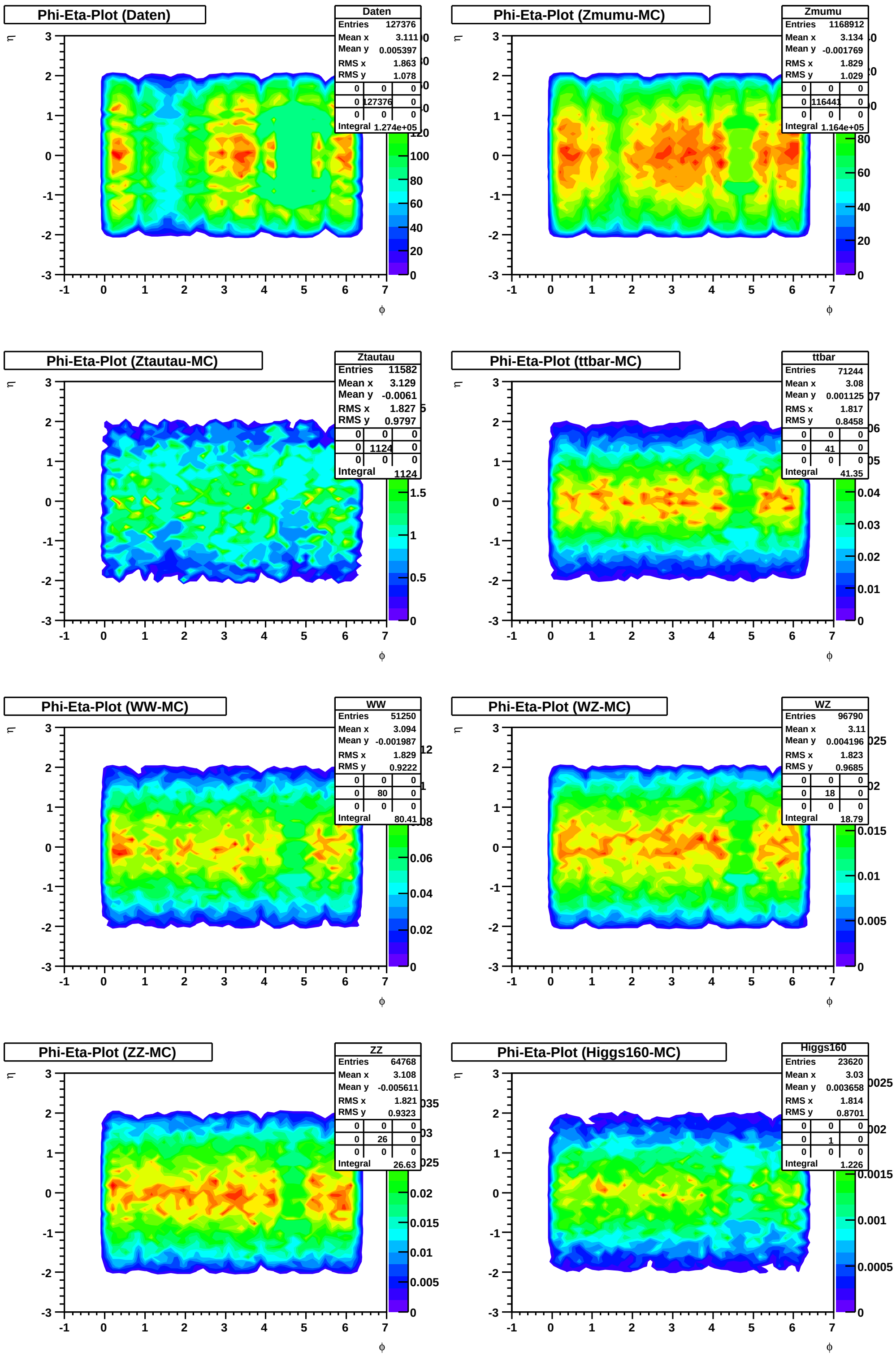

Abbildung 3.10: In dieser zweidimensionalen Darstellung sind für die in dieser Analyse verwendeten Monte-Carlo-Simulationen die ϕ- und die η-Koordinaten gegeneinander aufgetragen. Das linke obere Bild zeigt die Verteilung für die Daten, darunter sind im einzelnen die Untergrundprozesse dargestellt. Der Bereich, in dem nur eine geringe technische Auslese möglich ist ($|\eta| < 1.0$ und $4.25 < |\phi| < 5.15$), hebt sich mit weniger Einträgen deutlich ab.

reduziert werden müssen. Diese Rate kann letztendlich auf Magnetband zu weiteren Analysezwecken geschrieben werden.
Der DØ-Detektor ist dafür mit einem dreistufigen Ereignistriggersystem ausgestattet, welche nach identifizierten Teilchen oder Spuren selektieren. Die Steigerung in den Auswahlkriterien (zur genaueren Rekonstruktion) ist somit selbstverständlich und maßgebend. Die nächsten Abschnitte beschreiben die einzelnen Stufen (Level) des Triggersystems genauer, für zusätzliche Informationen sei [19] empfohlen.

Level-0-Trigger (L0)

In Kapitel 3.2.4 wurden bereits die Luminositätsmonitore vorgestellt, über welche die instantane Luminosität unter Berücksichtigung der $p\bar{p}$-Kollisionen ermittelt wird. Signale im Luminositätsdetektor sind nicht zwingend, aber weitestgehend grundlegend für eine Entscheidung der nächstfolgenden Trigger.

Level-1-Trigger (L1)

Level-1-Trigger beziehen ihre Informationen aus allen Detektorkomponenten jedoch nicht aus dem *Silicon Microstrip Tracker*. Sie sind Hard- und Firmwaresysteme, spezialisiert auf schnelle Auslese mit jedoch nur groben Messungen, welche die Eingangsrate von 2.5 MHz auf ca. 2.0 kHz reduzieren. L1 setzt sich aus dem sogenannten *L1 Trigger Framework* und diversen Untersystemen zusammen, die mit den jeweiligen Detektorkomponenten verkabelt sind. Das wären zum einen der L1-Spurtrigger, der eine ungefähre Impulsmessung liefert, der L1-Myontrigger, dessen Arbeit mindestens zwei Hits im Szintillator zugrunde liegen und zum anderen der L1-Kalorimetertrigger für Elektronen, Photonen und Jets aus Hadronen, mit dem eine Transversalenergieabschätzung möglich ist.

Level-2-Trigger (L2)

Nach voriger schneller Auslese wird diese in Kombination mit Informationen aus allen Subdetektorelementen betrachtet, so daß erstmalig ein Ereignis vollständig bewertet werden kann. Alle L2-Trigger erlangen ihre Erkenntnisse mit Hilfe von Firm- sowie Software. L2-Spurtrigger, bereitgestellt durch den CFT, knüpfen zusätzliche Bedingungen an die Größe der Impulse. Auch L2-Kalorimetertrigger sowie L2-Preshower-Trigger haben zum Ziel, die Energiemessung zu konkretisieren. L2-Myontrigger arbeiten auf dieser Ebene neben Szintillator- zusätzlich noch mit Driftröhreninformationen zur annähernden Impulsbestimmung. Im Anschluß werden die jeweiligen Daten der einzelnen Komponenten an ein zentrales System weitergereicht, dort kombiniert und innerhalb kürzester Zeit (100 μs) die Ereignisrate auf knapp 1.0 kHz herunter beschränkt.

Level-3-Trigger (L3)

Dieser Trigger ist ausschließlich softwarebasiert und kann somit die aus dem gesamten Detektor verfügbaren Erkenntnisse nutzen. Das Softwareprogramm stellt schon vorab eine einfachere Version der nachträglich offline angewendeten Rekonstruktionssoftware dar, in der nur Ereignisse, die bereits auf L1- und L2-Ebene bestanden haben, für Rekonstruktionszwecke verwertbar sind. Letztmalig wird die Ereignisrate auf 50 Hz reduziert. Auf Linux-Rechnern angewendet, werden zur Ereignisrekonstruktion lediglich 100 ms pro Event aufgewendet.

3.2.6 Die Myontrigger

Die in dieser Arbeit vorgestellte Analyse ist an eine Vielzahl von Ereignissen mit Myonen geknüpft. Das bedeutet, daß die verwendeten Daten einem dafür angepassten Myontriggersystem genügen müssen. Im konkreten heißt dies, Myonhits auf Trigger-Level 1 und 2 zu finden sowie Ereignisspuren mit hohem Transversalimpuls auf Trigger-Level 3.

Level-1-Myonen

Der hardwarebasierte Level-1-Myontrigger (DMU1) benutzt Szintillatorhits im Myondetektor, um seine Level-1-Objekte zu bilden und diese an die nächste Stufe weiterzureichen. Genauer gesagt, besteht ein Ansprechen des Triggers, sofern Koinzidenzen zweier Lagen derselben Region und desselben Oktanten vorhanden sind. Desweiteren wird die Szintillatorzeit genau so kalibriert, daß Myonen aus $p\bar{p}$-Kollisionen diesen zu einem Zeitpunkt $t = 0$ erreichen. Um dem Trigger zu genügen, werden deshalb nur Hits mit $|t| < 20$ ns erlaubt. Es werden lediglich Ereignisse mit zwei derartigen Myonen benötigt und dementsprechend getriggert. Teilchen, die diesem Kriterium nicht entsprechen, werden den *out-of-time*-Myonen zugeordnet, die mit großer Wahrscheinlichkeit der kosmischen Höhenstrahlung entstammen.

Level-2-Myonen

Das Trigger-Level 2 für Myonen stellt die komplette Zeitinformation und Kalibrationskonstanten der Proportionaldrahtkammern sowie des Szintillators zur Verfügung. Die Datensammlung ergibt sich einerseits direkt aus der Myondetektorauslesetechnik und andererseits aus den Level 1 Informationen, die zusätzlich Aussagen des *Central Fiber Trackers* beinhalten. Myonen entsprechend dieser Analyse passieren Level 2, sofern sie dem Qualitätskriterium 'medium' genügen. Die an diese Qualität gestellten Ansprüche sind wie folgt definiert:

- mindestens zwei Treffer in den MDT's, sowie mindestens ein Hit im Szintillator der A-Lage,
- mindestens zwei Treffer in der B- oder C-Lage,
- mindestens drei Treffer in den PDT's des zentralen Myonsystems mit einem in sogenannten *Look-up*-Tafeln für A- und BC-Lage gültigem Muster.

Level-3-Myonen

Mit voller Detektor- und Subdetektorinformation werden Level-3-Myonen softwarebasiert getriggert. Bedingt durch die steigende instantane Luminosität des Beschleunigers erweist sich der Trigger als unumgänglich. Auf die bereits vorgestellten Level-1- und Level-2-Myonen wurde bereits für solche Luminositäten ($40 - 50 \cdot 10^{30}\ cm^{-2}s^{-1}$) ein *pre-scaling* angewendet. Um nicht zu viele Ereignisse zu verlieren und Konflikte mit dem *pre-scaling* zu vermeiden, wird eine Spur mit hohem Transversalimpuls pro Event verlangt, welche auch von bereits getriggerten Myonen abgedeckt werden kann. Der Spuralgorithmus beginnt seine Arbeit mit der Suche nach axialen CFT-Tracks, die zu SMT- und CFT-Stereoclustern extrapoliert werden.
Diese Analyse verwendet letztendlich Ereignisse mit je einer globalen Spur mit einem p_T von 5 GeV oder 10 GeV abhängig von der jeweiligen Datennahmeperiode.

Kapitel 4

Rekonstruktion von Ereignissen

In diesem Kapitel sollen Mechanismen zur Identifikation und Rekonstruktion der für diese Analyse wichtigen Komponenten aufgezeigt werden. Darüber hinaus kann mit zusätzlichen Korrekturen das Übereinstimmen von Simulation und Daten verbessert und generell möglich gemacht werden.

4.1 Myonen

Myonen werden identifiziert, sofern Spuren im Spursystem des Detektors mit Spuren im Myonsystem in Bezug gesetzt werden können. Weiter muß der Impuls des Myons zu einer zentralen Spur gehören. Aufgrund ihrer unterschiedlichen Nachweiskriterien ist eine dementsprechende Charakterisierung sinnvoll. Diese Kategorisierung erfolgt in drei Stufen: *loose*, *medium* und *tight*.
Weist die Drahtkammer minimal zwei Treffer in der A- und drei Treffer in der BC-Lage auf, so sprechen wir von *tight* Myonen. Dieses Kriterium allein ist jedoch nicht bindend, vielmehr muß neben der Drahtkammer auch der Szintillator wenigstens einen Treffer in der A- und einen Treffer in der BC-Lage registriert haben, so daß letztenendes ein konvergierender Fit durch die Lagen des Myondetektors erstellt werden kann. Bei nur zwei Treffern in der BC-Lage der Drahtkammer und nur einem möglichen konvergierenden Fit, zu sonst gleichen Verhältnissen zur Kategorie *tight*, spricht man von *medium* Myonen. Schlägt nur einer dieser Punkte fehl, so handelt es sich um ein *loose* Myon, solange immerhin noch eine Spur im Szintillator nachzuweisen ist. Teilchen sind auch dann der Kategorie *loose* zuzuordnen, wenn Hits aus der BC-Lage des Szintillators (mindestens einer) sowie der BC-Lage der Drahtkammer (mindestens zwei) und eine der zentralen Spurkammer zuweisbare Spur nachvollzogen werden können.
Softwarepakete wie `muo_evt` $\mathtt{p14-br-05}$ und `muonid` $\mathtt{p14-br-14}$ werden zur Rekonstruktion der Myonen heran gezogen, siehe dazu [24].

4.2 Jets

Bei Kollisionen von Protonen und Antiprotonen wird, da beide zu den Hadronen gehören, hohe Jetaktivität erwartet. Jets entstehen durch Hadronisierung von Quarks und Gluonen. Zur Rekonstruktion der Jets werden die in den Kalorimeterzellen deponierte Energie mit Spurinformationen unter Zuhilfenahme des *Cone-*

Algorithmus[1] verglichen. Dieser faßt alle Teilchen innerhalb eines Konus mit Radius $R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.5$ zu einem Jet zusammen. Nachstehende Qualitätskriterien müssen erfüllt werden:

- $0.05 < EM\ fraction < 0.95$
  Der Anteil der Jetenergie im elektromagnetischem Teil des Kalorimeters soll in oben genannter Spanne liegen, um Elektron- und Myonjets auszuschließen.

- $Coarse\ Hadronic\ fraction\ (CHF) < 0.4$
  Der Wert des im grob-hadronischen Teil des Kalorimeters gemessenen Energiebruchteils darf einen Wert von 0.4 nicht überschreiten, damit ein einzelnes Myon nicht als Jet fehlinterpretiert wird.

- $Hot\ fraction < 10.0$
  Der Quotient aus energiereichstem und nächstenergiereichstem Turm (tower) des Jets im Kalorimeter soll kleiner 10 bleiben, um zu verhindern, daß ein einzelnes Photon oder Elektron als Jet klassifiziert wird.

- $N_{90} > 1$
  Die Anzahl der pseudoprojektiven Türme des Kalorimeters soll größer 1 sein. Mindestens 90% der Jetenergie sollen sich dabei in diesen Türmen befinden.

- $CHF < 0.1$ oder $f\_90 < 0.5$
  Mit dem Quotienten $f\_90 = \frac{N_{90}}{n_{itm}}$, wobei n_{itm} die Anzahl der Türme in einem Jet wiedergibt, soll die 0.5-Grenze nicht überschritten werden.

- $|\eta| < 2.5$
  Ein Raumwinkel dieses Bereiches ermöglicht eine präzisere Vermessung der Jetenergie innerhalb des Detektors.

***Jet Energy Scale* zur Korrektur der Jetenergie**

Vergleicht man die gemessene Jetenergie und die Jetenergie auf Teilchenebene miteinander, so werden Diskrepanzen offensichtlich. Gründe hierfür sind unter anderem der Algorithmus selbst, Collidereffekte und Unsicherheiten in den Messungen[2] im Kalorimeter zuzuschreiben. Zusätzliche Proton-Antiproton-Paare im System können weitere Wechselwirkungen ermöglichen, die dann wiederum die Energiemessungen im Kalorimeter verfälschen (*pile-up*-Effekt).
Die Jetenergie der Teilchen $E_{Jet}^{Teilchen}$ ermittelt sich aus der gemessenen Jetenergie E_{Jet}^{Kal} wie folgt:

$$E_{Jet}^{Teilchen} = \frac{E_{Jet}^{Kal} - E_{offset}}{R_{Jet} \cdot F_S} \tag{4.1}$$

Die Energie-Offset-Korrektur, dargestellt durch E_{offset}, ist abhängig vom Radius R des Konus, der Pseudorapidität η und der instantanen Luminosität. Sie wird bestimmt

[1] Der *Cone-Algorithmus* schließt lediglich Teilchen innerhalb $R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ ein, ohne grundsätzliche Beachtung solcher Teilchen, die auf oder wenig außerhalb des Radius liegen.

[2] Das System kann nicht alle Energiedeponierungen detektieren.

durch *underlying event*[3], *pile-up* und dem durch das Absorbermaterial des Kalorimeters vorhandenen Untergrund. Der deutlichst eingehende Beitrag zur Korrektur der Jetenergie wird durch die *Jet Response* R_{Jet} geliefert. 'Tote' Kalorimeterbereiche sowie nicht ausgelesene Regionen beschränken diese auf Werte < 1.0 . Durch F_S, der Schauerkorrektur, werden auch die Jet-Teilchenspuren berücksichtigt, die aufgrund von Wechselwirkung mit Detektormaterial oder durch das auf sie wirkende Magnetfeld eine Ablenkung aus ihrem Konus erfahren haben und somit für die eigentliche Messung der Jetenergie ausgeschlossen wären.
Mit Hilfe des Softwarepaketes `jetcorr` `v05 − 00 − 00` wird die Skalierung der Jetenergie vorgenommen [25].

4.3 Fehlende Transersalenergie $\not{E}_T$ und fehlender Transversalimpuls $\not{p}_T$

Die fehlende Transversalenergie $\not{E}_T$ im Detektor errechnet sich unter Verwendung der in den Kalorimeterzellen deponierten Energie. Sie ist um den Vertex herum, ihren Mutterteilchen entsprechend, verteilt und nur der Betrag des Leptons kann nachgewiesen werden. Der dazugehörige fehlende Betrag an transversaler Energie $\not{E}_T$ muß den schwer detektierbaren Neutrinos zugeordnet werden. Dabei sind Energiekorrekturen von Jets allgemein, Elektronenergien innerhalb von Jets sowie die auf Myonen anzuwendenden Korrekturen zu berücksichtigen. Die Berechnung erfolgt unter Ausschluß der grobhadronischen Lage, die eine Auflösung des $\not{p}_T$ verschlechtern würde.
Myonen, die unabhängig von ihrem Impuls lediglich um die 2.5 *GeV* Energie im Kalorimeter zurück lassen, erzwingen eine Korrektur des fehlenden Transversalimpulses durch Subtraktion des in der Spurkammer vermessenen p_T des Myons. Daraufhin kann der eigentliche Energieverlust, erhalten aus Parametrisierung unter Anwendung einer aus Teststrahldaten gewonnenen empirischen Funktion, addiert werden, sofern der Winkel berücksichtigt bleibt. Myonen, die dieser Berechnung stand halten müssen, haben mindestens *loose* Qualität, eine vollständig rekonstruierbare Spur und mindestens drei Treffer in den Siliziumkammern, deren maximaler Abstand zum Primärvertex 0.15 *cm* nicht überschreiten sollte.
Der fehlende Transversalimpuls $\not{p}_T$ wird mit Hilfe des Softwarepaketes `missingET` Version `v00 − 00 − 06` errechnet, Für die daran anschließende Korrektur kommt `d0correct` Version `v00 − 00 − 06` zum Einsatz. Weiterführende Informationen werden in [26] gegeben.

4.4 Monte-Carlo-Ereignissimulation

Analysen wie diese wollen Vorhersagen der Theorie mit experimentellen Daten vergleichen. Dabei sind als Hilfsmittel sogenannte Monte-Carlo-Ereignissimulationen sehr hilfreich. Zu ihrer Erzeugung sind drei Prozessierungsschritte notwendig.
An deren Anfang steht die Simulation von Proton-Antiproton-Kollision auf Partonniveau durch Ereignisgeneratoren[4] wie PYTHIA, die es ermöglichen eine Prozessaus-

[3] *Underlying event* ist die Wechselwirkung weiterer Partonen.
[4] Ereignisgeneratoren wie PYTHIA arbeiten auf der Basis von Zufallszahlengeneratoren.

wahl zu treffen. PYTHIA bestimmt hierbei für die zu erzeugenden Signal- und Untergrundereignisse die Matrixelemente in führender Ordnung von α_s, der starken Kopplungskonstante.
Im nächsten Abschnitt müssen zusätzlich noch Schauerkorrekturen (höherer Ordnung) für Gluon- als auch Photonabstrahlungen simuliert werden. Danach erfolgt die Hadronisierung der in diesen Prozessen entstandenen Quarks und Gluonen. Zum Schluß der Generierung steht in Form von `GEANT` eine Modellierung des DØ-Detektors, welche die hier simulierten Teilchen durchlaufen müssen, um am Ende auf einer Ebene vergleichbar mit den dem Experiment entstammenden Daten zu sein. `GEANT` berücksichtigt Effekte des magnetischen Feldes, Wechselwirkungen des Detektormaterials und Teilchenzerfälle ähnlich den im Detektor vorherrschenden Bedingungen. Grundsätzlich werden zu jedem Simulationsereignis 0.8 Proton-Antiproton-Kollisionen hinzugezählt, um zu berücksichtigen, daß bei der Strahlkreuzung durchaus mehr als nur diese eine Wechselwirkung stattfinden kann[5].
Zur möglichst realen Darstellung der Monte-Carlo-Simulationen wird im letzten Abschnitt das Ansprechverhalten des Detektors den Erwartungen entsprechend simuliert. Mit Hilfe digitalisierter Signale kann dieselbe Standardereignisrekonstruktion wie auch für Daten verwendet werden.
Um entsprechend genommene Daten mit Simulationsereignissen vergleichen zu können, ist es wichtig, sowohl Signal als auch Untergrund mit Hilfe eines PYTHIA-Ereignisgenerators zu produzieren[6]. Dafür sind vergleichbare Detektorzustände Bedingung, die die volle `GEANT`-Detektorsimulation benötigen.
In Tabelle 4.1 sind die verschiedenen für diese Analyse relevante Prozesse angegeben, daneben die zu ihnen gehörigen Wirkungsquerschnitte (in zweiter Ordnung Störungstheorie[7]) mal Verzweigungsverhältnis im angegebenen Zerfallskanal sowie die Anzahl ihrer generierten Ereignisse.
Da QCD-Mehrjet-Prozesse einen beachtlich großen Wirkungsquerschnitt aufweisen und zudem ihre Produktion ungenügend im Monte-Carlo-Ereignisgenerator beschrieben ist, ist es nicht ohne weiteres möglich, eine genügende Anzahl an Ereignissen zu simulieren, um detailliert genug den nach der Hauptselektion verbleibenden Rest beschreiben zu können. Es ist daher wichtig, diesen aufgrund der Diskrepanz zwischen Daten und Monte-Carlo-Simulationen abzuschätzen.
Zur Vertiefung dieser Thematik sei auf [27], [28] und [29] verwiesen.

4.5 Daten

Die vorliegende Analyse arbeitet vergleichend mit Monte-Carlo-Ereignissimulationen sowie Daten des DØ-Detektors. Diese Daten dokumentieren Kollisionen, welche über einen Zeitraum von April 2002 bis Dezember 2005 aufgezeichnet wurden. Der vollständige Datensatz dieser Periode entspricht in etwa 1 fb^{-1} an Ereignissen (Rekonstruktionssoftware p17). Informationen, welche aus diesen rekonstruierten Daten genutzt werden sollen, liegen in komprimierter Form, den sogenannten Thumbnails

[5] Zusätzliche Kollisionen, wie hier beschrieben, sind unter dem Namen *pile-up* und *underlying events* bekannt.
[6] Die dabei verwendete Partondichtestrukturfunktion (PDF) in führender Ordnung ist die `CTEQ6L`.
[7] Die zweite Ordnung Störungstheorie bezieht sich im Fachjargon auf die *next to leading order* (NLO).

(TMBs) vor.

Prozess	$\sigma \times BR$ **[pb]**	**Gen. Ereignisse**
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (120 GeV)	$9.8665 \cdot 10^{-4}$	25 750
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (140 GeV)	$2.2365 \cdot 10^{-3}$	26 500
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (160 GeV)	$2.7029 \cdot 10^{-3}$	26 250
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (180 GeV)	$1.8910 \cdot 10^{-3}$	27 250
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	5535	302 250
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	458.0226	2 156 250
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	245.68	1 029 250
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	1.883	191 000
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.1607	56 000
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< M_{Z/γ^*} < ∞	0.0063	36 500
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	458.0226	2 231 500
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	245.68	726 000
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	1.883	205 000
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.1607	58 000
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< M_{Z/γ^*} < ∞	0.0063	39 000
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$	1.2319	375 750
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$	0.1191	309 250
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$	0.1664	204 000
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$	0.695	603 500

Tabelle 4.1: Die für diese Analyse wichtigen Monte-Carlo-Datensätze sind mit ihrem Wirkungsquerschnitt mal Verzweigungsverhältnis sowie der Anzahl der generierten Ereignisse in der Tabelle angegeben.

Insgesamt arbeitet die Analyse hinsichtlich des zu untersuchenden myonischen Zerfallskanals mit zwei Datensätzen (dem `2MUhighpt`- und dem `MUinclusive`-skim), welche unterschiedlichen Aufzeichnungszeiträumen entsprechen. Diese stellen bereits eine erste Vorselektion der Daten dar, da an die Ereignisse in den skims Selektionsbedingungen gebunden sind (mindestens zwei *loose* Myonen, sowie zentrale Spuren mit mindestens 10 GeV). Die in den skims gespeicherten Daten dieser Analyse entsprechen einer integrierten Luminosität von 775 pb^{-1}, dies beeinhaltet nachfolgende Triggerbedingungen sowie sogenannte *good runs*. Ereignisse in den Datensätzen müssen zur weiteren Selektion hinsichtlich des untersuchten Zerfallskanals eine Reihe von Trigger durchlaufen. Auf Level 1 werden deshalb im gesamten Myonsystem ($|\eta| < 2.0$) mindestens zwei Szintillatorhits erwartet. Level-2-Myontrigger verlangen zusätzlich noch mindestens ein *medium* Myon, darüberhinaus werden auf Level 3 Ereignisse benötigt, die ein *loose* Myon mit zentraler Spur und einem $p_T > 6.0$ GeV aufweisen können (`DMU1_2LM6`). Andere Trigger (z.B. `DMU1_LM6_TK12`) verlangen auf Level 3 zusätzlich noch SMT- und CFT-Spuren

mit $p_T > 12.0\ GeV$. Darüberhinaus sind weitere Trigger im Einsatz. Der Unterschied zu den hier beschriebenen Triggern liegt lediglich in den Werten für die Transversalimpulse. Insgesamt sind alle Trigger, die in einer Datenperiode im Einsatz sind, über ein logisches ODER verknüpft.
Außerdem wurden alle problematischen Run-Nummern und Luminositätsblocks entfernt. Damit wird die Qualität der Daten gesichert und Detektor- und Rekonstruktionsfehler berücksichtigt, da die schlecht klassifizierten Runs bezüglich des Myonsystems, des Spurrekonstruktionssystems und des Kalorimeters ausgeschlossen werden.

4.6 Simulationsdatensätze

Die in dieser Analyse verwendeten Monte-Carlo-Simulationen sind in Tabelle 4.1 angegeben. Sie enthält die für den jeweiligen Prozess generierten Ereignisse sowie deren Wirkungsquerschnitte. Die Wirkungquerschnitte sind in Bezug auf die Zerfallskanäle mit den entsprechenden Verzweigungsverhältnissen gewichtet. Zum weiteren Vergleich mit Daten müssen die Monte-Carlo-Ereignissimulationen auf eine dem Datensatz entsprechende Luminosität normiert werden. Dieser Normierungsfaktor ergibt sich aus dem Quotienten des Wirkungsquerschnittes in Tabelle 4.1 und den angegebenen generierten Ereignissen, der Luminosität des Datensatzes sowie eines Effizienzenfaktors ε.
Für den Prozess $Z/\gamma^* \longrightarrow \mu^+\mu^-$, $5\ GeV < M_{Z/\gamma^*} < 15\ GeV$ ist noch zu beachten, daß die oben angegeben generierten Ereignisse dem Massenbereich entsprechen, jedoch auf Analyseebene noch mit einer Systemeffizienz (der `d0_mess` Software) zu gewichten sind. Monte-Carlo-Ereignisse dieser Kategorie werden demnach mit einem Faktor $1\,127/10\,000$ [30] neu gewichtet. Dieser Faktor ist anzuwenden, da bereits in der Prozessierung ein Schnitt auf den Transversalimpuls der Myonen mit $\not{p}_T > 3\ GeV$ verlangt wird.
Zur Anpassung der Monte-Carlo-Simulationen an die Daten müssen diese verschmiert werden. Verschmierungen werden allgemein innerhalb der Z-Resonanz anhand der $\frac{1}{p_T}$-Werte vorgenommen. Ziel ist es, mittels gaußischer Fits die Breite der Z-Resonanz der Monte-Carlo-Simulation der Breite der Z-Resonanz in den Daten anzugleichen, da auf Detektorebene die Simulation der Auflösung mittels `GEANT` zu optimistisch ausfällt. Dabei wird der Transversalimpuls jedes einzelnen Ereignisses durch

$$\frac{1}{p_T} \rightarrow \frac{1}{p_T} + C \cdot gauss(0., 1.)$$

ersetzt und neu angepaßt. In dieser Ersetzung beschreibt C eine Konstante, welche die Gauß-Verteilung schmaler oder breiter werden läßt. Für alle in dieser Analyse verwendeten Monte-Carlo-Simulationen wurde der Myontransversalimpuls bereits nach der Prozessierung verschmiert.
Um speziell die Z-Monte-Carlo-Simulationen genügend genau dem Datensatz anpassen zu können, welche im Ereignisgenerator (PYTHIA) unzureichend simuliert sind, ist eine Verschmierung allein nicht ausreichend. Hier wird zusätzlich das Verfahren der Neugewichtung (*reweighting*) angewendet. Anders als bei der Verschmierung, in der durch einen Zusatzfaktor dem Transversalimpuls ein neuer Wert gegeben wird, erfolgt hier die Korrektur der p_T der Z-Monte-Carlo-Simulation lediglich durch eine “Umschichtung” der Ereignisse, d.h. Überschüsse werden an Unterschüssen angeglichen.

Kapitel 5

Vorselektion

Die vorliegende Arbeit sucht nach Higgs-Zerfällen $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$. Der Schwerpunkt der Analyse liegt auf der Untersuchung des Prozesses für Higgs-Bosonen mit einer Masse von $m_{\mathrm{H}} = 160\ GeV$. Zerfälle dieser Topologie sind gekennzeichnet durch Ereignisse mit zwei Myonen, deren Transversalimpulse entsprechend der Higgsmasse relativ hoch sind. Zur Erklärung: Das Higgs-Boson mit 160 *GeV* zerfällt zunächst in ein W-Boson-Paar, deren Massen bei 80 *GeV* liegen. Danach zerfällt jedes W-Boson in ein Myon und sein Myon-Neutrino zerfallen. Dabei bleiben für beide Teilchen rund 40 *GeV* kinetischer Energie übrig. Analysen mit Higgs-Boson-Massen unter 160 *GeV* müssen berücksichtigen, daß kinematisch es immer schwieriger wird, reelle W-Boson-Paare aus Zerfällen zu erzeugen, da diese Prozesse zunehmend unterdrückt sind. D.h. die Wahrscheinlichkeit für diesen Zerfallskanal nimmt ab, ist jedoch nicht ganz auszuschließen. Es ist zu beachten, daß in diesem Fall, entsprechend der Masse des Higgs-Bosons, auch die Transversalimpuls-Verteilungen der Myonen sich zu kleineren Werten verschieben.
Weiter sind Prozesse der Art $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ bestimmt durch hohe fehlende Transversalenergie. Diese wird den im Zerfall produzierten Myon-Neutrinos zugeschrieben, für die, wie bereits erwähnt, eine Detektierung fast ausgeschlossen ist. Diese Ereignissignaturen gilt es im Vergleich zum Untergrund hervorzuheben und für eine Signal-zu-Untergrund-Diskriminierung effektiv zu nutzen. Als besondere Schwierigkeit sind hier Untergrundprozesse wie $Z/\gamma^* \longrightarrow \mu^+\mu^-$ und $Z/\gamma^* \longrightarrow \tau^+\tau^-$ zu nennen sowie $W^+W^- \longrightarrow \mu^+\nu\mu^-\bar{\nu}$, der sich als irreduzibel erweisen wird. Bevor jedoch die eigentliche Selektion mit Hilfe von Neuronalen Netzen erfolgen kann (Kapitel 6), müssen zu Beginn in der Vorselektion bereits einige Qualitätskriterien überprüft und festgehalten werden.

5.1 Myon-Qualitätsschnitte

Da diese Analyse den Zerfall in den myonischen Endzustand untersucht, ist es wichtig, Qualitätskriterien für die in diesem Kanal erzeugten Myonenspuren zu verlangen. So spielen beispielsweise Treffer und Spuren im zentralen Spur- wie auch Myonsystem eine Rolle. Ferner muß ein Ausschluß kosmischer Höhenstrahlung entstammender Myonen gewährleistet werden.

5.1.1 Nachweis in den Spursystemen

Der zu untersuchende Zerfallskanal $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ weist in seinem Endzustand unterschiedlich energetische Myonen auf. Diese in den Spursystemen des Detektors nachzuweisen, heißt Treffer in den Lagen des inneren Spurdetektors, Treffern in den Lagen des äußeren Myondetektors zuzuordnen. Dabei dürfen die Nachweiskriterien für Myonen nicht zu "streng" ausfallen, da sonst eine Vielzahl an Ereignissen verworfen werden würde. Eine Kategorisierung der Nachweiskriterien ist in Kapitel 4.1 zu finden. Nach dieser Einteilung wird das "einfachste" Nachweiskriterium für Myonen gewählt. Das heißt, daß Myonen im Zerfall $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ die Myonqualität *loose* (siehe Abschnitt 4.1) besitzen müssen.
Weiter werden Ereignisse mit mindestens zwei Myonen selektiert, von denen zwei entgegengesetzte Ladungen aufweisen. Dies folgt aus der Bedingung, daß, wenn ein neutrales Higgs-Boson in ein W^+W^--Paar zerfällt, aus den W-Bosonen lediglich Myonen unterschiedlicher Ladung hervorgehen können. Ist dieser Fall nicht gegeben, so kann das Ereignis nicht dem zu untersuchenden Prozess zugeordnet werden.

5.1.2 Ausschluß kosmischer Myonen

Neben den im Detektor stattfindenden Ereignissen aus Proton-Antiproton-Kollisionen ist stets mit Myonen der kosmischen Höhenstrahlung zu rechnen. Diese Teilchen werden als Untergrund gezählt und müssen daher separiert und ausgeschlossen werden. Dazu wird in einer speziellen Abfrage geprüft, wie der Weg des Myons durch den Detektor verlaufen ist. Myonen, welche aus Proton-Antiproton-Kollisionen stammen, müssen Spuren nachweislich zum Primärvertex besitzen, da sie dem Wechselwirkungspunkt entstammen. Betrachtet man Myonen der kosmischen Höhenstrahlung, welche durch den Detektor hindurch treten, so hinterlassen diese Teilchen Spuren im gesamten Querschnitt des Detektors, welche möglicherweiseweise außerhalb des Primärvertex verlaufen. Der Effekt wird als Möglichkeit zum Ausschluß kosmischer Myonen genutzt. Zusätzlich darf der Abstand der Myonen vom Primärvertex nicht größer als 0.16 *cm* sein.
Da es auch vorkommen kann, daß Myonen der kosmischen Höhenstrahlung in den Detektor eintreten und Spuren durch den Primärvertex hinterlassen, ist für diesen Fall ein zweites Kriterium anzuwenden. Dieses überprüft die Szintillatorzeit der Myonen im Detektor. Myonen aus Kollisionen bewegen sich mit annähernd Lichtgeschwindigkeit vom Wechselwirkungspunkt aus und erreichen die Szintillatoren spätestens 20 *ns* nach der Kollision, während kosmische Myonen unkorreliert zu den Kollisionen den Detektor treffen, also i.a. eine Szintillatorzeit größer 20 *ns* haben, so daß sie durch $|t| < 20\ ns$ effektiv verworfen werden können.

