



STUDIEN ZUR  
HIGGS-BOSON-SUCHE IM  
VEKTOR-BOSON-FUSION-KANAL  
MIT DEM CMS-EXPERIMENT

SANDRA KOWALCZYK

DIPLOMARBEIT

*bei Prof. Dr. W. Wagner  
Bergische Universität Wuppertal*

*Korreferent Prof. Dr. Th. Müller  
Institut für Experimentelle Kernphysik*

AN DER FAKULTÄT FÜR PHYSIK  
DER UNIVERSITÄT KARLSRUHE (TH)

30. JANUAR 2009



# Zusammenfassung

In der Teilchenphysik werden die fundamentalen Bestandteile der Materie, die Elementarteilchen, und ihre Wechselwirkungen untersucht, um ein tieferes Verständnis der zugrunde liegenden physikalischen Prozesse zu erlangen. Nach dem gegenwärtigen Wissensstand und mit experimentellen Untersuchungen weitestgehend vereinbar werden sowohl die elementaren Teilchen als auch ihre Wechselwirkungen durch das Standardmodell (SM) der Elementarteilchenphysik beschrieben. In diesem Rahmen wird zusätzlich ein weiteres Teilchen, das Higgs-Boson, vorhergesagt, das bis heute noch nicht entdeckt werden konnte. Mit dessen Hilfe wird eine elegante theoretische Formulierung ermöglicht, die zur Vereinheitlichung der schwachen mit der elektromagnetischen Wechselwirkung benötigt wird. Darüber hinaus werden durch Wechselwirkungen mit dem Higgs-Boson die Massen aller Teilchen erzeugt.

Einen Gegenstand der aktuellen Forschung stellen somit die Elementarteilchen dar, die jedoch subatomare Teilchen sind. Viele von ihnen treten unter normalen Gegebenheiten in der Natur nicht auf, so dass große Teilchenbeschleuniger benötigt werden, um die Teilchen in hochenergetischen Kollisionen zu erzeugen und schließlich mit Hilfe von Detektoren nachzuweisen. Derzeit ist das Tevatron, ein Proton-Antiproton-Beschleuniger in der Nähe von Chicago, mit 1.96 TeV der Teilchenbeschleuniger mit der weltweit höchsten Schwerpunktsenergie bei Teilchenkollisionen. Wenn jedoch diesen Sommer der Large Hadron Collider (LHC) am Europäischen Kernforschungszentrum (CERN) bei Genf in Betrieb genommen wird, wird es möglich werden, Schwerpunktsenergien bis zu 14 TeV bei einer Kollision zu erreichen. Beim LHC werden im Gegensatz zum Tevatron Proton-Proton-Kollisionen durchgeführt werden. Bei beiden Beschleunigern muss jedoch beachtet werden, dass sowohl Protonen als auch Antiprotonen zusammengesetzte Teilchen aus Quarks und Gluonen sind. Somit findet die eigentliche Kollision unter diesen Bestandteilen statt, die jedoch nur einen Bruchteil der kinetischen Protonenergie tragen. Dadurch steht nur ein geringerer Teil der genannten Schwerpunktsenergie bei einer Teilchenkollision zur Verfügung.

Der Trend zu Teilchenbeschleunigern mit immer höheren Schwerpunktsenergien wird dadurch begründet, dass Teilchen mit hohen Ruhemassen nur dann erzeugt werden können, wenn ausreichend Energie bei einer Kollision verfügbar ist. Dies ist auch von besonderer Bedeutung für die Erzeugung eines Higgs-Bosons, da die experimentell ermittelte Untergrenze für seine Masse bereits bei etwa  $115 \text{ GeV}/c^2$  liegt. Somit ist die Suche nach dem Higgs-Boson, dem letzten vom SM vorausgesagten und noch nicht beobachteten Teilchen, eines der Hauptziele für den LHC.

Auch wenn die Existenz des Higgs-Bosons bis jetzt noch nicht nachgewiesen wer-

den konnte, gibt es bereits genau berechnete Wirkungsquerschnitte für seine Erzeugung in verschiedenen, möglichen Kanälen und für unterschiedlich angenommene Higgs-Boson-Massen. Sowohl die Wirkungsquerschnitte als auch die Verzweigungsverhältnisse des Higgs-Bosons sind durch das SM bestimmt. In der vorliegenden Arbeit wird die Erzeugung des Higgs-Bosons im Vektor-Boson-Fusion-Kanal (VBF), also mittels zwei  $W$ - oder  $Z$ -Bosonen, untersucht. Für diesen Produktionskanal sind zwei vorwärts gerichtete Jets charakteristisch, was zu einer guten Untergrundunterdrückung führt, so dass der VBF-Prozess zu einem der wichtigsten Produktionsmechanismen am LHC wird, obwohl der Wirkungsquerschnitt des Gluon-Fusion-Kanals ungefähr zehn Mal größer ist. Außerdem wird durch den VBF-Mechanismus ein experimenteller Test des SM ermöglicht, da das Higgs-Boson direkt an die beiden schwachen Vektorbosonen koppelt und der Wirkungsquerschnitt somit proportional zum Quadrat dieser Kopplung ist. Durch eine präzise Messung des Wirkungsquerschnittes entsteht demnach ein direkter Zugang zu einer vom SM vorhergesagten Eigenschaft des Higgs-Bosons, nämlich der Stärke der Kopplung an die schwachen Vektorbosonen. Sollten größere Abweichungen zur theoretischen Vorhersage bei der experimentellen Bestimmung des Wirkungsquerschnittes gemessen werden, müsste die Theorie des SM mit dem involvierten Higgs-Boson neu überdacht werden.

Bei dieser Analyse wird der Zerfall des Higgs-Bosons in ein  $W$ -Bosonpaar betrachtet, bei dem beide  $W$ -Bosonen wiederum leptonisch zerfallen sollen, da Leptonen ein sauberes Detektorsignal liefern. Dadurch ist eine ausgeprägte Signalereignis-signatur gegeben, insbesondere wenn noch die beiden aus der VBF-Produktion stammenden Vorwärtsjets mit einbezogen werden. Dennoch existieren eine Reihe weiterer Prozesse, die den gleichen Endzustand vorweisen können und somit die gleiche Signatur im Detektor hervorbringen. Diese Problematik wird noch unterstützt durch den geringen Wirkungsquerschnitt des gesamten, betrachteten Prozesses  $qq' \rightarrow q''q'''H \rightarrow q''q'''WW' \rightarrow q''q'''l\nu'l'\nu'$ , der bei 0.12 pb für eine angenommenen Higgs-Bosonmasse von 130 GeV/c<sup>2</sup> liegt und auf 0.31 pb für 160 GeV/c<sup>2</sup> ansteigt. Diese deutliche Zunahme des Wirkungsquerschnittes bei höheren Higgs-Bosonmassen ist durch das steil ansteigende Verzweigungsverhältnis des Higgs-Bosons in ein  $W$ -Bosonpaar bedingt. Insgesamt liegt die Hauptaufgabe bei der Analyse darin, eine Methode zu entwickeln, die die Untergründe bestmöglichst reduziert, während das gesuchte Signal bestenfalls kaum verringert wird.

Allgemein ist die Analyse für das CMS-Experiment konzipiert, einem Vielzweckteilchendetektor am LHC. Hierfür werden Monte-Carlo-Ereignisse mit verschiedenen Generatoren sowohl für den Signalprozess als auch für verschiedene Untergründe erzeugt und mit der CMS-Software vollständig simuliert, so dass mit diesen Monte-Carlo-Daten eine Methode entwickelt werden kann. Für die Analyse wird eine Higgs-Bosonmasse von 130 GeV/c<sup>2</sup> angenommen. Durch den geringeren Wirkungsquerschnitt für kleine Higgs-Bosonmassen als für größere, stellt dieser gewählte Massenbereich in diesem Kanal einen schwierigen Fall dar. In einem ersten Schritt wurde eine schnittbasierte Vorselektion entwickelt, die dazu dient, in Frage kommende Kandidaten aus allen Prozessen zu identifizieren. Bedingt durch die Signalsignatur, werden mindestens zwei bis maximal fünf Jets, mindestens zwei isolierte geladene Leptonen sowie fehlende Energiedeposition in transversaler Richtung zur Strahlachse

gefordert. Letztere entsteht dadurch, dass die beiden erzeugten Neutrinos  $\nu$  nicht im Detektor nachgewiesen werden können, da sie nicht mit dem Detektormaterial wechselwirken. Hinzu kommen nur leichte kinematische Schnitte auf die genannten physikalischen Objekte, wodurch die Effizienz des Signalprozesses möglichst hoch gehalten wird. Alle selektierten Ereignisse erbringen letztendlich die geforderte Signatur und beinhalten ausschließlich klar definierte physikalische Objekte, was für die weitere Analyse von großer Bedeutung ist. Nach der Selektion der Higgs-Boson-Kandidaten wird dennoch nur ein Anteil von  $\sim 0.1\%$  unter allen Ereignissen erwartet, der dem wirklichen Signalprozess entspricht. Es ist also eine leistungsfähigere Methode zur Trennung von Signal- und Untergründereignissen für die weitere Analyse notwendig, so dass hierfür eine multivariate Analyseemethode eingesetzt wird.

In einem zweiten Schritt werden neuronale Netze benutzt, um eine endgültige Separation von Signalereignissen und Untergründen zu erzielen. Mittels eines Trainings sowohl mit Signal- als auch Untergründereignissen lernt das Netz die beiden Ereignistypen anhand verschiedener Merkmale zu unterscheiden und kombiniert die einzelnen Informationen der Eingabevariablen zu einer einzelnen Größe. In dieser Analyse werden insgesamt zehn Eingabevariablen benutzt, die kinematische Eigenschaften der auftretenden Objekte beinhalten. So ist zum Beispiel die Winkelverteilung zwischen den beiden Jets und auch die zwischen den beiden geladenen Leptonen eine sehr wichtige Variable. Jedoch muss beachtet werden, dass bei  $\sim 27\%$  aller Signal- und  $\sim 34\%$  der Untergründereignisse mehr als die zwei geforderten Jets vorhanden sind. Diese Ereignisse werden getrennt von denen mit genau zwei Jets behandelt, da keine eindeutige Zuordnung der vorhandenen zu den notwendigen zwei Jets möglich ist, welche durch den VBF-Produktionsprozess bedingt sind. Es wird demnach eine Methode benötigt, die jene zwei Jets auswählt, die am ehesten den VBF-Jets gerecht werden. Dafür wird wiederum ein eigenes neuronales Netz trainiert, bei dem die Eingabevariablen für das Signal mit der Kombination der richtigen beiden Jets gebildet werden, während für den Untergrund die gleichen Variablen jedoch aus den falschen Jetkombinationen konstruiert werden. Dieses Netz wird nur mit Higgs-Boson-Ereignissen trainiert und die drei Jets mit dem höchsten transversalen Impuls werden für die Selektion ausgewählt. Bei etwa 85-90% der entsprechenden Higgs-Bosonereignisse werden die richtigen Jets mit den VBF-Jets identifiziert. Nach der Jetauswahl mit Hilfe des neuronalen Netzes für alle dazugehörenden Prozesse wird ein zweites Training durchgeführt, das dem der Ereignisse mit genau zwei Jets entspricht. Daher, dass die erwartete Untergrundzusammensetzung für Ereignisse mit genau zwei Jets und denen mit mehreren eine unterschiedliche ist, werden zwei getrennte Netze trainiert. Insgesamt werden bei jeder Anwendung eines neuronalen Netzes jeweils zwei verschiedene Netzwerkpakete benutzt, die miteinander verglichen werden. Es stellt sich jedoch heraus, dass die Ergebnisse ungefähr einander entsprechen und somit im Rahmen dieser Analyse miteinander vergleichbar sind.

Schließlich wird die erwartete Signifikanz dieser Analyseemethode für eine integrierte Luminosität von  $10 \text{ fb}^{-1}$  ermittelt. Hierfür wird auf die jeweilige Netzwerkausgabe geschnitten, die entsprechende Signalanzahl  $s$  und Untergrundanzahl  $b$  wird bestimmt, so dass die erwartete Signifikanz  $\sigma$  mit  $s/\sqrt{b}$  berechnet werden kann. Mit

## VI

---

$\sim 1.5 \sigma$  für Ereignisse mit genau zwei Jets und  $\sim 0.35 \sigma$  für jene mit mehr als zwei wird es nicht möglich sein, einen signifikanten Beitrag durch Higgs-Bosonereignisse mit einer angenommenen Higgs-Bosonmasse von  $130 \text{ GeV}/c^2$  bei einer Datenmenge von  $10 \text{ fb}^{-1}$  zu messen.



Universität Karlsruhe (TH)  
Research University • founded 1825

---

IEKP-KA/2009-7

STUDIES FOR THE HIGGS BOSON SEARCH  
IN THE VECTOR BOSON FUSION CHANNEL  
WITH THE CMS EXPERIMENT

SANDRA KOWALCZYK

DIPLOMA THESIS

AT THE FACULTY FOR PHYSICS  
OF THE UNIVERSITY OF KARLSRUHE (TH)

*Referee: Prof. Dr. W. Wagner  
Bergische Universität Wuppertal*

*Co-Referee: Prof. Dr. Th. Müller  
Institut für Experimentelle Kernphysik*

JANUARY 30, 2009



# Introduction

In particle physics the fundamental constituents of matter, the elementary particles, are studied to gain a deeper understanding of the occurring physical processes. The current classification of these particles as well as their interactions are described by the Standard Model (SM) of particle physics, which is until now consistent with conducted experimental tests. As the research is focused on subatomic particles with many of them not appearing under normal circumstances, large particle accelerators are the instrument to create and detect these particles during high energetic collisions. Currently, the proton-antiproton collider Tevatron near Chicago is the particle accelerator with the highest centre of mass energy, namely 1.96 TeV, until the Large Hadron Collider (LHC) at CERN will be commissioned in summer 2009. The LHC is designed for centre of mass energies up to 14 TeV, whereby a higher mass region for the produced particles is reached. This is particularly important, as there is still one additional particle, the Higgs boson, which existence is predicted by the SM and has not yet been discovered. So, one of the main aims of the LHC will be the search for the Higgs boson, the last remaining particle of the SM.

Although the Higgs boson has not been observed yet, its production cross section, theoretically determined by the SM, is already well calculated. In this thesis the production channel of the Higgs boson via two weak vector bosons  $WW$  or  $ZZ$  is considered, referred to as vector boson fusion (VBF) process, which features a distinct signature with two forward jets. Thus, a good reduction of many background processes is achieved, so that the VBF process becomes one of the most important production modes at the LHC, although the cross section of the gluon fusion channel is  $\sim 10$  times higher. Additionally, the VBF process provides an experimental test of the SM, since the coupling between the Higgs boson and the two weak vector bosons influences directly the production cross section. So, more precisely, this production channel offers an access to the question, if the Higgs boson behaves as predicted by the SM. In the case of larger deviations, the theory behind the SM has to be reconsidered.

This analysis is prepared for the CMS experiment, which is a multi-purpose detector at the LHC. Monte Carlo generated events have been used, which are passed to a full CMS detector simulation to obtain a realistic detector response. For the generated Higgs boson events, the decay into a pair of  $W$  bosons is demanded with both of them decaying leptonically. The Higgs boson mass is assumed to be  $130 \text{ GeV}/c^2$ . Although these signal events provide a distinct signal signature, many background processes with much larger cross sections, such as  $t\bar{t}$  production, occur in proton-proton collisions, whose reduction is the main challenge for the analysis.

Therefore, an event preselection with several selection cuts is developed to enrich the fraction of signal events in the analysed sample, however, the expected fraction of Higgs boson events is still only  $\sim 0.1\%$ .

For the final separation between signal and background events neural networks are trained. Kinematic variables of the involved particles have been studied for the signal and background processes to find the best discriminating ones. These final variables are combined and used as inputs for the neural network to distinguish signal-like events from the backgrounds. With the output of the neural network the expected significance of the Higgs signal is determined.

The organisation of this thesis is given as follows. The first chapter introduces briefly the SM of particle physics, gives afterwards an overview of the theory of the Higgs mechanism and concludes with the presentation of production and decay processes of the Higgs boson and the measured limits to its mass. The experimental site is described in the second chapter, including the main characteristics of the LHC and the CMS experiment. The third chapter introduces the event modelling with Monte Carlo event generators and describes the used simulated samples in brief. Additionally, some kinematic distributions of the signal events are presented to show the main features of the signal signature, while in chapter four the selection of event candidates and the expected number of signal and background events is described. The fifth chapter explains briefly the functionality of neural networks in general and highlights the main differences of the two used neural network packages. Finally, the neural network analysis is presented in chapter six and the thesis is concluded with a summary of the results in the last chapter.

# Contents

<b>1</b>	<b>The Higgs Boson within the SM</b>	<b>7</b>
1.1	The Standard Model of Elementary Particle Physics . . . . .	7
1.2	Local Gauge Invariance in Quantum Field Theory . . . . .	9
1.3	Spontaneous Symmetry Breaking . . . . .	11
1.4	The Higgs Mechanism . . . . .	12
1.5	Yukawa Sector . . . . .	13
1.6	Higgs Boson Production and Decay at the LHC . . . . .	14
1.7	Bounds for the Higgs Boson Mass . . . . .	17
<b>2</b>	<b>The Compact Muon Solenoid at the LHC</b>	<b>21</b>
2.1	The Large Hadron Collider . . . . .	21
2.2	The Compact Muon Solenoid . . . . .	23
2.2.1	Tracking System . . . . .	25
2.2.2	Electromagnetic calorimeter . . . . .	26
2.2.3	Hadron Calorimeter . . . . .	28
2.2.4	Muon System . . . . .	29
2.2.5	Trigger and Data Acquisition . . . . .	30
<b>3</b>	<b>Monte Carlo Studies</b>	<b>33</b>
3.1	Monte Carlo Event Generators . . . . .	33
3.2	Monte Carlo Signal Event . . . . .	36
3.3	Generated Monte Carlo Event Data . . . . .	38
3.4	CMS Software Framework . . . . .	41
<b>4</b>	<b>Selection of Candidate Events</b>	<b>43</b>
4.1	Event selection . . . . .	43
4.1.1	Lepton Selection . . . . .	43
4.1.2	Jet Reconstruction . . . . .	44
4.1.3	Missing Transverse Energy . . . . .	45
4.1.4	Z Boson Veto . . . . .	46
4.2	Signal and Background Estimation . . . . .	47
<b>5</b>	<b>Neural Networks</b>	<b>51</b>
5.1	NeuroBayes <sup>®</sup> and MLP . . . . .	52
5.1.1	Network Structure . . . . .	52
5.1.2	Training the Network . . . . .	53

---

5.1.3	Variable Preprocessing . . . . .	54
5.1.4	Variable Ranking . . . . .	55
5.1.5	Training Result . . . . .	55
<b>6</b>	<b>Analysis with Neural Networks</b>	<b>57</b>
6.1	Input Variables . . . . .	57
6.2	Neural Network Training . . . . .	59
6.2.1	Events with Two Jets . . . . .	59
6.2.2	Three-to-Five-Jet Events . . . . .	62
6.3	Comparison of the Network Packages . . . . .	69
6.4	Network Output . . . . .	70
6.5	Expected Significance . . . . .	73
6.6	Comparison to a Recent Analysis . . . . .	75
<b>7</b>	<b>Conclusion</b>	<b>77</b>
7.1	Summary . . . . .	77
7.2	Outlook . . . . .	78

# List of Figures

1.1	The potential of the Higgs field . . . . .	11
1.2	Higgs boson couplings to fermions and weak gauge bosons . . . . .	14
1.3	The parton distribution function . . . . .	15
1.4	LO Feynman diagrams of the Higgs boson production . . . . .	15
1.5	Cross sections of the Higgs boson production processes . . . . .	16
1.6	Branching ratios of the Higgs boson decay . . . . .	17
1.7	Branching ratios in the range $100 \text{ GeV}/c^2 \leq m_{Higgs} \leq 200 \text{ GeV}/c^2$ . . . . .	17
1.8	Fit of the Higgs boson mass derived from electroweak measurements . . . . .	18
1.9	Limits on the production cross section of the Higgs boson . . . . .	19
2.1	CERN accelerator complex . . . . .	22
2.2	A perspective view of the CMS detector . . . . .	24
2.3	Schematic cross section through the CMS tracker. . . . .	25
2.4	Electromagnetic calorimeter . . . . .	27
2.5	Hadron calorimeter . . . . .	28
2.6	Muon system . . . . .	29
2.7	General architecture of the CMS DAQ system . . . . .	31
3.1	Basic structure in the generation of a simulated event . . . . .	34
3.2	Feynman diagram of the signal process . . . . .	36
3.3	Kinematics of the quarks after boson radiation . . . . .	37
3.4	$W$ boson mass distributions . . . . .	37
3.5	Distribution of pseudorapidity $\eta(W2)$ vs $\eta(W1)$ . . . . .	38
3.6	Higgs boson decay in its rest frame . . . . .	38
3.7	Distribution of $\cos(\theta)$ between the two charged leptons and transverse momentum of the sum of the two neutrinos . . . . .	39
3.8	Feynman diagrams of $W/Z$ +jets and $t\bar{t}$ . . . . .	40
3.9	Feynman diagrams of diboson production . . . . .	40
3.10	Feynman diagrams of single top and QCD events . . . . .	40
3.11	Framework of the event data model . . . . .	42
4.1	Construction of a cone . . . . .	45
4.2	Invariant Mass distribution of the sum of the two charged leptons . . . . .	46
4.3	Expected fractions of the different processes . . . . .	47
5.1	Network with one hidden layer . . . . .	51
5.2	Sigmoid activation function . . . . .	53

---

5.3	Illustration of gradient descent method . . . . .	54
5.4	Example of a neural network output . . . . .	56
6.1	Expected fractions of the different processes for two-jet events . . . . .	59
6.2	Shapes of jet variables (2 jets) . . . . .	60
6.3	Shapes of lepton variables (2 jets) . . . . .	61
6.4	Expected fractions of the different processes for 3-to-5-jet events . . . . .	63
6.5	Event displays with two and one matched jets . . . . .	63
6.6	Shapes of jet variables for the jet selection (3 to 5 jets) . . . . .	64
6.7	Shapes of the network output after the jet selection . . . . .	65
6.8	Shapes of jet variables ( $\geq 3$ jets) . . . . .	67
6.9	Shapes of lepton variables ( $\geq 3$ jets) . . . . .	68
6.10	Background rejection versus signal efficiency for the trained networks . . . . .	70
6.11	Shapes of the network output (2 jets) . . . . .	71
6.12	Shapes of the network output (3 to 5 jets) . . . . .	71
6.13	Expected events for the two-jet network output . . . . .	72
6.14	Expected events for the 3-to-5-jet network output . . . . .	72
6.15	Significance for all final networks . . . . .	74

# List of Tables

1.1	Properties of the fermions (spin- $\frac{1}{2}$ particles)	8
1.2	Properties of the gauge bosons (spin-1 particles)	9
2.1	Energy resolution of the ECAL	27
2.2	Transverse momentum resolution in the muon system	30
3.1	The Monte Carlo samples	39
4.1	Signal and background events after preselection	48
4.2	QCD events after preselection	49
6.1	Variable ranking of NeuroBayes (2 jets)	62
6.2	Variable ranking of MLP (2 jets)	62
6.3	Variable ranking of NeuroBayes for the jet selection ( $\geq 3$ jets)	65
6.4	Variable ranking of MLP for the jet selection ( $\geq 3$ jets)	65
6.5	Variable ranking of NeuroBayes for the final selection ( $\geq 3$ jets)	69
6.6	Variable ranking of MLP for the final selection ( $\geq 3$ jets)	69
6.7	Significance of the NeuroBayes two-jet network	73
6.8	Significance of the MLP two-jet network	74
6.9	Significance of the cut-based analysis	75



# Chapter 1

## The Higgs Boson within the Standard Model

The Standard Model (SM) of particle physics describes nearly all experimental results satisfactorily. The Higgs boson is also predicted by the SM and plays a decisive role in the electroweak theory. With the Higgs mechanism electroweak symmetry is broken and the  $W$  and  $Z$  bosons gain their masses by interacting with a new field, the Higgs field. The existence of a neutral scalar particle, the massive Higgs boson, is implied by the mechanism. The Higgs boson is the only predicted particle within the SM, which has not been observed yet. More detailed information can be found for example in reference [1] and [5].

The following sections provide an overview about the SM. To motivate the Higgs mechanism, the main components of the underlying theory will be first introduced before the theory itself is presented. The remaining sections will treat the production and decay modes of the Higgs boson as well as the current limits on its mass.

### 1.1 The Standard Model of Elementary Particle Physics

The SM is a theory for elementary particles and their interactions among each other. There are three kinds of particles in the SM: the fermions with spin  $\frac{1}{2}$ , the gauge bosons with spin 1, which mediate the three fundamental forces: the strong, the electromagnetic and the weak force, and the Higgs boson with spin 0.

The fermions are the constituents of matter and are divided into two classes, quarks and leptons, depending on their interactions. There are six quarks (up, down, charm, strange, top and bottom) as well as six leptons (electron, muon, tau and their corresponding neutrinos). All of them carry weak isospin, leading to an interaction via the weak force. In addition to these particles a respective antiparticle exists with the same mass and the opposite electric charge. Some properties of the fermions and their arrangement into three generations are shown in table 1.1. The first-generation particles build up all macroscopic objects, while the other heavier particles only result from high energy reactions and decay quickly into the lighter

	name	symbol	el. charge [ $e$ ]	mass [ $\text{MeV}/c^2$ ]
1st generation	up quark	$u$	$\frac{2}{3}$	$(1.5 - 3.0)$
	down quark	$d$	$-\frac{1}{3}$	$(3.0 - 7.0)$
	electron	$e$	$-1$	0.511
	$e$ -neutrino	$\nu_e$	0	$< 2 \cdot 10^{-6}$
2nd generation	charm quark	$c$	$\frac{2}{3}$	$(1.25 \pm 0.09) \cdot 10^3$
	strange quark	$s$	$-\frac{1}{3}$	$(95 \pm 25)$
	muon	$\mu$	$-1$	106
	$\mu$ -neutrino	$\nu_\mu$	0	$< 0.190$
3rd generation	top quark	$t$	$\frac{2}{3}$	$(172.5 \pm 2.3) \cdot 10^3$
	bottom quark	$b$	$-\frac{1}{3}$	$(4.20 \pm 0.07) \cdot 10^3$
	tau	$\tau$	$-1$	1777
	$\tau$ -neutrino	$\nu_\tau$	0	$< 18.2$

Table 1.1: Properties of the fermions (spin- $\frac{1}{2}$  particles) [2]. The electric charge is given in units of the positron charge, the mass in units of  $\text{MeV}/c^2$ .

ones.

The leptons only partly carry an elementary electric charge. The neutrinos are neutral and can only interact with other particles via the weak force, while the other ones also interact electromagnetically. In contrast to the leptons all quarks carry an electric charge and additionally a colour charge of the type red, blue or green. Therefore, they are also subject to the strong interaction. As coloured particles have not been observed as free objects, they have to form colourless, composite particles – the hadrons. Thus, a quark with one colour and an antiquark with the corresponding anticolour are confined into mesons, while three quarks, all with different colours, are confined into baryons.

