

Instituto de Física de Cantabria
(CSIC-Universidad de Cantabria)

y

Departamento de Física Moderna
(Universidad de Cantabria)



Búsqueda del Bosón de Higgs del Modelo
Standard en el canal de desintegración
 $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ en el experimento
CMS del LHC

Memoria de tesis presentada por
Rebeca González Suárez
para optar al grado de Doctor

Dirigida por
Dr. Javier Cuevas
Dr. Teresa Rodrigo Anoro^y

Santander, Junio de 2010

Instituto de Física de Cantabria
(CSIC-Universidad de Cantabria)
y
Departamento de Física Moderna
(Universidad de Cantabria)



Search for a SM Higgs Boson in the LHC
with the CMS experiment using the
 $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ decay channel

Memoria de tesis presentada por
Rebeca González Suárez
para optar al grado de Doctor

Dirigida por
Dr. Javier Cuevas
Dr. Teresa Rodrigo Anoro^y

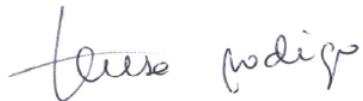
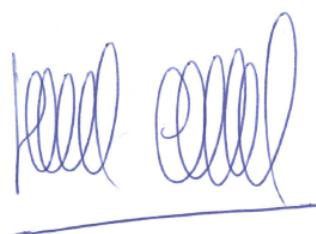
Santander, Junio de 2010

Dr. Javier Cuevas, profesor titular Física Atómica del departamento de Física de la Universidad de Oviedo, y **Dr. Teresa Rodrigo Anoro**, Catedrática de Universidad del área de Física Atómica, Molecular y Nuclear de la Facultad de Ciencias de la Universidad de Cantabria.

Certifican:

Que la presente memoria: **Búsqueda del Bosón de Higgs del Modelo Standard en el canal de desintegración $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ en el experimento CMS del LHC**, ha sido realizada bajo nuestra dirección en el Departamento de Física Moderna de la Facultad de Ciencias de la Universidad de Cantabria por **Rebeca González Suárez**, para optar al grado de Doctor en Ciencias Físicas.

Y para que así conste, en cumplimiento de la legislación vigente, firmamos el presente certificado:



Santander, Junio de 2010

Contents

1	Introduction	1
2	Standard Model Higgs Boson	5
2.1	The Standard Model	5
2.1.1	Elementary particles and interactions	5
2.2	Gauge Field Theories	8
2.2.1	Abelian Gauge Invariance (QED)	8
2.2.2	Gauge Theory of Strong Interactions (QCD)	10
2.2.3	The Gauge Theory of Electroweak Interactions	10
2.2.4	The Standard Model Lagrangian	11
2.3	The Higgs Mechanism	12
2.3.1	The SM Lagrangian with Higgs	14
2.4	The search for the SM Higgs Boson	15
2.4.1	Theoretical bounds on m_H	16
2.4.2	Experimental bounds on m_H	18
2.4.3	Higgs Physics at LEP	18
2.4.4	Higgs Physics at Tevatron	21
2.5	Beyond the Standard Model	23
2.5.1	Supersymmetry (SUSY)	24
2.5.2	The Minimal Supersymmetric Extension of the Standard Model (MSSM)	24
2.5.3	Higgs Bosons in the MSSM	25
3	The CMS experiment at LHC	27
3.1	The Large Hadron Collider	27
3.2	The CMS experiment	29
3.3	Detector Overview	30
3.3.1	Magnet	31
3.3.2	Inner Tracking System	32
3.3.3	Electromagnetic Calorimeter	34
3.3.4	Hadron Calorimeter	36
3.3.5	Muon System	38
3.3.5.1	Muon Alignment	40

CONTENTS

3.3.5.2	Alignment and Calibration scenarios	42
3.3.6	Trigger and data acquisition	43
3.4	CMS Software and Computing	45
3.4.1	CMS Remote Analysis Builder	48
3.4.2	ROOT	48
3.5	Workflow of the Analysis	49
4	Search for a SM Higgs in the WW^* decay channel	51
4.1	Higgs Phenomenology at the LHC	51
4.2	Higgs searches in CMS	54
4.3	The WW^* channel	56
4.3.1	Signature of Signal and Backgrounds	57
4.3.2	Description of the Search Process	59
4.4	LHC Center of mass Energy and Integrated Luminosity Scenarios	63
4.5	Monte Carlo Datasets	64
4.5.1	Monte Carlo Generators	64
4.5.2	Monte Carlo 7 TeV Production	65
4.6	Signal and Background Samples and Cross-Sections	68
5	Physics Observables	73
5.1	Muons	73
5.1.1	Muon Reconstruction	73
5.1.2	Muon Identification	77
5.1.3	Muon Isolation	81
5.1.4	Muon selection	83
5.2	Electrons	84
5.2.1	Electron Reconstruction	84
5.2.2	Electron identification	85
5.2.3	Electron selection	86
5.3	Jets	87
5.3.1	Jet Flavors	88
5.3.1.1	Calo Jets	88
5.3.1.2	Particle Flow Jets	88
5.3.1.3	Track Jets	89
5.3.2	Jet Algorithms	89
5.3.2.1	Iterative Cone	90
5.3.2.2	Midpoint Cone	90
5.3.2.3	Seedless Infrared-Safe Cone	91
5.3.2.4	k_t	91
5.3.2.5	Anti- k_t	91
5.3.3	Jet Energy Corrections	91
5.3.4	Jet Plus Tracks	93
5.3.5	Jets in the analysis	93
5.3.6	Jet Veto	96

5.4	missing E_T	98
5.4.1	Calorimeter missing E_T	99
5.4.2	Track Corrected missing E_T	99
5.4.3	Particle Flow missing E_T	100
5.4.4	P_T of the di- μ system	100
5.4.5	missing E_T in the analysis	101
6	Event Selection and Kinematics	109
6.1	Trigger Selection and Datasets	109
6.1.1	Skim and Pre-selection	112
6.2	Muon Selection	113
6.3	Central Jet Veto	113
6.4	Definition of the Variables	114
6.5	Sequential cuts	124
6.5.1	Mass dependent Optimization	124
6.6	Final Event Selection	125
6.6.1	Study at a centre of mass energy of 14 TeV	125
6.6.2	Study at a centre of mass energy of 10 TeV	127
6.6.3	Study at a centre of mass energy of 7 TeV	128
6.7	Signal and Background production rates	129
6.8	$\mu\mu$, $e\mu$ and ee final states	132
6.9	Complementary Strategies for the Analysis	132
6.9.1	Use of track Jets and b-tagging	133
6.9.1.1	Track Jets	133
6.9.1.2	Soft-muon and Life time based b-tagging	134
6.9.2	Multivariate Methods	136
7	Methods to estimate the Background from Data and Systematic Uncertainties	141
7.1	Control Regions to estimate the Background	141
7.1.1	$t\bar{t}$ Background Normalization	142
7.1.2	WW Background normalization	145
7.1.3	$Z + jets$ Background normalization	145
7.2	Control of the Muon Fake Rate in $W + jets$ events	146
7.3	missing E_T studies from data	148
7.4	WW Cross-section measurement	149
7.5	Systematic Uncertainties	154
7.5.1	Luminosity	154
7.5.2	Muon identification and Trigger efficiencies	154
7.5.3	Misalignment and Miscalibration scenarios	155
7.5.4	Missing E_T modeling	155
7.5.5	Jet reconstruction efficiency	155
7.5.6	Jet energy scale	156
7.5.7	ISR/FSR	158

CONTENTS

7.5.8	PDF estimation	158
7.5.9	Normalization for backgrounds	159
7.5.10	Monte Carlo statistics	159
8	Results	161
8.1	Limits and Significance	161
8.2	Results at a centre of mass energy of 14 TeV	162
8.3	Projections at 10 TeV	164
8.4	Projections at 7 TeV	166
9	Conclusions	171
10	Resumen del trabajo realizado y Conclusiones obtenidas	177
A	Signal and Background yields	185
A.1	Tables for 1 fb^{-1} at $\sqrt{s} = 14\text{ TeV}$	185
A.2	Tables for 200 pb^{-1} at $\sqrt{s} = 10\text{ TeV}$	188
A.3	Tables for 1 fb^{-1} at $\sqrt{s} = 7\text{ TeV}$	190
B	Results for the $\mu\mu$ final state at a centre of mass energy of 7 TeV	195
C	First LHC data analysis	199
C.1	Events with muons in the final state	200
C.2	Current LHC run conditions	201
C.3	Early data analysis	202
C.3.1	Muon Fake Rate	204
References		211

Para Meli, Jose y Fran.

CONTENTS

1

Introduction

The Standard Model(SM) of Particle Physics has shown a very good agreement with the experimental data up to now, but it is still incomplete. It fails to explain aspects like the breaking of the electroweak (EWK) symmetry, one of the most important open questions of contemporary particle physics. The minimal Higgs mechanism gives the simplest answer to the electroweak symmetry breaking leaving the SM mostly untouched. It requires the existence of a new scalar particle, the Higgs boson, that has to be detected. All the properties of the Higgs boson can be predicted in the SM except its mass, a free parameter of the theory.

Revealing the physical mechanism responsible for the breaking of the EWK symmetry is one of the main physics goals for the Large Hadron Collider (LHC). If the Higgs mechanism is valid, the LHC experiments have the potential to find the Higgs boson, closing a search that started years ago and in which several experiments at different colliders have participated. The large centre of mass energies produced in proton-proton collision at the LHC together with the high luminosities expected will allow the LHC to develop a physics program that will give the final answer to this search. The LHC will cover a wide range of Higgs searches, from the minimal SM Higgs, to beyond the SM through the Minimal Supersymmetric Standard Model(MSSM) extensions and other Super Symmetry (SUSY) models.

At the end of the Large Electron Positron (LEP) collider era, a solid exclusion was reached by combining the results of all the Higgs search channels performed by the four experiments, establishing a lower limit on m_H of $114.4 \text{ GeV}/c^2$ at 95% confidence level [1]. More recently, the direct Higgs searches performed in the CDF and D0 experiments at the Tevatron accelerator have reached better sensitivity, making possible the exclusion of a SM Higgs boson with a mass between 162 and $166 \text{ GeV}/c^2$ [2]. Constrains on m_H coming from EWK precision fits have also been derived. All the experimental and theoretical bounds together determine the mass range in which the search at the LHC will be mainly focused, and set the preferred mass for a SM Higgs boson in a narrow window: $115 \leq m_H \leq 148 \text{ GeV}/c^2$ [3]. The gluon fusion process is the predom-

1. INTRODUCTION

inant Higgs production mode at the LHC. In the already mentioned mass range, the main decay channel is $H \rightarrow WW^*$.

A prospective analysis for the search of a SM Higgs boson in the Compact Muon Solenoid (CMS) experiment at the LHC is performed in this thesis, focusing in the $gg \rightarrow H \rightarrow WW^* \rightarrow \mu\nu\mu\nu$ process. The search channel has a clear signature characterized by the presence of two high P_T muons in the final state and a significant amount of missing E_T due to the undetected neutrinos. No hard jet activity is expected.

The principal sources of background come from several SM processes that may have similar final states as those of the signal, mainly $W + jets$, $Z + jets$, $t\bar{t}$ and the almost irreducible WW . Other sources of background are WZ , ZZ and tW processes. Final states coming from QCD processes have different signatures but very large cross-sections and they have also been studied.

Several analysis techniques have been developed in this work. The main one is based on the application of a sequential selection criteria using a set of physical observables which characterize the main kinematic properties of the signal and discriminate with respect to the different backgrounds. Given the presence of neutrinos in the final state, that prevents the reconstruction of a clear Higgs mass peak, and the fact that the production cross-section for the Higgs signal is various orders of magnitude smaller than that of the backgrounds, the study relies in the control of the dominant sources of background.

The analysis has been designed to effectively reduce the different types of background processes. The muon reconstruction, identification and selection is optimized to select events with two muons coming from primary decays of W bosons and is the key to reject $W + jets$ events when the W decays into a muon. A detailed data-driven muon fake rate study has been designed to treat this background as it has a large cross-section and can be a source of background when a second muon is wrongly identified. The hadronic activity present in the $t\bar{t}$ events has to be controlled, and events with hard jets coming from b decays have to be vetoed. The proper estimation of the missing E_T in events with no genuine missing E_T is crucial to reduce the $Z + jets$ background to an acceptable level. The main variable to discriminate the signal from the WW background is the angle between the muons in the transverse plane to the beam, $\Delta\phi_{\mu\mu}$. The muons emitted by the W bosons coming from the decay of a Higgs scalar are close in the transverse plane due to spin correlations, while the angular separation of the muons coming from WW decays tend to be larger. This is the most characteristic feature of the signal, and plays a major role in the discovery potential of the Higgs using $H \rightarrow WW^*$ process.

The analysis developed in this thesis has been applied to different running scenarios and adapted to the evolving LHC running conditions. Initially, the design energy of the LHC of 14 TeV was considered, and the study was performed assuming an inte-

grated luminosity of 1 fb^{-1} . Later the search was optimized, first to a centre of mass energy of 10 TeV and to an integrated luminosity of 200 pb^{-1} and finally to 7 TeV, the current centre of mass energy at which the LHC is running. The study at 7 TeV is done for an integrated luminosity of 1 fb^{-1} , that corresponds to the total data expected to be collected by the end of 2011.

The decision of running the LHC at 7 TeV was announced at the Chamonix LHC workshop in 2010 [4]. The centre of mass energy determines the signal yield and the signal to background ratio, and more luminosity is needed in this case to achieve the same compared to a 14 TeV scenario.. Therefore, the direct Higgs search program with the CMS detector at the LHC will start to be significant in the last period of the 7 TeV run, and will probably be developed at the same time as the last run of the Tevatron.

During the first months of running the different background processes will start to be accessible as the recorded luminosity increases. The study of the WW background is particularly interesting in the early stages of the data taking given its similarities with the signal. Using a similar analysis as the Higgs to WW^* search with essentially the same observables, it will be possible to measure the cross-section of the WW processes at the LHC. Both studies are complementary as the best possible knowledge of WW background is essential for the performance of the Higgs to WW^* search and in the WW analysis, the Higgs activity could be observed as an excess with respect to the background.

The work is presented as follows: In chapter 2 a short theoretical introduction to the SM and the Higgs physics is offered, as well as the status of the Higgs searches performed up to now in the different accelerators. Then, in chapter 3 a comprehensive description of the CMS experiment and the LHC is presented.

In chapter 4 the main Higgs production and decay modes at the LHC are described and the signature of the signal and the main backgrounds is discussed in detail. In chapter 5 the reconstruction and identification of the different high level physics objects used in this study is described.

In chapter 6 the event selection is presented and the analysis is applied to the different scenarios of centre of mass energies and integrated luminosity.

In chapter 7 several data-driven techniques to estimate the yields coming from the main backgrounds are proposed. Control regions are determined for the estimation of the $t\bar{t}$, WW and $Z + jets$ processes. A fake rate study is performed to estimate the contribution from $W + jets$ events. An strategy for the measurement of the WW cross-section using a complementary study is developed. The relevant sources of systematic error are studied in detail and the most important ones are presented.

1. INTRODUCTION

The results of this study in the different scenarios corresponding to several centre of mass energies and integrated luminosity, focusing in the current benchmark scenario of 1 fb^{-1} at a centre of mass energy of 7 TeV are presented as a function of the Higgs mass m_H in terms of exclusion confidence levels and significances in chapter 8. The results of the work presented here is combined with the results provided by the rest of the Higgs search channels that have been performed by the CMS collaboration, especially the ZZ^* process, to estimate the potential of the CMS experiment for the search of a SM Higgs in the current centre of mass scenario of the LHC.

Chapter 9 contains the final conclusions extracted from the work described in the previous sections.

2

Standard Model Higgs Boson

2.1 The Standard Model

Theoretical and experimental studies in particle physics for more than fifty years have led to the development of the current Standard Model of Particle Physics. Although it is still incomplete, the Standard Model is widely accepted, and beautifully accords with almost all current experimental data, reaching in many cases a very precise agreement.

In the following, a concise introduction to the theoretical background of the Standard Model will be presented. For more details, extensive bibliography is available in this subject, starting with many introductory textbooks, as [5].

2.1.1 Elementary particles and interactions

The Standard Model (SM) [6; 7; 8] shows a bewitching simplicity when characterizing forces and matter at the most elementary level, relying on Elementary Particles.

Elementary Particles are point-like particles, not known to have sub-structure (up to the present limits of $10^{-18} - 10^{-19} \text{ m}$). The SM stands that there are two kinds of elementary particles; the ones that compose matter, and the ones that mediate the interactions.

That is: matter is composed by a limited set of primary blocks, and interactions are described as the exchange of force-carrier particles.

There are two types of elementary particles that conforms matter: leptons and quarks; and each type have three generations. The well-known electron, e^- , as well as its heavier copies, the muon, μ^- , and the tau, τ^- , which are massive and electrically charged, accompanied by their almost mass-less, electrically neutral neutrinos, ν_e , ν_μ , ν_τ ; plus their anti-particles, conform the group of leptons. Quarks (up u , down d , strange s , charm c , bottom b and top t ; plus their anti-particles), have mass and fractional electric

2. STANDARD MODEL HIGGS BOSON

charge. Both leptons and quarks are fermions, as they have spin $s = 1/2$. An sketched overview of the Fermions can be seen in table 2.1.

Quarks present an additional quantum number, the color, which can be of three

	Charge	First generation		Second generation		Third generation	
Leptons	-1	Electron	e^-	Muon	μ^-	Tau	τ^-
	0	Electron Neutrino	ν_e	Muon Neutrino	ν_μ	Tau Neutrino	ν_τ
Quarks	+2/3	Up	u	Charm	c	Top	t
	-1/3	Down	d	Strange	s	Bottom	b

Table 2.1: Overview of Fermions in the Standard Model.

types, generically denoted as q_i , $i = 1, 2, 3$. Color is not seen in Nature, and, because of that, it is not possible to individually observe quarks. However, their color-less bound states, called hadrons, can be observed. Hadrons are classified into baryons and *mesons*. Mesons are bosons made of one quark and one anti-quark, $q\bar{q}'$, like for instance the pions, π^+ , ($u\bar{d}$), and π^- , ($d\bar{u}$). The baryons are fermions composed by three quarks, being the proton, p , (uud), and the neutron, n , (udd), the most familiar examples.

The lightest and most stable particles make up the first generation, while the heavier and less stable particles belong to the second and third generations. All stable matter in the Universe is made from particles that belong to the first generation; as the atomic nucleus is formed by electrons, e^- , neutrons, n (udd), and protons, p (uud); any heavier particles quickly decay to the next most stable level.

The interactions of these 24 types of fermions (12 particles + 12 anti-particles) via three forces are enough to describe the elementary processes in particle physics. These forces are: The Electromagnetic Interaction (Quantum Electrodynamics, QED), corresponding to the classical Electrodynamics described by Maxwell; the Strong interaction, responsible for the atomic nuclei; and the Weak interaction, responsible for the nucleon decays. Other interactions, with special stress in the Gravitation, are also interesting but out of the scope of this thesis.

Interactions are mediated by the exchange of particles. Therefore, the second kind of elementary particles are the ones that intermediate these forces: bosons with spin $s = 1$.

The photon, γ , mass and charge-less, is exchanged in Electromagnetic interactions; eight gluons, g_α ($\alpha = 1, \dots, 8$), also mass and charge-less, but colored, mediate the Strong interactions amongst quarks; and the three weak bosons, the neutral Z^0 and the electrically charged W^\pm mediate the Weak interactions. A small summary of the Bosons is presented in table 2.2.

2.1 The Standard Model

Interaction	Particle	Charge	Self-interacting
Electromagnetic	Photon γ	0	No
Strong	Gluon g_α $(\alpha = 1, \dots, 8)$	0	Yes
Weak	Z^0 W^- W^+	0 -1 +1	Yes

Table 2.2: Overview of force-carrier Bosons in the Standard Model.

Theoretically, the SM is a quantum field theory based on the gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$. This gauge group includes the symmetry group of strong interaction $SU(3)_C$, and the symmetry group of Electroweak interactions $SU(2)_L \times U(1)_Y$, where $U(1)_{em}$ appears as a sub-group; and is in this sense that the weak and electromagnetic interactions are said to be unified.

Hence, the gauge sector of the SM is composed by twelve gauge bosons, which are the force carriers previously mentioned: the eight gluons g_α ($\alpha = 1, \dots, 8$) which are the gauge bosons of $SU(3)_C$; and the photon, γ , plus the three weak bosons, Z^0 , W^+ and W^- , which are the correspondent four gauge bosons of the $SU(2)_L \times U(1)_Y$ group.

Experimentally, it is found that the particles can have mass. The current knowledge of the masses of the fermions is presented in table 2.3.

First generation		Second generation		Third generation	
Particle	Mass [GeV/c^2]	Particle	Mass [GeV/c^2]	Particle	Mass [GeV/c^2]
e^-	$0.5 \cdot 10^{-3}$	μ^-	0.105	τ^-	1.777
ν_e	$< 10^{-6}$	ν_μ	$< 10^{-6}$	ν_τ	$< 10^{-6}$
u	1.5 to $3.3 \cdot 10^{-3}$	c	1.27	t	171.2
d	3.5 to $6.0 \cdot 10^{-3}$	s	0.104	b	4.20

Table 2.3: Masses of the fermions.

The range of the interactions depends on the inverse of the masses of the gauge bosons that mediate them. For that, the range of the Electromagnetic interaction, mediated by photons ($m_\gamma = 0$), is infinite. A summary on the masses of the boson and the range of the interactions can be seen in 2.4.

Masses and other properties of the different particles can be found in [9].

2. STANDARD MODEL HIGGS BOSON

Interaction	Particle	Mass [GeV/c ²]	Range of the Interaction [m]
Electromagnetic	γ	0	Infinite
Strong	g	< 0.20	10^{-15} (\sim nucleus diameter)
Weak	Z^0 W	91.19 80.40	10^{-18}

Table 2.4: Masses of the bosons.

2.2 Gauge Field Theories

As said above, The SM is described within the context of modern Relativistic Quantum Field Theories [10].

The quantum field theories of the electromagnetic, weak, and strong interactions of the fundamental particles, all belong to the class of local gauge theories.

In short, gauge theories are dynamical theories based on local invariance. The Lagrangian is invariant under a certain continuous group of local transformations called gauge transformations, which conform the symmetry group of the theory. For each symmetry group, there's an associated algebra of generators; and for each group generator there is a corresponding vector field called gauge field. Gauge fields are included in the Lagrangian to ensure the invariance. In the case of quantum field theories, the quanta of the gauge fields are called gauge bosons.

2.2.1 Abelian Gauge Invariance (QED)

Quantum Electrodynamics (QED) [11], the original and most successful gauge field theory, is the perfect example of the application of the gauge principle and its physical implications.

As a representative example of matter we can consider a fermion field ψ with $s = \frac{1}{2}$, mass m , and electric charge Q_e . The corresponding Lagrangian is well known,

$$\mathcal{L}_0 = \bar{\psi}(x)(i \not{\partial} - m)\psi(x) ; \quad \not{\partial} \equiv \partial_\mu \gamma^\mu \quad (2.1)$$

and the Dirac equation gives the corresponding equation of motion:

$$(i \not{\partial} - m)\psi(x) = 0 \quad (2.2)$$

It is immediate to show the invariance of the Lagrangian to the global $U(1)$ transformations which act on the fields and their derivatives like this:

$$\begin{aligned} \psi &\rightarrow e^{-iQ\omega}\psi \\ \partial_\mu\psi &\rightarrow e^{-iQ\omega}\partial_\mu\psi \end{aligned} \quad (2.3)$$

2.2 Gauge Field Theories

$Q\omega$ is the global phase, where ω is the transformation continuous parameter, an arbitrary constant number, independent of x .

By Noether's Theorem [12], this global invariance implies the conservation of the electromagnetic current J_μ ; and therefore, the conservation of the electromagnetic charge

$$J_\mu = \bar{\psi} \gamma_\mu e Q \psi ; \partial_\mu J^\mu = 0 ; eQ = \int d^3x J_0(x) \quad (2.4)$$

If the transformation is promoted to local, i.e., if the parameter θ is allowed to depend on the space-time point x , the transformation of the fields and their derivatives are like this:

$$\begin{aligned} \psi &\rightarrow e^{-iQ\omega(x)} \psi \\ \partial_\mu \psi &\rightarrow e^{-iQ\omega(x)} [\partial_\mu \psi(x) - iQ(\partial_\mu \omega(x))\psi] \end{aligned} \quad (2.5)$$

the Lagrangian acquires an extra term that spoils the invariance.

To solve this, a gauge vector boson field $A_\mu(x)$ is introduced; it interacts with the field ψ and transforms properly canceling the extra term

$$A_\mu \rightarrow A_\mu - \frac{1}{e} \partial_\mu \omega(x) \quad (2.6)$$

making the total Lagrangian gauge invariant.

The Lagrangian is easily built replacing the derivative ∂_μ by the covariant derivative D_μ ; where

$$D_\mu \psi \equiv (\partial_\mu + ieQ A_\mu) \psi \quad (2.7)$$

which transforms covariantly, i.e., as the field itself

$$D_\mu \psi \rightarrow e^{-iQ\omega(x)} D_\mu \psi \quad (2.8)$$

To include the propagation of the photon field, the kinetic term, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, written in terms of the field strength tensor and also gauge invariant, has to be added; giving as a result a final $U(1)$ gauge-invariant Lagrangian which looks like

$$\mathcal{L}_{QED} = \bar{\psi}(x)(i \not{D} - m)\psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) \quad (2.9)$$

2.9 is the QED Lagrangian, and it contains the interactions in the term $\bar{\psi}i \not{D}\psi$

$$\bar{\psi}ieQ A_\mu \gamma^\mu \psi \quad (2.10)$$

The gauge group for electromagnetism is the $U(1)_{em}$, with one generator, Q and one parameter θ . It has one gauge field, the Electromagnetic field, with the photon, γ , being the gauge boson. As the local gauge symmetry forbids mass terms for the gauge-field, the photons have to be mass-less to preserve the symmetry.

2. STANDARD MODEL HIGGS BOSON

2.2.2 Gauge Theory of Strong Interactions (QCD)

Quantum Chromodynamics (QCD) is the gauge theory for strong interactions, and is based on the local color transformations which leave its Lagrangian invariant. The gauge symmetry group that is generated by these color transformation is the non-abelian Lie group $SU(3)_C$. Where C refers to colors and 3 refers to the three possible color states of the quarks.

The building of the QCD Lagrangian is done by following the same steps as in the QED case. The gauge principle is applied taking into account the particularities of the non-abelian group $SU(3)$. The global symmetry is promoted to local by replacing the derivative of the quark by its covariant derivative, which, in the QCD case is like this:

$$D_\mu q = (\partial_\mu - ig_s \frac{\lambda_\alpha}{2} A_\mu^\alpha) q \quad (2.11)$$

Where $q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$, q_i are the quark fields, g_s is the strong coupling constant, $\frac{\lambda_\alpha}{2}$ are the $SU(3)$ generators and A_μ^α are the gluon fields.

The QCD Lagrangian is then written

$$\mathcal{L}_{QCD} = \sum_q \bar{q}(x)(i \not{D} - m_q)q(x) \frac{1}{4} F_{\mu\nu}^\alpha(x) F_\alpha^{\mu\nu}(x) \quad (2.12)$$

The gluon field strength is then

$$F_{\mu\nu}^\alpha(x) = \partial_\mu A_\nu^\alpha(x) - \partial_\nu A_\mu^\alpha(x) + g_s f^{\alpha\beta\gamma} A_{\mu\beta} A_{\nu\gamma} \quad (2.13)$$

it contains a bilinear term as it correspond to a non-abelian gauge theory with structure constants $f^{\alpha\beta\gamma}$ with $\alpha, \beta, \gamma = 1, \dots, 8$.

Similarly to QED, the gauge interactions among the quarks and gluons are contained in the $\bar{q}i \not{D}q$ term. There is, however, an important difference with the QED case. The gluon kinetic term $F_{\mu\nu}^\alpha F_\alpha^{\mu\nu}$ contains a three gluons term and a four gluons term. These are precisely, the self interaction gluon vertices, genuine of a non-abelian theory.

2.2.3 The Gauge Theory of Electroweak Interactions

The electroweak theory joins electromagnetism with the weak force in a single relativistic quantum field theory. As it was stated at the beginning of this chapter, Electromagnetism is a force of infinite range, while the influence of the charged-current weak interaction only spans distances shorter than about 10^{-15} cm . Two interactions so

different in range and apparent strength are described to a common gauge symmetry, namely, the symmetry $SU(2)_L \times U(1)_Y$ which is required to be a local symmetry of the electroweak Lagrangian.

$SU(2)_L$ is the weak isospin group, which acts on left-handed fermions, and $U(1)_Y$ is the weak hypercharge group. The $SU(2)_L \times U(1)_Y$ group has four generators, three of which are the $SU(2)_L$ generators, $T_i = \frac{\sigma_i}{2}$, with $i = 1, 2, 3$, and the fourth one is the $U(1)_Y$ generator, $\frac{Y}{2}$. The commutation relations for the total group are:

$$[T_i, T_j] = i\epsilon_{ijk}T_k; [T_i, Y] = 0; i, j, k = 1, 2, 3 \quad (2.14)$$

The left-handed fermions transform as doublets under $SU(2)_L$

$$f_L \rightarrow e^{i\vec{T}\cdot\vec{\theta}} f_L; f_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \dots \quad (2.15)$$

While the right-handed fermions transforms as singlets,

$$f_R \rightarrow f_R; f_R = e_R, u_R, d_R, \dots \quad (2.16)$$

The relation between the quantum numbers of the fermions, $Q = T_3 + \frac{Y}{2}$, is also incorporated to the SM. The number of associated gauge bosons, being equal to the number of generators, is four; three weak bosons of $SU(2)_L$, W_μ^i , with $i = 1, 2, 3$, and the hypercharge boson of $U(1)_Y$, B_μ .

The building of the SM Lagrangian is done by following the same steps as in any gauge theory. In particular, the $SU(2)_L \times SU(1)_Y$ symmetry is promoted from global to local by replacing the derivatives of the fields by the corresponding covariant derivatives. For a generic fermion field f this covariant derivative is:

$$D_\mu f = (\partial_\mu - ig\vec{T}\cdot\vec{W}_\mu - ig'\frac{Y}{2}B_\mu)f \quad (2.17)$$

where, g is the coupling constant corresponding to $SU(2)_L$ and g' is the one corresponding to $U(1)_Y$. For example, the covariant derivative for a left-handed and right-handed electron are respectively:

$$D_\mu e_L = (\partial_\mu - ig\frac{\vec{\sigma}}{2}\cdot\vec{W}_\mu - ig'\frac{1}{2}B_\mu)e_L; D_\mu e_R = (\partial_\mu + ig'B_\mu)e_R \quad (2.18)$$

As in the previous cases of QED and QCD, the gauge invariant electroweak interactions are generated from the $\bar{f}iDf$ term.

2.2.4 The Standard Model Lagrangian

To get the Lagrangian of the Electroweak Theory it is necessary to add the previous fermion terms, the gauge boson kinetic terms and the gauge boson self-interacting terms.

$$\mathcal{L}_{SM} = \mathcal{L}_f + \mathcal{L}_G + \mathcal{L}_{SBS} + \mathcal{L}_{YW} \quad (2.19)$$

2. STANDARD MODEL HIGGS BOSON

where the fermion Lagrangian is

$$\mathcal{L}_f = \sum_{l=l,q} \bar{f} i \not{D} f \quad (2.20)$$

and the Lagrangian for the gauge field is

$$\mathcal{L}_G = -\frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \mathcal{L}_{GF} + \mathcal{L}_{FP} \quad (2.21)$$

\mathcal{L}_{GF} and \mathcal{L}_{FP} are the gauge fixing and Faddeev Popov Lagrangians, needed in any gauge theory, previously omitted for QED and QCD; \mathcal{L}_{SBS} and \mathcal{L}_{YW} are the Symmetry Breaking Sector Lagrangian and the Yukawa Lagrangian, needed to provide masses to the weak bosons, m_W and m_Z and the fermions m_f .

The resulting Lagrangian contains the wanted self-interaction terms among the three electroweak bosons, and it can be shown that this is invariant under the $SU(2)_L \times U(1)_Y$ gauge transformations

The gauge bosons, W_μ^\pm , Z_μ and A_μ , are obtained from the electroweak interaction eigenstates

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2) \\ Z_\mu &= c_\omega W_\mu^3 - s_\omega B_\mu \\ A_\mu &= s_\omega W_\mu^3 + c_\omega B_\mu \end{aligned} \quad (2.22)$$

Note that the mass terms $m_W^2 W_\mu W^\mu$, $\frac{1}{2} m_Z^2 Z_\mu Z^\mu$ and $m_f \bar{f} f$ are forbidden by $SU(2)_L \times U(1)_Y$ gauge invariance. The needed gauge boson masses have to be generated in a gauge invariant way. The spontaneous breaking of the symmetry via the Higgs Mechanism can provide the mass generation.

2.3 The Higgs Mechanism

In the previous section, it has been shown that, to preserve the invariance of the Lagrangian, the mass terms of the gauge bosons are explicitly forbidden; but this is contradicted by the experimental evidence, as the masses of the weak bosons, W^\pm and Z^0 has been experimentally measured, and it is well-known that they are very heavy.

Hence, the Symmetry has to be broken. The Spontaneous Symmetry Breaking (SSB) is one of the key ingredients of the electroweak interactions. In Quantum Field Theory, a system is said to posses a symmetry that is spontaneously broken when the Lagrangian is invariant under the symmetry transformations, while the vacuum of the theory is not. The vacuum is the state in which the Hamiltonian expectation value is minimum.

2.3 The Higgs Mechanism

In the context of the SM, when the SSB refers to a gauge symmetry, the Higgs Mechanism [13; 14; 15; 16] has to operate. The symmetry-breaking is realized linearly by a scalar field which acquires a non-zero vacuum expectation value. As a result, not only the masses of the vector-bosons and the fermions arise, but also the Higgs particle, a neutral scalar field which was not experimentally observed so far.

Considering the Higgs Mechanism as a responsible for the symmetry breaking in the SM, a consistent formulation exists, and any observable can be calculated perturbatively in the Higgs self-coupling constant. However, the concept of symmetry breaking is more general, and even when the Higgs mechanism is the simplest one, there are countless alternatives to replace the standard Higgs sector, following just three basic requirements:

1. Electromagnetism has to remain unbroken
2. The full symmetry has to contain the electroweak gauge symmetry
3. The symmetry breaking has to occur at about the energy scale $v = (\sqrt{2}G_F)^{-\frac{1}{2}} = 256 \text{ GeV}$, being G_F the Fermi coupling constant.

Given that, the simplest choice for the SSB of the Electroweak theory in the SM is the Higgs Mechanism, in which an additional scalar field Φ is introduced:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^- \end{pmatrix} \quad (2.23)$$

Then, the Lagrangian of the SSB of the Electroweak theory will be:

$$\mathcal{L}_{SBS} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \quad (2.24)$$

the covariant derivative is then:

$$D_\mu = (\partial_\mu - \frac{1}{2}ig\vec{\sigma}\overrightarrow{W}_\mu - \frac{1}{2}ig'B_\mu) \quad (2.25)$$

Here, Φ is a fundamental complex $SU(2)$ doublet with hypercharge $Y(\Phi) = 1$ and $V(\Phi)$ is the simplest renormalizable potential:

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 ; \lambda > 0 \quad (2.26)$$

Depending on the mass parameter, $-\mu^2$, there are two possibilities that minimize the potential $V(\Phi)$, figure 2.1:

- If $(-\mu^2) > 0$, the minimum is at $\langle \Phi \rangle = 0$. The vacuum is $SU(2)_L \times U(1)_Y$ symmetric, and therefore, no symmetry breaking occurs.

2. STANDARD MODEL HIGGS BOSON

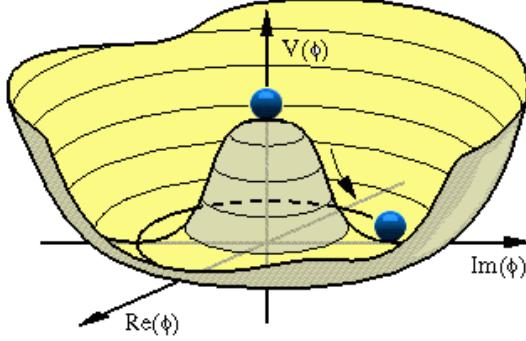


Figure 2.1: - Representation of the Higgs potential.

- If $(-\mu^2) < 0$, the minimum is at $\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$, for any arbitrary argument of Φ .

Therefore, there are infinite degenerate vacuum choices, either of them $SU(2)_L \times U(1)_Y$ non symmetric and $U(1)_{em}$ symmetric. The simplest one, where $\arg \Phi \equiv 0$, is taken.

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}; v = \sqrt{\frac{\mu^2}{\lambda}} \quad (2.27)$$

2.3.1 The SM Lagrangian with Higgs

To get the particles spectra and the particle masses; the $SU(2)_L \times U(1)_Y$ gauge invariant Lagrangian should be written:

$$\mathcal{L}_{SM} = \mathcal{L}_f + \mathcal{L}_G + \mathcal{L}_{SSB} + \mathcal{L}_{YW} \quad (2.28)$$

\mathcal{L}_f and \mathcal{L}_G have been given previously, and \mathcal{L}_{SSB} and \mathcal{L}_{YW} are the SSB and the Yukawa Lagrangians respectively,

$$\mathcal{L}_{SSB} = (D_\mu \Phi)^\dagger (D_\mu \Phi) + \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \quad (2.29)$$

$$\mathcal{L}_{YW} = \lambda_e \bar{l}_L \Phi_{eR} + \lambda_u \bar{q}_L \tilde{\Phi} u_R + \lambda_d \bar{q}_L \Phi d_R + h.c. + 2^{nd} \text{ and } 3^{rd} \text{ families} \quad (2.30)$$

Where:

$$l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}; q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}; \Phi = \begin{pmatrix} \Phi^+ \\ \Phi^- \end{pmatrix}; \tilde{\Phi} = i\sigma_2 \Phi^* = \begin{pmatrix} \Phi_0^* \\ -\Phi^- \end{pmatrix}$$

\mathcal{L}_{SSB} provides m_W and m_Z ; \mathcal{L}_f is needed to provide m_f , the masses of the fermions.

Then, in order to get the spectrum for \mathcal{L}_{SM} , a non-symmetric vacuum is fixed, $\langle \Phi \rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$. The physical spectrum is built performing small oscillations around the vacuum, parametrized by

$$\Phi(x) = e^{i \frac{\vec{\xi}(x) \cdot \vec{\sigma}}{v}} \begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix} \quad (2.31)$$

2.4 The search for the SM Higgs Boson

noting that $\vec{\xi}(x)$ and $H(x)$ are small fields.

Some gauge transformations are made to eliminate the un-physical fields $\vec{\xi}(x)$. Then, the weak eigenstates are rotated to the mass eigenstates which define the gauge boson fields 2.22. Once that is done, is possible to read the masses from the terms of the SM Lagrangian, and get the three level predictions for the masses of the gauge bosons:

$$m_W = \frac{gv}{2} ; m_Z = \frac{\sqrt{g^2 + g'^2 v}}{2} ; m_H = \sqrt{2}\mu \quad (2.32)$$

And for the fermions:

$$m_e = \lambda_e \frac{v}{\sqrt{2}} ; m_u = \lambda_u \frac{v}{\sqrt{2}} ; m_d = \lambda_d \frac{v}{\sqrt{2}} ; \dots \quad (2.33)$$

Finally, it is possible to re-write \mathcal{L}_{SSB} and \mathcal{L}_{YW} , after the application of the Higgs Mechanism, in terms of the physical scalar fields, and get, not only the masses, but also the kinetic and interaction terms for the Higgs sector.

With some calculations all masses can be given in terms of a unique mass parameter v and the coupling constants: g , g' , λ , λ_e , and so on. The interactions of H with fermions and gauge bosons are proportional to the gauge couplings and to the masses of the particles. The parameter v can be measured from μ -decay, identifying the predictions of the partial width $\Gamma(\mu \rightarrow \nu_\mu \bar{\nu}_e e)$ to low energies ($q^2 \ll m_W^2$) getting to:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} = \frac{1}{2v^2} ; v = (\sqrt{2}G_F)^{-\frac{1}{2}} = 246 GeV \quad (2.34)$$

The values of the masses of the W^\pm and Z^0 were successfully anticipated before being measured experimentally, being the input parameters: the fine structure constant α , G_F , and θ_w ¹; the best measured electroweak parameters before *LEP*.

It is interesting to point out that the Higgs boson mass m_H and the Higgs self-coupling constant λ are undetermined by the SM, being related at tree level by $\lambda = \frac{m_H^2}{2v^2}$. The hierarchy in the fermion masses is also undetermined by the SM.

2.4 The search for the SM Higgs Boson

The introduction of the Higgs Mechanism in the SM gives a very elegant answer to the practical needs of the theory, and goes in an excellent agreement with the experimental evidence. Bosons and fermions acquire mass that can be calculated, and can also be experimentally measured. With every piece of Nature's subatomic puzzle getting in its

¹ θ_w denotes the Weinberg mixing angle, $\tan \theta_w = \frac{g'}{g}$; which relates the masses of the weak bosons by $m_W = \cos \theta_w m_Z$.

2. STANDARD MODEL HIGGS BOSON

place, something is still currently missing: the Higgs Boson.

The existence of the Higgs Boson has to be proved or excluded. The search strategy has to be precisely designed considering the Higgs boson production mode and its decay products, and that depends strictly on its mass.

Unfortunately the Higgs mass, m_H , is not predicted in the SM. The Higgs mechanism provides m_H as a function of the Higgs self-coupling λ and $v = 246 \text{ GeV}$ in this way:

$$m_H = \sqrt{2\mu} = \sqrt{2v^2\lambda} \quad (2.35)$$

As λ is also unknown m_H can take any value in the SM. The unique restrictions come from the consistence of the theory and the experimental observation.

One of the main physics goals of the Large Hadron Collider (LHC) is the definitive search for a SM Higgs boson. Direct searches have been performed in past and recent experiments, see for example [17], but in this new era, all the possible mass spectrum will be covered. If the SM Higgs boson exist, the LHC experiments will find it. In this section, the current bounds on m_H both theoretical and experimental, are shown, that is, the opened window mass for the search of the SM Higgs Boson in the LHC, will be presented. In following chapters, the specific production and decay modes of a SM Higgs boson in the LHC will be amply discussed.

2.4.1 Theoretical bounds on m_H

There are three main theoretical bounds on the Higgs boson mass that can be calculated, being the upper bound that comes from Unitarity the first of them. In short, the Unitarity bound is related to the fact that the scattering probability can not be higher than one.

It is possible to find the unitarity conditions for particular scattering cases, like for example, $W_L^+ W_L^-$ SM scattering, giving as a result different upper bounds on the Higgs mass, that combined, give a rough.

$$m_H < 700 - 800 \text{ GeV}/c^2 \quad (2.36)$$

These bounds based on perturbative unitarity do not necessarily mean that the Higgs boson can't be heavier than these values.

In the case of large m_H , the perturbative approach is not valid anymore and non-perturbative techniques are required. For this heavy Higgs models, the Higgs self-interactions governed by the coupling λ become strong and new physics phenomena may appear at energies around 1 TeV. Hence, the existence of a non-perturbative regime for the scalar sector of the SM is still a possibility, this interesting phenomena can also be studied at the *LHC*, but is out of the scope of this thesis.

Then, the Higgs self-coupling has to be finite, and another upper bound can be derived considering Triviality. That means all the couplings being small enough for energies below some threshold. For example, if the energy scale is fixed to 10^{16} GeV , considering the top quark mass to be $m_t = 170 \text{ GeV}/c^2$, $m_H < 170 \text{ GeV}/c^2$ is derived following:

$$m_H^2 < \frac{4\phi^2\nu^2}{3\ln(\frac{\Lambda}{\nu})} \quad (2.37)$$

being Λ the cutoff scale.

A lower bound on the Higgs mass can be obtained from Vacuum Stability. The asymmetric vacuum of the $SU(2)_L \times U(1)_Y$ has to be stable under quantum corrections. The limit is resumed in this equation:

$$m_H^2 > \frac{4m_t^4}{\pi^2\nu^2} \ln\left(\frac{\Lambda}{\nu}\right) \quad (2.38)$$

Recent studies requiring vacuum stability at two-loop level up to energies of the order of 10^{16} GeV , considering $m_t = 170 \text{ GeV}/c^2$ and $\alpha_s = 0.117$ give an exclusion limit of $m_H > 132 \text{ GeV}/c^2$.

To summarize, requiring the Higgs coupling to remain finite and positive up to an energy scale Λ , constraints can be derived on the Higgs mass m_H , such analysis exist at the two-loop level for both lower and upper Higgs mass bounds, and can be plotted as a function of the cutoff energy Λ . Taking the top quark mass to be $175 \text{ GeV}/c^2$ and a QCD coupling $\alpha_s(m_Z) = 0.118$ the result is shown in 2.2. Lower bounds come from vacuum stability and upper bounds are derived from triviality.

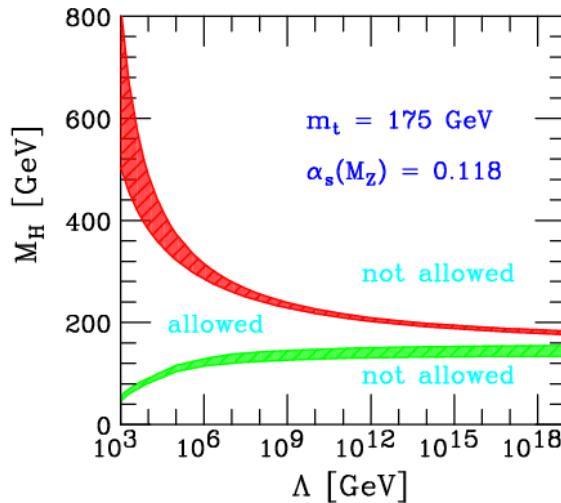


Figure 2.2: - Current theoretical uncertainties on the lower and upper m_H bounds when taking $m_t = 175 \text{ GeV}/c^2$ and $\alpha_s(m_Z) = 0.118$

2. STANDARD MODEL HIGGS BOSON

2.4.2 Experimental bounds on m_H

The hunt for the Higgs boson at present pp and $p\bar{p}$ colliders is a very exciting task, but it is also difficult, not only because of the fact that its mass is unknown and hence its preferred production and decay modes, but also because of its the small production cross-sections in the preferred mass range, explained by the small couplings of the Higgs to light fermions.

The search for the Higgs boson started years ago, but it is entering a new era in which it can finally be observed or excluded in the LHC. While triviality and vacuum stability only provide bounds on the Higgs mass as a function of the scale of new physics, more direct constraints can be derived from the experimental data collected up to now.

2.4.3 Higgs Physics at LEP

The Large Electron Positron Collider (LEP) at CERN, performed many analysis on the search for Higgs bosons of the SM and beyond [1]. It was designed with a possible Higgs boson discovery in mind, therefore, the Higgs searches were a very important part of the physics program at LEP. At the end of LEP-II they had made very detailed studies for all the possible relevant Higgs boson signatures, that allow to put a low exclusion limit on the mass of the SM Higgs boson. The LEP collaborations also carried out extensive searches for Higgs particles predicted by many scenarios beyond the Standard Model.

LEP was not the first experiment to look for the Higgs boson. The SINDRUM spectrometer at the Paul Scherrer Institute (PSI), a 590 MeV proton cyclotron, had already excluded m_H between 10 MeV/c² and 110 MeV/c² studying the decay of a pion to an electron, an electron neutrino, and a Higgs boson, decaying subsequently to a pair of electrons. The CERN-Edinburgh-Mainz-Orsay-Pisa-Siegen collaboration at the CERN Super Proton Synchrotron (SPS) had also studied the Higgs boson decay into a pair of electrons in the $K_L^0 \rightarrow \pi^0 H$. Also before LEP, the CLEO experiment, investigating decays of the Higgs boson into a pair of muons, pions and kaons produced via the flavor-changing neutral current $B \rightarrow K^0 H$, excluded the mass range 0.2 – 3.6 GeV/c⁰. Finally, the CUSB collaboration, through the search of the radiative decays of the Υ into a Higgs boson, led to exclude m_H below 5 GeV/ c^2 . As a result of these searches, the Higgs mass was already constrained when LEP started operation. In 1989, the Higgs mass was already known to be higher than 50 MeV/c².

At LEP, the production cross-sections for a low-mass Higgs boson in Z-decays were

2.4 The search for the SM Higgs Boson

very large, and that meant a lot of discovery potential for a wide mass range. The dominant production mechanism was the Bjorken process, $e^+e^- \rightarrow HZ^* \rightarrow Hff\bar{f}$; with minor contributions from other processes, like the Wilczek process $e^+e^- \rightarrow H\gamma$.

Concerning the decays, in the non-perturbative QCD domain, that is $m_H < 2 \text{ GeV}/c^2$, for Higgs masses $m_H < 2m_e$, the Higgs boson only can decay to a pair of photons via loop of W bosons. For masses below $2m_\mu$ it mostly decays to a pair of electrons, and between $2m_\mu$ and $2m_\pi$, to a pair of muons. Above that threshold, it became more difficult, as for masses below $2 - 3 \text{ GeV}/c^2$, the Higgs decays to a pair of hadrons via its interaction with two gluons through a top-quark loop or its interaction with quarks. The hadronization of these gluons gets more complex with the mass of the Higgs boson. The transition to the perturbative QCD domain comes above $\sim 2 \text{ GeV}/c^2$. In this re-

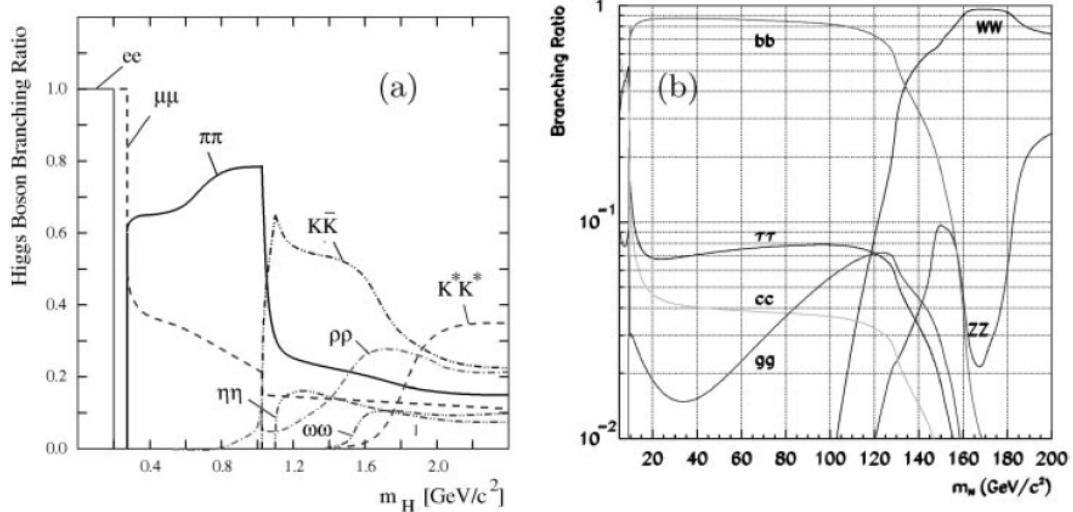


Figure 2.3: - Higgs boson branching ratios in the non-perturbative QCD low-mass region (a), and for larger masses (b).

gion, other decay modes start to gain importance. Below the $b\bar{b}$ threshold, the decays into the heaviest possible fermion pair ($c\bar{c}$ or $\tau^+\tau^-$) prevail; above the $b\bar{b}$ threshold, the dominant decay is $H \rightarrow b\bar{b}$ decay, as shown in figure 2.3.

It was possible for LEP (LEP1) to achieve more than a 95% C.L. exclusion for a SM Higgs boson with a mass below $20 \text{ GeV}/c^2$ using three topologies of the dominant $e^+e^- \rightarrow Z^* \rightarrow ZH$ production mode, one of them being background free, and the other, slightly affected by the $e^+e^- \rightarrow \gamma^*Z$ background. For higher masses, as the $b\bar{b}$ was the dominant Higgs decay, the huge amount of background coming from the hadronic decays of the Z and the small production cross-section for Higgs, make the study of Z^* decays involving hadrons or τ^\pm not suitable, hence the only studied channels were the decays to a pair of neutrinos or charged leptons (electrons or muons). A

2. STANDARD MODEL HIGGS BOSON

very small number of signal events was expected, and that, combined with the large Z decays collected, made necessary the use of more sophisticated techniques. The result of these studies led to a lower limit on the Higgs boson mass of $65.6 \text{ GeV}/c^2$.

In the second run of LEP (LEP2), up to 206.6 GeV of centre of mass energy, the dominant production process has the direct coupling of the Higgs to W or Z vector-bosons. The kinematic threshold for Higgs boson production quickly decreased around $\sqrt{s} - m_Z, \sim 10 \text{ GeV}$, as can be seen in 2.4. Higgs physics at LEP were strongly dependent on the luminosity and cross-section.

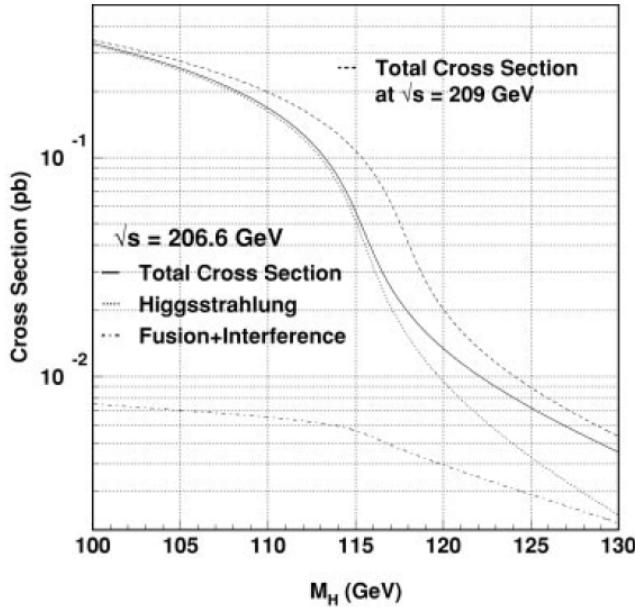


Figure 2.4: - Contribution to the total cross-section for Higgs boson production in e^+e^- collisions at $\sqrt{s} = 206.6 \text{ GeV}$

Finally, taking into account the results of the statistical combination of the searches carried out by the ALEPH, DELPHI, L3 and OPAL, LEP data set a lower limit in the Higgs boson mass excluding a signal with $m_h < 114.4 \text{ GeV}/c^2$ at 95% confidence level (CL) or higher.

Towards the end of the LEP running in the year 2000, hints for the direct observation of Higgs signal, corresponding to a Higgs-boson mass around $116 \text{ GeV}/c^2$, were detected; a mass value well compatible with the constraints derived from the precision electroweak measurements. At the LHC, this mass will be investigated again, as well as a very wide mass range, covering both, the excluded and not yet excluded masses, that will lead to the definitive discover of exclusion of the SM Higgs boson.

2.4.4 Higgs Physics at Tevatron

The Tevatron experiments CDF and D0, currently still running, are also performing SM Higgs physics. While at LEP, the discover of the Higgs boson could have been achieved for certain masses, the Tevatron can only reach exclusion for some particular mass values.

The Tevatron, a 1.96 TeV centre of mass energy $p\bar{p}$ collider, was not designed specifically for SM Higgs physics, but for top quark physics. During its twelve years of successful running not only the top quark was discovered [18; 19] but the amount of recorded collision data has been enough to exclude a Higgs boson mass range around $160 \text{ GeV}/c^2$ combining different channels and taking advantage of the on-going development of multidimensional analysis.

The main Higgs production modes at Tevatron are the ones associated with W and Z vector-bosons, and the main decay modes are the leptonic decays of the vector-bosons with H going to $b\bar{b}$, but many other signatures are being studied.

The last combined CDF and D0 Upper Limits on Standard-Model Higgs-Boson Production is shown in 2.5, giving a 95% C.L. exclusion for SM Higgs with mass m_H between 162 and 166 GeV/c^2 [2] using just the W^+W^- decay mode.

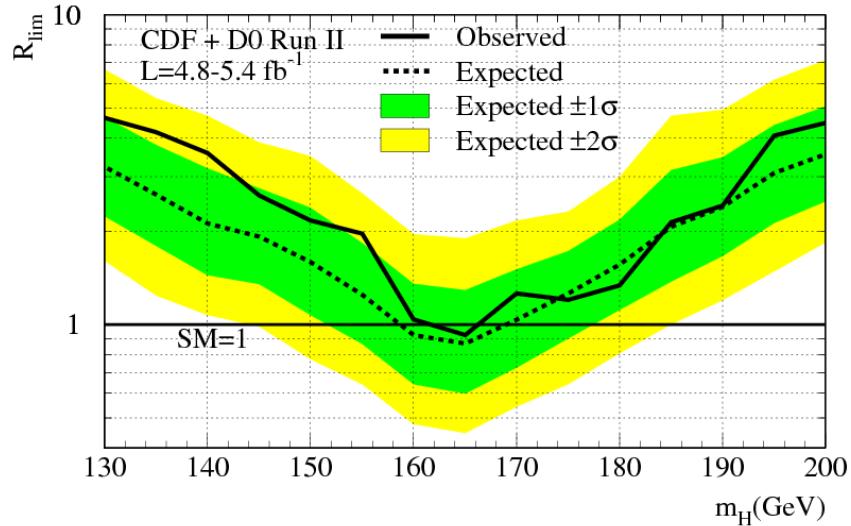


Figure 2.5: - Tevatron Higgs Combination with $L = 4.8 - 5.4 \text{fb}^{-1}$

A global fit of the data from the Z-pole experiments (ALEPH, DELPHI, L3, OPAL, SLD), from $n\bar{u} \rightarrow N$ deep inelastic scattering experiments (NuTeV, CHARM, CCFR, CDHS), from low energy $\nu - e$ scattering experiments (CHARM II) from atomic parity

2. STANDARD MODEL HIGGS BOSON

violation experiments (on cesium and thallium), and from top quark production (CDF, D0) provides very strong constraints on the range of m_H under the assumption that there is no physics other than that of the SM near to the weak scale. Assuming no physics beyond the SM, $m_H = (98^{+57}_{-38}) \text{ GeV}/c^2$ is obtained.

Only the direct observation of the Higgs boson will be the proof of its existence. The

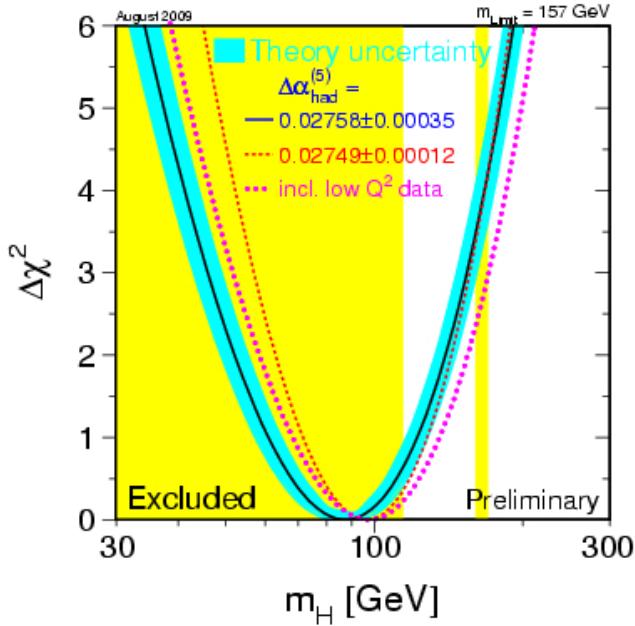


Figure 2.6: - $\Delta\chi^2$ curve derived from high- Q^2 precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs-boson mass, assuming the Standard Model to be the correct theory of nature.

figure 2.6 shows the $\Delta - \chi^2$ curve derived from high- Q^2 precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs-boson mass, assuming the Standard Model to be the correct theory of nature. The preferred value for its mass, corresponding to the minimum of the curve, is at $87 \text{ GeV}/c^2$, with an experimental uncertainty of $+35$ and $-26 \text{ GeV}/c^2$, at 68% confidence level derived from $\Delta\chi^2 = 1$ for the black line, thus not taking the theoretical uncertainty shown as the blue band into account. Precision electroweak measurements tell that the mass of the Standard-Model Higgs boson is lower than about $157 \text{ GeV}/c^2$, one-sided 95% confidence level upper limit derived from $\Delta\chi^2 = 2.7$ for the blue band, thus including both the experimental and the theoretical uncertainty. The limit increases to $186 \text{ GeV}/c^2$ when including the LEP-2 direct search limit of $114 \text{ GeV}/c^2$ shown in yellow. The LEP Electroweak Working Group keeps track of the new Higgs searches and updated results

and plots can be consulted in LEPEWWG.

An updated global analysis within the Standard Model (SM) of all relevant electroweak precision and Higgs boson search data can be consulted here [3], including the most recent results on the top quark and W boson masses, updated constraints on the strong coupling constant and other low energy measurement. Incorporating also the previously referenced latest results from the Tevatron together with the older constraints from LEP2, the result is a narrow 90% CL preferred mass window for the Higgs mass,

$$115 \leq m_H \leq 148 GeV/c^2 \quad (2.39)$$

The combination of experimental and theoretical bounds on m_H just discussed, set the pace for the upcoming Higgs physics studies, providing the mass range in which to search for the Higgs boson at the LHC.

2.5 Beyond the Standard Model

Despite the good agreement with most of the experimental evidence, the SM is unfortunately not sufficient to completely describe different aspects, like the mass of the neutrinos [20], or the dark matter, weakly interacting matter not described by the SM.

At theoretical level there are also a number of issues where the Standard Model is found not to be satisfactory:

- The quantization of charge.
- The observed gap between the masses of the quarks and leptons in the different families
- The considerable variation in the mixing angles between mass eigenstates and interaction eigenstates of the various quarks.
- The baryon number asymmetry.
- Gravity.
- Fine-tuning of the Higgs boson mass, also called the hierarchy problem.

To deal with these problems, a number of theories have been proposed, in the context of what is usually referred to as physics beyond the Standard Model. Examples of such theories include Technicolor [21], Little Higgs models [22] and theories involving extra dimensions [23].

2. STANDARD MODEL HIGGS BOSON

2.5.1 Supersymmetry (SUSY)

Supersymmetry [24] (SUSY) proposes a symmetry between bosonic and fermionic fields implying that each fermion has a corresponding bosonic partner and vice versa, so that the number of fermionic states is equal to that of bosonic states. If supersymmetry were an exact symmetry, then all supersymmetric particles would have the same properties as their ordinary partner except for the spin. This would solve the hierarchy problem, since for an unbroken Supersymmetry the quadratic divergencies caused by Standard Model particles to the Higgs mass would exactly cancel out with the corrections from their corresponding supersymmetric partners, which would have the same mass.

Part of the appeal of Supersymmetry is also that it opens the door for a Grand Unified Theory (GUT), that is a theory in which the electromagnetic, weak and strong forces all unite at a higher energy scale. Such a unification, not compatible with the experimental measurement of the coupling constants in the SM, is possible to achieve in the presence of Supersymmetry. Another interesting feature of many supersymmetric models is that they can provide a very good candidate for the dark matter in the universe.

Supersymmetry, however, cannot be an exact symmetry, as in that case supersymmetric particles would have the same mass as their Standard Model partners, and then they should already have been experimentally detected, which is not the case. If Supersymmetry exists is therefore said to be broken. It is out of the scope of this thesis to discuss supersymmetry in detail, but there is many bibliography on this topic that can be consulted, for example [25].

2.5.2 The Minimal Supersymmetric Extension of the Standard Model (MSSM)

The MSSM (Minimal Supersymmetric Extension of the Standard Model) is the supersymmetric version of the Standard Model with the minimum number of fields and couplings.

Still, it introduces new supersymmetric partners to all of the Standard Model particles. For the SM fermions, each helicity state has its own superpartner, such that for every fermion there exist two super-partners, yielding the selectrons ($\tilde{e}_{L,R}^\pm$), smuons ($\tilde{\mu}_{L,R}^\pm$), staus ($\tilde{\tau}_{L,R}^\pm$), sneutrinos ($\tilde{\nu}$) and squarks ($\tilde{q}_{L,R}$). On the other hand, the superpartners of the SM bosons mix, yielding two charginos ($\tilde{\chi}_{1,2}^\pm$) and four neutralinos ($\tilde{\chi}_{1,2,3,4}^0$), as well as gluinos (\tilde{g}).

The MSSM requires *two* Higgs doublets (as opposed to one in the SM), with hypercharge $Y = \pm 1$, as the mixing of left-handed chiral superfields with their conjugates is not allowed, and the sum of the hypercharge of all fermions in the Standard Model is zero, thereby exactly cancelling the Adler-Bell-Jackiw, or triangle, anomaly [26; 27]. With only one Higgs doublet in the MSSM, there would be only one higgsino with

$Y = 1$ and thus the sum of the hypercharges would not be zero, whereas two Higgs doublets imply the existence of a second higgsino with $Y = -1$ which would cancel out the first one. The Higgs sector of the MSSM is the so-called Type-II Two Higgs Doublet Model (2HDM).

In order to accommodate the above particle spectrum, 105 parameters need to be added to the existing 19 parameters of the Standard Model (21 masses, 36 real mixing angles and 40 CP-violating phases in the squark and slepton sector, 5 real parameters and 3 CP-violating phases in the gaugino-Higgsino sector). This yields a total of 124 parameters in the MSSM, if no assumptions at all are made as to the nature of supersymmetry breaking.

This has led to the emergence of more restricted versions of the MSSM where a number of these parameters are fixed, usually through considerations relating to the mechanism of supersymmetry breaking in a specific model. One of the most popular such models is the minimal supergravity model or mSUGRA. Here, the gaugino as well as all scalar masses and the trilinear couplings (which relate to the supersymmetry breaking) are assumed to be equal at the unification scale. With this assumption, only four free parameters and a sign remain, in addition to those of the Standard Model, making it very attractive, in particular when discussing supersymmetry phenomenology at colliders.

2.5.3 Higgs Bosons in the MSSM

The MSSM is described by the Type-II Two Higgs Doublet Model (2HDM). However, the 2HDM does not itself require the MSSM, in fact a Type-II 2HDM could be realized in nature even if Supersymmetry is not.

Three of the four degrees of freedom of the scalar Higgs doublet in the Standard Model are absorbed into the longitudinal components of the three massive vector fields in the unitary gauge, leaving a single massive scalar field, a single physical Higgs boson. Whereas in the 2HDM, having two Higgs doublets gives rise to eight degrees of freedom, of which three are absorbed by the massive vector fields giving five massive scalar fields, that is, five physical Higgs particles. For the CP-conserving case, these are:

- two neutral CP-even scalars, typically denoted h^0 and H^0 , where $m_H < m_h$,
- one CP-odd scalar A^0 ,
- two charged scalars H^\pm .

At tree-level in the MSSM only two parameters are needed to describe the properties of these five particles: $\tan\beta$ and the mass of one of them.

Higgs physics and SUSY studies will play a major role in high energy physics in the years to come. In the following, mainly SM Higgs physics will be discussed, but it is

2. STANDARD MODEL HIGGS BOSON

important to keep in mind the fact that, even with Higgs the SM is not complete, and many questions will remain open. A comprehensive discussion on Higgs physics in the SM and beyond is given in [28] amongst others.

In this thesis a detailed strategy on the search of a SM Higgs boson with a mass in the $120 \leq m_H \leq 200 \text{ GeV}/c^2$ in the $H \rightarrow WW^*$ channel using the CMS experiment at the LHC will be presented. Before that, in the next chapter, the Large Hadron Collider and the Compact Muon Solenoid will be introduced.

3

The CMS experiment at LHC

3.1 The Large Hadron Collider

The Large Hadron Collider, LHC [29], is the largest particle accelerator ever built, and also the one with the highest energy. It is a circular collider built at CERN under French and Swiss territory using the existing LEP tunnel. The circumference of the LHC ring is roughly 26.7 *Km* and it is located at a mean depth of 100 *m* underground, below the Jura mountains and the Lake Geneva. It was designed to provide proton-proton collisions with a center-of-mass energy of 14 TeV and a luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. It will also provide heavy ion collisions (Pb) with a center-of-mass energy of more than 1000 TeV and a luminosity exceeding 10^{27} . Some design parameters of the LHC are shown in table 3.1.

The first *pp* collisions, were recorded the 23rd of November 2009, successfully ending a very long journey and opening a new exciting page in the history of high energy physics.

The design of the started in 1981, but the project was not approved by CERN Council until December 1994. The building and commissioning of the LHC was a huge technological effort, with many ups and downs. The magnitude of the experiment, the biggest ever built, caused some unpredictable delays; and during fifteen years the design has considerably evolved and matured, but the original idea stayed untouched. In the years to come, LHC will reach its design energy and luminosity, and particle physics will enter new territories that have never been explored before.

The LHC is supplied with protons and *Pb* ions from the existing injector chain comprising Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS) and Super Proton Synchrotron (SPS). These accelerators were upgraded to meet the very stringent needs of the LHC. The figure 3.1 shows a schematic view of the accelerator complex, in which the LHC is the last element.

Inside the LHC the beam of particles circulate in a ultrahigh vacuum tube, accelerated

3. THE CMS EXPERIMENT AT LHC

Large Hadron Collider	
circumference	26659 m
depth	50 - 175 m
total number of magnets	9600
number of main dipoles	1232
number of quadrupoles	392
temperature	1.9 K (-271.3C)
beam vacuum pressure	10^{-13} atm
nominal p energy	7 TeV
center-of-mass energy	14 TeV
design luminosity	$10^{34} \text{ cm}^{-2}\text{s}^{-1}$
bunches per proton beam	2808
protons per bunch	1.1×10^{11}
turns per second	11245
collisions per second	600 million
length of each dipole	15 m
weight of each dipole	~ 35 t
dipole field	8.33 T

Table 3.1: Design parameters of the Large Hadron Collider.

and focused by magnets. There is a large variety of magnets in the LHC ring, dipoles, quadrupoles, sextupoles, decapoles, etc. The dipoles were the biggest technological challenge for the design of the LHC. The dipoles are superconducting electromagnets that provide a 8.33T magnetic field over their length, and for that, they have to be maintained at 1.9 K of temperature, therefore, the LHC has the largest cryogenic system in the world, with superfluid helium pumped into the magnet systems. The refrigeration process happens in a sequential way, first helium is cooled down to 4.5 K, then it is injected into the magnet cold masses, and, once the magnets are filled, the refrigeration units bring the temperature to 1.9 K .

