UNIVERSITEIT ANTWERPEN

Faculteit Wetenschappen

Departement Fysica

Search for a Brout-Englert-Higgs Boson decaying to $\mathrm{W}^+\mathrm{W}^- \to \ell^+\nu\ell^-\bar{\nu}$ with the CMS experiment

---

Zoektocht naar het Brout-Englert-Higgs Boson via het verval naar $\mathrm{W}^+\mathrm{W}^- \to \ell^+\nu\ell^-\bar{\nu}$ met het CMS experiment

Proefschift voorgelegd tot het behalen van de graad van doctor in de wetenschappen aan de Universiteit Antwerpen

Ph.D. Student:
Michele Selvaggi

Promotor:
Prof. Nick van Remortel

Zoektocht naar het Brout-Englert-Higgs boson via het verval naar $\mathbf{W^+W^-} \rightarrow \boldsymbol{\ell^+\nu\ell^-\bar{\nu}}$ met het CMS experiment

Op microscopisch niveau kunnen de natuurwetten worden geformuleerd op basis van een beperkt aantal elementaire deeltjes en hun onderlinge interacties. Hierbij is materie opgebouwd uit fermionen die interageren door het uitwisselen van bosonen. De elektrodynamica wordt op die manier beschreven door de uitwisseling van massaloze fotonen. De sterke kernkracht, die ondermeer verantwoordleijk is voor het binden van protonen en neutronen in de kern, wordt op een analoge manier uitgewerkt. Hierbij worden massaloze gluonen uitgewisseld tussen de quarks waaruit kerndeeltjes zijn opgebouwd. De zwakke kernkracht, die verantwoordelijk is voor ondermeer radioactief verval wordt gemedieerd door massieve bosonen: het neutrale Z deeltje en de geladen W^+ en W^-. Al deze interacties en fundamentele bouwstenen worden op een consistente manier behandeld in het zogenaamde Standaard Model van de elementaire deeltjes.

Ondanks haar groot voorspellend vermogen en talloze experimentele verificaties is het Standaard Model nog niet volledig begrepen. Verschillende vragen blijven onbeantwoord. Waarom heeft het model zoveel onverklaarbare vrije parameters? Waarom kan gravitatie niet op dezelfde manier worden beschreven? Waarom hebben sommige elementaire deeltjes een massa en anderen niet? Hoe ontstaan deze massa's en waarom zijn ze zo verschillend van elkaar? Indien de fundamentele symmetrieën die verantwoordelijk zijn voor het bestaan van de drie voorgenoemde interacties perfect behouden zijn, dan zouden alle elementarie deeltjes massaloos moeten zijn. Dit probleem werd in 1964 aangepakt door Brout, Englert en Higgs. Zij stelden een mechanisme voor dat de elektrozwakke symmetrie breekt door de introductie van een extra veld, het BEH veld wiens kwantum, het zogenaamde Higgs boson, nog steeds niet experimenteel is waargenomen. Om ondermeer deze vraag te beantwoorden zijn gedurende de laatste decennia subatomaire processen met steeds toenemende precisie onderzocht met behulp van grote deeltjesversnellers die botsingen met extreem hoge energiedichtheden realiseren.

De Large Hadron Collider (LHC) is de nieuwste en meest krachtige deeltjesversneller waar dit soort onderzoek gebeurt. De LHC is in staat om bundels van ultra-relativistische protonen frontaal te laten botsen met een massamiddelpuntsenergie van maximaal 14 TeV. Eén van de doelstellingen van het onderzoeksprogramma aan de LHC is het vinden of het ontkrachten van het bestaan van het Higgs boson. In vergelijking met andere processen is de productie van het Higgs boson bij de LHC zeer zeldzaam. Bovendien zal het, afhankelijk van zijn massa die theoretisch niet voorspeld is, kunnen vervallen in verschillende eindtoestanden. Het CMS experiment, dat is opgesteld in één van de vier botsingspunten

bij de LHC, is ontworpen met een maximale gevoeligheid voor de detectie van vervalproducten afkomstig uit vervallende Higgs deeltjes. Het experiment werd in 2009 in gebruik genomen en vergaarde tot en met juni 2011 1.1 fb^{-1} aan geïntegreerde luminositeit bij een massamiddelpuntsenergie van 7 TeV.

De studie beschreven in dit werk concentreert zich op het isoleren van een welbepaalde vervalmodus van het Higgs boson, namelijk het verval naar twee geladen ijkbosonen die op hun beurt weer vervallen in een lepton (muon of elektron) en hun bijbehorende neutrino: $\mathrm{W}^+\mathrm{W}^- \to \ell^+\nu\ell^-\bar{\nu}$. Deze modus is het meest gevoelig voor zowel de ontdekking als de uitsluiting van het bestaan van het Higgs deeltje in een ruim massagebied tussen 120 GeV/c^2 en 200 GeV/c^2. Na het verkrijgen van een gegevensstaal dat verrijkt is met potentiële Higgs gebeurtenissen zijn we in staat om een bovenlimiet te berekenen op de werkzame doorsnede voor Higgs productie bij de LHC. Wanneer deze bovenlimiet lager is dan de door het Standaard Model voorspelde waarde, dan kunnen we het bestaan van een Standard Model Higgs boson uitsluiten. Met 1.1 fb^{-1} aan vergaarde luminositeit sluiten we dan ook het bestaan van het Higgs deeltje binnen het Standaard Model uit in een massagebied tussen 145 GeV/c^2 en 190 GeV/c^2 met een vertrouwensinterval van 95%. Indien we gebruik maken van neurale netwerken om het gegevensstaal nog sterker te verzuiveren merken we op dat het uit te sluiten massagebied een beetje vergroot. Onze limieten zijn compatibel met concurrerende teams binnen en buiten het CMS experiment. Het bekomen resultaat wijst er op dat, indien het Higgs deeltje bestaat, zijn massa moet beperkt zijn tot het nauwe gebied $115\ \mathrm{GeV/c^2} < \mathrm{m_H} < 145\ \mathrm{GeV/c^2}$.

Tijdens het schrijven van dit werk vergaarde het CMS experiment naar het einde van 2011 meer gegevens en werd het gegevensstaal vijf keer groter. Hiermee was het CMS experiment in staat, door combinatie van alle vervalmodi, het bestaan van het Higgs deeltje in het Standaard Model uit te sluiten voor Higgs massa's groter dan 127 $\mathrm{GeV/c^2}$. Bovendien werd in de gegevens een overschot aan vervallen in het massagebied rond 125 $\mathrm{GeV/c^2}$ gesignaleerd. Dit is mogelijk een eerste hint voor het bestaan van het Higgs deeltje. De significantie van het geobserveerde signaal bedraagt 2.5 standaardafwijkingen, hetgeen betekent dat er nog een geringe kans bestaat dat we met een statistische fluctuatie van de achtergrond te maken hebben.

Sinds april 2012 realiseert de LHC proton-proton botsingen bij een iets hogere massamiddelpuntsenergie van 8 TeV met zeer hoge bundelintensiteiten. Tegen het einde van 2012 zal het CMS experiment minstens 15 fb^{-1} aan gegevens verzamelen, corresponderend met een verdrievoudiging van de huidige beschikbare gegevens een een vertienvoudiging van het staal gebruikt in dit werk. Deze gegevens zouden voldoende moeten zijn om het vermoeden van het bestaan van het Higgs deeltje met een massa rond 125 $\mathrm{GeV/c^2}$ te bevestigen of te ontkrachten.

Contents

Introduction

At the microscopic level the laws of nature can be formulated in terms of fundamental elementary particles. Ordinary matter, that virtually composes everything around us, is made of fermions that interact via the exchange of bosons. Three of the four known fundamental interactions can be explained in this way. The electromagnetic interaction happens via the exchange of massless photons. It is responsible for most types of chemistry, and electrodynamic phenomena. The weak and the strong force were discovered later as they are not as manifest for macroscopic objects. The weak interaction is responsible for radioactive decays of heavy nuclei and the strong interaction is what binds protons and neutrons together in atomic nuclei. The nuclear forces are mediated respectively via the exchange of massive W,Z bosons and massless gluons. All these interactions are quantitatively described in the so-called Standard Model of particle physics.

In spite of its highly predictive power and the numerous precise test that it passed, the Standard Model is far from being completely understood. Without even mentioning the gravitational interaction, for which we do not have a quantum description yet, several questions remain open. Why are many Standard Model parameters theoretically unexplained? Why are there three replicated generations of elementary particles? Why do some particles have masses and some other do not and how do these masses arise in the Standard Model? If the fundamental symmetries that are responsible for the existence of the three main interactions are conserved, all particles should be massless which is not the case experimentally. This is the so-called electroweak symmetry breaking problem. In 1964 Brout, Englert and Higgs proposed a mechanism capable of elucidating this last question. In this mechanism, the mass of a particle arises when it interacts with an additionnal field called the BEH field. The fundamental quantum, the so-called BEH or Higgs boson, has not been found so far.

In order to answer this and many other questions subatomic processes have been probed with increasing precision over the last century with the help of large accelerators producing high energy collisions. The Large Hadron Collider is the largest and most powerful accelerator in the world build up to date. At the LHC beams of ultra-relativistic protons collide at a center of mass energy up to 14 TeV. One of the main goals of the LHC physics program is to assert the existence or the absence of the BEH boson. Compared to other known particles, the BEH boson is expected to be produced at a much lower rate at the LHC. Moreover, depending on its mass, which is an unknown parameter, it will decay into different final states. The CMS detector was designed in order to achieve a very high sensitivity to the possible final states originating from BEH boson decays.

In the present study efforts were concentrated on one particular decay mode of the BEH boson. One year of LHC data was analyzed and an upper limit on the production rate could be set for a large range of BEH mass hypothesis, with an exclusion for the Standard Model BEH particle in the mass range $145 \leq m_H \leq 190$ GeV/c^2. This thesis is organized in the following way and divided mainly in two parts. In the first part we will

discuss general aspects of the BEH boson and the experimental apparatus. The second part of this work will deal in greater detail with the search of the BEH particle in a particular decay channel.

In the first Chapter an introduction on the Standard Model of particle physics is given. An emphasis is put on the BEH mechanism. In Chapters 2 and 3 the various decays and production modes of the BEH boson are discussed. A short review of results from previous experiments as well as short summary on alternative approaches to the electroweak symmetry breaking problem are given. In Chapter 4 we will discuss the characteristics of both the LHC and the CMS detector.

In Chapter 5 we look at the final state of a BEH boson decaying to two W bosons and their respective decays to two leptons and two neutrinos. We will discuss the final state topology of the signal and its main backgrounds. We will show that, although a mass peak reconstruction is impossible due to the lack of kinematic constraints on the final state, a clever selection of the kinematical variables characterizing the final state can help in discriminating the signal from the main irreducible WW background. In Chapter 6 we will describe how the reconstruction of the particles relevant to this analysis is performed in CMS. Finally, In Chapter 7 and 8 a phase-space region that enhances the signal-to-background ratio will be progressively isolated and final upper limits on the BEH production rate with one year of LHC data will be obtained.

The present thesis is based on the following list of publications:

- CMS Collaboration, "Measurement of the W^+W^-production and search for the Higgs boson in pp collisions at $\sqrt{s}$ = 7 TeV", PHYS. LETT. B 1-2 Vol. 699(2011) 25;
- CMS Collaboration, "Search for the Higgs Boson Decaying to W^+W^-in the Fully Leptonic Final State" , CMS Physics Analysis Summary, HIG-11-003;

and internal CMS publications:

- "Search Strategy for a Standard Model Higgs Boson Decaying to Two W Bosons in the Fully Leptonic Final State", CMS AN-2008/039;
- "Study of the $t\bar{t}$ process as a Background for Higgs Boson Decay to WW", CMS AN-2008/106;
- "Study of the uncertainties due to PDFs for Higgs boson decay to WW and backgrounds", CMS AN-2009/129;
- "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;
- "Search for Higgs Boson Decays to Two W Bosons in the Fully Leptonic Final State $\sqrt{s}$ = 7 TeV", CMS AN-2010/411;
- "Search for Higgs Boson Decays to Two W Bosons in the Fully Leptonic Final State $\sqrt{s}$ = 7 TeV with 2011 data of CMS detector.", CMS AN-2011/148;

List of abbreviations

AK	Anti-k_T (jet algorithm)
ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC Apparatus
b-tagging	Set of techniques for b-jet identification
BEH	Brout-Englert-Higgs
BR	Branching Ratio
CaloJet	Calorimeter based jet
CERN	Conseil européen pour la recherche nucléaire
CC	Charged Current
CKM	Cabibbo-Kobayashi-Maskawa (matrix)
CMS	Compact Muon Solenoid
CMSSW	CMS Software
CPT	Charge conjugation, reflection (parity) and time-reversal transformations
CT	Counter Terms
CTEQ	Coordinated Theoretical/Experimental Project on Parton Distribution Functions
DAQ	Data acquisition
DGLAP	Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (equations)
d.o.f	degree of freedom
DA	Deterministic Annealing (vertexing algorithm)
DT	Drift Tube
DY	Drell-Yan (Z+jets)
ECAL	Electromagnetic CALorimeter
E_T^{miss}	Missing Transverse Energy
EW	Electro-weak
ggF	gluon-gluon-Fusion
HCAL	Hadronic CALorimeter
HERWIG	LO MC-generator
HLT	High Level Trigger
HqT	algorithm calculating the Higgs q_T (transverse momentum) at NLO+NNLL accuracy
HS	Higgs-Strahlung
IC	Iterative Cone (jet algorithm)
IP	Impact Parameter
IR	Infra-Red

LEP	Large Electron-Positron collider
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LINAC	Linear Accelerator
L1	Level 1 Trigger
L1Offset	Correction method on jet energy for pile-up removal (old)
L1FastJet	Correction method on jet energy for pile-up removal (new)
LL	Leading Logarithm
LO	Leading Order
MC	Monte Carlo
MC@NLO	NLO Monte Carlo generator
ME	Matrix Element
MET	E_T^{miss}
min.E_T^{miss}	min(proj.trk E_T^{miss}, proj.PF E_T^{miss})
MIP	Minimum Ionizing Particle
MLP	Multi-Layer-Perceptron (see NN)
MSSM	Minimal Supersymmetric extension of the Standard Model
MSTW	Martin-Stirling-Thorne-Watt Parton Distribution Functions
NC	Neutral Current
NN	Neural Network
PD	Primary Dataset
PDF	Parton Distribution Function
PFJet	Particle-Flow based jet
PFE_T^{miss}	Particle-Flow based missing transverse energy
POWHEG	NLO Monte Carlo generator
PPL	Parton-Parton Luminosity
proj.E_T^{miss}	Projected Missing Transverse Energy
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
PU	Pile Up
PV	Primary Vertex
PYTHIA	LO Monte Carlo generator
Q	charge
Q	test-statistics of a given hypothesis
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
RPC	Resistive Plate Chamber
SC	Super-Cluster (array of 5×5 ECAL crystals)
SM	Standard Model
SPS	Super-Proton Synchrotron
SQED	Scalar Quantum Electrodynamics
SUSY	Super-Symmetry

TC	Techi-Color
TCHE	Track Counting High Efficiency (b-tagging)
TCHP	Track Counting High Purity (b-tagging)
trk.E_T^{miss}	Missing Transverse Energy calculated with charged particles
UV	Ultra-Violet
vev	vacuum expectation value
VBF	Vector Boson Fusion
Y	hypercharge

Part I

The Standard Model and the experimental apparatus

Chapter 1

Theoretical Background

Various high energy physics experiments have shown with a high degree of confidence that ordinary matter is made of fermions. They appear in nature in two distinct categories. Electrons (e), muons (μ), taus (τ) and their associated neutrinos (ν_e, ν_μ, ν_τ) are called leptons. They interact via the weak and, if charged, the electromagnetic force. On the other hand the quarks (u, d, c, s, t, b) can also interact strongly. These three forces have been successfully described within a unified approach in the theory by Glashow, Salam and Weinberg known today as the Standard Model of particle physics [1–3]. In the Standard Model, the electromagnetic, weak and strong interaction are mediated respectively by the photon, the W and Z bosons, and the gluons .

In this chapter we will see how the Standard Model is formulated. First, we introduce the basic rules needed in order to construct the Lagrangian of a theory of relativistic particles in the context of quantum mechanics (Quantum Field Theory). We will see that gauge invariance is necessary in order to describe interacting particles; we will treat first the simple case of the U(1) abelian group and then move to SU(N) (non-abelian case). In particular we will see that a fully gauge invariant theory prohibits the existence of massive particles. We will then see how spontaneous symmetry breaking of a gauge theory in the simple case of scalar electrodynamics can solve this issue. In the second part of this chapter we will show the Lagrangian of Standard Model, by concentrating mainly on the electroweak sector. After this insight, we will then describe how to spontaneously break the $\mathrm{SU(2)_L} \times \mathrm{U(1)}$ to $\mathrm{U(1)_{em}}$ via the Brout-Englert-Higgs-Guralnik-Hagen-Kibble mechanism [4–6], in order to give mass to particles. Finally, we will review all the interactions and their strength within the Standard Model framework.

In the following for simplicity, instead of referring to the BEHGHK, or BEH boson (or equivalently the mechanism), we will often refer to as simply the Higgs boson (mechanism) or as the scalar.

1.1 Ingredients

1.1.1 Basic principles

In quantum field theories, fields, and in general any operator (such as the Hamiltonian or the Lagrangian), are described in terms of creation and annihilation operators that act on multi-particles states. The complete derivation of the *second quantization* of fields is out of scope here and can be found in [7,8]. In such theories, a given state is described by the

number of particles, their momenta, and eventually other quantum numbers (such as their spin, charge ..). A n-particle state can be noted as:

$$|\mathbf{p_1}, \alpha_1; \mathbf{p_2}, \alpha_2; ...; \mathbf{p_n}, \alpha_n\rangle \tag{1.1}$$

where p_i label the particle momenta and α_i their additional quantum numbers. The reason why we need to deal with multi-particle states in a relativistic theory of quantum mechanics is that according to the uncertainty principle, large energy fluctuation can occur in a very short time, allowing for the creation of large number of particles (since E = m [1]). This means that the number of particles of a system cannot be fixed in the equation of motion but has to be derived dynamically.

A general theory of relativistic fields is described by its Lagrangian density $\mathcal{L}$. $\mathcal{L}$ is a function of the fields ϕ_α, their derivatives $\partial_\mu\phi_\alpha$ and eventually of space-time itself [2]:

$$\mathcal{L} = \mathcal{L}(x, \phi_\alpha(x), \partial_\mu\phi_\alpha(x)) \tag{1.2}$$

Poincaré invariance

For the theory to be able to describe relativistic fields the Lagrangian needs to be a Lorentz scalar, i.e its expression must involve only explicitly covariant terms [3]. Furthermore, translational invariance (that also implies energy conservation) restricts the general form of $\mathcal{L}$ to:

$$\mathcal{L} = \mathcal{L}(\phi_\alpha(x), \partial_\mu\phi_\alpha(x)) \tag{1.3}$$

Renormalizability

Saying that a theory is renormalizable is equivalent to the statement that the physics of interest at energies $E \ll \Lambda$ is largely insensitive to the higher-energy (shorter distances) physics appropriate to the scale Λ. Translated in the language of quantum field theory this is ensured if there are enough measurable parameters in the theory able to absorb the ultraviolet divergences that appear at high order calculations in perturbative series. This general condition may be summarized by the requirement that all the parameters that appear in the Lagrangian must have a dimension [4] in powers of mass equal or greater than 0. For instance a theory of complex scalar fields given by:

$$\mathcal{L} = -(\partial_\mu\phi)^\dagger(\partial^\mu\phi) - a\phi^\dagger\phi - b(\phi^\dagger\phi)^2 \tag{1.4}$$

is renormalizable, since $[\phi] = 1$, $[\mathcal{L}] = 4$ and therefore $[b] = 0$.

[1] We adopt the unit system: $\hbar = c = 1$
[2] We choose the metric $\eta_{\mu\nu} = diag[-, +, +, +]$
[3] It is also convenient to chose the fields $\phi_\alpha(x)$ to transform in representations of the Lorentz group
[4] The dimension in powers of mass of a given quantity is indicated by the notation $[\ldots]$

1.1.2 Gauge Fields

So far, we have introduced general rules that guide us in guessing the form of the Lagrangian of a non-specific theory. Let us consider now the Lagrangian of a free fermion field. The Dirac Lagrangian density is given by:

$$\mathcal{L}_{Dirac}^{Free} = -\bar{\psi}(\partial\!\!\!/ + m)\psi \tag{1.5}$$

where $\partial\!\!\!/ \equiv \gamma^\mu \partial_\mu$, [5] and $\bar{\psi} = \psi^\dagger i\gamma^0$. In the two following subsections we will show that the invariance of the Lagrangian under a group of local transformations implies the existence of a set of fields interacting with the Dirac field. We will first start with the simple abelian group U(1) and then move to the SU(N) non abelian case.

U(1) gauge interaction: Quantum Electrodynamics

First consider the following U(1) local transformation on the Dirac field:

$$\psi(x) \rightarrow e^{iq\alpha(x)}\psi(x) \tag{1.6}$$

The function $\alpha(x)$ is a real phase and depends on the space-time position (this is why the symmetry transformation is said to be *local*) and q is a small real number. Then the Lagrangian density in (1.5) is clearly not invariant under the transformation in (1.6) because of the term involving the derivative of ψ. The mass term is indeed invariant under such transformation. Consider now the existence of a vector field $A_\mu(x)$ (spin 1) and define the covariant derivative as:

$$D_\mu \equiv \partial_\mu - iqA_\mu(x) \tag{1.7}$$

Then, provided that while the Dirac field $\psi(x)$ transforms as 1.6 the vector field $A_\mu(x)$ transform as:

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \alpha(x) \tag{1.8}$$

the Lagrangian

$$\mathcal{L}_{Dirac}^{Int} = -\bar{\psi}(\not{D} + m)\psi \tag{1.9}$$

is invariant under the U(1) local transformation (1.6). Now, in order to have a full description of the dynamics of fields, we have to add a kinetic term for the gauge field which has to be a function of the Maxwell tensor $F_{\mu\nu}$, written in a covariant way. We can then write the Quantum Electrodynamics (QED) Lagrangian, in an expanded form:

$$\begin{aligned} \mathcal{L}_{QED} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \bar{\psi}(\partial\!\!\!/ + m)\psi + iq\bar{\psi}\gamma^\mu A_\mu \psi \\ &= \mathcal{L}_{Maxwell} + \mathcal{L}_{Dirac} + \mathcal{L}_{Int} \end{aligned} \tag{1.10}$$

[5] The γ^μ matrices satisfy the Clifford algebra anti-commutation relations $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$. The choice of the metric imposes the following choice for the γ^μ matrices: $\gamma^0 = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$, $\gamma^i = \begin{pmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{pmatrix}$, where σ_k are the Pauli matrices.

To summarize, imposing local U(1) gauge invariance is equivalent to adding a spin 1 massless field to the description of the system. This particle can be identified with the photon and appears to be coupled to the fermion with a coupling strength equal to the charge q.

It is worth stressing at this point that the gauge field (i.e the photon) appears to be massless in this framework. As a matter of fact, a mass term of the form:

$$\mathcal{L}_{Mass}^{Gauge} = \frac{1}{2} m_A^2 A_\mu A^\mu \tag{1.11}$$

would manifestly violate gauge invariance (to see this explicitly substitute A_μ by (1.8) in by (1.11)).

SU(N) gauge interaction: non-abelian case

The goal of this subsection is to explore the possibility of adding more vector particle interactions. The idea of non-abelian gauge symmetries, even though first proposed by Pauli, was formalized by Yang and Mills in 1954 [9]. Let us look at the case where the complex Dirac field in (1.6) becomes an n-plet and transforms as:

$$\psi(x) \rightarrow V(x)\psi(x) \tag{1.12}$$

where $V(x)$ is an element of the SU(N) group [6]. V can be expanded in the neighborhood of unity by using $N^2 - 1$ hermitian matrices called the *group generators* [7]:

$$V(x) = 1 + ig\alpha^a(x)t^a + \mathcal{O}(\alpha^2) \tag{1.13}$$

The commutation relations between the generators can be generally written as:

$$[t^a, t^b] = if^{abc}t^c \tag{1.14}$$

where f^{abc} are the *structure constants* of the group (this tensor is completely antisymmetric). Proceeding by analogy with the U(1) case we can define the covariant derivative as:

$$D_\mu \equiv \partial_\mu - igA_\mu^a(x)t^a \tag{1.15}$$

introducing as many vector fields as generators. If we write the infinitesimal transformation laws for ψ and A_μ^a as:

$$\begin{aligned} &\psi(x) \rightarrow (1 + ig\alpha^a(x)t^a)\psi(x) \\ &A_\mu^a(x) \rightarrow A_\mu^a(x) + \partial_\mu\alpha^a(x) - gf^{abc}A_\mu^b(x)\alpha^c(x) \end{aligned} \tag{1.16}$$

then we obtain, as in the previous subsection, that $D_\mu\psi$ transforms the same way as ψ. For the vector field tensor term in the Lagrangian to be invariant under (1.16) it has to be defined as:

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c \tag{1.17}$$

Finally, we can write down the gauge invariant Yang-Mills Lagrangian:

$$\mathcal{L}_{Yang-Mills} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} - \bar{\psi}(\not{\partial} + m)\psi + ig\bar{\psi}\gamma^\mu A_\mu^a t^a \psi \tag{1.18}$$

[6] SU(N) is a Lie group consisting of all $N \times N$ hermitian matrices satisfying $det(U) = 1$

[7] For N=2,3 the generators are respectively the 3 Pauli and the 8 Gell-Mann matrices

This Lagrangian describes the equations of motion of a fermion coupled to N^2-1 vector fields with a coupling strength g. The main difference with U(1) case, besides the number of gauge fields, is its non-abelian nature parameterized by (1.14). These additional terms in the field tensor $F^a_{\mu\nu}$ introduce cubic and quartic terms in A^a_μ in $\mathcal{L}$ leading to self-interacting gauge fields.

1.1.3 Spontaneous symmetry breaking

In the previous section we have learned how to build a theory of interacting fermions and vector fields by requiring gauge invariance. Nevertheless we have seen that in such theories the masses of gauge bosons vanish in order for the Lagrangian to stay symmetric under such gauge transformations. We will review here a mechanism that allows to *dynamically* generate massive bosons by spontaneously breaking the original invariance.

Goldstone theorem: global symmetry

The Goldstone theorem applies whenever a system is invariant under a given symmetry but the ground state of the theory is not. The symmetry in this case is said to be *spontaneously broken*. In this case, states the theorem, there must exist one massless boson, scalar or pseudo-scalar, associated to each generator that does not annihilate the vacuum. This theorem, first guessed by Nambu [10] and then formalized by Goldstone [11,12] is of crucial importance in several physical contexts, ranging from solid state physics, to fluid dynamics, to particle physics). We will see here how it applies to a simple example. Consider a theory of a complex scalar that is U(1) global invariant (similar to that given by (1.4)):

$$\mathcal{L} = -(\partial_\mu\phi)^*(\partial^\mu\phi) - \lambda\big[\phi^*\phi - \frac{\mu^2}{2\lambda}\big]^2 \tag{1.19}$$

The potential $V = \lambda\big[\phi^*\phi - \frac{\mu^2}{2\lambda}\big]^2$ is shown in Figure 1.1 for different sign combinations of the parameters λ and μ. In addition we require that the potential to be bounded from below ($\lambda > 0$). For the ground state not to be U(1) invariant μ^2 must also be positive. We end up with the *Mexican hat* potential shown in Fig. 1.1(a). The ground state defined by $|\phi| = \sqrt{\frac{\mu^2}{2\lambda}} = v$ or $\phi = ve^{i\theta}$ is clearly not U(1) invariant. Now consider a small perturbation around the vacuum parameterized respectively by two small real fields $\sigma(x)$ and $\eta(x)$ in the real and in the complex direction so that $\phi = v + \sigma + i\eta$. The new theory, neglecting the interaction terms, gets rephrased as:

$$\mathcal{L}_{broken} = -(\partial_\mu\sigma)(\partial^\mu\sigma) - 4\lambda v^2\sigma^2 - (\partial_\mu\eta)(\partial^\mu\eta) \tag{1.20}$$

We ended up with a massive field, σ, that corresponds to an excitation that *costs* energy (the radial direction in Figure 1.1(a)) and a massless field, η, that relates to the *cost-less* excitation around the vacuum minima. η is therefore the massless Goldstone boson associated to the U(1) symmetry breaking.

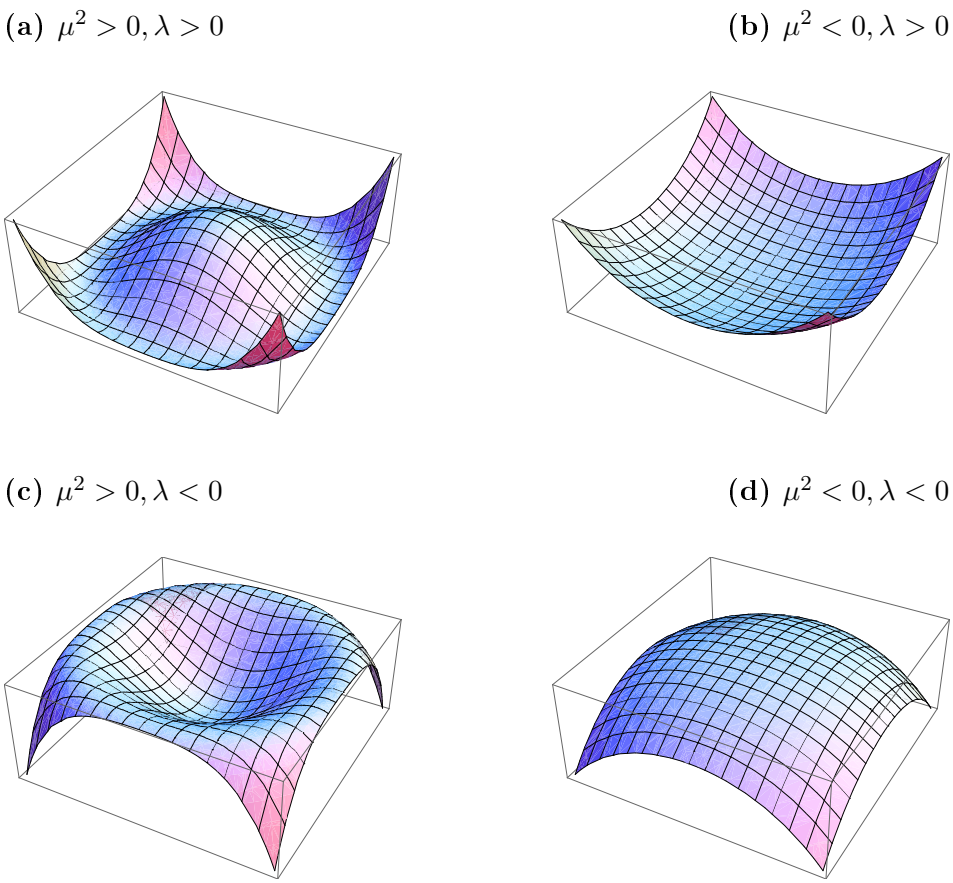


Figure 1.1: Shape of the Higgs potential as a function of the components of the complex scalar field. The four possible sign combinations for the parameters μ and λ. In (b) the ground state is trivial, in (a) there is an infinite number of them, and they all violate the Lagrangian U(1) symmetry. (c) and (d) are non-stable because the resulting Hamiltonian is not bound from below.

The BEHGHK mechanism: local symmetry

In 1964 BEHGHK [8] had the idea to apply the Nambu-Goldstone theorem to gauge symmetries [4–6]. They discovered that unlike in the global symmetry case the Goldstone bosons were *absorbed* by the longitudinal degrees of freedom of the vector bosons, allowing to provide them with a mass. We discuss here the instructive case of *scalar electrodynamics* (SQED) in order to prepare the ground for the electroweak symmetry breaking mechanism. SQED is the simplest extension of QED one can imagine. One single complex scalar is added to the theory, this field transforming under the U(1) local gauge group as:

$$\delta\phi = iq\alpha(x)\phi \tag{1.21}$$

while the gauge field A_μ transforms as previously as (1.8). The covariant derivative acting on ϕ is then given by

$$D_\mu\phi = \partial_\mu\phi - iqA_\mu\phi \tag{1.22}$$

[8] An acronym for: Brout, Englert, Higgs, Guralnik, Hagen, Kibble

The SQED Lagrangian is then written as (we do not consider the fermionic part here):

$$\begin{aligned}\mathcal{L}_{SQED} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - D_\mu\phi D^\mu\phi^* - V(\phi^*\phi) \\ &= \mathcal{L}_{QED} + \mathcal{L}_{Higgs}\end{aligned} \tag{1.23}$$

The vacuum expectation value v (vev) is, as in the previous case not invariant under the symmetry. However the difference with respect to the global symmetry case is that the covariant derivative acting on ϕ will mix up terms containing the Goldstone boson and the vector field A_μ. To see how the vector acquires a mass we choose the unitary gauge ($\phi^* = \phi$) and expand in powers of $h = \phi - v$. We then realize that all the cross terms vanish. We end up with (again neglecting higher order terms)

$$\mathcal{L}_{broken} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - (\partial_\mu h)(\partial^\mu h) - 4\lambda v^2 h^2 - q^2 v^2 A_\mu A^\mu \tag{1.24}$$

With this procedure the photon has become massive. We started with 2 degrees of freedom (d.o.f) for the photon (massless spin-1 particle) plus 2 d.o.f for the scalar (complex scalar). We ended up with a massive photon (3 d.o.f) plus 1 massive real scalar (1 d.o.f). We have *transfered* the d.o.f of the Goldstone boson to the longitudinal d.o.f of the photon. This example is purely instructive: indeed the photon is known to be massless, and as we will see later, the U(1) group stays unbroken in the Standard Model.

1.2 Building the Standard Model

In the previous section we have reviewed the general concepts that are necessary to build a quantum theory of interacting fermions and vector bosons, and studied a procedure that allows to generate a mass term for the gauge bosons. Here we will specifically write down what the field content of the Standard Model is. We will derive the most general set of interactions compatible with the elementary particles that have been observed up to now.

1.2.1 Particle Content

The three main forces [9] (the strong, the electroweak and the electromagnetic) are understood as due to the exchange of spin-1 particles between fermions. We start from the assumption that the gauge group is $SU(3)_c \times SU(2)_L \times U(1)_Y$. The eight spin-1 bosons associated with $SU(3)_c$ [10] are the gluons and are denoted by G^a_μ ($a = 1..8$). Gluons are thought to be massless, they carry a color index (not explicitated here) and they couple to quark and themselves, since $SU(3)$ is a non-abelian group. They are the bosons related to the strong interaction. The three spin-1 particles associated with $SU(2)_L$, W^b_μ ($b = 1..3$), and with $U(1)_Y$, B_μ, are related via the Brout-Englert-Higgs mechanism (detailed in the next section) to the electroweak and electromagnetic force. The subscript 'L' indicates that only the left-handed fermions transform under $SU(2)_L$, while the Y subscript is the so-called *hypercharge*, the quantum number conserved via the B_μ exchange, to be distinguished from the *electric charge* Q. They are in fact related via the Gell-Mann and Nishijma relation, $Q = I_3 + Y$, where I is the weak isospin (this equality will be derived later).

[9] We do not discuss the gravitational interaction here, which is an entire subject on its own

[10] The subscript c denotes *color*

	family I	family II	family III
leptons	$\nu_e \leq 3$ eV $e \sim 0.51$ MeV	$\nu_\mu \leq 0.19$ MeV $\mu \sim 106$ MeV	$\nu_\tau \leq 18$ MeV $\tau \sim 1.78$ GeV
quarks	$u \sim 7$ MeV $d \sim 3$ MeV	$c \sim 1.2$ GeV $s \sim$115 MeV	$t \sim 175$ GeV $b \sim 4.25$ GeV

Table 1.1: List of known fermions and their masses divided in three families

As stated before, the basic constituents of matter are fermions and they appear in nature in three "replicated" families [11]. Each family is divided in leptons and quarks. Leptons do not interact via the strong force (i.e they do not exchange gluons) while quarks do. Nevertheless both are sensitive to the electroweak interaction. The list of fermions and their measured mass is summarized in the Table 1.1.

Since the $SU(2)_L$ group *picks* only the left-handed component of a field, it is convenient to decompose the Dirac spinor into two Majorana spinors. So if e is the Dirac spinor of the electron field, then we can write:

$$e = \begin{pmatrix} e_L \\ e_R \end{pmatrix} \tag{1.25}$$

$$= P_L\mathcal{E} + P_R E \tag{1.26}$$

where P_L and P_R are the projectors on the left/right components of the spinor while $\mathcal{E}$ and E are respectively the Majorana spinors built from the left and right parts of e:

$$\mathcal{E} = \begin{pmatrix} e_L \\ i\sigma_2 e_L^* \end{pmatrix}, \quad E = \begin{pmatrix} -i\sigma_2 e_R^* \\ e_R \end{pmatrix} \tag{1.27}$$

On the other hand, since neutrinos show up in nature as pure left-handed particles, we label the corresponding Majorana spinor by ν (equivalent to $\mathcal{E}$ for the electron). We note by L_m and Q_m the left-handed $SU(2)_L$ doublets respectively corresponding to leptons and quarks:

$$L_m = \begin{pmatrix} \nu_m \\ \mathcal{E}_m \end{pmatrix}, \quad Q_m = \begin{pmatrix} \mathcal{U}_m \\ \mathcal{D}_m \end{pmatrix} \tag{1.28}$$

and by E_m, U_m, D_m the corresponding right-handed singlets. The transformation properties of the fermions under the different gauge groups can be summarized in a compact

[11] Families will be labeled by the index m=1..3

form by giving the dimension of their representation:

$$\begin{array}{rcll} P_L L_m & = & \begin{pmatrix} P_L\nu_m \\ P_L\mathcal{E}_m \end{pmatrix} & \rightarrow \quad (\mathbf{1},\mathbf{2},-\frac{1}{2}) \\ P_L Q_m & = & \begin{pmatrix} P_L\mathcal{U}_m \\ P_L\mathcal{D}_m \end{pmatrix} & \rightarrow \quad (\mathbf{3},\mathbf{2},+\frac{1}{6}) \\ P_R E_m & & & \rightarrow \quad (\mathbf{1},\mathbf{1},-1) \\ P_R U_m & & & \rightarrow \quad (\mathbf{3},\mathbf{1},-\frac{2}{3}) \\ P_R D_m & & & \rightarrow \quad (\mathbf{3},\mathbf{1},-\frac{1}{3}) \\ G^a_\mu & & & \rightarrow \quad (\mathbf{8},\mathbf{1},0) \\ W^b_\mu & & & \rightarrow \quad (\mathbf{1},\mathbf{3},0) \\ B_\mu & & & \rightarrow \quad (\mathbf{1},\mathbf{1},0) \end{array} \tag{1.29}$$

The first two numbers in parenthesis are the dimensions of the $SU(3)_c$ and $SU(2)_L$ representations, and the third number is the eigenvalue of the generator Y of the $U(1)_Y$ group [12]. This symbolic notation is a condensed way of writing down the full transformation of the fields (to be compared with (1.16)):

$$\begin{aligned} \delta L_m &= \big[(-\frac{i}{2}g_1\omega_1(x) + \frac{i}{2}g_2\omega_2^a(x)\sigma_a)P_L + (\frac{i}{2}g_1\omega_1(x) - \frac{i}{2}g_2\omega_2^a(x)\sigma_a^*)P_R\big]L_m \\ \delta Q_m &= \big[(\frac{i}{6}g_1\omega_1(x) + \frac{i}{2}g_2\omega_2^a(x)\sigma_a + \frac{i}{2}g_3\omega_3^a(x)\lambda_a)P_L \\ &\quad +(-\frac{i}{6}g_1\omega_1(x) - \frac{i}{2}g_2\omega_2^a(x)\sigma_a^* - \frac{i}{2}g_3\omega_3^a(x)\lambda_a^*)P_R\big]Q_m \\ \delta E_m &= [ig_1\omega_1(x)P_L - ig_1\omega_1(x)P_R]E_m \\ \delta U_m &= \big[(-\frac{2i}{3}g_1\omega_1(x) - \frac{i}{2}g_3\omega_3^a(x)\lambda_a^*)P_L + (\frac{2i}{3}g_1\omega_1(x) + \frac{i}{2}g_3\omega_3^a(x)\lambda_a)P_R\big]U_m \\ \delta D_m &= \big[(\frac{i}{3}g_1\omega_1(x) - \frac{i}{2}g_3\omega_3^a(x)\lambda_a^*)P_L + (-\frac{i}{3}g_1\omega_1(x) + \frac{i}{2}g_3\omega_3^a(x)\lambda_a)P_R\big]D_m \\ \delta G^a_\mu &= \partial_\mu\omega_3^a(x) - g_3 f^a_{bc}\omega_3^b(x)G^c_\mu \\ \delta W^a_\mu &= \partial_\mu\omega_2^a(x) - g_2\epsilon^a_{bc}\omega_2^b(x)W^c_\mu \\ \delta B_\mu &= \partial_\mu\omega_1(x) \end{aligned} \tag{1.30}$$

where ω_1, ω_2^a and ω_3^b are the real local phases introduced by a general $SU(3)\times SU(2)\times U(1)$ gauge transformation.

1.2.2 The Lagrangian

Now that the particle content and the field transformations have been shown, we can write down the gauge invariant, renormalizable Lagrangian involving those fields [13]:

$$\begin{aligned} \mathcal{L}_{f-g} &= -\frac{1}{4}G^\alpha_{\mu\nu}G^{\alpha\mu\nu} - \frac{1}{4}W^\alpha_{\mu\nu}W^{\alpha\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ &\quad -\frac{1}{2}\bar{L}_m\not{D}L_m - \frac{1}{2}\bar{E}_m\not{D}E_m - \frac{1}{2}\bar{Q}_m\not{D}Q_m - \frac{1}{2}\bar{U}_m\not{D}U_m - \frac{1}{2}\bar{D}_m\not{D}D_m \end{aligned} \tag{1.31}$$

[12] It is worth noting that since the left and right parts of a Majorana spinor are complex conjugate of one another (see 1.27) the right part transforms in the adjoint representation, for instance $P_R Q_m$ transforms as ($\bar{\mathbf{3}}$,$\mathbf{2}$,-$\frac{1}{6}$)

[13] Even though compatible with gauge invariance, we will ignore the terms $-\frac{g_3^2\Theta_3}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}G^{\alpha\mu\nu}G^{\alpha\lambda\rho} - \frac{g_2^2\Theta_2}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}W^{\alpha\mu\nu}W^{\alpha\lambda\rho} - \frac{g_1^2\Theta_1}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}B^{\mu\nu}B^{\lambda\rho}$, that are not relevant for the following discussion.

where:

$$
\begin{aligned}
G^{\alpha}_{\mu\nu} &= \partial_\mu G^{\alpha}_{\nu} - \partial_\nu G^{\alpha}_{\mu} + g_3 f^{\alpha}_{\beta\gamma} G^{\beta}_{\mu} G^{\gamma}_{\nu} && (1.32)\\
W^{\alpha}_{\mu\nu} &= \partial_\mu W^{\alpha}_{\nu} - \partial_\nu W^{\alpha}_{\mu} + g_2 \epsilon_{\alpha\beta\gamma} W^{\beta}_{\mu} W^{\gamma}_{\nu} && (1.33)\\
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu && (1.34)
\end{aligned}
$$

and the covariant derivatives are:

$$
\begin{aligned}
D_\mu L_m &= \partial_\mu L_m + \left[\frac{i}{2} g_1 B_\mu - \frac{i}{2} g_2 W^{\alpha}_{\mu} \sigma_a\right] P_L L_m \\
&\quad + \left[-\frac{i}{2} g_1 B_\mu + \frac{i}{2} g_2 W^{\alpha}_{\mu} \sigma^*_a\right] P_R L_m && (1.35)\\
D_\mu Q_m &= \partial_\mu Q_m + \left[-\frac{i}{6} g_1 B_\mu - \frac{i}{2} g_2 W^{\alpha}_{\mu} \sigma_a - \frac{i}{2} g_3 G^{\alpha}_{\mu} \lambda_a\right] P_L Q_m \\
&\quad + \left[\frac{i}{6} g_1 B_\mu + \frac{i}{2} g_2 W^{\alpha}_{\mu} \sigma^*_a + \frac{i}{2} g_3 G^{\alpha}_{\mu} \lambda^*_a\right] P_R Q_m && (1.36)\\
D_\mu E_m &= \partial_\mu E_m - i g_1 B_\mu P_L E_m + i g_1 B_\mu P_R E_m && (1.37)\\
D_\mu U_m &= \partial_\mu U_m + \left[-\frac{2i}{3} g_1 B_\mu - \frac{i}{2} g_3 G^{\alpha}_{\mu} \lambda_a\right] P_R U_m \\
&\quad + \left[\frac{2i}{3} g_1 B_\mu + \frac{i}{2} g_3 G^{\alpha}_{\mu} \lambda^*_a\right] P_L U_m && (1.38)\\
D_\mu D_m &= \partial_\mu D_m + \left[\frac{i}{3} g_1 B_\mu - \frac{i}{2} g_3 G^{\alpha}_{\mu} \lambda_a\right] P_R D_m \\
&\quad + \left[-\frac{i}{3} g_1 B_\mu + \frac{i}{2} g_3 G^{\alpha}_{\mu} \lambda^*_a\right] P_L D_m && (1.39)
\end{aligned}
$$

Note that this theory describes massless fermions and bosons. Unlike the case of equation (1.18) where gauge invariance only forbids a mass term for the gauge bosons, here there is neither a mass term for the fermions. Such a term would in fact also violate gauge invariance. To see why, let us write down what a mass term for an electron would be:

$$
\begin{aligned}
\mathcal{L}_{would-be} &= m_e \bar{e} e \\
&= m_e (\bar{e} P_R e + \bar{e} P_L e) \\
&= m_e (\bar{\mathcal{E}} P_R E + \bar{E} P_L \mathcal{E}) && (1.40)
\end{aligned}
$$

Let's look for instance at $\bar{\mathcal{E}} P_R E$. $P_R E$ is a singlet under $SU(2)_L$ and has $Y = -1$. On the other hand, $\bar{\mathcal{E}} P_R = \mathcal{E}^\dagger \beta P_L = \mathcal{E}^\dagger P_R \beta$ is the conjugate of $P_L \mathcal{E}$ so it has a non-trivial transformation under $SU(2)_L$ and $Y = +\frac{1}{2}$. The same argument applies for the other term in sum. Left and right components have different gauge transformation laws under $SU(2)_L \times U(1)_Y$ the mass term (1.40) cannot be gauge invariant. It is therefore clear at this point that the Lagrangian defined in (1.32) cannot be complete because it describes a theory of massless fermions and bosons.

We will use the idea introduced in Section 1.1.3 to generate masses. Therefore we postulate the existence of one $SU(2)_L$ doublet of complex scalar ϕ that transforms as (we note by $\tilde{\phi}$ its complex conjugate):

$$
\begin{aligned}
\phi &= \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} && \rightarrow \quad (\mathbf{1}, \mathbf{2}, \tfrac{1}{2}) \\
\tilde{\phi} &= i\sigma_2 \phi^* \quad \begin{pmatrix} \phi^{0*} \\ -\phi^{+*} \end{pmatrix} && \rightarrow \quad (\mathbf{1}, \mathbf{2}, -\tfrac{1}{2})
\end{aligned} \qquad (1.41)
$$

The terms that define the couplings with the gauge bosons and produce their mass terms are given by the covariant derivatives. In addition the scalar doublet needs to be coupled to the fermions (which will ultimately produce the masses for the fermions). We can therefore write the most general covariant and gauge invariant Higgs Lagrangian:

$$\begin{aligned}\mathcal{L}_{Higgs} &= -(D_\mu\phi)(D^\mu\phi^\dagger) - V(\phi^\dagger\phi) \\ &\quad -(f_{mn}\bar{L}_m P_R E_n\phi + h_{mn}\bar{Q}_m P_R D_n\phi + g_{mn}\bar{Q}_m P_R U_n\tilde{\phi} + h.c.) \end{aligned} \tag{1.42}$$

A few comments on the choice of the scalar field:

We know that we want to generate via the spontaneous symmetry breaking mechanism 3 masses for the gauge bosons $W^\pm$ and Z^0 and leave the photon massless. This means that the vacuum needs to break 3 generators of the $SU(2)_L \times U(1)_Y$ symmetry while leaving $U(1)_{em}$ unbroken. This also means that we need at least 3 d.o.f (the massless Goldstone bosons) to be absorbed by the longitudinal d.o.f of the $W^\pm$ and Z^0. Also, a gauge invariant term must be built in order to give masses to the fermions via the Yukawa terms. Now, in order to make a SU(2) singlet out of a doublet (the left-handed fermion) and a singlet (the right-handed fermion) the simplest choice appears to add in between a doublet [14]. We also want the Yukawa potential to be a singlet under $U(1)_Y$. This restricts the choice of the charges of the Higgs component. Consider for instance the term $\bar{L}_m P_R E_n\phi$. $P_R E_n$ has $Y = -1$ while $\bar{L_m} P_R = L_m^\dagger P_L\beta$ has $Y = \frac{1}{2}$. This implies that ϕ must be $Y = \frac{1}{2}$ and subsequently that its $SU(2)_L$ components must respectively charged and neutral. We also needed to define the conjugate $\tilde{\phi}$ such that its hypercharge is the opposite to ϕ. This is necessary for the term involving the up quarks to be hypercharge neutral as well.

Now we have to define the potential $V(\phi^\dagger\phi)$ such that the theory remains gauge invariant and renormalizable: [15]

$$V(\phi^\dagger\phi) = \lambda\big[\phi^\dagger\phi - \frac{\mu^2}{2\lambda}\big]^2 \tag{1.43}$$

$$= -\frac{\mu^2}{4\lambda} - \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{1.44}$$

To define completely the scalar sector we also need to know the gauge transformation laws:

$$\delta\phi = \frac{i}{2}\omega_2^a\sigma^a\phi + \frac{i}{2}\omega_1\phi \tag{1.45}$$

$$D_\mu\phi = \partial_\mu\phi - \frac{i}{2}g_1 B_\mu\phi - \frac{i}{2}g_2 W_\mu^\alpha\sigma_a\phi \tag{1.46}$$

We can finally write down the complete Standard Model Lagrangian:

$$\mathcal{L}_{SM} = \mathcal{L}_{f-g} + \mathcal{L}_{Higgs} \tag{1.47}$$

Note that $\mathcal{L}_{Higgs}$ parameterizes most of our ignorance in the theory. $\mathcal{L}_{f-g}$ contains four parameters (the 3 gauge couplings), while $\mathcal{L}_{Higgs}$ contains 10 masses and the 4 CKM matrix parameters (that will be defined later).

[14] Combine a spin $\frac{1}{2}$ particle with a spin 0. The only way to produce a spin 0 is to add another spin $\frac{1}{2}$. Of course one could make a more complicated choice, i.e adding 3 spin $\frac{1}{2}$ (adding 3 doublets), or one spin $\frac{1}{2}$ and one spin 1 (one doublet and one triplet)

[15] Gauge invariance imposes only terms of the form $f(\phi^\dagger\phi)$, renormalizability forbids terms with a power higher than 4.

1.2.3 The Electroweak symmetry breaking mechanism

The gauge invariance of $\mathcal{L}_{Higgs}$ allows us to choose a particular gauge to derive the spectrum of the theory. We choose the so-called unitary gauge, in which the field ϕ is defined by:

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{(2)}}(v + H(x)) \end{pmatrix} \tag{1.48}$$

where H is a real scalar field and v is the norm of ϕ that minimizes the potential. It is possible to show that any $SU(2)_L$ rotation on ϕ will produce a general $SU(2)_L$ doublet. We note by the way that no field configuration with $H(x) \neq -v$ is gauge invariant. The vacuum expectation value v is found by minimizing the potential V:

$$v^2 = \frac{\mu^2}{\lambda} \tag{1.49}$$

In order to identify the new theory spectrum and interactions, we need to expand $\mathcal{L}_{Higgs}$.The kinetic term equals to:

$$\begin{aligned} -(D_\mu\phi)^\dagger(D^\mu\phi) &= -\frac{1}{2}\partial_\mu H\partial^\mu H - \frac{1}{8}(v+H)^2 g_2^2(W_\mu^1 - iW_\mu^2)(W^{1\mu} + iW^{2\mu}) \\ &\quad -\frac{1}{8}(v+H)^2(g_1B_\mu - g_2W_\mu^3)(g_1B^\mu - g_2W^{3\mu}) \end{aligned} \tag{1.50}$$

the potential V to:

$$V = \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4}H^4 \tag{1.51}$$

and after expanding the Yukawa terms we can write the expanded $\mathcal{L}_{Higgs}$:

$$\begin{aligned} \mathcal{L}_{Higgs} &= -\frac{1}{2}\partial_\mu H\partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4 \\ &\quad -\frac{1}{8}(v+H)^2 g_2^2 |W_\mu^1 - iW_\mu^2|^2 \\ &\quad -\frac{1}{8}(v+H)^2(g_1B_\mu - g_2W_\mu^3)^2 \\ &\quad -\frac{1}{\sqrt{2}}(v+H)[f_{mn}\bar{\mathcal{E}}_m P_R E_n + h.c.] \\ &\quad -\frac{1}{\sqrt{2}}(v+H)[g_{mn}\bar{\mathcal{U}}_m P_R U_n + h.c.] \\ &\quad -\frac{1}{\sqrt{2}}(v+H)[h_{mn}\bar{\mathcal{D}}_m P_R D_n + h.c.] \end{aligned} \tag{1.52}$$

1.2.4 Mass spectrum

So far we have seen that the invariance of the Lagrangian under the gauge group transformations implies vanishing masses for fermions and bosons. In order to solve this issue, we have added to the theory an additional scalar doublet whose ground state violates gauge invariance. In this section we will see that this spontaneous symmetry breaking allows to generate the masses of bosons and fermions.

Bosons

The first boson emerging in this new theory is of course the real scalar H. Its mass can be easily read from $\mathcal{L}_{Higgs}$ by identifying a term of the form $\frac{1}{2}m_H H^2$. We find:

$$m_H^2 = 2\lambda v^2 = 2\mu^2 \tag{1.53}$$

Let us now consider the term $-\frac{1}{8}g_2^2 v^2 |W_\mu^1 - iW_\mu^2|^2$. By comparing it to $-\frac{1}{2}M_1^2 W_\mu^1 W^{1\mu}$ and $-\frac{1}{2}M_2^2 W_\mu^2 W^{2\mu}$ we have

$$M_1^2 = M_2^2 = \frac{g_2^2 v^2}{4} \tag{1.54}$$

It is natural to ask why these masses are equal, since this might be related to some symmetry that leaves the vacuum invariant. By solving:

$$\delta \begin{pmatrix} 0 \\ v \end{pmatrix} = (\frac{i}{2}\omega_2^a \sigma^a + \frac{1}{2}\omega_1) \begin{pmatrix} 0 \\ v \end{pmatrix} = 0 \tag{1.55}$$

We find $\omega_2^1 = \omega_2^2 = 0$ and $\omega_1 = \omega_2^3$. So the transformation that leaves the vacuum invariant is $I_3 + Y = Q$. It is said that the $SU(2)_L \times U(1)_Y$ is broken to the $U(1)$ electromagnetic group $U(1)_{em}$. Let us see how the W^a transform under $U(1)_{em}$. By using (1.30) we find that:

$$\delta \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \end{pmatrix} = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \end{pmatrix} = 0 \tag{1.56}$$

The eigenvectors of this transformations are the well known W^+, W^- bosons defined by:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{1.57}$$

the mass term now written as a function of W^+ and W^- is:

$$M_W^2 W_\mu^+ W^{-\mu} \quad \text{where} \quad M_W^2 = \frac{g_2^2 v^2}{4} \tag{1.58}$$

By looking at $-\frac{1}{8}v^2(g_1 B^\mu - g_2 W^{3\mu})^2$ we may now define the vector boson Z_μ as

$$\begin{aligned} Z_\mu &= \frac{g_2 W_\mu^3 - g_1 B_\mu}{\sqrt{g_1^2 + g_2^2}} \\ &= \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \end{aligned} \tag{1.59}$$

where θ_W is the angle of rotation from the original basis to the physical one. Similarly we identify the state orthogonal to the Z^0 boson as the massless photon A_μ:

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{1.60}$$

and the masses of the two bosons Z^0 and A_μ are:

$$M_Z^2 = \frac{1}{4}(g_1^2 + g_2^2)v^2 \tag{1.61}$$

$$M_A = 0 \tag{1.62}$$

We have found the symmetry that leaves the ground state invariant is the $U(1)_{em}$ group. This *unbroken* symmetry is responsible for the degeneracy of the W bosons masses and the masslessness of the photon γ. To summarize, the three "would be" Goldstone bosons associated with three broken generators of $SU(2)_L \times U(1)_Y$ were absorbed by the longitudinal d.o.f of the gauge bosons, "giving" them a mass. The subgroup $U(1)_{em}$ was unbroken in the process so the photon is massless. The electroweak symmetry breaking mechanism can therefore be understood schematically as three out of the four available degrees of freedom introduced by the doublet being *eaten* by the massive bosons W^+, W^-, Z^0.

Fermions

The terms that couple the scalar H to the fermions are in general not diagonal (see (1.50)). However it is always possible the rotate the fermions in a new basis where f_{mn}, g_{mn}, h_{mn} are diagonal. We can therefore write in this new basis (we use the same label for the fermions although they should be understood as primed, since they have been redefined):

$$\mathcal{L}_{mass}^{fermions} = -\frac{1}{\sqrt{2}} v \big[f_m \bar{\mathcal{E}}_m P_R E_m + g_m \bar{\mathcal{U}}_m P_R U_m + h_m \bar{\mathcal{D}}_m P_R D_m + \text{ h.c } \big] \tag{1.63}$$

Let us write down one of these terms:

$$\begin{aligned} \bar{\mathcal{E}}_m P_R E_m + \text{ h.c} &= \bar{\mathcal{E}}_m P_R E_m + \bar{E}_m P_L \mathcal{E}_m \\ &= \bar{e}_m e_m \end{aligned} \tag{1.64}$$

we then have:

$$\mathcal{L}_{mass}^{fermions} = -\frac{1}{\sqrt{2}} v \big[f_m \bar{e}_m e_m + g_m \bar{u}_m u_m + h_m \bar{d}_m d_m \big] \tag{1.65}$$

where the masses can be easily read. For instance the electron mass, $m_e = \frac{v f_e}{\sqrt{2}}$. Note that with this procedure the neutrinos ν_m remain massless.

1.3 The Interactions

In the last section we have derived the full particle content and mass spectrum of the Standard Model. Here we will extract the terms corresponding to the interactions.

1.3.1 The Higgs couplings

The Higgs couplings can be easily read from the $\mathcal{L}_{Higgs}$ equation 1.50. The Higgs self-couplings are given by:

$$\mathcal{L}_{H-H} = -\frac{m_H^2}{2v} H^3 - \frac{m_H^2}{8v^2} H^4 \tag{1.66}$$

the Higgs-Gauge interactions:

$$\mathcal{L}_{H-G} = -\frac{2M_W^2}{v} H W_\mu^+ W^{-\mu} - \frac{M_Z^2}{v} H Z_\mu Z^\mu \tag{1.67}$$

$$-\frac{M_W^2}{v^2} H H W_\mu^+ W^{-\mu} - \frac{M_Z^2}{2v} H H Z_\mu Z^\mu \tag{1.68}$$

and the Higgs-fermion couplings can directly be read from:

$$\mathcal{L}_{H-f} = -\sum_f \frac{m_f}{v}\bar{f}fH \tag{1.69}$$

The Feynman rules for some of these interactions are given in Figure 1.2 We note that every other particle that couples to the Higgs does it with a strength at least proportional to the particle mass (the photon is no exception, since it has no mass, it does not couples to the Higgs). For the fermions the coupling is $\frac{m}{v}$ where m is the mass of the particle. v is measured from experiment [13] to be $\approx 246GeV$. So the couplings to the fermions are in general very small (only the top quark has a big coupling and surprisingly close to 1). In general the heavier the particle, the higher its coupling to the Higgs.

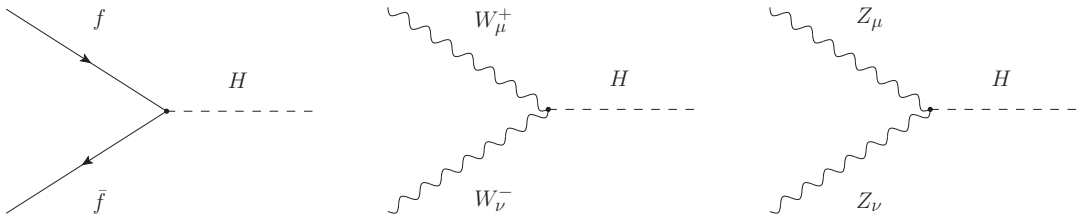


Figure 1.2: Feynman diagrams and rules for the couplings of the fermions and gauge bosons W's and Z to the Higgs boson

We note that the Higgs interactions preserve C, P, and T invariance [16]. Unlike all the other particles presented here, the Higgs is the only one that has yet to be discovered. The discussion on its existence being the subject of this thesis, the next chapter will be entirely dedicated to this undiscovered particle.

1.3.2 The Electroweak sector

The electroweak interactions are divided in gauge self couplings and gauge-fermions couplings. The difference is that they do not appear explicitly in the Lagrangian in mass eigenstate basis.

Bosons self-interactions

These are extracted from the $SU(2)_L$ sector only (since $U(1)_Y$ is abelian).

$$\begin{aligned}\mathcal{L}_{G-G} &= -\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} \\ &= -\frac{1}{2}g_2\epsilon_{abc}\mathcal{W}^a_{\mu\nu}W^{b\mu}W^{c\nu} - \frac{1}{2}g_2^2\epsilon_{abc}\epsilon_{ade}W^b_\mu W^c_\nu W^{d\mu}W^{e\nu} \\ &= \mathcal{L}_{cubic} + \mathcal{L}_{quartic}\end{aligned} \tag{1.70}$$

[16] C,P and T denote respectively charge conjugation, parity, and time reversal transformations.

where $\mathcal{W}^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu$. They can be written as:

$$\mathcal{L}_{cubic} = \mathcal{L}_{WW\gamma} + \mathcal{L}_{WWZ} \tag{1.71}$$

$$\mathcal{L}_{quartic} = \mathcal{L}_{WWWW} + \mathcal{L}_{WWZZ} + \mathcal{L}_{WW\gamma\gamma} + \mathcal{L}_{WWZ\gamma} \tag{1.72}$$

we will not write explicitly all these terms here. However, note that the $WW\gamma$ interaction is purely electromagnetic, so its coupling has to be the electric charge e. The useful relation between the g_1 and g_2 couplings and the electric charge e is derived by equating the coupling found by doing the explicit calculation of $\mathcal{L}_{WW\gamma}$ to e:

$$e = g_2 \sin\theta_W \tag{1.73}$$

Fermion gauge interactions

The fermion-bosons terms are derived from the covariant derivative terms in the Lagrangian:

$$\mathcal{L}_{G-f} = -\frac{1}{2}\bar{L}_m \not{D} L_m - \frac{1}{2}\bar{E}_m \not{D} E_m - \frac{1}{2}\bar{Q}_m \not{D} Q_m - \frac{1}{2}\bar{U}_m \not{D} U_m - \frac{1}{2}\bar{D}_m \not{D} D_m \tag{1.74}$$

Isolating only the terms involving the $W^\pm$ bosons produces the so-called *charged current* interactions (CC). They can be written as:

$$\begin{aligned}\mathcal{L}_{CC} = \frac{ig_2}{2\sqrt{2}}\Big[& W^+_\mu(\bar{\nu}'_m\gamma^\mu(1+\gamma_5)e'_m + V_{mn}\bar{u}'_m\gamma^\mu)(1+\gamma_5)d'_m) \\ & + W^-_\mu(\bar{e}'_m\gamma^\mu(1+\gamma_5)\nu'_m + (V^\dagger_{mn})\bar{d}'_m\gamma^\mu)(1+\gamma_5)u'_m)\Big]\end{aligned} \tag{1.75}$$

- The $W^\pm$ only couple to the left-handed fermions because of the presence of the $1+\gamma_5$ matrix ($\frac{1+\gamma_5}{2} = P_L$). This gives rise to the well known *V-A* (*Vector-Axial*) structure of the Standard Model.

- V_{mn} is the Cabibbo-Kobayashi-Maskawa (CKM) matrix. It parameterizes the rotation to the fermions mass eigenstate basis mentioned in Section 1.2.4. Only the quarks appear to be affected by this rotation. In reality the leptons are as well, but since there is no mass term for the neutrinos, the redefinition of ν_m can be chosen equal to that of e_m and since these matrices are unitary the factor becomes 1 for the leptons. But one has to keep in mind that the leptons appearing in this equation are rotated with respect to the ones in the original Lagrangian given by (1.74).

- The CKM matrix can be parameterized by 4 independent parameters (three angles and one phase).

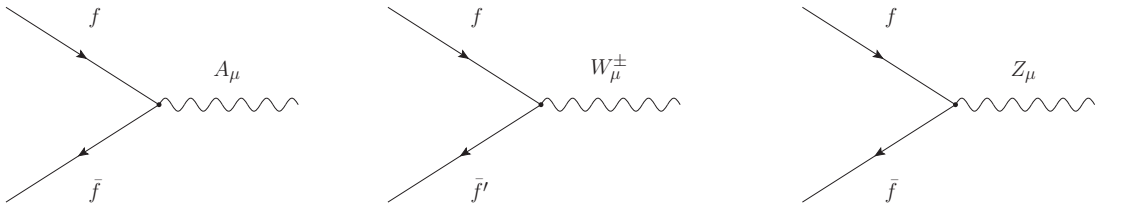
- One implication of the difference between quarks and leptons is that CC interactions do not change lepton flavor while they can change quark flavor. One consequence is that the couplings of $W^\pm$ to leptons are the same for each flavor, implying for instance that W's always decay in equal proportion to e, μ and τ.

- Since CC interactions pick up only the left-handed components of spinors, they explicitly violate charge conjugation (C) and parity (P). But since time-reversal (T) is conserved then (CP) is also conserved.

The other type of fermion gauge interactions in the electroweak sector are the *neutral current* interactions (NC) involving photons and Z's. They can be summarized as:

$$\begin{aligned}\mathcal{L}_{NC} &= \sum_f ieA_\mu \bar{f}\gamma^\mu Q f + \frac{ie}{\sin\theta_W \cos\theta_W}\sum_f Z_\mu \bar{f}\gamma^\mu(g_V + \gamma_5 g_A)f \\ &= \mathcal{L}_{em} + \mathcal{L}_{Z\bar{f}f} \end{aligned} \tag{1.76}$$

where the sum runs over the charged fermions in the electromagnetic term, and over all fermions in the Z boson term. Q is the electric charge of the fermion, and $g_V = \frac{1}{2}I_3 - Q\sin\theta_W{}^2$, $g_A = \frac{1}{2}I_3$.



$\equiv -Qe\gamma_\mu$ $\equiv -\frac{e}{2\sqrt{2}sin\theta_W}\gamma_\mu(1+\gamma^5)$ $\equiv -\frac{e}{sin\theta_W cos\theta_W}\gamma_\mu(g_V+\gamma^5 g_A)$

Figure 1.3: Feynman diagrams and rules for the couplings of the fermions to the photon, the W's and the Z bosons

- We see that the couplings of electromagnetic interactions is precisely the one derived in QED (see 1.10) and in the $WW\gamma$ case (see 1.72) and it couples as expected only to charged particles.
- NC interactions involving the Z boson are flavor diagonal (i.e. $Z \to \nu_e\nu_\tau$ is forbidden).
- NC interactions also violate both C and P but preserve CP.