5.1.3 Zentralität der Spur

Da die Impulsauflösung im Myonsystem sehr viel schlechter ist als in den zentralen Spursystemen, werden nur Ereignisse berücksichtigt, deren Treffer im Spursystem sich zu einer zentralen Spur zurückführen lassen. Diese Abfrage gewährleistet eine Myonselektion, welche auf Teilchen zurückgreift, deren Ursprung in einer Reaktion im oder zumindest sehr nahe des Wechselwirkungspunktes liegt. Die Verteilungen

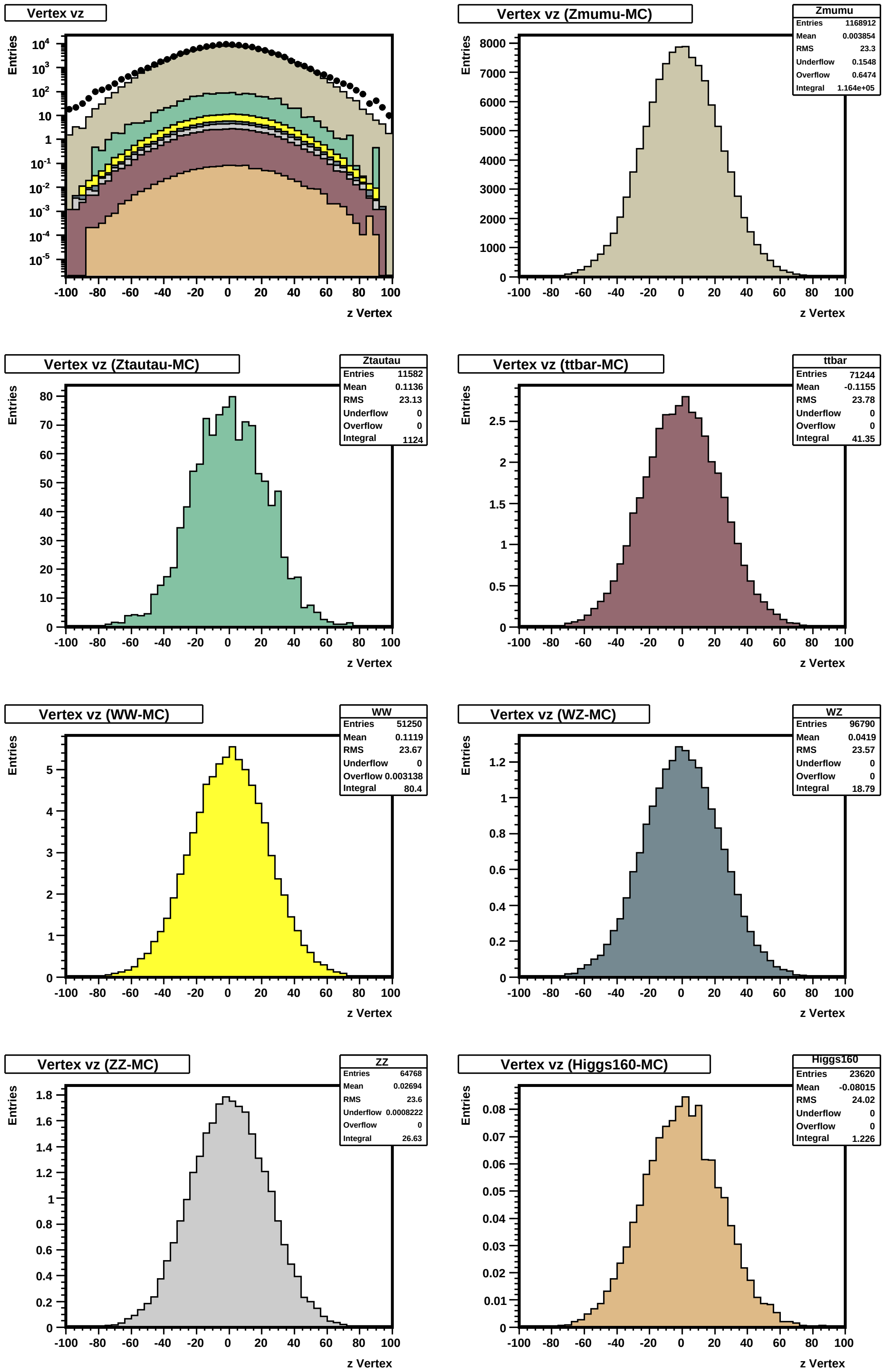

Abbildung 5.1: Darstellung der Verteilung der Ereignisse um den Vertex für Daten wie Untergrund gemeinsam (oben links) sowie für jeden Untergrundprozess getrennt. Die Abbildung bezieht sich nur auf die z-Komponente. Die x- und y-Komponente dieser Verteilung sind analog zu betrachten.

der Ereignisse um diesen Vertex sollten demnach symmetrischer Struktur sein. Die z-Komponente der Verteilung um den Vertex ist in Abbildung 5.1 dargestellt. Mit einer Abfrage nach der Zentralität der Teilchenspur kann die Impulsauflösung der Myonen entsprechend verbessert werden.

5.2 Spurisolationsschnitte

Weiter müssen auch Spurisolationsschnitte angewendet werden. Da neben den erwarteten Ereignissen sehr viele weitere möglich und sehr wahrscheinlich sind, besteht die Gefahr, Myonen aus anderen Zerfallsmechanismen als deren Sekundärprodukt nachzuweisen. Um sicher zu stellen, daß Myonen auch wirklich dem Higgs-Zerfall $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ entstammen, ist es wichtig einen Blick auf die Spuren dieser Myonen zu werfen. Sie sollten isoliert von Sekundärteilchen sein, um zu gewährleisten, daß es sich wirklich um ein primär erzeugtes Myon handelt. Stellt es sich heraus, daß weitere Spuren nachweisbar sind, muß davon ausgegangen werden, daß es sich um ein Myon aus dem Zerfall von Hadronen (hauptsächlich mit schweren Quarks) handelt, welches nicht dem zu untersuchenden Prozess zuzuordnen ist.

5.2.1 Transversale Energie im Kalorimeter-Halo

Um die transversale Energie im Kalorimeter-Halo bestimmen zu können, wird ein Hohlkegel mit $R_1 = 0.1$ und $R_2 = 0.4$ um die Myonspur gelegt. Dabei ist R folgendermaßen aus ϕ und η definiert:

$$R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \quad .$$

Die gemessene transversale Energie in diesem Hohlkegel um die gewählte Myonspur darf nicht größer als 2.5 *GeV* sein. Ein Schnitt dieser Größenordnung stellt ein Isolationskriterium der Myonspur dar und hilft diese Spur einem einzelnen Teilchen, dem Myon, zuzuordnen. Wie die Verteilungen nach einem solchen Schnitt aussehen, gibt Abbildung 5.2 wieder. Die Darstellung erfolgt für alle übereinander gelegten Prozesse, mit denen ein Abgleich auf den Datensatz möglich ist, sowie für Signal und Untergrund getrennt. Für erstere der eben genannten wurde eine logarithmische Skala gewählt.

5.2.2 Transversale Energie im Spur-Halo

Betrachtet man die transversale Energie im Spur-Halo, gibt auch diese eine Information zur Isolation der Myonspur wieder. Analog zu 5.2.1 kann mit dieser Größe eine Selektion auf Myonen mit Spurisolation erfolgen. Dazu wird in diesem Fall um die Myonspur ein Kegel mit $R = 0.5$ gelegt. Die Summe aller Transversalimpulse innerhalb des Konus, ausgenommen der Myonspur, soll in dieser Analyse die 2.5 *GeV*-Marke nicht überschreiten. Bei höheren Energiesummen besteht die Gefahr, daß Jetereignisse involviert sein könnten. Ähnlich den im letzten Abschnitt gezeigten Ergebnissen sind in Abbildung 5.3 die Verteilungen für den hier angewendeten Schnitt gezeigt. Für die Abbildung mit allen Verteilungen wurde eine logarithmische Skala gewählt.

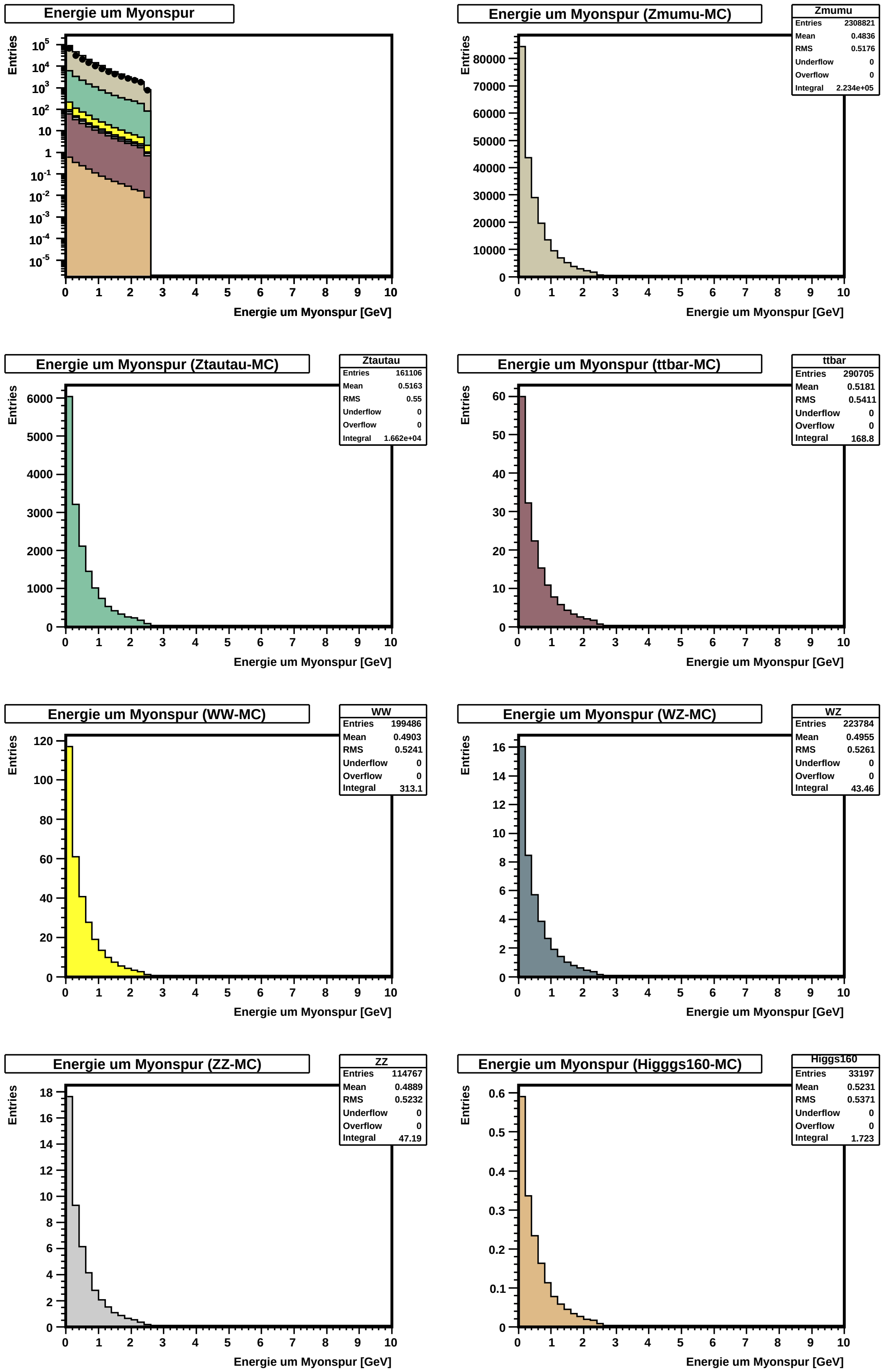

Abbildung 5.2: Darstellung der Verteilung der transversalen Energie im Kalorimeter-Halo nach Anwendung des Schnittes auf 2.5 *GeV* für Daten und Untergrung gemeinsam sowie getrennt.

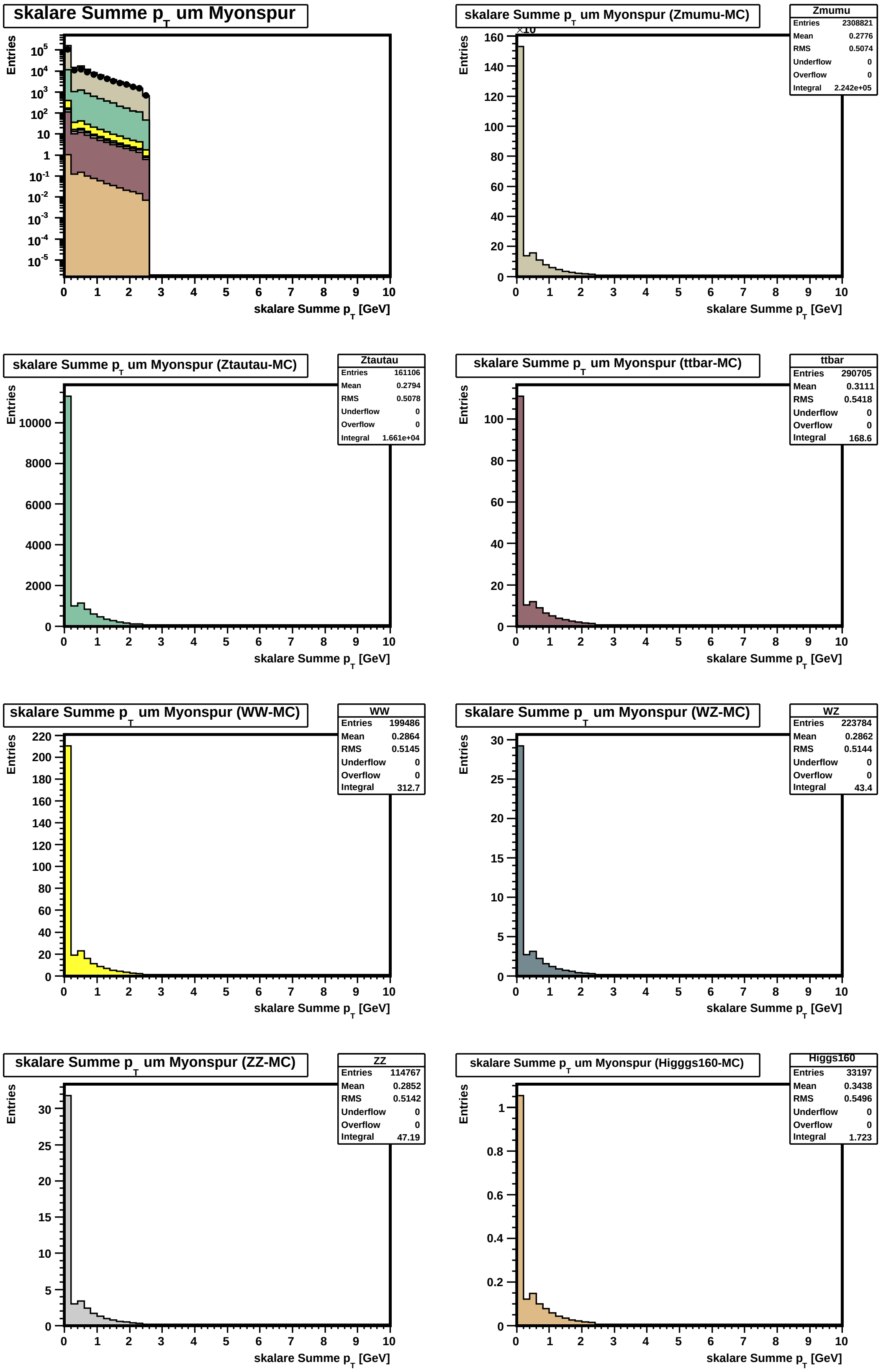

Abbildung 5.3: Darstellung der Verteilung der transversalen Energie im Spur-Halo nach Anwendung des Schnittes auf 2.5 *GeV*. Es werden wieder zum Vergeich des Signales mit Daten alle Untergrundereignisse aufsummiert, daneben aber auch ihre Einzeldarstellungen angegeben.

5.3 Schnitte für die Signal-zu-Untergrund-Diskriminierung

Mit möglichst einfachen Schnitten soll bereits in der Vorselektion eine erste Diskriminierung des Signals vom Untergrund erfolgen. Zu beachten ist, daß einerseits vom Signal nicht zuviel verworfen werden darf und andererseits die Schnitte so erfolgen, daß bereits hier ein Bruchteil an Untergrund absepariert werden kann.

5.3.1 Schnitte auf die Transversalimpulse

Ausgehend von einem 160 *GeV*-Higgs-Boson ist zu erwarten, daß die Impulse beider am Prozess beteiligten Myonen, der Zerfallskinematik entsprechend, relativ hoch sein müssen. Eine Verteilung der Transversalimpulse um die 40 *GeV* ist zu erwarten. Jedoch zeigt der überwiegende Teil der Untergrundprozesse, bei denen Zerfälle in Myonen angenommen werden, daß hier bereits sehr kleine Transversalimpulse möglich sind. Schnitte auf die Transversalimpulse beider Myonen sind demnach sinnvoll. Sie sind von besonderer Bedeutung für Untergrundprozesse, die Resonanzen bei sehr kleinen invarianten Massen aufzeigen, da hier die Transversalimpulse in einem kleinen p_T-Bereich dominieren. Folglich wirken sich Schnitte auf Transversalimpulse indirekt auch auf die Verteilung der invarianten Masse $M_{\mu\mu}$ (Abbildung 5.4) aus, da sich diese aus den Transversalimpulsen und den dazu gehörigen Energien berechnet. Die Schnitte müssen so gewählt werden, daß gerade diese Untergrundprozesse in sehr großem Umfang separiert werden können. Zu den hier besprochenen Prozessen zählen zum einen die Υ-Prozesse (Υ_{1s} und Υ_{2s}) sowie der Einfluß vom QCD-Mehrjet-Untergrund. Die Υ-Resonanzen befinden sich in einem Massenbereich um die 10 *GeV*. Der QCD-Mehrjet-Untergrund, welcher in Abb. 5.4 bei sehr kleinen Massen eingeordnet werden muß, liegt deutlich im Verteilungsbereich der invarianten Masse des Higgs-Signals, trotzdem sollte dennoch ein Großteil dieser Ereignisse verwerfbar sein. Darüberhinaus werden mit Hilfe der Schnitte Drell-Yan-2-Lepton-Ereignisse in $Z/\gamma^* \longrightarrow \mu^+\mu^-$ abesepariert. Sie sind damit für eine Separation vom Higgs-Signal-Prozess bedeutend günstiger als die weiteren Untergrund-Prozesse.
Im allgemeinen sind Zerfälle der Art $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ dadurch geprägt, daß das erste Myon energetischer ist als das ihm folgende Myon. Ein Schnitt auf den Transversalimpuls des ersten Myons kann dementsprechend höher ausfallen als ein Schnitt auf den Transversalimpuls des zweiten Myons. Die in dieser Analyse selektierten Ereignisse erfüllen die Bedingungen, daß

- der Transversalimpuls des führenden Myons größer 15 *GeV*

und

- der Transversalimpuls des folgenden Myons größer 10 *GeV*

sein soll.
Die Abbildungen 5.5 und 5.6 stellen die Verteilungen der Transversalimpulse beider Myonen dar. Zum Vergleich mit dem Datensatz wurden alle Monte-Carlo-Ereignissimulationen übereinander gelegt und logarithmisch aufgetragen. Daneben befinden sich auch die Verteilungen aller einzelnen Untergrundprozesse sowie des Signales.
Zum konsistenten Vergleich der hier verwendeten Monte-Carlo-Ereignissimulationen

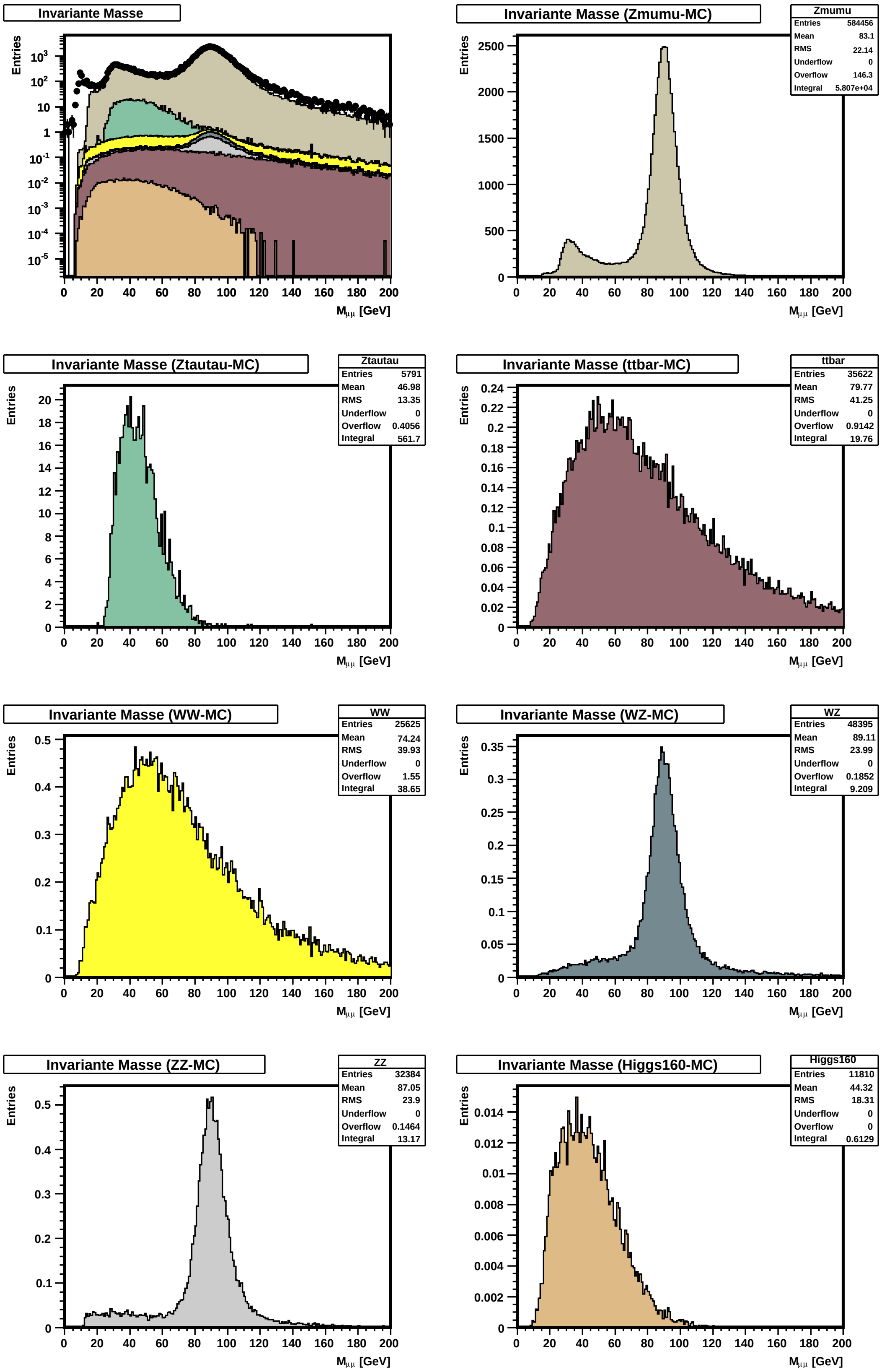

Abbildung 5.4: Die Abbildung zeigt die invarianten Massen für Daten (schwarze Punkte) und Monte Carlo (links oben, logarithmisch) wie auch für die einzelnen Monte Carlos. Deutlich erkennbar ist bei rund 90 GeV die Z^0-Resonanz. Resonanzen unterhalb des 10 GeV-Bereiches sind im Monte Carlo nicht mehr sichtbar, weil sie durch die p_T-Schnitte bereits absepariert wurden. Der Überschuß in den Daten im unteren Energiebereich zeigt die Υ-Resonanz. Das Monte-Carlo ist entsprechend der Daten-Luminosität normiert.

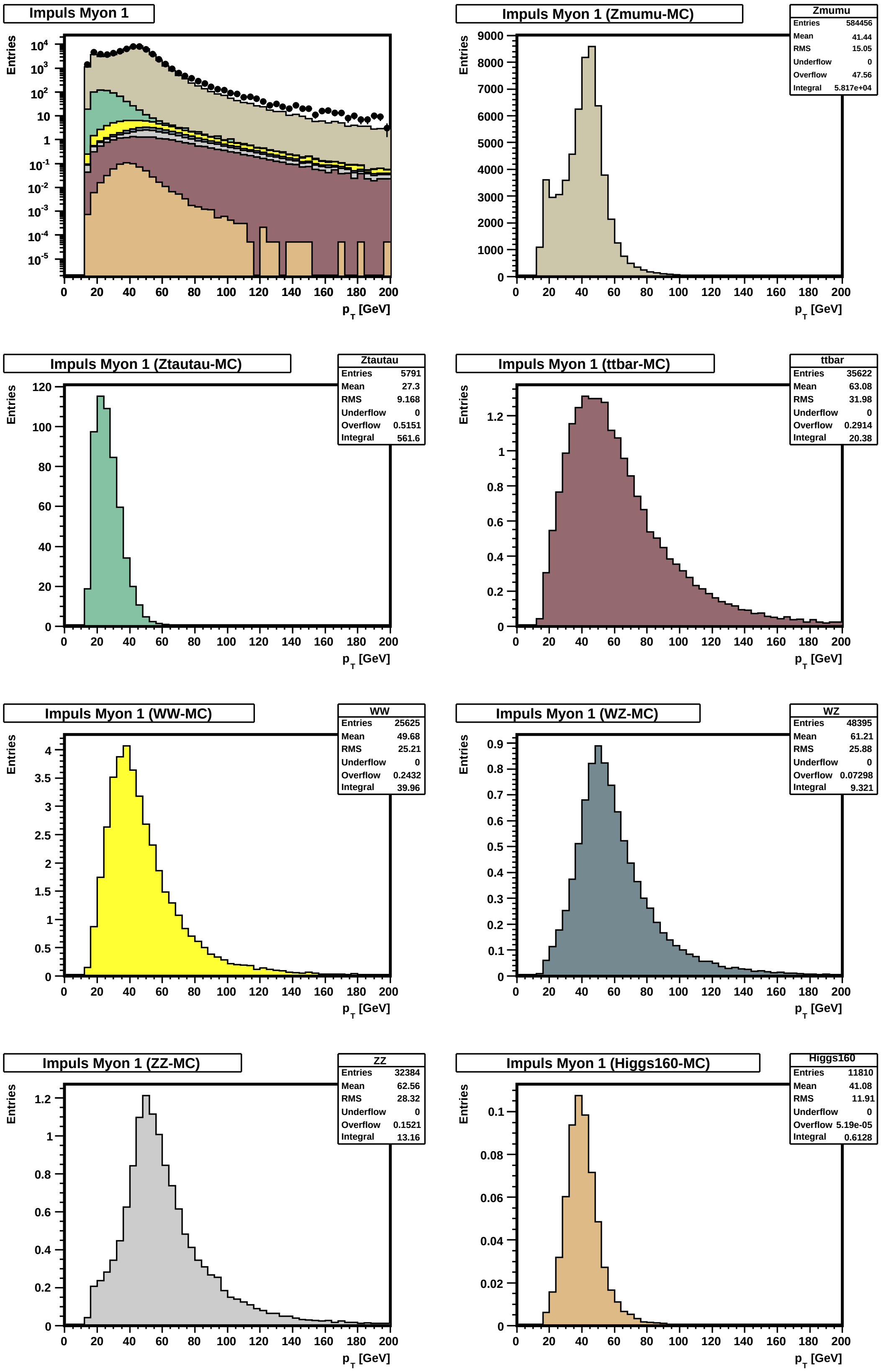

Abbildung 5.5: Die Abbildung zeigt die Verteilung der transversalen Impulse des führenden Myons für Daten und Monte Carlo (links oben) und die einzelnen Monte Carlos. Für einen Vergleich mit Daten wird eine logarithmische Skala gewählt (siehe links oben). Der Überschuß in den Daten zu hohen p_T-Werten kann durch die schlechtere Rekonstruktion der Ereignisse mit keinem Treffer im SMT erklärt werden.

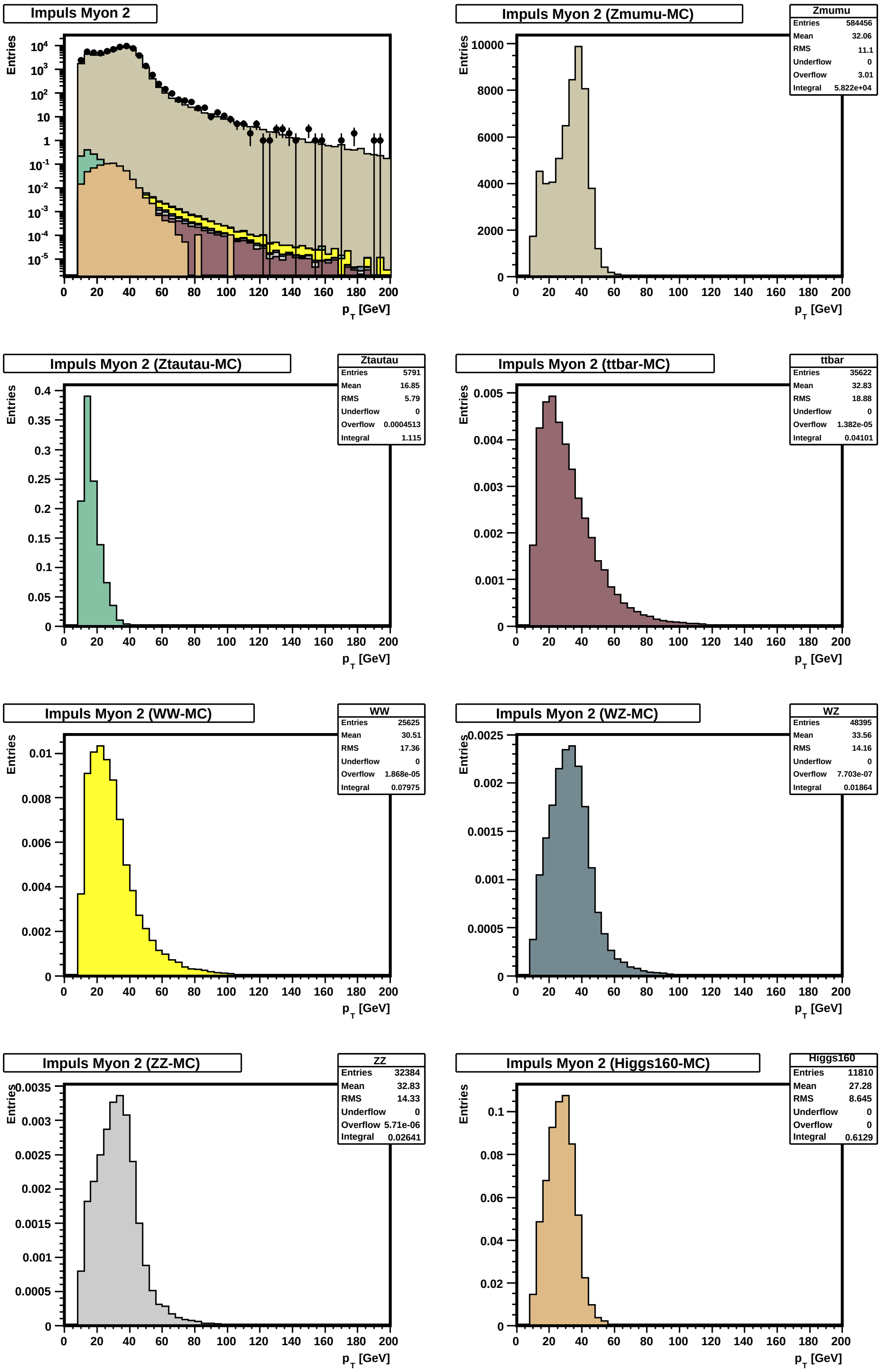

Abbildung 5.6: Die Abbildung zeigt die Verteilung der transversalen Impulse des zweiten Myons für Daten und Monte Carlo (links oben) und die einzelnen Monte Carlos. Auch hier wurde wieder für einen Datenvergleich eine logarithmische Skala gewählt (siehe erstes oberes Bild).

mit Daten, die selbst bereits über eine Vorauswahl selektiert werden, ist ein zusätzlicher Schnitt auf die Simulationsdatensätze notwendig. Der in der Vorselektion der Daten angewandte Schnitt erfolgt anhand des Transversalimpulses, welcher im Myon-Typ *Central* (zentrale Spuren müssen nachweisbar sein) festgehalten ist. Dementsprechend ist durch die Vorauswahl der Daten ein *Central*-Transversalimpuls größer 10.0 *GeV* für beide Myonen zu verlangen.

5.3.2 Schnitt auf die Pseudorapidität

Damit ein Myon überhaupt im Myondetektor wahrgenommen werden kann, muß es sich innerhalb eines bestimmten Raumwinkelbereiches bewegen. Daher ist es sinnvoll, Myonen aus diesem Raumwinkelbereich zu selektieren. Es wurde bereits erwähnt, daß Myonen rekonstruiert werden, indem ihre Treffer im Myonsystem mit Treffern und Spuren im zentralen Detektor zusammengeführt werden. Die Konstruktion des Detektors ist so ausgelegt, daß der *Central Fiber Tracker* (CFT) in einem Bereich $|\eta| < 1.6$ detektieren kann. Beim *Silicon Microstrip Tracker* (SMT) liegt dieser Bereich bei $|\eta| < 3.0$. Da Treffer und Spuren von Teilchen in diesen Detektorkomponenten, Treffern und Spuren im Myonsystem zugeordnet werden, sind diese mögliche Kandidaten zur Selektion. Konstruktionstechnisch deckt das Myonsystem einen Bereich von $|\eta| < 2.0$ ab, damit ist auch in der Vorselektion der Analyse ein Schnitt der gleichen Größe zu wählen.

5.4 Ergebnisse der Vorselektion

Zusammenfassend sind folgende Größen und Variablen der Vorselektion festzuhalten:

- Myonqualität *loose*
- zentrale Myonspuren, aus dem Primärvertex mit $p_T^{Central} > 10.0\ GeV$
- keine kosmischen Myonen
- isolierte Myonspuren
- mindestens zwei Myonen mit entgegengesetzter Ladung
- Impuls des ersten Myons mit $p_T^{\mu_1} > 15.0\ GeV$
- Impuls des zweiten Myons mit $p_T^{\mu_2} > 10.0\ GeV$
- Pseudorapidität $|\eta| < 2.0$

Wie sich die Schnitte dieser Größen auf Signal, Untergrund sowie Daten auswirken, ist in Tabelle 5.1 angegeben. Je nach Prozessart sind die nach der Vorselektion verbleibenden Ereignisse sowie deren statistische Fehler dokumentiert.

Prozess	**Verbleibende Ereignisse**
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (120 GeV)	0.176 ± 0.002
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (140 GeV)	0.476 ± 0.005
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (160 GeV)	0.613 ± 0.006
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (180 GeV)	0.442 ± 0.004
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	$1\,836 \pm 130$
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	$8\,208 \pm 30$
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	$49\,505 \pm 79$
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	460 ± 2
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	45.2 ± 0.3
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< $M_{Z/\gamma^*} < \infty$	1.62 ± 0.01
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	7.7 ± 0.9
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	543 ± 10
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	9.8 ± 0.2
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	1.32 ± 0.04
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< $M_{Z/\gamma^*} < \infty$	0.058 ± 0.002
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ (inkl.)	40.2 ± 0.2
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ (inkl.)	9.39 ± 0.04
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ (inkl.)	13.31 ± 0.08
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ (inkl.)	20.7 ± 0.1
$\sum_{Untergrund}$	$60\,702 \pm 156$
DATEN	63 688

Tabelle 5.1: In der Tabelle sind die nach der Vorselektion verbleibenden Ereignisse sowie ihre statistischen Fehler eingetragen. Diese sind bereits mit dem jeweiligen Faktor aus Luminosität (775 pb^{-1}), der Rekonstruktionseffizienz und dem Wirkungsquerschnitt zu generierten Ereignissen normiert.

Die Untergrundprozesse der Analyse wurden, entsprechend des verwendeten Datensatzes, auf eine Luminosität von 775 pb^{-1} normiert. Dieser Schritt ist für Aussagen zu Vergleichswerten notwendig.
Im direkten Vergleich des gesamten Untergrundes mit den Ereignissen in den Daten fällt auf, daß zwischen beiden eine Diskrepanz von rund 3000 Ereignissen liegt. Da die Unsicherheit auf die Luminosität zu 6.5% geschätzt wird, ist diese in einem Vergleich von Daten und Untergrund zu berücksichtigen. Zusätzlich muß noch erwähnt werden, daß nicht alle Monte-Carlo-Simulationen für die zu dieser Analyse beitragenden Untergrundprozesse aufgenommen wurden. Der wichtigste Anteil dieser Prozesse konnte mit Simulationsdatensätzen berücksichtigt werden, aber zur vollständigen Beschreibung des gesamten Datensatzes fehlt die Betrachtung von Υ-, QCD-Mehrjet- und

$W + jets$-Untergrund. Auf Υ wurde absichtlich verzichtet, da dieser in Studien eines kleineren Datensatzes als vollständig separabel eingestuft wurde, indem auch für diesen Untergrund ein Neuronales Netz herangezogen worden ist, mit welchem bereits bei sehr niedrigen Schnitten auf das Netz, Signal und Υ definitiv voneinander getrennt werden konnten. Da diese Analyse auf der Selektion mit Neuronalen Netzen basiert, ist es schwierig QCD-Mehrjet und $W + jets$ Untergrund einzubeziehen, da für beide Prozesse nicht ausreichende Monte-Carlo-Simulationen vorlagen. Sie erweisen sich jedoch laut [6] als vernachlässigbar nach Anwendung der Selektionskriterien.
Da die Υ-Resonanzen in einem sehr kleinen Massenbereich liegen (um die 10 GeV), sind in diesem Bereich und demnach auch bei kleinen Transversalimpulsen Abweichungen zwischen Daten und Untergrund zu erwarten. Obwohl der QCD-Mehrjet-Untergrund in dieser Analyse nicht explizit betrachtet werden kann, darf er vor einer Selektion nicht außer Acht gelassen werden. Deutliche Auswirkungen zeigen sich beispielsweise bei möglichen Schnitten auf die Transversalimpulse beider Myonen. Wird der Schnitt zu niedrig angesetzt, wäre der verbleibende Anteil an Ereignissen dieses Untergrundes zu hoch, da auch QCD-Mehrjet-Verteilungen in unteren Impuls- wie Massenregionen liegen. Folglich ist der in der vorliegenden Analyse angesetzte Wert auf p_T maßgeblich an der Reduzierung von QCD-Mehrjet-Untergrundereignissen beteiligt.
Darüberhinaus sind Unterschiede bei großen p_T- und $M_{\mu\mu}$-Werten sichtbar. Diese können aufgrund der Verschmierung erklärt werden. Eine Verschmierung wie in Kapitel 4.6 angegeben, ist erfolgt, aber nicht auf jede Datenperiode[1] optimiert. So wurde beispielsweise die letzte Datenperiode mit einem geringeren Magnetfeld aufgezeichnet [31].
In diesem Zusammenhang muß auch auf die Korrektur der Jetenergie eingegangen werden, welche nicht angewendet werden konnte, da diese zu der benutzten Release-Version noch nocht erhältlich war. Durch die nicht betrachtete *Jet Energy Scale* sind weitere Abweichungen zu erwarten.