The protons circulate around the ring in the LHC in well-defined bunches. Under nominal operating conditions, LHC will have 2808 bunches in each beam, with around 10^{11} protons in each bunch. The size of the bunch is not constant around the ring, it gets squeezed and expanded around the LHC; and, as they approach the collision points, the bunches get squeezed to about $16 \mu\text{m}$, to allow the proton-proton collision to happen with higher probability. The luminosity of the machine increases with the number of bunches. At full luminosity, LHC will have a bunch spacing of 25 ns , that is, around 7 m , corresponding to a frequency of 40 MHz . The average crossing rate is equal to the total number of bunches multiplied by the number of turns round the LHC per second: $2808 \times 11245 = 31.6MHz$. the particles are so tiny that the chance of any two colliding is very small. When the bunch cross, there will be a maximum of about 20 collisions between 200 billion particles. Bunches will cross about 30 million

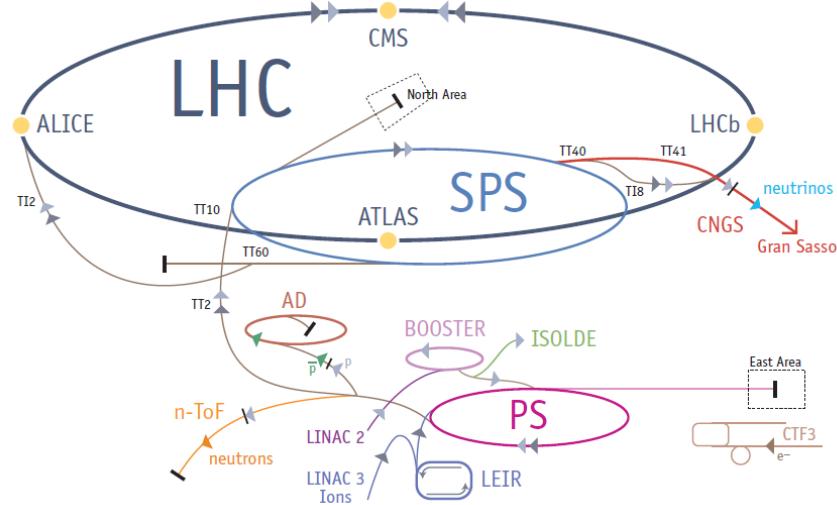


Figure 3.1: - Sketch of the accelerator complex at CERN

times per second, hence, at nominal conditions, LHC will generate up to 600 million particle collisions per second.

Detectors are installed in the collision points all around the LHC ring. There are six experiments installed at the LHC, in alphabetical order: ALICE (A Large Ion Collider Experiment) [30], ATLAS (A Toroidal LHC ApparatuS) [31], CMS (Compact Muon Solenoid) [32], LHCb (Large Hadron Collider beauty) [33], LHCf (Large Hadron Collider forward) and TOTEM (TOTal Elastic and diffractive cross-section Measurement). ALICE, ATLAS, CMS and LHCb are installed in four huge underground caverns build around the four collision points of the LHC, while and LHCf and TOTEM are close to the main detectors, ATLAS and CMS respectively.

ATLAS and CMS are the main experiments in the LHC, made for general purposes with the same physics goals but different technical solutions and design, the rest of the experiments are smaller and specialized in different topics. This thesis is being realized in the context of the CMS experiment. In the following, the main details of the CMS detector will be discussed.

3.2 The CMS experiment

The Compact Muon Solenoid, CMS [34], is a general-purpose detector built around a superconducting solenoid that generates a magnetic field of 4 T. CMS is with ATLAS one of the two main experiments at the LHC and it is designed to cover the widest possible range of physics. The CMS Collaboration is composed for more than 2000 people, from 181 institutions in 38 countries. Some general parameters of the CMS

3. THE CMS EXPERIMENT AT LHC

detector are shown in table 3.2.

Compact Muon Solenoid	
Length	21.6 m
Diameter	14.6 m
Weight	12500 t
Design	Barrel plus End-caps
Material cost	500 MCHF
Location	Cessy, France

Table 3.2: Design parameters of the Compact Muon Solenoid experiment.

The project of CMS was presented in October 1990 at the LHC workshop in Aachen, and totally defined in the following years. In the design of CMS, being the LHC a high luminosity proton-proton collider, it was natural to give priority to a very good muon detection system. That, together with the desire to make it compact, led to the choice of a strong magnetic field. The only practical magnet able to generate a strong magnetic field in a compact scenario is a solenoid. It was then decided for CMS to have a large superconducting solenoid, of about 14 meters long with a radius of 3 meters, providing a 4 T magnetic field. That guarantees a good momentum resolution for high momentum muons (> 1 TeV) up to rapidities of 2.5. The coil doesn't affect the performance of the calorimeter, as the big radius of the magnet allows the tracking and the calorimeters to be nicely placed inside the solenoid.

3.3 Detector Overview

The overall layout of CMS is shown in Figure 3.2. It is a cylindrical detector 13 meters long and 5.9 meters of inner diameter 4 T superconducting solenoid. The return field is large enough to saturate 1.5 m of iron, allowing 4 muon stations to be integrated, ensuring robustness and full geometric coverage. The magnet is large enough to accommodate the inner tracker and the calorimetry inside.

The coordinate system adopted by CMS, that will be used in the following sections, has the origin centered at the nominal collision point inside the experiment, the y -axis pointing vertically upward, the x -axis pointing radially inward toward the center of the LHC, and the z -axis along the beam direction towards the Jura mountains from the LHC Point 5. The azimuthal angle ϕ is measured from the x -axis in the $x - y$ plane, the polar angle θ is measured from the z -axis. Pseudorapidity η , is defined as $\eta = -\ln \tan(\frac{\theta}{2})$. The momentum and energy measured transverse to the beam direction, p_T and E_T , are computed from the x and y components.

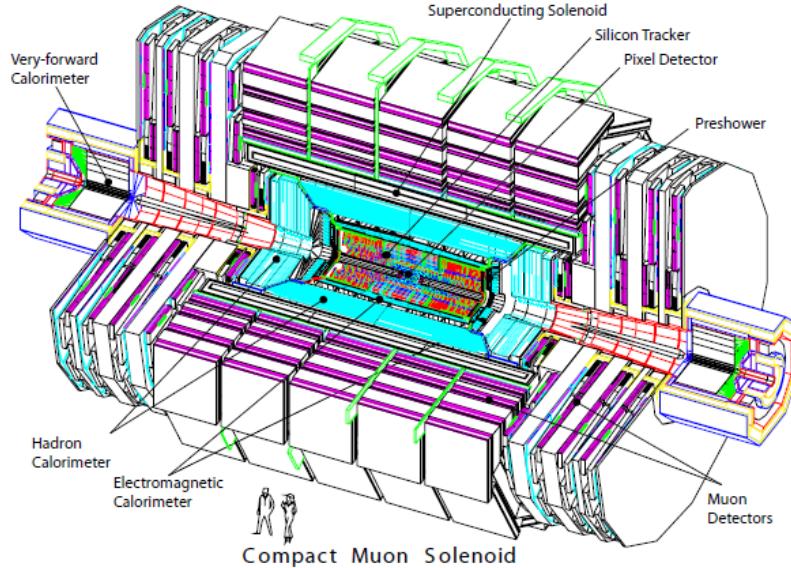


Figure 3.2: - A general view of the Compact Muon Solenoid

Details on the different components of CMS can be found in the next sub-sections.

3.3.1 Magnet

The CMS magnet [35] is the central device around which the experiment is built, the magnet is used to bend the paths of particles emerging from high-energy collisions in the LHC. To ensure the appropriate performance of the muon system, a momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1\text{TeV}/c$ is achieved by CMS with a large Superconducting Solenoid which parameters are given in table 3.3.

Superconducting Solenoid	
Field	4 T
Inner Bore	5.9 m
Length	12.9 m
Number of Turns	2168
Current	19.14 kA
Stored Energy	2.6 GJ
Hoop stress	64 atm

Table 3.3: Parameters of the CMS superconducting solenoid.

3. THE CMS EXPERIMENT AT LHC

The CMS solenoid was built using the same technique previously used in the construction of the large solenoids, ALEPH and DELPHI at LEP and H1 at HERA, using a high-purity aluminium-stabilized conductor and indirect cooling by thermosyphon, together with a full epoxy impregnation. However, in this case, the large increase in parameters such as magnetic field, Ampere-turns, forces and stored energy, led to major changes. A four-layer winding has been adopted using a novel conductor with a larger cross-section that can withstand a hoop stress of 64 atmospheres. The conductor carries a current of 20 kA , and has a compound structure. The Rutherford-type cable is co-extruded with pure aluminium, that acts as thermal stabilizer, this is then electron-beam-welded to 2 plates, made of high-strength aluminium alloy, for the mechanical reinforcement. The overall conductor cross-section is $64 \times 22\text{mm}^2$, and it was manufactured in 20 continuous lengths, of 2.65 Km each, 4 lengths were wound to make each of the 5 coil modules. These modules were assembled and connected together, as well as the whole magnet, in the surface experiment hall, SX5 at Point 5.

3.3.2 Inner Tracking System

The CMS tracker records the paths taken by charged particles, it reconstructs the paths of high-energy muons, electrons and hadrons, and see tracks coming from the decay of b quarks. For that, the tracker needs to record particle paths accurately yet disturbing the particle as little as possible. It is also the inner most layer of the detector and so receives the highest volume of particles: the construction materials were therefore carefully chosen to resist radiation. The final design consists of a tracker made entirely of silicon, pixels, located at the very core of the detector, and microstrip detectors surrounding it. As particles travel through the tracker, the pixels and microstrips produce electric signals that are amplified and detected.

Three regions can be delimitated on the inner Tracking System [36; 37] of CMS:

- The closest region to the interaction vertex, where the particle flux is the highest ($\approx 10^7/\text{s}$ at $r \approx 10\text{cm}$, nominal conditions), is where pixel detectors are placed. The size of a pixel is about $100 \times 150\mu\text{ m}^2$ giving an occupancy of $\approx 10^{-4}$ per pixel per LHC crossing.
- In the intermediate region, $20 < r < 55\text{ cm}$, the reduced particle flux allows the use of silicon microstrip detectors with a minimum cell size of $10\text{ cm} \times 80\text{ }\mu\text{m}$, which traduces on an occupancy of $\approx 2 - 3\%$ per LHC crossing.
- In the outermost region of the inner tracker, for $r > 55\text{ cm}$, larger-pitch silicon microstrips are used, with a maximum cell size of $25\text{ cm} \times 180\text{ }\mu\text{m}$, keeping the occupancy to about 1%.

The occupancy is expected to be at the level of 1% in the pixel detectors, and less than 20% in the outer silicon strip detectors, even for heavy-ion collisions, leading to a

3.3 Detector Overview

efficient track reconstruction in high density scenarios.

The layout of the CMS tracker is shown in figure 3.3. In the barrel region, close

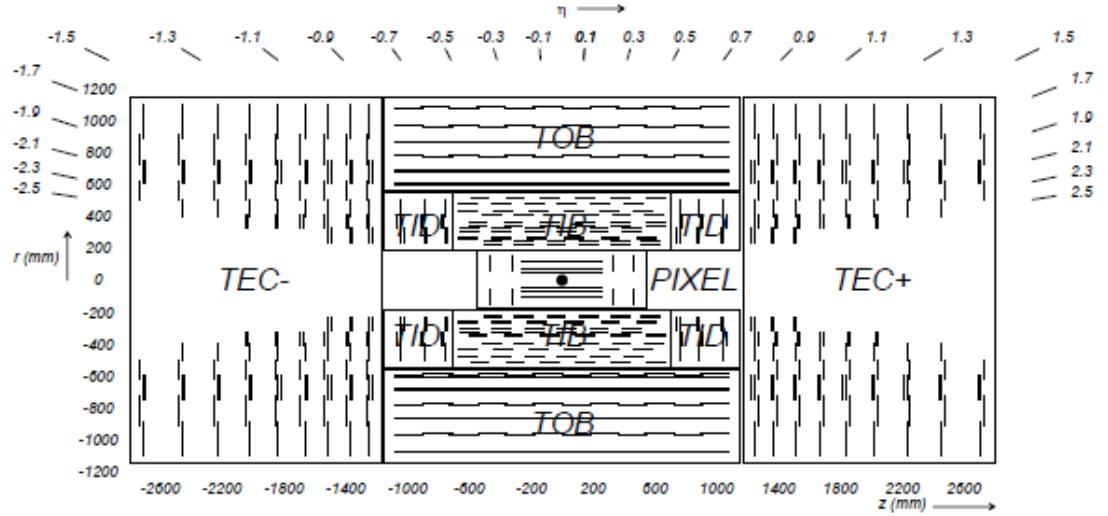


Figure 3.3: - Schematic cross-section through the CMS tracker, each line represents a detector module and double lines indicate back-to-back modules which deliver stereo hits.

to the interaction vertex, 3 layers of pixel detectors are placed at distances of 4.4, 7.3, and 10.2 cm, then, at r between 20 and 110 cm, silicon microstrips are placed. The forward region has 2 pixel and 9 microstrip layers in both endcaps. The barrel part is divided into the Inner and the Outer barrel, with the inner being shorter to avoid shallow crossing angles, and there are 3 additional inner disks in the transition region between barrel and endcaps on each side of the Inner barrel.

Inner Tracker	
Outer radius	110 cm
Length	540 cm
Pixel total area	$\approx 1 \text{ m}^2$
Silicon strips total area	$\approx 200 \text{ m}^2$
Pseudorapidity coverage	$ \eta < 2.4$
Number of pixels	66 million
Number of silicon strips	9.6 million

Table 3.4: Parameters of CMS tracking system

Some parameters of the inner tracking system of CMS are presented in 3.4.

The barrel tracker region is divided in two parts, the Tracker Inner Barrel, TIB, and the

3. THE CMS EXPERIMENT AT LHC

Tracker Outer Barrel, TOB. The TIB, made of 4 layers, covers up to $|z| < 65\text{ cm}$, using silicon sensors with a thickness of $320\text{ }\mu\text{m}$ and a strip pitch that varies from 80 to 120 μm . The first 2 layers are made with stereo modules that provide measurement in both $r - \phi$ and $r - z$ coordinates, the single point resolution in the $r - \phi$ direction is between 23 and 24 μm , and 230 μm in z . The TOB comprises 6 layers with a half-length of $|z| < 110\text{ cm}$. The radiation levels are smaller in this region, so thicker silicon sensors can be used to maintain a good S/N ratio for longer strip length and wider pitch. The strip pitch varies from 120 to 180 μm . The first two layers of the TOB also provide stereo measurements in both $r - \phi$ and $r - z$ coordinates, with a single point resolution between 35 and 52 μm for the $r - \phi$ direction and 530 μm in z .

The endcaps are divided into the Tracker End Cap, TEC, and the Tracker Inner Disks, TID. Each TEC comprises 9 disks that extend into the region $120 < |z| < 280\text{ cm}$, and each TID comprises 3 small disks that fill the gap between the TIB and the TEC. The TEC and TID modules are arranged in rings, centered on the beam line, and have strips that point towards the beam line, having a variable pitch. The first 2 rings of the TID and the innermost 2 rings and the fifth ring of the TEC have stereo modules. The thickness of the sensors is 320 μm for the TID and the 3 innermost rings of the TEC and 500 μm for the rest of the TEC.

The entire silicon strip detector consists of almost 15400 modules mounted on carbon-fibre structures and housed inside a temperature controlled outer support tube. The operating temperature is around -20°C .

The pixel detector has 3 layers in the barrel, located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, that have a length of 53 cm; and 2 disks in each endcap, extending from 6 to 15 cm in radius, placed on each side at $|z| = 34.5\text{ cm}$ and 46.5 cm .

To achieve an optimal vertex position resolution, a design with an almost square pixel shape of $100 \times 150\text{ }\mu\text{m}^2$ in both the (r, ϕ) and the z coordinates has been adopted. The barrel has 768 pixel modules arranged into half-ladders of 4 identical modules each. The large Lorentz effect (Lorentz angle is 23°) improves the $r - \phi$ resolution through charge sharing. The endcap disks are assembled in a turbine-like geometry with blades rotated by 20° to also benefit from the Lorentz effect. The endcap disks comprise 672 pixel modules with 7 different modules in each blade.

The spatial resolution is measured to be about 10 μm for the $r - \phi$ measurement and about 20 μm for the z measurement. The detector is readout using approximately 16000 readout chips, which are bump-bonded to the detector modules.

3.3.3 Electromagnetic Calorimeter

To measure the energies of emerging particles, in particular electrons and photons, CMS uses a electromagnetic calorimeter, ECAL [38; 39; 40]. The ECAL is composed

3.3 Detector Overview

by lead tungstate crystal primarily made of metal. Heavier than stainless steel, it is highly transparent and scintillates when electrons and photons pass through it producing light in fast, short, well-defined photon bursts that allow for a precise, fast and fairly compact detector. Photodetectors specially designed to work within the high magnetic field, are glued onto the back of each crystal to detect the scintillation light and convert it to an electrical signal that is amplified and sent for analysis.

The ECAL is hermetic and homogeneous, with a fine granularity, composed by 61200 lead tungstate ($PbWO_4$) crystals in the barrel part, and 7324 crystals in each of the 2 endcaps. These lead tungstate scintillating crystals have short radiation and Moliere lengths, and they are fast and resistant to radiation. Silicon avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum photodiodes (VPTs) in the endcaps. In addition, the sensitivity of both the crystals and the APD response to temperature changes requires a temperature stability. The layout of the ECAL is shown in figure 3.4.

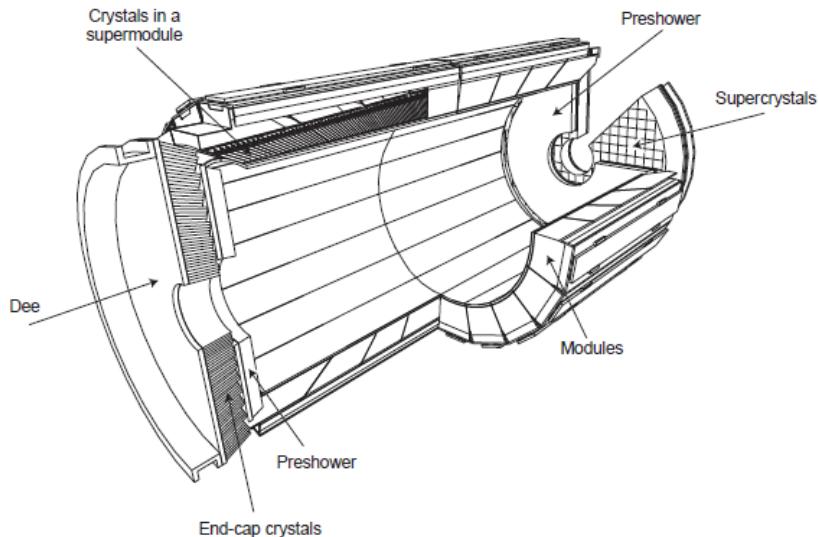


Figure 3.4: - Layout of the CMS electromagnetic calorimeter showing the arrangement of crystal modules, supermodules and endcaps, with the preshower in front.

The barrel section (EB) has an inner radius of 129 cm. It is structured as 36 identical super-modules, each covering half the barrel length, corresponding to $|\eta| < 1.479$. The crystals are quasi-projective and cover 0.0174 (that is 1°) in $\Delta\phi$ and $\Delta\eta$. The crystals have a front face cross-section of $\approx 22 \times 22 \text{ mm}^2$ and a length of 230 mm, corresponding to $25.8 \chi_0$. The endcaps section (EE) is located a distance of 314 cm from the vertex, covering a pseudorapidity range between $1.479 < |\eta| < 3.0$, are each structured as 2 dees, consisting of semi-circular aluminium plates from cantilevered structural units

3. THE CMS EXPERIMENT AT LHC

of 5×5 crystals, known as supercrystals. A preshower device is placed in front of the crystal calorimeter over much of the endcap pseudorapidity range, composed by 2 planes of silicon strip detectors, with a pitch of 1.9 mm , that lie behind disks of lead absorber at depths of $2\chi_0$ and $3\chi_0$.

3.3.4 Hadron Calorimeter

The Hadron Calorimeter (HCAL) [41] measures the energy of hadrons, like protons, neutrons, pions or kaons. Additionally it provides indirect measurement of the presence of non-interacting, uncharged particles such as neutrinos, which do not leave record of their presence in any part of the CMS detector. To spot these, the HCAL must be hermetic. If we see particles shoot out one side of the detector, but not the other, with an imbalance in the momentum and energy measured in the transverse plane, it indicates that invisible particles are being produced.

The HCAL is a sampling calorimeter, it finds a particles position, energy and arrival time using alternating layers of absorber and fluorescent scintillator materials that produce a rapid light pulse when the particle passes through. Optic fibres collect up this light and feed it into readout boxes where photodetectors amplify the signal. When the amount of light in a given region is summed up over many layers of tiles in depth, called a tower, this total amount of light is a measure of a particles energy.

The HCAL is massive and thick, made of staggered layers, with no gaps in direct lines. Fitting the HCAL into CMS was a challenge, as the showers, cascades of particles produced when a hadron hits the dense absorber material, are large. The fact that most of the CMS calorimetry is located inside the magnet coil, strongly influenced the design, very dependant of the choice of magnet parameters. An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity for the missing E_T measurement, relative to the invisible particles, that escape detection. The HCAL is organized into barrel (HB and HO), endcap (HE) and forward (HF) sections. There are 36 barrel wedges, of 26 tonnes each. These form the last layer of detector inside the magnet coil whilst a few additional layers, the outer barrel (HO), sit outside the coil, ensuring no energy leaks out the back of the HB undetected. Similarly, 36 endcap wedges measure particle energies as they emerge through the ends of the solenoid magnet. The two hadronic forward calorimeters (HF) are positioned at both ends of CMS, to pick up the particles coming out of the collision region at shallow angles. These receive the bulk of the particle energy contained in the collision so they must be radiation-resistant and use different materials to the other parts of the HCAL. In, figure 3.5, a longitudinal view of the different parts of the HCAL is shown.

Brass has been chosen as absorber material for its short interaction length, and for being easy to machine and non-magnetic. Maximizing the amount of absorber before the magnet requires keeping to a minimum the amount of space devoted to the active

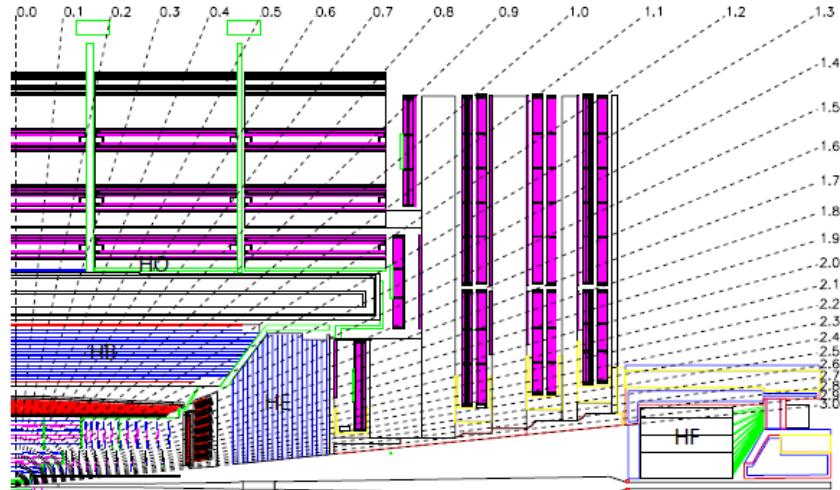


Figure 3.5: - Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters.

medium. The tile/fibre technology makes for an ideal choice. It consists of plastic scintillator tiles read out with embedded wavelength-shifting (WLS) fibres. The WLS fibres are spliced to high attenuation-length clear fibres outside the scintillator that carry the light to the readout system. The photodetection readout is based on multi-channel hybrid photodiodes (HPDs). The absorber structure is assembled by bolting together precisely machined and overlapping brass plates so as to leave space to insert the scintillator plates, which have a thickness of 3.7 mm. The overall assembly enables the HCAL to be built with essentially no uninstrumented cracks or dead areas in ϕ . The gap between the barrel and the endcap HCAL, through which the services of the ECAL and the inner tracker pass, is inclined at 53° and points away from the center of the detector.

In the barrel, the HB consists of 32 towers covering $-1.4 < \eta < 1.4$, resulting in 2304 towers with a segmentation $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The HB is constructed in 2 half barrels, and it is read out as a single longitudinal sampling. There are 15 brass plates, each with a thickness of about 5 cm, plus 2 external stainless steel plates for mechanical strength. Still in the barrel, the HO detector contains scintillators with a thickness of 10 mm, which line the outside of the outer vacuum tank of the coil and cover the region $-1.26 < |\eta| < 1.26$. The tiles are grouped in 30°-sectors, matching the ϕ segmentation of the DT chambers. They sample the energy from penetrating hadron showers leaking through the rear of the calorimeters and so serve as a tail-catcher after the magnet coil. They increase the effective thickness of the HCAL to over 10 interaction lengths, reducing the tails in the energy resolution function. The HO also

3. THE CMS EXPERIMENT AT LHC

improves the missing E_T resolution. It is located inside the barrel muon system and is hence constrained by the geometry and construction of that system, that will be largely discussed in the following sections. It is divided into 5 rings, in the ring 0 it has 2 scintillator layers on either side of an iron absorber with a thickness of about 18 cm, at radial distances of 3.850 m and 4.097 m, respectively. The other mobile rings have single layers at a radial distance of 4.097 m. Each ring covers 2.5 m in z . HO scintillators follow the HCAL barrel tower geometry in η and ϕ .

In the endcaps, each HE, consists of 14 η towers covering the region $1.3 < |\eta| < 3.0$. The segmentation varies in the outer and innermost parts, being the total number of HE towers is 2304. In the forward region, for pseudorapidities between 3.0 and 5.0, the steel/quartz fibre HF calorimeter operates.

3.3.5 Muon System

Muon detection is a main task for CMS, as they appear in many interesting decays of new physics or in Higgs boson decay signatures that can be produced by the LHC, like the one that is studied in this thesis. As muons can penetrate several meters of material without interacting, they are not stopped by any of CMS's calorimeters. Therefore, chambers to detect muons are placed at the very edge of the experiment where they are the only particles likely to register a signal conforming the Muon detection System [42].

Centrally produced muons are measured three times, in the inner tracker, after the coil, and in the return flux. In the muon system, a particle is measured by fitting a curve to hits among the four muon stations, which sit outside the magnet coil and are interleaved with iron return yoke plates. By tracking its position through the multiple layers of each station, combined with tracker measurements, the particle's path can be precisely traced.

The momentum is essentially determined by the muon bending angle at the exit of the 4T coil, taking the interaction point as the origin of the muon. The resolution of this measurement is dominated by multiple scattering in the material before the first muon station up to p_T values of 200 GeV/c, when the chamber spatial resolution starts to dominate. For low-momentum muons, the best momentum resolution is given by the resolution obtained in the silicon tracker. However, the muon momentum resolution for high p_T muons can be improved, extrapolating the muon trajectory beyond the return yoke back to the beam-line, combining the inner tracker and muon detector measurements.

There are three types of gaseous detectors used to identify and measure muons in CMS. The choice of the detector technologies has been driven by the very large surface to be covered and by the different radiation environments. In the barrel ($|\eta| < 1.2$), where the neutron induced background is small and the muon rate and the residual magnetic field in the chambers are low, drift tube chambers (DT) are used. In the

3.3 Detector Overview

endcaps, where both the muon rate and the neutron induced background rate are high, and the magnetic field is also high, cathode strip chambers (CSC) are located, covering the region up to $|\eta| < 2.4$. In addition to this, resistive plate chambers (RPC) are used in both the barrel and the endcap regions. The RPCs are operated in avalanche mode to ensure good operation at high rates (up to 10 kHz/cm^2) and have double gaps with a gas gap of 2 mm . RPCs provide a fast response with good time resolution but with a coarser position resolution than the DTs or CSCs. RPCs can unambiguously identify the correct bunch crossing. The DTs or CSCs and the RPCs operate within the first level trigger system, providing 2 independent and complementary sources of information. The complete system results in a robust, precise and flexible trigger device. The Layout of the muon system is presented in figure 3.6.

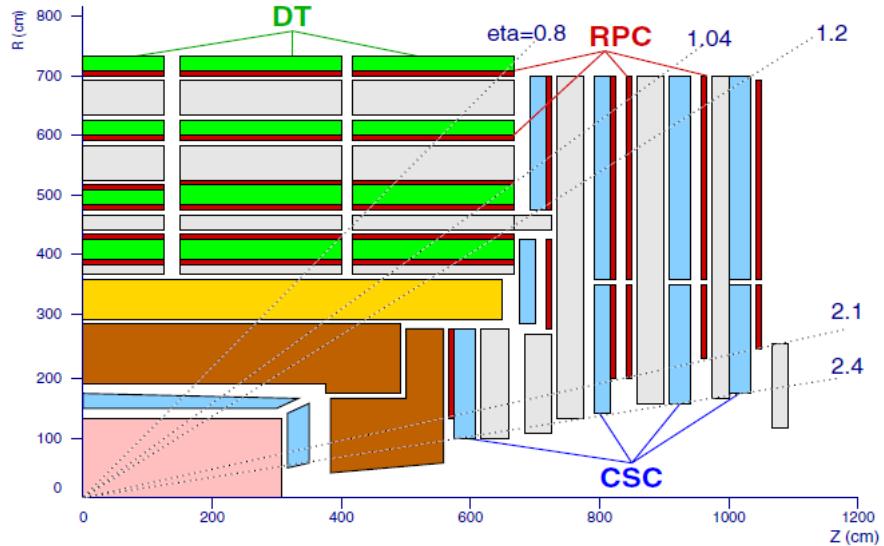


Figure 3.6: - Layout of the CMS muon system

In the Muon Barrel (MB) region, 4 stations of detectors are arranged in cylinders interleaved with the iron yoke. The segmentation along the beam direction follows the 5 wheels of the yoke (labeled YB-2 for the farthest wheel in $-z$, and YB+2 for the farthest is $+z$). In each of the endcaps, the CSCs and RPCs are arranged in 4 disks perpendicular to the beam, and in concentric rings, 3 rings in the innermost station, and 2 in the others. In total, the muon system contains of order 25000 m^2 of active detection planes, and nearly 1 million electronic channels.

The Barrel Detector, consists of 250 chambers organized in 4 layers, called stations, labeled MB1, MB2, MB3 and MB4, being the MB1 the closest to the beam axis, and the MB4 the outermost; within 5 wheels, divided in 12 sectors of 30° azimuthal angle each. Chambers in different stations are staggered so that a high- p_T muon produced near a sector boundary crosses at least 3 out of the 4 stations. There are 12 cham-

3. THE CMS EXPERIMENT AT LHC

bers in each of the 3 inner layers. In the 4th layer, the top and bottom sectors host 2 chambers each, thus leading to a total of 14 chambers per wheel in this outermost layer. The MB1, 2 and 3 chambers consist of 12 planes of aluminium drift tubes; 4 $r - \phi$ measuring planes in each of the 2 outermost superlayers, separated by about 20 cm and sandwiching a z -superlayer comprising 4 z -measuring planes. The MB4 station does not contain the z -measuring planes. The maximum drift length is 2.0 cm and the single point resolution is $\approx 200 \mu m$. Each station is designed to give a muon vector in space, with a ϕ precision better than 100 μm in position and approximately 1 mrad in direction. Each DT chamber has 1 or 2 RPCs coupled to it before installation, depending on the station. In stations MB1 and MB2, each package consists of 1 DT chamber sandwiched between 2 RPCs. In stations MB3 and MB4, each package comprises 1 DT chamber and 1 RPC, which is placed on the innermost side of the station. A high- p_T muon crosses up to 6 RPCs and 4 DT chambers, producing up to 44 measured points in the DT system from which a muon-track candidate can be built.

The Muon Endcap (ME) system comprises 468 CSCs in the 2 endcaps. Each CSC is trapezoidal in shape and consists of 6 gas gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. All CSCs except those in the third ring of the first endcap disk (ME1/3) are overlapped in phi to avoid gaps in the muon acceptance. There are 36 chambers in each ring of a muon station, except for the innermost ring of the second through fourth disks (ME2/1, ME3/1, and ME4/1) where there are 18 chambers. The gas ionization and subsequent electron avalanche caused by a charged particle traversing each plane of a chamber produces a charge on the anode wire and an image charge on a group of cathode strips. The signal on the wires is fast and is used in the Level-1 Trigger. However, it leads to a coarser position resolution. A precise position measurement is made by determining the center-of-gravity of the charge distribution induced on the cathode strips. Each CSC measures up to 6 space coordinates (r, ϕ, z). The spatial resolution provided by each chamber from the strips is typically about 200 μm (100 μm for ME1/1). The angular resolution in ϕ is of order 10 mrad.

In the forward region, the system comprises 4 stations covering the pseudorapidity region up to $|\eta| < 2.1$. However, a shortfall of funds has led to the staging of the chambers sitting beyond $|\eta| > 1.6$. RPCs in the first endcap station are also used to help resolve ambiguities in the CSCs. There are 36 chambers mounted in each of 2 rings in each of the endcap stations.

3.3.5.1 Muon Alignment

For optimal performance of the CMS muon spectrometer over the entire momentum range up to the TeV range, the different muon chambers have to be aligned with respect to each other and to the central tracking system to within a few microns in $r - \phi$. The required precision for the endcap is between 75 and 200 μm , while for the barrel the precision varies from 150 μm for the inner station to 350 μm for the outer station.

There are several potential sources of misalignment in the muon spectrometer, from chamber production to final detector operating conditions, including:

- Chamber construction tolerances: These are unavoidable geometrical tolerances in the production of the chamber components. The relative positioning of the different internal components of a chamber was measured during construction to be within the required tolerances.
- Detector assembly, closing tolerances: Gravitational distortions of the return yoke lead to static deformations of the steel support. This effect, together with the installation tolerances, results in displacements of the chambers in the different barrel wheels and endcap disks of up to several millimeters with respect to their nominal detector positions.
- Solenoid effects: Magnetic forces generated by the 3.8 T solenoid field lead to displacements and deformations of the return yoke which is at the same time the support structure of the muon chambers. This results in further displacements of the chambers with respect to their nominal positions.
- Time-dependent effects: During operation, thermal instabilities and other time-dependent factors can cause dynamic misalignments at the sub-millimeter level.

The strategy for the alignment of the CMS muon spectrometer is to combine different sources of information, from the production phase of the muon chambers to the final monitoring during operation. Using as input the quality control data recorded during the construction of the chambers, survey and photogrammetry measurements done at the different stages of chamber construction and detector assembly, optical data provided by the optical muon alignment system and information provided by the tracks (cosmic rays, beam halo, or collision tracks) crossing the detector.

The muon system of CMS comprises an optical alignment system which allows a fast and independent measurement of the misalignments. A network of laser lines, light detectors, distance meters and tiltmeters are extended over CMS in a redundant scheme able to determine movements of the different structures as can be seen in figure 3.7. The muon optical alignment system is divided into three subsystems, the Endcap and Barrel hardware alignment systems, providing internal alignment in the endcaps and barrel respectively, and the Link system, which connects endcap, barrel and the central silicon-strip tracker. The technologies for each subsystem vary depending on the specific needs and configurations. The distribution of the detectors is also characteristic of each subsystem. Nevertheless the three subsystems share the same reconstruction software named COCOA, integrated in the standard software of CMS.

To complement the optical system, tracks from physics-run collisions, cosmic-ray data and beam halo are used to measure the position and orientation of the muon chambers. This technique requires the accumulation of a high number of tracks which are

3. THE CMS EXPERIMENT AT LHC

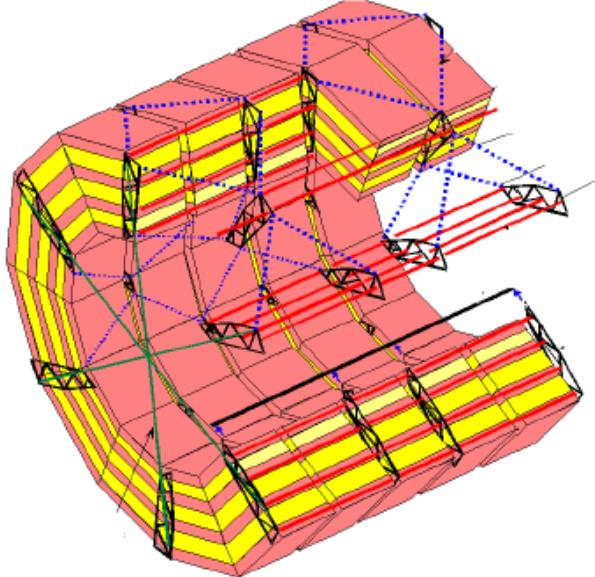


Figure 3.7: - Laser net intercalated between the barrel wheels.

processed offline to convert the redundancy in the track determination into information about the alignment. A CMS alignment framework has been developed specifically to the needs of tracking device. There are three track-based alignment algorithms for the muon system: the HIP and the Millepede approach algorithms, adapted to perform the alignment of the muon chambers with respect to the tracker, and the CSC Overlap Alignment Algorithm, more specific, focused in the alignment of the endcap rings using the existing overlap between chambers. More information in this subject can be consulted for example in [43] and [44].

3.3.5.2 Alignment and Calibration scenarios

Misalignment will affect the measurements and associated uncertainties of hits, tracks and all reconstructed objects depending on them. In 2007, a Monte Carlo study of some misalignment and miscalibration scenarios was performed to check the affect on the different physics analysis, as it is shown in [45]. Three scenarios were defined:

- **Start-up.** Also known as $0 pb^{-1}$ scenario. Tracker alignment precision resembling first collisions in CMS. Only information from survey measurements, Laser Alignment System and cosmic muon tracks used.
- **$10 pb^{-1}$.** Tracker can aligned by using cosmic muon data and a sample of collision tracks, mainly being isolated hadrons in minimum bias events and muons from the decays of low mass resonances like J/ψ and Υ . For the pixel detector, the alignment of its larger structures can be improved by a factor of 5, but no improvement is expected on the module-level alignment. For the strip tracker, the

sub-detector positions can be aligned with an accuracy of $100 \mu m$. Muon alignment not expected to improve with respect to the predefined geometry based on photogrammetry and pre-run cosmics. Numbers and correlations are taken from Magnet Test Cosmic Challenge (MTCC) [46] experience. Wheels and disks have $2 mm$ uncertainties in x and y coordinates, $3 mm$ in z in the barrel and $5 mm$ in z in the endcaps. Additionally, all chambers move with respect to the wheel/disk by $500 \mu m$.

- **$100 pb^{-1}$** . High P_T muons from Z and W decays available in significant quantity, in addition to the previously mentioned data. The misalignment of the pixel tracker is expected to be $O(20 \mu m)$, and that of the strip tracker $O(30\text{-}50) \mu m$. Additionally the whole muon system should be positioned with respect to the inner tracker with a precision of $1 mm$ and $250 \mu rad$. For the muon stations and the interleaving chambers on their turn an alignment precision of respectively $1 mm$ and $200 \mu m$ is foreseen.

Apart from the alignment of the tracking device, also the calibration constants employed in the various CMS sub-detector are determined only up to a certain level of accuracy. For this reason, in addition to the described misalignment shifts, several of the calibration constants were varied in their uncertainties. Examples are the intercalibration precision of the ECAL crystals (1.4 - 4%, depending on their location in η) and the gains for the HCAL (which are shifted by 10% in the $10 pb^{-1}$ scenario and smeared by 5% in the $100 pb^{-1}$ scenario).

In this study, the effects from the different misalignment and miscalibration scenarios are used for systematic studies. Figure 3.8 displays the fractional dimuon mass resolution as a function of mass for five resonances: J/ψ , Z , and a 1, 2, and 3 TeV Sequential Standard Model Z' .

3.3.6 Trigger and data acquisition

About one billion proton-proton interactions will take place every second inside the detector when performing at its peak. It is not possible to record all the data that is produced, and, as most of this data is unlikely to reveal new phenomena, it is not necessary either. It is a good idea then to select the potentially interesting events, and reduce the rate to just a few hundred events per second, which can be read out and stored for subsequent analysis.

But even recording just the potentially interesting events, many challenges arise. For example, with such a high collision rate, new particles are being generated before the ones from the last event have even left the detector so the data is stored in pipelines that can retain and process information from many interactions at the same time. To not confuse particles from two different events, the detectors must have very good time resolution and the signals from the millions of electronic channels must be synchronized so that they can all be identified as being from the same event. At design luminosity,

3. THE CMS EXPERIMENT AT LHC

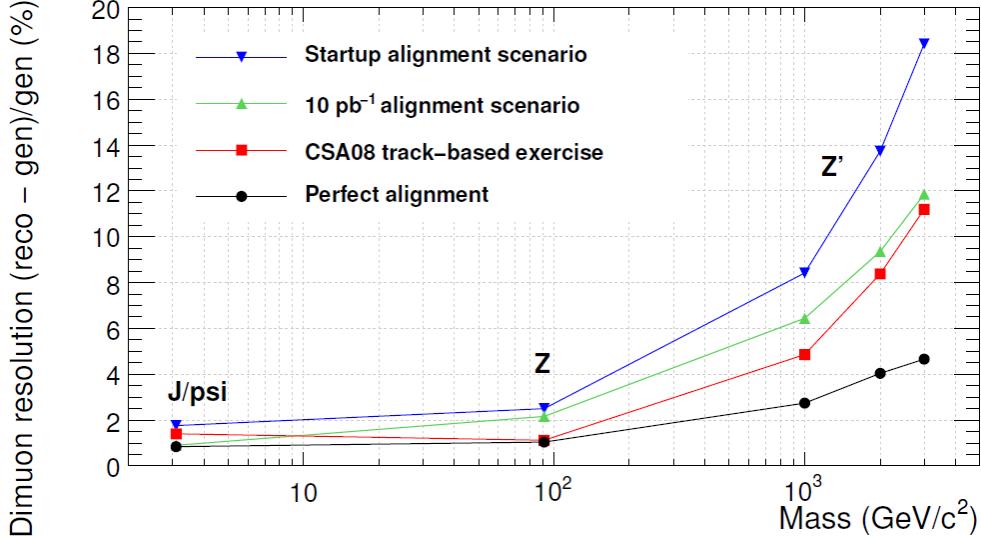


Figure 3.8: - Dimuon mass resolution for four alignment scenarios.

the LHC bunch crossing rate will be of 40 MHz , leading to $\approx 10^9$ interactions per second. Only 10^2 crossings at second can be stored; hence, the trigger system has to achieve a rejection factor of nearly 10^6 . The CMS trigger and data acquisition system [47; 48] consists of 4 parts: the detector electronics, the Level-1 trigger processor: calorimeter, muon, and global, the readout network, and an online event filter system that executes the software for the High-Level Triggers, HLT.

The minimum transit time for signals from the front-end electronics get to the services cavern housing the Level-1 trigger logic, reach a decision to keep or discard the event, and return back is around $3.2\text{ }\mu\text{s}$. During this time, the detector data must be held in buffers while trigger data is collected from the front-end electronics and the decision is made. Of the total latency, the time allocated to Level-1 trigger calculations is less than $1\text{ }\mu\text{s}$. The Level-1 triggers involve the calorimetry and muon systems, and its decision is based on the presence of objects such as photons, electrons, muons, and jets with a energy above some thresholds. It also employs global sums of E_T and missing E_T . Reduced-granularity and reduced-resolution data are used to form trigger objects. The overall scheme of the Level-1 trigger can be seen in figure 3.9.

Upon receipt of a Level-1 trigger, after the fixed time interval of $3.2\text{ }\mu\text{s}$, the data from the pipelines is transferred to front-end readout buffers. After further signal processing, zero suppression and/or data-compression, the data are placed in dual-port memories for access by the DAQ system. Each event, with a size of about 1.5 MB in the case of pp interactions, is contained in several hundred front-end readout buffers. Through the event building switch, data from a given event is transferred to a processor. Each processor runs the same high-level trigger (HLT) software code to reduce

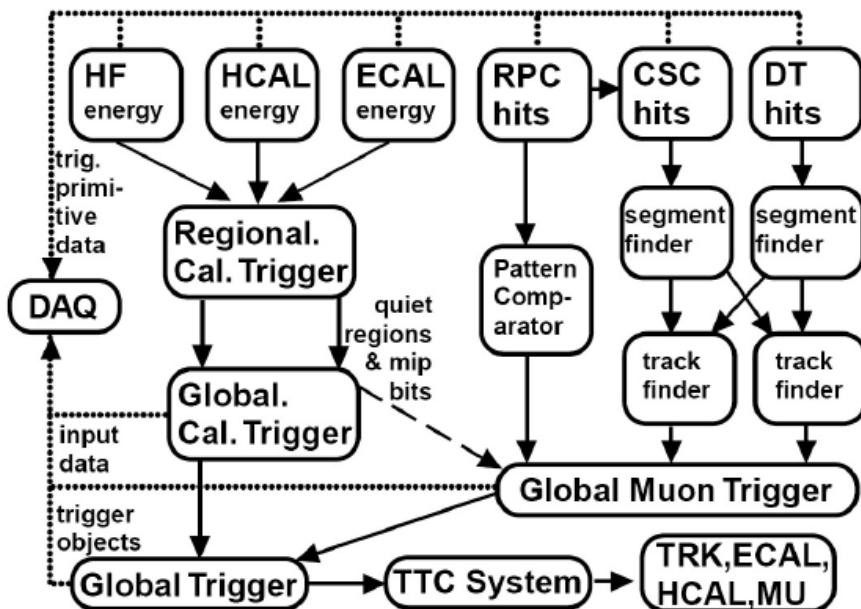


Figure 3.9: - Scheme of the Level-1 Trigger.

the Level-1 output rate in a factor of 10^3 . Various strategies guide the development of the HLT code. Rather than reconstruct all possible objects in an event, whenever possible only those objects and regions of the detector that are actually needed are reconstructed. Events are to be discarded as soon as possible. This leads to the idea of partial reconstruction and to the notion of many virtual trigger levels, e.g., calorimeter and muon information are used, followed by use of the tracker pixel data and finally the use of the full event information.

3.4 CMS Software and Computing

The CMS software and computing systems cover a broad range of activities including design, evaluation, construction, and calibration of the detector; the storage, access, reconstruction and analysis of data; and the support of a distributed computing infrastructure for physicists engaged in these tasks. The storage, networking and processing power needed to analyze these data is well in excess of todays facilities and exceed any reasonably projected capabilities of CERNs central computing systems.

After the trigger system, CMS still produces a huge amount of data that must be analyzed, more than five petabytes per year when running at peak performance. To meet this challenge, the LHC employs the Worldwide LHC Computing Grid (WLCG), a distributed computing and data storage infrastructure. Giving access to data to thousands of scientists all over the world in a highly distributed way. It starts in the Tier 0 center at CERN reconstructs first the full collision events and data is started

3. THE CMS EXPERIMENT AT LHC

to be analyzed; but there is still a long way to go. Once CERN has made a primary backup of the data it is then sent to large Tier 1 computer centers, located in seven countries around the world: France, Germany, Italy, Spain, Taiwan, the UK and the US. There events are reconstructed again, using information from the experiment to improve calculations using refined calibration constants. The most complex events are then sent to a number of Tier 2 facilities, around 40, for further specific analysis tasks. In this way information branches out from each tier, and the analysis of the final data can be performed locally all around the world.

The CMS computing model [49] is therefore, highly distributed, with a primary Tier-0 center at CERN being supplemented by Tier 1 and Tier 2 computing centers at national laboratories and universities worldwide. Much software was and is still being developed and verified for simulation and physics analysis, as well as common libraries, tools and frameworks.

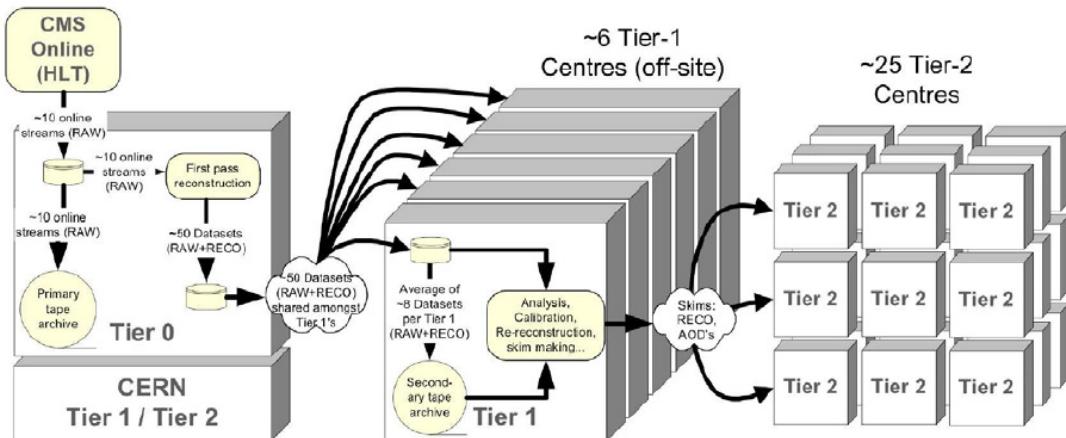


Figure 3.10: - Event data flow in the CMS Computing System.

In parallel to that, and before the data enters in the Tier system, calibration and express stream analysis are performed in the CERN Analysis Facility, CAF. The CAF is dedicated to latency critical activities like Calibration and Alignment, Detector/Trigger Commissioning or High Priority Physics Analysis. CAF access is restricted to users dedicated to these latency critical and high priority work-flows.

With respect to the software, the CMS Software, CMSSW, consists of more than 2 Million lines of code not including external third party packages and has a very active developer community composed by a growing group of geographically dispersed collaborators. It performs a variety of event processing, selection and analysis tasks, and is used in both offline and online contexts. It was designed to be sufficiently modular to facilitate the development and maintenance. The chosen architecture consists

of a common framework which is adaptable for each type of computing environment, physics modules which plug into the framework via a well-defined interface, and a service and utility toolkit which decouples the physics modules from details of event I/O, user interface, and other environmental constraints.

The central concept of the CMS data model is the Event. The Event provides access to the recorded data from a single triggered bunch crossing, and to new data derived from it. The Event also contains information describing the origin of the raw data, and the provenance of all derived data products. Events are physically stored as persistent ROOT files. The Event is used by a variety of physics modules, which may read data from it, or add new data, with provenance information automatically included. Each module performs a well-defined function relating to the selection, reconstruction or analysis of the Event. Several module types exist, each with a specialized interface. These include: event data producers, which add new data products into the event; filters used in online triggering and selection; analyzers, producing summary information from an event collection; and input and output modules for both disk storage and DAQ. As modules execute independently from one another, and communicate only through the Event; they are developed and verified independently. A complete CMS application is constructed by specifying to the Framework one or more ordered sequences of modules through which each Event must flow, along with the configuration for each. The Framework configures the modules, schedules their execution, and provides access to global services and utilities.

To maintain the required level of data reduction, keeping the flexibility, CMS uses several event formats with differing levels of detail and precision. Other specialized event formats are used for heavy-ion data. The process of data reduction and analysis takes place in steps, typically carried out at different computer centers. The main types are, RAW, RECO, and AOD. RAW events contain the full recorded information from the detector, plus a record of the trigger decision and other metadata. The RAW data is permanently archived in safe storage, and is designed to occupy around 1.5 MB/event (2 MB/event for simulated data, due to additional Monte Carlo truth information). The Reconstructed format, RECO data is produced by applying several levels of pattern recognition and compression algorithms to the RAW data. Reconstruction is the most CPU-intensive activity in the CMS data processing chain. The resulting RECO events contain high-level physics objects, plus a full record of the reconstructed hits and clusters used to produce them. Sufficient information is retained to allow subsequent application of new calibrations or algorithms without recourse to RAW data. RECO events are foreseen to occupy around 0.5 MB/event . Finally, AOD, Analysis Object Data is the compact analysis format, designed to allow a wide range of physics analyses whilst occupying sufficiently small storage so that very large event samples may be held at many centers. AOD data is produced by filtering of RECO data, and requires around 100 kB/event , small enough to allow a complete copy of the experimental data in AOD format to be held at computing centers outside CERN.

3. THE CMS EXPERIMENT AT LHC

To facilitate development, CMS consolidates its code base regularly into releases. Different releases are grouped into release cycles dedicated to a specific purpose (that is, data taking and reconstruction in LHC collision period, integration of a new ROOT version, etc.) to aggregate specific feature sets of the software stack. The rapid development cycle of CMS and the resulting high number of releases requires a thorough quality assurance process. To guarantee stable and performant releases while supporting a high turn around and diverse development, CMS implemented an advanced central release validation process.

3.4.1 CMS Remote Analysis Builder

CMS Remote Analysis Builder, known as CRAB, is a Python program intended to simplify the process of creation and submission of CMS analysis jobs into a grid environment. It is then used to run jobs on the grid.

In order to store and manage a huge quantity of data, while assuring data access to physicists of CMS collaboration and computing power for analysis and simulation, CMS uses a distributed architecture based on Grid infrastructure that guarantees resources and data availability to allowed users.

The users' point of view involves analyze the distributed data as they were in the local farm but the analysis in a distributed environment is a complex computing task, hence, the CMS collaboration had developed CRAB to simplify the work of users to create, submit and manage their analysis job in Grid environments, handling handles data discovery, resources availability, job creation and submission, status monitoring and output retrieval. CRAB is an user-friendly tool that helps users to run their analysis code on data available at remote sites, hiding Grid and CMS infrastructure details.

3.4.2 ROOT

ROOT [50] is an analysis package written in an object-oriented structure in C++. It uses built-in functions and user-compiled code to produce graphics and histograms, as well as trees with data objects. ROOT files can be easily navigated using a GUI Object Browser.

The CMS event data model (EDM) is based around the idea that people will use ROOT both in a "framework application", and by interactive manipulation of tree objects and histograms. Interactive ROOT sessions are conducted using an extension to C++ called CINT.

The backbone of the ROOT architecture is a layered class hierarchy with, currently, around 1200 classes grouped in about 60 frameworks (libraries) divided in 19 main categories (modules). This hierarchy is organized in a mostly single-rooted class library,

that is, most of the classes inherit from a common base class `TObject`. It enables the implementation of some essential infrastructure inherited by all descendants of `TObject`. However, there can also be classes not inheriting from `TObject` when appropriate.

3.5 Workflow of the Analysis

To perform the analysis described further in this thesis, the whole cms computing system was used. Monte Carlo simulated data and real data were equally treated. While the first is generated in the Tier 0, real data was recorded directly in the experiment and then transferred to the Tier 0. The format of the original samples (Monte Carlo and data) is RAW, so these samples are large and contain all the information possible. Afterwards the RAW data is transferred to the Tier-1 sites, in where the re-reconstruction, skimming and selection is performed. The resulting data, in RECO format, is then transferred to the Tier-2 in which it is re-processed to AOD format.

The analysis is done in AOD data, the lightest format. A dataset is chosen depending on the needs of the analysis, that is, for a Monte Carlo study, the datasets corresponding to the signal and the main backgrounds are chosen; for data-basis analysis, a primary dataset based in a certain HLT trigger path, or a secondary skim provided by the corresponding physics group, is selected.

The final physics analysis can be performed in many different ways, in the case of this thesis, it was done mainly in the so-called Analyzer mode. CMSSW provides different modules to perform, for example, the pre-selection of interesting events, sophisticated muon identification and isolation, Jet energy and missing E_T corrections, and so on. The Analyzer mode allows these modules to be added to the user code, in which a concrete kinematical selection oriented to, in this particular case, extract signal events from background in the most efficient way.

Once the Analyzer is tested, jobs are created and submitted using CRAB, accessing then the data stored in the Tier-2 sites worldwide. In this particular case, the Tier-2 site most accessed is the one built and maintained by IFCA (Instituto de Física de Cantabria), in which Higgs-related datasets are stored. As many as 25 different samples, 15 for background and 10 for signal can be studied in a single round of the Higgs to WW^* to $2\mu 2\nu$ analysis, some of them very large. Approximately each CRAB job analyzes 10.000 events, so, for samples of the order of 10^6 events, the order of 10^2 CRAB jobs are needed.

The analysis can be done locally in the Tier-3 machines up to a certain level. It was therefore, partially performed in the Tier 3 in Oviedo, where skimmings of the interesting data streams are stored. In that case, the analysis is done in an interactive way, using the `cmsRun` feature of CMSSW.

3. THE CMS EXPERIMENT AT LHC

The output of the analyzer is usually a rootfile with histograms, that is opened for further analysis with ROOT. Also some other studies outside the regular analysis chain, are done using ROOT locally.

4

Search for a SM Higgs in the WW^* decay channel

4.1 Higgs Phenomenology at the LHC

Higgs physics at the LHC will be different than in previous colliders due mainly to the larger centre of mass energy considered. The design collision energy for the LHC is 14 TeV, while the second most energetic accelerator built, the Tevatron, has an energy of about 2 TeV. Particle production in the LHC is in a mass range from 10^2 GeV and above and in a Feynman x region where the gluon density is much larger than the quark density. Therefore for Higgs and top physics it is useful to think of the LHC as a gluon collider. At the design centre of mass energy, the LHC plans to collect several hundred fb^{-1} of data. If the Higgs exists, the LHC will become a Higgs factory, producing hundreds of thousands or tens of thousands of Higgs bosons, depending on its mass.

The ratio between gluon fusion Higgs production and Higgsstrahlung at the LHC is

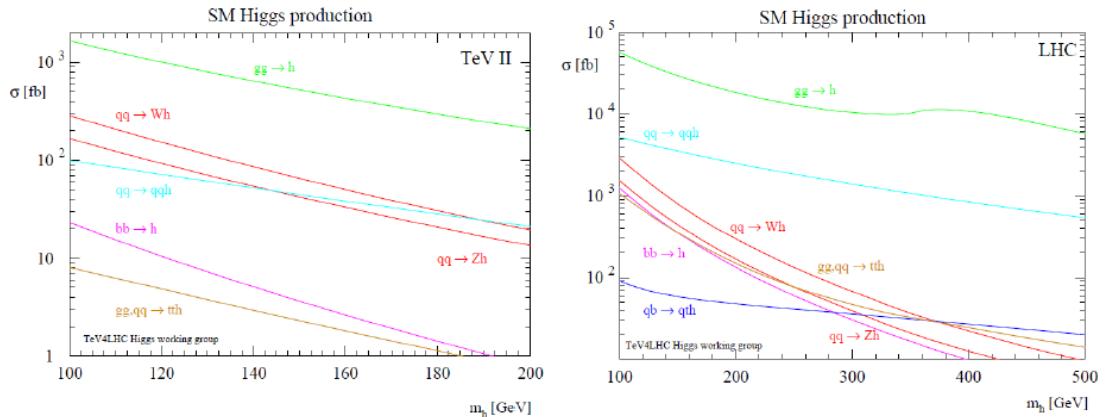


Figure 4.1: - Standard Model Higgs production cross-sections at the Tevatron (left) and at the LHC (right).

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

then larger than at the Tevatron. Figure 4.1 displays the various SM Higgs cross-sections for the Tevatron and the LHC at design energy, with a much larger range of m_H for the second. In the Tevatron the large m_H cross-sections are small and that does not happen in the LHC. Figure 4.2 shows the production cross-sections for different processes, Higgs included, against the centre of mass energy.

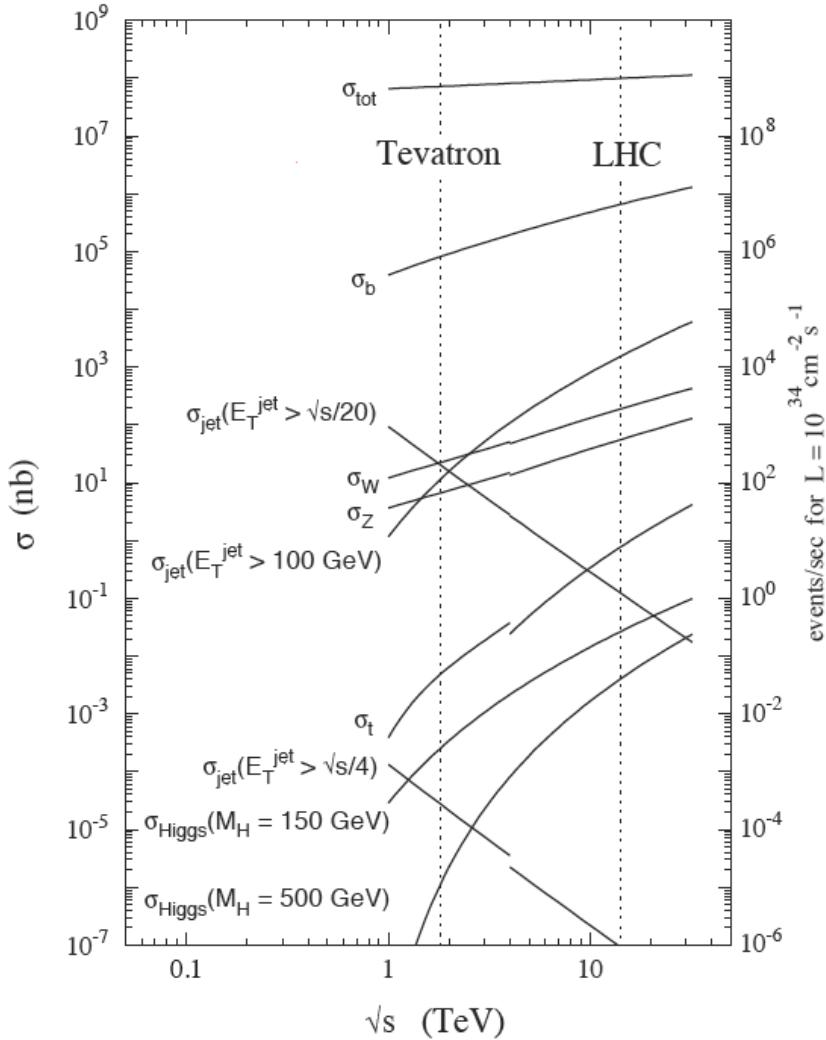


Figure 4.2: - Production cross-sections for different events at the Tevatron and at the LHC as a function of centre of mass energy.

When designing an analysis strategy for the search of a SM Higgs boson, the first thing to consider is the production mode. The dominating Higgs production mechanism at the LHC is the Gluon Fusion process [51], and it dominates for all possible Higgs masses. The process $gg \rightarrow h + X$ is an essentially strong-interaction process, that has

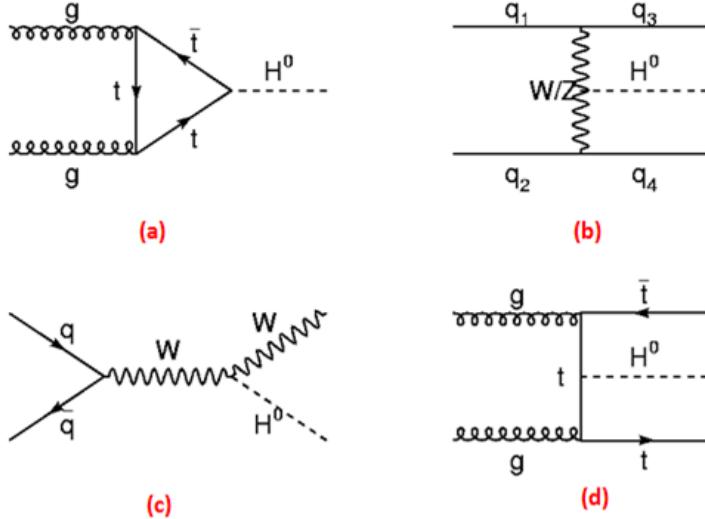


Figure 4.3: - Higgs production modes at hadron colliders: Gluon fusion (a), Vector Boson fusion (b), Associate Production with a W boson (c) and Associate production with a top pair (d).

attracted a large amount of theoretical work in recent years. Limits on the Higgs mass will rely upon QCD calculations of the cross-sections, and conversely, if a Higgs boson is discovered, discrepancies of its measured cross-section from QCD calculations may signal deviations of the Yukawa couplings from the Standard Model (SM) predictions [52].

The second most important production mode is the Vector Boson Fusion $qq \rightarrow qqh + X$, the inelastic scattering of two quarks (antiquarks), mediated by t -channel W or Z exchange, with the Higgs boson radiated off the weak bosons. In this case, the Higgs boson is produced via the fusion of the weak bosons W and Z , and is accompanied by two forward jets that carry large transverse momentum P_T in the final state.

The W,Z Associated Production $qq \rightarrow Wh + X$, $qq \rightarrow Zh + X$, also called Higgsstrahlung, is especially important for low masses, where also Bottom Fusion $bb \rightarrow h + X$ can be studied.

The $t\bar{t}$ Associated Production $qq, gg \rightarrow tth + X$ at the LHC is an important search channel for Higgs masses below $\approx 125 \text{ GeV}/c^2$. The single-top Associated Production $qb \rightarrow bth$ will also be accessible and studied in the LHC.

Different strategies for the Higgs identification can be developed depending on its mass, as not just the production but also the decaying modes change dramatically across the possible range of the Higgs mass. Figure 4.4 shows the branching ratios of the different Higgs decay channels depending on the mass.

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

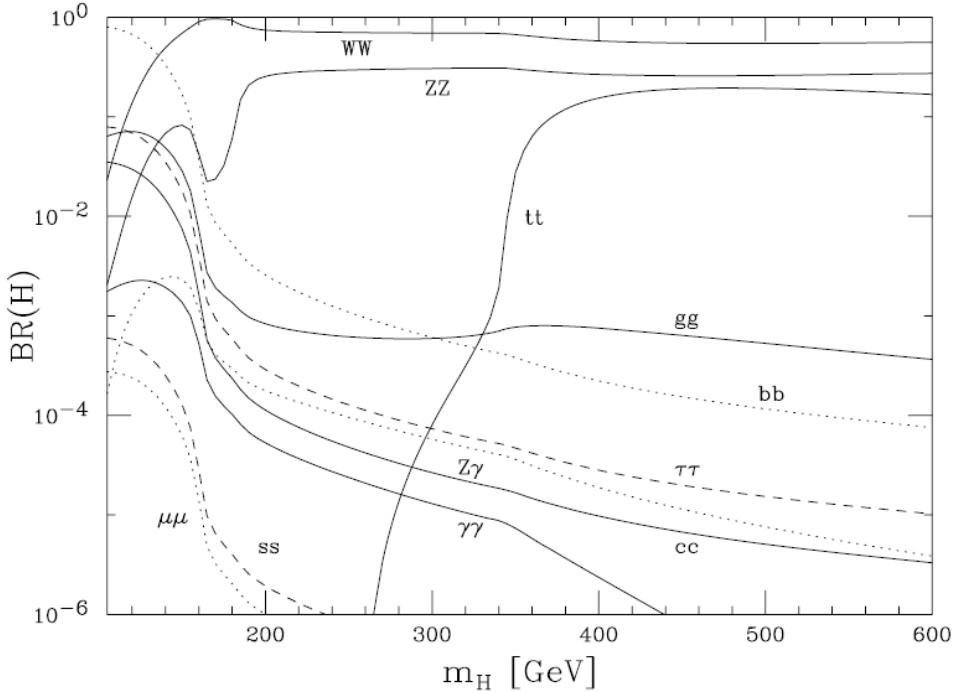


Figure 4.4: - Higgs Branching Ratios as a function of m_H .

The Higgs boson decays to the heaviest particles kinematically accessible for each mass. For masses up to about $\approx 140 \text{ GeV}/c^2$ the $b\bar{b}$ channel is the dominant decay, and also other channels contribute, like other fermion-antifermion pairs and even $\tau\tau$ decays, with $\sim 10\%$ of the BR. For higher masses, the vector boson channels open and the WW^* , ZZ^* (where one or both W/Z can be off shell) dominate. Loop mediated decays to two photons, through W and top quark loops, are much smaller but play a role. The two-photon channel $H \rightarrow \gamma\gamma$ is interesting in the mass range $120 - 130 \text{ GeV}/c^2$, as a narrow peak over the background could be reconstructed.

4.2 Higgs searches in CMS

Given the Higgs production and decay modes accessible at the LHC, many different analyses can be foreseen. The search strategy for several channels has been designed in CMS [53], see for example [54], oriented to the discovery or exclusion for a SM Higgs boson in a mass range that starts where LEP stopped, around $114 \text{ GeV}/c^2$, and with special stress in masses that are not accessible by the Tevatron. The main SM channels to be studied, amongst others, are:

- $H \rightarrow \gamma\gamma$

- $H \rightarrow WW^* \rightarrow 2l2\nu$
- $H \rightarrow ZZ^* \rightarrow 4l/2l2\nu/2l2b$
- $H2\tau$ ($H \rightarrow \tau\tau$; $H \rightarrow ZZ^* \rightarrow 2l2\tau/4\tau\dots$)
- VBF (VBF $H \rightarrow WW^* \rightarrow l\nu jj$, VBF $H \rightarrow ZZ^* \rightarrow 2ljj$, VBF $H \rightarrow invisible$)
- $t\bar{t}H$, $H \rightarrow b\bar{b}$

$H \rightarrow \gamma\gamma$ is a very important channel for low masses, $120 < m_H < 140$ GeV/c^2 . The small branching ratio is compensated with a clear signature with two photons in the final state. It can be produced in an inclusive way or in association with hadronic high P_T jets. Once the signal is identified the Higgs mass peak can be reconstructed. With a good control of the backgrounds, Higgs to $\gamma\gamma$ analysis can lead to a low mass Higgs discovery with significant luminosity. In [54] the estimated luminosity needed for discovery using this channel was ~ 30 fb^{-1} .