1.3.3 The Strong interaction

The strong force is governed by the $SU(3)_c$ sector of the Lagrangian:

$$\mathcal{L}_{QCD} = -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} - \sum_q \bar{q}(\not{D} + m)q \tag{1.77}$$

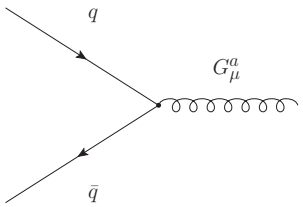
where the sum runs over the 6 quark flavors. The gluon field tensor terms describes the gluon self interactions. It can be expanded into a cubic and a quartic term, leading to 3-gluon and 4-gluon interaction:

$$\begin{aligned}\mathcal{L}_{gl-gl} &= -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} \\ &= -\frac{1}{2}g_3 f_{abc}\mathcal{G}^a_{\mu\nu}G^{b\mu}G^{c\nu} - \frac{1}{2}g_3^2 f_{abc}f_{ade}G^b_\mu G^c_\nu G^{d\mu}G^{e\nu}\end{aligned} \tag{1.78}$$

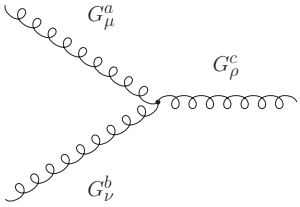
where $\mathcal{G}^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu$. On the other hand the gluon-quark couplings are described by the covariant derivative terms:

$$\mathcal{L}_{gl-f} = \frac{ig_3}{2} \sum_q G^\alpha_\mu \bar{q} \gamma^\mu \lambda_a q \tag{1.79}$$

The Feynman diagrams corresponding to these interaction are summarized in Figure 1.4. Note that a gluon emission never changes the quark flavor (similar to QED). Another thing to be noted, is that unlike the electroweak case where gauge self interaction and gauge fermion had different coupling constants, here the gluon-gluon and gluon quark vertices have the same coupling g_3.



$$\equiv -g_3 \gamma_\mu \frac{\lambda_a}{2}$$



$$\equiv g_3 f_{abc} (\eta^{\mu\nu} \delta p^\rho + perm(\mu, \nu, \rho))$$

Figure 1.4: Feynman diagrams and rules for the QCD couplings of gluon-quark interaction and gluon triple vertex self-interaction

Asymptotic freedom

In order to investigate the differences between the QCD and QED Lagrangian it is useful to ask how the coupling strength varies with the scale of observation. This can be done in perturbation theory by calculating the gauge bosons self-energy. The running of the coupling calculation is out of scope here, but the result is of crucial importance in order to understand QCD. By defining $\alpha_S = \frac{g_3^2}{4\pi}$ ("α strong"), for energies $m_b < \mu < M_Z$ (m_b is the mass of the b quark, see Tab. 1.1), the running of the strong coupling at leading order is found to be:

$$\frac{1}{\alpha_S(\mu)} = \frac{1}{\alpha_S(M_Z)} + \frac{1}{12\pi} \big[2n_q - 33\big] \log \frac{M_Z^2}{\mu^2} \tag{1.80}$$

where n_q is number of quark with a mass lower than m_b (i.e $n_q = 5$) and $\alpha_S(M_Z) = 0.1176 \pm 0.0020$ [13]. In particular it should be noted that the factor in front of $\log \frac{M_Z^2}{\mu^2}$ is *negative*. This implies that the coupling is a decreasing function of the energy. From (1.80) we see that there is a natural scale Λ_{QCD} at which the coupling diverges ($\Lambda_{QCD} \sim 200$ MeV)

At high energies ($\mu \geq \Lambda_{QCD}$), or equivalently low distance, quarks and gluons behave as free particles. In this regime the coupling constant is small enough to allow for calculations in perturbative series (Feynman diagrams). This phenomenon is called *asymptotic freedom* [14]. It is a general result that non-abelian gauge theories with small enough number of fermions are asymptotically free. It is indeed not the case of QED which displays the exact opposite behavior (becomes strongly coupled at very high energies and decouples at low energy).

Confinement

From equation (1.80) one can also predict what happens at low energy (large distance): decreasing the energy makes the coupling stronger. As a result of this behavior, unlike other elementary particles, quarks and gluons happen to behave as free particles as long as the typical energy involved is high enough. At low energies the particles that we observe are rather bound states of quarks and gluons. As a matter of fact, quarks and gluons have never been observed as *free* particles. It is also observed that bound states of quarks appear in nature only as color singlets, they are called hadrons [17]. Mesons are made of $q\bar{q}$ pairs. The lightest meson is the pion, made up of $u\bar{d}$ (π^+), $d\bar{u}$ (π^-) and $\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}$ (π^0). Baryons, on the other hand are made of 3 quarks. The lightest baryons are the well known proton (uud) and neutron (udd). Even though in principle allowed by QCD, more complicated structures such as tetraquarks or pentaquarks have never been observed.

To better understand the confinement feature of non perturbative QCD, we shall imagine, via a classical picture, to break a meson. If we pull apart the quark (and antiquark), when distance is yet small, we encounter a QED type potential $V(r) \sim \frac{1}{r}$ as we are in the regime of asymptotic freedom. When we reach distances $\sim \Lambda_{QCD}^{-1}$ the potential would start to rise linearly $V(r) \sim r$ (as a spring potential). Eventually, when enough energy is accumulated in between the $q\bar{q}$ pair, a $\bar{q'}q'$ would pop-up from the vacuum, leading to an additional meson. This intuitive picture explains why it is impossible to observe quarks and gluons as free particles at low energy. In high energy collisions, when a $q\bar{q}$ pair is created say from a Z boson decay $Z \to q\bar{q}$, the boosted quarks, after loosing enough energy by radiation (of gluons or photons) will combine with another $q\bar{q}$ created out of the vacuum. This will produce a cascade of mesons and baryons, leading to so-called *jets*. This process is called *hadronization* or *fragmentation*.

[17] Even though it is thought that this observation might be a consequence of $\mathcal{L}_{QCD}$ a formal prove does not exist, and up to this day it is only an hypothesis that has never been disproved

Chapter 2

Higgs Boson Decays and constraints

In the previous chapter we have outlined the general features of the Standard Model. Here we will focus more precisely on the Higgs sector. The electroweak plus Higgs sector introduce mainly the two gauge couplings g_1 and g_2 and the two parameters of the potential μ and λ (or alternatively v and λ) [1]. Measuring the $W^{\pm}$ and Z^0 masses and the electric charge e allows to fully determine g_1, g_2 and v. In addition we have the confirmation that, if the Higgs exists, the potential has to be the one given by Equation (1.44) from the experimental observation that $M_Z = M_W \cos\theta_W$. However, since the scalar particle H has not yet been observed, its mass is unknown and therefore the parameter λ remains undetermined in the Standard Model [2].

First some phenomenology on Higgs decays in a hadron colliders will be reviewed. After we will discuss which kind of constraints one can apply on the Higgs existence, and more precisely on its mass: on the one hand using theoretical arguments such as unitarity, triviality and vacuum stability, on the other hand the non-observation in previous experiments. Finally we will discuss the hierarchy problem and possible theoretical extensions and alternatives to the Standard Model.

2.1 Higgs decays and search modes

2.1.1 Formulae

The differential *decay rate* (or *width*) for a single particle decaying to X particles $A \to X$ is given by

$$d\Gamma(A \to X) = \frac{1}{2E_A}|\mathcal{M}|^2(2\pi)^4\delta^{(4)}(p_A - \sum_{f\in X} p_f)\prod_{f\in X}\frac{d^3p_f}{(2\pi)^3 2E_f} \tag{2.1}$$

where $\mathcal{M}$ is the quantum mechanical amplitude of the process under consideration. To obtain the total rate (probability of decay per unit time) one needs to integrate over the final state phase space $\prod_{f\in X}\frac{d^3p_f}{(2\pi)^3 2E_f}$. Notice that the decay rate is *not* Lorentz invariant due to the the factor $\frac{1}{2E_A}$. Indeed, the lifetime ($\tau = \frac{1}{\Gamma}$) of a moving particle is increased

[1] We are ignoring on purpose the Yukawa part of the Lagrangian here

[2] From now on the real scalar H will be called "the Higgs", BEH or simply H

by time dilation. Therefore the lifetime of particle is always calculated in its rest frame, $E_A = m_A$ and $p_A = (m_A, 0, 0, 0)$.

If instead we consider a process $A+B \to X$ the rate becomes proportional to the number density of target particles. To remove this dependence, the *cross section* is defined by dividing the rate by the flux F of incident particles. The flux can depend only on the relative velocity of A and B, which is a covariant quantity. The differential cross section, defined by

$$d\sigma(A + B \to X) = \frac{1}{4E_A E_B |v_A - v_B|} |\mathcal{M}|^2 (2\pi)^4 \delta^{(4)}(p_A - \sum_{f \in X} p_f) \prod_{f \in X} \frac{d^3 p_f}{(2\pi)^3 2E_f} \tag{2.2}$$

is therefore a Lorentz scalar and can therefore be calculated in any reference frame. The cross section has the dimensions of a surface and is typically measured in *barns* (1 b $= 10^{-24} cm^2$). The luminosity $\mathcal{L}$ of an accelerator is defined as the rate of incident particles per unit area of the beam. The product of the cross section σ times the luminosity then gives the number of events $\mathcal{N}$ that can be expected per unit time:

$$\mathcal{N} = \sigma \mathcal{L} \tag{2.3}$$

Let us consider now a process $A + B \to H \to X$. This process is enhanced when $s = (p_A + p_B)^2 \sim m_H^2$ (i.e when the Higgs particle is on-shell). In this particular case $\sigma(A + B \to H \to X)$ factorizes into:

$$\sigma(A + B \to H \to X) = \sigma_{tot} BR(H \to X) \tag{2.4}$$

where $\sigma_{tot} = \sigma(A + B \to H)$ is the Higgs production cross section and the *Branching Ratio* (BR) is the fraction between the decay rate into a particular mode X and the total decay rate:

$$BR(H \to X) = \frac{\Gamma(H \to X)}{\Gamma_{tot}} \tag{2.5}$$

In this chapter we will concentrate on reviewing the Higgs decays and branching ratios. In the next chapter the various production modes of the Higgs and the techniques to calculate σ_{tot} will be reviewed.

Without considering the higher order corrections to the self energy of the resonance, the propagator would simply be $\frac{1}{p^2 - m_H^2}$ and the amplitude would diverge for energies $\hat{s} \sim m_H^2$. The propagator gets therefore regularized by shifting the mass of the particle by a complex number:

$$\frac{1}{p^2 - m_H^2} \to \frac{1}{p^2 - m_H^2 - i m_H \Delta} \tag{2.6}$$

The time dependence of the particle in its rest frame is then be given by:

$$|H(t)\rangle = e^{-im_H t - \frac{1}{2}\Delta t}|H(0)\rangle \tag{2.7}$$

Δ can therefore be identified with the full decay rate Γ_{tot}. The Higgs is an unstable particle with a lifetime $\tau_H = \frac{1}{\Gamma_{tot}}$. The total Higgs width Γ_{tot} can be calculated as the sum of all the partial widths $\Gamma(H \to X)$:

$$\Gamma_{tot} = \sum_X \Gamma(H \to X) \tag{2.8}$$

The uncertainty principle allows us to reinterpret the width as an uncertainty on the mass itself, since $\Gamma \sim \frac{\hbar}{\tau} \sim \Delta m$. In cases where it is possible to measure the total invariant mass $\hat{s} = \left(\sum_{f \in X} p_f\right)^2$ in the final state, one can study the rate as a function of the invariant mass. Then the cross section follows a Breit-Wigner distribution (which is a direct consequence of Equation (2.6)):

$$\sigma(\hat{s}) \sim \frac{1}{(\hat{s} - m_H^2)^2 + (\frac{\Gamma}{2})^2} \tag{2.9}$$

This function has a peak at $\hat{s} \sim m_H^2$ and a width Γ. A fit to this peak is the way to experimentally determine the mass and the width of such an unstable particle.

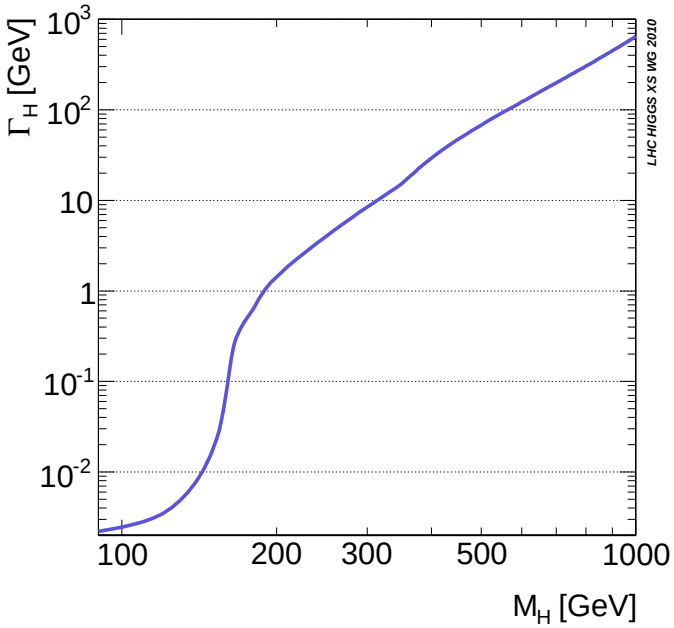


Figure 2.1: Higgs width as function of its mass

In Figure 2.1 the Higgs width is a plotted as a function of its hypothetical mass [15]. For low mass values ($m_H \sim 100$ GeV) the total width is small compared to the Higgs mass ($\Gamma \sim 1$ MeV), which implies that the peak is sharp and the mass will be perfectly measurable (modulo experimental resolution). However, for very high mass values ($m_H \sim 1$ TeV), the width is comparable to the mass itself. This will make it very difficult for an experiment to correctly measure the Higgs mass.

2.1.2 Decays

Once the Higgs is produced, it almost instantly decays to other stable or unstable particles. The condition for the decay to be possible is that the mass of the decay product must be less than half of the Higgs mass (since the Higgs boson couples to a massive particle and its antiparticle). If this is not the case it can still decay to heavier particles but these cannot be on shell; the decay can happen, even though kinematically disfavored (since suppressed by the propagator (2.6)), but one of the decay product has be to be virtual. For the decay to be kinematically allowed, it has therefore to depend on the Higgs mass. Naively, there are three possibilities:

- $\boldsymbol{m_H < 2M_W}$
  The heaviest available decay is $H \to b\bar{b}$ followed by $H \to \tau^+\tau^-$.
- $\boldsymbol{2M_W < m_H < 2m_t}$
  The favored decays here are $H \to W^+W^-$ and $H \to ZZ$. However since the Higgs coupling to $W^\pm$ is $\sim \times 2$ that to Z, even for $m_H > 2M_Z$ the decay to W^+W^- is preferred.
- $\boldsymbol{m_H > 2m_t}$
  Here the decay $H \to t\bar{t}$ is kinematically favored. However the Higgs coupling to fermions is proportional to the fermion mass, while the coupling to the gauge bosons $W^\pm$ and Z is proportional to their mass *squared*. Therefore, at these very high mass the Higgs will still decay mainly to bosons.

In the Higgs rest frame the products are emitted back-to-back and the decay is isotropic The decay rate of the Higgs in its rest frame into a particular flavor of fermion is given by

$$\begin{aligned} \Gamma(H \to f\bar{f}) &= \frac{3m_H}{8\pi}\left(\frac{m_f}{v}\right)^2\left[1-\frac{4m_f^2}{m_H^2}\right]^{\frac{3}{2}} \\ &\simeq \frac{3m_H}{8\pi}\left(\frac{m_f}{v}\right)^2 \quad \text{for } m_f \ll m_H \end{aligned} \tag{2.10}$$

As expected, the partial width to fermions is very sensitive to the fermion mass: the higher the fermion mass, (and its coupling to the Higgs) the higher the width. The width is also proportional to the Higgs mass. Instead, the rate of decay to W bosons is expressed as:

$$\begin{aligned} \Gamma(H \to W^+W^-) &= \frac{m_H}{16\pi}\frac{m_H^2}{v^2}\left[1-4\left(\frac{M_W^2}{m_H^2}\right)+12\left(\frac{M_W^2}{m_H^2}\right)^2\right]\left(1-\frac{4M_W^2}{m_H^2}\right)^{\frac{1}{2}} \\ &\simeq \frac{m_H}{16\pi}\frac{m_H^2}{v^2} \quad \text{for } M_W \ll m_H \end{aligned} \tag{2.11}$$

Here $\Gamma \propto \lambda m_H$ where λ is the Higgs self-coupling. So, quite differently from the fermion case, here the decay rate is not proportional to M_W (at least in the regime $M_W \ll m_H$). This reflects the fact that the gauge boson masses emerged in the first place as components of the Higgs doublet.

The formula for the decay into Z bosons can be easily guessed from $\Gamma(H \to W^+W^-)$ by noticing that the coupling is the same as that of the $H \to W^+W^-$ with a factor $\frac{1}{2}$. By substituting M_W by M_Z one has therefore:

$$\Gamma(H \to ZZ) = \frac{1}{2}\Gamma(H \to W^+W^-) \text{ with } M_W \leftrightarrow M_Z \tag{2.12}$$

The branching ratios (BR) of the Higgs as a function of its mass are shown in Figure 2.2. Naively one might think that the search mode is dictated by the decay rate of the Higgs boson for each mass hypothesis. This appears not to be always the best choice. As a matter of fact, in an experimental analysis, such as the one that will be presented in the present work, one wants to be able to separate the signal from the background events. This implies that the signature of a given search mode needs to be "clean". This means that the particles of a given decay need to be unambiguously identifiable and their kinematic properties to be well measured. For instance, fully hadronic Higgs decays are abundant at the LHC, but they need to be disentangled from the large QCD backgrounds, therefore one might prefer leptonic decays that are produced at a lower rate, but are easier to separate from the backgrounds.

With this in mind we can identify which decay channels are the most promising for Higgs searches depending on the mass hypothesis.

2.1.3 Search Modes

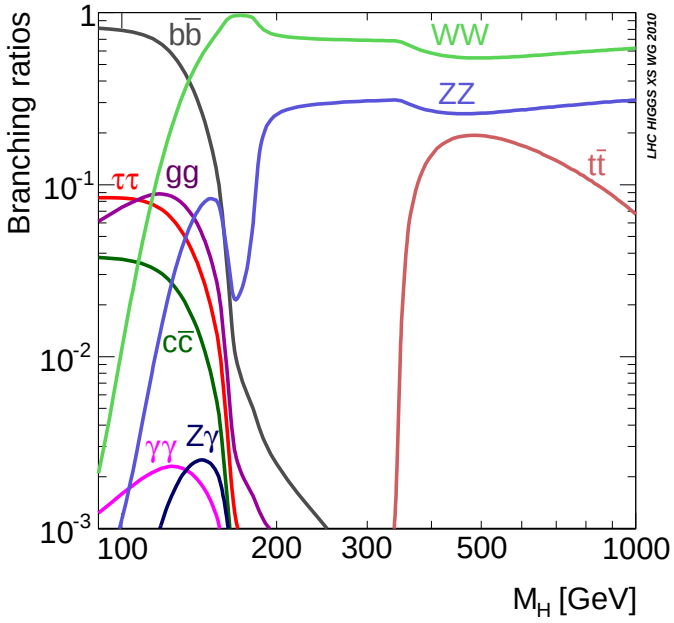


Figure 2.2: Higgs decay branching ratios as function of the Higgs mass [15]

Low Mass Higgs: $m_H \leq 140$ GeV

Although $H \to b\bar{b}$ is the dominant decay in this mass hypothesis range, it is almost unexploitable at the LHC given the huge amount of QCD background events. As will be

explained later, jets formed by b quark fragmentation are identifiable in the CMS experiment but the direct $b\bar{b}$ production cross section is ~ 10 orders of magnitude higher than the Higgs production cross section. One can search the Higgs in this channel when it is produced in association with other identifiable particles (i.e produced in VBF or Higgs-Strahlung mode). This helps in reducing the overwhelming QCD background. However as we will see in the next chapter the production rates in these modes are rather small at the LHC so a significant amount of integrated luminosity needs to be accumulated in order for it to become a competitive channel for Higgs searches.

The next channel that one might exploit at low mass is the $H\rightarrow\tau\tau$ mode. The τ leptons are not stable particles. They decay either hadronically ($\sim 85\%$), either to stable leptons e, μ, ν ($\sim 15\%$). Given the low Branching Ratio and the high amount of QCD and electroweak backgrounds Higgs searches in this channel are also difficult.

Finally, the main search mode at low mass is the Higgs decaying to two photons: $H\rightarrow\gamma\gamma$. The Higgs does not couple directly to photons since they are massless. The decay occurs via a loop of heavy particles that radiate two photons. Although the Branching Ratio is very small (see Figure 2.2) compared to $H\rightarrow b\bar{b}$ and $H\rightarrow\tau\tau$ this channel is very clean. The signature in this mode is of course two well identified photons. This leaves no indetermination in the final state. In CMS, the excellent resolution of the Electromagnetic Calorimeter allows to measure the photons kinematics very precisely, therefore making a mass peak reconstruction possible over the continuum $\gamma\gamma$ background. It is in this mode that a light Higgs boson will be most likely discovered.

Medium Mass Higgs: 140 GeV $\leq m_H \leq$ 180 GeV

In this mass range the BR of $H\rightarrow W^+W^-$ is almost one. This implies that the Higgs will almost exclusively decay in this mode. The W bosons are unstable particles, they instantly decay leptonically or hadronically. The fully hadronic decay is almost unexploitable again due to the huge QCD backgrounds. The semi-leptonic channel (meaning that one W will decay leptonically and the other hadronically) will have a signature consisting of 2 jets, one lepton and one neutrino, the neutrino staying undetected. In this mode, however, the discrimination against the electroweak W+jets production and QCD is very difficult. This channel can be exploited when the Higgs is produced in VBF mode (see next chapter), the additional forward jets in the events helping in discriminating against the backgrounds. The "golden" search mode in this mass range is when both W's decay leptonically $\mathrm{W^+W^-}\rightarrow\ell^+\nu\ell^-\bar{\nu}$ [3]. In this case the signature is 2 well identified leptons plus 2 neutrinos. As already stated neutrinos stay undetected in the final state, leading to a momentum imbalance in the plane transverse to the collisions. The main backgrounds in this search mode are almost exclusively electroweak processes such as direct WW production, W+jets, Z+jets and $t\bar{t}$. Since this particular analysis is the whole subject of this thesis, a more detailed discussion on this search channel will follow in the next chapters.

High Mass Higgs: $m_H \geq$ 180 GeV

At $m_H \sim 180$ GeV the Higgs is massive enough to be able to decay into two Z bosons. Although the BR is still lower that of $\mathrm{H\rightarrow W^+W^-}$ this decay is still very interesting because of how the Z bosons decay. They also, as the the W, decay either leptonically, either hadronically. However, since the Z boson couples, as the Higgs, only to a particle and

[3] In this notation $\ell = e, \mu$

its antiparticle they will decay into lepton, anti-lepton or $q\bar{q}$ pair. In this $H \rightarrow ZZ$ search, the most interesting case is when the Z's both decay to stable leptons, i.e $Z \rightarrow \ell^+\ell^-\ell^+\ell^-$ because one can fully reconstruct the invariant mass of the system, and as in the $H \rightarrow \gamma\gamma$ case, infer the Higgs mass.

2.2 Bounds on the Higgs Mass

2.2.1 From theory

Unitarity

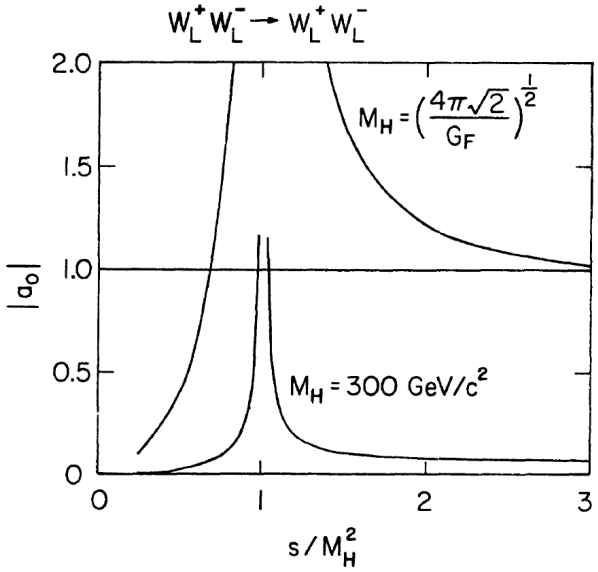


Figure 2.3: Amplitude of the $W_L^+W_L^-$ scattering as function of the energy in the center of mass frame for a Higgs boson mass below (m_H =300 GeV) and at the unitarity threshold.

Without the existence of a scalar, the amplitude of the $W_L^+W_L^- \rightarrow W_L^+W_L^-$ scattering diverges at high enough energies. By also including diagrams involving the exchange of a scalar unitarity can be restored as long as the the scalar is not too heavy. The *unitarity bound* can be expressed as [16–19]:

$$m_H < m_0 = \sqrt{8\pi v^2} = \sqrt{\frac{4\pi\sqrt{2}}{G_F}} \tag{2.13}$$

$$\simeq 1 \text{ TeV} \tag{2.14}$$

where $G_F = \frac{1}{v\sqrt{2}}$ is the Fermi constant [4]. The amplitude of the longitudinal scattering is shown is Figure 2.3. For $m_H < m_0$ the amplitude stays below 1. The apparent divergence at $s = m_H^2$ gets regularized with the Breit-Wigner prescription that was explained in the previous section. For values above the unitarity threshold, $m_H \geq m_0$, there is a range of energies for which the amplitude is always greater than 1 and unitarity is violated.

[4] $G_F = 1.1664 \times 10^{-5}\ (GeV)^{-2}$

Although indicative, this bound cannot be taken strictly. It does not mean that the Higgs mass *has to be* below m_0. Instead, it underlines the fact that at high energies the theory becomes strongly coupled and the perturbative approach is not valid anymore.

Bounds from the self-coupling running

In this section we will seen how the evolution of the self-couping λ (from Equation (1.44)) can be used to put bounds on the Higgs mass. The running of λ has been calculated up to 2-loops. However, for sake of simplicity, the 1-loop calculation will be discussed here, by considering only the dominant terms.

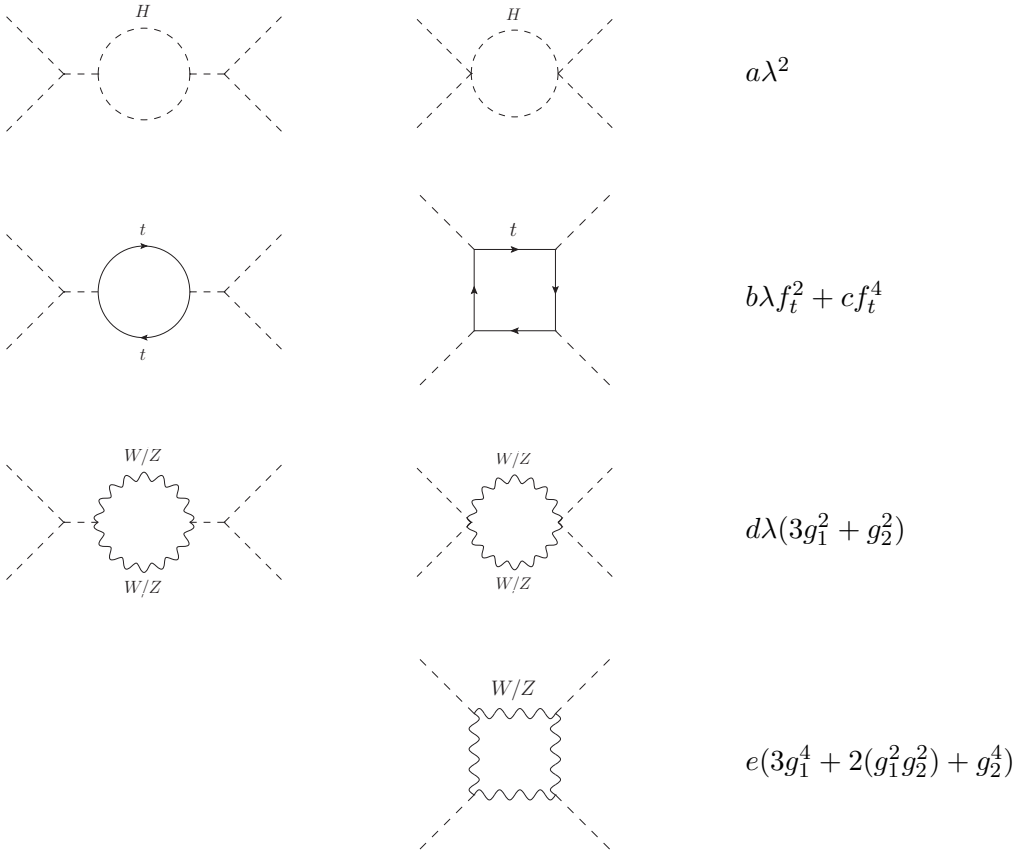


Figure 2.4: Some diagrams contributing to the Higgs self-coupling scale evolution.

The evolution of the self-coupling λ with the scale μ can be written:

$$\begin{aligned} \frac{d\lambda}{d\ln\mu^2} &= \beta_\lambda \\ &= a\lambda^2 + b\lambda f_t^2 - cf_t^4 - d\lambda(3g_1^2 + g_2^2) + e(3g_1^4 + 2(g_1^2 g_2^2) + g_2^4) \end{aligned} \tag{2.15}$$

where a, b, c, d, e are positive constants and f_t is the top Yukawa coupling. The corresponding diagrams are shown in Figure 2.4. Starting from a reference value λ_0, we can infer the value of λ at any scale μ. In the following $\lambda_0 = \lambda(v)$. The two regimes of weak and strong self coupling will be used to infer a lower and upper bound to the Higgs mass.

Triviality Here we look at the limit of strong self-coupling. Since $\lambda \propto m_H^2$, it corresponds to the high mass limit. Since $\lambda \gg f_t, g_1, g_2$ we can keep only the terms multiplied by the constant a. Solving Equation (2.14) in this case is trivial and leads to [5]:

$$\lambda(\mu) = \frac{\lambda(v)}{1 - a\lambda(v) \ln \frac{\mu^2}{v^2}} \tag{2.16}$$

At the scale $\Lambda_{triv} = ve^{2a\lambda(v)}$ the coupling diverges and the perturbative approach is not valid anymore. By requiring the Standard Model to be valid up a scale Λ_{triv}, i.e λ finite, we have:

$$\begin{array}{rrcl} & \mu & < & \Lambda_{triv} \\ \Leftrightarrow & m_H & < & m_H^{max} \text{ with } m_H^{max} = \sqrt{\frac{2v^2}{a \ln \frac{\Lambda_{triv}^2}{v^2}}} \end{array} \tag{2.17}$$

For each scale Λ_{triv} at which we believe the Standard Model might break down because new physics appears there is maximum allowed value for the Higgs mass. At a scale $\Lambda_{triv} = 10^{16}$ GeV, $m_H^{max} = 160$ GeV.

Stability Let us now study the case $\lambda \ll f_t, g_1, g_2$. Now the terms c and e dominate. One can then rewrite β_λ as [6]:

$$\beta_\lambda = \frac{3}{16\pi^2 v^4}\left(2M_W^4 + M_Z^4 - 4m_t^4\right) \tag{2.18}$$

Since β_λ is a constant, the solution of the differential Equation (2.16) is trivial in this scenario:

$$\lambda(\mu) = \lambda(v) + \beta_\lambda \ln \frac{\mu^2}{v^2} \tag{2.19}$$

By taking $m_t = 175$ GeV, $M_W = 80.4$ GeV and $M_Z = 91.2$ GeV, one finds that $\beta_\lambda < 0$. This implies that there are scales above which λ becomes negative. In such a case the Higgs potential becomes unbounded from below. The *stability* condition can therefore be summarized as $\lambda > 0$ and it translates into a lower bound on the Higgs mass:

$$\begin{array}{rrcl} & \mu & < & \Lambda_{stab} \\ \Leftrightarrow & m_H & > & m_H^{min} \text{ with } (m_H^{min})^2 = \frac{3}{8\pi^2 v^2}\left(2M_W^4 + M_Z^4 - 4m_t^4\right) \end{array} \tag{2.20}$$

The result makes apparently no physical sense since the factor $2M_W^4 + M_Z^4 - 4m_t^4$ is negative. Had we considered the calculation up two loops we would have found $(m_H^{min})^2 >$

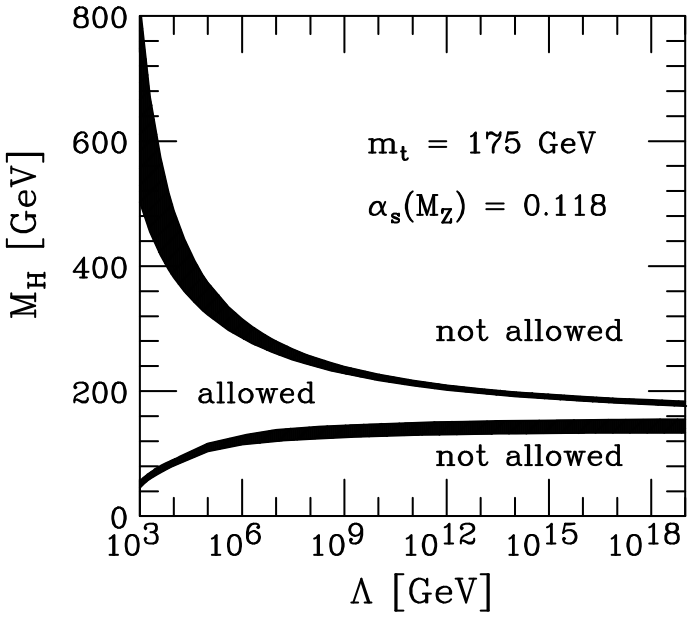


Figure 2.5: The upper (lower) curve is the upper (lower) theoretical bound on the Higgs mass resulting from the triviality (stability) requirement on the running of the Higgs self-coupling λ.

0. This procedure allows to set a lower bound on the Higgs mass for each scale up to which the Standard Model is believed to be correct.

The results of the last two sections can are summarized in Figure 2.5 [20]. For each scale Λ the range of allowed Higgs masses is between the boundaries set by the triviality and the stability conditions. If we believe that the Standard Model is the ultimate theory of elementary particles up to the Planck scale, we see that there is a very narrow range of possibilities for the Higgs mass ($m_H \sim 175$ GeV). On the other hand, if the Standard Model is only an effective theory valid up 1 TeV then m_H is allowed to be in the range 50 to 500 GeV. [7]

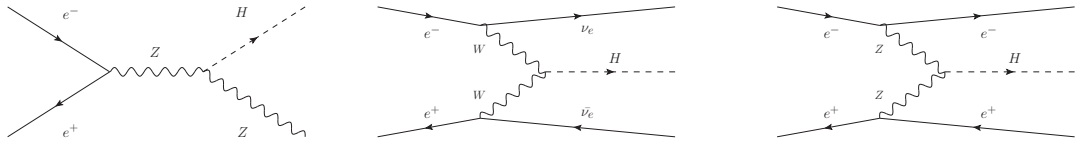
2.2.2 From Experiments

Direct searches

LEP The first dedicated searches of the Higgs boson were performed at the *Large Electron Positron Collider* (LEP) between 1989 and 2000. The e^+e^- collisions center of mass energy $\sqrt{\hat{s}}$ varied between 90 and 209 GeV. The dominant production mechanism at LEP was the so-called Higgs Strahlung production: e^+e^- produce a Z boson that radiates a Higgs boson (see Figure 2.6(a)). However, due to obvious kinematic constraints imposed by $\sqrt{\hat{s}}$ such a mechanism has a high rate only for Higgs masses $m_H \leq 120$ GeV.

As seen in the previous sections, in this low mass range the Higgs decays almost exclusively into a $b\bar{b}$ pair. The search channels depend therefore on the possible Z decays,

[5] $a = \frac{4\pi^2}{3}$
[6] $c = \frac{3}{16\pi^2}$ and $e = \frac{3}{256\pi^2}$
[7] However, we will see that some of these possible values are already excluded from previous experiments.



(a) Higgs-Strahlung (b) Vector-Boson Fusion (W) (c) Vector-Boson Fusion (Z)

Figure 2.6: Main Higgs production mechanisms at LEP

leading mainly to three possibilities: $HZ \to b\bar{b}q\bar{q}$ (4 jets final state), $HZ \to b\bar{b}l^+l^-$ (2 jets, 2 leptons final state), $HZ \to b\bar{b}\nu\bar{\nu}$ (2 jets , missing energy).

The non-observation of a statistically significant excess of signal candidates over the known Standard Model backgrounds led to establish lower limit to the possible Higgs masses. With 2.5 fb^{-1} of data the experiments at LEP [21] established that:

$$m_H \geq 114.6 \text{ GeV at } 95\% \text{ CL} \tag{2.21}$$

Tevatron At the Tevatron protons and anti-protons collide ($p\bar{p}$) a center of mass energy $\sqrt{s} = 1.96$ TeV. Here we present the results corresponding to the data collected up to February 2012, corresponding to an integrated luminosity of 10 fb^{-1}.

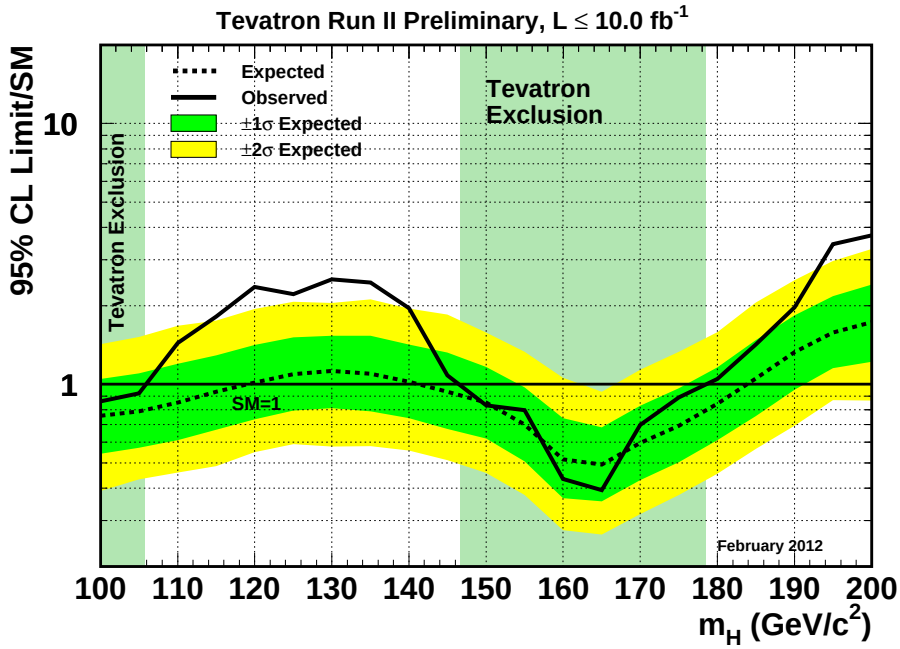


Figure 2.7: Expected (Montecarlo simulation) and observed (real data) exclusion limits at the 95 % confidence level corresponding to 10 fb^{-1} of data-taking at the Tevatron.

At $\sqrt{s} = 1.96$ TeV, the main production mechanisms are those of the LHC (they will be discussed extensively in the next chapter). Most of the final state topologies were studied independently to increase sensitivity at all possible masses and the results were combined [22]. As for the LEP studies, the outcome of these searches have been negative so far, allowing for the exclusion of some possible mass ranges.

The expected and observed exclusion limits (i.e. in the hypothesis of pure background) are given in Figure 2.7. The plot reads in the following way: if the value on the vertical axis (observed) is x, then a Higgs boson with a mass m_H produced with a rate of x times the rate predicted by the Standard Model is excluded at the 95 % confidence level. In short, Higgs masses for which the observed limit is below 1 are excluded [8].

Combining the LEP and Tevatron results, we then know today with 95% confidence that $m_H \geq 115$ GeV and 147 GeV$\leq m_H \leq 178$ GeV.

Indirect search

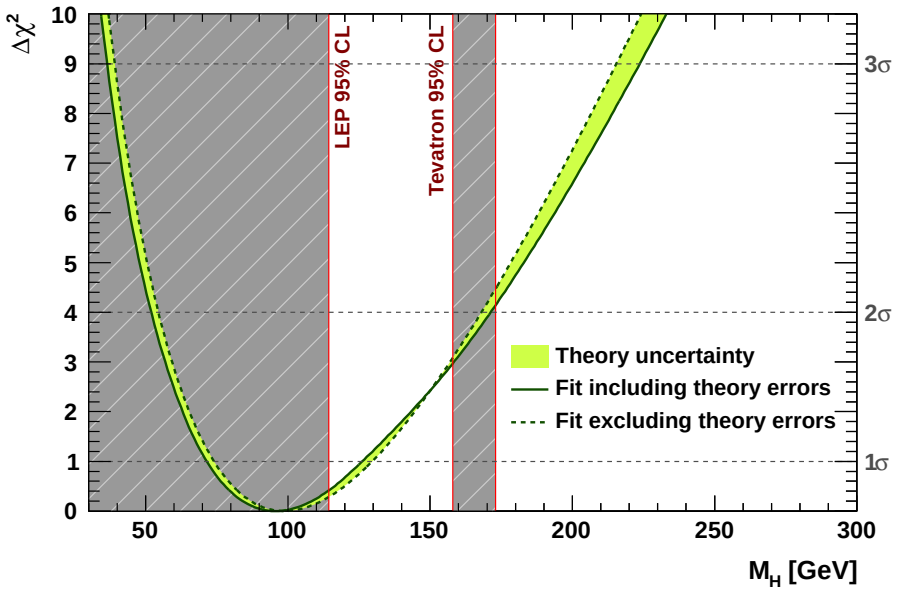


Figure 2.8: $-2\ln\mathcal{L} = \Delta\chi^2$ as a function of the possible Higgs mass. $\mathcal{L}$ is the likelihood that a given value of m_H is compatible with the other experimentally determined parameters of the Standard Model

[8] Also note that if an excess of events is observed, (the background only hypothesis is hard to validate), the observed limit would go up, much like the region $m_H \in [120; 140]$ GeV in Figure 2.7

The Fermi constant G_F is known experimentally with extremely high precision from the muon decay lifetime. At tree level:

$$G_F^{tree} = \frac{e^2}{4\sqrt{2}\sin^2\theta_W M_W^2} \tag{2.22}$$

Higher order corrections are taken into account by writing:

$$G_F = G_F^{tree}\left(1 + \Delta\alpha - \frac{\cos^2\theta_W}{\sin^2\theta_W}\Delta\rho\right) \tag{2.23}$$

where $\Delta\alpha$ is essentially the photon self-energy and $\Delta\rho$ are the high order corrections to the parameter $\rho = \frac{M_W}{M_Z\cos\theta_W}$ (=1 at tree level). The corrections to ρ are indeed from W boson self energy.

(a) (b)

Figure 2.9: Some Diagrams contributing to the gauge boson mass at 1-loop

By simple power counting, one can guess:

$$\Delta\rho = am_t^2 - b\log m_H \tag{2.24}$$

where a and b are well known calculable real parameters. By replacing (2.24) in (2.23) we have then a relation that connects experimentally determined quantities (such as G_F, M_Z, M_W, m_t) to the unknown parameter m_H.

The most likely value for m_H can then be obtained by performing a maximum likelihood estimation for each mass hypothesis [23, 24]. This is shown in Figure 2.8 and the result is:

$$m_H = 96^{+31}_{-24}\ \text{GeV} \tag{2.25}$$

The conclusion is that the experimental values of the measurable parameters of the standard model "prefer" a light Higgs boson. Notice however that the 1-σ error band is quite large. The is due both to the high experimental uncertainty on the top quark mass, and also to the fact that the m_H dependence in $\Delta\rho$ in only logarithmic, leading to a poor constraining power of the Higgs mass. Also we can note that the preferred value of m_H given in (2.25) is perfectly compatible with the non excluded region in Figure 2.7.

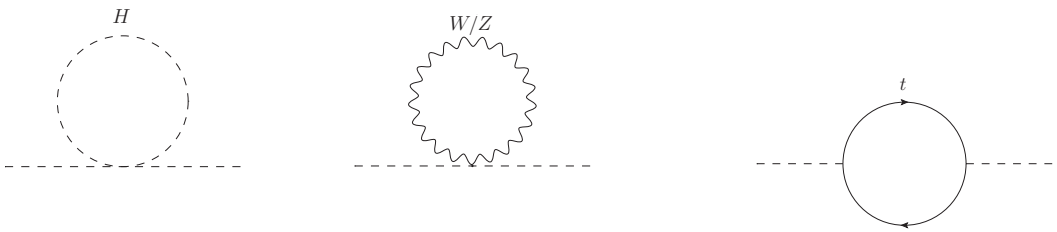
2.3 Standard Model Extensions

2.3.1 The hierarchy problem

In the previous section we have been looking at the implications the RGE [9]at 1-loop of the Higgs self-coupling λ have on the Higgs mass. It is natural to ask what consequences can be drawn from the RGE of the parameter μ. The parameter μ is the Higgs mass itself at tree level. At higher orders, the mass of the Higgs receives several corrections. Schematically, one can write:

$$\delta\mu^2 = \delta m_H^2 + \delta m_G^2 + \delta m_f^2 \tag{2.26}$$

where the terms in the sum correspond to the correction at 1-loop level to Higgs self-energy originating from a Higgs, a Gauge boson and fermions loops.



(a) Higgs (b) Gauge (c) Fermions

Figure 2.10: Some Higgs self-energy diagrams displaying a UV quadratic divergence

All of the corresponding diagrams are UV divergent. A naive power counting of divergence in diagrams in Figure 2.10 ,

$$\begin{aligned} \text{loop} &= \int^{\Lambda_{UV}} \frac{d^4k}{k^2 - M^2} \\ &\sim \int^{\Lambda_{UV}} kdk \\ &\sim \Lambda_{UV}^2 \end{aligned} \tag{2.27}$$

gives that the dominant UV divergence is quadratic [10]. The corrections to the Higgs mass at 1 loop are therefore quadratic:

$$\delta\mu^2 = \frac{1}{16\pi^2}\left[C_G g^2 + C_H\lambda - \sum_m C_m f_m\right]\Lambda_{UV}^2 \tag{2.28}$$

[9]RGE = Renormalization Group Equations

[10]There are also logarithmic divergence but these are neglectable at $\Lambda_{UV} \to +\infty$

where the three terms in the sum follow from Equation (2.28). The parameter g generically denotes the electroweak couplings g_1 and g_2. Also the constants C_G, C_H and C_m are perfectly calculable at 1 loop but we do not need to make them explicit for the present discussion. The standard renormalization procedure imposes that:

$$\mu^2_{bare} + \delta\mu^2 = \lambda v^2 \tag{2.29}$$

It is clear that for an arbitrary Λ_{UV}, this equality can hold only through a careful adjustment of the parameters μ_{bare} or λ. For $\Lambda_{UV} \sim v$ the correction is still acceptable, however, for scales as high as $\Lambda_{UV} \sim M_{Planck} \sim 10^{19}$ GeV, since the dependence is quadratic in Λ_{UV} this would require a *fine-tuning* of ~ 38 orders of magnitude!

Although it might look as just a technicality, some regard this fine-tuning of the bare parameters of the theory as unnatural. It appears as a direct consequence of the huge difference between the natural electroweak scale, v and M_{Planck} [11]. This is often referred as the *hierarchy* or the *naturalness* problem.

The hierarchy problems leads to think that the Standard Model cannot be a fundamental theory valid at energies up to the Planck scale M_{Planck}, unless we accept the unnatural fine-tuning of its parameters. There are two ways one can address this problem:

- A new fundamental scale Λ exists, $\Lambda \gtrsim v_{weak}$ corresponding to new dynamics (gravitation?) which would make the whole hierarchy problem non-relevant anymore since Λ would then be a natural cut-off on the extrapolations to higher scales of the Standard Model parameters (such as μ, or λ).
- The Higgs sector is fundamental, although more complex than in the Standard Model, but the perturbative series stays convergent up to any scale (no quadratic divergences). For this to happen new fundamental particles would be required to exist in order to provide the appropriate cancellations in the Higgs self-energy, and/or, the Higgs sector has to obey to different dynamics that do display well-behaved logarithmic correction at high energy.

More generally, other fundamental questions arise: Why nature would provide a scalar particle simply for the purpose of electroweak symmetry breaking? In other words, what is so-special about the weak scale $v_{weak} \sim G_F^{-\frac{1}{2}}$?

During the last decades theoretical physicists have tried to address these questions by building more fundamental models. Due to the proliferation of such models, we will just sketch the main ideas behind two quite different approaches: the Technicolor models (TC) and Supersymmetry (SUSY).

2.3.2 Technicolor

The starting point of Technicolor theories is the observation that QCD naturally provides a fundamental scale, Λ_{QCD}, that is generated dynamically by the running of α_S. Indeed Λ_{QCD} is not obtained as a rescaling of some other, more fundamental constant, instead, it arises from the pure quantum mechanics corrections of the $SU(3)_c$ gauge theory. The

[11] In other words, it reflects the fact the gravity is much weaker than the other forces.

insight here is that the strong dynamics of QCD naturally explain the large scale difference between λ_{QCD} and M_{Planck}.

Another key aspect of QCD, not discussed so far, is *chiral symmetry breaking*. Spontaneous symmetry breaking has already been observed in particle physics: due to confinement, light quarks form at low energy a condensate $q\bar{q}$ whose ground state energy is > 0. (i.e, it is massive).

$$\begin{aligned} \langle q\bar{q} \rangle &= \langle 0 | q_L \bar{q}_R + q_R \bar{q}_L | 0 \rangle \\ &\sim \Lambda^3_{QCD} \end{aligned} \tag{2.30}$$

Now, since the QCD Lagrangian with massless quarks is $SU(2)_L \times SU(2)_R$ (chiral symmetry) invariant, the equality (2.30) spontaneously breaks this chiral symmetry into the subgroup $SU(2)_{isospin}$. The broken generators corresponding to the axial current generate, via the Goldstone theorem, three massless Goldstone bosons: the pions. However since the quarks *do* have a small mass, the pions are actually the massive pseudo-Goldstone of the spontaneous chiral symmetry breaking.

The main idea for TC is to define some type of strong dynamics theory in which the weak scale v_{weak} naturally emerges as the scale where the Technicolor coupling α_{TC} diverges (as is Λ_{QCD} in QCD). In pure Technicolor theory, first postulated by Susskind [25] and Weinberg [26], the *techniquarks* and *techigluons* interact in this strong dynamics framework. The techniquarks condensates (the *technimesons*) break the generalized chiral symmetry, generating the associated Goldstones: the *technipions*. In this framework the longitudinal degrees of freedom of the W and Z bosons are simply the "eaten" technipions. Therefore the masses of the gauge bosons are generated dynamically: this phenomenon is called *dynamical symmetry breaking*.

Other theories postulate that the Higgs boson is itself the low-mass pseudo-Goldstone of the chiral symmetry breaking [27] (Little Higgs Models) or a heavy composite object [28,29] (Top-Color Assisted Technicolor in extra-dimensions). As TC theories will no longer be discussed in this thesis, for a complete review of TC models and related phenomenology see [30].

2.3.3 Supersymmetry

Supersymmetric theories are invariant under transformations that switch bosonic fields into fermionic fields and vice-versa. The nice feature about such theories is that the *supercharges*, i.e the generators of supersymmetric transformations are closely related to the space-time translation generators. Hence, if gauged, they provide a framework for building a quantum theory of gravity.

Many such theories exist, we will briefly discuss here the most simple one: the *minimal supersymmetric extension of the Standard Model* (MSSM). In the MSSM each fundamental fermion has its bosonic superpartner, the s-fermion. Likewise, every boson has its bosonic duplicate: the bosino.

The first welcome result about SUSY is that it naturally solves the hierarchy problem. Every quadratic contribution appearing in the scalar self-energy (Equation (2.28)) involves a likewise contribution with opposite sign and same coupling that cancels the quadratic term in the cut-off, hence recovering a well-behaved logarithmic behavior in the mass counterterms.

In MSSM, in order for the theory to be anomaly free, two Higgs doublets with opposite hypercharge need to be postulated. The electroweak symmetry breaking then occurs in the following way: the scalar doublets spontaneously acquire a vev, v_1 and v_2, with $v_{weak} = \sqrt{v_1^2 + v_2^2}$ (the ratio is denoted as $\tan\beta = \frac{v_1}{v_2}$). Three out of eight d.o.f. of the complex scalars are then "eaten" by the longitudinal d.o.f. out the electroweak bosons W and Z, leaving a total of five massive real scalars: h^0, H^0, A^0, H^+ and H^-.

Besides the fact that there are 2 Higgs doublets in MSSM, the main difference with the Standard Model is that the self coupling in the Higgs potential is related to the electroweak couplings g_1 and g_2:

$$\lambda_{MSSM} = \frac{g_1^2 + g_2^2}{4} \tag{2.31}$$

This feature has the dramatic consequence at tree level:

$$m_{h^0} \leq M_Z \tag{2.32}$$

However, if higher orders corrections are included, one gets $m_{h^0} \leq 200 - 300$ GeV. This is why the scalar h^0 is called the "light scalar" in MSSM. We note that this bound is perfectly compatible with current experimental bounds on the Higgs mass. Due to the shape of the MSSM Higgs potential, the MSSM parameter space is fully determined by the knowledge of $\tan\beta$ and m_{A^0} and the mixing angle α, which is a function of m_{h^0}, $m_{H^0}^2$ and $m_{A^0}^2$.

In the large m_{A^0} limit:

$$\cos^2(\beta - \alpha) = \frac{M_Z^2 \sin^2 4\beta}{4m_{A^0}^2} \tag{2.33}$$

It is interesting to study the behavior of the light (h^0) and heavy (H^0) CP-even Higgs coupling to Standard Model particles in the limit $m_{A^0} \to \infty$ (decoupling limit). The coupling to the gauge bosons are those of the Standard Model times a factor that depends on the mixing angle α:

$$\begin{aligned} h^0VV &\sim \sin(\beta - \alpha) \\ H^0VV &\sim \cos(\beta - \alpha) \end{aligned} \tag{2.34}$$

In the decoupling limit the light scalar coupling W and Z becomes SM-like and the heavy scalar completely decouples. This implies that if $m_{A^0} \to \infty$, it is impossible to discriminate between the light MSSM scalar and the SM scalar. However we also see that for arbitrary mixing value, the coupling of the light MSSM Higgs to a gauge boson is always smaller that in the SM.

For the fermions, the couplings depend on the isospin eigenvalue. By noting f_u the upper component of the SU(2) doublets and f_d the lower component one has:

$$\begin{aligned} h^0 f_d f_d &\sim \sin(\beta-\alpha) - \tan\beta\cos(\beta-\alpha) \\ h^0 f_u f_u &\sim \sin(\beta-\alpha) + \frac{\cos(\beta-\alpha)}{\tan\beta} \\ H^0 f_d f_d &\sim \cos(\beta-\alpha) + \tan\beta\sin(\beta-\alpha) \\ H^0 f_u f_u &\sim \cos(\beta-\alpha) - \frac{\sin(\beta-\alpha)}{\tan\beta} \end{aligned} \tag{2.35}$$

Here also in the large m_A regime the light Higgs becomes SM-like. However when m_A is small, its coupling to the lower components of the fermion doublets gets enhanced by a factor $\tan\beta$. This is also the case for the heavy scalar H^0 in the large m_A regime. Hence, the rate of decays such as $h^0 \to b\,\bar{b}$ or $h^0 \to \tau\,\tau$ are enhanced. This has crucial implications for searches of MSSM higgs at the LHC.

We have reviewed here only Higgs related aspect of a simple SUSY model (for an extensive review of a bottom-up SUSY construction, as well phenomenology, see [31]). We have seen that such theories can nicely solve the naturalness problem. There are also several aspects that make such theories appealing: the natural involvement of gravity, the unification of couplings at the Planck scale, a natural candidate for dark matter (the LSP). Finally, SUSY theories, unlike TC, are not strongly coupled and can therefore be treated perturbatively.

Chapter 3

Higgs Boson Production at the LHC

Having derived the full set of interactions in the Standard Model it should in principle be possible to compute the cross section of any process involving elementary particles. In Chapter 1, we have seen that the non abelian nature of QCD forbids a perturbative approach for physics involving energies $E \leq \Lambda_{QCD}$. This is typically the case for collisions involving hadrons [1] where the partons (gluons and quarks inside a hadron) are themselves interacting via the exchange of low momentum particles.

In the first section of this chapter a general review of the theoretical ingredients that enter a hadronic collision calculations will be given. We will then enumerate qualitatively the Higgs production mechanisms and their relative importance at the LHC. Finally, we will look more deeply into the dominant production mode, by reviewing the techniques that allow to calculate its total rate as well as the differential distributions in terms of Lorentz invariant quantities such as the transverse momentum and the rapidity.

3.1 Hadronic processes

3.1.1 Parton Distribution functions and Factorization Theorem

As a first approximation partons inside a hadron can be treated as a cloud of point-like non-interacting particles. This is the case in the so-called *parton model* [32] approximation. In this framework hadrons are boosted to a reference frame in which they have infinite momentum. In this case the inner-hadron partonic motion is "slowed" by time dilation and the partons within a same hadron act as they "do not see each other". The scattering occurs in an incoherent fashion. In this model each parton carries a fraction x of the total longitudinal hadronic momentum.

The functions that parameterize the x distribution are called *parton distribution functions* (PDF's). They cannot be calculated from first principles since they describe the non-perturbative properties of hadrons and they need to be measured from experiments. However, some rules can be imposed, which can help in constraining them. For instance, if $f_p(x)$ is the PDF describing a parton p inside a hadron [2], the sum of momenta needs to be equal to the total hadron momentum:

$$\sum_p \int_0^1 x f_P(x) dx = 1 \tag{3.1}$$

[1] At the Large Hadron Collider the collisions involve protons.

[2] $f_p(x)dx$ is the probability of "finding" a parton p inside the hadron, with a momentum fraction between x and $x + dx$.

and in a proton since the number of expected up(down) quarks is 2(1) one can require:

$$\int_0^1 (f_u(x) - f_{\bar{u}}(x))dx = 2 \tag{3.2}$$

$$\int_0^1 (f_d(x) - f_{\bar{d}}(x))dx = 1 \tag{3.3}$$

Even though very attractive for its simplicity this model has its limitations. It is unrealistic that the probability of "finding" a parton inside a hadron does not depend on the scale at which one is probing it. Intuitively, as one increases the scale at which one probes a gluon, the gluon would eventually start to "look like" 3 gluons or a gluon and a $q\bar{q}$ pair. This implies that the probability of finding a gluon or a quark has changed. That is, PDF's are in fact scale dependent, $f_p(x) \equiv f_p(x, \mu^2)$. The evolution of PDF's with the scale is governed by the DGLAP equations [33–35]. They can written in a general form as:

$$\mu^2 \frac{\partial f_p(x,\mu^2)}{\partial \mu^2} = \frac{\alpha_S(\mu^2)}{2\pi} \sum_{p'} \int_x^1 \frac{dz}{z} P_{pp'}(\frac{x}{z}, \alpha_S(\mu^2)) f_p(z, \mu^2) \tag{3.4}$$

where the functions $P_{pp'}$ are the so-called *Altarelli-Parisi* splitting functions. They describe the probability of the splitting of 1→2 partons and they can, in principle, be calculated at any order in perturbation theory. Experimentally one can measure the PDF's at a given scale (in Deep Inelastic Scattering experiments for instance) and via these equations extrapolate to the desired scale [3].

In Figure 3.1 the proton PDF's of gluon and light quarks are shown. At a low enough momentum fraction the gluon PDF's dominates over light quarks and antiquarks. The DGLAP equations are the analogues of the renormalization group equations for the coupling constant. Whereas in the case of the renormalization of the coupling one *absorbs* the UV divergences appearing in loop contributions of Feynman diagrams, here the collinear divergences (i.e when partons are emitted collinearly) appearing in the splitting functions are absorbed into a redefinition of the PDF's at a new scale called factorization scale. This procedure introduces, in addition to the *renormalization scale* μ_R (defined as the scale at which α_S absorbs the divergent terms), the *factorization scale* μ_F.

Schematically one can write:

$$\alpha_S(\mu_R) = \alpha_{bare} + CT(\frac{\mu_R}{\Lambda_{UV}}) \tag{3.5}$$

$$f(x, \mu_F) = f_{bare}(x) + CT(\frac{\mu_F}{\Lambda_{IR}}) \tag{3.6}$$

where the Λ parameters are symbolic cut-offs that regularize the divergence under consideration and the counter-terms CT depend of course on the order of perturbation theory and on the process one is considering.

Having all these ingredients one can then in principle compute the total cross-section for hadronic process by separating the short distance physics, i.e the hard scattering, described by perturbative QCD, from the long distance physics, described by the PDF's.

[3] Here it is implicitly assumed that the partons have no intrinsic transverse momentum. In general, transverse momentum dependent (TMD) PDF's can be defined, but the description of such PDF's is out of scope here

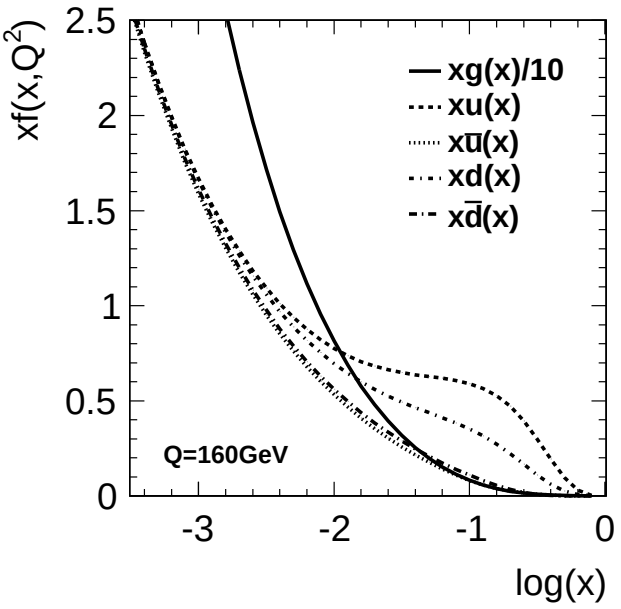


Figure 3.1: Parton distribution functions for gluons, and light quarks and antiquarks for a scale of 160 GeV. The CTEQ6.6 set was used [36]

$$\sigma_{had} = \sum_{p,p'} \int_0^1 dx_1 dx_2 f_p(x_1, \mu_F^2) f_{p'}(x_2, \mu_F^2) \hat{\sigma}(x_1, x_2, \mu_F, \mu_R) \tag{3.7}$$

This is the *factorization theorem* [37], it relates the total cross section $\sigma(PP \to X)$ to the partonic one $\hat{\sigma}$. More precisely, it explicitly states that the short distance processes are uncorrelated from the long distance ones. One can therefore calculate the hard scattering matrix element (ME) at a given order (however, as we will see in the next section, this is not always easy) and convolute it as in Equation (3.7) with PDF's that are *universal* functions, i.e they do not depend on the process under consideration. This decoupling of short range versus long range physics can qualitatively understood by the premise of this section. The partons inside the hadrons being boosted and therefore Lorentz contracted interact in a period of time that is much longer than that of the hard scattering. A schematic representation of hadronic collision is given in Figure 3.2.

For a general process the partonic cross section at the n-th order in α_S is given by:

$$\hat{\sigma} = \alpha_S^k \sum_{m=0}^{n} C_m \alpha_S^m \tag{3.8}$$

where the coefficients C_m depend on the kinematics of the process and the factorization and renormalization scales and can, at least in principle, be calculated for any order m. If $m = 0$ the process is said to be calculated at *leading order* (LO), if $m = 1$ at *next-to leading order* (NLO), etc .... If we note by Q the typical scale of the process under consideration, then the total hadronic (at the n-th order in perturbation series) cross section as :

$$\sigma_{had} = \sum_{p,p'} \int_0^1 dx_1 dx_2 f_p(x_1, \mu_F^2) f_{p'}(x_2, \mu_F^2) \sum_{m=0}^{n} \alpha_S(\mu_R^2)^{m+k} C_m(p_1, p_2, \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}) \quad (3.9)$$

It should be noted that the choice of μ_R and μ_F is completely arbitrary at this point [4]. The dependency on these scale should vanish if the calculation is made at all orders in perturbation theory. This can be summarized as the renormalization condition:

$$\frac{\partial \sigma}{\partial \mu_R} = 0 \quad (3.10)$$

$$\frac{\partial \sigma}{\partial \mu_F} = 0 \quad (3.11)$$

However, if one truncates the calculation at a given order, the dependency on μ_R and μ_F remains and a choice has to be made for these scales. Typically the choice is $\mu_R = \mu_F = Q$. A discussion on the uncertainty that derives from making such a choice will be given in the next sections.

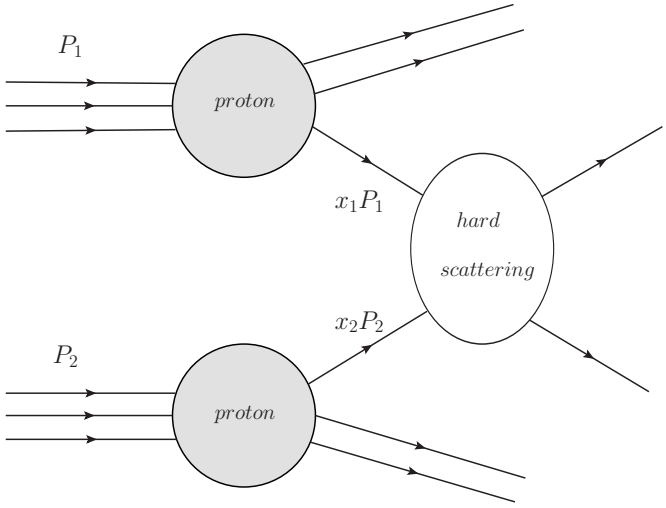


Figure 3.2: Schematic representation of a hadronic collision and of the factorization theorem

3.1.2 Kinematics in hadronic collisions

If we call $\sqrt{s}$ the energy in the center of mass frame of the two incoming protons, and $P_{1/2}$ the protons 4-momenta in the laboratory frame. We can write

$$s = (P_1 + P_2)^2 \quad (3.12)$$

In a similar way, we call $\sqrt{\hat{s}}$ the center of mass energy of the two partons [5], $p_{1/2}$ the partons 4-momenta, and $x_{1/2}$ the longitudinal fractions of the proton momenta, We have then:

$$\hat{s} = (p_1 + p_2)^2 = (x_1 P_1 + x_2 P_2)^2 \quad (3.13)$$

[4] Not exactly *any* choice is allowed. For instance μ_F has to be greater than Λ_{QCD} for obvious reasons.
[5] Here $\sqrt{\hat{s}}$ plays the role of Q in Equation (3.9)

Since $\sqrt{s} \gg m_P$ (≈ 1 GeV) the masses of the protons and partons can be neglected so that $P_{1/2}^2 = p_{1/2}^2 = 0$. We can therefore write:

$$x_1 x_2 = \frac{\hat{s}}{s} \tag{3.14}$$

Equation (3.14) implicitly states that the higher the energy of collisions [6] $\sqrt{s}$ the lower the possible values of x_i. This implies that at high energies gluon PDF's dominate [7].

By convention, the "z" axis is defined as the direction in which the collision occurs and "x","y" are the transverse directions. Rather than using the cartesian 4-component (or equivalently the polar coordinates) of momenta, it is convenient to operate with boost-invariant quantities. A 4-momentum $q_\mu = (E, q_x, q_y, q_z)$ is completely determined by two variables because of the on-shell condition and the rotational invariance around the z-axis. These are the transverse momentum q_T and the rapidity y.

$$q_T = \sqrt{q_x^2 + q_y^2} \tag{3.15}$$

$$y = \frac{1}{2} \ln \frac{E + q_z}{E - q_z} \tag{3.16}$$

$$= \frac{1}{2} \ln \frac{x_1}{x_2} \tag{3.17}$$

We also define the pseudo-rapidity as:

$$\eta = -\ln \tan(\frac{\theta}{2}) \tag{3.18}$$

where θ is the polar angle. Note that $\eta = y$ if the particle is massless (i.e $q^2 = 0$). Finally one can also note that:

$$x_1 = \sqrt{\frac{\hat{s}}{s}} e^{y} \tag{3.19}$$

$$x_2 = \sqrt{\frac{\hat{s}}{s}} e^{-y} \tag{3.20}$$

In particular, it is worth noting from (3.20) that $|y| \to \infty$ is equivalent to small momentum fractions. This means in practice that in order to probe such regions of the PDF's, one needs to study decays occuring in the forward region of the detector.

3.2 Higgs production at the LHC

3.2.1 Production mechanisms

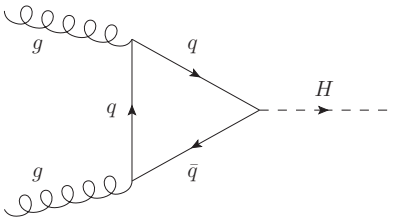
The set of interactions listed in Chapter 1 can help us in guessing which mechanisms for Higgs production are allowed in a hadron collider. Given that the initial state has to involve protons, the Higgs needs to be produced by an interaction of quarks or gluons. However, since it couples only to massive particles, it cannot *directly* be created by gluons.

Naively the first possibility is direct production from a quark (q) and an antiquark ($\bar{q}$). Since the $q\bar{q}H$ coupling is proportional to the mass of the quark m_q (see Equation (1.69))

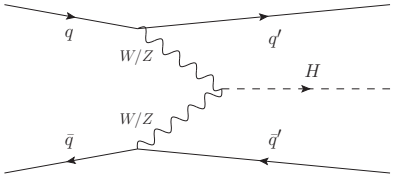
[6] At the LHC $\sqrt{s} = 7 - 14$ TeV.

[7] "low x" and high energies are therefore used interchangeably.

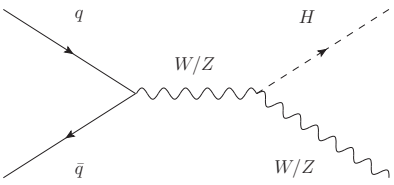
this mechanism would be eventually favored only for very massive quarks (top or bottom). However we know that the probability of "finding" a heavy quark inside a proton is very small. The heavy quarks parton distribution functions are not well-known (especially for quarks t and b) but from first principles one might guess that this probability is non-negligible only at very high energy. Therefore direct heavy quark higgs production cannot contribute much to the cross section. The Higgs has therefore to be produced by gluons or light quarks radiating heavy particles that couple substantially to the Higgs.