[1] Im TeVatron kollidieren nicht ununterbrochen Protonen und Antiprotonen, vielmehr umfaßt der Datensatz mehrere Datenperioden, welche durch sogenannte *Shutdowns* voneinander getrennt sind, in denen Systemkomponenten neu kalibriert werden.

Kapitel 6

Die Neuronalen Netze

In der Forschung werden künstliche Neuronale Netze (*Artificial Neural Networks* kurz *ANNs*) zur Datenanalyse und zu Klassifizierungszwecken eingesetzt. Diese Neuronalen Netze funktionieren und arbeiten nach Vorlage des menschlichen Gehirns und bestehen aus einer Vielzahl kleiner Elemente, den sogenannten Neuronen. Eingehende Informationen werden verarbeitet, indem sich Neuronen mittels gerichteter Verbindungen untereinander aktivieren, analog zu Vorgängen im Gehirn. Die Lernfähigkeit der Neuronalen Netze ist ihr wichtigstes Merkmal, durch sie sind sie in der Lage, Aufgaben anhand von Trainingsbeispielen zu erlernen, ohne explizit programmiert werden zu müssen. In der hohen Parallelität während der Verarbeitung von Informationen, der hohen Fehlertoleranz und dem relativ kleinen Wissensausfall bei Zerstörung von Neuronen liegen weitere Charakteristika, die für eine Benutzung von Neuronalen Netzen sprechen. Ein Überblick über die Thematik sowie die spezielle Anwendung in der vorliegenden Analyse sollen die nächsten Abschnitte vermitteln. Für weiterführende Informationen sei auf [32] und [33] verwiesen.

6.1 Neuronen - Grundlage künstlicher Neuronaler Netze

Dem biologischen Vorbild folgend, sind künstliche Neuronale Netze aus stark idealisierten Neuronen aufgebaut, welche aus drei Komponenten bestehen: dem Zellkörper, dem Dendriten und einem Axon. Dendriten sind stark verästelte Ausläufer der Neuronen. Ihre Aufgabe ist es Eingangssignale aufzunehmen. Das Axon hingegen dient der Weiterleitung von Signalen an andere Zellen. Im Gehirn sind sie normalerweise im Vergleich zu den Dendriten die längere Nervenfaser, welche am Ende Verästelungen aufweisen und Synapsen[1] für die Reizweiterleitung besitzen. Im Falle des ANN summieren die Dendriten die Eingaben des Netzes in die Zelle auf, übergeben die Summe dem Axon, der wiederum das Signal an Dendriten nachfolgender Synapsen weiterleitet (siehe dazu Abbildung 6.1). Dabei wird die Stärke der Synapsen durch einen numerischen Wert, dem Verbindungsgewicht angegeben. Eine direkte gewichtete Verbindung zwischen den Zellen i und j ist folglich darstellbar.

[1]Synapsen sind zwischen den Nervenzellen liegende Kontakstellen, an denen die Erregungsübertragung vom Axon einer Zelle auf den Dendriten einer anderen Zelle stattfindet.

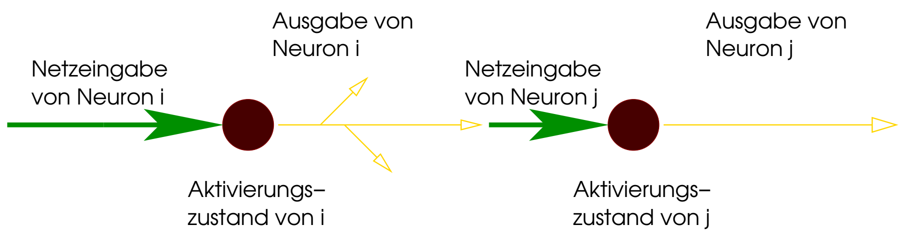

Abbildung 6.1: Dargestellt ist die schematische Signalleitung in einem Neuronalen Netz.

Betrachtet man ein einzelnes Neuron, so kann es wiefolgt definiert werden:

Ein künstliches Neuron stellt ein Tupel (x, w, f_a, f_o, o) dar, bestehend aus $x = (x_1, \ldots, x_n)$ dem Eingabevektor, $w = (w_1, \ldots, w_n)$ dem Gewichtsvektor, f_a der Aktivierungsfunktion und f_o der Ausgabefunktion. Über die Aktivierungsfunktion wird angegeben, wie sich ein neuer Aktivierungszustand (Aktivierungsgrad der Zelle) eines Neurons aus seiner alten Aktivierung, der Netzeingabe sowie dem Schwellenwert S dieses Neurons ermittelt. Aus der Aktivierung der Ausgabe des Neurons ist die Ausgabefunktion bestimmbar. Mittels $f_o(f_a(x, w)) = o$ wird der Ausgabewert des Neurons erzeugt. Die Weiterleitung dieses Wertes an nachfolgende Neuronen erfolgt über Axonkollaterale[2]. Eine Weiterleitung kann jedoch nur erfolgen, wenn ein bestimmter Schwellenwert S überschritten wurde. Die zugehörige Schwellenwertfunktion lautet:

$$f_o(\sum_{k=1}^{n} x_k w_k) = \begin{cases} 1 & \text{falls } \sum_{k=1}^{n} x_k w_k \geq S \\ 0 & \text{sonst} \end{cases} \tag{6.1}$$

Der zeitliche Abstand des Informationsflußes zwischen Neuronen ist nach unten beschränkt, daher sollte eine beschränkte Ausgabefunktion verwendet werden [32]. Sinnvoll dabei ist eine Beschreibung durch glatte, also differenzierbare Funktionen. Diese sind die *s-förmigen* oder *sigmoiden* Funktionen:

Die Funktion s_c heißt *sigmoid* oder *s-förmig*, wenn sie monoton steigend und differenzierbar ist. Hier muß

$$\lim_{x \to +\infty} s_c(x) = k_1 \;\text{ und }\; \lim_{x \to -\infty} s_c(x) = k_2 \;\text{ mit }\; k_2 < k_1 \tag{6.2}$$

gelten. Die folgenden Abbildungen in 6.2 stellen gebräuchliche Aktivierungs- bzw. Ausgabefunktionen dar. Am häufigsten verwendet werden die letzteren Beispiele.

[2]kollateral = parallel laufend

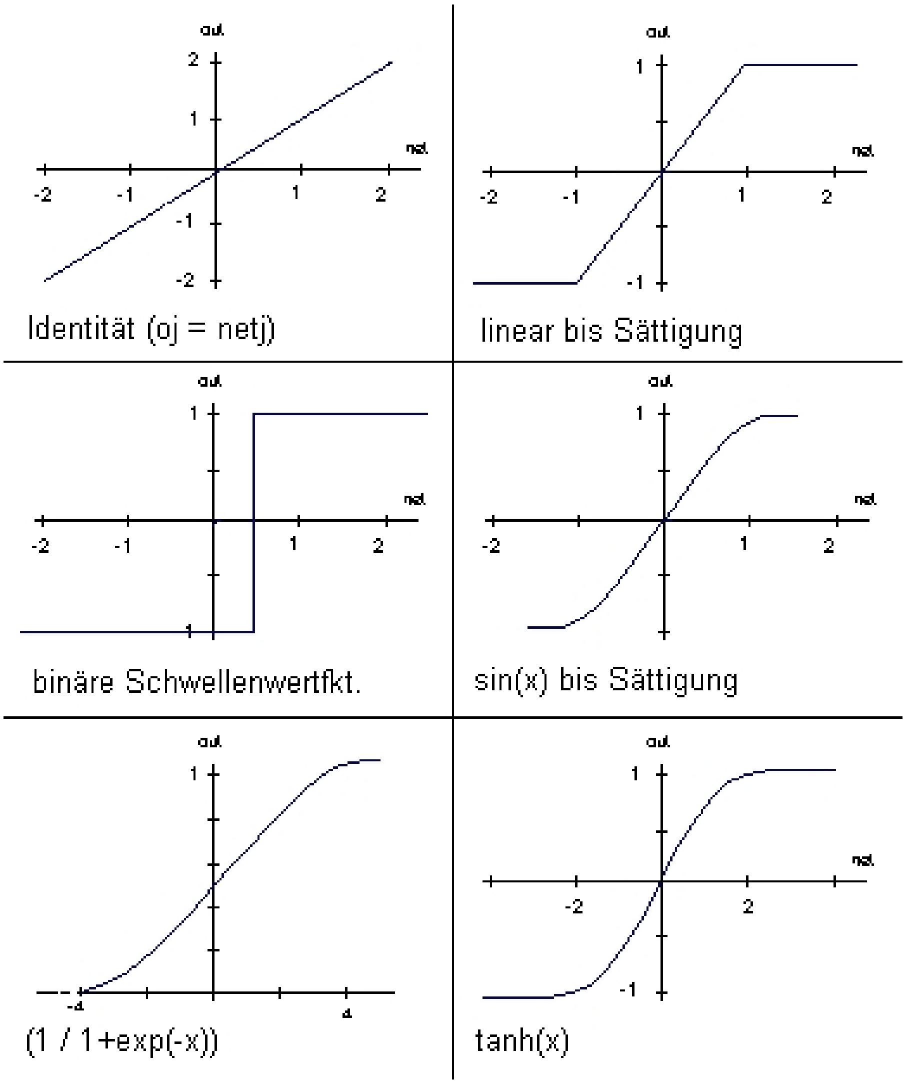

Abbildung 6.2: Mögliche *sigmoide* Aktivierungs- bzw. Ausgabefunktionen. Letztere sind am gebräuchlichsten. [32]

6.2 Vom Neuron zum Neuronalen Netz

Vorangehend wurden einzelne Neuronen beschrieben. Verbindet man diese miteinander, so entsteht ein Neuronales Netz. Ein Neuronales Netz wird gebildet aus einem Paar (N, V). Dabei beschreibt N die Menge von Neuronen und V die Menge von Verbindungen. Seine graphische Struktur ist in Abbildung 6.3 zu sehen, für welche noch folgende Zusätze und Einschränkungen zu beachten sind:

- Die Punkte in der Graphik symbolisieren die Neuronen.
- Die mit Pfeilen gekennzeichneten Linien stellen die gewichteten Verbindungen dar.

  - ◇ Positive Gewichte eines Neurons zu einem anderen üben hierbei einen exzitatorischen, also einen erregenden Einfluß auf diese aus.
  - ◇ Negative Gewichte zwischen Neuronen bedeuten einen inhibitorischen, demnach hemmenden Einfluß.
  - ◇ Ein Gewicht von Null heißt, daß momentan kein Einfluß des einen Neurons auf ein anderes zu erwarten ist.

- Neuronen einer Ebene werden zu einer Schicht zusammengefaßt.

- Jedes Neuronale Netz kann aus einer beliebigen Menge an Schichten aufgebaut sein, jedoch steht am Beginn immer eine Eingabeschicht und am Ende eine Ausgabeschicht. Dazwischen liegende Schichten werden als verdeckte Schichten bezeichnet.

- Jedem Neuron steht es offen, eine beliebige Menge von Verbindungen zu empfangen, über welche es die Eingaben erhält.

- Jedes Neuron kann lediglich eine Ausgabe über seine beliebige Menge von Verbindungen aussenden.

- Über von "außen" kommende Verbindungen erhält das Neuronale Netz seine Eingaben. Die Ausgaben werden über nach "außen" endende Verbindungen abgegeben.

Das Hauptinteresse eines Neuronalen Netzes liegt in der Gewichtung der Verbindungen der Neuronen untereinander. Über sie kann das Wissen Neuronaler Netze definiert werden, welches in ihren Gewichten gespeichert ist. Die Veränderung der Gewichte in einem Neuronalen Netz kann somit als Prozess des Lernens angesehen werden.

In der vorliegenden Analyse werden Netze ohne Rückkopplung (sogenannte *feedforward-Netze*) verwendet. In diesen existiert kein Pfad, der von einem Neuron i direkt oder indirekt über weitere Neuronen zurück zu diesem Neuron i führt. Somit ist der Datenfluß auf eine Richtung beschränkt. Darüberhinaus existieren lediglich Verbindungen von einer Schicht zur nächsten. Sie sind also ebenenweise verbunden (siehe Abbildung 6.3). Sehr oft wird in diesem Zusammenhang auch der Begriff des *Perzeptrons* anstelle des Neuronalen Netzes benutzt. Das *Perzeptron* besteht wie jedes Neuronale Netz aus einer Schicht von Eingabezellen, welche über gewichtete Verbindungen mit einer folgenden Schicht von Neuronen in Kontakt sind. Einfache Muster werden von diesen Neuronen erkannt, welche wieder gewichtet verbunden mit einer weiteren Schicht sind. Verbindungen zwischen verdeckten Schichten und der Ausgabeschicht sind variabel und demnach trainierbar. Ist mehr als eine verdeckte Schicht vorhanden, spricht man von Multilagen-Perzeptronen. Die Anzahl der Ausgabeneuronen bleibt vom Prinzip unabhängig.

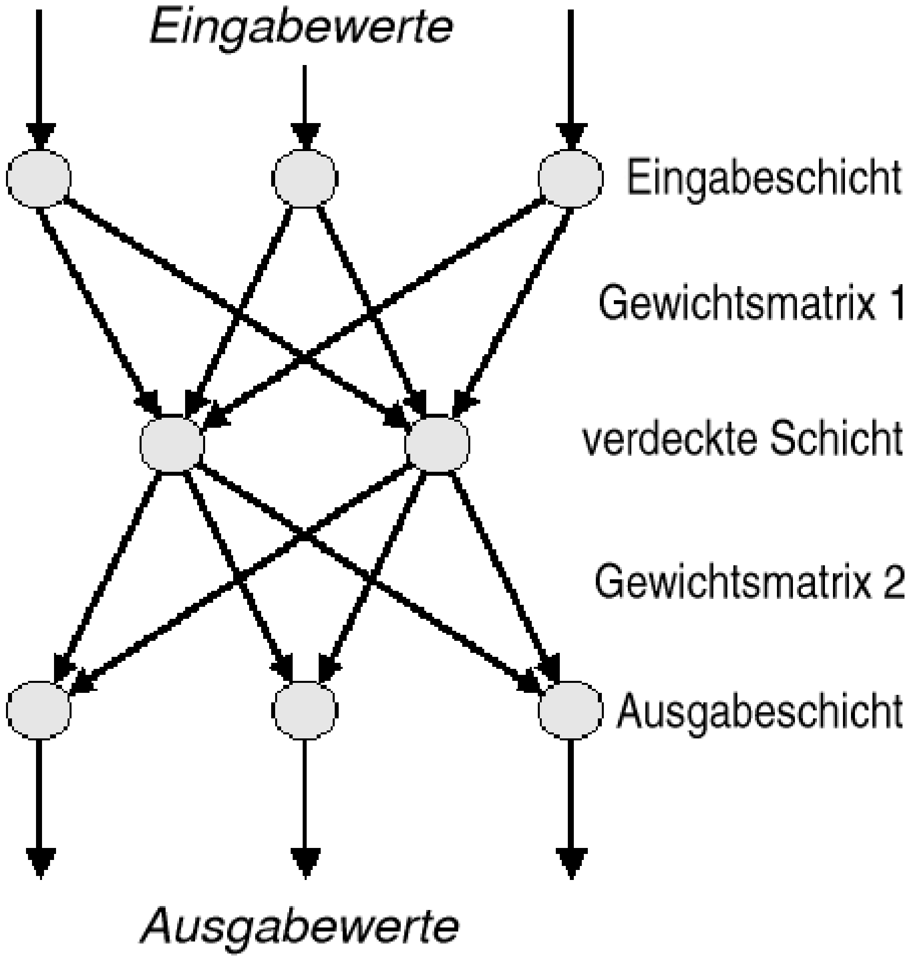

Abbildung 6.3: In der graphischen Darstellung ist der Aufbau eines Neuronalen Netzes zu sehen.

6.2.1 Die Gewichtsmatrix

Neuronale Netze sind auch als Matrizen darstellbar. Die mathematische Einheit einer Matrix ist genau wie eine einzelne Zahl. Die Matrix **W** setzt sich aus einer Menge von Elementen w_{ij} zusammen. Der Index i steht dabei für die Zeile eines Elementes innerhalb der Matrix, der Index j hingegen beschreibt seine Spalte. Indem Neuronale Netze die Verbindungen zwischen ihren Neuronen gewichten, lernen sie und speichern so ihr Wissen.
Eine sinnvolle Darstellung für die Netzgewichte bietet eine sogenannte Gewichtsmatrix. Ist keine verborgene Schicht vorhanden, so ist das Netz mit nur einer Matrix zu beschreiben. Diese Analyse verwendet Neuronale Netze, welche genau mit einer verdeckten Schicht arbeiten. Demnach sind zwei Gewichtsmatrizen notwendig zur Beschreibung dieser Neuronalen Netze (siehe Bild 6.3). In Abbildung 6.4 ist die beschriebene Verallgemeinerung dargestellt.

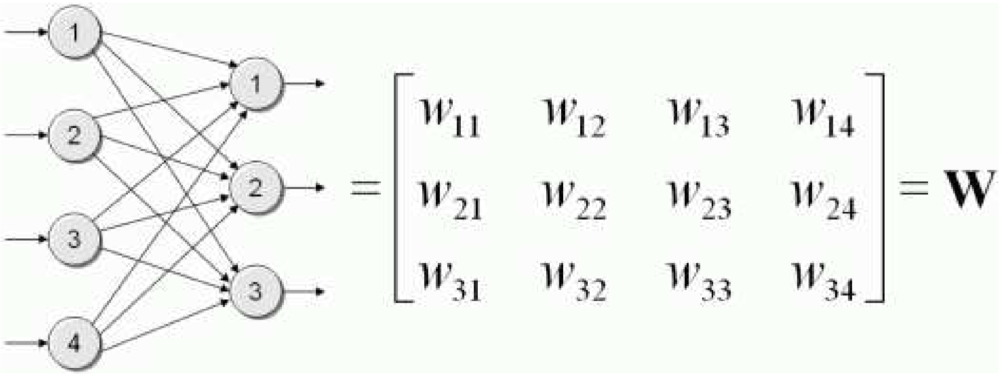

Abbildung 6.4: Darstellung der Äquivalenz zwischen schematischer Illustration (Abb. 6.3) und Matrizenschreibweise. Der erste Index $(1,2$ oder $3)$ bezieht sich auf die Ausgabeschicht $(1,2$ oder $3)$, Index zwei beschreibt entsprechend die Eingabeschicht (1 bis 4). Für Neuronale Netze mit verborgenen Schichten sind weitere, der Anzahl der verdeckten Schichten entsprechende Gewichtsmatrizen einzuführen.

6.2.2 Der Lernprozess

Neuronale Netze “lernen”, indem sie sich gemäß einer fest vorgegebenen Vorschrift selbst modifizieren. Diese Vorschrift wird als *Lernregel* (6.3.3) bezeichnet. Allgemein kann ein Lernprozess folgende Merkmale aufweisen:

- Entwicklung neuer Verbindungen
- Löschen existenter Verbindungen
- Veränderung der Gewichte (Modifikation der Verbindungsstärke)
- Modifikation des Schwellenwertes
- Modifikation der Aktivierungs- bzw. Ausgabefunktion
- Entwicklung neuer Zellen
- Löschen bestehender Zellen

Das Hauptinteresse liegt dabei auf dem Lernen durch eine Veränderung der Gewichte.

6.2.3 Hebb’sche Lernregel

Der Psychologe Donald Hepp formulierte im Jahre 1949 erstmalig eine Lernregel [32]. Sein Algorithmus versuchte die Lernfähigkeit des Gehirns zu erklären. Dieser Algorithmus auf das mathematische Modell übertragen, ergibt eine Regel mit dem Wortlaut:

Wenn Zelle j eine Eingabe von Zelle i erhält, so daß beide gleichzeitig stark aktiviert sind, dann erhöhe das Gewicht w_{ij} zwischen ihnen oder anders, verstärke die Verbindung von Zelle i zu Zelle j.

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \quad \text{mit} \quad \Delta w_{ij} = \xi \cdot o_i \cdot a_j \tag{6.3}$$

Dabei steht Δw_{ij} für die Änderung des Gewichtes w_{ij}, ξ ist die Konstante der Lernrate, o_i steht für die Ausgabe der Vorgängerzelle i und a_j für die Aktivität der Nachfolgerzelle j [33]. Die Gewichtsänderung ist demnach abhängig von einer konstanten Lernrate, der Ausgabe des Vorgängerneurons und dem aktuellen Wert der Aktivierungsfunktion des darauffolgenden Neurons. Problematisch zeigt sich hier ein Fall: Halten die Aktivitäten der Neuronen i und j länger an, so wachsen ihre Gewichte ins Unendliche, da die Zellen eine Funktion des "Vergessens" nicht kennen. Durch entsprechende Modifikationen läßt sich dieser Mangel jedoch beseitigen. Ein Großteil der in den letzten Jahren entwickelten Lernregeln basiert auf der Hebb'schen Lernregel [32].

6.3 Das Backpropagationsnetz

Die Neuronalen Netze der vorliegenden Analyse zeichnen sich durch ein weiteres Merkmal aus. Sie gehören zu den *Backpropagationsnetzen*, welche nach dem *Backpropagationslernverfahren* arbeiten. Diese Rechenvorschrift ermöglicht es, Verbindungen zu den verborgenen Schichten zu modifizieren [34]. Als iteratives Verfahren ermittelt es im Netz eine Konfiguration der Gewichte, bei der die Fehlersumme über alle Trainingsmuster minimal ist. Basis der *Backpropagationsnetze* sind die bereits vorgestellten *Feedforwardnetze* mit einer Eingabeschicht, einer Ausgabeschicht und mindestens einer verborgenen Schicht, d.h. für diese Analyse:

- In 3 Schichten sind die Neuronen angeordnet.
- Die erste Schicht U_0 ist aus n Eingabeneuronen ohne Eingangsgewichte aufgebaut. Sie reichen folglich eine unveränderte Eingabe an die zweite Schicht weiter.
- Daran anschließend ist eine verborgene Schicht U_1.
- In der dritten Schicht U_2 befinden sich m Ausgabeneuronen (m=1 in vorliegender Analyse)
- Als Aktivierungsfunktion außerhalb der Eingabeschicht dient im Regelfall die gewichtete Summe der Eingaben.
- In der verdeckten Schicht ist die Ausgabefunktion s_c *sigmoid* und nicht-linear.

$$s_c(x) = \frac{1}{1+e^{-x}} \tag{6.4}$$

- Die Ausgabefunktion der Ein- und Ausgabeschicht ist die Identität (siehe Abbildung 6.2)

Die *sigmoide* Ausgabefunktion s_c ist differenzierbar und ermöglicht, auch nicht-lineare Funktionen durch ein *Backpropagationsnetz* zu approximieren. Auch Treppenfunktionen sind mit ihr approximierbar. Sie beeinflußt desweiteren die Fehlerfunktion (Kapitel 6.3.2), die durch s_c geglättet wird. *Sigmoide* Ausgabefunktionen verstärken die Entstehung lokaler Minima.[3] Darüberhinaus besitzt die Fehlerfunktion in lediglich

[3]Ein lokales Minimum ist der Funktionswert einer Stelle, in deren Umgebung die Funktion keine kleineren Werte annimmt.

einem Minimum bzw. Maximum eine waagerechte Tangente. Der Gradient des Fehlers $-\nabla_{\vec{w}} F(\vec{w})$ zeigt dabei immer in Richtung des Minimums (Kapitel 6.3.2).
Das Lernverfahren bei *Backpropagationsnetzen* erfolgt durch überwachtes Lernen. In diesem Fall wird extern über einen "Lehrer" dem Netz zu jeder Eingabe die korrekte Ausgabe oder die Differenz zwischen tatsächlicher und korrekter Ausgabe angegeben. Diese Differenz geht in die Lernregel ein, so daß das Netz entsprechend modifiziert werden kann. Das Training des *Backpropagationsalgorithmus* durchläuft alternierend drei Phasen: den Vorwärtspass (Forward Pass), die Bestimmung des Fehlers und den Rückwärtspass (Backward Pass).

Forward Pass
Für jedes Neuron wird in der U_1 Schicht die Aktivierung berechnet. Mit Hilfe der Ausgabefunktion ermittelt sich daraufhin die Ausgabe von U_1, welche gleichzeitig wieder die Eingabe von U_2 darstellt. Demnach sind Ausgaben der Vorgängerschicht immer Eingaben der folgenden Schicht, so daß die Daten in Abbildung 6.3 von oben nach unten das Netz durchlaufen. Sobald der Forward Pass die Ausgabeschicht erreicht, wird die Ausgabe bestimmt.

Fehlerbestimmung
Jede Eingabe ist auch ein Element der Trainingsmenge. Zu jedem Element dieser Trainingsmenge ist die gewünschte Ausgabe des Netzes dann bekannt. Der Fehler des Netzes bestimmt sich dabei mittels einer Fehlerfunktion (Kapitel 6.3.2). Eine dabei vorgegebene Güteschwelle entscheidet über den weiteren Trainingsverlauf. Bei einem Fehler oberhalb dieser Schwelle wird das Netz folglich mit dem Backward Pass modifiziert. Für einen Fehler unterhalb der Schwelle ist das Training beendet, gegebenenfalls wird eine Testphase eingeläutet, um die Generalisierbarkeit (Kapitel 6.3.1) zu überprüfen.

Backward Pass
Sukzessive werden in diesem Schritt die Verbindungen zwischen den Neuronen im *Backpropagationsnetz* unter Zuhilfenahme einer Lernregel modifiziert. Als Grundlage dieser Lernregel dienen die Fehler des Netzes. Der Backward Pass arbeitet in entgegengesetzter Richtung zum Forward Pass, somit wird oft von einem rückwärtsverteilten Fehler gesprochen. Beim Backward Pass werden zuerst die Gewichte zwischen Ausgabeschicht U_2 und der verdeckten U_1 Schicht modifiziert. Anschließend erfolgt noch die Modifizierung der Verbindungen von U_1 zur Eingabeschicht U_0.

6.3.1 Die netzspezifische Generalisierungsfähigkeit

Backpropagationsnetze approximieren eine Funktion über Trainingsdaten, welche sozusagen Stützstellen symbolisieren, in denen das Netz genügend Präzision aufweisen kann. Diese Präzision soll aber auch außerhalb der Stützstellen gelten, welche als Generalisierungsfähigkeit des Netzes bezeichnet wird. Mittels einer Testmenge[4] wird die Generalisierungsfähigkeit verifiziert. Ist die gewünschte Genauigkeit erreicht, gilt

[4]Testmenge und Trainingsmenge dürfen in keiner Komponente übereinstimmen, sie sollen vielmehr orthogonal zueinander sein.

die Trainingsphase als beendet und die Testphase zur Verifizierung der Generalisierungsfähigkeit des Netzes kann eingeleitet werden [35].

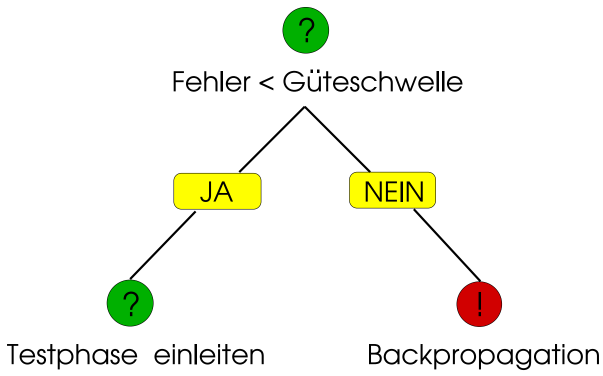

Abbildung 6.5: Prozessierung zur Einleitung der Testphase I.

Bei Abbruch des Trainings werden die Trainingsmuster perfekt erkannt. Dabei besteht jedoch keine Beurteilungsmöglichkeit ob das Netz lediglich einzelne Muster gelernt hat oder ob eine grundlegende Abbildungsvorschrift für die vorhandenen Trainingsmuster gefunden wurde.

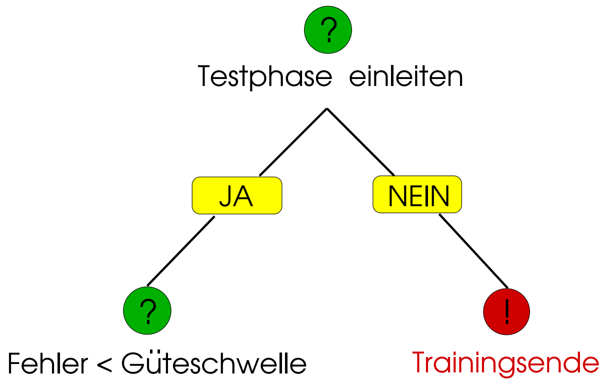

Abbildung 6.6: Prozessierung zur Einleitung der Testphase II.

Die Testmenge wird während der Testphase angewendet. Der Verlauf des weiteren Trainings entscheidet sich anhand der zur Testmenge gehörigen Fehler. Sofern dieser Fehler weiterhin unter der zuvor definierten Güteschwelle liegt, kann das Training nun endgültig abgeschlossen werden. Ist der Fehler aber immer noch zu groß, muß nachtrainiert werden.

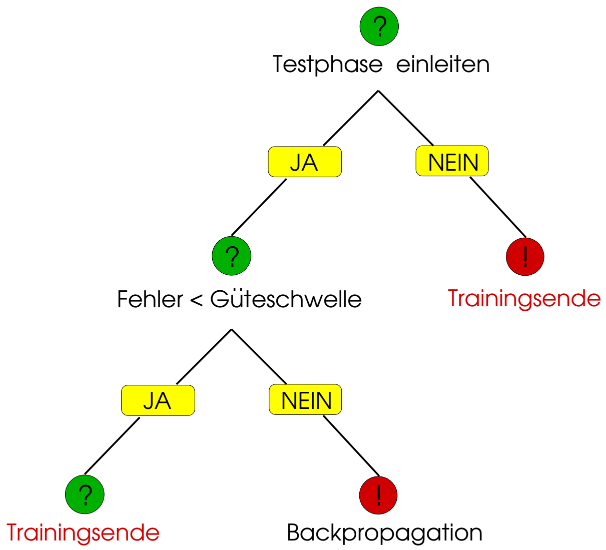

Abbildung 6.7: Prozessierung zur Einleitung der Testphase III.

Die Trainingsphasen in der vorliegenden Analyse werden auf 500 Trainingsepochen festgesetzt. Bei der Wahl der Trainingsepochen ist darauf zu achten, daß diese weder zu niedrig noch zu hoch angesetzt werden. Es sollte die Möglichkeit gewährleistet werden, daß das Neuronale Netz innerhalb der Trainingsepochen die Muster in seinem Algorithmus so gut lernt, daß diese danach auf einen neuen Datensatz ohne Probleme angewendet werden können. Neuronale Netze, die keine Funktion des Vergessens aufweisen, zeigen sich als problematisch bei sogenannten *Übertraining*. Tritt dieser Fall ein, spezialisiert sich das Netz so sehr auf das eingestellte Training, daß es seine Generalisierungsfähigkeit verliert. Der Effekt zeigt sich bei Anwendung auf die Test-Daten. Es ergibt sich ein Netz, welches nur "gut" innerhalb des eingestellten Trainings funktioniert, in Anwendung außerhalb des Trainings zeigen sich qualitativ minderwertige Ergebnisse.
Die Frage, wie rechtzeitig festgestellt werden kann, wann ein Netz seine Generalisierungsfähigkeit verliert, ist einfach zu klären. Dazu wird einfach während der Trainingsphase eine graphische Ausgabe erzeugt, welche in regelmäßigen Etappen aktualisiert wird. (Die graphischen Ausgaben sind in den Abbildungen in Kapitel 6.4 dargestellt.) In diesen Diagrammen werden die Fehler der trainierten und der getesteten Daten gegen die Trainingsepochen aufgetragen. Ihr zeitlicher Verlauf ist sozusagen während der gesamten Trainingsphase überschaubar. Im Regelfall sollen die Fehler für Test und Training mit zunehmender Trainingsepoche sinken. Steigt jedoch der Fehler der Testdaten mit der Epochenanzahl wieder an, so hat der Prozess des Übertrainierens begonnen. Die Zahl der Trainingsepochen sollte in jedem Fall vermindert werden.

6.3.2 Fehlerbestimmung

Über die zuvor erwähnte Güte können nur Aussagen gemacht werden mit Hilfe des Fehlers des Netzes. Seiner Bestimmung muß also entsprechende Beachtung

entgegengebracht werden. Zu einem Eingabevektor $\vec{x} = (x_1, \ldots, x_n)$ und im Netz fest vorgegebenen Gewichten $\vec{w} = (w_{U_1}, \ldots, w_{U_2})$ liefert das *Backpropagationsnetz* stets diesselbe Ausgabe. Daher ist das Ein- und Ausgabeverhalten des Netzes als Funktion o_{Netz} darstellbar. Für die Netzwerkausgabe gilt dann $\vec{o} = o_{Netz}(\vec{x}, \vec{w})$.
Die Güte dieser Netzwerkausgabe in Bezug auf Trainings- oder Testdaten $(\vec{x}, \vec{o}')$ wird dann beschrieben durch $\vec{o}' - \vec{o}$. Sei f die zu "lernende" Funktion, so gilt $\vec{o}' = f(\vec{x})$. Für die Güte des Netzes ergibt sich folglich der Ausdruck : $f(\vec{x}) - o_{Netz}(\vec{x}, \vec{w})$. Zwischen positiven und negativen Fehlern zu unterscheiden ist wenig sinnvoll, demnach werden die Fehler quadriert.

Zu einem gegebenen $\vec{x}$ und $\vec{w}$ des *Backpropagationsnetzes* ermittelt sich der quadratische Fehler zu

$$F(\vec{x}, \vec{w}) = |f(\vec{x}) - o_{Netz}(\vec{x}, \vec{w})|^2 \quad . \tag{6.5}$$

Er ist ein Maß der zum Neuronalen Netz gehörigen Güte.

Diese Definition bezieht sich dabei auf eine konkrete Eingabe $\vec{x}$. Um nun repräsentative Aussagen über die Güte machen zu können, benötigt man ein Fehlermaß bezüglich aller möglichen Eingabevektoren $\vec{x}_k$ mit $k = 1, \ldots, N$. Dazu wiederum wird der quadratische Fehler gebraucht, denn er wird von möglichst vielen $\vec{x}_k$ bestimmt. Anschließend erfolgt die Mittelung über die Anzahl der Muster.

Der mittlere quadratische Fehler definiert sich demnach durch

$$F(\vec{w}) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} F(\vec{x}_k, \vec{w}) \quad . \tag{6.6}$$

$F(\vec{w})$ kann für den gesamten Gewichtsraum $G_{\vec{w}}$ [5] bestimmt werden, indem man $\vec{w}$ variiert. Die strukturielle Fehleroberfläche der dabei entstehenden Fehlerfunktion wird bedeutend bestimmt durch die sigmoide Ausgabefunktion[6] (6.4), so daß der Gradientenabstieg (6.3.3) erleichtert wird, indem Punkte auf der Fehleroberfläche Neigungen in Richtung der nächstliegenden Täler erhalten. Der Gradientenabstieg führt in das nächstliegende Tal, wenn diese vorgegebene Richtung beibehalten wird, wenn also der Gewichtsvektor in Richtung $-\nabla_{\vec{w}} F(\vec{w})$ verschoben wird.
Variation der Gewichte zielt also darauf den Fehler zu minimieren. Letztenendes ist $\vec{w}$ so zu wählen, daß $F(\vec{w})$ ein absolutes Minimum $F_{min}(\vec{w})$ erreicht.

6.3.3 Die Backpropagationslernregel

Der Lernprozess des *Backpropagationverfahrens* basiert auf dem bereits erwähnten Verfahren des Gradientenabstieges. Dieser wird an der im Vektorraum $G_{\vec{w}}$ liegenden Fehlerfunktion ausgeführt. Dabei wird zu Beginn das Gewicht w_1 zufällig gewählt und der Gradient sowie die Tangente in diesem Punkt berechnet. Die fest vorgegebene Länge der Lernrate in Abhängigkeit von der Tangente gibt dabei an, wieviel sich der

[5] $G_{\vec{w}}$ sei als Vektorraum aller $\vec{w}$ definiert.
[6] Die Fehleroberfläche wird durch die sigmoide Ausgabefunktion geglättet.