For Higgs masses between 140 and 180 GeV/c^2 the channel that dominates is the $H \rightarrow WW^*$ that covers the most interesting SM Higgs mass range for a possible discovery [3]. The search strategy for this channel is the topic of this thesis, and for that, this decay will be largely discussed in the following starting in the next section 4.3. It will be mainly produced via gluon-fusion, but also the VBF Higgs to WW^* , that adds two forward jets to the final state, will be studied. In this case the cross-section is smaller than the gg -induced process but the signature is cleaner.

The $H \rightarrow ZZ^* \rightarrow 4l$ decay channel [55] was known as the Golden Higgs decay, as it brings the experimentally cleanest signature for discovery, with a narrow 4-lepton invariant mass peak on top of a smooth background. It is a powerful analysis in a wide mass range with a very high Branching ratio except for Higgs masses around two times the mass of the W , when the $H \rightarrow WW^*$ dominates. However, by indirect estimation of m_H based on EWK precision fits, see 2.4.4, the preferred mass for a SM Higgs boson is around 150 GeV/c^2 , and in that region, not only the WW Branching Ratio is higher, but also one of the Z bosons of the $H \rightarrow ZZ^*$ decay is off-shell. The analysis of this channel relies on the lepton identification and background control, and is sensitive up very high masses ~ 500 GeV/c^2 with a discovery potential for high energies of 5 fb^{-1} of integrated luminosity [55].

As mentioned in the previous section 4.1, the $H \rightarrow \tau\tau$ [56; 57] is a significant channel for low masses despite its small cross-section, as it can be used for studying the properties of a possible Higgs boson, like the couplings to leptons. Different final states with leptons and jets in the final state (Vector-Boson fusion induced) are studied. The Higgs mass peak is reconstructed but it is not as straightforward as in the case of Higgs to ZZ .

Higgs to $b\bar{b}$ is the only Higgs decay that can be studied at the LHCb experiment [33],

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

and it is also considered in the Higgs program of CMS. It presents a large combinatorial background with more than 5 jets, and for that requires an excellent b -tagging performance.

Concerning the MSSM, the most significant channels at the LHC centre of mass energies are $\tau\tau$ and $\mu\mu$. Charged Higgs is interesting at low ($m_{H^\pm} < m_{top}$, $H^\pm \rightarrow \tau\nu$) and high ($m_{H^\pm} > m_{top}$, $H^\pm \rightarrow tb$) masses.

In the case of bbH/A , the reach can be extended down to $m_A \sim 150 \text{ GeV}/c^2$, $\tan\beta \sim 10$. SM Higgs searches can be reinterpreted in MSSM ($h \rightarrow \gamma\gamma$, $qqh/H \rightarrow q\bar{q}\tau\tau$) covering the low and intermediate $\tan\beta$ region. These searches suffer from systematics, and many parameters and scenarios have to be studied.

4.3 The WW^* channel

In the Higgs mass region around $m_H = 160 \text{ GeV}/c^2$, the branching ratio to WW^* becomes close to 1. This decay is the most promising channel for masses above $120 \text{ GeV}/c^2$, it can be studied even for masses where one of the W bosons is off-shell and stays as the dominant mode for higher masses, being especially relevant up to $200 \text{ GeV}/c^2$.

The leptonic WW^* signature was originally proposed and shown to be useful in the mass range of $150 - 180 \text{ GeV}/c^2$ in [58]. Since then, the strategy went through subsequent updates and improvements [59; 60], and is currently considered the main SM Higgs search channel from 130 to $200 \text{ GeV}/c^2$. The strategy presented in [58] and further improved in [59; 60] is followed and developed in this thesis, and also in [61; 62; 63] and in the public documents [64; 65].

The leptonic decay of the W bosons, in figure 4.5, brings interesting final states that

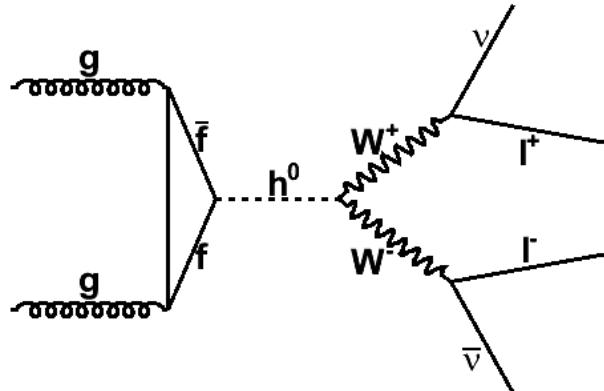


Figure 4.5: - Feynman diagram of the gluon-fusion induced Higgs to WW^* decaying to two leptons and two neutrinos.

contain two leptons and two neutrinos. Three signatures are currently under study for this decay in CMS, each one corresponding to a different final state: ee , $e\mu$ and $\mu\mu$. This thesis is centered on the study of the potential of the $gg \rightarrow H \rightarrow WW^* \rightarrow 2\mu 2\nu$ but the three final states are similar and complementary.

In this channel a clear Higgs mass peak reconstruction is not possible because of the presence of the undetected neutrinos in the final state. A counting experiment that requires an accurate estimation of the background contribution to the signal region is proposed.

4.3.1 Signature of Signal and Backgrounds

The $gg \rightarrow H \rightarrow WW^* \rightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$ decay channel has a clean signature. The final state presents two opposite charge high- P_T isolated muons and a significant amount of missing E_T due to the undetected neutrinos. No hard jet activity is expected.

The backgrounds considered will be all the possible processes that can yield similar signatures: real or fake multi-lepton final states plus missing E_T . The Feynman diagrams of some of the backgrounds are presented in figure 4.6.

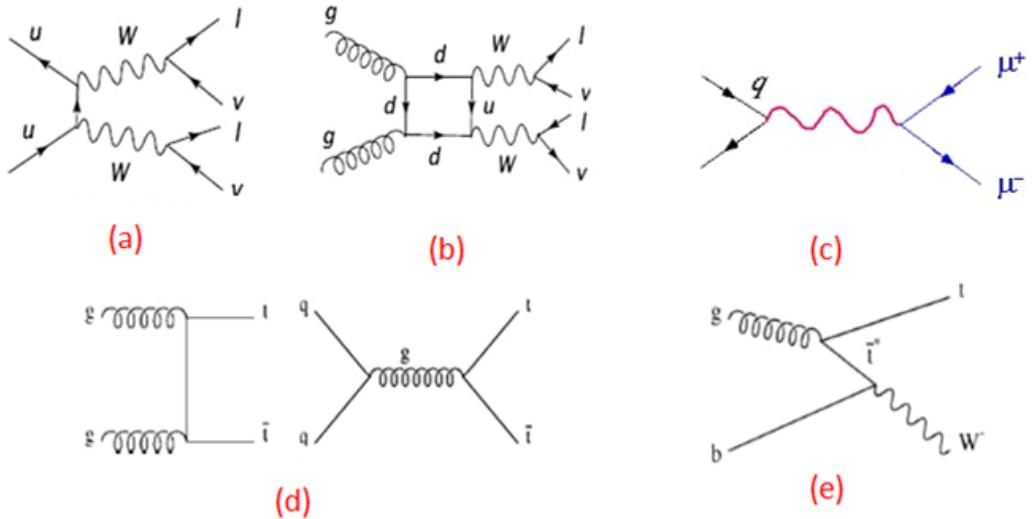


Figure 4.6: - Background processes: (a) $q\bar{q}$ induced WW production, (b) gg induced WW production, (c) Drell-Yan production, (d) $t\bar{t}$ production, (e) single-top tW production, others would be WZ , ZZ , $W + jets$ and QCD .

The continuum WW production is almost identical to the signal. The only way to disentangle WW from the signal is by means of the $\Delta\Phi_{\mu\mu}$ variable, the angle between the muons in the transverse plane to the beam. The difference in azimuthal angle in the production of the two leptons is the key of the Higgs to WW^* analysis. For signal,

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

as they come from the decay of a spin 0 particle, the angle between the muons in the transverse plane takes small values, as it is sketched in figure 4.7. That does not happen for WW background, were the muons tend to be emitted back-to-back.

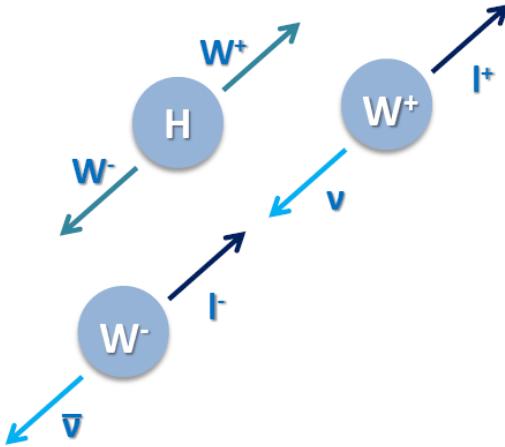


Figure 4.7: - Scheme of the leptonic decay of a Higgs to WW^* event.

The large $t\bar{t}$ production at the LHC is another important background characterized by the hadronic activity. This background is rejected by the application of a jet veto, as jets are expected in the final state of $t\bar{t}$ dileptonic events. Events with jets will be therefore rejected.

Another background with different characteristics is the one coming from $Z + jets$, when the Z boson decays into two muons $Z \rightarrow \mu\mu$. This background has no genuine missing E_T but can be dangerous when this quantity is not properly estimated. It has a large cross-section compared to the signal.

$W + jets$ can be a source of background when one isolated object, track/jet is misidentified as a second muon and two muons are reconstructed in the final state. It has an even larger cross-section than $Z + jets$, and requires a detailed control of the fake muons that can lead to di-muonic final states. The way to reject this background is then, to unambiguously identify prompt muons coming from W decays.

To visualize the kinematics of these events on Monte Carlo produced samples it is possible to use a dedicated tool available in CMS, the Physics Analysis Oriented Event Display Fireworks/cmsShow [66]. In figure 4.8 the displays for Higgs to WW^* to two muons, and WW to two muons are shown. Muons are presented in red, missing E_T as a yellow arrow and jets (if present) in purple. In these figures the different angular separation of the muons can be clearly seen. Muons are very close when coming from a Higgs decay and present a wider separation in the case of a typical WW event. In

figure 4.10, the display for a $t\bar{t}$ event, going to two muons and two jets is shown. In figure 4.11, the corresponding displays for $Z+jets$ and $W+jets$ are also presented.

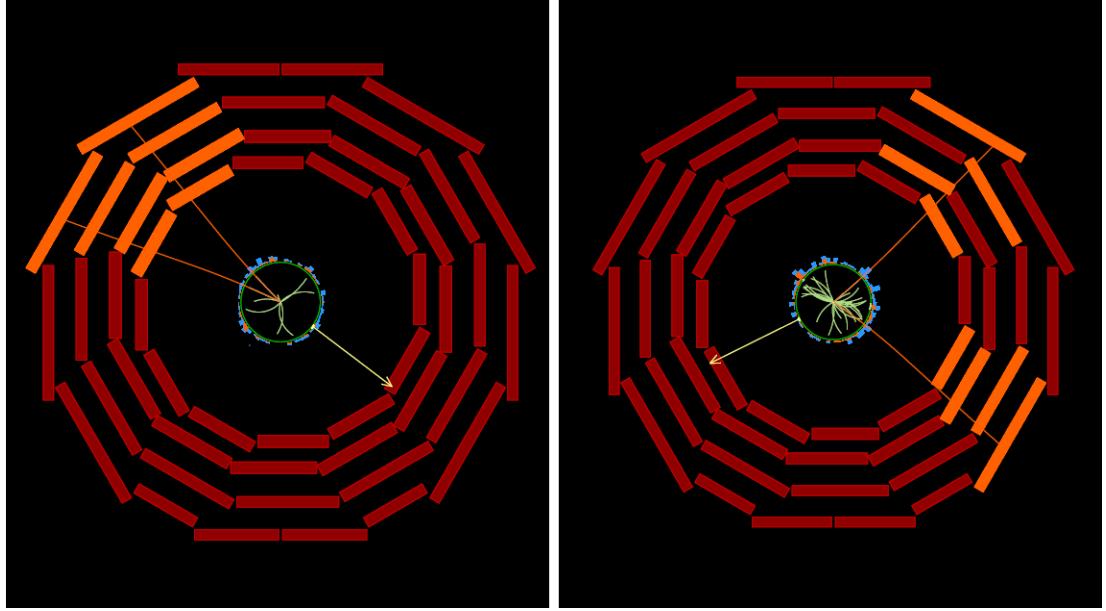


Figure 4.8: - Display of a Higgs to WW^* Monte Carlo event, with $m_H = 160 \text{ GeV}/c^2$, (left) and a WW event (right), $\sqrt{s} = 7 \text{ TeV}$.

More backgrounds that have to be considered in the analysis are other multiboson events as WZ and ZZ , single-top production tW or QCD . These present lower cross-sections and different kinematics, like having more than two leptons in the final state, for example $WZ \rightarrow \mu\nu\mu\mu$ or $ZZ \rightarrow 4\mu$ decays, or different hadronic content.

4.3.2 Description of the Search Process

Events with muons have to be selected in the analysis presented in this thesis, as the selected Higgs boson decay channel presents two muons in the final state. Therefore, a Muon primary dataset has to be used. The High Level Trigger path or paths have to be chosen carefully, as they will determine different kinematical conditions, such as the minimum P_T of the muons considered later.

The muons are selected if they carry a substantial P_T and are sufficiently isolated, that is, if they do not come inside a jet or accompanied by other tracks. More details on the muon selection will be given in the following sections 5.1.

The characteristics of the signal change depending on the Higgs mass. When the Higgs mass is below the $2 \times m_W$ threshold, one of the W bosons is off-shell, and that is reflected in the kinematics of the event. The effect is clearly visible in the P_T of

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

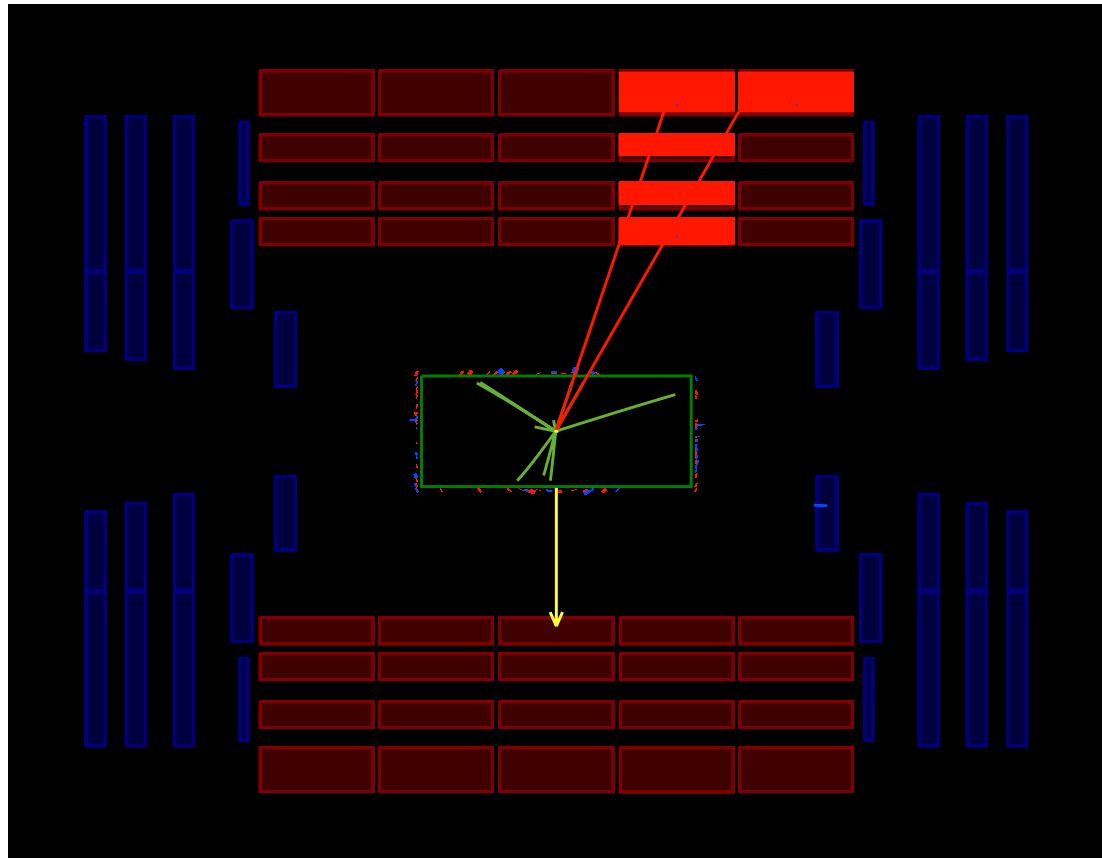


Figure 4.9: - r-Z view of a Higgs to WW^* Monte Carlo event, with $m_H = 160 \text{ GeV}/c^2$, $\sqrt{s} = 7 \text{ TeV}$.

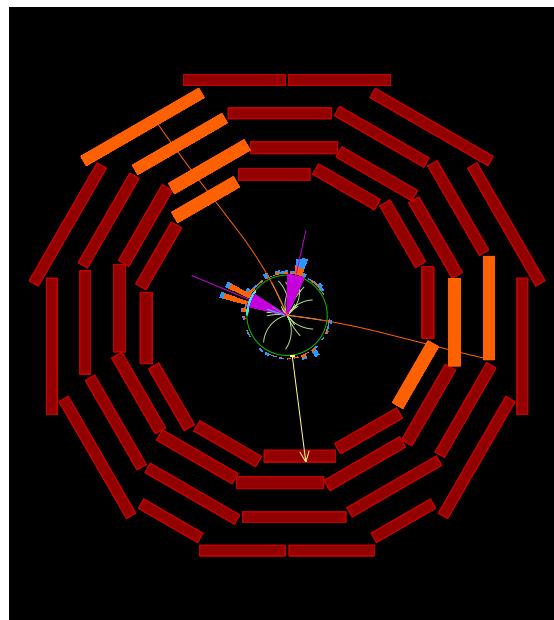


Figure 4.10: - Display of a $t\bar{t}$ Monte Carlo event, $\sqrt{s} = 7$ TeV.

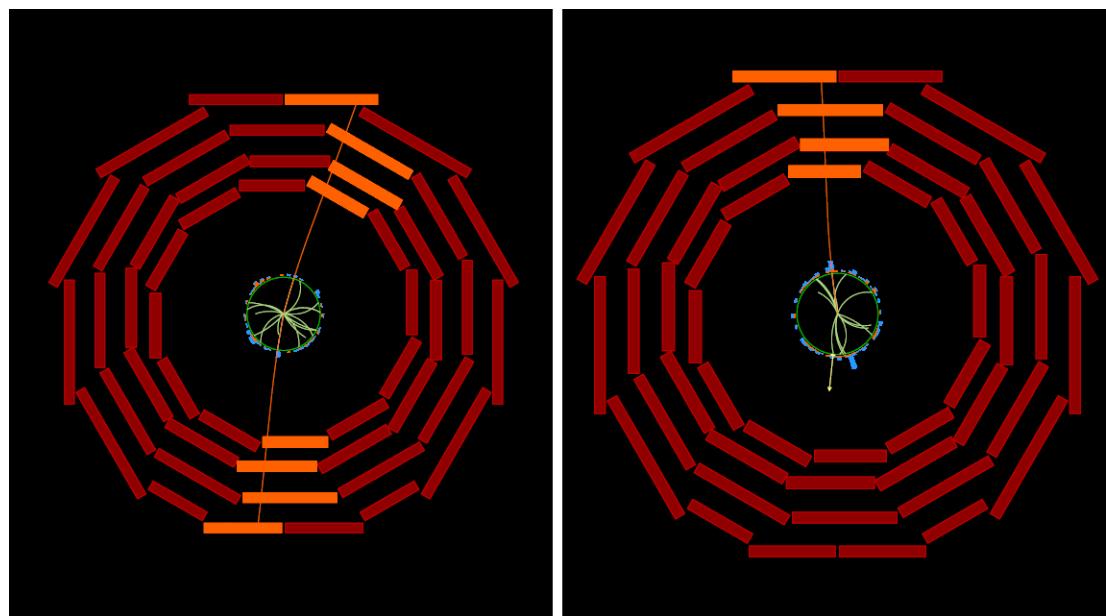


Figure 4.11: - Display of a $Z + jets$ Monte Carlo event(left) and a $W + jets$ event(right), $\sqrt{s} = 7$ TeV.

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

the second more energetic μ of the event, shown in figure 4.12. The P_T distribution of the first μ is also shown in figure 4.13. The analysis addresses the differences in the kinematics by applying specific sets of cuts for each Higgs mass hypothesis.

The analysis requires the best muon reconstruction, identification and isolation possible using all the information available in the muon chambers, tracker and calorimeters. The performance of the muon reconstruction is especially critical for the low Higgs mass range, around $130\text{ GeV}/c^2$, when low momentum muons have to be identified.

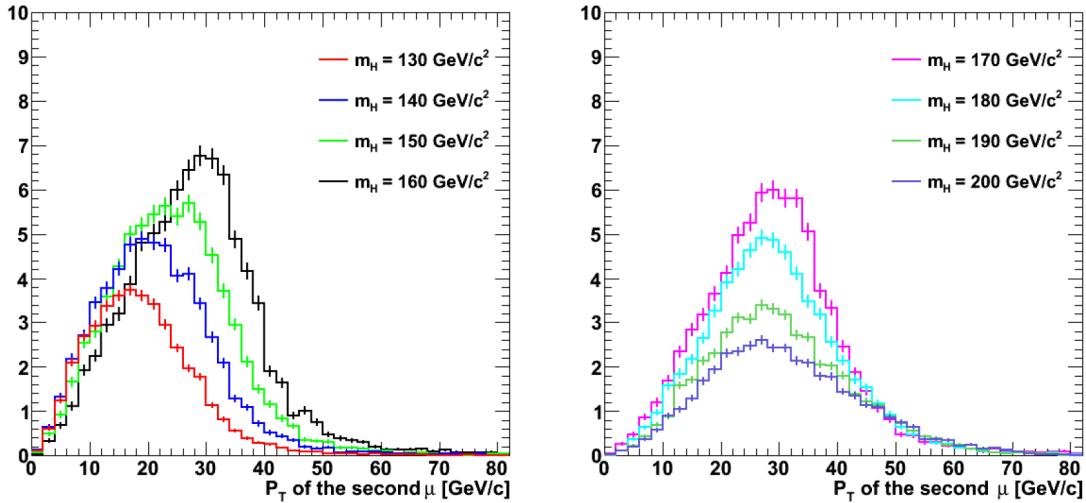


Figure 4.12: - P_T of the second μ for different Higgs masses, at 7 TeV of centre of mass energy assuming 1 fb^{-1} of luminosity as simulated by Pythia Monte Carlo.

The rejection of multi-jet background, mainly $t\bar{t}$, relies on a good jet identification. There are many different jet-reconstruction algorithms in CMS, and depending on the data-taking scenario (energy, luminosity, calibration) the algorithm choice and the E_T threshold vary. Essentially, a tight jet quality criteria is applied, strong enough to be algorithm-independent. The jet identification criteria can be loosened to adapt the analysis to different running conditions and to optimize the balance between signal selection and background rejection. Jet identification and reconstruction are mainly based on calorimetric information. Tracker information can also be used in jet identification to complement the calorimeters, and also to give an independent estimation of the jet content in the event. More details on the use of jets in the analysis will be given in section 5.3.

Signal events will present significant missing E_T . A cut on this variable helps to reduce background, especially $Z \rightarrow \mu\mu$. The missing E_T estimation depends mainly on the calorimeters, but tracker information can be also used to give better results. More details will be given in section 5.4.

4.4 LHC Center of mass Energy and Integrated Luminosity Scenarios

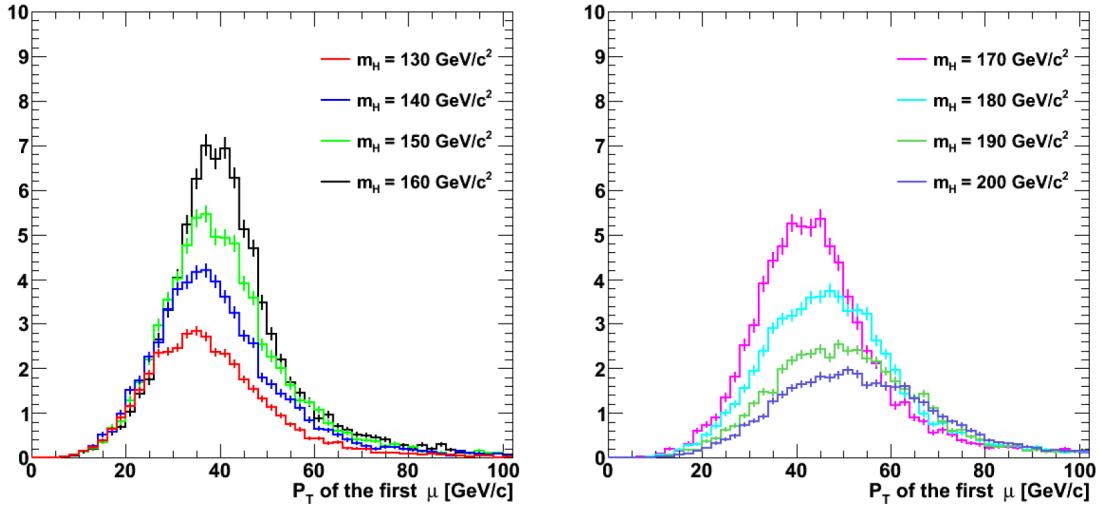


Figure 4.13: - P_T of the first μ for different Higgs masses, at 7 TeV of centre of mass energy assuming 1 fb^{-1} of luminosity as simulated by Pythia Monte Carlo.

The considered center-of-mass energy determines the cross-section and kinematics of signal and backgrounds. For example at low energies the gluon fusion events tend to be more central, a question that has to be contemplated in the analysis.

The analysis is performed on Monte Carlo generated samples. The samples used will be described in section 4.5.

4.4 LHC Center of mass Energy and Integrated Luminosity Scenarios

Before the start of the current run of the LHC at a centre of mass energy of 7 TeV, several benchmark energies and luminosities were planned. The detailed Monte Carlo-based study performed in this thesis has evolved and has been adapted to the needs of the CMS experiment related to the changing environment derived from the variation of the planned running conditions.

Initially, a high energy LHC start-up was expected, and the analysis was originally conceived for the LHC design centre of mass energy of 14 TeV assuming 1 fb^{-1} of integrated luminosity. A prompt Higgs discovery in the case of the original schedule of the first LHC run was intended.

After the incident happened in 2008 that delayed the start of the LHC, the plan was revisited. One of the benchmarks for early physics presented in the LHC Performance Workshop at Chamonix in 2009 [67] was an integrated luminosity of 200 pb^{-1}

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

at $\sqrt{s} = 10$ TeV. The analysis was optimized and performed for this scenario. Lower luminosities and energies were also considered.

Finally, following the definitive first run conditions announced in the LHC Performance Workshop at Chamonix in 2010 [4], the analysis was adapted to an integrated luminosity of 1 fb^{-1} at $\sqrt{s} = 7$ TeV, planned for the end of the present run, in 2011.

The implications of the different centre of mass energies will be largely discussed in the following sections and chapters. The signal to background ratio changes when lowering the energy as the cross-sections for signal and backgrounds are not equally affected. More luminosity is needed to achieve the same sensitivity.

4.5 Monte Carlo Datasets

The datasets for the presented study have been generated using various Monte Carlo Generators. The resulting events were propagated through a detailed simulation of the CMS detector response using different releases of the official CMS software.

Some details regarding the different Monte Carlo generators used in the different CMS productions as well as the cross-sections considered for signal and backgrounds will be presented.

4.5.1 Monte Carlo Generators

Event generators are indispensable tools for the modeling of complex physics processes, as the ones produced at the LHC. Generators are used to set detector requirements, formulate analysis strategies and calculate acceptance corrections.

There are many physics aspects that can determine the construction of an event generator, such as hard processes, initial- and final-state radiation, multiple interactions and beam remnants, hadronization and decays, and how all of them come together. Several generators are used in CMS to take advantage of their different features when generating events in different regions of the phase space.

In order to introduce the current main generators the first one must be PYTHIA [68]. PYTHIA is a program for the generation of high-energy physics events, that is, for the description of collisions at high energies between elementary particles such as e^+e^- , $p\bar{p}$ and, like in the case of the LHC, pp , in various combinations. It contains theory and models for a number of physics aspects, including hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation and decay.

MC@NLO [69] is a Fortran package that combines a Monte Carlo event generator

with Next-to-Leading-Order calculations of rates for QCD processes. It makes use of the Fortran HERWIG [70] event generator. The processes available in the latest version are those of Higgs boson, single vector boson, vector boson pair, heavy quark pair, single top (with and without associated W), lepton pair and associated Higgs+W/Z production.

Alpgen [71] is a generator for hard multiparton processes in hadronic collisions. It produces matrix element (ME) level events which are subsequently passed to a parton shower (PS) / hadronization code (like Pythia or HERWIG) for further event development. Using the MLM ME/PS matching procedure, one can combine the matrix element calculations with parton showers while avoiding double-counting. This gives the best Monte Carlo prediction of multi-jet final states at the moment. As other generators of interest to CMS, Alpgen itself is executed as a separate (pre-CMSSW) series of steps, to produce parton-level, unweighted Alpgen events.

TopReX [72] is a specialized event generator that provides the simulation of several important processes in pp collisions, not implemented in PYTHIA (yet). Some of these processes include t -quarks whose spin polarizations are taken into account in its subsequent decay. Several non-SM top quark decay channels are included, too. All calculated sub-processes can be accessed from PYTHIA as external processes. In addition, TopReX can be used as stand alone event generator, providing partonic final states before showering. In this mode the control of the event generation is taken by TopReX itself.

MadGraph/MadEvent [73] is a software that allows to generate amplitudes and events for any process (with up to 9 external particles) in any model. Implemented models are the Standard Model, Higgs effective couplings, MSSM, the general Two Higgs doublet model, and several minor models, and there exists an easy-to-use interface for implementing model extensions.

4.5.2 Monte Carlo 7 TeV Production

The Summer 2009 Full Simulation Production for Physics at 7 TeV centre of mass energy in CMS was the last used for the Standard Model Higgs Monte Carlo analysis performed in this thesis. It was produced with the CMSSW 3x [74] software release, and several generators were used.

The main background samples used are presented in table 4.1. Some of the signal samples used available in the same release are presented in table 4.2. Further 7 TeV re-processing was used as well and validated giving similar performance, as can be seen in figure 4.14 and in figure 4.15. Systematic validation of the samples were also made from 14 TeV to 10 TeV, and from 10 TeV to 7 TeV.

The Monte Carlo production used in CMS has been shown as stable, consistent be-

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

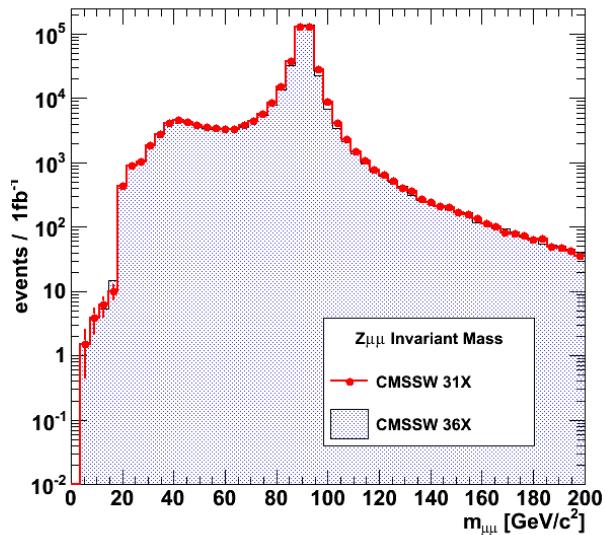


Figure 4.14: - Invariant mass of $Z + jets$ events produced with two different software releases.

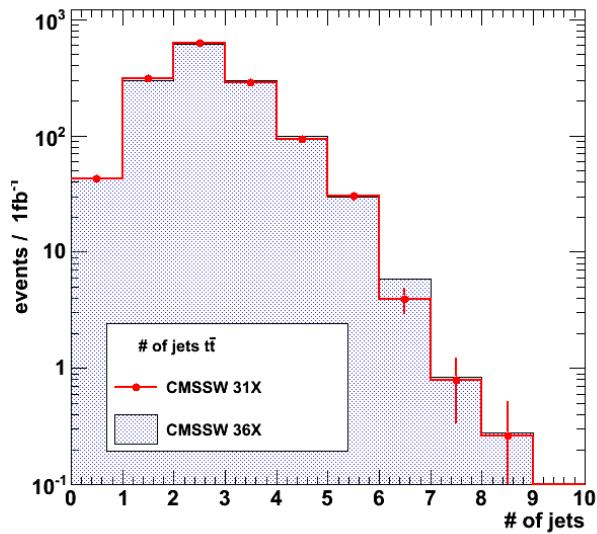


Figure 4.15: - Distribution of the number of jets in $t\bar{t}$ events produced with two different software releases.

4.5 Monte Carlo Datasets

Sample	number of events	Equiv. Lumi (pb^{-1})
$Z \rightarrow \mu\mu$	2085340	1314
$W \rightarrow \mu\nu$	2015438	291
$t\bar{t}$	587460	3830
SingleTop tW	472965	44423
WW	109480	2867
WZ	115040	6455
ZZ	147332	24639

Table 4.1: Background samples used in the analysis, Summer09 production at 7 TeV

Sample	number of events	Equiv. Lumi (pb^{-1})
HWW120 $H \rightarrow WW^* \rightarrow 2l2$ gluon fusion	103700	419.595
HWW130	108200	238.926
HWW140	109550	168.731
HWW150	109550	139.046
HWW155	97850	116.199
HWW160	109550	122.123
HWW165	101600	117.106
HWW170	108150	133.698
HWW175	117450	156.261
HWW180	105450	153.803
HWW190	103550	205.615
HWW200	99900	236.457

Table 4.2: Higgs samples used in the analysis, Summer09 production at 7 TeV

tween versions and up to now, consistent with the data. It is not expected to be the last production at this energy, but it is the last official production used in this study and the one that is being currently used for collision data. The Monte Carlo will evolve with the luminosity, and it will be tuned with data.

In the following section, more details on the Monte Carlo production of the signal and background samples will be given.

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

4.6 Signal and Background Samples and Cross-Sections

The Higgs signal samples $pp \rightarrow H \rightarrow WW^* \rightarrow l\bar{l}l\nu$, with $l = e, \mu, \tau$, for various Higgs boson masses between 120 and 200 GeV/c^2 used in thesis have been generated using PYTHIA [68]. All leptonic decays, including $W \rightarrow \tau\nu$, were included in the simulation.

The differential Higgs boson transverse momentum distribution, P_T^H , is sensitive to higher order corrections. To include these effects, a differential reweighting technique has been applied. For that, the leading order (LO) PYTHIA Higgs boson transverse momentum spectrum is reweighted to match the MC@NLO [75] differential distribution, as MC@NLO incorporates next-to-leading order (NLO) matrix element calculations. The reweighting is done by means of P_T dependant k-factors defined like this:

$$K(P_T^H) = \frac{\sigma^{PYTHIA}}{\sigma^{MC@NLO}} \left(\frac{\partial \sigma^{MC@NLO}}{\partial P_T^H} \Big|_{P_T^H} \right) \left(\frac{\partial \sigma^{PYTHIA}}{\partial P_T^H} \Big|_{P_T^H} \right)^{-1} \quad (4.1)$$

Since the differential distributions $\frac{\partial \sigma^X}{\partial P_T^H}$ are not known analytically, the K-factors are computed using binned distributions with a bin size of 2.5 GeV/c . The PYTHIA Higgs transverse momentum distribution for a Higgs mass of $m_H = 160$ GeV/c^2 before and after the application of the k-factors are shown in figure 4.16. The reweighting is only

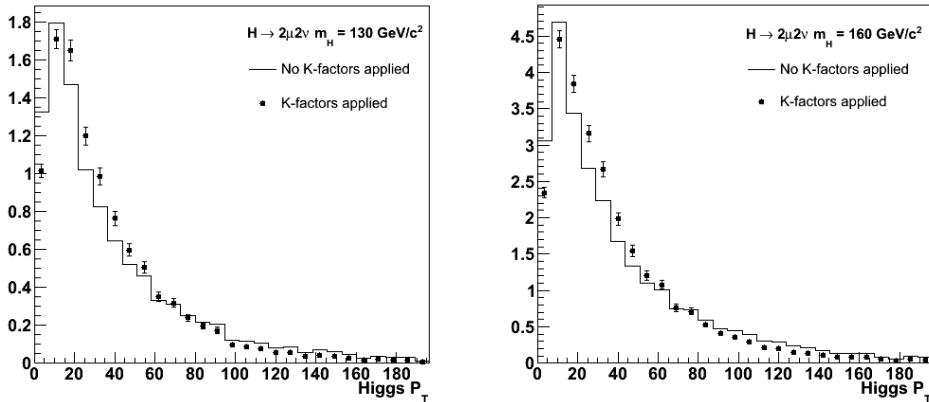


Figure 4.16: - Comparison between the PYTHIA distributions before (solid line) and after the reweighting (dots) for $\sqrt{s} = 10$ TeV, considering $m_H = 130$ GeV/c^2 (left) and $m_H = 160$ GeV/c^2 (right).

applied to Higgs events originating from gluon fusion processes. The cross-sections used for signal and backgrounds are calculated at NLO and PYTHIA is a LO generator. The ratio of gluon fusion to vector-boson fusion in the samples has to correspond to the ratio of their NLO cross-sections, and for that, each vector-boson fusion event is multiplied by the k-factor:

$$K_{VBF} = \frac{\sigma_{VBF}^{NLO} \times \sigma_{GF}^{LO}}{\sigma_{VBF}^{LO} \times \sigma_{GF}^{NLO}} \quad (4.2)$$

4.6 Signal and Background Samples and Cross-Sections

where the cross-sections have been computed using PYTHIA (LO) and the dedicated programs HIGLU [76; 77; 78] (gluon fusion at NLO) and VV2H [79] (vector-boson fusion at NLO). The resulting k-factor for all Higgs mass hypothesis under investigation is $K_{VBF} \approx 0.5$.

For the initial 14 TeV analysis, 9 Higgs samples, for masses from 120 to 200 GeV/c^2 were produced. The NLO Higgs cross-section [80] times $H \rightarrow WW^* \rightarrow l\nu l\nu$ branching ratio for the different Higgs masses considered in the analysis at 14 TeV are shown in table 4.3. HIGLU is used to calculate the NLO cross-sections, and the branching ratios

$m_H [GeV/c^2]$	$\sigma^{NLO} [pb]$	$BR(H \rightarrow WW^*)$	$\sigma^{NLO} \times BR(H \rightarrow WW^* \rightarrow l\nu l\nu)$
120	40.31	0.13	0.56
130	35.08	0.28	1.06
140	31.02	0.48	1.58
150	27.55	0.68	1.98
160	24.63	0.89	2.34
170	22.18	0.96	2.26
180	20.17	0.93	1.99
190	18.43	0.77	1.51
200	16.76	0.73	1.30

Table 4.3: NLO Higgs cross-section times $H \rightarrow WW^* \rightarrow l\nu l\nu$ branching ratio for the different Higgs masses at $\sqrt{s} = 14$ TeV

are computed with HDECAY [81; 82].

The analysis had to be adapted to lower centre of mass energies, first to 10 TeV and then to 7 TeV. The main impact of changing the centre of mass energy is the change in signal and background cross-sections. Processes with higher invariant masses naturally get hit harder; also, processes originating from primary gluons, like gluon fusion Higgs suffer more than those involving qq initial states, like most of predominant backgrounds for Higgs searches, as WW for example.

For the 10 TeV and 7 TeV analysis, the gluon fusion next-to-next-to-leading order (NNLO) cross-sections were also estimated. When using NNLO cross-sections, the k-factor has to be modified consequently:

$$K_{VBF} = \frac{\sigma_{VBF}^{NLO} \times \sigma_{GF}^{LO}}{\sigma_{VBF}^{LO} \times \sigma_{GF}^{NNLO}} \quad (4.3)$$

The cross-sections have been calculated at different orders in perturbative QCD for the four main Higgs production mechanism at the LHC, using dedicated tools.

- Associated production with Z/W (HV), calculated at NLO using V2HV [83].
- Associated production with a top pair (Htt), at LO using HQQ [84].

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

- Vector-Boson fusion (VBF), at NLO using VV2H.
- Gluon-Fusion (GF), the main production mode, calculated, as previously said, at (N)NLO using HIGLU and HggTotal [85].

As before, the Branching ratios of the SM Higgs are computed with the tool HDECAY.

Five gluon HWW signal samples from 130 to 200, plus one additional 500 GeV/c^2 sample, were produced first for 10 TeV and then for 7 TeV. Many other mass points were further produced for a centre of mass energy of 7 TeV as well as some Higgsstrahlung VBF samples. The NNLO and NLO Higgs cross-section times $H \rightarrow WW^* \rightarrow l\nu l\nu$ branching ratio for some of the masses considered are shown in tables 4.4 and 4.5.

m_H [GeV/c^2]	$\sigma^{NNLO} \times BR(H \rightarrow WWW \rightarrow l\nu l\nu)$
130	0.917
150	1.638
160	1.891
170	1.735
200	0.944
500	0.141

Table 4.4: NNLO Higgs cross-section times $H \rightarrow WW^* \rightarrow l\nu l\nu$ branching ratio for the different Higgs masses at 10 TeV

m_H [GeV/c^2]	$\sigma^{NNLO} \times BR(H \rightarrow WWW \rightarrow l\nu l\nu)$
130	0.516
150	0.897
160	1.022
170	0.926
200	0.487
500	0.053

Table 4.5: NNLO Higgs cross-section times $H \rightarrow WW^* \rightarrow l\nu l\nu$ branching ratio for the different Higgs masses at 7 TeV

The question of how precisely the signal cross-sections are known is difficult to address, especially when experimental cuts are applied on the hadronic structure of the final state. The efficiency of these cuts is very dependent on the modeling of the hadronic structure. In this thesis, PYTHIA is used to provide the mentioned structure, through a parton-shower algorithm. However, high precision, fixed-order theoretical calculations do usually not allow for an evolution of the partons in such a shower (an exception to this is MC@NLO); thus the hadronic final state structure is not perfect. A more extensive review on this subject can be found in [86] where it is discussed that if the restriction on the final state phase-space by cuts on the hadronic structure is not very severe, the efficiencies of fixed-order and parton-shower calculations agree very well. As

4.6 Signal and Background Samples and Cross-Sections

$\sigma_{GF}^{NLO} \times \text{BR}(H \rightarrow WWW \rightarrow l\nu l\nu)$			
m_H [GeV/c ²]	14 TeV	10 TeV	7 TeV
130	1.06	0.56	0.31
150	1.98	1.01	0.53
160	2.34	1.18	0.61
170	2.26	1.12	0.58
200	1.30	0.60	0.31

Table 4.6: NLO gluon fusion Higgs cross-section times $H \rightarrow WW^* \rightarrow l\nu l\nu$ branching ratio for the different Higgs masses at 14, 10 and 7 TeV

a consequence no additional systematic uncertainty on the signal cross-section due to the unknown hadronic structure is applied after the selection cuts.

Process	WW	WZ	ZZ	tW	$t\bar{t}$	$W \rightarrow \mu\nu$	$Z \rightarrow \mu\mu$
$\sigma^{NLO}[pb]$ 14 TeV	114.3	49.9	15.3	62	840	20283.7	3259.7
$\sigma^{NLO}[pb]$ 10 TeV	74.1	31.4	9.9	26	415	14253.7	2323.6
$\sigma^{NLO}[pb]$ 7 TeV	42.9	18.3	5.9	10.5	165	9679.9	1606.6

Table 4.7: NLO cross-sections for the different background processes at different centre of mass energies

For the considered backgrounds, the cross-sections at 14, 10, and 7 TeV are computed at NLO with MCFM [87]. The results for the different backgrounds at the considered centre of mass energies are listed in table 4.7.

In the case of 14 TeV samples, the sample for the WW continuum background was produced using MC@NLO, the matrix-element generator Alpgen was used to produce $t\bar{t}$, $W + jets$, and Drell-Yan (DY) samples, Wt was produced using TopRex, where the parton shower is provided by PYTHIA, while the rest of the backgrounds were generated with PYTHIA.

For 10 TeV, the WW continuum background was produced using PYTHIA for the $qq \rightarrow WW$, while the sample from $gg \rightarrow WW$, not present in the previous production, was produced using the $ggWW$ generator. The matrix-element generator Madgraph was used for $t\bar{t}$, $W + jets$, DY, $V + \gamma$ ($V = W/Z$), and single top processes. All other background samples were generated with PYTHIA.

In the case of the 7 TeV estimation, the Monte Carlo production was done in a similar way. Additional QCD and $\gamma + jets$ samples were also produced for lepton fake rate studies using PYTHIA for 14, 10 and 7 TeV.

The characteristics of signal and backgrounds for the analysis have been established in

4. SEARCH FOR A SM HIGGS IN THE WW^* DECAY CHANNEL

the present chapter. In the following chapter, the main physics observables that will be used to perform the analysis are presented and their reconstruction and identification methods in the CMS experiment will be detailed.

5

Physics Observables

As was stated in the previous chapter, the signature of the signal studied in this thesis, $gg \rightarrow H \rightarrow WW^* \rightarrow \mu^+\nu_\mu\mu^-\bar{\nu}_\mu$, is defined by a final state that contains two high- P_T isolated muons with opposite charge, no central hard jet activity and a significant amount of missing E_T .

The high level objects used in the analysis are therefore Muons, Jets and missing E_T . In the following sections, the Muon Reconstruction/Identification/Isolation, Jet Reconstruction/Identification and Jet E_T measurement and the missing E_T determination in the CMS experiment will be discussed.

5.1 Muons

5.1.1 Muon Reconstruction

The ability to identify and reconstruct muons with high efficiency over the whole kinematic range of the LHC is one of the key elements to the success of the CMS experiment and in particular for the $H \rightarrow 2\mu 2\nu$ analysis. Algorithms that are robust and flexible are required and the use of all the available detector information over the full geometrical acceptance of the CMS detector is recommended.

The reconstruction of muons in CMS [88; 89] combines tracking and calorimeter information. The high-level muon physics objects are reconstructed in a multi-faceted way, with the final collection being composed by three different muon types, Standalone, Global and Tracker muons.

The muon reconstruction chain starts with the local reconstruction of hits in the muon spectrometer. First, hits in DTs, CSCs and RPCs are reconstructed from digitized electronics signals. Hits within each DT and CSC chamber are then matched to form segments (track stubs).

5. PHYSICS OBSERVABLES

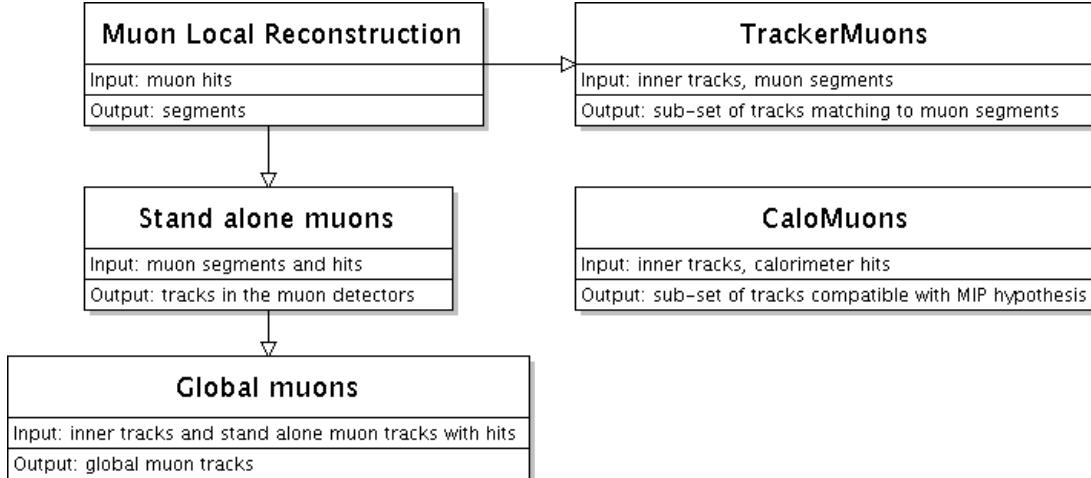


Figure 5.1: - Muon reconstruction overview.

In the offline reconstruction, these segments reconstructed in the muon chambers are used to generate seeds consisting of position and direction vectors and an estimate of the muon transverse momentum. These initial estimates are used as seeds for the track fits in the muon system, which are performed using segments and hits from DTs, CSCs and RPCs and are based on the Kalman filter technique. The result is a collection of track objects reconstructed in the muon spectrometer, which are referred to as Standalone Muons.

For each standalone muon track, a search for tracks matching it amongst those reconstructed in the inner tracking system (referred to as tracker tracks, inner tracks or silicon tracks) is performed, and the best-matching tracker track is selected. For each tracker track - standalone muon pair, the track fit using all hits in both tracks is performed, again based on the Kalman filter technique. The result is a collection of track objects referred to as Global Muons. References to the matching tracker tracks, standalone muon tracks, and Global muon tracks are assembled into one single collection of muon objects.

Complementary to the global-muon reconstruction, an alternative will be to consider all tracker tracks to be potential muon candidates and checking this hypothesis by looking for compatible signatures in the calorimeters and in the muon system. Tracker tracks identified as muons by this method are referred to as Tracker Muons.

The tracker-muon algorithm is particularly useful for the identification of low- P_T muons (with P_T of the order of several GeV/c), which may not leave enough hits in the muon stations for a standalone muon to be reconstructed. The default criteria for tagging a tracker track as Tracker muon are very loose (every track with $p > 3\text{GeV}/c$ and $P_T > 1.5 \text{ GeV}/c$ matched with at least one segment in the muon stations), so Tracker

muons should in general not be used without further requirements.

Calorimeter-based muons, or Calo muons for short, represent a subset of all tracker tracks reconstructed in the event, which includes tracks with energy depositions in the calorimeters compatible with those of a minimum-ionizing particle. The fake rate of these muon candidates is high and they should not be used when muon purity is essential. A typical use case for calo muons is the reconstruction of the J/ψ decaying to low-momentum muons that have little or no information in the muon system, thus improving signal to background ratio compared with the inner tracks.

The figure 5.2 shows a direct comparison of the reconstruction efficiencies as a function of P_T , η and ϕ . The values are almost the same for $P_T \geq 10 \text{ GeV}/c$. The structure visible in the efficiency as a function of ϕ , for the seed, standalone and global reconstruction, is due to the ϕ -acceptance of the muon system. The efficiency as a function of the muon transverse momentum is also shown. The muon reconstruction efficiency increases up to a plateau which is approximately constant from $8 \text{ GeV}/c$ to $1 \text{ TeV}/c$ (starting from $P_T = 5 \text{ GeV}/c$ more than 50% of muons are reconstructed). At TeV momenta the muon reconstruction efficiency decreases slowly, due to the effects of bremsstrahlung on finding correct seeds.

These algorithms fulfill all the necessary requirements for the reconstruction of single muons with full detector simulation. Muons are identified and reconstructed with efficiencies close to 99% with clearly defined and understood efficiency losses due to the CMS detector geometry. The final output from the algorithms is a muon physics object together with a compatibility value indicating the probability of the track being a muon. These algorithms satisfy all of the requirements for robust high efficiency reconstruction and are the foundation for the data taking and real analysis. They can also be tuned for the more complex environments of real events and actual detector inefficiencies.

For this study, Global muons are used. Each sub-detector is able to measure a part of the properties of the muon, the concept of a Global muon combines information from multiple sub-detectors in order to obtain the most accurate possible description of the muon, where the track parameters are measured in two sub-detectors: the inner tracker, and the muon system. The momentum resolution of muon tracks up to $P_T = 200 \text{ GeV}/c$ reconstructed in the muon system alone is dominated by multiple scattering. At low momentum, the best momentum resolution for muons is obtained from the silicon tracker. At higher momentum, however, the characteristics of the muon system allow the improvement of the muon momentum resolution by combining the muon track from the silicon detector, tracker track, with the muon track from the muon system, standalone muon, into a Global muon track. The reconstruction of Global muon tracks begins after the completion of the reconstruction of the central tracker tracks and the muon system tracks.

5. PHYSICS OBSERVABLES

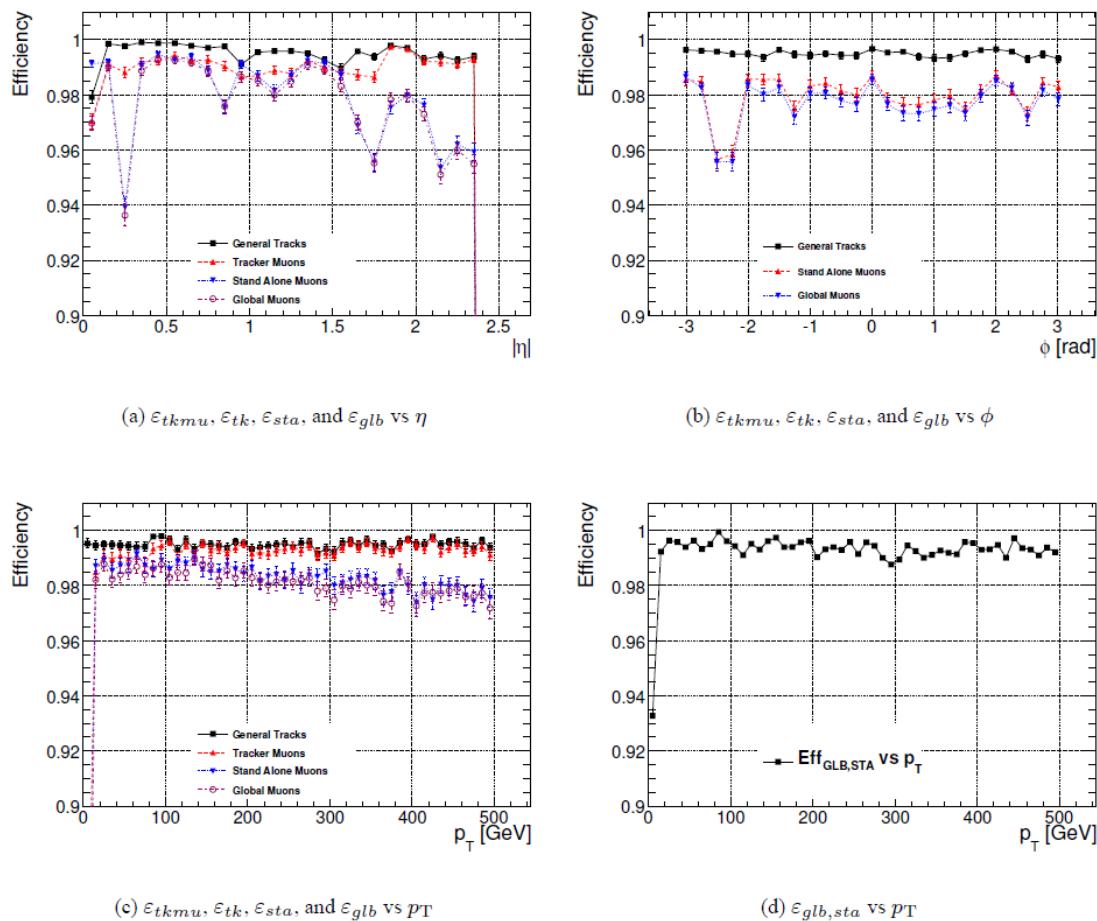


Figure 5.2: - Efficiencies of the different muon reconstruction steps as a function of η , ϕ and P_T .

5.1.2 Muon Identification

The selection criteria for the Muons that are used in the analysis is presented in this subsection. Several Cosmic ray data campaigns in CMS prior to LHC collisions led to well developed Muon identification methods that are already being useful in the data taking and that will be the key of many physics analysis. To illustrate the need for muon identification [90], figure 5.3 shows the probability that a generated kaon/pion in the $|\eta| < 2.4$ region results in a reconstructed muon, as a function of P_T .

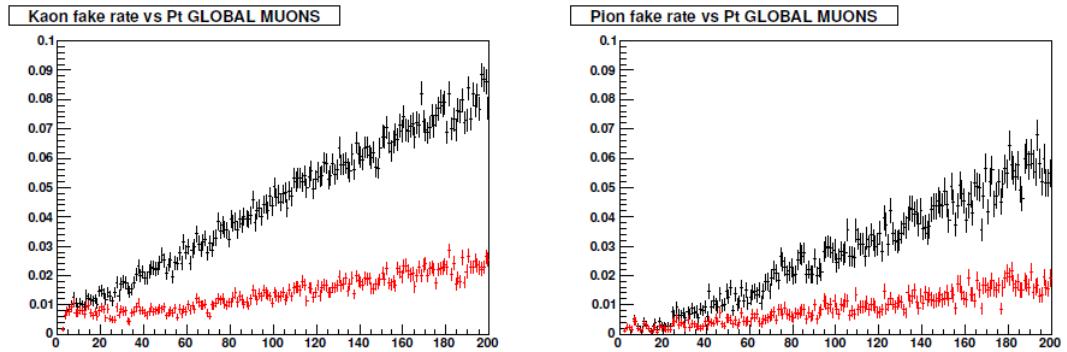


Figure 5.3: - Global Muon and Tracker Muon fake rates as a function of P_T . The total fake rate is presented in black and decays in flight are presented in red.

In the software of CMS there are several muon categories that correspond to different quality requirements. These selectors are based either on parameters used in the muon identification, like the impact parameter, or in the compatibility with different properties in various subdetectors. The following selection types are available:

- *AllGlobalMuons*: checks if the muon is a Global muon
- *AllStandAloneMuons*: checks if the muon is a Standalone muon
- *AllTrackerMuons*: checks if the muon is a Tracker muon
- *TrackerMuonArbitrated*: resolves ambiguity of sharing segments
- *AllArbitrated*: all muons with the Tracker muon arbitrated
- *GlobalMuonPromptTight*: Global muons with tighter fit requirements
- *TMLastStationLoose*: penetration depth loose selector
- *TMLastStationTight*: penetration depth tight selector
- *TM2DCompatibilityLoose*: likelihood based loose selector
- *TM2DCompatibilityTight*: likelihood based tight selector

5. PHYSICS OBSERVABLES

- *TMOneStationLoose*: requires one well matched segment
- *TMOneStationTight*: require one well matched segment with a tighter criteria
- *TMLastStationOptimizedLowPtLoose*: combination of *TMLastStation* and *TMOneStation*
- *TMLastStationOptimizedLowPtTight*: combination of *TMLastStation* and *TMOneStation*
- *GMTkChiCompatibility*: require tracker stub to have good χ^2 relative to global track
- *GMSaChiCompatibility*: require standalone stub to have good χ^2 compatibility relative to global track
- *GMTkKinkTight*: require a small kink value in the tracker stub

The selectors are used depending on the specific needs of each analysis, as they are optimized in a different way. The muon identification algorithm of choice for the analysis was designed to suppress hadronic punch-through and muons from decays in flight (GlobalMuonPromptTight). The Muon has to be reconstructed as a Global muon, with a normalized- $\chi^2 < 10$.

The Global muon normalized- χ^2 is powerful to reject both decays in flight and punch-through. When using this variable, the long tails present in the signal muon distribution, even in Monte Carlo events, have to be considered. The efficiencies for a cut on normalized- $\chi^2 < 10$ or < 5 are displayed in figure 5.4. In what follows a normalized- $\chi^2 < 10$ is required. The distribution of the normalized- χ^2 for events coming from different processes is presented in 5.5.

To further suppress non-prompt muons the following additional track-quality selection cuts using the tracker-track information are used:

A loose cut in the xy impact parameter of the silicon fit, $d0$. This is the distance of closest approach of a track to the primary vertex. A cut in this variable is efficient for prompt muons and rejects a significant fraction of decays in-flight. If the cut on $d0$ is tight enough, muons from b and c -quark decays could be rejected. A baseline cut $|d0| < 2\text{ mm}$ is applied, preserving the efficiency for muons from bottom and charm, then, a tighter cut can be used to eliminate these muons. Specifically for this analysis, a cut in the significance of the impact parameter, $\frac{|d0|}{\sigma_{d0}} < 3$ is also applied. In the figure 5.6, the impact parameter distribution and its significance are presented for muons coming from different kind of events. For muons coming from W decays, as in the case of the Higgs signal, the impact parameter is small, while it is larger in the case of muons coming from secondary vertex, in heavy quark decays.

The number of valid hits used for the silicon fit also shows discriminating power. For that, a cut in $N_{hits} \geq 11$ is applied. The distribution of the number of hits can be seen

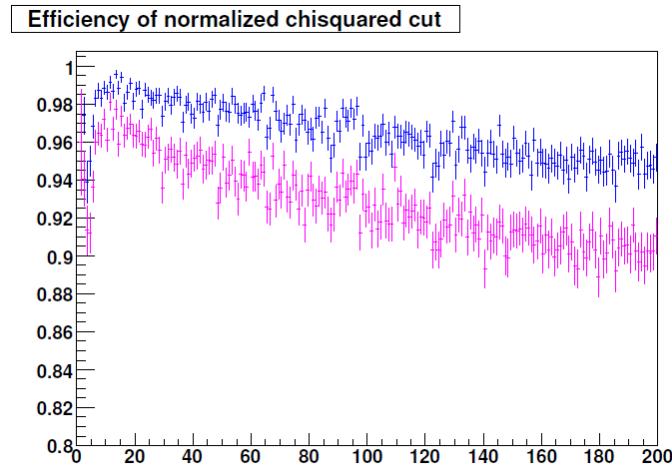


Figure 5.4: - Efficiency for muons requiring normalized $\chi^2 < 5$ (pink) and normalized- $\chi^2 < 10$ (blue) as a function of P_T .

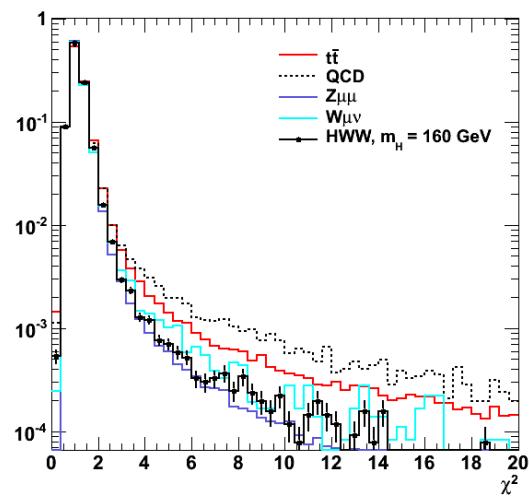


Figure 5.5: - Distribution of the normalized- χ^2 for Global Muons in $|\eta| < 2.4$ normalized to the unit area for different processes.

5. PHYSICS OBSERVABLES

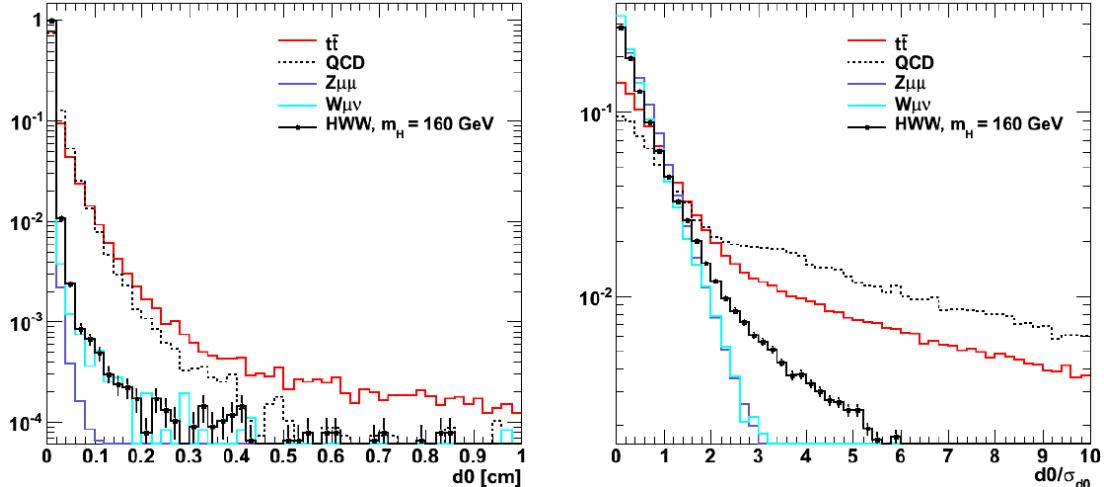


Figure 5.6: - Distributions of the Impact parameter $|d0|$ (left) and Impact parameter significance $\frac{|d0|}{\sigma_{d0}}$ (right) for Global Muons in $|\eta| < 2.4$ normalized to the unit area for different processes.

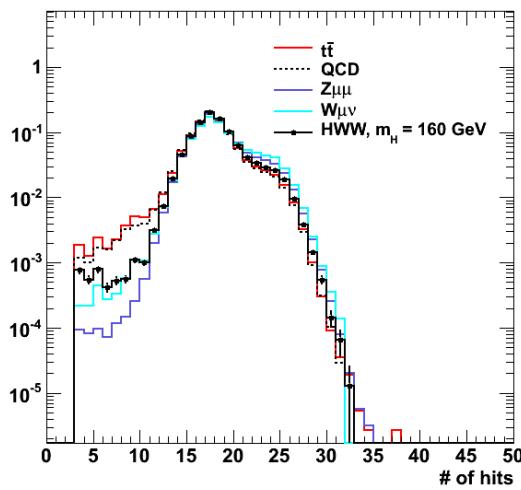


Figure 5.7: - Distribution of the number of valid hits for Global Muons in $|\eta| < 2.4$ normalized to the unit area for different processes.

in 5.7, muons coming from QCD and b -decays (top events) present fewer than 10 valid hits more often than prompt muons.

5.1.3 Muon Isolation

Muons coming from the signal are expected to be isolated with no substantial deposits in the tracker and in the calorimeters.

A cone around the muon is defined as shown in figure 5.8 and the tracker and calorimeter depositions within it are considered. The muon is considered isolated if these depositions do not exceed some thresholds. The application of the isolation in the muon selection helps to reduce background coming from many sources, especially QCD , but also $t\bar{t}$ and $W+jets$. The QCD background is expected to be negligible in the analysis, however, the cross-section of the QCD processes at the LHC is so large that, as it will be discussed in C.1, for $100 pb^{-1}$, ~ 30 events with at least one good muon are expected for a Higgs signal with a mass of $m_H = 160 GeV/c^2$ and of the order of 170M events are expected for QCD.

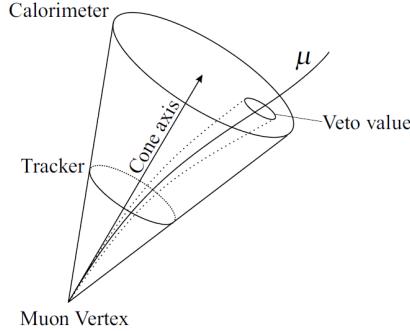


Figure 5.8: - Schematic illustration of the isolation cone. The muon direction at the vertex defines the cone axis.

Many isolation criteria have been studied, using information from the tracker and the calorimeters. The choice made for the analysis is an hybrid isolation criteria, P_T dependant, based on a combination of information in the tracker, HCAL and ECAL.

The tracker-based isolation variable is defined as the scalar sum of the P_T of the tracks in the $\eta \times \phi$ plane inside a cone of $\Delta R = 0.3$

$$\Delta R = \sqrt{\Delta\Phi^2 + \Delta\eta^2} \quad (5.1)$$

around the muon ($0.01 < \Delta R < 0.3$ to veto the muon track).

$$SumP_T = \sum_{0.01 < \Delta R < 0.3}^{tracker} P_T^{tracks} \quad (5.2)$$

5. PHYSICS OBSERVABLES

In the calorimeters, the $hadE_T$ and emE_T variables, corresponding to the HCAL and ECAL respectively, are calculated as the scalar sum of the transverse energy of the calorimeter towers in a cone of $\Delta R = 0.3$ around the muon ($0.01 < \Delta R < 0.3$).

$$hadE_T = \sum_{\substack{HCAL \\ 0.01 < \Delta R < 0.3}} E_T^{towers}$$

$$emE_T = \sum_{\substack{ECAL \\ 0.01 < \Delta R < 0.3}} E_T^{towers} \quad (5.3)$$

The tracker isolation is more important than the calorimetric, but a combination of ECAL, HCAL and Tracker isolation performs better. A combined isolation variable is built,

$$Iso_\mu = SumP_T + hadE_T + emE_T \quad (5.4)$$

After a study on the performance of cuts on this combined Isolation variable divided by the P_T of the muon, a better performance for low P_T muons (15–25 GeV/c) is found, as the cut becomes tighter as the P_T decreases. At higher P_T ($P_T > 50$ GeV/c), the cut becomes too loose, and for that region, a cut on absolute isolation is imposed.

The Iso_μ distributions as a function of P_T for muons coming from Higgs signal ($m_H = 160$ GeV/c 2) and for muons coming from QCD events are shown in figure 5.9, the black line shows the applied cut.