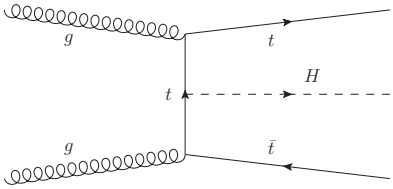
(a) Gluon-Fusion



(b) Vector-Boson-Fusion



(c) Higgs-Strahlung



(d) Associated production

Figure 3.3: Leading order Feynman diagrams of the main Higgs boson production processes at the LHC

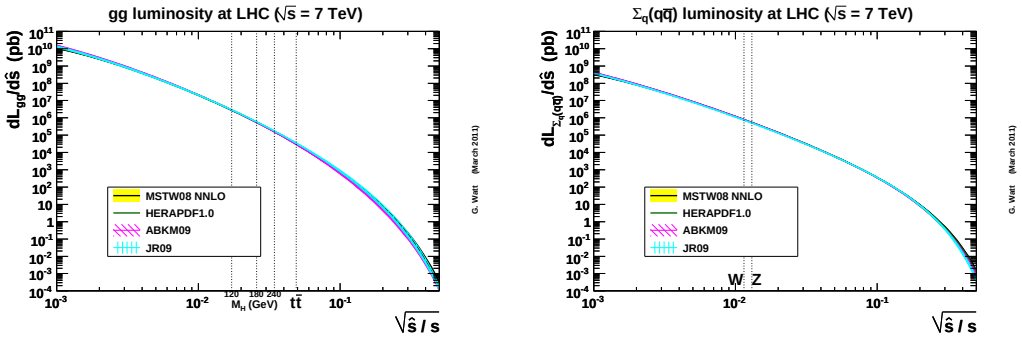
The main production mode at the LHC is the so-called *gluon fusion* mode (ggF). The diagram related to this process is shown in Figure 3.3(a). The incoming gluons couple to a quark loop that couples directly to the Higgs. One might think that this production mode is negligible since it involves a loop (and therefore the rate is $\mathcal{O}(\alpha_S^2)$). However two aspects need to be taken into account. First, the high top mass value ($m_t \sim 10^5 m_{u,d}$) substantially enhances this contribution.

The second reason is less obvious and can be understood by looking at *parton-parton luminosities* (PPL) that are defined as:

$$\mathcal{L}_{pp'}(\frac{\hat{s}}{s}) = \sum_{p,p'} \int_0^1 dx_1 dx_2 f_p(x_1) f_{p'}(x_2) \delta(x_1 x_2 - \frac{\hat{s}}{s}) \tag{3.21}$$

where $\sqrt{\hat{s}}$ is the parton-parton invariant mass and $\sqrt{s}$ is the energy in the center of mass frame of the proton-proton collisions. This function is therefore proportional to the joint probability of extracting two partons with a given a invariant mass. Quark-antiquark (light quarks) and gluon-gluon PPL's are shown in Figure 3.4(a) and (b) for $\sqrt{s} = 7$ TeV [38]. From this plot it is clear that for LHC energies, the gg luminosity exceeds $q\bar{q}$ luminosity over a wide range of partons-parton invariant masses. This has to do of course

with the fact the gluons dominate inside the proton at high energies (see Equation (3.14)). For values of $\sqrt{\hat{s}} \sim 100$ GeV, the gg luminosity exceeds $q\bar{q}$'s by a factor 10 while for $\sqrt{\hat{s}} \sim 1$ TeV they have a similar value. This argument explains why the gluon fusion Higgs production is enhanced at the LHC as long as $m_H \leq 1$ TeV.



(a) (b)

Figure 3.4: Parton-parton luminosities from different PDF sets at the LHC for a center of mass energy of 7 TeV for gluon-gluon (a) and quarks antiquarks (b) as a function of the momentum fractions product $x_1 x_2 = \frac{\hat{s}}{s}$. The value of $\frac{\hat{s}}{s}$ for some Standard Model processes is given as a reference by a vertical line.

The second most important contribution at the LHC is *vector boson fusion* (VBF). The Higgs is generated via a fusion of bosons $W^{\pm}$ or Z^0 that are themselves radiated off the incoming quarks (see Figure 3.3(b)). Even though disfavored by the fact that light quark parton luminosity is small at the LHC, this mechanism still contributes in a non negligible way to the total cross section ($\sim 5\%$). Moreover the presence of two forward jets in addition to the Higgs decay products can give a handle on discriminating this topology (the signal) from those electroweak processes that decay like the Higgs but do not have two forward jets in the final state (the backgrounds).

Finally, two other mechanisms are worth to mention: the *Higgs-Strahlung* (the Higgs is radiated off a $W^{\pm}$ or Z^0 boson) and the *Associated production* (similar to ggF, the difference being that the heavy quarks appear in the final state). These processes are shown in Figure 3.3(c) and (d). The rates of these production modes are one order of magnitude below VBF. The reason for this is that more than one very massive object are produced in the final state. This requires additional momentum carried by the initial state partons. As for the VBF case, additional particles are present in the final state (other than the decay products of the Higgs) allowing for better discrimination against backgrounds.

In Figure 3.5 the cross section and the uncertainties from theory for these processes are plotted as a function of the Higgs mass [15]. A more detailed discussion on the computation and the uncertainties will be given later in this Chapter.

3.2.2 Gluon Fusion: inclusive cross section

As we have argued in the previous section Higgs Gluon Fusion production is predominant over the other production mechanism at the LHC along a wide range of possible Higgs masses. We will therefore concentrate on this mechanism in the following discussions.

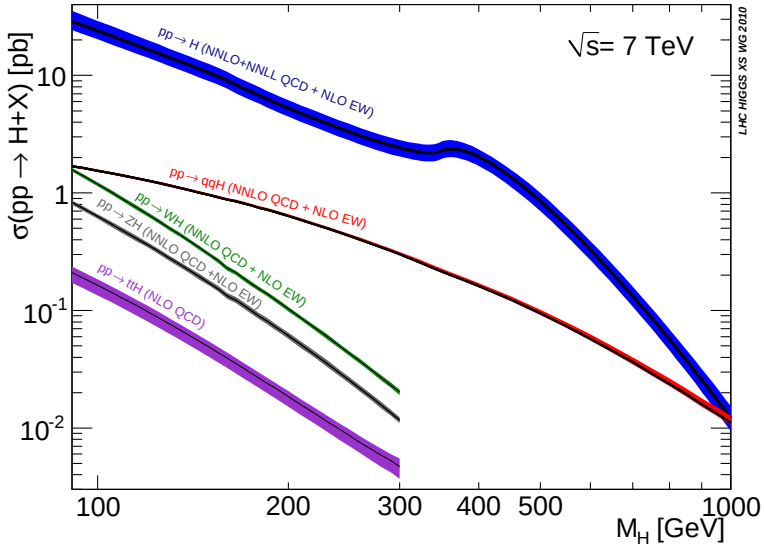


Figure 3.5: Cross section of the various Higgs production mechanisms at the LHC for a center of mass energy of 7 TeV as function of the Higgs mass. The error bands correspond the theoretical uncertainties. The process are, from top to to bottom, Gluon Fusion (blue), Vector Boson Fusion (red), Higgs Strahlung (green/grey) and Associated $t\bar{t}$ production (purple)

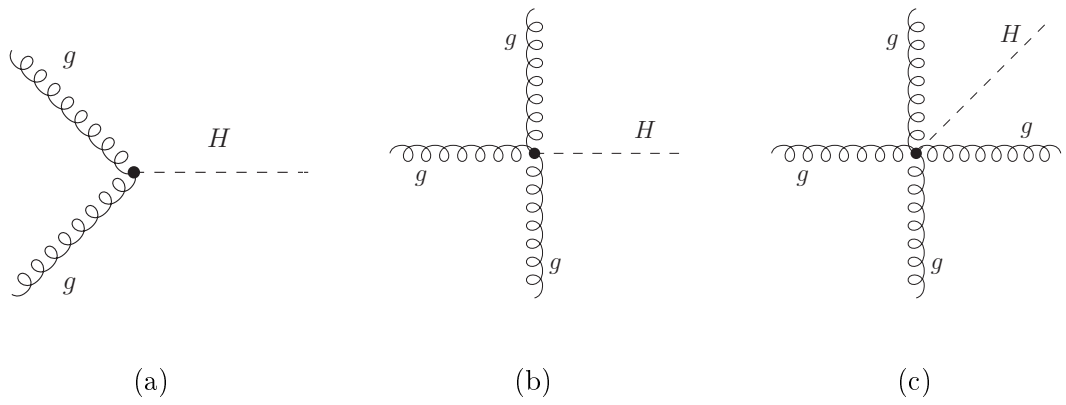
Leading Order

The leading order diagram (Figure 3.3(a)), although already containing a loop of heavy quarks, is easy to compute. The apparent UV divergence in the loop get regularized by working in dimensions $d = 4 - 2\epsilon$, and by taking $\epsilon \to 0$ at the end of the calculation. By considering only top quarks in the loop one gets the interesting result as a function of the gluon-gluon luminosity for the total hadronic cross section at LO:

$$\sigma(PP \to H) = \frac{\alpha_S^2}{64\pi v^2} \left| I\left(\frac{m_H^2}{m_t^2}\right) \right| \frac{m_H^2}{s} \mathcal{L}_{gg}\left(\frac{m_H^2}{s}\right) \tag{3.22}$$

Here I(x) is a complex function. $|I(x)|^2$ diverges at $x \to 0$ ($m_H \to 0$), has a maximum at $x \simeq 4$ ($m_H \simeq 2m_t$) and goes to 0 when $x \to \infty$ ($m_H \to \infty$). This result is very attractive qualitatively since it already describes accurately the shape of the cross section as a function of the Higgs mass, compared to higher order calculations (see Figure 3.5). Note that the cross section is already $\mathcal{O}(\alpha_S^2)$ at LO (corresponding to $k = 2$ in Equation (3.8)).

However this result cannot be used quantitatively. The reason is the following: since this production mechanism involves gluons the higher order QCD corrections can be very important opposed as if the initial state was made of light quarks. This is because of the many possible combinations of gluon-gluon and gluon-quark couplings. Moreover additional gluon radiation can come from the loop of top quarks. Therefore, for quantitative predictions higher order corrections are needed.



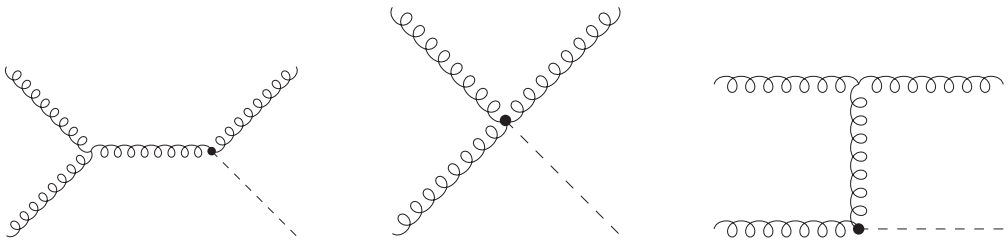
(a) (b) (c)

Figure 3.6: Vertices of the effective theory in the limit $m_t \to \infty$

Higher Orders

Many diagrams contribute to the partonic matrix element at high orders. In order to reduce them, one can make the approximation of the large top mass limit, $m_t \to \infty$ which is valid only in regions $m_H \ll m_t$. This leads to an effective theory in which the top loop loop gets reduced to a single vertex that directly couples the gluons to the Higgs. This theory is described by the following Lagrangian:

$$\mathcal{L}_{eff} = -\frac{1}{4}\left[1 - \frac{\alpha_S}{3\pi}\frac{H}{v}\right] G^a_{\mu\nu} G^{a\mu\nu} \tag{3.23}$$



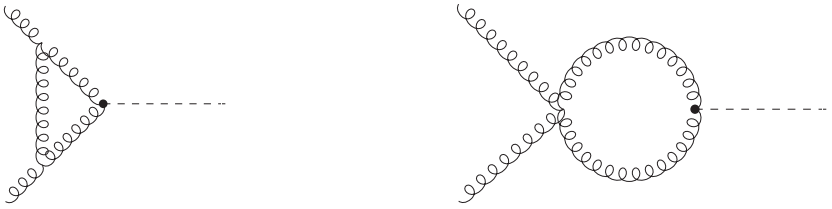
(a) (b) (c)

Figure 3.7: Some diagrams corresponding to the real emissions at NLO in the effective theory

The vertices of the theory at LO are given in Figure 3.6.

The high order corrections to the Born amplitude are given by virtual and real corrections. One needs to compute them separately and then add the corresponding amplitudes. However in these high order contributions new divergences appear when the partons are emitted collinearly (*collinear divergences*) and when their energy $E \to 0$ (*soft divergences*).

These divergences are not physical, they just reflect the breakdown of the perturbative approach in ranges of energy where it cannot apply anymore since confinement starts to play a role. Therefore they need to be regularized in some way. The way this is done is again with dimensional regularization , i.e computing the amplitudes in $4-2\epsilon$ dimensions.



(a) (b)

Figure 3.8: Some diagrams corresponding to the virtual corrections at NLO in the effective theory

By doing this, the *Infrared* (IR) divergence that is present in the virtual contribution (Figure 3.8) cancels with that of real emissions (Figure 3.7). In the real emission amplitudes, collinear divergences and UV divergences are still present and one has to absorb them into a redefinition of the PDF's and the coupling α_S, as shown by Equations (3.6), by introducing the two new scales μ_R and μ_F.

By following this procedure one can finally get a finite result for higher orders gluon fusion production cross sections. However as one increases the order in perturbation theory the computation becomes more and more complicated, due to the increasing number of diagrams, and the opening of new channels.

Sources of uncertainties

There are several sources of uncertainties entering these calculations and having an impact on the final result:

- As the order of the calculation increases, the number of diagrams involved increase exponentially and also the computation time. One needs therefore to stop the calculation at a given order. As of today the most accurate calculations reach the NNLO level. It is therefore necessary to define a procedure that allows to associate an uncertainty to the fact that higher orders are ignored. One might naively think that if the calculation is stopped at a given order n then correction need to be at most of a magnitude $\mathcal{O}(\alpha_S^{n+1})$. This is true only to certain extent, because as already mentioned, other channels can open up, increasing the number of diagrams substantially. In principle the only way to accurately estimate which fraction of the cross section is neglected by truncating the series at a given order, is to calculate it at the next order. When this is not possible one can vary the factorization and renormalization scales and see which impact this variation has on the final result. After all, if the result calculated at all orders in perturbation theory should not depend on these scales, then how much it depends on the values chosen for μ_R and μ_F at a given order should give a rough estimate of the higher orders contributions.

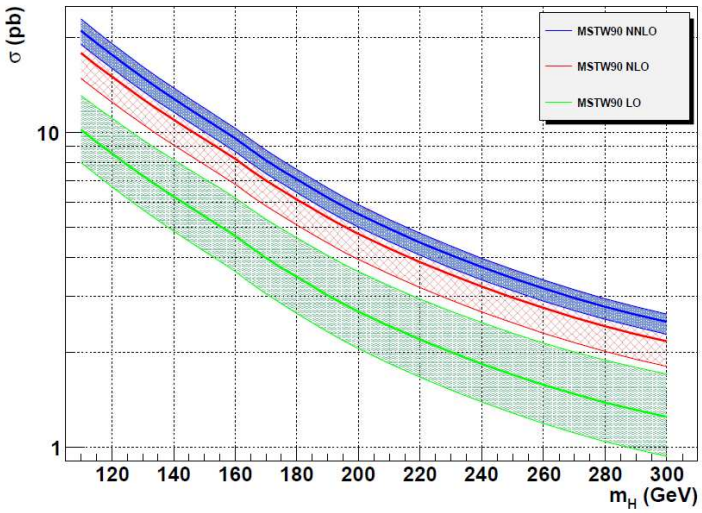


Figure 3.9: Cross section of gluon fusion Higgs production as a function of the Higgs mass at LO, NLO and NNLO. The uncertainty bands correspond to the variation of the renormalization and factorization scales in the range $\mu \in [\frac{\mu_0}{2}; 2\mu_0]$. The central value here was fixed at $\mu_R = \mu_F = \mu_0 = \frac{m_H}{2}$. This calculation uses the PDF set MSTW 2008.

In Figure 3.9 we show the LO, NLO, NNLO fixed order cross sections [39] . First we notice that the ratio between LO and NLO is almost a factor ~ 2 and it decreases between NLO and NNLO. We define the K-factor as the ratio between the cross section calculated at a given order n divided by the contribution at $n-1$.

$$K_{NLO} = \frac{\sigma_{NLO}}{\sigma_{LO}} \simeq 2 \tag{3.24}$$

$$K_{NNLO} = \frac{\sigma_{NNLO}}{\sigma_{NLO}} \simeq 1.2 \tag{3.25}$$

This indicates a good convergence of the perturbative series. To estimate the uncertainty bands the scale were varied to the extreme values $[\frac{\mu_0}{2}; 2\mu_0]$ with $\mu_R = \mu_F = \mu_0 = \frac{m_H}{2}$. This is the convention [8] used in [39]. What is also clear is that the dependence of the cross section on μ_R and μ_F substantially decreases with increasing the order of the calculation, validating the hypothesis that the dependence on these parameters controls the magnitude of higher order terms.

- Another important source of uncertainty is the choice of the PDF's. Many groups extract PDF's from data [36], but they use different conventions and methods, and not always the same data. Each group provides their own uncertainty (the procedure that allows these uncertainties to be estimated will be discussed later). Here we stress that the difference in the choice of the PDF set can have a big impact on the

[8] As mentioned earlier, the choice of the central value μ_0 is arbitrary. The other common choice is $\mu_0 = m_H$. The choice of varying the central scale between $[\frac{\mu_0}{2}; 2\mu_0]$ is also arbitrary. As it is clear from Figure 3.9, this choice underestimates the uncertainty of the LO, therefore some authors [40] argue that a more appropriate variation of the scales would be $[\frac{\mu_0}{3}; 3\mu_0]$.

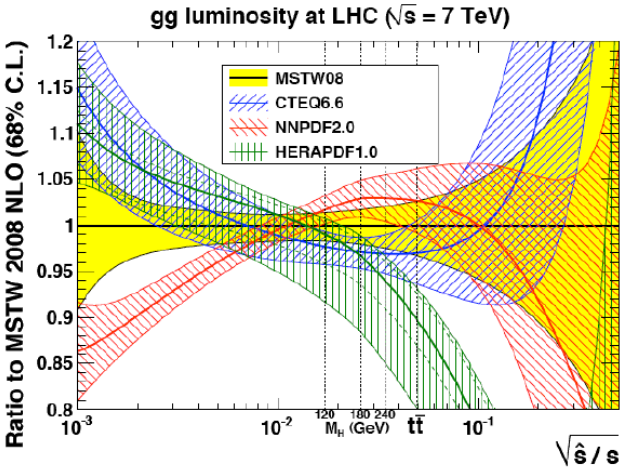


Figure 3.10: Gluon-gluon luminosities uncertainties normalized to MSTW2008 from different PDF sets.

inclusive cross section. In Figure 3.10 the uncertainties of the gluon gluon luminosity (normalized to the luminosity obtained with the reference PDF set MSTW2008) as a function of the normalized center of mass energy with different PDF's is shown. In the region of a light Higgs there is a quite good agreement among the different PDF sets.

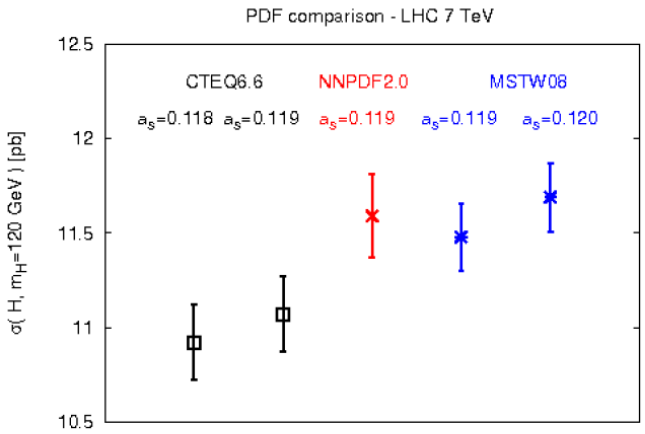


Figure 3.11: Higgs cross section uncertainty ($m_H = 120$ GeV) at NLO with different choices of α_S made by different groups.

Also, the uncertainty on the measurement of α_S have to be included. In Figure 3.11 the cross section value at NLO with different choices of α_S made by different groups is shown for $m_H = 120$ GeV is shown. The differences are of the order of a few percents.

Threshold Resummation and Electro-Weak corrections

As of today the most accurate calculation of gluon fusion Higgs production at fixed order reach the NNLO level in α_S. Rather than just "sizing" the higher order QCD corrections by varying the renormalization and factorization scales, one can observe that the most significant contribution from NLO to NNLO is obtained by accurately *resumming* soft gluons emissions. One can therefore apply this soft gluon approximation to estimate higher order terms. These soft gluons arise when the available energy in the center of mass frame is just above the Higgs mass threshold, i.e $\hat{s} = x_1 x_2 s \geq m_H^2$. In this case one can show that large logarithms of the form $\alpha_S^n \log^m(1-z)$ arise, where $z = \frac{m_H^2}{\hat{s}}$. Briefly, the resummation procedure consists in resumming a carefully chosen subset of these logarithms *at all orders* in perturbation theory (more details will be given later) so that the series become convergent. When one sums only the $\alpha_S^n L^{2n}$ terms [9] the resummation is said to be at the *leading logarithm* (LL), if one resums $\alpha_S^n L^{2n-1}$ terms the resummation is said to be at the *next-to leading logarithm* (NLL), etc .... The overall increase in the cross section at the NNLL level are of the order of $\sim 7-9\%$ [15].

Here we have discussed only QCD corrections, since they are the most relevant for ggF, but other corrections are also present. The quark loop can also be subject to emission of photons, or even electroweak bosons. These corrections are small compared to the gluon and quark emission, since the difference in value of the couplings ($\alpha_{em} \simeq \frac{1}{137}$ while $\alpha_S \simeq 0.12$) is important and the combinatorics favors QCD corrections, but in principle they still should be considered. The main difficulty in computing electroweak corrections comes from their mixing with QCD effects. In the *complete factorization* formalism [15] the two effects factorize. They modify the cross section by a factor that varies as a function of the Higgs mass ($\sim 5\%$ for $m_H = 120$ GeV and $\sim -2\%$ for $m_H = 300$).

3.2.3 Gluon Fusion: differential distributions

In the previous paragraph we have sketched the main ideas that allow to compute with high accuracy the gluon fusion integrated cross section. Once the Higgs is produced, one might also be interested in knowing how the rate of Higgs production is distributed geometrically. More specifically we have seen that the 4-momentum of a particle in high energy collisions can be fully specified by the boost invariant variables p_T and y.

Rapidity distribution

In Equation (3.17) we have related the pseudo-rapidity to the parton momenta. The pseudo-rapidity (at LO and in the limit of a massless Higgs) equals to:

$$\begin{aligned} \eta &\simeq 0 \quad \text{if} \quad x_1 \sim x_2 \\ \eta &\simeq \pm\infty \quad \text{if} \quad x_1 \gg x_2 \text{ or } x_2 \gg x_1 \end{aligned} \tag{3.26}$$

This implies that if the initial parton momenta are balanced, the Higgs will be decaying in the central part of the detector ($\theta = \frac{\pi}{2}$). On the other hand, if the parton are very unbalanced, the Higgs will be produced in the forward region ($\theta = 0, \pi$). However, at LO we know that initial parton momenta are constraint to satisfy $x_1 x_2 \hat{s}$. Qualitatively,

[9] Here $L \sim \log(1-z)$

the bulk of the contribution will come from momenta that satisfy $\hat{s} \simeq m_H^2$. Now, if we write $\tau = x_1 x_2$ and $g(x)$ the gluon PDF, then the product of gluons PDF's $g(x)g(\frac{\tau}{x})$ has a maximum at $x_0 = \sqrt{\tau}$. [10].This implies that bulk of the Higgs production will occur when $x_1 \simeq x_2 \simeq \sqrt{\tau} = \frac{m_H}{s}$. Therefore, the distribution $\frac{d\sigma}{d\eta}$ is expected to be peaking at values $\eta \simeq 0$ and decreasing as $\eta \rightarrow \pm\infty$.

Transverse momentum distribution

At the lowest order the Higgs is obviously produced with a vanishing transverse momentum because of the absence of radiation. The corresponding distribution is therefore $\frac{d\sigma}{dp_T^2} \propto \delta(p_T^2)$. In order to generate transverse momentum one needs therefore to consider higher order terms. A non-vanishing Higgs p_T corresponds to a final state containing additional partons, that eventually will hadronize into jets. From an experimental perspective having an accurate prediction for the Higgs p_T spectrum can be useful in order to understand the jet structure of Higgs events in a enhanced QCD environment such as the LHC.

The partonic (p_T) differential cross section can be written as an expansion in the α_S coupling:

$$\frac{d\hat{\sigma}}{dp_T^2} = \frac{\sigma_0}{\hat{s}}\left[\hat{s}\delta(p_T^2) + \frac{\alpha_S(\mu_R)}{2\pi}G^{(1)} + \left(\frac{\alpha_S(\mu_R}{2\pi}\right)^2 G^{(2)} + \ldots\right] \tag{3.27}$$

where σ_0 is the born cross section defined as

$$\sigma_0 = \frac{\pi}{64}\left(\frac{\alpha_S(\mu_R)}{3\pi v}\right)^2 \tag{3.28}$$

When dealing with the differential p_T distribution, the convention that is often used is LO=$\mathcal{O}(\alpha_S)$, NLO=$\mathcal{O}(\alpha_S^2)$, etc .... [11] This matrix element describes very well the spectrum as long as the Higgs $p_T \simeq m_H$. On the other hand, the terms $G^{(i)}$ contain logarithms of the form $\alpha_S^n \log^m(\frac{m_H^2}{p_T^2})$ that systematically diverge at low p_T.

To see how these logarithmic divergences emerge, let us consider the simple case of gluon emission out a $q\bar{q}$ final state (see Figure 3.12). The denominator of the quark propagator is $\propto \frac{1}{(p+k)^2} \sim \frac{1}{2pk} \propto \frac{1}{E_g(1-\cos\theta_{qg})}$. The integral over the final state space phase leads therefore to a divergence when $E_g \sim 0$ (soft divergence) and when $\theta_{qg} \sim 0$ (collinear divergence):

$$\alpha_S \int \frac{dE_g}{E_g} \int \frac{d\theta_{qg}^2}{\theta_{qg}^2} \sim \alpha_S \log^2(\ldots) \tag{3.29}$$

These are residual divergences that survived after the redefinition of the scales and the cancellation with the virtual contributions. In general these logs correspond to a gluon soft emission ($|p_T| \sim 0$) or collinear ($\theta \sim 0$). They come from a singularity in the propagator.

[10] One can easily show this by taking the following parameterization for the gluon density, $g(x) = \frac{(1-x)^5}{x^{0.2}}$

[11] This convention is somewhat confusing. What is called LO in the case of $\frac{d\sigma}{dp_T}$ corresponds to NLO in the case of the integrated cross section. This is simply because when dealing with $\frac{d\sigma}{dp_T}$ the actual LO is $\propto \delta(p_T^2)$ and therefore is not interesting.

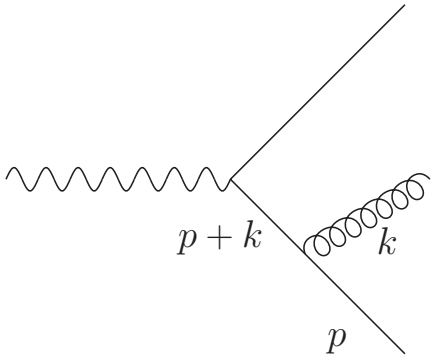


Figure 3.12: Diagram of gluon emission out of a $q\bar{q}$ final state.

Formally, when $p_T \to 0$, one can write an observable R, as [12]:

$$\begin{aligned} R &= R_0\big[1+\sum_{n=0}^{\infty}\alpha_S^n(a_{2n}L^{2n}+a_{2n-1}L^{2n-1}+\ldots)\big] \\ &\sim R_0\big[1+\alpha_S(L^2+L+1)+\alpha_S^2(L^4+L^3+L^2+L+1)+\ldots\big] \end{aligned} \tag{3.30}$$

where $L=\log(\frac{m_H^2}{p_T^2})$. At leading order, the term $\alpha_S L^2$ displays both a collinear and a soft emission, while $\alpha_S L$ only a soft. The ratio of two successive terms in the α_S expansion is always of the order $\sim \alpha_S L^2$ which is ~ 1 if L is large and α_S is small. Naively, adding all these terms of $\mathcal{O}(1)$ would spoil the convergence of the perturbative expansion. However, it has been shown [41] that (3.30) can be exponentiated in the following form:

$$\begin{aligned} R &= R_0\exp\big[\sum_{n=1}^{\infty}\alpha_S^n\sum_{m=0}^{m+1}G_{nm}L^m\big] \\ &\sim R_0\exp\big[\begin{array}{llll} (\alpha_S L^2+\alpha_S^2L^3+\alpha_S^3L^4+\ldots)+ & \to & Lg_1(\alpha_S L) & \text{(LL)} \\ (\alpha_S L+\alpha_S^2L^2+\alpha_S^3L^3+\ldots)+ & \to & g_2(\alpha_S L) & \text{(NLL)} \\ (\alpha_S^2 L+\alpha_S^3L^2+\alpha_S^4L^3+\ldots)+ & \to & \alpha_S g_3(\alpha_S L) & \text{(NNLL)} \\ \mathcal{O}(\alpha_S^3 L)\big] & & & \end{array} \end{aligned} \tag{3.31}$$

This advantage of rearranging the terms this way is that the ratio of two successive columns is $\alpha_S L \sim 1/L$, which implies that each line is formally a convergent series. The functions $Lg_1(\alpha_S L)$, $g_2(\alpha_S L)$ and $\alpha_S g_3(\alpha_S L)$ are said to respectively resum the leading logs (LL), the next-to leading logs (NLL) and the next-to-next-to leading logs (NNLL). When resumming soft gluons one needs to carefully avoid double counting with contributions already present in the matrix element.

This can be summarized in Table 3.1. Every line corresponds to a fixed order calculation (i.e perturbative expansion in α_S) while a column corresponds to a log expansion (i.e. expansion in $\alpha_S L$). A NLO+NNLL calculation would correspond therefore to formally add the first two lines (NLO) plus the first three columns(NNLL) and subtract the intersection.

[12] This formalism introduced here works for threshold resummation as well. Both methods resum emissions of soft gluons. When calculating $\frac{d\sigma}{dp_T}$, it is often referred as *recoil resummation*

	LL	NLL	$NNLL$	...
$LO\ \mathcal{O}(\alpha_S)$	$\alpha_S L^2$	$\alpha_S L$	...	...
$NLO\ \mathcal{O}(\alpha_S^2)$	$\alpha_S^2 L^4$	$\alpha_S^2 L^3$	$\alpha_S^2 L^2$	$\alpha_S^2 L$
...	...	...	...	...
$N^nLO\ \mathcal{O}(\alpha_S^n)$	$\alpha_S^n L^{2n}$	$\alpha_S^n L^{2n-1}$	$\alpha_S^n L^{2n-2}$	...

Table 3.1: Correspondence between the fixed order perturbative expansion and the expansion resulting from soft gluon resummation

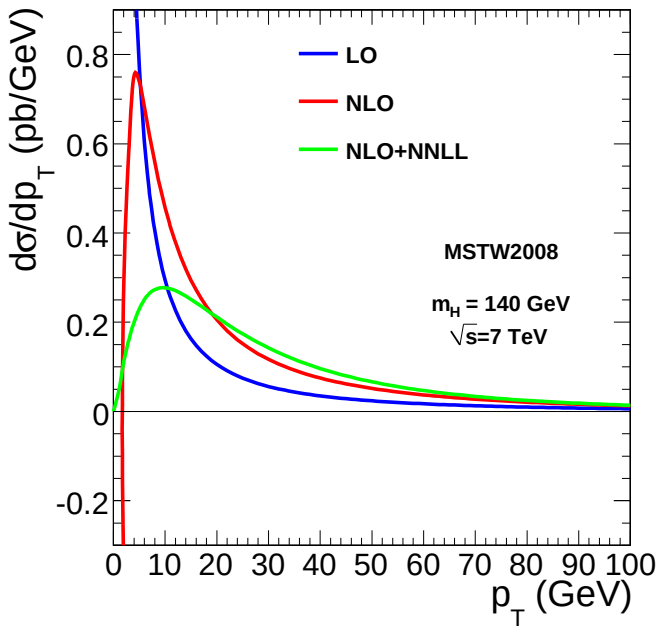


Figure 3.13: p_T spectrum comparison at LO, NLO and NLO+NNLL. At $p_T \to 0$ the LO calculation diverges to $+\infty$, the NLO diverges to $-\infty$ and the NLO+NNLL is finite. This plot was obtained with HqT2.0 [42]

By resumming these soft gluon emissions the p_T spectrum gets finite. In Figure 3.13 we show the p_T spectrum resulting from a fixed order calculation a LO and NLO compared to the NLO+NNLL resummed matched result. The first comment is that the resummation has an impact only a low p_T. The hard (high-p_T) part of the spectrum is well described by the fixed order NLO calculation. However, one can clearly see that the both the LO and NLO result without resummation diverge at low p_T and by resumming soft gluons a finite result is recovered. This is indeed what one would expect from first principles. The probability for a vanishing Higgs p_T (i.e no gluon radiation, initial partons stay collinear along the "x" axis) has to be zero.

As for the inclusive cross section, the p_T spectrum shape and normalization depends on the choice of the factorization and renormalization scales. The uncertainties are shown by the bands in Figure 3.14 where the scales have been varied in the range $[\frac{m_H}{2}, 2m_H]$. Both Figures 3.13 and 3.14 were obtained with HqT2.0 [42].

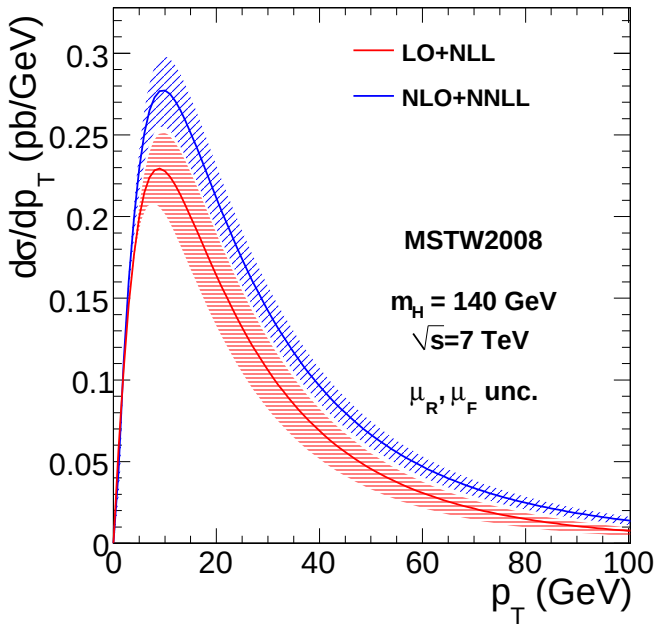


Figure 3.14: p_T spectrum comparison between NLO+NLL and NNLO+NNLL. The scale uncertainties are included.

Monte Carlo (MC) Event generators

In a high energy collider experiment, rather than observing total rates one observes events (which can be seen as "screenshots" of the detector at a given time). Distributions of kinematic variables of the decay products are then populated with data taking over a given period of time. In order to be able to compare this distributions with theory , one therefore needs to have an *event by event* description of the collision product at the parton level, then apply an algorithm that correctly treats the hadronization of radiated partons and ultimately a procedure that simulates the actual deposits of the final state particles into a real detector.

Moreover, not all differential distributions are analytically calculable. When many particles are present in the final state the phase space integration can become very difficult to compute; one needs therefore to fall back on Monte Carlo integration methods.

A Monte Carlo Event generator is a program that is able to reproduce particle physics events resulting from a collision with the same probability as they occur in nature. It is able to generate the 4-momenta of all the particles appearing in the final state. As explained in the previous section, when partons are well separated, i.e when no collinear of infrared divergences can occur, the matrix element (ME) description is good enough. However, we have seen that when this is not the case, large logarithms arise and need to be resummed.

Parton Shower The idea behind the parton shower formalism is that a parton emission will undergo several successive splitting 1→2 partons until the typical energy of the partons is $\mathcal{O}(\Lambda_{QCD})$. At that point the elementary partons will hadronize into QCD jets. The main ingredient of a parton shower description is the so-called *Sudakov form factor*:

$$\Delta_p(t) = exp\left[-\sum_p{}' \int_{t0}^{t} \frac{dt'}{t'} \int dz \frac{\alpha_S}{2\pi} P_{p'p}(z)\right] \tag{3.32}$$

where the $P_{p'p}(z)$ are again the *Altarelli-Parisi splitting functions*, $t = \mu^2$. The $P_{p'p}(z)$ are related to the probability that a parton p with momentum q radiates a parton with momentum zq. On the other hand the Sudakov form factor can be interpreted as the probability for a parton *not* to split when evolving from a scale t_0 to a scale t. [13] They can be used in an iterative Monte Carlo procedure, together with the Altarelli-Parisi equations to generate the sequential splittings of partons, i.e the parton shower. This method not only provides a full kinematic description of the jets in the final state, it also naturally includes the contribution resulting from the resummation of the IR divergent leading logarithms.

This procedure allows to generate radiation starting from the partons in the final state : they start with a high virtuality and decrease until the hadronization scale $\sim \Lambda_{QCD}$ is reached. This *forward evolution* generates final state radiation (FSR). On the other hand, *backwards evolution* generates Initial State Radiation (ISR).

The actual implementation of the algorithm can vary. In this thesis we mostly use PYTHIA [43] in which the ordering parameter is the virtuality, in contrast to other generators such as HERWIG [44] that use angular ordering.

NLO Monte Carlo Generators Both PYTHIA and HERWIG provide a good description of soft gluon emissions via the parton shower formalism. In addition, PYTHIA provides a good description of the underlying event (multi-parton interaction) and the hadronization (the LUND Model [45]). However, both these generators compute the hard scattering at LO only. This gives rise, in the case of the Higgs production, to a poor description of the hard part of the p_T spectrum. Ideally one would like to merge the parton shower approach valid for the soft p_T region, with high order matrix element calculation without double counting.

The most common NLO MC generators are MC@NLO [46] and POWHEG [47–49]. They have a similar approach in the sense that the $\mathcal{O}(\alpha_S)$ emission is always considered to be the hardest, therefore this contribution is removed from the parton shower since better described by the ME. While MC@NLO has the disadvantage to provide negative weighted events, POWHEG only provides positive weights. The combination of high order ME calculation with parton showers is a very active (and recent) research field. An example is the MadGraph [50] effort that provides a fully automated matrix element computation at NLO merged to showering and hadronization algorithms. An extensive review of Monte Carlo event generators is given in [51].

[13] The argument inside the exponential can be seen as the expected number of splittings between t_0 and t. The Sudakov is then interpreted as a Poissonian distribution with $n = 0$, i.e probability of non-emission.

Chapter 4

The LHC and the CMS experiment

The CERN ("Centre Europeen pour la Recherche Nucleaire") has played a major role in particle physics since its creation in 1954. Crucial discoveries such as the W and Z boson discoveries in the UA1 and UA2 experiments [52–55] have led to a Nobel Prize for Rubbia and Van der Meer in 1984. Other important results such as the determination of the number of neutrino families [56] or the discovery of direct CP-violation in neutral kaons decays [57, 58] are also important achievements of CERN experiments.

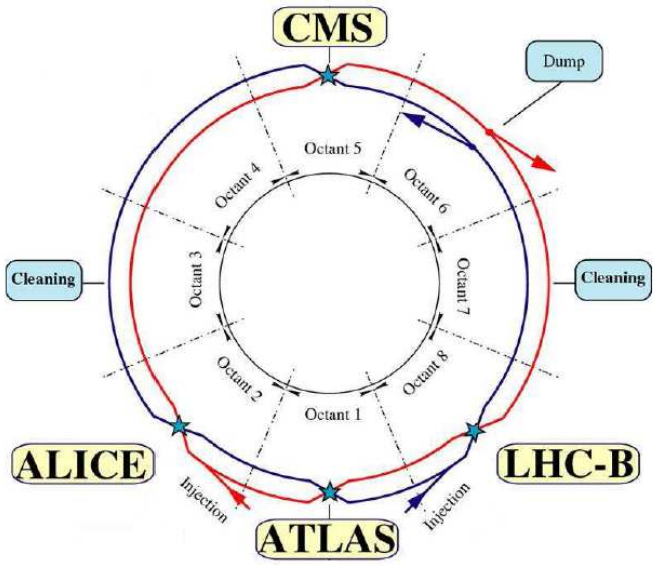


Figure 4.1: Layout of the LHC. The two beams are in red (clockwise) and blue (anti-clockwise). The experiments ATLAS, LHCb, CMS and ALICE are located at the four collision points.

Today CERN is hosting the most powerful hadron accelerator in world, the Large Hadron Collider (LHC) [59–61]. The LHC is an underground ring that accelerates protons and heavy ions and makes them collide in four distinct points, where are located the four

major LHC detectors: ATLAS [62, 63], LHCb [64], CMS [65–67] and ALICE [68] (see Figure 4.1).

ATLAS and CMS are multipurpose detectors as they serve the scope of probing new physics as well as performing precision measurements. The primary goal of these experiments, at least in the short term is to elucidate the nature of the electroweak symmetry breaking mechanism. On the long run, these experiments will be able to exclude or confirm new physics, as long as it manifests at the $\mathcal{O}(1$ TeV) scale. The LHCb was specifically designed for studying B-physics and CP-violation in general. On the other hand, Alice's purpose is to study the hot dense medium produced in heavy-ions collisions.

In this Chapter we will first describe the main LHC characteristics, and how it did operate since the first days of its operations. We will then focus on the CMS experiment, by explaining how in general the collision products can be analyzed with its various sub-detectors and how the data acquisition and the detector simulation is made.

4.1 The Large Hadron Collider

The LHC apparatus [59–61] is installed in the original cavity of LEP. The tunnel lies 100 m underground and is 27 km long. The design energy of the accelerated protons is 7 TeV, allowing for an available energy in center of mass frame (e.c.m) of 14 TeV. In 2010 and 2011, an energy of 3.5 TeV per beam, corresponding to 7 TeV in center of mass frame was achieved.

Center of mass energy $\sqrt{s}$	14 TeV
Instantaneous Luminosity	10^{34} cm^{-2} s^{-1}
Bunch crossing	25 ns
Number of bunches	2808
Number of protons per bunch	1.15×10^{11}
Beam Lifetime	15 hrs
Transverse Beam size	167 μm
Dipole Magnet Field	8.3 T

Table 4.1: List of the main LHC design parameters

The protons are first created out of hydrogen that gets ionized in a strong electric field. The initial protons are accelerated up to an energy of 50 MeV in the Linear Accelerator (LINAC) then injected into the Proton-Synchrotron-Booster (PSB) where they reach an energy of 1.4 GeV. Protons are then inserted into the the Proton Synchrotron (PS) where the beam gets subdivided into small bunches and further accelerated up to 25 GeV. The bunches are then injected into the Super-Proton-Synchrotron (SPS) where they reach an energy of 450 GeV (see Figure 4.2). Finally the 450 GeV proton bunches are inserted in two locations of the LHC where two separate beams will circulate in opposite directions and accelerated up to the nominal energy. Once the nominal energy is reached the beam will circulate and collide until the beam are believed to be deteriorated (i.e when the instantaneous luminosity has decreased substantially). At this point the beams are dumped (see Figure 4.1) and the whole process starts again with a new filling of proton bunches.

4.1.1 Parameters

The Magnets

Due to their high mass, protons have the property to lose a little amount of energy due to Bremsstrahlung radiation[1].

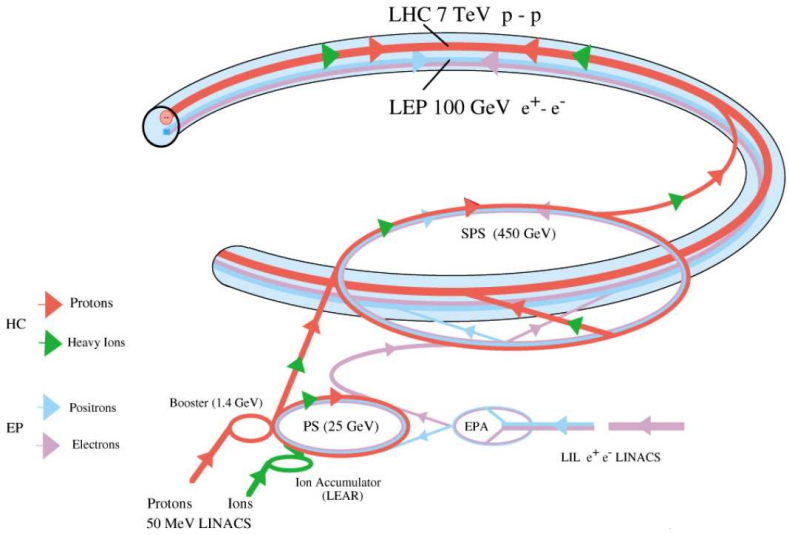


Figure 4.2: The LHC complex compared to the LEP complex

However, the higher the mass of a particle, the higher the magnetic field magnitude has to be in order to bend the trajectory. The limiting factor towards reaching high energy circular proton beams is therefore in being able to produce a strong enough magnetic field the keeps TeV protons in the trajectory constrained by the LHC tunnel.

The LHC uses a total of 9600 magnets. The 1232 dipole magnets are cooled down to 1.9 K with super-fluid helium. At such a low temperature the dipoles being in a super-conducting state, can provide a magnetic field of 8.3 T. The rest of the magnets are quadrupoles and octopoles. They are used mostly for the control of the protons trajectory and the focusing of the beams at the interaction points. A closer look at a transverse view of the beam pipe with the magnets is shown in Figure 4.3.

Instantaneous Luminosity

Experiments such as CMS or ATLAS aim at studying extremely rare processes. In order to accumulate a statistically significant number of such events, the LHC instantaneous luminosity has to be important (see Equation (2.3)). The instantaneous luminosity characterizing two identical colliding beams can be generally written as[2]:

[1] In a circular motion, the amount of radiation that an accelerated charged particle loses per unit time goes $P \sim m^{-4}$. In particular, this implies that a proton circular beam loses $\left(\frac{m_e}{m_p}\right)^4 \sim 10^{-13}$ less energy per unit time that an electron does. This is also the reason why it is impossible to reach the TeV scale with an electron-positron circular collider of the same size as the LHC.

[2] The integrated luminosity is given by $\mathcal{L} = \int L dt$ where L is the instantaneous luminosity.

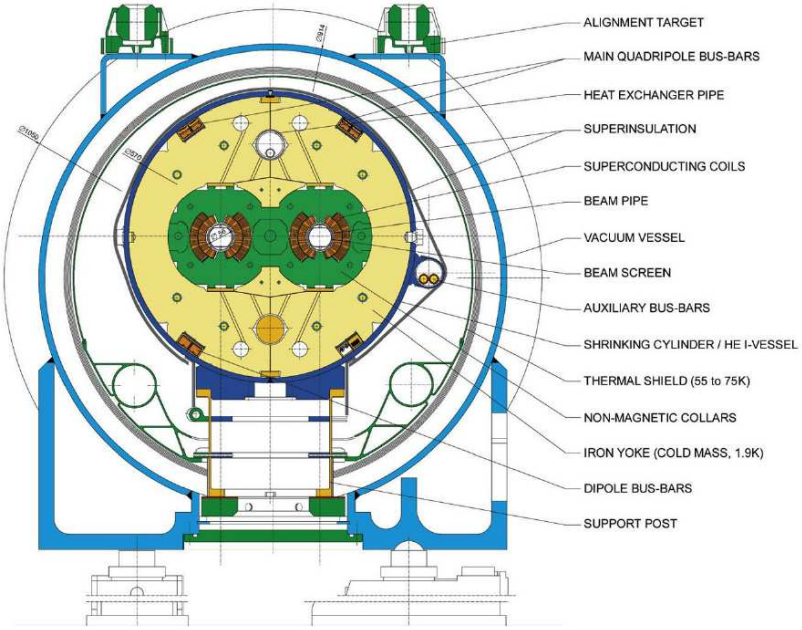


Figure 4.3: A transverse view of the beam-pipe and the magnets together with the holding cooling vessel.

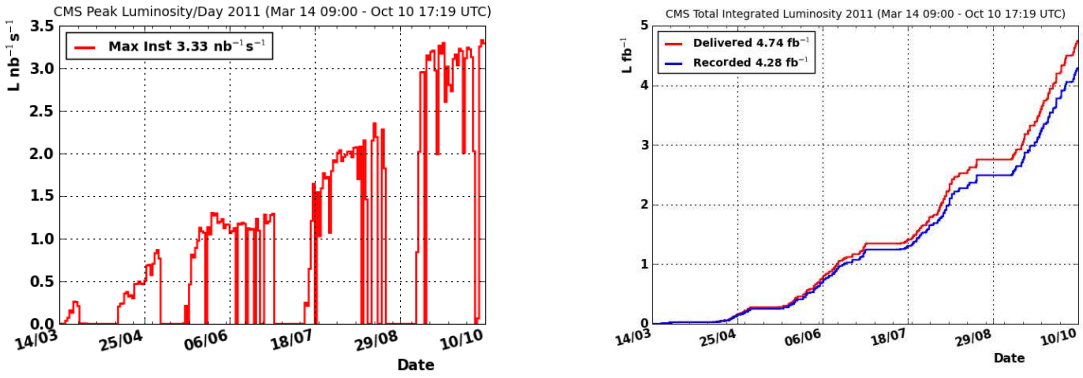
$$L = f n_b \frac{N_b^2}{A} \tag{4.1}$$

where f is the revolution frequency (fixed by the LHC length and the speed of light), n_b is the number of bunches of protons, N_b is the number of protons per bunch, and A is the effective overlap area of the beams (defined in the direction transverse to beam). The design luminosity of the LHC, $L = 10^{34} cm^{-2} s^{-1}$ is achieved by subdividing the beam in $n_b = 2808$ bunches with $N_b = 1.15 \times 10^{11}$ protons per bunch. Some of the parameters that characterize the LHC apparatus are summarized in Table 4.1.

At such a high luminosity, it is estimated that ~ 25 interactions between protons will occur during a bunch crossing. These *pile-up* events, corresponding mostly to low transverse momentum inelastic interactions, will overlap with the interesting high momentum hard scattering events. A high resolution detector is essential to separate these two very different type of events.

4.1.2 The LHC performance in 2011

The LHC operations started in September 2008. After a week of circulating beams a major incident occurred, causing serious damage to the machine that needed several months of stop in order to repair the damage. Due to this incident, LHC started again operations one year later, in November 2009, at half the nominal energy. Since then, the instantaneous luminosity has been gradually (and cautiously) increased (see Figure 4.4(a)). Up to today, the LHC program has succeeded in providing to CMS an integrated luminosity of approximately 5 fb^{-1}, most of which was recorded in 2011 with an instantaneous luminosity peaking at 3.3×10^{33} at the end of 2011, with 1000 proton bunches and 50 ns spacing.



(a) (b)

Figure 4.4: (a) Instantaneous luminosity peak value delivered per day. The unfilled periods correspond to technical stops of the machine. (b) Integrated luminosity accumulated until october 2011

The LHC will keep operating until the end of 2012. At least a full year of technical stop is then foreseen. In 2014 the LHC is expected to run close to the design energy ($\sqrt{s}$ =14 TeV).

4.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) [65–67] is a cylindrical shape detector located at the "Point 5" of the LHC beam-pipe (see Figure 4.1). CMS is 21m long and has a diameter of 15 m. The total inelastic proton-proton cross section is 100 $\sim$ mb at the LHC, which corresponds, at the design luminosity, to approximately 10^9 inelastic collisions per second. This means that during a bunch crossing (25 ns), there will be $\sim$ 25 collisions (pile-up), generating $\sim$ 1000 particles. These extreme conditions imply several requirements on the design and the timing response of the CMS detector. CMS also needs to be highly granular and to be able to resist to such a radiation heavy environment.

The topology of CMS is dictated by the type of stable particles that are identifiable: hadrons, electrons, muons[3] and photons. Muons play a special role as they are heavy and therefore their presence might indicate a decay of some heavy resonance. A similar role is played by heavy flavor quarks such as b and c, and τ leptons. Although these particles are unstable (they decay after traveling a small distance), they can still be identified via indirect techniques.

The first layer of the detector that is encountered by the collision products is a highly granular tracker, that allows to reconstruct the trajectory of charged particles without significantly affecting their momenta. A high magnetic field is applied in order to bend these particles and therefore to be able to infer their momenta by studying their trajectory. The second layer is the electromagnetic calorimeter (ECAL) that allows to measure the energy of electrons and photons. The next layer is the Hadronic Calorimeter (HCAL) that

[3]Ultra-relativistic muons can be considered to be stable, due to their Lorentz dilated lifetime

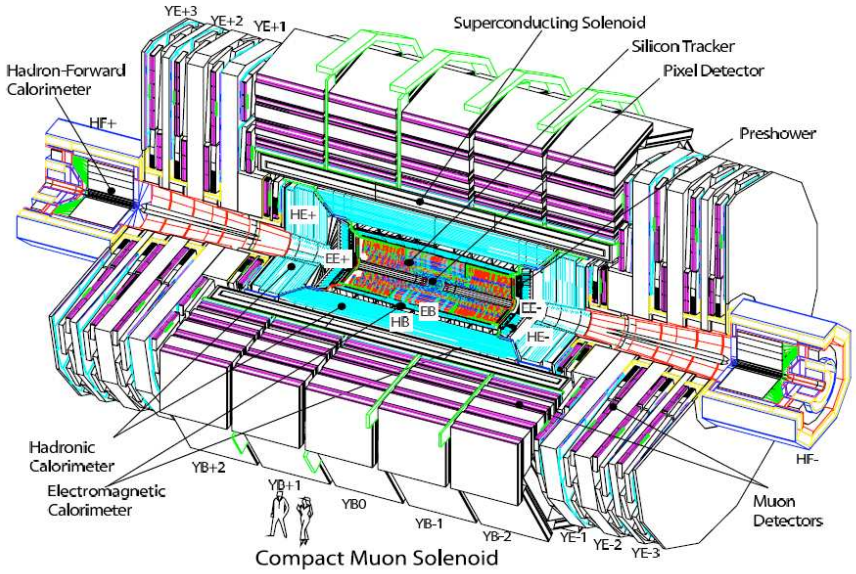


Figure 4.5: Perspective view of the CMS detector

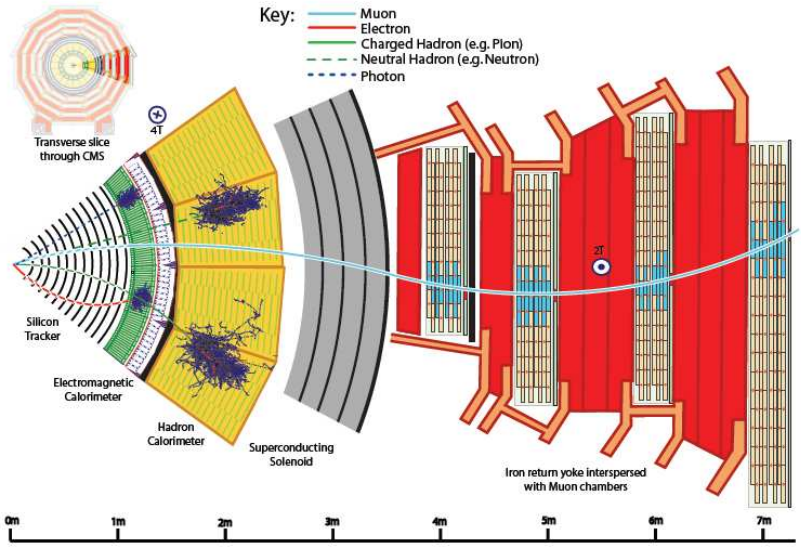


Figure 4.6: Transverse slice of the CMS detector. The particle-detector interactions are also represented

was designed to measure the energy of hadrons (charged pions, kaons, protons ...). Finally, the only particles surviving the passage through the calorimeters are the muons. The last layer is a dedicated muon detector (see Figure 4.5 and 4.6).

Before discussing in detail each sub-detector, we review here the coordinate convention in CMS. The z-axis is in the beam direction and the x,y direction are the transverse coordinates. The transverse momentum and pseudo-rapidity definition are the same given

in Chapter 3. There are particles, such as neutrinos that remain undetected (due to their weakly interacting nature). Although the proton-proton 3-momentum is conserved in a collision, most of the proton remnants are lost in the beam line. Therefore one cannot indirectly infer the 3-momentum sum of those particles that do not leave a signal in detector. However in the transverse plane the total 3-momentum is conserved and equal to 0. The vectorial momentum imbalance observed in the transverse plane has therefore to be equal to the total transverse momentum of the undetected particles. This quantity is called the *Missing Transverse Energy* ($\overrightarrow{E_T^{miss}}$)[4].

4.2.1 The Magnet

The magnetic field bends the charged particles coming from the interaction point, allowing to measure their momenta in the tracker and the muon chambers. It should be able to significantly deviate electrons and muons with energies as high as 1 TeV. In order to achieve a good momentum resolution ($\frac{\Delta p}{p} \sim 10\%$ for p = 1 TeV) the CMS collaboration has chosen to operate with a field B = 4T (now operating at 3.8T) [69].

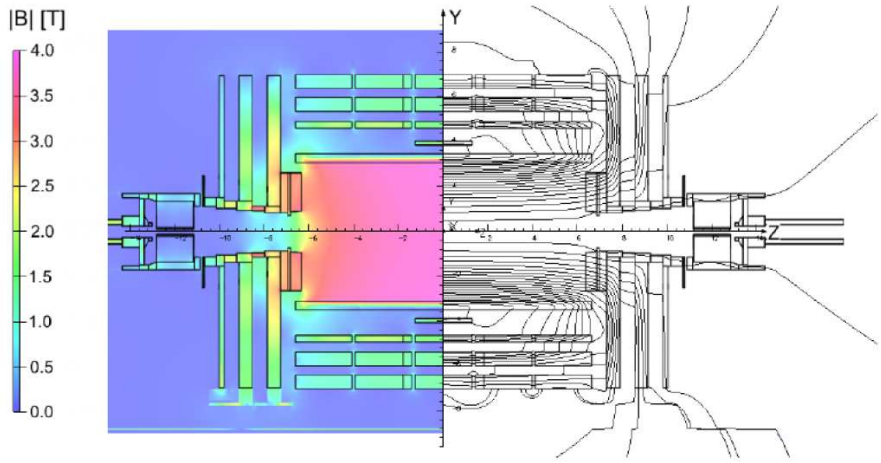


Figure 4.7: Norm of the Magnetic Field inside CMS.

The solenoid is located between the calorimeters and the Muon Chambers so that the field lines are not constant within CMS. Outside the solenoid volume the field direction is inverted. This location allows to have a strong and constant field inside the tracker and the calorimeter volume (the magnetic field reaches its maximal value inside the solenoid and, as shown in Figure 4.7, B = 2T in the muons volume). Also, an optimal length to diameter ratio was chosen in order to achieve a good momentum resolution in the forward region.

The solenoid is made of NbTi material, cooled down to 4.5K by a cryostat filled with liquid helium. Such a low temperature is necessary to achieve a super-conducting state. In Table 4.2 we give a list of the main features of the CMS solenoid.

[4]We will often refer to the Missing Transverse Energy as $|\overrightarrow{E_T^{miss}}| = E_T^{miss}$

Field strength	4 T
Inner Bore	5.9 m
Length	12.9 m
N_{turns}	2168
Current	19.5 kA
Stored Energy	2.7 GJ

Table 4.2: List of the CMS solenoid parameters

4.2.2 The Tracking System

The Tracker [70] is the sub-detector that lies the closest to the beam line. Its main purposes are to measure with high accuracy the momentum of charged particles and, via appropriate algorithms to determine the position of the interaction vertices (corresponding to the hard scattering and the pile-up vertices) as well as identifying the so-called "secondary" vertices. Secondary vertices are displaced vertices (with respect to the beam line). They originate from the decay of long lived particles such as b-quarks, K^0_S, Λ, Ξ.

The tracker is a silicon detector. The basic component of such a detector is a made of a bulk of an n-type semi-conductor, on top of which is placed an array of p-type pixels, or strips. When a charged particle passes through the bulk, it create a particle-hole pairs. The resulting charges migrate towards the pixel/strip with a certain angle (due to the presence of the magnetic field) and are read via a system of electronic readouts.

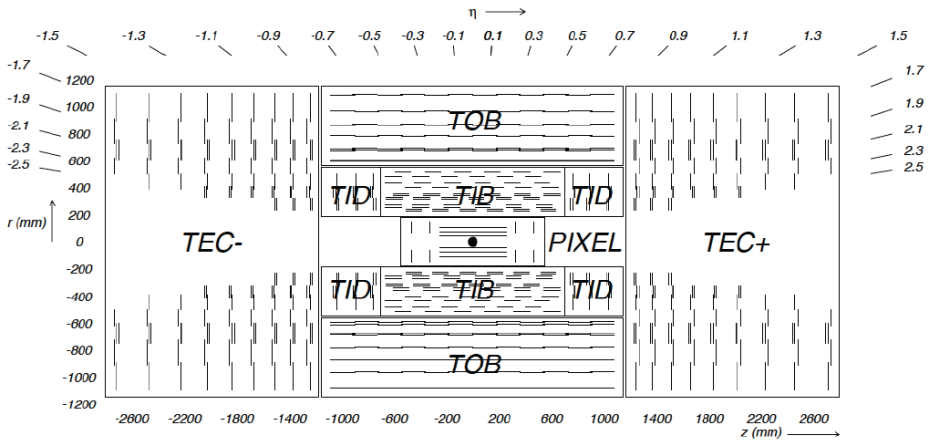


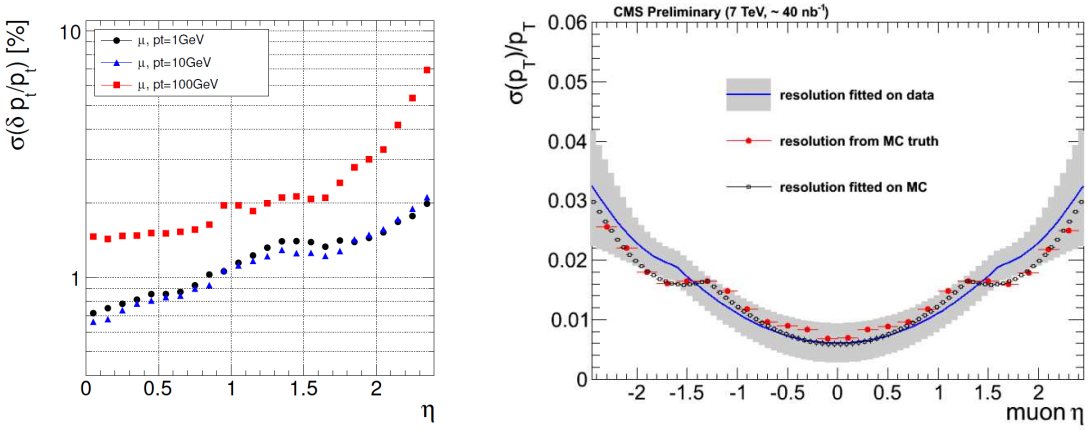
Figure 4.8: Longitudinal cross section of the CMS tracker. Single lines represent simple module, double lines represent stereo modules

Due to the high number of particles produced per bunch crossing, the tracker was designed with a high granularity. It consists of an inner part, the silicon pixel detector and an outer part, the silicon strip detector. It covers a total pseudo-rapidity range $|\eta| \leq 2.5$. A longitudinal view of the tracker volume is given in Figure 4.8.

- The pixel detector consists of three layers (the Barrel), located at the respective radii 4.4, 7.3 and 10.2 cm, and two disks that "close" the cylinder (the End-caps).

This corresponds to a total of 64 million 100×150 μm^2 pixels. The Pixel detector has therefore an extremely high resolution in the $r \times \phi$ (10 μm) plane and in the z-direction (20 μm) plays therefore a key role in the determination of the vertices position.

- The second component of the tracking system is the silicon strip detector. It consists of four subsystems: the Tracker Inner Barrel (TIB), Outer Barrel (TOB), the Inner Disk (TID) and the Outer Disks (TEC). The TIB has 4 layers and covers a radius from 20 to 55 cm and a longitudinal distance $|z| \leq 65$ cm. The TOB consists instead of 6 layers covering $|z| \leq 118$ cm and radii between 55 cm and 116 cm. In the TIB and TOB the first two layers are filled with *stereo modules* (meaning that each module consists in two series of silicon strips pointing in a different direction in order to increase the $r \times \phi$ resolution) opposed to the other layers that are made of simple modules. The overall achieved resolution is $\sim$ 30 (40) μm in the $r \times \phi$ plane and $\sim$ 230 (300) μm in the z direction for the TIB(TOB). The TID and TEC are rings centered on the z-axis. The TID(TEC) is composed of 3(9) rings. The first two layers of TID and layers 1,2 and 5 of TEC have stereo modules. The resolution is comparable to that of TIB and TOB.



(a) (b)

Figure 4.9: Resolution achieved with the tracker on muons transverse momentum in Monte Carlo simulation (a) and Data (b)

In Figure 4.9 the tracking momentum resolution is shown: in Monte Carlo simulation (a) for different average values of the muon transverse momentum and in data (b) as a function of the pseudo-rapidity. The p_T resolution is approximately $\sim$ 1% in the central region of the detector and $\sim$ 5% in the forward region (depending on the muon p_T) [71].

4.2.3 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [72] surrounds the tracking volume. Its purpose is to measure the energy of particle that interact electro-magnetically, mainly electrons (and positrons) and photons.

The principle is the following: at high energy ($\geq$ 100 MeV) electrons lose their energy primarily via bremsstrahlung emission of photons. In matter, the emitted high energy photons convert into electron-positron pairs that, in turn, will radiate photons, etc .... This process, known as electromagnetic shower, will continue until the energy of the e^+e^- pair becomes smaller than some critical energy E_c (defined as the energy regime at which bremsstrahlung emission is no longer the predominant mechanism for electrons energy loss). At energies $E \leq E_c$ electrons mainly lose their energy via inelastic collisions with the electrons present in the atoms of the material. If the material is chosen in an appropriate way[5], it is possible to *contain* the whole shower inside a small volume. The impact of particles resulting from the shower with the material's atom produces an amount of light that can be related to the total energy loss of the initial electron/photon.

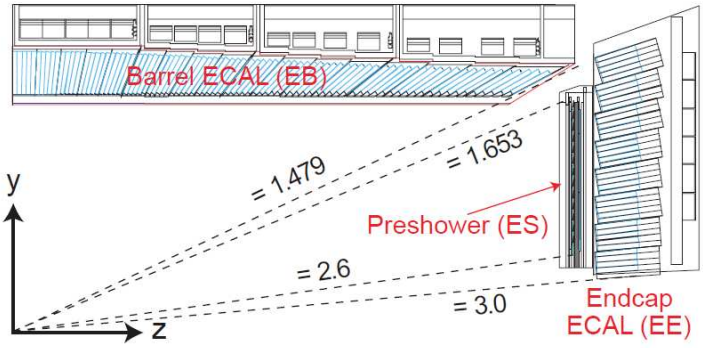


Figure 4.10: Longitudinal cross section of the CMS Electromagnetic Calorimeter

The CMS ECAL scintillator is the $PbWO_4$ crystal. The small radiation length (89 mm) and Moliere Radius (21.9 mm) allowed to build a compact, highly granular and fast response calorimeter able to collect the majority of the incoming photon/electron initial energy. The ECAL barrel covers a pseudo-rapidity $|\eta| \leq 1.479$. It consists of 61200 $PbWO_4$ crystals with dimensions 22×22×240 mm. The scintillated light is collected by Avalanche Photo-diodes (APDs) in the Barrel and Vacuum Photo-diodes (VPDs) in the End-caps. Due to the high rate of neutral pion production in the forward region[6], a lead preshower detector is installed in order to be able to discriminate the pion decay from the prompt photons[7] as the ECAL granularity is not fine enough in order to resolve two close-by photons. The preshower detector also helps in discriminating electrons from Minimum Ionizing Particles (MIP)[8] and helps in determining their position. A longitudinal cross-section of the ECAL detector is shown in Figure 4.10.

The energy resolution of the ECAL detector can be expressed as:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{4.2}$$

[5] That is, by choosing a material with the right radiation length (average distance over which the electron loses all but $\frac{1}{e}$ of its energy) and Moliere radius (transverse extension of the shower)

[6] Boosted pions mainly decay to two collinear, almost unresolvable photons

[7] We refer to prompt particles when they originate as the final state of the hard scattering

[8] A particle is in the MIP regime if the energy it loses by ionizing matter is minimum. Most of the stable particles (with the exception of the electron) are MIP at a high enough energy

where S is the stochastic term (photo-statistics and event by event fluctuations of the shower), N is the noise term (electronics, pile-up) and C is the constant term.

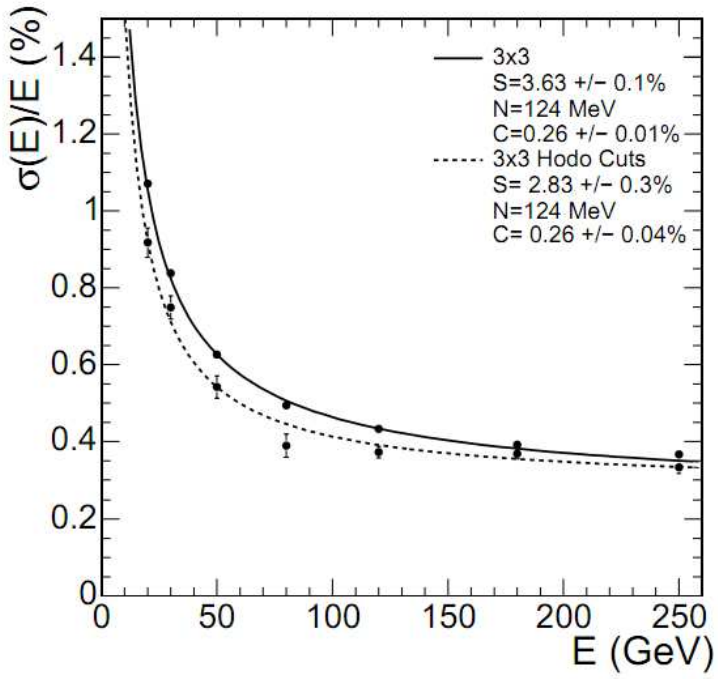


Figure 4.11: Energy resolution as a function of the electron energy obtained from a test beam. The energy here is the total energy deposited in 3×3 array of ECAL crystals.