Punkt abwärts in Richtung des Gradienten bewegt (sogenannte *Liniensuche*). Aus dieser Bewegung ermittelt sich eine Gewichtsänderung, die zum neuen Gewichtsvektor w_2 führt. Zu diesem Vektor wird der Fehler ermittelt. Liegt er oberhalb der bereits definierten Güteschwelle wird in diesem Punkt wiederum der Gradient und die Tangente bestimmt und das Verfahren iterativ in der beschriebenen Weise weitergeführt. Wird ein Minimum erreicht, endet die Iteration.
Als allgemeine Form für die Richtung des Anstiegs ist

$$\nabla_{\vec{w}} F(\vec{w}) := \left(\frac{\partial F(\vec{w})}{\partial w_1}, \ldots, \frac{\partial F(\vec{w})}{\partial w_q} \right)^T \tag{6.7}$$

anzugeben. Die Anzahl der Gewichte im Netz ist mit q bezeichnet. Die Vektoren $\vec{w}$ und auch der Gradient sind also q-dimensional. Der Gradient für den speziellen Fall ermittelt sich mit (6.5) und (6.6) zu

$$\nabla_{\vec{w}} F(\vec{w}) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \nabla_{\vec{w}} F_k(\vec{x}_k, \vec{w}) \quad . \tag{6.8}$$

Das Gradientenabstiegsverfahren ist der zentrale Punkt in Lernmethoden die auf Basis einer *Liniensuche* arbeiten. Die verschiedenen Algorithmen in diesen Liniensuchverfahren unterscheiden sich lediglich in der Art und Weise, wie die Richtung aus dem Gradienten berechnet wird. In dieser Analyse wird ausschlielich mit Netzen gearbeitet, die mit der *Broyden, Fletcher, Goldfarb, Shanno* (BFGS) Lernmethode ihr Gradientenabstieg vollziehen. Diese Methode impliziert die Berechnung von $n \times m$ Gewichtsmatrizen $\mathbf{W}_{nm}$, wobei $n \times m$ die Anzahl der Gewichte zwischen zwei benachbarten Schichten beschreibt (siehe Abb. 6.4).

6.3.4 Probleme der Backpropagation

Das *Backpropagationsverfahren* mit dem Gradientenabstieg verbirgt jedoch einige Gefahren. Ihre Ursache ist hauptsächlich in der Tatsache zu suchen, daß lediglich die lokale Umgebung (der Gradient) bekannt ist, aber auch die Lernrate nimmt maßgeblich Einfluß. Die folgenden Abschnitte sollen die wesentlichen Probleme kurz verdeutlichen.

Lokale Minima

Wird das Gradientenabstiegsverfahren betrachtet, kann nicht eindeutig festgelegt werden, ob das Verfahren in einem lokalen oder globalen (absolutem) Minimum endet, da der Startvektor $\vec{w}$ als eine zufällige Größe bestimmt wird. Sofern dieser in der Nähe eines absoluten Minimums liegt, ist auch das Ziel des Gradientenabstiegsverfahrens ein optimales Minimum. Wird jedoch in der Nähe eines lokalen Minimums gestartet, läuft der Iterationsprozess auf ein suboptimales Minimum heraus.
Die Anzahl der Verbindungen zwischen den Neuronen ist ein Maß für die Dimension des Netzes. Je größer diese wird, desto größer zeigt sich auch die Zerklüftung der Fehleroberfläche. Steigt die Zerklüftung, nimmt auch die Anzahl der lokalen Minima innerhalb der Oberfläche zu und es wird schwieriger, ein globales Minimum zu erreichen. Weiter besteht auch die Möglichkeit, daß tiefere Täler entstehen, die in ihrer Ausdehnung, aber so schmal sind, daß sie bei zu großer Lernrate übersprungen werden können und das globale Minimum schließlich überlaufen wird.

Flache Plateaus

Der Betrag des Gradienten bestimmt die Größe der Gewichtsveränderung. In flachen Plateaus jedoch ist dieser Gradient relativ klein und führt dazu, daß eine sehr große Anzahl an Iterationen nötig ist, um das Plateau wieder zu verlassen. Im Extremfall kann das Verfahren mit einem Nullvektor stagnieren. Das Neuronale Netz würde dann das Verhalten als ein Erreichen eines Minimums interpretieren.

Oszillationen

Ist der Gradient am Rande eines Tals so groß, um durch die Gewichtsveränderung einen Sprung auf die gegenüberliegende Seite des Tals zu realisieren, kann dort der Gradient den gleichen Betrag mit umgekehrten Vorzeichen besitzen. So wäre das Verfahren wieder am Startpunkt angelangt. In diesem Fall spricht man von Oszillationen.

Verlassen guter Minima

Ausgehend von einem globalen Minimum in einem steilen Tal besteht wieder die Gefahr, daß der Betrag des Gradienten zu groß wird. Die Gewichtsveränderung kann somit das Verfahren weg von einem optimalen Minimum zu einem suboptimalen Minimum bewegen.

6.3.5 Lösungen der Probleme

Wahl der Lernrate

Ein entscheidender Gesichtspunkt zum Verhalten beim Gradientenabstieg ist die Wahl der Lernrate ξ. Ist diese sehr ungünstig gewählt, sind die obigen Problemstellungen sehr wahrscheinlich. Zu einer optimalen Lernrate gehören diverse Faktoren u.a. das Problem an sich, die Wahl der Trainingsdaten sowie die Größe und Topologie des Netzes.
Lokale Minima können mit einer geschickten Wahl der Lernrate umgangen werden. Wird sie zu groß gewählt, läuft man Gefahr in engen Tälern Minima zu überspringen. Dem abzuhelfen, heißt eine kleinere Lernrate zu wählen. Auch Oszillationen kann damit begegnet werden, jedoch nicht ohne vollständigen Ausschluß. Dazu müßte die Lernrate extrem klein gewählt werden, was eine immense Trainingszeit zur Folge hätte. Bei flachen Plateaus (sie erfordern höhere Lernraten) wäre diese sogar inakzeptabel. Demnach darf die Lernrate generell weder zu groß noch zu klein gewählt werden.

Initialisierung der Gewichte

Auch die Initialisierung der Gewichte entscheidet über den Verlauf des Trainings. Durch diesen Schritt wird der Startpunkt für das Gradientenabstiegsverfahren festgelegt. Wie bereits beschrieben, ist der Startpunktpunkt entscheidend für lokale und globale Minima sowie die dargestellten Probleme. Treten Schwierigkeiten auf, empfiehlt es sich, zuerst die Initialisierung der Gewichte zu überprüfen.

6.4 Spezifizierung Neuronaler Netze anhand von Untergrundprozessen

Nachdem in den vorherigen Abschnitten bereits alle gemeinsamen Details (Lernmethode, Anzahl der verborgenen Schichten, etc.), der in dieser Analyse verwendeten Neuronalen Netze, abgehandelt worden sind, sollen die folgenden Abschnitte ihre speziellen Eigenschaften (Eingabevariablen, die der Anzahl der Eingabeneuronen entsprechen) etwas näher beleuchten. Insgesamt verwendet die Analyse sechs verschiedene, den jeweiligen Untergrundprozessen angepaßte Neuronale Netze.
Diese Analyse konnte sich nicht auf ein oder maximal zwei Netze zur einfacheren, hierarchisch strukturierteren Selektion beschränken. Die Begründung hierfür soll in den nächsten Kapiteln ersichtlich werden, da die möglichen Selektionsvariablen nicht für alle relevanten Untergrundprozesse gleich effektiv waren.
Damit ein Neuronales Netz zur Signal-Untergrund-Diskriminierung effizient eingesetzt werden kann, ist es notwendig, physikalisch relevante Variablen zu finden. Diese Variablen sollten im Signal-Untergrund-Vergleich erhebliche Unterschiede zwischen Signal und Untergrund aufweisen. Eine genaue Analyse der diversen Prozesse ist demnach Voraussetzung. Sie verhilft Mechanismen aufzudecken, welche die Selektion mittels Neuronaler Netze begünstigen.

6.4.1 Wahl geeigneter Variablen

Damit ein Neuronales Netz in seinem Training grundlegende Muster lernt, um später mit diesem Wissen das Signal vom Untergrund trennen zu können, müssen Variablen gefunden werden, an denen sich das Netz bei seinem Lernprozess orientieren kann. Diese Variablen sollten nach Möglichkeit in ihren Signal- und ihren Untergrund-Verteilungen deutlich variieren. Auch Separationsmöglichkeiten für korrelierte bzw. nicht korrelierte Variablen müssen in diesem Zusammenhang betrachtet werden. Weist z.B. der Signalprozess eine Korrelationen zweier Größen auf, während diese Größen im Untergrund unkorreliert erscheinen, ist der Effekt im Zusammenhang mit Neuronalen Netzen unbedingt zu nutzen.
Die folgenden Abschnitte sollen einen kurzen Überblick über mögliche Kandidaten für Neuronale Netze geben. Ihre Verteilungen im Vergleich von Untergrund zum Signal sind in den jeweiligen Abbildungen dargestellt. Auf die bekannten Größen der Vorselektion (die Transversalimpulse beider Myonen, siehe dazu Abb. 5.5 und Abb. 5.6) soll hier nicht noch einmal eingegangen werden. Sie sind auf jeden Fall als mögliche Variablen für Neuronale Netze festzuhalten.

Invariante Masse

Die invariante Masse $M_{\mu\mu}$ beider Myonen wurde bereits in Kapitel 5.3.1 Abbildung 5.4 dargestellt. Sie beschreibt eine der wichtigsten Variablen bei einer Selektion mit Neuronalen Netzen, da sich die Verteilungen dieser Größe hinsichtlich prozessbezogener Signaturen sehr gut untersuchen lassen. Das 160 GeV-Higgssignal zeigt eine charakteristische Ereignisverteilung um die 40 GeV. Besonders WW- und $t\bar{t}$-Produktionen aus Untergrundprozessen decken diesen Bereich mit einer ähnlichen

Signatur ab. Der Effekt, der hier von einem Netz ausgenutzt werden kann, liegt in der Breite der Verteilungen, da der Untergrund durch seine großen und kleinen Werte der invarianten Masse gekennzeichnet ist.
Speziell der Prozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$ ist hinsichtlich der Benutzung der invarianten Masse als mögliche Variable für ein Neuronales Netz problematisch, da dieser Untergrund mit dem Signal fast übereinstimmende Gemeinsamkeiten aufweist. Dagegen können in den übrigen Monte-Carlo-Simulationen, die eine Beteiligung des Z-Bosons aufweisen, die Effekte dessen Resonanz stark ausgenutzt werden. Die Resonanz des Z-Bosons liegt bei rund 90 GeV, so daß dieser Bereich und somit ein Großteil der Untergrundereignisse davon betroffener Monte-Carlo-Simulationen absepariert werden können.

Fehlende transversale Energie

Darüberhinaus kann als eine weitere Möglichkeit für die Signal-zu-Untergrund-Diskriminierung die fehlende transversale Energie $\not{E}_T$ angesehen werden. Speziell für das Signal ist bekannt, daß diese Größe, den nur schwer zu detektierenden Neutrinos zugeschrieben werden muß, weiterhin nehmen auch Auflösungseffekte maßgeblichen Einfluß.
Ähnlich der Impulsverteilungen, in denen die Impulse der Myonen um einen der Higgs-Masse entsprechenden chrakteristischen Wert verteilt sind (siehe Abb. 5.5 und 5.6), kann auch für die Myon-Neutrinos eine solche Verteilung für die fehlende transversale Energie angenommen werden.
Wie in Abbildung 6.8 zu sehen ist, verteilt sich die fehlende transversale Energie des Higgs-Signals um die 60 GeV. Somit eignet sich diese Größe, zumindest einen erheblichen Anteil an $Z/\gamma^* \longrightarrow \mu^+\mu^-$ und $Z/\gamma^* \longrightarrow \tau^+\tau^-$ abzuseparieren.

Öffnungswinkel zwischen den Myonen

Für Myonen, ausgehend vom Signal-Prozess, sind die Flugrichtungen miteinander korreliert. Diese Korrelation ist bei den Untergrundprozessen nicht zu beobachten. Darüberhinaus kann die Flugrichtung der Myonen, unabhängig von der Prozessart, über ihren Öffnungswinkel in der transversalen Ebene zueinander definiert werden. Diese Größe, im weiteren mit $\Delta\phi$ bezeichnet, bietet aufgrund der schon erwähnten Korrelation eine zusätzliche Separationsmöglichkeit, welche für ein Neuronales Netz verwendet werden soll.
Eine Darstellung der Winkelverteilungen ist in Abbildung 6.9 gegeben. Deutlich hebt sich das Signal vom Untergrund ab, da es sich hier um korrelierte Myonen handelt, deren Öffnungswinkel aufgrund der Korrelation eher kleine Werte annimmt. Mehrheitlich sind diese Winkel sehr viel kleiner als 90°.
Die Korrelation ist Folge des Zerfalls des Higgs-Bosons in ein W-Boson-Paar. Das Higgs-Boson, ein neutrales Spin-0-Teilchen, geht über in zwei W-Bosonen, geladene Spin-1-Teilchen, deren Spins somit in einer Korrelation zueinander stehen müssen. Diese Korrelation wird an die Myonen im folgenden Zerfall weitergegeben [6].

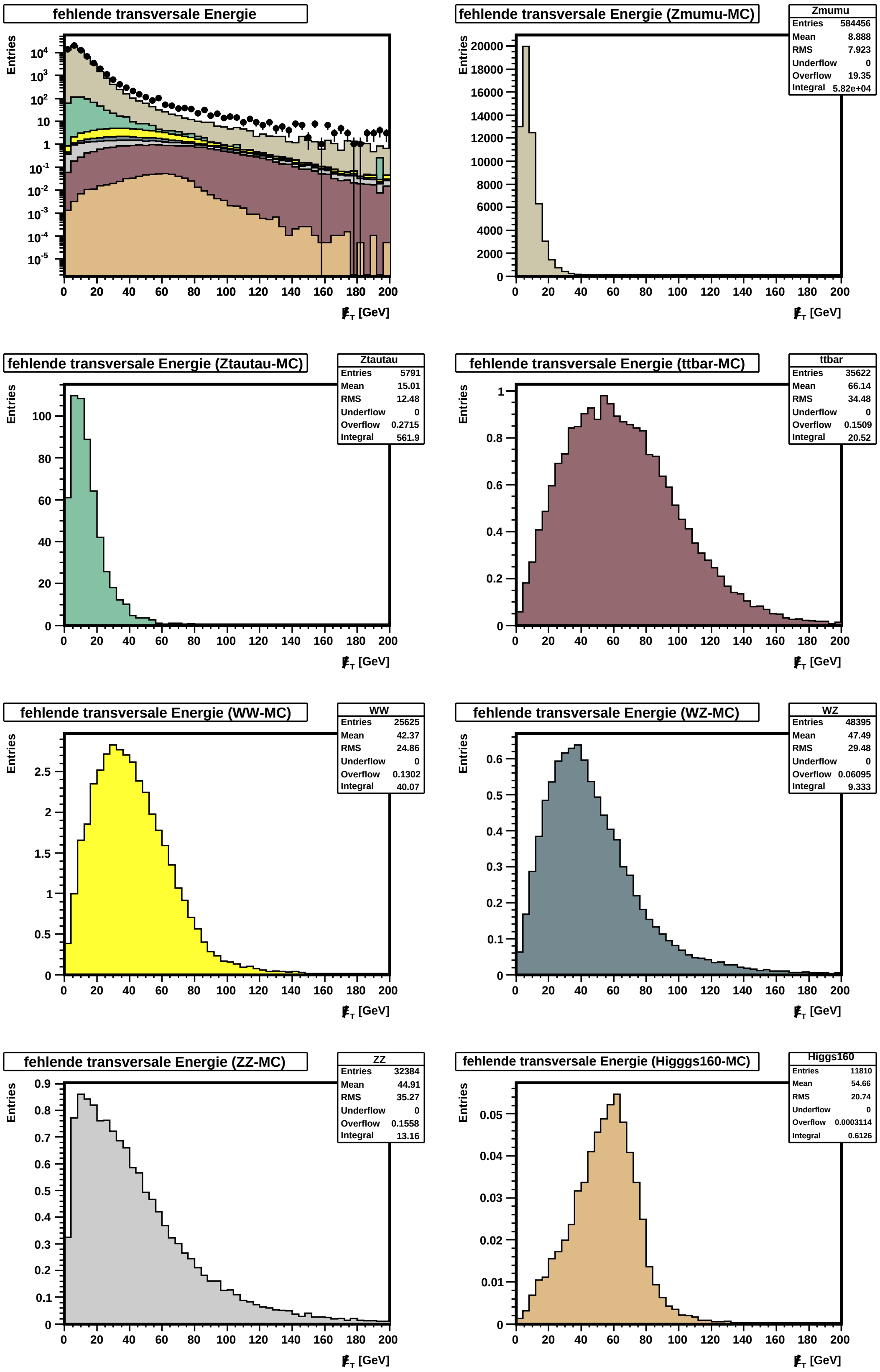

Abbildung 6.8: Die Verteilungen der fehlenden transversalen Energie werden in dieser Abbildung dargestellt. Zum direkten Vergleich von Signal und Untergrund mit den Daten sind diese logarithmisch aufgetragen.

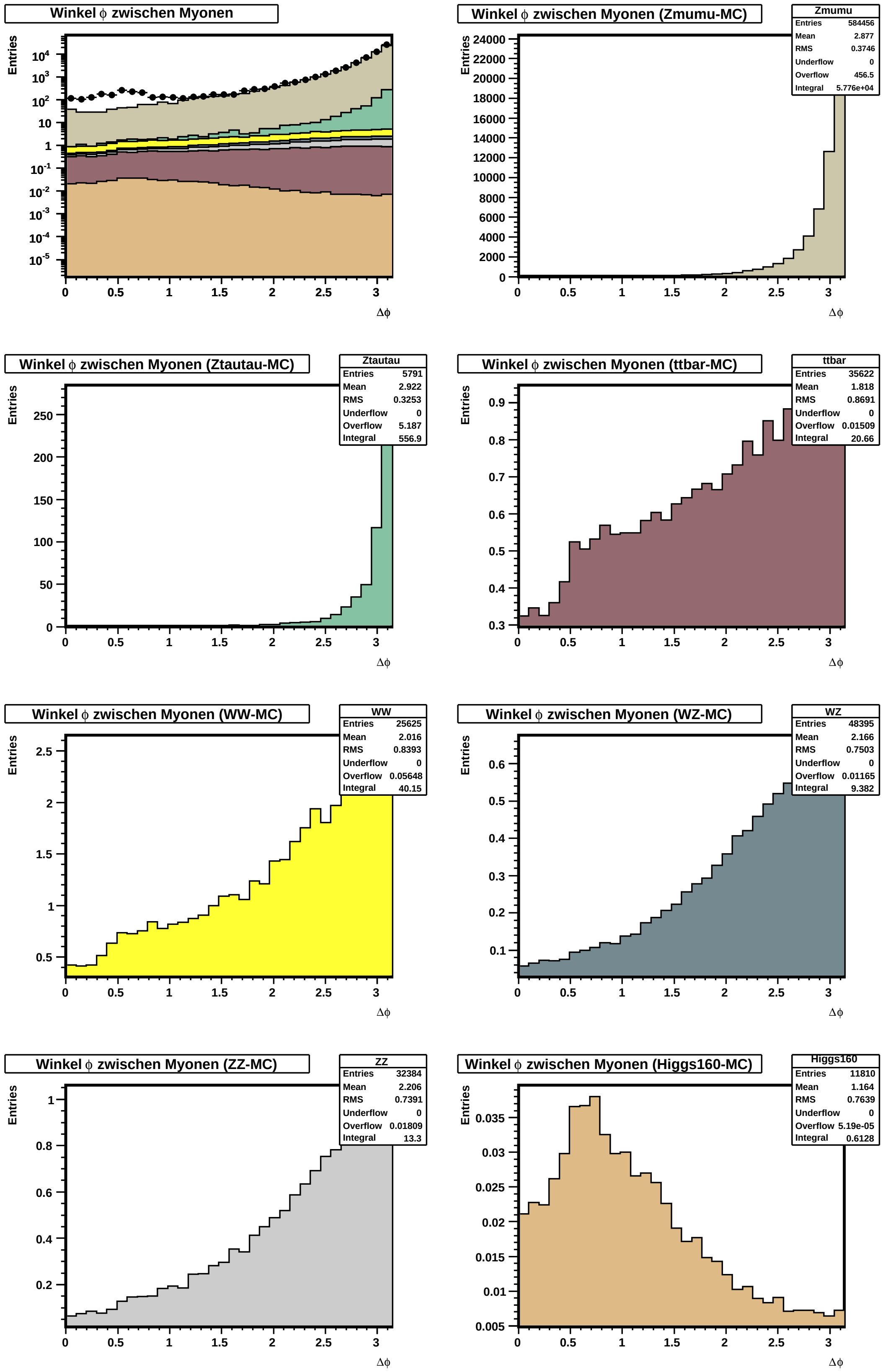

Abbildung 6.9: In dieser Abbildung sind die Winkelverteilungen $\Delta\phi_{\mu\mu}$ zwischen den Myonen aufgetragen. Deutlich erkennbar hebt sich die Gegenläufigkeit des Signals vom Untergrund ab. Der Überschuß in den Daten bei kleinen Öffnungswinkeln muß wieder dem Υ-Prozess zugeordnet werden, der in dieser Arbeit wegen seiner vollständigen Separierbarkeit nicht explizit implementiert wurde.

Transversale Masse

Die transversale Masse M_T ist eine weitere Größe, die hinsichtlich ihres Einsatzes als möglicher Lern- und Trainingskandidat für Neuronale Netze überprüft werden muß. Sie berechnet sich aus der Summe der Transversalimpulse beider Myonen, der fehlenden Transversalenergie sowie dem Öffnungswinkel, den die Ebene in der beide Myonen liegen mit der Ebene der fehlenden transversalen Energie einschließt.

$$M_T^{\mu\mu} = \sqrt{2 \cdot p_T^{\mu\mu} \cdot (1 - \cos\angle(\mu\mu, \not{E}_T\ ))}$$

Ihr Verlauf für Signal- und Untergrundprozesse, dargestellt in Abbildung 6.10, weist einige Charakteristika auf, die sich für eine Selektion ausnutzen lassen sollten. Dies gilt besonders für die Prozesse $Z/\gamma^* \longrightarrow \mu^+\mu^-$ und $Z/\gamma^* \longrightarrow \tau^+\tau^-$, die im Gegensatz zum Prozess $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ ihren erheblichen Ereignisanteil im Massenbereich bis 60 GeV aufweisen. Das Signal, welches in diesem Bereich langsam ansteigt und den Hauptteil der Ereignisse zwischen 100 GeV bis 160 GeV zeigt, fällt kurz vor der Higgs-Masse (160 GeV) steil ab. Das liegt daran, daß keine transversale Masse größer der Higgs-Masse kinematisch möglich sein kann, da die in die Berechnung eingehenden Größen, p_T und $\not{E}_T$, durch die an der Reaktion beteiligten Teilchen begrenzt sind.

Summe der transversalen Energie aller Jets

Die Summe der transversalen Energie über alle Jets H_T stellt, in besonderer Hinsicht auf den Untergrund aus $t\bar{t}$-Produktionen, eine überaus nützliche Netzvariable dar. Abbildung 6.11 kann dies weiter verdeutlichen. Zu sehen sind die Verteilungen von Signal und Untergrund, von denen lediglich die $t\bar{t}$-Monte-Carlo-Simulationen im Vergleich einen signifikant separablen Verlauf aufweist. Dieser eignet sich daher besonders, ihn als möglich Netzvariable zu implementieren, um das Signal von diesem speziellen Untergrund abzuheben.

Skalierte fehlende transversale Energie

Zu weiteren Selektionszwecken wurde auch die Größe der skalierten fehlenden transversalen Energie $\not{E}_T\ ^{skal}$ untersucht. Dazu betrachtet man die Größe $\not{E}_T$. Sofern diese einen erheblichen Beitrag an falsch gemessener Jetenergie aufweist, können die Fluktuationen der Jetenergiemessung in der transversalen Ebene über $\Delta E^{Jet} \cdot \sin\theta^{Jet}$ approximiert werden. Dabei ist ΔE^{Jet} proportional zu $\sqrt{E^{Jet}}$. Auch der Öffnungswinkel zwischen der projizierten Energiefluktuation und der fehlenden transversalen Energie $\angle(Jet, \not{E}_T\ )$ bietet sich an, um den Beitrag der Jets zur fehlenden transversalen Energie zu bestimmen.

Die Größe der skalierten fehlenden transversalen Energie definiert sich demnach wie folgt:

$$\not{E}_T\ ^{skal} = \frac{\not{E}_T}{\sqrt{\sum_{Jets}(\Delta E^{Jet} \cdot \sin\theta^{Jet} \cdot \cos\angle(Jet, \not{E}_T\ ))^2}} \quad .$$

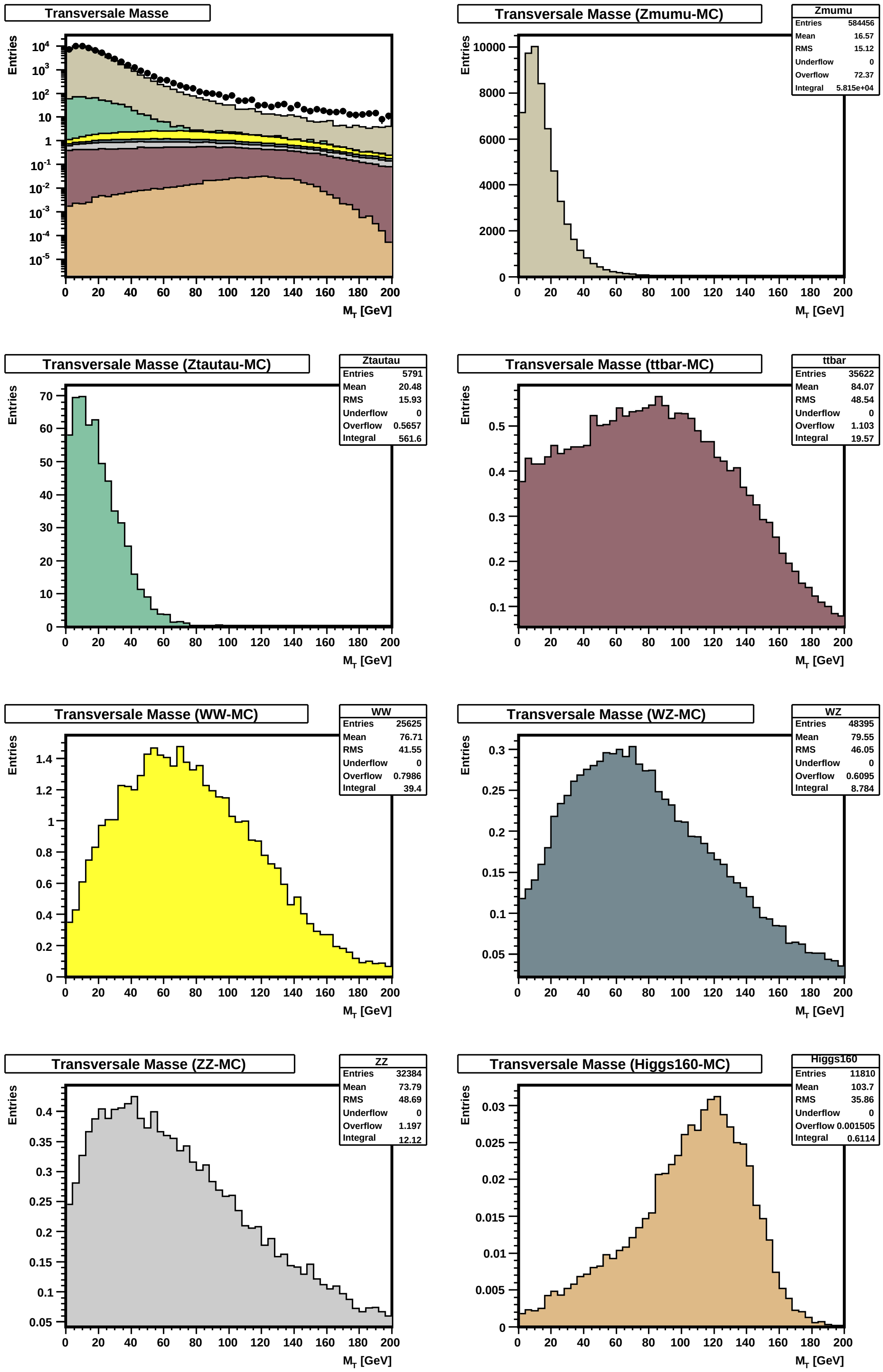

Abbildung 6.10: Die Abbildung zeigt die Verteilungen der transversalen Massen. Eine logarithmische Darstellung für die Darstellung aller Prozesse wurde gewählt. Auch hier ist der Überschuß bei hohen Werten der transversalen Masse der schlechteren Rekonstruktion der Ereignisse mit null SMT-Treffern zuzuordnen.

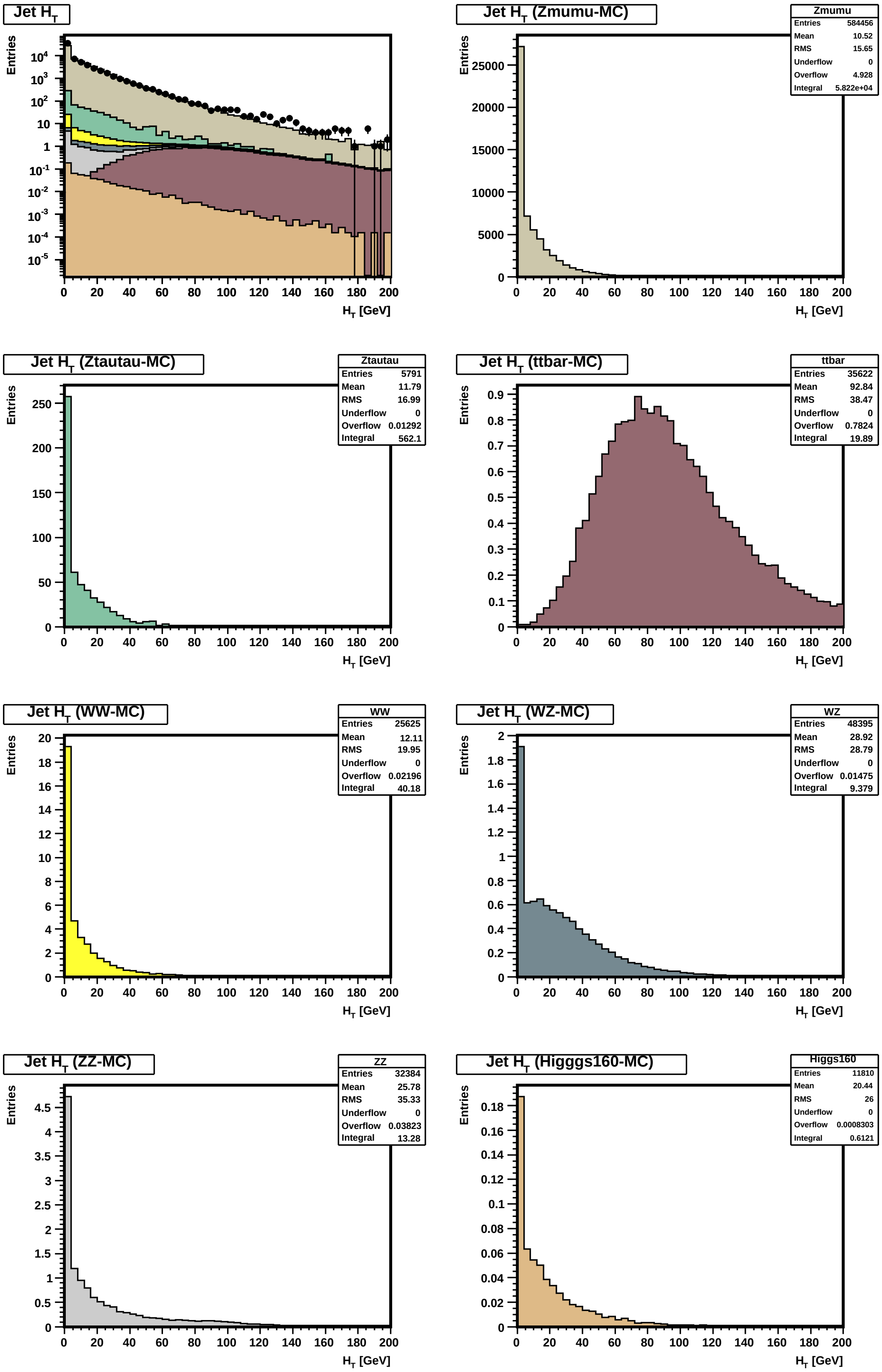

Abbildung 6.11: Die Verteilungen in dieser Abbildung zeigen den Verlauf der Summe der transversalen Energie über alle Jets. Deutlich hervorzuheben ist der Untergrund aus $t\bar{t}$-Produktionen.

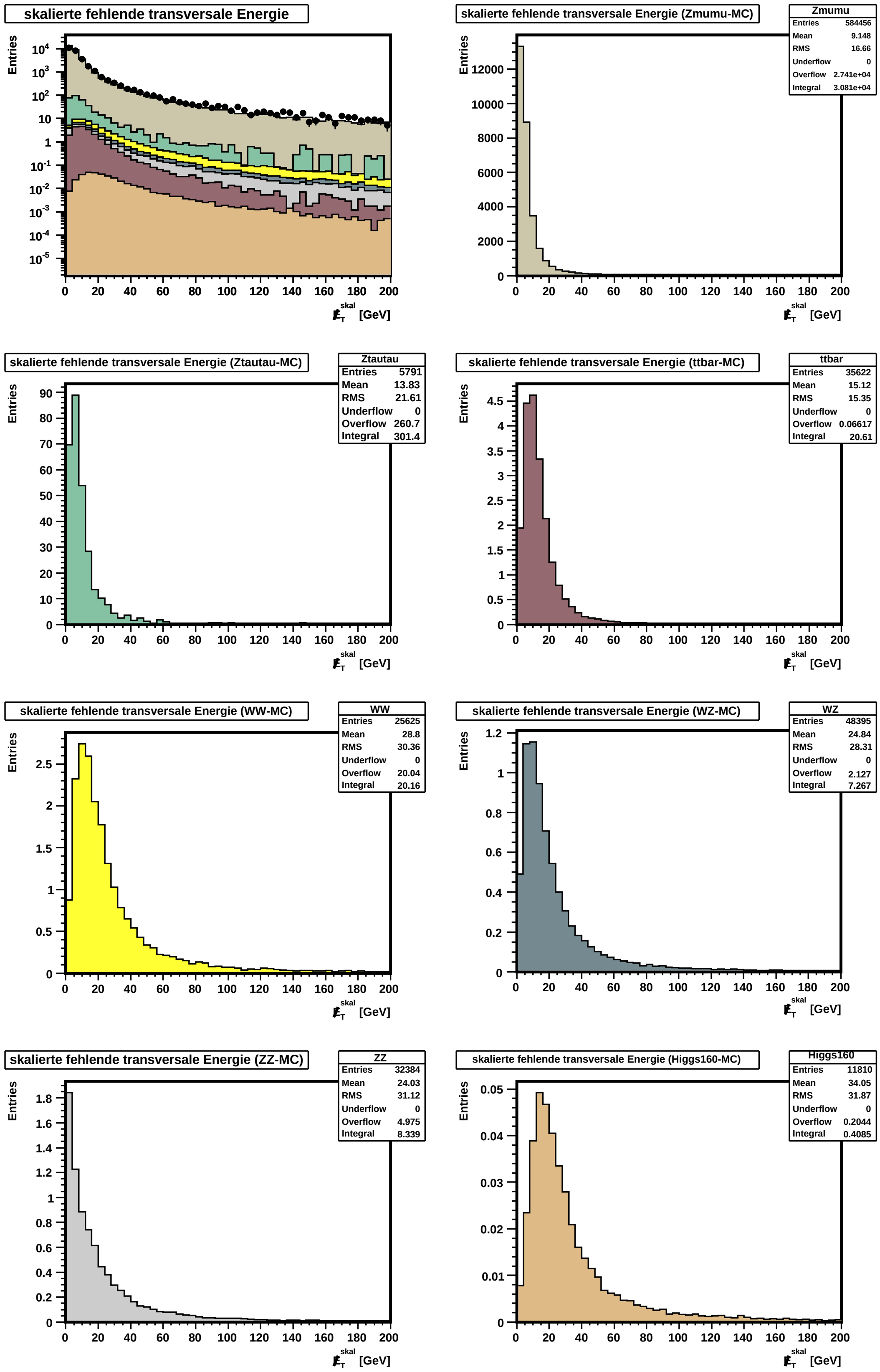

Abbildung 6.12: Die Abbildung zeigt die Verteilungen der skalierten fehlenden transversalen Energie, welche als weitere Netzvariable in Betracht gezogen wird. Für das erste Bild wurde die logarithmische Darstellungsskala gewählt.

In Abbildung 6.12 sind ihre Verteilungen für die jeweiligen Prozesse dargestellt. Es läßt sich erkennen, daß gerade im Energiebereich unter 15 *GeV* und über 120 *GeV* die Möglichkeiten zur Signalselektion gegeben sind.