The gauge bosons are the force mediating particles of the strong and the electromagnetic force, described by the theory of quantum chromodynamics (QCD) and quantum electrodynamics (QED), respectively, and the weak force. They and their interactions are all described by gauge theories with the mediating particles represented by quantum fields. Furthermore, the electromagnetic and the weak force can be unified to one single theory, the electroweak theory. Within this mathematical framework of the SM, Lagrangians are used for the description and have to be invariant under local gauge transformations. This leads to gauge symmetries, which induce conserved quantities as it is specified in Noether's Theorem. So the composite gauge group  $SU(3)_C \times SU(2)_L \times U(1)_Y$  relates to the three forces and induces the electric charge, weak isospin and colour charge as conserved quantities. But the SM is not a complete theory of fundamental interactions because it does not implicate gravity, the fourth known fundamental force.

The gauge bosons and some of their properties are listed in table 1.2. The gluons, mediating the strong interaction between colour charged particles, are massless and carry colour charges. As there are eight independent colour states, eight types of gluons exist. Due to their own colour charge they can interact among each other.

name	force	symbol	el. charge [ $e$ ]	mass [ $\text{MeV}/c^2$ ]
gluon	strong	$g$	0	0
photon	electromagnetic	$\gamma$	0	0
W boson	weak	$W^\pm$	$\pm 1$	$80.403 \pm 0.029$
Z boson	weak	$Z^0$	0	$91.188 \pm 0.002$

Table 1.2: Properties of the gauge bosons (spin-1 particles) [2]. The electric charge is given in units of the electron charge, the mass in units of  $\text{GeV}/c^2$ .

The gauge boson of the electromagnetic force is the massless and neutral photon. It only interacts with electrically charged particles, so it does not interact with itself. In contrast to these gauge bosons, the ones of the weak force, the  $W$  and the  $Z$  bosons, are massive. The  $W$  bosons additionally carry an electric charge.

This noticeable difference between the gauge bosons, that there are not only massless but also massive ones, is not predicted by the gauge theory itself. This concerns especially the theory of electroweak unification as there are massless gauge bosons, the photons, as well as massive ones, the  $W$  and  $Z$  bosons, in one single theory. Therefore, the last particle of the SM, the Higgs boson, is essential for giving masses to the massive gauge bosons and also to the other massive particles. For this purpose a scalar field, the Higgs field, has to be introduced with the massive Higgs boson as quantum of this field. A particle then gets its mass through interactions with this field. According to the intensity of the interaction the mass is determined. This process is explained by the Higgs mechanism and will be described in more detail in the following sections.

## 1.2 Local Gauge Invariance in Quantum Field Theory

The SM is a gauge theory, so the mathematical framework is the quantum field theory with Lagrangians, that are invariant under certain transformations. QED, for example, is the gauge theory with the electromagnetic field as gauge field, the photon as massless gauge boson and the symmetry group  $U(1)$ . The Dirac Lagrange density function of a fermion field  $\psi$  is

$$\mathcal{L} = i\hbar c \bar{\psi} \gamma^\mu \partial_\mu \psi - mc^2 \bar{\psi} \psi \quad (1.1)$$

with the Dirac spinor  $\psi$  and the Dirac matrices  $\gamma_\mu$ .  $\mathcal{L}$  is invariant under global gauge transformations, but the invariance under a local gauge transformation

$$\psi \rightarrow e^{i\alpha(x)} \psi \quad (1.2)$$

is required. This can be achieved by replacing the derivative  $\partial_\mu$  by the covariant derivative  $\mathcal{D}_\mu$ . Therefore, a vector field  $A_\mu$  has to be introduced. After the variable transformation  $\lambda(x) = -\frac{\hbar c}{q} \alpha(x)$ , where  $q$  is the charge of the involved particle, the

new field  $A_\mu$  transforms to

$$A_\mu \rightarrow A_\mu + \partial_\mu \lambda \quad (1.3)$$

and the covariant derivative is defined by

$$\mathcal{D}_\mu \equiv \partial_\mu + i \frac{q}{\hbar c} A_\mu. \quad (1.4)$$

Due to the addition of the gauge field  $A_\mu$  to preserve local gauge invariance of  $\mathcal{L}$ , also the "free" term for the field

$$\mathcal{L}_A = -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{8\pi} \frac{m_A c^2}{\hbar} A^\nu A_\nu \quad (1.5)$$

has to be taken into account. The first part of equation 1.5, including the field strength tensor  $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$  is local gauge invariant in contrast to the second one, the mass term of the field. Therefore, the gauge field has to be massless. The vector field  $A_\mu$  can be identified as the photon field and equation 1.5 corresponds to its kinetic energy. This leads to the complete Lagrange density function of QED

$$\mathcal{L} = i\hbar c \bar{\psi} \gamma^\mu \mathcal{D}_\mu \psi - mc^2 \bar{\psi} \psi - \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} \quad (1.6)$$

with the photon as massless gauge particle. In summary, it describes the interaction of the fermion field with the photon field.

As all fundamental interactions in the SM can be described by gauge theories, the principle of local gauge invariance can be applied to the other gauge groups:  $SU(3)$  for the strong interaction, with massless gluons as gauge bosons and  $SU(2)$  for the weak interaction, where three fields have to be added to achieve local gauge invariance. In the electroweak theory developed by Glashow, Weinberg and Salam the gauge group is the  $SU(2)_L \times U(1)_Y$ . In weak interactions the quantum number is the weak isospin  $(\vec{T}^2, T_3)$  with  $\vec{T} = \frac{1}{2} \vec{\tau}$  the Pauli matrices  $\vec{\tau}$ . Relative to it all fermions are grouped into left-handed doublets and right-handed singlets, while only the left-handed particles take part in weak interactions. This is indicated by  $SU(2)_L$ . The hypercharge  $Y$ , given by  $Q = T_3 + \frac{Y}{2}$ , is the quantum number of  $U(1)_Y$ . For local gauge invariance four gauge fields have to be added, the field  $B_\mu$  which corresponds to the generator  $Y$  of the  $U(1)_Y$  group and the three fields  $\vec{W}_\mu$  corresponding to the generators  $\vec{T}$  of the  $SU(2)_L$  group. Linear combinations of those fields build the charged W bosons, the Z boson ( $Z_\mu$ ) and the photon ( $A_\mu$ ):

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2) \quad (1.7)$$

$$Z_\mu = W_\mu^3 \cos(\theta_W) - B_\mu \sin(\theta_W) \quad (1.8)$$

$$A_\mu = W_\mu^3 \sin(\theta_W) + B_\mu \cos(\theta_W) \quad (1.9)$$

with the Weinberg angle  $\theta_W$ .

Nevertheless, only massless gauge bosons like the photon are predicted by the theory. To solve this problem the concept of spontaneous symmetry breaking of the  $SU(2)_L \times U(1)_Y$  symmetry is used to give masses to the W and Z bosons.

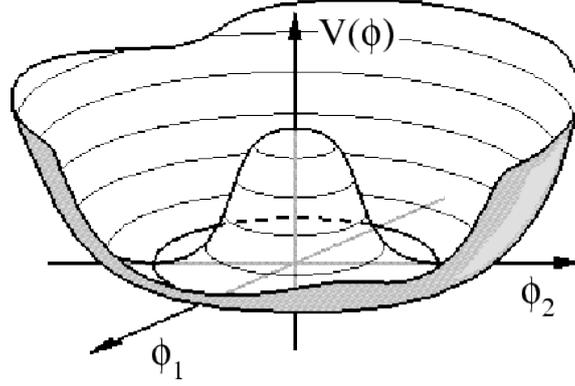


Figure 1.1: Illustration of the Higgs potential. The minima of the potential are lying on a circle.

### 1.3 Spontaneous Symmetry Breaking

Having a Lagrangian  $\mathcal{L}$  of a system that is symmetric with respect to a symmetry group, the symmetry is spontaneously broken if the vacuum state is not invariant under a corresponding transformation. Then according to Goldstone's theorem, for each generator of the broken symmetry new massless scalar particles, the Goldstone bosons, appear. For example the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi_1)(\partial^\mu\phi_1) + \frac{1}{2}(\partial_\mu\phi_2)(\partial^\mu\phi_2) + \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2 \quad (1.10)$$

with the potential

$$\mathcal{U}(\phi) = -\frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2 \quad (1.11)$$

shown in figure 1.1, is invariant under rotations in  $(\phi_1, \phi_2)$ -space. The minima of equation 1.11 lie on a circle with radius  $\mu/\lambda$ . By choosing one particular ground state, the original symmetry is broken. To parametrise the fluctuations about the chosen ground state, two new fields are introduced:

$$\eta \equiv \phi_1 - \frac{\mu}{\lambda}, \quad \xi \equiv \phi_2. \quad (1.12)$$

The Lagrangian 1.10 expressed by these new field variables changes to

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2 + \frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi) + \mu\lambda(\eta^3 + \mu\xi^2) - \frac{\lambda^2}{4}(\eta^4 + \xi^4 + 2\mu^2\xi^2) + \frac{\mu^2}{4\lambda^2}. \quad (1.13)$$

The field  $\eta$  now carries the mass  $m_\eta = \sqrt{2}\mu\frac{\hbar}{c}$  while the field  $\xi$  stays massless. The other terms represent the different couplings between the two fields and the final term is just a constant. Summarised, the mechanism of spontaneous symmetry breaking gives mass to the field  $\eta$ , but also produces one massless scalar particle, the Goldstone boson, corresponding to the field  $\xi$ , which cannot be identified with any known elementary particle. For this purpose the Higgs mechanism is presented in the following section.

## 1.4 The Higgs Mechanism

The Higgs mechanism can be considered as union of gauge invariance and spontaneous symmetry breaking. As masses for the  $W$  and  $Z$  bosons are required in the electroweak theory, the Higgs mechanism will be shown for the group  $SU(2)_L \times U(1)_Y$ . To break this symmetry group, the Higgs field  $\phi$  has to be at least an  $SU(2)$  doublet with four real components :

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (1.14)$$

The important part of the Lagrange density function is the one of the interaction of the Higgs field with the four gauge fields

$$\left| \left( i\partial_\mu - g\vec{T} \cdot \vec{W}_\mu - g'\frac{Y}{2}B_\mu \right) \phi \right|^2 - \mathcal{U}(\phi) \quad (1.15)$$

with the couplings  $g$  and  $g'$  according to the gauge fields and the Higgs potential

$$\mathcal{U}(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (1.16)$$

with  $\mu^2 < 0$  and  $\lambda > 0$ , so that the mechanism of spontaneous symmetry breaking can be used. Therefore, one certain minimum of  $\mathcal{U}(\phi)$  has to be chosen such as

$$\phi_1 = \phi_2 = \phi_3 = 0 \quad \phi_4^2 = -\frac{\mu^2}{\lambda} \equiv v^2 \quad (1.17)$$

with  $v^2$  as the vacuum expectation value. The expected value of the field at this ground state can be taken as

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.18)$$

As before, the fluctuations around the ground state are parametrised by the use of new fields, here four real fields  $\theta_1, \theta_2, \theta_3$  and  $h$  are introduced, so that

$$\phi(x) = \begin{pmatrix} \theta_2 + i\theta_1 \\ \frac{1}{\sqrt{2}}(v + h(x)) - i\theta_3 \end{pmatrix} = e^{i\theta_a(x)\tau^a(x)/v} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + h(x)) \end{pmatrix}. \quad (1.19)$$

The factor  $e^{i\theta_a(x)\tau^a(x)}$ , including the three fields  $\theta_a$  which correspond to the massless Goldstone bosons, just turns out to be a local  $SU(2)$  transformation. So the Goldstone bosons can be "gauged away" and only the Higgs field  $h(x)$  remains:

$$\phi(x) \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ (v + h(x)) \end{pmatrix}. \quad (1.20)$$

This new field is used to expand the Higgs interaction term of the Lagrangian, given by equation 1.15, what leads to

$$\begin{aligned} & \left| \left( -ig\frac{\tau_a}{2}W_\mu^a - i\frac{g'}{2}B_\mu \right) \phi \right|^2 \\ &= \left( \frac{1}{2}vg \right)^2 W_\mu^+ W_\mu^- + \frac{1}{8}v^2 (W_\mu^3, B_\mu) \begin{pmatrix} g^2 & -gg' \\ gg' & g'^2 \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \end{aligned} \quad (1.21)$$

with the fields  $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$  as introduced before. So with the Higgs mechanism, the gauge bosons of the weak interaction become massive and equation 1.21 contains the corresponding mass terms. The mass of the  $W$  bosons can be specified to

$$M_W = \frac{1}{2}vg. \quad (1.22)$$

To determine the masses of the other bosons, the matrix in equation 1.21 has to be diagonalised to identify the mass eigenstates. The fields  $A_\mu$  for the photon and  $Z_\mu$  for the uncharged  $Z$  boson are introduced with

$$A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}} \quad \text{with} \quad M_A = 0 \quad (1.23)$$

$$Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \quad \text{with} \quad M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \quad (1.24)$$

or by using the weak mixing or Weinberg angle  $\theta_W$  with  $\tan(\theta_W) = \frac{g'}{g}$  the transformation becomes as introduced in equation 1.8 and 1.9.

But not only the  $W$  and  $Z$  bosons gained their masses, also the Higgs boson, one new massive scalar particle, appeared with the mass  $M_H = \sqrt{2}\lambda v$ . Unfortunately, it is not possible to predict its mass, because the free parameter  $\lambda$  is involved. So one of the main aims of the new particle collider Large Hadron Collider (LHC) at CERN is the search for this last fundamental particle of the SM, as it has not been observed yet.

## 1.5 Yukawa Sector

The fermions gain their masses through interactions with the same scalar field  $\phi(x)$  as it has been introduced for the weak gauge bosons. After the weak theory, all left-handed states of the fermions can be grouped into  $SU(2)_L$  doublets,  $L$  and  $Q$  for the leptons and quarks, respectively, while the right-handed states remain in  $SU(2)_L$  singlets. For any fermion generation, the Yukawa Lagrangian, which is invariant under local  $SU(2)_L \times U(1)_Y$  transformation, is introduced with

$$\mathcal{L}_F = -\lambda_e \bar{L}\phi e_R - \lambda_d \bar{Q}\phi d_R - \lambda_u \bar{Q}\phi u_R + h.c. \quad (1.25)$$

and the same procedure is repeated as above. In the case of an electron, for example, the Lagrangian  $\mathcal{L}_F$  becomes

$$\mathcal{L}_F = -\frac{1}{\sqrt{2}}\lambda_e(\bar{\nu}_e, \bar{e}_L) \begin{pmatrix} 0 \\ v+h \end{pmatrix} e_R + \dots = -\frac{1}{\sqrt{2}}\lambda_e(v+h)\bar{e}_L e_R + \dots \quad (1.26)$$

The fermion mass corresponds to the constant term in front of  $\bar{f}_L f_R$ :

$$m_e = \frac{\lambda_e v}{\sqrt{2}}, \quad m_u = \frac{\lambda_u v}{\sqrt{2}}, \quad m_d = \frac{\lambda_d v}{\sqrt{2}}. \quad (1.27)$$

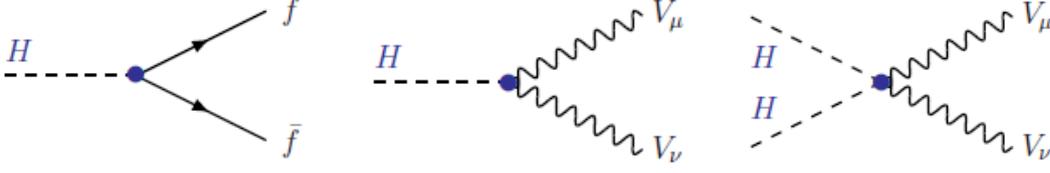


Figure 1.2: Higgs boson couplings to fermions  $f$  and massive weak gauge bosons  $V$ .

Hence, through the introduction of the Higgs field  $\phi(x)$ , it is possible to generate the masses of the weak gauge bosons  $W$  and  $Z$ , below labelled with  $V$ , as well as the masses of the fermions  $f$ , while the  $SU(2) \times U(1)$  gauge symmetry is still preserved. The corresponding Higgs boson couplings can be identified with

$$g_{Hff} = i\frac{m_f}{v}, \quad g_{HVV} = -2i\frac{M_V^2}{v} \quad \text{and} \quad g_{HHVV} = -2i\frac{M_V^2}{v^2} \quad (1.28)$$

represented in figure 1.2.

## 1.6 Higgs Boson Production and Decay at the LHC

The cross section  $\sigma$  for a process depends on the collider and its centre of mass energy. The expected number of events  $N$  during a certain period of time is given by

$$N = \mathcal{L}_{int} \cdot \sigma \quad (1.29)$$

with  $\mathcal{L}_{int}$  the luminosity integrated over this time interval. The luminosity is characteristic for a collider and depends only on the beam parameters. To achieve a high luminosity, high beam intensities are needed.

The LHC is a proton-proton collider and as protons are compound objects, the fundamental interaction occurs between the constituents of the proton, the partons. There are three valence quarks ( $uud$ ) bound by gluons which again can split into quark-antiquark pairs, the sea quarks. Each of them carries a certain momentum fraction  $x_i = \frac{p_i}{p_p}$  of the proton momentum  $p_p$ . Therefore, the parton distribution function (PDF)  $f_{i,p}(x_i, \mu^2)$  is needed to describe this momentum distribution. So the parton  $i$  carries the momentum fraction  $x_i$  with the given probability by the PDF. Figure 1.3 shows the PDF for a factorisation scale of  $\mu^2 = (175 \text{ GeV})^2$ , which describes the energy scale of the process.

The Higgs boson prefers to couple to heavy particles like the  $W$  and  $Z$  bosons or the top and bottom quarks, but to the latter more rarely. This leads to the four main production modes for Higgs bosons at the LHC: the gluon-gluon fusion and the weak vector boson fusion (VBF) process, the associated production with  $W$  or  $Z$  bosons and also with heavy top or bottom quarks. The corresponding Feynman

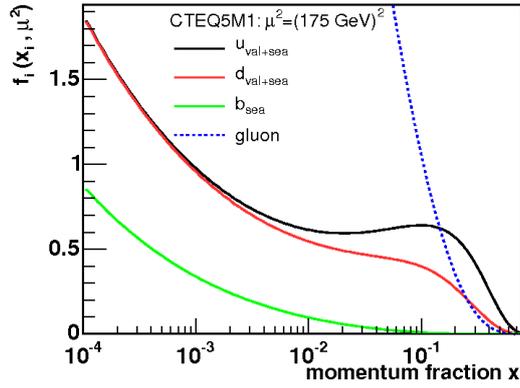


Figure 1.3: The CTEQ5M1 parton distribution function at  $\mu^2 = (175 \text{ GeV})^2$  [8].

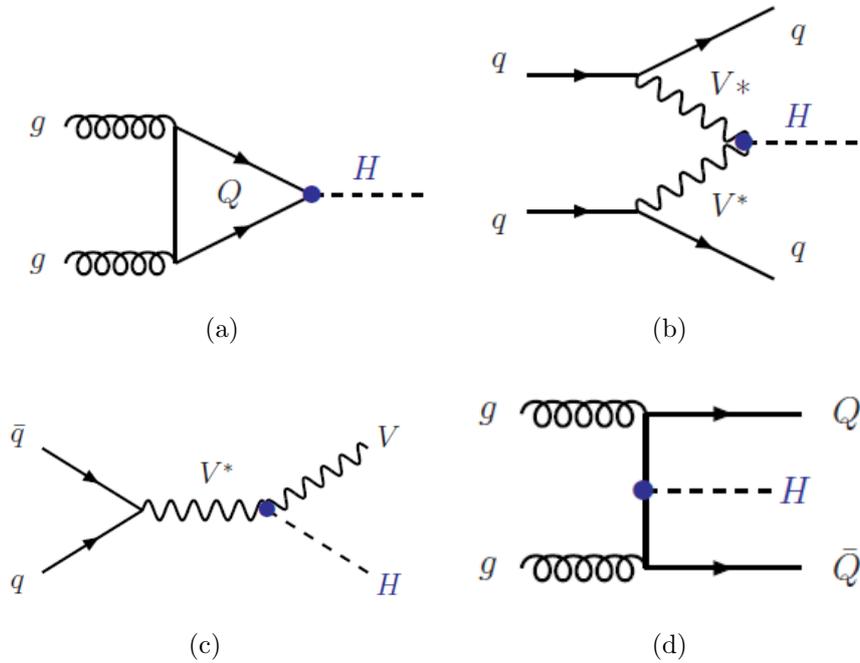


Figure 1.4: Dominant Higgs production mechanisms at leading order: gluon-gluon fusion with a heavy quark loop (a), vector boson fusion (b), associated production with W or Z bosons (c) or with heavy quarks (d).

diagrams are shown in figure 1.4 and the cross sections at the LHC in dependence of the Higgs boson mass in figure 1.5.

The gluon-gluon fusion production mode has the largest cross section over the whole Higgs mass range. Gluons couple mainly via a heavy-quark loop to the Higgs boson. Anyhow, VBF is also an important production process even though its cross section is smaller by a factor of 10. The event signature with two forward directed jets is very characteristic for these events and makes it possible to reduce many background processes. As the LHC is a proton-proton collider, antiquarks are only

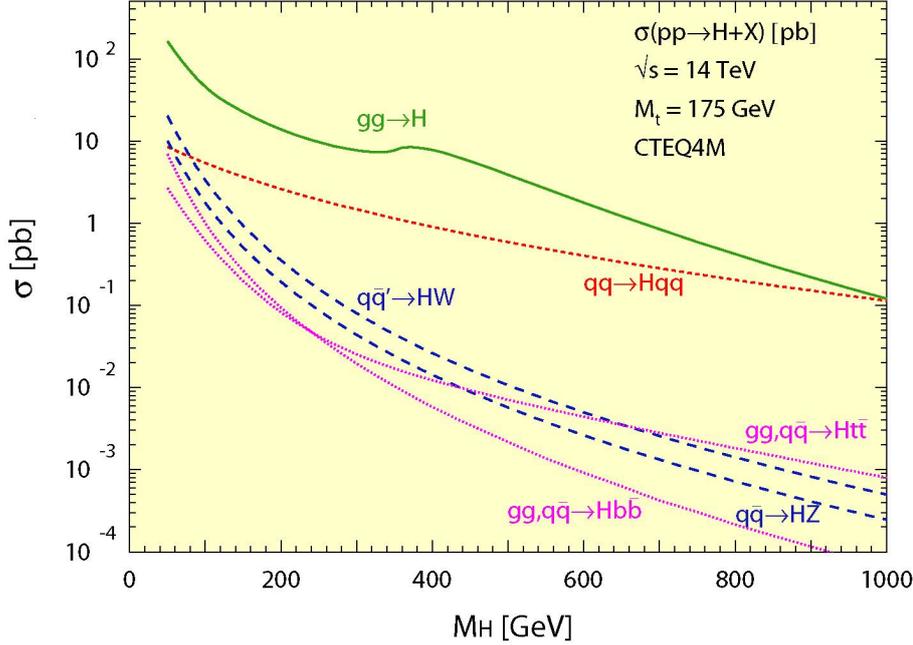


Figure 1.5: Cross sections of the Higgs production in dependence of the Higgs mass at the LHC [9].

present as sea quarks. Therefore, the associated production mechanism with  $W$  or  $Z$  bosons has only low cross sections.

For the Higgs boson decay, heavy particles are again preferred. The decay into massless particles is mediated by a massive particle loop. So the decay into photons is preferably mediated by top quarks or  $W$  bosons, as the particles have to be charged to couple to the photons, while the decay into gluons can only be mediated by coloured particles, what leads to the top quarks as preferred ones because of their high mass.

The branching ratios of the main decay channels are shown in figure 1.6. In the low mass range with Higgs boson masses  $m_H \leq 130 \text{ GeV}/c^2$  the dominant decay of the Higgs boson is the one into a  $b\bar{b}$  pair. The decay into heavier particles like  $t\bar{t}$  is highly suppressed for  $m_H$  below twice the corresponding particle mass. But in the mass range with  $m_H \geq 130 \text{ GeV}/c^2$  the decay is dominated by the mode into weak vector boson pairs. If  $m_H < 2m_{W,Z}$  one of the bosons is produced off-shell. In the interval  $150 \text{ GeV} < m_H < 190 \text{ GeV}/c^2$  the branching ratio to  $Z$  bosons has an unusual deformation. It comes from a threshold effect, because the decay to two  $W$  bosons on the mass shell gets possible, while the decay to two  $Z$  boson is still connected with one off-shell produced boson. At higher Higgs boson masses both bosons are produced on the mass shell. Then the branching ratio to  $W$  boson pairs is circa twice the one to  $Z$  bosons, as there are two possible combinations of  $WW$  production,  $W^+W^-$  or  $W^-W^+$ , in contrast to the neutral  $Z$  boson.

In this analysis the process  $qqH \rightarrow qqWW \rightarrow qql\nu l\nu$  is discussed. The Higgs boson is produced via the VBF mechanism and decays into a  $W$  boson pair. Each of them is assumed to decay into a charged lepton and a neutrino. The Higgs mass

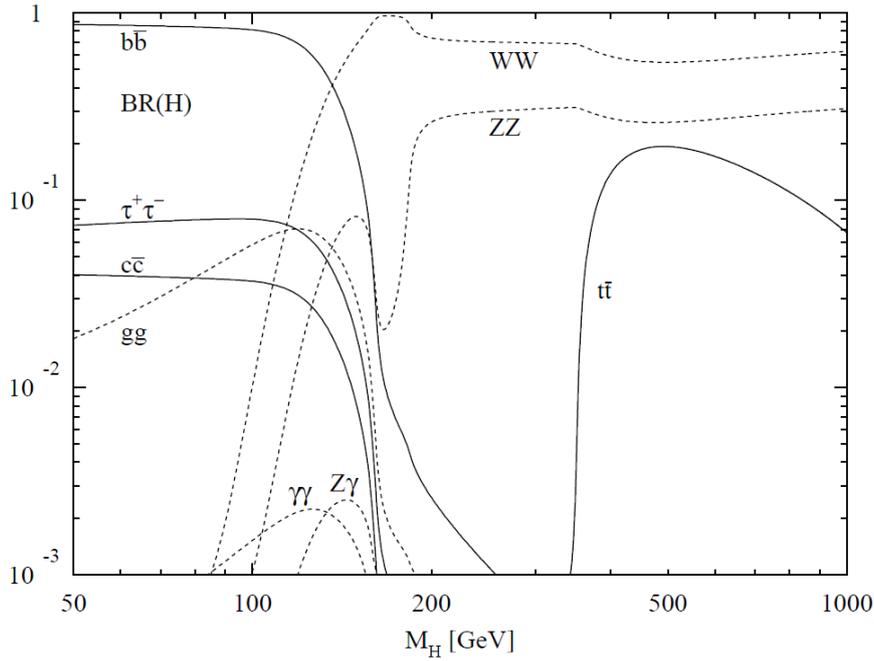


Figure 1.6: Branching ratios of the main decay channels of the Higgs boson [9].

is supposed to be  $130 \text{ GeV}/c^2$ . The cross section multiplied with the corresponding branching ratios for this process is  $0.12 \text{ pb}$ . A detailed list is given in reference [11]. In figure 1.7 the branching ratios in this mass region are shown.