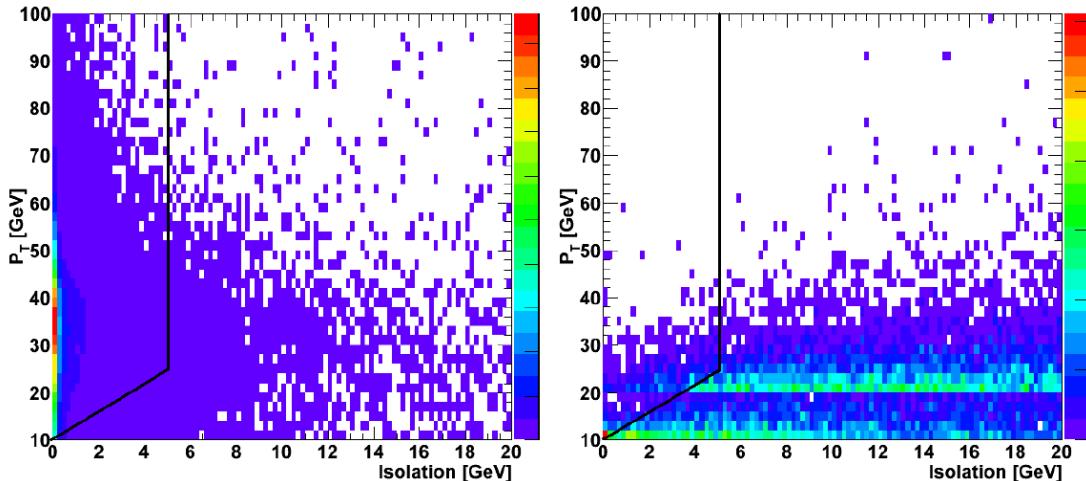


Figure 5.9: - Isolation variable vs. P_T distributions for real muons from $H(160) \rightarrow WW \rightarrow 2\mu 2\nu$ (left) and from fake muons from QCD events (right). The black line indicates the cut position.

The isolation criteria for the analysis is, therefore, the one that can be derived from the S to B relation in the 2-dimensional plot:

- Muons with $P_T > 25 \text{ GeV}/c$, are considered isolated if $Iso_\mu < 5 \text{ GeV}/c$
- Muons with $P_T < 25 \text{ GeV}/c$, however, are considered isolated if $Iso_\mu < \frac{P_T^\mu - 10.0}{3}$

5.1.4 Muon selection

In addition to the muon identification and isolation criteria previously discussed, the two muons in the analysis are required to be within the acceptance $|\eta| < 2.4$ with at least one of them inside $|\eta| < 2.1$ to ensure that one of the muons corresponds with the one that fires the trigger, as the Resistive Plate Chambers (RPC), dedicated to the trigger, cover in both barrel and endcaps the region up to $|\eta| < 2.1$.

The P_T and η distributions for all reconstructed muons without any kind of selection are presented for $t\bar{t}$ background and signal in figure 5.10.

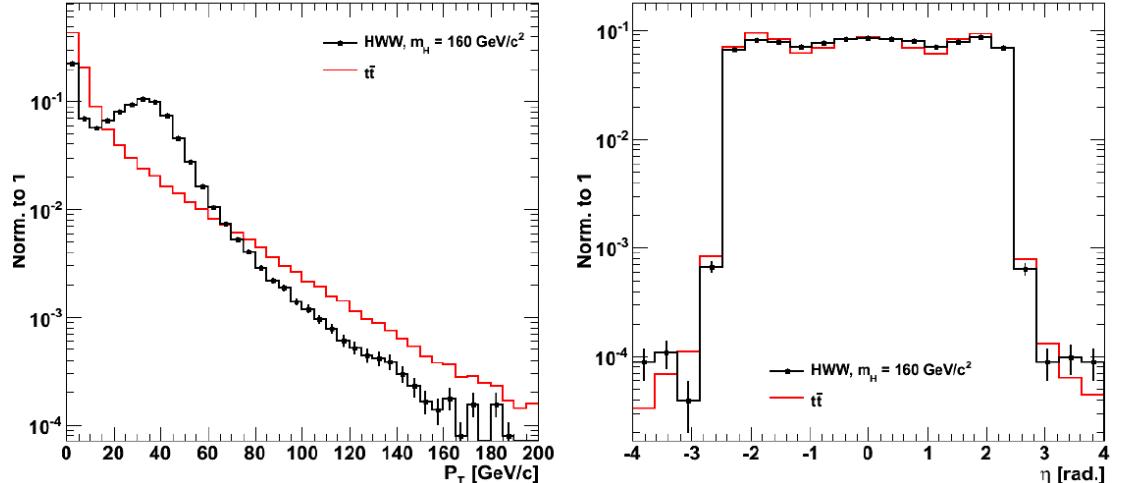


Figure 5.10: - P_T (left) and η (right) distributions for muons coming from $t\bar{t}$ (solid line) and Higgs to WW^* signal for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ (dots).

The muons used in the analysis are required to have a transverse momentum higher than 10 and 20 GeV/c . If more than two muons of this kind are found, the event is rejected. This measure efficiently reduces the contamination of WZ and ZZ backgrounds. Additionally, as they come from the decay of a neutral Higgs into two opposite-signed W bosons, they also must have opposite charge.

The use of the proposed muon identification plus the hybrid isolation when compared

5. PHYSICS OBSERVABLES

with a baseline muon identification as for example just Global muons, with an absolute isolation criteria, leads to a substantial improvement in terms of background reduction, especially in the case of $W + jets$, potential source of fake muons, as can be seen in table 5.1.

Process	WW	WZ	ZZ	$t\bar{t}$	$W + jets$	$Z + jets$
Basic Muon ID/iso	454	474	19.62	3216	139	430874
Current ID/iso	405	451	18.43	2774	4	404683

Table 5.1: Number of expected events at Muon selection level considering $1fb^{-1}$ of integrated luminosity for different backgrounds at a centre of mass energy of 14 TeV.

To summarize the muon selection used in this analysis, events are selected if they have two muons satisfying:

- Global muons, built using all the information possible in muon chamber and tracker,
- GlobalMuonPromptTight, normalized- $\chi^2 < 10$,
- $P_T > 20$, $P_T > 10$ GeV/c
- $|\eta| < 2.4$ (at least one of them within $|\eta| < 2.1$)
- $|d0| < 0.2$ cm,
- $\frac{|d0|}{\sigma_{d0}} < 3$,
- $N_{hits} > 10$,
- Isolated (see 5.1.3).

5.2 Electrons

Final states with electrons have to be addressed in the work as they complement the study and contribute to the sensitivity in the Higgs to WW^* channel. In this section the electron reconstruction, identification and isolation in CMS will be introduced.

5.2.1 Electron Reconstruction

The reconstruction of electrons in CMS [91; 92] uses information from the pixel detector, the silicon strip tracker and the ECAL. Electron energy is detected in the process of electromagnetic shower of the calorimeter. The shower shapes differentiate between electrons and hadrons using the different EM and hadronic shower development. Electron energy is deposited mostly in the ECAL. The measurement of the electron energy in the ECAL is hampered by the amount of tracker material which is distributed in front

of the ECAL, and by the presence of a strong magnetic field aligned with the collider beam axis. Electrons traversing the silicon layers of the tracker radiate bremsstrahlung photons and the energy that reaches the ECAL has a significant spread in the azimuthal direction ϕ .

The electromagnetic showers initiated by electrons or photons deposit their energy in several crystals of the ECAL. For a single electron or photon reaching the ECAL, most of the energy is collected in a small number of crystals. To obtain a measurement of the electron energy at primary vertex and minimize the cluster containment variations, it is essential to collect the bremsstrahlung photons produced. This is the purpose of the super-clustering algorithms, that build clusters of clusters to take into account the ϕ spread and collect the bremsstrahlung energy.

The electron reconstruction relies on a dedicated Gaussian Sum Filter (GSF) method [93; 94] using track and clusters in the ECAL. The track can be either seeded by an ECAL-driven approach where super-clusters are used to filter the collection of standard trajectory seeds or by the reconstruction of tracks matched with ECAL ParticleFlow clusters in a tracker-driven approach. CMS provides reconstruction algorithms for electron objects that define a standard preselection of electrons suitable for most analyses involving electrons, as well as the resolution of multiple electron candidates and the algorithms to define the final electron momentum and energy.

5.2.2 Electron identification

For electron selection [95], the following variables are used to discriminate between real and fake electrons:

- Super-cluster energy / track momentum at vertex
- Delta Eta between Super-cluster position and track direction at vertex extrapolated to ECAL assuming no radiation
- Delta Phi between Super-cluster position and track direction at vertex extrapolated to ECAL assuming no radiation
- Ratio of energy in HCAL behind Super-cluster to Super-cluster energy
- Energy in 3×3 crystals / energy in 5×5 crystals
- Energy of closest basic cluster to track impact point at ECAL / outermost track momentum
- Energy of closest basic cluster to track impact point at ECAL / innermost track momentum
- Delta Phi between track impact point at ECAL and closest basic cluster

5. PHYSICS OBSERVABLES

- $1/E(\text{Super-cluster}) - 1/p(\text{track at vertex})$
- Bremsstrahlung fraction = $(\text{track momentum at vertex} - \text{track momentum at ECAL}) / (\text{track momentum at vertex})$
- Sigma Eta Eta cluster shape covariance
- Sigma Phi Phi cluster shape covariance

A selected subset of the above variables can be combined to perform electron ID using cuts, or multivariate techniques such as a likelihood, or neural networks.

The electron identification criteria for analysis in CMS is based on the three cut-based standard selections defined by the Electron and Gamma Physics Object Group (egamma POG), robust, loose and tight. These levels of identification are optimized for different target levels of efficiency and purity. The loose category was tuned for high efficiency in the Higgs to ZZ to 4e analysis [96]; the medium was tuned for the intermediate case [91], while the tight was tuned for high purity in the low Higgs mass region for the Higgs to WW^* analysis with two electrons in the final state [97].

5.2.3 Electron selection

A baseline for electron selection in the analysis can be proposed using the same P_T and η requirements as for muons:

- GSF Electrons
- Loose ID
- $P_T > 20, P_T > 10 \text{ GeV}/c$
- $|\eta| < 2.4$
- $N_{\text{hits}}^{\text{tracker}} > 10$
- $|d_0| < 0.04 \text{ cm}$
- isolated

In a similar way as for muons, an isolation variable that takes into account the electron deposits in tracker and calorimeters is built:

$$Iso_e = \text{Sum}P_T + \text{had}E_T + \text{em}E_T \quad (5.5)$$

Then, an electron is considered isolated if:

$$\frac{Iso_e}{P_T^e} < 0.1 \quad (5.6)$$

The use of electrons in the analysis requires a slightly different optimization of the cuts and brings different systematic errors and issues to the analysis that need to be addressed in an optimal way. For example, the control of the conversions, the jet cleaning and the treatment of the fakes are different than in the muon channel. In the case of the study of final states concerning electrons, all these aspects have been carefully taken care of and the details can be consulted in [61; 62; 63].

5.3 Jets

Jets are a main ingredient in this study to reconstruct the hadronic content in the event. As was discussed in the first chapter quarks and gluons cannot be observed directly although they can be detected in the tracking and calorimeter systems as they fragment into stable hadrons. They appear in the detector as jets of particles [98] .

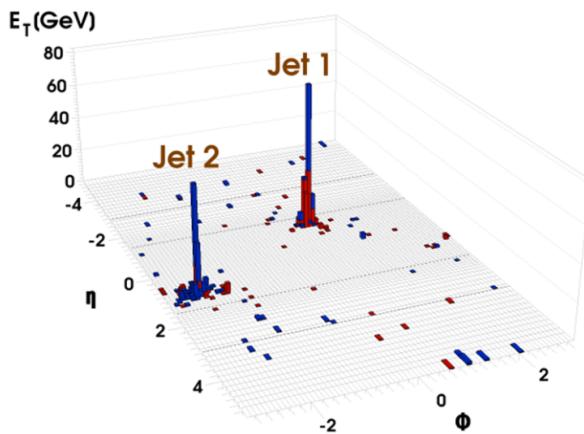


Figure 5.11: - View of the first di-jet event recorded in CMS with $\sqrt{s} = 7$ TeV collision data.

The signal events are characterized by the absence of central hard jet activity, but this is not the case for all the backgrounds, in particular for $t\bar{t}$ events.

The $t\bar{t} \rightarrow b l^+ \nu \bar{b} l^- \bar{\nu}$ background has an inclusive production cross-section more than 40 times larger than the signal at the design centre of mass energy of the LHC. At design energy, of the order of 840 $t\bar{t}$ will be produced in the LHC for each pb^{-1} of luminosity, leading to more than 1.5 million of $t\bar{t}$ events with $2 fb^{-1}$. That characterizes the LHC as a *top*-factory. The main way to reduce this background is by means of the application of a central jet veto, that is, vetoing events with central jets in the analysis.

The kinematics of the $t\bar{t}$ processes slightly favors central jets. In the case of the signal, only vector-boson induced events present jets in the final state, and in that case such

5. PHYSICS OBSERVABLES

jets are emitted with high rapidity. Therefore, only central jets, $|\eta| < 2.5$, are considered for the jet veto.

The central jet veto also suppresses QCD , $Z + jets$, $W + jets$, and other sources of background with hadronic activity, and avoids potential problems with fake jets in the forward region, where the jet reconstruction is expected to be more difficult.

The Jet identification and reconstruction in CMS is a complicated and sensitive task that involves the calorimeters and the tracking systems. Two jet flavors were studied in this analysis, *CaloJets*, that is, jets constructed using information from CaloTowers in the ECAL and HCAL, and Particle Flow (PF) Jets, coming from Particle Flow candidates. The possibility of using TrackJets, especially in the early data-taking scenario, is considered. Different Jet algorithms are also studied. The Jet Plus Tracks (JPT) corrections are also studied, as another level between Calo and Track Jets.

5.3.1 Jet Flavors

5.3.1.1 Calo Jets

The calorimeter and tracking systems participate in the detection of the stable hadrons produced in the fragmentation of the partons. There are several algorithms that can be used to cluster energy deposits in the CMS calorimeters into collimated collections of stable particles [99], jets.

Calorimeter jets, CaloJets in short, are reconstructed using energy deposits in calorimeter towers as inputs. They are composed of one or more HCAL cells and corresponding ECAL crystals. The unweighted sum of energy deposits of one single HCAL cell and 5×5 ECAL crystals form a projective tower in the barrel region, $|\eta| < 1.4$. A more complex association between HCAL cells and ECAL crystals is required in the forward region.

Calorimeter jets bring a good description of both the parton-level and the hadron showers emerging from the hard interaction.

5.3.1.2 Particle Flow Jets

The particle-flow event reconstruction [100], which will be discussed also in following sections, aims to reconstruct and identify all stable particles in the event, that is, electrons, muons, photons, charged hadrons and neutral hadrons, combining all CMS sub-detectors towards an optimal determination of their nature, direction and energy. A list of these individual particles can then be used, as if they came from a Monte Carlo event generator, to determine the missing E_T , reconstruct and identify taus from their decay products, quantify charged lepton isolation with respect to other particles, tag b -jets, and also to build jets, amongst others.

The typical jet energy fractions carried by charged particles, photons and neutral hadrons are 65%, 25% and 10% respectively. These fractions ensure that 90% of the jet energy can be reconstructed with good precision by the particle-flow algorithm, both in value and direction. Only 10% of the energy is affected by the poor hadron calorimeter resolution and by calibration corrections of the order of 10 to 20%. As a natural consequence, it is expected that jets made of reconstructed particles be much closer to jets made of Monte Carlo generated particles than jets made from the sole calorimeter information, in energy, direction and content.

5.3.1.3 Track Jets

The jet reconstruction can be performed using only charged tracks, giving as a result the jet objects known as TrackJets [101].

The calorimeters, in which energy is deposited, collect both the charged and the neutral component of the parton hadronization, however, the strong inner CMS magnetic field, causes the CaloTowers to fail in collecting the energy of low momentum charged particles which do not reach the calorimeters.

The tracking momentum measurements are more accurate than the calorimeter measurements for charged particles with energies up to several hundreds of GeV, and the direction of charged particles at the interaction point is extremely well determined by the track reconstruction. For these reasons a multi-jet event in CMS is expected to be cleaner when looking at the tracks instead of the CaloTowers, with less overlap and interference, and less background (e.g. pile-up events). On the other hand, as just $\frac{2}{3}$ of the energy of the jet are carried by charged particles, TrackJets are not optimal to measure jet energies, leading to a worse resolution than Particle Flow or Calo Jets.

Jet finding with charged tracks only is independent from jet finding with calorimetric towers and gives an excellent alternative to identify jets and determine their directions. They also have independent systematic uncertainties that have to be evaluated separately.

5.3.2 Jet Algorithms

Two broad classes of jet algorithms exist, cone and sequential recombination algorithms. In the Cone algorithms, jets are defined as dominant directions of energy flow. A stable cone is defined as a circle of fixed radius R in the $\eta - \phi$ plane such that the sum of all the momenta of the particles within the cone points in the same direction as the center of the circle. Cone algorithms attempt to identify all the stable cones in the event, and most implementations use a seeded approach to do so, in which iterations are done starting from one seed for the center of the cone until the cone is stable. The set of seeds can be taken as the set of initial particles, like over a P_T threshold or as the midpoints between previously-found stable cones. This iterative method fails to

5. PHYSICS OBSERVABLES

identify all the stable cones, leading to infrared or collinear unsafety in the perturbative computations. Another class of jet algorithms, sequential recombination jet algorithms, define the distance between pairs of particles, performing successive recombinations of the pair of closest particles and stops when all resulting objects are too far apart. Algorithms within that class differ by the definition of the distance. By design, these clustering algorithms are infrared and collinear safe to all orders of calculation.

Several basic algorithms are currently implemented and used in CMS:

- Midpoint Cone algorithm
- Iterative Cone algorithm
- Seedless Infrared Safe Cone algorithm
- k_T algorithm
- Anti- k_T algorithm

In this analysis, Iterative, Seedless Infrared Safe and Anti- k_T algorithms are studied.

5.3.2.1 Iterative Cone

The Iterative Cone algorithm is a simple cone-based algorithm. It has a short and predictable execution time in CMS. Calorimeter towers and particles with $E_T > 1$ GeV are considered in descending order as starting points, seeds, for an iterative search for stable cones such that all inputs with $\Delta R \leq$ that certain value are considered. The cone size parameter, ΔR , is defined as follows:

$$\Delta R = \sqrt{\Delta\Phi^2 + \Delta\eta^2} \quad (5.7)$$

A cone is considered stable if its geometric center agrees with the (Φ, η) location of the sum of the constituent four vectors within a certain tolerance. Once a stable cone is found, it is declared a jet and its constituents are removed from the remaining inputs. The algorithm is neither collinear- nor infrared-safe.

The High Level Trigger in CMS uses iterative Cone Calo jets of $\Delta R = 0.5$.

5.3.2.2 Midpoint Cone

The Midpoint Cone algorithm is based as well on an iterative procedure to find stable cones. Unlike in the Iterative Cone algorithm, the infrared safety is addressed by considering the midpoints between each pair of proto-jets which are closer than twice the cone radius R as additional seeds. Each input object, like a tower, particle, etc. can initially be associated with several proto-jets, and a splitting and merging algorithm is applied afterwards to ensure each input object appears in one jet only. Despite these improvements to the cone-based clustering procedure, the algorithm has been shown not to be infrared-safe for QCD orders beyond NLO.

5.3.2.3 Seedless Infrared-Safe Cone

The Seedless Infrared-Safe Cone (SISCone) [102], jet algorithm is the first cone algorithm that is at the same time infrared and collinear safe to all orders of QCD , while fast enough to be used in experimental analysis.

5.3.2.4 k_t

The k_t algorithm [103; 104; 105] is a recombination algorithm that successively merges pairs of nearby objects, namely partons, particles, or calorimeter towers, in order of increasing relative transverse momentum. A single parameter, D , which approximately characterizes the size of the resulting jets, determines when this merging stops. No splitting or merging is involved because each object is uniquely assigned to a jet.

In the case of the k_t , the definition of the distance between the objects that determines the algorithm is given by:

$$d_{ij}^2 = \min(k_{t,i}^{2p}, k_{t,j}^{2p})(\Delta y_{ij}^2 + \Delta\Phi_{ij}^2) \quad (5.8)$$

where $p = 1$. If $p = 0$, it results in the Cambridge/Aachen [106; 107] algorithm.

5.3.2.5 Anti- k_t

The Anti- k_t algorithm [108] is a third variation of the k_t . Considering again the distance:

$$d_{ij}^2 = \min(k_{t,i}^{2p}, k_{t,j}^{2p})(\Delta y_{ij}^2 + \Delta\Phi_{ij}^2) \quad (5.9)$$

in the Anti- k_t , $p = -1$.

Here, soft particles are recombined with hard ones before recombining among themselves, resulting in regular, softresilient, hard jets. This the hallmark of the iterative cone and, in that respect, the Anti- k_T can be seen as an infrared and collinear safe replacement.

The Anti- k_t is the current CMS recommendation for physics analysis with jets, and therefore will be used by default in this study.

5.3.3 Jet Energy Corrections

As the calorimeter response to particles is not linear, the translation from the measured jet energy to the true particle or parton energy is not straightforward. CMS has developed a set of jet corrections [109; 110] that allows the proper mapping of the measured jet energy deposition.

CMS has adopted a factorized solution to the problem of jet energy corrections, where each level of correction takes care of a different effect, shown in Figure 5.13. Each

5. PHYSICS OBSERVABLES

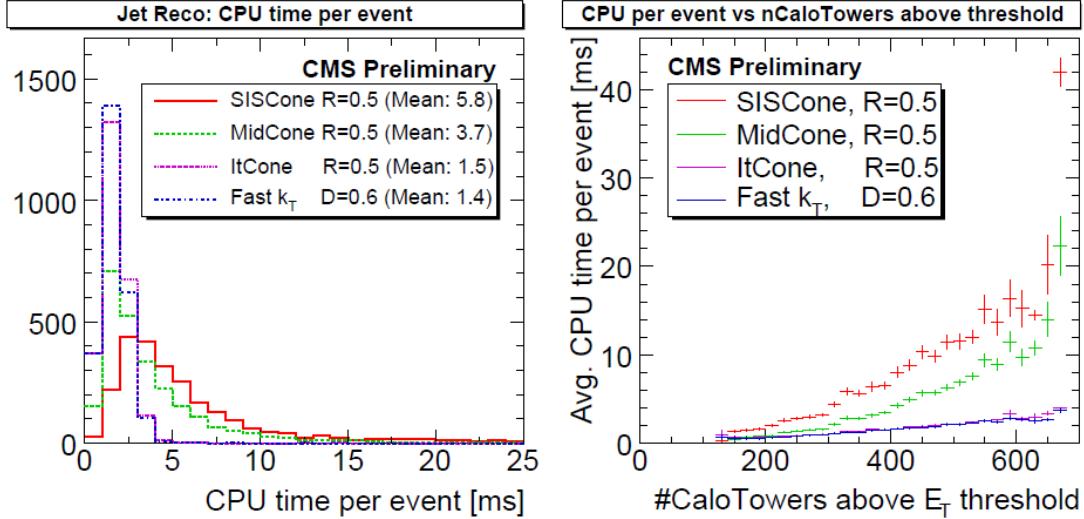


Figure 5.12: - CPU time required for each jet algorithm to cluster all CaloTowers above the E_T threshold of 0.5 GeV into jets (left), and average CPU time as a function of the number of CaloTowers above E_T threshold (right)

level of correction is essentially a scaling of the jet four momentum with by a factor depending on various jet-related quantities, such as P_T , η , flavor, etc., and they are applied sequentially with a fixed order.

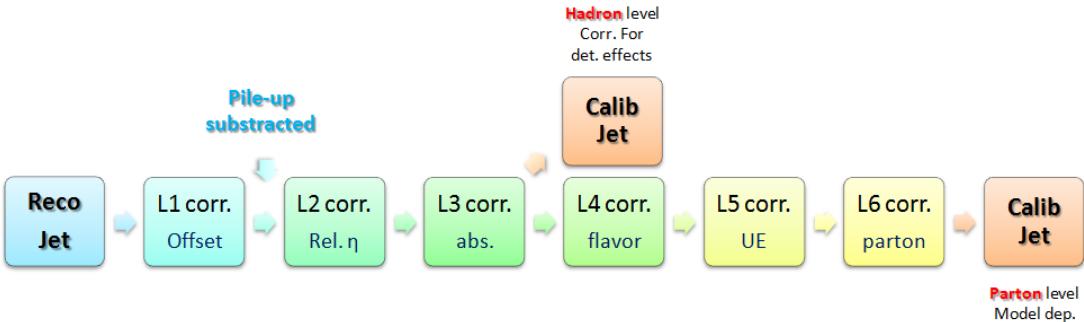


Figure 5.13: - Sequential application of the different levels of correction from Reconstructed to Calibrated Jet in CMS.

The energy of jets considered for the analysis developed in this thesis, as well as for any physics analysis to be performed in CMS, has to be corrected. The combination of L2(Relative) + L3(Absolute) jet corrections is currently the default correction in CMS, and is therefore required here.

The L2 Relative correction makes the jet response flat vs η . The uniformity in pseudorapidity is achieved by correcting a jet in arbitrary η relative to a jet in the central

region ($|\eta| < 1.3$). The derivation of the Relative correction is done either by using MC truth or by employing a data driven method (di-jet balance).

The L3 Absolute correction makes the jet response flat vs P_T . Once a jet has been corrected for eta dependence (L2 relative correction), it is corrected back to particle level (this means that the corrected CaloJet pt is equal on average to the Generated Jet pt). The derivation of the Absolute correction is also done either by using MC truth information or by employing data driven techniques ($Z/\gamma + jet$ balance).

The effect of the Jet Energy Corrections in the analysis can roughly be described as $E_T^{jet} \approx (1.5 - 2) \cdot E_T^{RAW}$ depending slightly on η .

Further corrections can be applied but are not necessary as the motivation of the jet identification in this study is not oriented to jet measurements.

5.3.4 Jet Plus Tracks

The CMS JetPlusTracks [111] algorithm covers the space between Calo and Track Jets providing energy corrections to calorimeter jets using tracks. It is a simple and robust algorithm that improves the jet energy scale and resolution in the transverse energy interval of 20 – 120 GeV and in the barrel region $|\eta| < 1.5$.

Jet Plus Tracks Jets are seeded on CaloJets. The algorithm subtracts the expected, average calorimeter response of charged particles which enter into the jet reconstruction cone at the surface of the electromagnetic calorimeter from the calorimeter jet energy, adding the momentum of these particles measured with the tracker to the jet energy.

The JetPlusTracks corrections are accompanied by the Zero Suppression (ZSP) corrections. The complete energy correction chain (ZSP+JPT) is made then in two steps, in the first, the calorimeter jet energy is corrected for the effect of the ZSP on the calorimeter readouts forming ECAL+HCAL tower, and in the second, the energy after ZSP corrections is corrected again with the JPT algorithm.

The *JPT* corrections can be applied at the same time as the *L2 – L3* corrections discussed before, and JPT Jets can be understood as a complementary step between Track and CaloJets

5.3.5 Jets in the analysis

All the previously discussed methods of jet reconstruction and identification have been contemplated, and four kinds of jets were studied in detail during the work presented in this thesis,

- Calorimeter Jets Anti- k_T L2+L3 corrected

5. PHYSICS OBSERVABLES

- Particle Flow Jets Anti- k_T L2+L3 corrected
- Track Jets Anti- k_T
- Calorimeter Jets Anti- k_T JPT/L2+L3+JPT corrected

SisCone and *IterativeCone*, amongst others, were also studied at different stages of the work.

In any analysis that concerns $t\bar{t}$, either as a signal or as a background, the definition of the jet object is something that has to be understood very carefully. The determination of what we denote as the 0-jet bin in the analysis will define the content of the sample in which the kinematical cuts are applied.

It is necessary to determine a *jet* to make a robust jet veto, that will lead to similar results when using different jet definitions. It is a conservative view that may not be enough to get rid of all the $t\bar{t}$ background, and for that, looser energy thresholds and additional requirements can be studied and will be discussed in the following sections.

As was previously discussed, are accounted for by the Jet Veto. The first requirement will be $|\eta|_{jet} < 2.5$.

The E_T threshold for a jet definition is an important requirement, however, it is very difficult to properly establish. Lowering the energy threshold implies more background rejection, but on the other hand it means large uncertainties in the jet energy scale (JES) [112] and also theoretical uncertainties. In the low P_T region, below 50 GeV/c, JES uncertainty increases rapidly with decreasing P_T , as shown in figure 5.14. The

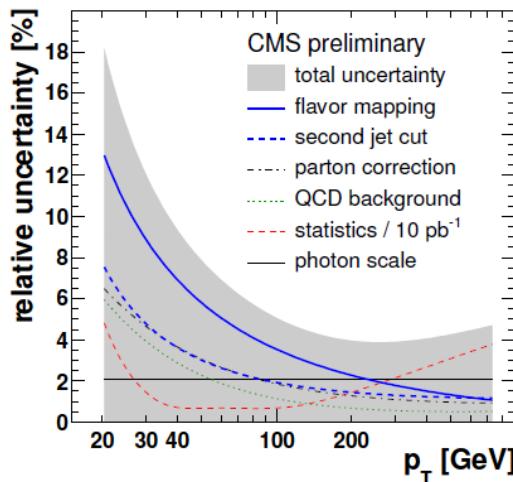


Figure 5.14: - Total relative uncertainty for measured response in photon+jet (one of the main calibration samples).

JES error more than doubles from 7% at 50 GeV to 18% at 20 GeV. For that, an E_T threshold of 30 GeV is chosen.

The E_T resolution for Anti- k_T CaloJets L2+L3 corrected alone and together with the resolution for uncorrected (RAW) jets for $t\bar{t}$ events is shown in figure 5.15. The plots are produced using the latest Jet Energy corrections corresponding to the 7 TeV scenario and using Monte Carlo information. The same information for Particle Flow jets

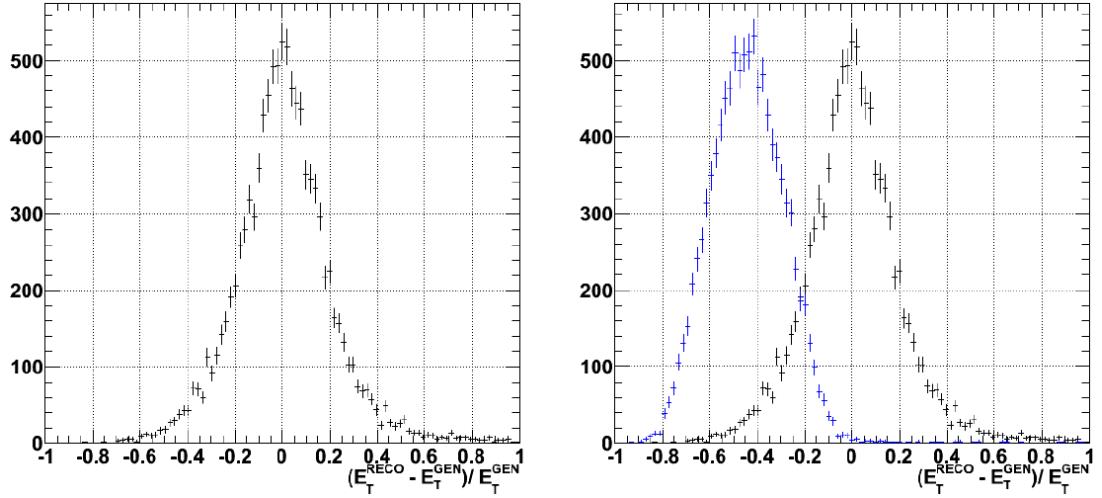


Figure 5.15: - Jet Energy resolution for L2L3 corrected Anti- k_T CaloJets in $t\bar{t}$ events, alone (left) or together with the resolution for RAW jets in blue (right).

is shown in figures 5.16. As expected, for the Particle Flow jets, the energy corrections are not so important as in the case of Calorimeter Jets. One interesting variable that was also studied is the α parameter. A useful variable to complement the CaloJets using track information, α is defined as the sum of the transverse momentum of the tracks divided by the transverse energy of the jet,

$$\alpha = \frac{\sum P_T^{tracks}}{E_T^{jet}} \quad (5.10)$$

It helps to reduce the contamination from fake jets coming from underlying event and pile up. In an ideal detector, $\alpha \approx 0.66$, as two thirds of the jet energy is carried by charged particles. For a fake jet, this value tends to be smaller. The many low P_T particles that populate underlying events, leave energy in the ECAL if they are neutral, while the charged one are curled up in the magnetic field and usually do not reach the calorimeter. The α ratio is smeared and reduced by the detector energy resolution and by inefficiencies in the charged particle reconstruction in the tracker leading to values of $\alpha < 0.4$. A baseline quality criteria for the estimation of α would be the following:

- $|\eta| < 2.4$

5. PHYSICS OBSERVABLES

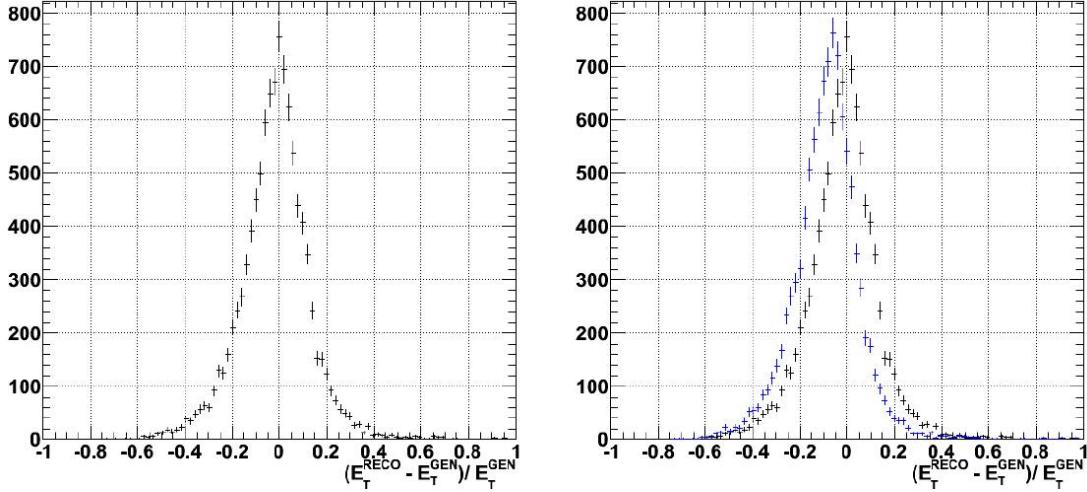


Figure 5.16: - Jet Energy resolution for L2L3 corrected Particle Flow Anti- k_T Jets in $t\bar{t}$ events, alone in the plot of the left or together with the resolution for RAW jets (in blue) in the right.

- $P_T > 2 \text{ GeV}/c$
- $|d0| < 0.3 \text{ cm}$
- $n_{\text{hits}} > 5$
- $\Delta R_{\text{track-jet}} < 0.5$

In figure 5.17, the distribution of α and E_T resolution as a function of the E_T of the L2L3 corrected Anti- k_T Calo Jets is presented. This parameter is not needed when the Jet reconstruction algorithm already uses tracker information, as for Particle Flow or JPT jets.

5.3.6 Jet Veto

The Jet Veto in the analysis is defined following the recommendation of the JetMET group of CMS. If at least one:

- Anti- k_T CaloJet, with a size of 0.5, L2 + L3 corrected
- $|\text{eta}| < 2.5$
- $E_T^{\text{corrected}} > 30 \text{ GeV}$

is found, the event is rejected.

Similar performance was achieved with Calo Jets, when using Iterative Cone, Sis Cone or Anti- k_t . That can be seen in figure 5.18, where the number of central jets reconstructed in signal and $t\bar{t}$ background, for different jet definitions is presented.

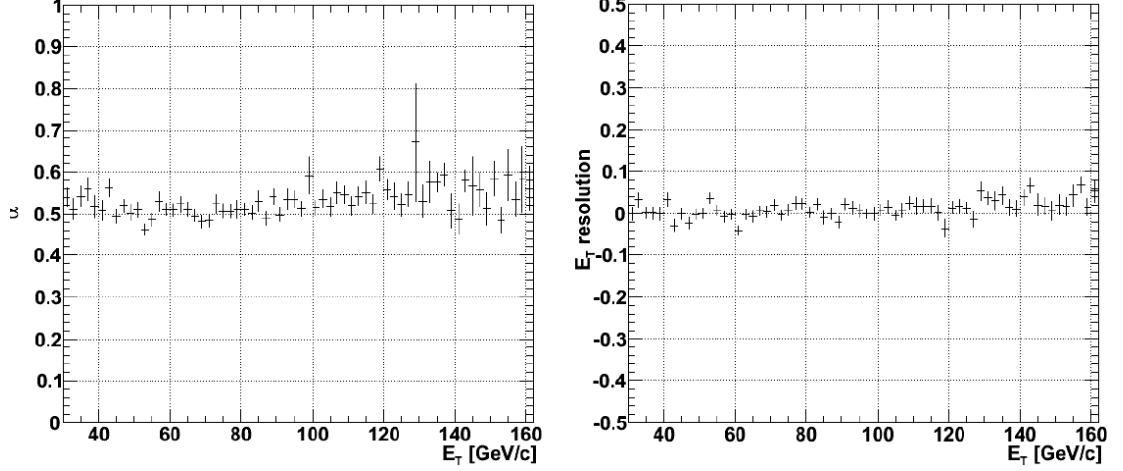


Figure 5.17: - Distributions of α (left) and E_T resolution (right) as a function of the E_T for the L2L3 corrected Anti- k_T Calo Jets in $t\bar{t}$ events.

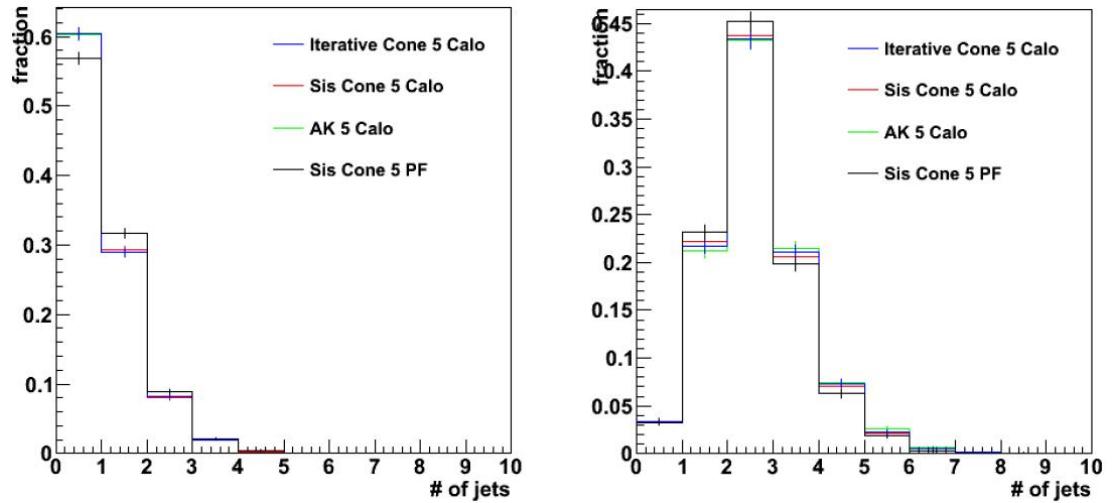


Figure 5.18: - Number of reconstructed jets in the signal (right) for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ and for $t\bar{t}$ (left) for different algorithms: Iterative Cone in blue, Sis Cone in red and Anti- k_T Calo Jets in green, Particle flow SisCone in black.

5. PHYSICS OBSERVABLES

The use of JPT or Particle Flow Jets is contemplated in the analysis. In that case a re-optimization of the jet veto parameters can be performed, taking into account the systematic error on the corresponding JES. Such optimization would probably be in favor of a lower P_T cut that could lead to a better veto sensitivity.

To complement the Central Jet Veto (CJV) based on Calo Jets by using TrackJets or by the application of a Track Counting Veto (TCV) will help to further reduce $t\bar{t}$ events. The distribution of the number of tracks, excluding the muons, is shown for $t\bar{t}$, $Z + jets$ and WW events in figure 5.19. The track quality parameters are the same as in 5.3.5.

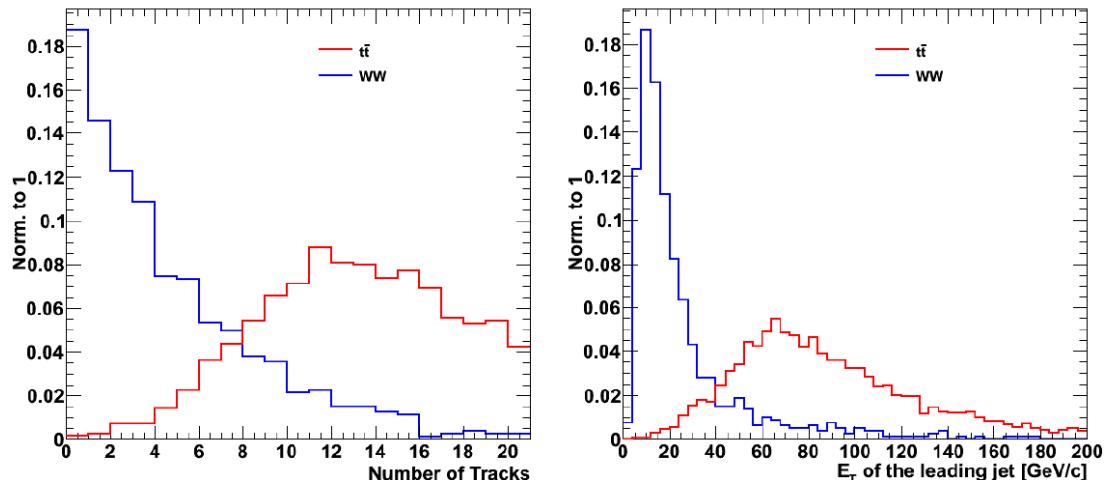


Figure 5.19: - Distribution of the number of tracks and of the E_T of the leading Anti- k_T Calo Jet for $t\bar{t}$ and WW normalized to one.

5.4 missing E_T

The presence of two neutrinos in the final state characterizes the signature of the signal and makes the missing E_T estimation a crucial factor in the analysis. The neutrinos pass undetected through the detector, and their presence is only observed by a momentum imbalance in the transverse plane. The ability to measure this imbalance will determine how well the signal events can be identified [113].

There are several methods to estimate the missing E_T in CMS, the standard calorimeter (CaloMET), the Track-corrected (tcMET), and the Particle Flow (PFMET). In the following the characteristics and main differences of each one will be presented.

5.4.1 Calorimeter missing E_T

The calorimeter missing E_T (CaloMET) is determined from the transverse vector sum over energy deposits in uncorrected projective Calorimeter Towers, in both the ECAL and the HCAL.

The Calorimeter Tower based missing E_T calculation can be improved by correcting it for several effects. For this analysis, the most important concerns the muons. Muons are minimum ionizing particles that deposit very little energy in the calorimeters. As they can significantly contribute to the total missing E_T in the event, their footprint have to be taken into account. This is done by adding their deposits in the calorimeter to the missing E_T sum and then subtracting their transverse momentum, precisely measured in the tracker and in the muon chambers:

$$E'_T = E_T + \sum_{\mu} \vec{E}_{T,calo}^{\mu} - \sum_{\mu} \vec{P}_T^{\mu} \quad (5.11)$$

doing that, the muon momentum measurement from the central tracker and muon system is used to replace the energy measured along the muon trajectory in the calorimeter.

The electron energy scale can be used as well to adjust the missing E_T . Additionally isolated taus lead to jets that differ substantially from average QCD jets. Specific corrections for τ -jets based on particle flow methods can be employed in this case.

There are many other corrections to the calorimeter missing E_T . The Jet Energy Scale corrections, often called Type I corrections, are also very important. These corrections take the measured raw energy values and adjust them for the difference between the raw jet energy and the true jet energy, as defined by the Jet Energy Scale (JES) group.

Corrections for mismeasurements due to uninstrumented regions and for effects due to the soft underlying event, pile-up etc. are the so called Type II corrections, that improve the bias and resolution in several benchmark processes studied, contemplated but not yet developed in CMS.

5.4.2 Track Corrected missing E_T

Using not just the calorimeters, but other CMS subdetectors, fundamental improvements to calorimeter-based missing E_T can be made. That is the case of the Track corrected missing E_T (tcMET) [114], in which tracker info is used.

The Track Corrected missing E_T improves the resolution and tails replacing the expected energy deposition in the calorimeter of all well reconstructed tracks by their measured momentum in the tracker. Since at the momentum scale of the order of 1 GeV the tracker has a superior resolution compared to the calorimeter, this leads to significant improvements. The missing E_T is computed by replacing the expected

5. PHYSICS OBSERVABLES

energy deposition of good tracks (assumed to be pions) with the corresponding momentum. This has to be done taking care to avoid double-counting by including energy from both the calorimeter and the tracker.

Track-corrected missing E_T removes the expected energy of charged hadron deposits in the calorimeter. The expected deposition is determined from a response function derived from a single pion Monte Carlo sample. This is one of the key differences between this approach and Particle Flow; tcMET relies on the expected deposition rather than the actual energy contained in a calorimeter cluster.

5.4.3 Particle Flow missing E_T

A more refined approach consists of using the individual particles that have been reconstructed using the Particle Flow technique [100].

As in the Jet reconstruction, the particle-flow missing E_T reconstruction algorithm is intended to provide a complete and unique event description at the level of individually reconstructed particles, with an optimal combination of information across all CMS sub-detectors. The reconstructed and identified individual particle list includes muons, electrons (with individual Bremsstrahlung photons), photons (either unconverted or converted), charged hadrons (without or with a nuclear interaction in the tracker material), as well as stable and unstable neutral hadrons. These particles are not necessarily isolated, and may even originate from an intricate overlap of reconstructed charged particles, ECAL and HCAL energy clusters, and hits in the muon chambers.

Determining the missing E_T using particle-flow is in principle simple: the transverse vector sum over all reconstructed particles in the event results in an azimuthal, momentum two-vector. In this sense, missing E_T from particle flow is analogous to the one that would have been determined at generator level via all visible Monte Carlo truth particles, except that the particles used to determine particle-flow missing E_T are reconstructed with some efficiency and fake rate, with a finite momentum resolution. Hence, the particle-flow algorithm is fully general and neither specific nor tuned to missing E_T determination. Consequently, it has to be calibrated and verified using a data-driven manner analogous to the calorimeter-based MET determination.

5.4.4 P_T of the di- μ system

Due to the topology of the signal, and for an early data-taking scenario where the muon system is expected to work better than the calorimeters, another estimation of the missing E_T could be the opposite of the transverse momentum of the muon pair.

For signal events this estimation leads to reasonably good results, close to the Calo estimation, but this is not always the case for the backgrounds. The use of this vari-

able as a Calorimeter-independent estimation, accounting for the fact that it is not the optimal, should not be discarded.

5.4.5 missing E_T in the analysis

In a first approach to this analysis, CaloMET corrected by muons was considered. It gives a very good estimation of the missing E_T for signal events and for those background events which have genuine missing E_T . On the other hand, for background events with no genuine missing E_T , like for example $Z \rightarrow \mu\mu$, CaloMET overestimates the missing E_T giving as a result an increase of the background which is not desirable, as it has a cross-section orders of magnitude larger than the signal.

This background can be better controlled when using either Track Corrected missing E_T or Particle Flow missing E_T . A comparison of the missing E_T distribution after muon selection can be seen for signal and the main backgrounds in figures 5.20, 5.21 and 5.22.

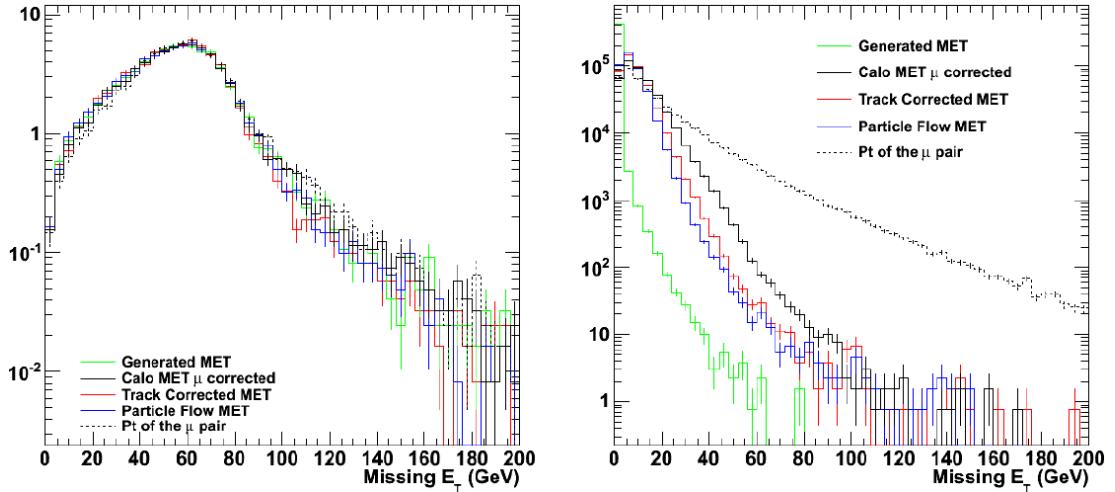


Figure 5.20: - Distributions for different missing E_T estimators for Higgs signal with $m_H = 160 \text{ GeV}/c^2$ (left) and $Z + \text{jets}$ (right)

The choice of the missing E_T in the analysis is not straightforward, both the use of Track Corrected or Particle Flow significantly help in the correct estimation of the missing E_T , and neither one or the other gives a perfect performance. However, the Particle Flow missing E_T gives the best resolution, as can be seen in figures 5.23 and 5.24.

5. PHYSICS OBSERVABLES

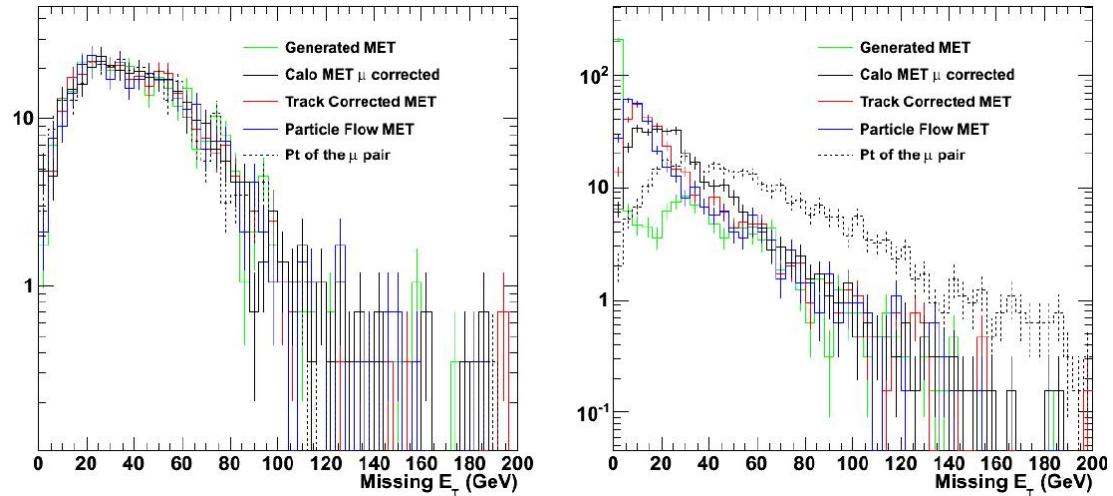


Figure 5.21: - Distributions for different missing E_T estimators for WW (left) and WZ (right)

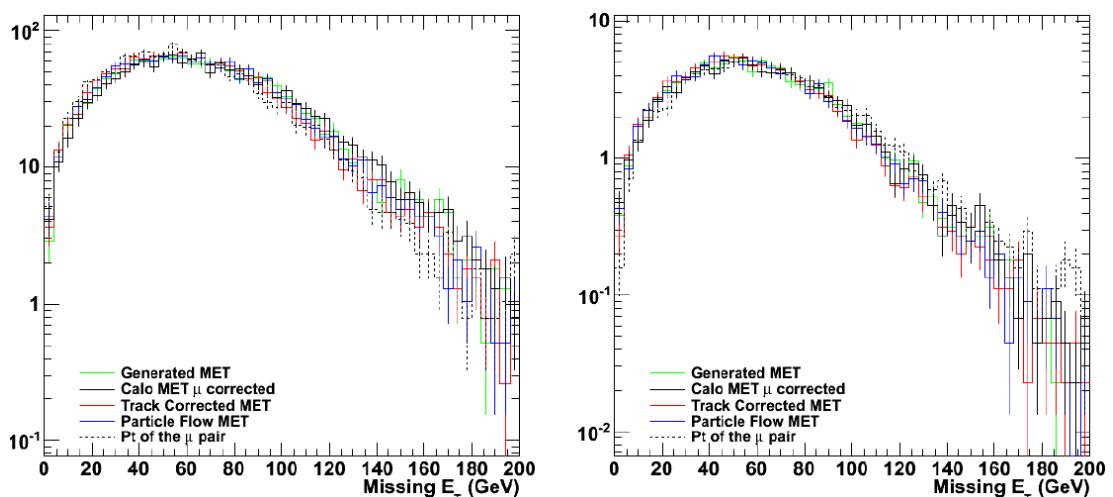


Figure 5.22: - Distributions for different missing E_T estimators for $t\bar{t}$ (left) and tW (right)

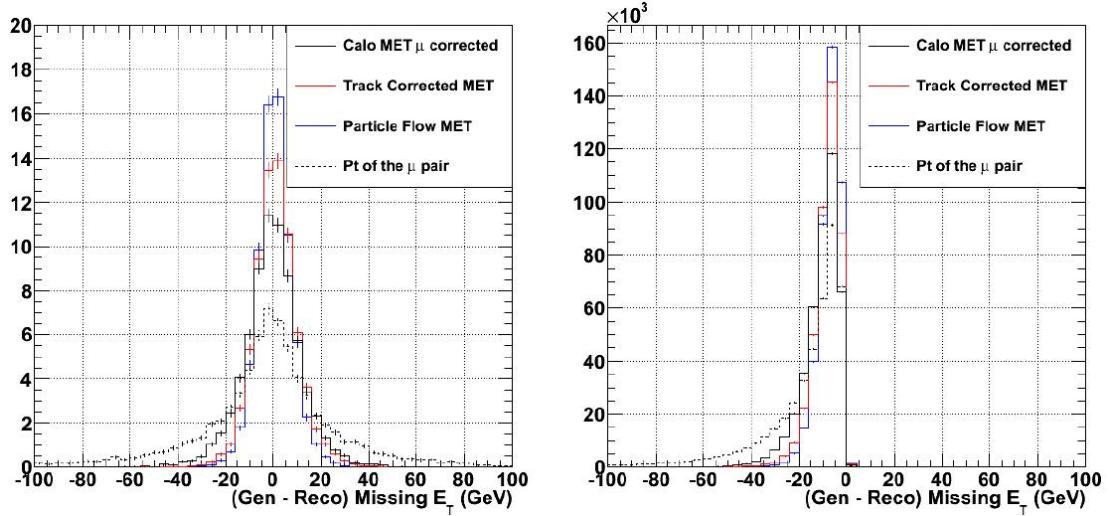


Figure 5.23: - Difference between reconstructed and generated missing E_T for higgs signal with $m_H = 160 \text{ GeV}/c^2$ (left) and $Z + \text{jets}$ (right)

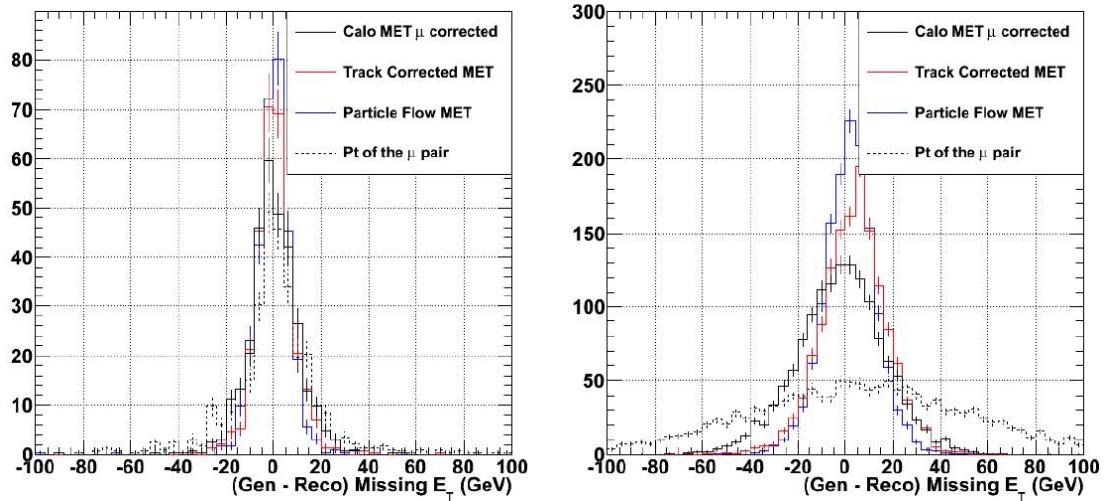


Figure 5.24: - Difference between reconstructed and generated missing E_T for WW (left) and $t\bar{t}$ (right)

5. PHYSICS OBSERVABLES

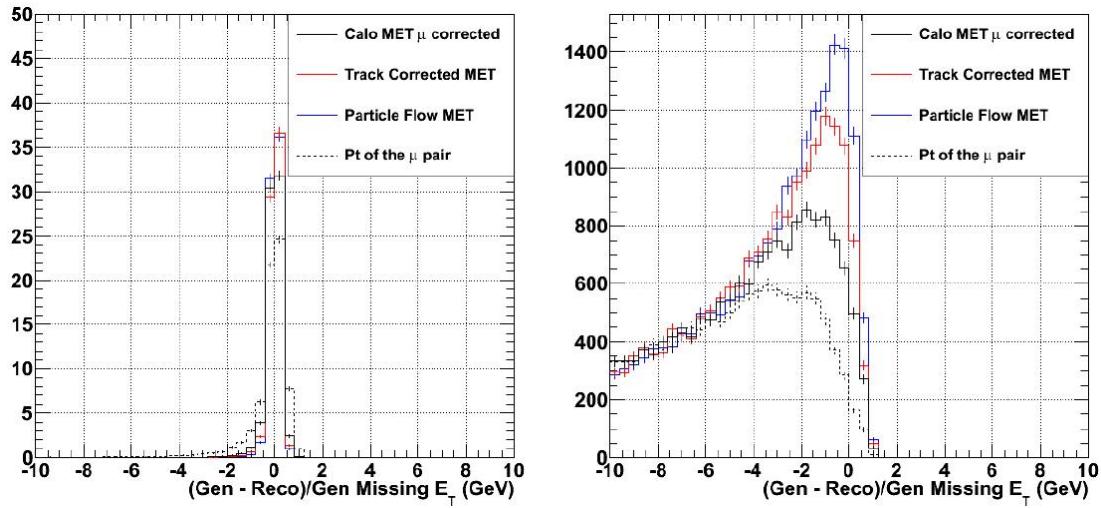


Figure 5.25: - Difference between reconstructed and generated missing E_T divided by the generated for Higgs signal with $m_H = 160 \text{ GeV}/c^2$ (left) and $Z + \text{jets}$ (right)

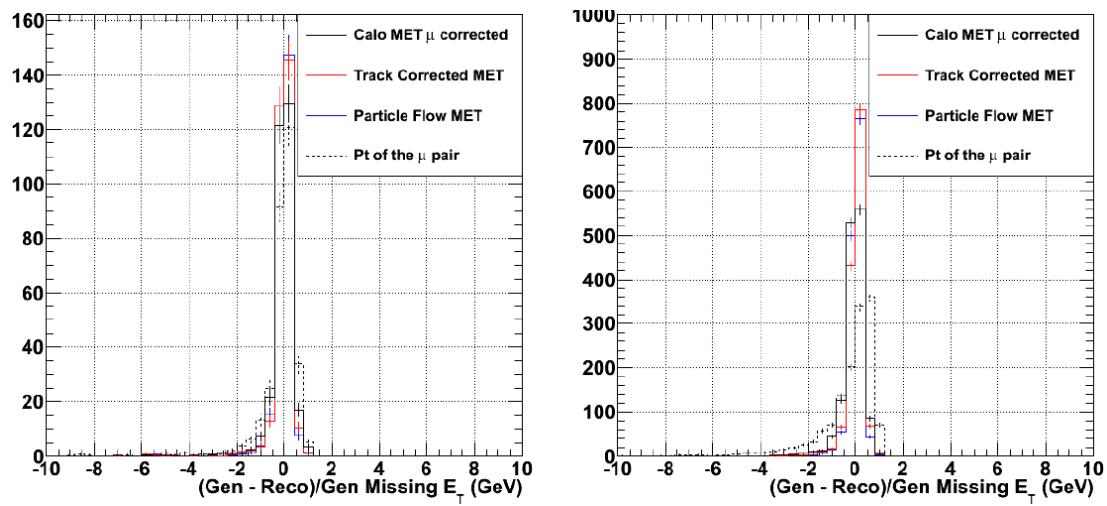


Figure 5.26: - Difference between reconstructed and generated missing E_T divided by the generated for WW (left) and $t\bar{t}$ (right)

Track Corrected missing E_T has been used in the Higgs to WW^* analysis in CMS since it was available, leading to a better control of the $Z + jets$ background. To determine if this is the best choice for the analysis will be only possible with more collected collision data, in the future.

The main concern on the use of missing E_T in the analysis is related to the tails present in the $Z + jets$ background at large values of the missing E_T . To control this background variables based on the angle between the missing E_T and the muons and jets can be used. In the $Z + jets$ events, the missing E_T tends to go in the same direction as one of the muons. The angle between the missing E_T and the closest μ hence has some discrimination power that can be exploited. This variable, as well as the Projected missing E_T , $E_T^{miss} \cdot \sin(\min(\Delta\phi_{E_T^{miss}}, \frac{\pi}{2}))$, is presented in figure 5.27. The

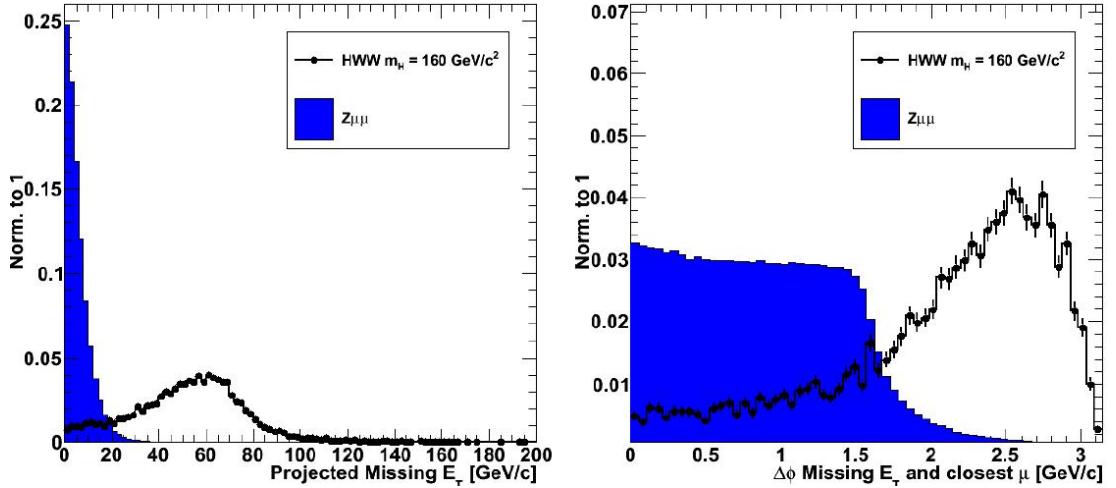


Figure 5.27: - Projected missing E_T (left) and $\Delta\phi$ between missing E_T and the closest muon (right) normalized to unit area for signal and $Z + jets$ background.

different missing E_T projections with respect to the leptons or jets are useful to identify events with no real missing E_T .

The P_T of the muon pair can be seen in figure 5.28. The angle between the missing E_T and the muon pair $\pi - \Delta\Phi_{(MET,pair)}$ is presented in figure 5.29, as well as the projected missing E_T , in figure 5.30.

The analysis is a challenge for the detector as it makes use of all the sub-detectors and requires a good performance of every one of them. From the interaction point outwards, it uses tracker information in many ways: the muon leaves a track in the tracker that is combined with other information to reconstruct it, quality parameters of the muon track help further in the muon identification and selection, tracker info is also essential for the isolation and is also used for the missing E_T estimation and Jet identification and reconstruction. ECAL and HCAL are used mainly for missing E_T

5. PHYSICS OBSERVABLES

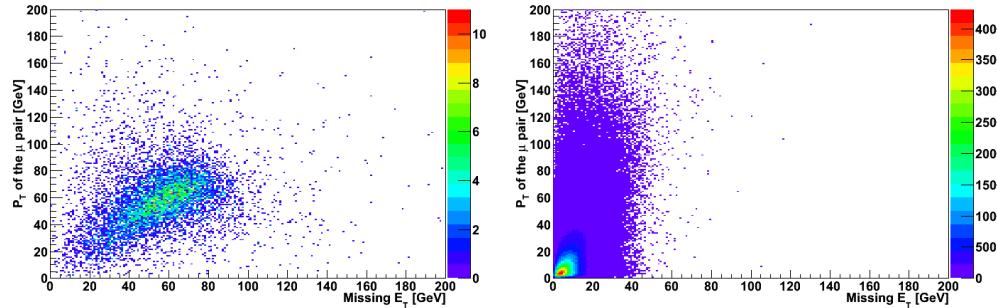


Figure 5.28: - P_T of the muon pair Vs. missing E_T for signal (left) and $Z + jets$ (right)

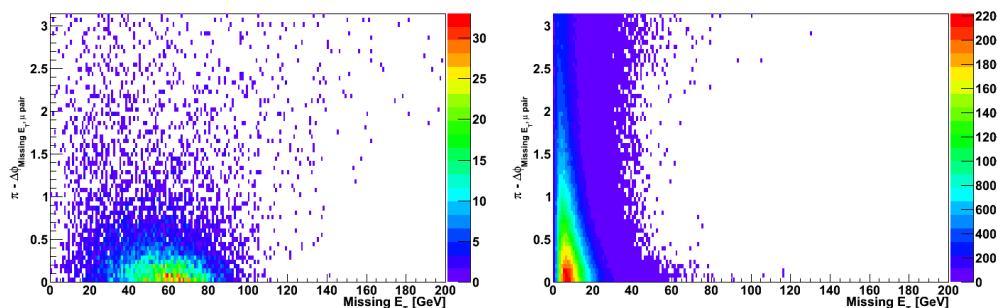


Figure 5.29: - $(\pi - \Delta\Phi_{(MET,pair)})$ Vs. missing E_T for signal (left) and $Z + jets$ (right)

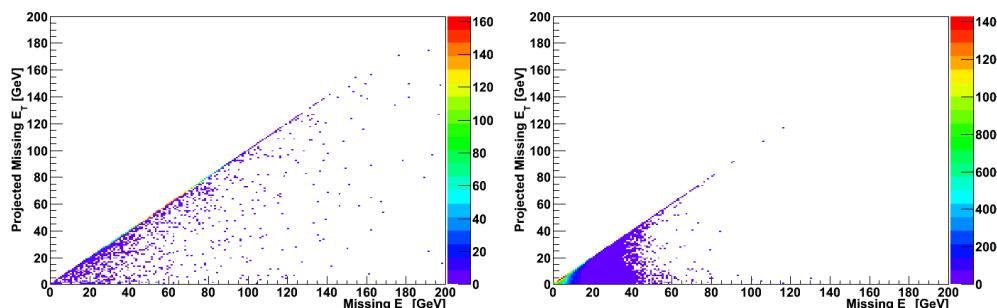


Figure 5.30: - Projected missing E_T Vs. missing E_T for signal (left) and $Z + jets$ (right)

estimation and Jet identification and reconstruction, and also in the muon isolation. Then, the muon system is used in the full muon reconstruction/identification/selection sequence.

The $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ analysis is one of the most, if not the most, promising towards a SM Higgs boson discovery in the LHC, and it is also interesting in terms of knowledge of the detectors. With a good knowlegde of every subdetector and an optimized performance of each part, the analysis could be almost background free.

But in scenarios of early data taking, when the performance of different parts of the detector is not optimal or not perfectly calibrated, the detector can become a challenge for the analysis. The analysis the can be adapted to perform relying on just the muon chambers and the tracker, using the tracker to isolate the muons and to determine the jet content of the event, and using the muons to make a rough estimation of the missing E_T .

The reconstruction and determination of the main physics observables used in the analysis were introduced. For an optimal performance, the whole CMS detector is needed. Eventually, in early scenarios where the detector is not fully commisioned, the analysis could also be performed relying on the muon system and tracker only. In the next chapter, the analysis chain is detailed and the different centre of mass energies and luminosity regimes is explained.

5. PHYSICS OBSERVABLES

6

Event Selection and Kinematics

An analysis for the search of a SM Higgs boson in the CMS experiment at the LHC is performed using the Higgs to WW^* decay channel. The Higgs mass range studied starts around $120\text{ GeV}/c^2$, close to the limit set by LEP2, up to $\sim 200\text{ GeV}/c^2$, covering the theoretically preferred m_H value and masses already studied at the Tevatron. Gluon fusion is the dominant production mode for the Standard Model Higgs in the LHC, and in most of the considered mass range, the decay to WW has a Branching Ratio close to 1. The subsequent decay to two leptons, in particular muons, and two neutrinos gives a clean signature.

The physical observables that have been used to study this channel are determined by the chosen final state: two high P_T isolated Muons and missing E_T to identify the signal, and Jets to reject backgrounds. The choice of muons in the final state is ideal for the CMS experiment, designed for an optimal muon detection. Signal events present a substantial missing E_T due to the presence of neutrinos in the final state, and, as no hard jet activity is expected to be present in the signal, events with Jets have to be identified and vetoed.

The characteristics of the LHC make possible a discovery-oriented analysis based on a single channel counting experiment. It consist of a limited set of kinematical cuts designed to take advantage of the characteristics of the signal to extract it from a vast background. The analysis uses information from every sub-detector of CMS and requires the best calibration and performance. and an efficient reconstruction of the physical objects and offline analysis methods. The details of the event selection will be discussed in the following sections.

6.1 Trigger Selection and Datasets

The Trigger selection is the first step in the analysis flow [115]. CMS has designed a redundant trigger to collect potential signal events for the different physics analyses. Trigger tables are designed in CMS to optimize the trigger performance depending on

6. EVENT SELECTION AND KINEMATICS

the expected luminosity and expected pile-up conditions, as well as on the expected detailed response of all sub-detectors. Depending of the characteristics of the final states considered, the corresponding trigger paths are chosen. The analysis performed in this thesis makes use of Muon trigger paths to selects events with muons in the final state.