Winkel zwischen fehlender transversaler Energie und der ϕ-Koordinate des zweiten Myons

Daß der Winkel zwischen $\not{E}_T$ und der ϕ-Koordinate des zweiten Myons sich in

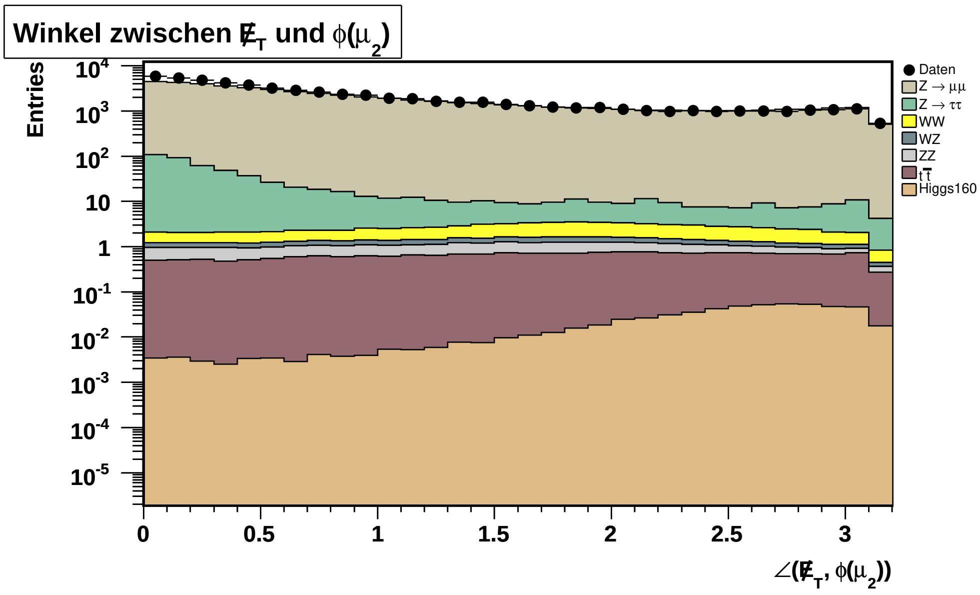

Abbildung 6.13: Die Verteilungen der Abbildung geben den Öffnungswinkel zwischen $\not{E}_T$ und ϕ-Koordinate des zweiten Myons wieder. Die Darstellung erfolgt mit logarithmisch und zeigt eine deutliche Signatur zwischen Signal und Untergrund.

derselben Art und Weise als Netzvariable eignet, wie dies bereits im Abschnitt über den Öffnungswinkel zwischen Myonen verdeutlicht wurde, kann in Abbildung 6.13 verglichen werden. Die Werte dieser Größe zeigen eindeutige gegenläufige Tendenzen, die für ein Neuronales Netz ein sehr gutes Lern- und Trainingsbeispiel darstellen. Um die Abhängigkeit, also Korrelationen verschiedener Größen untereinander näher darzustellen, ist in Abbildung 6.15 diesselbe Größe lediglich mit der ϕ-Komponente des führenden Myons aufgetragen. Beide Verteilungen sind mit einer logarithmischen Skala abbgebildet. Im Signal-Monte-Carlo wird ersichtlich, daß beide Winkel unabhängig vom Myon sehr weit geöffnet sind. Dieser Effekt kann jedoch nur für das zweite Myon genutzt werden, da nur für dieses Myon die Produktionen der Untergrundprozesse eher kleine Winkel aufweisen.
Zur graphischen Veranschaulichung soll die Korrelation zwischen der ϕ-Koordinate von $\not{E}_T$ und der ϕ-Koordinate des zweiten Myons, deren Einfluß sich ähnlich der eben beschriebenen Größe äußert, in einer zweidimensionalen Darstellung gezeigt werden

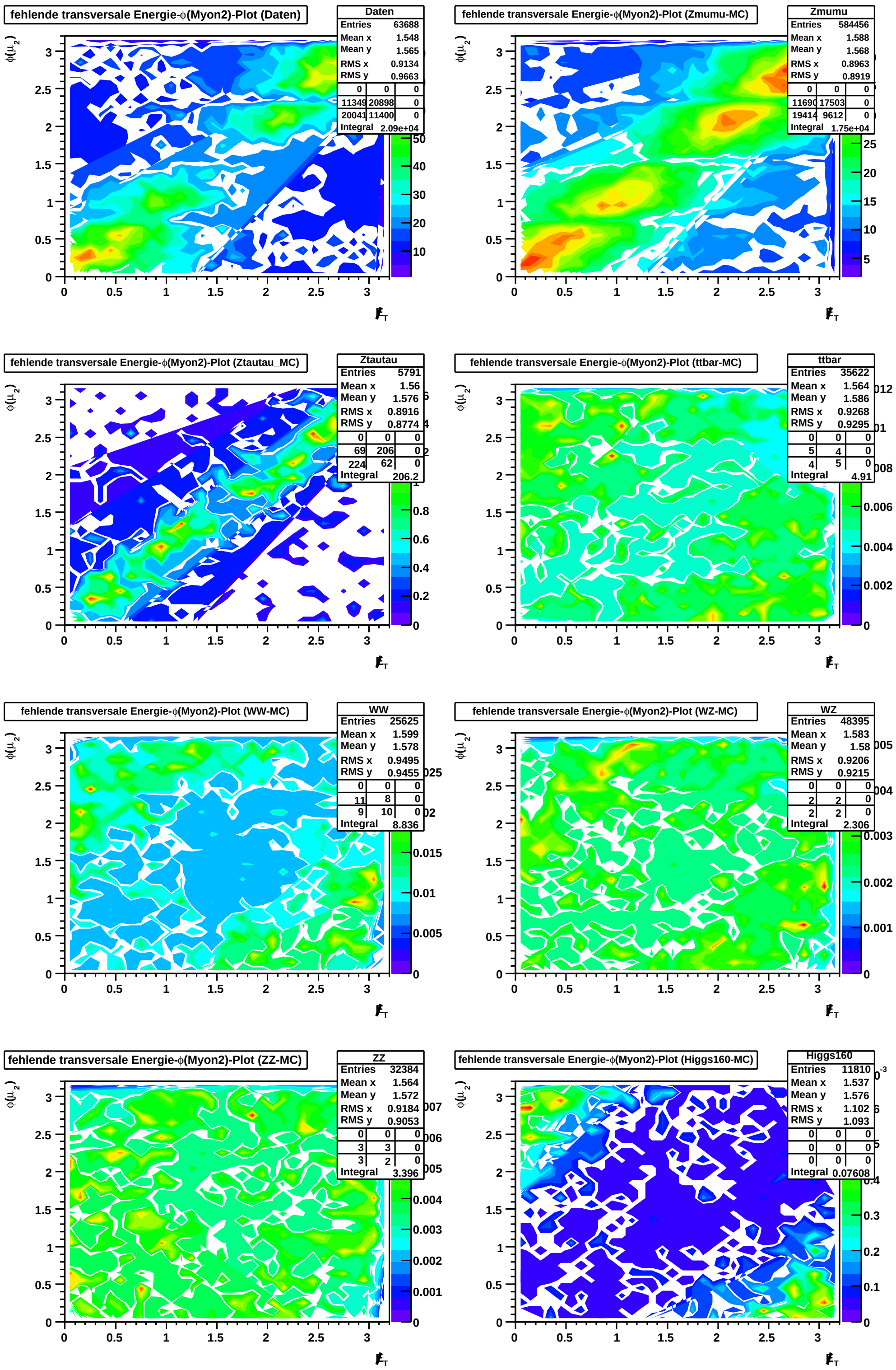

Abbildung 6.14: In dieser zweidimensionalen Darstellung sind die ϕ-Koordinaten der fehlenden transversalen Energie und des zweiten Myons gegeneinander für die Daten (oben links), das Signal (unten rechts) sowie den hier betrachteten Untergrundprozessen aufgetragen.

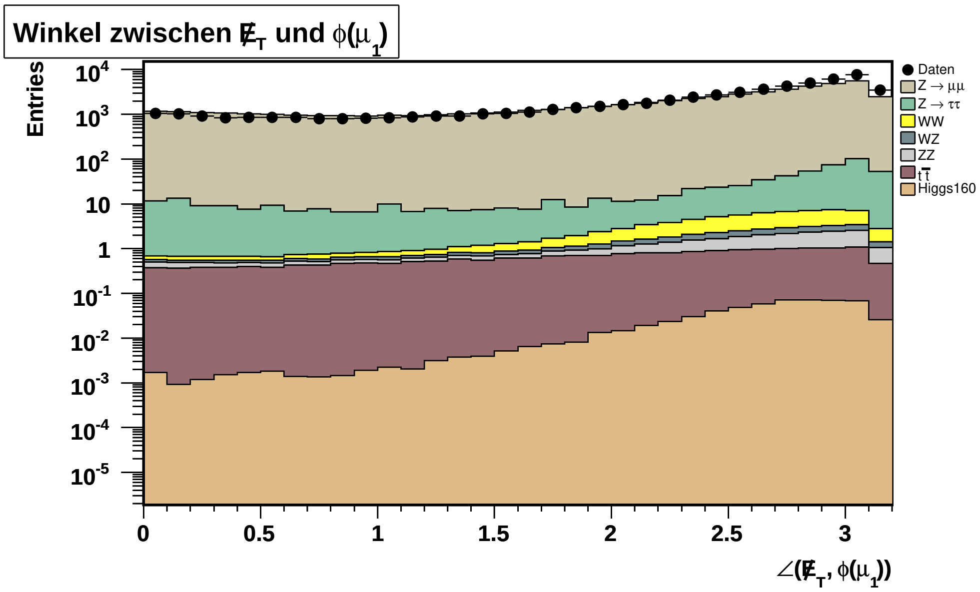

Abbildung 6.15: Die Verteilungen der Abbildung geben den Öffnungswinkel zwischen $\not\!\!E_T$ und ϕ-Koordinate des ersten Myons wieder. Die Darstellung erfolgt logarithmisch und zeigt im Gegensatz zur vorigen Abbildung keine nennenswerten Signaturen zur Ereignisselektion.

(Abbildung 6.14). Auch hier wird deutlich, daß die Verteilungen zwischen Signal und Untergrund variieren. Korrelationen dieser Art sind von Bedeutung, da sich das Neuronale Netz aus ihnen gewinnbringende Informationen ziehen kann.

Beginnend ist zu allen hier vorgestellten und verwendeten Neuronalen Netzen zu bemerken, daß darauf geachtet wurde, daß Trainings- und Testereignisse orthogonal zueinander sind. Dazu wurde jedes zweite Ereignis für das Training benutzt und die verbleibenden restlichen Ereignisse zum Testen herangezogen. Folglich teilte sich die zur Verfügung stehende Ereignismenge zu je 50% auf Trainings- und Testmege auf.

6.4.2 Das WW-Netz

Nachdem einige Variablen vorgestellt wurden, die als potentielle Netzvariablen in Betracht kommen, soll in den folgenden Abschnitten näher auf die einzelnen Netze an sich eingegangen werden. Da sich herausgestellt hat, daß aufgrund der Fülle an Informationen ein einziges Neuronales Netz die Selektion dieser Analyse nicht effizient bewältigen kann, werden entsprechend der eingebundenen Reaktionen aus Untergrundprozessen mehrere Neuronale Netze entworfen und ihrem jeweiligen Simulationsdatensatz gemäß aufgebaut und trainiert.
Mit der Darstellung des Neuronalen Netzes, speziell entwickelt hinsichtlich des WW-Untergrundes, soll hier begonnen werden. Aufgrund der kinematischen und topologischen Vergleichbarkeit von Signal-Prozess und Untergrund ist es in diesem Netz besonders interessant, beide zu separieren. Dies wird später auch im Neuronalen Netz

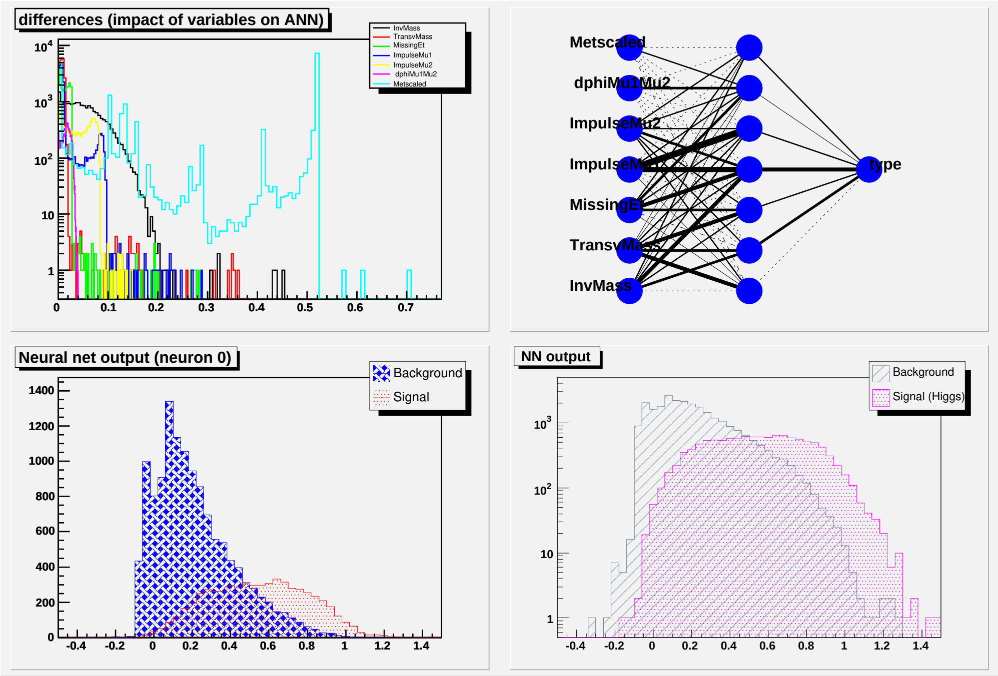

Abbildung 6.16: In dieser Abbildung wird ein Überblick über verschiedene Netzkomponenten des WW-Netzes gegeben. Links oben ist der Einfluß der Variablen auf das Netz gezeigt, er bestimmt sich aus den betragsmäßig größten Differenzen der Gewichte untereinander. Diese Differenz wird für jedes Ereignis neu berechnet. Je kleiner ihr Wert, desto weniger trägt diese Variable zum Netz bei. Daneben befindet sich ein Bild der Netzstruktur mit sieben Eingabeneuronen, einer verdeckten Schicht mit 7 Neuronen und der Ausgabeschicht mit dem Ausgabeneuron, welches die Ereignisse typspezifisch sortiert (mit 1 für das Signal (Higgs) und 0 für den Untergrund (hier WW)). Desweiteren sind auch die gewichteten Verbindungen zwischen den Neuronen dargestellt. Die Stärke der Linien im Bild symbolisiert die Stärke der Gewichtung. Darunter sind die Netzausgaben zu sehen linkerhand für die trainierten und rechterhand (mit logarithmischer Skala) für trainierte und getestete Ereignisse.

deutlich werden, denn die Möglichkeiten einer Selektion mittels des WW-Netzes sind besonders in Hinsicht auf 2-dimensionale Variabeln sehr aussichtsreich. Es besteht die Aussicht, daß gerade gegenüber 1-dimensionalen Variablen das WW-Netz die Korrelation in zwei Dimensionen besser erkennen soll. Es ist zu erwarten, daß sich damit Erfolge in einer Selektion erzielen lassen, der WW-Untergrund aber trotzdem eine irreduzible Größe dieser Analyse bleibt.
Nachdem die in Kapitel 6.1 beschriebenen Variablen in Bezug auf eine Verwertung geprüft und getestet wurden, konnten genau die Variablen für ein späteres Netz gewählt werden, deren Einfluß auf dieses Netz innerhalb des Trainings am größten gewesen ist (siehe dazu Abbildung 6.16 oberes linkes Bild). Für jedes Ereignis wird demnach die dem Betrage nach größte Differenz zwischen den gewichteten Verbindungen der Neuronen untereinander gebildet. Je höher die Differenzen ausfallen, desto stärker beeinflußt die zugehörige Variable das Neuronale Netz. Für den Lernprozess des Neuronalen Netzes werden entsprechend die Variablen verwendet, für die über mehrere Trainings- und Testphasen die Differenz einen, im Vergleich zu allen Variablen, deut-

lichen Beitrag liefert. Die gewählten Variablen setzen sich aus

- invarianter Masse $M_{\mu\mu}$
- transversaler Masse M_T
- fehlender transversaler Energie $\not{E}_T$
- Impuls des ersten Myons $p_T^{\mu_1}$
- Impuls des zweiten Myons $p_T^{\mu_2}$
- Öffnungswinkel zwischen den Myonen $\Delta\phi_{\mu\mu}$

und

- skalierter fehlender transversaler Energie $\not{E}_T{}^{skal}$

zusammen. Die Werte dieser Größen werden an die Eingabeneuronen übergeben, folglich befinden sich in der Eingabeschicht sieben Eingabeneuronen. Das Neuronale Netz durchläuft dann seinen Lernalgorithmus, wie er in Kapitel 6.3 beschrieben ist. In der verdeckten Schicht sind in diesem Netz sieben verdeckte Neuronen implementiert, da sich gerade diese Anzahl in einem Testalgorithmus als besonders effektiv erwiesen hat. Eine Selektion mit Neuronalen Netzes zielt letztendlich darauf, eine Typisierung der Ereignisse vornehmen zu können. Dafür kann mit Hilfe zweier Kategorien die Einteilung nach Signal (rote bzw. magenta-farbene Verteilungen) und Untergrund (blau bzw. grau dargestellt) erfolgen. Zuordnungen in Richtung 1 werden für signaltypische Ereignisse verwendet, die 0 weist auf Ereignisse mit Untergrundsignatur hin. Je nachdem wie sehr ein Ereignis signal- oder untergrundartig war, wird es auf der Skala eher rechts bzw. eher links eingestuft. Daß auch Werte kleiner 0 und größer 1 in den Verteilungen zustande kommen, liegt daran, daß die Variablen, welche den Lernprozess maßgeblich beeinflußen weder beschränkt noch normiert sind. Da das Neuronale Netz jedes zweite Ereignis dem Trainingsprozess unterzieht und die dazwischenliegenden Ereignisse für seine Testphase benutzt, kann es vorkommen, daß einzelne Ereignisse des Trainings- und des Testalgorithmus so unterschiedliche Werte in ihren Variablen aufweisen, daß diese entsprechend ihrer Gewichtungen im Netz zu Werten über 1 und unter 0 führen.

Während der gesamten Lern- und Testphase kann das Neuronale Netz überwacht werden, indem der Verlauf der Fehlerkurven (siehe Kap. 6.3.2) für Trainings- und Testsatz beobachtet wird. In dieser Ausgabe werden alle 10 Epochen die Fehlerwerte aktualisiert. Das Neuronale Netz will diesen Fehler minimieren, möglichst so, daß dabei ein globales Minimum gefunden wird. Dazu werden Kurven mit stetig sinkendem Fehler erwartet. Dabei sollte sich der Verlauf von Trainingsfehlerkurve und Testfehlerkurve sehr stark annähern und auf keinen Fall wieder divergieren, denn das wäre ein Anzeichen beginnenden Übertrainings. Speziell für den WW-Untergrund ist in Abbildung 6.17 der Verlauf beider Kurven dargestellt. In blau werden die trainierten Ereignisse beschrieben, in rot die getesteten Ereignisse. Die Minimierung des Fehlers ist in den Kurven erkennbar, jedoch liegen beide Kurven am Ende des Trainings in wesentlich schlechterer Übereinstimmung als zu ihrem Beginn. Daraus kann geschlußfolgert werden, daß die Generalisierung des Netzes nicht eineindeutig auf Testereignisse angewendet werden konnte. Damit können besondere Kennzeichen zur

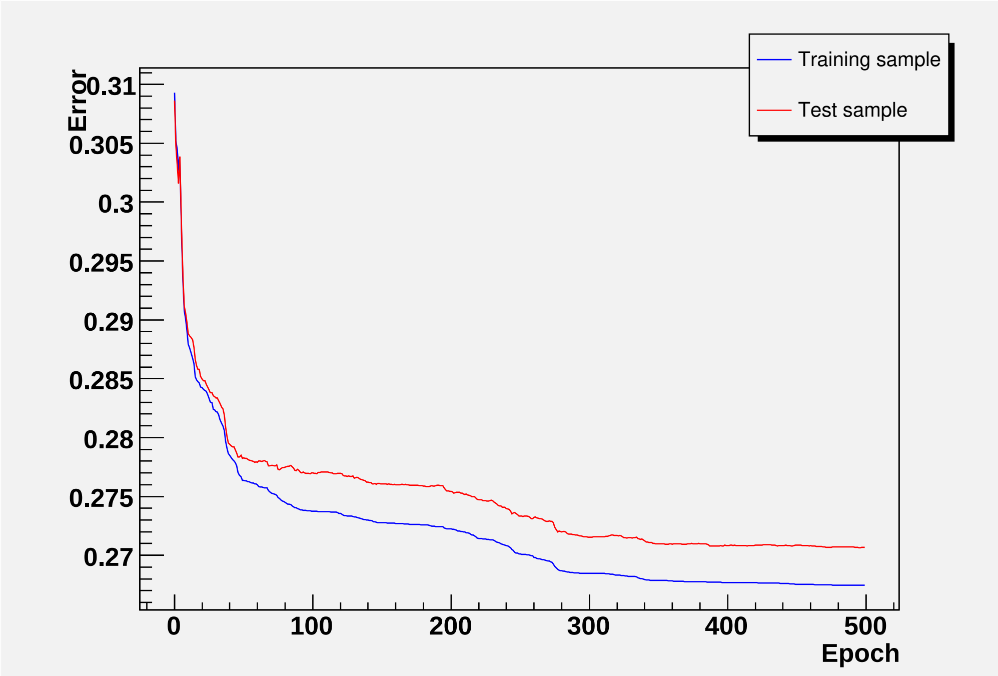

Abbildung 6.17: Die Graphik zeigt den zeitlichen Verlauf der Fehlerkurven für Trainings- (blau) und Testdaten (rot). Die Daten des WW-Netzes wurden dieser Graphik zugrunde gelegt.

Generalisierungsfähigkeit des Netzes aus dem Kurvenverlauf gezogen werden, denn je eindeutiger die Übereinstimmung im zeitlichen Verlauf, desto größer die Generalisierungsfähigkeit und spätere Selektion. Erweist sich der Verlauf der Fehlerkurven als stabil, so liefern auch die Netzausgaben sinnvolle Ergebnisse und die Trainings- und Testphase ist abgeschlossen. Die Daten des Netzes (Werte der Neuronen, Synapsen und Gewichte) werden extern in einem Programm gespeichert, welches für die eigentliche Selektion herangezogen werden kann.

6.4.3 Das WZ-Netz

Die Prozedur, nach welcher in der Trainings- und Testphase verfahren wird, wurde bereits im vorigen Abschnitt erläutert. Das Verfahren zur Bestimmung der für das Netz geeigneten Variablen, die Lernphase an sich sowie die gesamte Testphase werden untergrundübergreifend auf alle Neuronalen Netze in gleicher Weise angewandt. Dabei können und werden sich die einzelnen Ergebnisse voneinander unterscheiden.
Im Rahmen dieses Abschnitts soll das WZ-Netz näher beleuchtet werden. Anders als bei WW-Prozessen, deren Zerfallsstruktur in überdeckender Ähnlichkeit zum Signalprozess analysiert wurde, sind bei WZ-Reaktionen aus Untergrundprozessen zwar noch gemeinsame kinematische wie topologische Sequenzen erkennbar, jedoch auch Unterschiede zu verdeutlichen. Diese kristallisieren sich besonders durch das WZ-Paar heraus, welches nicht dem gleichen Endzustand wie der Signalprozess entspricht. Ein Z-Boson im Vergleich zum W-Boson ist ein elektrisch neutrales Teilchen, dessen Zerfall in ein Lepton-Antilepton-Paar gerade für diese Analyse von Bedeutung ist. In die-

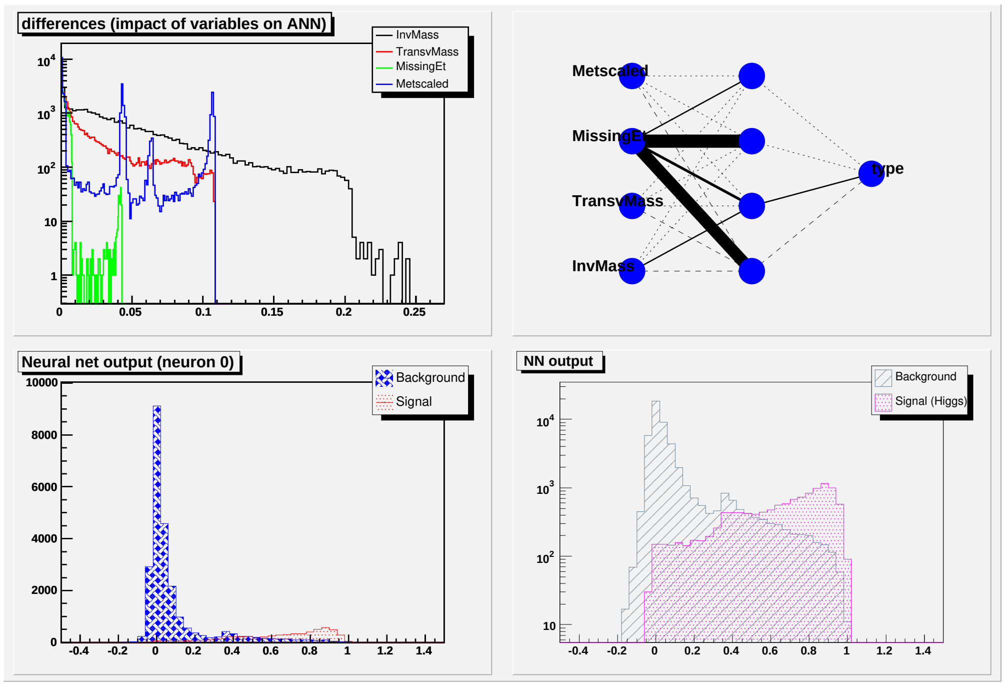

Abbildung 6.18: Die Abbildungen zeigen Aufbau und Verhalten des WZ-Netzes. Links sind die Variablen zu sehen, deren Einfluß auf das Netz als signifikant betrachtet wurde. Ihr Einfluß während der Lernphase ist schematisch dargestellt. Daneben ist der strukturielle Aufbau des Netzes abgebildet. Deutlich erkennbar sind die einzelnen Neuronen sowie die gewichteten Verbindungen unter ihnen. Die Eingabeschicht besteht diesmal aus den vier Neuronen für die Eingabevariablen. In der verborgenen Schicht sind weitere vier Neuronen untergebracht, bevor abschließend in der Ausgabeschicht mittels des Ausgabeneurons die Typisierung in Untergrund bzw. Signal vorgenommen wird. Die Ausgaben des Netzes sind darunter für den Trainingssatz (links) und für den Satz aus Trainings- und Testdaten (rechts mit logarithmischer Skala) dargelegt.

sem Fall weist der Endzustand dieser Untergrundproduktion nur ein Neutrino aus dem W-Zerfall auf. Daraus ist zu schlußfolgern, daß weniger fehlende transversale Energie $\not{E}_T$ im Vergleich zum Signal zu erwarten ist. Diese Variable und Variablen, die sich aus $\not{E}_T$ berechnen, sollten demnach für ein Netztraining mit Ereignissen dieses Untergrundes geeignet sein (siehe dazu Abbildung 6.18). Der speziell für dieses Netz gewählte Satz an Variablen setzt sich aus

- invarianter Masse $M_{\mu\mu}$
- transversaler Masse M_T
- fehlender transversaler Energie $\not{E}_T$

und

- skalierter fehlender transversaler Energie $\not{E}_T{}^{skal}$

zusammen. Ihr Einfluß auf das Netz stellt die linke obere Abbildung in 6.18 dar. Daneben kann der Aufbau des Netzes betrachtet werden. Aufgrund der vier Variablen muß

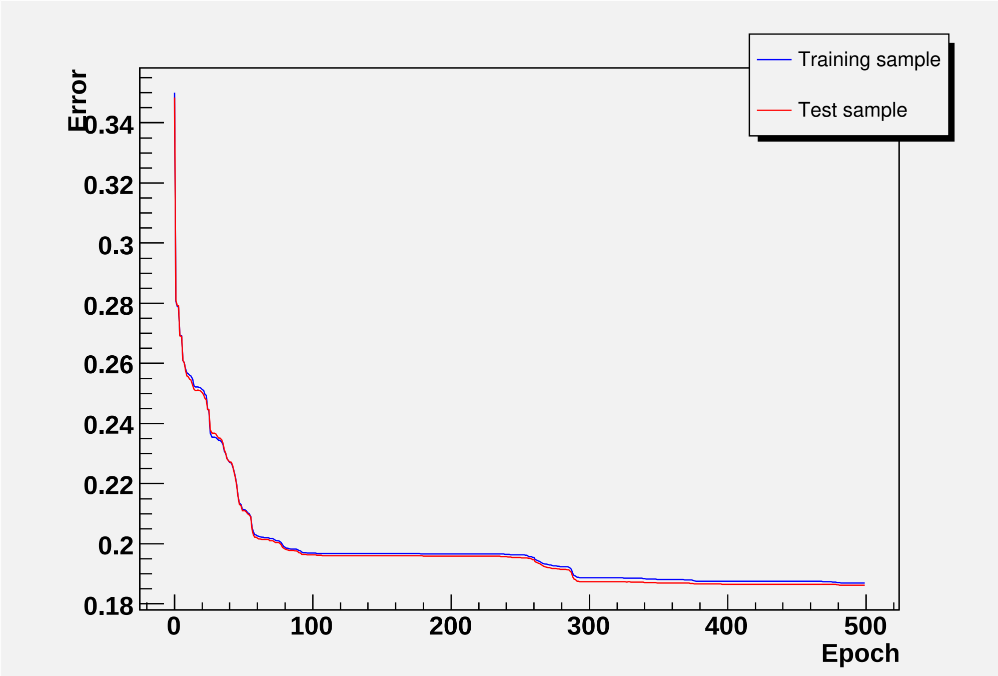

Abbildung 6.19: Im zeitlichen Verlauf dieser Abbildung wird die Minimierung der Fehler für Trainingsdaten (blau) und Testdaten (rot) dargestellt. Die Graphik basiert auf Angaben des WZ-Netzes.

die Eingabeschicht aus vier Neuronen bestehen, welche die Werte der Variablen an die verdeckte Schicht, entsprechend der Gewichte zwischen den Neuronen, weiterleiten. Die verborgene Schicht des WZ-Netzes wurde mit vier verborgenen Neuronen als besonders effektiv getestet.
Da das Neuronale Netz sein Wissen in den gewichteten Verbindungen speichert und durch ihre Aktualisierung der Lernprozess fortschreitet, ist der Erfolg einer Selektion entscheidend von den gewichteten Verbindungen zwischen den Neuronen der einzelnen Schichten abhängig. Wie stark diese im Falle des WZ-Netzes ausgeprägt sind, zeigt in 6.18 die Graphik rechts oben. Je stärker die Verbindung zwischen zwei Neuronen dargestellt ist, desto stärker wird das anknüpfende Neuron mit der Information zum Handeln aktiviert. Damit kann aber noch keine Aussage zum Beitrag der zugehörigen Variable getroffen werden. Die Abbildung 6.18 zeigt lediglich wie das Gewicht zwischen zwei Neuronen geprägt ist (z.B Variable *MissingEt*), will man wissen wie sich ihr Einfluß auf das gesamte Netz gestaltet, müssen die Differenzen zwischen den Gewichten dieser Variable gebildet werden. Für *MissingEt* ist diese Differenz kleiner als z.B. für *InvMass*, demnach übt die invariante Masse (*InvMass*) einen stärkeren Einfluß auf das WZ-Netz aus als die fehlende transversale Energie (*MissingEt*).
Die Ergebnisse des Neuronalen Netzes werden in den Netzausgaben wiedergegeben, welche die Fallunterscheidung zwischen Untergrund und Signal beeinhalten. Wie sich der Verlauf des aktiven Trainings, also die Minimierung des Fehlers in Abhängigkeit von der Trainingsdauer gestaltet hat, kann Abbildung 6.19 entnommen werden. Es ist deutlich zu erkennen, daß der Fehler über die 500 Trainingsepochen hinweg stetig klei-

ner wird und letzenendes einen minimalen Wert erreicht. Besonders günstig in diesem Trainings- und Testalgorithmus ist der sehr gut übereinstimmende Verlauf beider Kurven, welche fast identisch liegen. Das bedeutet, daß die Trainingsereignisse (etwa die Hälfte aller dem Netz zur Verfügung stehenden Ereignisse) des Neuronalen Netzes auf Basis der WZ-Untergrundproduktionen diese Untergrundprozesse im gesamten sehr gut beschreiben, so daß die Generalisierung auf die Testdaten besonders gut ausfällt. Dieser Effekt wird auch in den Verteilungen der Signal- und Untergrundereignisse deutlich. Hier sind die Verteilungen für Trainingsdaten (Bild unten links in Abb. 6.18, lineare Darstellung) und die Verteilungen für Trainings- und Testdaten (unten rechts in Abb. 6.18, logarithmische Darstellung) kaum unterscheidbar, so daß dies ein Indiz für sehr gut übereinstimmende Ereignisse beider Mengen ist.
Nach Abschluß der Trainings- und Testphase müssen auch von diesem Netz die numerischen Werte der Neuronen und Synapsen zu späteren Selektionszwecken festgehalten werden. Sie werden in eine Datei exportiert. Auf die Erfahrungen und Ergebnisse des Netzes kann so jederzeit wieder zugegriffen werden.

6.4.4 Das ZZ-Netz

Im folgenden Abschnitt wird ein Neuronales Netz hinsichtlich des ZZ-Untergrundes erläutert. Wie bereits im vorigen Abschnitt festgestellt wurde, ist der für diese Analyse relevante Zerfall eines neutralen Z-Bosons in ein Lepton und sein dazugehöriges Antilepton (die Ladungssumme beider Teilchen muß neutral, also 0 bleiben) bestimmt. Z-Zerfälle in Quarks, welche daneben sehr häufig auftreten sind hier nicht betrachtet worden. ZZ-Produktionen aus Untergrundprozessen zerfallen demnach in zwei Leptonen und zwei Antileptonen. Demnach sind in erster Zerfallsstufe keine Neutrinos über die fehlende transversale Energie nachweisbar. Diese Größe sollte sich dementsprechend stark von der des Signalprozesses unterscheiden. Wird rückblickend noch einmal Abbildung 6.8 betrachtet, ist festzustellen, daß gerade im Energiebereich bis 60 GeV ein erheblicher Anteil an fehlender transversaler Energie rekonstruiert werden kann.

Eine Antwort findet sich, wenn weitere Zerfälle der Leptonen in Betracht gezogen werden. Beispielsweise gibt es die Möglichkeit, daß ein Z-Boson in ein $\tau^+\tau^-$-Paar zerfällt. Dies ist das schwerste der drei Leptonen und auch dasjenige, welches am schnellsten weiter zerfällt. Durch den Sekundärzerfall des τ-Leptons wiederum können Myon, Myon-Neutrino und Tau-Neutrino entstehen. Das bedeutet, daß die in diesem Untergrundprozess rekonstruierte $\not{E}_T$ aus Neutrinos des Sekundärzerfalls stammt. Aufgrund der Zerfallshierarchie läßt sich weniger fehlende transversale Energie der Neutrinos aus dem ZZ-Zerfall nachweisen, als dies bei $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ gegeben ist. Die Maxima der $\not{E}_T$ -Verteilungen zwischen $ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ und dem Signal sind, wie in Abb. 6.8 ersichtlich gegeneinander verschoben. Somit stellt die fehlende transversale Energie auch hier eine Möglichkeit hinsichtlich eines Neuronalen Netzes dar. Auch Größen, die von $\not{E}_T$ abhängen, sind in Betracht zu ziehen. Darüberhinaus, und das gilt auch für die bereits vorangestellten Netze, sollte in diesem Zusammenhang immer auch an die invariante Masse (Abb. 5.4) gedacht werden. Ganz besonders für Prozesse mit Beteiligung von Z-Bosonen ist die Resonanz dieses Teilchens um die 90 GeV zu beobachten. Diese Resonanz liegt energetisch gesehen deutlich hinter dem Maximum der invarianten Masse $M_{\mu\mu}$ des Signalprozesses. Somit

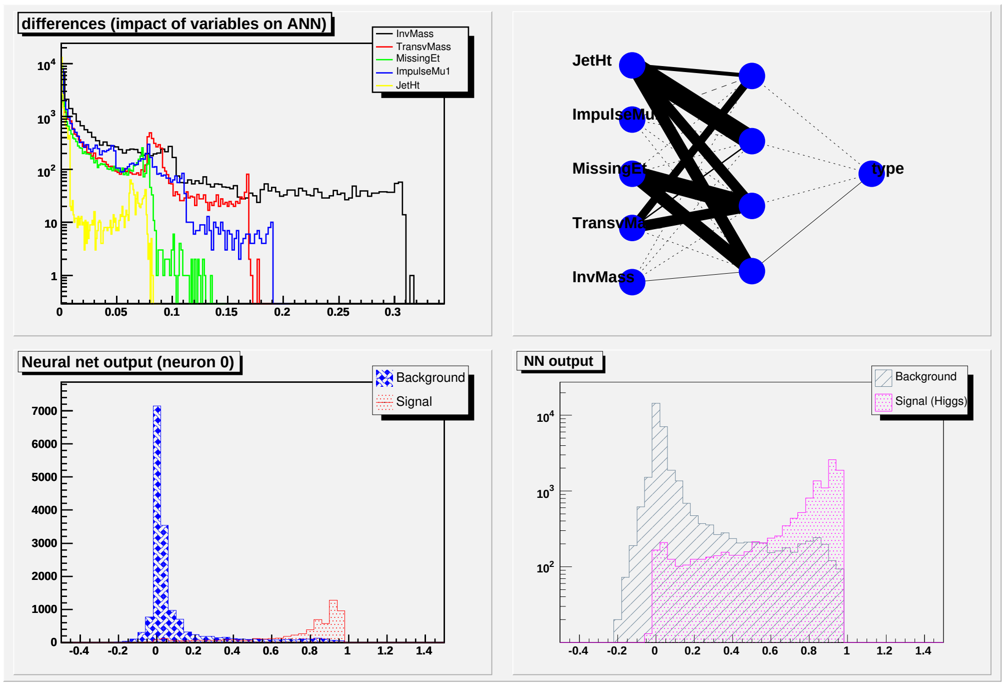

Abbildung 6.20: Diese Abbildungen geben Auskunft über Aufbau und Verhalten des *ZZ*-Netzes. Linkerhand ist der Einfluß der verwendeten Variablen auf das Netz gezeigt. Rechterhand daneben ist der strukturielle Aufbau des Netzes abgebildet. Deutlich erkennbar sind die einzelnen Neuronen sowie die gewichteten Verbindungen unter ihnen. Die Eingabeschicht besteht aus fünf Neuronen für die Eingabevariablen. In der verborgenen Schicht sind weitere vier Neuronen untergebracht, bevor abschließend in der Ausgabeschicht mittels des Ausgabeneurons die Typisierung in Untergrund bzw. Signal vorgenommen wird. Die Ausgaben des Netzes sind darunter für den Trainingssatz (links) und für den Satz aus Trainings- und Testdaten (rechts mit logarithmischer Skala) dargelegt.

sollte auch der Einfluß der invarianten Masse auf das Netz ausgenutzt werden. Für das *ZZ*-Netz sind insgesamt die Variablen

- invariante Masse $M_{\mu\mu}$
- transversale Masse M_T
- fehlende transversale Energie $\not{E}_T$
- Impuls des führenden Myons $p_T^{\mu_1}$

und

- die Summe transversaler Energie aus allen Jets H_T

für ein Netztraining ausgewählt worden. Die linke obere Abbildung in 6.20 stellt dar, welchen Einfluß die einzelnen Variablen auf das Neuronale Netz ausüben. Daneben befindet sich der Netzaufbau. Das *ZZ*-Netz setzt sich aus der Eingabeschicht mit fünf Neuronen (entsprechend der verwendeten Variablen), einer verdeckten Schicht mit vier

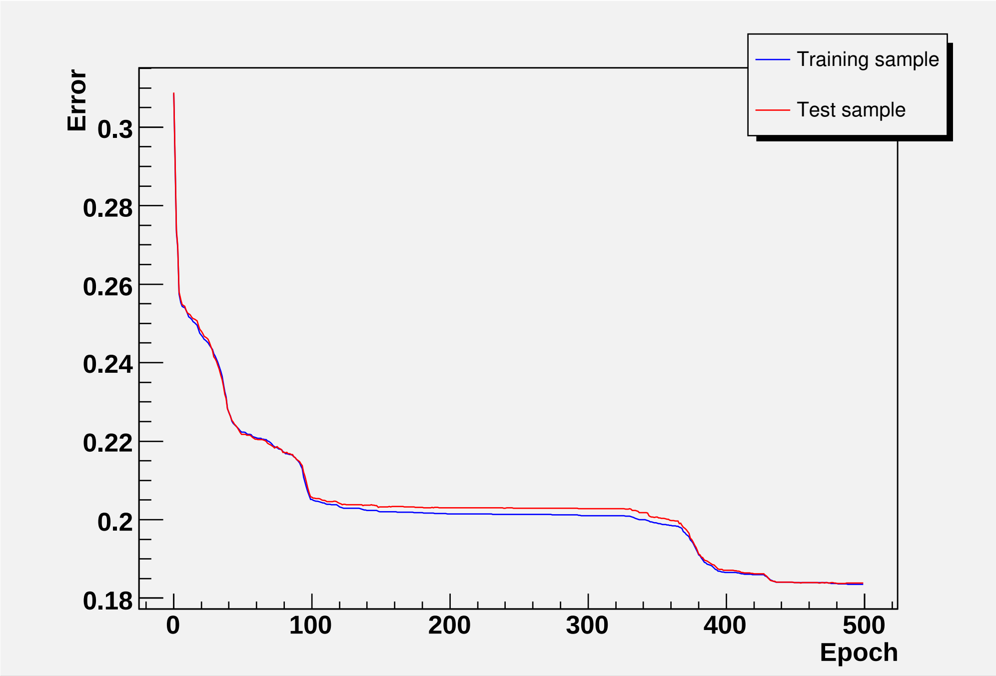

Abbildung 6.21: Im zeitlichen Verlauf dieser Abbildung wird die Minimierung der Fehler für Trainingsdaten (blau) und Testdaten (rot) dargestellt. Die Graphik basiert auf Angaben des ZZ-Netzes.