At high energies, only the constant contribution is relevant. The constant term includes errors from the inter-calibration, energy leakages from the back of the crystals and non-uniformity of the light collection in the longitudinal direction. During test beams the CMS ECAL resolution has been found to be [72] (see Figure 4.11):

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2 \tag{4.3}$$

4.2.4 The Hadronic Calorimeter

A Hadronic Calorimeter detects and measures the energy of long lived hadrons. The interaction of hadrons with nuclei produces a hadronic shower. Typically hadronic showers are much larger than the electromagnetic ones, thus the HCAL occupies a much larger volume than ECAL.

The large pseudo-rapidity coverage $|\eta| \leq 5.2$ allows to measure the energy and the direction of hadrons produced during the collision over a wide geometrical acceptance [73]. Morevover the HCAL provides a measure of the E_T^{miss}, since hadrons contribute, in proton-proton collisions, to a significant fraction of the total produced particles, and thus of the total deposited energy.

HCAL has four subcomponents: the Barrel (HB), the EndCaps (HE), the outer barrel (HO), and the forward calorimeter (HF) as shown in Figure 4.12.

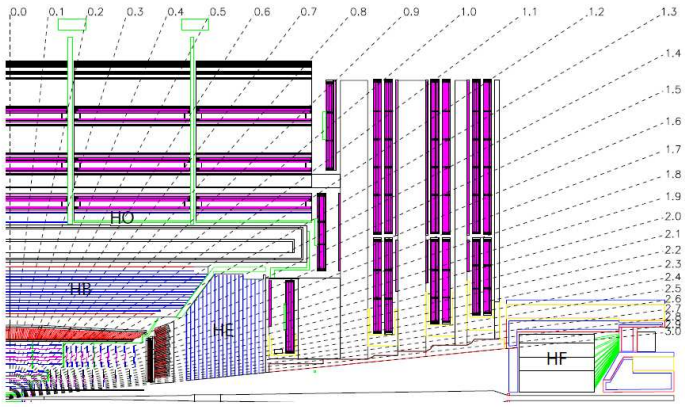


Figure 4.12: Longitudinal cross section of the CMS Hadronic Calorimeter.

The HCAL barrel covers a pseudorapidity range $|\eta| \leq 1.3$ and a radial extension 1.71 m $\leq$ R $\leq$ 2.95 m. HCAL consists of alternating layers of 5 cm brass absorber (70% Cu + 30% Zn) and 4 mm plastic scintillator. The readout is made of optical fibers and hybrid photo-detectors. In order to ensure enough penetration depth, the outer calorimeter (HO) is placed after the magnet (that also serves as a additional absorber). The end-caps (HE) cover the region $1.3 \leq |\eta| \leq 3$. The absorber used in (HE) is different due to the high radiation and high particles rates: C260000 cartridge brass was used. Finally we have HF, in the forward region $|\eta| \leq 5.2$, the material must survive to even higher level of radiation. Therefore the choice was made to use a Cherenkov detector with quartz fibers.

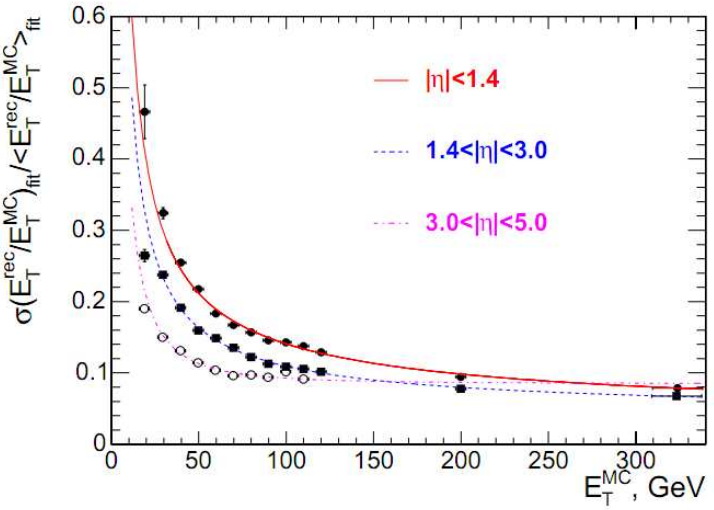


Figure 4.13: Jet transverse energy resolution as a function of the simulated jet transverse energy in different pseudo-rapidity ranges

The jet energy resolution of the HCAL in different pseudo-rapidity range is shown in Figure 4.13.

4.2.5 The Muon Detector

As the CMS acronym indicates, a high focus was put on muon detection while designing CMS. These particle play indeed a special role in Higgs and new physics searches (as it is the first heavy stable particles that heavy resonances will decay into). At high energies, muons are *minimum ionizing particles* (MIPs), that is, they lose very little of their initial energy by traversing matter. Muons are, with neutrinos, the only (known up to date) particles capable of reaching the last layer of the CMS detector: the Muon Chambers [74].

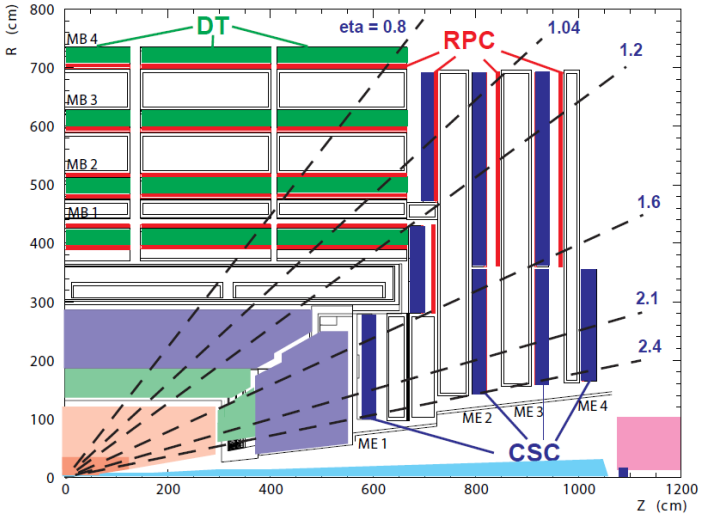


Figure 4.14: Longitudinal cross section of the CMS Muon Detector.

A muon chamber consists of a volume filled with a gas and a conducting wire. The volume delimiter is also conducting so that when a muon passes through volume, the gas gets ionized, and the charges that are released drift towards the wire (an electric field is applied). The induced electric signal is then recorded by appropriate electronic readouts.

The muon system in CMS consists of three types of gaseous chambers (see Figure 4.14):

- The Drift Tubes (DTs) chambers covers a pseudo-rapidity range $|\eta| \leq 1.2$. The chambers have a cross section of 4.2 $\times$ 1.3 cm and a length of 2.5 m. They are filled with $Ar/C0_2$ gas. The container is aluminum based and a stainless steel wire acts as the anode. The DTs are particularly suited for the central part of detector, where muon rates are small, the magnetic field is uniform and the neutron-induced background is small.

- In the End-caps where rates are higher and the magnetic field is non-uniform and large it was chosen to use Cathode Strip Chambers (CSCs). Each CSC has a trapezoidal shape and consists of six cathode strips alternated with gas gaps.

- The Resistive Plate Chambers (RPCs) are placed both in the Barrel and in the End-caps. They are used to help in matching the muon tracks with the beam crossings and to complement the Level 1 Trigger (see next section).

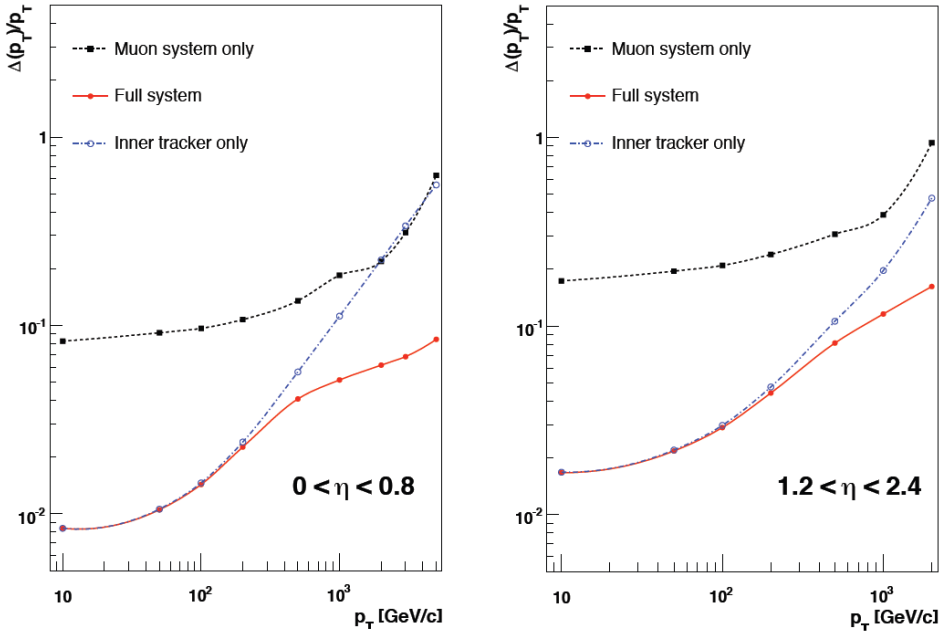


Figure 4.15: Muon transverse momentum resolution as a function of its transverse momentum using only the muon system, only the tracker system, and both in the central region of the detector (left) and in the forward region (right).

The DTs and CSCs provide a similar spatial resolution of $\sim$ 100 μm. Used in combination with the tracker information, they provide a precise measurements of the muon trajectories and of their momenta. As an illustration Figure 4.15 shows the muon transverse momentum resolution with and without the use of the muon chambers. The exploitation of the muon chambers allows to significantly decrease the uncertainty on the muons momenta at high energies, where the curvature induced by the magnetic field on the muons trajectory is small.

4.2.6 The Trigger System

Under nominal conditions the LHC delivers bunch crossings at a rate of 40 MHz. Given the huge amount of information in one single event it is impossible to store on tape every event.

A dedicated trigger system has therefore been designed in order to keep the data acquisition rate at a sustainable level. Storing only "interesting events" can be achieved by a fast decision making system in situ. CMS uses a two level trigger system: the Level 1 Trigger (L1), and the High Level Trigger (HLT).

L1 Trigger The L1 trigger can achieve an event rate reduction down to $\mathcal{O}$(100 KHz). This reduction rate is performed in three basic steps:

- At the local level, the Trigger Primitive Generators (TPGs) use information from the Calorimeters (energy deposits) and Muon system (track segments in the muon chambers) to form *primitive objects*.

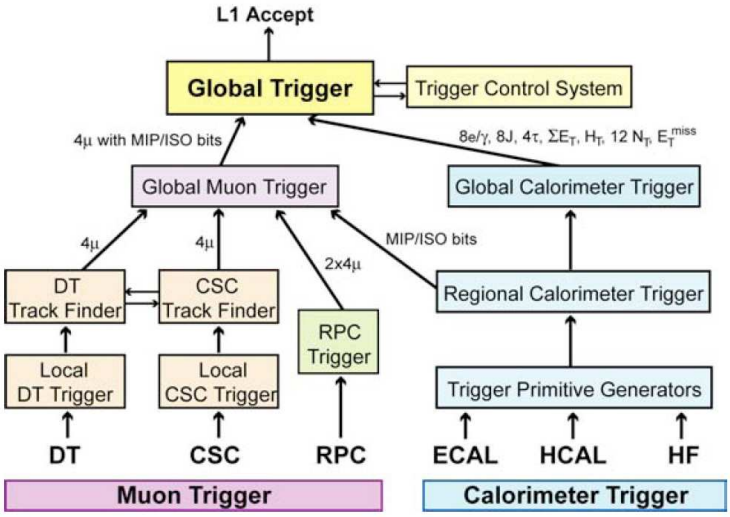


Figure 4.16: Architecture of the Level-1 Trigger

- Regional Triggers (RTs) then combine the Trigger Primitives information to form particle candidates (such as photons, muons, electrons) with a given rank based on the quality of the deposits.
- Finally the Global Calorimeter/Muon Triggers combine the various regional triggers outputs and determine the highest ranks. The event is kept or rejected if the particles fulfill some characteristic requirements and if certain general detector criteria are properly satisfied.

If the event is accepted, the information is communicated to the various sub-detectors via the Timing Trigger and Control System (TTC). A final L1 decision takes approximately 3 μs to be taken. During this decision making the data is temporarily written in memory buffers located in the front-end electronics and collected centrally if the event is accepted by the L1 trigger.

The High Level Trigger The High Level Trigger (HLT) is basically a farm of CPUs performing algorithmic operations aimed at checking particular requirements on one or more particle candidates in order to select interesting events for physics analysis. The HLT decision is subdivided in two steps:

- Muons and Calorimeter information is used is order to construct more elaborated objects than L1.
- Tracker information (including heavy flavor identification) is included in order to make use of the full detector information

The results are stored in Primary Datasets (PDs) by making sure that little or no overlap occurs among various PDs. With this system, CMS is able to reduce the initial event rate by a factor 10^5. The final rate after the HLT is $\mathcal{O}(100\ Hz)$.

4.2.7 Simulation Chain and the CMS Software

A proper simulation of the interactions of the collision products through the CMS detector is crucial in order to be able to interpret the data and to test detector response. Also, a unified framework capable of handling all the aspects of such a multi-purpose experiment in a user friendly environment is necessary. We will review the basic steps that characterize event simulation and stress the major capabilities of the CMS software.

Event simulation

We have seen in the previous Chapter the importance of being able to produce kinematic distributions characterizing a given process (such as Gluon Fusion Higgs production) and moreover, to have at our disposal an event-by-event description of such processes. In this context, Monte Carlo programs are able to produce an *exclusive* description of the final state in terms of partons (quarks and gluons) and leptons. When a hadronization algorithm is added to such a procedure, we have then a full description of the final state in terms of stable colorless hadrons (and stable leptons and photons). The outcome of this event generation can then be exploited by a program that simulates the passage of such particles through the CMS detector.

The GEANT4 [75, 76] (GEometry ANd Tracking) packages serves this task. Based on their 4-momenta, particles are traced through a simulated version of the CMS detector. The passage of the particles through matter is properly modeled. The result of the particles interaction with the CMS material is recorded as a *simulated hits*. The GEANT toolkit is also able to generate so called *secondary particles*, that originate from the interaction of *primary* or *prompt* particles with the detector material (such as electron/positron conversion into photons). The simulated hits are then used as inputs by *emulators*, that simulate the digitization of the deposits in the various sub-detectors, including the trigger decisions.

Computing in CMS

The CMS software (CMSSW) [77] is an object oriented framework based on the programing languages C++ and Python. The data formats are structured in "trees" that contain most of the event information that is needed and they can be analyzed with the ROOT package [78].

The CMSSW code covers a wide variety of operations within CMS. It is used for the Monte Carlo event simulation (i.e all the chain described in the previous paragraph) for objects reconstruction (see next Chapter) and for the data analysis. In addition it is also used to organize the storage and distribution of the data. CMSSW is being constantly updated in new releases in order to account for changes in the calibration constants of a particular subdetector, or the definition of a new relevant variable.

The LHC is expected to produce ~ 5 petabytes of data per year. In order to cope with such a high data volume the LHC experiments make use of a data storage infrastructure and a distributed computing system called the Worldwide LHC Computing Grid (WLCG). After data is recorded, calibration and express stream analysis are performed in the CERN Analysis Facility (CAF). The data is then distributed into the *Tier 0*, located at CERN, then to the *Tier 1* centers (in seven countries) and finally to the *Tier 2* centers. This flow of data allows the various institutions around the world to perform data analysis locally in their nearest Tier 2 center.

Part II

Analysis and Results

Chapter 5

The $H \rightarrow W^+W^- \rightarrow \ell^+ \nu \ell^- \bar{\nu}$ signal and its main backgrounds

As already introduced in Chapter 2 the decay modes of the Standard Model Higgs are strongly dependent on its mass. Indirect constraints from the electroweak global fit point to a low Higgs mass [23]. Although a Higgs mass $m_H = 155 - 165$ GeV has already been excluded at 95 % confidence level by the Tevatron [22], there is still a challenging opportunity for discovery (or exclusion) in the decay channel $H \rightarrow W^+W^-$ in the mass range $m_H = 120 - 200$ GeV because in this region $m_H \approx 2m_W$ and the branching ratio (BR) of $H \rightarrow W^+W^-$ decay is close to 1. Therefore this channel is of particular importance in the first years of running of the LHC.

The W boson can decay into two quarks or into a lepton and a neutrino [1] (e ν_e, μ ν_μ and τ ν_τ), leading to three main possibilites for the $H \rightarrow W^+W^-$ final state:

- fully hadronic: ≥ 4 jets in the final state
- semi-hadronic(leptonic): ≥ 2 jets, 1 lepton and missing energy in the final state
- fully leptonic: ≥ 0 jets, 2 leptons in the final state and large missing energy in the final state.

Although the cross-section of the fully hadronic and semi-leptonic decays are higher than that of the fully leptonic decay, these searches are very challenging at the LHC due to the high production rate of QCD processes that constitute a background for these decay modes. On the other hand the fully leptonic search presents a much cleaner final state since the backgrounds in this case are mainly electroweak processes with a much lower cross section. The present work will study exclusively the fully leptonic final state.

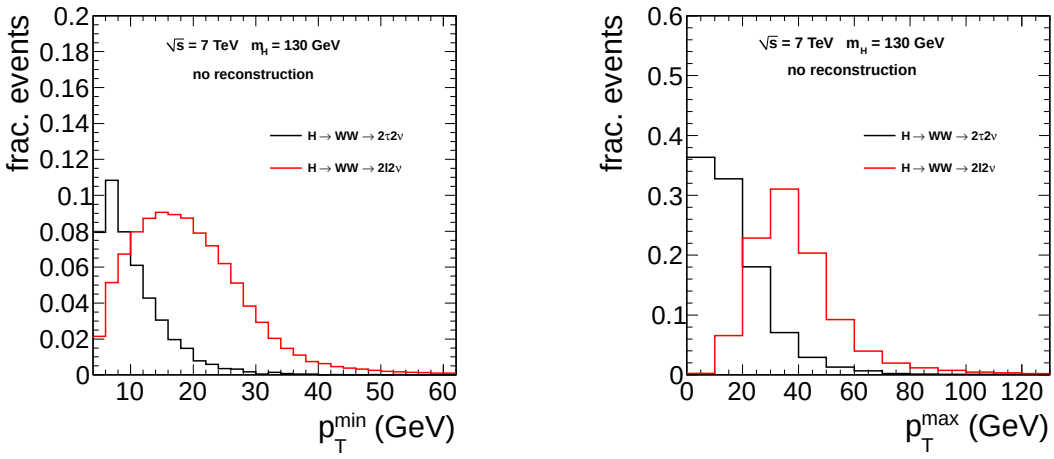
First, a short introduction on the signal and the major backgrounds is given. We will also briefly introduce the most significant signatures discriminating between the signal and the backgrounds. We then discuss at higher detail the signal and the major irreducible background W^+W^- topology. We will see how the events topology can be inferred from simple qualitative arguments.

[1] In the following we use the convention that a lepton refers to a massive lepton. The notation "l" refers to a stable lepton e or μ unless stated otherwise.

5.1 Signal and Backgrounds

5.1.1 The Signal

We have already discussed in Chapter II the main production mechanisms and decays of the Higgs particle at the LHC. We recall that at the LHC the Higgs production originates from gluon gluon fusion (ggH), vector boson fusion (VBF), associated Production with Z/W and top pair. We recall that the ggH mechanism is predominant at the LHC (approximately 95%). When $m_H \sim 160$ GeV $\sim 2M_W$ the Higgs is produced on-shell and the W$^+$W$^-$ branching ratio is one, meaning the Higgs boson exclusively decays via this channel. If $m_H < 160$ GeV the Higgs is produced off-shell, and this decay is gradually suppressed as $m_{\rm H}$ becomes smaller. However if the Higgs mass is in the range $m_H = 120 - 200$ GeV this channel is still highly explopitable (see Figure 2.2).



(a) (b)

Figure 5.1: Comparison of the distribution of the tranverse momenta of the leptons originating from the H→W$^+$W$^-$ decay in the case where the W$^+$W$^-$ system decays directly to stable leptons e or μ (red curve) or to tau leptons that decay leptonically (black curve). The tranverse momentum of the less (most) energetic stable lepton is shown in 'a' ('b').

When both the W bosons decay leptonically, the signal will be characterized by the presence of two leptons with opposite charge and a relatively high transverse momentum (depending on the Higgs mass), large missing transverse energy due to the neutrinos which can escape detection. There are then 9 equally likely possibilities involving every possible lepton pair (ee, eμ, eτ, $\mu\mu$, μe, $\mu\tau$, $\tau\tau$, $\tau\mu$, τe).

Unlike electrons and muons, τ leptons are unstable and cannot be detected directly. They decay electroweakly into a W boson that in turn decays into hadrons ($\sim 65\%$) or stable leptons ($\sim 35\%$). In the former case this leads to a final state different from the "two leptons and missing transverse energy" signature. In the latter case the final state is apparently similar to that of a W$^+$W$^-$ → eμ, ee, $\mu\mu$ decay. However the stable lepton originating from a τ decay has a lower momentum than that of a lepton originating directly

from a W boson (since the available phase space for such a decay is reduced) as is shown in Figure 5.1.

A low fraction of these events will also present one or more jets in the final state, when gluons are radiated off the initial state (see Chapter 3). The number of jets in the final state (and their respective momenta) is also a variable that can be exploited. As we will see in greater detail later, these variables will help in discriminating between the signal and the backgrounds.

5.1.2 The Backgrounds

We will briefly introduce here the main processes that have the potential to mimic the signature of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ decay. These processes involve the presence of two well identified leptons plus contribution of E_T^{miss} (and eventually one or more jets in the final state). They are the W^+W^- irreducible background, $t\bar{t}$ pair production, single top production and Drell-Yan and W+jets plus other minor contributions.

Process	$\sigma\times$ BR (pb)
$gg \to H \to WW \to 2\ell 2\nu$ (m_H = 130 GeV)	0.45
$gg \to H \to WW \to 2\ell 2\nu$ (m_H = 160 GeV)	0.87
$gg \to H \to WW \to 2\ell 2\nu$ (m_H = 190 GeV)	0.49
$W \to \ell\ \nu$ + jets	31314
$qq \to WW \to 2\ell 2\nu$	4.51
$t\bar{t} \to 2\ell 2\nu 2b$	157.5
$Z \to \ell\ell$ + jets ($m_{\ell\ell} > 20\ GeV$)	4998
$tW \to 2\ell\nu b$	10.9

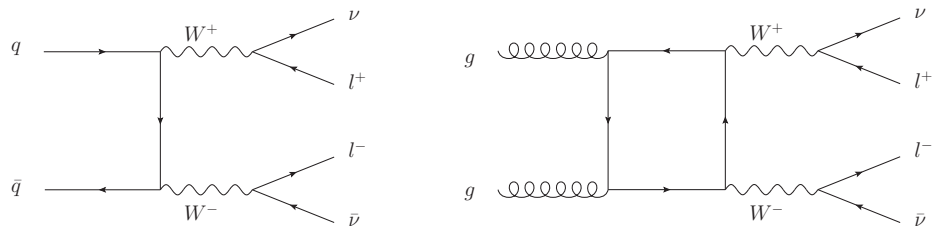
Table 5.1: Most significant backgrounds cross-sections at NLO times Branching ratio at the center of mass energy $\sqrt{s} = 7$ TeV in pb compared to the signal production. Here l = e, μ, τ(from [15]).

The Table 5.1 summarizes the magnitudes of these processes at the center of mass energies $\sqrt{s} = 7$ TeV. The Higgs gluon fusion cross section times branching ratio is also shown as a reference for the three mass points $m_H = 130, 160, 190$ GeV. With 1 fb^{-1} of data taking (approximately 1 year of early LHC running) we expect ~ 450 Higgs bosons decaying into leptons compared to 30 $\times 10^6$ W+jets, 5×10^6 Drell-Yan, 150 $\times 10^3$ $t\bar{t}$ pairs and 4×10^3 W^+W^- events. Discovering a small signal contribution hidden in such an overwhelming background necessitates to define kinematical variables that behave differently between the signal and the various backgrounds, and exploit these differences by selecting particular phase-space region in which the ratio signal over background (S/B) is enhanced.

WW continuum

A W^+W^- pair can be produced by a $q\bar{q}$ scattering or by gluon-gluon fusion. Figure 5.2(a) shows one of the leading order diagrams for this process. The gluon induced process involves a light quark loop and is therefore highly suppressed and is shown in 5.2(b) . For the qualitative discussion of this chapter we are going to consider only quark induced W pair production as it is the dominant production mode.

The W^+W^- background is an *irreducible* background in the sense that a W pair alone is produced in this process, which makes it almost identical to a $H \to W^+W^-$ decay. We



(a) (b)

Figure 5.2: Example of Feynman diagrams of the diboson W$^+$W$^-$ production at the LHC. In (a) the dominant quark induced production mode is shown. In (b) the sub-dominant gluon induced production, involving at light quark loop, and therefore suppressed.

will see later that the crucial discriminant for its reduction is the opening angle between the leptons due to different spin correlation in the signal and the background. A detailed comparison with the signal will be performed in the next paragraph.

Top production

The t$\bar{\text{t}}$ pair production is a major background for H→W$^+$W$^-$ but unlike the W$^+$W$^-$ continuum it is a reducible background. The top quark decays almost exclusively (98%) to a W boson and a b quark leading, in addition to two leptons and large missing energy, to the presence of two central b-jets (see the corresponding diagrams in Figure 5.3). These central b-jets can be identified with the b-tagging technique (this techniquewill explained in Chapter 6). A veto on the presence of such jets can lead to a considerable reduction of this major background. A top quark can also be produced alone, predominantly via the so-called tW channel (see Figure 5.4). The final state is in this case two leptons plus missing energy and one b-jet. If one requires a low number a reconstructed jets (0 or 1) the yields of top induced background can be sensibly reduced.

Drell-Yan (Z+jets)

The Drell-Yan (DY) [2] process, $Z/\gamma^* \rightarrow \ell\ell$, has a huge cross-section (see Table 5.1). As Figure 5.5 shows, the Z/γ boson decay to two leptons potentially mimics the signal. In principle this background affects only the the same flavor channels H→W$^+$W$^-$ → ee and H→W$^+$W$^-$ → $\mu\mu$, although a $Z/\gamma \rightarrow \tau\tau$ can also affect the mixed flavor channel. This background contribution can be reduced by requiring no significant E_T^{miss} is present in the final state. With high luminosity conditions however, the high amount of simultaneous interactions per bunch-crossing can severely affect the E_T^{miss} resolution having the effect for the DY to display significant missing energy. One can also reduce this background by applying a veto to events that present two opposite sign and same flavor leptons with an invariant mass matching the Z boson mass.

[2] Sometimes the Drell-Yan refers to W or Z production. Hereafter Drell-Yan will be referring exclusively to Z/γ + jets production.

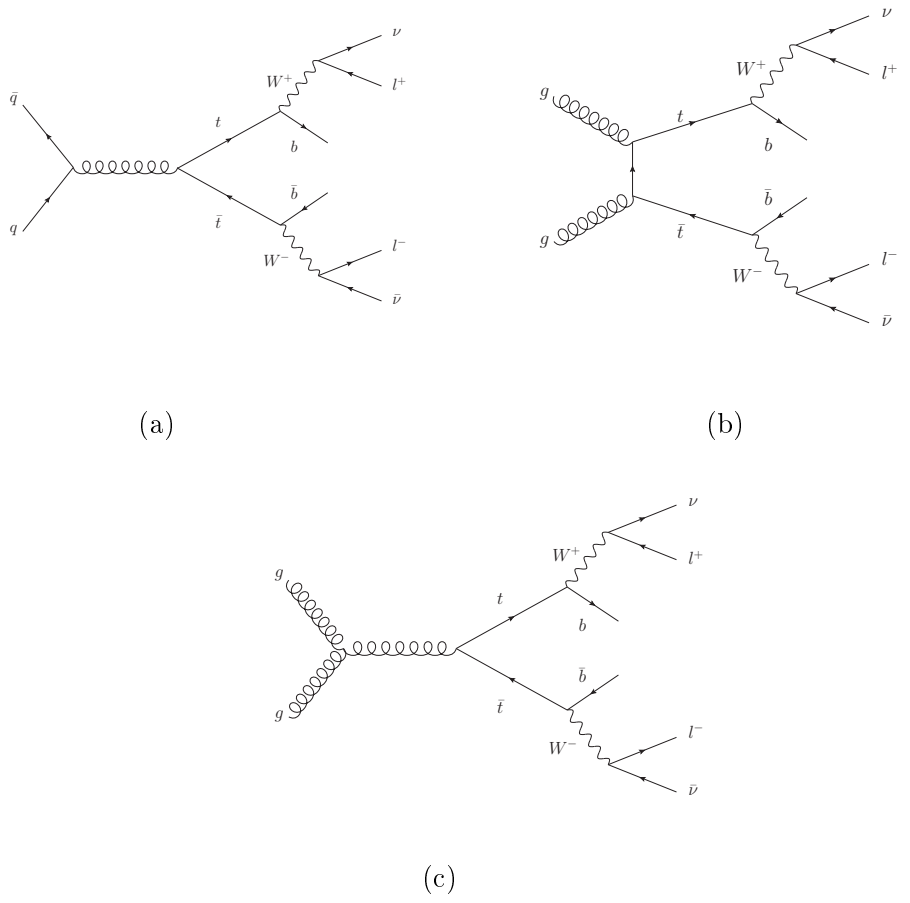


Figure 5.3: (a) Feynman diagrams of $t\bar{t}$ pair production by quark production (a), gluon fusion (b) and three gluon vertex (c) at LO

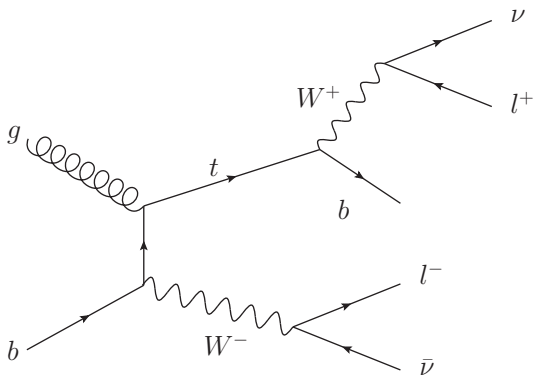


Figure 5.4: Single Top production via tW channel at LO

W+jets

The W+jets process is produced at an extremely high rate at the LHC. When the W decays leptonically the final state consists of one lepton, missing energy plus 0 or more reconstructed jets. Sometimes the jets can be misidentified as leptons and this background can then potentially mimic the 2 leptons plus missing energy final state. The "fake-rate" i.e. the probability that a jet can fake a stable lepton depends on whether the lepton is an electron or a muon (the fake rate is higher for electrons) and on the momentum of the jet (the lower the jet momentum, the higher the fake-rate probability). This background can then be reduced by:

- requiring a minimum threshold on the identified lepton momenta
- applying identification and isolation criteria of on the identified leptons

The first listed requirement is done at the expense of significantly killing the signal in the case of a final state involving one or two τ leptons in the final state. The latter criterium relies on the fact that a jet faking a lepton usually present a high activity in a region that surrounds it. Requiring that a lepton is isolated is asking that little or no high momentum particles surround the lepton.

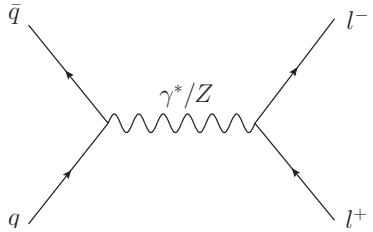


Figure 5.5: Diagram of Drell-Yan production at LO

Other backgrounds

Other backgrounds may mimic the signal such as the diboson production WZ and ZZ, and Wγ. Despite not producing exactly 2 leptons in the final state, the dibosons can fake the signal due to possible misidentification of leptons or loss of the leptons out of the detector acceptance. These backgrounds are essentially reduced by requiring exaclty two opposite sign leptons in the final state. In the Wγ case, the photon might convert into an electron-positron pair by interacting with the detector material and can be reduced by using conversion rejection techniques.

5.2 Topology of the Signal and the Irreducible W$^+$W$^-$ pair background

For simplicity only the signal and the W$^+$W$^-$ background will be considered here. We will first talk about the boost of the W$^+$W$^-$ system and what can be concluded for the signal and background topology. Next, it will be seen how spin correlations arise, and how they can be used as an effective discriminant between the signal and the background. This discussion largely follows [79].

5.2.1 Constraint on the W^+W^- system from parton distribution functions

In Chapter 3, Section 3.1.2 we have established that the product of the momentum fractions x_1 and x_2 of the two partons entering the hard interaction is fixed by the ratio $\frac{\hat{s}}{s}$. We recall that the likelihood of a parton to enter the hard scattering with a characteristic energy scale Q with a momentum fraction x is described by the parton distribution functions $f(x, Q^2)$.

With the constraint on the product $x_1x_2 = \frac{\hat{s}}{s}$ it is possible to infer the most probable values of x_1 and x_2 for a particular process.

For Higgs production at 7 TeV with $m_H = 160$ GeV, we have $\sqrt{\hat{s}} \approx 160$ GeV. :

$$x_1x_2 \approx \frac{(160\text{GeV})^2}{(7\text{TeV})^2} \approx 5.2\ 10^{-4} \tag{5.1}$$

Similarly, for W^+W^- production, $\sqrt{\hat{s}} \approx 2m_W \approx 160$ GeV which gives the same value for x_1x_2 at 7 TeV.

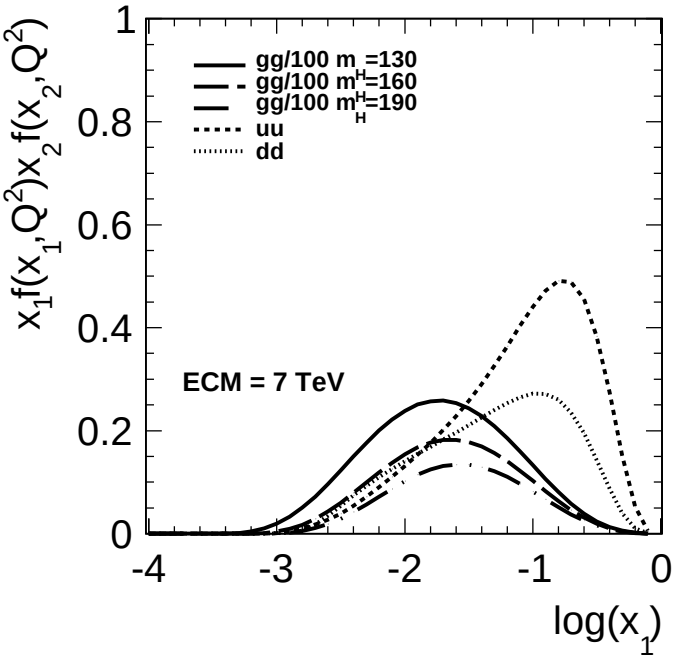


Figure 5.6: Products of PDFs a function of as a function of x_1, the product x_1x_2 being fixed to $\frac{\hat{s}}{s}$. The value of $\sqrt{\hat{s}}$ is taken equal to m_H for gluons products and to $2m_W$ for quark product.The center of mass energy is $\sqrt{s} = 7$ TeV

In Figure 5.6 we show the pdf's products as a function of one of the two partons momentum fraction (x_1) ,the other (x_2) being constrained by Equation 5.1 at 7 TeV. The maxima of these curves indicate the most probable values of x_1 and x_2. The peak of the $q\bar{q}$ products lies in the high x region, resulting in a sizable imbalance between the momenta of the two partons entering the hard process. A rough estimate of the *Lorentz boost* of the W^+W^- system that comes out of the hard scattering is easy to infer:

Let us call $\sqrt{s_{lab}}$ the energy in the laboratory frame, and γ the Lorentz factor. We have:

$$\sqrt{s_{lab}} = \gamma\sqrt{\hat{s}} \tag{5.2}$$

Since we placed ourselves in the limit where protons are massless,

$$E_i = P_i = x_i\frac{\sqrt{s}}{2} \tag{5.3}$$

so,

$$\sqrt{s_{lab}} = E_1 + E_2 = (x_1 + x_2)\frac{\sqrt{s}}{2} \tag{5.4}$$

where E_i is the energy of the i-th proton. And finally,

$$\gamma = \frac{(x_1 + x_2)}{2\sqrt{x_1 x_2}} \tag{5.5}$$

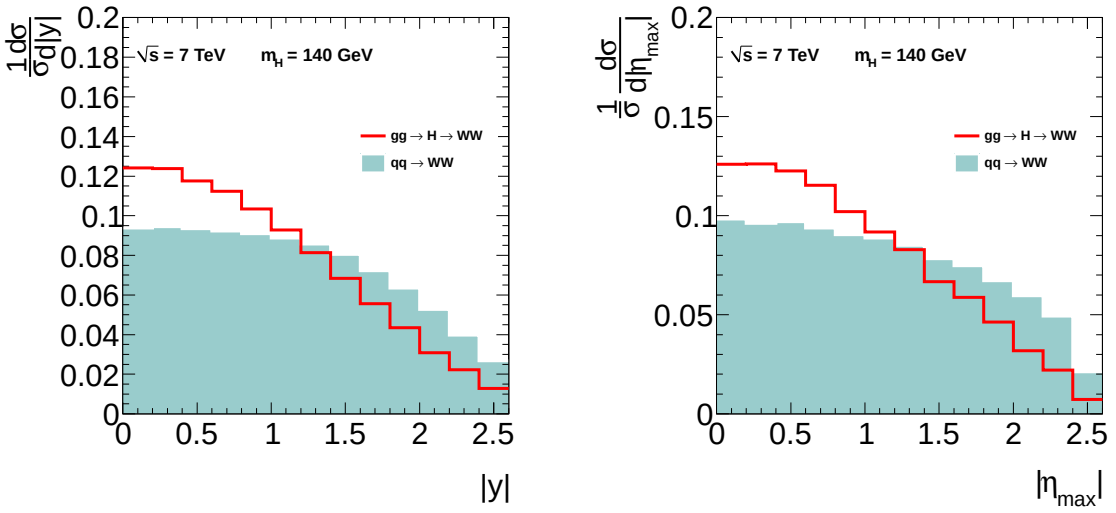
For the Higgs, regardless of the collision energy and its mass, we find that the system will be produced in average at rest, which is consistent with the fact that the same PDF describes the gluon in the two protons. On the other hand, since the proton is made of u and d valence quarks while $\bar{u}$ and $\bar{d}$ appear only via gluon exchange (sea quarks) their probability densities are suppressed at energies $x = 10^{-3}$ - 10^{-1}. This asymetry between q and $\bar{q}$ PDFs is responsible for an average momentum imbalance in W$^+$W$^-$ production. Using the formula 5.5 we can give a rough estimate of the Lorentz boost of the produced system.

$\sqrt{s}$	7 TeV	10 TeV	14 TeV
γ ($u\bar{u}$)	3.9	5.3	7.3
γ ($d\bar{d}$)	2.5	3.4	4.4

Table 5.2: Approximate boost values for the W$^+$W$^-$ background at different center of mass energies

We have summarized in Table 5.2 the approximate values of the boosts for the W$^+$W$^-$background at different center of mass energies. First we observe that systems produced by $d\bar{d}$ collisions are less boosted that those resulting from $u\bar{u}$. This results from the fact that the peak of $d\bar{d}$ in Figure 5.6 lie at lower x values than the one from $u\bar{u}$. Second, and more important, the W$^+$W$^-$ system gets more and more boosted as the energy of the collisions increase. This means in practice that with increasing energy the background events will be less and less central, resulting in a better separation from the signal, that lies in the central region of the detector.

In Figure 5.7(a) we show the rapidity distribution normalized to unity of the signal and background obtained from MonteCarlo simulation without reconstruction (at the generator level). As anticipated, the W$^+$W$^-$ events populate higher rapidity values while the signal events lie in more central regions. We also show in Figure 5.7(b) the pseudo-rapidity distribution of the most energetic lepton that comes from a W decays in the signal and background scenario. The very same effect is apparent here: the leptons coming from a Higgs decay are more central. We conclude from this simple argument that the signal over background ratio can be enhanced by rejecting events with leptons at high values of pseudorapidity. It is worth to stress again here that the arguments are rather qualitative, since we have been talking only about the most probable value for the boost and we have neglected the fact that other production mechanisms (gg fusion) are involved in W$^+$W$^-$ pair production.



(a) (b)

Figure 5.7: (a) Rapidity distribution of the W^+W^- system and (b) pseudo-rapidity distribution of the highest p_T lepton coming from the W^+W^- decay system in the case of the signal and background

5.2.2 Spin-Correlation of the W^+W^- system

We will see here how the V-A structure of the electroweak theory affects the decay of the W bosons, and how this can be used to infer properties of the angular distributions of the final state leptons.

Let us consider the Signal case first. The Higgs boson is a spin-0 particle while the W boson has a spin-1. We assume for simplicity that the Higgs boson is produced at rest in the lab frame.

Then the two W's are emitted back to back, as shown in Figure 5.8. By angular momentum conservation the spin vector has to be conserved in the process. This implies that the z- component of the spin needs to be conserved as well (the z-axis being defined along the direction of propagation of the W's). Thus we have $S_z(W^+W^-) = 0$. The W boson being a spin-1 particle we have $S_z(W^\pm) = 0, \pm 1$. We call these states W_T if $S_z(W) = \pm 1$ and W_L if $S_z(W) = 0$. Therefore, by spin conservation there are only two ways the Higgs can decay:

$$H \to W_T^+ W_T^-, W_L^+ W_L^- \tag{5.6}$$

In the Standard Model, the electroweak sector only couples to the "left-handed" component of the fermionic Dirac particles (or equivalently right-handed antiparticles). This means here that W^- decays into a "left-handed" electron and a "right-handed" anti-neutrino. Similarly W^+ decays into a "right-handed" positron and a "left-handed" neutrino.

We first look at the case $H \to W_T^+ W_T^-$. As shown in Figure 5.8 there are two ways for the W_T spins to have anticorrelated projections along the "z" axis. Let us look at the configuration displayed in the left part of Figure 5.8. $S_z(W_T^+) = -1$, so, for the spin to be conserved in this decay, both the electron and the neutrino must have $S_z = -\frac{1}{2}$. In the

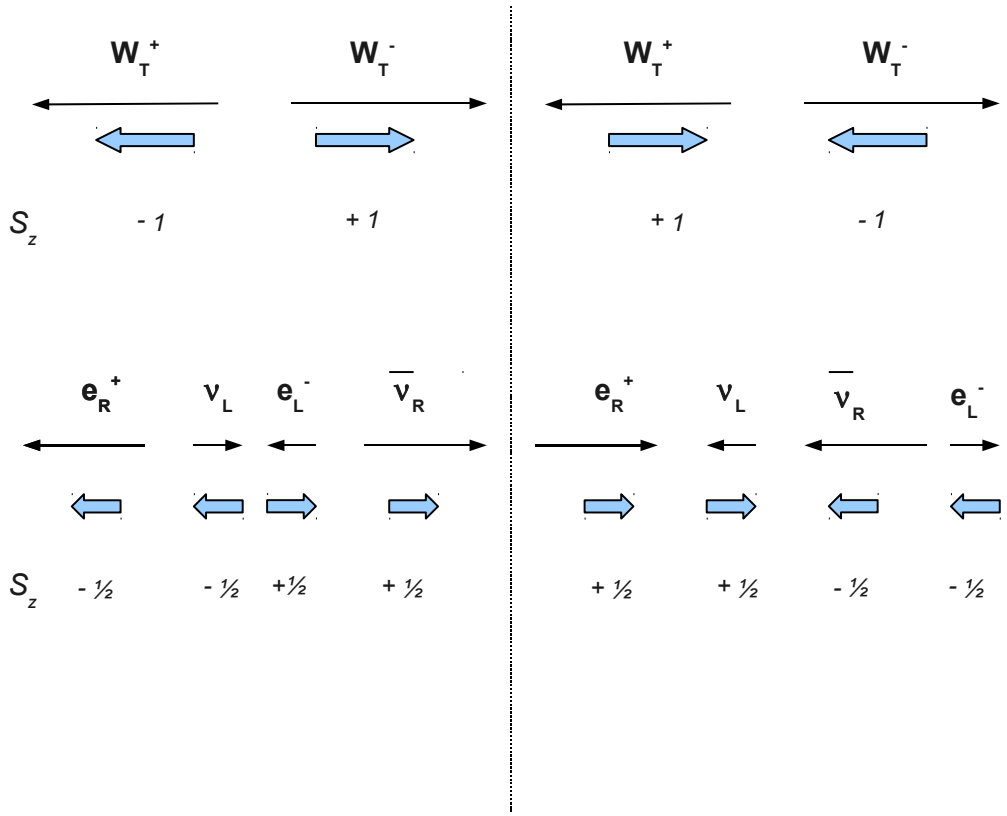


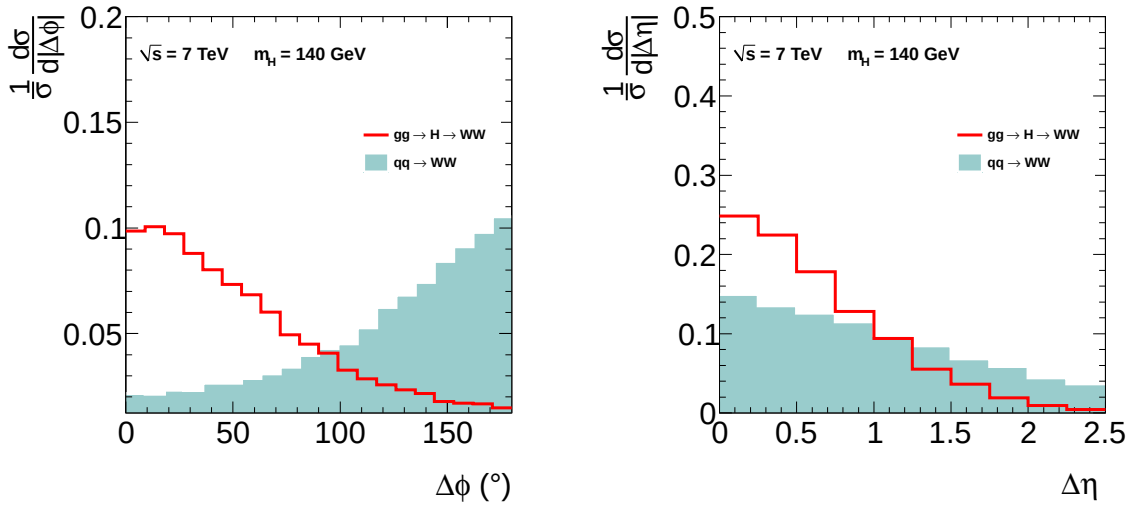
Figure 5.8: Simplified picture showing how the spin anti-correlation of the W$^+$W$^-$ system implies a small opening angle between the leptons

massless limit the chirality of a fermion is equivalent to its helicity. So, since the neutrino is left-handed (in the Dirac sense, but also in the helicity sense [3]) it must fly in the opposite direction of its spin. By momentum conservation the electron must fly in the opposite direction. In the massless limit of the positron we find that its helicity is indeed "right-handed". The same argument applies for the W$_T^-$: the right-handed antineutrino "forces" the electron backwards. By using a similar argument, it is easy to show that the other configuration (shown in the right part of the figure) also produces two leptons flying in the same direction. The same happens for the W$_L^+$W$_L^-$ case. This can be shown by taking another axis "i", orthogonal to "z" along which S$_i$(W$^\pm$) = ±1 and then the same argument applies. We have shown that the spin anticorrelation of the W's, induced by the spin-0 nature of the Higgs, implies that the two leptons will preferably decay along the same direction. What happens in the W$^+$W$^-$ background case? Consider here only the major contribution for W$^+$W$^-$ production, the quark induced process. The quarks being spin $\frac{1}{2}$ particles, they can form either spin-0 or spin-1 systems. Therefore, the W$^+$W$^-$ background can be produced in three different configurations of polarization:

$$q\bar{q}\rightarrow W_T^+W_T^-, W_L^+W_L^-, W_T^\pm W_L^\mp \tag{5.7}$$

[3] For the present argument it is perfectly fine to assume that neutrinos are massless.

The first two $W_T^+W_T^-$ and $W_L^+W_L^-$ are equivalent to the signal case. On the contrary, the $W_T^\pm W_L^\mp$ corresponds to a spin-1 configuration (z-projected) in which the two W's have an uncorrelated spin. Therefore the W's do not impose any correlation between the leptons momenta and thus allow for a wider opening angles between the leptons. It has been shown [79] that the magnitude of $W_T^\pm W_L^\mp$ production is approximately half of the total W^+W^- production for $m_{WW} \approx 160$ GeV.



(a) (b)

Figure 5.9: (a) Opening angle and (b) pseudo-rapidity gap between the two leptons in the case of the signal and background

We show in Figure 5.9 (a) the opening angle distribution of the visible leptons for the signal and the background. These plots were produced at the generator level, without any preselection or detector effects included. As showed the signal peaks at small values of the angle. For the background, we see indeed a peak at high values of angle in the tranverse plane ($\Delta\phi$ distribution, in Figure 5.9(a)) while we have a flatter distribution in the longitudinal plane Figure 5.9(b)) due to the longitudinal boost of the W^+W^- system (shown in the last section). We can therefore conclude that the opening angle between the leptons is a good discriminating variable between the signal and background: as it will be confirmed in the next chapter, rejecting events with high $\Delta\phi$ or $\Delta\eta$ values should increase the signal to background ratio.

5.3 Conclusion

In this chapter we have presented the phenomenology of the signal and the major backgrounds. After presenting the topology and production mechanisms of $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$, the main background mechanism were shown: the W^+W^- irreducible background and $t\bar{t}$ as well as minor contributions such as single-top, Drell-Yan, and W+jets. Later, efforts were concentrated on showing how to deal with the irreducible W^+W^- background. First, from considerations on parton densities, we showed that the background events can be rejected by cutting events at large rapidities due to a high boost of the W^+W^- system

compared to the signal. Second, we showed how spin anti-correlation emerge due to the $SU(2)_L$ structure of the couplings terms in the standard model and how these spin correlations are suppressed in the background case. These considerations showed that the Higgs will be decaying in two leptons with a small opening angle providing us another crucial discriminating variable for the suppression of the irreducible W^+W^- background.

Chapter 6

Object Reconstruction in CMS

In this chapter we will review the main reconstructed objects that are relevant for detecting the final state of the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ signal. We will start from the "low-level" reconstruction objects, such as vertices and tracks down to "high-level" objects such electrons, muons, jets and missing energy. We will focus on how the pile-up affects their performance. Finally we briefly introduce the b-tagging algorithm which helps in reducing the top induced backgrounds.

6.1 Track reconstruction

The determination of charged particle trajectories, along with the precise measurement of their direction and momentum plays a crucial role for this analysis, as these particles will finally be identified as electrons, muons, or simply particles that form a jet. The full track reconstruction algorithm consists of five main steps :

- local hits reconstruction
- seed generation
- pattern recognition (trajectory building)
- final track fitting
- Quality filtering

Local hits reconstruction The track reconstruction starts by clustering adjacent deposits of charges in the pixel or strip detector in so-called "hits". Hits are determined by their position in the $(\eta \times \phi)$ plane, the tracker layer, and the uncertainty on their position. These variables are used as inputs for the track reconstruction chain.

Seed generation Seeds are formed by associating a pair or a triplet of tracker hits in the innermost tracker layers (the pixel layers and the double-sided strip layers). These layers provide maximal information as :

- the occupancy rates are low, leading to lower uncertainty in the position determination.

- the amount of traversed material is low, lowering the probability that the particles loose energy by brehmstralung emissions or conversions in the case of electrons.

Seeds are constrained by the beam-spot region (defined as the profile of the most luminous collision region) and a simplistic estimate of the vertices position obtained only from pixel information. Seeds provide the starting estimate of charged particle trajectory.

Pattern recognition The track pattern recognition uses the so-called Combinatorial Track Finding algorithm (CTF) [80]. The seeds are extrapolated "inside-out" compatible with the equation of motion of a charged particle in the CMS magnetic field. When compatible hits are found, they are added to the trajectory and the uncertainties on the track parameters are updated. The propagation continues until no compatible hits are found or the tracker material has ended. This procedure is made flexible by allowing the user to specify parameters such as the minimum track transverse momentum, the minimum number of tracker hits or the maximum number of consecutive invalid hits. In order to reduce the bias introduced by choosing the inner tracker layer as the starting point, the procedure is reversed starting from the outside layers ("outside-in") checking if the trajectory is compatible with the initial one. Double counting is reduced by discarding tracks that share a significant number of common hits.

Track fitting The final collection of hits compatible with a trajectory is then re-fitted in order to assert a final estimate of the track parameters. For each hit associated to the track, the position and uncertainty are re-estimated by taking into account the energy loss and multiple scatterings occurring at each layer.

Quality filtering At high luminosity conditions secondary charged particles can be created by the interaction of prompt particle with the tracker material [1]. The high combinatorial possibilities involved are also responsible for fake tracks. The "fake rate" [2] can be reduced by keeping tracks that satisfy only certain criteria. The variables that are used at this scope are (among others):

- χ^2/ndof, the chi-square of the track per number of degree of freedom
- $\mathrm{d_0}$, $\delta(\mathrm{d_0})$ the distance in the transverse plane of the track to the primary vertex and its associated uncertainty
- $\mathrm{d_z}$, $\delta(\mathrm{d_z})$ the distance in the longitudinal direction of the track to the primary vertex and its associated uncertainty

At the end of this procedure, the track is defined by its momentum vector $\vec{p} = (\vec{p_T}, p_z)$ in the laboratory frame and its impact parameters to the nominal interaction point ($\mathrm{d_0}$,$\mathrm{d_z}$) of the CMS detector (and their associated errors).

[1] A particle that generates from the proton-proton scattering is called a prompt particle.
[2] To be distinguished from the probability jets have to fake electrons or muons.

6.2 Vertex reconstruction

Two kinds of vertices exist: primary and secondary vertices. The primary vertex is the interaction point of a hard proton-proton collision (there can be several reconstructed primary vertices per bunch-crossing, in the case of high luminosity). Secondary vertices are generated when long-lived particles decay in the detector volume. In order to reconstruct the vertex position (and uncertainty) reconstructed tracks are used. Mainly two algorithms are used in CMS, depending on the instantaneous luminosity delivered by the LHC:

- The standard clustering (low luminosity)
- The deterministic annealing clustering (high luminosity)

Standard Clustering First the reconstructed tracks in the event undergo a selection, involving requirements on the χ^2 of the fit, on the number of pixel and strip hits, and on the distance of closest approach to the beam-spot. The selected tracks are sorted according to their "z" coordinate and then clustered together if their z-distance is smaller than a parameter $z_{sep} \sim 2$ mm. These track clusters define the vertex candidates. The AVF (Adaptive Vertexing Fitting) algorithm iteratively minimizes the distance between the tracks and the vertex position hypothesis [81]. Along this procedure, a dynamic weight is associated to each track. The weight is constructed in such a way that it is close to 1 if the track is close to the vertex (good tracks), and close to 0 if it is far away (an outlier). It can therefore be interpreted as the probability that a track belongs to the vertex. The number of degrees of freedom of the vertex found by the AVF algorithm can then be written as:

$$\mathrm{ndof} = 2 \sum_{tracks} w_i - 2 \tag{6.1}$$

where the w_i are the weights associated to each track. The hard interaction vertex is then usually found by selecting the vertex candidate with the highest ndof, and sum of momenta of tracks associated to it.

Deterministic annealing (DA) clustering At high instantaneous luminosity, simultaneous proton-proton collision are frequent. The clustering procedure explained is consistent when the uncertainty on the "z" position of the tracks is small compared to the separation z_{sep} required for the clustering. If z_{sep} is chosen to be too small (which would be the natural thing to ask when several simultaneous interactions occur at the same time) low resolution tracks tend to be split off. One can apply a tighter selection to the tracks (for example selecting tracks with a small uncertainty on the z-position). However this can have the effect of increasing the probability that two close-by vertices get merged together.

The DA clustering procedure has the goal of finding a collection of vertex candidates and associating the vertices to the reconstructed tracks. Such an assignment can be quantified by:

$$\chi^2 = \sum_{ik} p_{ik}(z_i - z_k)^2 \tag{6.2}$$

where the index i (k) runs over tracks (vertices). Rather than finding the configuration (p_{ik}, z_k) that minimizes χ^2, the DA algorithm treats the $<\chi^2>$ as the mean energy of a thermodynamical system and looks for the most likely set of parameters that maximizes the entropy for a given mean energy value [82]. The mean energy ($<\chi^2>$) is then gradually decreased while keeping the entropy maximal. This procedure allows to find the final optimal set of parameters (p_{ik}, z_k) by reducing the bias introduced by initial parameters such as z_{sep} and avoiding local minima issues that would naturally occur by simply minimizing the χ^2 in Equation (6.2).

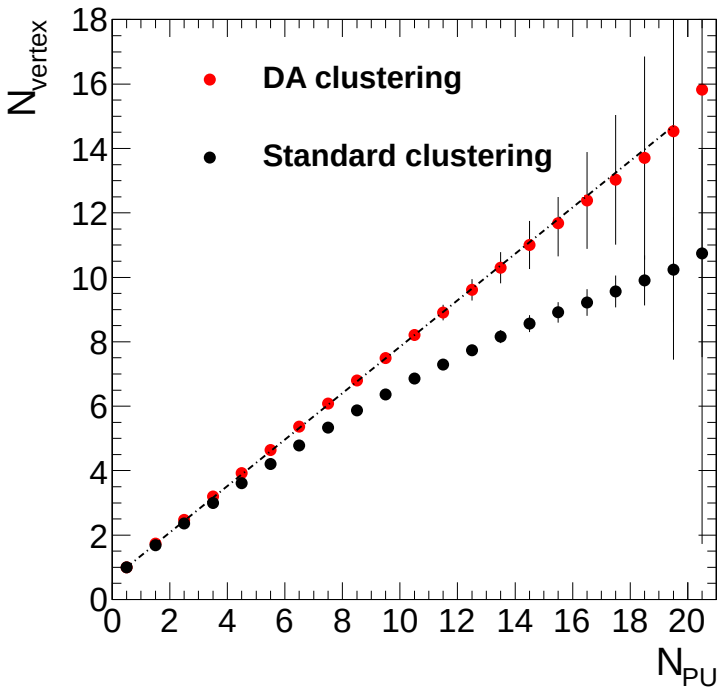


Figure 6.1: Average number of reconstructed vertices (N_{vtx}) as a function of the true number of hard interaction in a high luminosity scenario (N_{PU}).

The performance of DA clustering compared to the standard clustering procedure is shown in Figure 6.1. When the number of hard scatterings per event is high the DA algorithm performs better as no efficiency loss occurs. Vertices will be reconstructed using the DA procedure in the present study.

6.3 Jets

Partons originating from the hard interaction cannot be observed as isolated particles as these are not color singlets and confinement rather produces a spray of hadrons that are ultimately collected and measured by the hadronic calorimeter. When these particles are charged they also leave a measurable track in the tracker material. The most simplistic definition of a jets is a cone containing such collimated hadrons.

A jet algorithm defines a set of rules for grouping these particles together. It can be defined at two levels of the reconstruction chain:

- genJets , jets formed at the simulation (partonic) level , i.e no detector simulation nor reconstruction is involved

- recoJets , jets formed at the reconstruction level , i.e formed directly from the energy deposits in the detector material.

RecoJets can be formed out of the purely hadronic/electromagnetic deposits (CaloJets) or from so-called Particle-Flow candidates, which are higher-level objects obtained from the combination of the tracker, ECAL and HCAL information as well as the muon chambers (Particle-Flow Jets, or PFjets).

6.3.1 Jet Algorithms

Several ways to define a jet exist. A general consensus has been reached in 1990 in the "Snow-mass" accord [83]. A jet algorithm needs:

- be easily implementable both in theoretical calculations and in a experiment.
- must lead to a finite result at any order of perturbation theory

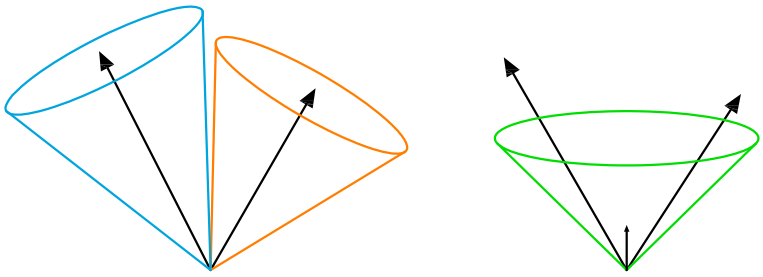


Figure 6.2: Left: two jets are reconstructed by the jet algorithm. Right: A soft emission modifies the jet structure as only one jet is reconstructed. This algorithm is not infrared safe

The first aspect is of particular importance in a QCD "rich" environment such as the LHC, where decisions have to be taken fast. A performing jet algorithm in this context is one that is able to cluster the particles in a jet in a short timescale. The second aspect is related to collinear and infrared safety. We have extensively treated this issue in Chapter 3. These divergences reflect the fact that the perturbation theory language breaks down for low-energy emitted partons. In the jet language, one has to make sure that collinear and soft partons emissions do not affect the definition of a jet. In other words a jet algorithm needs to be infrared and collinear "safe". This is illustrated in Figure 6.2 and 6.3.

We will now review the two most commonly used in CMS: the Iterative Cone, and the Ant-k_T. All these jet algorithm have in common a user configurable parameter: the jet cone size R.

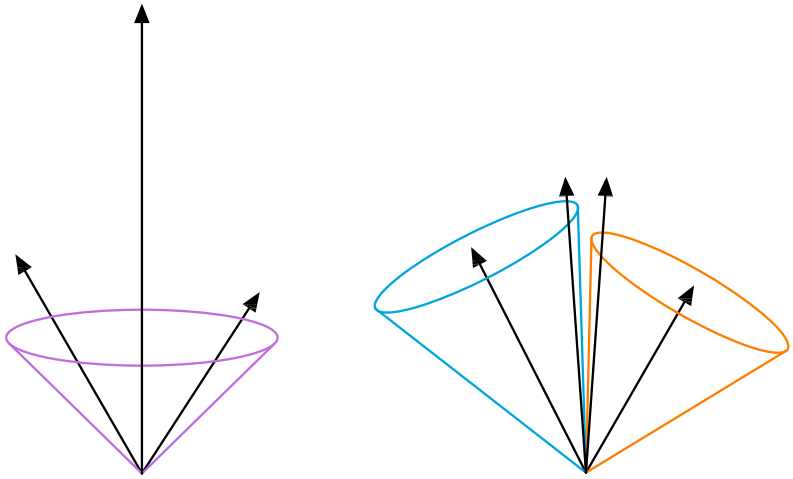


Figure 6.3: Left: One jet is reconstructed by the jet algorithm. Right: The central parton splits collinearly and modifies the jet structure as the two jets are now reconstructed. This algorithm is not collinear safe

Iterative Cone

The iterative cone algorithm (IC) is mainly used in the High Level Trigger since it is very fast. It proceeds in the following steps:

- A list of p_T ordered objects is formed (these can be tower deposits, or simply partons if GenJets are being created)
- A cone of size R, defined in the $(\eta \times \phi)$ plane, is drawn around the seed (the most energetic object). All objects contained in the cone define the protojet.
- Properties of the protojet are calculated, such as the total momentum and the $(\eta \times \phi)$ position:

$$p_T = \sum_i p_{T_i} \tag{6.3}$$

$$\eta = \frac{1}{p_T} \sum_i p_{T_i} \eta_i \tag{6.4}$$

$$\phi = \frac{1}{p_T} \sum_i p_{T_i} \phi_i \tag{6.5}$$

- The newly calculated η, ϕ coordinates define the new seed. The procedure is iterated until the jet is stable, that is, when the variation of its center between two iteration is below a certain threshold.
- The objects that formed the jet are removed from the list and the procedure starts again until no new objects are found.

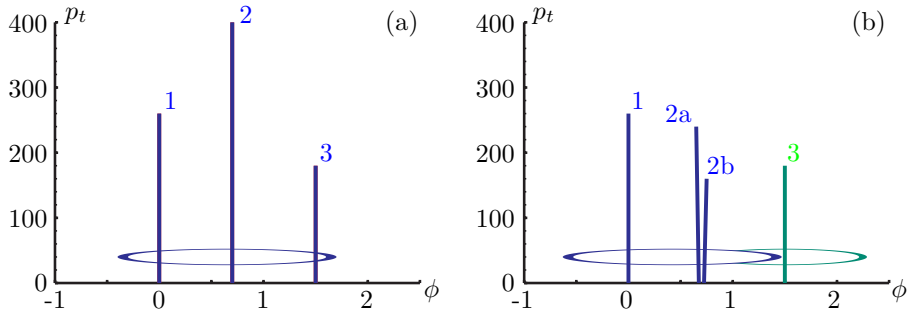


Figure 6.4: The IC cone algorithm in a simple example [84]

Although simple and fast, this algorithm is not collinear safe. This is well explained by Figure 6.4. To simplify, the detector is made 1-dimensional. In (a) 1 jet is found in a 3 particles event. If particle 2 splits collinearly the IC algorithm now finds 2 jets in the event. The IC algorithm is therefore no longer used in physics analyses. It has in most cases been replaced by the anti-k_T algorithm.

The anti-k_T algorithm (AK)

Rather than combining particle inside a cone and checking whether the cone is stable, the AK algorithm rather tries to reconstitute the parton shower backwards [84].

- the algorithm starts by computing the distance among all the available objects in the event d_{ij} and d_{iB}:

$$d_{iB} = \frac{1}{p_{Ti}^2} \tag{6.6}$$

$$d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta\phi^2\Delta\eta^2}{R^2} \tag{6.7}$$

$$\tag{6.8}$$

  where R is a cone size parameter. The smallest distance among the d_{ij} and d_{iB} is found.

- if d_{iB} is the smallest, the particle i is called a jet and removed from the list. If a given d_{ij} is found to be the smallest, the particle i and j are combined together and form a new particle with p_T, η and ϕ calculated as in Equation (6.5). The original particles i and j are removed from the list.

- The procedure is repeated until no particle are left.

As the algorithm forces hard particles and soft ones to be combined together first as well as collinear ones, the AK algorithm is by construction infrared safe and collinear safe. Also, the fact that soft particles recombine with hard ones before recombining among themselves, result in very regular jets, as one would expect. This is shown in Figure 6.5.

These properties make the AK algorithm the preferred one for physics analyses. In the present study, jets are constructed with the anti-k_T with a cone $R = 0.5$.

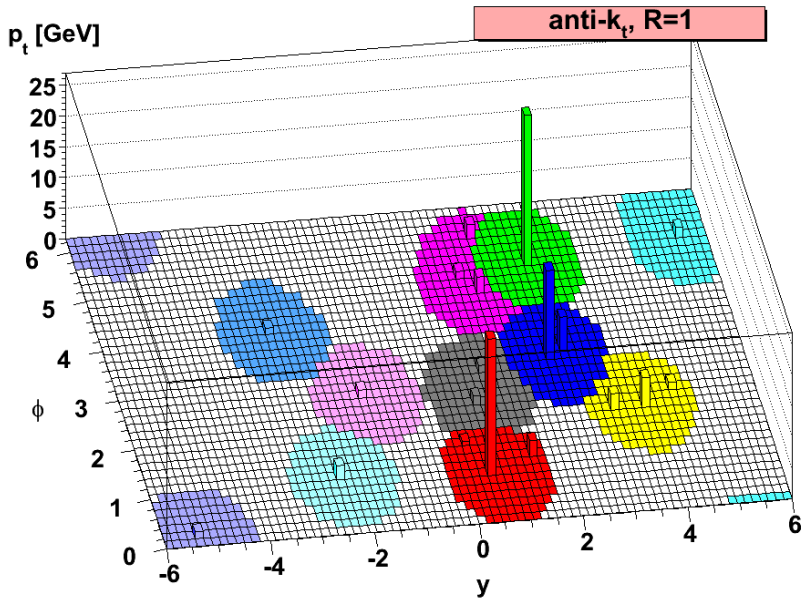


Figure 6.5: Illustration of the regularity of jets obtained with the anti-k_T algorithm [84]

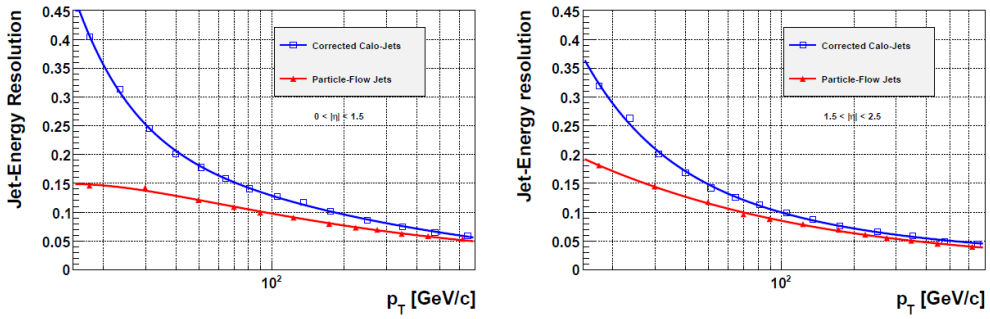
6.3.2 Jet Reconstruction and Corrections

Several types of reconstructed jets exist within the CMS framework, based on which information from the detector is chosen to be used. The CaloJets and Particle-Flow jets definition and their calibration will be discussed here.