The choice of the Muon path or paths used in the analysis had evolved within the different releases, from the combination of the single relaxed and isolated paths used in the 14 TeV study to the relaxed single muons path used in the 10 and 7 TeV studies when a single relaxed path, *HLT_Mu9*, is required (*HLT_Mu15* being also studied). The selected paths with their respective thresholds are presented in table 6.1. The isolation criteria employed in the case of the single isolated muon path is both calorimetric and pixel track based. In the case of the non-isolated triggers, the additional requirement of the two-dimensional Delta-*xy* between the Level 3 muon and the beam spot $< 2\text{ cm}$ is applied. As the signal has two well isolated high P_T muons, the expected

HLT path	Level-1 threshold [GeV/c]	HLT threshold [GeV/c]
Single Isolated μ	7	11
Single Relaxed μ	7	16
<i>HLT_Mu9</i>	7	9

Table 6.1: Level-1 and HLT P_T thresholds for the muon HLT paths used in the analysis.

efficiency for single-muon paths, always assuming a good performance of the trigger, is high, and in that case, as long as the offline selection is tight enough, the overall performance is not expected to change much when using one muon path or other. Using Monte Carlo simulation the trigger efficiency in this channel takes values higher than 90% in all the considered Higgs mass range.

The trigger menus for the luminosities of $8 \cdot 10^{29}$ (STARTUP scenario) and 10^{31} (IDEAL conditions) have been studied in the analysis. The efficiencies for signal of *HLT_Mu9* are presented versus P_T and versus η in figure 6.1 and 6.2 for a couple of Higgs masses, $m_H = 130\text{ GeV}/c^2$ and $m_H = 160\text{ GeV}/c^2$.

The HLT efficiencies for some Higgs signal samples are presented in table 6.2, and the efficiencies for some of the main backgrounds are shown in table 6.3. Within

HLT path	$H \rightarrow WW^*$ $m_H = 120\text{ GeV}/c^2$	$H \rightarrow WW^*$ $m_H = 130\text{ GeV}/c^2$	$H \rightarrow WW^*$ $m_H = 160\text{ GeV}/c^2$	$H \rightarrow WW^*$ $m_H = 200\text{ GeV}/c^2$
<i>HLT_Mu3</i>	99.76	99.62	99.76	99.77
<i>HLT_Mu5</i>	98.96	98.86	98.88	99.07
<i>HLT_Mu9</i>	98.62	98.49	98.49	98.75

Table 6.2: HLT efficiencies of Muon paths for different Higgs signal samples at $\sqrt{s} = 7\text{ TeV}$.

6.1 Trigger Selection and Datasets

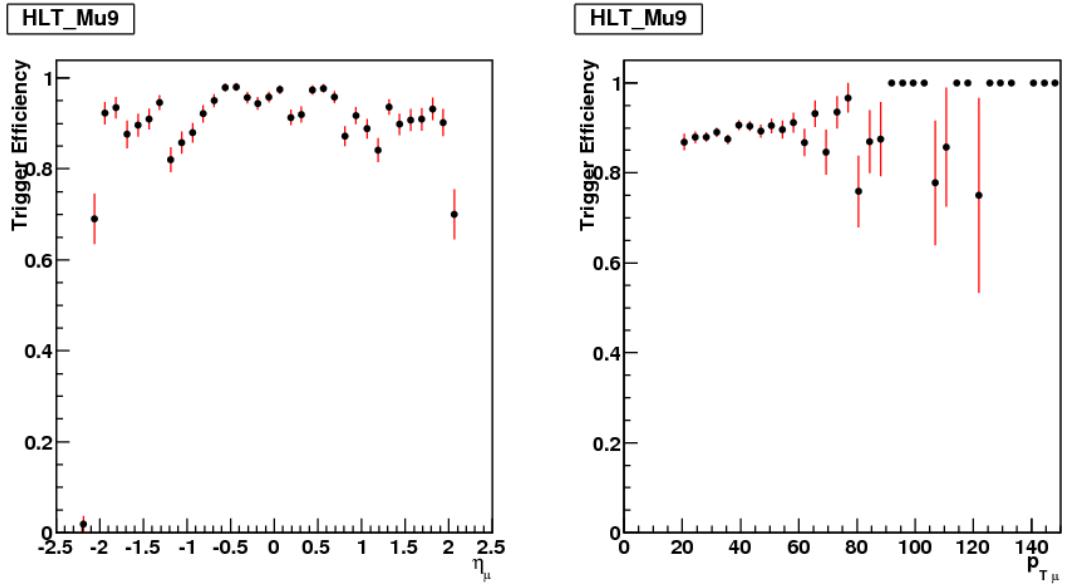


Figure 6.1: - Efficiency as a function of η and P_T of the highest P_T muon reconstructed in $|\eta| < 2.1$ for HLT_Mu9 considering the Higgs mass hypothesis $m_H = 130 \text{ GeV}/c^2$.

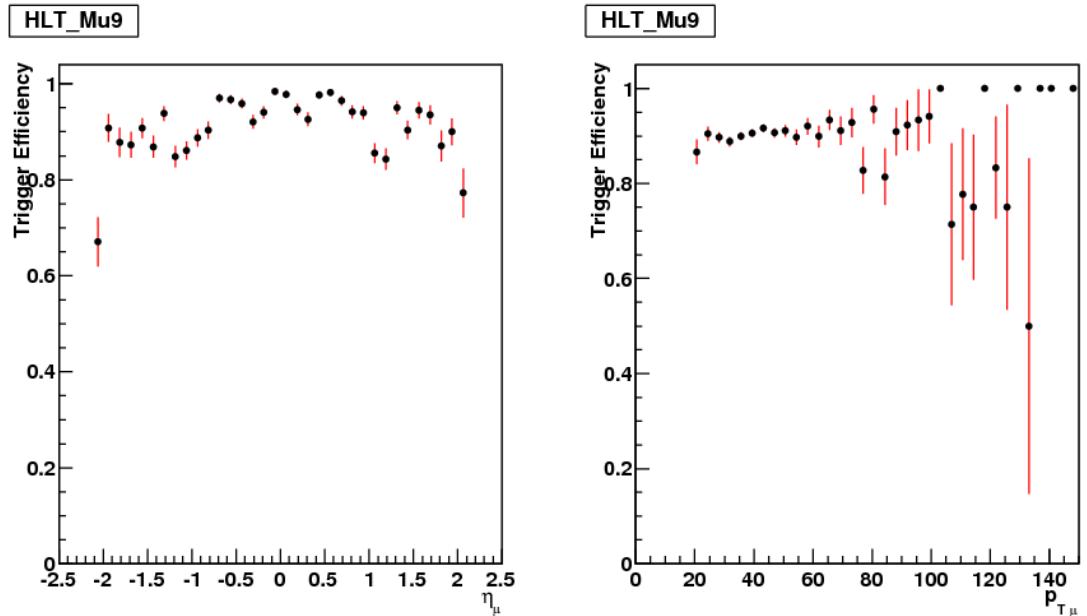


Figure 6.2: - Efficiency as a function of η and P_T of the highest P_T muon reconstructed in $|\eta| < 2.1$ for HLT_Mu9 considering the Higgs mass hypothesis $m_H = 160 \text{ GeV}/c^2$.

6. EVENT SELECTION AND KINEMATICS

HLT path	WW	$t\bar{t} \rightarrow 2l$	$Z + jets$
<i>HLT_Mu3</i>	99.88	99.81	99.76
<i>HLT_Mu5</i>	98.77	99.34	98.73
<i>HLT_Mu9</i>	98.77	99.00	98.34

Table 6.3: HLT efficiencies of Muon paths for some of the backgrounds of the analysis at $\sqrt{s} = 7$ TeV.

fluctuations, the trigger efficiency for the studied paths is expected to be around 95% or higher. However, the estimation of the trigger efficiencies can only be done using data, and more luminosity will be needed in order to determine the real values.

In order to apply the study to real data, the trigger path will determine the dataset to perform the analysis. The collision data is organized in Primary [116] and Secondary Datasets [117] according to trigger selections with minimal overlap. The Primary Datasets are produced at the Tier-0 and re-reconstructed at Tier-1 sites. Then, the Secondary Datasets are extracted from them with more restrictive trigger selections. More details on the datasets used for the physics analysis in CMS can be found in [116]. The analysis presented in this thesis will be performed in data in the Muon (Mu) Secondary Dataset. This secondary dataset uses a logical OR of several HLT muon paths, in which *HLT_Mu9* is included as well as alternatives like *HLT_Mu3* or *HLT_Mu5*.

6.1.1 Skim and Pre-selection

With the objective of set upsetting up the initial sample in which to perform the event selection, additional selection cuts can be applied in order to reduce background without affecting the signal. A baseline skim and pre-selection have been designed for all the final states in the $H \rightarrow WW^*$ decay channel in high luminosity data-taking scenarios. It is intended to provide background reduction together with the highest efficiency possible for the signal under consideration. Events are retained if they pass an OR of the main single muon HLT paths, and if they contain at least two reconstructed leptons with $P_T > 10$ and $P_T > 20$ GeV/c .

A kinematic pre-selection can be applied to further reduce data. Events are required to have a minimal missing E_T , and a loose cut is proposed to clean $b\bar{b}$ -resonances: that is, the invariant mass of the system of the two leptons has to be higher than 12 GeV/c and the missing E_T higher than 20 GeV/c . For the current 7 TeV data-taking, the application of a skim or pre-selection is still not completely defined. It will be necessary to collect more data to design the most efficient way to select the events for the analysis. In any case, the application of a skim or pre-selection will only determine the size of the initial sample in which the analysis will be performed and will not have an effect in the final selection.

6.2 Muon Selection

Once the initial sample in which the analysis will be performed is defined, the muon selection is applied to select events with two muons in the final state. The muon selection criteria proposed in this thesis was discussed in 5.1.4. It was designed using Monte Carlo and applying the knowledge acquired in the study of months of data recorded from cosmic-rays, to efficiently select prompt muons coming from W decays. The first studies of muons with collision data in CMS show that this muon identification performs well in real data, with a good agreement with the Monte Carlo, and some examples will be shown in chapter C.

To extract the signal, events with two high P_T isolated muons are selected. Muons have to be identified and selected as described in section 5.1.4, and isolated as described in section 5.1.3, using tracker and calorimetric information. Signal muons will come from the decay of oppositely charged W bosons, originated from a neutrally-charged Higgs boson, therefore, the muons are also required to have opposite charge. If more than two of such muons are found, the event is rejected. That helps to reduce background coming from WZ events, in which the Z and W bosons decay into muons, and from $t\bar{t}$ and ZZ events, prone to multi-leptonic final states.

The contamination of the $W + jets$ background, that has a large cross-section, is determined by the performance of the muon selection. $W + jets$ events that go through further steps of the analysis present 'fake' muons that do not come from the prompt W decay and can be misidentified as signal. Data-driven techniques will be proposed in the next chapter to control the muon fake rates, that have been already tested in real data, as it will be shown in chapter C.

6.3 Central Jet Veto

The application of the Central Jet Veto (CJV) is the next step after events with two muons are selected. The jet veto was already proposed and discussed with detail in section . The jet veto helps to get rid of an important part of the background and defines the sample in which the cuts in the kinematical variables of the event are applied. It is mandatory to get rid of the background coming from $t\bar{t}$ events, where b -jets are present. The ability of identification of such jets in CMS will determine how the selection will be affected by this background.

The distribution of the number of jets for Higgs signal and $t\bar{t}$ background, normalized to unit area, is presented in figure 6.3. The Jet Veto will reject background coming from all the background sources that present hadronic activity such as $W + jets$, $Z + jets$ or tW .

6. EVENT SELECTION AND KINEMATICS

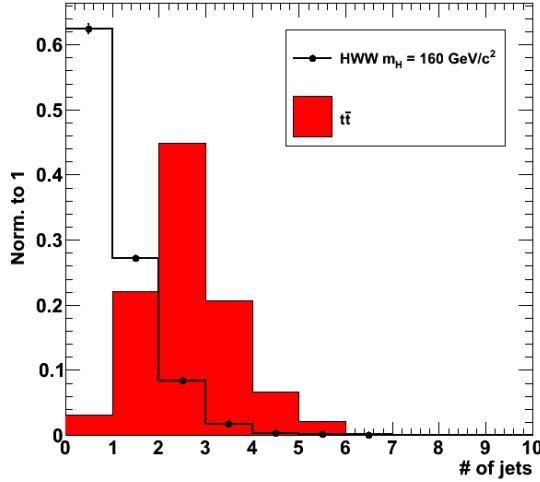


Figure 6.3: - Distribution of the number of jets in the event for Higgs to WW^* signal, for a Higgs mass of $160 \text{ GeV}/c^2$, and $t\bar{t}$ background.

6.4 Definition of the Variables

After the Central Jet Veto, the analysis applies a sequential set of kinematical cuts designed to extract signal events reducing as much as possible the background contribution. For that, a limited group of variables are chosen by means of their discriminating power against the different sources of background.

The variables used in the analysis are presented here in the same order in which the cuts are applied in the analysis:

- **missing E_T :** genuine missing E_T is expected for the signal due to the two neutrinos in the final state. A lower cut in the missing E_T will not affect much the signal while reducing significantly the background coming from sources with no real missing E_T , especially $Z \rightarrow \mu\mu$. The missing E_T estimation, as discussed in 5.4 is affected by many factors, and is very sensitive to the performance of the detector. The distribution of the missing E_T for Higgs signal and $Z + jets$ background, normalized to unit area, is presented in figure 6.4.
- **Angle between the muons in the transverse plane to the beam, $\Delta\Phi_{\mu\mu}$:** $\Delta\Phi_{\mu\mu}$ is the main variable against the potentially irreducible WW background. As previously discussed in 4.3.1, muons coming from signal tend to be emitted close together in the transverse plane due to spin correlations, while WW events tend to be emitted back-to-back, that is, with a large $\Delta\Phi_{\mu\mu}$ angle. This is the most characteristic feature of the signal, and plays a major role in the discovery potential of the $H \rightarrow WW^*$ analysis. The distribution of $\Delta\Phi_{\mu\mu}$ for Higgs signal and WW background, normalized to unit area, is presented in figure 6.5.

6.4 Definition of the Variables

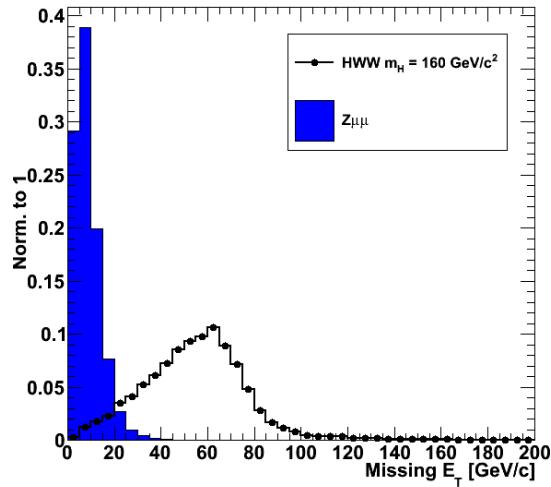


Figure 6.4: - Distribution of the missing E_T for signal, for a Higgs mass of $160 \text{ GeV}/c^2$, and $Z + jets$ background.

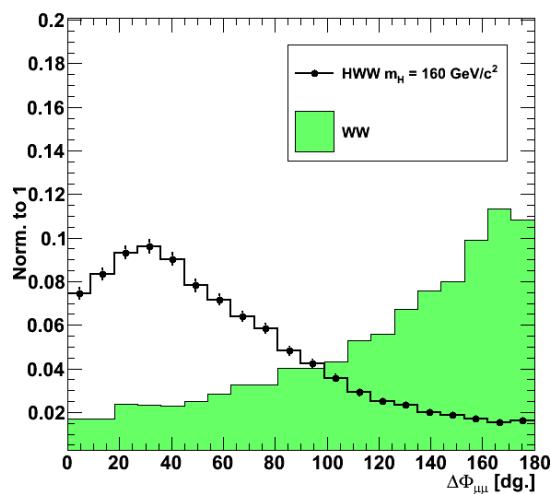


Figure 6.5: - Distribution of $\Delta\Phi_{\mu\mu}$ for signal, for a Higgs mass of $160 \text{ GeV}/c^2$, and WW background.

6. EVENT SELECTION AND KINEMATICS

- **Invariant mass of the muon pair, $m_{\mu\mu}$:** An upper cut is applied in the invariant mass of the muon pair to reduce the contamination of muons coming from Z decays. The distribution of this variable for signal is expected to peak around 40 GeV and if the cut is tight, the Z mass peak ($m_Z = 91.1876 \text{ GeV}/c^2$) can be removed. Other backgrounds with high $m_{\mu\mu}$ values are reduced as well by the application of this cut. The distribution of $m_{\mu\mu}$ for Higgs signal and WW background, normalized to unit area, is presented in figure 6.6.

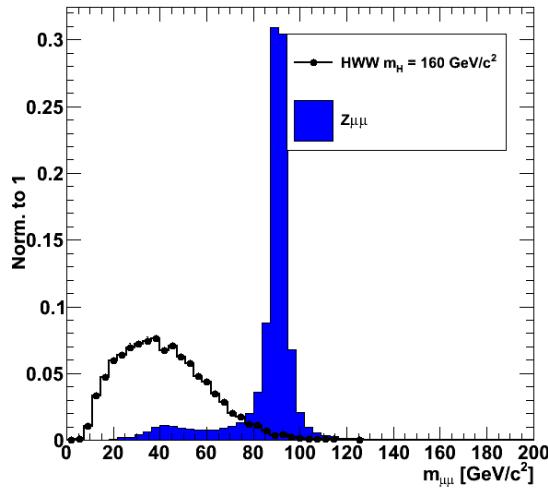


Figure 6.6: - Distribution of $m_{\mu\mu}$ for signal, for a Higgs mass of $160 \text{ GeV}/c^2$, and $Z + jets$ background.

- **Transverse momentum of the muons, P_T^{min} and P_T^{max} :** As was stated in 4.3.2, the Higgs boson mass hypothesis considered determines the P_T of the muons coming from its decays. This is used to further reduce background. The distribution of P_T^{min} and P_T^{max} for Higgs signal and ZZ and WZ background, normalized to unit area, are presented in figure 6.7.

Additional variables can be studied and have been used in more sophisticated approaches, like the multivariate analysis based on neural networks or boosted decision trees that will be discussed later. Some of these variable are:

- Angle between the muons, $\Delta\eta_{\mu\mu}$, that behaves similarly to $\Delta\Phi_{\mu\mu}$.
- Transverse mass of both muon-missing E_T pairs, that helps against non- W background.
- Pseudorapidity η of the muons, because muons from signal are more central.
- Angle between the closest muon and the missing E_T in the transverse plane, that discriminates between events with real missing E_T like the signal and those without.

6.4 Definition of the Variables

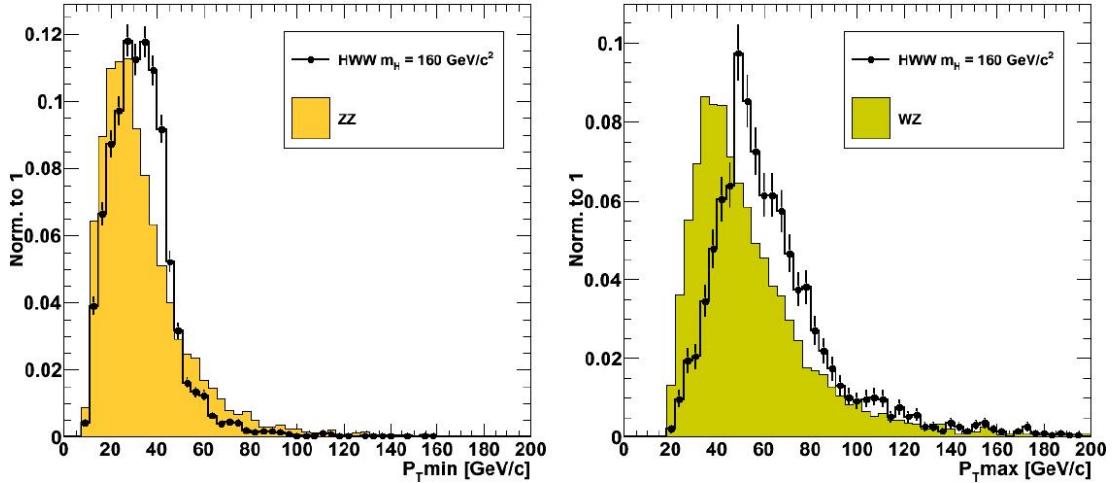


Figure 6.7: - Left: Distribution of P_T min for signal, for a Higgs mass of $160 \text{ GeV}/c^2$, and ZZ background. Right: Distribution of P_T max distribution for signal and WZ background.

- Projected missing E_T , used for the same purposes.

The number of jets present in the event affect the distributions of the kinematic variables. The distributions of the variables before the application of the jet veto for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ and the main backgrounds, assuming 1 fb^{-1} of integrated luminosity at a centre of mass energy of 7 TeV are presented in figures 6.8, 6.9 and 6.10.

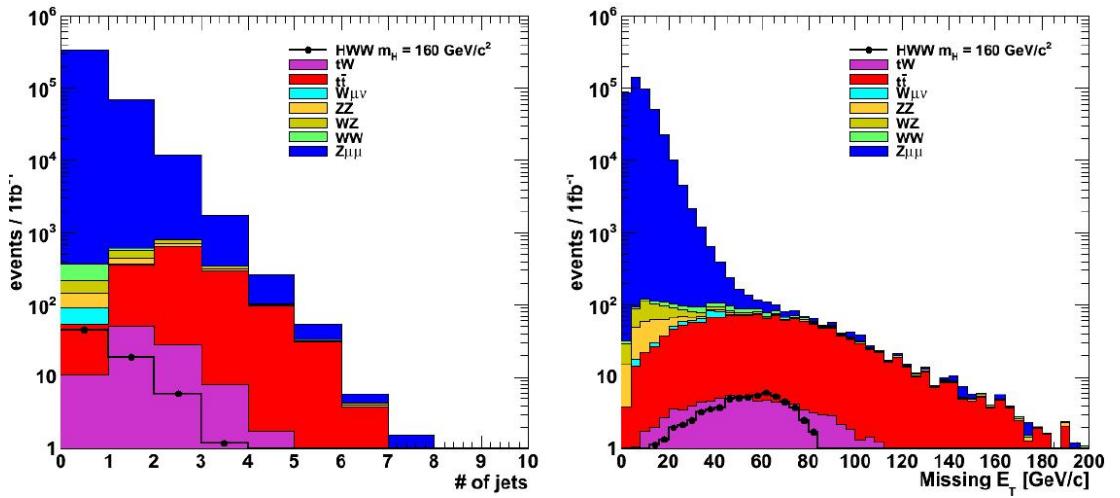


Figure 6.8: - Distribution of the number of jets in the event (left), and missing E_T (right), before the jet veto, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

The distributions of the kinematical variables after the application of the jet veto are

6. EVENT SELECTION AND KINEMATICS

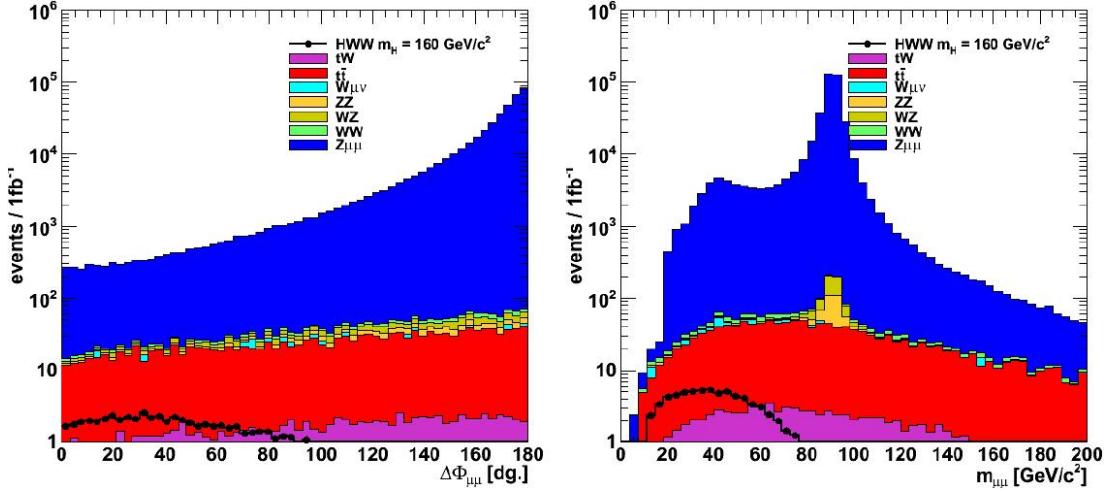


Figure 6.9: - $\Delta\Phi_{\mu\mu}$ (left) and Invariant mass of the muon pair, $m_{\mu\mu}$ (right), before the jet veto, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

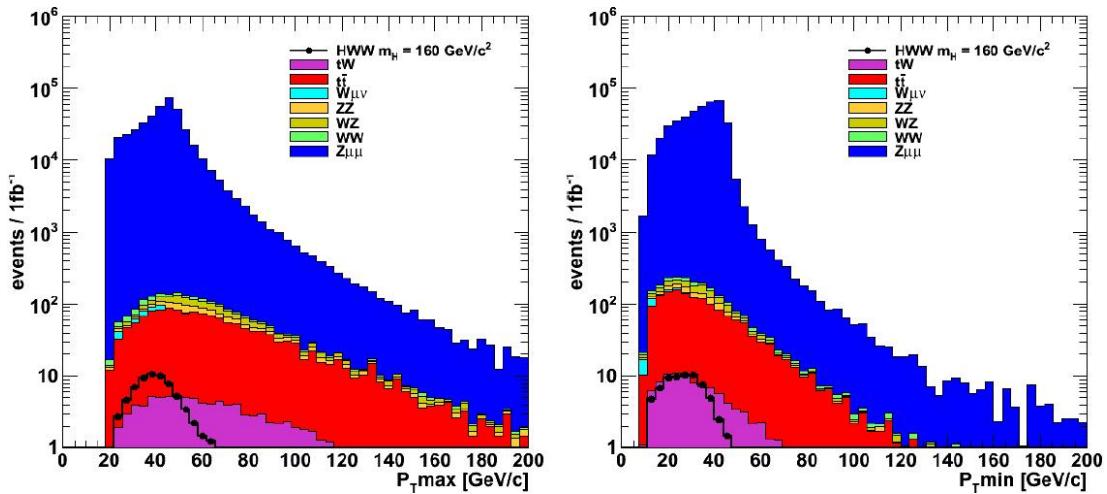


Figure 6.10: - $P_T(\text{highest})$ (left) and $P_T(\text{lowest})$ (right), before the jet veto, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

6.4 Definition of the Variables

presented in figures 6.11, 6.12 and 6.13. This is the phase space in which the analysis is performed.

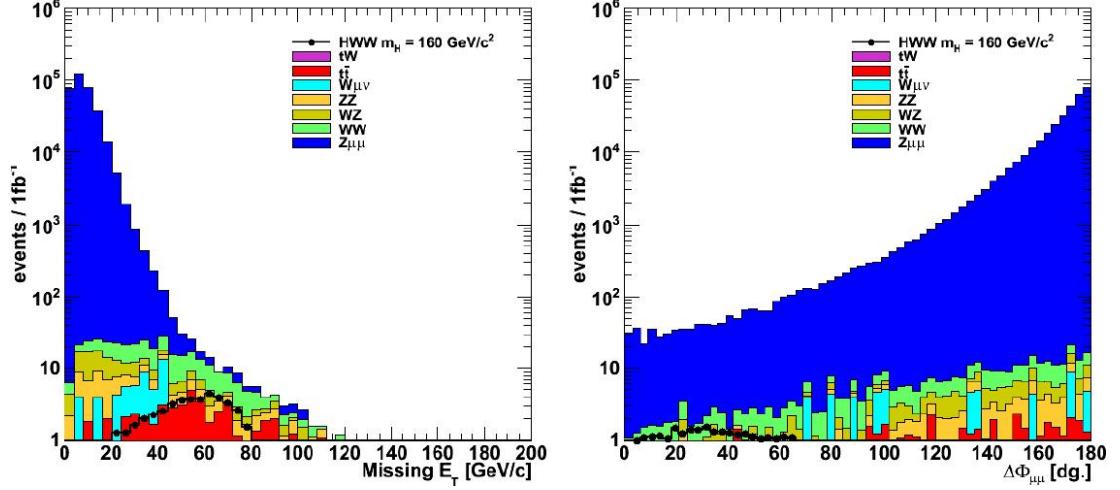


Figure 6.11: - missing E_T (left) and $\Delta\Phi_{\mu\mu}$ (right) after the jet veto, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

The distributions of the kinematical variables in events with 1 jet are presented in figures 6.14, 6.15 and 6.16.

And finally, the distributions for the kinematical variables used in the analysis in events with 2 or more jets are presented in figures 6.17, 6.18 and 6.19. The study of events with one jet can be proposed but needs more luminosity and a different optimization.

6. EVENT SELECTION AND KINEMATICS

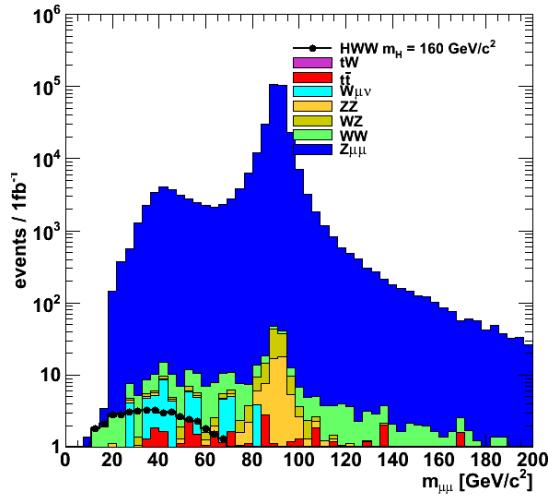


Figure 6.12: - Invariant mass of the muon pair, $m_{\mu\mu}$, after the jet veto, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

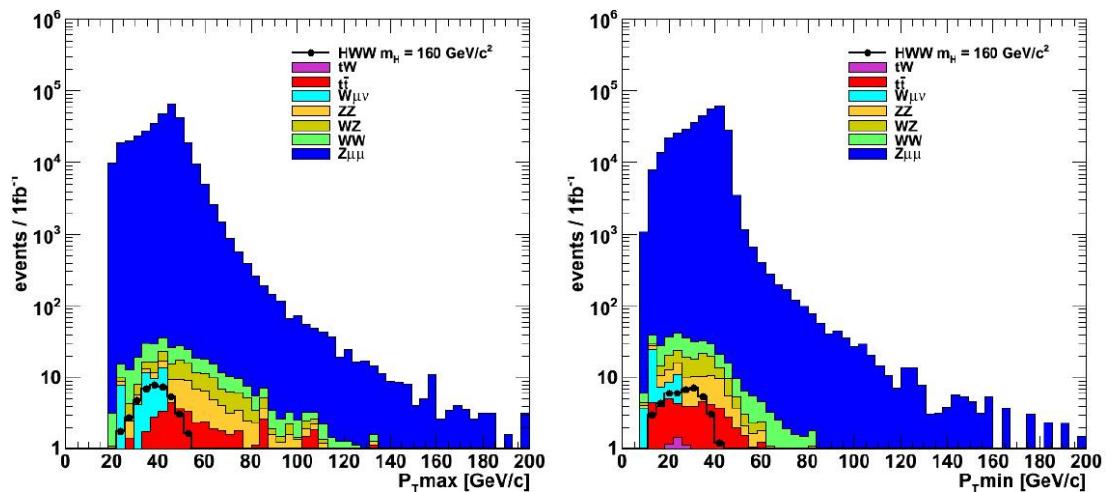


Figure 6.13: - $P_T(\text{highest})$ (left) and $P_T(\text{lowest})$ (right), after the jet veto, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

6.4 Definition of the Variables

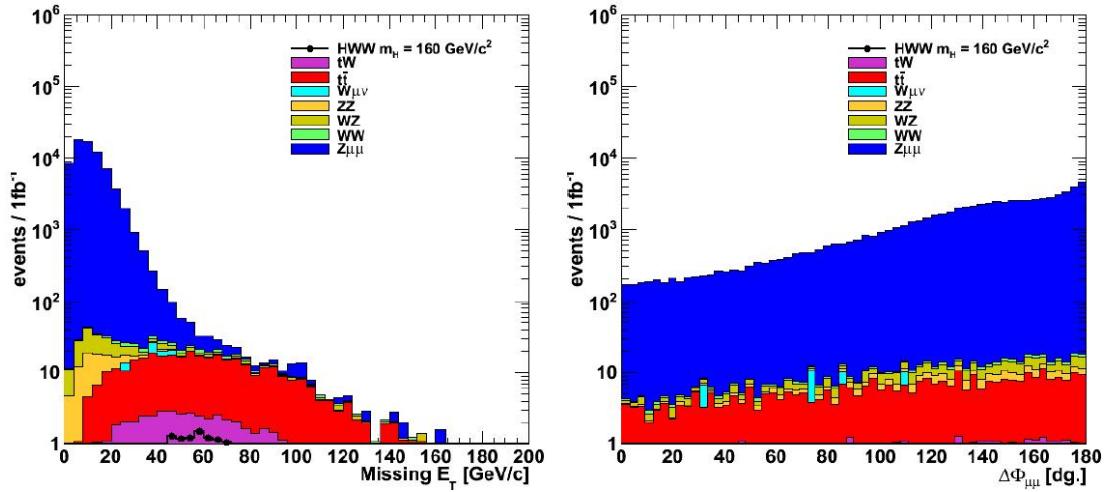


Figure 6.14: - missing E_T (left) and $\Delta\Phi_{\mu\mu}$ (right), asking for 1 jet in the event, $1 fb^{-1}$ of integrated luminosity at 7 TeV assumed.

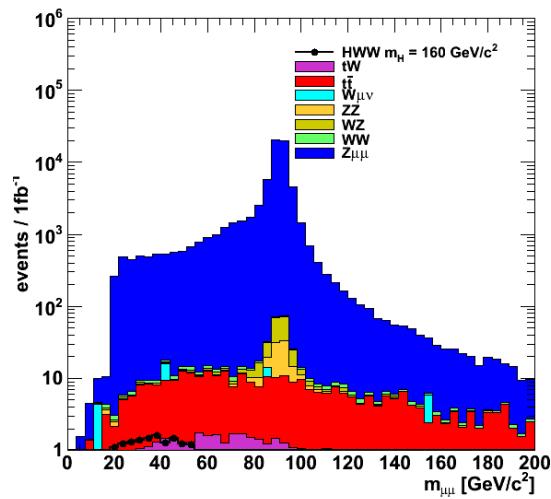


Figure 6.15: - Invariant mass of the muon pair, $m_{\mu\mu}$, asking for 1 jet in the event, $1 fb^{-1}$ of integrated luminosity at 7 TeV assumed.

6. EVENT SELECTION AND KINEMATICS

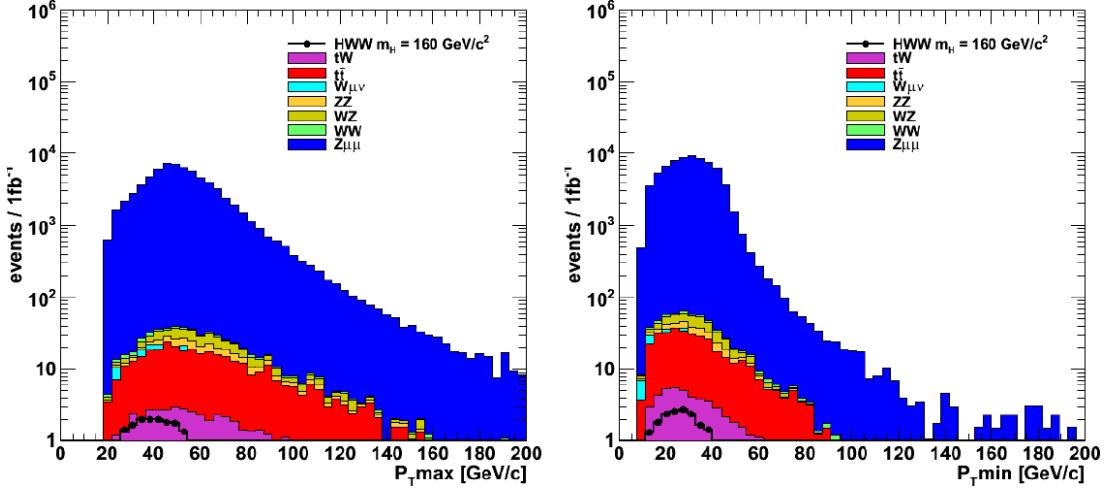


Figure 6.16: - $P_T(\text{highest})$ (left) and $P_T(\text{lowest})$ (right), asking for 1 jet in the event, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

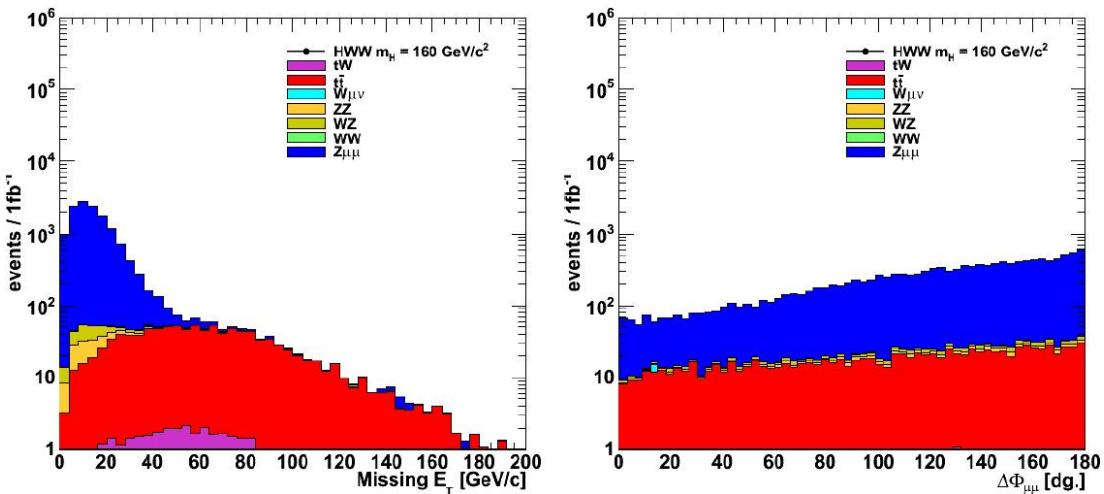


Figure 6.17: - missing E_T (left) and $\Delta\Phi_{\mu\mu}$ (right), asking for more than 1 jet in the event, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

6.4 Definition of the Variables

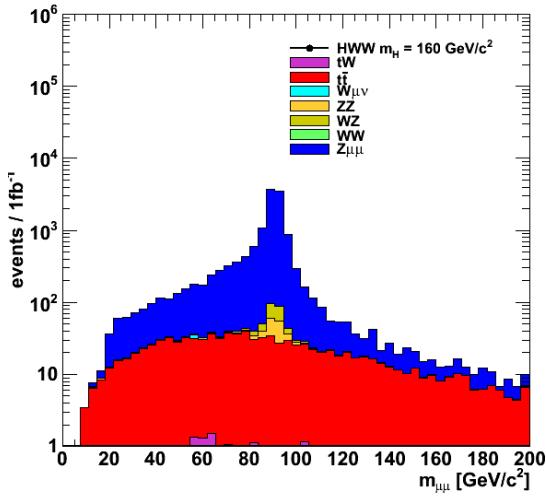


Figure 6.18: - Invariant mass of the muon pair, $m_{\mu\mu}$, asking for more than 1 jet in the event, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

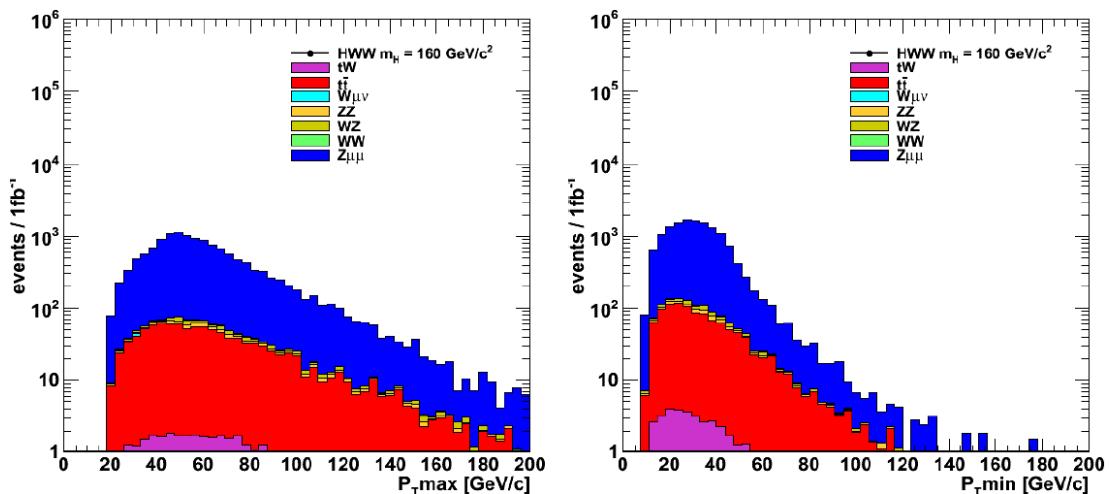


Figure 6.19: - $P_T(\text{highest})$ (left) and $P_T(\text{lowest})$ (right), asking for more than 1 jet in the event, 1 fb^{-1} of integrated luminosity at 7 TeV assumed.

6. EVENT SELECTION AND KINEMATICS

6.5 Sequential cuts

The analysis consists of the combination of sequential cuts applied after jet veto in order to separate the signal from the background events using the variables that have been presented in the previous section, missing E_T , $\Delta\phi_{\mu\mu}$, $m_{\mu\mu}$, P_T^{max} and P_T^{min} .

The three studied leptonic final states of the $H \rightarrow WW^*$ analysis, ee , $e\mu$ and $\mu\mu$, present differences, not only in the experimental identification of muons and electrons, but also because some of the backgrounds do not affect them in the same way. For example, $Z + jets$, that in the case of ee and $\mu\mu$ is a very important background, is considered negligible in the $e\mu$ channel. The same variables are then optimized in a different way and each final state has a different set of cuts.

For the $H \rightarrow WW^* \rightarrow \mu\mu$ analysis a first set of cuts is proposed in [118]. An optimization is done for $m_{\mu\mu}$, P_T^{max} and P_T^{min} to find the combination that gives the best significance, using as input the number of signal and background events, and treating properly the statistical uncertainties and the theoretical systematics errors in the background. The optimization is done using the signal dataset with $m_H = 165 \text{ GeV}/c^2$, that has the higher cross-section, and using all the background contributions, properly normalized considering their production cross-sections for $1-2 \text{ fb}^{-1}$ of integrated luminosity at a centre of mass energy of 14 TeV . The considered systematic error is set around 10-15 %. The cuts in missing E_T and $\Delta\phi_{\mu\mu}$ are chosen before the optimization. That gives as a result the following initial collection of cuts:

Variable	Cut
missing E_T	$> 50 \text{ GeV}$
$\Delta\phi_{\mu\mu}$	$< 0.8 \text{ rad.}$
$m_{\mu\mu}$	$> 50 \text{ GeV}$
P_T^{max}	between 35 and 55 GeV/c
P_T^{min}	$> 25 \text{ GeV}/c$

6.5.1 Mass dependent Optimization

However, to cover the mass range going from 120 to $200 \text{ GeV}/c^2$ and even higher with just a set of cuts is found to not be completely optimal. In particular, the cut on the P_T of the muons has to be adapted to be sensitive to the signal also in the mass region around $120 \text{ GeV}/c^2$. The different kinematics previously discussed in 4.3.2 below the W mass threshold and above, have to be addressed and a second optimization is performed.

In the $H \rightarrow WW^* \rightarrow \mu\nu\mu\nu$, three Higgs mass regions were considered for a mass dependent optimization of the analysis, with a different set of cuts proposed for each one. The set of cuts for the low mass region, to be applied in masses around $130 \text{ GeV}/c^2$, was optimized using the signal dataset with $m_H = 130 \text{ GeV}/c^2$; for the intermediate mass region, the signal sample used was the corresponding to $m_H = 160$

6.6 Final Event Selection

	missing E_T (GeV) >	$\Delta\Phi_{\mu\mu}$ (deg.) <	P_T^{max} (GeV)	P_T^{min} (GeV/c) >
$m_H = 130 \text{ } GeV/c^2$	44	69	(20, 48)	10
$m_H = 160 \text{ } GeV/c^2$	48	57	(28, 50)	25
$m_H = 200 \text{ } GeV/c^2$	55	57	(36, 70)	25

Table 6.4: Summary of the kinematical cuts applied for the event selection in the analysis

GeV/c^2 for masses around $160 \text{ } GeV/c^2$ and finally for the high mass region, for masses around $190 \text{ } GeV/c^2$ and higher, the signal dataset with $m_H = 190 \text{ } GeV/c^2$ was used. All the variables were studied again and the different methods to estimate them optimized accordingly. Given the considered Higgs mass (130, 160 and $200 \text{ } GeV/c^2$) in each case, the optimization was done by maximizing the statistical significance for each cut, defined like:

$$n_\sigma(\vec{sel}) = \frac{N_S}{\sqrt{N_B + (\Delta N_B)^2}} \quad (6.1)$$

where \vec{sel} denotes a certain set of requirements, and N_S and N_B the number of signal and background events normalized to their respective cross-sections assuming $1 \text{ } fb^{-1}$ of luminosity. The considered error for background was a conservative value of 35%, leading to an optimization function of:

$$n_\sigma(\vec{sel}) = \frac{N_S}{\sqrt{N_B + (0.35 \times N_B)^2}} \quad (6.2)$$

The cuts applied in this thesis, which are presented in table 6.4, are the result of an optimization done using MINUIT [119] and the Monte Carlo 14 TeV samples.

To check the robustness of the optimization, these cuts have been compared with the ones that can be derived using a multivariate *BDT* approach, see section 6.9.2. For that, a *BDT* was built using the same variables, and trained against all backgrounds using the three different Higgs signal samples. The results were in a good agreement. However, the optimization performed here is theoretical and done in Monte Carlo. To perform the analysis in data a re-optimization might be necessary, as the proper estimation of some of the backgrounds depends of data-driven methods that can not be applied in Monte Carlo.

6.6 Final Event Selection

6.6.1 Study at a centre of mass energy of 14 TeV

In the original plan for the first physics run of the LHC, the expected centre of mass energy was $\sqrt{s} = 14 \text{ } TeV$. An analysis for the $H \rightarrow WW^* \rightarrow \mu\nu\mu\nu$ was proposed for this scenario considering an integrated luminosity of the order of $1 \text{ } fb^{-1}$ [61; 62].

6. EVENT SELECTION AND KINEMATICS

The muon identification and isolation were not optimized. In the muon selection step of the analysis, events with:

- Two reconstructed Global muons,
- $P_T > 20, 10 \text{ GeV}/c$
- $|\eta| < 2.4$
- Isolated in the tracker and in the calorimeters

were selected. The isolation in this case was absolute. The sum of the P_T of the tracks around each muon was required to be less than $2 \text{ GeV}/c$ and the sum of the transverse energy deposited in the calorimeters less than $5 \text{ GeV}/c$.

To define the initial sample, a loose pre-selection was applied before the kinematical cuts with a similar criteria to the one proposed in 6.1.1. Calorimeter missing E_T corrected for muons was used in the analysis and for the jet veto, uncorrected (RAW) CaloJets were considered. If at least one Iterative Cone Calo Jet with $\Delta R 0.5$ was found in $|\eta| < 2.5$, with an uncorrected energy higher than 15 GeV , the event was rejected. After the central jet veto the sequential cuts on the main variables of the analysis were applied depending on the Higgs mass hypothesis considered in each case to the signal and backgrounds properly normalized to their correspondent cross-sections.

The Monte Carlo samples used to perform this analysis were simulated considering a centre of mass energy of 14 TeV and ideal conditions. Additionally, some background samples were produced with different miss-alignment and miss-calibration scenarios 4.4 and were used to calculate the systematics.

The expected number of events and the relative efficiencies after each cut are presented for signal and backgrounds for each Higgs mass hypothesis studied in the appendix A.1. The results for $m_H = 130 \text{ GeV}/c^2$ and $m_H = 160 \text{ GeV}/c^2$ are presented in tables 6.5 and 6.6.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. $\&& \mu\mu$	48.9	459	226	184	461	5128	177	40447
Jet Veto	23.2(47%)	309 (67%)	58(25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
MET	16.1(69%)	200 (64%)	32 (55%)	39 (67%)	44 (79%)	157 (83%)	53 (40%)	1553 (15%)
$\Delta\Phi_{\mu\bar{\mu}}$	10.3(64%)	64 (32%)	8 (25%)	10 (24%)	7.1 (16%)	46 (29%)	17 (33%)	218 (14%)
$m_{\mu\bar{\mu}}$	9.9 (96%)	44 (69%)	1.9 (24%)	1.3 (13%)	3.1 (44%)	28 (62%)	15 (85%)	27 (12%)
$P_T(\text{highest})$	8.7 (88%)	35 (78%)	1.5 (79%)	0.6 (50%)	1.3 (41%)	16 (57%)	14 (98%)	19 (56%)
$P_T(\text{lowest})$	8.7 (100%)	35 (100%)	1.5 (100%)	0.6 (100%)	1.3 (100%)	16 (100%)	14 (100%)	19 (100%)

Table 6.5: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 130 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

6.6 Final Event Selection

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	143	459	226	184	461	5128	177	39814
<i>Jet Veto</i>	86 (60%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10045 (25%)
<i>MET</i>	66 (78%)	174 (56%)	28 (48%)	35 (60%)	41 (74%)	143 (76%)	33 (25%)	1116 (11%)
$\Delta\phi_{\mu\mu}$	42 (63%)	46 (26%)	6.1 (22%)	(19%)	5.1 (12%)	32 (22%)	10 (29%)	122 (11%)
$m_{\mu\mu}$	38 (90%)	32 (71%)	1.3 (20%)	0.8 (13%)	2.5 (50%)	21 (65%)	8 (85%)	6 (5%)
$P_T(max)$	32 (83%)	23 (72%)	1.0 (78%)	0.3 (37%)	1.0 (39%)	9 (45%)	4 (46%)	1 (27%)
$P_T(min)$	21 (67%)	10 (43%)	0.1 (14%)	0 (0%)	0.3 (29%)	6 (59%)	0.07 (2%)	0.6 (40%)

Table 6.6: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

6.6.2 Study at a centre of mass energy of 10 TeV

The nominal design energy of 14 TeV of the LHC will not be reached until the following years. as discussed in 4.4, in the LHC workshop in Chamonix 2009 [67], a run at a centre of mass energy of 10 TeV was intended for the first run of the LHC. For that, a new study for an integrated luminosity of 200 pb^{-1} and at a centre of mass energy of 10 TeV, one of the planned benchmarks for physics in the LHC, was performed.

The analysis had to address the new centre of mass energy and the low luminosity considered. The differences in terms of cross-sections for signal and backgrounds when lowering the centre of mass energies were discussed in 4.6. Less events are expected, but the calibration and commissioning of the detector plus all the experience with cosmic data, led to a better identification and description of all the physical objects used in the analysis: Muons, Jets and missing E_T . The optimized muon selection and isolation criteria presented in 5.1.4 and 5.1.3 were applied in the study at 10 TeV, meaning an important reduction of the contribution of the $W + jets$ background with no effect on the signal efficiency. For the Jet Veto, the analysis incorporated the L2 + L3 energy corrections and moved to the infra-red safe Anti- k_t algorithm. Corrected Anti- k_t 5 CaloJets were considered. If at least one jet with a corrected energy higher than 30 GeV was found in $|\eta| < 2.5$, the event was rejected. Another improvement with respect to the 14 TeV study concerns the missing E_T estimation. The Track Corrected missing E_T was available and could be used in the 10 TeV analysis, leading to a significant reduction of the $Z + jets$ background.

A new optimization of the cuts was found not to be necessary. After the central jet veto the kinematical cuts on the main variables of the analysis were applied according to the Higgs boson mass hypothesis considered in the correspondent Monte Carlo production at 10 TeV. No miss-alignment or miss-calibration scenarios were considered, ideal conditions were assumed.

The expected number of events and the relative efficiencies after each cut are pre-

6. EVENT SELECTION AND KINEMATICS

sented for signal and backgrounds for each of Higgs mass hypothesis studied in the appendix A.2. The results for $m_H = 130 \text{ GeV}/c^2$ and $m_H = 160 \text{ GeV}/c^2$ are shown in tables 6.7 and 6.8.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$Z + jets$
Pre-sel. && $\mu\mu$	17.9	56.9	25.7	9.24	37	465	1272
Jet Veto	4.65 (60%)	38.4 (67%)	7.9 (30%)	6.10 (66%)	11 (2%)	3.7 (9%)	375 (29%)
MET	2.87 (61%)	24.7 (64%)	4.6 (57%)	4.55 (74%)	9.0 (79%)	2.8 (77%)	24 (6%)
$\Delta\Phi_{\mu\bar{\mu}}$	2.02 (70%)	8.8 (35%)	1.1 (23%)	0.95 (20%)	2.5 (27%)	0.8 (27%)	4 (14%)
$m_{\mu\bar{\mu}}$	1.94 (96%)	5.8 (65%)	0.16 (14%)	0.16 (16%)	1.4 (58%)	0.4 (47%)	0.7 (19%)
$P_T(\text{highest})$	1.71 (88%)	4.3 (74%)	0.10 (66%)	0.086 (54%)	0.7 (47%)	0.23 (63%)	0.7 (100%)
$P_T(\text{lowest})$	1.71 (100%)	4.3 (100%)	0.10 (100%)	0.086 (100%)	0.7 (100%)	0.23 (100%)	0.7 (100%)

Table 6.7: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 130 \text{ GeV}/c^2$, 200 pb^{-1} at 10 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$Z + jets$
Pre-sel. && $\mu\mu$	22.9	56.9	25.7	9.24	465	37	1272
Jet Veto	13.1 (57%)	38.4 (67%)	7.9 (30%)	6.10 (66%)	11 (2%)	3.7 (9%)	375 (29%)
MET	9.7 (73%)	21.2 (55%)	3.7 (47%)	4.14 (67%)	8.4 (74%)	2.7 (74%)	13 (3%)
$\Delta\Phi_{\mu\bar{\mu}}$	6.4 (66%)	6.2 (29%)	0.6 (15%)	0.72 (17%)	2.1 (24%)	0.6 (20%)	1 (11%)
$m_{\mu\bar{\mu}}$	5.9 (91%)	4.3 (68%)	0.08 (13%)	0.13 (17%)	1.2 (55%)	0.4 (64%)	0.7 (50%)
$P_T(\text{highest})$	5.1 (86%)	2.9 (68%)	0.05 (66%)	0.056 (44%)	0.5 (42%)	0.23 (63%)	0.7 (100%)
$P_T(\text{lowest})$	3.6 (71%)	1.3 (44%)	0.03 (50%)	0.016 (28%)	0.2 (33%)	0.07 (28%)	≤ 1.4

Table 6.8: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$, 200 pb^{-1} at 10 TeV

6.6.3 Study at a centre of mass energy of 7 TeV

The LHC is expected to run at $\sqrt{s} = 7 \text{ TeV}$ until 1 fb^{-1} of data is collected or the end of 2011, whichever happens first [4]. The analysis was consequently adapted to the current plans of CMS for the physics run at 7 TeV, and tested in the correspondent Monte Carlo production assuming 1 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$ of centre of mass energy.

The improved methods for the identification and isolation of prompt W muons, for the missing E_T estimation and for the Jet identification that were successfully used in the 10 TeV analysis 6.6.2 were still valid for the current study. The analysis was consistently reproduced in parallel using Anti- k_t 5 Particle Flow Jets and Particle Flow missing E_T , and Anti- k_t 5 JPT Jets and Track Corrected missing E_T giving a good agreement. After the central jet veto the kinematical cuts were applied according to the Higgs boson mass hypothesis considered. In this case, the 7 TeV Monte Carlo production was used, where several samples of Higgs signal for different mass hypothesis

6.7 Signal and Background production rates

and production modes were simulated. No miss-alignment or miss-calibration scenarios were considered, ideal conditions were assumed.

The expected number of events and the relative efficiencies after each cut are presented for signal and backgrounds for each of Higgs mass hypothesis studied in the appendix A.3. The results for $m_H = 130 \text{ GeV}/c^2$ and $m_H = 160 \text{ GeV}/c^2$ are shown in tables 6.9 and 6.10. One of the benchmark luminosities for the 7 TeV run will be 100

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	25.3	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	15.1 (59%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	7.4 (49%)	97 (40%)	17 (24%)	18.6 (34%)	6.9 (67%)	28 (64%)	≤ 14	55 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	5.1 (68%)	33 (33%)	4.1 (25%)	4.0 (21%)	2.0 (29%)	6 (22%)	—	5 (8%)
$m_{\bar{\mu}}$	4.9 (96%)	18 (55%)	1.1 (26%)	0.7 (17%)	1.3 (63%)	4 (69%)	—	≤ 4
P_T (highest)	4.5 (92%)	14 (78%)	0.6 (57%)	0.4 (58%)	0.5 (42%)	2.2 (50%)	—	—
P_T (lowest)	4.5 (100%)	14 (100%)	0.6 (100%)	0.4 (100%)	0.5 (100%)	2.2 (100%)	—	—

Table 6.9: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 130 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	67.5	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	37.6 (55%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	23.9 (63%)	83 (34%)	15 (21%)	16.5 (30%)	6.2 (61%)	26 (59%)	≤ 14	20 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	15.8 (65%)	18 (22%)	2.5 (17%)	2.9 (17%)	1.4 (23%)	6 (21%)	—	0.8 (3%)
$m_{\bar{\mu}}$	14.7 (92%)	9 (51%)	1.0 (37%)	0.6 (19%)	0.9 (61%)	4 (64%)	—	≤ 1.5
P_T (highest)	13.0 (88%)	7 (75%)	0.5 (49%)	0.2 (42%)	0.4 (47%)	0.8 (23%)	—	—
P_T (lowest)	9.3 (71%)	4 (61%)	0.2 (33%)	0.12 (49%)	0.16 (36%)	0.6 (66%)	—	—

Table 6.10: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

pb^{-1} , expected at the end of 2010. In table 6.11, the expected number of events for a Higgs signal with $m_H = 160 \text{ GeV}/c^2$ assuming an integrated luminosity of 100 pb^{-1} is shown. The expected number of events for signal will be, for the highest cross-section of the order of 1 ± 1 , and less than 1 event is expected for background. The Higgs search in this channel will therefore not be exploited until the end of 2011, when ten times the luminosity is expected to be collected.

6.7 Signal and Background production rates

The analysis was designed for an integrated luminosity of 1 fb^{-1} at 14 TeV and subsequently adapted to the changing running conditions of the LHC to 200 pb^{-1} at 10 TeV

6. EVENT SELECTION AND KINEMATICS

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$ selection	6.7	31	30	22	10	139	2	42027
all cuts	0.93	0.4	0	0	0	0	0	0

Table 6.11: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$, 100 pb^{-1} at 7 TeV

and finally to 1 fb^{-1} at 7 TeV. It has evolved integrating all the possible improvements in the identification and proper estimation of the physics objects used, that come from the increasing knowledge of the CMS detector acquired during commissioning and many campaigns of cosmic data-taking. The analysis based on sequential cuts was applied to different Monte Carlo generated samples of all the possible backgrounds and to a wide spectrum of Higgs signal samples with different mass hypothesis. The expected number of events for different masses and 1 fb^{-1} at 14 TeV is presented in table 6.12.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$m_H = 120 \text{ GeV}/c^2$	4.9	35	1.5	0.6	1.3	16	14	19
$m_H = 130 \text{ GeV}/c^2$	8.7	35	1.5	0.6	1.3	16	14	19
$m_H = 140 \text{ GeV}/c^2$	18	35	1.5	0.6	1.3	16	14	19
$m_H = 150 \text{ GeV}/c^2$	9	10	0.1	0	0.3	6	0.07	0.6
$m_H = 160 \text{ GeV}/c^2$	21	10	0.1	0	0.3	6	0.07	0.6
$m_H = 170 \text{ GeV}/c^2$	17	10	0.1	0	0.3	6	0.07	0.6
$m_H = 180 \text{ GeV}/c^2$	11	7	0.1	0	0.8	5	0.01	0.6
$m_H = 190 \text{ GeV}/c^2$	5.6	7	0.1	0	0.8	5	0.01	0.6
$m_H = 200 \text{ GeV}/c^2$	4.6	7	0.1	0	0.8	5	0.01	0.6

Table 6.12: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis for 1 fb^{-1} at 14 TeV

The expected number of events for 200 pb^{-1} at 10 TeV for different signals and the main backgrounds are presented in table 6.13.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$Z + jets$
$m_H = 130 \text{ GeV}/c^2$	8.55	21.5	0.5	0.43	1.15	3.5	3.5
$m_H = 150 \text{ GeV}/c^2$	8.75	6.5	0.15	0.08	0.35	1	-
$m_H = 160 \text{ GeV}/c^2$	18	6.5	0.15	0.08	0.35	1	-
$m_H = 170 \text{ GeV}/c^2$	13.9	6.5	0.15	0.08	0.35	1	-
$m_H = 200 \text{ GeV}/c^2$	2.94	4.5	0.15	0.14	0.5	1	-

Table 6.13: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis for 1 fb^{-1} at 10 TeV

6.7 Signal and Background production rates

And finally, the expected number of events for 1 fb^{-1} at 7 TeV for the backgrounds and all the Higgs mass points available are presented in table 6.14. The expected number

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$m_H = 120 \text{ GeV}/c^2$	1.80	14	0.6	0.4	0.5	2.2	—	—
$m_H = 130 \text{ GeV}/c^2$	4.5	14	0.6	0.4	0.5	2.2	—	—
$m_H = 140 \text{ GeV}/c^2$	7.5	14	0.6	0.4	0.5	2.2	—	—
$m_H = 150 \text{ GeV}/c^2$	4.8	4	0.2	0.12	0.16	0.6	—	—
$m_H = 155 \text{ GeV}/c^2$	5.7	4	0.2	0.12	0.16	0.6	—	—
$m_H = 160 \text{ GeV}/c^2$	9.3	4	0.2	0.12	0.16	0.6	—	—
$m_H = 165 \text{ GeV}/c^2$	9.0	4	0.2	0.12	0.16	0.6	—	—
$m_H = 170 \text{ GeV}/c^2$	7.4	4	0.2	0.12	0.16	0.6	—	—
$m_H = 175 \text{ GeV}/c^2$	5.3	4	0.2	0.12	0.16	0.6	—	—
$m_H = 180 \text{ GeV}/c^2$	4.5	1.6	0.2	0.2	0.18	1.7	—	—
$m_H = 190 \text{ GeV}/c^2$	2.1	1.6	0.2	0.2	0.18	1.7	—	—
$m_H = 200 \text{ GeV}/c^2$	1.17	1.6	0.2	0.2	0.18	1.7	—	—

Table 6.14: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis for 1 fb^{-1} at 7 TeV

of events for signal and backgrounds for the main Higgs mass hypothesis at the three different centre of mass energies for 1 fb^{-1} are presented in table 6.15.

	14 TeV		10 TeV		7 TeV	
	N_S	N_B	N_S	N_B	N_S	N_B
$m_H = 130 \text{ GeV}/c^2$	8.7	87.4	8.55	30.58	4.5	17.7
$m_H = 150 \text{ GeV}/c^2$	9	17.07	8.75	8.08	4.8	5.08
$m_H = 160 \text{ GeV}/c^2$	21	17.07	18	8.08	9.3	5.08
$m_H = 170 \text{ GeV}/c^2$	17	17.07	13.9	8.08	7.4	5.08
$m_H = 200 \text{ GeV}/c^2$	4.6	13.51	2.94	6.29	1.17	3.88

Table 6.15: Number of expected events for signal and backgrounds for different Higgs mass hypothesis and centre of mass energies assuming 1 fb^{-1} of integrated luminosity

6. EVENT SELECTION AND KINEMATICS

6.8 $\mu\mu$, $e\mu$ and ee final states

The sensitivity of the $H \rightarrow WW^*$ depends on the three final states. The study to the $e\mu$ and ee final states is complementary to the $\mu\mu$ and the final result combines the three analysis.

During the evolution of the analysis, from 14 TeV the analysis based on sequential cuts for the three final states have been performed in parallel. All the analyses can start from the same initial sample, lepton-triggered (with an OR of electron and muon triggers). Then, the lepton selection is done. Each channel treats the leptons in the appropriate way, selecting either two μ , two e or a pair $e\mu$. The three final states use the same Jet definition and missing E_T estimation. After the lepton selection the jet veto is applied, and then the kinematical cuts are applied on the same variables following the same order. The optimization of the cuts was done independently for each final state.

In table 6.16, the expected number of events after all cuts for 14 TeV are presented. In ideal conditions, the $e\mu$ channel would have of the order of twice the signal than ee or $\mu\mu$, and less background, but differences arise due the optimization of the kinematical cuts.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$ final state	21	10	0.1	0	0.3	6	0.07	0.6
ee final state	20	13	1.6	0	0	4	2.03	0.9
ee final state	32	16	0.3	0	1.4	6.8	6.25	0

Table 6.16: Expected number of events after all cuts, mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

6.9 Complementary Strategies for the Analysis

As the analysis can be complicated for low masses and in low luminosity scenarios, two complementary strategies can be designed in order to maximize the sensitivity for the considered search channel. More strategies can be further defined when the analysis is performed with real data.

Detailed studies concerning the use of the track content of the event meant a possible improvement in the control of the $t\bar{t}$ background. A multivariate approach can be applied when the cross-section of the signal is so small that the proper event selection in the sequential cuts analysis becomes difficult.

6.9.1 Use of track Jets and b-tagging

The Monte Carlo truth shows that after the application of the Central Jet Veto there still remain some b -quarks coming from top decays inside the acceptance that can go through the kinematical cuts sequence. Some of them even carry a significant transverse momentum, as it is shown in figure 6.20. The identification of the b -quarks that are not reconstructed as Jets represents a challenge for the analysis. The aim is to properly identify all the b -quarks that, within the acceptance, carry enough momentum to pass the jet threshold. For that, the Track content of the events can be studied. Also a

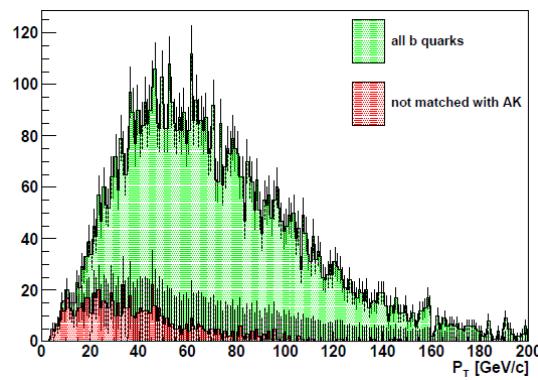


Figure 6.20: - Monte Carlo truth distribution of the P_T of all the generated b -quarks in green and the ones not matched to Anti- k_t Calo Jets, in red, for $t\bar{t}$ events.

Particle Flow based study would entail improvements in the b -jets identification, but the study is not considered here.

6.9.1.1 Track Jets

The Track Jets, defined in section 5.3, have been studied in parallel in order to complement the calorimetric jet reconstruction and to evaluate if the analysis could be performed with, if not just tracker and muon chambers information, as little calorimetric information as possible.

The tracker information is always valuable in the analysis, especially to identify low P_T b -jets coming from $t\bar{t}$ events that can pass the Jet Veto with the actual requirements. Track Jets give a better estimation of the original b -quark direction.

For that a TrackJets veto can be used just after the Central Jet Veto, considering Track Jets

- $E_T > 15 \text{ GeV}/c^2$
- $|\eta| < 2.5$

6. EVENT SELECTION AND KINEMATICS

As an example of the use of track jets, in table 6.17 the analysis based on sequential cuts using CaloJets is compared with the analysis based on sequential cuts using Track-jets veto at CJV cut level.

	HWW	WW	WZ	ZZ	$t\bar{t}$	tW	$Z + jets$	$W + jets$
CJV (caloJets)	37.65	246.1	106.9	64.2	63.75	13.25	350680.45	31.75
CJV (trackJets)	41.85	244.9	96.55	60.65	46.35	11.55	347889.85	28.85

Table 6.17: Results of the application of the Central Jet Veto using Calo or Track jets, for signal with a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ and backgrounds, 1 fb^{-1} at 7 TeV

In table 6.18, the central jet veto using Track Jets is applied not in substitution but as a complement of the CJV applied in the analysis. The analysis can be performed

	HWW	WW	WZ	ZZ	$t\bar{t}$	tW	$Z + jets$	$W + jets$
CJV (caloJets)	37.65	246.1	106.9	64.2	63.75	13.25	350680.45	31.75
CJV (trackJets)	34.5	229.25	70.65	47.65	24.45	7.2	326873.7	28.85

Table 6.18: Results of the application of the Central Jet Veto using Track Jets right after the CJV using Calo jets, for signal with a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ and backgrounds, 1 fb^{-1} at 7 TeV

using Track Jets only with no dramatic effect on the signal to background ratio, and can obtain some improvement with the use of Track Jets as a complement of the Calo Jets.

6.9.1.2 Soft-muon and Life time based b-tagging

Two features of the b -quark can be used for its identification: the b -quark has a large lifetime, that allows the hadrons containing b -quarks to travel some distance before decaying; and a large mass, that leads to high transverse momentum in its decay products. This causes b -jets to be displaced from the primary vertex and to be wider, have higher multiplicities and invariant masses, and also to contain low-energy leptons with momentum perpendicular to the jet. These properties can be measured, and jets that have them are more likely to be b -jets. Two kind of b -tagging methods [120] were studied in order to identify jets coming from decays of b -quarks in this study, soft-muon, and lifetime based b -tagging.