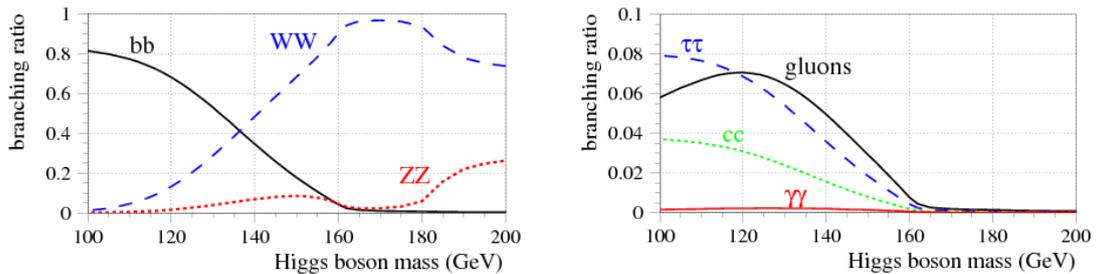


Figure 1.7: Branching ratios of the main decay channels of the Higgs boson in the mass range  $100 \text{ GeV}/c^2 \leq m_{Higgs} \leq 200 \text{ GeV}/c^2$  [10].

## 1.7 Bounds for the Higgs Boson Mass

The SM does not predict the Higgs boson mass, but it is possible to set bounds for it. First direct searches have been conducted at the Large Electron Positron (LEP) collider without any observations of a Higgs signal. The four LEP experiments have

been combined and the Higgs boson mass is constrained to

$$m_H > 114.4 \text{ GeV}/c^2 \quad \text{at 95\% confidence level (C.L.).} \quad (1.30)$$

More precise information is given in reference [12] and [13].

The Higgs boson also makes a contribution by loop corrections to the propagators and so to the masses of the weak gauge bosons. High precision electroweak measurements are sensitive to the Higgs mass. The LEP Electroweak Working Group has made the fit shown in figure 1.8 with the resulting upper limit

$$m_H < 154 \text{ GeV}/c^2 \quad \text{at 95\% C.L.} \quad (1.31)$$

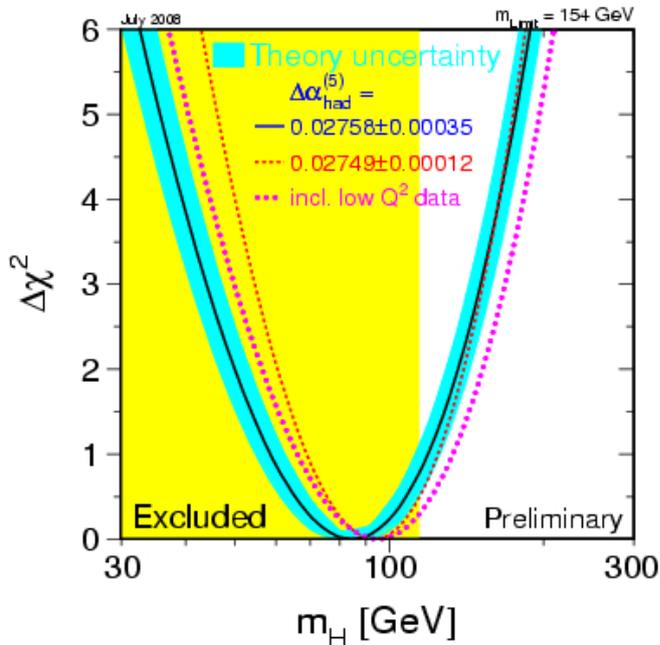


Figure 1.8: Fit of the Higgs boson mass derived from high precision electroweak measurements. Theoretical uncertainties are not taken into account and are assigned by the blue band. The direct search limit is shown in yellow [13].

A combination of the results of the experiments CDF and DØ at Tevatron delivers upper limits on the production cross section which are shown in figure 1.9. The Higgs boson mass  $m_H = 170 \text{ GeV}/c^2$  is excluded at 95% C.L.. For more details see reference [14].

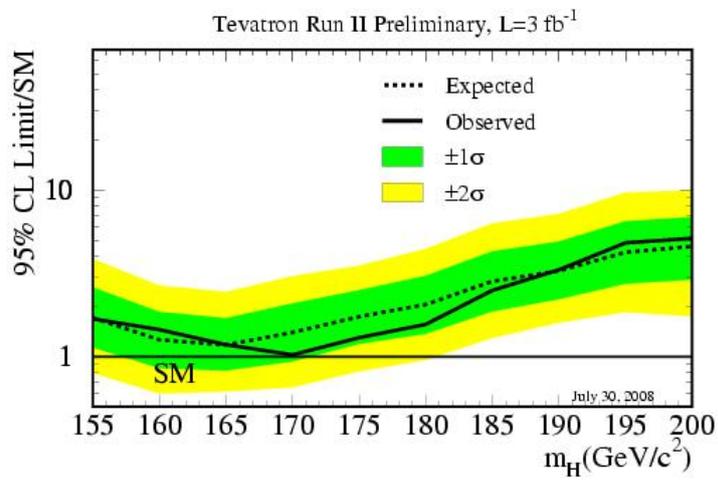


Figure 1.9: The ratios of the 95% C.L. expected and observed limits to the SM cross section, as functions of the Higgs boson mass for the combined CDF and  $D\bar{O}$  analyses. A value of the ratio which is less or equal to one indicates the exclusion at 95% C.L. of the particular Higgs boson mass [14].



# Chapter 2

## The Compact Muon Solenoid Experiment at the Large Hadron Collider

This chapter focuses on the Compact Muon Solenoid (CMS) experiment conducted at the Large Hadron Collider (LHC) which is situated near Geneva at the French-Swiss border. The LHC reuses the 27 km tunnel, originally built for the Large Electron Positron Collider (LEP) at the particle physics laboratory CERN. The following contains an introduction to the proton-proton collider LHC before proceeding to the main part, the description of the CMS experiment.

### 2.1 The Large Hadron Collider

The LHC has a design centre of mass energy of 14 TeV and a peak luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . The machine luminosity, already mentioned in section 1.6, depends on the beam parameters and such a high one can only be obtained by high beam intensities. So the bunch structure of the LHC beam plays a decisive role. In one proton beam there are 2808 bunches, each with  $1.15 \cdot 10^{11}$  protons and a bunch spacing of 25 ns. The advantage of the high luminosity is, that also rare events can be explored. With equation 1.29 it is understandable that only then a sufficient number of events can be obtained, as their cross sections are low.

With 14 TeV the LHC is the first collider exploring the TeV energy scale. Therefore, it is possible to verify the consistency of the Standard Model and maybe to reveal new physics beyond it. But of course, one of the most important aims is to explain the origin of mass and electroweak symmetry breaking, assumed to be caused by the Higgs mechanism.

Compared to the LEP, which used electrons instead of protons for collisions, there is a big reduction of the energy loss in  $\mathcal{O}(10^{13})$  through synchrotron radiation as there is a dependency on  $(E/m)^4$ . The mass of protons is about 2000 times bigger than the one of electrons.

Commissioning of the LHC with proton beams was first attempted in September 2008 and started with an injection energy of 0.45 TeV. But unfortunately collisions

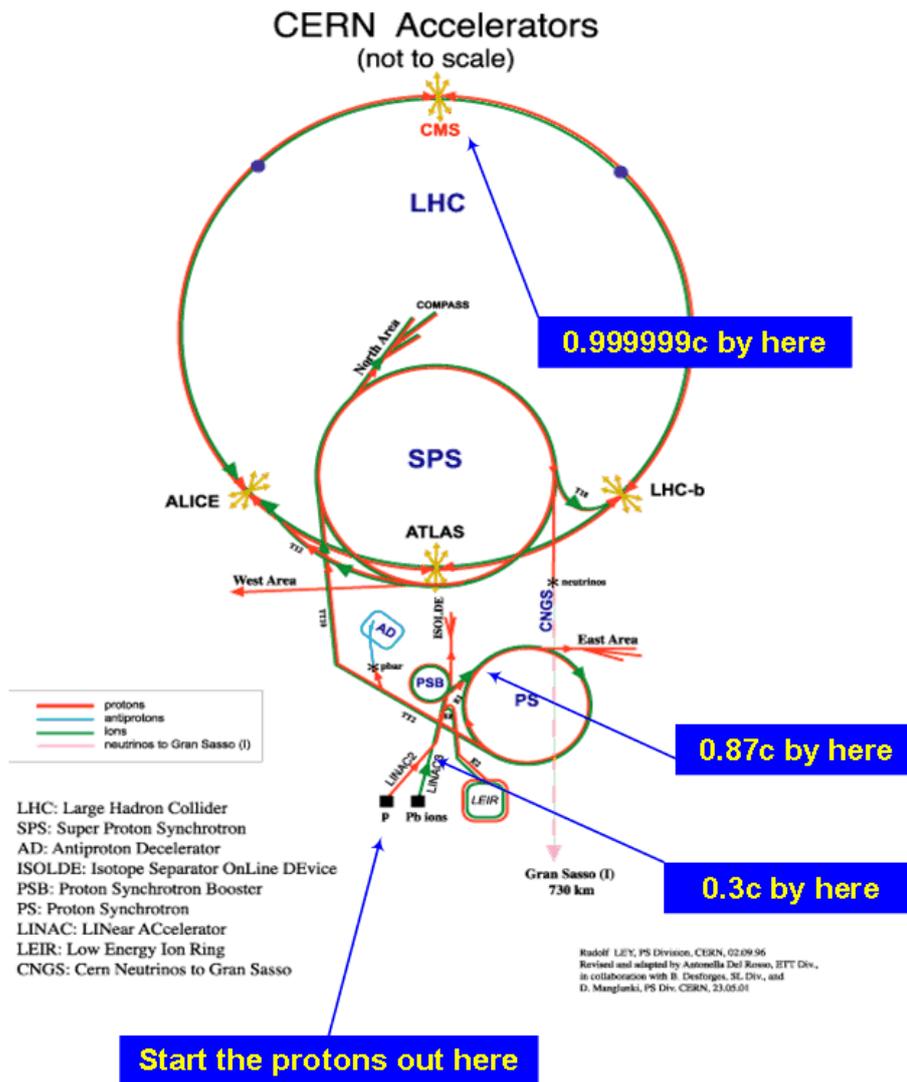


Figure 2.1: Schematic overview of the CERN accelerator complex, with proton velocities.

could not be performed because there was a malfunction caused by a faulty electrical connection between two of the accelerator's magnets. With the superconducting dipole magnets the protons are kept on their circular track which implies that the peak beam energy of 7 TeV needs a peak dipole field of 8.33 T. But the outcome of the incident is that 53 magnets have to be cleaned or replaced which will delay the restart of the LHC, planned for summer 2009. Reference [18] goes into details of the incident and of the foreseen schedule.

The whole accelerator complex is shown in figure 2.1. It includes the linear accalerator Linac2, where the beam is produced, followed by four ring accelerators. At the Linac2 duoplasmatron source hydrogen gas enters a cathode chamber with electrons. It dissociates, becomes a plasma and the protons leave the chamber in a canal through the anode, expand and form the proton beam. Then the beam is focused and formed to bunches by a Radio Frequency Quadrupole, which also

accelerates the protons to an energy of 750 keV and injects them to three Linac tanks, which increase the energy up to 50 MeV. The ejected beam from the Linac2 is characterised by a current of 180 mA and a pulse length of 30  $\mu$ s. For more details see reference [17].

An energy of 1.4 GeV is reached through the Proton Synchrotron Booster (PSB) which is composed of four PSB rings, each containing two proton bunches. But only six of them are injected to the Proton Synchrotron (PS). Here the energy increases to 25 GeV and the 25 ns bunch spacing is provided. First each of the six bunches are split into three, then boosted to 25 GeV and then split twice in two. So 72 bunches are recaptured and shortened to 4 ns to fit into the Super Proton Synchrotron (SPS) acceleration system. Through the ejection of only six instead of eight bunches from the PSB, the bunch train is composed of 72 bunches and, additionally, a gap of  $\sim$ 320 ns for the ejection kicker rise time of the PS.

Three or four PS pulses are injected into the SPS and fill  $\sim$ 27% (3/11) or  $\sim$ 36% (4/11) of its beamline. The LHC receives 12 SPS cycles, already accelerated to 450 GeV, to build the final proton beam filling 2808 bunches of 3564 possible bunch places. Two counter-rotating beams are needed whose energy is then boosted to 7 TeV. Assuming some minutes for machine settings and other checks, the total turnaround time amounts to 70 minutes. The luminosity then decays during a physics run principally due to the beam loss from collisions, what leads to a net estimate of the luminosity lifetime of  $\sim$ 15 hours. More detailed information about the LHC can be found in reference [16].

The operation with heavy ions e.g. lead instead of protons is also planned. Therefore the A Large Ion Collider Experiment (ALICE) has been constructed. It is one out of six experiments located at the LHC. For the proton collisions there are two multi-purpose detectors, CMS and A Toroidal LHC Apparatus (ATLAS), designed to operate at the peak luminosity of the LHC. The other three experiments are smaller and have been designed for low luminosity. The Large Hadron Collider beauty (LHCb) experiment is used for b-physics, the Large Hadron Collider forward (LHCf) for studies of particles generated in the forward region and the TOTal Elastic and diffractive cross section Measurement (TOTEM). As shown in figure 2.1 there are four collision points of the LHC beam with large underground caverns for the experiments ALICE, ATLAS, CMS and LHCb. LHCf and TOTEM are each arranged near ATLAS and CMS.

## 2.2 The Compact Muon Solenoid

The CMS detector has been assembled in a cavern 100 metres underground in France between Lake Geneva and the Jura mountains. As the name suggests, its main features are the compact size, compared to the other multi-purpose detector ATLAS, the specialisation for muon tracking and the huge superconducting solenoid magnet. The CMS detector is 21.6 m long and with a diameter of 14.6 m its total weight is 12500 t. An overview of the detector is given in figure 2.2.

To meet the demands of the LHC, its motivation, physics at the TeV energy scale, predefines the requirements for CMS. The high design luminosity implies huge event

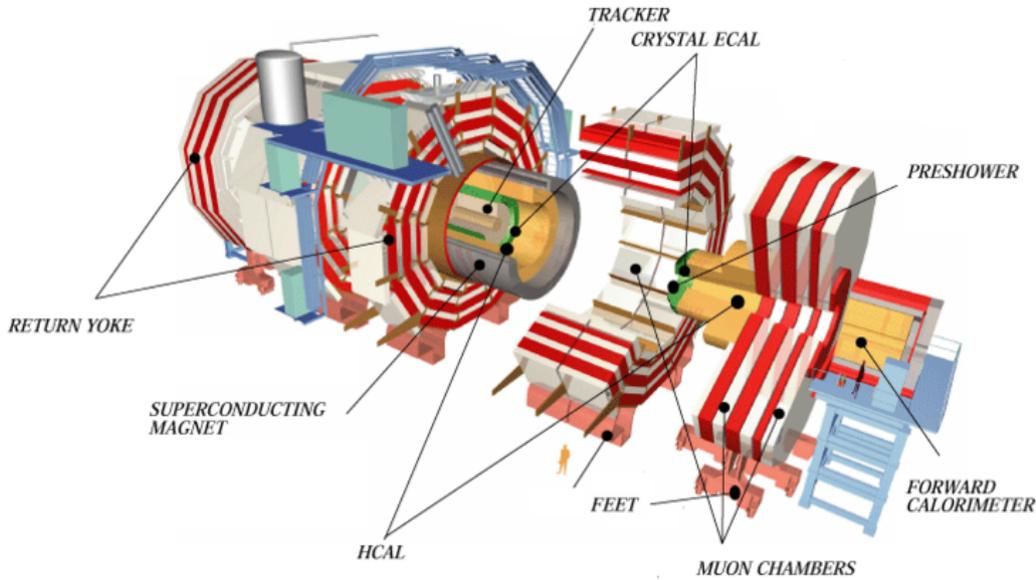


Figure 2.2: A perspective view of the CMS detector [21].

rates and the short bunch spacing of only 25 ns reinforces the need of a detector with a short response time. This leads to a radiation-hard, high-granularity detector with a good time resolution and the focus on the following main detector requirements:

- a good muon identification and momentum resolution over a wide geometric area;
- for the inner tracker: pixel detectors close to the interaction region to achieve a good momentum resolution and reconstruction efficiency of charged particles;
- for the electromagnetic calorimeter: an excellent diphoton and dielectron mass resolution,  $\pi^0$  rejection and photon and lepton isolation;
- for the hadron calorimeter: an almost hermetic coverage and a fine lateral segmentation to get a good missing transverse energy and dijet mass resolution;

The outcome of these needs is the first essential component of CMS, a strong magnet, to measure accurately the momentum of even high energetic, charged particles. The solenoid, described in reference [23], is a coil of superconducting niobium-titanium wire and achieves a 3.8 T magnetic field. With its overall length of 13 m and an inner diameter of 6 m, the bore of the coil is large enough to house the tracker and the calorimetry. Outside of the coil is the muon detector, which consists of four layers and is interleaved with three layers of a 12-sided iron structure, forming the return yoke of the magnet.

Reference [20] gives an overview of the CMS experiment, while reference [19] provides a more detailed description.

The coordinate system, used by CMS, has the origin centred at the nominal collision point. The  $x$ -axis is horizontal, pointing radially inward to the centre

of the LHC, the  $y$ -axis is vertical pointing upwards and the  $z$ -axis is horizontal pointing west along the beam direction. The azimuthal angle  $\phi$  is measured in the  $x$ - $y$  plane from the  $x$ -axis and the polar angle  $\theta$  in the  $y$ - $z$  plane with respect to the  $z$ -axis. Pseudorapidity is defined as  $\eta = -\ln(\tan(\theta/2))$ . The momentum and energy transverse to the beam direction,  $p_T$  and  $E_T$ , respectively, are computed from the  $x$  and  $y$  components.

### 2.2.1 Tracking System

The primary purpose of the tracker is the accurate momentum measurement of charged particles. Hence, a precise reconstruction of the particle's trajectory, as they spiral in the magnetic field, is needed. For identifying heavy flavours, the second goal is a good secondary vertex resolution.

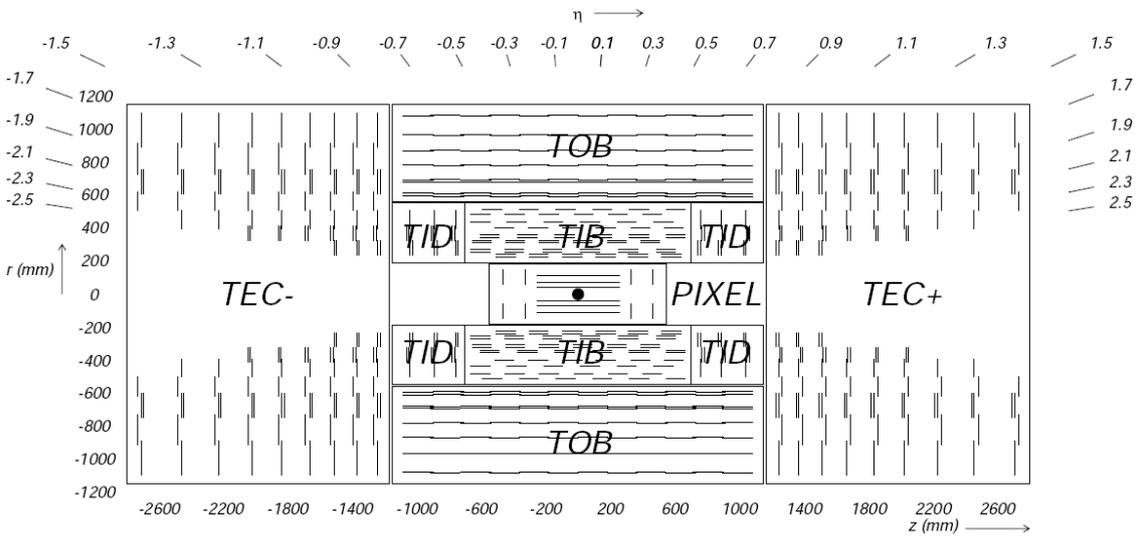


Figure 2.3: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits [19].

The CMS tracker, detailed in reference [27], is shown in figure 2.3. It is the innermost layer of the detector, so it is located close to the interaction region. Therefore it has to be radiation-hard, as it deals with the highest particle density. Its dimensions are 5.8 m in length and 2.6 m in diameter and its acceptance extends up to  $|\eta| < 2.5$ . It is made completely of silicon, that covers an area over  $200 \text{ m}^2$ . The inner component of the tracker is a pixel detector with 66 million silicon pixels on a  $1 \text{ m}^2$  surface, arranged on 1440 pixel detector modules. These are surrounded by 9.3 million silicon detector strips.

#### Silicon Pixel tracker

The pixel tracker is especially used for secondary vertex reconstruction. With a pixel cell size of  $100 \times 150 \mu\text{m}^2$  an almost similar position resolution in the  $r$ - $\phi$  and

$z$  direction with 10 and 20  $\mu\text{m}$ , respectively, can be obtained. The pixel detector is composed of three cylindrical layers with a length of 53 cm at radii of 4.4, 7.3 and 10.2 cm surrounding the interaction point. In addition there are two endcap disks on each side at  $|z|=34.5$  and  $|z|=46.5$  cm. Using this arrangement, the pixel detector delivers three high precision space points for every charged particle over almost the full  $\eta$ -range.

### Silicon Strip Tracker

The silicon strip tracker assures at least 9 measurements, 4 of them are two-dimensional, in the range of  $|\eta| < 2.4$ . This is achieved by the following tracker layout.

First, there is the Tracker Inner Barrel (TIB) with four barrel layers in the radial region between 20 and 55 cm and  $|z| < 65$  cm. In the first two layers the single point resolution is 23  $\mu\text{m}$ , while in the other two it goes up to 35  $\mu\text{m}$ . The TIB is completed with two inner endcaps, the Tracker Inner Disks (TID), composed of three small disks with strips radial to the beam axis. This subsystem measures up to 4  $r$ - $\phi$  points on a trajectory.

Second, there is the Tracker Outer Barrel (TOB), reaching an outer radius of 116 cm and covering  $|z| < 118$  cm. It is made up of six barrel layers, which leads to another 6  $r$ - $\phi$  measurements with a single point resolution of 53  $\mu\text{m}$  on the first four layers and 35  $\mu\text{m}$  on the other two.

At last, the Tracker EndCaps (TEC), one on each side, close off the tracker. Each consists of nine disks, extending from  $124 \text{ cm} < |z| < 282 \text{ cm}$  and  $22.5 \text{ cm} < |r| < 113.5 \text{ cm}$  and carrying up to seven rings of silicon microstrip detectors. The TEC delivers up to 9  $\phi$  measurements per trajectory.

The silicon strip tracker comprises 15 148 highly sensitive strip detector modules, each with a different design, depending on the position within the detector. Each module is composed of sensors, its mechanical support structure and the readout electronics.

### 2.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) described in reference [26], surrounds the tracking system and is supposed to make a precision energy measurement of electromagnetic interacting particles like electrons and photons. This is achieved by a hermetic homogeneous calorimeter with high density crystals what leads to a fast, radiation resistant calorimeter with fine granularity.

It is composed of a cylindrical barrel section, consisting of 61200 lead tungstate ( $\text{PbWO}_4$ ) crystals, and two endcaps with 7324 crystals in each of them. Lead tungstate has a density of  $8.28 \text{ g/cm}^3$ , a short radiation length of 0.89 cm and a small Molière radius of 2.2 cm. When particles pass through the calorimeter, the crystals produce blue-green scintillation light in fast, short, well-defined photon showers. 80% of the light is emitted in 25 ns which copes with the LHC bunch crossing time. The amount of light is proportional to the energy deposit of the particle in the crystal, but the yield is relatively low so that highly sensitive photodetectors are necessary to measure and convert it to an electrical signal.

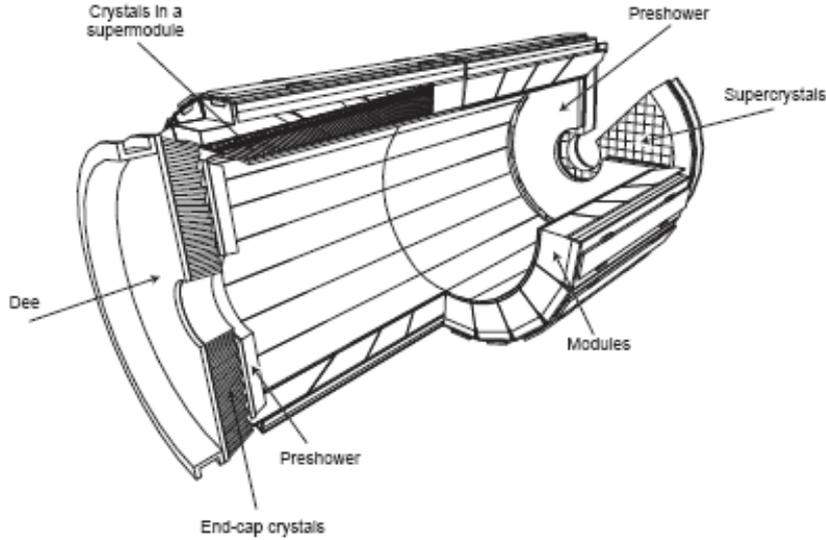


Figure 2.4: Layout of the ECAL with the arrangement of crystal modules, supermodules and endcaps [19].

The crystals are grouped into different modules, depending on the position in  $\eta$ . Four modules build a supermodule, which is composed of 1700 crystals. The arrangement in the ECAL is shown in figure 2.4.

The barrel section is made up of 36 supermodules, weighs 67.4 t and covers  $|\eta| < 1.479$ . The crystals have a front face cross section of  $22 \times 22 \text{ mm}^2$  and measure 230 mm in length. The endcaps, each divided into two halves, consist of supercrystals, units of  $5 \times 5$  crystals. They each weigh 24.0 t and cover the range of  $1.479 < |\eta| < 3.0$ . Here the crystals cross section is  $28.62 \times 28.62 \text{ mm}^2$  and their length amounts to 220 mm. For a better spatial precision a preshower detector sits in front of the endcaps. The main aim is the identification of neutral pions, decaying into two closely-spaced low energy photons. As the preshower has a much finer granularity, it is possible to distinguish between those two low energy photons and single high energy ones. The preshower ranges from  $1.656 < |\eta|$  to 2.6 and has a total thickness of 20 cm. It is composed of two detector layers, each with lead radiators followed by silicon strip sensors.

The expected energy resolution of the ECAL barrel and the endcaps is given in table 2.1.

Contribution	Barrel ( $\eta = 0$ )	Endcap ( $\eta = 2$ )
stochastic term	$2.7\%/\sqrt{E}$	$5.7\%/\sqrt{E}$
constant term	0.55%	0.55%
noise (low luminosity)	155 MeV	770 MeV ( $E_T = 205 \text{ MeV}$ )
noise (high luminosity)	210 MeV	915 MeV ( $E_T = 245 \text{ MeV}$ )

Table 2.1: Energy resolution of the ECAL barrel and endcap [26].