Jets definitions

Calorimeter Jets (CaloJets) Calorimeter jets are reconstructed using the energy deposits in the ECAL and the HCAL cells. Cells are combined in "towers". In the barrel region ($|\eta| < 1.4$) of the calorimeters, a tower is formed from the sum of one HCAL cell and a 5×5 array of ECAL crystals. In the endcap region ($1.4 < |\eta| < 3.0$) the tower definition is more complex [85]. When building towers, thresholds on the deposited energies of single cells are applied in order to suppress the noise from readout electronics. Once towers are computed, they are used as input for one of the jet algorithms explained above.

Particle Flow Jets (PFJets) The Particle flow algorithm combines all the CMS subdetector information in order to form particle "candidates" such as electrons, muons, photons, charged and neutral hadrons. In the present study PF objects will be used only for jets and occasionally E_T^{miss} . We will therefore concentrate on hadronic objects. Charged hadrons are first reconstructed using the tracker information while neutral hadrons and photons are reconstructed from the energy deposits in ECAL and HCAL. Typically, neutral hadrons are identified by requiring that no track extrapolation to calorimeter cells matches with the energy deposits. When such matching occurs, the energy sum in the cells can be compared to the momentum measured from the track. An energy deposit excess in the cells can be interpreted as a neutral hadron. Once all the particle candidates are listed, they serve as an input to a jet algorithm in order to form Particle Flow jets.



(a) Barrel (b) End-Caps

Figure 6.6: Comparison of the jet energy resolution of Particle Flow Jets and CaloJets, in the barrel (left) and in the endcaps (right) [86]

Since the measurement of the energy of charged hadrons and photons is given respectively by the tracker and ECAL with this procedure, and these particles constitute ~ 90 % of a typical jet's energy, the momentum and spatial resolution of a PFjet are expected to be much higher than in CaloJets. This is clearly illustrated by Figure 6.6.

Jet Corrections

CMS has adopted a factorized multi-level approach for correcting the jet energy of raw jets [3]. In principle seven levels of corrections can be applied to a jet [87]. However for the present analysis only three are relevant:

- Level 1: Offset correction. This step mostly corrects for pile-up and electronic noise.
- Level 2: correction for variations of the jet response with pseudo-rapidity with respect to a reference control region
- Level 3: correction for variations of the jet response with momentum with respect to a reference control region

As the Level 1 correction will extensively be treated in the next paragraph, here we briefly review the principle of Level 2 and Level 3 corrections.

Level 2 The goal of Level 2 corrections on jets is to make the detector response to jets uniform in pseudo-rapidity. The central part of the detector, the barrel, gives the most precise response in measuring the energy of the jet since granularity is higher than in endcaps and the whole region is uniform. The barrel is therefore taken as the reference region in the Level 1 correction procedure. Rather than relying on simulation the correction needs to be inferred directly from Data. The idea is to use a physical process, abundantly produced at the LHC, for which the final state is perfectly balanced in the transverse region. This technique is often referred as "Tag and Probe". In this case CMS uses di-jet events that fulfill both these requirements.

[3] A raw object is in general defined as an object to which no detector correction are applied.

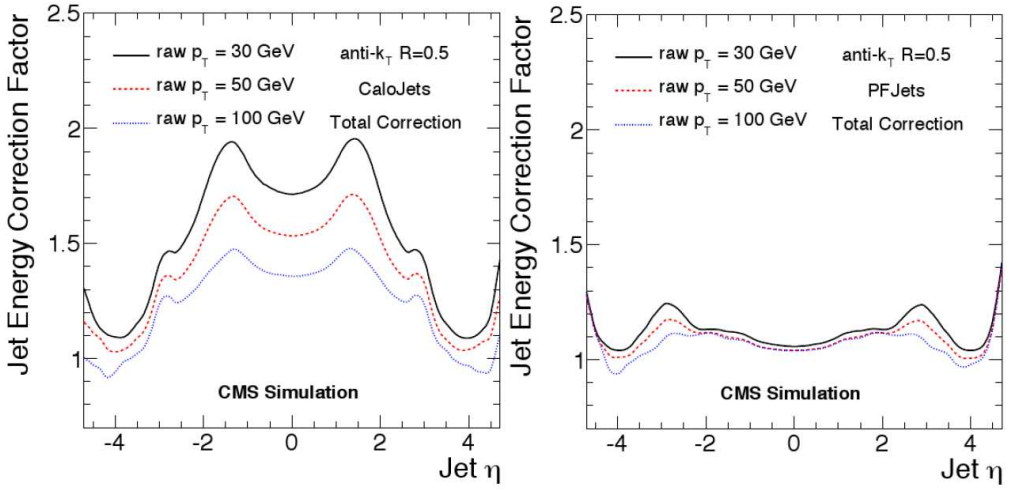


Figure 6.7: Total energy correction factor (L2+L3) derived from simulation as a function of pseudo-rapidity and for different jet momentum values for CaloJets (left) and Particle Flow jets (right)

If one jet is reconstructed in the barrel (the tag), one can probe the energy of the other jet (the probe) and map the imbalance as a function of pseudo-rapidity. The correction factor as function of η can then easily be derived from the jet imbalance [87].

Level 3 Once a flat response in η is obtained one has to correct for the fact that the jet response also varies with the momentum of jet. In order to achieve this, a process that produces a final state with one jet and another object needs to be used. The other must have the characteristics of having a good momentum resolution, and, most important, its momentum determination should not depend on the hadronic calorimeter response. Two processes that fulfil these requirements have been identified in CMS: γ+jet and Z+jet. The calibration with the former can be achieved with low integrated luminosity, as the cross section is large at the LHC, although QCD backgrounds are very significant and can bias the measurement. Z+jets on the other hand is a perfect candidate, as one can use the the $Z \rightarrow \mu\mu$ decay and exploit the excellent muon chambers performances of CMS.

In Figure 6.7 we show the total jet correction factor versus η for CaloJets and PFJets. One can easily see that PFJets are much better calibrated already before the L2 correction, as they use more detector information. Moreover, the p_T correction is very significant for CaloJets and almost negligible for PFJets. For the present analysis only Particle Flow jets will be used.

6.3.3 Pile-Up Subtraction

During a bunch-crossing several proton-proton interactions can happen (pile-up). We have already seen how the pile-up can have a detrimental impact on the vertex reconstruction efficiency if precautionary measures are not taken. This is also the case for jet reconstruction and energy resolution. The reason is the following. Pile-up interactions originate from minimum-bias (low momentum transfer) interactions. If the vertexing algorithm is well tuned, most of the vertices related to those interaction can be reconstructed (see Figure

6.1). Charged particles (in this case hadrons) originating from pile-up can easily be associated to a particular vertex by requiring that the distance between these two object is below a certain threshold. The problem comes however when one has to deal with neutral particles since they do not leave any track in the tracker volume.

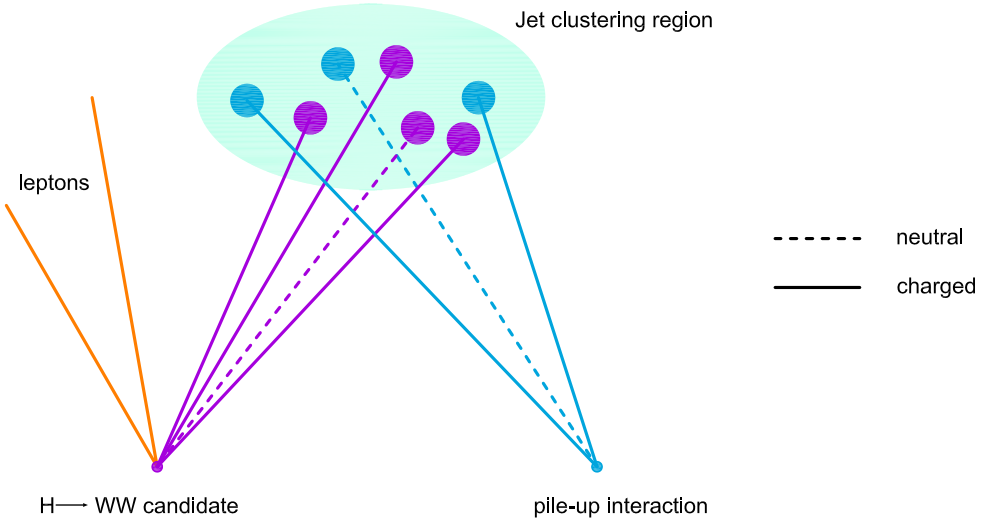


Figure 6.8: Schematic explanation on how the pile-up affects the jet reconstruction. While charged hadrons originating from a pile-up vertex can be identified with their track (solid line), neutral hadrons cannot (dashed line)

When an excess deposit is seen in the Calorimeter cell, the jet algorithm has no way of knowing that the neutral candidate originate from one vertex or another. If a neutral particle comes from a pile-up vertex, the particle gets clustered, along with other genuine prompt particles, to form a jet that will have therefore a sensibly higher energy than the original jet that originates purely from the hard interaction (see Figure 6.8). This offset can be very significant with 20 - 50 pile-up interactions present at the LHC in a high luminosity regime and needs to be corrected. This is the goal of the Level 1 jet energy corrections (L1). We will review here two classes of methods that were designed to deal with this issue.

L1 Offset

The L1 Offset is the method used at the Tevatron for correcting for pile-up biases in the jet energy scale. The principle is straightforward: one parameterizes the pile-up energy offset as a function of the number of reconstructed vertices and the detector region (in this case divided in η region), as one expect the pile-up activity to be symmetrically distributed in

the transverse plane. This parameterization has the advantage of being relatively simple to estimate since it can be measured directly from a minimum bias sample. The main disadvantage is that is estimated a priori, and does not take into account the dependence of the particular physics analysis one is considering. Moreover, it is an average correction, in the sense that two jets reconstructed in the same region in η from two different events, will essentially receive the same correction provided that the number of reconstructed vertices in the event are equal. This correction does not take into account the particular jet properties such as the momentum or the area (see later for a definition of the jet area). Finally, the parameterization needs to be updated as soon as the beam conditions change, as out-of-time pile-up can change through time [4]. Nevertheless it is a stable method that has been repeatedly tested in the past.

L1 Fastjet Area

Several definitions of a jet area exist for an infrared and collinear safe jet algorithm [88,89]. The definition used here is the so called *active area* that is estimated by the FastJet algorithm.

The area of a jet can be understood as a measure of the jet susceptibility of being contaminated by uniformly distributed soft particles. A grid of infinitely soft particles (typically 10^{-100} GeV) called "ghosts" is generated in a given $|\eta|$ range; the ghosts that get clustered within a given jet form the jet area (A). If a jet is purely coming from a PU interaction, the level of contamination is p_T/A, where p_T is the transverse momentum of the jet. The density of PU contamination per event is then computed as:

$$\rho = \text{median}\left[\frac{p_{Tj}}{A_j}\right] \tag{6.9}$$

where j runs over all the jets in the event within a given region set by the user. The median (and not the average) is taken so that the PU density ρ is less sensitive to genuine high p_T non-pile-up jets and low-area jets.

In Figure 6.9 the p_T/A is shown for all jets in a high pile-up event. Besides a few genuine hard jets, most jets display a $p_T/A \sim 25$ GeV that will dominate the pile-up contamination estimate.

In every event, the pile-up contamination can then be estimated with this method and the jets momenta can be corrected according to the following formula:

$$p_{T_i}^{corr} = p_{T_i} - \rho A_i \tag{6.10}$$

where ρ is calculated for each event and is common for every jet.

In Figure 6.10 the average jet offset, calculated as the difference between the reconstructed jet p_T and the jet p_T obtained from simulation without pile-up, is shown as a function of the number of pile-up interaction N_{PU} in a $H\rightarrow W^+W^-\rightarrow \ell^+\nu\ell^-\bar{\nu}$ simulation sample. The black dots show that if no correction is applied, as pile-up increases the true jet momentum is not properly determined. By comparing the L1Offset and L1 Fastjet methods we see that the latter does better in recovering the original jet momentum.

Overall the L1FastJet method provides a better estimate of the jet originating from the hard scattering. It will therefore be used as the standard L1 correction method throughout this work.

[4] Typically, when the read-out timing of a particular subdetector (in this case the HF calorimeter cells) is higher than that of the bunch crossing, can result in recording an event from previous/next bunch-crossing

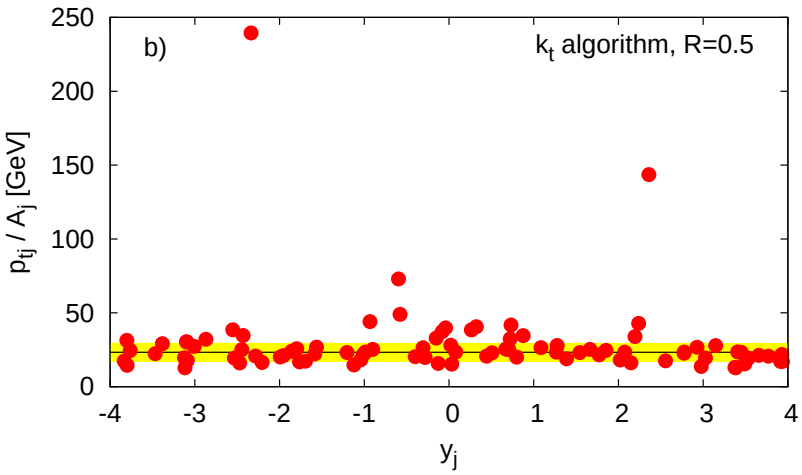


Figure 6.9: Typical values of p_T/A in a high pile-up environment with many jets. The red dots represent the jets in the event, from [89].

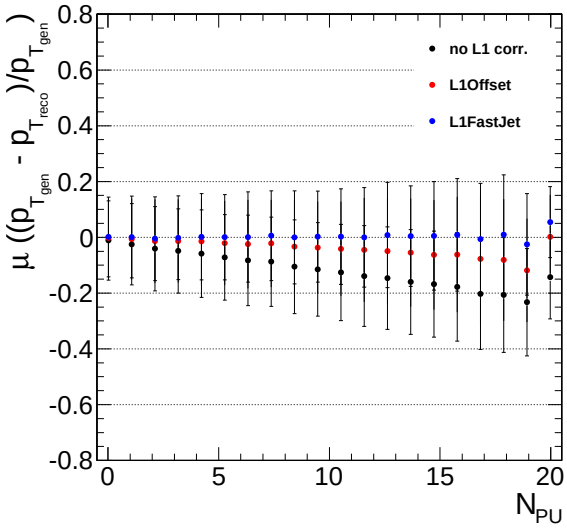


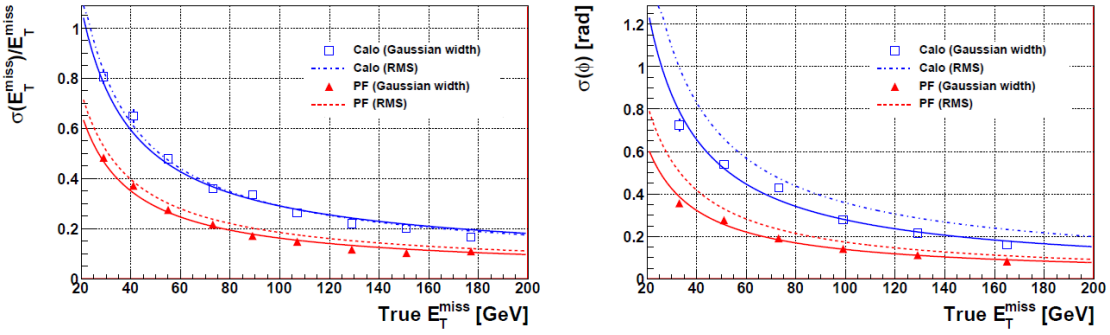
Figure 6.10: Comparison of the jet p_T resolution with pile-up, after no correction is applied (black), L1Offset applied (red), L1FastJet Area (blue) as a function of the number of pile-up interactions

6.4 Missing Transverse Energy

The missing transverse is a crucial component of the present analysis because of the presence of neutrinos in the signal final state. The E_T^{miss} is defined as the norm of the vector obtained by projecting in the transverse plane the negative sum of all the vectorial components of the detected particles. As for jets there, we are interested in two definitions of E_T^{miss} :

- Calorimeter E_T^{miss}
- Particle Flow E_T^{miss}

The former is calculated from the calorimeter towers and the latter is simply obtained by adding the contribution of all the reconstructed particle flow candidates in the event. The performance and resolution of the E_T^{miss} is usually tested from physical processes that display no-real E_T^{miss}, such as pure QCD samples, or Z bosons decaying to visible leptons.



(a) (b)

Figure 6.11: Comparison of the E_T^{miss} resolution of Particle Flow Jets and CaloJets of the norm (left) and the angle (right) [86]

As for the jets, the Particle Flow based E_T^{miss} has a much better resolution than the Calorimeter based one, as can be seen in Figure 6.11.

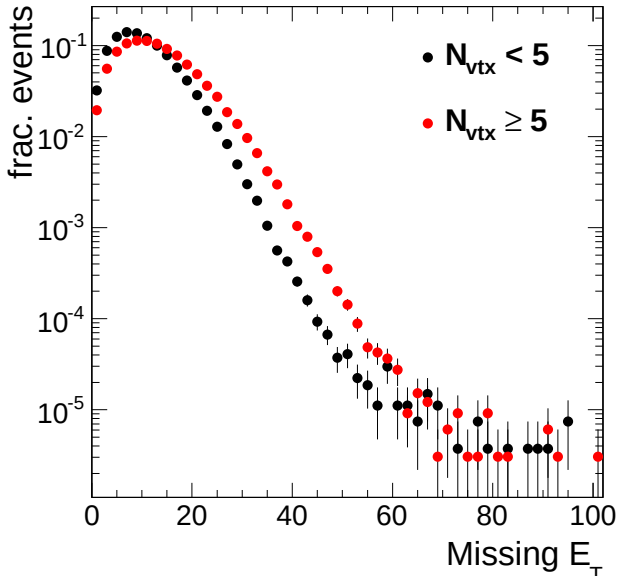


Figure 6.12: E_T^{miss} distribution in simulated Z $\rightarrow \mu\mu$ events with high/low pile-up. N_{vtx} is the number of reconstructed vertices and as it is strongly correlated with the number of true pile-up, we consider that it is a reasonable measure of the amount of pile-up interactions in the event.

Similarly, problems arise when high pile-up is involved. When no real E_T^{miss} is expected, high pile-up has a severe impact of the E_T^{miss} resolution. Unlike jets, one can-

not simply correct for the upward bias introduced by pile-up contamination, when fake E_T^{miss} is involved, since in this case pile-up does not introduce and upward bias on a localized vectorial quantity. Rather, since the contamination is diffused within the entire geometrical acceptance the effect will be to severely affect the resolution on events with fake E_T^{miss}. This is easily observed in Figure 6.12 on Z $\rightarrow\mu\mu$ events. As pile-up increases, the E_T^{miss} distribution gets smeared towards higher values. The solution to this problem is analysis dependent, and will be addressed in the next chapter, when details of the signal extraction will be given.

6.5 Leptons

The presence of high energy leptons are usually an indication of a weak $W^{\pm}$ or Z decay. Their reconstruction and identification is therefore very important for a $H\rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}$ search. The reconstruction of electrons and muons involves in both cases tracking. A matching between the track and a deposit in the ECAL (for electrons) and Muon Chambers (for muons) is then required for a complete reconstruction of electrons and muons. τ lepton reconstruction is not treated in this study.

6.5.1 Electrons

An electron that decays in the central part of the detector produces hits in the silicon tracker and gets ultimately stopped in the ECAL crystals. By interacting with the tracker material, the electron produces a shower of particles (electron/positron pairs and photons) that will follow its trajectory and hit the ECAL crystals. Because of this, a typical hit in the ECAL is spread and several crystals (a "cluster") are hit. The 'brehmstrahlung recovery' clustering algorithm is a dedicated procedure developed by CMS that builds and collects single clusters and merges them into the prototype of an electron candidate called a 'supercluster'. The 'supercluster' (SC) contains all the energy of the electron plus its shower products. The higher the momentum of the initial electron, the better the electron description as a supercluster.

Reconstruction

Two different reconstruction methods are available in CMS for electrons:

- The tracker driven seeding (low p_T)
- the ECAL driven seeding (high p_T)

The first method is suitable for low momentum electrons that cannot reach the ECAL. The second method concerns high momentum electrons that have reached the ECAL. As leptons originating from a $W^{\pm}$ or Z decay have generally a high momentum ($p_T > 20$ GeV), we will discuss only the ECAL driven reconstruction here.

The electron reconstruction follows these steps:

- The starting point in order to trigger an ECAL seeding is the presence of a supercluster (SC) with an energy deposit > 4 GeV. A supercluster is typically a matrix of 5×5 ECAL cells.

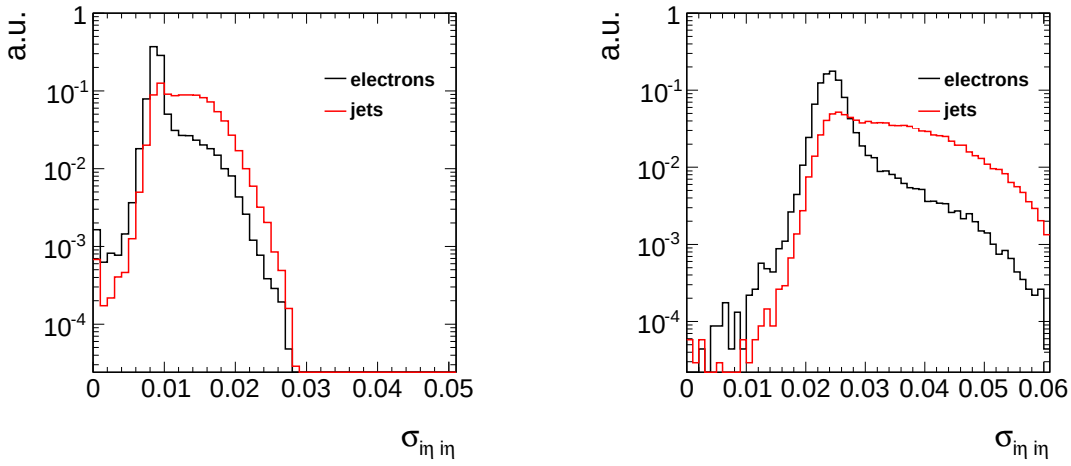
- The SC is then matched to a prototrack formed by two hits in the tracker that are compatible with the beam-spot. This requirement ensures a high purity of electron candidates and enhances the reconstruction efficiency. The tracking procedure starts here.
- The track is propagated to outer layers with the help of a Bethe Heitler modeling for energy loss from electrons from bremsstrahlung and conversions (tracks suffer from non-Gaussian fluctuations due to Brehmstrahlung emissions).
- The tracking ends when the last layer is reached. The best candidate having a trajectory that matches the SC is kept (based on the χ^2 of the fit). The charge of the electron is determined by the curvature of the track.

Idenfication

In the previous section we have seen how an electron candidate is formed. Depending on the particular analysis and the backgrounds to reject other criteria can be applied to electron candidates in order to increase the signal purity. Here we are just going to list the most relevant variables that are used for electron identification

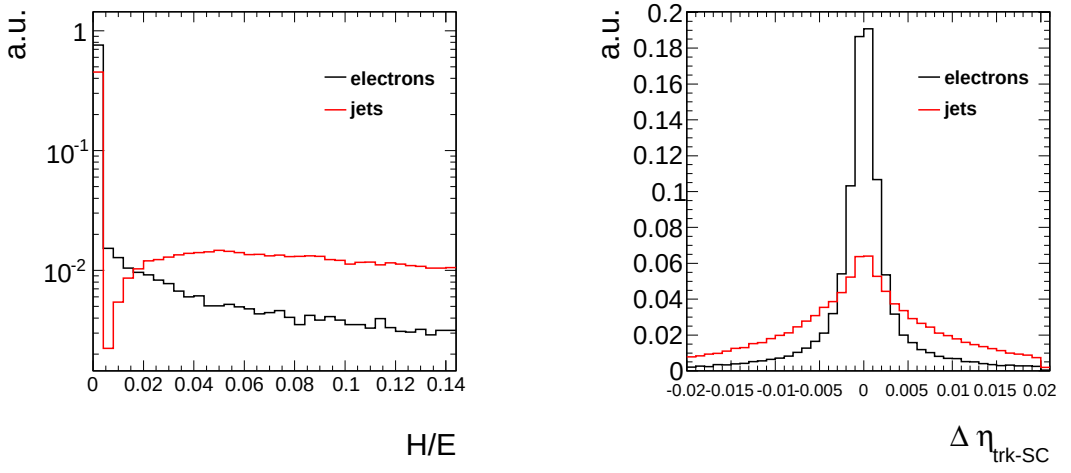
- $\Delta\phi_{\mathrm{trk-SC}}$ and $\Delta\eta_{\mathrm{trk-SC}}$. The difference in angle and pseudo-rapidity of the track position (at the vertex) and the supercluster. These variables describe how well the track matches the supercluster position.
- H/E is the energy deposited is the hadronic calorimeter divided by the energy deposited in the electromagnetic calorimeter. This variable is mostly used in order to reject jets faking electrons. A high H/E ratio indicates that substantial fraction of the energy was deposited in HCAL, which might be an indication of a fake electron (=a jet).
- $\sigma_{\mathrm{i}\eta\mathrm{i}\eta} = \sqrt{\sum_{\mathrm{i}} \frac{\mathrm{E_i}}{\mathrm{E_{center}}}(\eta_{\mathrm{i}} - \eta_{\mathrm{center}})^2}$ measures the standard deviation of the SC cell position in eta (weighted by the energy fraction of the cell) in the 5×5 matrix of crystals centered on the seed crystal. Since brehmstrahlung occurs mainly in the ϕ direction (because of the magnetic field direction) this variable is insensitive to it and essentially measures the spread of the electron deposit. It helps in discriminating real electrons against jets that leave a more spread deposit than "good quality" electrons.
- $\mathrm{f_{brem}} = \frac{\mathrm{p_{in}} - \mathrm{p_{out}}}{\mathrm{p_{in}}}$ is the fraction of the electron energy lost by bremsstrahlung in the tracker volume. A real energetic electron typically loses more energy than a jet that fakes an electron.
- $\mathrm{E_{SC}/P_{in}}$ The energy measured in the supercluster divided by the momentum measured by the track momentum before entering the tracker volume. It also a measure of the fraction of the energy lost by brehmstrahlung.

These variables are shown in Figure 6.13, 6.14, 6.15.



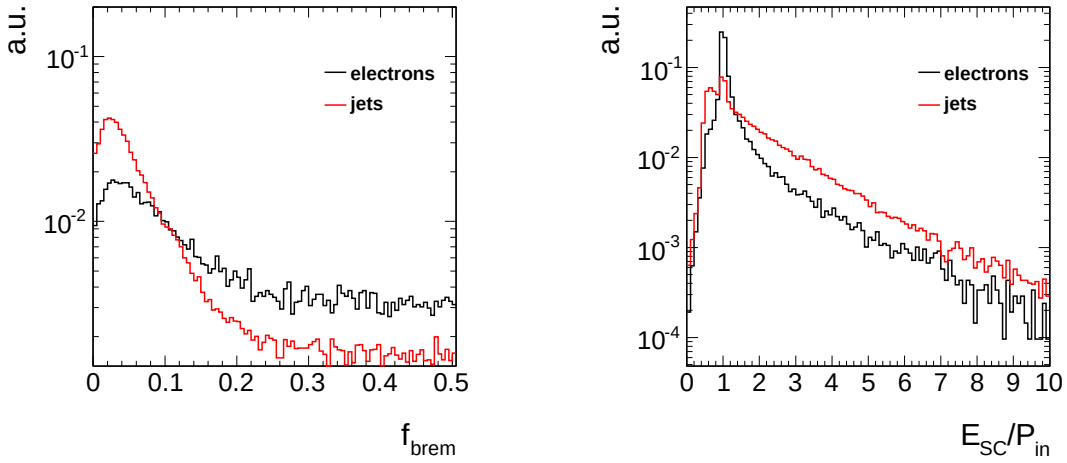
(a) (b)

Figure 6.13: $\sigma_{i\eta i\eta}$ in the barrel (a) and in the endcaps (b) for electrons and jets (fake electrons) in simulation. The electron candidates are required to fulfill the requirement $p_T > 10$ GeV



(a) (b)

Figure 6.14: $\Delta\eta_{\mathrm{trk-SC}}$ (a) and H/E (b) for electrons and jets (fake electrons) in simulation. The electron candidates are required to fulfill the requirement $p_T > 10$ GeV



(a) (b)

Figure 6.15: (a) f_{brem} and E_{SC}/P_{in} (b) distributions for electrons and jets. The electron candidates are required to fulfill the requirement $p_T > 10$ GeV

6.5.2 Muons

Reconstruction

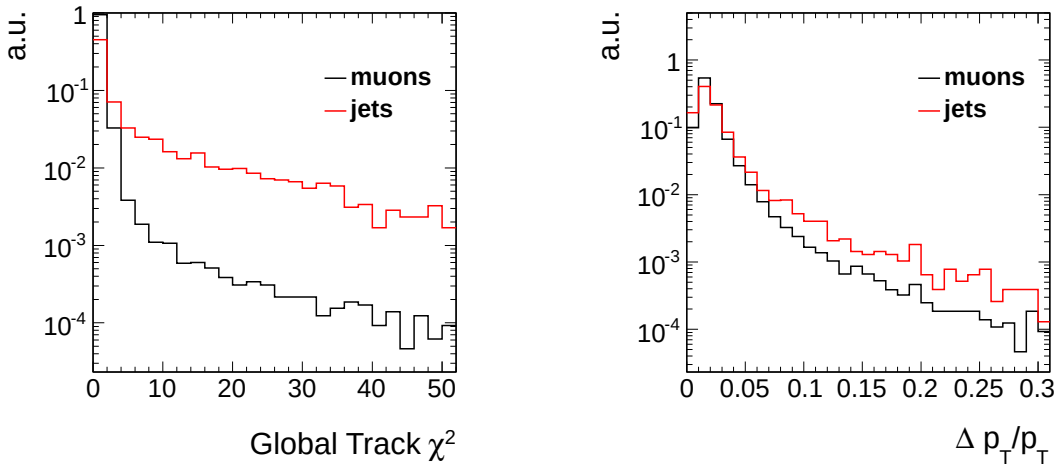
As muons produce a track both in the tracker and in the muon spectrometer (RPC, CSC, DT) they are reconstructed by making use of both these subdetectors. The former track is called a *tracker track*, the latter is a *standalone-muon track*. Three reconstruction approaches are used:

- Global Muons (outside-in): The starting point here is a standalone-muon track. If it matches a tracker track a global fit including both hits from the tracker and the muon spectrometer is made. The fit leads to a *global-muon track*.
- Tracker Muon (inside-out): The starting objects are tracker tracks. Possible muon candidates are required to have a $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV. The candidates are then extrapolated to the muon system by taking into account the expected energy loss and the uncertainty due to multiple scattering. If the extrapolated track matches a muon segment (a short track of DT or CSC hits), the tracker track qualifies as a *tracker-muon track*.

High momentum muons reach the muon system more often than soft muons. Thus, global muon reconstruction is more appropriate for analyses involving high energy muons, while tracker muons suit better low momentum muon based searches. Thanks to the high reconstruction efficiency of the tracker, 99% of the reconstructed muons are either global or tracker muons.

Identification

Given that muons are the only detectable particles able to reach the muon chambers, the misidentification rate is very low. Also muons are much less subject than electrons to brehmstrahlung radiation which makes them much cleaner object. The muon identification variables are essentially track related variable (both in the tracker and in the muon spectrometer). Typically these are:



(a) (b)

Figure 6.16: (a) χ^2 and (b) $\Delta p_T/p_T$ distributions for muons and jets. The muon candidates are required to fulfill the requirement $p_T > 10$ GeV

- The χ^2 of the global fit (for global muons)
- The number of "good" hits both in the tracker and in the muon detector. A hit is labeled as valid or good when the distance on surface between that track and the hits is lower than some threshold that varies depending on the tracker layer/muon chamber
- The relative error on the transverse momentum , $\Delta p_T/p_T$.

The variables χ^2 and $\Delta p_T/p_T$ are shown in Figure 6.16 for prompt muons and fake muons. Other variables combining muon segment information can be used. The algorithm that is used in this analysis for muon identification checks whether the candidate satisfies the following requirements:

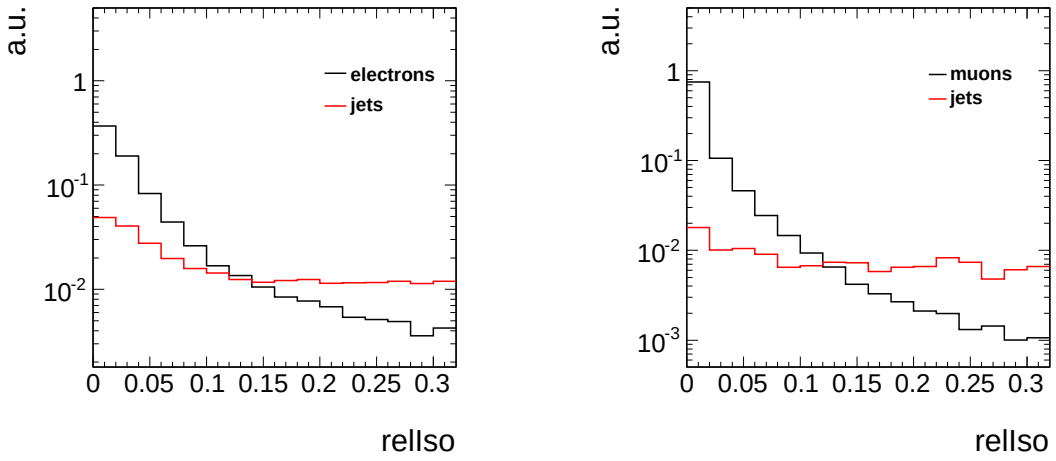
- At least two segments in the muon chamber need to be matched (the distance between the track and the segment is below a certain threshold) to the extrapolation of the track
- At least one of these two segments must belong to the last muon station that is crossed by the track extrapolation.

We will refer to the outcome of such requirement as *Tracker Muon Last Station*.

6.5.3 Lepton isolation

In the previous sections we have briefly reviewed the variables responsible for electron and muon identification. An efficient fake lepton rejection can be achieved by requiring *isolation*. Jets that fake a lepton, (i.e. that are accepted as muon on electron candidates) are usually surrounded by a high amount of activity (neutral or charged). Requiring that a lepton is isolated is equivalent to asking that such activity is absent (or small) and can sensibly reduce the fake rate. Quantitatively, the relative isolation variable can built as:

$$\text{relIso} = \frac{\sum p_T^{tracks} + \sum p_T^{em} + \sum p_T^{had}}{p_T^{lepton}} \tag{6.11}$$



(a) (b)

Figure 6.17: (a) The relative isolation variable for electrons (a) and muons (b) compared to jets faking the leptons in simulation

where the sums in the numerator run over all the particle flow candidates that are contained in a cone of size R around the lepton and p_T^{lepton} is the lepton nominal transverse momentum. Both R and the maximum value that one allows for the isolation are parameters that are chosen according to the rejection power one wants to achieve against the fake rate. The isolation variable with a cone size of 0.3 are shown in Figure 6.17 for electrons and for muons. Requiring relIso to be below a certain value is then equivalent to increase the purity of real leptons with respect to fakes.

The performance of isolation can be seriously degraded with pile-up. As particles originating from pile-up vertices can end up in the isolation cone, the value of the numerator of Equation (6.11) increases, and signal leptons can be rejected. A naive solution would simply be to increase the threshold for the isolation, but in that case one accepts too many fakes when the number of pile-up interaction is small. The problem comes from the fact that the number of pile-up interactions is not constant among different events/data taking period.

Pile-up subtraction with FastJet One method for subtracting pile-up contribution from the isolation cone is therefore the Fastjet area method, as it provides an estimate of the pile-up activity on an event by event basis. With pile-up, the isolation variable can be defined as:

$$\text{relIso}^{\text{PU}} = \text{relIso} - \frac{\rho \pi R^2}{p_T^{lepton}} \tag{6.12}$$

where ρ is the pile-up density measured in momentum per unit area and R is the isolation cone size. However it has been shown [90] that this method is not able to perform an adequate rejection of fakes when the pile-up is high.

Pile-up subtraction with Particle Flow A more intuitive procedure would consist in calculating the isolation variable by counting in the cone only particles candidates that are attached to the primary vertex that is associated with the lepton. As neutral hadrons and photons cannot directly be associated to the vertex, a threshold on their momenta can be imposed, so that the probability that they originate from minimum-bias vertex is reduced. The particular implementation of this variable will be discussed in the next Chapter.

6.6 B-tagging

Top induced processes represent a significant fraction of the background for this analysis. Hence the reconstruction and identification of jets that originate from the showering/hadronization of heavy flavor quarks (mostly b-quarks) is of particular importance. B hadrons have the feature of having a longer lifetime , compared to light hadrons, [5] This often results in a displaced vertex. Other features such as the high mass of B hadrons and their semi-leptonic decays products can also be used as a discriminating factors.

6.6.1 Ingredients

The inputs needed for a performant b-tagging algorithm are jets and tracks. The high granularity of the CMS tracker makes the tracking the most powerful tool for b-tagging. The identification of b-jets relies primarily on hits in the pixel detector, as it is the closest to the interaction point. Tracks are then reconstructed (see Section 6.1) and selected if they satisfy the following requirements:

- total number of hits (pixel+strips) in the tracker > 7
- number of hits in the pixel detector > 2
- transverse impact parameter to beam axis $d_{xy} < 0.2$ cm
- longitudinal impact parameter to the beam-spot $d_z < 17$ cm
- transverse momentum of the track $p_T > 1$ GeV
- $\chi^2/\text{ndof} < 5$ (for the track fit)
- distance to the jet axis $\Delta R < 0.5$

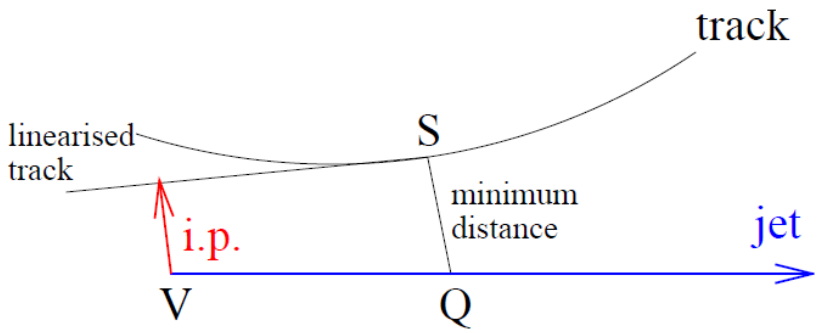


Figure 6.18: Schematic representation of the track distance with respect to the vertex

Another important ingredient to b-tagging are soft muons. B-hadrons tend to decay leptonically, with the lepton decaying perpendicular to jet axis. As muons display a better purity than electrons, their presence is often used as an additional discriminant for b-tagging.

6.6.2 Algorithms

The most powerful single track discriminant for the association of a track to a displaced vertex is the distance between the track and the vertex at the point of closest approach: the impact parameter (IP) (see Figure 6.18). It can be defined either in the transverse plane or in 3 dimensions. As the uncertainty σ_{IP} associated with the IP can be high, one rather uses the significance as a discriminator:

$$\mathrm{sign}_{\mathrm{IP}} = \frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}} \tag{6.13}$$

The simplest way of producing a b-tagging discriminator out of tracks IP significance is the so-called *track counting* algorithm. This approach identifies a jet as b-jet if there are at least N tracks with $\mathrm{sign}_{\mathrm{IP}} > \mathrm{S}$. The single discriminant is then obtained by fixing the value of N, and consider as discriminating variable the significance of the Nth track (the tracks are ordered in decreasing significance). Two choices are possible in CMS for this class of algorithms:

- TCHE (Track counting high efficiency) If one is interested in a high b-tagging efficiency , the number of tracks are fixed to N=2. This is obtained at the cost of increasing the mis-tag rate [6]

[5] Usually measured by the Lorentz invariant distance $c\tau \sim 480\ \mu$ m (τ here is the particle lifetime).

[6] The mis-tag rate is defined as the probability that a jet which does not originate from a b-quark gets tagged as a b-jet.

- TCHP (Track counting high purity) If one is interested in high purity, N=3. This is obtained at the cost of lowering the b-tagging efficiency.

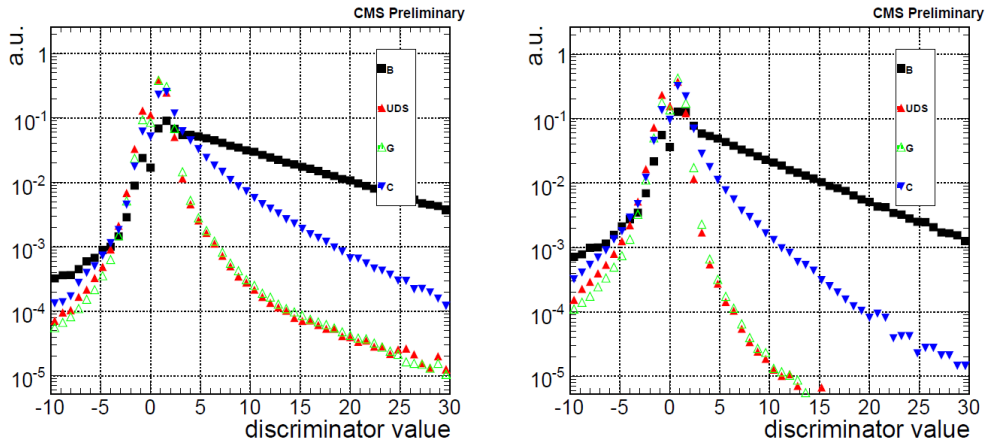


Figure 6.19: Distribution of the the TCHE (left) and TCHP (right) discriminators for light and heavy quarks in simulated QCD events. ([91])

The previous discriminators are shown in Figure 6.19 in QCD simulated events. The black dots indicate b-jets. The b-tagging efficiency if higher for THCE (left) than for TCHP (right) as a higher fraction of b-jet have a high discriminator value for TCHE. Conversely more light jets (red/green) have an higher discriminator value for TCHE than for TCHP, which result in a higher mis-tag rate for TCHE. The choice of the algorithm depends of the relative weight one associates to efficiency versus mis-tag and is analysis dependent.

Other algorithms exist in CMS, but as they are not used in the present work, the reader can refer to [91] for a complete review.

Chapter 7

The W^+W^- common preselection

In this Chapter will first explain the general philosophy of this analysis. We have seen in Chapter 5 that the final state we are interested in consists of two highly energetic leptons (electrons or muons), two neutrinos and eventually 0 or 1 jet. As the final state involves neutrinos (and therefore missing energy) one cannot reconstruct the full invariant mass of the Higgs boson via the momenta of the particles in the final state. Therefore the strategy of this search will be to design a particular phase space region in which the signal is minimally reduced and the backgrounds maximally reduced and eventually observe and quantify an excess with respect to the Standard Model backgrounds expectation.

We first introduce the samples that we will be analyzing both in simulation and in data. Building the phase space is mainly the subject of Section 3 and 4 of this Chapter. A first effort will be dedicated to kill all the reducible backgrounds by sequential selection criteria on the most relevant variables. After this stage, mainly the irreducible WW background will be left.

7.1 The Simulation and Data samples

7.1.1 Simulation samples

General Features

We have already listed in Chapter 5 the main backgrounds for this analysis. In Table 7.1.1 the background and signal samples are given. In the first column the process name is listed. The second column contains the MonteCarlo generator that was used while the third column contains the cross section times branching ratio for the corresponding process.

Only official CMS simulation samples from are used throughout the analysis. Three different generators were used for the various samples: Higgs signal samples were generated with Powheg while background samples were generated with either Pythia or Madgraph [43] [47] [50]. The only exception is the sample corresponding to W^+W^- by gluon fusion ($gg \rightarrow W^+W^-$) for which a dedicated generator was used and WZ and ZZ processes where Pythia is used, since the MadGraph samples are mixed with $qq \rightarrow W^+W^-$ in a single VV sample. All generated events were passed to Pythia for fragmentation and hadronization before simulation and reconstruction in the CMS detector. For simulation, reconstruction and analysis of these samples the CMS software was used.

Process	Generator	cross-section (pb)
$qq \to WW$	madgraph	43.0
$gg \to WW \to 2\ell 2\nu$	gg2ww/pythia	0.153
$t\bar{t}$	madgraph	157.5
Single Top	madgraph	32.9
Z ($m_{\ell\ell} > 10$ GeV)	powheg/pythia	16677
W/Z+γ	madgraph	165.0
$W \to \ell\nu$	madgraph	-
WZ	pythia	18.2
ZZ	pythia	5.9
$gg \to H \to WW \to 2\ell 2\nu$	powheg/pythia	vary
qqH, $H \to WW \to 2\ell 2\nu$	powheg/pythia	vary

Table 7.1: Summary of Monte Carlo datasets used. The cross sections for a SM Higgs boson is taken from the LHC Higgs cross-section working group [15] and are Higgs mass dependent

	$gg \to H \to WW$			$qq \to H \to WW$		
m_H	$2\ell 2\nu$	$\ell\nu\tau\nu$	$2\tau 2\nu$	$2\ell 2\nu$	$\ell\nu\tau\nu$	$2\tau 2\nu$
120	11.1	11.1	2.77	0.84	0.84	0.21
130	20.11	20.11	5.03	1.63	1.63	0.41
140	28.55	28.55	7.14	2.45	2.45	0.61
150	34.31	34.31	8.58	3.10	3.10	0.78
160	38.51	38.51	9.63	3.69	3.69	0.92
170	34.94	34.94	8.73	3.64	3.64	0.91
180	29.39	29.39	7.35	3.21	3.21	0.8
190	21.71	21.71	5.43	2.51	2.51	0.63
200	18.22	18.22	4.55	2.18	2.18	0.54
250	10.86	10.86	2.71	1.39	1.39	0.35
300	7.81	7.81	1.95	0.96	0.96	0.24
350	7.19	7.19	1.8	0.67	0.67	0.17
400	5.51	5.51	1.38	0.44	0.44	0.11
450	3.52	3.52	0.88	0.32	0.32	0.08
500	2.20	2.20	0.55	0.24	0.24	0.06
550	1.38	1.38	0.35	0.19	0.19	0.05
600	0.87	0.87	0.22	0.15	0.15	0.04

Table 7.2: Production cross-sections multiplied by branching ratio for Higgs produced via vector-boson fusion and gluon fusion while decaying to W^+W^- with the W^+W^- system decaying to various final states (in fb).

For the majority of the backgrounds the most accurate cross-section available at NLO are taken from MCFM [92], while FEWZ [93] is used for W and Z cross-sections at the NNLO level. The program ggtoWW [94] is used to evaluate the cross-section of the gluon-induced W-boson pair production. For the signal samples, the NNLO cross-sections and

the branching-ratios are those provided by the LHC Higgs Cross Section Working Group in [15]. The signal cross section depending on the Higgs mass and the final state are given in Table 7.2. As we are probing the existence of the Higgs particle in a mass range m_H = 120 - 600 GeV, we need a simulation sample for each of the mass points in Table 7.2.

A few comments:

- in order to limit the size of the sample, the Drell Yan is generated only for di-lepton invariant masses $m_{\ell\ell} > 10$ GeV.
- the Single Top includes all production modes: tW, s and t channel.
- for Higgs production only the gluon fusion ($gg \rightarrow H$) and vector boson fusion ($qq \rightarrow H$) are used here. All the decays are included (e,μ and τ)

Higgs p_T reweighting

For the signal produced in gluon fusion mode, the total cross section was estimated at NNLO+NNLL accuracy. In addition, the p_T spectrum is taken from the code HqT introduced in Chapter 3, at NLO+NNLL accuracy. In Figure 7.1 the comparison of the p_T spectrum obtained from HqT and the Powheg MonteCarlo is shown. The latter produces a significantly harder spectrum. Some of the observables used in the analysis, the number of jets e.g., are sensitive to the p_T(H) spectrum (as the energy scale of the recoil jet(s) is determined by the Higgs momentum in the transverse plane)

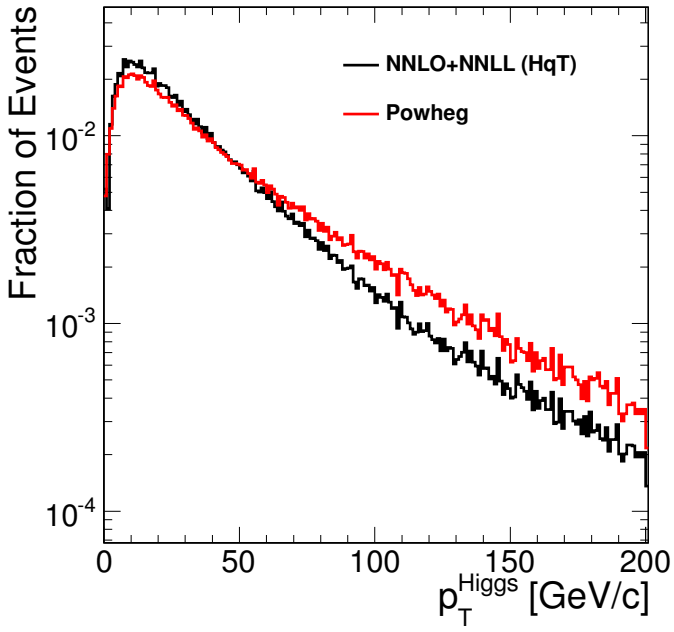


Figure 7.1: Higgs transverse momentum spectrum as predicted by Powheg and the NLO+NNLL calculation.

As the results obtained from HqT are more accurate, the events generated with Powheg (that are the ones ultimately used to guide the selection strategy) will be reweighted in order to match the p_T spectrum of HqT. This procedure assumes that within any given p_T bin the kinematics of the events are well described by the Powheg simulation.

7.1.2 Data Samples

The datasets used for this analysis correspond to the 2011 data taking period. They are organized in so-called primary datasets (PD) which are essentially fractions of the total datasets where only the interesting events are kept.

Given the e^+e^-, $\mu^+\mu^-$ and $e^\pm\mu^\mp$ final states considered in this analysis, the following four data samples have been used for the signal extraction. The DoubleElectron Primary Dataset (PD) serves the e^+e^-final state, while the DoubleMuon PD is used for the $\mu^+\mu^-$ channel. The cross-lepton MuEG (Muon-ElectronGamma) PD serves the $e^\pm\mu^\mp$ final states. Finally, extra events that do not fire the double lepton triggers, but fire the single muon trigger, are collected in the SingleMuon PD and are used for the modes containing a muon in the final state. We consider only the subset of runs which have passed all the data quality tests of the Physics Validation Team of CMS. The analyzed data corresponds to an integrated luminosity of 1.1 ± 0.1 fb^{-1}.

7.2 Signal selection

7.2.1 Introduction

In this selection we will explain how the signal extraction will be performed. The idea is to first apply an event selection that is independent of the the Higgs mass parameter m_H, aimed at substantially reducing all the background contributions that are not the WW irreducible background. The type of requirements that we are going to apply were already sketched in Chapter 5. They involve a selection on the leptons identification and isolation variables, and on their momenta, on the jets (such as the jet count and top-tagging discrimination). This will be the first part of the signal selection. We will refer to this selection as the common preselection, or the WW selection. As different backgrounds contribute differently to the 4 possible final states ee, $\mu\mu$, μe, and eμ the signal selection will be final state state dependent [1].

The second part will be aimed at specifically reducing the WW continuum and will be the subject of next chapter. In this phase the selection will be Higgs mass dependent, since the signal distributions that we a going to use to discriminate between the signal and the WW background change in shape for different Higgs masses. This selection block will be referred as "mass-dependent" (m.d.)

In short:

1. We select events that pass pre-defined lepton triggers.
2. We then select those events with two oppositely charged high p_T isolated leptons (ee, $\mu\mu$, eμ) requiring:
   - $p_T > 20$ GeV for the leading lepton (highest p_T);
   - $p_T > 10$ GeV for the trailing lepton (lowest p_T);
   - identification and isolation requirements on both leptons.
3. We apply a common W^+W^- preselection, which requires in brief:

[1] The final state eμ and μe are treated separately as the lepton momentum thresholds are different for electrons and muons, as we will explain later in this section. We use therefore the convention that when we speak about the final state $\ell_1\ell_2$, ℓ_1 is the most energetic lepton of the pair.

- require the number of reconstructed jets with a $p_T > 30$ GeV to be 0 or 1;
- exactly two high p_T leptons that are inconsistent with a Z or top decay;
- large transverse missing energy due to the neutrinos.

4. Finally, we perform two *Higgs mass dependent* event selections, both based on selecting a particular phase space region. One uses purely kinematical cuts, the other uses a more advanced multivariate technique (Neural Network)

The goal of the present study is to measure, or to derive an upper limit (in case of no excess with respect to background expectation) on the signal cross section for different possible Higgs masses. When we are probing a particular Higgs mass for a given final state, if the signal is produced accordingly to the standard model expectations, we expect, for a given integrated luminosity $\mathcal{L}$:

$$N_{\text{before-selection}}(m_H) = \sigma(m_H) \cdot BR(W^+W^- \to \ell\ell\nu\nu) \cdot \mathcal{L} \quad (7.1)$$

$$N_{\text{after-selection}}(m_H) = \sigma(m_H) \cdot BR(W^+W^- \to \ell\ell\nu\nu) \cdot \mathcal{L} \cdot \epsilon_{\text{tot}}(m_H) \quad (7.2)$$

where $\epsilon_{\text{tot}}(m_H) = \epsilon_{\text{common}} \cdot \epsilon_{\text{m.d.}}(m_H)$ is total signal selection efficiency. That includes the geometrical detector acceptance, the trigger, etc ... :

$$\epsilon_{\text{common}} = \epsilon_A \cdot \epsilon_{\text{trigger}} \cdot \epsilon_{\text{leptons}} \cdot \epsilon_{Z-\text{veto}} \cdot \epsilon_{\text{top-veto}} \cdots \quad (7.3)$$

A few remarks:

- Both in Equations 7.2 and 7.3 the selection is actually dependent on the final state and the number of reconstructed jets. The latter will help in dealing with top backgrounds rejection.
- Ideally, in order to minimize the dependence of the final number of expected signal events on the MonteCarlo simulation after the full selection, the selection efficiency needs to be, when possible estimated from data.
- The background estimates, when possible, will have to be determined from data.

We will now explain each step of the signal selection.

7.2.2 Triggers

Triggering on Higgs boson decays in the dilepton final state increases in difficulty with increasing instantaneous luminosity since the rates become high. Most of the requirements that will be applied on the leptons at the "analysis" stage need to be applied at the HLT level in order to keep the rates sustainable. A low mass Higgs preferably decays to low momentum leptons, but the rate of such leptons can be sensibly high at the LHC. Single lepton triggers can only be sustained with very tight identification and isolation requirements and large transverse momentum thresholds.

This means that double lepton triggers are the only viable option to maintain sensitivity to a low mass Higgs boson, where the leptons transverse momentum can be small. These dilepton triggers have a high efficiency to collect Higgs boson events

- **Double electron** The main dielectron triggers, require two HLT electron candidates with loose shower shape and calorimeter isolation requirements on both electrons and a match to a Level-1 seed for the leading electron. To be consistent with the signal selection of $p_T > 20, 10$ GeV for the leading and trailing electrons, $p_T > 17, 8$ GeV is required at the HLT level. Controlling the total trigger rate is most challenging in the dielectron channel, due to large fake electron background rates. Additional requirements must be added to the track-to-cluster matching and track isolation to control the total trigger rate at instantaneous luminosities above 1×10^{33} $cm^{-1}s^{-1}$.

  The identification and isolation requirements are described in Table A.1 in Appendix A. Because the electron HLT uses simplified algorithms compared to the offline selection, the HLT paths in simulation and in data do not always correspond exactly. We have chosen therefore to apply no HLT selection on simulation, and correct for the HLT efficiencies in data as measured in [90].

- **Dimuon** The main dimuon triggers require two HLT muon candidates with transverse momentum greater than 7 GeV and a match to a Level-1 seed is required for both muons. These are described in Table A.4 in Appendix A.

- **Mixed lepton** In the electron-muon channel we use two complementary triggers, which require both muon and electron HLT candidates. These are summarized in Table A.1 in Appendix A.

- **Single Muon** Finally, to recover any residual inefficiency, we also allow events that passed only the single electron or single isolated muon triggers summarized in Table A.1 in Appendix A.

The efficiencies of the HLT paths measured in data with are summarized in Appendix A. The efficiency is parameterized in bins of p_T and η in order to facilitate an event-by-event re-weighting procedure on the simulation. The double lepton triggers are highly efficient ($\geq 99\%$).

7.2.3 Primary Vertex

Vertex Selection

Selecting the vertex corresponding to the hard interaction is crucial in a high pile-up environment. In this analysis we use the primary vertices that are reconstructed with the Deterministic Annealing algorithm, introduced in Chapter 6 Section 6.2. With the DA algorithm, nearby interactions can be resolved and correctly reconstructed as two (or more) separate vertices, and, therefore, the position of the hard-scattering collision is determined with higher precision. This has a direct impact on observables that depend on the signal vertex position, like the impact parameter of tracks. The impact parameter (IP) of a track with respect to the primary vertex position is a quantity used to select signal leptons and reject backgrounds leptons, as will be discussed later. If the position of the signal interaction is incorrectly determined, the IP of the track of a genuinely prompt particle can be artificially large. This effect can be responsible for a significant drop in the efficiency of selecting signal prompt leptons.

We demand that the reconstructed primary vertices fulfill the following requirements:

- $d_z < 24$ cm (distance in the z-direction to the nominal detector center)

- $\rho < 2$ cm (radial position with respect to the beamspot).
- ndof ≥ 4 (number of degrees of freedom in the fitted vertex).

From the set of primary vertices in the event passing these selection cuts, the vertex with the largest summed squared-p_T of the associated tracks is chosen as the event primary vertex. Reconstructed leptons will be required to have small impact parameters with respect to this vertex.

Pile-up re-weighting

Due to the fast evolution of the LHC machine, with a rapid rise in the instantaneous luminosity, the data taking conditions have changed rapidly. In particular it is difficult to exactly reproduce the number of overlapping events (i.e. pileup) between data and simulation, and thus there are differences in the number of reconstructed primary vertices. Moreover, the pile-up distribution in simulation is flat. [2]

To make sure that the simulation reproduces the data, we have to re-weight the number of pile-up interactions in the simulation in order to match the distribution observed in the data. The problem is however that we cannot measure the pile-up distribution in data, since we only know the number of reconstructed vertices.

In order to infer the true pile-up distribution in data we can use a method called Bayesian unfolding [95], based on Bayes' theorem. We want to derive the true values (the causes) of pile-up in data N_{PU} $(i = 1, ..., n_{max})$ when knowing the effects, i.e. the number of reconstructed vertices N_{vtx} $(j = 1, ..., n_{max})$, that we can measure in data. [3].

We can then write a migration matrix that contains the conditional probabilities $P(N_{vtx} = j|N_{PU} = i)$ that a given number of pile-up interactions $N_{PU} = i$ will result in $N_{vtx} = j$ reconstructed vertices. This matrix is shown is Figure 7.2(a). Using Bayes' theorem it is possible to compute, under a certain prior probability distribution for the "cause" $P_0(N_{PU} = i)$ the conditional probability that N_{vtx} =j originates from N_{PU} =i.

$$P(N_{PU} = i|N_{vtx} = j) = \frac{P(N_{vtx} = j|N_{PU} = i)P_0(N_{PU} = i)}{\sum_{k=1}^{n_{max}} P(N_{vtx} = j|N_{PU} = k)P_0(N_{vtx} = j)}. \tag{7.4}$$

The $P(N_{PU} = i)$ distribution is then:

$$P(N_{PU} = i) = \frac{1}{N_{tot}} \sum_{j=1}^{n_{max}} n(N_{vtx} = j)P(N_{PU} = i|N_{vtx} = j) \tag{7.5}$$

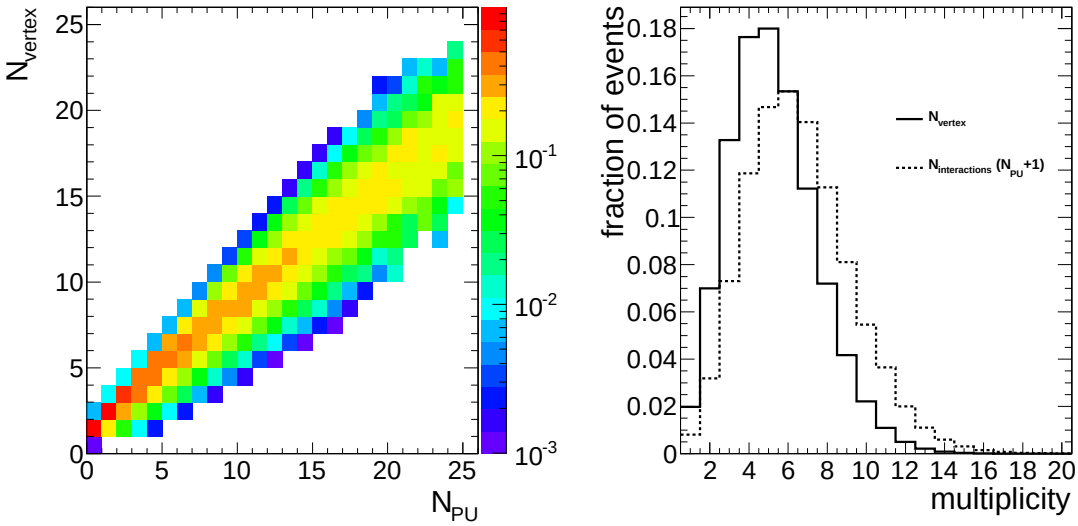
The advantage of this procedure is that it is iterative, one can then insert the newly calculated probability in Equation 7.5 into Equation 7.4 as a new prior P_1. The procedure finally converges when no significant change occurs in the final distribution anymore [4]. We can then re-weight our simulation in order to match the "unfolded" pile-up distribution in data. The procedure goes as follows:

[2] It is actually flat up to 10 interactions, with a poissonian tail from 10 to 20.

[3] Here we fix $n_{max} = 25$ as the maximum number of reconstructed vertices in data, compatibly with the observation in data.

[4] That is, when $P_N \sim P_{N-1}$ according to some χ^2 test for example.

- extract the migration matrix from simulation
- derive the true pile-up distribution in Data iteratively with Equations 7.4 and 7.5
- re-weight the simulation in order to match N_{PU} in data
- check that the N_{vtx} distributions match in data and simulation



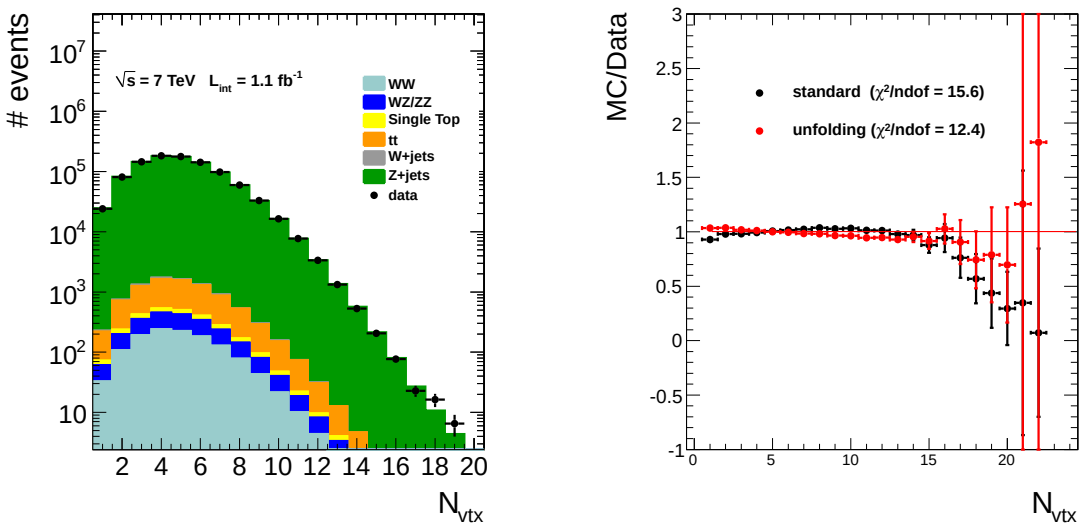
(a) (b)

Figure 7.2: (a) The migration matrix that serves as an input to the unfolding procedure (columns are normalized to 1) (b) Comparison between the number of reconstructed vertices in data and the true number of interactions.

The final number of interactions obtained with this method is shown in Figure 7.2(b), compared to the number of reconstructed vertices. The average value is 6.5 interactions per bunch crossing, while we observe on average 5.3 reconstructed primary vertices, leading to $\sim$ 80% efficiency. This can be well seen in Figure 6.1. The efficiency is much higher however for reconstructing the hard interaction vertex ($\sim$ 99.9 % from simulation).

In Figure 7.3(a) we show the number of reconstructed vertices in simulation and data, in order to check the consistency of the method and we see a good agreement. This distribution was produced at the level of the two lepton selection, that will be illustrated in the next section. In Figure 7.3(b) we show the ratio MC(reweighted)/Data as a function of the number of reconstructed vertices. This is done in order to check how well this unfolding procedure is able to reproduce the correct distribution of vertices in Data. We compare our method with the official CMS reweighting procedure [5]: in both cases the ratios are close to 1. By fitting the ratio to a flat function, we obtain a smaller χ^2/ndof with the unfolding procedure, indicating a better fit. Both methods give however very comparable results.

[5]The standard CMS procedure is based on directly predicting the average number of pile-up interactions per bunch crossing, given the instantaneous luminosity.



(a) (b)

Figure 7.3: (a) Distribution of the number of reconstructed vertices in simulation (after pile-up re-weighting) and data. (b) Ratio of the N_{vtx} distribution between simulation and data. The goodness of fit is given for the unfolding procedure (red) and the standard CMS reweighting procedure (black).

7.2.4 Lepton selection

Once the hard interaction is defined, we want to select the leptons that are most likely to originate from it. Since we want to keep most of the signal events and reject the background (mainly W+jets), we ask that the leptons fulfill a set of isolation and identification requirements, as well as some minimum transverse momentum.

Electrons

Electron Selection Since we allow the trailing electron leg to have low p_T threshold (in order to increase the signal efficiency), we need to make strict requirements on the identification variables in order to reduce the jet induced background. The specific requirements to select good prompt electrons are the following:

- Basic acceptance cuts are imposed, $p_T > 10$ GeV and $|\eta| < 2.5$;
- The following electron identification is applied to barrel (endcap) electrons when $p_T > 20$ GeV:
  - The lateral shower shape, $\sigma_{i\eta i\eta} < 0.01$ (0.03)
  - The track-cluster matching in the ϕ-direction, $\Delta\phi_{\mathrm{trk-SC}} < 0.06$ (0.03)
  - The track-cluster matching in the η-direction, $\Delta\eta_{\mathrm{trk-SC}} < 0.004$ (0.007)
  - The relative hadronic activity, H/E < 0.04 (0.1)
- The following electron identification is applied to barrel (endcap) electrons when $10 < p_T \leq 20$ GeV:

- The lateral shower shape, $\sigma_{i\eta i\eta} < 0.01$ (0.03)
- The track-cluster matching in the ϕ-direction, $\Delta\phi_{\mathrm{trk-SC}} < 0.03$ (0.02)
- The track-cluster matching in the η-direction, $\Delta\eta_{\mathrm{trk-SC}} < 0.004$ (0.005)
- The relative hadronic activity, H/E < 0.025 (0.1)
- The track bremsstrahlung fraction, fbrem > 0.15 OR ($|\eta| < 1.0$ AND $\mathrm{E_{SC}/p_{IN}} > 0.95$)

The identification variables above were defined in Section 6.5.1. The H/E requirement is looser in the endcap since the tighter values introduce noticeable signal reconstruction efficiency drop with large pileup.

Isolation requirements are then imposed by computing the particle flow isolation, defined as the scalar sum of the p_{T} of the particle flow candidates satisfying the following requirements (we call this variable $\mathrm{Iso_{PF}}$):

- $\Delta\mathrm{R} < 0.4$ to the electron in the $\eta \times \phi$ plane,
- for neutral hadron PF candidates, require that it is outside the footprint veto region of $\Delta\mathrm{R} < 0.07$,
- for photon and electron PF candidates, require that it is outside the footprint veto region of $|\Delta\eta| < 0.025$,
- $|\mathrm{d_z(PF\ Candidate)} - \mathrm{d_z(electron)}| < 0.1$ cm, if the PF candidate is charged,
- $p_{\mathrm{T}} > 1.0$ GeV, if the PF candidate is classified as a neutral hadron or a photon.