The soft-muon method [121] is used in many analyses [122; 123] to get rid of $t\bar{t}$ events in which one of the two b -quarks decays into a muon. In this case, as it is not isolated, this muon will not be selected in the first step of the analysis, and the event

6.9 Complementary Strategies for the Analysis

can go further in the analysis chain. A soft-muon veto can be applied to suppress these events. The soft muon can be defined as follows:

- A Muon failing the muon selection criteria (see 5.1.4),
- with at least one segment in the muon chamber and some penetration requirements (`TMLastStationOptimizedLowPtTight`),
- $P_T > 3 \text{ GeV}/c$,
- $|\eta| < 2.4$,
- $|d0| < 0.2 \text{ cm}$,
- $N_{Hits} > 10$

The soft-muon tagging can be used to estimate $t\bar{t}$ background contribution in data by counting the number of soft muons in the final state without jets. To estimate the background coming from this source the following strategy can be used. Soft muon tagging efficiency can be obtained from Monte Carlo simulation where more statistics and a better control of systematics are available, leading to a expected tagging efficiency of the order of 20%. This value can be compared with the tagging efficiency measured directly from data in events having one or two jets. Finally the top background can be obtained using the tagging efficiency derived from Monte Carlo in cases where no jets are present. Different Monte Carlo samples can be used to estimate systematic errors of this method.

The fraction of events where at least one soft muon is found as a function of the jet multiplicity, after the muon selection, for Higgs signal and for all the backgrounds stacked is shown in figure 6.21.

The simplest lifetime-based b-tags [124; 125; 126; 127] are based on the three dimensional signed impact-parameter $d0$. It is signed positive/negative if the track passes closest to its associated jet-axis down/up-stream of the primary vertex. Quality cuts, like $|d0| < 2 \text{ mm}$ and jet-track miss-distance less than 0.7 mm , are applied to reject tracks that are badly reconstructed or originate from conversions etc. Track Counting b-tag algorithms order the tracks in a jet by decreasing impact-parameter significance $\frac{d0}{\sigma d0}$. Its output is the $\frac{d0}{\sigma d0}$ of typically, the second or third track, that has a large, positive tail for b-jets. The discriminant is built around the requirement of a minimum number of good quality tracks with an impact parameter significance exceeding a given threshold.

From the different b -tagging algorithms available in CMS, a safe choice would be the track counting algorithm, `TrackCountingHighEffBJetTags`. Anti b -tagging can be applied in the analysis by vetoing events with at least one CaloJet with $E_T^{\text{corr.}} > 20 \text{ GeV}$ and a b -tagging discriminant greater than 1.75, which is the current working point in

6. EVENT SELECTION AND KINEMATICS

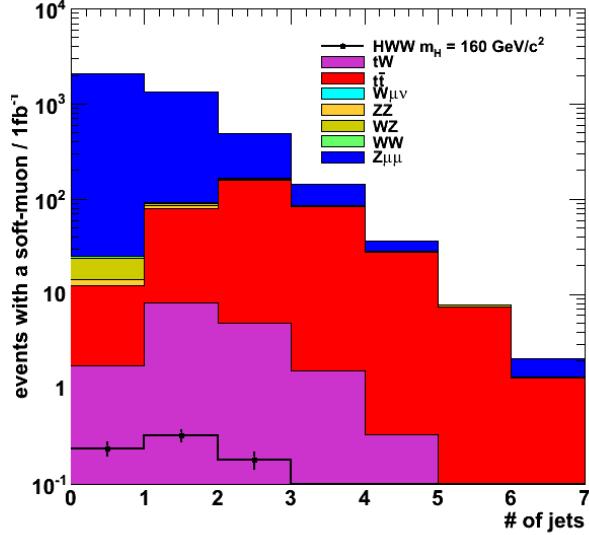


Figure 6.21: - Number of events with at least one soft-muon versus the number of jets in the event for Higgs to WW^* signal, with $m_H = 160 \text{ GeV}/c^2$ on top of the stacked backgrounds in different colors, considering 1 fb^{-1} .

the b -tag working group of CMS. In table 6.19, the result of the application of additional cuts using b -tagging and soft-muon in the analysis after jet veto are shown. A 69 % of the remaining $t\bar{t}$ background can be achieved loosing just 6 % of signal in that level.

	HWW	WW	WZ	ZZ	$t\bar{t}$	tW	$Z + jets$	$W + jets$
CJV (caloJets)	37.65	246.1	106.9	64.2	63.75	13.25	350680.45	31.75
IP based btagging	35.4	237.45	99.9	58.05	23.05	7.1	338864.9	28.85
Soft muon veto	35.15	236.3	91.15	55.8	19.4	6.4	337536.8	28.85

Table 6.19: Results of the application of IP-based b -tagging and Soft-muon veto for signal with a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ and backgrounds, 1 fb^{-1} at 7 TeV

6.9.2 Multivariate Methods

In the low mass region for a Higgs boson with a mass $\sim 130 \text{ GeV}/c^2$, the analysis is more complicated, even for 14 TeV data. It has to search for a very small signal in large datasets. In order to extract more signal, the acceptance has to be higher for background too. Multivariate classification methods based on machine learning techniques can then be a fundamental ingredient to the search. In recent years, these methods have significantly evolved and have been successfully used in other experiments.

6.9 Complementary Strategies for the Analysis

A multivariate analysis based on Boosted Decision Trees, BDT, has been carried out in the $\mu\mu$ channel to improve the signal to background discrimination. The TMVA [128] toolkit integrated in ROOT was used. Several approaches have been tested, although up to now a bi-dimensional BDT is the method being used in the analysis, since it is found to provide the best results. This BDT relies on the observation that the most important backgrounds after reducing others are WW and $t\bar{t}$, and aims at the rejection of as much as possible of these backgrounds.

Some of the most powerful variables to discriminate against these backgrounds present different shapes, so a BDT trained against the sum of them, that is, a plain BDT one-dimensional approach, might not learn well enough the properties of each one and may therefore not be optimal. Then, the idea of this method is to develop two, and even three independent BDTs, one trained against each background alone, which are optimized separately.

The sample used for training and testing has been selected after the HLT requirement, the standard di-muon selection, the common preselection cuts and the central jet veto. In addition to the main discriminating variables used in the analysis based on sequential cuts, some more have been used to build the BDT. The additional variables used are, in the case of a bi-dimensional BDT trained against $t\bar{t}$ and WW :

- $\eta^{\mu\mu}$
- $\min(\Delta\phi_{\mu^1, E_T}, \Delta\phi_{\mu^2, E_T})$
- $P_T^{\mu\mu}$

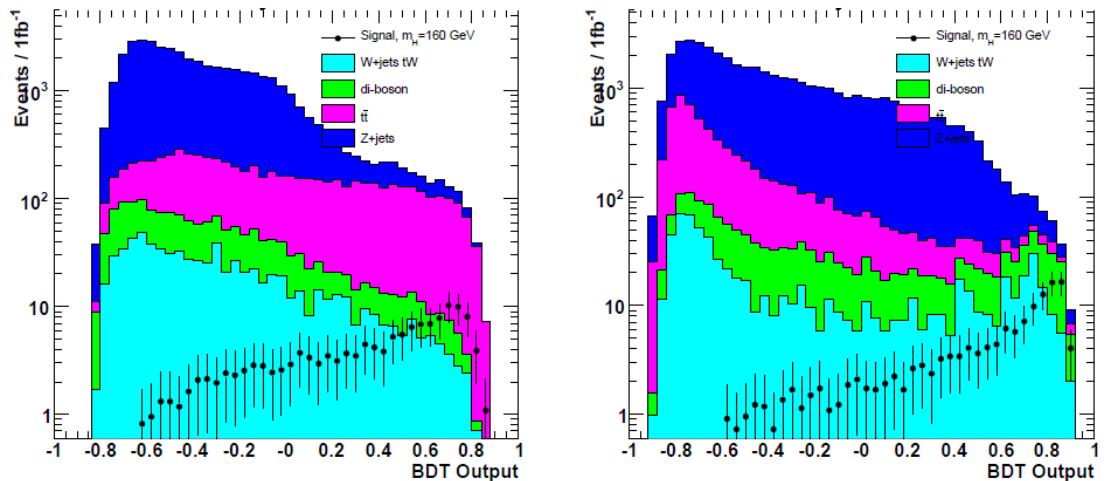


Figure 6.22: - Muon multivariate analysis. Distributions of the two BDT discriminant variables for the training against WW (left) and $t\bar{t}$ (right)

After the training, the two resulting functions, that can be seen in figure 6.22, against WW and $t\bar{t}$, are combined into a bi-dimensional one. The discrimination power of

6. EVENT SELECTION AND KINEMATICS

this variable is evident in figure 6.23, where the distributions of the 2D variable are presented applied to WW , $t\bar{t}$, $Z + jets$, and signal with mass $160 \text{ GeV}/c^2$.

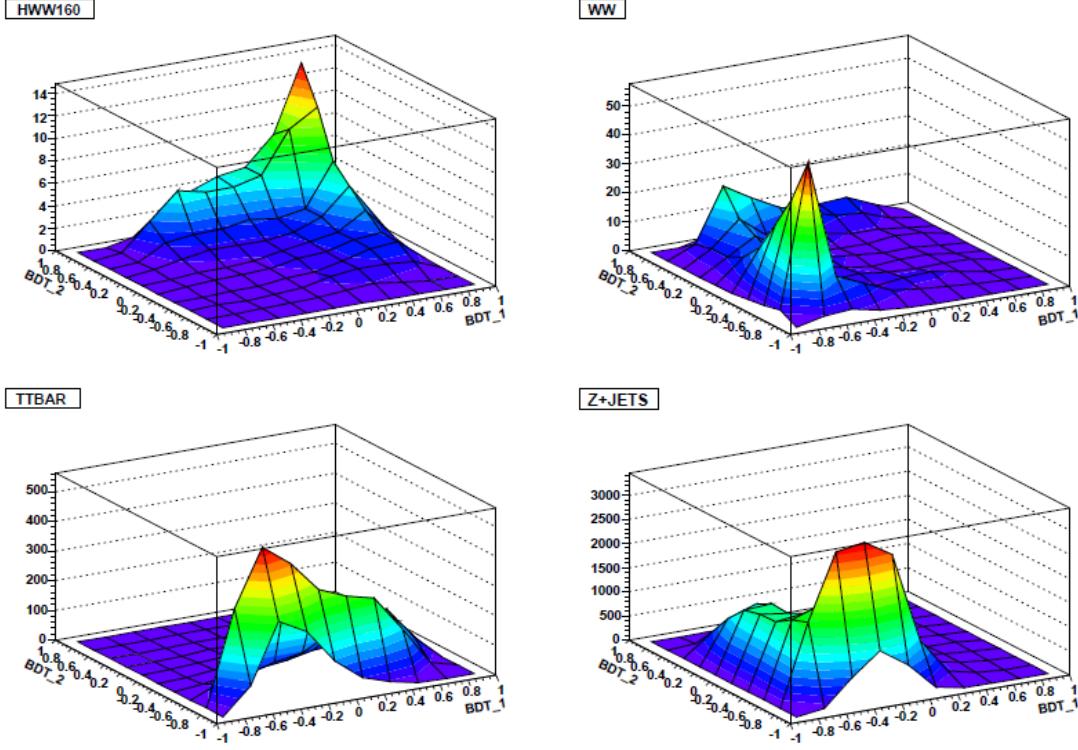


Figure 6.23: - Muon multivariate analysis. Distributions of the 2D variable constructed from the two BDT trained against WW and $t\bar{t}$ when applied to WW , $t\bar{t}$, $Z + jets$, and signal with mass $160 \text{ GeV}/c^2$

A circular cut is applied around the region with highest signal content, the center of the circle and the radius are chosen optimizing the significance. This radius is the final discriminating variable. A substantial improvement can be achieved when using multivariate methods in the low mass region, for example, table 6.20, with results for 14 TeV considering 1 fb^{-1} . On the other hand, the training of the BDT requires very high statistics and it is very sensitive to fluctuations. a perfect understanding of the additional variables used and its correlations is needed is needed. For these reasons it will never be a first day analysis.

The analysis of the Higgs to WW^* decay channel to two muons and two neutrinos was presented and performed in Monte Carlo considering different centre of mass scenarios an integrated Luminosities. Event yields are calculated for signal and backgrounds for

6.9 Complementary Strategies for the Analysis

	Cut-based		BDT	
	N_S	N_B	N_S	N_B
$m_H = 130 \text{ GeV}/c^2$	8.7	87.4	6.9	38.0
$m_H = 150 \text{ GeV}/c^2$	9	17.07	13.0	27.0
$m_H = 160 \text{ GeV}/c^2$	21	17.07	16.4	6.0
$m_H = 170 \text{ GeV}/c^2$	17	17.07	11.8	6.0
$m_H = 200 \text{ GeV}/c^2$	4.6	13.51	8.2	25.0

Table 6.20: Number of expected events for signal and backgrounds for different Higgs mass hypothesis and centre of mass energy of 14 TeV assuming 1 fb^{-1} of integrated luminosity

different Higgs mass hypothesis. However, as this is a Monte Carlo exercise the main backgrounds will not be completely understood until CMS collects more collision data. Data driven methods proposed to be applied in the future to estimate $t\bar{t}$, WW , $Z+jets$ and $W+jets$ backgrounds are described in the following chapter.

6. EVENT SELECTION AND KINEMATICS

7

Methods to estimate the Background from Data and Systematic Uncertainties

As discussed in previous chapters, the performance of the analysis presented in this thesis will depend on how well the signal will be identified in data, with the appropriate Muon selection, missing E_T estimation and jet identification, and how the different backgrounds will be controlled and reduced, eventually to zero. The main backgrounds are $t\bar{t}$, $Z + jets$ and $W + jets$, that present very large cross-sections, and WW , which is almost identical to the signal. Many different studies have been proposed to estimate and control them, and several of them have been performed in the context of this thesis. All the studies are performed for a centre of mass energy of 7 TeV, and an integrated luminosity of 1 fb^{-1} .

7.1 Control Regions to estimate the Background

A strategy to define control regions for $t\bar{t}$, WW and $W + jets$ backgrounds in real data is presented in this section.

A normalization region in the phase-space is defined for each case to measure the background rate. This region has to be close to the signal region, but still with a minimum presence of signal events. Then, the background measured in that region is used as the reference to estimate the magnitude of the background in the signal region. For that the measured events in the background region, N_{bkg}^N , are multiplied by the ratio of the efficiencies:

$$N_{bkg}^S = \frac{\varepsilon_{bkg}^S}{\varepsilon_{bkg}^N} N_{bkg}^N \quad (7.1)$$

If the normalization region is close enough to the signal region, most of the systematic errors will cancel in the ratio of the efficiencies, giving as a result a reliable estimation of

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

the background contamination in the signal region. The precise definition of a control region depends of the definition of the signal region.

7.1.1 $t\bar{t}$ Background Normalization

The definition of the control region of the $t\bar{t}$ background is done by explicitly asking for the presence of at least two jets in the final state. The projected missing E_T variable, $E_T^{miss} \cdot \sin(\min(\Delta\phi_{(E_T^{miss}, \mu_{closest})}, \frac{\pi}{2}))$, which helps against $Z + jets$ background contamination, especially decaying to τ leptons, has been used to clean the sample of this kind of background. The set of proposed cuts is therefore:

- $P_T^{min} > 15 \text{ GeV}/c$
- $\Delta\phi_{\mu\mu} < 150 \text{ (deg.)}$
- $|m_{\mu\mu} - m_Z| > 30 \text{ GeV}$
- $E_T^{miss} \cdot \sin(\min(\Delta\phi_{(E_T^{miss}, \mu_{closest})}, \frac{\pi}{2})) > 50 \text{ GeV}$
- Number of jets in the event > 1

The cuts are applied after the muon selection step in the analysis. The jets are defined in the same way as described in section 6.3, and several jet reconstruction algorithms have been tested leading to comparable results.

In table 7.1 the expected number of events in the $t\bar{t}$ control region for $t\bar{t}$ and the rest of the backgrounds is presented. The number of expected events for signal considering a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ is ~ 5 . The ratio R is then defined

	$t\bar{t}$	WW	WZ	ZZ	tW	$W + jets$	$Z + jets$
Number of expected events (all cuts)	200	3	0.2	0.12	7.4	0	5

Table 7.1: Number of expected events in the $t\bar{t}$ control region. The number of expected events for signal considering a Higgs mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ is 4.6.

as $R = \frac{\epsilon_{CJV}^S}{\epsilon_{2jets}^C}$, leading to:

$$\frac{\Delta N_{t\bar{t}}^S}{N_{t\bar{t}}^C} = \frac{\Delta R}{R} \bigoplus \frac{\Delta N_{t\bar{t}+background}^C}{N_{t\bar{t}}^C} \bigoplus \frac{\Delta N_{background}^C}{N_{t\bar{t}}^C} \quad (7.2)$$

where, $N_{t\bar{t}}^S$ is the number of $t\bar{t}$ events in the signal region, and $N_{t\bar{t}}^C$ is the number of $t\bar{t}$ events in the control region. Three terms contribute, the first one comes from the Jet Energy Scale (JES), the second from the statistical fluctuation of the observed number of events in the real experiment, and the last one coming from the fractional uncertainty on the contamination due to the rest of the backgrounds in the control region.

7.1 Control Regions to estimate the Background

To estimate the effect of JES uncertainty on R, the 4-momentum of the jet is re-scaled by the transformation $E_T^{jet} \rightarrow E_T^{jet}(1 + \lambda)$. The effect of this shift is presented in figure

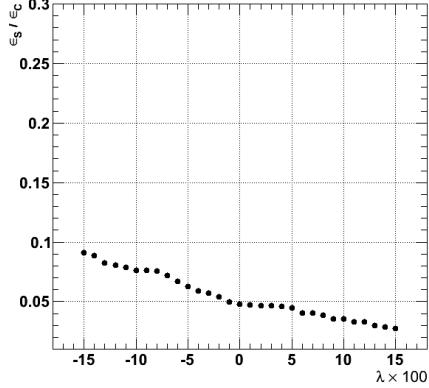


Figure 7.1: - Ratio of the efficiencies in the 0-jet bin and in the 2 or more for $t\bar{t}$ as a function on λ . The E_T of the jet is multiplied by $(1 + \lambda)$.

7.1. It is possible to compute a mean value of $\frac{\Delta R}{R}$ from the average of $\frac{\Delta R}{R}(+)$ and $\frac{\Delta R}{R}(-)$. Considering the expected systematic error for the jet identification as 7%, that leads to $\sim 28\%$. The statistical uncertainty for $t\bar{t}$ events will be $\sim 7\%$. The uncertainty on the background is set to be around 20% and $\frac{\sum N_{B,i}^C}{N_{t\bar{t}}^C} \sim 8\%$, leading to a contribution of about 2.5%. The total uncertainty in the control region is then computed giving $\sim 29\%$.

To model the jet multiplicity for $t\bar{t}$ events, as the signal region corresponds to cases where no reconstructed jets are found in the central region, it is possible to look at the number of events with 2 or more than 2 jets, compute the ratio, and extrapolate the agreement between data and simulation to the events in which no jets are present. In figure 7.2, the distribution of the number of jets after the application of the rest of the selection. The expected number of events for two jets is ~ 144 and ~ 103 for more than two jets, which gives a relative error when extrapolating to the 0 jet bin of about 7% since the number of events on that region is about 5 times smaller.

In addition to that, it is useful to look at several jet angular distributions, like the η distribution for the two most energetic jets, or the ΔR and $\Delta\eta$ distributions between the two most energetic jets, as shown in figure 7.3. With a good agreement between data and simulation for these distributions the correction of the theoretical jet angular predictions will be ensured. The overall error is about 29.8%.

As discussed in section 6.9.1.2 the possibility of identifying top quark decays by the presence of soft muons from the semi-leptonic decays of the two b-quarks present in the event can also be studied.

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

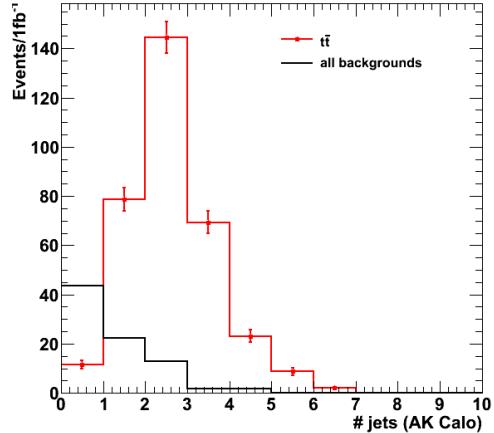


Figure 7.2: - Distribution of the number of jets in the events in the control region defined by all the cuts except the one in the number of jets, for $t\bar{t}$ and the backgrounds.

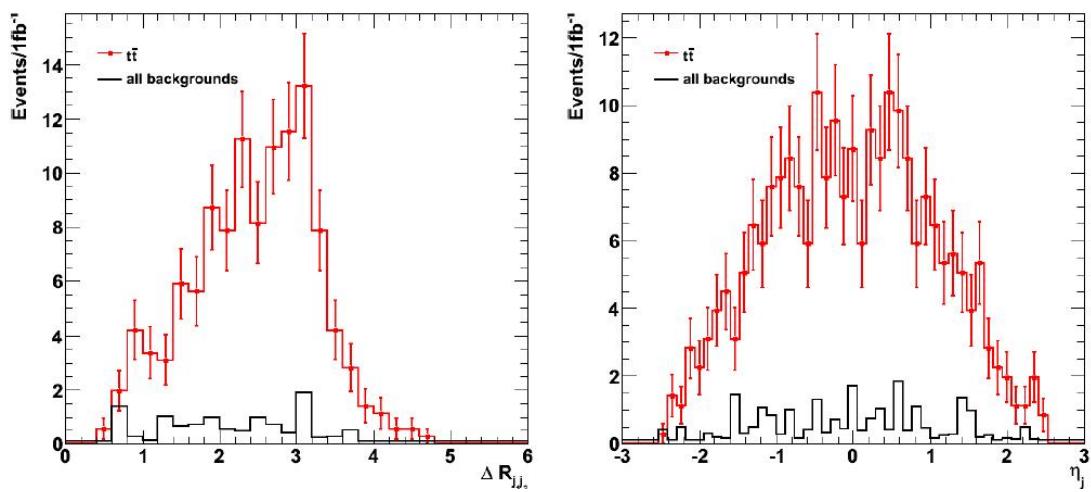


Figure 7.3: - ΔR (left) and distribution of the pseudorapidity, η , (right) of the two most energetics jets in the $t\bar{t}$ control region, for $t\bar{t}$ and the rest of backgrounds together.

7.1 Control Regions to estimate the Background

7.1.2 WW Background normalization

The WW normalization region can be defined in a similar way as the $t\bar{t}$ region, but applying a jet veto instead of selecting events with two jets, and applying an additional cut in the invariant mass to avoid Higgs signal contamination. The set of proposed cuts for the WW control region is the following:

- $P_T^{min} > 15 \text{ GeV}/c$
- $\Delta\phi_{\mu\mu} < 150 \text{ (deg.)}$
- $|m_{\mu\mu} - m_Z| > 30 \text{ GeV}$
- $m_{\mu\mu} > 115 \text{ GeV}$
- $E_T^{miss} \cdot \sin(\min(\Delta\phi_{(E_T^{miss}, \mu_{closest})}, \frac{\pi}{2})) > 50 \text{ GeV}$
- Number of jets in the event = 0

The selection is applied after the muon identification step in the signal and background samples. In 7.2, the expected number of events for WW and the different backgrounds is presented, no events are expected for Higgs signal considering $m_H = 160 \text{ GeV}/c^2$.

	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Number of expected events (all cuts)	14	0.6	0.3	1.1	5	0	0.8

Table 7.2: Number of expected events in the WW control region. The number of expected events for Higgs signal with mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ is 0.

Assuming a 15% error on the background normalization, the precision on the WW normalization region is about 28%.

It is possible to go further in the study of the WW within the context of the Higgs to WW^* analysis, as the same objects and strategy can be used to measure the WW cross-section, as will be discussed in 7.4.

7.1.3 $Z + jets$ Background normalization

Another important source of background comes from $Z + jets$, when the Z boson decays into two muons. As discussed before, this background is controlled by means of an accurate estimation of the missing E_T .

A control region for this background is designed requiring the di-lepton invariant mass to be inside the Z mass region. This $Z + jets$ selection can be used for many studies, namely to test the lepton ID modeling as well as trigger efficiencies, and the missing E_T estimation. After the muon selection, the Central Jet Veto is applied using the definition presented in 6.3 followed by two kinematical requirements:

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

- $15 < E_T^{miss} \cdot \sin(\Delta\phi(E_T^{miss}, \mu_{closest})) < 25 \text{ GeV}$
- $m_{\mu\mu} < 70 \text{ and } m_{\mu\mu} > 110 \text{ GeV}/c^2$.

The expected number of events for $Z + jets$ and the different backgrounds is presented in table 7.3, around 8 events of Higgs signal with a mass of $m_H = 160 \text{ GeV}/c^2$ are expected to be present in the control region.

	$Z + jets$	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$
Number of expected events (all cuts)	1652	36	0.9	1.4	0.9	4	0

Table 7.3: Number of expected events in the $Z + jets$ control region. The number of expected events for Higgs signal for a mass hypothesis of $m_H = 160 \text{ GeV}/c^2$ is 8.0.

7.2 Control of the Muon Fake Rate in $W + jets$ events

The $W + jets$ background becomes a dangerous background when one jet or track is misidentified as a lepton. In that case, two leptons are present in the event, one prompt lepton from the W boson decay, well isolated; and a fake non-prompt lepton from the hadronic decay of short lived mesons (D, B) or the decay in the detector of long lived particles(pions, kaons), or from misidentified hadrons due to, for example, electron conversions or $\pi^\pm\pi^0$ overlap [65; 129; 130].

The proposed data-driven method to estimate the contribution of the $W + jets$ background for the $H \rightarrow WW^* \rightarrow 2l2\nu$ signal is developed in two consecutive steps. In the first step, the probability of a jet faking a lepton is measured from the rate by which jets are identified as leptons.

$$f_{jet \rightarrow \text{fake lepton}} = \frac{\#\text{of jets matching lepton}(l)}{\#\text{reconstructed jets}(d)} \quad (7.3)$$

The numerator, l , is defined as the number of reconstructed jets that match a reconstructed, and eventually, identified lepton with $P_T > 10 \text{ GeV}/c$, called l -objects. The denominator, d , are reconstructed jets, called d -objects. In the second step, the standard selection is relaxed by asking only one fully identified lepton(l) with $P_T > 20 \text{ GeV}/c$ and an additional jet, d , with $P_T > 10 \text{ GeV}/c$. All possible $(l+d)$ -combinations in the event that pass the relaxed selection but fail the standard selection are retained and used to estimate this background after being reweighted according to the lepton faking jet probability of the corresponding d -object previously measured in the QCD multi-jet data. The samples, $W + jets$, Wbb , Wcc and QCD , have been used with a minimal event selection consistent of loose trigger paths.

There are differences in the implementation of this method for muons or for electrons.

7.2 Control of the Muon Fake Rate in $W + jets$ events

In the case of the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$, the muon fake rate is estimated. Because most of the muons are not accompanied by hadronic activity, the use of jets as d -objects may not be the most appropriate choice. For muons, therefore, isolated tracks with $|\eta| < 2.5$ and $P_T > 10$ GeV/c are used instead of jets to calculate the muon faking loosely isolated track probability:

$$f_{\text{loosely isolated track} \rightarrow \text{fake muon}} = \frac{\#\text{muons}(l)}{\#\text{reconstructed loosely isolated tracks}(d)} \quad (7.4)$$

where the denominator, d , has to satisfy the same loose tracker isolation defined for muon and some additional quality requirements, namely: $\chi^2/\text{dof} < 40$, number of hits > 8 and transverse impact parameter < 0.1 .

The probability of a loosely isolated track faking a muon according to definition in 7.4,

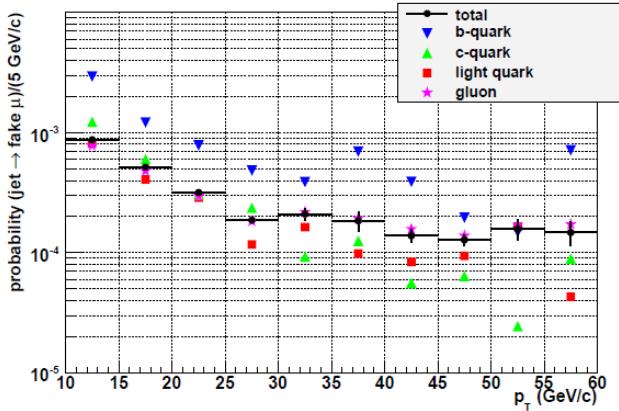


Figure 7.4: - Muon faking loosely isolated track probability as a function of p_T after lepton reconstruction and identification. The black dashed line is obtained with QCD multi-jet events and the red full line with $W + jets$ events.

is given in Fig. 7.4. The distributions provide the weights necessary in the reweighting of the $(l + d)$ selected $W + jets$ events. Due to the limited statistics, this was not possible and a simple definition was used to define the denominator-objects.

Using a *QCD* sample, the fake rate is calculated as follows: once a l -object, that is, a properly identified muon fulfilling all the quality cuts of the analysis, including isolation, (see 5.1.4) is found, the event is scanned for all the possible d -objects, defined as above, in order to make $l + d$ pairs. Some of the $l + d$ pairs will be, indeed, $l + l$ pairs, and with that, the fake rate is calculated as:

$$f = \frac{l + l}{l + d} \quad (7.5)$$

This method will be used and further developed in real data. A preliminary study has been already performed on real collision data.

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

7.3 missing E_T studies from data

A very extended method to control the missing E_T modeling can be done in data using $Z + jets$ events. The method has been tested in Monte Carlo samples and it consists of the subtraction of one of the two muons in $Z \rightarrow \mu\mu$ as if it was a real neutrino. The P_T distribution of the dropped muon is then compared with the missing E_T distribution for $W \rightarrow \mu\nu$. The muon selection and isolation criteria is the same used in the Higgs search analysis, detailed in 5.1.4. To select the $Z + jets$ events, the following selection is applied:

- Two identified isolated muons in the event,
- $P_T > 25$, $P_T > 10$ GeV/c,
- no hard jets, the Central Jet Veto is applied,
- events are required to be in the Z mass window, $|m_{\mu\mu} - 90| < 20$.

For the selection of $W + jets$, events are selected if they pass this requirements:

- One identified and isolated muon,
- $P_T > 25$ GeV/c,
- $|\pi\Delta\phi(\mu, MET)| < 1$, so-called acoplanarity cut.

This cuts for the selection of Z and W events are proposed as baseline and they will be tuned with a dedicated study in data.

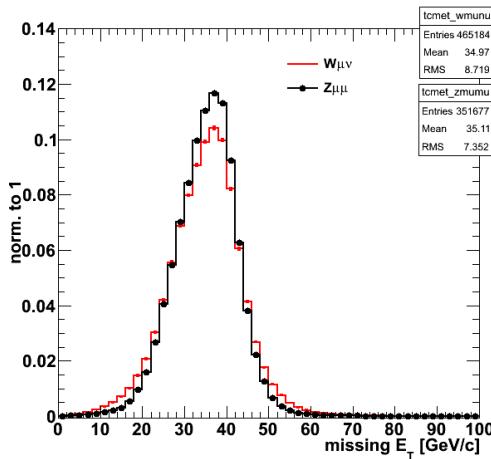


Figure 7.5: - Track Corrected missing E_T distribution from $W \rightarrow \mu\nu$ Monte Carlo events and $Z \rightarrow \mu\mu$ 'data' events where one of the muons has been dropped.

The missing E_T distribution of the $W + jets$ events is then compared to the one

calculated from $Z \rightarrow \mu\mu$ events as follows:

$$MET_Z = \sqrt{(P_X^\mu + MET_X)^2 + (P_Y^\mu + MET_Y)^2} \quad (7.6)$$

The Z kinematics are further re-scaled by the ratio of the invariant masses, m_W/m_Z . Figure 7.5 shows a fair agreement between the two resulting distributions, proving that the $Z + jets$ sample can be used as a benchmark to characterize the missing E_T from data. The study was performed using also Calo MET and Particle Flow MET, and different Monte Carlo generators leading to comparable results.

Another method to check the behavior of the missing E_T in events with real missing E_T in order to validate the Monte Carlo consist of selecting $t\bar{t}$ events going to $2be\mu2\nu$, asking explicitly for two tight jets, making the selection almost background free. From the distribution of real and simulated missing E_T a systematic error can be estimated. In figure 7.6, the missing E_T distribution for $t\bar{t}$ events decaying to $e\mu$ and two jets, and the rest of backgrounds is shown [45].

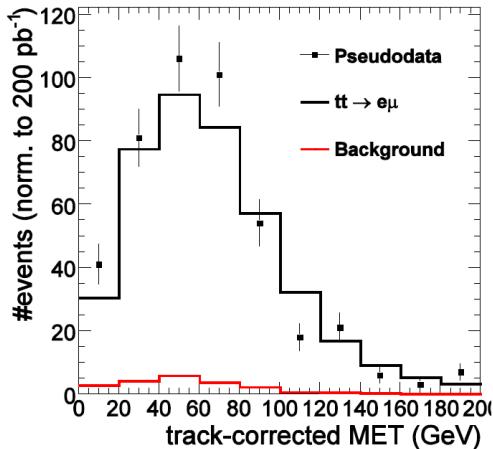


Figure 7.6: - missing E_T distribution for $t\bar{t}$ events (black) and backgrounds (red) in the dileptonic $e\mu$ channel, number of events normalized to $200 pb^{-1}$

7.4 WW Cross-section measurement

In the present run of the LHC at $\sqrt{s} = 7$ GeV, different physics channels will become accessible as the luminosity increases. At this centre of mass energy, Higgs studies can not be performed until a substantial amount of luminosity is collected. However, the continuum WW production decaying into two muons and two neutrinos has a larger cross-section and it can be studied with luminosities of the order of $100 pb^{-1}$. Using the same physics objects and a similar final state, the cross-section of the WW can be determined prior to the application of the $H \rightarrow WW^*$ analysis in real data.

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

The $WW \rightarrow 2\mu 2\nu$ and $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ processes present the same final state at leading order, with two prompt isolated muons coming from W decays, no hard jet activity and missing E_T . The differences arise in the production mechanisms, as the Higgs to WW^* is mainly produced via gg -fusion, while WW is mostly qq -induced. This difference is reflected in the kinematics and in the initial and final state radiation. It was discussed previously in 4.3 how the muons coming from Higgs decays tend to be close together in the transverse plane due to spin correlations, whereas muons coming from WW tend to be back to back. The study of the WW background can be based on the Higgs to WW^* analysis using the same physical observables, using its decay to two muons. This kind of approach allows for a better knowledge of the WW background to complement the Higgs study and has been extensively studied [122; 123], performed using exactly the same physics objects and samples that the Higgs to WW^* .

In figures 7.7 and 7.8, the kinematical variables for HWW and WW are presented after jet veto.

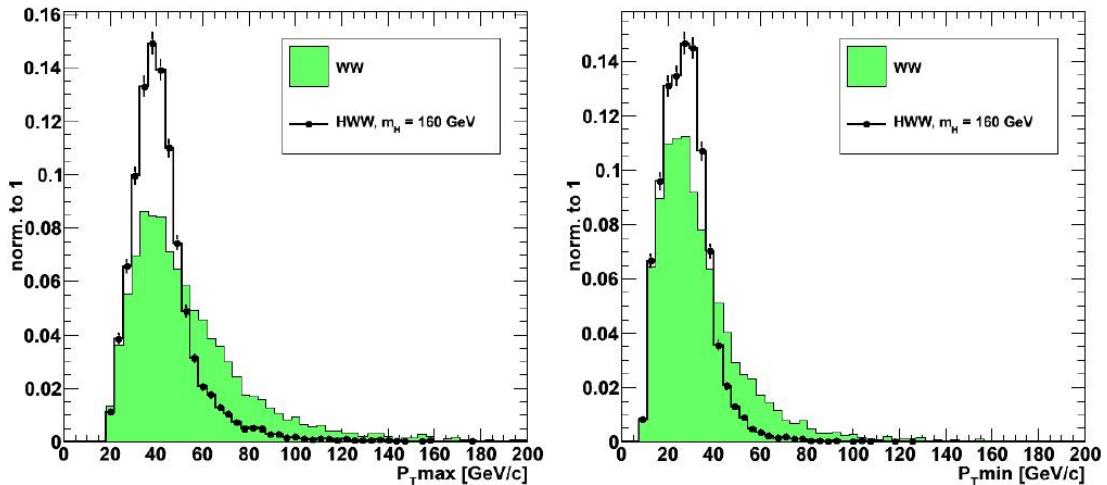


Figure 7.7: - P_T^{max} (left) and P_T^{min} (right) for Higgs to WW^* and WW , normalized to unit area, $m_H = 160 \text{ GeV}/c^2$.

In the analysis performed for the WW channel events are selected if they present two isolated muons (see 5.1.4), asking for the P_T of the second muon to be higher than 20 GeV/c , no jets are present in the final state (see 6.3) and:

- Missing $E_T > 50 \text{ GeV}$
- $m_{\mu\mu} < 76 \text{ OR } m_{\mu\mu} > 106 \text{ GeV}$

Additionally a soft-muon veto to reduce $t\bar{t}$ is also applied, following the definition given in 6.9.1.2, without effect in the signal efficiency. The expected number of events after the selection are presented in table 7.4.

If the Higgs exists, the signal to background ratio of the WW analysis will be affected.

7.4 WW Cross-section measurement

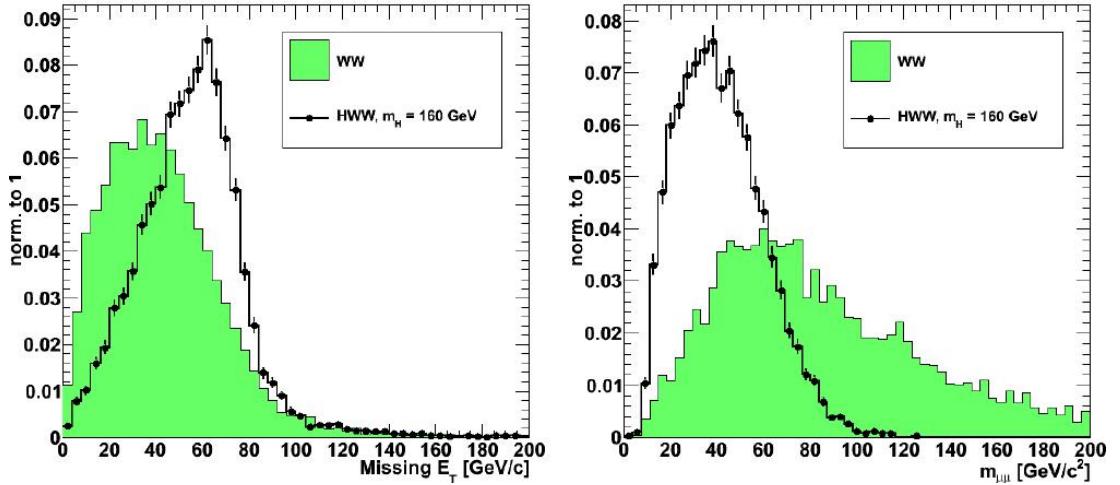


Figure 7.8: - missing E_T (left) and Invariant mass of the muons (right) for Higgs to WW^* and WW , normalized to unit area, $m_H = 160 \text{ GeV}/c^2$.

	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Number of expects events (all cuts)	47	1.2	1.6	3.3	12	0	17

Table 7.4: Number of expected events for the WW analysis based on sequential cuts.

The effect will be more evident for Higgs with high cross-section, for instance, for a mass of $m_H = 160 \text{ GeV}/c^2$, around 22 events are expected at this luminosity. The expected number of events for the WW analysis, with and without Higgs, $m_H = 160 \text{ GeV}/c^2$ is presented in table 7.5. It will be possible then to observe the Higgs boson

	Signal	Backgrounds	Backgrounds + Higgs
Number of expected events (all cuts)	49	35.1	57.3

Table 7.5: Number of expected events for the WW analysis based on sequential cuts, with and without Higgs, $m_H = 160 \text{ GeV}/c^2$.

in the WW analysis as an excess in the background, that can be up to two times for some masses.

The final distribution of the $\Delta\phi_{\mu\mu}$ variable is presented in Figure 7.9 after all cuts for the rest of the backgrounds, WW and Higgs to WW^* . Two mass hypothesis are shown: $m_H = 160 \text{ GeV}/c^2$, and $m_H = 130 \text{ GeV}/c^2$ in red.

The contribution of the gluon-fusion induced WW processes, $gg \rightarrow WW$, has been also studied. Produced by the same mechanism, the gluon-fusion WW events present

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

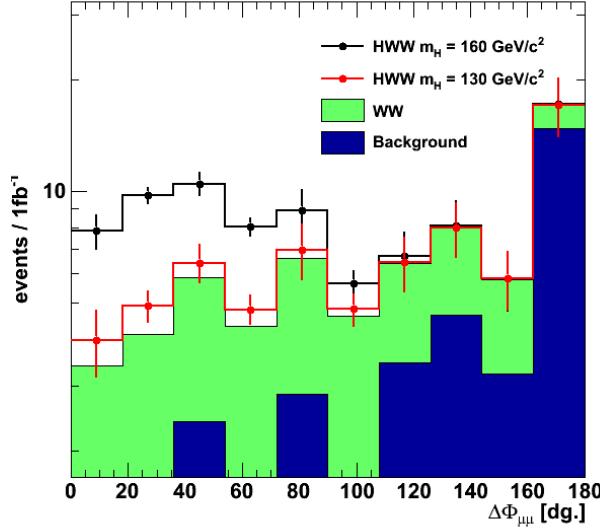


Figure 7.9: $-\Delta\phi_{\mu\mu}$ after all cuts for WW and the rest of backgrounds including Higgs, $m_H = 160 \text{ GeV}/c^2$ and $m_H = 130 \text{ GeV}/c^2$ (red).

kinematics more similar to the Higgs to WW^* events, as can be seen in figures 7.10 and 7.11. However, the cross-section for dileptonic WW produced via gluon fusion is $\sim 0.15 \text{ pb}$, that is, around 4 % of the qq .

Including the gg -induced process, the expected number of events after the WW cuts can be consulted in table 7.6. The distribution of $\Delta\phi_{\mu\mu}$ including the gg contribution

	$qq \rightarrow WW$	$gg \rightarrow WW$	WW_{total}
Number of expected events (all cuts)	47	2.96	49.96

Table 7.6: Number of expected events for the WW sequential analysis.

is presented in Figure 7.12.

With 100 pb^{-1} in the 7 TeV run, the expected number of events for qq and gg WW signal is ~ 5 , and the number of expected background events is ≤ 4 , in the presence of a Standard Model Higgs boson with a mass of $160 \text{ GeV}/c^2$ the number of background events goes up to ~ 6 . With a perfect control of the background contributions in the WW analysis, this channel could be used to observe the presence of Higgs to WW^* processes.

7.4 WW Cross-section measurement

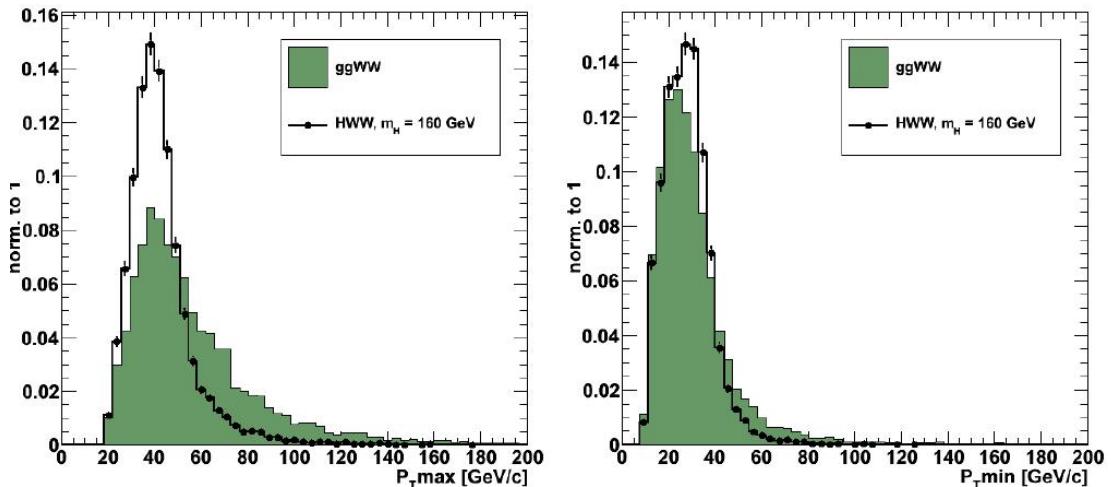


Figure 7.10: - P_T^{\max} (left) and P_T^{\min} (right) for gluon fusion WW and Higgs to WW^* normalized to unit area, $m_H = 160 \text{ GeV}/c^2$.

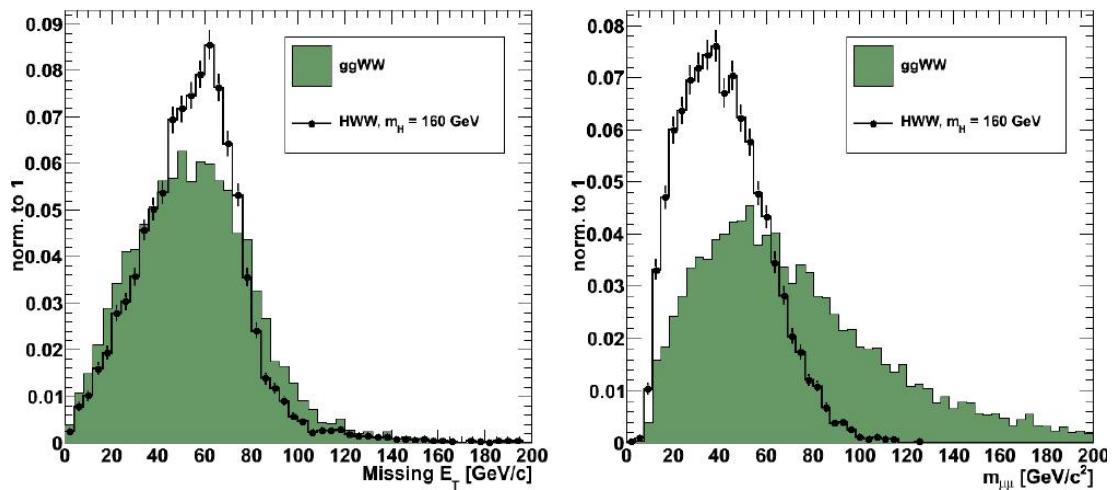


Figure 7.11: - missing E_T (left) and Invariant mass of the muons (right) for gluon fusion WW and Higgs to WW^* normalized to unit area, $m_H = 160 \text{ GeV}/c^2$.

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

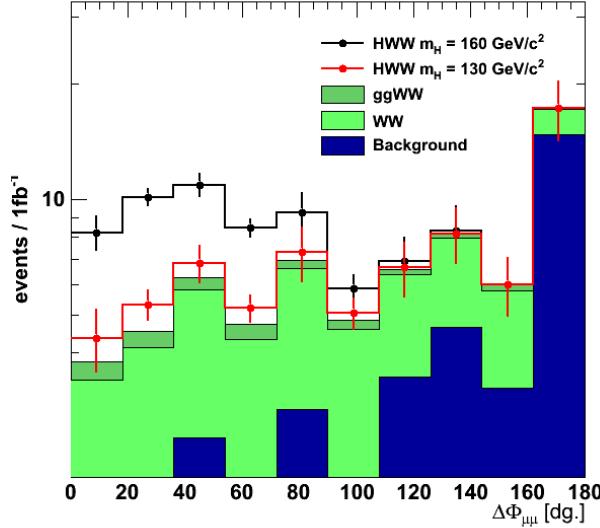


Figure 7.12: - $\Delta\phi_{\mu\mu}$ after all cuts for gg and qq induced WW and the rest of backgrounds including Higgs, $m_H = 160 \text{ GeV}/c^2$ and $m_H = 130 \text{ GeV}/c^2$ (red).

7.5 Systematic Uncertainties

Several sources of systematic uncertainties have been studied. The use of Monte Carlo to estimate the expected signal and background rates led to two types of systematic errors, experimental and theoretical, which may affect the normalization and shapes of signal and backgrounds. The systematic sources that were taken into account in the analysis will be described in the following sections. Most of them will be estimated in real data, and are expected to improve with the luminosity collected.

7.5.1 Luminosity

The uncertainty related to the luminosity coming from the LHC is set to about 20%. In the case of CMS, the systematic related to the luminosity can be studied from data using the $Z \rightarrow \mu\mu$ and ee decays of $Z + jets$. Different studies establish this value around 5% for this systematic uncertainty [62].

7.5.2 Muon identification and Trigger efficiencies

Lepton identification and Trigger efficiencies can be measured from data using for instance the tag and probe method. The $Z \rightarrow \mu\mu$ decay provides an unbiased and high-purity muon sample in which trigger and kinematical cuts efficiencies can be studied. One of the muons, the tag, has to pass all the muon identification criteria, while the other, the probe, just has to pass some loose identification criteria related to the efficiency under study. Apart from that, to ensure the purity of the sample, the invariant

mass of the two muons has to be inside a relatively narrow window around the mass of the Z boson. This is one of the studies that can be performed in the early data-taking scenario, with the first collected pb^{-1} of luminosity, and has already started in data.

The error coming from this source is expected to be small, just a few percent, contributing slightly to the total uncertainty.

7.5.3 Misalignment and Miscalibration scenarios

The effect of misalignment and miscalibration on the analysis in Monte Carlo is studied by repeating the analysis in misaligned and miscalibrated scenarios. Such a study was done as described in section 4.4, using special samples of $t\bar{t}$ events in the three misalignment and miscalibration scenarios provided by CMS: 0, 10 and 100 pb^{-1} , leading to around 2% of uncertainty in the muon reconstruction.

7.5.4 Missing E_T modeling

Concerning the modeling of the missing E_T , a 5% or better precision is expected after the detector comes to stable operations, as was shown in [131]. The effect in the final analysis should be relatively small providing that $Z + jets$ is a small component of the final background.

7.5.5 Jet reconstruction efficiency

The systematic uncertainty associated to the jet veto can be estimated using $Z + jets$ events. Using two simulated samples produced with different generators, treating one as it is, for example PYTHIA $Z + jets$ events, and the other one as if it was real data, in this case $Z + jets$ POWHEG or Madgraph. The method consist on the study of the variation of the relative efficiency of the Jet Veto as a function of the E_T threshold of the considered jets. The relative efficiencies are presented in table 7.7. The systematic

Jet E_T threshold (GeV)	$Z + jets$ (PYTHIA)	$Z + jets$ (Powheg)	$Z + jets$ (MadGraph)	$Z + jets$ (Sherpa)
[0, 10]	0.24	0.28	0.25	0.34
[0, 15]	0.46	0.49	0.48	0.52
[0, 20]	0.62	0.62	0.64	0.62
[0, 25]	0.72	0.71	0.73	0.69
[0, 30]	0.79	0.78	0.79	0.74
[0, 35]	0.83	0.82	0.83	0.79
[0, 40]	0.87	0.86	0.87	0.82
[0, 45]	0.90	0.88	0.89	0.85

Table 7.7: Relative efficiencies for different E_T thresholds and Monte Carlo generators.

uncertainty obtained, between 1% and 6%, is be extracted from the difference of the

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

relative efficiencies of the Central Jet Veto for jets with $E_T > 30$ GeV.

7.5.6 Jet energy scale

The uncertainty due to jet energy scale, JES, has been studied using $t\bar{t}$ events, as a substantial hard jet activity is expected in these processes. The systematic is calculated by varying the energy of all jets in the event by $+/- 7\%$ [132]. It is also possible to combine Jets with missing E_T , to determine how the missing E_T affects the systematic uncertainty. Calorimetric Jets and missing E_T are used in the study, and for each energy shift, the systematic error is calculated for events with a cut in the missing E_T or in the missing E_T corrected with jets (Type I). The Type I missing E_T is directly affected by the energy shift as it is corrected by all the jets in the event with $E_T^{RAW} > 20$ GeV.

Three kind of events are then selected:

- events with 2 muons
- events with 2 muons and Missing $E_T > 50$ GeV
- events with 2 muons and Missing $E_T^{TypeI} > 50$ GeV

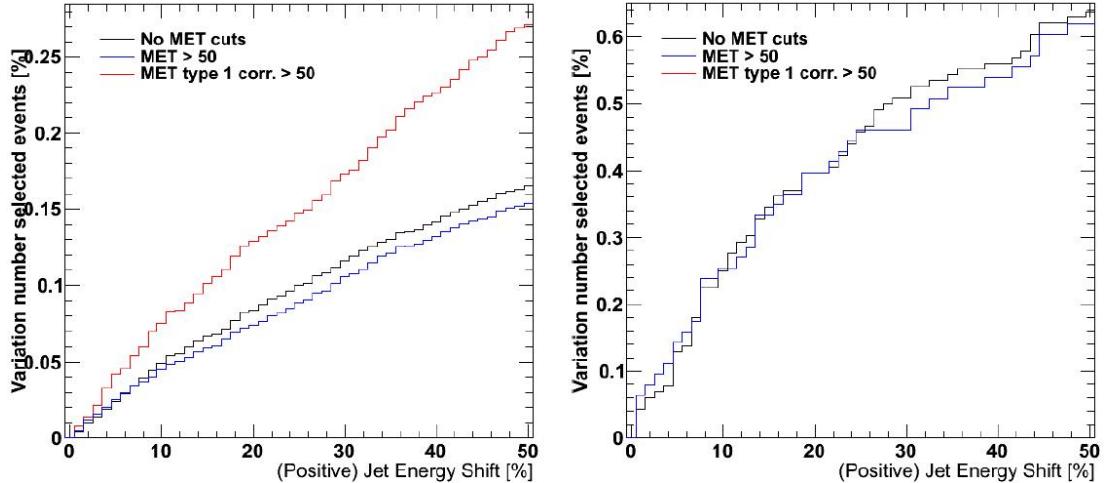


Figure 7.13: - Systematic associated to JES (%) as a function of the Jet energy shift, for positive shifts, from 0 to 50%, for $t\bar{t}$ events with two jets (left) and no jets (right), the blue line represents the systematic considering a MET cut in 50 GeV and the red line represents a cut in the Jet-Corrected missing E_T (Type I corrections)

This method has been used for studies of the $t\bar{t}$ background, requiring two or more hard jets in the event, and also for Higgs studies, when events with jets are vetoed. When no jets are present in event the systematic behaves in a different way, and the effect of the Type 1 corrections, is as expected, negligible.

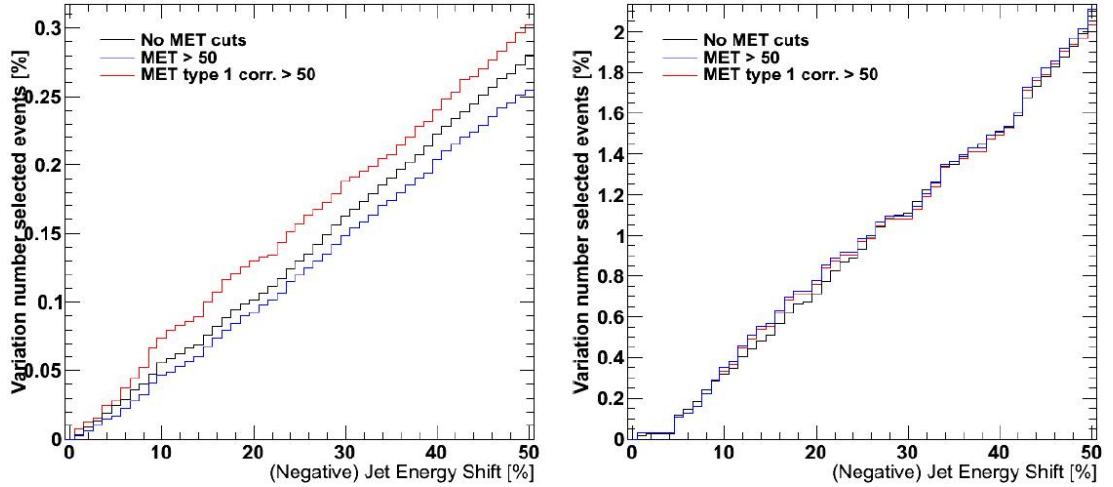


Figure 7.14: - Systematic associated to JES (%) as a function of the Jet energy shift, for negative shifts, from -50% to 0% , for $t\bar{t}$ events with two jets (left) and no jets (right), the blue line represents the systematic considering a MET cut in 50 GeV and the red line represents a cut in the Jet-Corrected missing E_T (Type I corrections)

The effect can be seen in figures 7.13 and 7.14, where the systematic is shown as a function of the Jet energy shift for $t\bar{t}$ events with two jets ($t\bar{t}$ signal region) and no jets (Higgs signal region), with and without cuts on the missing E_T . The different values of the systematic for a 7% energy shift for events with two or more jets are given in table 7.8: And the different values of the systematic for events with no central jets

2-jet			
sys	-	$MET > 50$	$MET (TypeI) > 50$
-7%	0.03	0.02	0.04
$+7\%$	0.03	0.03	0.05

Table 7.8: Systematic uncertainties for an energy shift of 7% in events with two or more jets.

are presented in table 7.9: The effect of the energy shift is stronger for events with no central jets than in the $t\bar{t}$ signal region. For high energy shifts the influence of the negative shift is stronger than the influence of the positive. The systematic has been estimated as the highest value for the given shift. Another estimation of the systematic uncertainty associated to the JES would be for instance the study of the how the acceptance is affected by the variation of the JES by $\pm 1\sigma$.

The total uncertainty obtained from this method is $\sim 15\%$, and this value will be propagated in the selection. Note that the effect on the background is larger than on

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

0 – jet			
sys	-	$MET > 50$	MET (<i>TypeI</i>) > 50
−7%	0.15	0.13	0.13
+7%	0.14	0.15	0.15

Table 7.9: Systematic uncertainties for an energy shift of 7 % in events with no central jets.

	$t\bar{t}$ Small ISR/FSR	$t\bar{t}$ Default ISR/FSR	$t\bar{t}$ Large ISR/FSR
$\mu\mu$	1404	1402	1398
<i>Jet Veto</i>	30.4 (2%)	30.3 (2%)	30.6 (2%)
<i>MET</i>	17.5 (57%)	17.3 (57%)	17.0 (55%)
$\Delta\Phi_{\mu\bar{\mu}}$	3.7 (21%)	3.3 (19%)	3.1 (18%)
$m_{\bar{\mu}}$	2.1 (56%)	2.2 (65%)	1.9 (61%)
P_T (highest)	0.7 (36%)	1.0 (46%)	0.8 (40%)
P_T (lowest)	0.29 (39%)	0.45 (44%)	0.42 (53%)

Table 7.10: Expected number of events after each cut of the analysis selection for $t\bar{t}$ samples with three different values of ISR/FSR.

the signal.

7.5.7 ISR/FSR

To study the contribution of the Initial and Final state radiation (ISR/FSR) to the uncertainty, three $t\bar{t}$ samples were produced using a parametrized detector simulation assuming:

- Small ISR/FSR: ISR=FSR=1.
- Large ISR/FSR: ISR=FSR=4.
- Default ISR/FSR: ISR=2.5, FSR=2.5.

While the default values of ISR and FSR used for the Monte Carlo production in CMS are $ISR = 2.5$ and $FSR = 4$. In table 7.10 the expected number of events after each cut of the selection for the three different samples are presented. The effect of the small ISR/FSR is more important, but overall the expected systematic is of the order of 1% at jet veto level.

7.5.8 PDF estimation

The uncertainty used in this work has been obtained from results described in [133; 134; 135].

Source	Signal (%)	Background (%)
Luminosity	10	10
Lepton & trigger efficiencies	4	4
Muon Miscalibration and misalignment	2	2
Electron Miscalibration and misalignment	4	4
missing E_T modeling	1	2
Jet reconstruction	3	8
PDF uncertainties	5	5
$t\bar{t}$ cross-section	—	18
WW cross-section	—	15
$WZ/ZZ/Wt$ cross-sections	—	10
DY cross-section	—	5
$W+jets$ cross-sections	—	100
MC statistics	5	10

Table 7.11: Summary of all systematic uncertainties. Relative contribution for signal and background.

7.5.9 Normalization for backgrounds

The systematic uncertainty related to the background normalization is relevant in this analysis as the background contribution is relatively high at the final level. The normalization of three main backgrounds, $t\bar{t}$, WW and $Z+jets$, has been addressed in section 7.1. The normalization of the $W+jets$ background is affected by the limited amount of Monte Carlo QCD events, as also discussed in section 7.1. Finally 5% of uncertainty for the Drell-Yan background and 10% for the di-boson have been obtained [65].

7.5.10 Monte Carlo statistics

The limited Monte Carlo statistics induces a systematic error that has to be taken into account 4.5.2. The contribution on is estimated to be around of $\sim 5\%$ for the signal and roughly $\sim 10\%$ for the backgrounds.

A summary of the different systematic uncertainties is shown in table 7.11. The overall relative error depends on the considered Higgs mass, but it is set to be around 15% for the signal and 30% for the background. The study of the systematics will be then done in real data and the values are expected to improve with the luminosity collected.

To model and control the main backgrounds, specific data driven methods have been proposed and tested in this chapter. The main sources of systematic errors that will affect the analysis were studied.

In the next chapter, significance and exclusion limits for different centre of mass ener-

7. METHODS TO ESTIMATE THE BACKGROUND FROM DATA AND SYSTEMATIC UNCERTAINTIES

gies are estimated using the signal and background yields obtained in chapter 6 and the systematic uncertainties discussed here, to estimate the sensitivity of this channel.

8

Results

A search for the SM Higgs boson using the $gg \rightarrow H \rightarrow WW^* \rightarrow \mu\nu\mu\nu$ channel has been presented and applied to different scenarios of centre of mass energy and integrated luminosities at the LHC. As a result, signal and background yields were obtained, and considering the systematic uncertainties discussed in the previous chapter, significance and exclusion limits on the production cross-section of the SM Higgs can be derived.

8.1 Limits and Significance

A large variety of methods exists for expressing the statistical significance of a possible signal in the presence of background. The choice is to use the same likelihood ratio [136] technique as at LEP and the Tevatron [137; 138], the Modified Frequentist approach. Hence, to evaluate the 95% confidence level (C.L.) exclusion limit, 3σ evidence and 5σ discovery potential, the likelihood-ratio estimator S_L is used:

$$S_L = 2\ln(L_{s+b}/L_b), \quad (8.1)$$

where L_{s+b} is the value obtained in the full signal-plus-background confidence level, and L_b is the value from the background-only confidence level. Systematics are taken into account by performing a bayesian pseudo-integration assuming they are Gaussian distributed.

The predicted distributions of Q are integrated to obtain the p -values: $1 - CL_b = 1 - \mathcal{P}_b(Q \leq Q_{observed})$ and $CL_{s+b} = \mathcal{P}_{s+b}(Q \leq Q_{observed})$. These measure the compatibility of the observed data with the two hypotheses, where \mathcal{P}_b and \mathcal{P}_{s+b} are the probabilities for a single experiment to obtain a value of Q smaller than or equal to the observed value, given the background or the signal-plus-background hypotheses. The estimator of the expected 95% CL upper limit is calculated performing toy Monte Carlo experiments.

The significance as well as the upper limits on the Higgs production cross-section, together with the luminosity needed for exclusion have been calculated. The Bayesian

8. RESULTS

method [139] was also exercised, found to agree with the Modified Frequentist within 10 % or better.

8.2 Results at a centre of mass energy of 14 TeV

The results of the analysis performed in this thesis have been approved several times by the collaboration. Public results for 14 TeV can be consulted in [64] and [65].

The number of signal and background events are presented in this section for var-

	HWW ($160 \text{ GeV}/c^2$)	WW	WZ	ZZ	tW	$t\bar{t}$	$W + \text{jets}$	DY
ee preselection	81	258	277	122	215	2372	150	15340
Final selection	20.1	13.7	1.7	0.0	0.0	4.3	2.0	1.0
$\mu\mu$ preselection	143	459	226	184	461	5128	177	39814
Final selection	21.0	10.0	0.1	0.0	0.3	6.0	0.1	0.6
$e\mu$ preselection	195	632	42	4	604	6008	453	471
Final selection	31.5	16.0	0.3	0.0	1.4	7.0	6.3	0.0

Table 8.1: Number of expected events for an integrated luminosity of 1 fb^{-1} after preselection and after the final kinematical selection, for the signal and for the backgrounds.

ious Higgs mass values. Results are normalized to 1 fb^{-1} of integrated luminosity. For the $m_H = 160 \text{ GeV}/c^2$ hypothesis, the number of signal and background expected events for the three final states in the sequential cut based analysis after the preselection and after the final kinematical selection are shown in Table 8.1. The number of expected events for various representative Higgs mass values is shown in Table 8.2.

$H \rightarrow WW \rightarrow 2l2\nu$	m_H [GeV/c^2]				
	130	150	160	170	190
ee preselection	24	58.7	81	82.7	58
Final selection	6.6	12.5	20.1	21.2	11.0
$\mu\mu$ preselection	49	108	143	144	96
Final selection	8.7	9	21.0	17	5.6
$e\mu$ preselection	61	150	195	202	139
Final selection	11.0	20.6	31.5	33.1	12.2

Table 8.2: Number of expected signal events for an integrated luminosity of 1 fb^{-1} after preselection and after the final kinematical selection for different masses of the Higgs boson in the three final states.

The expected significance of an event excess under the assumption of the presence of a Higgs boson are shown in figure 8.1 for an integrated luminosity of 1 fb^{-1} for the sequential cut based and for a multivariate analysis, based on a Neural Network.

In Figure 8.2 the median of the 95% CL exclusion limit distributions for every mass hypothesis is plotted versus the mass of the Higgs boson considered. The error bars

8.2 Results at a centre of mass energy of 14 TeV

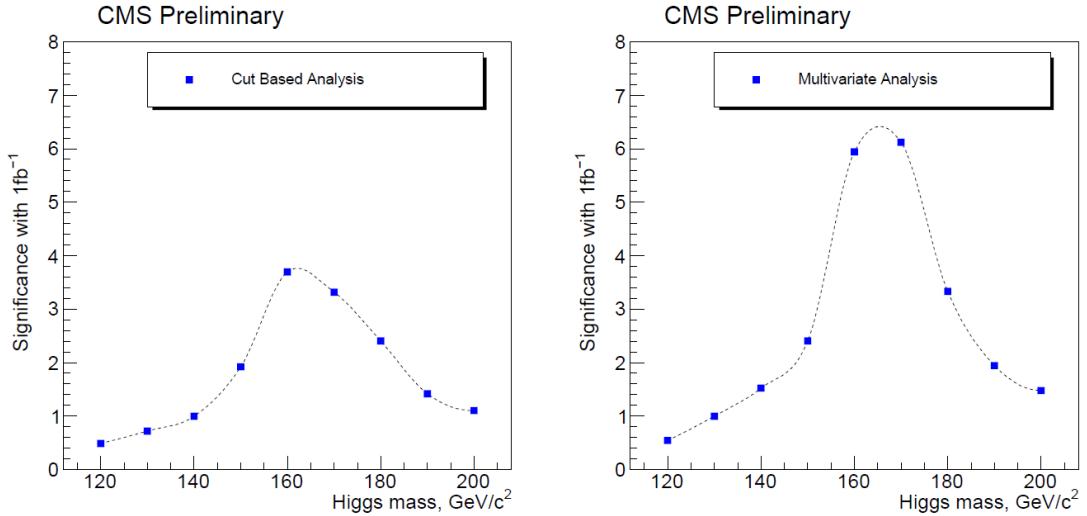


Figure 8.1: - Left: The expected significance of an event excess under the assumption of a Higgs boson presence for the cut based analysis and for an integrated luminosity of 1fb^{-1} . Right: The expected significance of an event excess under the assumption of a Higgs boson presence for the NN analysis and for an integrated luminosity of 1fb^{-1} .

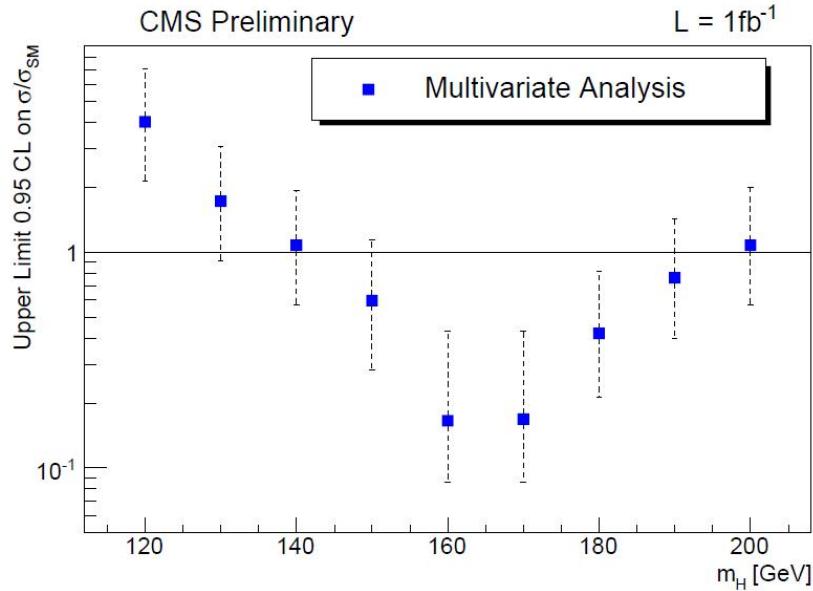


Figure 8.2: - The 95% confidence level upper exclusion limits obtained for each of the Higgs mass considered. No signal is assumed to be present

8. RESULTS

are obtained by searching for the two values which bound the 16% of the upper limit distribution on the right and on the left of the median.