Neuronen (aufgrund der besonderen Effektivität dieser Anzahl an Neuronen) und der Ausgabeschicht mit nur einem Neuron zusammen. In der Ausgabeschicht werden wieder die Zuordnungen getroffen, d.h. es wird verglichen, ob ein Ereignis eher signaltypisch oder untergrundtypisch eingeordnet werden kann. Wie diese Zuordung letzenendes aussieht, kann den unteren Diagrammen in Abb. 6.20 entnommen werden. Links erfolgte die Zuordnung nur mit den Ereignissen des Trainingssatzes, rechts wurden die Daten von Training und Test verwendet. Die farbliche Darstellung der Untergrundprozesse erfolgte in blau bzw. grau, Signalereignisse sind in rot bzw. in magenta aufgetragen. Die Zuordnung in Richtung 0 bedeutet, daß ein Ereignis eher untergrundartiger Natur ist. Erfolgt eine Einteilung in Richtung 1, spricht man von signalähnlicher Signatur.
Der allgemeine Verlauf der Trainings- und Testphase, also der Verlauf der Fehlerminimierung im Training sowie im Test zeigen die Kurven in Abbildung 6.21. Zu sehen sind die zwei stetig fallenden Kurven der Testdaten (rot) und der Trainingsdaten (blau). Der anfänglich starke Abfall schwächt sich über die Trainingsepochen immer mehr ab, so daß dies als Indiz für das Erreichen eines Minimums angenommen werden kann. Dabei erfolgt die Minimierung des Fehlers, wie in Abbildung 6.21 ersichtlich, stufenweise. Hier werden durch das Netz andere Richtungen im Gradientenabstiegsverfahrens getestet, die in diesem Fall zu positiven Ergebnissen führen, indem der Fehler danach weiter minimiert werden kann. Die Trainingsergebnisse des Netzes werden für eine weitere Verarbeitung, speziell für die spätere Selektion gespeichert.

6.4.5 Das $t\bar{t}$-Netz

Dieser Abschnitt erläutert das Neuronale Netz, welches speziell auf $t\bar{t}$-Reaktionen aus Untergrundprozessen zugeschnitten ist. Auch für dieses Netz wurden die bereits vorgestellten Variablen getestet und ihr Einfluß geprüft. In dieser Analyse wurde bereits hervorgehoben, daß $t\bar{t}$-Produktionen und ihre anschließenden Zerfälle stark durch Jetaktivität gekennzeichnet sind. Diese Tatsache und die damit einhergehenden physikalischen Signaturen gilt es für eine Signalseparation von $t\bar{t}$-Untergrund auszunutzen.

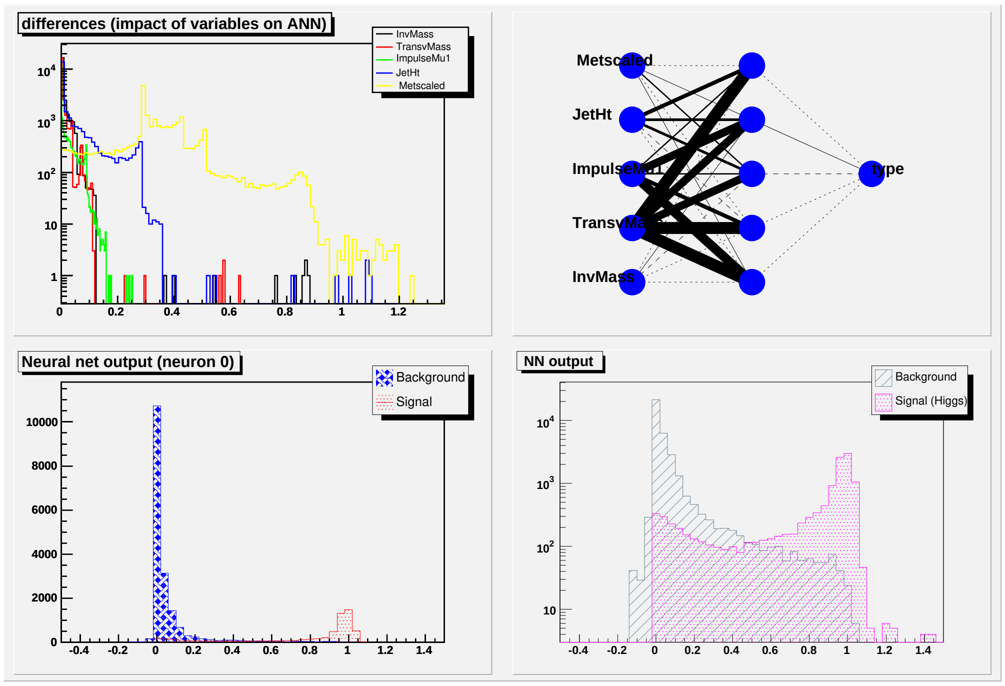

Abbildung 6.22: Struktur, einflußnehmende Variablen und die Ergebnisse nach Trainings- und Testabschluß werden in dieser Abbildung gezeigt. Links oben werden die verwendeten Variablen und ihre Wirkung auf das Netz gezeigt, daneben zeichnet sich der strukturielle Netzaufbau ab. Eingabe- und verborgene Schicht arbeiten im Falle des $t\bar{t}$-Netzes mit fünf Neuronen. Die Ausgabe erfolgt in Analogie zu den weiteren Netzen über ein Ausgabeneuron (in der Ausgabeschicht), welches die Einordnung in einen Bereich zwischen 0 und 1 vornimmt (0 steht für Untergrund, 1 für Signal). Diese Einordnungen sind darunter dargestellt: linkerhand für Trainingsdaten, rechterhand für trainierte wie auch getestete Daten.

Gekennzeichnet von dieser Jetaktivität wurde bereits in Kapitel 6.1 hervorgehoben, daß die Summe der transversalen Energie über alle Jets H_T ein sehr wichtiges Selektionskriterium in Bezug auf $t\bar{t}$-Prozesse darstellt. Aufgrund der Menge beteiligter Jets steigt in Folge die von diesen Jets im Kalorimeter deponierte transversale Energie. Da der Signalprozess nicht so stark durch Jetaktivität geprägt ist, wird die Summe der vorhandenen transversalen Jetenergien dementsprechend gering ausfallen. Betrachtet man Abbildung 6.11, kann dieser Unterschied zwischen $t\bar{t} \longrightarrow \ell\nu\bar{\ell}\bar{\nu}b\bar{b}$ und $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ deutlich wahrgenommen werden. Diese Unterschiede in den Verteilungen können im Training des Neuronalen Netzes stark ausgenutzt werden, so daß der Einfluß von H_T auf das Netz besonders stark ausgeprägt sein sollte.

Dies gilt nicht nur für diese Variable an sich, sondern insbesondere für alle die Größen, in die Jetenergien in die Berechnung eingehen. Nennenswert ist hier die Größe der skalierten fehlenden transversalen Energie, deren Einfluß auf das Netz ebenfalls besonders positiv erwartet wird. Abbildung 6.22 kann die Vermutungen bestätigen (linkes oberes Bild). Daneben sind auch Variablen wie invariante Masse, transversale Masse und die Impulse der Myonen betrachtet worden. Dabei haben sich letzendlich die Größen

- invariante Masse $M_{\mu\mu}$
- transversale Masse M_T
- Impuls des führenden Myons $p_T^{\mu_1}$
- Summe transversaler Energie über alle Jets H_T

und

- skalierte fehlende transersale Energie $\not{E}_T{}^{skal}$

als optimale Parameter für das $t\bar{t}$-Netz erwiesen. Ihr Verhalten auf das Netz wird in Abb. 6.22 gezeigt.
Strukturiell setzt sich dieses Neuronale Netz in der Eingabeschicht aus fünf Neuro-

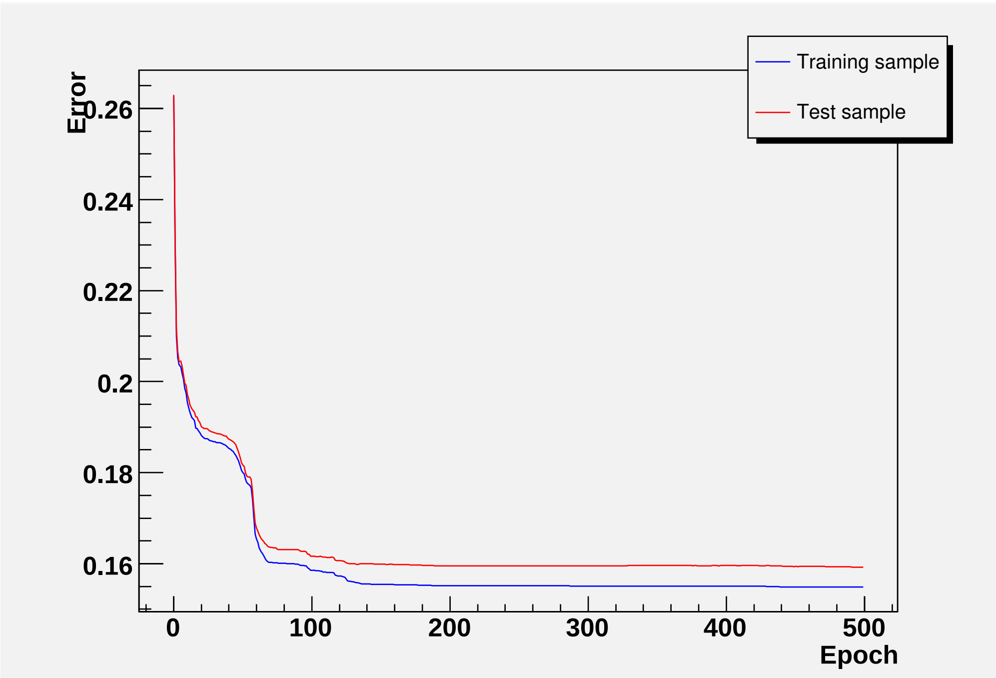

Abbildung 6.23: In dieser Abbildung werden die Fehler der Trainingsdaten (blau) und die Fehler der Testdaten (rot) gegen die Trainingsepochen aufgetragen ($t\bar{t}$-Netz). Die Minimierung der Fehler mit der Zeit ist deutlich erkennbar.

nen, entsprechend der genannten Variablen, zusammen. Diese geben die Information bzw. “ihr Wissen” an die fünf Neuronen der verdeckten Schicht weiter. (In der verdeckten Schicht sind gerade deshalb fünf Neuronen implementiert, da diese Zahl sich

im Vergleich mit mehr oder weniger Neuronen am effizientesten abheben konnte.) Am Ende des Prozesses entscheidet ein Neuron in der Ausgabeschicht, ob das Ereignis tendenziell eher untergrund- (0) oder signalartig (1) war. Die Verteilungen, wie die Ereignisse eingeordnet wurden, sind in den unteren Bildern in 6.22 dargestellt. Links wurden nur die trainierten Ereignisse für eine Darstellung zugrunde gelegt, während rechts zusätzlich auch die getesten Daten in die Verteilungen eingegangen sind.
Der Prozess der Fehlerminimierung während des gesamten Trainingszeitraums kann in Abbildung 6.23 verfolgt werden. In blauen und roten Kurven sind die Fehler auf das Training und auf den Test gegen die Trainingsepochen aufgetragen. Die fortlaufende Minimierung ist deutlich ersichtlich. Inwiefern sich der konstante Unterschied des Fehlers im Netz bemerkbar macht, kann im Abschnitt über das WW-Netz nachgelesen werden. Auch hier ist es wieder wichtig, die Ergebnisse, also das "erlernte Wissen" des Netzes für spätere Selektionszwecke zu sichern.

6.4.6 Das $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz

An dieser Stelle soll auf das Neuronale Netz zu Untergrundprozessen der Art $Z/\gamma^* \longrightarrow \tau^+\tau^-$ eingegangen werden. Betrachten wir den Prozess etwas näher, besonders die Produkte, in die das Z-Boson bzw. das γ^* in diesem Untergrund zerfällt, stellt man fest, daß mit zwei τ-Leptonen eine dem Signalprozess $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ sehr ähnliche Topologie vorhanden ist. Im Sekundärzerfall kann dann das τ weiter in ein μ sowie sein ν_μ und das zum Tau-Lepton gehörige Tau-Neutrino ν_τ übergehen, so daß weitere, sich ähnelnde Merkmale auftreten. Somit müssen sich prozesstypische Eigenheiten in Erinnerung gerufen werden. Es sollte ausgenutzt werden können, daß für Myonen (mindestens zwei) aus Tau-Zerfällen für diesen Untergrundprozess immer noch ein Öffnungswinkel $\Delta\phi_{\mu\mu}$ von annähernd 180° bezüglich ihrer Flugrichtung zueinander zu erwarten ist. In Abbildung 6.9 kann dies noch einmal rückblickend betrachtet werden. Die Myonen bewegen sich demnach voneinander weg, im Vergleich zum Signal, wo die Flugbahn beider Teilchen zu einer gemeinsamen Richtung tendiert. Das bedeutet, daß auf die Variable $\Delta\phi_{\mu\mu}$ besonderer Wert gelegt werden soll. Auch die fehlende transversale Energie darf nicht außer Acht gelassen werden, aufgrund von Neutrinos in Untergrundproduktionen dieser Art. Die Liste der Variablen, die sich für ein Netztraining bezüglich des $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Untergrundes anbieten, umfaßt letztendlich folgenden Größen:

- transversale Masse M_T
- fehlende transversale Energie $\not{E}_T$
- Impuls des ersten Myons $p_T^{\mu_1}$
- Öffnungswinkel zwischen beiden Myonen $\Delta\phi_{\mu\mu}$

und

- Winkel zwischen fehlender transversaler Energie und der ϕ-Koordinaten des zweiten Myons $\angle(\not{E}_T\ ,\phi(\mu_1))$.

Wie groß ihr Einfluß auf das Neuronale Netz ist, kann der linken oberen Darstellung in Abbildung 6.22 entnommen werden. In seinem Aufbau ist auch das Netz zum Prozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$ in drei Schichten gegliedert. In der Eingabeschicht werden die

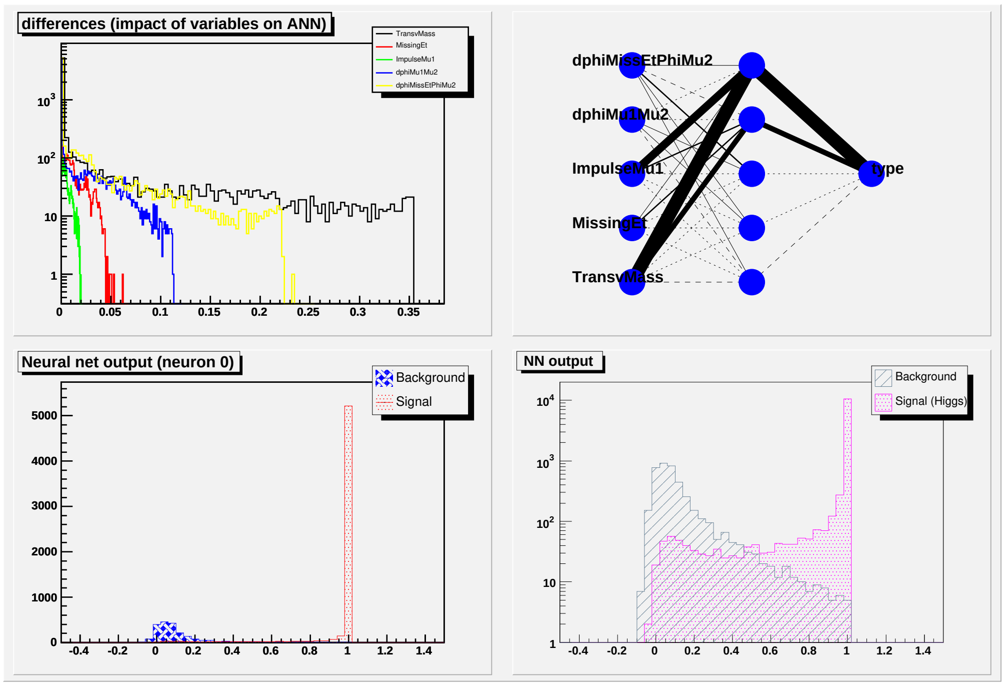

Abbildung 6.24: In dieser Abbildung werden die Netzstruktur (rechts oben), auf sie einflußnehmende Variablen (links oben) sowie die Ergebnisse der Trainings- und Testphase (untere Bilder) präsentiert. Auch das $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz ist wieder aus drei Schichten aufgebaut. Dazu gehören die Eingabeschicht, in diesem Fall mit fünf Eingabeneuronen, die verborgene Schicht mit ebenfalls fünf Neuronen und als Abschluß noch die Ausgabeschicht, welche auch diesmal wieder über ihr Ausgabeneuron die Untergrund- bzw. die Signalzuordung übernimmt. Wie diese Zuordnungen im einzelnen aussehen, zeigen die Netzausgabebilder darunter. Das linke Bild basiert auf den Trainingsdaten, das rechte Bild betrachtet zusätzlich auch die getesteten Ereignisse.

Eingaben über fünf Eingabeneuronen empfangen und an die fünf Neuronen der verborgenen Schicht über gewichtete Verbindungen weitergleitet. Der Prozess der Fehlerminimierung kann beginnen, an dessen Ende in der Ausgabeschicht vom Ausgabeneuron entschieden wird, inwiefern das betrachtete Ereignis eher dem Untergrund oder eher dem Signal zugeordnet werden kann. Die Verteilung, wie die Ereignisse eingeordnet wurden, ist auf den beiden unteren Bildern dargestellt. Es werden die trainierten Ereignisse berücksichtigt (links), aber auch ein Satz trainierter und getesteter Ereignisse zusammen (rechts). Daß die Signalverteilungen in diesen Bildern mehr Einträge zählen als bei anderen bereits vorgestellten Verteilungen, liegt an der dem Netz zur Verfügung stehenden Anzahl an Trainings- und Testdaten, welche zu je 50% auf beide Mengen aufgeteilt werden. Besonders die Prozesse $Z/\gamma^* \longrightarrow \tau^+\tau^-$ und auch später $Z/\gamma^* \longrightarrow \mu^+\mu^-$ sind hinsichtlich ihrer Monte-Carlo-Simulationen in unterschiedlichen Bereichen der invarianten Masse zu erwähnen. Die Netzbetrachtung erfolgt für diese Untergrundprozesse über den gesamten Massenbereich. Dementsprechend sind die Beiträge hinsichtlich ihres Hauptbeitrages zu normieren. Für das $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz können nur dementsprechend wenige Ereignisse für ein Neuronales Netz herangezogen werden. Die Anzahl dieser $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Ereignisse ist deshalb geringer

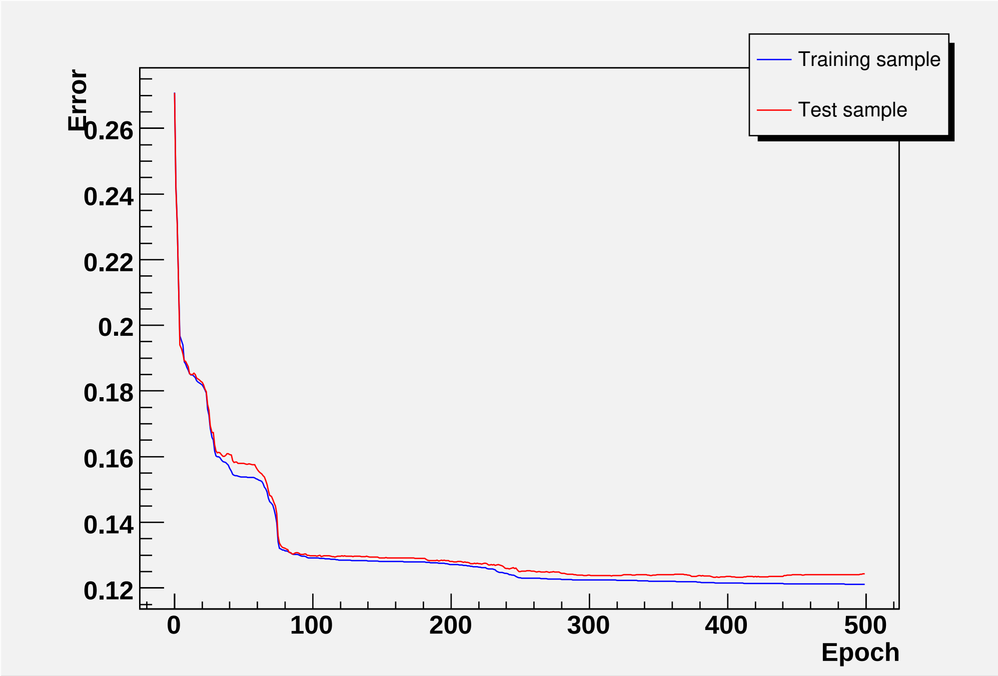

Abbildung 6.25: In dieser Abbildung werden die Fehler der Trainingsdaten (blau) und die Fehler der Testdaten (rot) gegen die Trainingsepochen aufgetragen ($Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz). Die Minimierung der Fehler mit der Zeit ist deutlich erkennbar.

als die der Signalereignisse.
Wie sich zeitlich gesehen der Fehler verändert hat, ist Abbildung 6.25 zu entnehmen. Daß die Fehler mit anhaltendem Training immer kleiner werden, ist deutlich zu erkennen, damit kann das Netzverhalten als realistisch eingeschätzt werden. Für das Training dieses Netzes ist weiterhin zu beachten, daß die zur Verfügung stehenden Trainings- und Testereignisse auf den am stärksten beitragenden Untergrund normiert sind, welcher durch das Monte-Carlo-Simulationsbeispiel von 60 bis 130 *GeV* beschrieben wird. Diese Normierung ist von Bedeutung für einen konsistenten Vergleich entsprechend einer bestimmten Luminosität. Zur weiteren Verarbeitung sollten die gewonnenen Sachverhalte in einem der vorliegenden Analyse gerechten Code exportiert werden, um bei der Signal-zu-Untergrund-Diskriminierung auf die “gelernten” Inhalte dieses Neuronalen Netzes zurückgreifen zu können.

6.4.7 Das $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz

Abschließend soll das Neuronale Netz in Bezug auf Untergrundprozesse der Art $Z/\gamma^* \longrightarrow \mu^+\mu^-$ besprochen werden. In Analogie zum vorigen Kapitel in dem bereits auf ähnliche Signaturen eingegangen wurde, kann auch hier von topologischen Gemeinsamkeiten zwischen Signal und Untergrund gesprochen werden. Das Z/γ^* zerfällt in diesem Monte-Carlo in ein Myon-Antimyon-Paar. Somit ist dieser Prozess genau wie $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ gekennzeichnet durch zwei Myonen entgegengesetzter Ladung. Aufgrund der Verteilungen soll auch die invariante

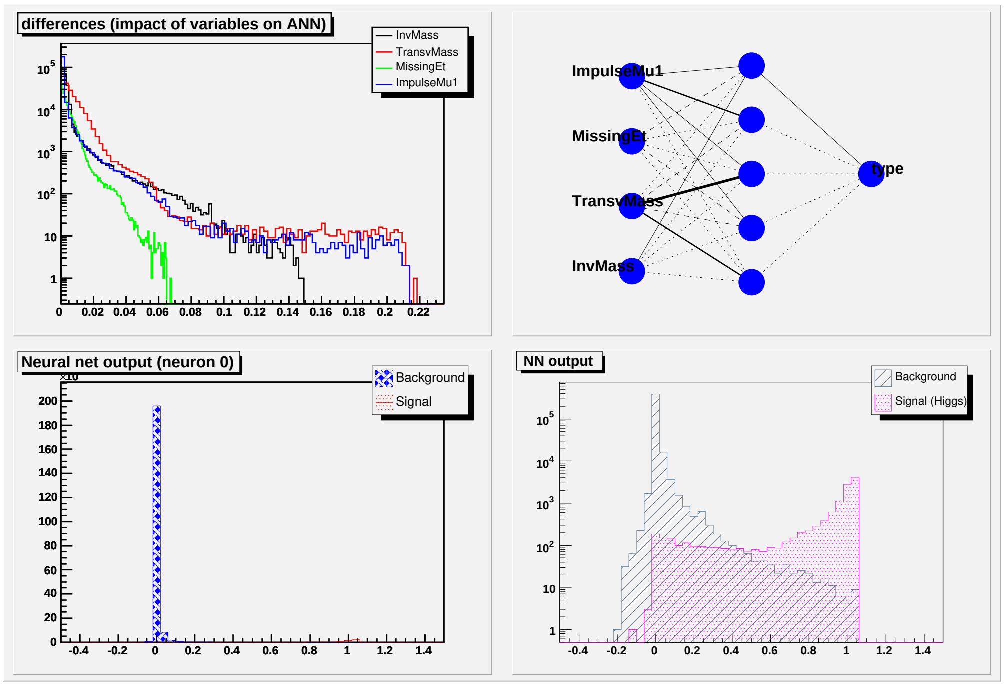

Abbildung 6.26: Dargestellt sind der verwendete Satz an Variablen sowie deren Einfluß auf das Netz im linken oberen Bild. Daneben wurde der Netzaufbau veranschaulicht. Deutlich zu sehen, sind der Aufbau aus drei Schichten. Die Eingabeschicht ist stark durch die verwendeten Variablen geprägt, so daß diese aus vier Eingabeneuronen besteht. In der verborgenen Schicht sind weitere fünf Neuronen angeordnet, bevor in der Ausgabeschicht über das Ausgabeneuron die Kategorisierung bezüglich Untergrund und Signal vorgenommen wird. Wie diese Kategorisierung ausgefallen ist, zeigen die beiden unteren Bilder: links wurden nur trainierte Ereignisse berücksichtigt, rechts zusätzlich auch die getesten Ereignisse.

Masse $M_{\mu\mu}$ für dieses Neuronale Netz berücksichtigt werden. Im Untergrundprozess $Z/\gamma^* \longrightarrow \mu^+\mu^-$ ist deutlich in Abbildung 5.4 die Z-Resonanz um die 90 GeV zu erkennen. Diese liegt verschoben zur Verteilung der invarianten Masse im Signalprozess $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$. Demnach werden durch die Unterschiede beider Verteilungen sehr gute Selektionsmöglichkeiten für das Neuronale Netz dieser Monte-Carlo-Ereignissimulation bereit gestellt. Insgesamt belaufen sich die für ein Netztraining benutzten Variablen auf

- invariante Masse $M_{\mu\mu}$
- transversale Masse M_T
- fehlende transversale Energie $\not{E}_T$

und

- Impuls des ersten Myons $p_T^{\mu_1}$.

Ihr Verhalten und ihren Einfluß auf das Neuronale Netz kann Abbildung 6.26 entnommen werden. Je weiter rechts die Graphen der einzelnen Variablen im linken oberen

Bild liegen, desto stärker ist ihre Wirkung auf das Netz. Daneben befindet sich der strukturielle Aufbau des Netzes. Den Eingangsparametern entsprechend sind in der Eingabeschicht vier Eingabeneuronen, welche ihre Eingabewerte an die fünf Neuronen in der verdeckten Schichten weitergeben. In der Ausgabeschicht wird abschließend über das Ausgabeneuron entschieden, wie ein trainiertes bzw. getestetes Ereignis einzuordnen ist (siehe dazu Abb. 6.26 unten). Signalähnlichkeit entspricht der Zuordnung in Richtung 1, Untergrundähnlichkeit wird mit Werten zu 0 dokumentiert. Wie bereits beim $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz beschrieben, sind die Drell-Yan-Prozesse anders zu behandeln, aufgrund der Unterteilungen der Monte-Carlo-Simulationen in verschiedene Bereiche von invarianten Massen. Da der Untergrund $Z/\gamma^* \longrightarrow \mu^+\mu^-$ den Hauptbeitrag aller Untergrundprozesse liefert, bleiben nach der Vorselektion sehr viele dieser Ereignisse erhalten, welche prinzipiell für die Trainings- und Testphase eines Neuronalen Netzes zur Verfügung stehen würden. Dieser Anteil ist bedeutend höher als im Signalprozess. Auch der $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Untergrundprozess wurde anhand seines invarianten Massenbereiches von 60 *GeV* bis 130 *GeV* normiert.

Die Lernphase von Neuronalen Netzen in dieser Analyse definiert sich über die

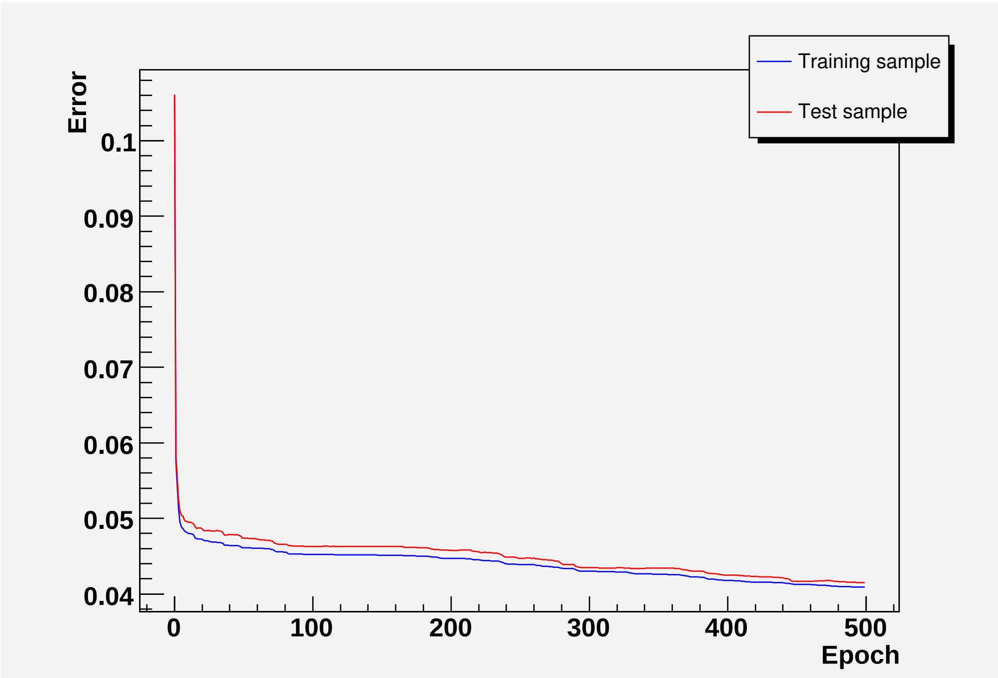

Abbildung 6.27: Die Abbildung zeigt den Verlauf beider Fehlerkurven (blau für trainierte Ereignisse, rot für getestete Ereignisse) die sich auf die Trainingsepochen des $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netzes beziehen.

Minimierung des Fehlers. Wie der Verlauf der Fehler in Abhängigkeit der Trainingsepochen innerhalb des $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netzes aussieht, ist in Abbildung 6.27 beschrieben. Die blaue Kurve steht dabei für Trainingsdaten, anhand der roten Kurve kann der Fehler der Testdaten analysiert werden. Beide zeigen einen abfallenden Verlauf zu kleiner werdenden Fehlerwerten. Um auch für dieses Neuronale Netz einen konsistenten Vergleich der verschiedenen Massenbereiche miteinander und später auch auf Daten zu gewährleisten, müssen die zum gesamten $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Prozess beitragenden Monte-Carlo-Simulationen entsprechend des Hauptbeitrages ($60\ GeV - 130\ GeV$) normiert werden. Dementsprechend stehen jeweils nur eine begrenzte Anzahl an Ereignissen für Training und Test zur Verfügung.

Zur Veranschaulichung werden noch einmal die Eingabevariablen aller Netze dargestellt. Sie sind Tabelle 6.1 zu entnehmen.

	WW-**Netz**	WZ-**Netz**	ZZ-**Netz**	$t\bar{t}$-**Netz**	$Z\tau\tau$-**Netz**	$Z\mu\mu$-**Netz**
$M_{\mu\mu}$	X	X	X	X		X
M_T	X	X	X	X	X	X
$\not{E}_T$	X	X	X		X	X
$p_T^{\mu_1}$	X		X	X	X	X
$p_T^{\mu_2}$	X					
$\Delta\phi_{\mu\mu}$	X				X	
$\not{E}_T{}^{skal}$	X	X		X		
H_T			X	X		
$\angle(\not{E}_T, \phi(\mu_1))$					X	

Tabelle 6.1: Der Tabelle zeigt eine Übersicht der in den Netzen verwendeten Eingabevariablen.

Variablen, welche mit einem X gekennzeichnet sind, werden für das entsprechend angegebene Netz als Eingabevariablen benutzt. Die Eingaben hinsichtlich des WW-Netz liegen deutlich am höchsten, WZ- und $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netze dagegen brauchen nur einen Bruchteil an Variablen.

6.5 Wahl der Schnittpunkte

Die Signal-Untergrund-Selektion der vorliegenden Analyse soll mittels Neuronaler Netze erfolgen. Die verwendeten Netze sind im einzelnen in Kapitel 6.4 aufgelistet und erläutert. Bevor sie jedoch zur Selektion angewendet werden können, müssen die Schnitte auf die verschiedenen Netze bestimmt werden.
Da für die Signal-zu-Untergrund-Diskriminierung alle Neuronalen Netze gleichzeitig angewendet werden, muß beachtet werden, da die Schnitte auf die Netze so abgestimmt sind, daß ein vernünftiges Signal-Untergrund-Verhältnis erhalten bleibt. Würde eine Optimierung der Schnittpunkte nicht erfolgen, könnte es passieren, daß nach der Selektion mit Neuronalen Netzen alle Ereignisse in Signal und Untergrund separiert sind.
Zu Beginn dieses iterativen Prozesses steht jedoch die Analyse der Schnittpunkte der Netze bezüglich des Untergrundes, anhand dessen sie konzipiert worden sind. Dazu wird nach dem bestmöglichen Verhältnis von Signal (S) und Untergrund (B) ge-

sucht. Das bedeutet, daß bei einem Durchlaufverfahren über mögliche Schnittpunkte der Quotient

$$\frac{S}{\sqrt{B}} \tag{6.9}$$

maximal werden soll. Dazu werden Schnitte zwischen 0 und 1 durchlaufen und ihre Verhältnisse aus (6.9) verglichen. Der Punkt an dem dieser Quotient maximal ist, wird als untergrundspezifischer Schnittpunkt festgehalten. Im einzelnen haben sich damit folgende Schnittpunkte bezüglich der diversen Untergrund-Monte-Carlo-Simulationen ergeben:

Neuronales Netz	**Untergrund**	**Schnittpunkt**
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	$Z/\gamma^* \longrightarrow \tau^+\tau^-$	0.98
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	$Z/\gamma^* \longrightarrow \mu^+\mu^-$	0.80
WZ-Netz	$WZ \longrightarrow \mu\nu\ell\bar{\ell}$	0.70
$t\bar{t}$-Netz	$t\bar{t} \longrightarrow \ell\nu\bar{\ell}\bar{\nu}b\bar{b}$	0.98
WW-Netz	$W^+W^- \longrightarrow \mu^+\nu\mu^-\bar{\nu}$	0.62
ZZ-Netz	$ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$	0.88

Tabelle 6.2: Der Tabelle sind mögliche Schnittpunkte für eine nichtoptimierte Selektion zu entnehmen. Das bedeutet, daß die oben angegebenen Schnitte jeweils nur für den angegebenen Untergrund ein maximales $\frac{S}{\sqrt{B}}$-Verhältnis erzeugen.