The second and third criteria are applied in order to make a distinction between the probed electron (that lies in the center of the cone and the particles that surround it). The fourth criterium requires that the particles that enter the isolation cone are attached to the primary vertex (hard interaction). This can be done only for charged particles. This is why we veto low energy neutrals, assuming that these mainly originate from pile-up interactions.

We require $\mathrm{relIso_{PF}} = \frac{\mathrm{Iso_{PF}}}{\mathrm{p_T}} < 0.13$ (0.09) for electrons in the barrel (endcap). Additional details of the algorithm are given in 6.5.3.

In order to veto fake electrons from converted photons, we look for a reconstructed conversion vertex where one of the two tracks is compatible with the electron. The vertex fit probability is required to be greater than 10^{-6}. We then require that there are no missing expected hits forming the electron track (this requirement reduces substantially the $\mathrm{W/Z}\gamma$ backgrounds).

Finally to reduce fake electrons from non-prompt sources, we require the transverse and longitudinal impact parameters with respect to the primary vertex to be less than 0.02 and 0.1 cm respectively.

Electron efficiencies The electron selection efficiency can be factorised into two contributions, the efficiency from the electron reconstruction and from the additional analysis selections (acceptance, trigger, identification and isolation).

The electron reconstruction efficiency is defined as the efficiency for a supercluster to be matched to a reconstructed ECAL driven electron (see Section 6.5.1). This was measured and is known to give perfect agreement between data and simulation [96].

The electron selection efficiency was measured in [90]. The values for the Trigger efficiencies are summarized in Appendix A. The resulting data to simulation scale factors, given in Table 7.3 show that the electron efficiency is slightly overestimated in simulation (by approximately 5%, depending on the momentum and pseudo-rapidity range).

Measurement	Barrel ($\|\eta\| < 1.479$)	Endcap ($\|\eta\| > 1.479$)
$10 < p_T < 15$	0.93 ± 0.06	0.86 ± 0.05
$15 < p_T < 20$	0.92 ± 0.01	0.92 ± 0.03
$p_T > 20$	0.96 ± 0.00	0.96 ± 0.00

Table 7.3: Data to simulation scale factors for electrons parameterized as a function of the electron transverse momentum and pseudo-rapidity.

Muons

Muon Selection Most good quality muons are reconstructed as both types at the same time (the two muon classes were defined in Section 6.5.2) and the momentum resolution is dominated by the inner tracker system up to about 200 GeV in transverse momentum. We require the muon to be reconstructed as GlobalMuon, with $\chi^2/\text{ndof} < 10$ on the global fit, with at least one good muon hit, and at least two matches to muon segments in different muon stations. In case of a TrackerMuon it must satisfy the "Tracker Muon Last Station" defined in Chapter 6. This allows the recovery of events in which the two muons are close by in the spectrometer, for which the *global muon* reconstruction is known to be inefficient [97]

In addition, the following specific requirements to select good prompt isolated muons are the following:

- transverse momentum p_T greater than 10 GeV;
- more than 10 hits in the inner tracker;
- at least one pixel hit;
- impact parameter in the transverse plane $|d_0| < 0.02$ (0.01) cm for muons with p_T greater (smaller) than 20 GeV, calculated with respect to the primary vertex;
- longitudinal impact parameter $|d_z| < 0.1$ cm, calculated with respect to the primary vertex;
- pseudo-rapidity $|\eta|$ must be smaller than 2.4;
- relative p_T resolution is better than 10%.

We require this last point also in order to increase the resolution in the missing energy (see next section).

Furthermore, the particle flow candidate-based isolation variable is used to reduce the contamination from the non-isolated muons originating from jets.

- Iso_{PF}: defined as the scalar sum of the p_{T} of the particle flow candidates satisfying the following requirements:
  - $\Delta\text{R} < 0.3$ to the muon in the $\eta \times \phi$ plane,
  - $|\text{d}_\text{z}(\text{PFCandidate}) - \text{d}_\text{z}(\text{muon})| < 0.1$ cm, if the PF candidate is charged,
  - $p_{\text{T}} > 1.0$ GeV, if the PF candidate is classified as a neutral hadron or a photon.

We require $\frac{\text{Iso}_{\text{PF}}}{\text{p}_\text{T}} < 0.13$ (0.06) for muons in the barrel with p_T greater (smaller) than 20 GeV. For muons in the endcap, we require $\frac{\text{Iso}_{\text{PF}}}{\text{p}_\text{T}} < 0.09$ (0.05) for muons with p_T greater (smaller) than 20 GeV.

Measurement	Barrel ($\lvert\eta\rvert < 1.479$)	Endcap ($\lvert\eta\rvert > 1.479$)
$10 < \text{p}_\text{T} < 15$	0.93 ± 0.02	0.95 ± 0.02
$15 < \text{p}_\text{T} < 20$	0.96 ± 0.01	0.93 ± 0.01
$\text{p}_\text{T} > 20$	1.00 ± 0.00	0.98 ± 0.00

Table 7.4: Data to simulation scale factor for muons parameterized as a function of the muon transverse momentum and pseudo-rapidity.

Muon Efficiencies As for the electron case the muon reconstruction efficiency is known to be close to 1. Therefore we give here only the selection efficiency specific to the analysis, as calculated in [90]. The data to simulation scale factors are summarized in Table 7.4.

Lepton Pairs

The electrons and muons that pass the previous selection in a given event, become candidates to form a lepton pair. This object is a pair of leptons that satisfy the following requirements:

- the two leptons must have an opposite charge
- at least one of the two must have a transverse momentum p_T greater than 20 GeV (both already have $\text{p}_\text{T} > 10$ GeV)

Signal events from simulation are re-weighted by the scale factor defined in Tables 7.4 and 7.3. We call the criteria applied up this point the "Lepton Selection".

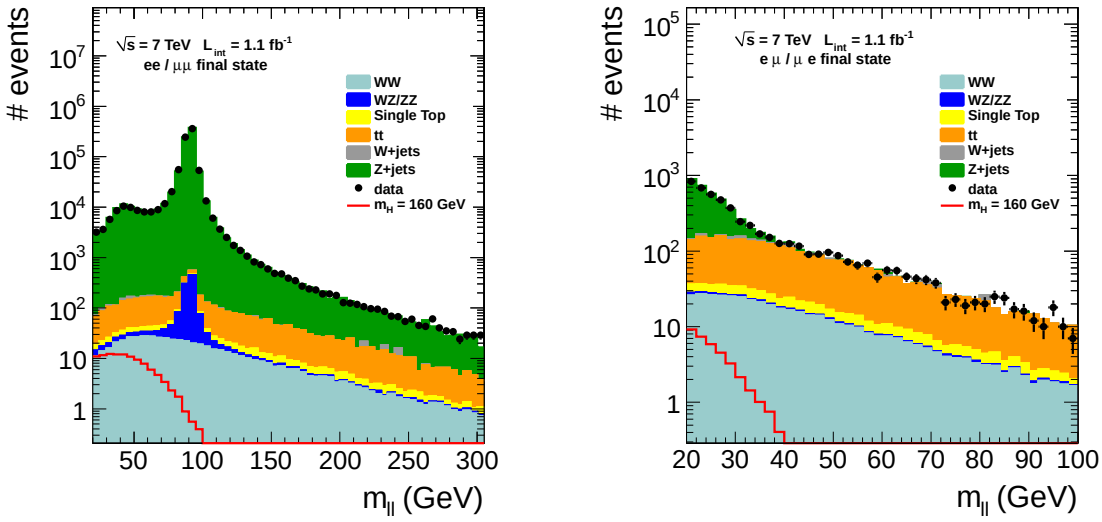
7.2.5 Extra-lepton rejection

To further reduce the background from diboson processes, we veto events containing an additional lepton meeting the previously described selection requirements (in Section 7.2.4) with $p_T > 10$ GeV. This removes $\sim 60\%$ of the WZ component but only $\sim 10\%$ on the ZZ component, which is dominated by $ZZ \to 2l2\nu$ decays after the full event selection and surviving the Z veto. The efficiency of this particular selection for $WW \to 2l2\nu$ events is $\sim 99.9\%$.

7.2.6 The invariant mass of the leptons

To reduce the Drell–Yan background in the e^+e^- and $\mu^+\mu^-$ final states, we reject events with a dilepton invariant mass within 15 GeV of the Z, where the invariant mass is defined by

$$m_{\ell\ell} = \sqrt{(p_{\ell_1} + p_{\ell_2})^2} \tag{7.6}$$



(a) (b)

Figure 7.4: Invariant mass distribution of the two selected leptons at the lepton selection level in the same-flavor channel ee/$\mu\mu$ (a) and opposite flavor channel eμ/μe (b)

To motivate this cut we show in Figure 7.4 the lepton invariant mass (for the same and opposite flavor final states).

We also reject events with a dilepton invariant mass below 12 GeV to suppress contributions from low mass resonances. Unfortunately no simulation sample was available for these low resonances: they are not included in Figure 7.4. Their contribution is however known to be negligible.

7.2.7 Missing Transverse Energy

The missing transverse energy is used to reject background events where there is no natural source of missing energy, like in Drell-Yan and QCD events. In the $Z/\gamma^* \to \tau\tau$ process there is a large difference in the masses of τ and Z. The taus are produced with large boost and their decay products, including neutrinos, are aligned with the leptons. Therefore a transverse component of missing energy with respect to the leptons is a better measure of true missing energy in the event. In order to reject such background events with a small opening angle between E_T^{miss} and one of the leptons, we used the projected E_T^{miss} for event selection, defined as:

$$\text{proj.}E_T^{miss} = \begin{cases} E_T^{miss} & \text{if } \Delta\phi_{min} > \frac{\pi}{2} \\ E_T^{miss}\sin(\Delta\phi_{min}) & \text{if } \Delta\phi_{min} < \frac{\pi}{2} \end{cases} \tag{7.7}$$

$$\text{with } \Delta\phi_{min} = \min(\Delta\phi(\ell_1, E_T^{miss}), \Delta\phi(\ell_2, E_T^{miss})) \tag{7.8}$$

where $\Delta\phi(\ell_i, E_T^{miss})$ is the angle between E_T^{miss} and lepton i in the transverse plane.

As mentioned in Chapter 6 Section 6.4 in the presence of high pile-up, the tail of the E_T^{miss} distribution in $Z/\gamma^* \to \ell\ell$ events increases significantly.

To suppress the dependence of the E_T^{miss} selection on the amount of pile-up, an algorithm was developed [90] called trkE_T^{miss}(track missing transverse energy), constructed from charged particles originating from the primary vertex.

The event trk.E_T^{miss} is defined as

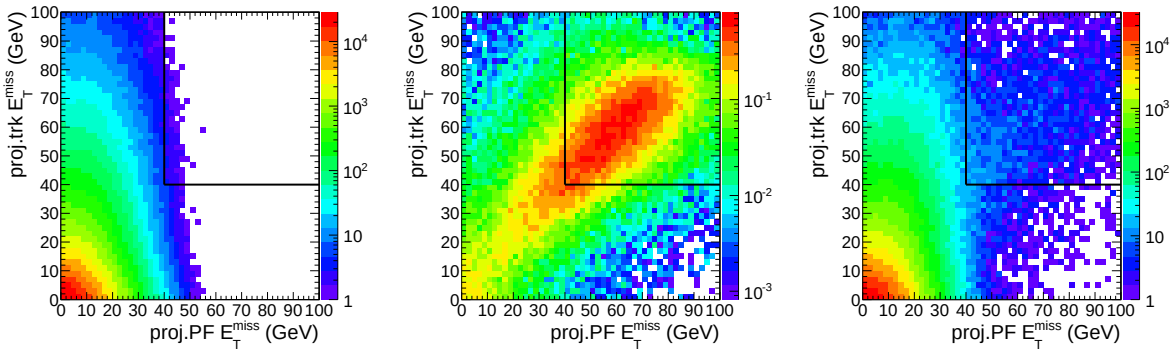
$$\text{trk.}E_T^{miss} = -\overrightarrow{p_T}(l_1) - \overrightarrow{p_T}(l_2) - \sum_i \overrightarrow{p_T}(i) \tag{7.9}$$

where $\overrightarrow{p_T}(l_1)$ and $\overrightarrow{p_T}(l_2)$ are the transverse momentum vectors of the two leptons passing the lepton selections described in Section 7.2.4, and $\overrightarrow{p_T}(i)$ represents the transverse momentum vectors of the charged Particle Flow candidates satisfying the following requirements:

- the track matched to PFCandidate has $\Delta z < 0.1$ cm with respect to the signal primary vertex;
- the track has $\Delta R > 0.1$ with respect to both leptons, to avoid double-counting of the leptons.

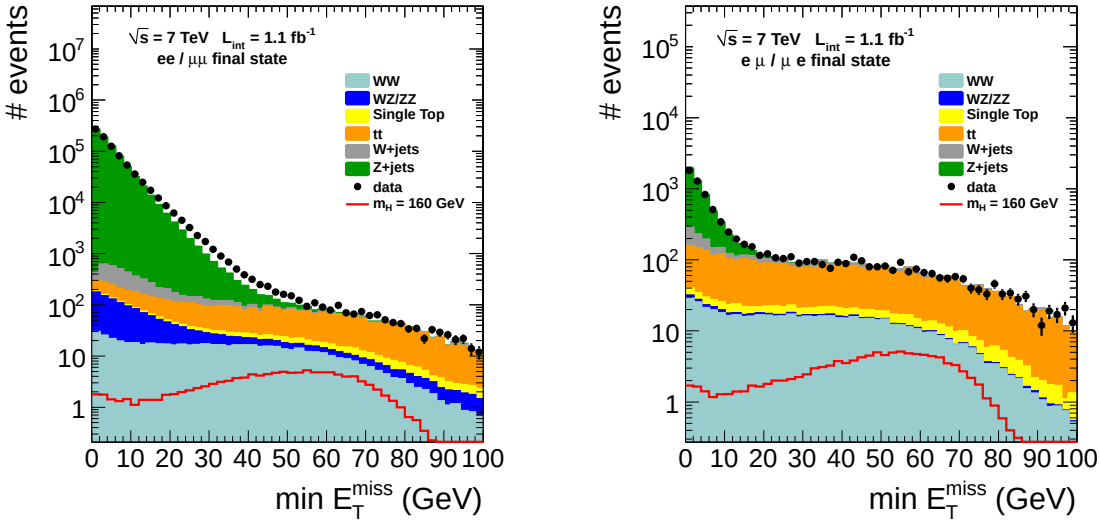
Compared to the projected Particle Flow E_T^{miss} (proj.PFE_T^{miss}), we observe that the projected track E_T^{miss} (proj.trk.E_T^{miss}) has a slightly larger tail in $Z/\gamma^* \to \ell\ell$ background events. However these two E_T^{miss} values are weakly-correlated in backgrounds with no genuine E_T^{miss}, and strongly correlated for the signal processes with genuine E_T^{miss}, as shown in Figure 7.5. Moreover, for the signal both variables peak at high values. Therefore the signal over background ratio can be improved if we select the events with the mininum of these two projected E_T^{miss} values, min E_T^{miss}= min(proj.trk E_T^{miss}, proj.PF E_T^{miss}) greater than a given threshold.

In Figure 7.6 we show the minE_T^{miss} variable at the lepton selection level for the same-flavor ($ee/\mu\mu$) and opposite flavor final states ($e\mu$). In the same flavor final state we observe a significant discrepancy at large E_T^{miss} values (where the signal peaks) between



(a) (b) (c)

Figure 7.5: Distributions of proj.PFE_T^{miss} vs. proj.trkE_T^{miss} in simulation Drell-Yan (a), Signal (m_H = 160 GeV) (b), and Data (c). The figures were obtained by omitting the Z-veto selection of Section 7.2.6



(a) (b)

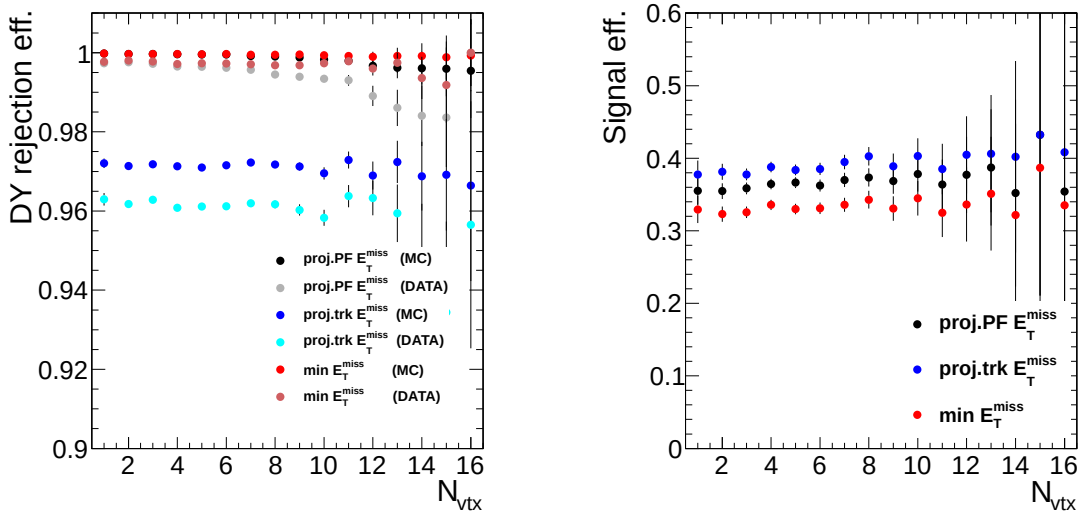
Figure 7.6: Distribution of the minE_T^{miss} variable at the lepton selection level for the same-flavor ($ee/\mu\mu$) and opposite flavor final states ($e\mu/\mu e$)

data and simulation. We believe that this difference is due to poor out-of-time pile-up (and hence E_T^{miss}) modeling in simulation which can lead to an excess of Drell-Yan events with high E_T^{miss} that are not correctly predicted by simulation. This stresses the importance of determining the Drell-Yan background from fully data-driven techniques, in order to reduce the dependency on the simulation (see the final discussion in Chapter 8).

We apply different selection requirements between $ee/\mu\mu$ and $e\mu/\mu e$ final states since Drell-Yan mostly contributes to ee and $\mu\mu$ channels. The selection requirements are:

- minE_T^{miss} > 20 GeV for opposite flavor

- minE_T^{miss} > 40 GeV for same flavor



(a) (b)

Figure 7.7: (a) Rejection efficiency of the Drell-Yan background after the selection MET ≥ 40 GeV for same flavor events with MET = proj.PFE_T^{miss} (black), proj.trk.E_T^{miss}, (blue), min E_T^{miss} (red) in data and simulation as a function of the number of reconstructed vertices in the event. In order to increase the statistics, the Z-veto selection of Section 7.2.6 was inverted here, and only same flavor events were used. (b) Signal selection efficiency after the same MET selection in simulation.

This variable also helps in designing a selection criterium that is independent on the amount of pile-up in the event. This is shown in Figures 7.7(a) and (b). We see that the proj.trk E_T^{miss} gives a stable efficiency versus pile-up. On the other hand, due the its large tail it performs worse than proj.PFE_T^{miss} in terms of Drell-Yan rejection efficiency. By combining these two variables in the min E_T^{miss} we can achieve the stability vs. pile-up offered by the proj.trk E_T^{miss} and the better Drell-Yan rejection obtained with the proj.PFE_T^{miss}.

The efficiency on the selection that was just exposed is believed to be well described by simulation for the signal events, that display a real E_T^{miss}. In the background data estimate section we will see however that we cannot rely on simulation for estimating the Drell-Yan contribution.

7.2.8 Jet Counting

Jet Definition

Jets in this analysis are reconstructed by combining the energy measured in the calorimeters and tracks from charged particles on basis of the standard CMS particle flow algorithm and using the Anti-k_T clustering algorithm with R = 0.5. Events will be studied in two classes: one with no counted jets and one with exactly one counted jet within $|\eta| < 5$. In the zero jet bin, the so-called jet veto aims at reducing the contribution of the $t\bar{t}$ background with

one or more jets above a certain p_T threshold. In this analysis we use high p_T jets (counted jets) to define the analysis jet bin and low p_T jets for b-tagging (see later). We define:

- counted jet: a reconstructed jets with $p_T > 30$ GeV within $|\eta| < 5.0$;
- low p_T jet: a reconstructed jets with $7 < p_T < 30$ GeV within $|\eta| < 5.0$

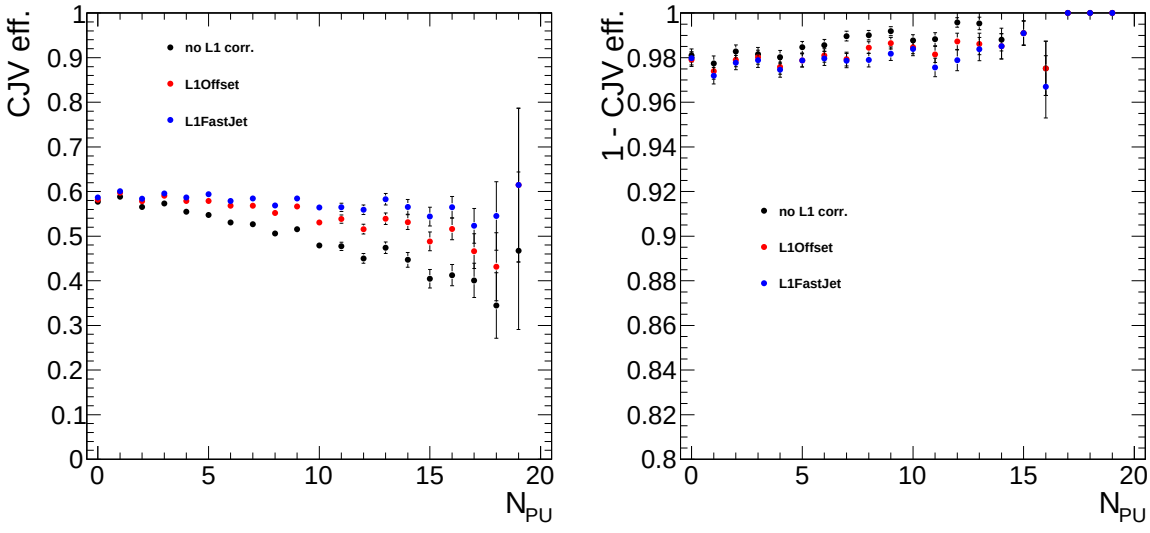
We analyze the events separately based on the number of counted jets in the event. In this analysis we treat only the 0 and 1 jet bin cases.

Pile-up subtraction

With the presence of pile-up, the momenta of the jets are overestimated due to the contamination of particles originating from secondary interactions, leading to a lower jet veto efficiency for both the Higgs signal and the $t\bar{t}$ background in the zero jet bin as well as extra contaminations in the one jet bin.

The two methods available in CMS that correct for the pile-up contribution to the jet energy scale were examined already in Chapter 6: L1Offset and and L1Fastjet. The standard L2 and L3 jet energy scale factors are applied on top of these L1 corrections.

As for the missing energy, due to the large variations in the number of pile-up interactions (N_{PU}) per run and even per lumi-section, we require the methods to give stable results with respect to number of simultaneous interactions. In Figure 6.10, we had already shown that the L1 FastJet method appears to be more stable as the number of pile-up interactions (N_{PU}) increases while a small fraction of the needed correction seemed to be lost with L1Offset at high N_{PU}.



(a) (b)

Figure 7.8: Efficiency of (a) selecting the signal, (b) rejecting the $t\bar{t}$ background when events with 0 jets with a $p_T > 30$ GeV are selected.

Here we look at the L1 correction robustness against pile-up is studied in Monte Carlo for the Higgs and $t\bar{t}$ processes. In Figure 7.8 the signal selection efficiency (a) and the

background rejection efficiency (b) are shown in simulation when 0 jets are found above 30 GeV as a function of the number of reconstructed primary vertices. We see that for signal (background) selection (rejection) the efficiency is flat with respect to pile-up when the L1 Fastjet correction is applied, while a fraction of the efficiency is lost when L1Offset is applied. This confirms what we saw in Figure 6.10.

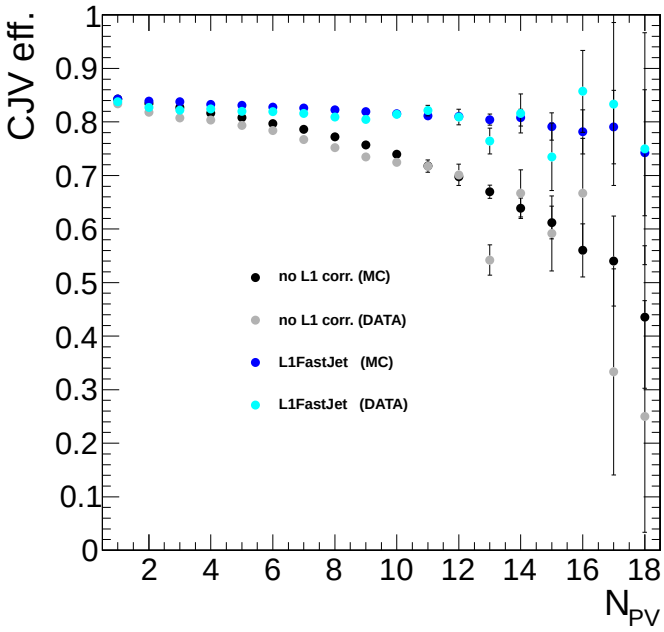


Figure 7.9: Efficiency of (a) selecting Drell-Yan events when events with 0 jets with a p_T > 30 GeV are selected.

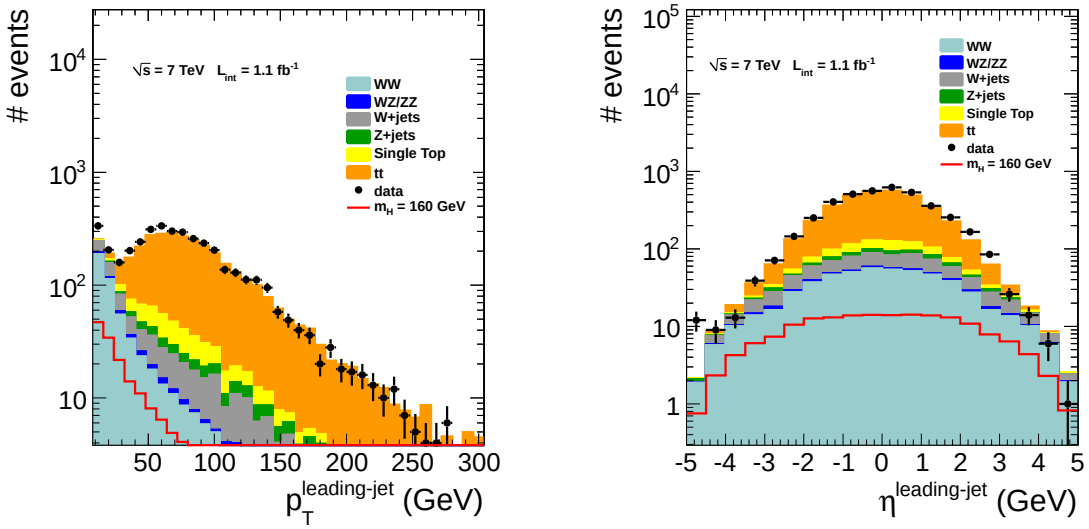
We have also tested the performance of the L1 FastJet algorithm on data. Z+jets events (playing the role of the signal here) were selected by requiring two opposite sign, same flavor, well isolated and identified electrons or muons with an invariant mass consistent with the Z mass ($|m_{\ell\ell} - m_Z| < 15$ GeV) and the primary vertex. In Figure 7.9 we show the jet veto efficiency (0 jet selection) on Drell-Yan. This is not meant to be a measurement of the 0-jet selection efficiency for the signal, as the jet structure of gg → H and qq → Z processes are very different. However, we can:

- validate the method on Data, as the L1 Fastjet method is stable versus pile-up.
- deduce that the simulation efficiency for L1 Fastjet corrected jets is well described by simulation (as there is almost perfect agreement)

Given the second point, we can confidently take as the efficiency of selecting 0 and 1 jet events for signal events the value given by simulation. To confirm the good agreement between data and simulation in Figure 7.10 we show the p_T and η distribution of the most energetic jet in the event (with $p_T > 7$ GeV).

7.2.9 Jet Lepton Angle

We require the angle in the transverse plane between the dilepton system and the most energetic jet with $p_T^{jet} > 15$ GeV to be smaller than 165 degrees in the ee/$\mu\mu$ final states.



(a) (b)

Figure 7.10: p_T (a) and η (b) distribution of the most energetic jet in the event (with p_T > 7 GeV)

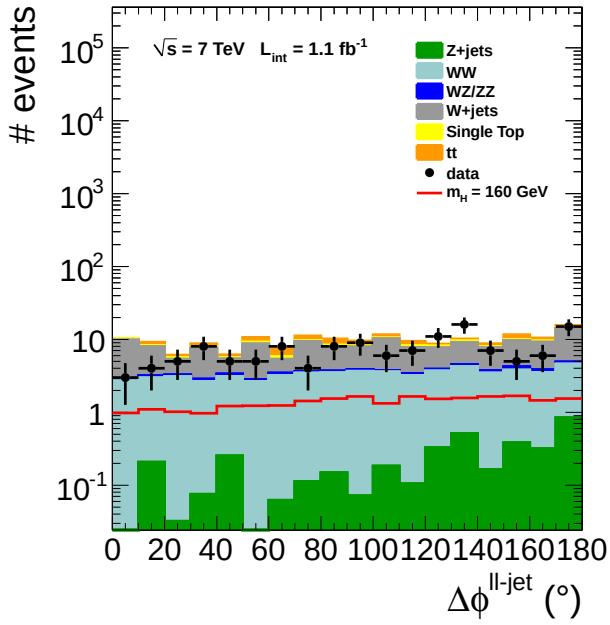


Figure 7.11: Opening angle distribution between the di-lepton system (in the same flavor case) and the leading jet (p_T^{jet} > 15 GeV)

This requirement rejects $Z/\gamma^* \rightarrow \ell\ell$+1jet events, where the Z boson recoils against a jet (see Figure 7.11).

7.2.10 Top tagging

At this point of the selection most backgrounds have been drastically suppressed. The final step of the WW selection, before addressing the WW rejection is the top rejection.

Top backgrounds pose a significant challenge, because the production cross-section is substantially higher than the W^+W^- cross-section. To reduce the top background, we use two top tagging methods. Both methods rely on the fact that top quarks decay to Wb with almost 100% probability.

Soft Muon rejection The first method vetoes events containing soft muons from the b-quark decays. We define soft muons as:

- $p_T > 3$ GeV;
- reconstructed as a TrackerMuon;
- meets Tracker Muon Last Station muon identification requirements;
- number of valid inner tracker hits > 10;
- transverse impact parameter with respect to the Primary Vertex, $|d_0| < 0.2$ cm;
- longitudinal impact parameter with respect to the Primary Vertex $|d_z| < 0.1$ cm;
- non-isolated $\mathrm{relIso_{PF}} > 0.1$ if $p_T > 20$GeV.

b-tagging The second method uses Track Counting High Efficiency (TCHE) [91] b-jet tagging. In this method, events containing low p_T jets tagged with the TCHE algorithm with a discriminator value of greater than 2.1 are vetoed. The algorithm is applied also in the case of zero counted jets, which can still contain low p_T tagged jets.

The Higgs signal efficiency versus the $t\bar{t}$ efficiency in events with no counted jets according for different standard b-tagging algorithms is shown in Figure 7.12. For a top selection efficiency below 60%, the TCHE (red points) tagger performs better than the others. Our current cut value has a signal efficiency of about 98.1% and a $t\bar{t}$ efficiency of about 50%.

As we will in see in the Top background data estimate section, the b-tagging efficiencies are compatible between data and simulation. In the 0-jet bin analysis we require that no jets (low p_T as defined in Section 7.2.8) are b-tagged and that no soft muons are found. In the 1-jet bin we require that the hard jet is not-bagged and also ask for no extra soft-muons in the event.

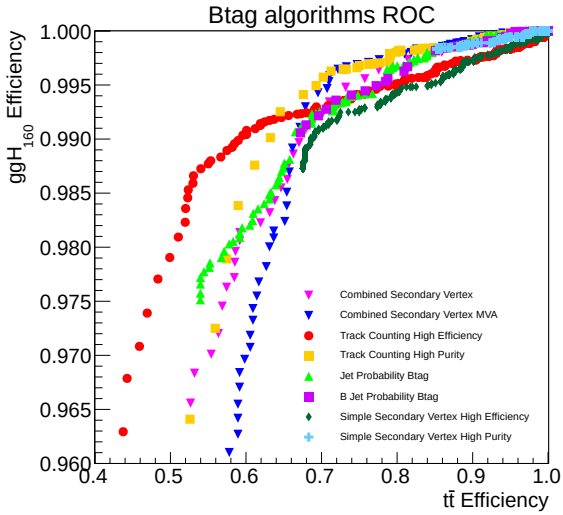


Figure 7.12: Higgs signal efficiency versus $t\bar{t}$ efficiency in events with no counted jets for different standard b-tagging algorithms.

	Z→ $\ell\ell$	Z→ $\tau\tau$	$t\bar{t}$	Single Top	W+jets	WZ/ZZ	WW	W/Z+γ	H_{140}
$\mu\mu$	2.5 ± 0.8	0.0 ± 0.0	5.1 ± 1.0	2.4 ± 0.3	5.8 ± 4.0	4.3 ± 0.2	70.0 ± 0.8	0.1 ± 0.1	14.5 ± 0.3
μe	1.0 ± 0.3	0.7 ± 0.3	6.9 ± 1.1	4.0 ± 0.4	26.5 ± 9.3	3.0 ± 0.2	120.6 ± 1.0	5.2 ± 1.1	17.8 ± 0.3
eμ	0.6 ± 0.2	1.1 ± 0.3	8.6 ± 1.4	4.2 ± 0.4	29.5 ± 10.1	3.8 ± 0.2	131.9 ± 1.1	4.1 ± 1.0	20.6 ± 0.3
ee	1.6 ± 0.8	0.0 ± 0.0	3.0 ± 0.8	2.2 ± 0.3	15.2 ± 7.4	2.4 ± 0.1	43.5 ± 0.6	0.4 ± 0.3	8.9 ± 0.2
all	5.7 ± 1.1	1.8 ± 0.5	23.7 ± 2.2	12.8 ± 0.7	77.0 ± 16.1	13.5 ± 0.4	366.0 ± 1.8	9.8 ± 1.5	61.7 ± 0.5

Table 7.5: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the WW common selection in the 0 jet bin

	Z→ $\ell\ell$	Z→ $\tau\tau$	$t\bar{t}$	Single Top	W+jets	WZ/ZZ	WW	W/Z+γ	H_{140}
$\mu\mu$	1.5 ± 0.3	0.0 ± 0.0	14.0 ± 1.7	4.6 ± 0.4	0.0 ± 0.0	1.5 ± 0.1	22.1 ± 0.4	0.0 ± 0.0	5.2 ± 0.1
μe	2.4 ± 0.5	4.8 ± 0.7	24.7 ± 2.2	6.9 ± 0.5	2.3 ± 1.7	2.5 ± 0.2	36.6 ± 0.6	0.9 ± 0.4	7.6 ± 0.2
eμ	0.7 ± 0.2	6.6 ± 0.9	29.5 ± 2.5	8.5 ± 0.5	6.5 ± 4.9	3.9 ± 0.2	41.1 ± 0.6	0.8 ± 0.3	8.0 ± 0.2
ee	1.5 ± 0.3	0.0 ± 0.0	10.2 ± 1.5	2.6 ± 0.3	6.0 ± 4.7	1.6 ± 0.1	13.5 ± 0.3	0.1 ± 0.1	3.1 ± 0.1
all	6.1 ± 0.7	11.4 ± 1.2	78.4 ± 4.0	22.6 ± 0.9	14.7 ± 7.0	9.5 ± 0.3	113.3 ± 1.0	1.8 ± 0.5	23.9 ± 0.3

Table 7.6: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the WW common selection in the 1 jet bin

7.2.11 Summary of the WW common preselection

Here we summarize the expectation of the signal and background at WW selection selection level.

In Appendix B we show the background and signal (for some Higgs masses) expectations and the data yields at each step of the WW common preselection. In Table 7.5 and 7.6 the final yields at the WW preselection are given for each final state. At this point the backgrounds are purely simulation expectations, while the signal has been corrected for the data estimated efficiencies (essentially trigger and lepton selection efficiencies). The first number given corresponds to the expected number of events for an integrated luminosity of 1.1 ± 0.1 fb^{-1}. The second entry is the statistical uncertainty due to limitations on the size of the simulation samples.

7.3 Data driven background estimates at the WW selection level

7.3.1 Introduction

In order to have a reliable estimate of the number of Signal events after the 0 and 1-jet bin selections introduced in the previous sections we need to assess how many background events are still present in the sample. A naive approach would simply consist in counting how many background events survive after the selection cuts that define the Signal region directly from MonteCarlo simulation. This approach presents however several limitations. Consider that we want to measure the background B in the Signal region. Then, if we rely on simulation, the number of B events left after the signal selection can be written as:

$$N_B^{sr} = \sigma_B \mathcal{L} \prod_i \epsilon_i \tag{7.10}$$

where N_B^{sr} is the number of backgrounds left after the Signal selection, σ_B is the total cross section of the background B, $\mathcal{L}$ is the integrated luminosity and the ϵ_i are the efficiencies of selecting this particular background at each step i of the selection (this includes also the geometrical acceptance) inferred directly from simulation. The estimate of N_B^{sr} using the previous formula suffers from the uncertainty on each parameter in the right-hand side of the equation: the cross section, the luminosity and each individual selection efficiency described in the previous section, such as the triggers efficiencies, the leptons selection, isolation and identification, the E_T^{miss} resolution, the top-tagging efficiencies, etc ... Some of these can be reliably extracted from simulation, but some others cannot because simulation is known to not be particularly reliable. Such is the case for b-tagging efficiencies for instance, where the details of b-jet fragmentation cannot be reliably simulated at low energy. Another example is the E_T^{miss} distribution with the presence of pile-up interactions. In addition, given that the total selection was designed in the first place such that the background yield is small, the statistical uncertainty on N_B^{sr} is expected to be important. Moreover the systematic uncertainties on some of the efficiencies simply cannot be estimated by simulation. So given the complexity of the signal selection, and the low background yield in the signal region, the total resulting uncertainty on N_B^{sr} can be very significant, if estimated via Equation 7.10.

An alternative method needs therefore to be defined. Before any selection the number of B events is N_B^0. The standard selection "sr" (the signal region) enhances the Signal yields and reduces strongly the background and the corresponding efficiency is ϵ^{sr}. An alternative selection "cr" (control region) can also be defined so that the background under study is kept high but the other backgrounds and the signal are significantly reduced. We can always write:

$$N_B^0 = \frac{N_B^{sr}}{\epsilon^{sr}} = \frac{N_B^{cr}}{\epsilon^{cr}} \tag{7.11}$$

where the ϵ^{sr} and ϵ^{cr} are the total efficiencies of respectively selecting the background in the signal and the control region and N_B^{sr} and N_B^{cr} are the background yield in the signal and control regions. Relying on the good knowledge of the considered background in the control phase-space region, the number of background events measured in the control region can then extrapolated to the signal region using the selection efficiencies obtained by MonteCarlo simulation, or if possible, measured from Data:

$$N_B^{sr} = \frac{\epsilon^{sr}}{\epsilon^{cr}} N_B^{cr} \tag{7.12}$$

Ideally, if we know that the control region is populated only by one type of background events, it is enough to count the number of events in the control region N_{all}^{cr}, and use Equation 7.12 to estimate the background yield in the signal region. In practice however it is not always possible to define a control region in which the background purity is 100%, therefore the other backgrounds need to be subtracted from N_{all}^{cr}. Finally, the number of events B in the signal region can be written as:

$$N_B^{sr} = \frac{\epsilon^{sr}}{\epsilon^{cr}} \left(N_{all}^{cr} - N_{other}^{cr} \right) \tag{7.13}$$

This approach is attractive mainly for two reasons:

- If the control region is properly chosen the event count of the other backgrounds N_{other}^{cr} is low, while N^{cr} is high. The statistical uncertainty on the number of observed events N^{cr} is therefore low. On the other hand N_{other}^{cr} can have a big uncertainty, especially if these events are counted directly from MonteCarlo simulation. However, since the control region is chosen such that N_{other}^{cr} is low the impact on the total uncertainty will not be significant.
- In practice the control region is chosen such that it differs from the signal region by only one specific selection requirement. This implies that all the common selection cuts cancel out in the ratio and the uncertainties on quantities involved in the common selection can be neglected. So one only needs to consider the uncertainties on this ratio whereas in the MonteCarlo counting method (Equation 7.10) one had to estimate the impact of all kinematic cuts on the final yield.

In the following section we will apply this method to the measurement of the main backgrounds of the $H \rightarrow W^+W^-$ analysis at the common preselection level.

7.3.2 The Drell-Yan background

In this section we will determine the background that originates from a decay of a Z boson into a pair of same flavor leptons ($e^{\pm}e^{\mp}$ and $\mu^{\pm}\mu^{\mp}$): the Drell-Yann Z /$\gamma^{\star}$. In theory since no neutrinos are present in the final-state, we expect the contribution of this background to be negligible in the signal region. However, in practice at high luminosity, with the presence of several simultaneous interactions per bunch crossing (pile-up), the E_T^{miss} resolution gets smeared toward high values, resulting in a non-negligible tail in the E_T^{miss} distribution of the Drell-Yan [6]. Moreover, the Z boson decaying to a pair of τ leptons, and both the taus decaying leptonically produces non-zero E_T^{miss} in the event.

In principle this background could simply be inferred in the signal region by counting the number of same flavor lepton $N(ll)^{sr}$ and subtracting the contamination from other backgrounds. The contamination can be classified in two categories:

- The backgrounds that equally decay into same-flavor and opposite-flavor final state such as the top backgrounds, W^+W^-, and W+jets. They can be determined by simply counting the opposite flavor events with a proper factor that takes into account the difference in efficiency in selecting electrons and muons (k_{ee} and $k_{\mu\mu}$).
- The backgrounds that decay mostly into a same flavor state via a Z boson resonance such as the Diboson backgrounds (WZ,ZZ). As these backgrounds are both characterized by the presence of real E_T^{miss} we assume that they are well modeled by simulation. Their contamination will therefore be calculated from simulation.

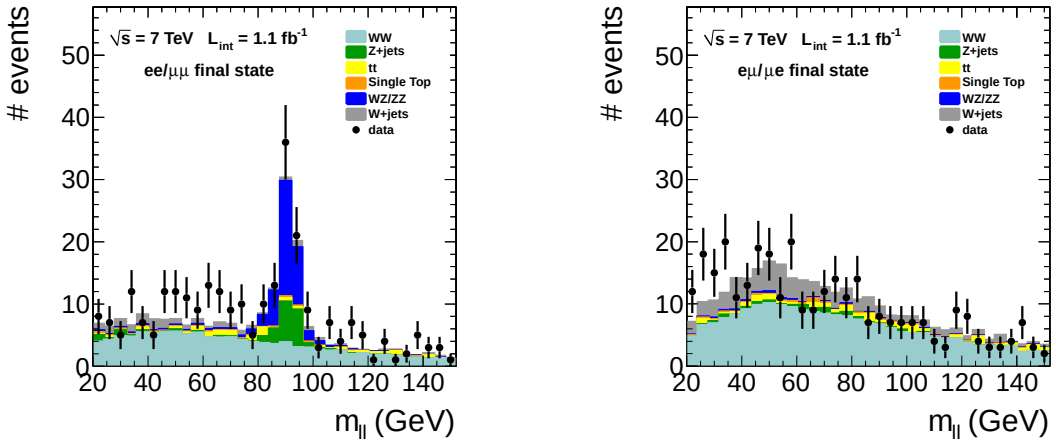
We can write:

$$\begin{aligned} N(Z\to ee)^{sr} &= N_{ee}^{sr} - \tfrac{1}{2}k_{ee}N_{mix}^{sr} - N_{ee}^{sr}(WZ, ZZ) \\ N(Z\to \mu\mu)^{sr} &= N_{\mu\mu}^{sr} - \tfrac{1}{2}k_{\mu\mu}N_{mix}^{sr} - N_{\mu\mu}^{sr}(WZ, ZZ) \end{aligned} \tag{7.14}$$

where N_{mix}^{sr} is the number of opposite flavor final state and the factor $\frac{1}{2}$ accounts for the fact that there as twice as many opposite flavor events than same flavor for non-Z decays (such as W^+W^- or $t\bar{t}$). This approach works if $N_{ee/\mu\mu} \gg N_{mix}$, or in other words if the statistical uncertainty on $N_{ee/\mu\mu}$ is small. The signal was designed in the first place to specifically suppress same flavor events originating from Z, so this approach cannot be used as such in the signal region.

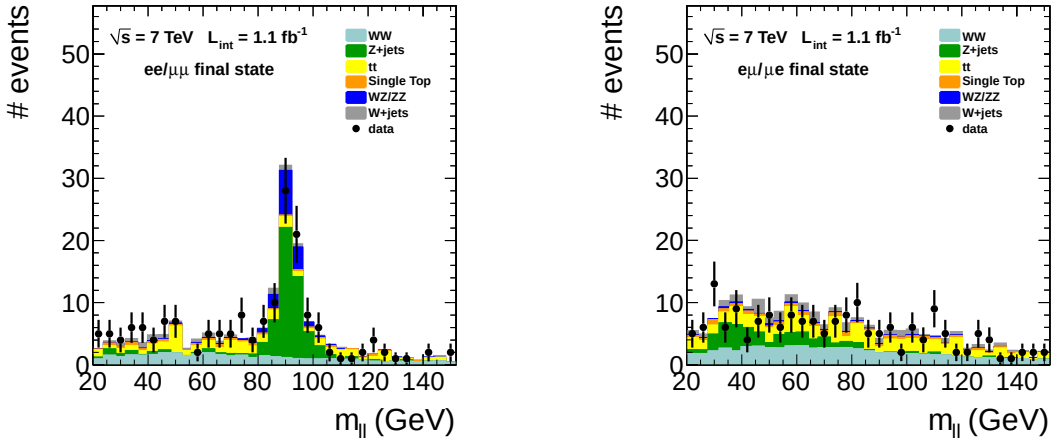
A region in which the same flavor contribution is enhanced needs therefore to be selected. The invariant mass of the di-lepton for same-flavor and opposite-flavor events is shown after the common W^+W^- 0-jet selection (Figure 7.13) and 1-jet selection (Figure 7.14) The natural choice is the control region is $|m_{\ell\ell} - m_Z| \leq 15$ GeV and the total signal selection aside from the Z-veto cut. From both figures it is clear that this control region is dominated by Drell-Yann and Diboson events for the same-flavor final state events. The number of same flavor events originating from Z decays in the signal region can then be written as:

[6] This effect is mitigated by the decision of using the min. proj. E_T^{miss} variable previously defined, but is still present



(a) (b)

Figure 7.13: Invariant mass of the di-lepton system for (a) same flavor final state and (b) opposite flavor final state after the W^+W^- selection in the 0-jet bin (no Z-veto is applied)



(a) (b)

Figure 7.14: Invariant mass of the di-lepton system for (a) same flavor final state and (b) opposite flavor final state after the W^+W^- selection in the 1-jet bin (no Z-veto is applied)

$$\begin{aligned} N(Z\to ee)^{sr} &= R^{ee}_{out/in}\left(N^{cr}_{ee} - \tfrac{1}{2}k_{ee}N^{cr}_{mix} - N^{cr}_{ee}(WZ,ZZ)\right) \\ N(Z\to \mu\mu)^{sr} &= R^{\mu\mu}_{out/in}\left(N^{cr}_{\mu\mu} - \tfrac{1}{2}k_{\mu\mu}N^{cr}_{mix} - N^{cr}_{\mu\mu}(WZ,ZZ)\right) \end{aligned} \quad (7.15)$$

where $k_{ee} = \frac{1}{k_{\mu\mu}} = \sqrt{\frac{N_{ee}^{loose}}{N_{\mu\mu}^{loose}}}$. N_{ee}^{loose} and $N_{\mu\mu}^{loose}$ are the number of same flavor event in "loose" control region defined by $|m_{\ell\ell} - m_Z| \leq 15$ GeV and $\text{min}E_T^{miss} \geq 40$ GeV. If we do so the control region is very similar to the signal region, with the exception of the invariant mass cut. The dilepton invariant mass is expected to be well modelled by simulation.

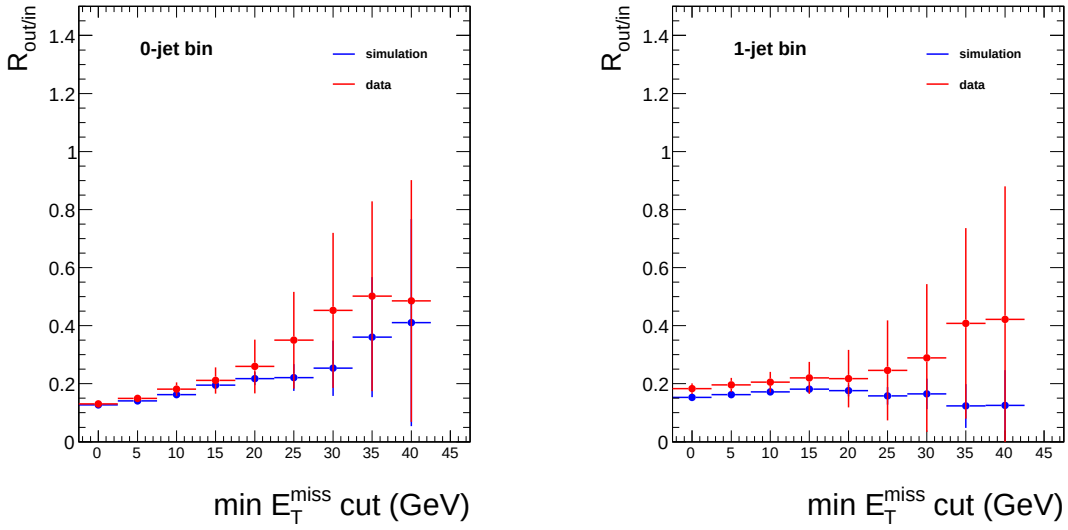


Figure 7.15: $R_{out/in}$ in simulation and data in the (a) 0-jet and (b) 1-jet bin

The ratio $R_{out/in}$ is defined as the ratio of efficiencies of selecting event out of the Z peak and in the Z peak. The Drell-Yan and Diboson yield in the same flavor final state and $R_{out/in}$ strongly depends on the E_T^{miss} (and the min. proj. E_T^{miss}). In Figure 7.6 (a) and (b) the min. proj. E_T^{miss} distribution we can deduce that min. proj. E_T^{miss} is poorly modeled by simulation. The E_T^{miss} is expected to be mismodelled at high luminosity especially for events with non-real E_T^{miss}.

final state	ee	$\mu\mu$
$k_{ll}(0-\text{jet})(\text{Data})$	0.76 ± 0.04	1.31 ± 0.07
$k_{ll}(1-\text{jet})(\text{Data})$	0.72 ± 0.05	1.37 ± 0.10

Table 7.7: Factors that account for the difference in efficiency between the *ee* and $\mu\mu$ final states

$R_{out/in}$ can be either estimated from simulation or from data. In order to increase the statistics we chose to estimate $R_{out/in}$ by slightly relaxing the min E_T^{miss} cut in both the signal and the control region to 20 GeV. By doing doing so we are implicitly assuming that the variable $R_{out/in}$ is independent from the E_T^{miss} cut. If the fake E_T^{miss} introduced by the Drell Yan background was to be perfectly modeled by simulation this would be the case. For the reason already stated however, this is not the case: we decide to associate to $R_{out/in}$ a systematic error that accounts for this dependence. We take $\Delta R_{out/in}(syst.) = \frac{1}{2}|\text{max}_{i,j}(R_{out/in}(\text{min.proj.}E_T^{miss} > i) - R_{out/in}(\text{min.proj.}E_T^{miss} > j))|$, where i,j run over all possible $\text{min.proj.}E_T^{miss}$ cuts. In Figure 7.15, $R_{out/in}$ is shown as a function of the min.

proj. E_T^{miss} cut value for the ee and $\mu\mu$ channels combined in data and simulation.

In Table 7.7 the values k_{ll} are summarized. In order to increase the statistics, we have chosen to combine the ee and $\mu\mu$ channels in one single measurement. The yield of the Z background decaying to same flavor leptons as measured in data with the procedure that was just explained are summarized in the 0 and 1 jet bin in Table 7.8 and 7.9. In the signal region there is 1-to-3 times more Drell-Yan background in the same-flavor final state than was expected by simulation . To account for this discrepancy, the normalization factor given in the last line of Table 7.8 and 7.9 will be applied to the Drell-Yan simulated events. For the opposite flavor final state, corresponding to $Z \to \tau\tau$ decays, we will simply use the yields obtained from simulation.

Parameter	Value
Number of obsverved events in the control region (Data)	305
Number of opposite flavor events in the control region	226 ± 15
Number of diboson events in the control region (sim.)	46.0 ± 4.7
$R_{out/in}$	0.19 ± 0.04 ± 0.08
number of Z events in the signal region (data)	6.4 ± 4.0 ± 2.7
number of Z events in the signal region (sim.)	4.1 ± 1.0
data/simulation scale factor	1.56

Table 7.8: Summary of measurements that enter in the 0-jet bin Drell-Yan contribution estimate. When two uncertainties are given, the first corresponds to the statistical uncertainty and the second to the systematics.

Parameter	Value
Number of obsverved events in the control region (Data)	201
Number of opposite flavor events in the control region	113 ± 10.6
Number of diboson events in the control region (sim.)	18.7 ± 1.9
$R_{out/in}$	0.15 ± 0.04 ± 0.05
number of Z events in the signal region (data)	10.0 ± 3.0 ± 4.2
number of Z events in the signal region (sim.)	3 ± 1
data/simulation scale factor	3.3

Table 7.9: Summary of measurements that enter in the 1-jet bin Drell-Yan contribution estimate. When two uncertainties are given, the first corresponds to the statistical uncertainty and the second to the systematics.

7.3.3 Top backgrounds

Backgrounds with top quarks in the final state are produced at high rate at the LHC. The cross-section for such processes is substantially higher than that of the irreducible W^+W^- background. As seen in the previous sections the contribution of these backgrounds at W^+W^- selection level is reduced with the help of a b-tagging discriminator and with the requirement that no soft muons should be present in the final state.

0 jet bin

We recall that in the 0-jet bin the dedicated anti-top tagging cuts are defined in the following way: jets with $7 < p_T < 30$ GeV are required to not be tagged as b-jets, and no extra soft muons should be present in the event. A top enhanced region in the 0-jet bin can then be defined as: at least one jet is b-tagged or at least one soft muon is found. We call the efficiency of selecting such events in the 0-jet bin $\epsilon_{\mathrm{soft-tag}} = \epsilon^{\mathrm{cr}}$.

In Figure 7.16 the distribution of the number of jets is shown in simulation compared to data. Given that the signal selection strategy depends on the number of reconstructed jets with $p_T > 30$ GeV, the approach to the estimate top backgrounds differs between the 0 and 1 jet bin.

Given that the signal and the control region are complementary in this case, we have $\epsilon^{\mathrm{sr}} = 1 - \epsilon_{\mathrm{soft-tag}}$. We can estimate the number of top events in the signal region with:

$$N_{\mathrm{top}}^{\mathrm{sr}} = \frac{1 - \epsilon_{\mathrm{soft-tag}}}{\epsilon_{\mathrm{soft-tag}}} \left(N_{\mathrm{all}}^{\mathrm{soft-tag}} - N_{\mathrm{other}}^{\mathrm{soft-tag}} \right) \tag{7.16}$$

where $N^{\mathrm{soft-tag}}$ is the total number of soft-tagged events in data, while $N_{\mathrm{other}}^{\mathrm{soft-tag}}$ is the number of "contaminating" events in the control region (i.e. other backgrounds that fall in the control region) and is inferred from MonteCarlo simulation. As pointed out earlier, the details of b-jet fragmentation are poorly modeled by simulation (especially at low transverse momentum), therefore the efficiency $\epsilon_{\mathrm{soft-tag}}$ needs to be measured directly from data.

We define an event sample where exactly one jet > 30 GeV was reconstructed and we require this jet to be tagged as b-jet. This sample is almost a clean $t\bar{t}$ sample (95% pure in simulated events). We know that the hard b-tagged jet is one leg of the $t\bar{t}$ decay. We also know that the other top has decayed "at best" into a jet with a momentum $p_T < 30$ GeV. So the probability for a top quark to decay into a low p_T b-tagged jet (or no jet) or a soft muon is given by:

$$\epsilon_{\mathrm{b\,or}\,\mu} = \frac{N_{\mathrm{b\,or}\,\mu}}{N_{\mathrm{1b}}} \tag{7.17}$$

where the denominator N_{1b} is the number of events with exactly one b-tagged jet, and the numerator is the subset of these events with at least one soft b-tagged jet or a soft muon. We have measured $\epsilon_{\mathrm{b\,or}\,\mu}$ in Data and MonteCarlo and we find a good agreement: we find $\epsilon_{\mathrm{b\,or}\,\mu} = 41.7 \pm 4.1\%$ ($41.5 \pm 4.1\%$) in Data (MonteCarlo). The quoted uncertainty includes the statistical error (10%) and the small bias (5%) due to non-purity of $t\bar{t}$ in the N_{1b} sample. This agreement validates the method.

We have now to distinguish between Single Top and $t\bar{t}$ events. The efficiency for a Single Top event to be soft muon or b-tagged is simply $\epsilon_{\mathrm{b\,or}\,\mu}$. For a $t\bar{t}$ event however, it is the probability that one of top quark, or both are tagged, that is one minus the probability that neither of them is tagged:

$$\begin{aligned} \epsilon_{\mathrm{soft-tag}}(\,\mathrm{SingleTop}\,) &= \epsilon_{\mathrm{b\,or}\,\mu} \\ \epsilon_{\mathrm{soft-tag}}(\,t\bar{t}\,) &= 1 - (1 - \epsilon_{\mathrm{b\,or}\,\mu})^2 \end{aligned} \tag{7.18}$$

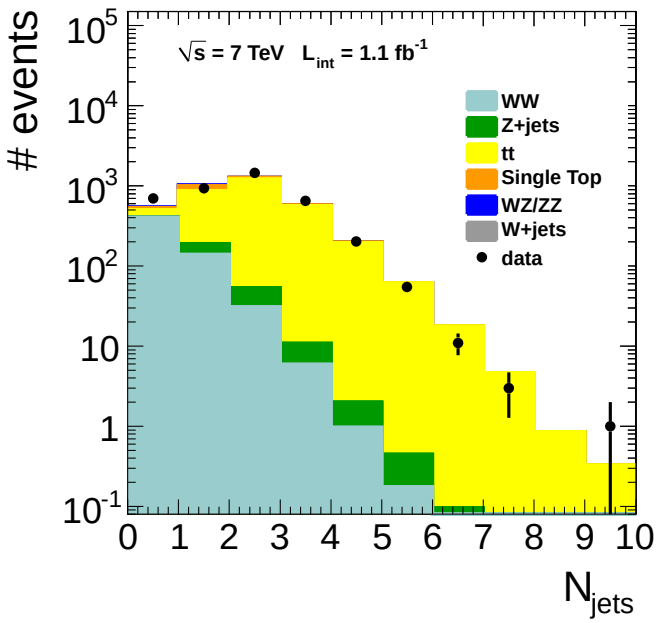


Figure 7.16: Contribution of the backgrounds in different jet bins

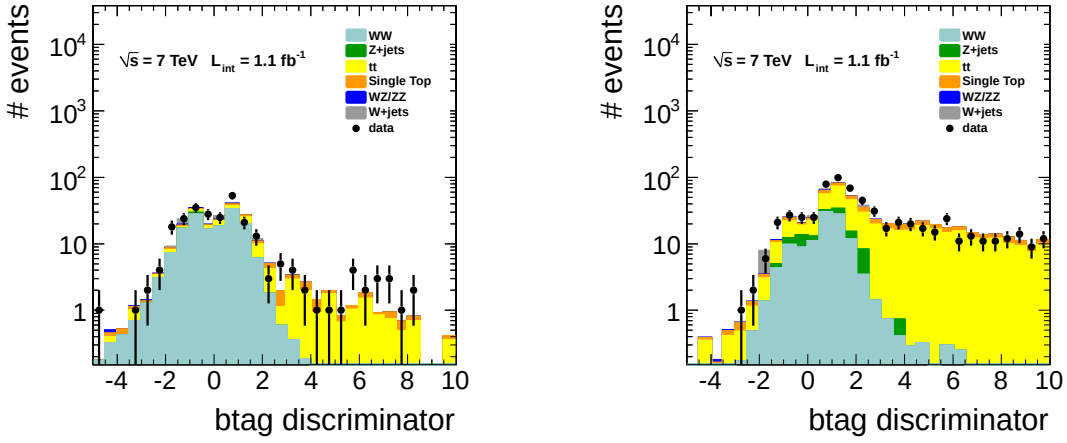


Figure 7.17: Distribution of the leading jet btag discriminator in the 0-jet bin (a) and in the 1-jet bin (b)

and finally taking into account the relative fraction of $t\bar{t}$ in the 0-jet bin $F_{t\bar{t}} = 1 - F_{\mathrm{SingleTop}}$:

$$\epsilon_{\mathrm{soft-tag}} = F_{t\bar{t}}(1 - (1 - \epsilon_{\mathrm{b\,or}\,\mu})^2) + F_{\mathrm{SingleTop}}\epsilon_{\mathrm{b\,or}\,\mu} \quad (7.19)$$

The $t\bar{t}$ fraction is measured in the 0 jet bin from simulation ($F_{t\bar{t}} = 72\%$). This number is subject to a large uncertainty however, mainly due to the fact that the ratio of $t\bar{t}$ and Single Top production at the LHC is known with an uncertainty[7] of 17% [98]. We find $\epsilon_{\mathrm{soft-tag}}$ $= 59.1 \pm 7.6\%$ ($58.1 \pm 7.6\%$) in Data (MonteCarlo) for the soft-tagging efficiency in the 0-jet bin. Here again the agreement is good between Data and MonteCarlo simulation. In Figure 7.17(a) the btag discriminator of the leading jet in the 0-jet bin is shown.

Parameter	Value
$t\bar{t}$ fraction in the 1-jet btag sample	95.0%
numerator for soft mu or b tagging efficiency ($N_{\mathrm{b\,or}\,\mu}$)	192 ± 14
denominator for soft mu or b tagging efficiency ($N_{1\mathrm{b}}$)	460 ± 21
soft mu or b tagging efficiency ($\epsilon_{\mathrm{b\,or}\,\mu}$)	$41.7 \pm 4\%$
$t\bar{t}$ fraction in the 0-jet bin	71.7%
soft-tagging efficiency combined ($\epsilon_{\mathrm{soft-tag}}$)	$59.1 \pm 7.6\%$
number of contaminating events in control region	12.3 ± 3.5
number of observed events in the control region (data)	88
number of observed events in the control region (sim)	56.5
number of top events in the signal region (data)	52.2 ± 16
number of top events in the signal region (sim)	31.1 ± 10.0
number of top events in the signal region (sim, counting)	37.4 ± 6.1
data/simulation scale factor	1.39

Table 7.10: Summary of measurements that enter in the 0-jet bin top contribution estimate

Finally, using Equation (7.17), we can estimate the residual number of top events in the 0-jet bin after the common W^+W^- selection. From simulation we find the number of contaminating events in the control region $N_{\mathrm{other}}^{\mathrm{soft-tag}} = 12.3 \pm 3.5$, the main contribution coming from the W^+W^- and W+jets backgrounds. When we apply this method to simulated events we find $N_{\mathrm{top}}^{\mathrm{sr}}(\mathrm{sim}) = 31.1 \pm 10.0$ events, which is in perfect agreement with the number of simulated top events that fall in the signal region (i.e from pure MonteCarlo counting), $N_{\mathrm{top}}^{\mathrm{sr}}(\mathrm{sim, counting}) = 37.4 \pm 6.1$. In Data, we find an excess of top events with respect to simulation, $N_{\mathrm{top}}^{\mathrm{sr}}(\mathrm{data}) = 52.2 \pm 16$ In order to account for this discrepancy, top-related distributions (and yields) taken from simulation will be renormalized by a scale factor $\frac{N_{\mathrm{top}}^{\mathrm{sr}}(\mathrm{data})}{N_{\mathrm{top}}^{\mathrm{sr}}(\mathrm{sim})} = 1.39$. All these results are summarized in Table 7.10.

1 jet bin

The top background rejection in the 1-jet bin is performed by requiring that the leading jet is not b-tagged, that no other (soft) jets are b-tagged and no soft muons are present in the

[7] In addition, some NLO Single Top diagrams overlap with LO $t\bar{t}$ diagrams. This interference is not taken into account in MonteCarlo simulation

event. A control region can thus be defined by requiring the leading jet to be b-tagged, and that no other soft muons and soft b-tagged jets are present in the event. The signal and the control region then differ by the requirement that the leading jet is (is not) b-tagged. We want therefore to determine the efficiency of b-tagging a hard jet $\epsilon_{\mathrm{hard-tag}}$. The residual number of events in the signal region is then simply:

$$\mathrm{N}_{\mathrm{top}}^{\mathrm{sr}} = \frac{1-\epsilon_{\mathrm{hard-tag}}}{\epsilon_{\mathrm{hard-tag}}}\left(\mathrm{N}_{\mathrm{all}}^{\mathrm{hard-tag,\,no\,soft\,b\,or}\,\mu} - \mathrm{N}_{\mathrm{other}}^{\mathrm{hard-tag,\,no\,soft\,b\,or}\,\mu}\right) \tag{7.20}$$

with $\mathrm{N}_{\mathrm{other}}^{\mathrm{hard-tag,\,no\,soft\,b\,or}\,\mu}$ to be determined as before from simulation.

In order to determine $\epsilon_{\mathrm{hard-tag}}$ from data a sample dominated by top events with hard jets has to be defined. We define this control sample as containing events with exactly two hard jets ($\mathrm{p_T} \geq$ 30 GeV). In order to increase the $\mathrm{t\bar{t}}$ purity we also require at least one of the two jets to be b-tagged. From simulation we check that this sample is indeed dominated by $\mathrm{t\bar{t}}$ events: we find 99% purity. We can assess the average hard-jet b-tagging efficiency by calculating:

$$\epsilon_{\mathrm{hard-tag}} = \frac{\mathrm{N_{bb}}}{\frac{\mathrm{N_{1b-leading}}+\mathrm{N_{1b-trailing}}}{2}} \tag{7.21}$$

where $\mathrm{N_{bb}}$ is the number of events where both the hard jet are b-tagged, and $\mathrm{N_{1b-leading}}$ ($\mathrm{N_{1b-trailing}}$) are respectively the number of events where the leading (trailing) jet is b-tagged. In data (simulation) we find $\epsilon_{\mathrm{hard-tag}} = 60.7 \pm 4\%$ (65.3 ± 3). Within the quoted statistical uncertainties the agreement is good.

Parameter	Value
$\mathrm{t\bar{t}}$ fraction in the 2-jet btag sample	99%
numerator for hard jet b-tagging effciency	759 ± 18
denominator for hard jet b-tagging effciency	461 ± 19
hard jet b-tagging efficiency in simulation ($\epsilon_{\mathrm{hard-tag}}$)	$65.3 \pm 3\%$
hard jet b-tagging efficiency in data ($\epsilon_{\mathrm{hard-tag}}$)	$60.7 \pm 4\%$
number of contaminating events in control region	19.2 ± 4.6
number of observed events in the control region (data)	320
number of observed events in the control region (sim)	320.9
number of top events in the signal region (data)	159.6 ± 18
number of top events in the signal region (sim)	160 ± 25
number of top events in the signal region (sim, counting)	122 ± 4
data/simlation scale factor	1.30

Table 7.11: Summary of measurements that enter in the 1-jet bin top contribution estimate

We can then use Equation (7.20) to estimate the number of top events in the signal region in the 1-jet bin. The number of other background events in the control region estimated from simulation is 19.2 ± 4 coming mainly from Drell-Yan, W+jets and $\mathrm{W^+W^-}$. As a closure test we apply this method to estimate the number of top events in simulation

and we find 160.5 ± 29 which is in agreement with 122.6 ± 4 found from counting of $t\bar{t}$ and Single Top events. In data, we find $N^{sr}_{top}(sim) = 159.6 \pm 18$, leading to a scale factor of 1.30. The results are summarized in Table 7.11. In Figure 7.17(b) the btag discriminator of the leading jet in the 1-jet bin is shown.

7.3.4 WW background

Although the cross section of the W^+W^- is sensibly lower than that of top and Z production at the LHC, the residual contribution of this background after the full $H \rightarrow W^+W^-$ selection is the highest. Following the same approach, a region where the W^+W^- contribution is enhanced needs to be defined. Naively the W^+W^- common selection provides by construction such a control region. However in this same region the signal is enhanced as well, so an additionnal selection needs to be defined such that the new selection is almost "signal free".