The discovery of a SM Higgs boson with $m_H \sim 160 - 170 \text{ GeV}/c^2$ is possible for the CMS experiment using the Higgs to WW^* decay with 1 fb^{-1} of collected data at a centre of mass energy of 14 TeV. A 95% CL exclusion is possible in the mass range $140 - 200 \text{ GeV}/c^2$ using the same luminosity.

8.3 Projections at 10 TeV

Responding to the new considered centre of mass energy of the LHC announced in the Chamonix workshop on LHC startup in 2009 [67], the CMS collaboration made a combination and extrapolation to lower LHC energies in [140] using the $H \rightarrow WW^* \rightarrow 2l2\nu$ [65] and $H \rightarrow ZZ^* \rightarrow 4l$ [55] results at 14 TeV.

Figure 8.3 shows how cross-sections scale down for the main processes coming into

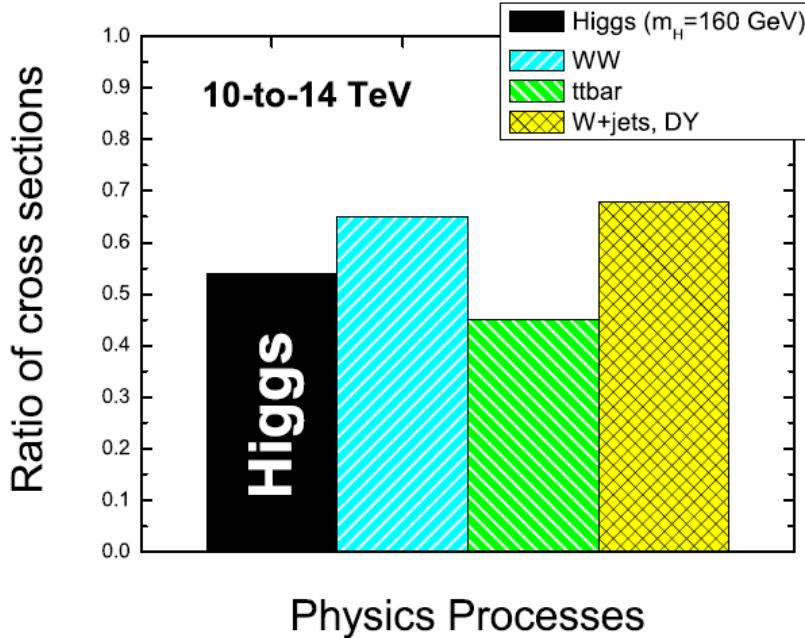


Figure 8.3: - The change in cross-sections for signal and main backgrounds for the $H \rightarrow WW^* \rightarrow 2l2\nu$ channel from 14 to 10 TeV.

the $H \rightarrow WW^*$ analyses. The signal is more affected than the main background WW . The $t\bar{t}$ background also falls, but its contribution in the end of the analysis chain is not dominant, 30% or less. The W production is the least affected. Although the net contribution of $W + jets$ is small, it nevertheless presents a large systematic error of about a factor of two.

8.3 Projections at 10 TeV

The change in the detector acceptance has to be taken into account. The acceptance is expected to be better with decreasing centre of mass energy of as heavy objects tend to be less boosted. The effect was studied for the $H \rightarrow ZZ^*$ channel at the generator level since the acceptance is largely driven by the kinematic cuts on leptons. The effect was found to be of a second-order scale and amounts to only a few percent. The jet veto and missing E_T cuts applied in the analysis cannot be realistically evaluated at the generator level and would require a full detector simulation. However, based on the ZZ^* results, the acceptance cut changes are expected to be small and then neglected in the first approximation.

To evaluate the impact of switching the LHC energy from 14 to 10 TeV, the signal and background event counts from the 14 TeV analysis were scaled by the ratio of the cross-sections. Then, the exclusion limits were re-calculated with the new event count numbers. The final result is shown in Figure 8.4.

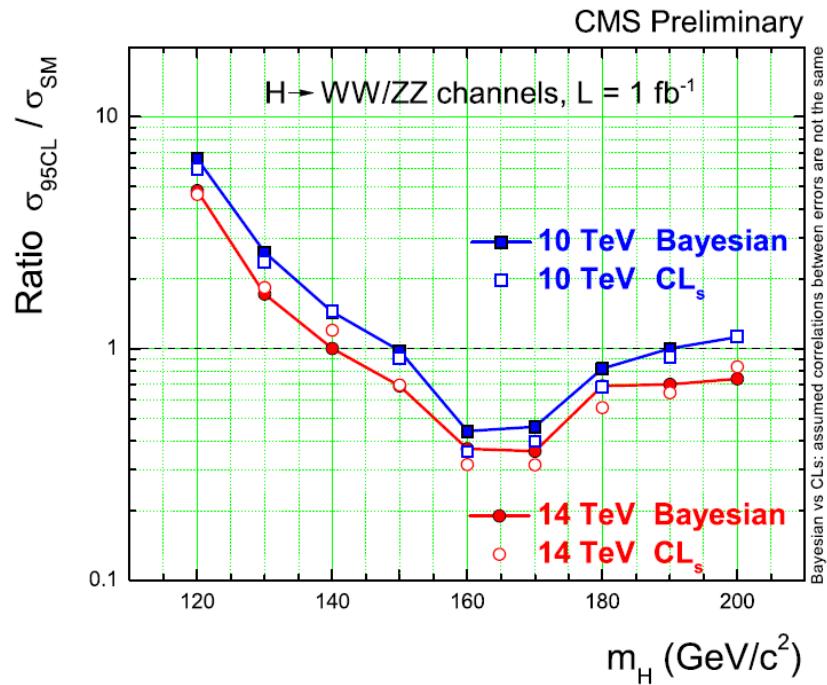


Figure 8.4: - Projected exclusion limits for the SM Higgs boson at 14 and 10 TeV centre of mass energies. $H \rightarrow WW^* \rightarrow 2l2\nu$ and $H \rightarrow ZZ^* \rightarrow 4l$ channels are included. Two different methodologies, Bayesian and CLs, were used for the combination and show a very reasonable agreement.

The effect of changing from 14 to 10 TeV roughly corresponds to a loss of a factor of 1.5 in sensitivity at a given luminosity, or equivalently, running at 10 TeV requires about 2 times larger integrated luminosity to reach the same sensitivity for SM Higgs. At 10 TeV, an exclusion of the SM Higgs boson with $m_H \sim 160 - 170 \text{ GeV}/c^2$ would

8. RESULTS

require 200 pb^{-1} .

8.4 Projections at 7 TeV

The latest results concerning the sensitivity of CMS to a potential SM Higgs discovery/exclusion [141] at 7 TeV are the projections made using the public results corresponding to the 14 TeV analysis, that were presented for this channel in chapter 6. The projections are based on re-scaling the event yields for signal and background by the ratio of the cross-sections from 14 to 7 TeV, including the changes in acceptance when applicable and with no modifications in the reconstruction performance and considering all the available Higgs channels. Complementary, the significance and exclusion limits have been calculated for the dimuon final state using simulated data at 7 TeV in B.

The ratio of parton luminosities for the LHC operating at 7 TeV and 10 TeV, as a function of the invariant mass of the produced final state can be seen in Figure 8.5.

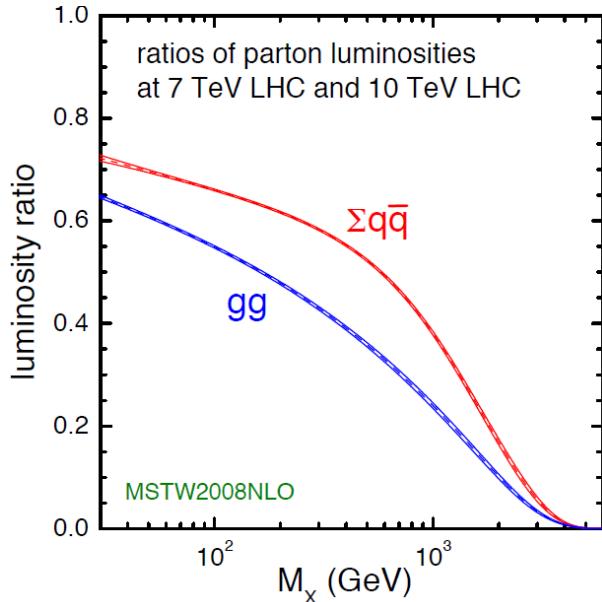


Figure 8.5: - Ratio of parton luminosities for the LHC operating at 7 TeV and 10 TeV, as a function of the invariant mass of the produced final state.

These projections were done for Standard Model Higgs analyses using the results from 7 channels,

- $H \rightarrow \gamma\gamma$
- $H \rightarrow WW^*$ ($\rightarrow ee, e\mu$ and $\mu\mu$)
- $H \rightarrow ZZ$ ($\rightarrow 4e, 4\mu$ and $2e2\mu$)

and the combination $H\gamma\gamma + HWW + HZZ$. A combination was also done including the MSSM channels, $bb\Phi \rightarrow \tau\tau$ and light charged Higgs $tt \rightarrow (H^\pm b)(Wb) \rightarrow (\tau_{had}\nu b)(l\nu b)$, but these results will not be discussed in this thesis.

The 95% Confidence Level for the HWW channel was calculated using the combination of the three channels, ee , $e\mu$ and $\mu\mu$ for a multivariate approach. The expected exclusion limit as well as the $1-\sigma$ and $2-\sigma$ bands is presented in figure 8.6.

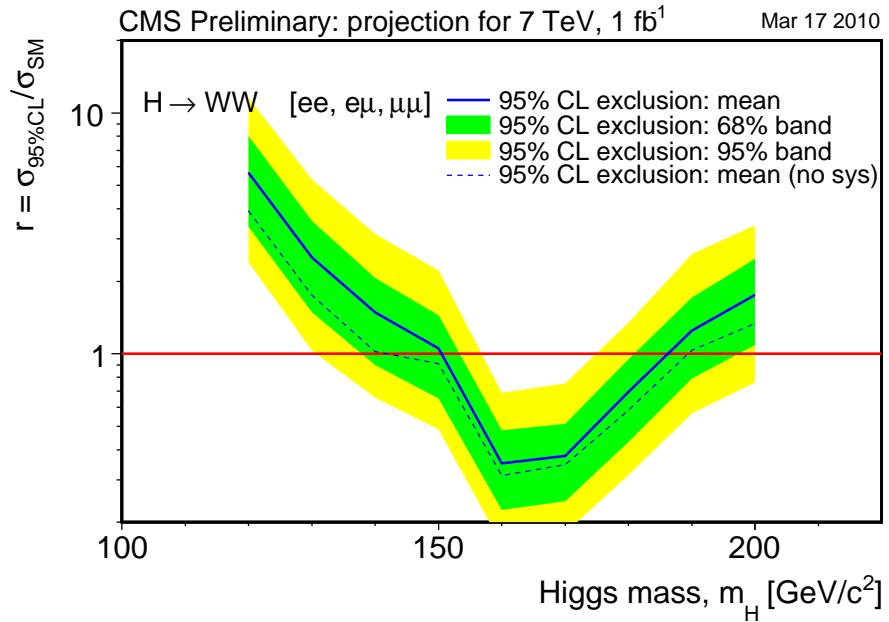


Figure 8.6: - Projected exclusion limit for the combination of the three final states in the multivariate $H \rightarrow WW^*$ analysis, from 14 to 7 TeV.

As seen in Figure 8.7, the search using the $H \rightarrow WW^*$ channel is expected to reach a discovery level sensitivity for the SM Higgs boson in the mass range $160 < m_H < 170$ GeV/c^2 .

The projected exclusion limit for the combination of $H\gamma\gamma + HWW + HZZ$, 7 channels, can be seen in figure 8.8, and in figure 8.9 focusing in the mass range between 100 and 200 GeV/c^2 .

An estimation of the sensitivity of a possible combination ATLAS-CMS can be done multiplying by a factor of 2 the number of expected events for signal and backgrounds, and that is shown in figure 8.10.

The expected range of exclusion for $1fb^{-1}$ of integrated luminosity at a centre of mass energy of 7 TeV coming from the projections made by scaling the 14 TeV yields by the

8. RESULTS

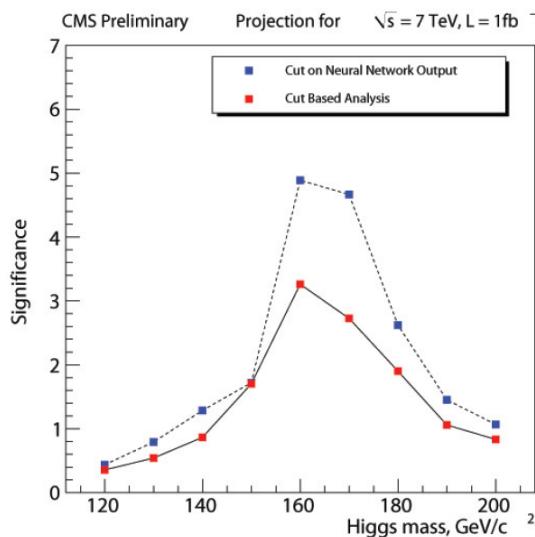


Figure 8.7: - Expected significance for the $H \rightarrow WW^*$ search, as a function of the SM Higgs boson mass.

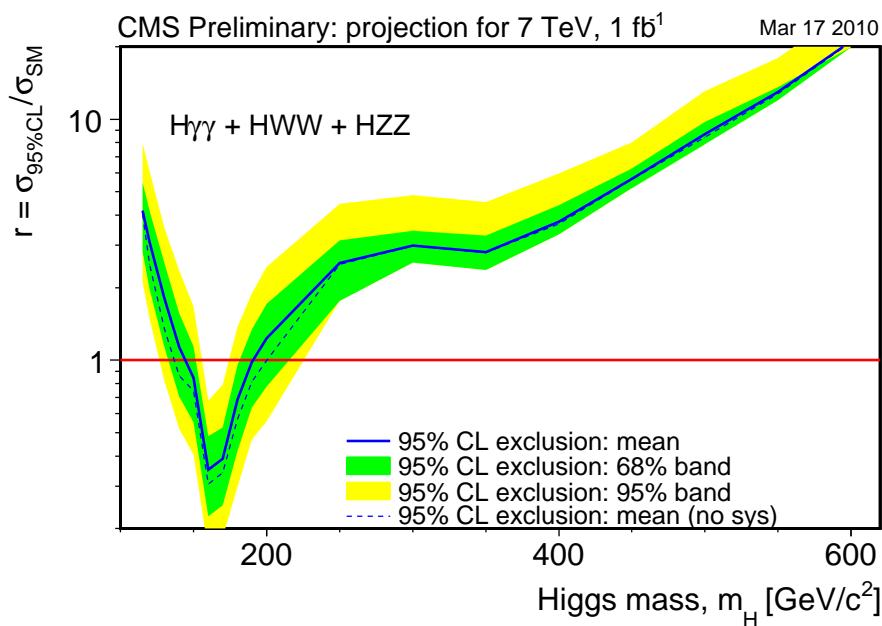


Figure 8.8: - Projected exclusion limit for the combination of the main Higgs analysis from 14 to 7 TeV in the whole Higgs mass range.

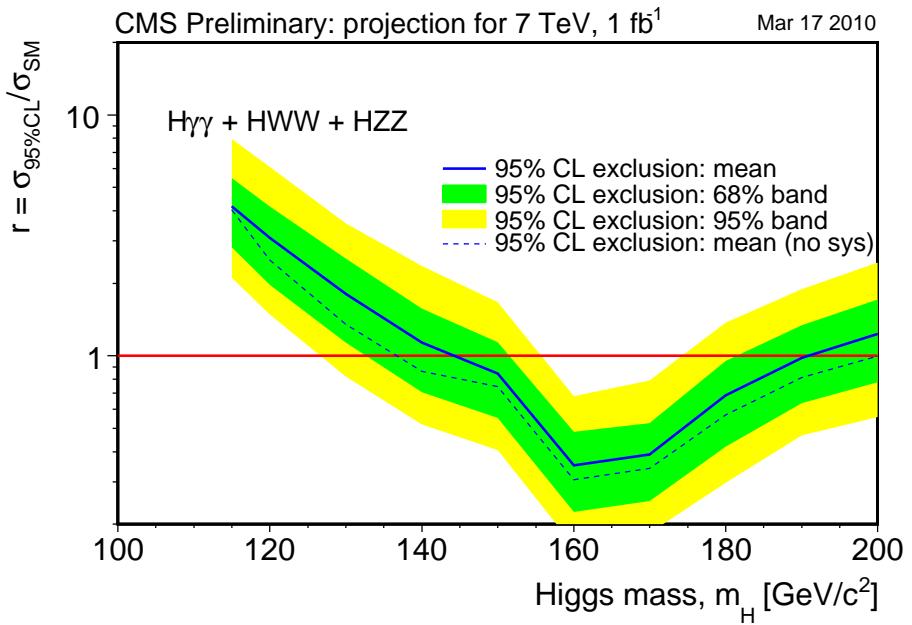


Figure 8.9: - Projected exclusion limit for the combination of the main Higgs analysis from 14 to 7 TeV in the mass range between 100 and 200 GeV/c^2 .

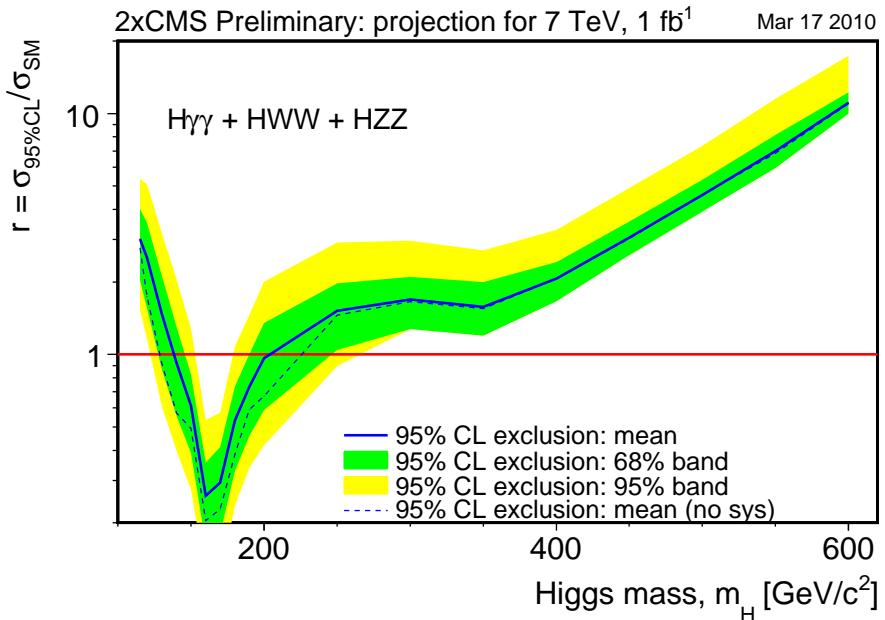


Figure 8.10: - Projected exclusion limit for the combination of the main Higgs analysis from 14 to 7 TeV in the whole Higgs mass range multiplied by a factor of 2.

8. RESULTS

ratio of the cross-sections for the $H \rightarrow WW^*$ channel is $150 < m_H < 185 \text{ GeV}/c^2$, and the search is expected to reach discovery level sensitivity for the SM Higgs boson in the mass range $160 < m_H < 170 \text{ GeV}/c^2$.

When combined together with the rest of channels studied within the CMS experiment, the expected exclusion limit for the SM Higgs at 7 TeV and with 1 fb^{-1} extends to $145 < m_H < 190 \text{ GeV}/c^2$.

Signal hypothesis significance and exclusion limits were calculated and the sensitivity to a possible SM Higgs boson discovery or exclusion is determined. For the design energy of 14 TeV with 1 fb^{-1} of integrated luminosity exclusion will be possible for a SM Higgs in the mass range $140 - 200 \text{ GeV}/c^2$, and discovery will be possible for $m_H \sim 160 - 170 \text{ GeV}/c^2$. At a centre of mass energy of 10 with 200 pb^{-1} of integrated luminosity, exclusion will be possible for a SM Higgs boson with $m_H \sim 160 - 170 \text{ GeV}/c^2$. Finally, with the expected integrated luminosity at the end of 2011, 1 fb^{-1} , at 7 TeV, it will be possible to exclude a mass range around $160 \text{ GeV}/c^2$, and when combining the results obtained for this channel with the rest of the Higgs searches performed in CMS, exclusion will be possible in the mass range $145 < m_H < 190 \text{ GeV}/c^2$

9

Conclusions

In this thesis, a complete strategy for the search of the Standard Model Higgs boson using the $H \rightarrow WW^* \rightarrow \mu\nu\mu\nu$ process in the CMS experiment at the LHC has been developed. The study was done using simulated data. The data collected recently by the CMS experiment has been also analyzed here, and some preliminary results have been obtained showing an the readiness and excellent behavior of the complete analysis chain. The clean signature of this channel is especially sensitive for Higgs masses around $160 \text{ GeV}/c^2$, and it is considered the most promising search channel at the LHC for the Standard Model Higgs boson in the mass range favored by the theory and previous data [3].

In the physics program of the CMS experiment at the LHC, the search for the Higgs boson is a central component, as this particle is the last missing piece of the Standard Model of particle physics. The experimental information gathered in the past few decades combined with precise predictions of the theory, have led to exclusion of the Higgs boson in the mass range up to $114.4 \text{ GeV}/c^2$, and around $160 \text{ GeV}/c^2$, and define a narrow mass window preferred: $115 \leq m_H \leq 148 \text{ GeV}/c^2$ [3].

The analysis presented in this work has been driven mainly for the need to extract a small signal from a large background. This background is composed by many physics processes that can be potentially misidentified as the signal: WW , $t\bar{t}$, $Z + jets$ and $W + jets$. A complete strategy has been specifically developed to understand and reduce the presence of events coming from those process in the final state. Other sources of background that have been considered and treated are WZ , ZZ and tW . The production cross-section for the background processes is several orders of magnitude larger than for Higgs signal, however, after the application of the different proposed selection, it has been shown that a signal to background ratio larger than one can be achieved for all the Higgs masses considered.

The work performed was done within the Higgs group of the CMS collaboration, and it has been adapted to the various expected LHC running conditions. Several studies

9. CONCLUSIONS

considering different centre of mass energies and integrated luminosities were carried out. The increase in luminosity needed to reach the same level of sensitivity for lower centre of mass energies have been determined. The cross-sections for signal and backgrounds are different for a centre of mass energy lower than the design energy of the LHC. Consequently, the muon selection and the determination of the physics observables have been optimized in order to allow this analysis to be robust and reach enough significance even with the expected integrated luminosity at the 7 TeV run.

All the available information provided by the CMS subdetectors: tracker, calorimeters and muon system has been used to select events that show two isolated muons in the final state with a high transverse momentum. The muon selection has been optimized to identify muons coming from W decays and the isolation criteria has been developed to mainly reject the QCD background. It has been shown as well that the muon selection can reduce the $W + jets$ background to a negligible rate at the end of the selection process.

A careful study of the different jet reconstruction algorithms and energy corrections in CMS has been performed, leading to a robust jet definition. By means of the Jet Veto, most of the $t\bar{t}$ background, characterized by jets coming from b decays, can be rejected as well as other backgrounds that present hadronic activity.

A set of kinematical variables have been selected due to their discriminating power against various background: the amount of missing E_T in the event, the angle between the muons in the transverse plane to the beam, $\Delta\phi_{\mu\mu}$, the invariant mass of the system of the muons, and the transverse momentum of the two muons. The missing E_T cut has been applied to reject events with no genuine missing E_T . The invariant mass requirement gets rid of events coming from Z decays, and the key variable to extract the small Higgs to WW^* signal from the vast background, especially WW , is $\Delta\phi_{\mu\mu}$. To account for the different kinematics of the signal events, determined by the Higgs mass hypothesis considered, the selections have been independently optimized for three different mass regions by maximizing the statistical significance in each step of the analysis.

Many other physics tools have been studied. The use of TrackJets and b-tagging was shown to improve the control of the $t\bar{t}$ events that remain after the Jet Veto, based essentially on calorimetric information. The use of multivariate analysis techniques has been proposed as a powerful alternative, especially in the low mass region, around $130\text{ GeV}/c^2$, when the signal extraction is more complicated due to the low cross-section and the challenging kinematics.

Control regions have been established to estimate the contribution of the main backgrounds in the different signal regions defined by the various Higgs mass hypothesis. The presence of two hard jets in the final state has been used to define the control region for $t\bar{t}$. For the definition of the control region for $Z + jets$, events with two muons

with an invariant mass compatible with the mass of the Z boson have been selected. In the case of the WW background, the definition of a control region is more difficult and has been determined by means of the $\Delta\phi_{\mu\mu}$ variable.

To better model the $W + jets$ background, a data-driven method has been designed and tested with Monte Carlo simulated events. The fake rate for events with one well identified muon and a 'fakeable' object is calculated in QCD and extrapolated to $W + jets$ final states. The result has been propagated to the rest of the selection process leading to a much better understanding of the background coming from this kind of sources. The performance of this method has been tested already in real data, showing a good agreement with the expectations.

The measurement of the WW cross-section has been proposed to be performed as the initial step of the study of the Higgs decay into WW^* . The WW processes have a higher cross-section and therefore will be accessible before. The understanding of this background is a main ingredient in the analysis and the similarities between the two processes allow to define a complementary study using the same physical objects. It has been shown that if the Higgs boson exists with a mass in the range under study in this work, the analysis for the measurement of the EWK WW cross-section will allow to observe the presence of the Higgs boson as an excess on the background, and for that, the $\Delta\phi_{\mu\mu}$ variable can be used at the end of the analysis chain, where the Higgs presence will be visible.

To conclude, a comprehensive study of all the sources of systematic uncertainties has been performed and several methods to determine them in real data have been developed.

	HWW ($160 \text{ GeV}/c^2$)	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	DY
ee Final selection	20.1	13.7	1.7	0.0	0.0	4.3	2.0	1.0
$\mu\mu$ Final selection	21.0	10.0	0.1	0.0	0.3	6.0	0.1	0.6
$e\mu$ Final selection	31.5	16.0	0.3	0.0	1.4	7.0	6.3	0.0

Table 9.1: Number of expected events for an integrated luminosity of 1 fb^{-1} and a centre of mass energy of 14 TeV after applying the final selection for a Higgs signal with $m_H = 160 \text{ GeV}/c^2$ signal and for the backgrounds.

The signal and background yields for the three final states of the Higgs to WW^* channel, including the $\mu\mu$ results obtained directly in the work developed in this thesis for an integrated luminosity of 1 fb^{-1} at a centre of mass energy of 14 TeV are shown in table 9.1. Significance and exclusion limits have been calculated using the final expected events for signal and background together with the corresponding uncertainties obtained in this study.

9. CONCLUSIONS

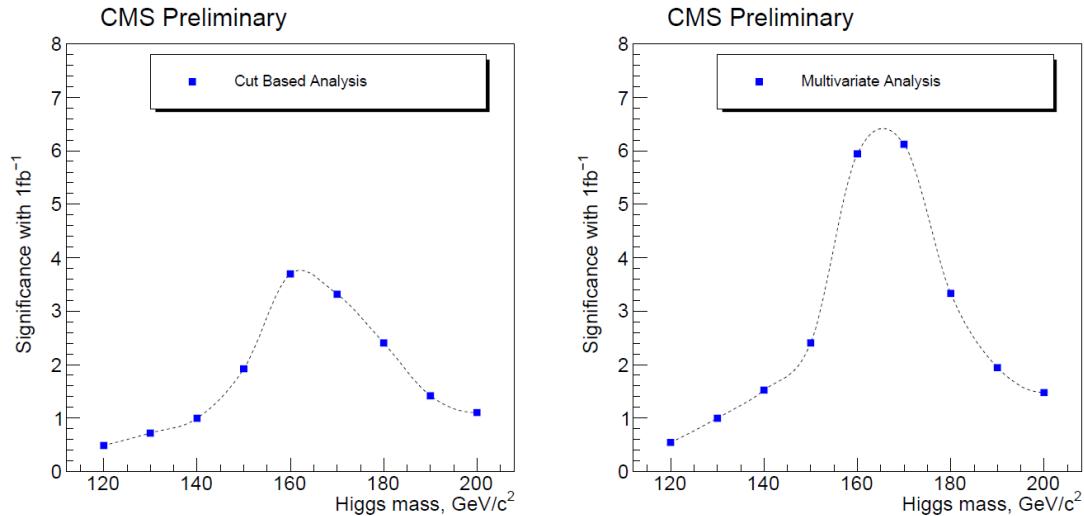


Figure 9.1: - Left: The expected significance of an event excess under the assumption of a Higgs boson presence for the cut based analysis and for an integrated luminosity of $1 fb^{-1}$ at a centre of mass energy of 14 TeV. Right: The expected significance of an event excess under the assumption of a Higgs boson presence for the NN analysis and for an integrated luminosity of $1 fb^{-1}$ at a centre of mass energy of 14 TeV.

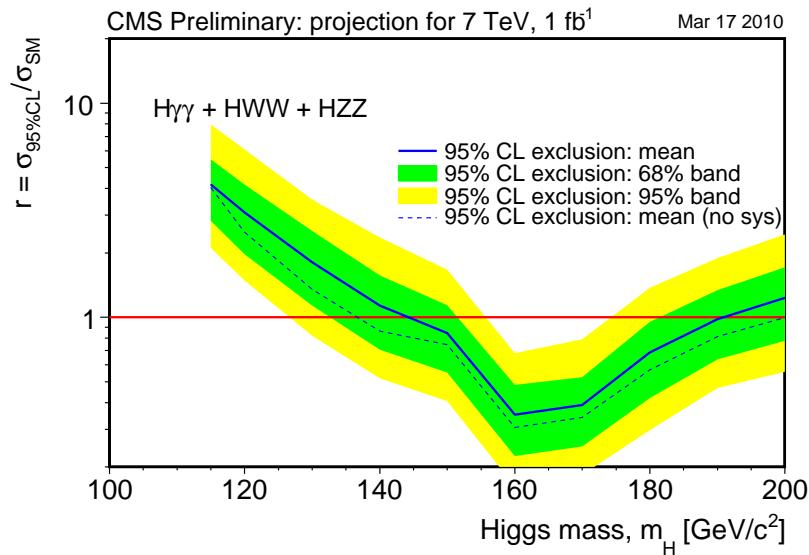


Figure 9.2: - Projected exclusion limit for the combination of the main Higgs analysis from 14 to 7 TeV in the mass range between 100 and 200 GeV/c^2 .

Sensitivity for discovery is reached for masses between 160 and 180 with an integrated luminosity of 1 fb^{-1} at the design collision energy of the LHC, as it is shown in figure 9.1. With the same integrated luminosity a solid exclusion can be reached in a wide mass range from 140 to 190 GeV/c^2 when combining the results obtained in this study with the other final states of the Higgs to WW^* decay mode. For a centre of mass energy of 10 TeV, the optimization done in the different aspects of the selection allows for an exclusion of the SM Higgs boson in the interval $m_H \sim 160 - 170 \text{ GeV}/c^2$ with 200 pb^{-1} of integrated luminosity.

The Higgs decay into WW^* is the most powerful search channel for the Higgs boson during the low energy run of the LHC. The analysis that has been presented in this thesis, optimized for a low centre of mass energy, can exclude a Higgs boson with a mass close to $160 \text{ GeV}/c^2$, with approximately 1 fb^{-1} of integrated luminosity at 7 TeV. When combining the result of the search presented here with the rest of the final states of the Higgs to WW^* decay, the exclusion extends from 150 to 185 GeV/c^2 and the channel reaches a discovery level sensitivity for the SM Higgs boson in the mass range $160 < m_H < 170 \text{ GeV}/c^2$ [141].

The results of this work combined together with the rest of Higgs search channels, led to an expected exclusion limit for the SM Higgs at 7 TeV and with 1 fb^{-1} that covers the Higgs mass range from 145 to 190 GeV/c^2 . Figure 9.2 shows the projected exclusion limit for the combination of $H\gamma\gamma + HWW + HZZ$ processes in the mass range between 100 and 200 GeV/c^2 .

Finally, combined projections for the three Standard Model signatures of the Higgs to $WW^* \rightarrow 2l2\nu$, $ZZ^* \rightarrow 4l$ and $\gamma\gamma$ have been determined, reaching a discovery level sensitivity for $m_H = 160 - 170 \text{ GeV}/c^2$ and an exclusion level sensitivity for the $m_H = 145 - 190 \text{ GeV}/c^2$ mass range. A Higgs boson mass $m_H \leq 500 \text{ GeV}/c^2$ can be excluded in models with four generations of quarks, regardless of how heavy their masses might be.

9. CONCLUSIONS

10

Resumen del trabajo realizado y Conclusiones obtenidas

El Modelo Estándar en Física de Partículas ha mostrado un excelente acuerdo con los resultados experimentales hasta la fecha pero aún está incompleto. No es capaz de dar explicación a fenómenos tales como la rotura de la simetría electrodébil, una de las cuestiones no resueltas de mayor importancia en la física de partículas contemporánea. El mecanismo de Higgs ofrece la respuesta más sencilla a la rotura de la simetría electrodébil dejando el Modelo Standard casi intacto e introduciendo una nueva partícula que debe ser detectada, el bosón de Higgs. Todas las propiedades del bosón de Higgs pueden ser predichas dentro del Modelo Estándar con la excepción de su masa, parámetro libre de la teoría.

Revelar el mecanismo físico responsable de la rotura de la simetría electrodébil, es decir, descubrir el Higgs, es uno de los principales objetivos del acelerador LHC (Large Hadron Collider). Si el mecanismo de Higgs es correcto y el bosón de Higgs existe, los experimentos del LHC tienen potencial para encontrarlo, cerrando así una larga búsqueda que comenzó años atrás y en la que han participado varios experimentos y aceleradores. La alta energía del centro de masas producida en colisiones protón-protón en el LHC junto con la elevada luminosidad esperada permitirán dar una respuesta definitiva a esta búsqueda. El programa de física del LHC cubrirá un amplio programa de Higgs, desde el Higgs del Modelo Estándar a sus extensiones en el MSSM, otros modelos de SUSY y búsquedas más allá del Modelo Estándar.

En la última etapa de LEP (Large Electron Positron), la combinación de los resultados de las búsquedas directas de Higgs en sus cuatro experimentos permitió establecer una sólida exclusión para masas por debajo de $114.4 \text{ GeV}/c^2$ a un nivel de confianza superior al 95% [1]. En la actualidad, las búsquedas de Higgs llevadas a cabo por los experimentos CDF y D0 del Tevatrón han alcanzado sensibilidad suficiente como para excluir el bosón de Higgs del Modelo Standard en un rango de masa entre 162 y 166 GeV/c^2 [2]. Un ajuste utilizando medidas de precisión electrodébiles permite establecer

10. RESUMEN DEL TRABAJO REALIZADO Y CONCLUSIONES OBTENIDAS

otras restricciones en el valor de m_H . La combinación de los límites teóricos y experimentales permite determinar el rango de masa en el que se desarrollarán las búsquedas en el LHC, y define el valor preferido por la teoría, entre 115 y leq 148 GeV/c^2 [3]. En este rango de masa, el modo dominante de producción de Higgs en el LHC es mediante fusión de gluones, y el principal modo de desintegración es $H \rightarrow WW^*$. Modos de producción y desintegración estudiados aquí.

En esta tesis se presenta un análisis prospectivo para la busca del bosón de Higgs de acuerdo con el Modelo Standard en el experimento CMS del LHC, centrándose en el proceso $gg \rightarrow H \rightarrow WW^* \rightarrow \mu\nu\mu\nu$. El canal de búsqueda presenta un estado final caracterizado por la presencia de dos muones aislados de alto momento y energía transversa faltante debido a los dos neutrinos que no son detectados. La ausencia de actividad hadrónica hace que en la señal no se esperen jets.

Las principales fuentes de fondo proceden de los diferentes procesos capaces de producir estados finales similares a los de la señal, especialmente $W + jets$, $Z + jets$, $t\bar{t}$ y el fondo, casi irreducible, de producción de WW . Otras fuentes de fondo que también han sido consideradas en esta tesis son los procesos WZ , ZZ y tW . Los procesos de QCD, a pesar de presentar estados finales diferentes, han sido igualmente tenidos en cuenta al presentar secciones eficaces muy elevadas.

Diversas técnicas de análisis se han estudiado en el contexto de esta tesis, la principal se basa en la aplicación secuencial de cortes en un grupo de observables físicos, elegidos por caracterizar a la señal y discriminar con respecto a los fondos. La presencia de neutrinos en el estado final, que impide la reconstrucción de un pico de masa clara, además del hecho de que las secciones eficaces de señal son varios órdenes de magnitud menores que las de los fondos, hace que el éxito del análisis dependa de la apropiada determinación y control de fondos.

En el análisis se han diseñado diferentes estrategias para reducir cada uno de los fondos. La reconstrucción, identificación y selección de muones está optimizada para seleccionar sucesos con dos muones procedentes de la desintegración de bosones W y es clave para rechazar sucesos de $W + jets$ en los que el bosón W decae a muón y un segundo objeto es erróneamente identificado dando lugar a estados finales con dos muones. Un método basado en datos para la determinación de este fondo consistente en estimar la probabilidad de que una traza aislada sea identificada como muón procedente de W se ha desarrollado y probado en Monte Carlo y en datos de colisión reales.

La actividad hadrónica que caracteriza procesos de fondo, principalmente $t\bar{t}$, ha de ser controlada en el análisis. Sucesos con jets procedentes de desintegraciones de quarks b han de ser vetados. Regiones de control y estudios complementarios se proponen para la correcta estimación de este fondo. El uso de información del tracker para complementar a los calorímetros se plantea para mejorar la determinación de la energía

y dirección de los jets. La adecuada estimación de la energía transversa faltante en sucesos en los que no hay neutrinos es crucial para reducir el fondo de $Z + jets$ a niveles aceptables. La variable principal para separar la señal del fondo de WW es el ángulo entre los muones en el plano transverso, $\Delta\phi_{\mu\mu}$. Las correlaciones de espín hacen que los muones emitidos en la desintegración de los dos bosones W provenientes de un bosón de Higgs escalar tengan la misma dirección en dicho plano, mientras que los que provienen de la desintegración del fondo de WW tienden a tener un mayor valor angular de separación. Esta característica de la señal juega un papel primordial en el potencial del canal $H \rightarrow WW^*$ ante un posible descubrimiento. Regiones de control y otros métodos también son establecidos para la estimación de los fondos de $Z + jets$ y WW .

El análisis se ha adaptado a diferentes escenarios de energía del centro de masas y luminosidades integradas en el LHC. Originalmente se ideó para la energía nominal del LHC de 14 TeV asumiendo una luminosidad integrada de $1fb^{-1}$ y más tarde se optimizó para una energía del centro de masas de 10 TeV y una luminosidad integrada de $200pb^{-1}$, para finalmente adaptarse a la actual energía del centro de masas del LHC, 7 TeV y una luminosidad integrada de $1fb^{-1}$, correspondiente a las previsiones del LHC para finales de 2011.

La decisión de que el LHC funcionara a una energía de 7 TeV se anunció en el workshop de Chamonix 2010 [4]. A esta energía del centro de masas la relación entre el número de sucesos de señal y el número de sucesos de fondo se degrada ligeramente con respecto a lo esperado a 14 TeV y se necesita una mayor luminosidad para llegar a tener el mismo potencial. La búsqueda directa de Higgs comenzará por lo tanto en la segunda mitad del run a 7 TeV y continuará durante los años siguientes. La primera etapa de búsquedas de Higgs en el LHC irá en paralelo con el final de la era del Tevatron.

Los diferentes procesos de señal y fondo se irán haciendo accesibles a medida que la luminosidad integrada aumente y podrán ser estudiados. El estudio del fondo WW es particularmente interesante dadas sus similitudes con la señal. Es posible establecer una estrategia para la medida de la sección eficaz del fondo WW utilizando un análisis similar al de la búsqueda de Higgs utilizando los mismos observables. Ambos estudios son complementarios ya que para la búsqueda del Higgs en el canal WW es imprescindible un conocimiento detallado del fondo WW . En este trabajo se muestra que sucesos de señal de $H \rightarrow WW^*$ podría detectarse en el contexto del análisis de WW como un exceso de fondo para determinadas masas.

El trabajo de esta tesis está organizado de la siguiente manera, en el capítulo 2 se presenta una breve introducción de la física del Modelo Standard y el mecanismo de Higgs junto con los resultados de las diferentes búsquedas realizadas en diversos experimentos. El complejo acelerador LHC junto con el experimento CMS y sus subdetectores son presentados en el capítulo 3.

10. RESUMEN DEL TRABAJO REALIZADO Y CONCLUSIONES OBTENIDAS

Los principales modos de producción y desintegración de Higgs en el LHC son discutidos en el capítulo 4, y el canal de búsqueda es presentado. La reconstrucción e identificación de los diferentes observables utilizados en el análisis se describe en el capítulo 5.

La selección de los sucesos se detalla en el capítulo 6 para las diferentes luminosidades integradas y energías del centro de masas.

Los métodos de estimación de fondo a partir de los datos se presentan en el capítulo 7. Regiones de control son diseñadas específicamente para los fondos de $t\bar{t}$, WW y $Z + jets$. Para la estimación del fondo de $W + jets$ se desarrolla un estudio de la tasa de fakes de muones. Adicionalmente se ha diseñado un análisis orientado a la medida de la sección eficaz del fondo de WW , estudio complementario a la búsqueda de Higgs WW^* que utiliza los mismos objetos y estudia similares estados finales. Asimismo, las diferentes fuentes de incertidumbre sistemática son analizadas y expuestas.

Los resultados de la aplicación del análisis en los diferentes escenarios de energía del centro de masas y luminosidad integrada en el LHC, especialmente $1 fb^{-1}$ de luminosidad a una energía de 7 TeV, referente para los próximos meses, se presentan en el capítulo 8 en forma de niveles de confianza en exclusión y significancia en función de la masa considerada. Los resultados de este trabajo se combinan con el resto de canales de búsqueda del Higgs en CMS, principalmente con $H \rightarrow ZZ^*$, para estimar el potencial de CMS ante el posible descubrimiento o exclusión del bosón de Higgs en el primer periodo de toma de datos en el LHC.

Las conclusiones finales del trabajo realizado en esta tesis están presentadas detalladamente en 9.

El estudio ha sido realizado en datos simulados pero la luminosidad recientemente recogida en el experimento CMS permite comenzar ya al estudio de los diferentes observables físicos y los primeros fondos electrodébiles, así como comprobar el estado de excelente preparación del análisis en datos reales. Resultados preliminares pueden consultarse en el anexo C

El canal de búsqueda del bosón de Higgs estudiado, $H \rightarrow WW^*$, ofrece un estado final claro y una elevada sección eficaz, comparado con otros posibles canales de desintegración, en el rango de masa $160 GeV/c^2$ favorecido por la teoría y los resultados experimentales [3]. Por lo tanto, es considerado el más prometedor de cara a la búsqueda del bosón de Higgs del Modelo Estándar en el LHC

El trabajo realizado en esta tesis se ha desarrollado dentro del grupo de Higgs de la colaboración CMS, y ha sido convenientemente adaptado a los diferentes regímenes

de luminosidad integrada y energía del centro de masas esperados para el LHC. Se ha estimado el aumento de la luminosidad necesaria para alcanzar el mismo grado de sensibilidad a energías de centro de masas menores a la de diseño del LHC. Asimismo, teniendo en cuenta que las secciones eficaces tanto de señal como de fondos disminuyen con la energía, la selección de muones y otros aspectos relativos a la determinación de los diferentes observables físicos utilizados en el análisis han sido optimizados consecuentemente, permitiendo al análisis ser lo suficientemente robusto como para alcanzar significancia incluso en escenarios de baja energía y luminosidad.

Ha sido utilizada toda la información disponible proveniente de los subdetectores para seleccionar sucesos con dos muones aislados en el estado final, de alto momento transverso y carga opuesta. La selección de muones ha sido especialmente optimizada para rechazar fondo proveniente de sucesos de QCD. Se ha mostrado que la selección de muones puede reducir la contribución del fondo de $W + jets$ casi completamente. El detallado estudio de los diferentes algoritmos de reconstrucción de jets y las correcciones de energía ha permitido llegar a una robusta definición del objeto jet en el análisis. La mayor parte del fondo de $t\bar{t}$ es eliminado mediante la aplicación del Jet Veto, al igual que otros procesos de fondo que presentan actividad hadrónica. La cinemática de los sucesos de señal Higgs depende de la masa considerada y por eso la selección de sucesos mediante cortes en las diferentes variables ha sido optimizada, maximizando la significancia, para tres regiones de masa de Higgs.

El uso de otras herramientas ha sido también estudiado. Los Trak Jets, reconstruidos a partir de trazas y técnicas de b-tagging, es decir, identificación de jet provenientes de quarks b, han sido utilizados para estimar adecuadamente el fondo de $t\bar{t}$. También se han aplicado técnicas de análisis multidimensional, especialmente útiles en la región de masa baja de Higgs, en torno a $130 \text{ GeV}/c^2$, donde la extracción de la señal es compleja debido a la menor sección eficaz y a la cinemática de este tipo de sucesos.

Se han establecido regiones de control para los principales fondos. La región de control de fondo $t\bar{t}$ se ha definido mediante la presencia de dos o más jets en el estado final. Para la región de control de $Z + jets$, se han seleccionado sucesos para los que la masa invariante de los dos muones es compatible con la masa del bosón Z . Y en el caso del fondo de WW , la región de control se ha definido utilizando la variable $\Delta\phi_{\mu\mu}$. Para la apropiada estimación de la contribución del fondo de $W + jets$, se ha aplicado el estudio de la tasa de 'fakes' de muones realizado en muestras de QCD.

El análisis para la medida de la sección eficaz del proceso WW se ha propuesto como primer paso en el estudio de la desintegración de Higgs a WW^* . Este proceso, de mayor sección eficaz que los procesos de Higgs, es accesible con menor luminosidad y por lo tanto puede estudiarse con antelación. La compresión de este fondo es un ingrediente principal en el análisis y las similitudes entre los estados finales permiten hacer uso de los mismos observables para el estudio de ambos procesos. Se ha determinado que, si

10. RESUMEN DEL TRABAJO REALIZADO Y CONCLUSIONES OBTENIDAS

el bosón de Higgs existe con una masa dentro del rango estudiado en esta tesis, podría observarse en análisis para la medida de la sección eficaz de la producción de WW como un exceso de fondo y para ello, es idóneo el uso de la variable $\Delta\phi_{\mu\mu}$ después de la selección.

Finalmente se ha realizado un exhaustivo estudio de todas las fuentes de incertidumbre sistemática y se han desarrollado diversos métodos para su estimación en datos reales.

	HWW ($160 \text{ GeV}/c^2$)	WW	WZ	ZZ	tW	$t\bar{t}$	$W + \text{jets}$	DY
ee Final selection	20.1	13.7	1.7	0.0	0.0	4.3	2.0	1.0
$\mu\mu$ Final selection	21.0	10.0	0.1	0.0	0.3	6.0	0.1	0.6
$e\mu$ Final selection	31.5	16.0	0.3	0.0	1.4	7.0	6.3	0.0

Table 10.1: Número de sucesos esperados a nivel final en el análisis para una luminosidad integrada de 1 fb^{-1} a 14 TeV para señal, Higgs con masa $m_H = 160 \text{ GeV}/c^2$ y fondos.

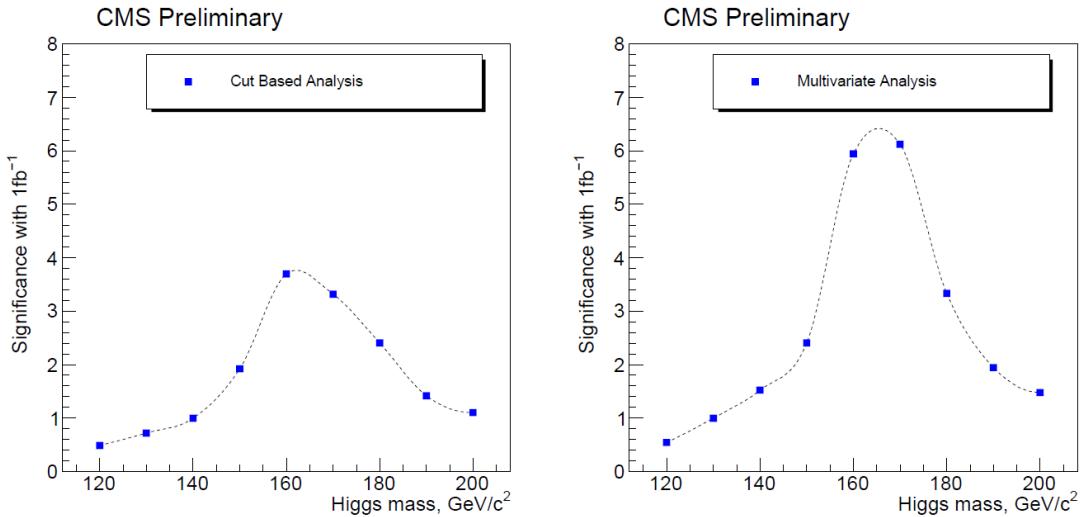


Figure 10.1: - Izquierda: Significancia esperada para el análisis con cortes para 1 fb^{-1} de luminosidad integrada a una energía del centro de masas de 14 TeV. Derecha: Significancia esperada para el análisis basado en una red neuronal para 1 fb^{-1} de luminosidad integrada a una energía del centro de masas de 14 TeV.

El número de sucesos esperados para señal y fondo para los tres estados finales estudiados en el canal Higgs yendo a WW^* , incluyendo los resultados para $\mu\mu$ obtenidos directamente en el trabajo realizado en esta tesis se muestran en la tabla 10.1. Significancia y límites de exclusión han sido calculados utilizando el número de sucesos de señal y fondo y la incertidumbre sistemática previamente estudiada.

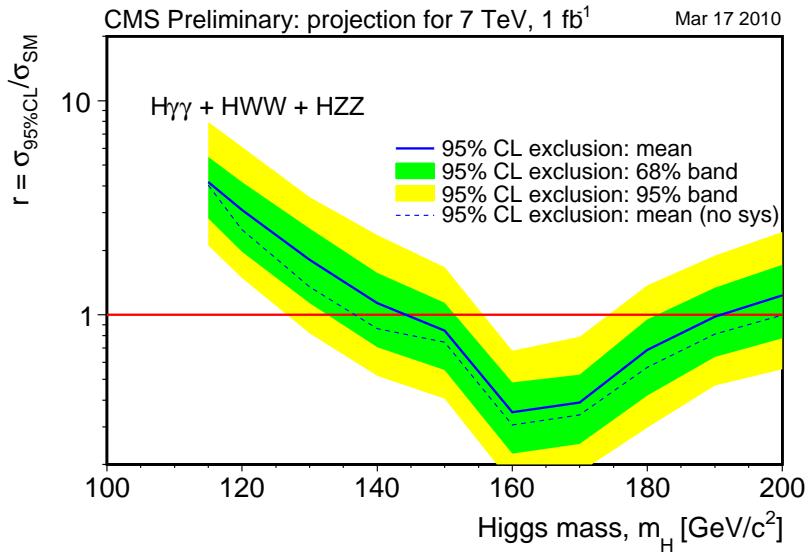


Figure 10.2: - Límite exclusión proyectado para la combinación de los principales canales de búsqueda de Higgs a 7 TeV en la región de masa entre 100 y 200 GeV/c^2 .

Se alcanza sensibilidad suficiente para realizar un posible descubrimiento del bosón de Higgs en este canal para masas entre 160 y 180 con una luminosidad integrada de $1 fb^{-1}$ a 14 TeV [10.1]. En el mismo escenario de luminosidad, se consigue una sólida exclusión del rango de masa entre 140 y 190 GeV/c^2 mediante la combinación de los resultados obtenidos en este estudio con el resto de estados finales del canal. Para 10 TeV, la optimización realizada para baja energía en diversos aspectos de la selección permite la exclusión del bosón de Higgs del modelo estándar en el intervalo $m_H \sim 160 - 170 GeV/c^2$ con una luminosidad integrada de $200 pb^{-1}$.

Este es el canal más potente para la búsqueda del bosón de Higgs durante el periodo de toma de datos a 7 TeV. El análisis presentado en esta tesis permitiría la exclusión del bosón de Higgs con una masa en torno a $160 GeV/c^2$, con una luminosidad integrada de aproximadamente $1 fb^{-1}$ a 7 TeV. Al combinar los resultados con el resto de estados finales, la exclusión ocupa las masas de 150 a $185 GeV/c^2$ y el canal adquiere sensibilidad para descubrimiento en el rango de masa de $160 < m_H < 170 GeV/c^2$ [141].

Combiando los resultados del estudio aquí presentado con el resto de búsquedas de Higgs en CMS, se alcanza nivel de exclusión para todo el rango de masa entre 145 y $190 GeV/c^2$, como puede verse en la figura 10.2, en la que se ha considerado la combinación de $H\gamma\gamma + HWW + HZZ$. En las mismas condiciones, se tendría sensibilidad suficiente para descubrir el Higgs en una masa $m_H = 160 - 170 GeV/c^2$.

10. RESUMEN DEL TRABAJO REALIZADO Y CONCLUSIONES OBTENIDAS

Appendix A

Signal and Background yields

A.1 Tables for 1 fb^{-1} at $\sqrt{s} = 14 \text{ TeV}$

The expected number of events and the relative efficiencies after each cut are presented for signal and backgrounds for each of Higgs mass hypothesis studied in tables A.1, 6.5, A.2, A.3, 6.6, A.4, A.5, A.6 and A.7.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	22.0	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	13.5 (61%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	7.9 (59%)	200 (64%)	32 (55%)	39 (67%)	44 (79%)	157 (83%)	53 (40%)	1553 (15%)
$\Delta\Phi_{\mu\bar{\mu}}$	5.6 (70%)	64 (32%)	8 (25%)	10 (24%)	7.1 (16%)	46 (29%)	17 (33%)	218 (14%)
$m_{\mu\bar{\mu}}$	5.5 (99%)	44 (69%)	1.9 (24%)	1.3 (13%)	3.1 (44%)	28 (62%)	15 (85%)	27 (12%)
$P_T(\text{highest})$	4.9 (89%)	35 (78%)	1.5 (79%)	0.6 (50%)	1.3 (41%)	16 (57%)	14 (98%)	19 (56%)
$P_T(\text{lowest})$	4.9 (100%)	35 (100%)	1.5 (100%)	0.6 (100%)	1.3 (100%)	16 (100%)	14 (100%)	19 (100%)

Table A.1: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 120 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	77	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	48 (62%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	35 (74%)	200 (64%)	32 (55%)	39 (67%)	44 (79%)	157 (83%)	53 (40%)	1553 (15%)
$\Delta\Phi_{\mu\bar{\mu}}$	22 (63%)	64 (32%)	8 (25%)	10 (24%)	7.1 (16%)	46 (29%)	17 (33%)	218 (14%)
$m_{\mu\bar{\mu}}$	20 (92%)	44 (69%)	1.9 (24%)	1.3 (13%)	3.1 (44%)	28 (62%)	15 (85%)	27 (12%)
$P_T(\text{highest})$	18 (87%)	35 (78%)	1.5 (79%)	0.6 (50%)	1.3 (41%)	16 (57%)	14 (98%)	19 (56%)
$P_T(\text{lowest})$	18 (100%)	35 (100%)	1.5 (100%)	0.6 (100%)	1.3 (100%)	16 (100%)	14 (100%)	19 (100%)

Table A.2: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 140 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

A. SIGNAL AND BACKGROUND YIELDS

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	108	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	66 (61%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	43 (56%)	174 (56%)	28 (48%)	35 (60%)	41 (74%)	143 (76%)	33 (25%)	1116 (11%)
$\Delta\Phi_{\mu\bar{\mu}}$	26 (59%)	46 (26%)	6.1 (22%)	6.6 (19%)	5.1 (12%)	32 (22%)	10 (29%)	122 (11%)
$m_{\mu\bar{\mu}}$	24 (92%)	32 (71%)	1.3 (20%)	0.8 (13%)	2.5 (50%)	21 (65%)	8 (85%)	6 (5%)
$P_T(\text{highest})$	18 (78%)	23 (72%)	1.0 (78%)	0.3 (37%)	1.0 (39%)	9 (45%)	4 (46%)	1 (27%)
$P_T(\text{lowest})$	9 (51%)	10 (43%)	0.1 (14%)	0 (0%)	0.3 (29%)	6 (59%)	0.07 (2%)	0.6 (40%)

Table A.3: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 150 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	144	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	82 (57%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	65 (79%)	174 (56%)	28 (48%)	35 (60%)	41 (74%)	143 (76%)	33 (25%)	1116 (11%)
$\Delta\Phi_{\mu\bar{\mu}}$	40 (61%)	46 (26%)	6.1 (22%)	6.6 (19%)	5.1 (12%)	32 (22%)	10 (29%)	122 (11%)
$m_{\mu\bar{\mu}}$	35 (88%)	32 (71%)	1.3 (20%)	0.8 (13%)	2.5 (50%)	21 (65%)	8 (85%)	6 (5%)
$P_T(\text{highest})$	25 (72%)	23 (72%)	1.0 (78%)	0.3 (37%)	1.0 (39%)	9 (45%)	4 (46%)	1 (27%)
$P_T(\text{lowest})$	17 (66%)	10 (43%)	0.1 (14%)	0 (0%)	0.3 (29%)	6 (59%)	0.07 (2%)	0.6 (40%)

Table A.4: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 170 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	130	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	71 (54%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	47 (67%)	131 (42%)	23 (39%)	30 (51%)	35 (62%)	123 (66%)	14 (10%)	742 (7%)
$\Delta\Phi_{\mu\bar{\mu}}$	28 (59%)	39 (30%)	5.5 (24%)	6.2 (20%)	4.5 (13%)	29 (23%)	5 (33%)	107 (14%)
$m_{\mu\bar{\mu}}$	23 (81%)	27 (69%)	1.0 (17%)	0.7 (12%)	2.3 (50%)	18 (62%)	4 (85%)	4 (4%)
$P_T(\text{highest})$	17 (75%)	17 (64%)	0.7 (71%)	0.5 (71%)	1.7 (75%)	8 (47%)	0.5 (12%)	0.6 (15%)
$P_T(\text{lowest})$	11 (64%)	7 (38%)	0.1 (20%)	0 (0%)	0.8 (50%)	5 (61%)	0.01 (3%)	0.6 (100%)

Table A.5: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 180 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

A.1 Tables for 1 fb^{-1} at $\sqrt{s} = 14 \text{ TeV}$

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	97	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	53 (54%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	37 (70%)	131 (42%)	23 (39%)	30 (51%)	35 (62%)	123 (66%)	14 (10%)	742 (7%)
$\Delta\Phi_{\mu\bar{\mu}}$	17 (46%)	39 (30%)	5.5 (24%)	6.2 (20%)	4.5 (13%)	29 (23%)	5 (33%)	107 (14%)
$m_{\mu\bar{\mu}}$	12 (73%)	27 (69%)	1.0 (17%)	0.7 (12%)	2.3 (50%)	18 (62%)	4 (85%)	4 (4%)
$P_T(\text{highest})$	9.6 (75%)	17 (64%)	0.7 (71%)	0.5 (71%)	1.7 (75%)	8 (47%)	0.5 (12%)	0.6 (15%)
$P_T(\text{lowest})$	5.6 (58%)	7 (38%)	0.1 (20%)	0 (0%)	0.8 (50%)	5 (61%)	0.01 (3%)	0.6 (100%)

Table A.6: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 190 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
Pre-sel. && $\mu\mu$	89	459	226	184	461	5128	177	40447
<i>Jet Veto</i>	48 (54%)	309 (67%)	58 (25%)	58 (32%)	56 (12%)	188 (4%)	132 (75%)	10290 (25%)
<i>MET</i>	31 (66%)	131 (42%)	23 (39%)	30 (51%)	35 (62%)	123 (66%)	14 (10%)	742 (7%)
$\Delta\Phi_{\mu\bar{\mu}}$	13 (42%)	39 (30%)	5.5 (24%)	6.2 (20%)	4.5 (13%)	29 (23%)	5 (33%)	107 (14%)
$m_{\mu\bar{\mu}}$	10.1 (75%)	27 (69%)	1.0 (17%)	0.7 (12%)	2.3 (50%)	18 (62%)	4 (85%)	4 (4%)
$P_T(\text{highest})$	7.2 (71%)	17 (64%)	0.7 (71%)	0.5 (71%)	1.7 (75%)	8 (47%)	0.5 (12%)	0.6 (15%)
$P_T(\text{lowest})$	4.6 (63%)	7 (38%)	0.1 (20%)	0 (0%)	0.8 (50%)	5 (61%)	0.01 (3%)	0.6 (100%)

Table A.7: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 200 \text{ GeV}/c^2$, 1 fb^{-1} at 14 TeV

A. SIGNAL AND BACKGROUND YIELDS

A.2 Tables for 200 pb^{-1} at $\sqrt{s} = 10 \text{ TeV}$

The expected number of events and the relative efficiencies after each cut for each Higgs mass hypothesis are presented in tables 6.7, A.8, 6.8, A.9 and A.10.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$Z + jets$
Pre-sel. && $\mu\mu$	7.8	56.9	25.7	9.24	465	37	1272
<i>Jet Veto</i>	10.4 (57%)	38.4 (67%)	7.9 (30%)	6.10 (66%)	3.7 (9%)	11 (2%)	375 (29%)
<i>MET</i>	6.5 (62%)	21.2 (55%)	3.7 (47%)	4.14 (67%)	2.7 (74%)	8.4 (74%)	13 (3%)
$\Delta\Phi_{\mu\bar{\mu}}$	4.0 (61%)	6.2 (29%)	0.6 (15%)	0.72 (17%)	0.6 (20%)	2.1 (24%)	1 (11%)
$m_{\mu\bar{\mu}}$	3.7 (93%)	4.3 (68%)	0.08 (13%)	0.13 (17%)	0.4 (64%)	1.2 (55%)	0.7 (50%)
$P_T(\text{highest})$	3.1 (82%)	2.9 (68%)	0.05 (66%)	0.056 (44%)	0.23 (63%)	0.5 (42%)	0.7 (100%)
$P_T(\text{lowest})$	1.75 (56%)	1.3 (44%)	0.03 (50%)	0.016 (28%)	0.07 (28%)	0.2 (33%)	≤ 1.4

Table A.8: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 150 \text{ GeV}/c^2$, 200 pb^{-1} at 10 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$Z + jets$
Pre-sel. && $\mu\mu$	23.2	56.9	25.7	9.24	37	465	1272
<i>Jet Veto</i>	12.7 (54%)	38.4 (67%)	7.9 (30%)	6.10 (66%)	3.7 (9%)	11 (2%)	375 (29%)
<i>MET</i>	9.7 (76%)	21.2 (55%)	3.7 (47%)	4.14 (67%)	2.7 (74%)	8.4 (74%)	13 (3%)
$\Delta\Phi_{\mu\bar{\mu}}$	6.1 (63%)	6.2 (29%)	0.6 (15%)	0.72 (17%)	0.6 (20%)	2.1 (24%)	1 (11%)
$m_{\mu\bar{\mu}}$	5.4 (88%)	4.3 (68%)	0.08 (13%)	0.13 (17%)	0.4 (64%)	1.2 (55%)	0.7 (50%)
$P_T(\text{highest})$	4.1 (75%)	2.9 (68%)	0.05 (66%)	0.056 (44%)	0.23 (63%)	0.5 (42%)	0.7 (100%)
$P_T(\text{lowest})$	2.78 (67%)	1.3 (44%)	0.03 (50%)	0.016 (28%)	0.07 (28%)	0.2 (33%)	≤ 1.4

Table A.9: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 170 \text{ GeV}/c^2$, 200 pb^{-1} at 10 TeV

A.2 Tables for 200 pb⁻¹ at $\sqrt{s} = 10$ TeV

	<i>HWW</i>	<i>WW</i>	<i>WZ</i>	<i>ZZ</i>	<i>tW</i>	<i>t̄t</i>	<i>Z + jets</i>
Pre-sel. && $\mu\mu$	13.4	56.9	25.7	9.24	37	465	1272
<i>Jet Veto</i>	6.8 (51%)	38.4 (67%)	7.9 (30%)	6.10 (66%)	3.7 (9%)	11 (2%)	375 (29%)
<i>MET</i>	4.4 (63%)	15.5 (40%)	3.0 (38%)	3.48 (56%)	2.3 (63%)	7.3 (64%)	5 (1%)
$\Delta\Phi_{\mu\bar{\mu}}$	1.94 (44%)	4.9 (31%)	0.5 (18%)	0.67 (19%)	0.5 (23%)	1.8 (24%)	0.7 (14%)
$m_{\mu\bar{\mu}}$	1.39 (71%)	3.2 (65%)	0.08 (14%)	0.099 (14%)	0.3 (62%)	1.1 (61%)	0.7 (100%)
$P_T(\text{highest})$	1.07 (77%)	2.3 (71%)	0.05 (66%)	0.065 (65%)	0.20 (59%)	0.5 (46%)	0.7 (100%)
$P_T(\text{lowest})$	0.59 (55%)	0.9 (38%)	0.03 (50%)	0.028 (42%)	0.10 (50%)	0.2 (33%)	≤ 1.4

Table A.10: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 200$ GeV/c², 200 pb⁻¹ at 10 TeV

A. SIGNAL AND BACKGROUND YIELDS

A.3 Tables for 1 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$

The expected number of events and the relative efficiencies after each cut are presented in tables A.11 6.9, A.12, A.13, A.14, 6.10, A.15, A.16, A.17, A.18, A.19 and A.20.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	10.7	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	6.4 (59%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	2.70 (42%)	97 (40%)	17 (24%)	18.6 (34%)	6.9 (67%)	28 (64%)	≤ 14	55 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	1.94 (71%)	33 (33%)	4.1 (25%)	4.0 (21%)	2.0 (29%)	6 (22%)	—	5 (8%)
$m_{\bar{\mu}}$	1.91 (98%)	18 (55%)	1.1 (26%)	0.7 (17%)	1.3 (63%)	4 (69%)	—	≤ 4
$P_T(\text{highest})$	1.80 (94%)	14 (78%)	0.6 (57%)	0.4 (58%)	0.5 (42%)	2.2 (50%)	—	—
$P_T(\text{lowest})$	1.80 (100%)	14 (100%)	0.6 (100%)	0.4 (100%)	0.5 (100%)	2.2 (100%)	—	—

Table A.11: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 120 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	41.1	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	23.9 (58%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	13.1 (54%)	97 (40%)	17 (24%)	18.6 (34%)	6.9 (67%)	28 (64%)	≤ 14	55 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	9.1 (69%)	33 (33%)	4.1 (25%)	4.0 (21%)	2.0 (29%)	6 (22%)	—	5 (8%)
$m_{\bar{\mu}}$	8.3 (91%)	18 (55%)	1.1 (26%)	0.7 (17%)	1.3 (63%)	4 (69%)	—	≤ 4
$P_T(\text{highest})$	7.5 (90%)	14 (78%)	0.6 (57%)	0.4 (58%)	0.5 (42%)	2.2 (50%)	—	—
$P_T(\text{lowest})$	7.5 (100%)	14 (100%)	0.6 (100%)	0.4 (100%)	0.5 (100%)	2.2 (100%)	—	—

Table A.12: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 140 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

A.3 Tables for 1 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	55.3	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	31.4 (56%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	16.6 (52%)	83 (34%)	15 (21%)	16.5 (30%)	6.2 (61%)	26 (59%)	≤ 14	20 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	10.5 (63%)	18 (22%)	2.5 (17%)	2.9 (17%)	1.4 (23%)	6 (21%)	—	0.8 (3%)
$m_{\bar{\mu}}$	9.8 (92%)	9 (51%)	1.0 (37%)	0.6 (19%)	0.9 (61%)	4 (64%)	—	≤ 1.5
$P_T(\text{highest})$	8.3 (84%)	7 (75%)	0.5 (49%)	0.2 (42%)	0.4 (47%)	0.8 (23%)	—	—
$P_T(\text{lowest})$	4.8 (58%)	4 (61%)	0.2 (33%)	0.12 (49%)	0.16 (36%)	0.6 (66%)	—	—

Table A.13: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 150 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	59.9	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	34.0 (56%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	19.4 (56%)	83 (34%)	15 (21%)	16.5 (30%)	6.2 (61%)	26 (59%)	≤ 14	20 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	11.8 (61%)	18 (22%)	2.5 (17%)	2.9 (17%)	1.4 (23%)	6 (21%)	—	0.8 (3%)
$m_{\bar{\mu}}$	10.9 (92%)	9 (51%)	1.0 (37%)	0.6 (19%)	0.9 (61%)	4 (64%)	—	≤ 1.5
$P_T(\text{highest})$	9.1 (83%)	7 (75%)	0.5 (49%)	0.2 (42%)	0.4 (47%)	0.8 (23%)	—	—
$P_T(\text{lowest})$	5.7 (63%)	4 (61%)	0.2 (33%)	0.12 (49%)	0.16 (36%)	0.6 (66%)	—	—

Table A.14: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 155 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	66.6	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	37.1 (55%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	25.0 (67%)	83 (34%)	15 (21%)	16.5 (30%)	6.2 (61%)	26 (59%)	≤ 14	20 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	16.5 (66%)	18 (22%)	2.5 (17%)	2.9 (17%)	1.4 (23%)	6 (21%)	—	0.8 (3%)
$m_{\bar{\mu}}$	14.8 (89%)	9 (51%)	1.0 (37%)	0.6 (19%)	0.9 (61%)	4 (64%)	—	≤ 1.5
$P_T(\text{highest})$	12.5 (84%)	7 (75%)	0.5 (49%)	0.2 (42%)	0.4 (47%)	0.8 (23%)	—	—
$P_T(\text{lowest})$	9.0 (71%)	4 (61%)	0.2 (33%)	0.12 (49%)	0.16 (36%)	0.6 (66%)	—	—

Table A.15: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 165 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

A. SIGNAL AND BACKGROUND YIELDS

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	62.7	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	34.5 (54%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	23.6 (68%)	83 (34%)	15 (21%)	16.5 (30%)	6.2 (61%)	26 (59%)	≤ 14	20 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	14.9 (62%)	18 (22%)	2.5 (17%)	2.9 (17%)	1.4 (23%)	6 (21%)	—	0.8 (3%)
$m_{\bar{\mu}}$	13.2 (88%)	9 (51%)	1.0 (37%)	0.6 (19%)	0.9 (61%)	4 (64%)	—	≤ 1.5
P_T (highest)	10.1 (77%)	7 (75%)	0.5 (49%)	0.2 (42%)	0.4 (47%)	0.8 (23%)	—	—
P_T (lowest)	7.4 (72%)	4 (61%)	0.2 (33%)	0.12 (49%)	0.16 (36%)	0.6 (66%)	—	—

Table A.16: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 170 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	58.7	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	31.3 (53%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	21.2 (67%)	83 (34%)	15 (21%)	16.5 (30%)	6.2 (61%)	26 (59%)	≤ 14	20 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	12.4 (58%)	18 (22%)	2.5 (17%)	2.9 (17%)	1.4 (23%)	6 (21%)	—	0.8 (3%)
$m_{\bar{\mu}}$	10.4 (84%)	9 (51%)	1.0 (37%)	0.6 (19%)	0.9 (61%)	4 (64%)	—	≤ 1.5
P_T (highest)	7.6 (72%)	7 (75%)	0.5 (49%)	0.2 (42%)	0.4 (47%)	0.8 (23%)	—	—
P_T (lowest)	5.3 (69%)	4 (61%)	0.2 (33%)	0.12 (49%)	0.16 (36%)	0.6 (66%)	—	—

Table A.17: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 175 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	53.7	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	28.8 (53%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	16.1 (55%)	57 (23%)	11 (16%)	13.4 (24%)	5.3 (51%)	20 (46%)	≤ 14	5 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	9.6 (59%)	14 (24%)	2.1 (18%)	2.6 (19%)	1.3 (23%)	4 (22%)	—	≤ 4
$m_{\bar{\mu}}$	7.7 (80%)	6 (45%)	0.5 (23%)	0.4 (15%)	0.8 (59%)	3.1 (68%)	—	—
P_T (highest)	6.4 (82%)	4 (56%)	0.3 (66%)	0.3 (69%)	0.5 (61%)	2.2 (72%)	—	—
P_T (lowest)	4.5 (70%)	1.6 (44%)	0.2 (50%)	0.2 (85%)	0.18 (38%)	1.7 (75%)	—	—

Table A.18: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 180 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

A.3 Tables for 1 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	40.5	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	21.1 (52%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	11.5 (54%)	57 (23%)	11 (16%)	13.4 (24%)	5.3 (51%)	20 (46%)	≤ 14	5 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	5.9 (50%)	14 (24%)	2.1 (18%)	2.6 (19%)	1.3 (23%)	4 (22%)	—	≤ 4
$m_{\bar{\mu}}$	4.4 (75%)	6 (45%)	0.5 (23%)	0.4 (15%)	0.8 (59%)	3.1 (68%)	—	—
$P_T(\text{highest})$	3.6 (80%)	4 (56%)	0.3 (66%)	0.3 (69%)	0.5 (61%)	2.2 (72%)	—	—
$P_T(\text{lowest})$	2.1 (59%)	1.6 (44%)	0.2 (50%)	0.2 (85%)	0.18 (38%)	1.7 (75%)	—	—

Table A.19: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 190 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$\mu\mu$	34.3	313	304	216	98	1386	25	420268
<i>Jet Veto</i>	17.6 (51%)	239 (76%)	68 (22%)	54 (25%)	10.2 (10%)	43 (3%)	14 (57%)	330362 (78%)
<i>MET</i>	9.6 (54%)	57 (23%)	11 (16%)	13.4 (24%)	5.3 (51%)	20 (46%)	≤ 14	5 (0%)
$\Delta\Phi_{\mu\bar{\mu}}$	4.3 (44%)	14 (24%)	2.1 (18%)	2.6 (19%)	1.3 (23%)	4 (22%)	—	≤ 4
$m_{\bar{\mu}}$	3.1 (71%)	6 (45%)	0.5 (23%)	0.4 (15%)	0.8 (59%)	3.1 (68%)	—	—
$P_T(\text{highest})$	2.3 (76%)	4 (56%)	0.3 (66%)	0.3 (69%)	0.5 (61%)	2.2 (72%)	—	—
$P_T(\text{lowest})$	1.17 (50%)	1.6 (44%)	0.2 (50%)	0.2 (85%)	0.18 (38%)	1.7 (75%)	—	—

Table A.20: $\mu\mu$ channel, for mass-dependent analysis. Optimized results for a Higgs mass hypothesis of $m_H = 200 \text{ GeV}/c^2$, 1 fb^{-1} at 7 TeV

A. SIGNAL AND BACKGROUND YIELDS

Appendix B

Results for the $\mu\mu$ final state at a centre of mass energy of 7 TeV

The 95% confidence level exclusion using this channel with 1 fb^{-1} of integrated luminosity at 14 TeV has been calculated and this result has been projected at 7 TeV using the ratio of the cross sections. In Figure B.1, the 95% confidence level with the 1σ and 2σ bands for 14 TeV is presented together with the projection that can be obtained scaling the event yields at 7 TeV.