### 2.2.3 Hadron Calorimeter

The hadron calorimeter (HCAL) measures the energy of hadron jets and their position as well as the resulting missing transverse energy. For more details see reference [24]. It uses non-magnetic brass as absorber material with a short interaction length of 1.49 cm, assembled with 70000 plastic scintillator tiles. The energy resolution is estimated to be constant at 5%. The HCAL is composed of four subdetectors

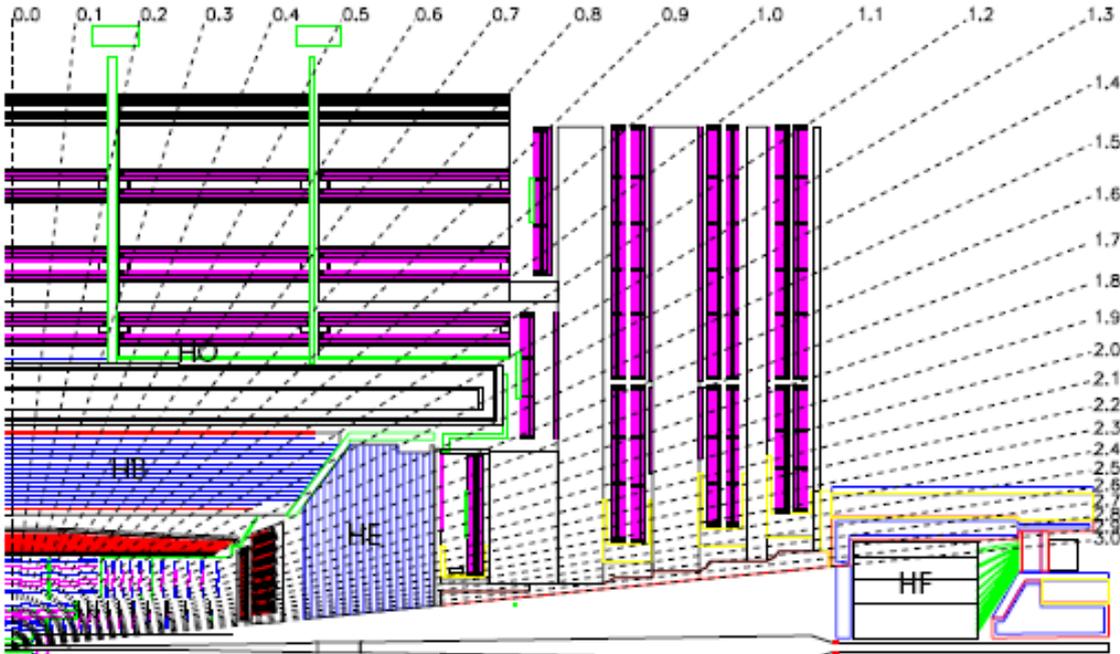


Figure 2.5: Longitudinal view of a quarter of the CMS detector, showing the location of the hadron calorimeter [19].

(figure 2.5):

- The hadron calorimeter barrel (HB) is the last layer inside the magnetic coil and ranges radially from  $1.77 \text{ m} < r < 2.95 \text{ m}$  and in pseudorapidity up to  $|\eta| < 1.3$ . There are 36 identical wedges, aligned parallel to the beam axis, forming the barrel. Each wedge weighs 26 t. The plastic scintillator is divided into 16  $\eta$  sectors.
- As the stopping power of the HB is not sufficient for hadron showers, the HB is completed by a tail catcher, the outer hadron calorimeter (HO). It has the same pseudorapidity range as the HB and is the first layer behind the iron yoke. So, the HO uses the solenoid coil as absorber and measures late starting showers and the energy deposit after having passed the HB.
- The hadron calorimeter endcaps (HE) sit behind the electromagnetic calorimeter endcaps. They cover  $1.3 < |\eta| < 3.0$  and have to be high radiation tolerant, as about 34% of the final state particles will pass this region. Their geometric

design is influenced by the requirement to minimize the cracks between HB and HE.

- The forward hadron calorimeters (HF), situated 11.2 m from the interaction point away, enlarge the pseudorapidity range to  $|\eta| = 5.2$ . As the main need in this region is a radiation hard active material, quartz fibres are used for the HFs. The signal is generated with a Cherenkov-based technology.

### 2.2.4 Muon System

One of the main features of the CMS experiment is the precise and robust muon measurement, including the identification, momentum measurement and triggering. Muons occur in many interesting processes and have a clear signature. As they are highly penetrating, the muon system [25] constitutes the most outer layer of the detector. The four muon stations are interleaved with the iron yoke plates. An overview is given on figure 2.6.

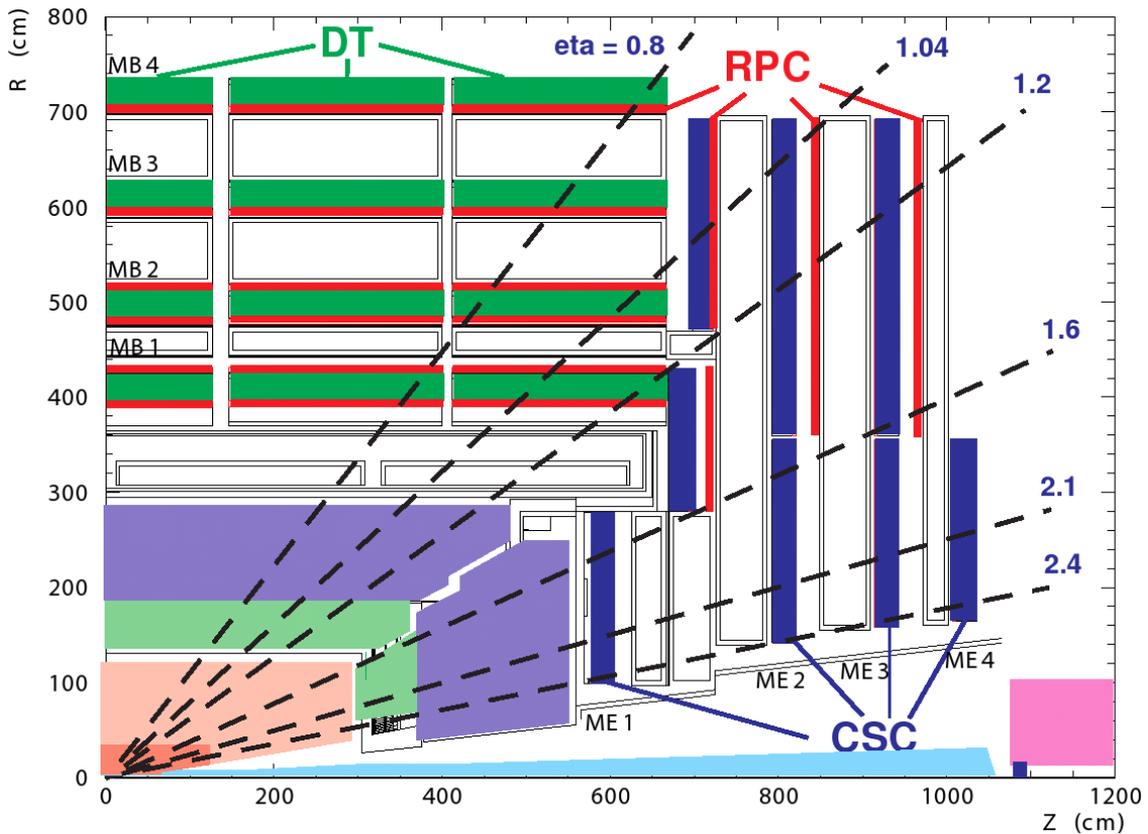


Figure 2.6: Quarter-view of the CMS detector. The muon system is labeled [22].

The muon system consists of about  $25000 \text{ m}^2$  of detection area, composed of three types of gaseous particle detectors:

- In the barrel region four stations of concentric cylinders, containing 250 drift tube (DT) chambers, are installed, 60 in each of the three inner ones and 70 drift chambers in the outer cylinder. They cover the region of  $|\eta| < 1.2$  and provide measurements in the  $r$ - $\phi$  plane as well as in the  $z$  direction. They are also used for triggering with a Level-1 trigger  $p_T$  resolution of about 15%.
- In the two planar endcap regions 540 trapezoidal cathode strip chambers (CSC) measure muons between  $0.9 < |\eta| < 2.4$ . The CSCs are arranged in four stations in each endcap and consist of arrays of anode wires crossed with cathode strips, leading to precise  $r$ - $\phi$  and  $\eta$  measurements. As they have a fast response time and a fine segmentation, they are also suitable for triggering on the  $p_T$  of muons with a Level-1 trigger resolution about 25%.
- To form a fast, independent and highly-segmented trigger system, 610 resistive plate chambers (RPCs) are integrated in the muon system, reaching to  $|\eta| < 1.6$ . There are 6 layers embedded in the barrel region and three in each endcap region, delivering both a good time resolution of just one nanosecond, much shorter than the bunch crossing time, and an acceptable spatial resolution.

Finally, the muon system delivers an independent muon momentum measurement additionally to the inner tracking system, leading to an improvement of the momentum resolution. In table 2.2 the transverse momentum resolution in the whole muon system is summarized.

$p_T$ [GeV]	barrel region	endcap region
	$\Delta p_T/p_T$ [%]	$\Delta p_T/p_T$ [%]
1000	18.6	37.9
500	15.2	-
300	12.3	-
100	9.3	19.4
10	8.9	15.2

Table 2.2: Transverse momentum resolution in the muon system [25].

### 2.2.5 Trigger and Data Acquisition

The CMS Trigger and Data Acquisition (DAQ) system, specified in reference [28], has to solve the important task, to select the interesting events, like energetic, head-on collisions, at very high interaction rates. At the design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , the proton collision rate will be 20 per bunch crossing with a bunch crossing frequency of 40 MHz. This high rate of events has to be reduced to  $\sim 100$  Hz, which is reached by the performance of the Level-1 (L1) Trigger, followed by the High-Level Trigger (HLT).

The L1 Trigger analyses for every bunch crossing coarsely segmented data from the calorimeters and the muon system. It is made up of custom-designed, largely programmable electronics and has to decide in only  $3.2 \mu\text{s}$  if an event is accepted or rejected. During this time, the whole detector data is stored in pipelined memories. The L1 Trigger reduces the data rate to maximal 100 kHz.

The selected data is then pushed into the DAQ system. Implemented in a computer farm, the HLT reduces the L1 output rate to 100 Hz for mass storage. The HLT is a software system, using faster versions of the offline reconstruction software to run a physics selection. An overview of the architecture of the DAQ system is presented in figure 2.7.

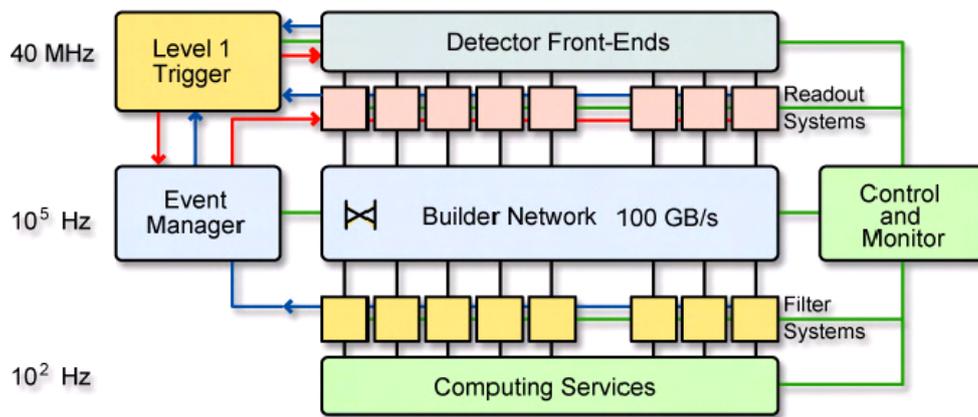


Figure 2.7: General architecture of the CMS DAQ system [19].



# Chapter 3

## Monte Carlo Studies

The interactions between two protons are very complex and it is impossible to select a pure signal sample, as there are always produced events, which result in the same signature in the detector as the signal, the so called background processes. So it is of paramount importance to understand both the signal and background processes as good as possible. Therefore, Monte Carlo (MC) event generators are used to simulate the different processes. In this chapter the concept of event generators is described and those, used in this analysis, are briefly introduced. Furthermore, the main characteristics of the signal event signature are presented and an overview of the used MC samples is given. Finally, the used software framework is described in brief.

### 3.1 Monte Carlo Event Generators

The generated Monte Carlo events should describe the real data as realistic as possible. The underlying theory for interactions between particles is quantum mechanics, where many properties of the particles, like their momentum, are stochastically distributed. So, also the output of a MC event generator should have some fluctuations, which leads to the basic feature of every MC event generator: random numbers are thrown to obtain candidate events, while all significant variables with the predicted probability distributions are taken into account. But as proton-proton collisions have a complex nature, the event generator subdivides the process into several components shown in figure 3.1. The basic steps in the event generation are:

- **Parton Distribution**

For collisions of composed particles like protons, it is important to know the momentum fraction of the partons for the simulation. The parton distribution function is described in section 1.6.

- **Hard Subprocess**

The hard process characterises fundamentally the event and is well described by perturbation theory. It has to be considered, that it is possible to produce heavy particles with a subsequent decay to partons, in a time shorter than the

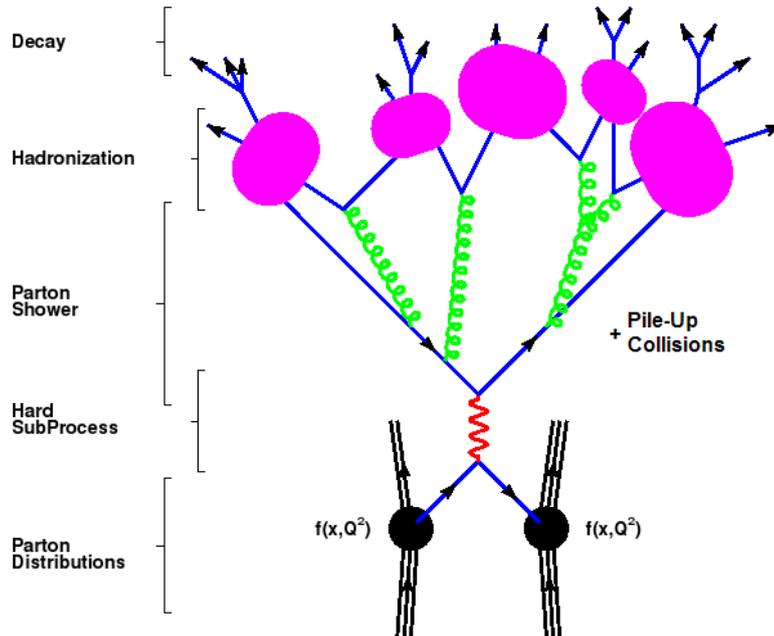


Figure 3.1: Schematic overview of the basic structure of a showering and hadronization generator event. [30]

begin of the parton shower. These short-lived resonances are close connected to the hard process itself.

- **Parton Shower**

The parton shower is linked to higher order QCD effects. Partons can radiate for example gluons before or after the hard process took place, the initial- or final-state radiation, respectively. Additionally, all partons branch, what is described by an event generator with a splitting function. The branching or showering continues until a certain cutoff point is reached. Due to the showering, jets of quarks and gluons are produced in the direction of the primary parton.

- **Hadronisation**

Coloured particles do not exist. So the quarks and gluons, produced in the shower cannot be understood as free particles. The confinement of the strong interaction leads to new quark-antiquark pairs, if two coloured particles separate. So the partons in the jets have to be grouped to colourless objects, forming hadrons which can be seen by the detector. Hadronisation proceeds at a scale of low momentum transfer where the strong coupling constant becomes large. The outcome of this is, that perturbation theory breaks down. The generators use different hadronisation models to simulate this process.

- **Decay**

The formed hadrons often have only short lifetimes and decay into long living

lighter hadrons.

- **Underlying Event and Pile-Up**

The initial partons for the hard process stem from the primary hadrons, which are left over as coloured hadron remnants. But these partons of the remnant can also interact with other partons, leading to multiple parton interactions, which is called the underlying event. Furthermore, it is possible, that other proton-proton collisions appear during the same bunch crossing. This pile-up as well as the underlying event has to be added to the event simulation, too.

A more detailed overview about MC generators for hadron colliders and their methods can be found in reference [30]. Showering and hadronisation event generators are based on the presented structure and generate a complete event. But it exists also another generator type, the matrix element event generators. These generators are also called parton-level generators, since the final state consists of bare partons. They compute all relevant tree-level matrix elements for the hard process with a fixed parton number in the final state in the lowest order of perturbation theory. The hadronisation process or next to leading order effects are not considered. To obtain a full description of the generated process, the final state partons have to be passed subsequently to a showering and hadronisation generator.

The MC event generators used for this analysis to model the signal and backgrounds are now introduced.

### **PYTHIA**

PYTHIA [31] is a showering and hadronisation event generator, which can be used for general purpose to obtain a full event simulation of high energetic collisions of electrons, positrons, protons or antiprotons. But it is also possible to process events, produced with another matrix element generator. The hadronisation model is based on the Lund string model, where a linear confinement between two quarks is assumed. When they move apart, a colour flux tube is stretched until the stored energy in this tube or string is high enough to produce a new quark-antiquark pair.

### **ALPGEN**

ALPGEN [33] is a parton-level generator for multiparton hard processes at hadron colliders. The matrix elements are evaluated in leading order in QCD and electroweak interactions for many Standard Model parton processes. Spin correlations are taken into account as well as the full information on the colour and flavour structure is given for the parton-level events. The generator is designed especially with the focus on multi-jet final states.

### **TopReX**

TopReX [34] is like ALPGEN a parton-level generator, but it is specialised on heavy particle production processes in hadronic collisions. The generated events of processes such as single top quark or Higgs production contain information about the differential cross section, colour flow and the flavours and momenta of the partons in the initial and final state. Also the spin polarisations are considered.

## 3.2 Monte Carlo Signal Event

For this analysis only those events are considered as signal, where the Higgs boson is produced via the vector boson fusion (VBF) process and decays into a  $W$  boson pair. For these  $W$  bosons only the leptonic decay is examined. Figure 3.2 shows the Feynman diagram with all final state particles. For this study a sample of  $\sim 45000$

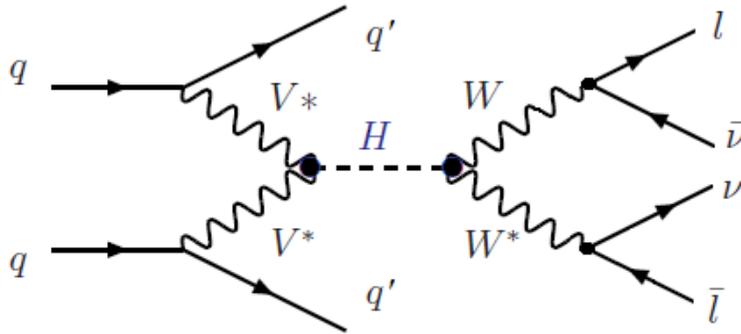


Figure 3.2: Feynman diagram of the signal process. The Higgs boson is produced by the VBF process and decays into two  $W$  bosons, each of which decays subsequently into a charged lepton and the corresponding neutrino.

events has been generated and showered with PYTHIA version 6.409.

Even though the cross section for the VBF production channel is roughly 10 times smaller at the LHC than the gluon initiated one, this production mode is nevertheless important to study. Due to the  $WWH$  or  $ZZH$  vertex for the Higgs boson production, leading to a cross section proportional to the squared coupling between the Higgs and the weak vector bosons, it is possible to proof directly the existence of such tree level couplings and thus, if the Higgs boson behaves as predicted by the standard model or not. This becomes apparent by an observation of the VBF production process. In addition, it features a distinct signature.

Characteristic for the VBF production mode are two very forward jets with maxima in the region  $|\eta| \simeq 3$ . Each of the two initial quarks, originating from the protons, emit a weak vector boson, for example one a  $W^+$  and the other a  $W^-$  boson. Thereby, the quarks are scattered to high pseudorapidities and due to the confinement, they form these two forward jets with the characteristics adopted from the primary quarks. As there is no colour exchange between the two initial quarks, there should not occur any jets in the central region. In figure 3.3 the distributions of the transverse momentum and pseudorapidity of the scattered initial quarks are shown.

The produced Higgs boson subsequently decays into a  $W^+W^-$  pair. As the mass of the Higgs boson is assumed to  $m_H = 130 \text{ GeV}/c^2$ , at most one  $W$  boson can be produced on-shell. In figure 3.4 the mass distributions of the two produced  $W$  bosons are pictured. The distribution of the heavier boson, shown in (a), is fitted with a Breit-Wigner distribution. The measured  $W$  mass  $m_W = 80.266 \pm 1.560 \text{ GeV}/c^2$  is compatible with the assumption, that one  $W$  boson is produced on-shell.

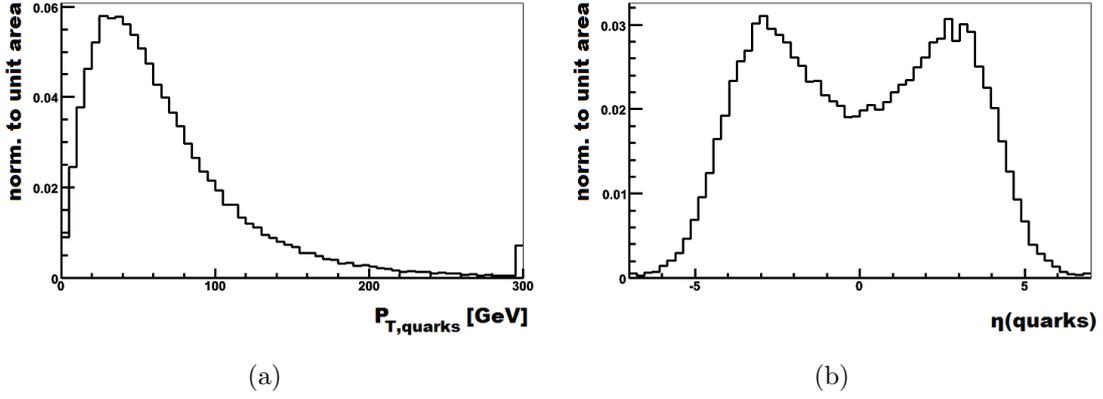


Figure 3.3: Kinematics of the initial quarks after the vector boson radiation. The distribution of the transverse momentum  $p_T$  (a) and pseudorapidity  $\eta$  (b) are presented.

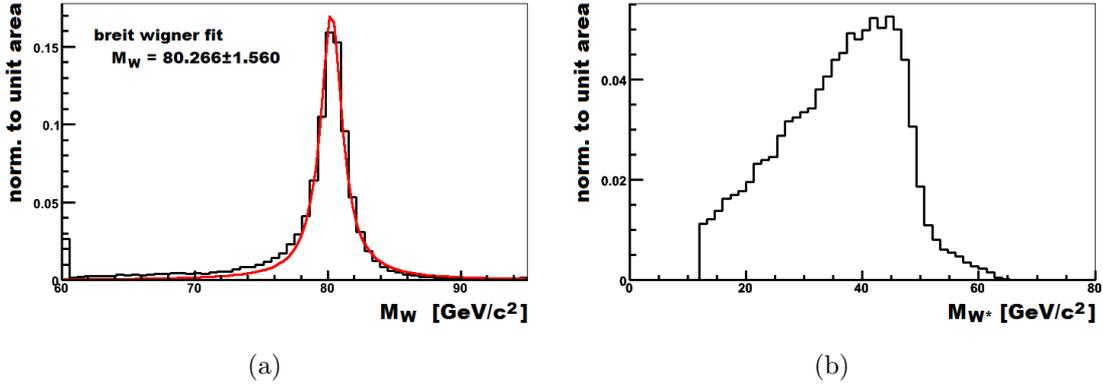


Figure 3.4: Mass distributions of the two produced  $W$  bosons. The boson with higher mass is shown in (a) and the other one in (b).

The two  $W$  bosons have to be produced back to back in the rest frame of the Higgs boson, since the momentum has to be conserved. This is shown in figure 3.5. Furthermore, as the Higgs boson is a particle with spin 0 in contrast to the  $W$  bosons with spin 1, also the spins of the  $W$  bosons have to point into the opposite direction. Thus, two left-handed or two right-handed  $W$  bosons are produced in the Higgs boson decay, each of which decays leptonically, for example into  $e^-$  and  $\bar{\nu}_e$ . Due to the V-A theory of the weak interaction, the  $W$  bosons only interact with left-handed fermions and right-handed antifermions, so that the  $e^-$  has to be left-handed and the  $\bar{\nu}_e$  right-handed. A well-defined state only exists for massless particles, but compared to the much larger mass of the  $W$  or Higgs boson, the lepton mass can be neglected in this treatment. In figure 3.6 the whole Higgs decay is shown in its rest frame. The red arrows above indicate the spin direction of the particles.

This explains another feature of the signal event. The polar angle  $\theta$  between the two charged leptons is very small, which means, that they almost point into the same direction, as it is expected from the theory. The distribution of the cosine of

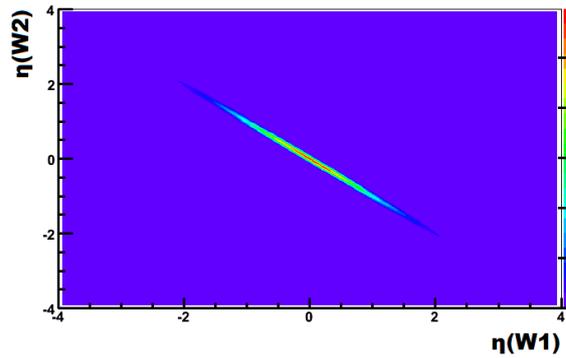


Figure 3.5: Distribution of pseudorapidity  $\eta(W2)$  vs  $\eta(W1)$  in the rest frame of the Higgs boson. The two  $W$  bosons are produced back to back.

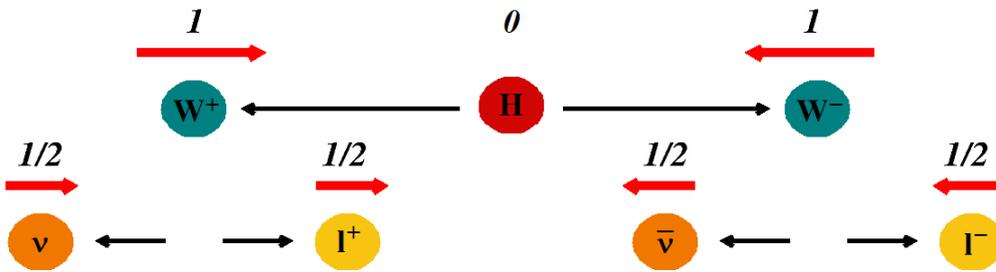


Figure 3.6: Example of Higgs boson decay in its rest frame. The spin configuration for a production of two left-handed  $W$  bosons and their decays is demonstrated by the red arrows above the particles. The two leptons  $l^+$  and  $l^-$  must always fly into the same direction due to the V-A theory.

$\theta$  is given in figure 3.7(a). In addition to these leptons, two neutrinos are produced from the  $W$  boson decays. It is not possible to detect the neutrinos directly, as they do not interact with the detector material, but they appear as missing transverse energy  $\cancel{E}_T$ . To obtain the expected transverse momentum distribution of  $\vec{\cancel{E}}_T$  of the generated MC event, the momenta of the neutrinos have to be added before the transverse momentum is determined. The resulting distribution is shown in figure 3.7(b).

### 3.3 Generated Monte Carlo Event Data

As already mentioned, it is not sufficient to understand only the signal events, also the backgrounds have to be examined, consisting of several different processes. The used samples with the generators, which were used to produce them, are listed in table 3.1.