Rein optisch kann dieser Schnittpunkt anhand der Netzabbildungen im letzten Kapitel kontrolliert werden, da in den unteren Abbildungen die Verteilungen von Signal und Untergrund aufgetragen sind. Legt man virtuell den oben ermittelten Schnittpunkt in diese Grafiken, kann abgeschätzt werden, ob der Schnitt an sich im physikalischen Sinn ein zufriedenstellendes Ergebnis liefern könnte.

Neuronales Netz	**Schnittpunkt (neu)**
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	0.98
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	0.90
WZ-Netz	0.34
$t\bar{t}$-Netz	0.72
WW-Netz	0.54
ZZ-Netz	0.76

Tabelle 6.3: In der Tabelle sind die endgültigen Schnittpunkte nach einer Optimierung festgehalten. Diese sind in chronologischer Reihenfolge aufgelistet, in der die Schnittpunkte optimiert wurden.

Die einfache Schnittanalyse der in Tab. 6.2 dargestellten Art birgt die Gefahr einer

Überselektion. D.h. es wäre denkbar möglich, soviel zu selektieren, daß am Ende keine Ereignisse für eine Analyse zur Verfügung stehen. Um dies zu verhindern, wird die Beeinflussung der Schnitte untereinander studiert und optimiert. Dazu werden sukzessive die Schnitte auf immer mehr Untergründe zusammen untersucht. Den Startpunkt stellt genau der Untergrund dar, welcher die größte Signaleffizienz nach einem untergrundspezifischen Schnitt aufweisen kann, d.h. je mehr Signal nach einem Schnitt erhalten bleibt, desto besser ist der Schnittpunkt als Ausgangspunkt für die Optimierung geeignet. In der vorliegenden Analyse wurden die Schnitte ausgehend vom $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Untergrund verbessert, da dieser die höchste Signaleffizienz von 88% [7] erbringen konnte.
Additiv wurde der nächste Untergrund hinzugenommen und wieder das maximale Verhältnis von $\frac{S}{\sqrt{B}}$ gesucht. Zu diesem Wert wird der Schnittpunkt übernommen und für eine weitere Optimierung angewendet. Iterativ wird dieser Prozess so lange fortgeführt, bis alle Untergründe zusammen betrachtet werden konnten und dementsprechend auch für diese ein maximales Signal-über-Untergrund-Verhältnis gefunden wurde. In Tabelle 6.3 sind die Ergebnisse der Schnittoptimierung festgehalten. Sie sind desweiteren in der Reihenfolge angegeben, in der die einzelnen Netze bearbeitet wurden.
Im Vergleich mit Tab. 6.2 sieht man, daß besonders bei den letzteren Prozessen WZ, $t\bar{t}$, ZZ und WW die Schnittstellen deutlich herabgesenkt werden mußten, um ein vernünftiges Verhältnis zwischen Signal und Untergrund zu erhalten. Demnach werden diese Schnittwerte stark durch die Schnittwerte der zuvor optimierten Netze beeinflußt. Die Schnittpunkte, wie sie in Tabelle 6.3 eingetragen sind, werden für die folgende Hauptselektion eingesetzt.

6.6 Signal und Untergrund vor Anwendung der spezifischen Schnitte

Um einen optischen Eindruck der Verteilungen der Ereignismengen von Signal und Untergrund zu gewinnen, sollen diese im folgenden Abschnitt vorgestellt werden. Eine erste Möglichkeit zur Einschätzung der Separation durch Neuronale Netze kann somit gegeben werden.
Die Abbildung 6.28 zeigt die Verteilungen von Untergrund und Signal vor Anwendung der optimierten Schnitte. Es sind im einzelnen die Verteilungen für alle Netze aufgetragen. Im Untergrund sind für diese Abbildung alle Ereignisse der Untergrundprozesse zusammengefaßt, die auch dem Neuronalen Netz zur Verfügung stehen, diese sowie das Signal selbst wurden entsprechend ihres Wirkungsquerschnittes (mal Verzweigungsverhältnis), ihren generierten Ereignissen und der Luminosität normiert. Das Higgssignal in den Verteilungen des $t\bar{t}$-Netzes, des $Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netzes und des $Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netzes ist um Faktor 1000 hochskaliert um einen besseren Vergleich der Verteilungen von Signal und Untergrund zu ermöglichen.
Man sieht, daß außer dem WW-Netz die normierten Verteilungen sehr deutliche Unterschiede aufweisen. Demnach sollten gerade für diese Untergrundereignissimulatio-

[7] Diese Signaleffizienz berechnet sich anhand der Ereignisse nach der Vorselektion und den Ereignissen nach einem Schnitt auf das Netz.

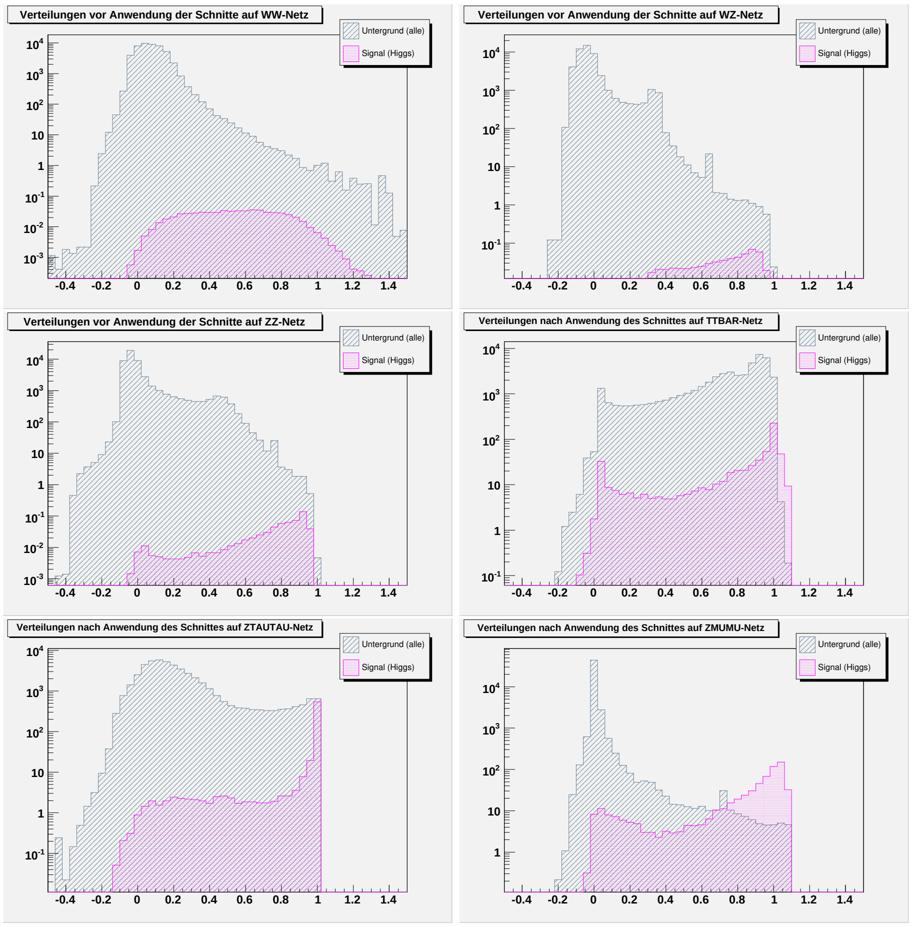

Abbildung 6.28: Die Abbildung zeigt die Verteilungen der Ereignisse der jeweiligen Netze für alle Untergrundprozesse (grau) wie für den Signalprozess (magenta) vor der Anwendung der optimierten Schnitte. Je nachdem wie stark ein Ereignis in die Richtung des Untergrundes oder in die Richtung des Signals tendiert, fällt seine Einteilung auf der Skala aus.

nen die Selektionsergebnisse entsprechend positiv ausfallen sollten. Zwischen 0 und 1 liegen besonders für das WW-Netz die Ereignisse von Signal und Untergrund fast deckungsgleich. Deutlich erkennbar hebt sich die Untergrundverteilung unter der Signalverteilung hervor, da hier zahlenmäßig mehr Ereignisse zur Verfügung standen.

Kapitel 7

Ergebnisse der Selektion

In diesem Kapitel werden die Ergebnisse der Analyse präsentiert. Da der Schwerpunkt auf der Suche einer Higgsmasse von 160 *GeV* liegt, werden diese Ergebnisse zunächst diskutiert. Hierzu wird die Selektion mittels Neuronaler Netze mit einer schnittbasierten Analyse [6] verglichen. Der Vergleich beschränkt sich auf eine Higgsmasse von 160 *GeV*, bei welcher der größte Wirkungsquerschnitt mal Verzweigungsverhältnis erwartet wird. Danach werden die Ergebnisse für 120 *GeV*-, 140 *GeV*- und 180 *GeV*-Higgs-Bosonen präsentiert. Diese beschränken sich auf die ANN-Selektion.

7.1 Analyse der 160 *GeV*-Higgsmasse mit Neuronalen Netzen

Nachdem in Abschnitt 6.5 die optimalen Schnittpunkte der Neuronalen Netze vorgestellt wurden, erfolgt die Selektion. Für eine Higgsmasse von 160 *GeV* werden die Ergebnisse in Tabelle 7.1 präsentiert.
Auffällig ist der Untergrundprozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$, welcher vollständig separiert werden konnte. Auch $Z/\gamma^* \longrightarrow \mu^+\mu^-$ muß in diesem Zusammenhang erwähnt werden, da sich in diesem Prozess der überwiegende Teil des Untergrundes entfernen ließ. Lediglich bei Restanteilen in einem Bereich der invarianten Masse von 60 *GeV* bis 250 *GeV* war es nicht möglich, diese mittels Neuronaler Netze als untergrundartig zu identifizieren. Auch die W- und Z-Bosonzerfälle und Untergrundereignisse aus $t\bar{t}$-Produktionen erwiesen sich als nicht vollständig separabel. Eine besondere Betrachtung richtet sich hierbei auf Untergrundprozesse der Art $p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$. Diese wurden bereits zu Beginn der Analyse hinsichtlich einer Selektion dieser Art als extrem schwierig eingestuft. Dies spiegelt sich im Ergebnis mit der größten Anzahl an verbleibenden Ereignissen wider. Sieht man sich das Resultat für $p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ an, stellt man fest, daß auch für diesen Untergrundprozess die eingangs festgestellten Selektionsmöglichkeiten ein beachtliches Ergebnis liefern, welches mit nur mehreren Hundertsteln verbleibender Restereignisse sehr gut durch die Neuronalen Netze reduziert werden konnte. In der Summe verbleiben demnach mit der hier vorgestellten Analyse mittels Neuronaler Netze im gesamten Untergrund lediglich 3.17 ± 1.04 Ereignisse.

Prozess	**Ereignisse ANN-Selektion**
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	0 ± 1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.2 ± 0.1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.013 ± 0.008
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< $M_{Z/\gamma^*} < \infty$	0.00000 ± 0.00009
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.0 ± 0.2
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.000 ± 0.005
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< $M_{Z/\gamma^*} < \infty$	0.00000 ± 0.00008
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ (inkl.)	2.66 ± 0.07
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ (inkl.)	0.122 ± 0.005
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ (inkl.)	0.106 ± 0.007
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ (inkl.)	0.068 ± 0.006
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (**160 GeV**)	$\mathbf{0.235 \pm 0.004}$
$\sum_{\textbf{Untergrund}}$	$\mathbf{3.17 \pm 1.04}$
Signaleffizienz	17%
Verhältnis von $\frac{S}{\sqrt{B}}$	**0.133**

Tabelle 7.1: Die Tabelle gibt die verbleibenden Ereignisse nach einer Selektion mittels Neuronaler Netze für eine integrierte Luminosität von 775 pb^{-1} an. Das Signal-Untergrund-Verhältnis sowie die Signaleffizienz werden ebenfalls präsentiert. Die angegebenen Fehler beziehen sich lediglich auf die statistischen Unsicherheiten aufgrund der Monte-Carlo-Simulationen.

Der angegebene Fehler bezieht sich hierbei lediglich auf den Fehler, der durch die Statistik der Monte-Carlo-Simulationen folgt. Systematische Unsicherheiten aufgrund der *Jet Energy Scale*, der Impulsauflösung der Myonen, etc. werden sich laut [6] auf etwa $10-30\%$ belaufen, wurden aber hier nicht berücksichtigt.
Um eine endgültige Beziehung für das Signal-Untergrund-Verhältnis ableiten zu können, muß auch auf das Ergebnis des Signals eingegangen werden. Nach der Vorselektion verbleiben für mögliche Higgs-Bosonen mit einer Masse von 160 *GeV* noch 0.613 ± 0.006 Ereignisse. Dies entspricht 45% des Anfangssignals. Werden die Neuronalen Netze in die Analyse implementiert, selektieren diese schließlich 0.235 ± 0.004 Ereignisse. Ausgehend vom Ergebnis der Vorselektion können mit Hilfe der ANN-Selektion 38.5% der Signalereignisse identifiziert werden. Die Signaleffizienz der gesamten Selektion liegt damit bei 17%.

Für das Verhältnis von Signal und Untergrund ergibt sich schließlich ein Wert von $\frac{S}{\sqrt{B}} = 0.133$. Wie gut bzw. wie schlecht dieses Ergebnis ist, kann erst eingeschätzt werden, wenn die Selektion mit einer schnittbasierten Analyse erfolgt ist. Für den Vergleich wird auf Kapitel 7.3 verwiesen.

7.2 Schnittbasierte Analyse der 160 GeV-Higgsmasse

Zur Einschätzung der Sensitivität der Selektion mittels Neuronaler Netze muß eine Vergleichsbasis gefunden werden. Diese Vergleichsbasis stellt in der vorliegenden Arbeit die schnittbasierte Analyse [6] dar. Dazu werden Signal, Untergrund und Daten spezifischen 1-dimensionalen Schnitten unterzogen, um letzendlich eine Möglichkeit zum Vergleich zu schaffen.
Auch für diese Art von Selektion ist eine Vorselektion der hier bereits vorgestellten Art notwendig. Sie hilft den für diesen Zerfallskanal nötigen Qualitätsrahmen zu definieren. D.h. es müssen Ereignisse mit mindestens zwei isolierten Myonen entgegengesetzter Ladung und bestimmter Myon-Qualität selektiert werden. Da die Bedingungen für die schnittbasierte Analyse identisch zu den Bedingungen der Netz-Analyse sind, wird auf die Ergebnisse in Kapitel 5 zurückgegriffen. Hierfür sei auf Tabelle 5.1 verwiesen.

	Selektionskriterium	Schnittwert
Schnitt 1	Vorselektion	siehe Kapitel 5
Schnitt 2	fehlende transv. Energie $\not{E}_T$	$\not{E}_T > 20\ GeV$
Schnitt 3	skalierte $\not{E}_T$	$\not{E}_T{}^{Skal} > 15\ GeV$ (für $N_{Jet} > 0$)
Schnitt 4	invariante Masse $M_{\mu\mu}$	$20\ GeV < M_{\mu\mu} < 80\ GeV$
Schnitt 5	Summe der p_T's und $\not{E}_T$	$100\ GeV < p_T^{\mu_1} + p_T^{\mu_2} + \not{E}_T < 160\ GeV$
Schnitt 6	transversale Masse M_T	$80\ GeV < M_T < 150\ GeV$
Schnitt 7	Summe E_T der Jets	$H_T < 100\ GeV$
Schnitt 8	Öffnungswinkel Myonen $\Delta\phi_{\mu\mu}$	$\Delta\phi_{\mu\mu} < 2.0$

Tabelle 7.2: In der Tabelle sind die Schnittvariablen sowie ihre Schnittwerte für eine Selektion mit 160 GeV-Higgs-Bosonen angegeben. Sie wurden [6] entnommen.

Ähnlich der Variablen eines Netztrainings muß auch hier geschaut werden, welche Parameter sich für Schnitte eignen. Da dies bereits ausführlich in anderen Analysen zum Higgs-Boson untersucht wurde, soll im Rahmen dieser Arbeit auf die Schnittpunkte einer dieser Analysen (nämlich [6]) zurückgegriffen werden. Die Dissertation [6] untersucht ähnlich der vorliegenden Arbeit mehrere Higgs-Bosonmassen, legt aber besonderes Interesse auf das 160 GeV-Higgs. Da der Vergleich beider Methoden ausschließlich für diesen Massenpunkt erfolgen soll, werden die dementsprechenden Schnitte übernommen. Diese wurden nicht neu auf die integrierte Luminosität von

775 pb^{-1} dieser Analyse optimiert. Die Kriterien und ihre veranschlagten Werte sind Tabelle 7.2 zu entnehmen.
Schnitt 1 stellt die bereits vorgestellte Vorselektion dar. Ihre Ergebnisse sind im angegebenen Kapitel erläutert. Danach werden die übrigen sieben Schnitte in einer eigenen Analyse angewendet. Die Ergebnisse sind in Tabelle 7.3 eingetragen.

Prozess	**Ereignisse Schnitt-Selektion**
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	0 ± 1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.3 ± 0.2
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.3 ± 0.2
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.000 ± 0.005
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< M_{Z/γ^*} < ∞	0.00000 ± 0.00009
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.2 ± 0.2
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.000 ± 0.005
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< M_{Z/γ^*} < ∞	0.00000 ± 0.00008
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ (inkl.)	6.1 ± 0.1
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ (inkl.)	0.346 ± 0.008
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ (inkl.)	0.29 ± 0.01
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ (inkl.)	0.70 ± 0.02
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (**160 GeV**)	**0.303 ± 0.004**
$\sum_{\textbf{Untergrund}}$	**8.24 ± 1.05**
Signaleffizienz	22%
Verhältnis von $\frac{S}{\sqrt{B}}$	**0.106**

Tabelle 7.3: Die Tabelle gibt die verbleibenden Ereignisse nach einer Selektion mittels Schnitten für eine Higgsmasse von 160 *GeV* an. Das Signal-Untergrund-Verhältnis sowie die Signaleffizienz werden ebenfalls präsentiert. Die angegebenen Fehler beziehen sich lediglich auf statistische Unsicherheiten.

Hier sind bereits erste Unterschiede zur ANN-Selektion (vgl. Tab. 7.1) ersichtlich. Beispielsweise konnte die schnittbasierte Selektion den Prozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$ nicht vollständig separieren. Für den Massenbereich von 60 *GeV* bis 130 *GeV* verbleiben demnach 0.2 ± 0.2 Ereignisse. Im Untergrundprozess $Z/\gamma^* \longrightarrow \mu^+\mu^-$ ist lediglich eine Verschiebung zu niedrigeren Massenbereichen erkennbar, so daß bei 15 *GeV* bis 60 *GeV* 0.3 ± 0.2 Ereignisse nicht durch Schnitte vom Signal separiert werden konnten. Dagegen liegt die Sensitivität im Bereich 130 *GeV* bis 250 *GeV* deutlich

besser, da sich dieser vollständig mit den angegebenen Schnitten entfernen ließ.
Weiterhin fällt auf, daß in dieser Art von Analyse die verbleibenden Ereignisse der restlichen Untergrundprozesse höher liegen. Insgesamt belaufen sich damit die verbleibenden Ereignisse auf 8.24 ± 1.05. Für das Signal-Untergrund-Verhältnis $\frac{S}{\sqrt{B}}$ bedeutet das, auch wenn vom Signal 29% mehr Ereignisse selektiert werden konnten, was einer Signaleffizienz von 22% entspricht, daß es mit $\frac{S}{\sqrt{B}} = 0.106$ niedriger als in der ANN-Selektion liegt.

7.3 Vergleich von Neuronalen Netzen und Schnitten

Da beide Analysen vorangehend präsentiert wurden, kann ein direkter Vergleich erfolgen. Dazu werden die Ergebnisse im einzelnen wie auch im gesamten näher betrachtet. Die schnittbasierte Analyse hinsichtlich der Luminosität von 775 pb^{-1} des hier verwendeten Datensatzes wurde nicht neu optimiert. Die Schnitte sind demnach nur in [6] als optimal zu betrachten.
Vergleicht man die Ereigniszahlen der unterschiedlichen Prozesse, wird man schnell feststellen, daß diese bei der schnittbasierten Analyse mehrheitlich höher liegen als bei der Analyse mit Neuronalen Netzen. Dies gilt im besonderen nicht nur für den Untergrund, sondern auch für das Signal selbst. Deshalb ist es wichtig, das Verhältnis $\frac{S}{\sqrt{B}}$ für beide Selektionsmöglichkeiten zu berechnen.
Die beste Effizienz der Analyse ist erreicht, wenn dieses Verhältnis maximal wird. Da hier zwei Arten der Analyse verglichen werden, ist diejenige hervorzuheben, welche das größere Verhältnis von Signal-zu-Wurzel-aus-Untergrund erbringen kann. Wie bereits angegeben liegt der Wert der Selektion mittels Neuronaler Netze bei 0.133. Dagegen erhält man für die schnittbasierte Analyse 0.106 als Ergebnis. Das bedeutet für einen Vergleich untereinander, daß die Anwendung Neuronaler Netze ein um 25% besseres Ergebnis bei einer Optimierung von $\frac{S}{\sqrt{B}}$ liefert. Damit kann diese Methode als die effizientere Methode eingestuft werden.
Somit wird festgehalten, daß mittels künstlicher Neuronaler Netze die Bedingungen des Untergrundes und die Bedingungen des Signales besser zueinander approximiert werden können als es mit Schnitten möglich wäre. Mit Neuronalen Netzen ist die Möglichkeit gegeben, durch ihren Lern- und Trainingsprozess die Signaturen unterschiedlicher Prozesse so stark zu abstrahieren, daß anschließend die zur Verfügung stehenden Separationsparameter genutzt werden können. Ziel des Netzes ist die Einordnung der Ereignisse auf einer Skala zwischen 0 und 1. Dabei steht die 0 für untergrundartige und die 1 für signalartige Ereignisse. Je besser diese ausfällt, desto deutlicher kann mit einem Schnitt auf den Netz-Ausgabewert die Selektion erfolgen. Im Vergleich zur Analyse mit Schnitten konnte gezeigt werden, wie sensitiv die Arbeit mit Neuronalen Netzen ist, was durch das 25% höhere Signal-Untergrund-Verhältnis bestätigt wird.

7.4 Analyse der 120 *GeV*-Higgsmasse mit Neuronalen Netzen

Nachdem der Schwerpunkt der Arbeit mit der Suche nach 160 *GeV*-Higgs-Bosonen präsentiert wurde, folgen in den nächsten Abschnitten die Ergebnisse der Analyse anderer Higgsmassen. Beginnend soll hier die Auswertung der 120 *GeV*-Massen vorgestellt werden. Bevor die Ergebnisse nach der Selektion angegeben werden können, müssen zunächst die Neuronalen Netze an sich sowie die spezifischen Schnittpunkte auf das Netz betrachtet werden.

Neuronales Netz	Untergrund	Schnittpunkt
WZ-Netz	$WZ \longrightarrow \mu\nu\ell\bar{\ell}$	0.40
WW-Netz	$W^+W^- \longrightarrow \mu^+\nu\mu^-\bar{\nu}$	0.18
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	$Z/\gamma^* \longrightarrow \tau^+\tau^-$	0.98
ZZ-Netz	$ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$	0.60
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	$Z/\gamma^* \longrightarrow \mu^+\mu^-$	0.88
$t\bar{t}$-Netz	$t\bar{t} \longrightarrow \ell\nu\bar{\ell}\bar{\nu}b\bar{b}$	0.98

Tabelle 7.4: Der Tabelle sind mögliche Schnittpunkte für eine nichtoptimierte Selektion für 120 *GeV*-Higgs-Bosonen zu entnehmen. Das bedeutet, daß die oben angegebenen Schnitte jeweils nur für den angegebenen Untergrund ein maximales $\frac{S}{\sqrt{B}}$-Verhältnis erzeugen (S = Higgssignal, B = Untergrund). Sie müssen mit einem weiterführenden Algorithmus optimiert werden.

Neuronales Netz	Schnittpunkt (neu)
WZ-Netz	0.40
WW-Netz	0.00
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	0.96
ZZ-Netz	0.24
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	0.66
$t\bar{t}$-Netz	0.50

Tabelle 7.5: In der Tabelle sind die endgültigen Schnittpunkte nach einer Optimierung für eine 120 *GeV*-Higgsmasse festgehalten. Diese sind in der Reihenfolge aufgelistet, in der die Schnittpunkte optimiert wurden.

An den Neuronalen Netzen hinsichtlich der Analyse der 120 *GeV*-Higgsmasse hat sich im Vergleich zur 160 *GeV*-Higgsmasse nichts wesentliches geändert. Es wird mit demselben Satz an Variablen gearbeitet. Damit erfolgt eine neue Trainings- und Testphase des Netzes. Lediglich ein diesem Signal entsprechender Higgssignal-

Simulationsdatensatz wurde ersetzt. Alles, den Aufbau und die Schnittvariablen betreffend, kann in Kapitel 6.4 nachgelesen werden. Die trainierten Netze an sich unterscheiden sich natürlich von denen der bereits vorgestellten Analyse. Demzufolge müssen auch die Schnittpunkte auf die Netze neu optimiert werden. Dafür wird dasselbe Verfahren wie bereits zuvor angewendet, indem zuerst die optimalen Schnittpunkte untergrundspezifisch bestimmt wurden. Das Ergebnis ist Tabelle 7.4 zu entnehmen.

Prozess	**Ereignisse ANN-Selektion**
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	0 ± 1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.2 ± 0.1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.013 ± 0.008
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< M_{Z/γ^*} < ∞	0.00000 ± 0.00009
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.0 ± 0.2
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.000 ± 0.005
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< M_{Z/γ^*} < ∞	0.00000 ± 0.00008
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ (inkl.)	2.67 ± 0.07
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ (inkl.)	0.122 ± 0.005
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ (inkl.)	0.106 ± 0.007
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ (inkl.)	3.11 ± 0.04
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ **(120 GeV)**	$\mathbf{0.0090 \pm 0.0004}$
$\Sigma_{\textbf{Untergrund}}$	$\mathbf{6.22 \pm 1.04}$
Signaleffizienz	2%
Verhältnis von $\frac{S}{\sqrt{B}}$	**0.004**

Tabelle 7.6: Die Tabelle gibt die verbleibenden Ereignisse nach einer Selektion für die 120 *GeV* Masse eines Higgs-Bosons mittels Neuronaler Netze an. Das Signal-Untergrund-Verhältnis sowie die Signaleffizienz werden ebenfalls präsentiert. Die angegebenen Fehler beziehen sich lediglich auf statistische Unsicherheiten aufgrund der Monte-Carlo-Ereignissimulationen.

Danach wurden iterativ die Schnittpunkte neu bestimmt unter Hinzunahme immer weiterer Untergrundprozesse. Startpunkt war in diesem Fall der Prozess $WZ \longrightarrow \mu\nu\ell\bar{\ell}$, da hier die meisten Signalereignisse selektiert werden konnten (85%). Wie sich die Punkte diesbezüglich geändert haben, kann in Tabelle 7.5 nachgelesen

werden. Die Reihenfolge in der die Netze aufgelistet sind, entspricht der Reihenfolge in welcher die Schnittpunkte nacheinander optimiert wurden. Auf den ersten Blick fällt besonders der Schnittpunkt auf das WW-Netz ins Auge, welcher bei 0.0 nach der Optimierung liegt. Da der Schnittpunkt vor der Optimierung auf den gesamten Untergrund bei 0.18 lag und bereits vor dem WW-Netz der Schnitt auf das WZ-Netz erfolgt, ist dieser Wert mit dem Schnittpunkt des WZ-Netzes korreliert. D.h. durch das WZ-Netz wurde bereits bei sehr niedrigen Netzwerten soviel Untergrund reduziert, daß dieser bei der Optimierung des WW-Schnittwertes nicht mehr beitragen konnte und ein dementsprechend niedriger Schnittwert zum maximalen $\frac{S}{\sqrt{B}}$-Verhältnis bestimmt werden konnte.
Im Vergleich zum 160 GeV-Higgs-Boson werden deutliche Unterschiede offensichtlich. Das liegt daran, daß kinematisch gesehen die Verteilungen der in die Netze eingehenden Größen verschoben sind. So müssen beispielsweise Variablen wie invariante Masse oder die Impulse der Myonen zu kleineren GeV-Einheiten verschoben sein. Die Auswirkungen werden in den Netzen deutlich, die daraufhin sehr viel schwieriger abstrahieren können, inwiefern ein Ereignis eher signal- oder eher untergrundartig war. Daraufhin werden die Verteilungen von Signal und Untergrund nach dem Netztraining eher ineinander übergehen, so daß sich die Optimierung der Schnittpunkte zu kleineren Werten verschiebt. Besonders deutlich wird dies anhand des $t\bar{t}$- und anhand des WW-Netzes, so daß eine starke Abhängigkeit der Netze von der Masse des Higgs-Bosons angenommen wird.
Diese Schnittpunkte werden letztendlich für die Selektion herangezogen. Ihre Ergebnisse sind in Tabelle 7.6 aufgeführt. Auffällig und besonders hervorstechend ist dabei der Untergrundprozess mit $t\bar{t}$-Produktionen, von dem um Faktor 50 mehr Ereignisse selektiert wurden. Betrachtet man weiter das Verhältnis $\frac{S}{\sqrt{B}}$ mit 0.004, ist festzustellen, daß für eine eventuelle Higgsmasse von 120 GeV diese Art der Analyse wenig Aussicht auf Erfolg läßt. Dies wird durch eine Signaleffizienz von nur 2% bestätigt.

7.5 Analyse der 140 GeV-Higgsmasse mit Neuronalen Netzen

Eine weitere Higgsmasse, die es zu untersuchen gilt, ist die 140 GeV-Masse. Demnach müssen die Trainings- und Testphase der Neuronalen Netze mit einer dieser Masse entsprechender Signal-Monte-Carlo-Simulation neu durchgeführt werden.
Die zusätzlichen Parameter wie strukturieller Aufbau und Trainingsvariablen können, wie beim 120 GeV-Higgs, den bereits vorgestellten Netzen in Kapitel 6.4 entnommen werden, da diese Größen identisch ünernommen worden sind. Die Beschränkung liegt lediglich darin, daß mit dem neuen Signal die Netze neu angepaßt werden müssen. Das wiederum wirkt sich auf die möglichen Schnittwerte der diversen Netze aus, so daß diese wieder neu bestimmt und optimiert werden müssen. Die Werte für die Schnitte auf Netze bezüglich des zugrundegelegten Untergrundprozesses werden in Tabelle 7.7 angegeben.

Neuronales Netz	**Schnittpunkt (neu)**
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	0.98
WW-Netz	0.26
WZ-Netz	0.40
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	0.76
ZZ-Netz	0.54
$t\bar{t}$-Netz	0.26

Tabelle 7.8: In der Tabelle sind die endgültigen Schnittpunkte nach einer Optimierung für eine 140 GeV-Higgsmasse festgehalten. Diese sind in der Reihenfolge aufgelistet, in der die Schnittpunkte optimiert wurden.

Neuronales Netz	**Untergrund**	**Schnittpunkt**
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	$Z/\gamma^* \longrightarrow \tau^+\tau^-$	0.98
WW-Netz	$W^+W^- \longrightarrow \mu^+\nu\mu^-\bar{\nu}$	0.24
WZ-Netz	$WZ \longrightarrow \mu\nu\ell\bar{\ell}$	0.46
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	$Z/\gamma^* \longrightarrow \mu^+\mu^-$	0.76
ZZ-Netz	$ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$	0.68
$t\bar{t}$-Netz	$t\bar{t} \longrightarrow \ell\nu\bar{\ell}\bar{\nu}b\bar{b}$	0.98

Tabelle 7.7: Der Tabelle sind mögliche Schnittpunkte für eine nichtoptimierte Selektion für 140 GeV-Higgs-Bosonen zu entnehmen. Das bedeutet, daß die oben angegebenen Schnitte jeweils nur für den angegebenen Untergrund ein maximales $\frac{S}{\sqrt{B}}$-Verhältnis erzeugen. Sie müssen mit einem weiterführenden Algorithmus optimiert werden.

Diese Werte stellen jeweils nur optimale Schnittpunkte den entsprechend angegebenen Untergrundprozessen dar. Zur Optimierung müssen wieder iterativ die restlichen Untergrundproduktionen in den Algorithmus aufgenommen werden, um nach und nach durch maximale Signal-Untergrund-Verhältnisse die neuen Schnittpunkte zu erhalten. Ähnlich der 160 GeV-Analyse stellt auch hier den Startpunkt zur Optimierung der Prozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$ dar. Für diesen war die Selektion an Higgssignal-Ereignissen mit 80% deutlich am höchsten.
Wie sich die Schnittpunkte durch den Optimierungsprozess verändert haben, kann Tabelle 7.8 entnommen werden. Wichtig ist auch hier wieder, daß die Reihenfolge in der die Darstellung der Netze erfolgt, der Reihenfolge entspricht, in der die Neuronalen Netze optimiert wurden. Diese Reihenfolge wird hierbei von der Signalselektion der einzelnen Netze entsprechend ihres Untergrundes bestimmt.
Diese Schnitte liegen im Vergleich zum 120 GeV-Higgs-Boson höher, aber mehrheitlich immer noch niedriger als im Vergleich zu 160 GeV-Higgs-Bosonen. Werden diese beiden Beispiele als Richtlinien für einen Vergleich herangezogen, so erwartet man ein besseres Ergebnis als in Kapitel 7.4, aber ein immer noch schlechteres Ergebnis

als es in Kapitel 7.1 präsentiert werden konnte. Wie die Selektion der 140 *GeV*-Masse ausfällt, ist in Tabelle 7.9 angegeben.
Deutlich erkennbar ist die schlechtere Separation im Bereich des Prozesses $Z/\gamma^* \longrightarrow \mu^+\mu^-$, in welchem diesmal noch insgesamt 2.3 ± 0.5 Ereignisse nach der Selektion mittels Neuronaler Netze verbleiben. Auch der Untergrundprozess der $Z/\gamma^* \longrightarrow \tau^+\tau^-$-Produktionen kann nicht mehr vollständig separiert werden, so daß im Massenbereich von 60 *GeV* bis 130 *GeV* noch 0.2 ± 0.2 Ereignisse durch die Netze selektiert wurden. Besonders auffällig sind auch die Ergebnisse der W- und Z-Bosonzerfälle. Gerade der Untergrundprozess $W^+W^- \longrightarrow \mu^+\nu\mu^-\bar{\nu}$, weist, wie eingangs erwähnt, die höchste Schwierigkeitsstufe in Analysen, die diesen Zerfallskanal bearbeiten, auf. Hier verbleiben nach Anwendung der Neuronalen Netze 9.0 ± 0.1 Ereignisse. Auch die Zerfälle WZ und ZZ zeigen deutlich höhere Ereigniszahlen. Insgesamt verbleiben vom gesamten Untergrund 12.9 ± 1.1 Ereignisse.

Prozess	**Ereignisse ANN-Selektion**
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	0 ± 1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	1.8 ± 0.4
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.5 ± 0.2
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.013 ± 0.008
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< $M_{Z/\gamma^*} < \infty$	0.00000 ± 0.00009
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.2 ± 0.2
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.000 ± 0.005
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< $M_{Z/\gamma^*} < \infty$	0.00000 ± 0.00080
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ (inkl.)	9.0 ± 0.1
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ (inkl.)	0.403 ± 0.009
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ (inkl.)	0.38 ± 0.01
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ (inkl.)	0.62 ± 0.02
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ (**140 GeV**)	**0.265 ± 0.003**
$\sum_{\textbf{Untergrund}}$	**12.9 ± 1.1**
Signaleffizienz	24%
Verhältnis von $\frac{S}{\sqrt{B}}$	**0.074**

Tabelle 7.9: Die Tabelle gibt die verbleibenden Ereignisse nach einer Selektion für die 140 *GeV* Masse eines Higgs-Bosons mittels Neuronaler Netze an. Das Signal-Untergrund-Verhältnis sowie die Signaleffizienz werden ebenfalls präsentiert. Die angegebenen Fehler beziehen sich lediglich auf statistische Unsicherheiten der Monte-Carlo-Ereignissimulation.

Die Anzahl der selektierten Ereignisse des Signalprozesses liegt für 140 GeV-Higgs-Bosonen um 0.030 $\pm$ 0.003 Ereignisse höher als bei 160 GeV. Da dies einer Signaleffizienz von 24% entspricht, kann diese Größe besser durch die Neuronalen Netze separiert werden. Dennoch ergibt sich für das Signal-Untergrundverhältnis $\frac{S}{\sqrt{B}}$ lediglich ein Wert von 0.074. Damit werden die Erwartungen bestätigt, daß für 140 GeV-Higgsmasse die Analyse besser ausfällt als für 120 GeV. Sie kann aber nicht die Ergebnisse der Selektion der 160 GeV-Higgs-Bosonen erzielen.