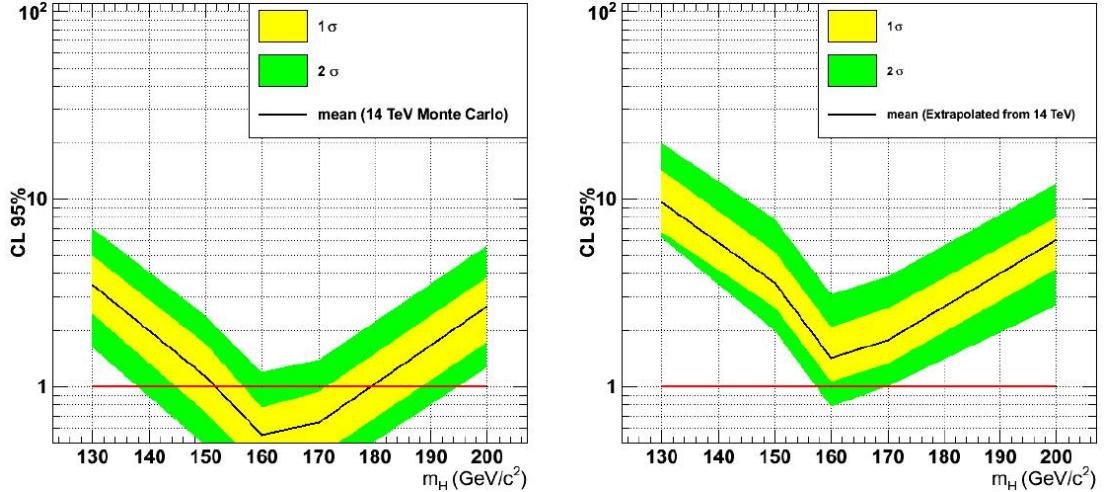


Figure B.1: - Left: Expected 95% confidence level, with the 1σ and 2σ bands in the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ channel with 1 fb^{-1} of integrated luminosity at 14 TeV. Right: Expected 95% confidence level, with the 1σ and 2σ bands for the projection made at 7 TeV.

In figure B.2, the expected 95% confidence level, with the 1σ and 2σ bands is presented using the Monte Carlo simulated at a centre of mass energy of 7 TeV and assuming 1 fb^{-1} of integrated luminosity.

B. RESULTS FOR THE $\mu\mu$ FINAL STATE AT A CENTRE OF MASS ENERGY OF 7 TeV

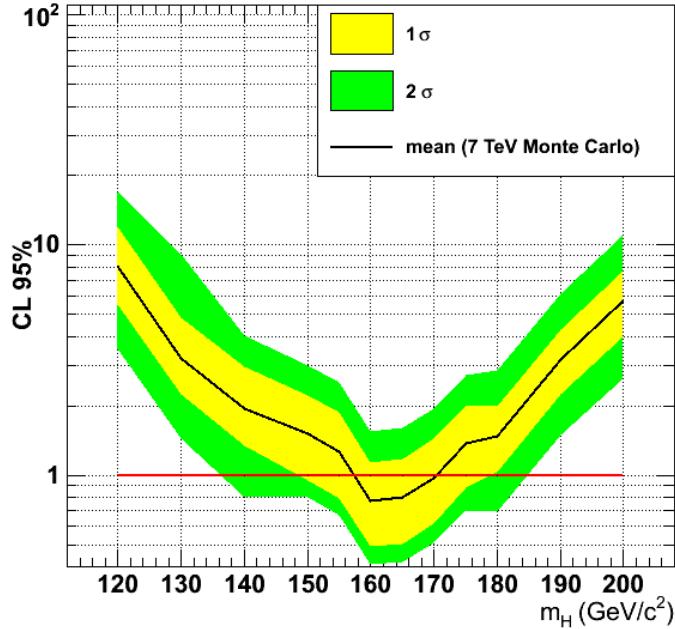


Figure B.2: - Expected 95% confidence level, with the 1σ and 2σ bands in the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ channel, $1 fb^{-1}$ of 7 TeV data assumed.

In figure B.3, the expected significance estimated using simulated events at 7 TeV is also presented with $100 pb^{-1}$ and $1 fb^{-1}$ of integrated luminosity respectively.

With $1 fb^{-1}$ of integrated luminosity and using the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ decay channel more than 2σ of sensitivity can be achieved and it is possible to set a exclusion limit on the SM Higgs boson mass for masses in the range of $\sim 158 - 170$ GeV/c^2 . The combination of the $\mu\mu$ final state with ee and $e\mu$ will improve the sensitivity of the channel in a substantial way. In figure B.4, the expected 95% confidence level, with the 1σ and 2σ bands is presented using the yields obtained in simulated data multiplied by a factor of two or by factor of three, to estimate the effect of the inclusion of the three final states. An effort was done to validate all the statistical tools available, for example [142; 143], giving as a result a very reasonable agreement, with differences smaller than 10%.

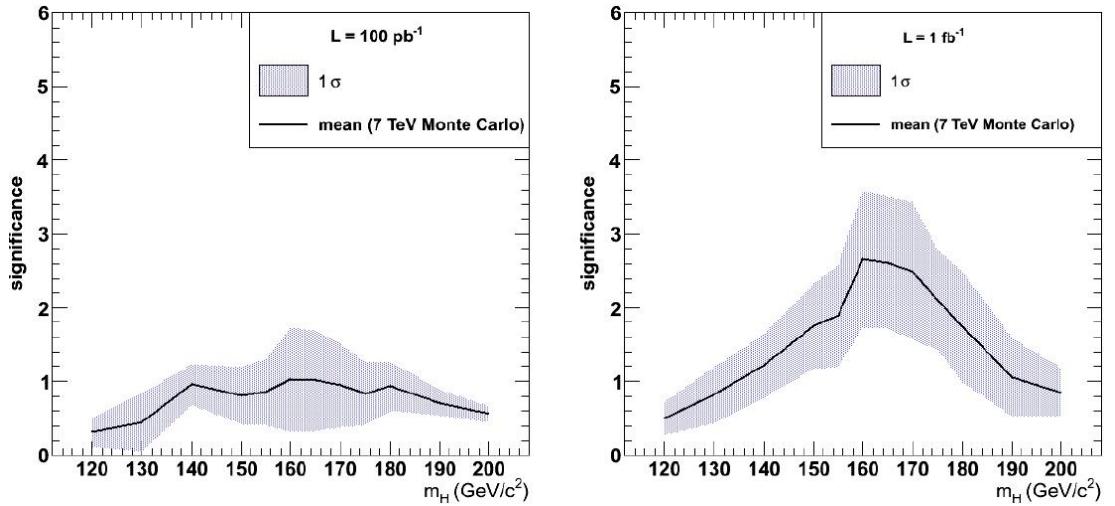


Figure B.3: - Expected significance in the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ channel for 100 pb^{-1} (left) and 1 fb^{-1} (right) of integrated luminosity at a centre of mass energy of 7 TeV.

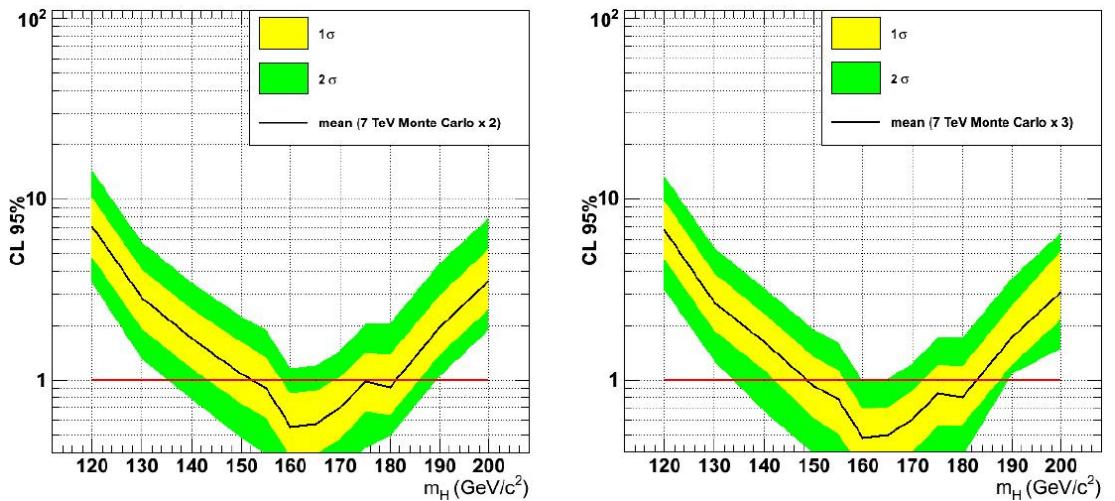


Figure B.4: - Expected 95% confidence level, with the 1σ and 2σ bands in the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$ channel, 1 fb^{-1} of 7 TeV data assumed, multiplying the signal and background yields by a factor of two (left) and three (right).

B. RESULTS FOR THE $\mu\mu$ FINAL STATE AT A CENTRE OF MASS ENERGY OF 7 TEV

Appendix C

First LHC data analysis

The LHC circulated its first proton beams on 10 September 2008 suffering a serious malfunction nine days later. A failure in an electrical connection led to serious damage, and CERN spent over a year repairing and consolidating the machine to ensure that it will not happen again.

After the recommissioning of the machine that began in the summer, on November 20th 2009 a clockwise circulating beam was established again. Just three days later, on November 23rd the LHC circulated two beams simultaneously for the first time, allowing the operators to test the synchronization of the beams and giving the experiments their first chance to look for proton-proton collisions. The injection energy was 450 GeV.

Less than one month later, on the 30th of November 2009, the LHC became the highest energy accelerator of the world, having accelerated the proton beams to an energy of 1.18 TeV, exceeding the previous world record of 0.98 TeV, which had been held by the Tevatron collider since 2001.

The first physics program was designed for 7 TeV collision energy, and consequently the 19th of March 2010 the beams were accelerated up to 3.5 TeV each. The first collisions of beams at 7 TeV followed the 30th of March 2010.

Since then the LHC has been running and CERN has announced that it will keep running for 18-24 months, allowing the experiments to record data. During this first period, 1 fb^{-1} of data is expected to be collected.

CMS has been producing results with collision data from the start, publishing a first paper on collision data using the 900 GeV and 2.36 TeV data [144].

C. FIRST LHC DATA ANALYSIS

C.1 Events with muons in the final state

In order to start the study of the backgrounds, all events with muons in the final state are interesting, especially in the early data-taking period. As the integrated luminosity increases, more processes are accessible.

At the present early stage of the LHC data taking, when investigating events with one muon in the final state, the study of $W + jets$ with $W \rightarrow \mu\nu$ is especially important. It has common features with the $H \rightarrow WW^* \rightarrow 2\mu 2\nu$: one isolated muon and missing E_T , providing an ideal scenario to define a good muon identification and missing E_T estimation, due to its large cross-section. In table C.1, the expected number of events having at least one muon passing the same requirements as the ones in the standard analysis, for different integrated luminosities at 7 TeV are presented for backgrounds and for Higgs with a mass hypothesis of $m_H = 160 \text{ GeV}/c^2$.

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$	QCD
1 pb^{-1}	0.3	5.31	1.66	0.38	1.62	24.83	4246.85	678.54	1693.19
10 pb^{-1}	3.2	53.1	16.6	3.81	16.23	248.3	43205.3	6897.0	16920.5
50 pb^{-1}	16.0	265.7	83.0	19.1	81.3	1242.7	213375	33828.2	84619.1
100 pb^{-1}	32.0	531	166	38.1	162.3	2485	426750	67656	169238

Table C.1: Number of expected events with at least one μ for Higgs signal, $m_H = 160 \text{ GeV}/c^2$, and backgrounds considering different integrated luminosities at 7 TeV

The first accessible electroweak background with two muons in the final state is $Z + jets$ with $Z \rightarrow \mu\mu$. A large and clean dimuon sample will be extracted and will be useful in the determination of systematic errors related to muon identification, and estimation of the missing E_T . When around 50 pb^{-1} will be collected, the control region of the $t\bar{t}$ background can be investigated 7.1, and studies on jets and missing E_T estimation can also be performed. With a luminosity around $50 - 100 \text{ pb}^{-1}$, the study of the WW background can be performed. It is especially important as it is, as stated before, the irreducible background for the Higgs to WW^* and the only way to control it, as was stated before, relies on the kinematics of the muons, determined by spin correlations.

With around 1 fb^{-1} of integrated luminosity the full analysis chain for Higgs presented in this thesis can be performed, using all the previously acquired knowledge of the detector, the physics objects and backgrounds. In table C.2, the expected number of events at di-muon selection level for different luminosities at 7 TeV are shown for the backgrounds and for the Higgs signal with the highest cross-section in the considered mass range, $m_H = 160 \text{ GeV}/c^2$.

Additionally to the final states with one or two muons, the study of backgrounds with three or even more muons could be also considered. Particularly interesting is the WZ final state which is a source of background when the final state has two muons. The

C.2 Current LHC run conditions

	HWW	WW	WZ	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$1 pb^{-1}$	0.0	0.3	0.3	0.2	0.1	1.4	0.0	415.9
$10 pb^{-1}$	0.7	3.1	3.0	2.1	1.0	13.8	0.2	4213.8
$50 pb^{-1}$	3.4	15.6	15.2	10.8	4.9	69.3	1.2	20913.8
$100 pb^{-1}$	6.7	31.3	30.4	21.6	9.8	138.5	2.4	41827.6
$200 pb^{-1}$	13.5	62.5	60.8	43.2	19.5	277.1	5.0	83655.1
$500 pb^{-1}$	33.7	156.3	152.0	108.0	48.9	693.0	12.5	210134
$1 fb^{-1}$	67.5	312.7	304.0	216.0	97.8	1385.6	24.9	420268

Table C.2: Number of expected events at the $\mu\mu$ selection level for Higgs signal, $m_H = 160$ GeV/c^2 , and backgrounds considering different luminosities at 7 TeV

final state of the WZ when both the W and the Z bosons decay to muons, brings a very clean final state in which to measure its cross-section. In the table C.3, the expected number of events after the selection of events with three muons at 7 TeV is shown for the backgrounds with different luminosities, no Higgs signal is selected with this requirement.

	WZ	WW	ZZ	tW	$t\bar{t}$	$W + jets$	$Z + jets$
$50 pb^{-1}$	1.5	0	0.2	0.0	0.0	0	0.3
$100 pb^{-1}$	3.0	0	0.4	0.0	0.0	0	0.7
$200 pb^{-1}$	6.0	0	0.8	0.0	0.1	0	1.4
$500 pb^{-1}$	15	0	2.1	0.0	0.3	0	3.5
$1 fb^{-1}$	30	0	4.2	0.0	0.5	0	7

Table C.3: Number of expected events with three μ for backgrounds considering different luminosities at 7 TeV

Final states with muons can also correspond to new physics beyond the standard model, having a principal role, for example, SUSY processes. Multi-muonic signatures can be present in SM Higgs, MSSM, SUSY and exotica searches, but the natural way to follow is to control first what is already known, i.e., SM processes. With a good understanding of the latter, the path to the discovery of new physics can be safely opened.

C.2 Current LHC run conditions

The LHC has been running since the 30th of March 2010. Its status is presented weekly, see for example [145]. The LHC is still in a phase of early commissioning. The current (June 2010) parameters of collision are:

$$N = 2 \times 10^{10}, k = 6, \beta* = 2m - 3 \text{ colliding pairs per IR} \quad (\text{C.1})$$

C. FIRST LHC DATA ANALYSIS

the peak luminosity being $L \sim 6 \times 10^{28} \text{ cm}^{-2} \text{s}^{-1}$. The delivered integrated luminosity as a function of the fill number and of the days of running is presented in figure C.1. Despite the fact that the luminosity seems very moderate compared to design the LHC

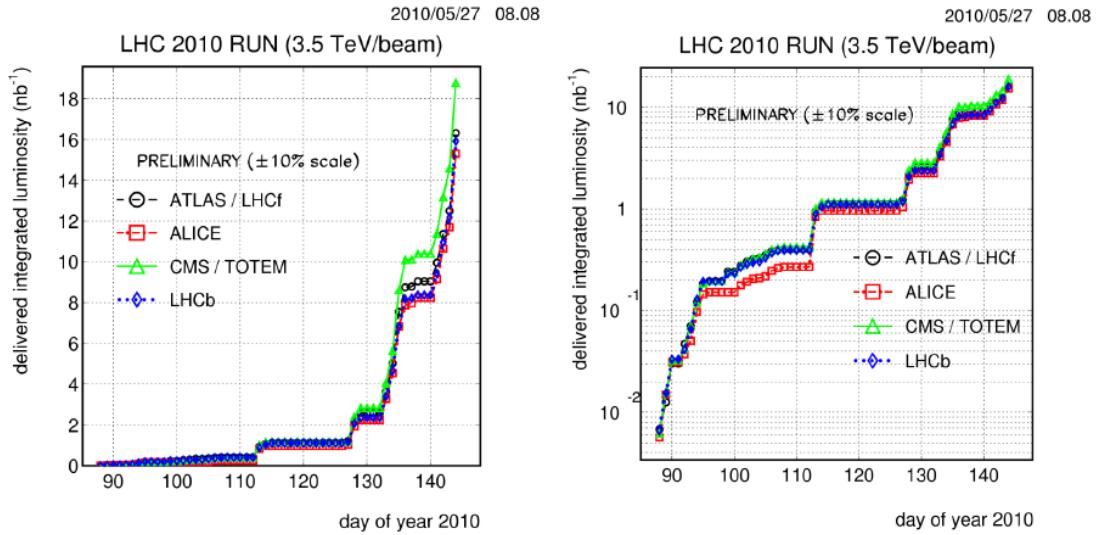


Figure C.1: - Delivered integrated luminosity in the LHC 2010 run as in the 18th of may 2010.

is progressing quickly. Most of the machine protection commissioning steps have been completed and the beam intensities are starting to increase.

By the summer of 2010, more than 1 pb^{-1} of data is expected to be collected, by the end of the year, 100 pb^{-1} and if everything goes well, by the fall of 2011, 1 fb^{-1} of data is expected. The Higgs program will start at around 500 pb^{-1} .

C.3 Early data analysis

Before the collision data, CMS recorded $\sim 1B$ of cosmic events in the CRAFT 08/09 campaigns, where CRAFT stands for Cosmic Run At Four Tesla. As a result, 23 performance papers were submitted to JINST, for example [34] or [89].

In the 900 GeV run of November, approximately 300k events were recorded. 900 GeV data was extremely useful for calibrations, and, as a result, the first physics paper on collision data [144] was accepted for publication. Around 20k events were recorded in December with a collision energy of 2.36 TeV, that meant the first peek at high energies for CMS. In [144], the inclusive particle P_T , and pseudorapidity $|\eta|$ at 900 GeV and 2.36 TeV were presented, see Figure C.2.

Since March 30th, CMS has been continuously recording 7 TeV collision data. CMS is

C.3 Early data analysis

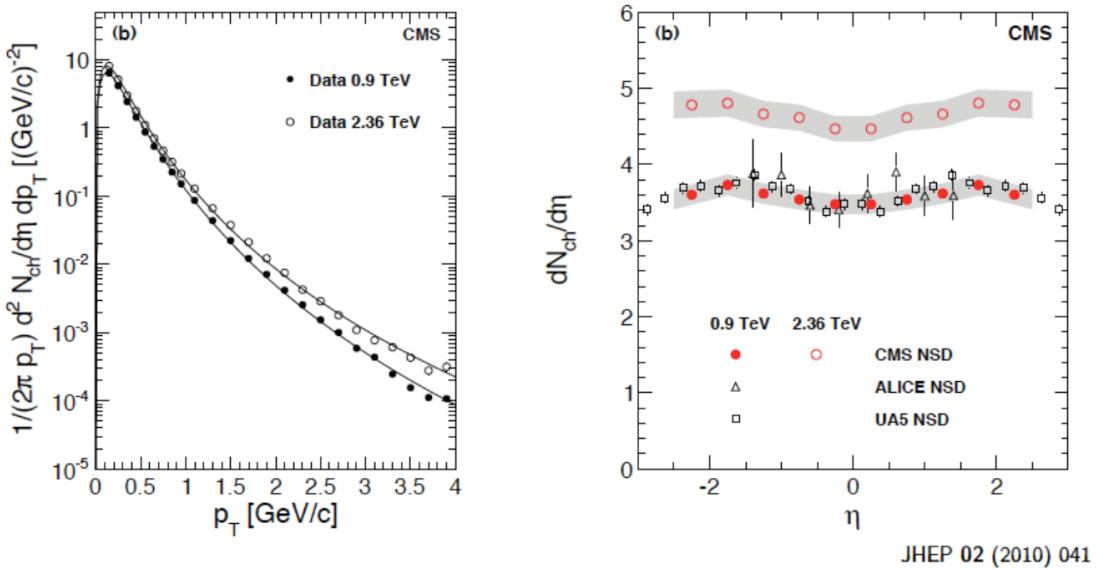


Figure C.2: - Inclusive particle P_T , and pseudorapidity $|\eta|$ at 900 GeV and 2.36 TeV in CMS [144].

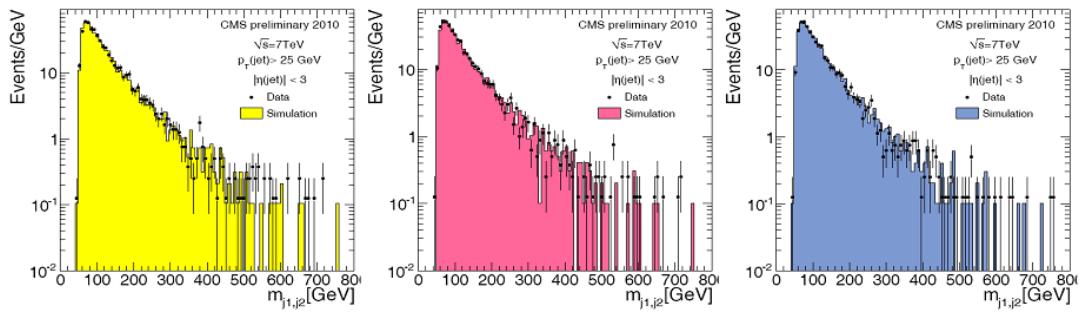


Figure C.3: - Di-jet mass, $m_{j1,j2}$ for several Jet reconstruction techniques in CMS at 7 TeV: Calorimeter Jets, JPT and Particle Flow Jets.

C. FIRST LHC DATA ANALYSIS

overall performing well, Level 1 triggers are efficient and High Level Trigger menus are working. The Tier 1 and Tier 2 centers have been reliable. The CMS operation at the start of the 7 TeV run showed all the subsystems were performing higher than 96 % and since then they have already improved. The machine is still running at low luminosity, but some resonances have already been seen, such as K_S , Λ , $\phi \rightarrow KK$ or J/ψ . Electron, Photon and Muon reconstruction are behaving as expected, but what can mainly be studied are jets. With the collision data, Calo, Particle Flow and JPT Jets are tested using the Anti- k_T 5 algorithm. Concerning the missing E_T , Particle Flow, Track corrected and Calo missing E_T are considered. The Monte Carlo (corresponding to Minimum Bias) describes the data well over 5 orders of magnitude. There are some high tails subject to ongoing study. In figure C.3, an example of the jet performance in collision data is presented for the different jet algorithms, an example of the missing E_T performance is presented in figure C.4

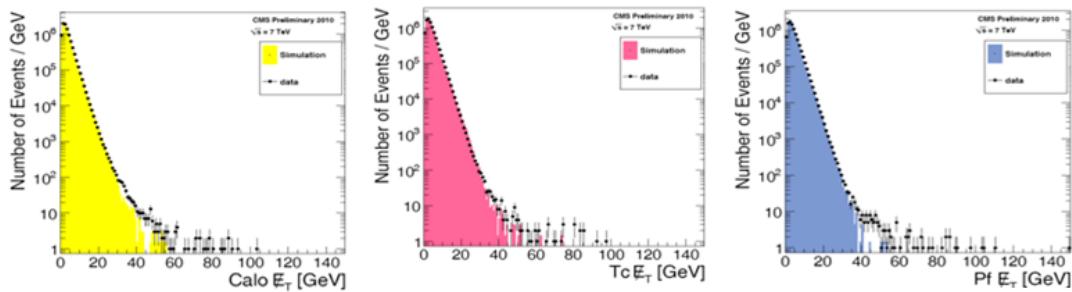


Figure C.4: - Missing Transverse Energy for an inclusive jet selection in CMS at 7 TeV using different estimators: Calorimeter, Track Corrected and Particle Flow missing E_T .

The integrated luminosity collected up to date is $\sim 16 nb^{-1}$, and already some candidates of W and Z have been selected. In figure C.5 the display for the first $Z \rightarrow \mu\mu$ candidate found in CMS is shown.

C.3.1 Muon Fake Rate

The data-driven study of the muon fake rate proposed in 7.2 is oriented to dileptonic final states and requires a relatively large muon enriched QCD sample. For that, it is not yet possible to perform it on real collision data. However, preliminary estimations of the muon fake rate can be done, selecting fakeable objects with a loose identification criteria and determining how many of them pass a tight muon selection criteria close to the one which will later be used in the analysis, and the comparing what is obtained in real data with the Monte Carlo, to check the robustness of the definitions. The loosely identified objects that have been used for the study are muons, and tracks.

One of the proposed loose muon, d , identification criteria is as follows:

- $P_T > 10 \text{ GeV}/c$, $|\eta| < 2.1$,

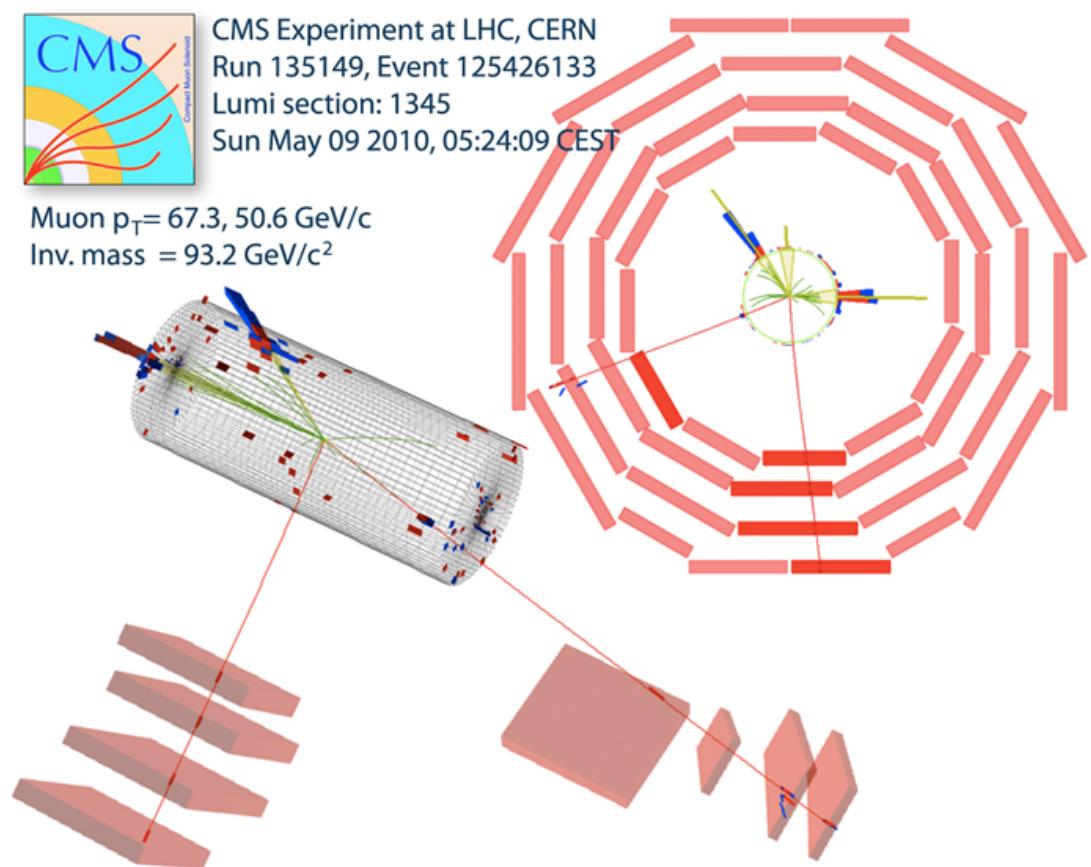


Figure C.5: - Event display of the first $Z \rightarrow \mu\mu$ candidate in CMS.

C. FIRST LHC DATA ANALYSIS

- Tracker Muon,
- $|d0| < 0.2$.

Muons passing this requirements will account for the denominator of the fake rate estimation. The tight selection criteria to determine if the muon will be flagged as 'good', l muon, will be for example:

- $P_T > 10 \text{ GeV}/c$, $|\eta| < 2.1$,
- Tracker Muon, Global Muon, Global Muon Prompt Tight,
- $|d0| < 0.02$,
- $N_{hits} > 10$,
- isolated.

Muons that pass this quality cuts will account for the numerator of the fake rate estimation. The fake rate is then calculated in the following way:

$$f_\mu = \frac{l}{d} \quad (\text{C.2})$$

The fake rate is factorized in P_T and η bins. For the study presented here, a data sample has been selected asking for a low energy trigger path and EWK events coming from W and Z were cleaned in order to extract a QCD data sample. The Monte Carlo samples used for comparison were Minimum Bias and QCD, normalized to the corresponding cross-sections. The integrated luminosity for these plots is of the order of 12 nb^{-1} .

The use of an isolation criteria in the denominator of the fake rate estimation is still under discussion. The study can be performed without the application of the isolation, but that leads, as expected, to a substantial increase of the rate. In figure C.6, the fake rate for Tracker muons without the application of isolation in either the numerator or the denominator is presented.

For the study of the muon fake rate using isolation, a combined variable built using the deposits in the tracker and in the calorimeters has been used, and the selection cut is relative. Therefore, if:

$$\frac{Iso_{tracker} + Iso_{ECAL} + Iso_{HCAL}}{P_T^\mu} < 0.2 \quad (\text{C.3})$$

muons are considered as isolated. In figure C.7, the fake rate for Tracker muons applying isolation in the denominator is presented.

For the early data studies, where the collected integrated luminosity is still not large,

C.3 Early data analysis

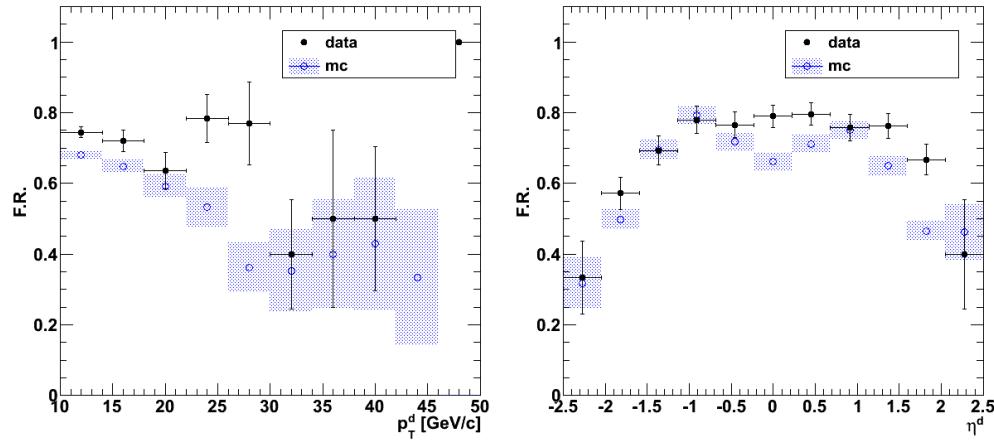


Figure C.6: - Muon Fake Rate for Tracker Muons as a function of P_T (left) and η (right) without applying isolation.

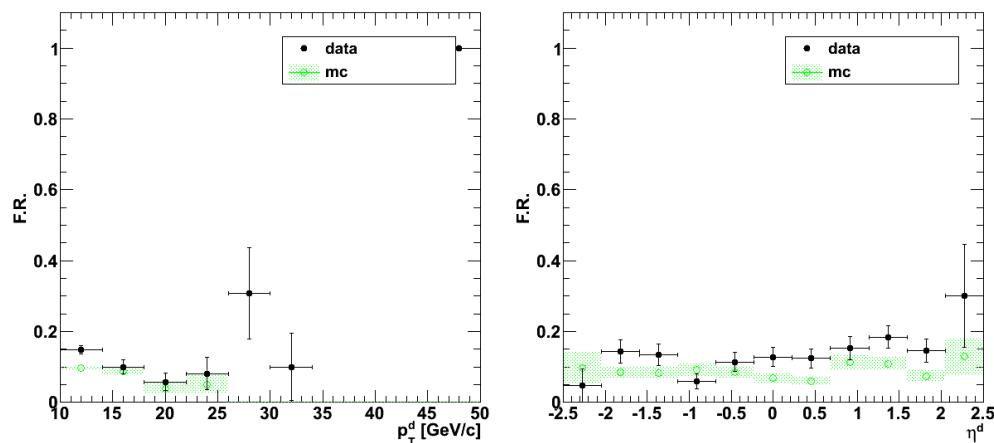


Figure C.7: - Muon Fake Rate for Tracker Muons as a function of P_T (left) and η (right) applying isolation in the denominator.

C. FIRST LHC DATA ANALYSIS

and the underlying event and pileup may not be properly modeled in the MC simulation, a very loose alternative isolation based on the tracker could be used to work with a larger sample:

$$Iso_{tracker} < 10 \quad (C.4)$$

In figure C.8, the fake rate for Tracker muons applying loose tracker isolation in the denominator is presented. With more luminosity, the di-lepton fake rate method

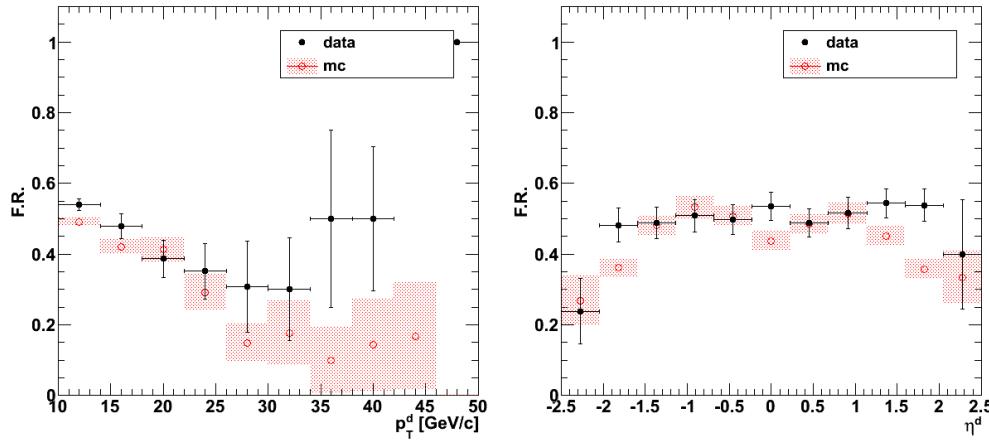


Figure C.8: - Muon Fake Rate for Tracker Muons as a function of P_T (left) and η (right) applying loose tracker isolation in the denominator.

proposed and developed in 7.2 will be applied.

Another more general definition of the fakeable object, 'd', object is done for tracks. Tracks are selected with a general identification criteria and it is possible to determine how many of them are identified as muons passing the tighter 'I' selection criteria. The proposed track, d , identification is:

- $P_T > 10 \text{ GeV}$, $|\eta| < 2.1$,
- General Track,
- $|d0| < 0.2$,
- $N_{hits} > 10$.

All the tracks passing this requirements are taken into account in the denominator of the fake rate estimation. In Figure C.9 the fake rate for tracks with no isolation requirements presented.

If the tracks are required to be isolated, the fake rate obtained is the one in Figure C.10. These preliminary studies show a very good agreement between data and Monte-Carlo, but more luminosity is needed to properly apply this fake rate to estimate the $W + jets$ contribution in the analysis.

C.3 Early data analysis

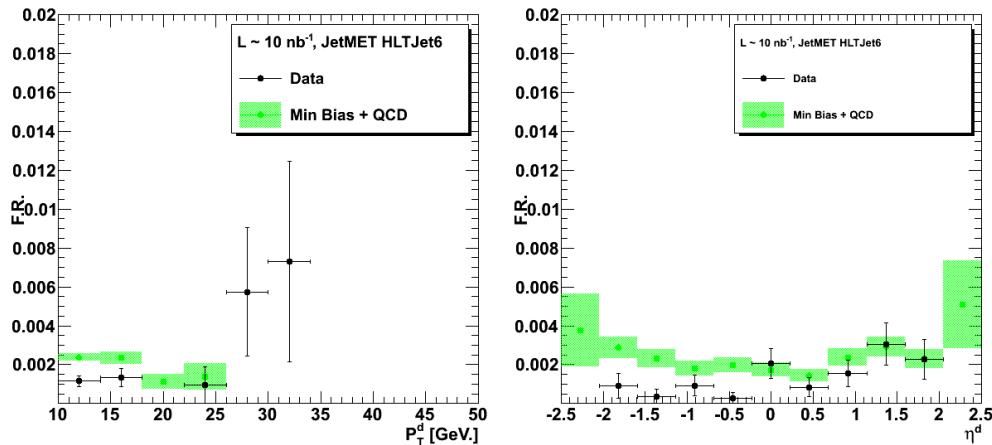


Figure C.9: - Muon Fake Rate for Tracks as a function of P_T (left) and η (right) without applying isolation in the denominator.

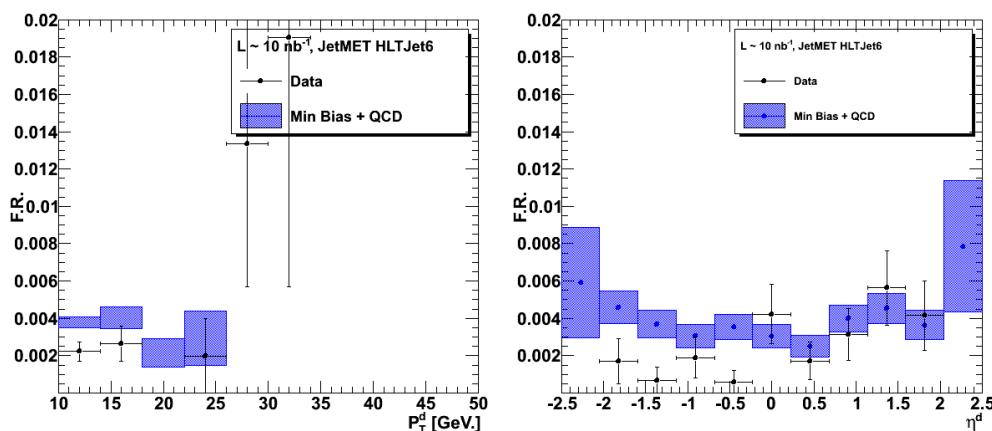


Figure C.10: - Muon Fake Rate for Tracks as a function of P_T (left) and η (right) applying tracker isolation in the denominator.

C. FIRST LHC DATA ANALYSIS

The current LHC data-taking scenario has been presented as well as the roadmap for the first analysis with muons in the final state and the estimation of the luminosity needed to perform the analysis developed in this thesis. With tens of nb^{-1} of collision data collected, some EWK candidates of $W + jets$ and $Z + jets$ have been found and the study of the muons can start. Fake rates have been calculated from Tracker muons and tracks in QCD showing a reasonable agreement with the Monte Carlo.

References

- [1] M. M. Kado and C. G. Tully, *The searches for Higgs bosons at LEP*, *Annu. Rev. Nucl. Part. Sci.* **52** (2002) 65–113. 1, 18, 177
- [2] **The CDF and D0 Collaborations** Collaboration, e. a. T. Aaltonen, *Combination of Tevatron searches for the standard model Higgs boson in the W+W- decay mode*, *Phys. Rev. Lett.* **104**, 061802 (2010) [[hep-ex/1001.4162v3](#)]. 1, 21, 177
- [3] J. Erler, *The Mass of the Higgs Boson in the Standard Electroweak Model*, *Phys.Rev.D81:051301* (2010) [[hep-ph/1002.1320v1](#)]. 1, 23, 55, 171, 178, 180
- [4] LHC, “Summary of the LHC Performance Workshop - Chamonix 2010.” 3, 64, 128, 179
- [5] D. Perkins, *Introduction to High Energy Physics*. Addison-Wesley, USA, 1987. 5
- [6] S.L.Glashow, *Partial Symmetries of weak interactions*, *Nucl.Phys.* **22** (1961) 579–588. 5
- [7] S. Weinberg, *A model of leptons*, *Phys.Rev.Lett.* **19** (1967) 1264–1266. 5
- [8] A.Salam, *Weak and electromagnetic interactions*, in *Elementary Particle Theory, Proceedings of the Nobel Symposium held 1968 at Lerum, Sweden* (N.Svartholm, ed.), pp. 367–377. Almqvist & Wiksell, 1968. 5
- [9] **Particle Data Group** Collaboration, C. Amsler *et. al.*, *Review of particle physics*, *Phys. Lett.* **B667** (2008) 1. 7
- [10] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Addison-Wesley, USA, 1995. 8
- [11] R. Feynman, *QED: the strange theory of light and matter*. Princeton University Press, 1985. 8
- [12] E. Noether, *M. A. Tavel's English translation of Noether's Theorems (1918), reproduced by Frank Y. Wang*, [physics/0503066v1](#). 9

REFERENCES

- [13] C. R. H. G. S. Guralnik and T. W. B. Kibble, *Global Conservation Laws and Massless Particless*, *Phys. Rev. Lett.* **13** (1964) 585587. 13
- [14] P. Higgs, *Broken Symmetries and the Mass of the Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509. 13
- [15] P. Higgs, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Lett.* **12** (1964) 132–133. 13
- [16] F. Englert and R. Brout, *Global Conservation Laws and Massless Particless*, *Phys. Rev. Lett.* **13** (1964) 321–323. 13
- [17] **LEP Working Group for Higgs boson searches** Collaboration, R. Barate *et. al.*, *Search for the standard model Higgs boson at LEP*, *Phys. Lett.* **B565** (2003) 61–75, [[hep-ex/0306033](#)]. 16
- [18] **CDF** Collaboration, F. A. et al, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, *Phys. Rev. Lett.* **74** (1995) 2626 – 2631. 21
- [19] **D0** Collaboration, S. A. et al., *Observation of the Top Quark*, *Phys. Rev. Lett.* **74** (1995) 2632 – 2637. 21
- [20] **SNO** Collaboration, Q. R. Ahmad *et. al.*, *Measurement of the charged current interactions produced by B-8 solar neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001) 071301, [[nucl-ex/0106015](#)]. 23
- [21] K. D. Lane, *Technicolor 2000*, [hep-ph/0007304](#). 23
- [22] M. Schmaltz and D. Tucker-Smith, *Little Higgs Review*, *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229–270, [[hep-ph/0502182](#)]. 23
- [23] M. Shifman, *Large Extra Dimensions: Becoming acquainted with an alternative paradigm*, [arXiv:0907.3074](#). 23
- [24] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*, *Nucl.Phys.* **B70** (1974) 39–50. 24
- [25] P. Binetruy, *Supersymmetry: Theory, Experiment, and Cosmology*. 2006. Oxford University Press. 24
- [26] S. L. Adler, *Axial-vector vertex in spinor electrodynamics*, *Phys. Rev.* **177** (Jan, 1969) 2426–2438. 24
- [27] J. S. Bell and R. Jackiw, *A PCAC puzzle: $\pi 0 \rightarrow \gamma\gamma$ gamma in the sigma model*, *Nuovo Cim.* **A60** (1969) 47–61. 24
- [28] J. Gunion, H. Haber, G. Kane, and S. Dawson, *The Higgs Hunter's Guide*. Addison-Wesley, Reading, UK, 1990. 26

REFERENCES

- [29] *LHC Yellow Book*. CERN, 1995. CERN-AC-95-05. 27
- [30] **ALICE** Collaboration, *ALICE: Technical proposal for a large ion collider experiment at the CERN LHC*. CERN-LHCC-95-71. 29
- [31] **ATLAS** Collaboration, G. Aad *et. al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003. 29
- [32] **CMS** Collaboration, G. L. Bayatian *et. al.*, *CMS physics: Technical design report*. CERN-LHCC-2006-001. 29
- [33] **LHCb** Collaboration, *LHCb technical design report: Reoptimized detector design and performance*. CERN-LHCC-2003-030. 29, 55
- [34] **CMS** Collaboration, S. C. et al, *The CMS Experiment at the CERN LHC*, *Journal of Instrumentation (JINST)*, **3** S08004 (2008). 29, 202
- [35] **CMS** Collaboration, *The CMS magnet project : Technical Design Report*, 1997. 31
- [36] **CMS** Collaboration, *The CMS tracker system project : Technical Design Report*, 1997. 32
- [37] **CMS** Collaboration, *The CMS tracker : addendum to the Technical Design Report* , 2000. 32
- [38] **CMS** Collaboration, *A lead/scintillator electromagnetic calorimeter for CMS*, 1993. 34
- [39] **CMS** Collaboration, *The CMS electromagnetic calorimeter project : Technical Design Report* , 1997. 34
- [40] **CMS** Collaboration, *Changes to CMS ECAL electronics : addendum to the Technical Design Report*, 2002. 34
- [41] **CMS** Collaboration, *The CMS hadron calorimeter project : Technical Design Report*, 1997. 36
- [42] **CMS** Collaboration, *The CMS muon project : Technical Design Report* , 1997. 38
- [43] M. S. S. nudo, *Geometría del detector cms reconstruida con el sistema de alineamiento link.*, *PhD Thesis* (2009). Universidad de Cantabria. 42
- [44] P. M. R. del Árbol, *Development and application of alignment algorithms for optimization of the muon detection at the cms experiment at the lhc*, *PhD Thesis* (2010). Universidad de Cantabria. 42

REFERENCES

- [45] J. M. V. García, *Medida de la sección eficaz y masa del quark top en las desintegraciones dileptónicas de pares top-antitop en el experimento cms del lhc*, *PhD Thesis* (2009). Universidad de Oviedo. 42, 149
- [46] CMS Collaboration, *The CMS Magnet Test and Cosmic Challenge (MTCC Phase I and II) Operational Experience and Lessons Learnt*, . CERN-CMS-NOTE-2007-005. 43
- [47] CMS Collaboration, *CMS TriDAS project : Technical Design Report; 1, the trigger systems*, 2000. 44
- [48] CMS Collaboration, *CMS trigger and data-acquisition project : Technical Design Report*, 2002. 44
- [49] CMS Collaboration, *The CMS Computing Model*, 2004. 46
- [50] R. Brun and F. Rademakers, *Root - an object oriented data analysis framework*, *Inst. Meth. in Phys. Res. A* **389**, 81 (1997). 48
- [51] M. E. M. H. M. Georgi, S. L. Glashow and D. V. Nanopoulos, *Higgs Bosons from Two-Gluon Annihilation in Proton-Proton Collisions*, *Phys. Rev. Lett.* **40** (1978) 692–694. 52
- [52] M. G. P. N. S. Catani S, D. de Florian, *Soft-gluon resummation for Higgs boson production at hadron colliders*, *JHEP* **0307** (2003) 028, [hep-ph/0306211]. 53
- [53] CMS Collaboration, “CMS Higgs public results.” 54
- [54] CMS Collaboration, *CMS Physics Technical Design Report, Volume II: Physics Performance*, *J. Phys. G: Nucl. Part. Phys.* **34** (2007) 995. 54, 55
- [55] S. Baffioni *et. al.*, *Search for the higgs boson in the zz^* decay channel with the cms experiment*, . CMS-PAS-HIG-08-003 / CMS AN-2008/050. 55, 164
- [56] M. V. Acosta *et. al.*, *Towards the search for the standard model higgs boson produced in vector boson fusion and decaying into a τ pair in cms with 1 fb^{-1} : τ identifications studies*, . 55
- [57] M. V. Acosta *et. al.*, *Search for the standard model higgs boson produced in vector boson fusion and decaying into a tau pair in cms with 1 inverse femtobarn*, . CMS-PAS-HIG-08-008 /CMS AN-2007/035. 55
- [58] H. D. M. Dittmar, *How to find a Higgs Boson with a Mass between 155–180 GeV at the LHC*, *Phys. Rev. D* **55** (1997) 167–172, [hep-ph/9608317]. CMS NOTE-1997/083 - LHC Higgs Search with l+nul- nubar Final States. 56
- [59] M. D. M. G. F. P. G. Davatz, G. Dissertori, *Effective K-factors for $gg \rightarrow H \rightarrow WW \rightarrow l\nu l\nu$ at the LHC*, *JHEP 0405* **009** (2004) [hep-ph/0402218]. 56

REFERENCES

- [60] M. D. G. Davatz and A. Giolo-Nicollerat, *Standard Model Higgs discovery potential of CMS in the $H \rightarrow WW \rightarrow l\nu l\nu$ channel*, *J. Phys. G: Nucl. Part. Phys.* **34** (2007). N 85. 56
- [61] C. C. et al, *Search for the Higgs boson in the WW decay channel with the CMS experiment*, . CMS AN-2007/037. 56, 87, 125
- [62] C. C. 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. 56, 87, 125, 154
- [63] G. B. et al, *Search Strategy for a Standard Model Higgs Boson Decaying to Two W Bosons in the Fully Leptonic Final State at $\sqrt{s} = 10$ TeV*, . CMS AN 2009/139. 56, 87
- [64] CMS Collaboration, C. C. et al, *Search for the Higgs boson in the WW decay channel with the CMS experiment*, . PAS HIG 07 001. 56, 162
- [65] CMS Collaboration, C. C. et al, *Search Strategy for a Standard Model Higgs Boson Decaying to Two W Bosons in the Fully Leptonic Final State*, . PAS HIG 08 006. 56, 146, 159, 162, 164
- [66] CMS Collaboration, “Physics Analysis Oriented Event Display (Fireworks / cmsShow).” 58
- [67] LHC, “Summary of the LHC Performance Workshop - Chamonix 2009.” 63, 127, 164
- [68] P. S. T. Sjöstrand, S. Mrenna, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605** (2006) 026, [[hep-ph/0603175](#)]. 64, 68
- [69] S. Frixione and B. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **0206** (2002) 029, [[hep-ph/0204244](#)]. 64
- [70] G. Corcella *et. al.*, *HERWIG 6.5*, *JHEP* **0101** **010** (2001) [[hep-ph/0011363](#)]. 65
- [71] F. P. R. P. A. P. M.L. Mangano, M. Moretti, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **0307** **001** (2003) [[hep-ph/0206293](#)]. 65
- [72] L. S. S.R. Slabospitsky, *TopReX generator (version 3.25). Short manual*, *Comput.Phys.Commun.* **148** (2002) 87 – 102, [[hep-ph/0201292](#)]. 65
- [73] J. Alwall *et. al.*, *MadGraph/MadEvent v4: The New Web Generation*, *JHEP* **0709:028** (2007) [[hep-ph/0706.2334](#)]. 65
- [74] CMS Collaboration, “CMSSW.” 65

REFERENCES

- [75] P. N. S. Frixione and B. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, *JHEP* **0308** (2003) 007, [[hep-ph/0305252](#)]. 68
- [76] M. Spira, *HIGLU: A Program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections*, [hep-ph/9510347](#). 69
- [77] M. Spira, *HIGLU and HDECAY: Programs for Higgs Boson Production at the LHC and Higgs Boson Decay Widths*, *Nucl.Instrum.Meth.* **A389** (1997) 357 – 360, [[hep-ph/9610350](#)]. 69
- [78] M. Spira, “HIGLU.” 69
- [79] M. Spira, “VV2H.” 69
- [80] D. G. M. Spira, A. Djouadi and P. Zerwas, *Higgs boson production at the LHC*, *Nucl. Phys. B* **453** (1995) 17–82. 69
- [81] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, *Comput. Phys. Commun.* **108** (1998) 56–74, [[hep-ph/9704448](#)]. 69
- [82] M. Spira, “HDECAY.” 69
- [83] M. Spira, “V2HV.” 69
- [84] M. Spira, “HQQ.” 69
- [85] F. P. Ch. Anastasiou, R. Boughezal, *Mixed qcd-electroweak corrections to higgs boson production in gluon fusion*, *JHEP* **0904:003** (2009) [[hep-ph/0811.3458](#)]. 70
- [86] F. S. B. R. W. C. Anastasiou, G. Dissertori, *Qcd radiation effects on the $h \rightarrow ww \rightarrow ll\nu\nu$ signal at the lhc*, *JHEP* **0803:017** (2008) [[hep-ph/0801.2682](#)]. 70
- [87] K. E. J. Campbell, “MCFM.” 71
- [88] **CMS** Collaboration, G. Abbiendi *et. al.*, *Muon Reconstruction in the CMS Detector*, . CMS AN 2008/097. 73
- [89] **CMS** Collaboration, *Performance of CMS Muon Reconstruction in Cosmic-Ray Events*, *Journal of Instrumentation (JINST)* **5 T03022** (2010) [[ins-det/0911.4994](#)]. 73, 202
- [90] **CMS** Collaboration, M. Mulders *et. al.*, *Muon Identification in CMS*, . CMS AN 2008/098. 77
- [91] W. Adam *et. al.*, *Electron reconstruction in cms*, . CMS AN -2009/164. 84, 86

REFERENCES

- [92] S. Baffioni *et. al.*, *Electron reconstruction in cms*, . CMS NOTE 2006/40. 84
- [93] R. Frühwirth, *Track fitting with non-gaussian noise*, *Computer Physics Communications* **100** (1997) 1–16. 85
- [94] W. Adam *et. al.*, *Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc*, . CMS NOTE-2005/001. 85
- [95] S. Baffioni *et. al.*, *Electron identification in cms*, . CMS AN -2009/178. 85
- [96] S. Baffioni *et. al.*, *Discovery potential for the sm higgs boson in the $h \rightarrow zz^* \rightarrow e+e-e+e-$ decay channel*, . CMS NOTE-2006/115. 86
- [97] F. Beaudette *et. al.*, *Search for a light standard model higgs model in the $h \rightarrow ww^* \rightarrow ee\nu\nu$ channel*, . CMS NOTE-2006/114. 86
- [98] G. P. Salam, *Towards Jetography*, hep-ph/0906.1833v2. 87
- [99] **CMS** Collaboration, *Performance of Jet Algorithms in CMS*, . CMS PAS JME-07-003; CMS AN-2008/001. 88
- [100] **CMS** Collaboration, *Particle Flow Reconstruction of Jets, Taus, and MET*, . CMS PAS PFT-09-001; CMS AN-2009/039. 88, 100
- [101] **CMS** Collaboration, *Performance of Jet Reconstruction with Charged Tracks only*, . CMS PAS JME-08-001; CMS AN-2008/041. 89
- [102] G. S. G. P. Salam, *A practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **0705:086** (2007) [hep-ph/0704.0292]. 91
- [103] S. Catani *et. al.*, *New clustering algorithm for multi - jet cross-sections in $e+e-$ annihilation*, *Phys.Lett. B* **269** (1991) 432–438. 91
- [104] S. Catani *et. al.*, *Longitudinally-invariant k-clustering algorithms for hadron-hadron collisions*, *Nucl. Phys. B* **406** (1993) 187–224. 91
- [105] D. E. S. S. D. Ellis, *Successive Combination Jet Algorithm For Hadron Collisions*, *Phys.Rev. D* **48** (1993) 3160–3166, [hep-ph/9305266]. 91
- [106] S. M. B. W. Yu.L. Dokshitzer, G.D. Leder, *Better Jet Clustering Algorithms*, *JHEP* **9708:001** (1997) [hep-ph/9707323]. 91
- [107] T. W. M. Wobisch, *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering*, hep-ph/9907280. 91
- [108] G. S. M. Cacciari, G. P. Salam, *The anti- k_t jet clustering algorithm*, *JHEP* **0804:063** (2008) [hep-ph/0802.1189]. 91
- [109] **CMS** Collaboration, R. Harris *et. al.*, *Plans for Jet Energy Corrections at CMS*, . CMS-PAS-JME-07-002 / CMS AN-2008/003. 91

REFERENCES

- [110] **CMS** Collaboration, A. Santocchia *et. al.*, *Jet Corrections to Parent Parton Energy*, . CMS-PAS-JME-08-002 / CMS AN 2008/084. 91
- [111] **CMS** Collaboration, R. Bainbridge *et. al.*, *The Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS*, . CMS-PAS-JME-09-002 / CMS AN 2008/084. 93
- [112] **CMS** Collaboration, D. del Re; M. Voutilainen, *Jet energy calibration with photon+jet events*, . CMS-PAS-JME-09-004 /CMS AN-2009/012. 94
- [113] **CMS** Collaboration, D. G. J. R. C. T. H. Pi, P. Avery, *Measurement of Missing Transverse Energy With the CMS Detector at the LHC*, . CMS-NOTE-2006-035 ; CERN-CMS-NOTE-2006-035. 98
- [114] **CMS** Collaboration, F. Golf *et. al.*, *Performance of Track-Corrected Missing Transverse Energy in CMS*, . CMS-PAS-JME-09-010/CMS AN-2009/022. 99
- [115] **CMS** Collaboration, A. B. T. Virdee, A. Petrilli, *CMS High Level Trigger*, . LHCC-G-134 ; CERN-LHCC-2007-021. 109
- [116] **CMS** Collaboration, “Physics Dataset Working Group .” 112
- [117] **CMS** Collaboration, “Physics Secondary Datasets.” 112
- [118] E. T. M. Z. V. Drollingher, U. Gasparini, *Physics study of the higgs decay channel $h \rightarrow ww^(*) \rightarrow 2\mu 2\nu$* , . CMS NOTE 2006/055. 124
- [119] F. James and M. Roos, *Minuit: Function minimization and error analysis*, . 125
- [120] T. Speer *et. al.*, *Algorithms for b jet identification in cms*, . BTV-09-001 / CMS AN-2009/085. 134
- [121] P. Demin *et. al.*, *Tagging b jets with electrons and muons at cms*, . CMS NOTE 2006/043. 134
- [122] R. G. S. C. Jorda *et. al.*, *Analysis strategy for the measurement of the $ww \rightarrow \mu\nu\mu\nu$ cross-section*, . CMS AN-2009/025. 134, 150
- [123] R. G. S. C. Jorda *et. al.*, *Analysis strategy for the measurement of the $ww \rightarrow \mu\nu\mu\nu$ cross-section with $200 pb^{-1}$ at center of mass energy of 10 tev*, . CMS AN-2009/185. 134, 150
- [124] A. Rizzi *et. al.*, *Track impact parameter based b-tagging with cms*, . CMS NOTE 2006/019. 135
- [125] A. Rizzi *et. al.*, *A combined secondary vertex based b-tagging algorithm in cms*, . CMS NOTE 2006/014. 135
- [126] T. Muller *et. al.*, *Inclusive secondary vertex reconstruction in jets*, . CMS NOTE 2006/027. 135

REFERENCES

- [127] T. Muller *et. al.*, *Refined analysis of tth with hbb at cms*, . CMS AN 2007/001. 135
- [128] A. Hoecker *et. al.*, *Tmva - toolkit for multivariate data analysis, PoSACAT 040* (2007) [[physics.data-an/0703039](#)]. 137
- [129] **CMS** Collaboration, G.Bauer *et. al.*, *Study of data-driven methods for estimation of fake lepton backgrounds*, . CMS AN-2009/120. 146
- [130] **CMS** Collaboration, W. Andrews *et. al.*, *Data-driven methods to estimate the electron and muon fake contributions to lepton analyses*, . CMS AN-2009/041. 146
- [131] V. Büge *et. al.*, *Prospect for the precision measurement of the w mass with the cms detector at the lhc*, . CMS AN-2006/061. 155
- [132] J. Andrea *et. al.*, *Measurement of the top quark pair production cross section with $l = 100 \text{ pb}^{-1}$ in dileptonic final states using the cms detector*, . CMS AN-2008/016. 156
- [133] S. Baffioni *et. al.*, *Measurement of inclusive $w \rightarrow e\nu$, $z \rightarrow ee$ cross section in pp collisions at $\sqrt{s} = 14 \text{ tev}$* , . CMS AN-2007/06. 158
- [134] N. Adam *et. al.*, *Towards a measurement of inclusive $w \rightarrow e\mu$, $z \rightarrow \mu\mu$ cross section in pp collisions at $\sqrt{s} = 14 \text{ tev}$* , . CMS AN-2007/031. 158
- [135] F. Olness *et. al.*, *$W/Z/Higgs$ production at LHC and PDF uncertainties*, . HERA-LHC Workshop. 158
- [136] T. Severini, *Likelihood methods in statistics*. Oxford University Press, 2000. 161
- [137] A. L. Read, *Modified frequentist analysis of search results (the cls method)*, . CERN-OPEN-2000-205. 161
- [138] A. Read, *Presentation of search results: the cls technique, Journal of Physics G: Nucl. Part. Phys. 28* (2002) 2693–2704. 161
- [139] A. OHagan, *Kendalls advanced theory of statistics, Volume 2B: Bayesian inference*. Edward Arnold, London, 1994. 162
- [140] **CMS** Collaboration, S. Baffioni *et. al.*, *Projected exclusion limits on the sm higgs boson cross sections obtained by combining the $h \rightarrow ww^*$ and zz^* decay channels*, . CMS AN-2009/020. 164
- [141] **CMS** Collaboration, *The CMS physics reach for searches at 7 TeV*, . CMS NOTE -2010/008. 166, 175, 183
- [142] T. Junk, “mclimit.” 196

REFERENCES

- [143] G. Q. G. Schott, D. Piparo, “RooStatsCms.” 196
- [144] CMS Collaboration, V. K. et al, *Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 tev*, *Journal of High Energy Physics.* **2** (2010) 1–35. 199, 202, 203
- [145] J. Wenninger, *LHC Report*, Friday 11 June 2010. 201

Acknowledgements

First and foremost, I would like to thank Javier Cuevas, that guided my work during this four years of PhD, for his excellent advice, continuous support and extraordinary patience. He has been the best supervisor I could possibly have hoped for, and by far the most fun. He introduced me to high energy physics and I will never be able to thank him enough for the great opportunity he gave me. Together with Javier I also want to thank Teresa Rodrigo for being an inspiration and because without her support, this thesis would not have been possible.

I want to express my deepest gratitude to Javier Fernández Menéndez, who contributed to this work in many ways, for his kindness and great help during all these years; and also to Isidro González Caballero, that gave me some valuable suggestions and help, and always had sense of humor.

My sincere thanks are due to Guillermo Gómez Ceballos, which was always there for me, for his useful comments, endless advice and help.

In four years working in the University of Oviedo and Instituto de Física de Cantabria I had the chance to meet different people for whom I have great regard, and I wish to extend my thanks to all those who have helped me in one way or another, like Alberto Ruiz, Francisco Matorras, Jesús Marco de Lucas, Celso Martínez Rivero, Iban Cabrillo, Rocío Vilar, Gervasio Gómez, Iván Vila and Alicia Calderón amongst others. And to my current office mate, Shan-Huei Chuang, Sunny, for all the hours we spent together in the last months.

My past and present fellow PhD students deserve a special mention here from Mar Sobrón, one of the nicest persons I've met at CERN, to Pablo Martínez Ruiz del Árbol, that was my family in Geneva and with whom I shared moments of all sorts and a lot of laughs; also thanks to my good friend Lara Lloret Iglesias, for making my life better in general and easier at work; also to the legendary Clara Jordá Lope, that never fails to make me laugh, and to her 'probe', Jordi Duarte; and of course thanks to Patricia Lobelle, Santiago Folgueras, Jesús Vizán and Bárbara Álvarez, with whom I spent many days working in Oviedo and helped me in many ways.

Especial appreciation goes to the great colleagues based at CERN, Luca Scodellaro and Marcos Fernández, for all the support and the laughs, that make my everyday a bit more entertaining.

And of course to all my friends in Switzerland, beginning with María Cepeda, my brilliant flat-mate, for all the patience and the good moments, and for supporting me with a smile in every random thing I do. Thanks to my dear friend Lola Saiz Santos, for just one thing: being AWESOME. And thanks to Paloma Quiroga, the last to arrive and the first to go, the summer of 2009 was one the best of my life, and that was entirely your fault, thanks for that, and thanks for all the AdR memories. Thanks also to Ricardo Vázquez Gómez, Ricci, for the beautiful decoration of my house, amongst other things you do, and thanks to MedievÁlex Iribarren, María Moreno, Carmen Díez and Jorge Casalderrey, you guys are the best. Thanks also to my friends from Spain, especially to Paula Barrio and Ana Fernández Trapiello, I'm glad for having you both still in my life after all these years.

I would like to thank many other people that shared different experiences with me during my PhD, but I will just mention Fabian Stöckli, because knowing him changed me and made me want to be better.

And of course thanks to Elias Coniavitis for all the nice and happy moments, for the trips, the games, and the most important thing: the Kaviar.

Last but not least, thanks to my all family for being so weird and so cool at the same time, and to Doctor Oetker, that fed me during the summer of 2008.