The dominant background processes for this analysis are  $W$  or  $Z$ +jets as well as  $t\bar{t}$  production. A  $W$  or  $Z$  boson is produced via quark-antiquark annihilation and due to gluon radiation, both, heavy and light flavour jets emerge. For  $t\bar{t}$  events gluon fusion is the main production mode at the LHC. These events are of particular

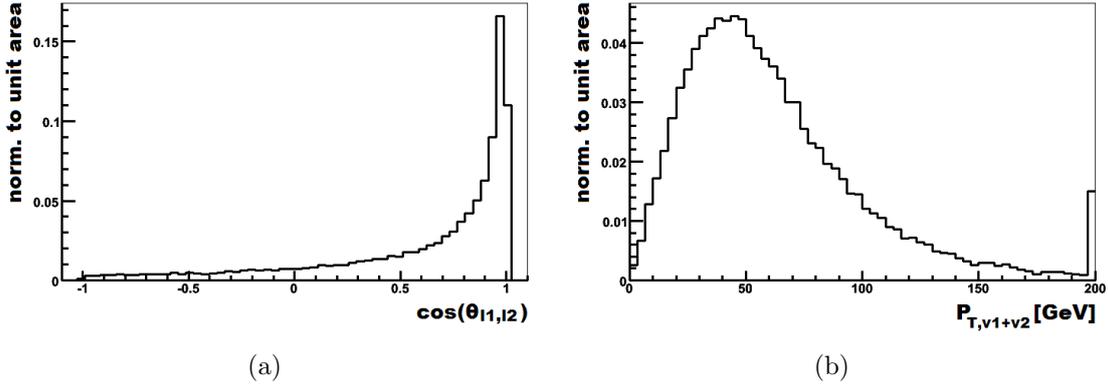


Figure 3.7: Distribution of cosine of the polar angle  $\theta$  between the two charged leptons (a) and transverse momentum of the sum of the two neutrinos.

process	hard process	shower
$qqH$		PYTHIA
$t\bar{t}$	ALPGEN	PYTHIA
$tW$	TopRex	PYTHIA
$WW$		PYTHIA
$WZ$		PYTHIA
$ZZ$		PYTHIA
$W$ +jets	ALPGEN	PYTHIA
$Z$ +jets	ALPGEN	PYTHIA
QCD		PYTHIA

Table 3.1: The Monte Carlo samples and used generators for the production. PYTHIA package version 6.409 and ALPGEN version 2.12 were used.

importance, if both  $W$  bosons, originating from the top quark decays, decay leptonically. Then the signal event signature is faked, as there are two charged leptons, missing transverse energy caused by the neutrinos and two jets, arising from the  $b$  quarks or, for example, from the underlying event to provide more forward directed ones. Examples of Feynman diagrams for these processes are shown in figure 3.8.

Further backgrounds are the diboson production process, consisting of  $WW$ ,  $WZ$  and  $ZZ$  production, and the single top production process, named as  $tW$ , with the corresponding Feynman diagrams in figure 3.9 and 3.10(a), respectively.

Events without any electroweak gauge bosons involved, are called QCD events and consist of many different processes. Since it is possible, that hadronic jets are misidentified as charged leptons and that mismeasured transverse energy fakes the neutrino signature, these events also have to be considered as background source. So, the QCD background is an instrumental background which results from the detector simulation and not from the same signal signature. Therefore, it is hardly possible to model it sufficiently well with MC generators. An example for a QCD event is illustrated in figure 3.10(b).

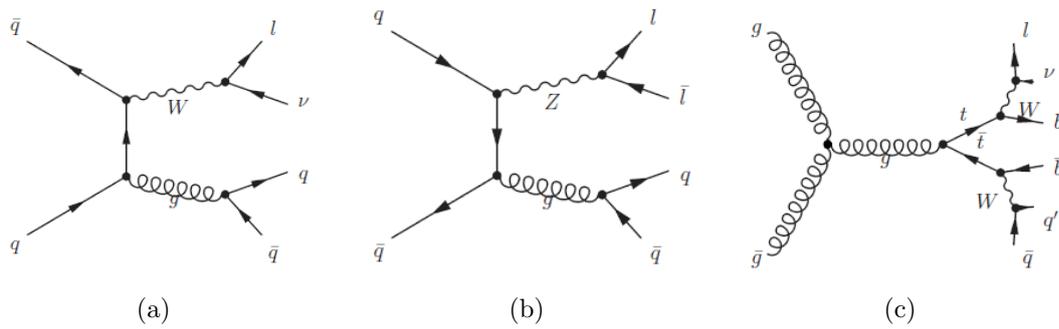


Figure 3.8: Examples of the production of a  $W$  boson (a) or a  $Z$  boson (b) plus jets and a  $t\bar{t}$  event (c).

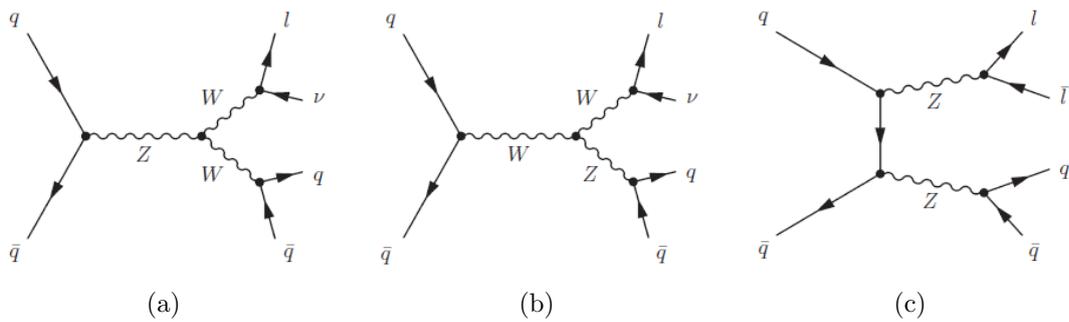


Figure 3.9: Examples of diboson production. There are three different processes:  $WW$  (a),  $WZ$  (b) and  $ZZ$  (c) production.

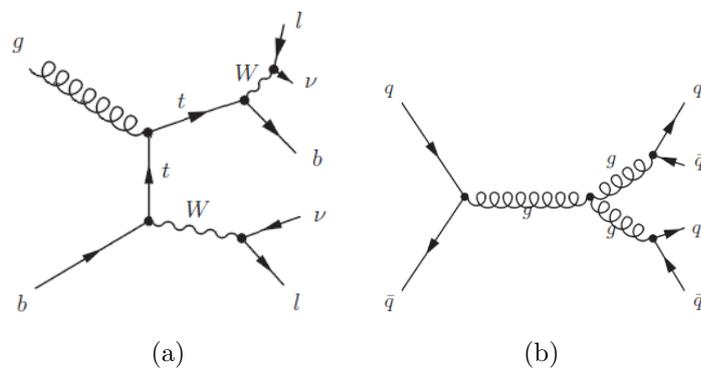


Figure 3.10: Examples of  $tW$  (a) and QCD (b) events.

Afterwards, all MC samples are subjected to a full CMS detector simulation, based on the software package GEANT4 [36]. For the simulation the particles of the generated MC events propagate through the detector and due to their interaction with the detector material, a detector response is simulated. These obtained signals from the different detector components are in a comparable format to those signals, which will be produced from real events when the experiment is running. Finally, more complex objects like jets or tracks are reconstructed.

## 3.4 CMS Software Framework

All simulations of the samples rely on the CMS software framework CMSSW and have been generated within the production of the so-called "CSA07 data" using CMSSW versions 1.4.X and 1.6.X. CMSSW is a collection of software with a modular architecture, which allocates all necessary modules for data analysis with the CMS experiment like for example event simulation and reconstruction. The core within the framework is the Event Data Model (EDM). Different types of data pertaining to the MC event, the raw detector output as well as the reconstructed data, are stored in the EDM and are passed through the different modules. These can be dynamically loaded and also individually adjusted or extended, which facilitates the development of the analysis software. A framework diagram of the EDM is shown in figure 3.4.

The event selection, applied on the events coming from the detector simulation, has been performed with CMSSW version 1.6.12. More information about the framework and the used samples are given in reference [37].

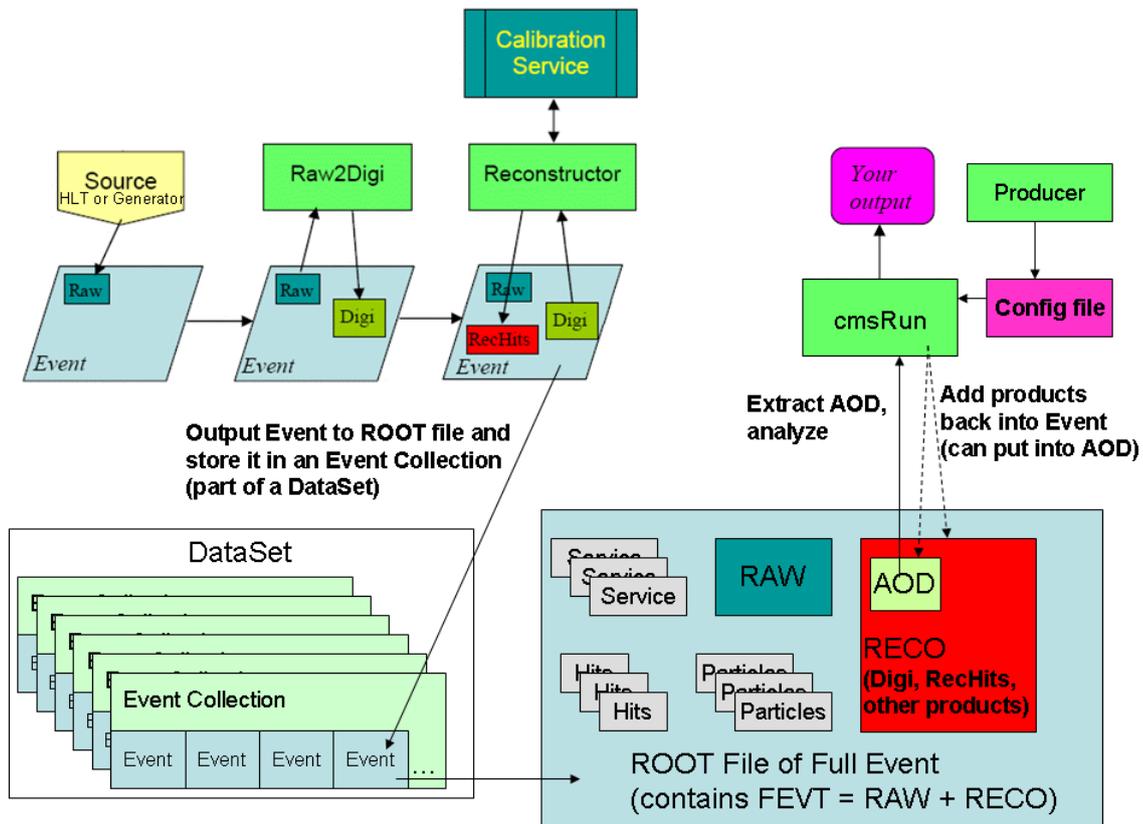


Figure 3.11: Framework diagram of the event data model (EDM). The output from a detector or a MC event is stored in the Event as a collection of the RAW data. As the event data is processed, products (of producer modules) are additionally stored in the Event as reconstructed (RECO) data objects. A subset of the RECO data is the Analysis Object Data (AOD), which is sufficient for most kinds of physics analysis [37].

# Chapter 4

## Selection of Candidate Events

To increase the ratio of signal to background events in the analysed Monte Carlo samples, a preselection has to be performed. Therefore, after having reconstructed the physical objects out of the detector information, some kinematic cuts are applied on these objects. Finally, at the end of this chapter, an overview about the expected signal and background events for  $10 \text{ fb}^{-1}$  is given.

### 4.1 Event selection

In section 3.2 the main characteristics of the signal process have been presented and the Feynman diagram with all final state particles is shown in figure 3.2. Summarised the signal signature comprises:

- two isolated charged leptons,
- two forward jets
- and missing transverse energy originating from the neutrinos.

The following sections describe the reconstruction requirements for these physical objects.

#### 4.1.1 Lepton Selection

In this analysis only those events are considered, where the  $W$  boson decays into an electron or a muon and the corresponding neutrino. The decay into taus is not included. As both leptons originate from a  $W$  boson decay, they have to be isolated. Therefore, at the primary vertex an isolation cone with radius  $\Delta R = 0.3$  is placed around the lepton and within this cone tracks are disallowed, except from the leptons track itself. For this purpose only those tracks are considered, which are associated with the primary vertex of the lepton by considering the transverse impact parameter. The closest point between the track and primary vertex with respect to the  $z$ -axis has to be below 1.5 mm and the impact parameter to any other interaction vertex has to be above the previous one. Tracks originating from the underlying

event, for example, are expected to be neglected. Due to this isolation criterion leptons, which are for example faked by jets, should be rejected. Additionally, there are some selection requirements that are specific to the lepton type.

### Electron Identification

- The main energy deposit of an electron is in the electromagnetic calorimeter (ECAL). Therefore, the ratio between the energy deposit in the hadronic calorimeter (HCAL) and the ECAL has to be below 0.1.
- The energy loss caused by bremsstrahlung shall not be distinctive. Affecting the electron reconstruction close to the vertex, the ratio between the energy of the electron super cluster of the ECAL and its momentum at the inner layer of the tracker is demanded close to one. The accepted range is between 0.5 and 1.5. Related to the reconstruction close to the ECAL, the ratio between the energy of the basic cluster and the momentum at the outer layer of the tracker is taken and again demanded close to one, namely above 0.8.

### Muon Identification

- To test the muon hypothesis compatibility, a relative likelihood is calculated based on the energy deposit in the ECAL, HCAL and its outer barrel. It is defined as  $L_{\mu}/(L_{\mu}+L_{not-\mu})$  and has to exceed 0.8 for accepting a muon.
- The number of valid hits of the muon track in the tracker is required to be equal or above 8.
- Muons penetrate through the whole muon detectors, so that the extrapolated silicon track of a muon can be matched with a segment of the outer station of the muon detector. The TMLastStation algorithm is based on this coherence and at least one matched segment of the outer station is demanded.
- Not less than two matched segments of any station of the muon detector are required, matched with the muon track in the tracker.

More detailed information about the electron and muon identification can be found in reference [39] and [40], correspondingly. An event is only selected, if all involved leptons fulfil the presented lepton identity cuts for electrons or muons, respectively, and additionally, if there are at least two selected leptons with a transverse momentum of  $p_T \geq 15 \text{ GeV}/c$ .

### 4.1.2 Jet Reconstruction

The measurement of jets is basically a clustering of energy deposits in the ECAL and HCAL. For this analysis the Iterative Cone algorithm with a cone size  $R = 0.5$  is used for the jet reconstruction. One or more HCAL cells are combined with the corresponding ECAL crystals to calorimeter towers, which are sorted according

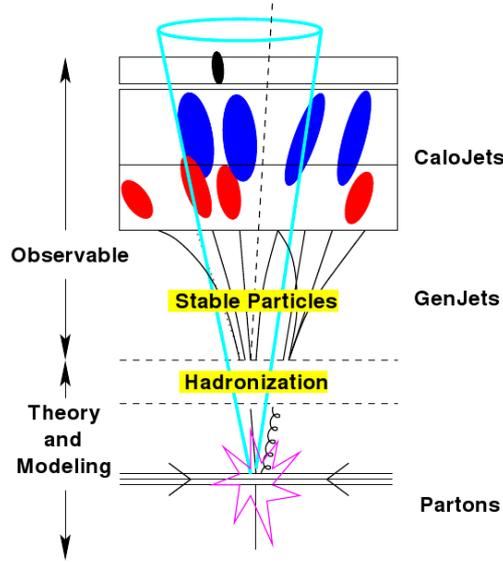


Figure 4.1: Illustration of a cone for the jet reconstruction [45].

to their transverse energy  $E_T$  in a list, used as input. Only those towers with  $E_T > 1$  GeV are considered. Starting with the tower with the highest  $E_T$ , called seed, a cone of the size  $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$  is constructed around the tower. All other towers within this cone are associated with a proto-jet, whose energy and direction is determined by the sum of its constituent parts. The calculated jet direction is used as new seed for the described procedure. This algorithm is repeated until the proto-jet energy changes less than 1% and its direction change drops below  $\Delta R < 0.01$ . Then all towers enclosed by the cone of the stable proto-jet are removed from the input list, as they have been already allocated to a jet and the method is performed until no further towers are listed as input. The construction of a cone for the jet reconstruction is schematically illustrated in figure 4.1. For more details on jet algorithms see for example reference [42].

Advanced jet corrections on the energy, necessary because of calorimeter effects, for instance, are not applied in this analysis. To prevent the reconstruction of an isolated electron as a jet, since they also leave energy deposit in the HCAL, it is needed to remove the corresponding jets from the jet list. Therefore, a cone with  $R = 0.5$  is constructed around the isolated electron and all jets within this cone are omitted. This procedure is called jet cleaning.

For the event selection a minimum of two jets is required and up to five are accepted. The transverse momentum of the selected jets has to fulfil  $p_T \geq 20$  GeV/ $c$  and jets in the full calorimeter range  $|\eta| \leq 5$  are accepted.

### 4.1.3 Missing Transverse Energy

In the event topology two neutrinos appear because of the claimed leptonic decay of both  $W$  bosons. As already mentioned in section 3.2, neutrinos cannot be measured

directly, since they do not interact with the detector material. The answer to this problem is that the incoming beam particles do not have any transverse momentum, so that after an interaction the transverse momentum still has to be zero. Hence, the neutrinos can be measured as missing transverse energy  $\cancel{E}_T$ . As it is described in reference [41], with the sum over all calorimeter towers having an energy  $E_n$ , pseudorapidity  $\eta_n$  and azimuthal angle  $\phi_n$ , the vector of the missing transverse energy can be calculated as follows:

$$\vec{\cancel{E}}_T = - \sum_n \left( \frac{E_n \cos \phi_n}{\cosh \eta_n} \hat{x} + \frac{E_n \sin \phi_n}{\cosh \eta_n} \hat{y} \right). \quad (4.1)$$

The total scalar missing transverse energy is defined as magnitude of the vector:  $\cancel{E}_T = |\vec{\cancel{E}}_T|$ . Since muons are also not stopped in the detector, they have to be considered for the calculation, too. Their expected energy deposit in the calorimeter, which is about 4 GeV, has to be replaced by their reconstructed transverse track momentum  $p_T$ . For the event selection this muon-corrected missing transverse energy is required to be  $\cancel{E}_T \geq 20$  GeV.

#### 4.1.4 $Z$ Boson Veto

To reduce those backgrounds where both selected charged leptons stem from a  $Z$  boson decay, such as in the  $Z$ +jets production, the  $Z$  boson mass can be reconstructed by calculating the invariant mass of the sum of the two leptons. The invariant mass of a particle with an energy  $E$  and momentum  $\vec{p}$  is defined as

$$m_{inv} = \sqrt{\frac{E^2}{c^4} - \frac{\vec{p}^2}{c^2}}. \quad (4.2)$$

Only those events are selected with an invariant two-lepton mass of  $m_{inv} \leq 85$  GeV/ $c^2$ . This cut is expected to affect only barely the signal sample as can be seen in figure 4.2, where the distribution of the invariant mass of the generated leptons is presented. The used Monte Carlo sample is the same as in section 3.2.

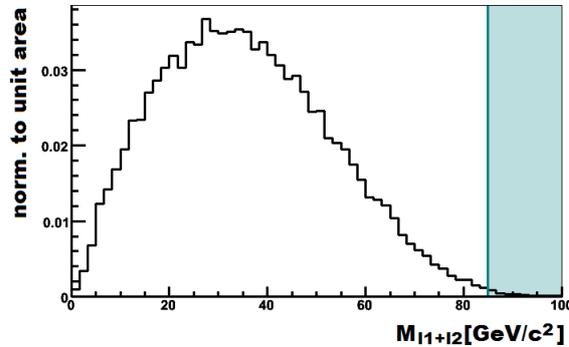


Figure 4.2: Invariant mass distribution of the vectorsum of the two charged leptons in the signal sample.

## 4.2 Signal and Background Estimation

To calculate the expected number of candidate events, theoretical predictions as well as detector effects and efficiencies have to be taken into account. In general, the expected number of events  $N_{exp}$  can be estimated with

$$N_{exp} = \sigma \epsilon_{evt} \mathcal{L}_{int}, \quad (4.3)$$

where  $\sigma$  is the theoretically predicted cross section of the process,  $\epsilon_{evt}$  the event detection efficiency and  $\mathcal{L}_{int}$  the integrated luminosity. For the signal process, named as qqH, since the Higgs boson is produced via the vector boson fusion process, the decay into a  $W$  boson pair is demanded with both bosons decaying leptonically and also for the  $W/Z$ +jets production the specific decay mode into leptons is requested. Therefore, for these processes the cross section is already multiplied with the corresponding branching ratios.

The event detection efficiency is composed of the event detection efficiency  $\epsilon_{evt}^{MC}$ , the efficiency correction factor  $\epsilon_{corr}$  and the trigger efficiency  $\epsilon_{trigger}$ :

$$\epsilon_{evt} = \epsilon_{evt}^{MC} \epsilon_{corr} \epsilon_{trigger}. \quad (4.4)$$

The event detection efficiency  $\epsilon_{evt}^{MC}$  can be obtained from the Monte Carlo sample by calculating the ratio of the number of events after the event selection to the one before, while the correction factor  $\epsilon_{corr}$  adjusts differences between the simulated Monte Carlo samples and real data.

For this analysis the expected number of signal and background events for an integrated luminosity  $\mathcal{L}_{int} = 10 \text{ fb}^{-1}$  is calculated, considering only the one factor of the event detection efficiency  $\epsilon_{evt}^{MC}$ . An overview about the resulting composition of the selected data set is given in figure 4.3 for  $10 \text{ fb}^{-1}$ . Table 4.1 provides in addition the number of estimated events, the assumed cross sections and the numbers of simulated events before and after the preselection for all samples under consideration.

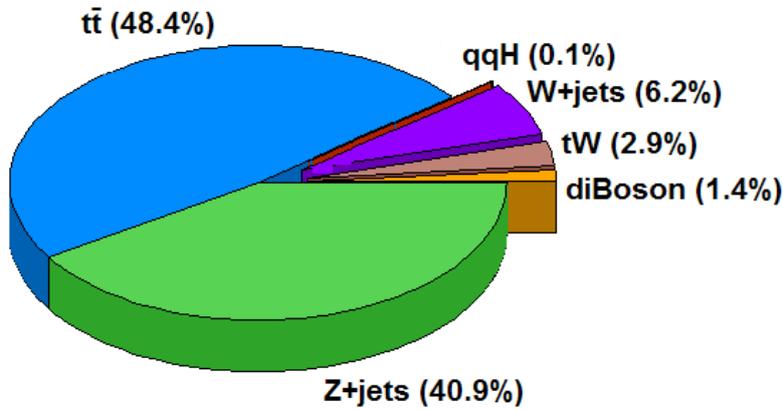


Figure 4.3: Expected fractions of the different processes after the event selection for  $10 \text{ fb}^{-1}$ .

process	$\sigma$ [pb]	tot. events	sel. events	eff [%]	est. events
qqH	0.07	45317	3032	6.691	46.8
tt0jet	619	1456646	3084	0.212	13105.4
tt1jet	176	361835	882	0.244	4290.1
tt2jet	34	81215	183	0.225	766.1
tt3jet	6	14036	26	0.185	111.1
tt4jet	1.5	5352	2	0.037	5.6
tW	62	443791	790	0.178	1103.7
WW	114.3	845261	283	0.033	382.7
WZ	0.585	58897	176	0.299	17.5
ZZ	16.1	143113	117	0.082	131.6
W0jet	45000	8796412	0	0.000	0.0
W1jet lowPt	9200	9088026	2	0.000	20.3
W1jet highPt	250	247013	0	0.000	0.0
W2jet lowPt	2500	2380315	21	0.001	220.6
W2jet highPt	225	287472	8	0.003	62.6
W3jet lowPt	590	352855	46	0.013	769.2
W3jet highPt	100	117608	27	0.023	229.6
W4jet lowPt	125	125849	24	0.019	238.4
W4jet highPt	40	39719	20	0.050	201.4
W5jet lowPt	85	62238	28	0.045	382.4
W5jet highPt	40	43865	24	0.055	218.9
Z0jet	4400	3251851	17	0.001	230.0
Z1jet lowPt	935	944726	125	0.013	1237.1
Z1jet highPt	30	36135	22	0.061	182.7
Z2jet lowPt	271	289278	568	0.196	5321.1
Z2jet highPt	28	35285	268	0.760	2126.7
Z3jet lowPt	68	73182	334	0.456	3103.5
Z3jet highPt	13	24316	227	0.934	1213.6
Z4jet lowPt	14	33083	169	0.511	715.2
Z4jet highPt	4.3	6616	75	1.134	487.5
Z5jet lowPt	8.8	12136	73	0.602	529.3
Z5jet highPt	4.9	5966	41	0.687	336.7

Table 4.1: Number of estimated signal and background events after event preselection for an integrated luminosity of  $\mathcal{L} = 10 \text{ fb}^{-1}$ . The  $W/Z$ +jets samples are sorted according to the  $p_T$  of the boson. LowPt ranges from  $0 \text{ GeV}/c \leq p_T \leq 100 \text{ GeV}/c$  and highPt from  $100 \text{ GeV}/c < p_T \leq 300 \text{ GeV}/c$ .

In table 4.2 the same values are given for the QCD background. As already mentioned in section 3.3, the QCD background is only unsatisfactorily described, in consequence of the multiplicity of different QCD events. Additionally, the selection efficiency is very low, the selected events have been only chosen as a result of mismeasurements in the detector, and their cross section is very large. For a well described QCD background, real data has to be studied. Finally, this is the reason, why in the following chapters the QCD background is not anymore included.

Nevertheless, it is interesting to study briefly these selected QCD events, for example, to improve the event selection over again. All together 46 events pass the event selection and only two of these events contain muons as isolated leptons. So it could be advantageous to investigate the electron identification requirements once more.

process	$\sigma$ [pb]	tot. events	sel. events	eff [%]	est. events
QCD Pt 0-15	$55 \cdot 10^9$	733104	0	0.000	0.0
QCD Pt 15-20	$1.46 \cdot 10^9$	1300976	0	0.000	0.0
QCD Pt 20-30	$6.3 \cdot 10^8$	2511934	0	0.000	0.0
QCD Pt 30-50	$1.63 \cdot 10^8$	2441441	0	0.000	0.0
QCD Pt 50-80	$21.6 \cdot 10^6$	2456689	0	0.000	0.0
QCD Pt 80-120	$3.08 \cdot 10^6$	1178357	0	0.000	0.0
QCD Pt 120-170	$4.94 \cdot 10^5$	1263201	5	0.000	19553.5
QCD Pt 170-230	$1.01 \cdot 10^5$	1160565	8	0.001	6962.1
QCD Pt 230-300	24500	1202951	12	0.001	2444.0
QCD Pt 300-380	6240	1183344	8	0.001	421.9
QCD Pt 380-470	1780	1189350	7	0.001	104.8
QCD Pt 470-600	683	1229378	5	0.000	27.8
QCD Pt 600-800	204	727670	1	0.000	204
QCD Pt 800-1000	35.1	1114443	0	0.000	0.0
QCD Pt 1000-1400	10.9	763147	0	0.000	0.0
QCD Pt 1400-1800	1.6	723170	0	0.000	0.0
QCD Pt 1800-2200	0.145	772233	0	0.000	0.0
QCD Pt 2200-2600	$2.38 \cdot 10^{-2}$	753720	0	0.000	0.0
QCD Pt 2600-3000	$4.29 \cdot 10^{-3}$	751840	0	0.000	0.0
QCD Pt 3000-3500	$8.44 \cdot 10^{-4}$	705554	0	0.000	0.0
QCD Pt 3500-inf	$1.08 \cdot 10^{-4}$	749522	0	0.000	0.0

Table 4.2: Number of estimated QCD events after event preselection for an integrated luminosity of  $\mathcal{L} = 10 \text{ fb}^{-1}$ . The samples are sorted according to the labelled  $p_T$  interval, given in  $\text{GeV}/c$ .