7.6 Analyse der 180 GeV-Higgsmasse mit Neuronalen Netzen

Abschließend soll auch die letzte untersuchte Higgsmasse von 180 GeV mit ihren Ergebnissen präsentiert werden. Wie schon bei allen Beschreibungen der anderen Massen voraus, werden beginnend bei den Neuronalen Netzen die Veränderungen beschrieben.

Neuronales Netz	**Untergrund**	**Schnittpunkt**
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	$Z/\gamma^* \longrightarrow \tau^+\tau^-$	0.98
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	$Z/\gamma^* \longrightarrow \mu^+\mu^-$	0.96
WW-Netz	$W^+W^- \longrightarrow \mu^+\nu\mu^-\bar{\nu}$	0.46
ZZ-Netz	$ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$	0.80
$t\bar{t}$-Netz	$t\bar{t} \longrightarrow \ell\nu\bar{\ell}\bar{\nu}b\bar{b}$	0.98
WZ-Netz	$WZ \longrightarrow \mu\nu\ell\bar{\ell}$	0.76

Tabelle 7.10: Der Tabelle sind mögliche Schnittpunkte für eine nichtoptimierte Selektion für 180 GeV-Higgs-Bosonen zu entnehmen. Das bedeutet, daß die oben angegebenen Schnitte jeweils nur für den angegebenen Untergrund ein maximales $\frac{S}{\sqrt{B}}$-Verhältnis erzeugen. Sie müssen mit einem weiterführenden Algorithmus optimiert werden.

Um das hier vorgestellte Higgssignal untersuchen zu können, müssen die Neuronalen Netze entsprechend des zu diesem Massenpunkt gehörigen Simulationsdatensatz modifiziert werden. Dies beeinhaltet eine neue Trainings- und Testphase. Die dafür verwendeten Variablen sowie der weitere Aufbau des Netzes, also die Anzahl der Neuronen und Schichten werden hinsichtlich der 160 GeV-Analyse beibehalten. Nach Abschluß des Lernprozesses gilt es, die Schnittpunkte auf die Netze über ein maximales $\frac{S}{\sqrt{B}}$-Verhältnis zu bestimmen. Dabei werden anfänglich alle Netze bezüglich ihres jeweiligen Untergrundprozesses auf den bestmöglichen Schnitt untersucht. Die Werte dieser Algorithmusphase sind in Tabelle 7.10 angegeben.

Neuronales Netz	**Schnittpunkt (neu)**
$Z/\gamma^* \longrightarrow \tau^+\tau^-$ -Netz	0.98
$Z/\gamma^* \longrightarrow \mu^+\mu^-$ -Netz	0.90
WW-Netz	0.50
ZZ-Netz	0.16
$t\bar{t}$-Netz	0.18
WZ-Netz	0.06

Tabelle 7.11: In der Tabelle sind die endgültigen Schnittpunkte nach einer Optimierung für eine 180 GeV-Higgsmasse festgehalten. Diese sind in der Reihenfolge aufgelistet, in der die Schnittpunkte optimiert wurden.

Besonders die Schnittpunkte der W- und Z-Netze fallen niedriger aus, als die in der 160 GeV-Analyse. Da diese jedoch noch nicht auf das beste Signal-Untergrund-Verhältnis bezüglich aller Untergrundprozesse optimiert worden sind, muß dies in einem weiteren Algorithmus noch nachgeholt werden. Dazu wird der Untergrundprozess herangezogen, dessen Ereignisse bezüglich des Signals am höchsten waren.
Im Falle der 180 GeV-Higgs-Bosonmasse wird der Startpunkt wieder durch den Prozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$ dargestellt. Das Netz zu diesem Untergrund kann immerhin noch 84% des Signals nach der Vorselektion separieren. Iterativ werden dann nach ihren Signaleffizienzen entsprechend die weiteren Netze zur Schnittoptimierung herangezogen. Die Ergebnisse zur Optimierung der Schnittpunkte sind Tabelle 7.11 zu entnehmen.
Da besonders die Werte für ZZ-, $t\bar{t}$- und WZ-Netz in dieser Tabelle sehr niedrig im Vergleich zur 160 GeV-Analyse ausfallen, liegen die Erwartungen an die Signalsignifikanz $\frac{S}{\sqrt{B}}$ entsprechend geringer. Inwiefern sie besser oder schlechter im Vergleich zur 120 GeV- bzw. 140 GeV-Analyse sind, kann allerdings erst mit dem Signal-Untergrund-Verhältnis $\frac{S}{\sqrt{B}}$ festgestellt werden. Die Ergebnisse der Selektion der 180 GeV-Higgs-Bosonen werden in Tabelle 7.12 präsentiert.
Die Grundtendenz im Vergleich zu beiden vorangestellten Higgsmassen liegt in dieser Analyse deutlich positiver. Für den Untergrundprozess $Z/\gamma^* \longrightarrow \mu^+\mu^-$ konnten demnach lediglich 0.5 ± 0.2 Ereignisse selektiert werden. Der Prozess $Z/\gamma^* \longrightarrow \tau^+\tau^-$ erwies sich für diesen Massenpunkt als vollständig reduzibel. Und auch die restlichen Untergrundprozesse zeigen niedrigere Ereigniszahlen. Vergleicht man diese Tabelle mit der Analyse der 160 GeV-Higgs-Bosonen stellt man schnell fest, daß sie diesbezüglich schlechtere Tendenzen aufweist. Summiert man die Untergrundprozesse auf, so erhält man ein Ergebnis von verbleibenden 5.1 ± 0.2 Ereignissen. Vom Signal werden noch 0.177 ± 0.002 Ereignisse selektiert. Das entspricht einer Signaleffizienz von 19%. Ein Vergleich der Signal-Untergrund-Verhältnisse bestätigt mit 0.078, daß die Analyse mit 180 GeV-Higgs-Bosonen zwar hinsichtlich der Massen von 120 GeV und 140 GeV qualitativ besser scheint, jedoch in Bezug auf die Analyse eines 160 GeV-Higgs-Bosons ein um 41% schlechteres Ergebnis liefert.

Prozess	**Ereignisse ANN-Selektion**
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 5 GeV< M_{Z/γ^*} <15 GeV	0 ± 1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.1 ± 0.1
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.3 ± 0.2
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.013 ± 0.008
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \mu^+\mu^-$, 500 GeV< M_{Z/γ^*} < ∞	0.00000 ± 0.00009
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 15 GeV< M_{Z/γ^*} <60 GeV	0.0 ± 0.1
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 60 GeV< M_{Z/γ^*} <130 GeV	0.0 ± 0.2
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 130 GeV< M_{Z/γ^*} <250 GeV	0.000 ± 0.005
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 250 GeV< M_{Z/γ^*} <500 GeV	0.000 ± 0.001
$Z/\gamma^* \longrightarrow \tau^+\tau^-$, 500 GeV< M_{Z/γ^*} < ∞	0.00000 ± 0.00008
$p\bar{p} \longrightarrow W^+W^- \longrightarrow \ell^+\nu\ell^-\bar{\nu}$ (inkl.)	3.71 ± 0.08
$p\bar{p} \longrightarrow WZ \longrightarrow \ell\nu\ell\bar{\ell}$ (inkl.)	0.200 ± 0.006
$p\bar{p} \longrightarrow ZZ \longrightarrow \ell\bar{\ell}\ell\bar{\ell}$ (inkl.)	0.176 ± 0.009
$p\bar{p} \longrightarrow t\bar{t} \longrightarrow \ell^+\nu\ell^-\bar{\nu}b\bar{b}$ (inkl.)	0.54 ± 0.02
$H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ **(180 GeV)**	$\mathbf{0.177 \pm 0.002}$
$\sum_{\textbf{Untergrund}}$	$\mathbf{5.1 \pm 1.1}$
Signaleffizienz	19%
Verhältnis von $\frac{S}{\sqrt{B}}$	**0.078**

Tabelle 7.12: Die Tabelle gibt die verbleibenden Ereignisse nach einer Selektion für die 180 GeV Masse eines Higgs-Bosons mittels Neuronaler Netze an. Das Signal-Untergrund-Verhältnis sowie die Signaleffizienz werden ebenfalls präsentiert. Die angegebenen Fehler beziehen sich lediglich auf statistische Unsicherheiten der Monte-Carlo-Ereignissimulation.

Abschließend sollen die wichtigsten Ergebnisse dieses Kapitels noch einmal für alle Higgsmassen und Analysemethoden tabellarisch dargestellt werden. In der Tabelle 7.13 können die verbleibenden Ereignisse des Signals sowie des gesamten Untergrundes verglichen werden. Daneben sind auch die einzelnen Signaleffizienzen und das Verhältnis aus Signal und der Wurzel des Untergrundes $\frac{S}{\sqrt{B}}$ angegeben. Die Auflistung erfolgt für alle in dieser Analyse behandelten Higgsmassen und Selektionsmethoden.

Higgssignal	**Signal**	$\sum_{Untergrund}$	$\frac{S}{\sqrt{B}}$	**Signaleffizienz**
120 *GeV* **(ANN)**	0.0090 ± 0.0004	6.22 ± 1.04	0.004	2%
140 *GeV* **(ANN)**	0.265 ± 0.003	12.9 ± 1.1	0.074	24%
160 *GeV* **(ANN)**	0.235 ± 0.004	3.17 ± 1.04	0.133	17%
160 *GeV* **(Schnitte)**	0.303 ± 0.004	8.24 ± 1.05	0.106	22%
180 *GeV* **(ANN)**	0.177 ± 0.002	5.1 ± 1.1	0.078	19%

Tabelle 7.13: In der Tabelle können Signal- sowie Untergrundereignisse und Signaleffizienzen sowie $\frac{S}{\sqrt{B}}$-Verhältnisse verglichen werden. Sie werden für alle Higgsmassenpunkte und Analysemethoden (in den Klammern angegeben) gezeigt.

7.7 Anwendung auf Daten

Nachdem die Methode an sich sowie ihre Ergebnisse für die verschiedenen Higgsmassen vorgestellt wurden, könnte die Selektion mittels Neuronaler Netze auch auf den Datensatz angewendet werden. Wie bereits gezeigt werden konnte, liegen die Ergebnisse dieser Analyse (160 *GeV*-Higgsmasse) im Vergleich zu einer schnittbasierten Analyse in der Optimierung des Verhältnisses $\frac{S}{\sqrt{B}}$ um 25% besser. Daraus kann geschlußfolgert werden, daß die hier beschriebene Analysemethode sehr gut funktioniert und entsprechend für eine Suche nach einem hypothetischen Higgs-Boson eingesetzt werden kann.
Gegen eine Anwendung auf Daten ist hinsichtlich dieser Selektionsmethode nichts einzuwenden, da ihre Funktionalität eindeutig überprüft werden konnte. Jedoch wurde bereits während der Vorselektion darauf aufmerksam gemacht, daß die Betrachtung der Daten nicht als vollständig anzusehen ist, da die Beschreibung dieser durch Untergrund-Monte-Carlo-Simulationen nicht eindeutig erfolgen konnte (siehe Kapitel 5). Es sei nochmals daran erinnert, daß für die Analyse nicht alle notwendigen Untergrundprozesse involviert werden konnten aufgrund der unzureichenden Beschreibung durch Monte-Carlo-Simulationen. Dazu gehören der Υ-, der QCD-Mehrjet- und der $W + jets$-Untergrund. Auf Υ braucht in diesem Zusammenhang nicht näher eingegangen werden, da aufgrund der absoluten Separierbarkeit im Endergebnis der Anteil dieser Ereignisse zu 0 erwartet wird [31]. Demhingegen gelten QCD-Mehrjet wie auch $W + jets$ nicht als vollständig separabel, so daß durchaus mit selektierten Ereignissen zu rechnen ist.
Ein weiteres problematisches Kriterium ist die anfänglich erwähnte *Jet Energy Scale* zur Korrektur der Jetenergie. Die Korrektur kann in Kapitel 4 nachgelesen werden. Zum Zeitpunkt der Analyse war sie bezüglich der hier benutzten Release-Version noch nicht erhältlich und konnte demnach nicht in die Arbeit aufgenommen werden. Zur Verdeutlichung dieser Problematik sei auf Abbildung 7.1 verwiesen, in welcher nochmal die Verteilungen der fehlenden transversalen Energie von Untergrund und Daten aufgetragen sind. Deutlich heben sich die Daten vom Untergrund bei hohen $\not{E}_T$ -Werten ab. Diese Abweichungen können auf die nicht durchgeführte Jetkorrektur zurückgeführt werden (vgl. [6]).

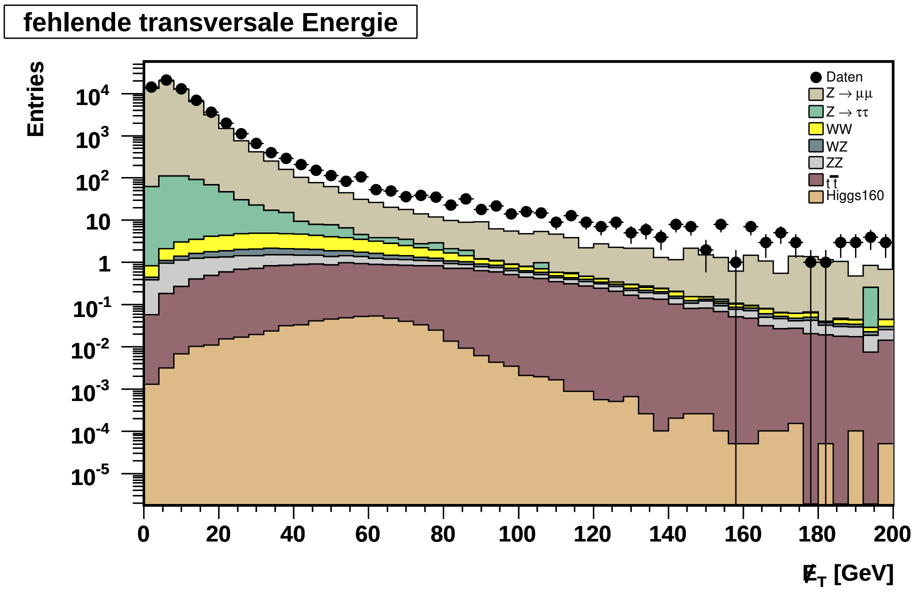

Abbildung 7.1: Die Verteilungen dieser Abbildung zeigen die fehlende transversale Energie für Untergrund, Daten und Signal. Deutlich zeichnet sich ein erheblicher Überschuß an Ereignissen bei hohem $\not{E}_T$ in den Daten ab. Diese Diskrepanz wird der nicht durchgeführten Jetenergiekorrektur zugeschrieben, welche für den Überschuß in den Daten verantwortlich sein kann.

Auch die Verteilung des Transversalimpulses des führenden Myons (siehe Abb. 5.5) im Vergleich mit den Daten zeigt eine Diskrepanz zwischen beiden. Als Ursache für diesen Unterschied kann angegeben werden, daß kein expliziter Schnitt auf die Treffer im SMT innerhalb der Vorselektion verlangt wird. Besonders problematisch sind hierbei Ereignisse, die keinen Treffer im *Silicon Microstrip Tracker* aufweisen können, ihre Rekonstruktion gestaltet sich fehlerhaft und muß in diesem Rahmen hier zusätzlich erwähnt werden.
Darüberhinaus ist die hier durchgeführte Verschmierung der Transversalimpulse in den Monte-Carlo-Simulationen noch nicht optimal. Aufgrund der verschiedenen Datennahmeperioden durch die *Shutdowns* ist eine exakte Verschmierung der p_T's sehr schwierig. D.h. durch die Unterbrechungsphasen der Proton-Antiproton-Kollisionen, in denen diverse Systemkalibrierungen vorgenommen werden (beispielsweise wurde im letzten Shutdown die Stärke des Magnetfeldes reduziert aufgrund von Hardwareproblemen [31]), kann die Auflösung in den Daten stark variieren, so daß die Verschmierung in den Monte-Carlo-Simulationen zur Anpassung an die Daten nicht für alle Perioden optimal erscheint. Hierin ist eine weitere Fehlerquelle zur Daten-Monte-Carlo-Beschreibung zu sehen.
Da eine angewendete Selektion auf die Daten am Ergebnis bezüglich des Vergleiches beider Methoden nichts ändern würde, kann in diesem Zusammenhang auf die genaue Anzahl verbleibender Ereignisse in den Daten verzichtet werden. Ein Unterschied zwischen Daten und Monte-Carlo-Simulationen würde für beide Selektionsmethoden aufgrund der unzureichenden Beschreibung der Daten durch die Simulationen erwartet werden, da für die Netz-Analyse wie auch für die Schnitt-Analyse derselbe Datensatz

verwendet wird.

Kapitel 8

Zusammenfassung und Ausblick

Die vorliegende Arbeit beschreibt die Suche nach dem Standardmodell-Higgs-Boson im Zerfallskanal $H \longrightarrow W^+W^- \longrightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ mit Hilfe Neuronaler Netze. Ausgehend von Proton-Antiproton-Kollisionen einer Schwerpunktsenergie von $\sqrt{s} = 1.96\ TeV$ erfolgte die Suche mit Daten des DØ-Detektors. Diese Daten wurden von April 2002 bis Dezember 2005 aufgezeichnet und entsprechen bezüglich dieser Analyse einer integrierten Luminosität von 775 pb^{-1}.
Die Entdeckung des Higgs-Bosons, dessen Existenz nach dem jetzigen Kenntnisstand durch das Standardmodell gefordert wird, stellt eine besonders große Herausfoderung für die moderne Teilchenphysik dar. Denn nur mit seiner Entdeckung, welche notwendig zum Verständnis des Massenursprungs von Fermionen und Bosonen wäre, kann unser heutiges Standardmodell verifiziert werden. Da die Theorie die Masse des Higgs-Bosons selbst als einen freien Parameter beschreibt, der proportional zur seiner Selbstkopplungskonstante ist, sind theoretisch nur Eingrenzungsgebiete zur Suche gegeben. Bisher konnten lediglich Massen unter 114.4 GeV und über 290 GeV ausgeschlossen werden.
In der vorliegenden Analyse wurde der Schwerpunkt auf eine hypothetische Masse des Higgs-Bosons von 160 GeV gesetzt. Darüberhinaus konnte der Algorithmus auf weitere hypothetische Higgsmassen (120 GeV, 140 GeV und 180 GeV) angewendet werden. Ziel war es, die Selektion mittels Neuronaler Netze bestmöglichst zu optimieren. Dazu wurde eine Kombination mehrerer Neuronaler Netze geschaffen, deren Selektivität im einzelnen den zur Verfügung stehenden Untergründen angepaßt wurde. Dabei weist speziell der erste Schnitt der Optimierung noch Verbesserungsmöglichkeiten auf, da dieser hinsichtlich der bereits optimierten Schnittpunkte abschließend neu zu studieren wäre.
Die Anzahl der Neuronalen Netze richtete sich in dieser Analyse nach der Anzahl der implementierten Untergrundprozesse. Um die Qualität dieser Selektionsmethode realistisch einschätzen zu können, wurde eine Vergleichsbasis geschaffen. Deshalb war es notwendig nicht nur mittels Neuronaler Netzen zu selektieren, sondern darüberhinaus auch eine Selektion, welche auf Schnitten aufgebaut ist, durchzuführen. Für beide Methoden sind die Verhältnisse von Signal und Untergrund ($\frac{S}{\sqrt{B}}$) sowie die Signaleffizienzen berechnet worden. Je besser diese Relationen ausfielen, desto besser war die Methode einzustufen. Mit dem Verhältnis $\frac{S}{\sqrt{B}}$ konnte gezeigt werden, daß eine Selektion, welche durch Neuronale Netze erfolgt, eine bessere Sensitivität besitzt.

Die restlichen drei Higgsmassen wurden mit derselben Methode untersucht, jedoch verzichtete sie, einen Vergleich mit einer schnittbasierten Analyse anzustreben. Den neuen Signalen entsprechend wurden die Neuronalen Netze nochmals einer Trainings- und Testphase unterzogen, um für eine Selektion heranziehbar zu sein. Die Ergebnisse dieser Massenpunkte konnten nicht die Ergebnisse der 160 GeV-Masse erzielen. Besonders das 120 GeV-Higgs-Boson ist in diesem Zusammenhang zu nennen, da es sich mit einer sehr niedrigen Signaleffizienz von 2% und einem Verhältnis von $\frac{S}{\sqrt{B}} = 0.004$ deutlich von den anderen Selektionen abhebt.
In diesem Zusammenhang wird noch einmal erwähnt, daß die Optimierung der vorliegenden Arbeit auf Basis des 160 GeV-Higgs-Bosons erfolgte und demnach gerade für diesen Massenpunkt die Struktur und auch die verwendeten Variablen dementsprechend effektiv sind. Zur Untersuchung weiterer möglicher Massen eines Higgs-Bosons wird empfohlen, die Analyse unter Verwendung Neuronaler Netze nicht nur hinsichtlich neuer Trainings- und Testereignisse zu optimieren, sondern vielmehr von grundauf neu zu gestalten. Dies beeinhaltet sowohl den Aufbau hinsichtlich der Anzahl von Neuronen und Schichten sowie die in das Netz eingehenden Variablen neu zu studieren. Möglichkeiten zur Verbesserung sind insofern gegeben, daß einzelne Variablen anderer Higgsmassen so stark im Vergleich zum 160 GeV-Higgs variieren, daß der Einfluß und die Effektivität dieser Variablen auf das Netz gänzlich verloren gehen kann, währenddessen andere neue Separationsparameter unberücksichtigt blieben. Diese Abhängigkeiten gilt es, für eine optimale Analyse weiterer Higgsmassen zu studieren.
Hinsichtlich dieser Analyse lagen die wichtigen Erkenntnisse in der Effizienz der Methode an sich, d.h. die Frage, wie besser selektiert werden kann, stand hierbei im Vordergrund. Es konnte gezeigt werden, daß diese Selektion mittels Neuronaler Netze zu geringeren Signaleffizienzen (5% weniger, aufgrund der hier vorgestellten Optimierung des Signal-zu-Untergrund-Verhältnisses) als im Vergleich zu der schnittbasierten Analyse führt, aber letzendlich ein besseres $\frac{S}{\sqrt{B}}$-Verhältnis (25% mehr) erzielt. Jedoch erfolgt eine Optimierung der Methode nur für die Netz-Analyse, hinsichtlich der Schnitt-Selektion wurden Schnitte aus [6] übernommen, welche entsprechend des Datensatzes dieser Analyse angepaßt sind und demzufolge nicht auf die Luminosität des hier verwendeten Datensatzes optimiert sein können. Mit einer nachträglich durchgeführten Optimierung der Schnitte sind Möglichkeiten zur Verbesserung der Schnittselektion gegeben.
Im weiteren muß die Daten- und Monte-Carlo-Beschreibung weiter verbessert werden, da diese zu große Diskrepanzen aufweist. In weiterführender Arbeit sollten in Bezug auf diese Analyse, die noch nicht betrachteten Untergrundprozesse einbezogen werden, damit eine Abschätzung dieser Ereigniszahlen gegeben werden kann. Besonders QCD-Mehrjet und $W + jets$-Prozesse sind hier zu nennen. Auch zusätzliche Zerfallsmöglichkeiten von ZZ-Untergrundereignissen ($ZZ \rightarrow q\bar{q}q\bar{q}$ und $ZZ \rightarrow \ell\bar{\ell}\nu\bar{\nu}$) müssen studiert werden.
Darüberhinaus muß zum konsistenten Vergleich von Untergrund und Daten die *Jet Energy Scale* zur Verbesserung der Jetenergie sowie die Myonimpulsauflösung näher betrachtet werden aufgrund nicht korrigierter Jetenergie- und Myonimpulswerte. Weitere Optimierungsmöglichkeiten werden auch durch die Verschmierung der p_T der Monte-Carlo-Simulationen gegeben. Diese Verschmierung sollte nicht einheitlich für

alle Transversalimpulse erfolgen, sondern vielmehr nach den Auflösungen in den Datenperioden unterschieden und dementsprechend kombiniert werden.
Abschließend wird festgehalten, daß auf der Suche des Higgs-Bosons des hier vorgestellten Zerfallskanal und der hier vorgestellten Selektionsmethode Verbesserungen hinsichtlich traditioneller Methoden erreicht werden konnten. Auch wenn die Analyse nicht bis ins Detail optimierbar war, sollte eines offensichtlich geworden sein: Die Ergebnisse der Selektion mit Hilfe Neuronaler Netze im Vergleich zu Schnitt-Selektionen liegen um 25% besser. Durch die Anwendung der oben genannten Verbesserungsvorschläge sollte sich das Potential der Selektionsmethode weiter steigern lassen. Auch Ergebnisse in den Daten sollten sich unter Berücksichtigung der noch anzustrebenden Erweiterungen erzielen lassen, so daß letztendlich eine Vergleichsbasis zwischen Daten, Signal und Untergrund geschaffen werden kann. Mit dieser können Aussagen bezüglich der möglichen Existenz eines Higgs-Bosons getroffen und darüberhinaus die Ausschlußgrenzen weiter verfeinert werden.

Anhang A

Die Pseudorapidität

Durch die Pseudorapidität η kann die im Experiment schwieriger zu bestimmende Rapidität y leicht approximiert werden, welche sich als sehr hilfreich in der Beschreibung inklusiv gemessener Reaktionen erwiesen hat. Die Rapidität definiert sich zu

$$y = \frac{1}{2} \cdot \ln \frac{E + p_z}{E - p_z} \quad .$$

Umformungen ergeben

$$\tanh(y) = \frac{p_z}{E}$$

Der longitudinale Impuls entlang der Einfallsrichtung[1] des Teilchens ist mit p_z gegeben, die Energie wird durch E dargestellt.
Sofern die Masse des einfallenden Teilchens unbekannt ist, ist eine Ersetzung der Rapidität durch die Pseudorapidität möglich, die sich als Winkelvariable wie folgt definiert:

$$\eta = -\ln(\tan \frac{\theta}{2}) \quad .$$

Die inverse Funktion dieser Gleichung ermittelt sich zu

$$\theta = 2 \cdot \tan^{-1}(e^{-\eta}) \quad .$$

Der Polarwinkel θ ist in diesem Fall der Winkel zwischen Teilchen und Strahlachse und mit einfachem Aufwand innerhalb des Detektors messbar.
Für eine vernachlässigbare Masse, sprich also wenn $p \gg m$, geht y gegen η. Sie stimmen exakt überein, wenn

$$m = 0$$

ist. In der experimentellen Anwendung findet die Pseudorapidität jedoch weitaus größeren Anklang als der Polarwinkel, da Differenzen in η im Gegensatz zu solchen in θ annähernd lorentzinvariant sind. Außerdem ist bekannt, daß das Schwerpunktsystem der Parton-Parton-Streuung bezüglich des der einfallenden Teilchen geboostet ist, was die Verwendung solcher, unter diesen Transformationen invarianter oder unter longitudinalen Boosts leicht transformierbarer Variablen motiviert.

[1] hier die z-Achse

Um den Zusammenhang zwischen η und θ zu verdeutlichen sei auf Tabelle A.1 verwiesen. Diese orientiert sich am experimentellen Arbeitsbereich des DØ-Detektors, der einen Pseudorapiditätsbereich von $-4 < \eta < +4$ abdeckt. Dementsprechend ist in Tab. A.1 abzulesen, daß nur Teilchen mit einer Mindestentfernung von 2.1° von der Strahlachse im Detektor nachzuweisen sind. [36]

θ [in °]	90	45	40.4	15	10	2.1
η	0	0.88	1	2.03	2.44	4

Tabelle A.1: Zusammenhang zwischen η und θ.

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[1] G. HACKER, *Grundlagen der Teilchenphysik*. URL: `http://www.didaktik.physik.uni-erlangen.de` (Stand: 13.März 2006)

[2] L. FINKE, *Das Standardmodell der Teilchenphysik - Eine Einführung*. Moderne Methoden/Experimente der Teilchen- und Astroteilchenphysik (Wintersemester 01/02), Physikalische Institute, RWTH Aachen

[3] D. PERKINS, *Introduction to High Energy Physics*. (Addison-Wesley, New York), (1987), ISBN 0-521-62196-8

[4] D.J. GRIFFITHS, *Introduction to Elementary Particles*. (John Wiley & Sons, Inc.), (1987), ISBN 0-471-60386-4

[5] P. SCHMÜSER, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*. Lecture notes in physics; Vol.295, (Springer-Verlag, Berlin Heidelberg New York London Paris Tokyo), (1995), ISBN 3-540-58486-2 NE:GT

[6] J. ELMSHEUSER, *Search for Higgs Boson in $H \rightarrow WW$ Decays at the DØ Experiment and Precise Muon Tracking*. Dissertation, Ludwig-Maximilians-Universität München, (Mai 2004)

[7] W.N. COTTINGHAM und D.A. GREENWOOD, *An Introduction to the Standard Model of Particle Physics*. XVIII, hardback, (Cambridge University Press, Cambridge), (1998), ISBN 0-521-58832-4

[8] S. WEINBERG, *The Quantum Theory of Fields Vol. II*. XXI, hardback, (Cambridge University Press, Cambridge), (1996), ISBN 0-521-55002-5

[9] F. HALZEN und A.D. MARTIN, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*. (John Wiley & Sons, Inc.), (1984), ISBN 0-471-88741-2

[10] A.S. INDENHUCK, *Das Standardmodell der Teilchenphysik*. (Februar 2004)

[11] THE LEP und SLD COLLABORATION, *A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model*. hep-ex/0312023 (2003)

[12] R.K. ELLIS, W.J. STIRLING und B.R. WEBBER, *QCD and Collider Physics*. (Cambridge University Press, Cambridge), (1996), ISBN 0-521-58189-3

[13] A. DJOUADI, J. KALINOWSKI und M. SPIRA, *HDECAY: A Programm for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension*. Comput. Phys. Commun. 108:56-74 (1998), DESY-97-079, IFT-96-29, PM-97-04

[14] THE DØ COLLABORATION, *The Upgraded DØ Detector: Nuclear Instruments and Methods*. physics/10507191, Fermilab-Pub-05/341-E. (2005)

[15] THE CDF COLLABORATION F. ABE et al., *Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab*. Phys. Rev. Lett. **74** (1995), 2626

[16] THE DØ COLLABORATION S. ABACHI et al., *Observation of the Top Quark*. Phys. Rev. Lett. **74** (1995), 2632

[17] G. JACKSON, *The Fermilab Recycler Ring Technical Design Report*. FERMILAB-TM-1991 (1994)

[18] J. THOMPSON, *Introduction to Colliding Beams at Fermilab*. FERMILAB-TM-1909 (1994)

[19] THE DØ COLLABORATION S. ABACHI et al., *The DØ Upgrade: The Detector and its Physics*. FERMILAB Pub 96/357-E (1996)

[20] U. BAUR, R.K. ELLIS und D. ZEPPENFELD, *QCD and Weak Boson Physics in Run II*. Fermilab-Pub-00/297 (2000)

[21] T. LECOMPTE und H.T. DIEHL, *The CDF and DØ Upgrades for Run II*. Ann. Rev. Nucl. Part. Sci. 50:71-77 (2000), 0163-8998/00/1201-0071

[22] S. KLIMENKO, J. KONIGSBERG und T.M. LISS, *Averaging of the Inelastic Cross Sections measured by the CDF and the E811 Experiments*. FERMILAB-FN-0741 (2003)

[23] B. CASEY et al., *The Upgraded DØ Luminosity Determination* (2004), DØ Note 4328-PHYS

[24] THE MUON-ID GROUP, *Muon-ID certification v3.0*. URL: `http://www-d0.fnal.gov/phys_id/muon_id/d0_private/muon_id.html`
(Stand: 12.Januar 2006)

[25] THE JET ENERGY SCALE GROUP, *Certified Jet Energy Scale v5.3*. URL: `http://www-d0.fnal.gov/phys_id/jes/d0_private/certified/certified.html`
(Stand: 13.Januar 2006)

[26] THE DØ COLLABORATION, *DØ Jet and MET ID Properties in p11 and p13*. URL: `http://www-d0.fnal.gov/d0upgrad/d0_private/software/jetid/certification/v2_0/jetid_p13.html`
(Stand: 15.Januar 2006)

[27] T. SJØSTRAND. Comput. Phys. Commun. 135, 238-259, (2001)

[28] M. MANGANO et al. JHEP07, 001, (2003)

[29] R. BRUN, *GEANT 3*. CERN-DD/EE/84-1, (1984)

[30] MONTE CARLO WORKING GROUP, *Sample Summary Table*. URL: `http://d0-france.in2p3.fr/WORKING_GROUPS/MONTECARLO/Catalog/current/section_summary.php` (Stand: 13.März 2006)

[31] J. ELMSHEUSER, *Private Communications*. (2005)

[32] PROF. DR. W.-M. LIPPE, *Interaktive Einführung in Neuronale Netze*. URL: `http://wwwmath.uni-muenster.de/SoftComputing/lehre/Material/wwwnnscript`
(Stand: 27.Januar 2006)

[33] R. ROJAS, *Neural Networks - A Systematic Introduction*. (Springer Verlag, Berlin Heidelberg New York London Paris Tokyo), (1996), ISBN 3-540-60505-3

[34] J. MÖCK, *Einsatz Neuronaler Netze als intelligente Trigger im H1-Experiment*. Diplomarbeit, TU München, Werner-Heisenberg-Institut, (März 1994)

[35] U. LÄMMEL und J. CLEVE, *Lehr- und Übungsbuch künstliche Intelligenz*. 2. vollständig überarbeitete Auflage, (Fachbuchverlag Leipzig im Carl Hanser Verlag, München, Wien), (2004), ISBN 3-446-22574-9

[36] M. ERLEBACH, *Studien zu Multijet Ereignissen am TeVatron*. Diplomarbeit, LMU München, (Juli 2004)

D A N K S A G U N G

Ich möchte an dieser Stelle die Gelegenheit nutzen und den Menschen danken, die maßgeblich zum Gelingen dieser Arbeit beigetragen haben, deshalb geht mein herzlicher Dank an ...

- ... Prof. Dr. Otmar Biebel, der mir diese Analyse ermöglicht hat und der mir immer helfend zur Seite stand, bei allem, was im vergangenem Jahr an Schwierigkeiten angefallen ist.

- ... Prof. Dr. Gerhard Buchalla für seine Erstellung des Zweitgutachtens.

- ... Prof. Dr. Dorothee Schaile für den Arbeitsplatz an ihrem Lehrstuhl.

- ... Dr. Johannes Elmsheuser und Dr. Thomas Nunnemann, für die Betreuung meiner Arbeit sowie ihre wertvollen Hinweise und Ratschläge zu dieser Analyse.

- ... Marion Lambacher und Britta Tiller, die durch ihre offene und freundliche Art mir jederzeit ihre Hilfe, ihr Verständnis und ihren Zuspruch hinsichtlich dieser Arbeit und hinsichtlich privater Belange zukommen ließen.

- ... Philippe Calfayan und Alexander Grohsjean, meine Anlaufpunkte bei programmiertechnischen Schwierigkeiten, mit denen ich auch darüberhinaus sehr anregende und informative Unterhaltungen führen konnte.

- ... Dr. Johannes Elmsheuser und Raphael Mameghani, mit denen ich mir ein Büro geteilt habe und mit denen es in diesem Büro nie langweilig wurde.

- ... Herta Franz, unserer Sekretärin, für ihr offenes Ohr und die kleinen Gespräche zwischendurch.

- ... den Lehrstuhl der Elementarteilchenphysik mit allen Mitgliedern, die ich während meiner Zeit kennlernen durfte: Otmar Biebel, Meta Binder, Philippe Calfayan, Günter Duckeck, Johannes Elmsheuser, Frank Fiedler, Herta Franz, Alexander Grohsjean, Petra Haefner, Ralf Hertenberger, John Kennedy, Tatjana Kennedy, Gernoth Krobath, Marion Lambacher, Raphael Mameghani, Doris Merkl, Alexander Mlynek, Thomas Müller, Thomas Nunnemann, Matthias Obermaier, Felix Rauscher, Benjamin Ruckert, Dorothee Schaile, Philipp

Schieferdecker, Matthias Schott, Cedric Serfon, Raimund Ströhmer, Britta Tiller, Attila Varga und Fritz Vollmer.

- ... Martina Lange, Melanie Pippig, Nadine Pippig und Claudia Schmidt sowie Johannes Ferber, Christian Lojewski und Martin Pregler für ihre unverzichtbare und wertvolle Freundschaft.

- ... meine Schwester Manuela, dafür, daß sie nicht nur meine Schwester ist, sondern vielmehr auch meine Freundin.

- ... meine Eltern Gudrun und Klaus, die mir ein wundervolles Elternhaus waren und immer sein werden, die mich stets in allem unterstützt haben und die mir durch ihren Zuspruch und durch ihre Kritik gezeigt haben, wie sehr sie mich lieben.

- ... meine Großeltern sowie dem Rest meiner Familie für ein wundervolles Leben mit ihnen.

- ... meinen Freund und Verlobten Andreas, der bedingungslos hinter mir steht und dessen Liebe und Zuneigung mir den wirklichen Sinn meines Lebens gezeigt haben. In ihm habe ich den Menschen gefunden, der mich den Wert dieser Arbeit und die Zeit zu ihrer Vollendung erst richtig schätzen lassen. Ihm und unserer gemeinsamen Zukunft widme ich diese Arbeit.

σ' αγαπώ

Selbstständigkeitserklärung

Ich versichere hiermit, die vorliegende Arbeit selbstständig verfaßt zu haben
und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Daniela Görisch

München, 24. März 2006