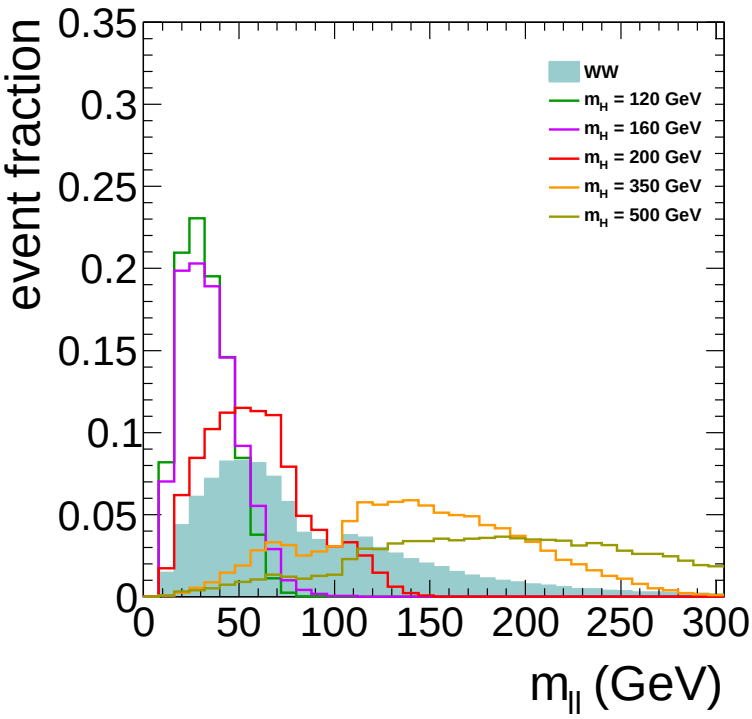


Figure 7.18: Distribution of the invariant mass of the two leptons. In order to see the difference in shape between the signal (for different Higgs masses) and the W^+W^- background the distributions are normalized to unity

In Figure 7.18 the invariant mass of the two leptons for the W^+W^- background and for different values of the Higgs mass is shown. We see that for $m_H \geq 200$ GeV the background and signal distribution overlap significantly. In this case it is difficult to define an efficient control region and therefore the W^+W^- yield will be taken directly from simulation. For $m_H \leq 200$ GeV however the leptons invariant mass distribution in the W^+W^- case is pretty

well separated than that of the signal. The control region in this case can be defined as $m_{\ell\ell}$ $\geq$ 100 GeV. This region is appropriate for both 0 and 1-jet bin analysis since the kinematics of the di-lepton system do not depend on the jet structure of the event. The signal region in this case is the common W^+W^- selection. We can write:

$$N^{sr}_{W^+W^-} = \frac{1}{\epsilon_{m_{\ell\ell}\geq 100}} \left(N^{m_{\ell\ell}\geq 100}_{all} - N^{m_{\ell\ell}\geq 100}_{other} \right) \tag{7.22}$$

The procedure is as follows:

- We measure the total number of events in the control region from data $N^{m_{\ell\ell}\geq 100}_{all}$ and subtract the events of the other backgrounds that fall in this region $N^{m_{\ell\ell}\geq 100}_{other}$. This last number is taken as before from simulation.

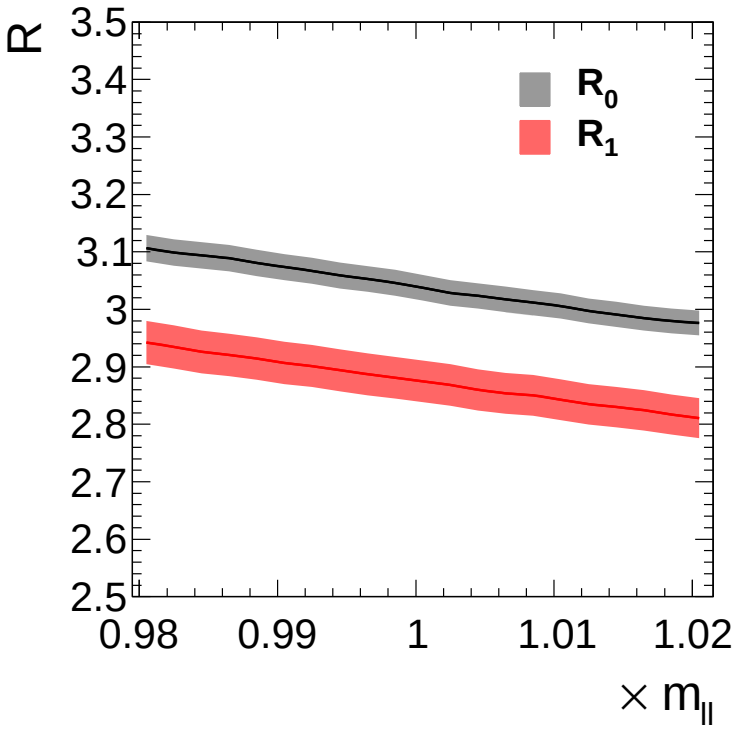


Figure 7.19: Extrapolation signal/control region factor $R_{0,1jet} = \frac{1}{\epsilon_{m_{\ell\ell}\geq 100}}$ in the 0 and 1 jet case as a function of the invariant mass resolution of the two lepton. The solid line represents the central whereas the light-shaded area represent the statistical uncertainty obtained by simulation

- We estimate $\epsilon_{m_{\ell\ell}\geq 100}$ from simulation as we believe that the dilepton invariant mass distribution is well modeled by simulation. In Figure 7.19 we show the effect of varying the invariant mass of the leptons on the ratio $R_{0,1jet} = \frac{1}{\epsilon_{m_{\ell\ell}\geq 100}}$. Assuming an average conservative uncertainty of 3% on the p_T resolution of leptons (which

implies $\frac{\Delta m_{\ell\ell}}{m_{\ell\ell}} \sim 2\%$) we deduct from Figure 7.19 that such an uncertainty on the the invariant has an effect on $R_{0,1jet}$ comparable to the statistical uncertainty.

- We extrapolate the control region estimate to the signal region (at the W^+W^- selection level)using Equation (7.22).

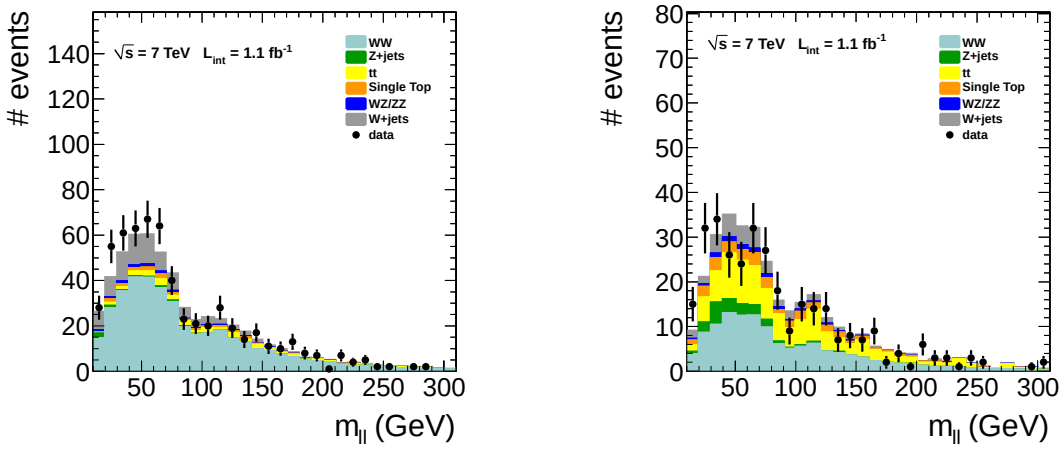


Figure 7.20: Distribution of the invariant mass of the di-lepton system in the 0-jet bin (a) and in the 1-jet bin (b)

Parameter	0-jet bin	1-jet bin
R	3.03± 0.06	2.87 ± 0.1
number of contaminating events in control region	47 ± 11	64±6
number of observed events in the control region (data)	176	107
number of observed events in the control region (sim)	169 ±8	103±5
number of WW events in the signal region (data)	392 ± 40 ± 35	124 ± 18 ± 30
number of WW events in the signal region (sim, counting)	369 ± 19	114±11
data/simlation scale factor	1.06	1.09

Table 7.12: Summary of measurements that enter WW background contribution estimate. When two uncertainties are quoted, the first is statistical, the second is systematic.

In Figure 7.20 the invariant mass distribution at the W^+W^- selection level is shown in the 0 and 1 jet bin. The highest contamination in the control region comes from W+jets background and top in the 0 jet bin (while in the 1 jet bin the top background dominates). The resulting W^+W^- background at common selection level is summarized in Table 7.12 for the 0 and 1 jet bin. The uncertainty on the final yield is highly dominated by the statistical uncertainty of the other background contaminations. We quote as statistical uncertainty the uncertainty on the number of observed events in the control region and as systematics the combination of the errors on the background contaminations and on the

extrapolation factor R. The data/simulation scale factor being close to 1, we conclude that this background is well modeled by simulation.

7.3.5 Fake leptons: W+jets and QCD

The main sources of fake leptons are $W + jets$ and QCD events, where at least one of the jets or a constituent is misidentified as an isolated lepton. The dominant background is $W + jets$ because there is already one prompt, well isolated, lepton from the W boson decay. Fake non-prompt leptons arise from the leptonic decay of heavy quarks, misidentified hadrons or electrons from photon conversions.

A set of loosely selected lepton-like objects, referred to as the "fakeable object", is defined in a sample of events dominated by dijet production. The efficiency for these denominator objects to pass the full lepton selection critera is measured. This background efficiency, typically referred to as the "fake rate", is parameterized as a function of the p_T and η in order to capture any dependence on kinematic and geometric quantities. The jet induced background study was not specifically adressed in this thesis. A detailed explanation can be found in [90]. We will use the results from [90] for this background.

Since the requirements that are applied are specifically designed to radically suppress this background, the amount of simulated events that pass the WW selection are limited (see the big statistical uncertainty on the W+jets yields in Table 7.5). The fake rates can be used as weights to extrapolate the background yield from a sample of loose denominator objects to the sample of fully selected leptons. These events can be used to form a "template" that we can use instead of the W+jets simulation sample. This will increase the number of events, and it has the advantage of being directly a direct data sample. We will therefore use this template instead of the simulation sample in what follows.

7.3.6 Other backgrounds

There are four processes which need to be estimated from Monte Carlo simulation, after applying the proper data corrections for lepton, trigger and jet veto efficiencies: WZ, ZZ, $W + \gamma$, $Z/\gamma^* \to \tau\tau$. All these backgrounds contaminations are very small.

7.4 Results at the WW preselection level with Data

7.4.1 Yields

We report here the final yields at the WW preselection level. The $Z \to ee, \mu\mu$, Top, W^+W^- and W+jets yields are estimated directly from data (see previous section). The other backgrounds are estimated from simulation.

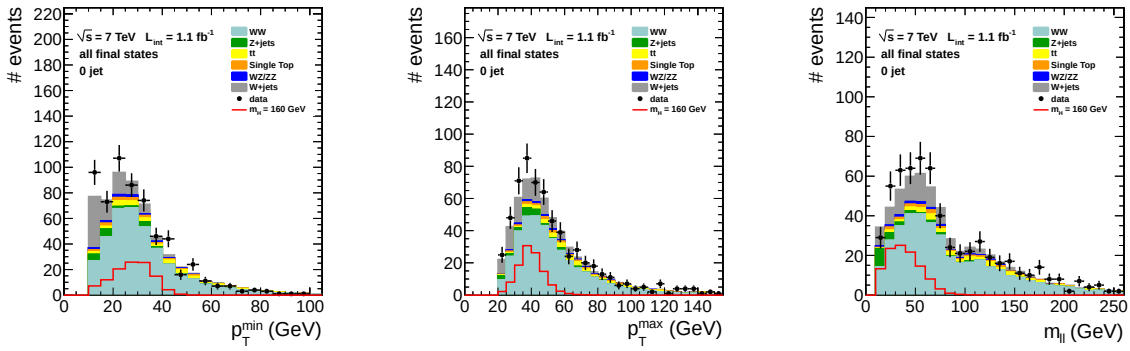
For backgrounds estimates taken from simulation, the uncertainty corresponds to the statistical uncertainty due to limited MonteCarlo sample, and the systematic uncertainty on the selection efficiency which has been estimated to 9.5 %. From those estimated from data, the total uncertainty has been computed in the previous section. The yields are summarized in Table 7.13.

	data	all bkg.	W^+W^-	$t\bar{t}+tW$	$W+\gamma$
0-jet	622	580 ± 67	392 ± 53	52.2 ± 16	9.8 ± 1.8
1-jet	340	355 ± 42	124 ± 35	159.6 ± 18	1.8 ± 0.5
	WZ/ZZ	$Z\rightarrow ee, \mu\mu$	$Z\rightarrow e\mu, \mu e$	$W+$jets	
0-jet	13.5 ± 1.4	6.4 ± 4.9	3.4 ± 0.7	103.0 ± 37.5	
1-jet	9.5 ± 1.0	10.0 ± 5.2	14.5 ± 2.0	35.9 ± 13.4	

Table 7.13: Expected number of signal and background events from the data-driven methods for an integrated luminosity of 1.1 fb^{-1}after applying the W^+W^- selection requirements. Statistical and systematic uncertainties on the processes are reported. For the yields derived from simulation, the uncertainty is that on the selection efficiency.

7.4.2 Control Variables

The most natural variables that define our final state at this point are the lepton tranverse momenta (p_T^{min} and p_T^{max}) and the missing transverse energy that is indirect measure of the neutrinos energy. For reasons already stated, the projected transverse missing energy (min $E_T{}^{miss}$) is more appropriate for disentangling the signal from particular backgrounds (rather than the $E_T{}^{miss}$).

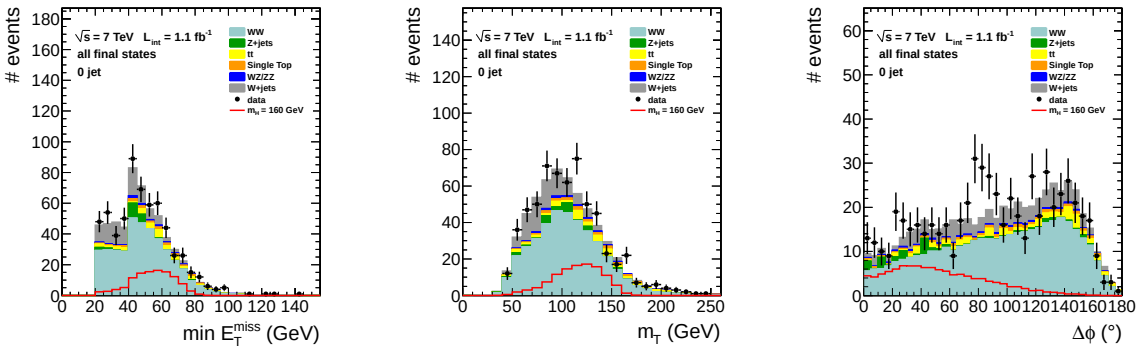


(a) (b) (c)

Figure 7.21: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 0-jet bin, of the trailing lepton p_T (a), leading lepton p_T (b) and dilepton invariant mass (c). All final states are included. Each component in simulation is scaled to data-driven estimates.

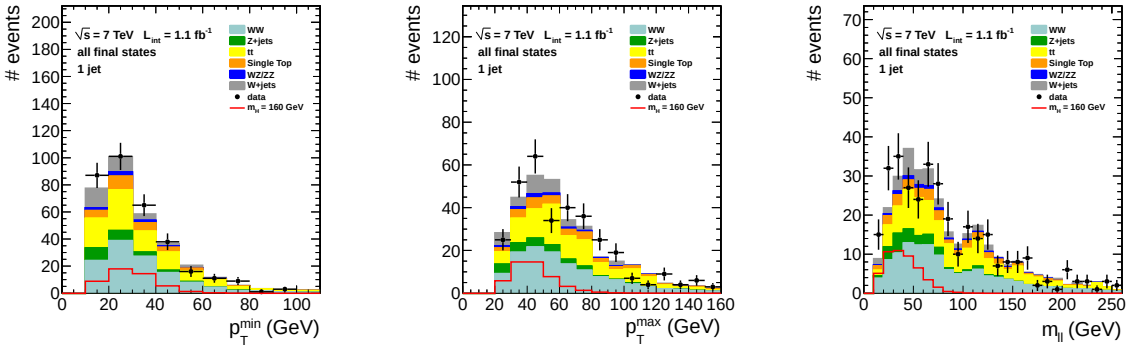
We have seen how the dilepton opening angle in the transverse plane ($\Delta\phi$) can play an important role in discriminating the signal from the W^+W^- continuum (see Chapter 5). As we are not able to reconstruct the full invariant mass of the final state, we also need variables that are sensitive to changes in the Higgs mass. The first natural choice is the invariant mass of the two leptons ($m_{\ell\ell}$).

A more sophisticated variable is the so-called "transverse mass". As the name indicates, it is simply the invariant mass of the system in transverse plane:



(a) (b) (c)

Figure 7.22: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 0-jet bin, of the min $E_T{}^{miss}$ (a), transverse mass (b) and dilepton $\Delta\phi$ (c). All final states are included. Each component in simulation is scaled to data-driven estimates.



(a) (b) (c)

Figure 7.23: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 1-jet bin, of the trailing lepton p_T (a), leading lepton p_T (b) and dilepton invariant mass (c). All final states are included. Each component in simulation is scaled to data-driven estimates.

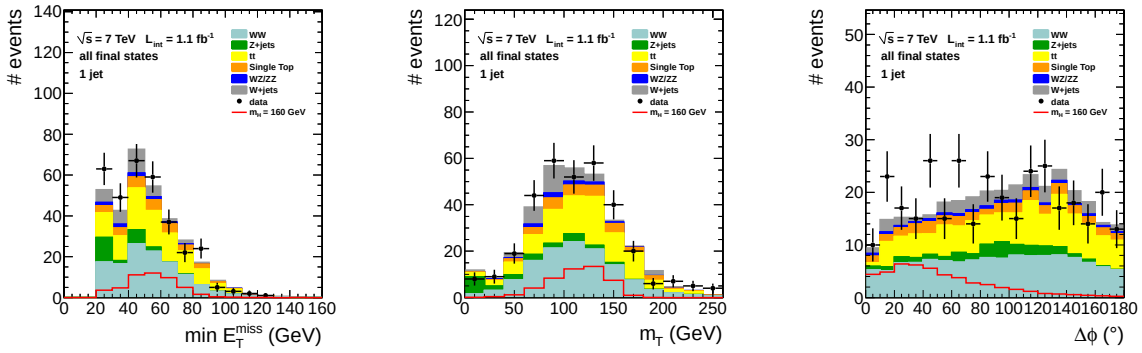
$$m_T = (p_T^{\ell\ell} + E_T{}^{miss})^2 - (\overrightarrow{p_T^{\ell\ell}} + \overrightarrow{E_T{}^{miss}})^2 \quad (7.23)$$

$$= \sqrt{2 p_T^{\ell\ell} E_T{}^{miss}(1 - \cos\Delta\phi_{\ell\ell - E_T{}^{miss}})} \quad (7.24)$$

These six distributions are shown in the 0 (Figure 7.21-7.22) and in the 1-jet bin (Figure 7.23-7.24). In Appendix C the same distributions are shown divided in same flavor and opposite flavor final states for $m_H = 160$ GeV.

7.4.3 W^+W^-cross section estimate

With these numbers we can give a rough estimate of the W^+W^- cross section. By treating the W^+W^- as the signal, we have



(a) (b) (c)

Figure 7.24: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 1-jet bin, of the min E_T^{miss} (a), transverse mass (b) and dilepton $\Delta\phi$ (c). All final states are included. Each component in simulation is scaled to data-driven estimates.

$$\sigma_{W^+W^-} = \frac{N_{data} - N_{bkg}}{\epsilon \cdot \mathcal{L} \cdot BR(W^+W^- \to \ell\ell\nu\nu)} \tag{7.25}$$

where $N_{data} - N_{bkg}$ is our best estimate of the W^+W^- yield after the selection. The uncertainties that enter this calculation are:

- luminosity: we take 4 % uncertainty for the integrated luminosity, this number being provided by the CMS luminosity group [90]
- efficiency: this uncertainty comes from the data estimate of the trigger and lepton selection efficiencies plus the E_T^{miss} resolution
- background dependent systematic and statistical uncertainty.

The dominant systematics are the background related uncertainties. In the 0 jet bin, we have $N_{bkg} = 188 \pm 40$. The signal (W^+W^-) efficiency is $\epsilon = 6.7 \pm 0.5$ %. Finally the branching ratio, from the particle data group [13] is $BR(W^+W^- \to \ell\ell\nu\nu) = 0.104$. We finally get:

$$\sigma(W^+W^-) = 53.4 \pm 2.9 \text{ (stat.)} \pm 7.3 \text{ (syst.)} \pm 2.1 \text{ (lumi.)} \text{ pb} \tag{7.26}$$

which makes our estimate compatible with latest calculations [99] give $\sigma(W^+W^-) = 47.0 \pm 2.0$ pb within 1 standard deviation. Please note that this is not meant to be a measurement of the WW cross section, which is not the subject of the present work. This was only made in order to check the consistency of the WW preselection. In order to make a precise measurement of the WW cross section, one needs to take into account the difference in the $gg \to W^+W^-$ and $qq \to W^+W^-$ selection efficiency [8].

[8] As a cross-check, we have also measured the WW cross section from the 1 jet-bin events. We found $\sigma(W^+W^-) = 44.1 \pm 7.3$ (stat.) ± 10.3 (syst.) ± 1.7 (lumi.) pb, which is compatible with 0 jet bin measurement.

7.5 Conclusion

In this Chapter we have described the selection requirements that allow to suppress a high fraction of the initial backgrounds. Data driven methods for estimating the contributions of such background at the end of the W^+W^- selection have been described in details. This allowed for a rapid estimate of the W^+W^- cross section and, most important, prepares the ground for the subject of next chapter: the Higgs signal extraction.

Chapter 8

Results

In the previous Chapter we have explained all the steps necessary to reduce most of the backgrounds affecting the H $\rightarrow$ WW search. At this point we are left mainly with the irreducible WW background. This background is difficult to reject, since its final state is almost identical to that of the signal. At the W^+W^-preselection level the signal to background ratio is $\sim$ 0.1 in the 0 jet bin and $\sim$ 0.05 in the 1 jet bin. We have seen however, in Chapter 5 that the signal to background ratio can be further improved by making use of the kinematical differences between these two processes. As we will see in the first section of this Chapter, these differences can be exploited in two ways: the first being an intuitive one, and the second being a more advanced technique. The result of these selections will be our final result, complemented with the inclusion of the uncertainties that affect the final yield estimates. We will then make use of these results to set upper limits on the mass dependent Higgs cross section.

8.1 Signal Extraction Strategy

We present here two complementary approaches in order to enhance the sensitivity to the signal:

- the cut-based approach
- the multivariate approach

8.1.1 Cut-based selection

This approach is simply an extension of the common preselection approach. We exploit the difference in the kinematic behavior of the signal and the background to derive a selection that enhances the sensitivity to our signal. The variables that will be used at this scope were presented in Section 7.4.2. We briefly recall them here:

- the transverse momenta of the leading p_T^{max} and trailing p_T^{min} leptons
- the dilepton invariant mass $m_{\ell\ell}$
- the angular distance in the transverse plane between leptons $\Delta\phi$
- the Higgs transverse mass m_T

m_H(GeV)	p_T^{max} (GeV)	p_T^{min} (GeV)	$m_{\ell\ell}$ (GeV)	$\Delta\phi$ (dg.)	m_T (GeV)
	>	>	<	<	[,]
120	20	10	40	115	[70,120]
130	25	10	45	90	[75,125]
140	25	15	45	90	[80,130]
150	27	25	50	90	[80,150]
160	30	25	50	60	[90,160]
170	34	25	50	60	[110,170]
180	36	25	60	70	[120,180]
190	38	25	80	90	[120,190]
200	40	25	90	100	[120,200]
250	55	25	150	140	[120,250]
300	70	25	200	175	[120,300]
350	80	25	250	175	[120,350]
400	90	25	300	175	[120,400]
450	110	25	350	175	[120,450]
500	120	25	400	175	[120,500]
550	130	25	450	175	[120,550]

Table 8.1: Final event selection requirements for the cut-based analysis for the zero and one jet bin anaysis.

As the distributions of the variables consistently change by varying the Higgs mass, the selection that we apply has to be Higgs mass dependent. The values of the applied rectangular cuts are shown in Table 8.1. The same selection is applied in the 0 and 1 jet bin as the signal and the WW background behave similarly in these two cases. Also, every final state is treated in the same way since most of the final state dependent background (such as Drell-Yan and Diboson) have been rejected already at this point and essentially, and only the WW remains.

A few comments have to be made on the selection that was adopted:

- As the hypothetical Higgs mass increases the selection on p_T^{max}, p_T^{min}, m_T requires the threshold on these variables to be higher and higher. This is obviously due to the fact that as the Higgs boson becomes heavier, the decay products are more and more Lorentz boosted, with respect to the WW background. However since the cross section is very low at high mass $m_H \geq 200$ GeV in order to not reject most of the signal, the lower bound on the m_T cut is kept constant.
- Due to the tremendous Lorentz boost of the W's at high Higgs mass, the leptons are folded along their direction and the angle $\Delta\phi$ between the leptons opens up, decreasing the discriminating power of this variable against the WW background.
- Finally, note that the missing energy was not used as an additional discriminating variable, as it is very much correlated with the m_T variable.

8.1.2 Neural Networks selection

In the previous section we have described the most simple approach to signal extraction, based on the selection of a particular phase space in which the signal over background ratio

is enhanced. Although very intuitive this approach is far from being optimal. In order to make use of all the available information present in the input discriminant variables more advanced techniques can be used. One among such techniques is the Neural Network discriminant. Neural Networks (NN) belong to a general class of multivariate statistical techniques. The detailed discussion of such methods is beyond the scope of this thesis, so for a complete review of such multivariate techniques and their application in high energy physics, please see [100].

8.1.3 Description

An artificial neural network can be seen as surjective mapping of the space of input variable of dimension N (x_1, ... , x_N) into a 1 dimensioned space, the final output discriminant. In our particular case the input random variables can be taken as the main discriminant variables of the analysis, such as those that are used in the cut-based selection.

In the present work, the so-called multi-layer perceptron (MLP) model was used. A typical MLP architecture is shown is Figure 8.1. It is implemented in the following way:

- the first layer has a number of "neurons" equal to the number of input variables, chosen by the user
- the number and the size of intermediary layers are chosen by the user.
- the final layer corresponds to the final output of the MLP.

A typical neuron "j" in the hidden layer is characterized by the synapse function κ and the activation function α:

$$\kappa \quad : \quad (\overrightarrow{y_j}, \overrightarrow{w_j}) \rightarrow \overrightarrow{y_j} \cdot \overrightarrow{w_j} \tag{8.1}$$

$$\alpha \quad : \quad x \rightarrow \tanh x \tag{8.2}$$

where the w_j (y_j) are the weights (outputs) associated to each neuron that connects to the neuron j (see Figure 8.1). The output function of the neuron is then simply defined as $\rho(\overrightarrow{y_j}, \overrightarrow{w_j}) = \alpha(\kappa(\overrightarrow{y_j}, \overrightarrow{w_j}))$.

A typical learning process works as follows:

- n events characterized by N variables are given as input.
- for every event the output of the final layer y_{NN} is computed. The input variable are normalized to be in the range [0,1] so that y_{NN} is the same range.
- the output is compared to a desired value: 0 if the background is being trained, 1 if the signal is being trained.
- the set of weights that minimizes the distance between the desired output and the y_{NN} value are computed.
- the whole procedure starts again with a new event until every event has been processed.

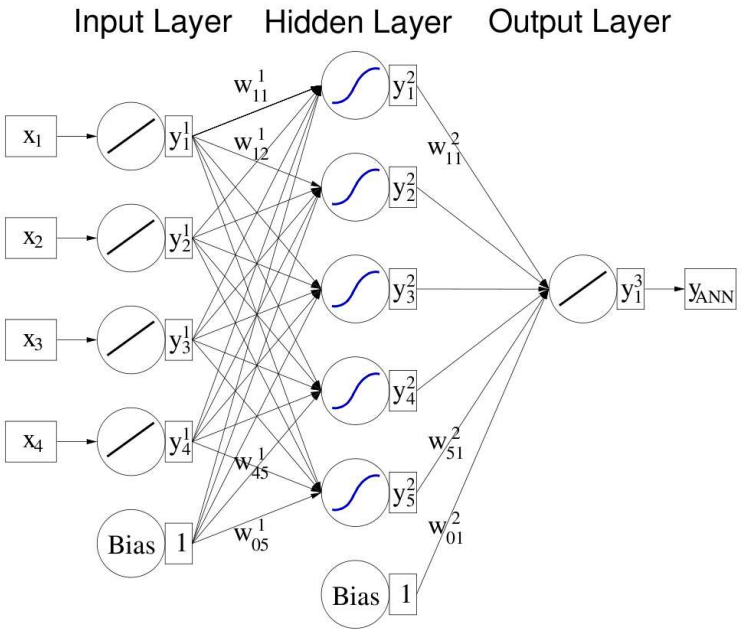


Figure 8.1: A typical multi-layer perceptron architecture. The first layer corresponds to the input variables and the final layer consisting of 1 neuron produces the final output. The intermediary layers are not fixed and can be set the user.

8.1.4 Procedure

The training is indeed performed on simulated events. Provided that enough events are used, the NN [1] can "learn" the differences between the signal and the background, and can build a single discriminant out of the learning process.

To simplify to the task to the NN, a preselection was applied after the common preselection in order to further reduce residual backgrounds other than the WW continuum. The following conditions were asked:

- an upper bound on the dilepton invariant mass $m_{\ell\ell}$
- an lower/upper bound on the transverse mass m_T

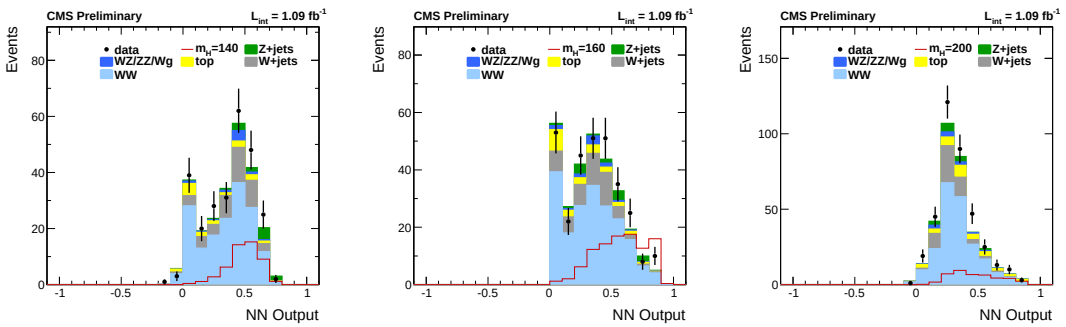
The cut values are mass dependent and are summarized in Table 8.2. As most of the other backgrounds have been almost completely suppressed at this point, the Signal is trained against the WW background only. For better comparison with the cut based results, we use as input variables the same variables that were used in that case: $m_{\ell\ell}$, m_T, p_T^{max}, p_T^{min}, and $\Delta\phi$.

Rather than fixing the architecture of the Neural Network, for every mass point we have trained the signal against the background for several different Network topologies. The number of internal layers was allowed to vary between 1 and 4, and the number of

[1] Enough here means that the number of events used to train the NN is consistently higher than the number of adjustable weights. If not "overtraining" can occur, which means that the NN learns statistical fluctuations, rather than discerning a global trend in the input variables. Overtraining issues can be checked by dividing the training sample into two independent samples (the training and the testing sample)

m_H(GeV)	$m_{\ell\ell}$ (GeV)	m_T (GeV)
	<	[,]
120	70	[80,120]
130	80	[80,130]
140	90	[80,140]
150	100	[80,150]
160	100	[80,160]
170	100	[80,170]
180	110	[80,180]
190	120	[80,190]
200	130	[80,200]
250	250	[80,250]
300	300	[80,300]
350	350	[80,350]
400	400	[80,400]
450	450	[80,450]
500	500	[80,500]
550	550	[80,550]

Table 8.2: Cuts applied after the WW common preselection in order to further reduce the non-WW backgrounds. The output of this selection serves as an input for the Neural Networks

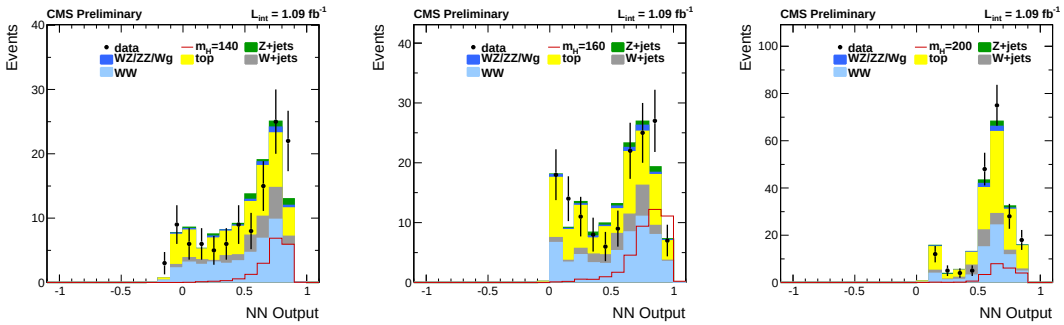


(a) (b) (c)

Figure 8.2: Distribution of the best Neural Network output for $m_H = 140, 160, 200$ GeV in the 0 jet bin.

neurons per layer between 5 (the number of input variables) and 9. This resulted in 16 trained networks per mass point. As the amount of simulation events is small, we have decided to train the NN network without making a distinction among different final states.

For each Network we have looked for the cut value on the NN output that maximizes the exclusion potential of the Higgs Signal. The best NN was then chosen as the configuration that provided the best exclusion limit for the signal (see later for a definition of the exclusion limit). The distributions of the output the best Neural Network are shown in Figure 8.2 and 8.3 for $m_H = 140, 160, 200$ GeV. In Appendix D the output is shown for every mass



(a) (b) (c)

Figure 8.3: Distribution of the best Neural Network output for $m_H = 140, 160, 200$ GeV in the 1 jet bin.

point and jet bin optimization. The selection criteria on the NN output are summarized in Table 8.3.

m_H(GeV)	0 jet	1 jet
	>	>
120	0.32	0.60
130	0.35	0.70
140	0.37	0.67
150	0.46	0.71
160	0.61	0.82
170	0.60	0.81
180	0.56	0.70
190	0.49	0.76
200	0.43	0.67
250	0.39	0.63
300	0.51	0.68
350	0.57	0.73
400	0.67	0.87
450	0.73	0.88
500	0.73	0.87
550	0.74	0.93

Table 8.3: Selection applied on the Neural Network output for every Higgs mass dependent optimization.

8.2 Summary of systematic uncertainties

After the full selection discussed in the previous section, we end up for each given final state, jet-bin and mass point with some number of MonteCarlo predicted signal events s and background events b_i, some of which are data-driven estimated. Besides the statistical

uncertainties on s and b_i, we have to quantify the experimental and theoretical biases that affect them.

8.2.1 Uncertainties on the signal yield

Every step of the selection starting from the triggers down to the final selection (see Equation 7.3) involves an systematic effect either due to the method used to determine the efficiency, when the efficiency is taken from simulation, and from the uncertainty on physical parameters involved in that particular selection.

- Lepton acceptance

  The energy scale for electrons and the momentum scale for muons is known with some uncertainty. We assign a systematic uncertainty by varying the transverse momentum of the muons by 1%, and 2% and 5% for electrons in the barrel and the endcap, respectively and measure the effect of the signal selection efficiency. The contribution to the uncertainty on the dilepton efficiency is about 1.5%.

- Lepton identification and trigger efficiencies

  The corresponding efficiencies were measured in data using the tag and probe technique on Z events. The estimated uncertainty is about 2% per lepton. Details can be found in [90].

- E_T^{miss} modeling: Events with neutrinos giving real E_T^{miss} in the final state (such as the signal) have a small uncertainty. However, since the E_T^{miss} is known to be mismodelled in a high pile-up regime we have addressed this uncertainty on the event selection efficiency by varying the E_T^{miss} in signal events by 10%. We find an uncertainty on the event selection efficiency of around 2.5%.

- Jet counting

  Since the analysis is divided in different jet bins, we must make sure that the jet multiplicity is well reproduced by our simulation. As said in Section 7.2.8 one possible approach would be to measure the jet counting efficiency in Drell-Yan events in data, and assign an uncertainty to discrepancy observed with simulation. However the jet content is very different for the signal and Drell-Yan events, so we prefer to just take the value from simulation, and address this uncertainty by varying the jet energy scale by 3%. We find an overall effect on the jet selection efficiency of 2%.

  The dominant effect on the jet counting efficiencies comes from theory since the jet momenta are strongly dependent on the Higgs momentum. Mainly three effects have to be estimated: those from parton-distribution functions, missing higher order corrections, and hadronization and parton shower modelling. Details of the following can be found in [90].

  – The uncertainties from parton-distribution function and α_S were estimated via the PDF4LHC prescription [15]: three PDF sets were used. Each PDF set leads to an uncertainty band and the total envelope is taken as the total uncertainties. For ggH and quark VBF induced processes the total uncertainty was found to be respectively 8% and 2%. The method used to estimate the uncertainty on a single PDF set is illustrated in Appendix E. The total uncertainty from PDFs+α_S was however not estimated directly here. The corresponding values are taken from [90].

  - Inclusive Higgs cross sections are subject to large uncertainties due to missing higher order corrections, as discussed in Chapter 3. Here we are mainly interested on the uncertainty on the exclusive 0 and 1 jet cross-section (σ_0 and σ_1). Knowing the uncertainty on the inclusive more than 0 jet ($\sigma_{\geq 0}$), more than 1 jet ($\sigma_{\geq 1}$) and more than 2 jets with a ($\sigma_{\geq 2}$) p_{T} threshold of 30 GeV, we can estimate the uncertainty on the exclusive 0 and 1 cross sections. The uncertainties due to missing higher orders terms on the inclusive can be estimated via the program MCFM [92] by varying the renormalization and factorization scales as explained in Chapter 3. Typical values are of the order of 20% for the exclusive 0 jet and 30% for the 1 jet bin at $\mathrm{m_H} = 160$ GeV (these numbers depend on the Higgs mass).
  - A sub-dominant uncertainty on the jet bin fraction comes from the mismodeling of the multi-parton interactions, the parton shower and the hadronization. In order to estimate this uncertainty the relevant parameters are varied in Pythia and the effect on the jet bin fraction is found to be typically of 5% in the 0 jet bin and 5-10% in the 1 jet bin, depending on the Higgs mass [90].

- Luminosity

  We assume an uncertainty of 4% on the integrated luminosity of 1.1 fb^{-1}, as adviced by the CMS luminosity group.

- Cross Section

  The uncertainty on the inclusive Higgs cross section is already included in the propagation of the exclusive binned cross-section. For simulation purposes, MonteCarlo programs tend to assume that the Higgs is produced on-shell, which is equivalent to say that its width is 0 [2]. Although valid at low mass ($\mathrm{m_H} \leq 200$ GeV), this approximation is not anymore justified for high masses (see Figure 2.1). When the width is large, destructive interference with other processes such as continuum WW production cannot be neglected. Therefore we attribute to the total Higgs cross section an additional uncertainty ranging from 1% at $\mathrm{m_H} = 200$ GeV up to 25% at $\mathrm{m_H} = 550$ GeV.

- Monte Carlo statistics

  Due to the finite size of the signal simulation sample, we have to consider the statistical uncertainty on the final Higgs selection. This contributes an uncertainty of about $2-3\%$ to the signal efficiencies depending on the Higgs mass.

8.2.2 Uncertainties on the backgrounds

The treatment of background uncertainties is different depending on how the background yield was estimated:

Data-driven backgrounds The methods used to estimate the different backgrounds are explained in Section 7.3. These background have been estimated at the WW selection level. Therefore we should consider, in addition to uncertainty on the yield at the WW preselection level, the eventual bias introduced by the successive selection.

The number of background "B" after the full NN or cut based selection equals to:

[2] The Breit-Wigner distribution is then simply a delta function.

$$N_{B_{full}} = R_{full} \cdot N_B^{cr} \tag{8.3}$$

$$\text{where } R_{full} = R_{in/out} \cdot \epsilon_{after-WWsel.} \tag{8.4}$$

where the $R_{in/out}$ and N_B^{cr} were estimated in Section 7.3. We have:

- statistical uncertainty on the number of observed events N_B^{cr} in the control region specifically designed for each background
- systematic uncertainty on the extrapolation factor R_{full} that includes both the uncertainty on the extrapolation at the WW level $R_{in/out}$ and the uncertainty on the efficiency of the selection from the WW level down to the final level $\epsilon_{after-WWsel.}$.
- The WW and the Top background

  The uncertainty on the selection, whether cut-based or NN output [3], involves only well modeled kinematic variables, and in principle there is some uncertainty associated to them. In this case we have simply taken the uncertainties estimated at the WW level and added an additional 10% uncertainty on the successive selection $\epsilon_{after-WWsel.}$ both for the cut-based and the neural network selection.
- The Drell-Yan background

  As the number of events in simulation surviving after the WW selection is extremely small for this background, we have decided to estimate its contribution at the final selection level. The control region was defined to be the same as that of Section 7.3.2 by removing the cuts on the transverse mass and the invariant mass. Although this leads to substantially large uncertainties (dominated by the low number of observed events in the control region defined), it is a robust method since the Drell-Yan background is estimated directly after the final selection. For the neural network selection, we have simply taken the uncertainties calculated at the WW level, supplemented with an additional 50% on the final yields, accounting for the error due to poor statistics on the numerator and denominator of the efficiency of the transverse mass, invariant mass cut and the neural network final selection.
- W+jets

  As the procedure is different for the W+jets background we just take the uncertainty on the fake rate estimate of 35% estimated in [90], and add the uncertainty on $\epsilon_{after-WWsel.}$ that is dominated by statistics in this case, leading to an uncertainty on the final yield up to 60%.

Other Background All the other backgrounds (including the WW background at $m_H > 200$ GeV) are estimated directly from simulation. As such, their final yield suffers from the same uncertainties as the signal. In some cases the statistical uncertainties can be much larger, such as the $Z \to \tau\tau$ background. Also, the uncertainty on the inclusive cross-section has been included for each of those background which can be as big as 50% for W+γ and as small as 3% for the diboson.

We show the final yields of the signal and backgrounds in Tables 8.4, 8.5, 8.6, 8.7 and 8.8, 8.9 respectively for the cut-based and the neural network selection. The total uncertainty on each process yield is included.

[3] Consider that the NN selection also includes a pre-kinematical selection on the invariant and the transverse mass.

mass (GeV)	H	WW	WZ/ZZ	Top	Z+jets	W+jets	Wγ	all. bkg	Data
120	6.0 ± 1.4	27.8 ± 5.3	0.7 ± 0.1	2.4 ± 1.3	0.0 ± 0.0	15.0 ± 8.8	3.0 ± 1.8	48.9 ± 10.6	54.0
130	11.1 ± 2.5	30.3 ± 5.8	0.7 ± 0.1	2.8 ± 1.4	0.0 ± 0.0	11.2 ± 6.6	1.6 ± 1.0	46.6 ± 9.3	51.0
140	14.5 ± 3.3	26.6 ± 5.1	0.6 ± 0.1	2.7 ± 1.4	0.0 ± 0.0	4.8 ± 2.9	1.3 ± 0.9	36.0 ± 6.9	36.0
150	13.6 ± 3.1	17.9 ± 3.4	0.4 ± 0.1	2.1 ± 1.3	0.0 ± 0.0	2.0 ± 1.3	0.3 ± 0.4	22.7 ± 5.0	21.0
160	20.2 ± 4.6	12.2 ± 2.4	0.3 ± 0.1	1.4 ± 1.0	0.0 ± 0.0	1.1 ± 0.7	0.0 ± 0.0	15.0 ± 5.4	15.0
170	15.5 ± 3.6	9.5 ± 1.8	0.2 ± 0.1	1.0 ± 0.6	0.0 ± 0.0	1.0 ± 0.7	0.0 ± 0.0	11.6 ± 4.1	13.0
180	11.5 ± 2.7	10.9 ± 2.1	0.2 ± 0.1	1.9 ± 1.2	0.0 ± 0.0	0.6 ± 0.5	0.0 ± 0.0	13.7 ± 3.7	14.0
190	10.1 ± 2.3	17.0 ± 3.3	0.3 ± 0.1	2.6 ± 1.4	0.0 ± 0.0	1.1 ± 0.7	0.0 ± 0.0	21.1 ± 4.3	22.0
200	7.6 ± 1.8	18.4 ± 3.5	0.4 ± 0.1	3.4 ± 1.7	0.0 ± 0.0	1.1 ± 0.7	0.0 ± 0.0	23.3 ± 4.4	26.0
250	4.0 ± 0.9	14.7 ± 1.5	0.5 ± 0.1	5.9 ± 2.5	0.0 ± 0.0	1.4 ± 0.9	0.0 ± 0.0	22.6 ± 3.2	21.0
300	3.1 ± 0.7	12.0 ± 1.3	0.4 ± 0.1	3.6 ± 1.6	0.0 ± 0.0	1.6 ± 1.0	0.3 ± 0.4	17.9 ± 2.4	17.0
350	3.2 ± 0.8	10.1 ± 1.1	0.3 ± 0.1	3.4 ± 1.5	0.0 ± 0.0	1.3 ± 0.8	0.3 ± 0.4	15.5 ± 2.2	15.0
400	2.5 ± 0.6	8.2 ± 0.9	0.2 ± 0.1	3.8 ± 1.7	0.0 ± 0.0	0.9 ± 0.6	0.3 ± 0.4	13.5 ± 2.1	12.0
450	1.4 ± 0.4	5.0 ± 0.6	0.1 ± 0.0	2.4 ± 1.2	0.0 ± 0.0	0.7 ± 0.4	0.0 ± 0.0	8.2 ± 1.5	6.0
500	1.0 ± 0.3	4.0 ± 0.5	0.1 ± 0.0	2.4 ± 1.3	0.0 ± 0.0	0.7 ± 0.5	0.0 ± 0.0	7.2 ± 1.5	4.0
550	0.6 ± 0.2	3.2 ± 0.4	0.1 ± 0.0	1.8 ± 1.2	0.0 ± 0.0	0.7 ± 0.5	0.0 ± 0.0	5.8 ± 1.3	2.0

Table 8.4: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the Cut-Based selection in the 0-jet bin in the opposite flavor state.

mass (GeV)	H	WW	WZ/ZZ	Top	Z+jets	W+jets	Wγ	all. bkg	Data
120	4.0 ± 0.9	19.8 ± 3.8	1.0 ± 0.1	0.8 ± 0.4	1.1 ± 2.9	7.2 ± 4.3	0.1 ± 0.1	30.0 ± 6.5	30.0
130	8.6 ± 2.0	24.8 ± 4.7	1.2 ± 0.1	0.7 ± 0.3	0.9 ± 2.8	6.2 ± 3.8	0.1 ± 0.1	33.7 ± 7.0	38.0
140	12.1 ± 2.8	23.9 ± 4.6	0.9 ± 0.1	1.0 ± 0.5	0.4 ± 1.5	4.7 ± 2.8	0.0 ± 0.0	30.8 ± 6.3	33.0
150	12.8 ± 2.9	16.6 ± 3.2	0.6 ± 0.1	0.6 ± 0.5	0.7 ± 2.5	0.7 ± 0.5	0.0 ± 0.0	19.2 ± 5.1	26.0
160	18.6 ± 4.3	11.5 ± 2.2	0.5 ± 0.1	0.9 ± 0.6	1.3 ± 3.6	0.0 ± 0.1	0.0 ± 0.0	14.1 ± 6.1	19.0
170	15.8 ± 3.6	9.3 ± 1.8	0.4 ± 0.1	1.3 ± 0.7	0.6 ± 2.7	0.2 ± 0.2	0.0 ± 0.0	11.4 ± 4.9	16.0
180	11.1 ± 2.6	10.5 ± 2.0	0.4 ± 0.1	2.2 ± 1.1	0.5 ± 1.9	0.0 ± 0.1	0.0 ± 0.0	13.6 ± 4.0	8.0
190	8.8 ± 2.0	15.2 ± 2.9	0.6 ± 0.1	3.4 ± 1.6	0.0 ± 0.1	0.7 ± 0.5	0.0 ± 0.0	19.9 ± 3.9	20.0
200	5.9 ± 1.4	14.8 ± 2.8	0.7 ± 0.1	3.3 ± 1.6	0.0 ± 0.2	0.9 ± 0.6	0.0 ± 0.0	19.7 ± 3.6	18.0
250	2.1 ± 0.5	8.7 ± 0.9	1.1 ± 0.1	4.3 ± 2.1	0.0 ± 0.1	0.9 ± 0.6	0.0 ± 0.0	15.0 ± 2.5	11.0
300	2.2 ± 0.5	8.2 ± 0.9	1.0 ± 0.1	3.9 ± 1.9	0.0 ± 0.1	0.9 ± 0.6	0.0 ± 0.0	14.0 ± 2.3	14.0
350	2.4 ± 0.6	7.6 ± 0.8	0.9 ± 0.1	3.7 ± 1.9	0.0 ± 0.1	1.0 ± 0.7	0.0 ± 0.0	13.3 ± 2.3	11.0
400	2.0 ± 0.5	6.2 ± 0.7	0.7 ± 0.1	3.1 ± 1.7	0.0 ± 0.0	0.9 ± 0.6	0.0 ± 0.0	10.9 ± 2.0	10.0
450	1.1 ± 0.3	3.4 ± 0.4	0.4 ± 0.1	2.9 ± 1.6	0.0 ± 0.0	0.6 ± 0.4	0.0 ± 0.0	7.3 ± 1.7	5.0
500	0.7 ± 0.2	2.6 ± 0.3	0.3 ± 0.1	1.9 ± 1.1	0.0 ± 0.0	0.4 ± 0.3	0.0 ± 0.0	5.2 ± 1.2	3.0
550	0.4 ± 0.1	2.2 ± 0.3	0.2 ± 0.0	1.5 ± 1.0	0.0 ± 0.0	0.4 ± 0.3	0.0 ± 0.0	4.3 ± 1.0	2.0

Table 8.5: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the Cut-Based selection in the 0-jet bin in the same flavor final state.

mass (GeV)	H	WW	WZ/ZZ	Top	Z+jets	W+jets	Wγ	all. bkg	Data
120	2.0 ± 0.8	7.1 ± 2.0	0.7 ± 0.1	5.0 ± 1.5	0.0 ± 0.0	2.2 ± 1.5	0.2 ± 0.2	15.1 ± 3.1	22.0
130	4.2 ± 1.7	7.7 ± 2.2	0.6 ± 0.1	6.5 ± 1.8	0.0 ± 0.0	2.2 ± 1.4	0.0 ± 0.0	17.1 ± 3.6	21.0
140	5.6 ± 2.3	6.9 ± 2.0	0.4 ± 0.1	5.3 ± 1.5	0.0 ± 0.0	1.7 ± 1.1	0.0 ± 0.0	14.3 ± 3.5	15.0
150	6.4 ± 2.6	6.1 ± 1.8	0.4 ± 0.1	5.1 ± 1.5	0.0 ± 0.0	0.8 ± 0.6	0.0 ± 0.0	12.4 ± 3.6	10.0
160	9.5 ± 3.9	5.0 ± 1.5	0.2 ± 0.1	4.7 ± 1.4	0.0 ± 0.0	0.8 ± 0.5	0.0 ± 0.0	10.7 ± 4.5	9.0
170	7.1 ± 2.9	4.0 ± 1.2	0.2 ± 0.0	4.2 ± 1.5	0.0 ± 0.0	0.4 ± 0.3	0.0 ± 0.0	8.7 ± 3.5	7.0
180	5.6 ± 2.3	4.7 ± 1.4	0.1 ± 0.0	4.9 ± 1.8	0.0 ± 0.0	0.5 ± 0.4	0.0 ± 0.0	10.2 ± 3.2	9.0
190	4.8 ± 2.0	6.6 ± 1.9	0.3 ± 0.1	10.0 ± 2.8	0.0 ± 0.0	0.8 ± 0.6	0.0 ± 0.0	17.7 ± 4.0	15.0
200	4.3 ± 1.8	7.9 ± 2.3	0.3 ± 0.1	11.6 ± 3.1	0.0 ± 0.0	1.0 ± 0.7	0.0 ± 0.0	20.7 ± 4.3	19.0
250	2.6 ± 1.1	8.4 ± 1.9	0.4 ± 0.1	18.6 ± 4.2	0.0 ± 0.0	1.3 ± 0.8	0.0 ± 0.0	28.7 ± 4.9	27.0
300	2.3 ± 0.9	7.7 ± 1.8	0.3 ± 0.1	16.9 ± 4.0	0.0 ± 0.0	1.2 ± 0.8	0.0 ± 0.0	26.1 ± 4.5	25.0
350	2.2 ± 0.9	6.9 ± 1.6	0.3 ± 0.1	13.7 ± 3.4	0.0 ± 0.0	1.0 ± 0.7	0.0 ± 0.0	21.9 ± 3.9	22.0
400	1.9 ± 0.8	5.9 ± 1.4	0.3 ± 0.1	14.5 ± 3.6	0.0 ± 0.0	1.3 ± 0.8	0.0 ± 0.0	21.9 ± 4.1	17.0
450	1.3 ± 0.5	3.7 ± 0.9	0.1 ± 0.0	8.0 ± 2.4	0.0 ± 0.0	1.0 ± 0.7	0.0 ± 0.0	12.9 ± 2.7	9.0
500	0.8 ± 0.4	2.9 ± 0.7	0.1 ± 0.0	4.7 ± 1.7	0.0 ± 0.0	0.9 ± 0.6	0.0 ± 0.0	8.6 ± 1.9	8.0
550	0.6 ± 0.3	2.4 ± 0.6	0.1 ± 0.0	4.4 ± 1.6	0.0 ± 0.0	0.8 ± 0.5	0.0 ± 0.0	7.8 ± 1.9	5.0

Table 8.6: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the Cut-Based selection in the 1-jet bin in the opposite flavor state.

mass (GeV)	H	WW	WZ/ZZ	Top	Z+jets	W+jets	Wγ	all. bkg	Data
120	1.2 ± 0.5	4.3 ± 1.2	0.3 ± 0.1	3.3 ± 1.5	2.8 ± 1.8	2.7 ± 1.7	0.1 ± 0.1	13.6 ± 3.2	17.0
130	2.4 ± 1.0	5.0 ± 1.4	0.4 ± 0.1	3.2 ± 1.3	2.4 ± 1.6	1.8 ± 1.2	0.1 ± 0.1	12.9 ± 3.0	20.0
140	3.8 ± 1.6	4.6 ± 1.3	0.4 ± 0.1	3.6 ± 1.4	1.6 ± 1.4	1.0 ± 0.7	0.1 ± 0.1	11.1 ± 3.0	20.0
150	4.5 ± 1.9	4.2 ± 1.2	0.2 ± 0.1	2.8 ± 1.0	1.3 ± 0.9	0.7 ± 0.5	0.0 ± 0.0	9.2 ± 2.7	18.0
160	7.4 ± 3.0	3.7 ± 1.1	0.2 ± 0.1	2.2 ± 0.8	0.7 ± 0.8	0.5 ± 0.4	0.0 ± 0.0	7.3 ± 3.4	17.0
170	6.1 ± 2.5	3.3 ± 1.0	0.2 ± 0.0	2.9 ± 1.0	0.2 ± 0.6	0.3 ± 0.3	0.0 ± 0.0	6.9 ± 3.0	16.0
180	5.0 ± 2.1	3.8 ± 1.1	0.1 ± 0.0	3.7 ± 1.3	0.1 ± 0.4	0.6 ± 0.5	0.0 ± 0.0	8.4 ± 2.7	15.0
190	4.1 ± 1.7	5.5 ± 1.6	0.2 ± 0.1	4.9 ± 1.6	0.1 ± 0.4	0.8 ± 0.5	0.0 ± 0.0	11.6 ± 2.9	20.0
200	3.1 ± 1.3	5.6 ± 1.6	0.3 ± 0.1	5.4 ± 1.9	0.1 ± 0.5	0.9 ± 0.6	0.0 ± 0.0	12.4 ± 2.9	21.0
250	1.3 ± 0.5	4.5 ± 1.0	0.5 ± 0.1	7.3 ± 2.4	0.1 ± 0.7	0.5 ± 0.4	0.0 ± 0.0	12.8 ± 2.8	16.0
300	1.4 ± 0.6	4.2 ± 1.0	0.4 ± 0.1	10.2 ± 2.9	0.1 ± 0.4	0.5 ± 0.4	0.0 ± 0.0	15.4 ± 3.2	18.0
350	1.6 ± 0.7	3.7 ± 0.9	0.4 ± 0.1	8.5 ± 2.6	0.2 ± 0.4	0.8 ± 0.6	0.0 ± 0.0	13.5 ± 2.9	19.0
400	1.4 ± 0.6	3.2 ± 0.7	0.3 ± 0.1	7.2 ± 2.4	0.2 ± 0.2	0.8 ± 0.5	0.0 ± 0.0	11.7 ± 2.6	15.0
450	0.8 ± 0.3	1.8 ± 0.4	0.2 ± 0.0	4.0 ± 1.7	0.2 ± 0.2	0.6 ± 0.4	0.0 ± 0.0	6.7 ± 1.8	8.0
500	0.6 ± 0.3	1.6 ± 0.4	0.1 ± 0.0	3.3 ± 1.6	0.0 ± 0.0	0.6 ± 0.4	0.0 ± 0.0	5.6 ± 1.7	6.0
550	0.4 ± 0.2	1.4 ± 0.3	0.1 ± 0.0	1.9 ± 1.0	0.0 ± 0.0	0.5 ± 0.4	0.0 ± 0.0	3.9 ± 1.1	3.0

Table 8.7: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the Cut-Based selection in the 1-jet bin in the same flavor final state.

mass (GeV)	H	WW	WZ/ZZ	Top	Z+jets	W+jets	Wγ	all. bkg	Data
120	10.4 ± 2.4	46.3 ± 9.0	1.9 ± 0.2	3.4 ± 1.7	2.6 ± 2.9	20.7 ± 12.1	1.6 ± 0.9	76.5 ± 15.7	85.0
130	24.6 ± 5.6	68.4 ± 11.0	2.6 ± 0.2	4.4 ± 1.8	3.5 ± 2.8	25.4 ± 14.8	2.1 ± 1.2	106.4 ± 19.6	114.0
140	42.6 ± 9.8	85.8 ± 11.3	3.1 ± 0.3	5.7 ± 2.0	3.7 ± 1.5	28.6 ± 16.6	2.5 ± 1.4	129.4 ± 22.6	141.0
150	48.5 ± 11.1	66.6 ± 8.2	2.3 ± 0.2	4.5 ± 1.8	2.6 ± 6.6	12.6 ± 7.3	0.3 ± 0.2	88.9 ± 17.0	106.0
160	46.0 ± 10.6	25.5 ± 4.7	0.6 ± 0.1	2.3 ± 1.5	0.9 ± 1.3	2.5 ± 1.5	0.0 ± 0.0	31.8 ± 11.8	41.0
170	42.1 ± 9.7	25.4 ± 3.9	0.7 ± 0.1	2.5 ± 1.4	0.8 ± 1.2	1.7 ± 1.0	0.0 ± 0.0	31.1 ± 10.6	42.0
180	33.3 ± 7.6	29.8 ± 4.5	1.0 ± 0.1	5.7 ± 2.4	0.9 ± 1.4	2.0 ± 1.2	0.0 ± 0.0	39.5 ± 9.4	36.0
190	27.1 ± 6.2	45.6 ± 6.8	1.7 ± 0.1	9.1 ± 3.1	0.5 ± 0.8	3.9 ± 2.2	0.0 ± 0.0	60.9 ± 10.0	61.0
200	21.8 ± 5.0	48.4 ± 7.0	2.3 ± 0.2	11.7 ± 3.5	0.5 ± 0.3	5.1 ± 3.0	0.0 ± 0.0	68.0 ± 9.7	78.0
250	12.6 ± 2.9	52.5 ± 5.1	3.0 ± 0.3	19.1 ± 5.0	0.1 ± 0.1	6.7 ± 3.9	0.3 ± 0.2	81.6 ± 8.6	89.0
300	7.9 ± 1.8	29.0 ± 2.8	1.7 ± 0.1	10.4 ± 3.6	0.0 ± 0.0	3.6 ± 2.1	0.3 ± 0.2	45.1 ± 5.3	45.0
350	7.5 ± 1.8	23.0 ± 2.2	1.3 ± 0.1	9.4 ± 3.5	0.0 ± 0.0	2.5 ± 1.5	0.3 ± 0.2	36.5 ± 4.7	39.0
400	5.2 ± 1.3	14.3 ± 1.4	0.8 ± 0.1	6.1 ± 3.3	0.0 ± 0.0	2.3 ± 1.3	0.0 ± 0.0	23.5 ± 4.0	28.0
450	3.2 ± 0.9	10.0 ± 1.0	0.8 ± 0.1	5.4 ± 2.8	0.0 ± 0.0	1.4 ± 0.8	0.0 ± 0.0	17.7 ± 3.2	21.0
500	2.1 ± 0.6	8.5 ± 0.8	0.5 ± 0.0	4.3 ± 2.4	0.0 ± 0.0	1.4 ± 0.8	0.0 ± 0.0	14.7 ± 2.7	14.0
550	1.3 ± 0.4	7.8 ± 0.8	0.4 ± 0.0	4.2 ± 2.1	0.0 ± 0.0	1.0 ± 0.6	0.0 ± 0.0	13.4 ± 2.4	12.0

Table 8.8: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the Neural Network selection in the 0-jet bin .

mass (GeV)	H	WW	WZ/ZZ	Top	Z+jets	W+jets	Wγ	all. bkg	Data
120	3.2 ± 1.4	11.7 ± 3.3	1.0 ± 0.1	9.3 ± 3.0	1.4 ± 1.8	6.0 ± 4.4	0.1 ± 0.1	29.5 ± 6.6	36.0
130	7.1 ± 3.0	12.3 ± 3.6	1.0 ± 0.1	10.0 ± 3.1	1.6 ± 1.6	5.4 ± 3.9	0.1 ± 0.1	30.4 ± 7.1	38.0
140	14.2 ± 6.0	18.4 ± 3.9	1.4 ± 0.1	16.7 ± 3.0	1.8 ± 1.4	7.5 ± 5.4	0.1 ± 0.1	45.7 ± 9.6	51.0
150	19.2 ± 8.1	18.0 ± 3.6	1.3 ± 0.1	14.7 ± 2.6	1.4 ± 3.6	4.5 ± 3.2	0.1 ± 0.1	40.0 ± 10.4	47.0
160	21.4 ± 9.0	9.7 ± 2.6	0.4 ± 0.0	9.6 ± 2.3	0.8 ± 1.2	1.3 ± 1.0	0.0 ± 0.0	21.9 ± 9.8	28.0
170	22.4 ± 9.5	11.9 ± 2.5	0.4 ± 0.0	12.7 ± 2.6	0.9 ± 1.4	1.5 ± 1.1	0.0 ± 0.0	27.5 ± 10.3	33.0
180	24.8 ± 10.5	26.1 ± 4.3	1.2 ± 0.1	29.4 ± 3.5	1.5 ± 2.3	4.4 ± 3.2	0.0 ± 0.0	62.6 ± 12.5	57.0
190	9.3 ± 3.9	11.0 ± 3.5	0.4 ± 0.0	16.0 ± 4.4	0.7 ± 1.0	1.8 ± 1.3	0.0 ± 0.0	29.8 ± 7.0	37.0
200	13.5 ± 5.7	25.5 ± 5.0	1.5 ± 0.1	42.4 ± 5.4	1.9 ± 0.9	4.3 ± 3.1	0.1 ± 0.0	75.6 ± 9.9	72.0
250	7.8 ± 3.3	23.5 ± 5.3	1.7 ± 0.2	51.0 ± 7.1	0.9 ± 0.4	4.3 ± 3.1	0.1 ± 0.0	81.4 ± 10.0	69.0
300	6.1 ± 2.6	17.2 ± 3.9	1.2 ± 0.1	41.3 ± 7.1	0.5 ± 0.3	3.6 ± 2.6	0.0 ± 0.0	63.9 ± 8.9	41.0
350	5.8 ± 2.5	12.3 ± 2.8	0.9 ± 0.1	29.4 ± 6.1	0.4 ± 0.4	2.9 ± 2.1	0.0 ± 0.0	45.8 ± 7.5	31.0
400	3.6 ± 1.6	5.4 ± 1.2	0.3 ± 0.0	13.5 ± 5.9	0.0 ± 0.0	1.3 ± 0.9	0.0 ± 0.0	20.5 ± 6.3	20.0
450	2.6 ± 1.2	5.1 ± 1.2	0.3 ± 0.0	12.2 ± 4.1	0.0 ± 0.0	1.4 ± 1.0	0.0 ± 0.0	19.1 ± 4.5	18.0
500	2.0 ± 0.9	5.5 ± 1.2	0.3 ± 0.0	13.4 ± 3.3	0.0 ± 0.0	1.3 ± 0.9	0.0 ± 0.0	20.5 ± 3.7	18.0
550	1.1 ± 0.5	2.8 ± 0.6	0.1 ± 0.0	5.4 ± 2.7	0.0 ± 0.0	0.8 ± 0.6	0.0 ± 0.0	9.2 ± 2.8	9.0

Table 8.9: Expected number of signal and background events for an integrated luminosity of 1.1 fb^{-1} after applying the Neural Network selection in the 1-jet bin .

8.3 Limit Setting

When establishing an upper limit on a signal cross-section we adress the following question: how likely is the observation of n events after a particular phase space selection to originate from the existence of the signal, or conversely to result from a statistical fluctuation of the background. This problem can be addressed in the framework of statistical hypothesis testing.

8.3.1 Hypothesis testing

The starting point is the definition of the *background* (b) hypothesis and the *signal plus background* (s+b) hypothesis. Consider for instance that the Standard Model prediction is to expect 10 background and 5 signal events and we observe 14 events, the observation is *more* compatible with the (s+b) hypothesis. Conversely if one observes 11 events the (b) hypothesis is favored. Although very intuitive, we need to make the statement *more or less compatible* quantitative.

The test statistics

The probability or likelihood that the outcome of a random LHC experiment resulting in n when s or s + b events are expected is described by a Poisson distribution. We can write:

$$L(n|s+b) = \frac{(s+b)^n e^{-(s+b)}}{n!} \tag{8.5}$$

$$L(n|b) = \frac{b^n e^{-b}}{n!} \tag{8.6}$$

The building block of the problem of hypothesis testing is the definition of a "test statistics". A test statistic is some function of the observation that is able to rank an observation as signal or background-like. In this case, a convenient test is the likelihood-ratio Q', defined as:

$$Q' = \frac{L(n|s+b)}{L(n|b)} \tag{8.7}$$

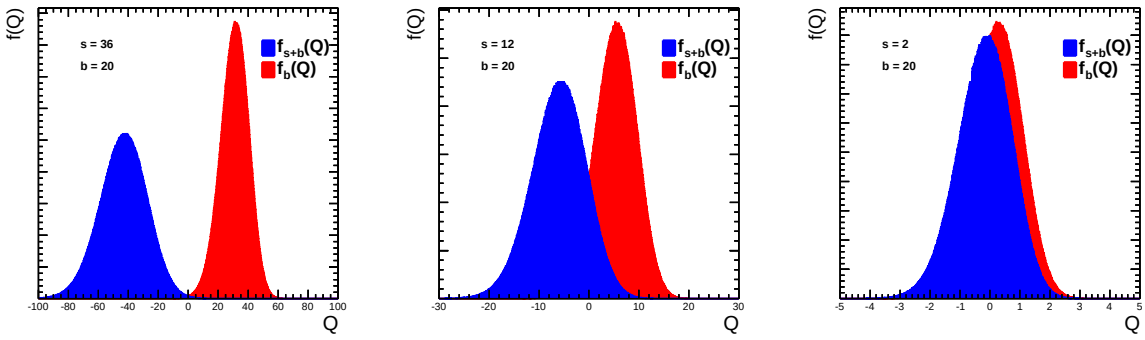
Rather than using Q' it is common to use the test statistics Q:

$$Q = -2 \ln Q' \tag{8.8}$$

$$= 2s - 2n \ln(1 + \frac{s}{b}) \tag{8.9}$$

This test statistics Q has to advantage of being negative (positive) for signal (background)-like outcomes. This can be easily seen by expanding the logarithm in the limit where $s \ll b$. The intermediary situation, when Q=0, indicates that no clear separation can be made between the two hypothesis.

In practice the probability distribution for the two hypothesis can be obtained by generating a high number of pseudo-experiments via a MonteCarlo random number generator. The outcomes of such experiments correspond to the number of observed events that follow the poissonian distributions from Equations (8.5) and (8.6). With this procedure the distribution of n can be generated both in the background only and in the signal plus background case hypothesis. The distribution of the test statistics Q can then be easily obtained via the Equation 8.9. The probability distributions for the background obtained



(a) (b) (c)

Figure 8.4: Example of distributions of the test statistics Q for the signal plus background hypothesis (blue) and background only hypothesis (red) obtained with a high number of pseudo-experiments. From left to right the signal standard model prediction decreases in this examples.

with the "b" (red) and "b+s" (blue) hypothesis are shown in Figure 8.4 for 3 different scenarios from a large signal cross section (a) to a low one (c). In (c) the test distributions are not so well separated, leading to more ambiguity in the distinction of the two alternative hypothesis.

Confidence levels

Frequentist intervals We now have to define the rules for the exclusion or a discovery of the signal. We have:

$$CL_{s+b} = \int_{Q_{obs}}^{\infty} f_{s+b}(Q)\, dQ. \tag{8.10}$$

$$CL_{b} = \int_{Q_{obs}}^{\infty} f_{b}(Q)\, dQ. \tag{8.11}$$

We say that the (s+b) hypothesis is excluded at the α confidence level if $CL_{s+b} \leq 1 - \alpha$. In other words we are falsely excluding a real signal with a probability α. Conversely, the background hypothesis is rejected (in favor of the signal plus background hypothesis) at the β confidence level if $1 - CL_b \leq 1 - \beta$. This is equivalent to say that the probability for the background yield to statistically fluctuate up to n_{obs} observed events is smaller than $1 - \beta$.