# Chapter 5

## Neural Networks

The inspiration of artificial neural networks, in the following just called neural network (NN), originates from biological NNs, consisting of real neurons which are chemically connected and build a complex network for example in the central nervous system. Hence, also NNs are composed of interconnected artificial neurons, or more simply nodes, which use a simple mathematical model to process information. Additionally, the connections between the nodes are weighted, representing the strength of the corresponding contacts, to produce a specific signal flow which is determined by these nodes, the processing elements. With this structure and technique, NNs are suitable to model complex relationships between the inputs and outputs and are especially used for non-linear classification problems, such as in this analysis to decide if a given candidate event belongs to the Higgs signal or to one of the background processes. The basic structure of such a NN is the composition of different layers: the first one is the input, the last the output layer and the ones in between are the hidden layers, with all of them having an individual number of nodes. The connections are only established between nodes of two adjacent layers, so that all nodes of the same layer are independent of each other. For a classification problem the output layer consists typically of only one single node. This network topology with one hidden layer is shown in figure 5.1. It is a feed-forward net.

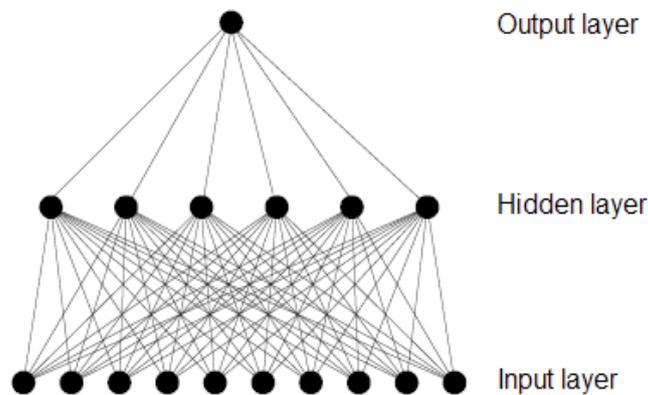


Figure 5.1: Neural network with one hidden layer and output node in the output layer.

## 5.1 NeuroBayes<sup>®</sup> and MLP

For this analysis, two different NN packages have been used: the NeuroBayes package and MLP. Both NNs have been used as a plugin within the framework of the Toolkit for Multivariate Data Analysis (TMVA), which is included in ROOT, a high energy physics data analysis environment. MLP is already implemented in TMVA, while NeuroBayes has to be additionally inserted. More detailed information about the framework TMVA and the NN provided by MLP can be found in reference [49], while information about NeuroBayes are given in reference [50].

### 5.1.1 Network Structure

Both used NNs feature the same basic network structure. For the classification of the events the structure of a feed-forward net with one hidden layer and one single output node has been chosen. The input layer consists of one node for each input variable, in total  $N$ , and, additionally, one bias node. Since the NN shall learn to distinguish between signal and background events based on the variables information, it is important to use discriminating variables. The hidden layer can be composed of an arbitrary number of nodes  $H$ . The input of each node  $j$  in the hidden layer can be calculated via the biased weighted sum of the values of the input layer  $x_i$ :

$$o_j(\mathbf{x}) = \sum_{i=1}^N \omega_{ij} x_i + \eta_{0j} \quad (5.1)$$

with  $\omega_{ij}$  denoting the weight between the node  $i$  from the input to the node  $j$  from the hidden layer and with  $\eta_{0j}$  implementing the threshold for the node  $j$  given by the weight between this and the bias node of the input layer. The output of each node results from the transfer of the weighted sum to a sigmoid function as neuron activation function, which maps the interval  $[-\infty, \infty]$  to  $[-1,1]$  or  $[0,1]$ . Therefore, an often chosen function is the hyperbolic tangent and two further sigmoid functions, used in the case of NeuroBayes or MLP, are given in equation 5.2 or 5.3, respectively.

$$S(x) = \frac{2}{1 + e^{-k(x)}} - 1 \quad (5.2)$$

$$S(x) = \frac{1}{1 + e^{-k(x)}} \quad (5.3)$$

Both functions are shown in figure 5.2.

Finally, the output of the one single node in the output layer is determined in an analogous manner as the nodes of the hidden layer, so that it is calculated with

$$o = S\left(\sum_{j=1}^H \omega_j S\left(\sum_{i=1}^N \omega_{ij} x_i + \eta_{0j}\right)\right) \quad (5.4)$$

where  $H$  is the number of hidden nodes and  $\omega_j$  the weight between the hidden node  $j$  and the output node.

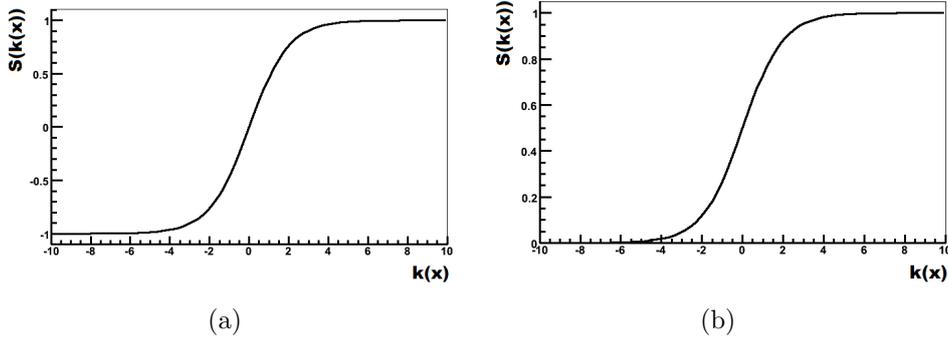


Figure 5.2: Possible sigmoid activation function  $S(x)$  as given in equation 5.2 (a) and 5.3 (b).

### 5.1.2 Training the Network

In this study, NNs are used to classify events, hence, historical or simulated data with a known classification, meaning a specified target, is needed for the network training. With every training iteration the weights between the network nodes  $\omega_{ij}$  are readjusted in order to minimise the deviation between the network output and the target. Therefore, in the case of MLP the quadratic error function  $E$  is minimised:

$$E = \sum_i \frac{1}{2} (o_i - T_i)^2, \quad (5.5)$$

where  $o_i$  denotes the network output for a training event  $i$  as it results from equation 5.4 and  $T_i$  represents the target value for the event classification, while for NeuroBayes the entropy loss function  $E_D$  is used:

$$E_D = \sum_i \log\left(\frac{1}{2}(1 + T_i o_i + \epsilon)\right). \quad (5.6)$$

The constant  $\epsilon$  is a small numerical regularisation factor, which will be equal to zero after a few training iterations. The shape of both functions is very complex in a multidimensional space, where the number of weights in the network predict the number of dimensions, so that the minimum cannot be found analytically. For this purpose, the algorithm of gradient descent is used, illustrated in figure 5.3. During the training the weights are updated after every iteration step with a change proportional to the gradient of the corresponding error function  $E$  or  $E_D$

$$\Delta\omega_{ij} = -\eta \frac{\partial E}{\partial \omega_{ij}} \quad (5.7)$$

where  $\eta$  is the step width or learning rate of the minimum search. If  $\eta$  is too large, the algorithm might overrun the searched minimum, while for  $\eta$  chosen too small, the algorithm takes very long. Finally, this method is often combined with the algorithm of backpropagation. Therefore, a local error is introduced to all nodes, resulting from the deviation between the network output and the target. This error

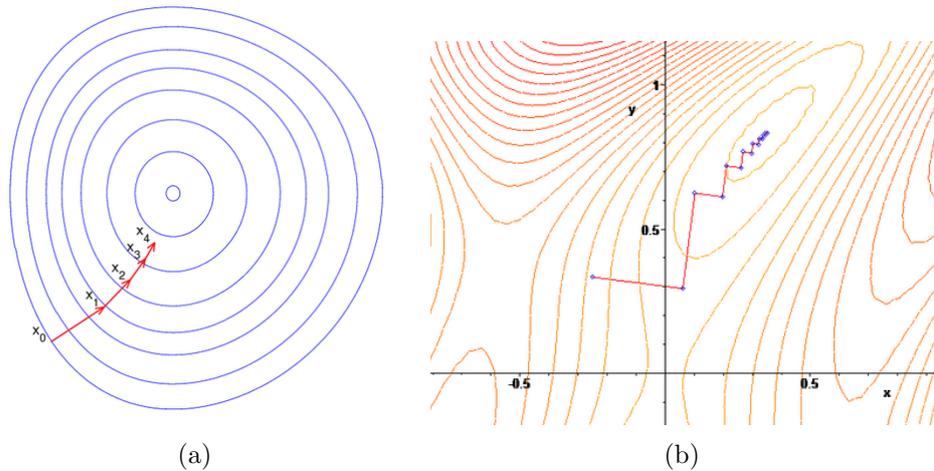


Figure 5.3: Illustration of gradient descent method [51]. As it is visible in figure (b), the step width plays a significant role for a successful search.

of the output node is propagated backwards providing the local errors, which are then minimised by adjusting the weights.

Within the NeuroBayes package the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, another method for the minimisation problem, is additionally available. This method implements a sophisticated line search algorithm to find the best step width for each iteration step.

### 5.1.3 Variable Preprocessing

The preprocessing procedure has a wide influence and is applied before the beginning of the training. For this purpose all variables are prepared in a way that they are easier to handle for the network, as it can be very useful for correlated input variables. Especially NeuroBayes provides an extensive preprocessing.

For all variables a global preprocessing is performed. Accordingly, they are equalised by flattening and scaling their distributions between -1 and 1, so that the influence of outliers diminishes. These flattened distributions are converted into Gaussians with a mean of zero and a standard deviation of one. By this transformation, a saturation of the nodes is avoided as all variables lie in the central region of the sigmoid activation function, as visible in figure 5.2. Optionally, an individual variable preprocessing is available to alter the global one. So it is for example possible to replace the transformation to a Gaussian distribution by the application of a spline fit to the flattened distribution and, furthermore, it is also possible to handle variables, which are only present for a subset of events. In those events, where the variable is undefined, the corresponding input is not used for the training.

The further procedure for the input variables implements a decorrelation of all variables from each other. For this purpose the covariance matrix has to be calculated for the preprocessed variables and, to find the decorrelated ones, the matrix has to be diagonalised wherefore Jacobi rotations are used in this case.

For MLP the preprocessing is more compact. There is the option to normalise all input variables, so that they lie within the interval  $[0,1]$  as well as to apply a transformation to eliminate linear correlations for Gaussian variables. Therefore, the symmetric covariance matrix, i.e. it is the squared one, is diagonalised and the square root of this matrix is computed. Besides, it is possible to apply the principal component analysis (PCA) which is not very different from the described linear decorrelation.

### 5.1.4 Variable Ranking

For both NNs a variable ranking is automatically accomplished. NeuroBayes ranks the variables on the base of their significance of their correlation to target. Their significance is calculated in an iterative way. First, the correlation matrix and the total correlation of all variables to the target is calculated and all variables are one after the other removed and the correlation to target is calculated again. The variable causing the smallest loss of correlation by its exclusion is discarded and listed as the least significant one. The whole procedure is repeated with the remaining  $N - 1$  variables and so on until, finally, the most significant variable remains. For the ranking the significance of each variable is calculated by multiplying the corresponding loss of correlation to target with the square root of the training sample size. Additionally, it is possible to cut on the determined significance to retain only the most important variables for the training. This is the pruning method, available in NeuroBayes.

The variable ranking implemented by the MLP NN is based on the sum of the squared weights of the connections between the corresponding variable input node  $i$  and the hidden nodes  $j$ . This total weight for the variable is multiplied with its squared sample mean  $\bar{x}_i$  to determine its importance  $I_i$ :

$$I_i = \bar{x}_i^2 \sum_{j=1}^H w_{ij}^2. \quad (5.8)$$

### 5.1.5 Training Result

For a network with only one output node, as it is used for the classification into signal or background events in this analysis, the network output after a successful training should offer a good separation of both classes. An example for the distribution is shown in figure 5.4. In this case signal events pile up near  $+1$  while the background close to  $-1$ .

An information about the performance of a NN delivers its purity. It is defined as the ratio between the number of selected signal events and all selected events lying in each output bin. For a well trained network the purity should be proportional to the output.

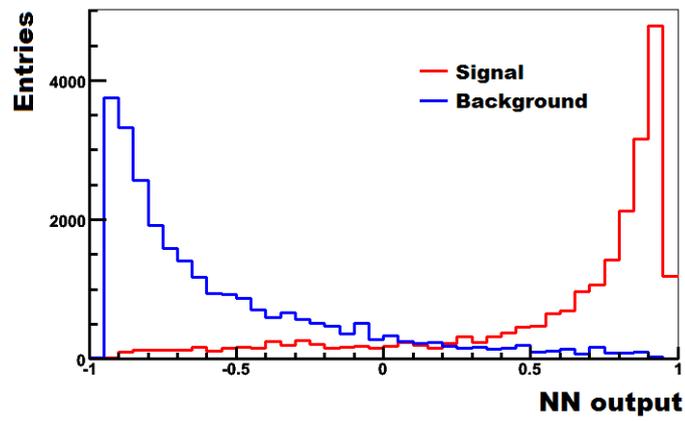


Figure 5.4: Illustration of a neural network output distribution for signal and background events.

# Chapter 6

## Analysis with Neural Networks

The aim of this analysis is to find an efficient method to distinguish signal from background events. Therefore, different neural networks are trained with Monte Carlo generated events and used for the event classification. All generated events, which have passed the preselection criteria (see table 4.1) are used for the training and are weighted in such a way that the training sample consists of one half signal and one half background events. The background sample itself is additionally weighted that the composition corresponds with the expected background as described in section 4.2. Tests have been accomplished with different signal to background ratios and also with diverse background compositions to improve the training, but this division delivered the best results.

### 6.1 Input Variables

To obtain a good separation of Higgs signal and background events, discriminating physical variables have to be delivered to the network as input, ideally, with each variable providing a different distribution for the two kind of classes, which leads, consequently, to new information for the network. The good separation power results from the combination of the input variables. Due to the distinct signal signature with two forward jets and two isolated charged leptons, kinematic variables of these physical objects are used as input for the network training.

#### Jet Variables

- $P_{Tj1}$ , the transverse momentum of the jet with highest energy. This as well as the following  $P_T$  variables for the jets and especially their invariant mass distribution show slightly larger values for Higgs-signal events than for all backgrounds.
- $P_{Tj2}$ , the transverse momentum of the jet with second highest energy.
- $M_{j1+j2}$ , the invariant mass of the two jets. The momenta of the jets are added before the invariant mass is determined.

- $P_{T,j1+j2}$ , the transverse momentum of the two jets. The calculation is carried out as described for the invariant mass.
- $\cos(\theta_{j1,j2})$ , where  $\theta$  is the polar angle between the two jets. In the case of signal-like jets, they point in the opposite direction, leading to an expected peak at -1 in the distribution of  $\cos(\theta)$ .
- $|\Delta\eta(j1, j2)|$ , with the pseudorapidities  $\eta$  of the two jets. Only if the two jets have nearly opposite pseudorapidities and are forward directed, large values are reached in the distribution as for the signal events. This variable is correlated to  $\cos(\theta_{j1,j2})$ .

### Lepton Variables

- $|\Delta\eta(l1, l2)|$ , where the pseudorapidities  $\eta$  of the two charged leptons are used. As the two leptons should point into the same direction for signal-like events, a peak around 0 is expected for the distribution. A more detailed explanation is given in section 3.2.
- $|\Delta\phi(l1, l2)|$ , with the azimuthal angle  $\phi$  of the two charged leptons. As explained for  $|\Delta\eta(l1, l2)|$ , the distribution should peak around 0 for signal-like events.
- $M_{l1+l2}$ , the invariant mass of the two charged leptons. This variable discriminates especially  $Z$ +jets events, if both charged leptons originate from the  $Z$  boson decay, as described in section 4.1.
- $|\Delta\phi(l1 + l2, \text{MET})|$ , with the azimuthal angle  $\phi$  of the summed charged lepton vectors and the missing transverse energy (MET) direction.

Further variables such as the transverse momentum distribution of  $\vec{E}_T$  or the two leptons have been tested and also different variable combinations have been taken into account, but finally, it has turned out, that too many highly correlated variables with only low discrimination power make the neural network even worse. So only the presented variables are used for the training.

After the event selection at least two isolated charged leptons and two up to five jets are required. There are only rare signal events with more than two present leptons, so that always the two leptons with highest transverse momentum are assumed to be the correct ones. This leads to a possible error due to the lepton selection in the case of less than 0.5% of the signal events. In contrast, more than a quarter of the 3032 selected signal events contains between three and five jets. For these events it is unknown, which jets correspond to the scattered quarks of the VBF process. Thus, all samples are separated depending on their jet number and independent neural networks are trained for events with exactly two jets and for those with three to five jets.

## 6.2 Neural Network Training

As already mentioned in chapter 5, two different neural network packages, NeuroBayes and MLP, are used for this analysis. For all neural networks the same topology has been chosen with one node for each variable and additionally one bias node in the input layer, one node less in the hidden one and one single output node in the final layer, since the network is applied to an event classification. Also trainings with more or less nodes in the hidden layer have been tested, but the training result was very stable under this modification. An important adjustment for the neural network is the modulation of the step width for the training, for which a smaller value has to be chosen in the case of a small training sample and the other way round, but it should be readjusted individually for every new training.

### 6.2.1 Events with Two Jets

Altogether 2208 signal and 5091 background events, weighted as explained before, are used for the training. The expected fractions of the different samples for the events containing exactly two jets are shown in figure 6.1. For these events the allocation of the physical objects is well-defined to form the corresponding variables, as there are only two leptons as well as two jets present in these events. So, all introduced variables serve as input for the neural network training. The shapes of all variables are shown in figure 6.2 and 6.3. It becomes apparent that in figure ?? the variables  $\cos(\theta)$  in (c),  $\Delta\eta$  in (d) and  $M_{j_1+j_2}$  in (e) show a good separation between signal and background events, while the other three variables provide only less discrimination power. For the lepton variables in figure 6.3 all variable distributions feature a visible separation, but none of them outperforms the others clearly.

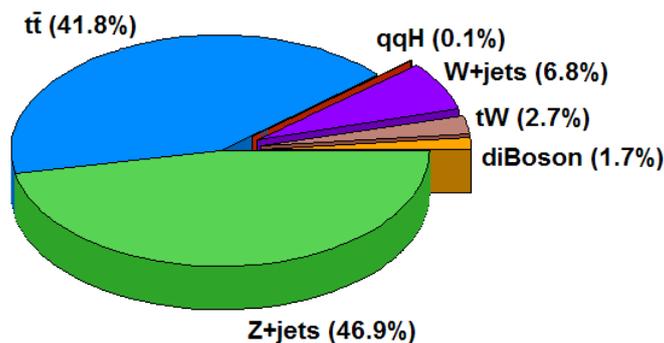
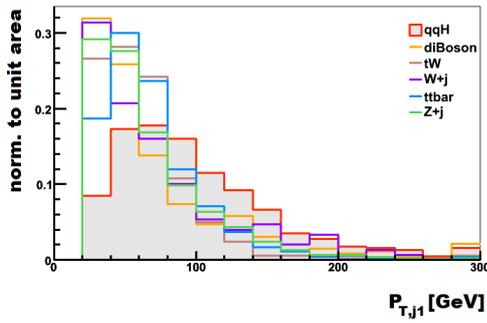
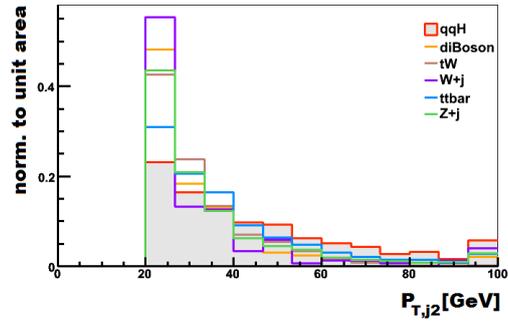
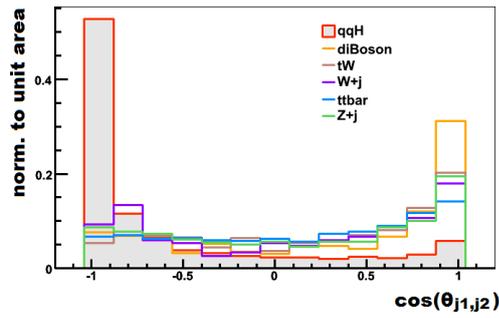
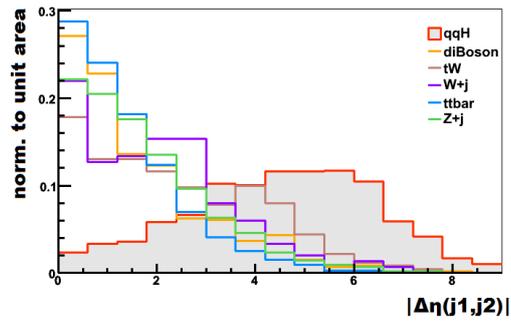
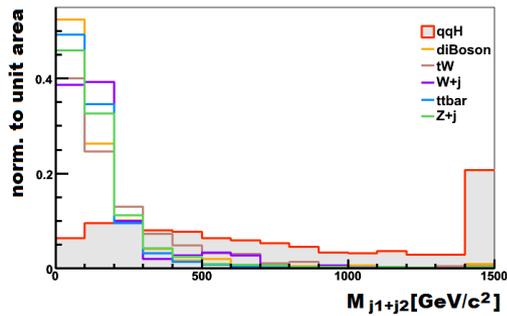


Figure 6.1: Expected fractions of signal and background processes for events containing exactly two jets.

(a)  $P_T$  of jet 1.(b)  $P_T$  of jet 2.(c)  $\cos(\theta)$  between jet 1 and jet 2.(d)  $\Delta\eta$  between jet 1 and jet 2.

(e) Invariant mass of jet 1+jet 2.

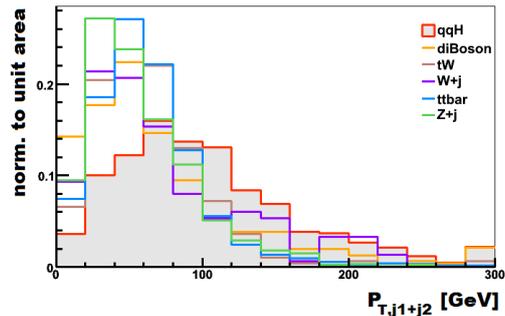
(f)  $P_T$  of jet 1+jet 2.

Figure 6.2: Shapes of jet variables for the two-jet network.

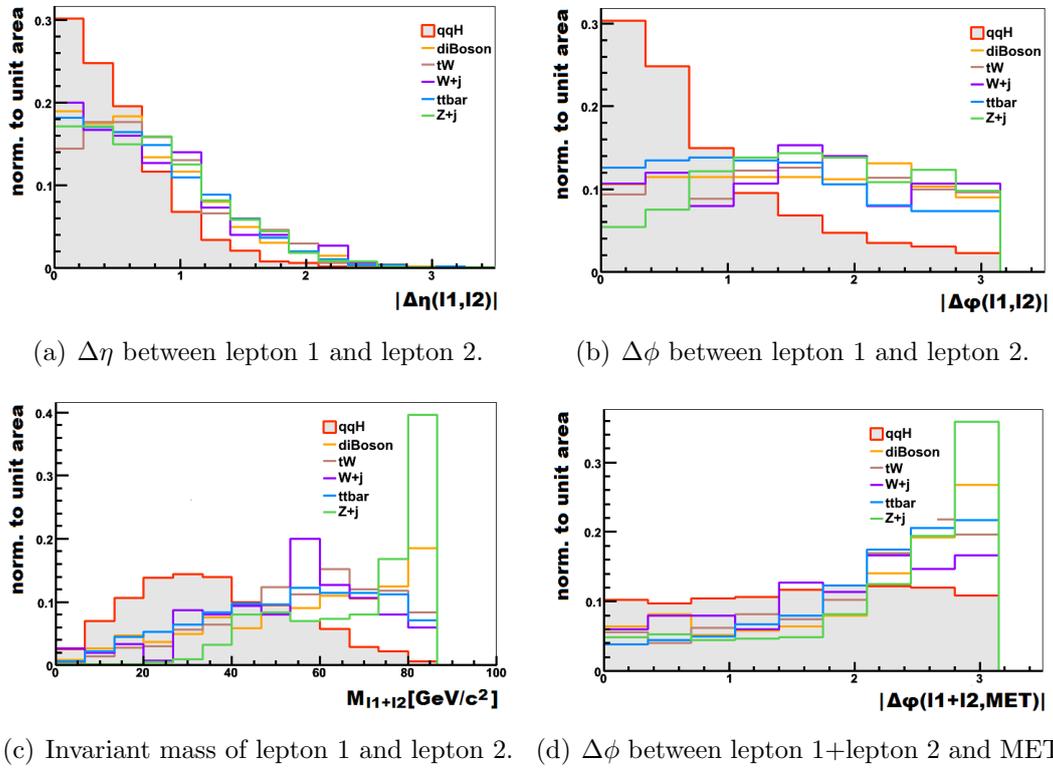


Figure 6.3: Shapes of lepton variables for the two-jet network.

Both neural networks establish an automatic variable ranking using the significance or importance, for NeuroBayes or MLP, respectively, as described in section 5.1. The result for the two-jet network is shown in table 6.1 and 6.2. Since both network packages provide an individual variable preprocessing and rank the variables according to different parameters, it is comprehensible, that the ranking can show very different results. Especially in the case of strongly correlated variables, only one variable is really significant. This can be seen for example in the ranking of the two variables  $|\Delta\eta(j1, j2)|$  and  $\cos(\theta_{j1, j2})$ . Both variables perform a good discrimination between signal and background events, but only one variable is important after the ranking. Which one is finally selected to be more relevant depends on the preprocessing. While  $|\Delta\eta(j1, j2)|$  comes up to the third place and  $\cos(\theta_{j1, j2})$  only to the ninth for the NeuroBayes ranking, it is just the opposite for the MLP ranking, where the latter achieves the first rank and the former only the seventh. In general it seems, that the different transverse momentum distributions tend to have more importance for the MLP network than for the NeuroBayes one, while for the latter the invariant mass distributions play a major role.

rank	variable	sig in $\sigma$
1	$M_{j1+j2}$	47.01
2	$M_{l1+l2}$	20.37
3	$ \Delta\eta(j1, j2) $	6.22
4	$P_{T, j1+j2}$	5.30
5	$ \Delta\eta(l1, l2) $	1.24
6	$ \Delta\phi(l1 + l2, MET) $	1.21
7	$ \Delta\phi(l1, l2) $	1.62
8	$P_{T, j1}$	1.38
9	$\cos(\theta_{j1, j2})$	1.28
10	$P_{T, j2}$	0.72

Table 6.1: Variable ranking of the NeuroBayes two-jet network according to the calculated significance sig.

rank	variable	imp
1	$\cos(\theta_{j1, j2})$	7.427
2	$P_{T, j1+j2}$	3.862
3	$ \Delta\phi(l1, l2) $	3.732
4	$M_{j1+j2}$	0.862
5	$P_{T, j2}$	0.647
6	$ \Delta\eta(l1, l2) $	0.603
7	$ \Delta\eta(j1, j2) $	0.401
8	$ \Delta\phi(l1 + l2, MET) $	0.339
9	$P_{T, j1}$	0.142
10	$M_{l1+l2}$	0.005

Table 6.2: Variable ranking of the MLP two-jet network depending on the calculated importance imp of each variable.

### 6.2.2 Three-to-Five-Jet Events

Figure 6.4 presents the expected fractions of the different samples for all events containing three to five jets. The background is dominated by the  $t\bar{t}$  process. Although only two jets are demanded for the signal process due to the two scattered quarks of the VBF process, additional jets can appear as a result of hard gluon radiation. Since rather high energetic jets are characteristic, only the three most energetic jets are considered for those events with four or five jets, whose fraction amounts to  $\sim 21\%$  and  $\sim 5\%$ , respectively, of all three to five jet events. Altogether, there are several possibilities for the jet allocation to the corresponding variables for this event type, as only two of the three available jets are needed to build the jet variables for the neural network training. Finally, a method has to be found to choose the best

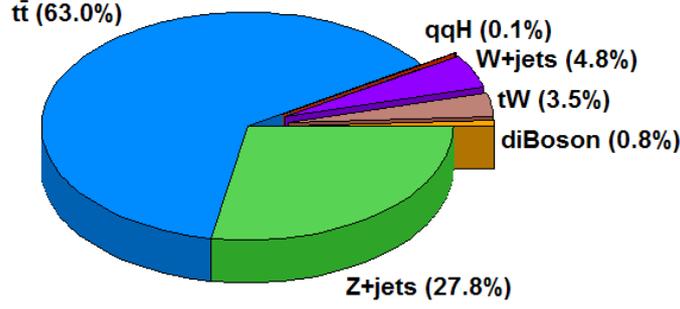


Figure 6.4: Expected fractions of signal and background processes for events containing three to five jets.

fitting jet combination out of the three possible candidates. Therefore, a specific neural network is trained in a first step.

For this purpose the Monte Carlo information of the Higgs signal events is used and only those events are chosen for the network training, for which both outgoing quarks can be matched with two of the three possible reconstructed jets. Figure 6.5 illustrates two event displays, containing the generated and reconstructed physical objects. One event (left) features two matched jets, the other event (right) has only one matched jet. The two leptons are mostly very well reconstructed.

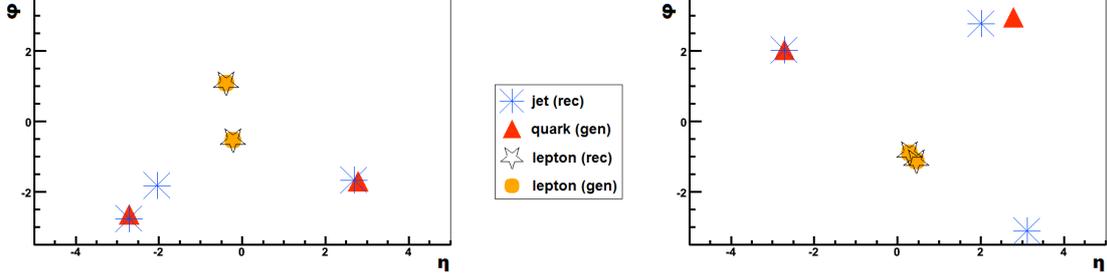
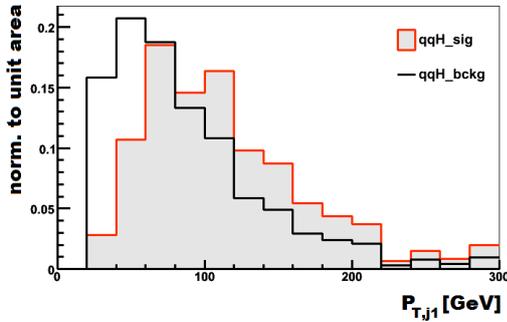
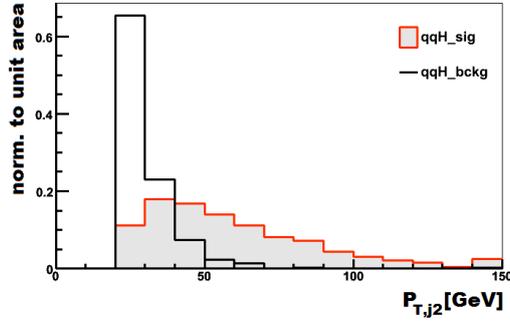
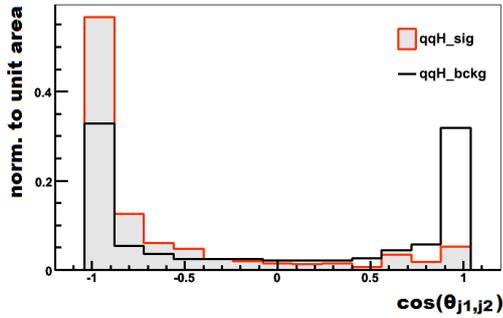
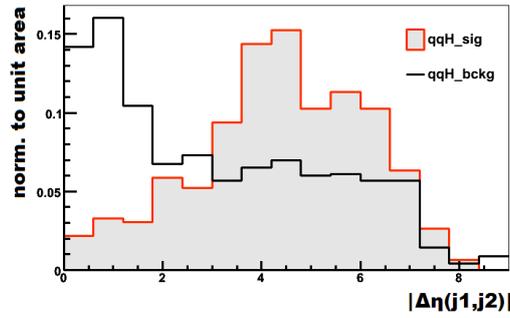
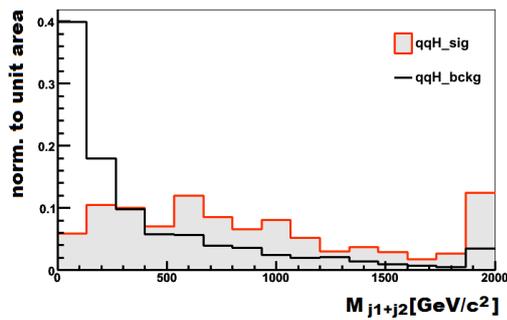


Figure 6.5: Position of the generated (gen) and reconstructed (rec) four-vectors of the two leptons, the generated outgoing quarks and the three reconstructed jets in the  $\eta$ - $\phi$  plane. On the left an event with two matched jets is shown, while on the right hand side only one jet is matched.

For the matching, a cone with  $R = 0.5$  is placed around the quark and a jet within this cone is identified as matched jet. Only these two matched jets provide the real signal signature. The other two wrong jet combinations are used as background events for the training. Thus, there are twice as much background combinations as signal ones. The background is correspondingly weighted down. After this matching criterion only 459 out of 824 signal events are suitable for the training. The resulting variable distributions are shown in figure 6.8. Also for this network the distributions of  $\cos(\theta)$  and  $\Delta\eta$  provide a strong discrimination power. Additionally, the transverse momentum distribution of the second jet, shown in figure (b), becomes more important for the jet selection. Overall, in many events, the two jets with highest transverse momentum can be identified with the two matched jets.

(a)  $P_T$  of jet 1.(b)  $P_T$  of jet 2.(c)  $\cos(\theta)$  between jet 1 and jet 2.(d)  $\Delta\eta$  between jet 1 and jet 2.

(e) Invariant mass of jet 1+jet 2.

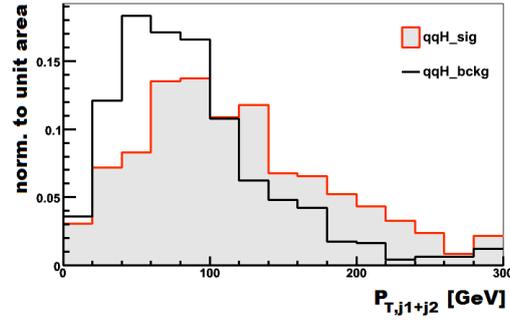
(f)  $P_T$  of jet 1+jet 2.

Figure 6.6: Shapes of variables for the jet selection for signal events containing three to five jets. The signal distribution stems from the two matched jets, while the background comes about the two remaining wrong combinations. For events with four or five jets, the three most energetic jets have been chosen.

The ranking for both neural network packages is presented in table 6.3 and 6.4. For these two networks the ranking result is quite similar. Only the variable  $P_{T,j1+j2}$  is considerably more important for the MLP network. As the last two variables provide a significance lower than  $0.5 \sigma$  after the NeuroBayes preprocessing, only the first four input variables are taken into account for the NeuroBayes training.

rank	variable	sig in $\sigma$
1	$P_{T,j2}$	18.92
2	$\cos(\theta_{j1,j2})$	9.31
3	$M_{j1+j2}$	2.77
4	$ \Delta\eta(j1, j2) $	1.09
5	$P_{T,j1+j2}$	0.44
6	$P_{T,j1}$	0.05

Table 6.3: Variable ranking of the NeuroBayes network according to the calculated significance sig. The network is used for the jet selection.

rank	variable	imp
1	$\cos(\theta_{j1,j2})$	1.483
2	$P_{T,j1+j2}$	0.866
3	$P_{T,j2}$	0.511
4	$ \Delta\eta(j1, j2) $	0.172
5	$M_{j1+j2}$	0.069
6	$P_{T,j1}$	0.009

Table 6.4: Variable ranking of the MLP network according to the importance imp of each variable. The network is trained for the jet selection.

In general, it is worth mentioning, that a powerful tool has been developed for the jet selection. Despite that there are only  $\sim 450$  available signal events for the training, the NeuroBayes network chooses in over 87% of the cases the correct jet combination, while the MLP net in even more than 90%.

In a second step all signal as well as background events are passed through the trained net. So, for each event all three jet combinations are tried and induce a network output. The jet combination with the maximum network output value is used for the second neural network, processing the final event classification. The resulting distributions are shown in figure 6.7.

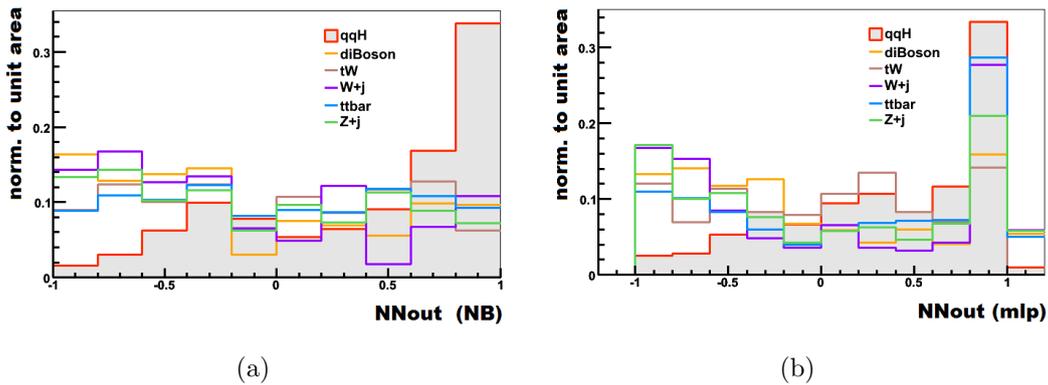


Figure 6.7: Shapes of the first network output used for the jet selection of NeuroBayes (a) and MLP (b).

Although the developed network delivers a satisfying result for the used signal events with two matched jets, there is still a quite big fraction of Higgs events which is classified as background-like, as can be seen in the shape of the network

output. This originates from the large fraction of events with only an incomplete jet matching, since not both jets feature the characteristics of the scattered quarks of the VBF process. In general, the NeuroBayes network seems to cope better with the backgrounds than the MLP network.

The second network training, used for the final Higgs signal event selection, is comparable to the one used for the two-jet events, since only the two jets, selected by the first neural network, are considered. Altogether 824 signal and 2571 background events are available for the second neural network training. The shapes of the corresponding variables are presented in figure 6.8 and 6.9. It is visible, that the two chosen jets, building the jet variables in figure 6.8, still provide a good discrimination power between signal and background events. However, compared to those jet variables used for the two-jet net (see figure 6.2), they seem to lose a little power for the separation of the two event classes. By contrast, these distributions of the lepton variables in figure 6.9 correspond to the ones of the two-jet network, shown in figure 6.3, as expected, since the jet number of an event should not affect the lepton distributions.

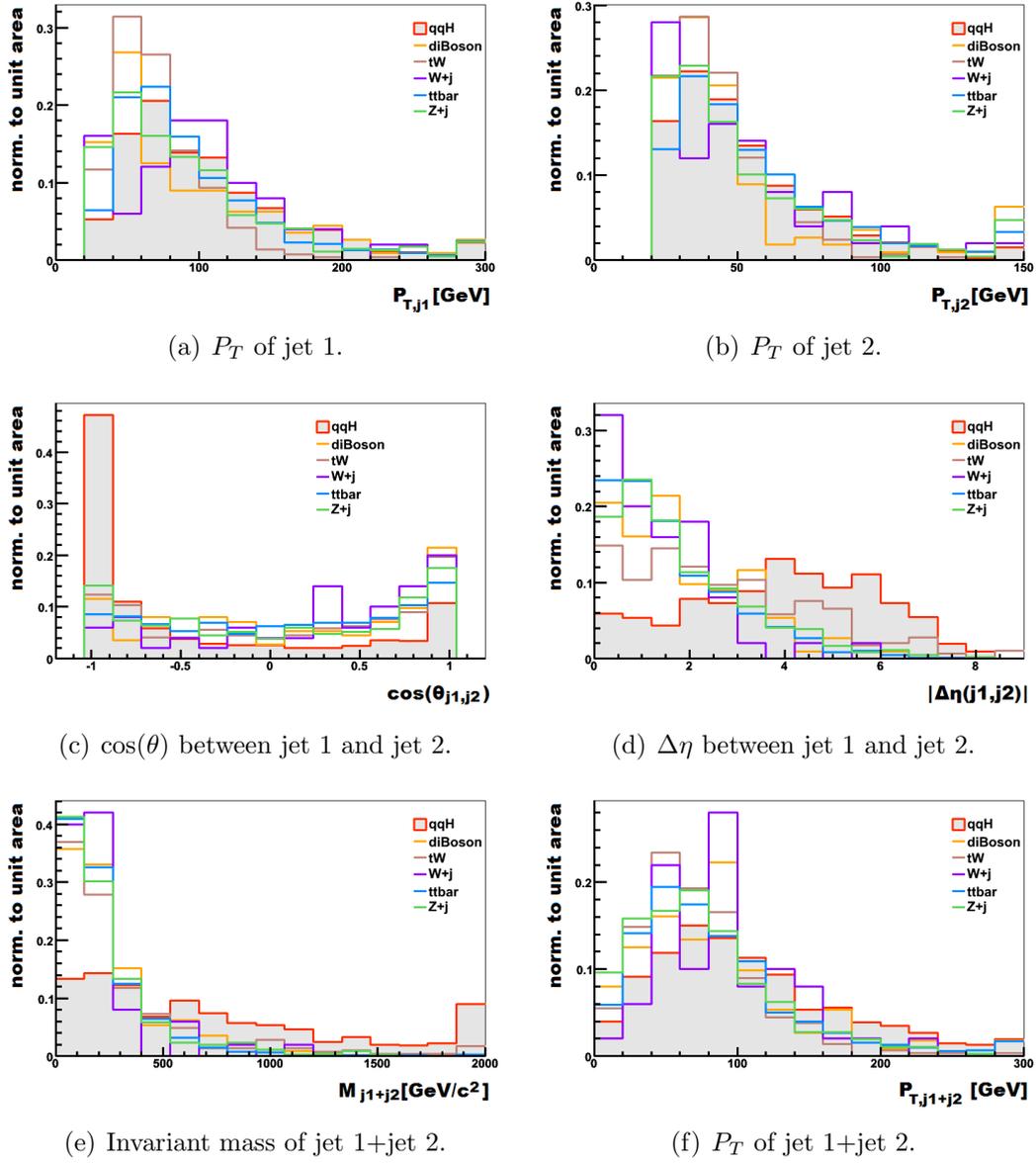
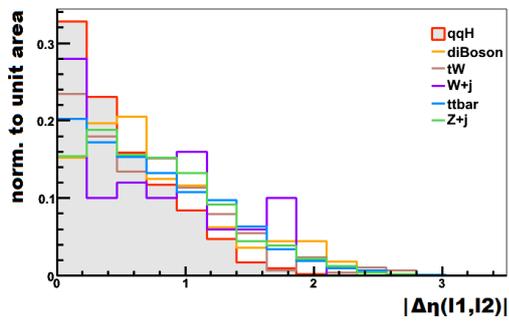
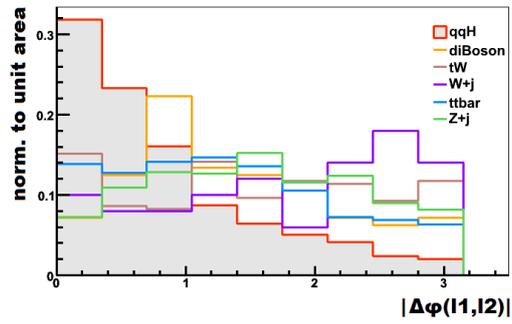
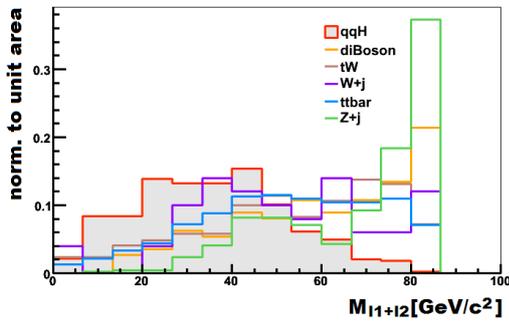


Figure 6.8: Shapes of jet variables for the three-to-five-jet network.

(a)  $\Delta\eta$  between lepton 1 and lepton 2.(b)  $\Delta\phi$  between lepton 1 and lepton 2.

(c) Invariant mass of lepton 1 and lepton 2.

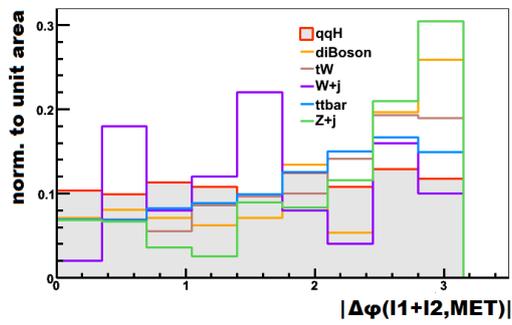
(d)  $\Delta\phi$  between lepton 1+lepton 2 and MET.

Figure 6.9: Shapes of lepton variables for the three-to-five-jet network.

The ranking is given in table 6.5 and 6.6. These results comply essentially with those of the two-jet net, but there is a trend of the variables to have lower significance or importance. For the NeuroBayes network only the eight most significant variables are used for the training, since the remaining two variables provide a significance lower than  $0.5 \sigma$ .

rank	variable	sig in $\sigma$
1	$ \Delta\eta(j1, j2) $	24.58
2	$M_{l1+l2}$	14.01
3	$P_{T,j1+j2}$	4.96
4	$ \Delta\phi(l1, l2) $	2.23
5	$M_{j1+j2}$	1.78
6	$ \Delta\eta(l1, l2) $	1.72
7	$P_{T,j2}$	1.64
8	$ \Delta\phi(l1 + l2, MET) $	0.91
9	$\cos(\theta_{j1,j2})$	0.46
10	$P_{T,j1}$	0.13

Table 6.5: Variable ranking of the NeuroBayes network according to the calculated significance sig. The network is used for the final selection.

rank	variable	imp
1	$P_{T,j1+j2}$	3.958
2	$\cos(\theta_{j1,j2})$	2.187
3	$ \Delta\phi(l1, l2) $	1.523
4	$ \Delta\eta(l1, l2) $	1.497
5	$P_{T,j2}$	0.588
6	$M_{j1+j2}$	0.419
7	$ \Delta\phi(l1 + l2, MET) $	0.058
8	$ \Delta\eta(j1, j2) $	0.035
9	$P_{T,j1}$	0.010
10	$M_{l1+l2}$	0.001

Table 6.6: Variable ranking of the MLP network according to the calculated importance imp. The network is used for the final event selection.

### 6.3 Comparison of the Network Packages

To compare each of the three trained neural networks with the corresponding one of the other network package, the background rejection, defined as  $1 - (\text{background efficiency})$ , versus signal efficiency is plotted and shown in figure 6.10. It is visible, that for the two-jet network in (a) the MLP net provides a slightly higher background rejection for high signal efficiencies than the NeuroBayes one, but this effect is not very distinct. The trained networks for those events containing three to five jets are comparable, the one for the jet selection as well as the one for the final signal event classification.

With regard to the analysis, both network packages provide a corresponding result for all different networks, since the region of high background rejection is focussed, where both curves are on top of each other.

However, there is a great advantage in the expenditure of time for the NeuroBayes network training. Due to the highly developed preprocessing leading to well accessible information by the input variables, the NeuroBayes net needs only few iterations for the whole network training (for example  $\sim 20$  for the two-jet network with 10 input variables), while the MLP network needs at least  $\sim 500$  iterations to come to a satisfying result. This is of course reflected in the duration of the training. For an assumed training with a total of 3283 events and five input variables, NeuroBayes provides a training which is already  $\sim 12$  times faster than the one by MLP. With an increase of training events and input variables the effect even becomes a

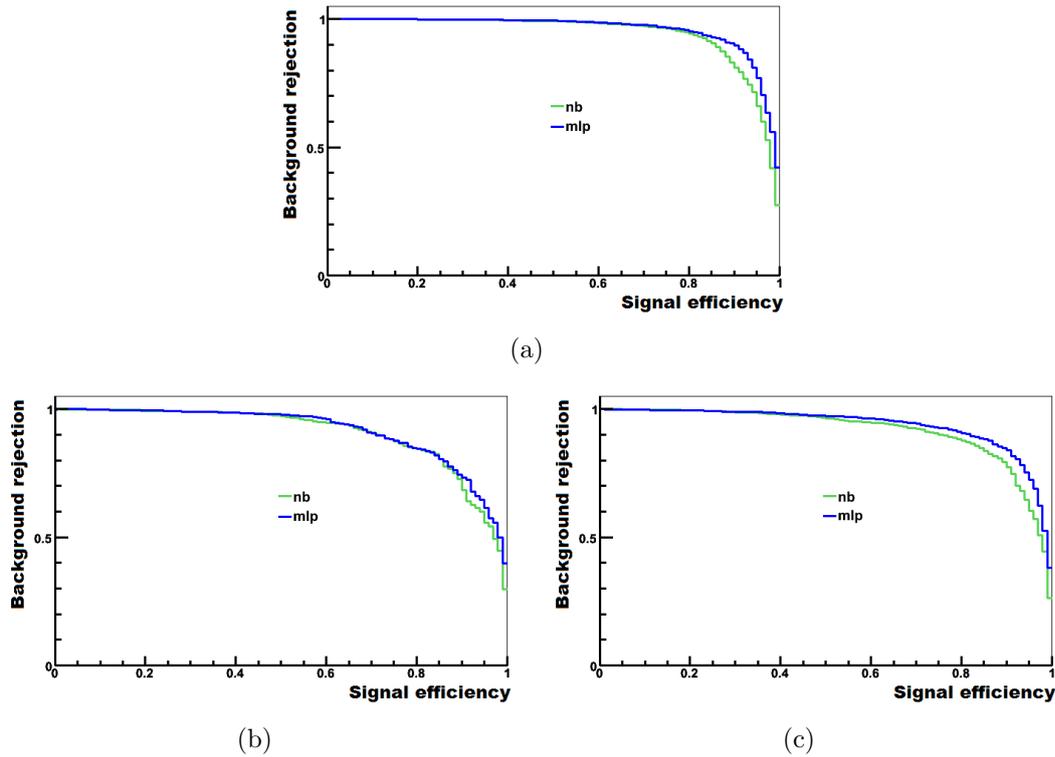


Figure 6.10: Background rejection versus signal efficiency for the trained neural networks with the distribution of the two-jet network in (a), the one of the three-to-five-jet net for the jet selection in (b) and the for the final event selection in (c).

lot larger, so that for the classification of the two-jet events, trained with altogether 7299 events and ten input variables, the NeuroBayes training is even  $\sim 33$  times faster with nearly the same result. This item is especially important for analyses with many variables and facilitates the adjustment of the network configurations a lot.

## 6.4 Network Output

The shapes for all processes of the final network outputs of the two-jet events and three-to-five-jet events are shown in figure 6.11 and 6.12, respectively. In figure (a) the results from NeuroBayes are presented, while in (b) those from MLP are given. As expected, the shape for the two-jet net delivers a very good separation between signal and background events, while the other one provides a slightly worse classification result. Due to very low statistics of these Monte Carlo events, the distributions of some background processes are not as smooth as for the two-jet network. An improvement of the three-to-five-jet network result could may be achieved with better statistics. In summary, good discrimination power between signal and background processes is achieved by both neural network packages for both events types.

The result becomes even more accessible, if the network output is weighted in

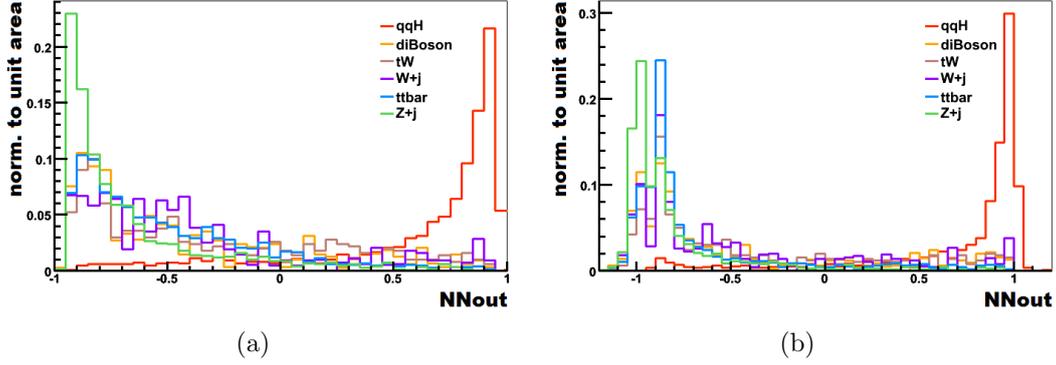


Figure 6.11: Shapes of the network output of the two-jet net in (a) and (b) with the NeuroBayes and MLP output, respectively.

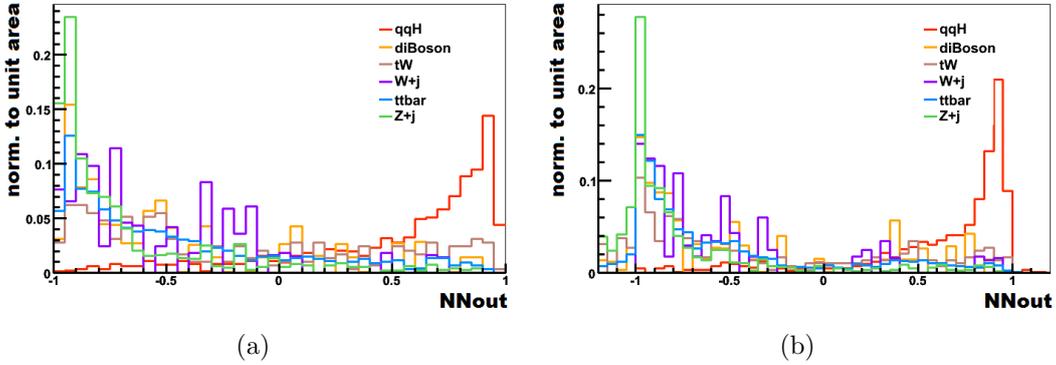


Figure 6.12: Shapes of the network output of the three-to-five-jet net with the NeuroBayes output in (a) and the MLP one presented in (b).

such a way, that the event number of the different processes corresponds to the expectation value, shown in figure 6.13 and 6.14 depending on the jet number of the event. In figure (c) and (d), respectively, only the last bins are shown to enable a closer look to the signal region. For the two-jet network the large number of  $W$ +jets events in the signal region is very conspicuous, although the expected fraction is only about 6.8% of all expected events of this type. It could be possible, that a boosting of the network to the  $W$ +jets background improves the result. For the latter net, the signal events are barely visible as only  $\sim 13$  events at all are expected for an integrated luminosity of  $\mathcal{L} = 10 \text{ fb}^{-1}$  for this event type, while for the two-jet net  $\sim 34$  events should be observable.

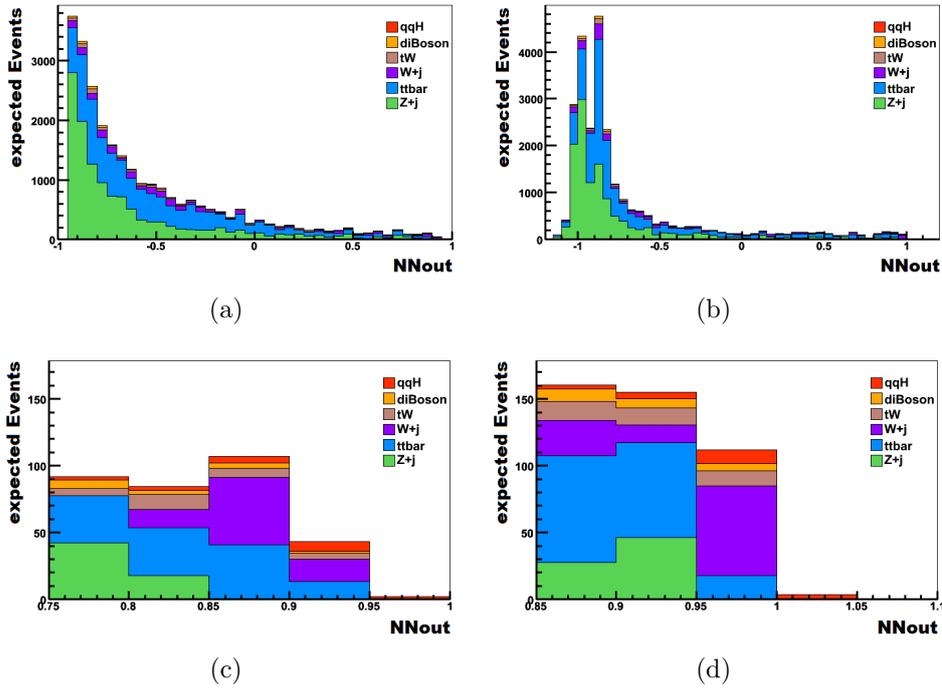


Figure 6.13: Expected events for the two-jet network output for NeuroBayes (a) and MLP (b) for an integrated luminosity of  $\mathcal{L} = 10 \text{ fb}^{-1}$ . In (c) and (d) only the distributions of the last bins are shown.

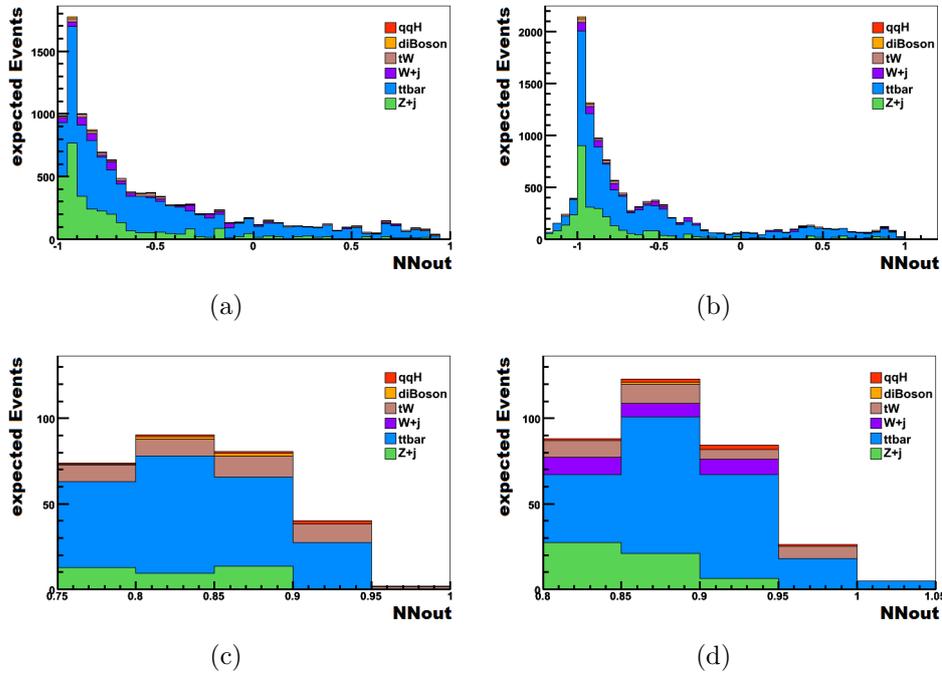


Figure 6.14: Expected events for the three-to-five-jet network output for NeuroBayes in (a) and MLP in (b) for an integrated luminosity of  $\mathcal{L} = 10 \text{ fb}^{-1}$ . The last bins of the distributions are shown in (c) and (d).

## 6.5 Expected Significance

With these obtained distributions the significance  $\sigma$ , defined as  $s/\sqrt{b}$ , can be calculated for each bin  $x$  of the network output distribution, where  $s$  is the number of signal and  $b$  the total number of background events. Therefore, all events are selected in the interval  $[x, x_{max}]$ .

To obtain additionally the statistical error due to the used Monte Carlo samples, first the absolute error for all samples has to be estimated for each bin  $x$  of the neural network output. Therefore, the number of events  $e_x$  in the interval  $[x, x_{max}]$  is determined for the unweighted sample and provide the corresponding absolute error with  $\sqrt{e_x}$ . Since all samples are weighted for the significance calculation, the absolute error  $\Delta e_{x,w}$  for the weighted sample for each bin  $x$  is estimated with the product of the expected number of events  $e_{x,w}$  (after the weighting) in the interval  $[x, x_{max}]$  with the relative error due to the unmodified sample:

$$\Delta e_{x,w} = e_{x,w} \frac{\sqrt{e_x}}{e_x}. \quad (6.1)$$

The resulting errors for the total background as well as the significance are approximately calculated using the Gaussian error propagation formula

$$\Delta y(x_1, x_2, \dots) = \sqrt{\left(\frac{\partial y}{\partial x_1} \Delta x_1\right)^2 + \left(\frac{\partial y}{\partial x_2} \Delta x_2\right)^2 + \dots} \quad (6.2)$$

for uncorrelated errors  $\Delta x_i$ . These determined errors are additionally plotted in the distributions for the significance, shown in figure 6.15. The exact values for the last bins of the two-jet net distributions shown in figure 6.13 are listed in table 6.7 and 6.8. As visible, comparable results are obtained with both networks.

bin	exp. signal	exp. background	$s/\sqrt{b}$	error
0.75	19.6	323.3	1.09	0.087
0.80	17.4	227.4	1.15	0.111
0.85	14.1	143.5	1.18	0.149
0.90	9.2	37.2	1.51	0.385
0.95	1.8	0	-	-

Table 6.7: Expected events for  $10 \text{ fb}^{-1}$  and significance for the last bins of the NeuroBayes network for events containing two jets.

bin	exp. signal	exp. background	$s/\sqrt{b}$	error
0.85	21.6	430.7	1.04	0.072
0.90	18.8	264.0	1.16	0.103
0.95	13.7	107.2	1.32	0.206
1.00	3.5	0	-	-

Table 6.8: Expected events for  $10 \text{ fb}^{-1}$  and significance for the last bins of the MLP network for events containing two jets.

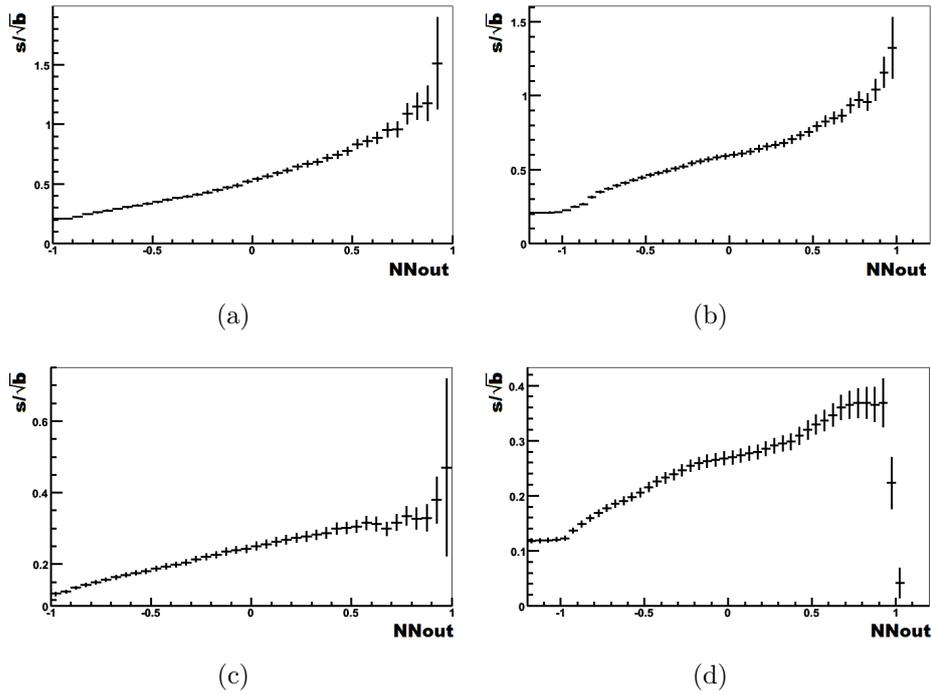


Figure 6.15: Significance, calculated as  $s/\sqrt{b}$ , for the two-jet network for NeuroBayes (a) and MLP (b) and for the final network for events containing three to five jets in (c) and (d) for NeuroBayes and MLP, respectively.

## 6.6 Comparison to a Recent Analysis

In reference [55] a cut-based analysis is presented for the same Higgs boson channel as studied in this thesis, but with an assumed Higgs boson mass of  $160 \text{ GeV}/c^2$ . Thus, more than twice the number of signal events are expected due to the high branching ratio of the Higgs boson into a  $W$  boson, namely  $\sim 0.9$  in comparison to  $\sim 0.3$  for a Higgs mass of  $130 \text{ GeV}/c^2$ . For the analysis three different channels are distinguished depending on the lepton flavour of the two involved, charged leptons and the significance, also calculated with  $s/\sqrt{b}$ , is determined for each channel for  $10 \text{ fb}^{-1}$ . The results are summarised in table 6.9.

channel	exp. signal	exp. background	$s/\sqrt{b}$
$ee$	6.5	113.0	0.61
$e\mu$	6.1	15.6	1.55
$\mu\mu$	7.5	14.9	1.95

Table 6.9: Expected events for  $10 \text{ fb}^{-1}$  and significance for the cut-based analysis depending on the lepton flavour of the two charged leptons.

With the presented neural network analysis a significance of  $\sim 1.3$  or  $\sim 1.5 \sigma$  for the two-jet network is reached depending on the used network package. This is a comparable result although a lower Higgs boson mass is considered. So, an extension to higher Higgs boson masses would be very interesting and would allow for a direct statement of the benefit of the developed method.



# Chapter 7

## Conclusion

### 7.1 Summary

In this thesis studies on the search for a Standard-Model-like Higgs boson with the CMS experiment are presented, where the Higgs boson mass is assumed to be  $130 \text{ GeV}/c^2$ . For the production the weak vector boson fusion (VBF) mechanism is examined, while the decay of the Higgs boson into a  $W$  boson pair with both  $W$  bosons decaying leptonically is considered. The cross section of this channel for the considered mass is only  $0.12 \text{ pb}$ , very low compared to the background processes, which have much larger cross sections. So the main challenge for an analysis is the improvement of the background reduction, while keeping the signal as high as possible. This is achieved in the presented analysis in two steps: first, several cuts are applied, followed by a neural network analysis.

The whole analysis is based on Monte Carlo generated events, which are passed to a full CMS detector simulation. To obtain preselected samples of candidate events of the signal as well as background processes, the event signature of the selected Higgs boson process is studied using pure Monte Carlo generated events. As a consequence, a potential candidate has to feature two isolated charged leptons and missing transverse energy originating from the two leptonically decaying  $W$  bosons as well as two forward jets due to the VBF Higgs production. This signal event signature is exploited by the application of several cuts on the fully simulated events. In this first step only those events pass the preselection criteria, which provide the demanded signature and comprise only well defined physical objects. The expected numbers of signal and background events are estimated for an integrated luminosity of  $10 \text{ fb}^{-1}$ . The fraction of expected signal events constitutes  $\sim 0.1\%$ , so that a more powerful discrimination method is needed.

In a second step, two different neural network packages are used and compared with each other for the final separation between signal and background events. In the subsequent analysis, all selected events are divided into two subsamples according to their jet number with events containing exactly two and those with more than two jets. On each of the obtained samples individual neural network trainings are applied, using several input variables with preferably large discrimination power between signal and background events. Therefore, ten variables which deliver for

example kinematic information about the charged leptons or jets are used. One of the most important input variables for the separation is the distribution of the angle between the two jets as well as the one between the two charged leptons.

Furthermore, within this analysis a method for those events which contain more than two jets had to be found for the best possible jet selection, since only two jets are demanded due to the event signature. The additional jets occur in consequence of hard gluon radiation in the initial or final state. For the jet selection the three jets with the highest transverse momentum are considered leading to three possible combinations, while only one combination corresponds to the two characteristic jets as a result of the VBF production. The correct jet combination can be found due to the Monte Carlo truth of each event by allocating the two scattered quarks to the corresponding reconstructed jets. So, an own neural network was trained, whose input variables are arranged by all possible jet combinations, leading to three possibilities for each variable. For the signal only the true combination is utilised, while for the background the two wrong remaining ones are used. This network is trained with Higgs boson events where a jet matching is possible, so that the correct jet allocation is known. The trained network shows a good performance. In 85-90% of the corresponding signal events a correct jet allocation could be achieved.

Finally, the expected significance  $\sigma$  of this analysis method is determined for an integrated luminosity of  $10 \text{ fb}^{-1}$ . A cut on the neural network output is applied and the number of signal events  $s$  as well as background events  $b$  is determined. The significance is calculated with  $s/\sqrt{b}$ . For events with exactly two jets about  $1.5 \sigma$  can be achieved with a cut at 0.9 of the neural network output. Then approximately 9 signal events and 37 background events are expected. For those events containing more than two jets a significance of only  $0.35 \sigma$  can be attained. Altogether, it will not be possible to measure a significant signal contribution in the first years of data taking.

## 7.2 Outlook

Although the presented analysis will not lead to an observation or exclusion of a Standard-Model-like Higgs boson during the first years after commissioning the Large Hadron Collider, anyhow, it will be interesting if the results are consistent with the Standard Model and if the Higgs mechanism with the predicted couplings is realised in nature. Furthermore, the VBF channel will supplement other channels and will contribute to the combination of all Higgs channels and thereby to the overall sensitivity of the CMS experiment. A significant result with  $4\text{-}5 \sigma$  could be obtained with an integrated luminosity of  $\sim 100 \text{ fb}^{-1}$  considering only the VBF channel.

To improve the significance a separation according to the type of the two involved charged leptons could be advantageous. Especially the decay of both  $W$  bosons into an isolated muon and the corresponding neutrino features a distinct signature, which is hardly to fake and should provide lower background levels. In this context, it could be profitable to study the isolation criteria for the electron selection once more to exclude faked electron candidates even more efficiently. Due to the limitation of

signal Monte Carlo samples, it has not been possible to perform a more specific analysis than presented, for example with the focus on the pure muon decay channel.

Furthermore, an extension of the analysis towards higher Higgs boson masses should also improve the result of the presented method, as the sensitivity is expected to be better. Not only the branching ratio into a pair of  $W$  bosons increases steeply with higher Higgs boson masses, but also the detection of the physical objects should change for the better, since the leptons have higher momenta. Finally, this leads to more expected signal events, too.

Aside from the final analysis, it would be interesting to improve the intermediate trained network for the vector boson fusion jet selection. Although the training is executed with very low statistics, it seems, that a powerful tool has been developed, which could be also useful for other VBF Higgs analyses. Therefore, VBF produced Higgs events with diverse decay channels could be used for the neural network training to increase the fraction of signal events containing more than two jets.

Altogether it is sure, that data taken by the Large Hadron Collider will certainly reveal the answer if the Higgs boson exists as predicted by the Standard Model of particle physics or if new physics can be observed.



# Bibliography

- [1] D. Griffiths, "Introduction to Elementary Particles", WILEY-VCH (2008).
- [2] W.-M. Yao *et al.*, "Review of Particle Physics", *J. Phys. G: Nucl. Part. Phys.* **33**, 1 (2006).
- [3] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19**, 1264 (1967).
- [4] S.L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* **22**, 579 (1961).
- [5] A. Djouadi, "The Anatomy of Electro-Weak Symmetry Breaking", hep-ph/0503172 (2005).
- [6] P.W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13**, 508 (1964).
- [7] P.W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.* **145**, 1156 (1966).
- [8] Parton Distribution Generator  
<http://durpdg.dur.ac.uk/hepdata/pdf3.html>
- [9] M. Spira *et al.*, "Electroweak Symmetry Breaking and Higgs Physics", hep-ph/9803257 (1998).
- [10] M. Spira, "Higgs branching ratios and widths"  
<http://diablo.phys.northwestern.edu/pc/brs.html>
- [11] M. Spira, Table of cross section and Branching Ratios at NLO  
<http://twiki.cern.ch/twiki/bin/view/CMS/SMHiggsBoson>
- [12] The LEP Electroweak Working Group  
<http://lepewwg.web.cern.ch>
- [13] ALEPH, DELPHI, L3, OPAL Collaborations and LEP Working Group for Higgs Boson Searches, "Search for the Standard Model Higgs boson at LEP", *Phys. Lett.* **565** (2003) 61, CERN-EP/2003-011.
- [14] The TEVNPH Working Group, "Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production at High Mass with  $3 \text{ fb}^{-1}$  of data", hep-ex/0808.0534-

- 
- [15] LHC Design Report  
<http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html>
- [16] L. Evans, P. Bryant, "LHC Machine", (2008) JINST **3** S08001.
- [17] M. Benedikt(ed.), "The PS Complex as Proton Pre-Injector for the LHC Design and Implementation Report", CERN 2000-03 (2000).
- [18] CERN Press Release  
<http://press.web.cern.ch/press/PressReleases/Releases2008/List.html>
- [19] CMS Collaboration, "The CMS experiment at the CERN LHC", (2008) JINST **3** S08004.
- [20] CMS public information Webpage  
<http://cms-project-cmsinfo.web.cern.ch/cms-project-cmsinfo/index.html>
- [21] Compact Muon Solenoid  
[http://en.wikipedia.org/wiki/Compact\\_Muon\\_Solenoid](http://en.wikipedia.org/wiki/Compact_Muon_Solenoid)
- [22] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 1", CERN/LHCC **2006-001** (2006). CMS TDR 8.1.
- [23] CMS Collaboration, "The Magnet Project Technical Design Report", CERN/LHCC **97-010** (1997). CMS TDR 1.
- [24] CMS Collaboration, "The Hadron Calorimeter Technical Design Report", CERN/LHCC **97-031** (1997). CMS TDR 2.
- [25] CMS Collaboration, "The Muon Project Technical Design Report", CERN/LHCC **97-32** (1997). CMS TDR 3.
- [26] CMS Collaboration, "The Electromagnetic Calorimeter Technical Design Report", CERN/LHCC **97-033** (1997). CMS TDR 4.
- [27] CMS Collaboration, "The Tracker Project Technical Design Report", CERN/LHCC **98-006** (1998). CMS TDR 5.
- [28] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems", CERN/LHCC **2000-38** (2000). CMS TDR 6.1.
- [29] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger," CERN/LHCC **2002-26** (2002). CMS TDR 6.2.
- [30] M. A. Dobbs, S. Frixione, E. Laenen, K. Tollefson, "Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics", hep-ph/0403045v2 (2004).

- [31] T. Sjöstrand, S. Mrenna, P. Skands, "PYTHIA 6.4 physics and manual", JHEP05 (2006). 026
- [32] Pythia Webpage  
<http://www.thep.lu.se/~torbjorn/Pythia.html>
- [33] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions", hep-ph/0206293v2 (2002).
- [34] S.R. Slabospitsky, L. Sonnenschein, "TopReX generator (version 3.25) short manual", hep-ph/0201292v1 (2002).
- [35] D. Chang, W.-Y. Keung, "CP Violation in the Decay of Neutral Higgs Boson into  $t - \bar{t}$  or  $W^+ - W^-$ ", hep-ph/9301265v1 (1993).
- [36] Geant4 Collaboration, S. Agostinelli et al., "Geant4 - a simulation toolkit", Nucl. Instrum. and Meth. **A 506** (2003) 250-303.
- [37] CMS Wiki Home Page  
<https://twiki.cern.ch/twiki/bin/view/CMS/WebHome>
- [38] PGS Webpage  
<http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs.html>
- [39] Electron Identification  
<https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideElectronID>
- [40] M. Mulders *et al.*, "Muon Identification in CMS", CMS NOTE **2008/098**.
- [41] P. Avery *et al.*, "Measurement of Missing Transverse Energy With the CMS Detector at the LHC", CMS NOTE **2006/035**.
- [42] A. Heister *et al.*, "Measurement of Jets with the CMS Detector at the LHC", CMS NOTE **2006/036**.
- [43] S. V. Chekanov, "Jet algorithms: a minireview", hep-ph/0211298v1 (2002).
- [44] CMS Collaboration, "Performance of Jet Algorithms in CMS", CMS NOTE **2008/001**.
- [45] C. Sander, "Performance of Jet Reconstruction"  
[www.desy.de/~csander/Talks/JetsCMS\\_080527.pdf](http://www.desy.de/~csander/Talks/JetsCMS_080527.pdf)
- [46] CMS Collaboration, "The CMS Physics Technical Design Report, Volume 2", CERN/LHCC **2006-021** (2006). CMS TDR 8.2
- [47] CMS Software Webpage  
[http://cms.cern.ch/iCMS/jsp/page.jsp?mode=cms\&action=url\&urlkey=CMS\\_OFFLINE](http://cms.cern.ch/iCMS/jsp/page.jsp?mode=cms\&action=url\&urlkey=CMS_OFFLINE)

- [48] TMVA Webpage  
<https://twiki.cern.ch/twiki/bin/view/TMVA/WebHome>
- [49] A. Höcker, P. Speckmayer, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, "TMVA Users Guide", physics/0703039 (2007).
- [50] M. Feindt, "A Neural Bayesian Estimator for Conditional Probability Densities", physics/0402093v1 (2004).
- [51] Gradient descent  
[http://en.wikipedia.org/wiki/Gradient\\_descent](http://en.wikipedia.org/wiki/Gradient_descent)
- [52] V. Blobel, E. Lohrmann, "Statistische und numerische Methoden der Datenanalyse", Teubner, (1998).
- [53] N. Kauer, T. Plehn, D. Rainwater and D. Zeppenfeld, " $H \rightarrow WW$  as the discovery mode for a light Higgs boson", hep-ph/0012351v1 (2000).
- [54] D. Rainwater and D. Zeppenfeld, "Observing  $H \rightarrow W^{(*)}W^{(*)} \rightarrow e^{\pm}\mu^{\mp}p_T$  in weak boson fusion with dual forward jet tagging at the CERN LHC", Phys. Rev. D, Volume 60, 113004 (1999).
- [55] C. Charlot *et al.*, "Search Strategy for a Standard Model Higgs Boson Decaying to Two W Bosons in the Fully Leptonic Final State", CMS AN-**2008/039** (2008).
- [56] T. Peiffer, "Search for the Higgs Boson in the WH Channel with the CDF II Experiment", IEKP-KA/2008-4 (2008).
- [57] G. Sartisohn, "Studies for the Measurement of the W Boson Helicity in Top Quark Decays with the CMS Experiment", IEKP-KA/2008-1 (2008).
- [58] Ph. Sturm, "Studies for the Measurement of Single-Top-Quark Events with the CMS Experiment", IEKP-KA/2008-8 (2008).
- [59] M. Renz, "B-Jet and C-Jet identification with neural networks as well as combination of multivariate analyses for the search for single top-quark production", IEKP-KA/2008-17 (2008).
- [60] E. Yazgan *et al.*, "Search for a Standard Model Higgs Boson in CMS via Vector Boson Fusion in the  $H \rightarrow WW \rightarrow l\nu l\nu$  Channel", CMS NOTE **2007/011** (2007).

# Danksagung

An aller erster Stelle möchte ich mich bei Herrn Prof. Dr. Wolfgang Wagner bedanken, für die Übernahme des Referats, die hervorragende Betreuung und für die vielen hilfreichen Verbesserungsvorschläge.

Des weiteren danke ich Herrn Prof. Dr. Thomas Müller für die Aufnahme in seine Top-Arbeitsgruppe und die Übernahme des Korreferats.

Ein ganz besonderer Dank gebührt Georg Sartisoehn für die Unterstützung und Hilfe mit einer immerwährenden Geduld, sowie das intensive Korrekturlesen dieser Arbeit.

Weiterhin sei Frau Dr. Jeannine Wagner-Kuhr für das aufmerksame Korrekturlesen und die detaillierten Verbesserungshinweise gedankt. Allen Mitgliedern der Top-Gruppe danke ich für die sehr angenehme Arbeitsatmosphäre und die vielen Diskussionen auch mit Themen jenseits der Physik. Philipp Sturm möchte ich auch für das Korrekturlesen und die Hilfe besonders zu Beginn der Arbeit danken. Allen Mitgliedern des EKP, besonders auch dem Admin-team und dem Sekretariat danke ich für die gute Zusammenarbeit.

Ich möchte mich auch bei meinen WG-Mitbewohnern Melanie Senn und Jan Bernhardt, sowie bei Christoph Sprunk für die Verbesserungsvorschläge bei dieser Arbeit im Bereich jenseits der Physik bedanken.

Abschließend danke ich meiner Schwester und meinen Eltern, ohne die das gesamte Studium nicht möglich gewesen wäre, für ihre grenzenlose Unterstützung und dafür, dass sie immer für mich da sind.

Hiermit versichere ich, die vorliegende Arbeit selbständig verfasst und nur die angegebenen Hilfsmittel verwendet zu haben.

Sandra Kowalczyk

Karlsruhe, 30. Januar 2009