This is shown in Figure 8.5. The blue shaded area corresponds to the false exclusion rate CL_{s+b} and the red shaded area is the significance $(1 - CL_b)$. In high energy physics experiments we use the convention that a signal is excluded if $CL_{s+b} \leq 5\%$. In this case we call it exclusion at the 95% confidence level. We speak about discovery when β is equal to the fraction of the area of a gaussian distribution obtained by integrating over 5 standard deviations.

$$\beta = \mathrm{erf}^{-1}(2 \cdot CL_b - 1) \tag{8.12}$$

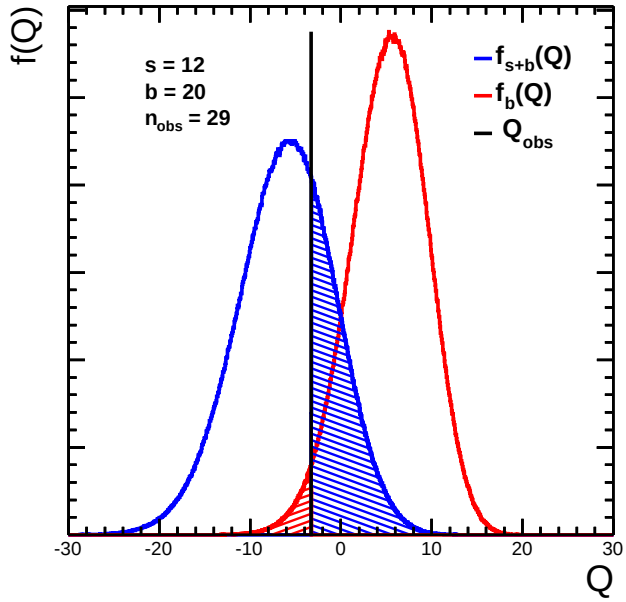


Figure 8.5: Test statistics for the signal and background hypothesis. CL_{s+b} is the dashed blue area and CL_b the dashed red. The observation Q_{obs} is represented by the black solid line.

$$\text{where erf(x)} = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-x^2}\, dx \tag{8.13}$$

β is usually referred as the significance of the observation. Observing a "5σ" excess in the number of observed events is then usually understood as the discovery a new particle.

The modified frequentist CL_s method Since no significant excess was observed we restrict our discussion to the signal exclusion only. Consider two experiments leading to the same expected signal but very different amount of expected background. For instance:

- A) s = 2, b = 6
- B) s = 2, b = 30

A small background fluctuation in case 2 can easily lead to a CL_{s+b} smaller in 2 than in 1, resulting in a better exclusion limit for the experiment 2. This outcome is quite undesirable as the exclusion limit would depend more on the sensitivity of the experiment itself rather than on the existence of the signal. Moreover, a downward fluctuation of the background in case 2 could lead to the (unphysical) conclusion that the signal cross section is negative.

One possible way to deal with this kind of situation is to normalize the confidence level CL_{s+b} by CL_b [101],

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{8.14}$$

case	s	b	n_{obs}	CL_b	CL_{s+b}	CL_s
A	2	6	6	0.60	0.31	0.52
B	2	30	28	0.40	0.27	0.68

Table 8.10: Illustration of the CL_s method.

By choosing to use CL_s, we are obliged to incorporate in the final limit both the knowledge on $f_{s+b}(Q)$ and $f_b(Q)$. When the two hypothesis are well separated, the CL_s and CL_{s+b} methods give the same result and they diverge substantially when the hypothesis (b) and (s+b) overlap. The CL_s approach is indeed more conservative, since $CL_b \leq 1$. This approach can be seen as an approximation of a confidence level built obtained in the *signal only* hypothesis [4].

As an illustration of this procedure we show in Table 8.10 a test on the CL_s and CL_{s+b} methods in the two simple cases mentioned above.

Parameter dependent exclusion Rather than simply testing the hypothesis (s+b) with the Higgs production rate predicted by the Standard Model we can let the quantity $\frac{\sigma}{\sigma_{SM}}$ be a free parameter. By operating this way, given an observation, we can always exclude a Higgs whose production mode exceeds a given value.

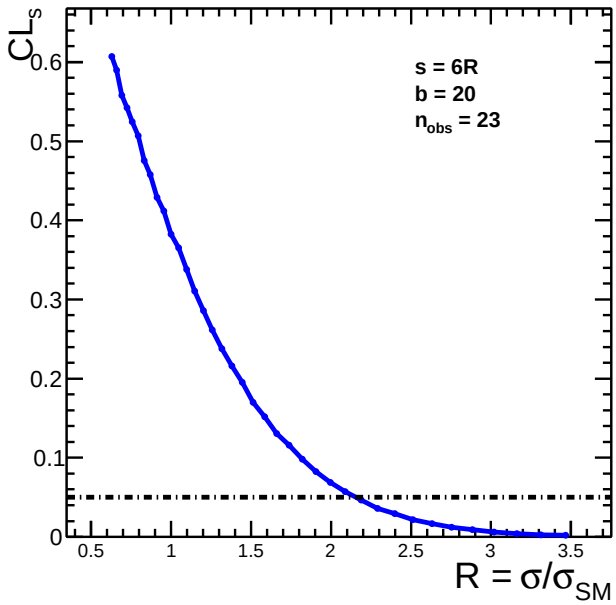


Figure 8.6: CL_s values shown as function of the parameter $R = \frac{\sigma}{\sigma_{SM}}$ for a given observation. In this toy example value a Higgs produced at a rate $\sigma > 2.2\ \sigma_{SM}$ would be excluded at the 95% confidence level.

This is shown in Figure 8.6 where we have generated several test-statistics distribution for different values of the ratio $R = \frac{\sigma}{\sigma_{SM}}$.

[4] This can be easily checked by looking at the case n = 0 where $CL_s = e^{-(s+b)}/e^{-b} = e^{-s}$

Including uncertainties So far, in our discussion we have assumed that s and b are fixed numbers. As we saw previously these numbers are subject to several uncertainties, both statistical and systematic. If after all the selection in given channel "i" (same-flavor, 0 jet bin for instance) we expect s_i and b_i signal events, in reality these numbers correspond to the best estimate for the background and the signal yield in that particular channel. In our model we should really treat b_i and s_i as a random variables.

As some of the considered uncertainties can be very large (especially in some background cases), we avoid using the normal distribution [5]. Instead every uncertainty is modeled via a log-normal distribution [6], that is always positive and in the case of small uncertainties is asymptotically identical to a gaussian. The only exception is that for statistical uncertainties, we use the "Gamma" distribution.

In practice, every time a pseudo-experiment is generated via Equation 8.6 and 8.9 the values s and b are themselves random number previously generated. For instance, in the simple case of a single channel, for log-normally distributed errors, the signal and background values are generated before each pseudo-experiment according to

$$s(\theta_1) = s \cdot \kappa_1^{\theta_1} \tag{8.15}$$

$$b(\theta_2) = b \cdot \kappa_2^{\theta_2} \tag{8.16}$$

where,

$$\kappa_1 = 1 + \frac{\delta s}{s} \tag{8.17}$$

$$\kappa_2 = 1 + \frac{\delta b}{b} \tag{8.18}$$

and θ_1,θ_2 are random variables generated from a log-normal distribution with respective parameters κ_1 and κ_2. Uncertainties have the overall effect, as expected, of spreading the (b) and (s+b) test-statistics, thus increasing the overlap (see Figure 8.7). As a result the exclusion limit of observations gets worse.

Combining different channels For every mass point (m_H), final state (fs) and jet bin (N_{jet}) optimization we have:

- expected number of signal events, $s(m_H,fs,N_{jet})$ and several sources of uncertainties $\delta\ s_j(m_H, fs, N_{jet})$
- expected number of background events, $b_i(m_H,fs,N_{jet})$ and several sources of uncertainties $\delta\ b_{ij}(m_H, fs, N_{jet})$

As we have seen in Section 8.2 some uncertainties are common for the s and b_i quantities. These uncertainties are then assumed to be fully correlated for the signal and the backgrounds that are affected by them. Such is the case for the luminosity and the efficiency (for non-data driven backgrounds). Uncertainties that are correlated are modelled by Equation 8.16 by taking the same value of θ in such a way that variations of s and b "happen together". Uncorrelated uncertainties are simply modelled by 8.16 with different random numbers.

When combining different final states or jet bins, the expected signal and backgrounds events yields are simply added and their corresponding uncertainties, when they belong to the same source, are also assumed to be fully correlated.

[5] As this can lead to unphysical downward fluctuations and negative signal/background

[6] A variable "θ" is log-normally distributed if $\ln(\theta)$ is normally distributed

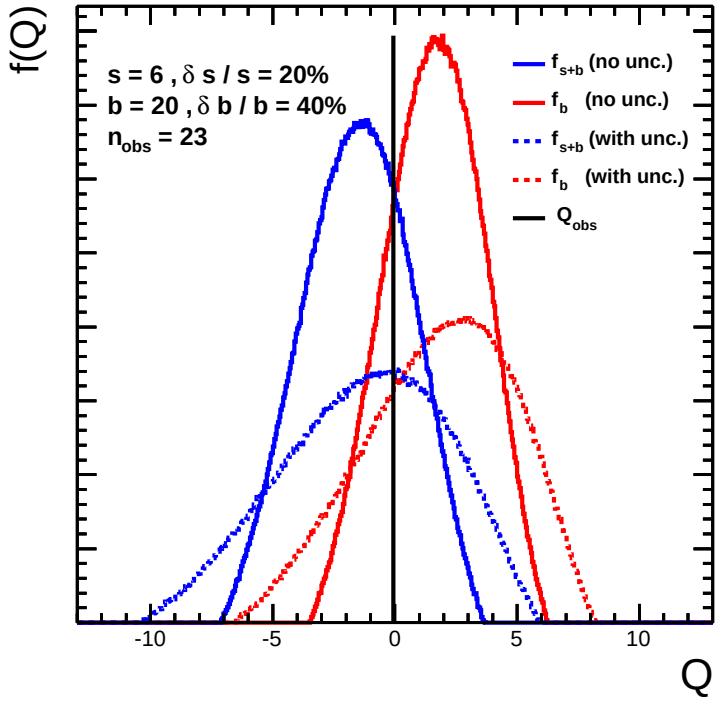


Figure 8.7: Comparison of test statistics with or without systematics

8.4 Results

The final upper limit exclusions are presented in the following way. For each mass point we give the expected (R_{exp}) and the observed exclusion limits (R_{obs}).

- The observed limit is simply computed with the observed number of events.
- The expected limit is obtained by setting the number of observed events equal to the total background expectation. This means that for a given value R_{exp}, cross sections $\sigma \geq R \cdot \sigma_{SM}$ would be excluded at the 95% confidence level if the observation would match the background expectation. This quantity serves as a baseline to determine the sensitivity we have for each mass dependent optimization. The lower it is, the higher is the sensitivity. We also compute the 1 and 2 σ bands. These are obtained by simply varying the number observed events (in this case equal to the number of background events) respectively by 1 or 2 standard deviations with respected to the nominal background expectation [7].

For each mass point we also show the value $R = 1$. Values of R exceeding 1 mean that a Higgs boson produced at the rate predicted by the Standard Model cannot be excluded. Conversely for R below 1, we can affirm that the Higgs is excluded, for that particular mass m_H. In Figures 8.8 (Cut-based 0 and 1 jet), 8.9 (NN 0 and 1 jet) and 8.10 (0 and 1 jet combined in Cut-Based and NN) we show the upper limits obtained with 1.1 fb^{-1}. The dotted line corresponds to the expected limit, the green and yellow bands are the 1

[7] The standard deviation is defined according to the test statistic for the background hypothesis.

Mass	Observed	Median Expected	68% probability band	95% probability band
120	3.9	2.9	[2.1, 4.0]	[1.5, 5.3]
130	2.0	1.4	[1.0, 2.0]	[0.7, 2.6]
140	1.0	0.8	[0.6, 1.1]	[0.4, 1.5]
150	0.8	0.6	[0.4, 0.8]	[0.3, 1.1]
160	0.5	0.3	[0.2, 0.4]	[0.1, 0.6]
170	0.6	0.4	[0.2, 0.5]	[0.2, 0.7]
180	0.4	0.5	[0.3, 0.7]	[0.2, 0.9]
190	1.0	0.8	[0.6, 1.1]	[0.4, 1.5]
200	1.5	1.1	[0.8, 1.6]	[0.6, 2.2]
250	1.9	2.3	[1.6, 3.2]	[1.2, 4.2]
300	2.4	2.5	[1.8, 3.4]	[1.3, 4.6]
350	2.2	2.2	[1.6, 3.1]	[1.2, 4.1]
400	2.2	2.5	[1.8, 3.4]	[1.3, 4.6]
450	2.3	3.3	[2.3, 4.6]	[1.8, 6.1]
500	3.0	4.5	[3.3, 6.3]	[2.4, 8.4]
550	3.4	6.5	[4.6, 9.0]	[3.5, 12.0]

Table 8.11: Cut based analysis expected and observed Upper Limits with 1.1 $fb^{-1}$of data for the 0 and 1 jet bin combined.

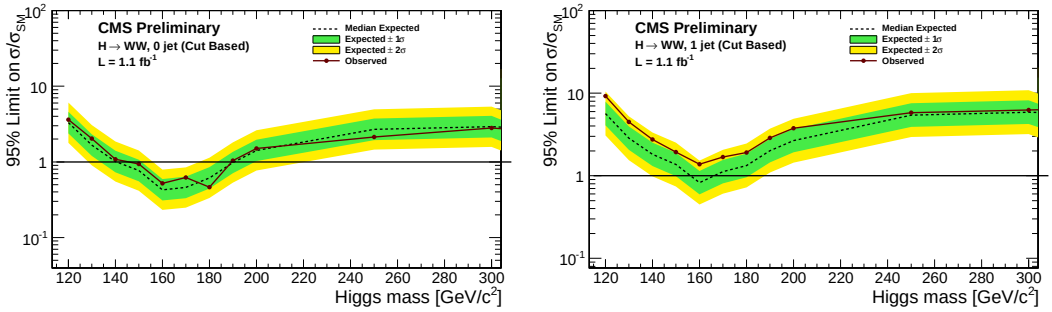
Mass	Observed	Median Expected	68% probability band	95% probability band
120	3.6	2.8	[2.0 , 4.0]	[1.5 , 5.3]
130	1.7	1.3	[1.0 , 1.9]	[0.7 , 2.5]
140	1.0	0.8	[0.6 , 1.2]	[0.4 , 1.5]
150	0.8	0.5	[0.3 , 0.7]	[0.2 , 0.9]
160	0.5	0.3	[0.2 , 0.4]	[0.1 , 0.5]
170	0.5	0.3	[0.2 , 0.4]	[0.1 , 0.5]
180	0.3	0.4	[0.2 , 0.5]	[0.2 , 0.7]
190	0.8	0.7	[0.5 , 1.0]	[0.3 , 1.3]
200	1.1	0.9	[0.6 , 1.2]	[0.4 , 1.6]
250	1.7	1.7	[1.2 , 2.3]	[0.9 , 3.1]
300	1.2	1.8	[1.3 , 2.6]	[1.0 , 3.4]
350	1.4	1.7	[1.2 , 2.4]	[0.9 , 3.2]
400	2.6	2.1	[1.5 , 2.9]	[1.1 , 3.8]
450	3.3	2.8	[2.0 , 3.9]	[1.5 , 5.2]
500	3.3	3.8	[2.7 , 5.3]	[2.0 , 7.0]
550	5.0	5.6	[4.0 , 7.8]	[3.0 , 10.4]

Table 8.12: Neural Network analysis expected and observed Upper Limits with 1.1 $fb^{-1}$of data for the 0 and 1 jet bin combined.

and 2 σ bands and the red-line is the observed limit. The results are summarized for each mass point in Table 8.11 and 8.12.

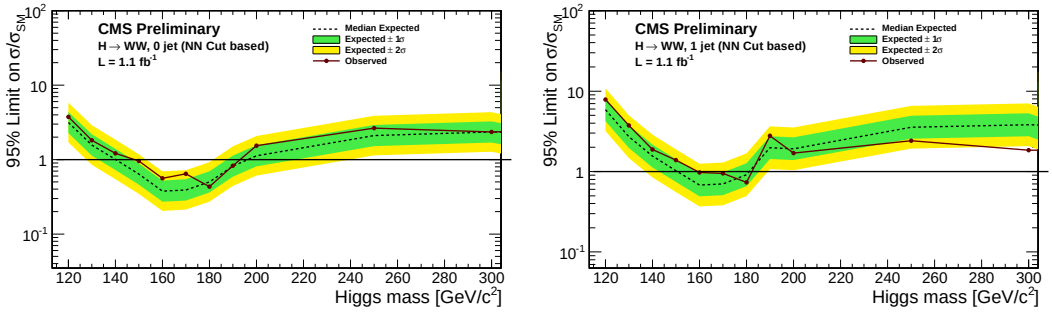
8.4.1 Cut-based results

The dotted line shows that, if our observation matched exactly the background expectation, we would be able to exclude Higgs boson masses $137 \leq m_H \leq 195$ GeV (Figure 8.10(a)) with 1.1 fb^{-1}. However, the observed upper limit (red line) lies above the expected one in the mass range $120 \leq m_H \leq 220$ GeV, indicating an overall excess observed in data. For low masses $m_H \leq 160$ GeV the excess corresponds to approximately 1 standard deviation with respect to the background expectation. This results in a slightly worse observed exclusion range of $145 \leq m_H \leq 190$ GeV.



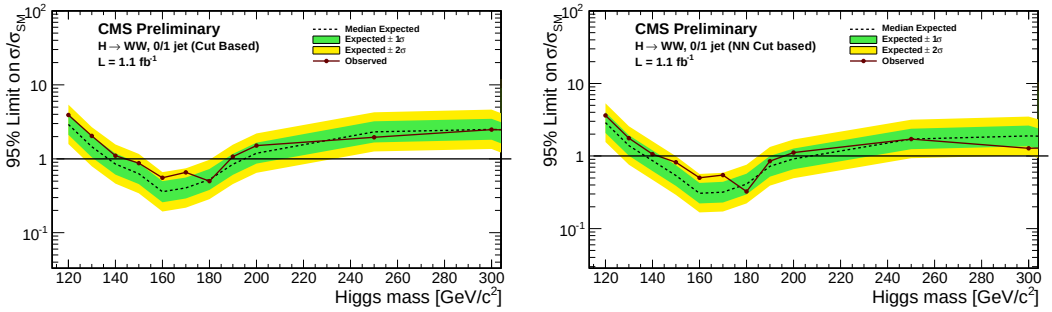
(a) (b)

Figure 8.8: Cut-based analysis upper limits at the 95% confidence level using 1.1 fb^{-1} of data for the 0-jet bin (a) and the 1-jet bin (b).



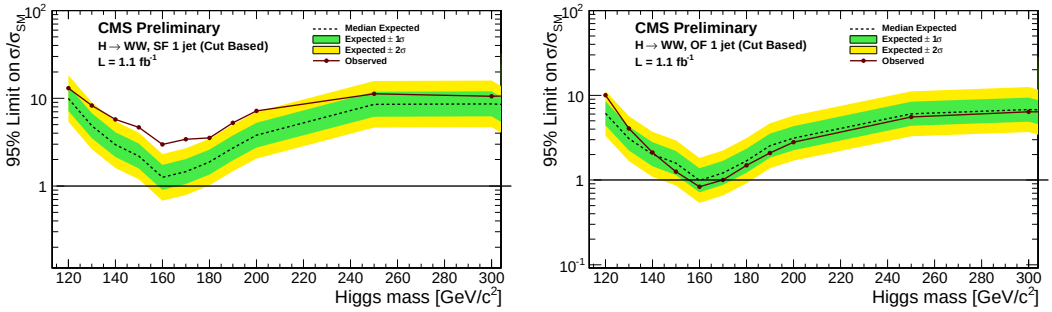
(a) (b)

Figure 8.9: Neural Network analysis upper limits at the 95% confidence level using 1.1 fb^{-1} of data for the 0-jet bin (a) and the 1-jet bin (b).



(a) (b)

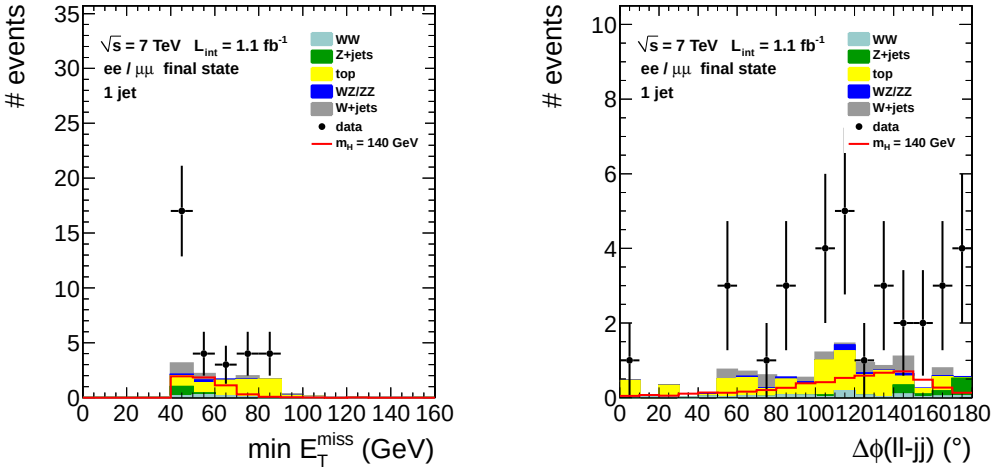
Figure 8.10: Combined upper limits at the 95% confidence level using 1.1 fb^{-1} of data for the cut-based (a) and the neural network (b) analysis.



(a) (b)

Figure 8.11: Cut-based analysis upper limits at the 95% confidence level using 1.1 fb^{-1} of data in the 1 jet bin for the same flavor (a) and opposite flavor final state (b).

The observed excess is driven both by the 0 and 1 jet bin analysis. In the latter case, the excess corresponds to 1.5 σ in the whole mass range. In Figure 8.11 we show the upper limits in the 1 jet bin for the same and opposite flavor final states. We see a significant excess (2 σ) in the same-flavor final state events confirmed by Table 8.7.



(a) (b)

Figure 8.12: Events surviving the final $m_H = 140$ GeV cut-based selection (omitting the m_T cut)shown in the angle between the di-lepton and the di-jet (of the two most energetic jets) system (a) and the minE_T^{miss} (b) distributions.

Due to the broad nature of this excess, it is hard to interpret it as a possible signal. Also, it appears to be present only in the same-flavor events, which might be a hint that we are rather looking at a possibly underestimated Drell-Yan contribution. Finally this excess is equally distributed among the ee and $\mu\mu$ final state, which confirms the Drell-Yan hypothesis. As we are looking at the one jet bin, these could be Z+2jets events, with the second jet being close to the 30 GeV threshold. We expect a typical Z+2jets event to have the di-jet system recoiling against the di-lepton system with a big opening angle. This can

be checked in the corresponding distribution in Figure 8.12(a). Finally we can also see that the excess event have minE_T^{miss} close to the 40 GeV cut (Figure 8.12(b)). This last observation further supplements the hypothesis of an underestimate of this background.

8.4.2 Neural Network results

The Neural Network based analysis leads to a slightly better sensitivity with respect to the cut-based analysis. The expected exclusion range obtained by combining both the 0 and 1 jet analysis is larger: $137 \leq m_H \leq 205$ GeV(Figure 8.10(b)). Starting from middle-range masses ($m_H \geq 150$ GeV) the expected upper limit exclusion is slightly smaller than in the cut-based case, indicating a better separation of the signal and background hypothesis. A clear advantage of this method can be seen in the 1-jet bin analysis where the expected exclusion is $150 \leq m_H \leq 180$ GeV, whereas in the cut-based case we are barely able to exclude $155 \leq m_H \leq 165$ GeV. Also a significant improvement is observed in the 0-jet bin analysis where the upper bound of the exclusion is extended up to 200 GeV. However, no significant improvement is observed at low masses ($m_H \leq 140$ GeV).

In data we observe an overall 1σ excess. The reason for this excess was investigated in the cut based analysis case. The observed upper limit is also slightly better than in the cut-based scenario. We are are able to exclude with 1.1 fb^{-1} a Standard Model Higgs boson with a mass $145 \leq m_H \leq 195$ GeV showing a very small improvement with respect to the cut-based analysis. As for the expected limit, no improvement is seen at low masses.

8.5 Conclusion

In this Chapter two approaches for signal extraction were presented. The cut-based analysis is a simpler approach, based on the selection of a particular phase space in which the signal over background ratio is enhanced. The Neural Network optimization relies on the construction of a single discriminant output that is able to separate the signal and background hypothesis via a learning process. Although slightly more performant, we believe the second approach to be less robust due to its non-linear nature. Both approaches show similar performances with the current data gathered with the CMS detector. As the final result we quote the exclusion range obtained with the cut-based selection rather than the one obtained with the Neural Network analysis. The latter can be seen as a cross-check of the final result. As shown in Figure 8.13, with the data gathered by the CMS experiment corresponding to 1.1 fb^{-1}, we are able to exclude a Standard Model Higgs Boson with a mass $145 \leq m_H \leq 190$ GeV.

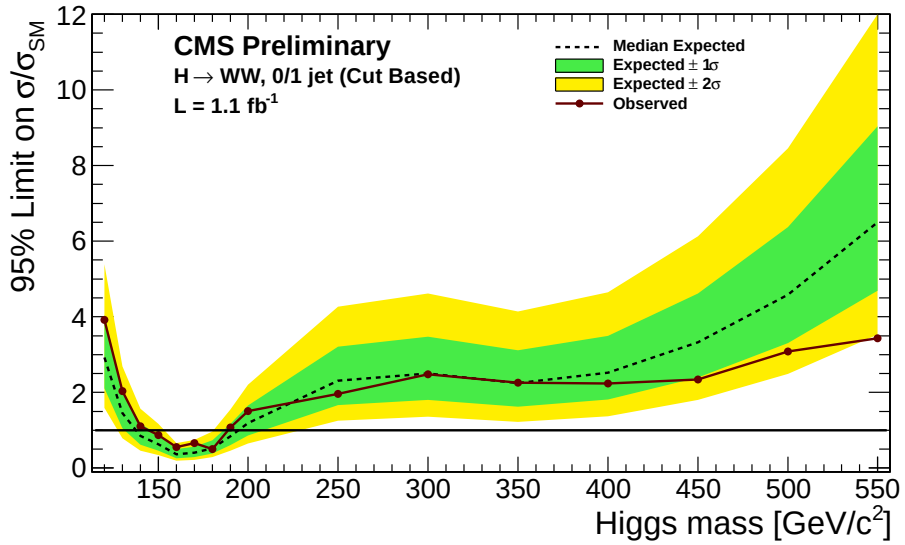


Figure 8.13: Combined upper limits at the 95% confidence level using 1.1 fb^{-1} of data for the cut-based analysis in the full mass range.

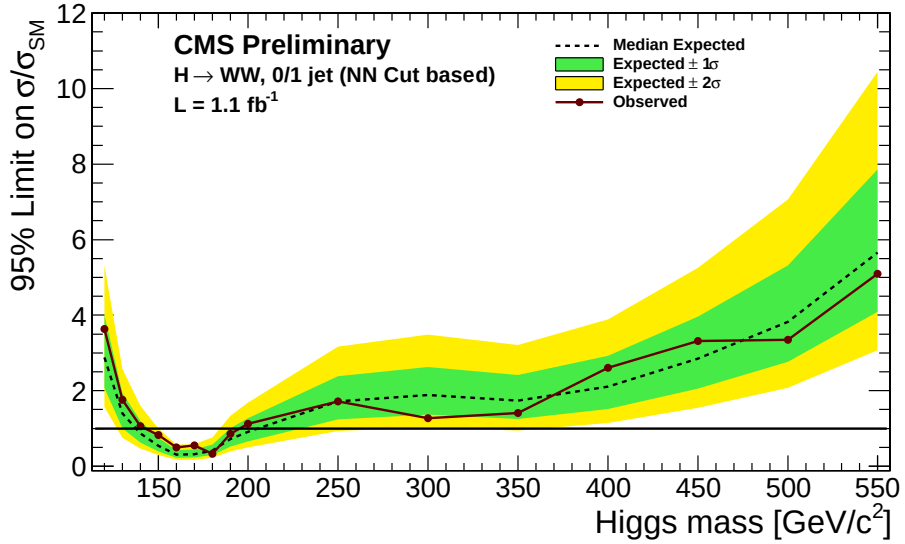


Figure 8.14: Combined upper limits at the 95% confidence level using 1.1 fb^{-1} of data for the neural network analysis in the full mass range.

Conclusions

In this thesis we perform a search of the Brout-Englert-Higgs Boson using 1.1 fb^{-1} of LHC data collected at a center of mass energy $\sqrt{s} = 7$ TeV with the CMS detector. The main possible Higgs decay channels are introduced. The presented search is performed considering the $H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}$ decay mode, which is the most sensitive in terms of discovery and exclusion potential in the mass range $m_H = 120 - 200$ GeV/c^2. A short introduction on the signal and its main background is given. In particular it is shown that a given definition of a phase space volume can sensibly increase the signal-over-background ratio. The full analysis is performed both with sequential selection criteria and with neural networks. Final upper limits on the Higgs production cross section are obtained in these two scenario. We conclude that, with 1.1 fb^{-1} of data, a BEH boson of mass $145 \leq m_H \leq 190$ GeV/c^2 is excluded at the 95% confidence level.

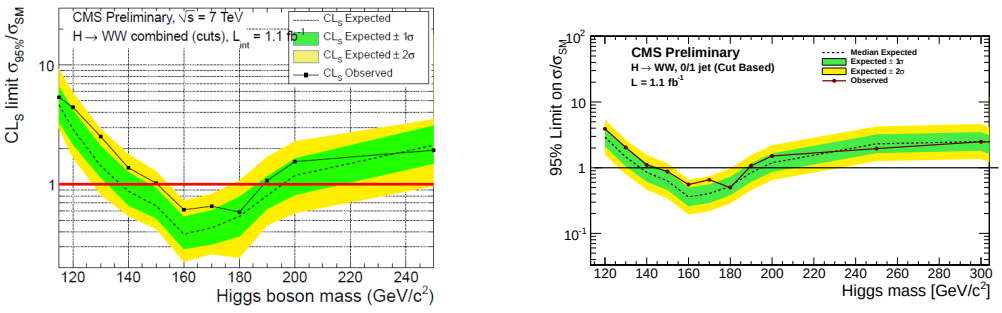


Figure 8.15: Official CMS combined upper limits at the 95% confidence level using 1.1 fb^{-1} of data for the cut-based analysis (left) compared to the upper limits obtained in this thesis (right).

If we compare our results, obtained independently, with the official CMS results, we observe a slight difference in the exclusion limit. The CMS exclusion limit as function of the Higgs mass is shown in Figure 8.15. The obtained exclusion is in that case $150 \leq m_H \leq 190$ GeV/c^2. This slight difference was carefully studied, and no major cause could be found. The difference is likely to be driven by a slightly different estimate ($\sim$ 5%) of the WW background in the opposite flavor channel, 0 jet bin for the $m_H = 140$ GeV/c^2 optimization. However, overall the results are compatible, if one considers the 1 and 2 σ bands.

A Standard Model BEH boson is more likely to have a low mass, if one believes indirect constraints from measured Standard Model parameters. If we take into account exclusion

from the LEP accelerator, a low mass Higgs is then expected to be discovered in the mass range $m_H = 115 - 145$ GeV/c^2.

In this thesis we have analyzed 1.1 fb^{-1}, a part of the total dataset corresponding to 4.8 fb^{-1} of integrated luminosity. The final CMS combination, including all BEH searches is shown in Figure 8.16. These results have further constrained the Higgs mass parameter. A Higgs Boson with a mass $m_H \geq 127$ GeV/c^2 is excluded, leaving a small range of possible masses. Furtheremore a clear 2.5σ excess is observed at approximately $m_H = 125$ GeV/c^2 , which is also seen in Tevatron data, shown in Figure 2.7.

In April 2012, the LHC has started taking data at $\sqrt{s} = 8$ TeV and high instantatenous luminosity. By the end of 2012, CMS will have collected appriximately an additional 15 fb^{-1} of integrated luminosity. This amount of data should be sufficient to finally either confirm the observed excess or show that this was the result of a misunderstood fluctuation. Although the H→W$^+$W$^-$ can add some sensitivity at low mass, the main search mode will be the H → $\gamma\gamma$ channel. In the near future we will then have a clearer picture on the nature of the electroweak symmetry breaking mechanism, and hopefully, on many other opened questions.

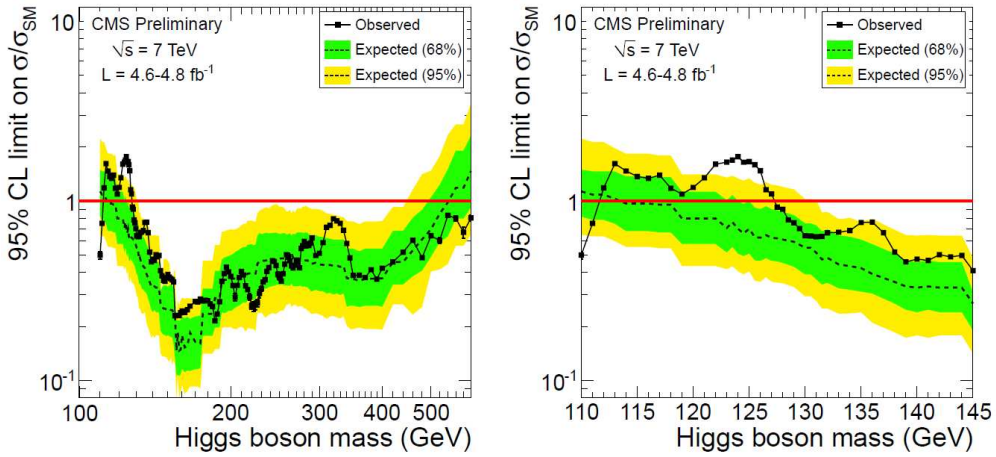


Figure 8.16: Combined upper limits at the 95% confidence level using 4.8 fb^{-1} with the CMS detector (left) with a zoom in the low mass region (right) showing an excess at $m_H = 125$ GeV/c^2.

Acknowledgements

This work was carried out over a period of four years at the University of Antwerp, with periodic stays at CERN. I would first like to thank all the people in the Elementary Particle Physics department for providing a pleasant working environment. In particular, I am grateful to Prof. Pierre Van Mechelen and Prof. Eddi de Wolf for accepting me in their group, despite my unexperience in experimental particle physics.

During this period I have had the chance to work in a mutable environment. Many people have left in these last years. In particular I would like to thank Majid Hashemi for helping me at the beginning of my PhD. Thanks for your patience and your advices. I wish you the best. I wish also to thank my office mate Krzysztof Kutak, for friendly and fruitful discussion about QCD. Thanks also to Damiano Tommasini, who helped me in understanding the subtelties of resummation.

I am particularly grateful to Giuseppe Zito and Suchandra Dutta, for giving me the opportunity to work in the DQM tracker community.

A special word goes to my promotor Prof. Nick van Remortel, for giving me the opportunity to work under his supervision. Besides always being present and friendly, I have had the great opportunity to profit from his experience. Thanks for everything. A special thanks goes to Xavier Janssen, who I have had the chance to work with in the last two years. His dedication and thoroughness have been very motivating for me. I would also like to thank all the persons I have had a chance to work with, directly or not in the Higgs to WW subgroup. Thanks to Guillelmo Gomez Ceballos, who always answered with great patience to my several questions. Also I would like to say thanks the Latinos group, with whom I have worked in 2011.

I am grateful to Romain, for always patiently solving my coding issues. I would also like to thank Luca and Zlate, with whom I have shared the writing experience, for their friendship, help and support. Thanks to Annachiara for her continuous support.

I am grateful to my parents, for supporting and encouraging my choices, and for being an example of intellectual curiosity and perseverance. Finally thanks to all my friends, for always being present.

Appendix A

Trigger tables and efficiency

A.1 Trigger definitions

Here we give the definitions of the trigger used in the analysis. In Table A.1 the HLT path names a given for Dielectron triggers. The number after the label "Ele" indicates the momentum threshold on the electron. The other labels that follow indicate additional isolation and identification requirements. They are explained in Table A.1. The following tables list the DiMuon, Mixed, and Single Muon triggers.

Dataset	Trigger name	L1 sel.	HLT sel.
DoubleElectron	HLT_Ele17_CaloIdL_CaloIsoVL_ Ele8_CaloIdL_CaloIsoVL	$p_T > 12$	$p_T > 17,8$ GeV/c
DoubleElectron	HLT_Ele17_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL_ Ele8_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL	$p_T > 12$	$p_T > 17,8$ GeV/c

Table A.1: Analysis triggers for the ee final state. The identification and isolation requirements are described in Table A.1.

A.2 Double lepton trigger efficiencies

The following efficiencies are used to correct the signal MonteCarlo by applying an event-by-event weight that p_T and $|\eta|$ dependent. Details on how these were estimated can be found in [90]. The overall effect of this reweighting procedure amounts to $\sim 1\%$.

name	criterion
CaloId_L	H/E $<$ 0.15(0.10) $\sigma_{\eta\eta}<$ 0.014 (0.035)
CaloId_VT	H/E $<$ 0.05(0.05) $\sigma_{\eta\eta}<$ 0.011 (0.031)
TrkId_VL	$\|\Delta\eta\|<$ 0.01 (0.01) $\Delta\phi<$ 0.15 (0.10)
TrkId_T	$\|\Delta\eta\|<$ 0.008 (0.008) $\Delta\phi<$ 0.07 (0.05)
CaloIso_VL	ECalIso/E_T $<$ 0.2 (0.2) HCalIso/E_T $<$ 0.2 (0.2)
CaloIso_T	ECalIso/E_T $<$ 0.15 (0.075) HCalIso/E_T $<$ 0.15 (0.075)
CaloIso_VT	ECalIso/E_T $<$ 0.05 (0.05) HCalIso/E_T $<$ 0.05 (0.05)
TrkIso_VL	TrkIso/E_T $<$ 0.2 (0.2)
TrkIso_T	TrkIso/E_T $<$ 0.15 (0.075)
TrkIso_VT	TrkIso/E_T $<$ 0.05 (0.05)

Table A.2: Summary of requirements applied to electrons in the triggers used for this analysis. The selection requirements are given for electrons in the barrel (endcap). L=Loose, VL=Very loose, T=Tight, VT=Very Tight.

Dataset	Trigger name	L1 sel.	HLT sel.
MuEG	HLT_Mu17_Ele8_CaloIdL	$p_T > 3,5$ GeV/c	$p_T > 17,8$ GeV/c
MuEG	HLT_Mu8_Ele17_CaloIdL	$p_T > 3,5$ GeV/c	$p_T > 8,17$ GeV/c

Table A.3: Analysis triggers for the $e\mu$ final state. The identification and isolation requirements for electrons are described in Table A.1.

Table A.4: Analysis triggers for the $\mu\mu$ final state. Triggers marked (*) are also used for efficiency studies.

Dataset	Trigger name	L1 sel.	HLT sel.
DoubleMu	HLT_DoubleMu6	$p_T > 3,3$ GeV/c	$p_T > 6,6$ GeV/c
DoubleMu	HLT_DoubleMu7	$p_T > 3,3$ GeV/c	$p_T > 7,7$ GeV/c
DoubleMu	HLT_Mu13_Mu8	$p_T > 3,3$ GeV/c	$p_T > 13,8$ GeV/c
DoubleMu	HLT_Mu17_Mu8	$p_T > 3,3$ GeV/c	$p_T > 17,8$ GeV/c

Dataset	Trigger name	L1 sel.	HLT sel.
SingleEle	HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT	L1_SingleEG15	$p_T > 27$ GeV/c
SingleMu	HLT_IsoMu12	L1_SingleMu7	$p_T > 12$ GeV/c
SingleMu	HLT_IsoMu17	L1_SingleMu10	$p_T > 17$ GeV/c
SingleMu	HLT_Mu15	L1_SingleMu10	$p_T > 15$ GeV/c

Table A.5: Single lepton triggers to recover lost efficiency. These trigges are also used for efficiency studies. The identification and isolation requirements for electrons are described in Table A.1.

Measurement	$0.00 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2.50$
$10 < p_T < 15$	0.99 ± 0.00	0.99 ± 0.01
$15 < p_T < 20$	1.00 ± 0.00	1.00 ± 0.00
$p_T > 20$	1.00 ± 0.00	1.00 ± 0.00

Table A.6: Overall double electron trigger efficiency as a function of p_T and $|\eta|$.

Measurement	$0.00 < \|\eta\| < 0.80$	$0.80 < \|\eta\| < 1.20$	$1.20 < \|\eta\| < 2.40$
$10 < p_T < 15$	0.94 ± 0.02	0.92 ± 0.02	0.94 ± 0.01
$15 < p_T < 20$	0.97 ± 0.01	0.96 ± 0.01	0.96 ± 0.00
$p_T > 20$	0.97 ± 0.00	0.95 ± 0.00	0.95 ± 0.00

Table A.7: Overall double muon trigger efficiency as a function of p_T and $|\eta|$.

Appendix B

Tables for the common preselection

B.1 Selection down to the jet count

In Table B.1- B.6 we show the background and signal (for some Higgs masses) expectations and the data yields at each step of the WW common preselection up to the number of jets selection.

selection	Z+jets	$t\bar{t}$	Single Top	W+jets
HLT + lepton sel.	771125.1 ± 253.3	5256.4 ± 32.9	387.4 ± 3.7	320.3 ± 32.8
$m_{\ell\ell} > 12$	770179.4 ± 252.3	5243.1 ± 32.9	386.3 ± 3.6	312.4 ± 32.4
Z veto	107056.6 ± 101.2	4677.0 ± 31.0	345.8 ± 3.4	300.0 ± 31.9
min E_T^{miss}	83.1 ± 3.2	2635.1 ± 23.3	203.9 ± 2.6	118.5 ± 20.2

Table B.1: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity before the jet bin definition. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	WZ/ZZ	gg → WW	qq → WW	W/Z+γ
HLT + lepton sel.	985.6 ± 3.1	43.2 ± 0.3	1091.6 ± 3.2	6946.7 ± 39.7
$m_{\ell\ell} > 12$	985.2 ± 3.1	42.9 ± 0.3	1088.1 ± 3.2	6937.3 ± 39.7
Z veto	136.5 ± 1.1	38.9 ± 0.3	972.5 ± 3.0	4699.6 ± 32.7
min E_T^{miss}	30.0 ± 0.6	27.4 ± 0.2	518.2 ± 2.2	14.8 ± 1.8

Table B.2: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity before the jet bin definition. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{120}	H_{130}	H_{140}	H_{150}
HLT + lepton sel.	45.9 ± 0.3	99.2 ± 0.5	162.0 ± 0.8	216.0 ± 1.0
$m_{\ell\ell} > 12$	44.2 ± 0.3	96.5 ± 0.5	158.6 ± 0.8	212.5 ± 1.0
Z veto	44.1 ± 0.3	95.5 ± 0.5	155.4 ± 0.8	206.1 ± 1.0
min E_T^{miss}	25.2 ± 0.2	58.3 ± 0.4	102.1 ± 0.6	145.7 ± 0.8

Table B.3: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity before the jet bin definition. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{160}	H_{170}	H_{180}	H_{190}
HLT + lepton sel.	259.9 ± 1.2	250.1 ± 1.1	208.2 ± 0.9	159.6 ± 0.7
$m_{\ell\ell} > 12$	256.4 ± 1.2	247.2 ± 1.1	206.5 ± 0.9	158.6 ± 0.7
Z veto	250.4 ± 1.1	237.6 ± 1.1	190.1 ± 0.9	140.7 ± 0.6
min E_T^{miss}	193.4 ± 1.0	190.2 ± 1.0	152.4 ± 0.8	109.8 ± 0.6

Table B.4: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity before the jet bin definition. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{200}	H_{250}	H_{300}	H_{350}
HLT + lepton sel.	144.5 ± 0.6	94.6 ± 0.4	71.5 ± 0.3	67.1 ± 0.3
$m_{\ell\ell} > 12$	143.7 ± 0.6	94.4 ± 0.4	71.4 ± 0.3	67.1 ± 0.3
Z veto	125.0 ± 0.5	81.4 ± 0.3	63.8 ± 0.3	61.7 ± 0.2
min E_T^{miss}	93.5 ± 0.5	52.9 ± 0.3	38.6 ± 0.2	36.0 ± 0.2

Table B.5: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity before the jet bin definition. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{400}	H_{450}	H_{500}	H_{550}
HLT + lepton sel.	52.0 ± 0.2	34.4 ± 0.1	22.2 ± 0.1	14.5 ± 0.1
$m_{\ell\ell} > 12$	51.9 ± 0.2	34.4 ± 0.1	22.2 ± 0.1	14.5 ± 0.1
Z veto	48.9 ± 0.2	32.8 ± 0.1	21.3 ± 0.1	14.0 ± 0.1
min E_T^{miss}	28.0 ± 0.1	18.4 ± 0.1	11.8 ± 0.1	7.6 ± 0.0

Table B.6: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity before the jet bin definition. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

B.2 Selection down to the WW preselection

B.2.1 0 jet bin

In Table B.7- B.12 we show the background and signal (for some Higgs masses) expectations and the data yields at each step of the WW common preselection up to the number of jets selection.

cut level	Z+jets	$t\bar{t}$	Single Top	W+jets
0 jet count	9.7 ± 1.3	58.0 ± 3.5	22.9 ± 0.9	77.0 ± 16.1
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	8.2 ± 1.2	56.6 ± 3.5	22.3 ± 0.9	77.0 ± 16.1
soft muon	8.1 ± 1.2	46.6 ± 3.1	20.2 ± 0.8	77.0 ± 16.1
b-tag discr. > 2.1	7.5 ± 1.2	23.7 ± 2.2	12.8 ± 0.7	77.0 ± 16.1

Table B.7: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 0 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

cut level	WZ/ZZ	gg → WW	qq → WW	W/Z+γ
0 jet count	15.2 ± 0.4	18.9 ± 0.2	360.8 ± 1.8	10.1 ± 1.6
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	14.7 ± 0.4	18.5 ± 0.2	355.9 ± 1.8	10.1 ± 1.6
soft muon	13.8 ± 0.4	18.5 ± 0.2	355.0 ± 1.8	10.1 ± 1.6
b-tag discr. > 2.1	13.5 ± 0.4	18.1 ± 0.2	347.9 ± 1.8	9.8 ± 1.5

Table B.8: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 0 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

cut level	H_{120}	H_{130}	H_{140}	H_{150}
0 jet count	16.7 ± 0.2	37.5 ± 0.3	64.4 ± 0.5	89.7 ± 0.7
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	16.4 ± 0.2	36.9 ± 0.3	63.3 ± 0.5	88.1 ± 0.7
soft muon	16.4 ± 0.2	36.8 ± 0.3	63.1 ± 0.5	87.8 ± 0.7
b-tag discr. > 2.1	16.1 ± 0.2	36.1 ± 0.3	61.7 ± 0.5	86.2 ± 0.7

Table B.9: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 0 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

cut level	H_{160}	H_{170}	H_{180}	H_{190}
0 jet count	116.1 ± 0.9	111.3 ± 0.8	86.7 ± 0.6	61.0 ± 0.5
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	113.6 ± 0.8	108.6 ± 0.8	84.6 ± 0.6	59.9 ± 0.5
soft muon	113.3 ± 0.8	108.2 ± 0.8	84.3 ± 0.6	59.6 ± 0.5
b-tag discr. > 2.1	110.8 ± 0.8	106.0 ± 0.8	82.7 ± 0.6	58.4 ± 0.5

Table B.10: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 0 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

cut level	H_{200}	H_{250}	H_{300}	H_{350}
0 jet count	49.2 ± 0.4	25.6 ± 0.2	17.4 ± 0.1	15.7 ± 0.1
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	48.3 ± 0.4	25.2 ± 0.2	17.0 ± 0.1	15.4 ± 0.1
soft muon	48.1 ± 0.4	25.0 ± 0.2	17.0 ± 0.1	15.3 ± 0.1
b-tag discr. > 2.1	47.0 ± 0.4	24.5 ± 0.2	16.6 ± 0.1	14.9 ± 0.1

Table B.11: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 0 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

cut level	H_{400}	H_{450}	H_{500}	H_{550}
0 jet count	11.6 ± 0.1	7.1 ± 0.1	4.4 ± 0.0	2.5 ± 0.0
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	11.3 ± 0.1	7.0 ± 0.1	4.3 ± 0.0	2.5 ± 0.0
soft muon	11.3 ± 0.1	6.9 ± 0.1	4.3 ± 0.0	2.5 ± 0.0
b-tag discr. > 2.1	11.0 ± 0.1	6.8 ± 0.1	4.2 ± 0.0	2.4 ± 0.0

Table B.12: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 0 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

B.2.2 1 jet bin

In Table B.13- B.18 we show the background and signal (for some Higgs masses) expectations and the data yields at each step of the WW common preselection up to the number of jets selection.

selection	Z+jets	$t\bar{t}$	Single Top	W+jets
1 jet count	44.5 ± 2.2	519.2 ± 10.3	102.9 ± 1.9	23.3 ± 8.9
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	20.8 ± 1.5	507.3 ± 10.2	99.5 ± 1.8	18.7 ± 7.8
soft muon	20.4 ± 1.5	416.4 ± 9.3	87.2 ± 1.7	18.5 ± 7.8
b-tag discr. > 2.1	17.5 ± 1.3	78.4 ± 4.0	22.6 ± 0.9	14.7 ± 7.0

Table B.13: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 1 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	WZ/ZZ	$gg \to WW$	$qq \to WW$	W/Z+γ
1 jet count	11.0 ± 0.4	7.0 ± 0.1	120.6 ± 1.1	2.7 ± 0.7
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	10.6 ± 0.4	6.8 ± 0.1	117.3 ± 1.0	2.1 ± 0.6
soft muon	10.0 ± 0.3	6.7 ± 0.1	115.9 ± 1.0	1.8 ± 0.5
b-tag discr. > 2.1	9.5 ± 0.3	6.3 ± 0.1	107.0 ± 1.0	1.8 ± 0.5

Table B.14: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 1 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{120}	H_{130}	H_{140}	H_{150}
1 jet count	6.1 ± 0.1	14.7 ± 0.2	27.1 ± 0.3	39.7 ± 0.4
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	6.0 ± 0.1	14.4 ± 0.2	26.3 ± 0.3	38.5 ± 0.4
soft muon	5.9 ± 0.1	14.2 ± 0.2	26.0 ± 0.3	38.0 ± 0.4
b-tag discr. > 2.1	5.5 ± 0.1	13.0 ± 0.2	23.9 ± 0.3	35.2 ± 0.4

Table B.15: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 1 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{160}	H_{170}	H_{180}	H_{190}
1 jet count	54.8 ± 0.5	55.2 ± 0.5	45.0 ± 0.4	33.1 ± 0.3
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	52.9 ± 0.5	53.1 ± 0.5	43.4 ± 0.4	31.9 ± 0.3
soft muon	52.2 ± 0.5	52.4 ± 0.5	42.8 ± 0.4	31.5 ± 0.3
b-tag discr. > 2.1	48.4 ± 0.5	48.2 ± 0.5	39.4 ± 0.4	29.1 ± 0.3

Table B.16: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 1 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{200}	H_{250}	H_{300}	H_{350}
1 jet count	29.1 ± 0.2	17.3 ± 0.1	13.4 ± 0.1	12.8 ± 0.1
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	28.0 ± 0.2	16.7 ± 0.1	12.9 ± 0.1	12.3 ± 0.1
soft muon	27.6 ± 0.2	16.5 ± 0.1	12.7 ± 0.1	12.1 ± 0.1
b-tag discr. > 2.1	25.4 ± 0.2	15.0 ± 0.1	11.6 ± 0.1	11.1 ± 0.1

Table B.17: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 1 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

selection	H_{400}	H_{450}	H_{500}	H_{550}
1 jet count	10.2 ± 0.1	6.7 ± 0.1	4.3 ± 0.0	2.9 ± 0.0
$\Delta\phi_{\ell\ell-\text{jet}} < 165$	9.8 ± 0.1	6.5 ± 0.1	4.2 ± 0.0	2.8 ± 0.0
soft muon	9.6 ± 0.1	6.4 ± 0.1	4.1 ± 0.0	2.7 ± 0.0
b-tag discr. > 2.1	8.8 ± 0.1	5.8 ± 0.1	3.7 ± 0.0	2.5 ± 0.0

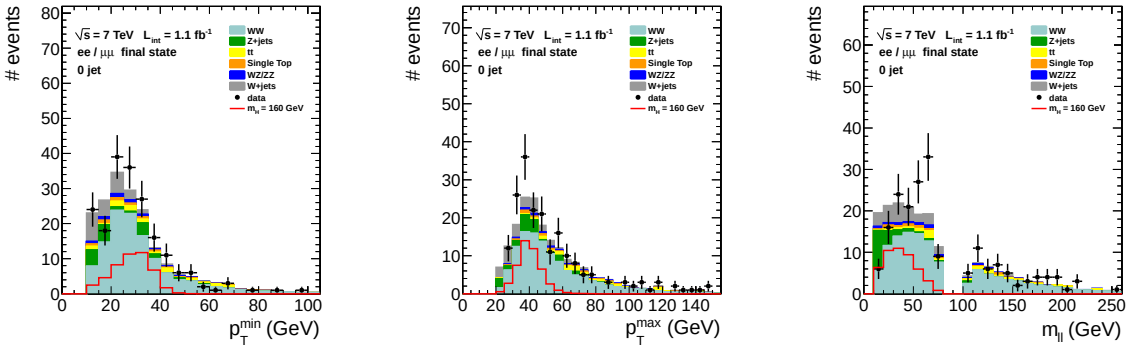
Table B.18: Cut-by-cut breakdown for the background and signal expected yields for 1.1 fb^{-1} of integrated luminosity in the 1 jet bin. Statistical uncertainty are included. The signal yields account for the data driven efficiency estimates.

Appendix C

Control Distributions at the WW preselection level

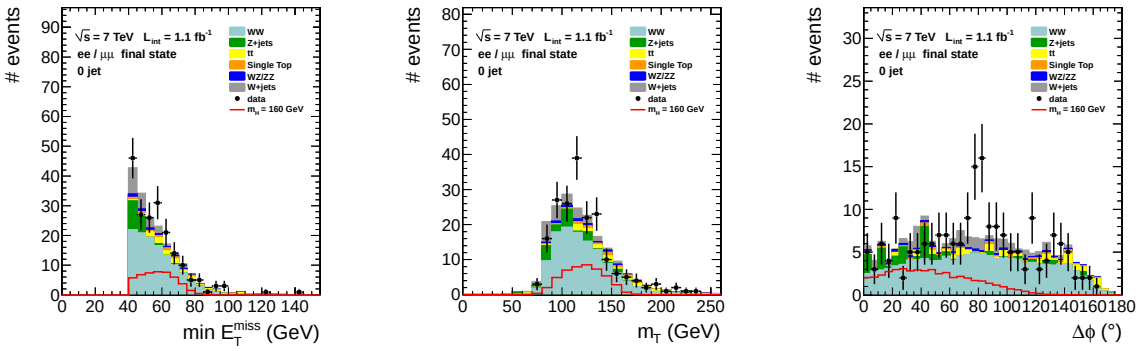
We show here the control distributions (p_T^{min},p_T^{max}, $m_{\ell\ell}$, min $E_T{}^{miss}$, m_T, $\Delta\phi$) are shown divided in same flavor and opposite flavor final states for $m_H = 160$ GeV in 0 and 1 jet bin.

C.1 0 jet bin



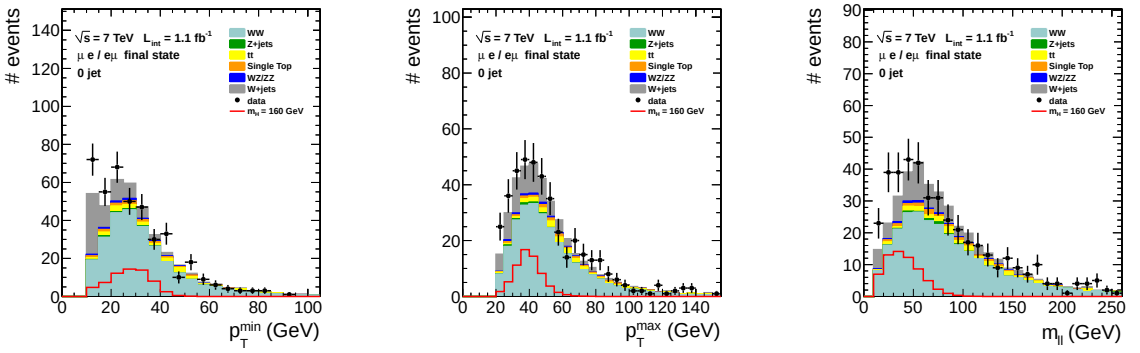
(a) (b) (c)

Figure C.1: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 0-jet bin, of the trailing lepton p_T (a), leading lepton p_T (b) and dilepton invariant mass (c) with $m_H = 160$ GeV for the same flavor final state. Each component in simulation is scaled to data-driven estimates.



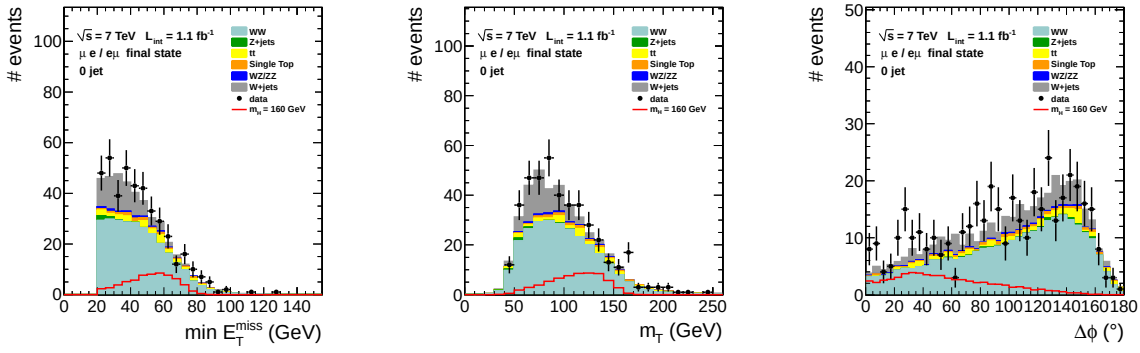
(a) (b) (c)

Figure C.2: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 0-jet bin, of the min $E_T{}^{miss}$ (a), transverse mass (b) and dilepton $\Delta\phi$ (c) with $m_H = 160$ GeV for the same flavor final state. Each component in simulation is scaled to data-driven estimates.



(a) (b) (c)

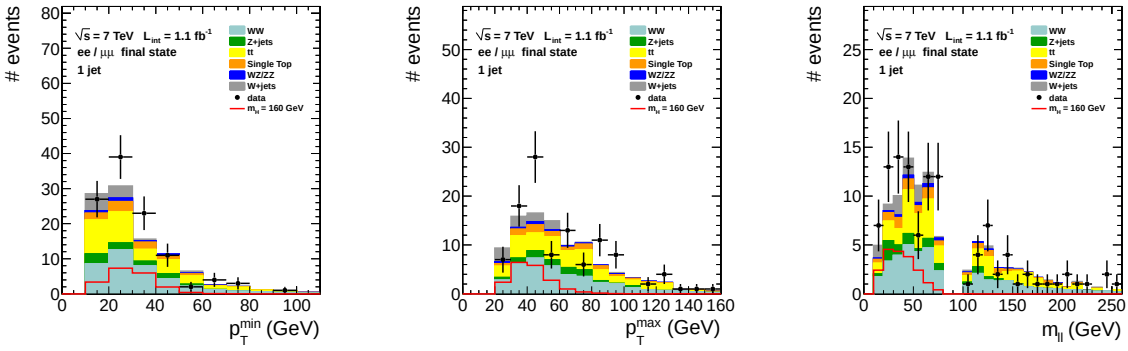
Figure C.3: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 0-jet bin, of the trailing lepton p_T (a), leading lepton p_T (b) and dilepton invariant mass (c) with $m_H = 160$ GeV for the opposite flavor final state. Each component in simulation is scaled to data-driven estimates.



(a) (b) (c)

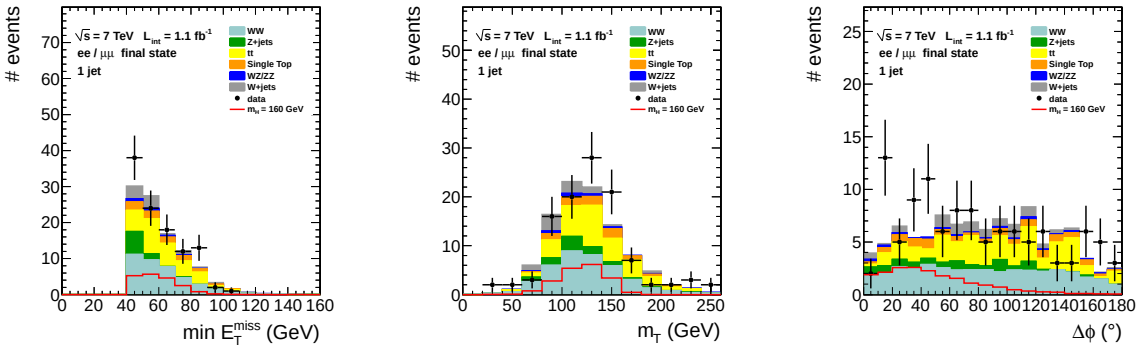
Figure C.4: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 0-jet bin, of the min $E_T{}^{miss}$ (a), transverse mass (b) and dilepton $\Delta\phi$ (c) with $m_H = 160$ GeV for the opposite flavor final state. Each component in simulation is scaled to data-driven estimates.

C.2 1 jet bin



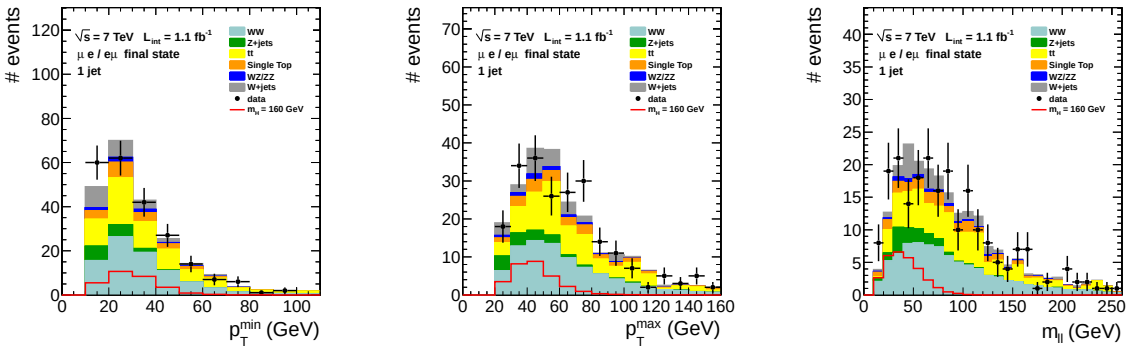
(a) (b) (c)

Figure C.5: Distribution, after WW selection for 1.1 fb^{-1}of data in the 1-jet bin, of the trailing lepton p_T (a), leading lepton p_T (b) and dilepton invariant mass (c) with $m_H = 160$ GeV for the same flavor final state. Each component in simulation is scaled to data-driven estimates.



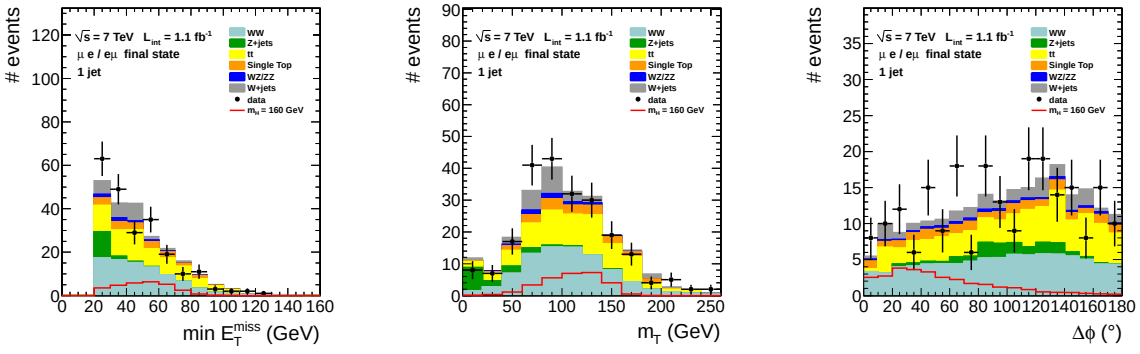
(a) (b) (c)

Figure C.6: Distribution, after WW selection for 1.1 fb^{-1}of data in the 1-jet bin, of the min E_T^{miss} (a), transverse mass (b) and dilepton $\Delta\phi$ (c) with $m_H = 160$ GeV for the same flavor final state. Each component in simulation is scaled to data-driven estimates.



(a) (b) (c)

Figure C.7: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 1-jet bin, of the trailing lepton p_T (a), leading lepton p_T (b) and dilepton invariant mass (c) with $m_H = 160$ GeV for the opposite flavor final state. Each component in simulation is scaled to data-driven estimates.



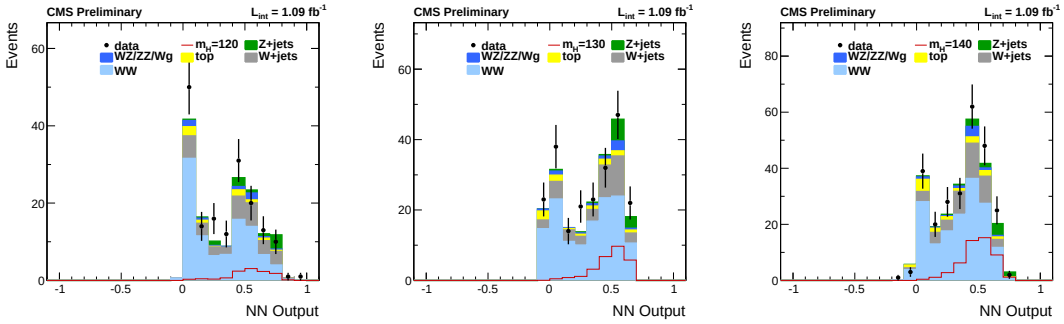
(a) (b) (c)

Figure C.8: Distribution, after WW selection for 1.1 $fb^{-1}$of data in the 1-jet bin, of the min E_T^{miss} (a), transverse mass (b) and dilepton $\Delta\phi$ (c) with $m_H = 160$ GeV for the opposite flavor final state. Each component in simulation is scaled to data-driven estimates.

Appendix D

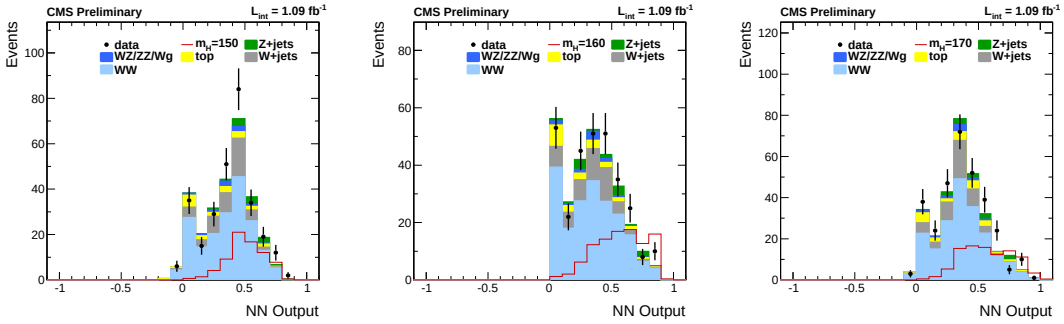
Neural Network distributions

D.1 O jet bin



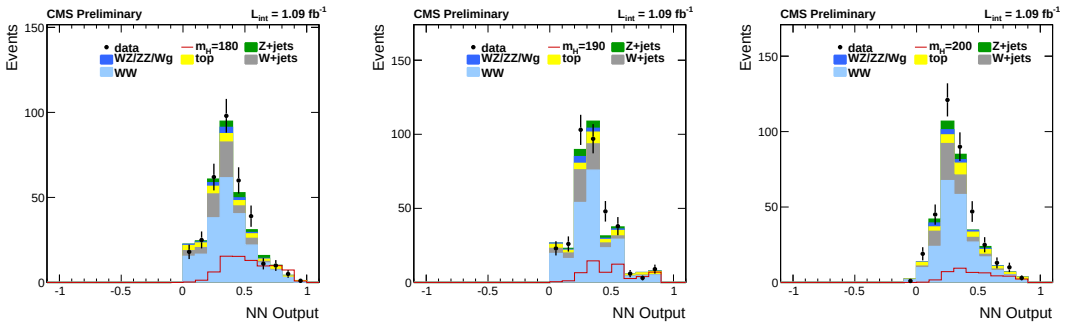
(a) (b) (c)

Figure D.1: Distribution of the best Neural Network output for $m_H = 120, 130, 140$ GeV in the 0 jet bin.



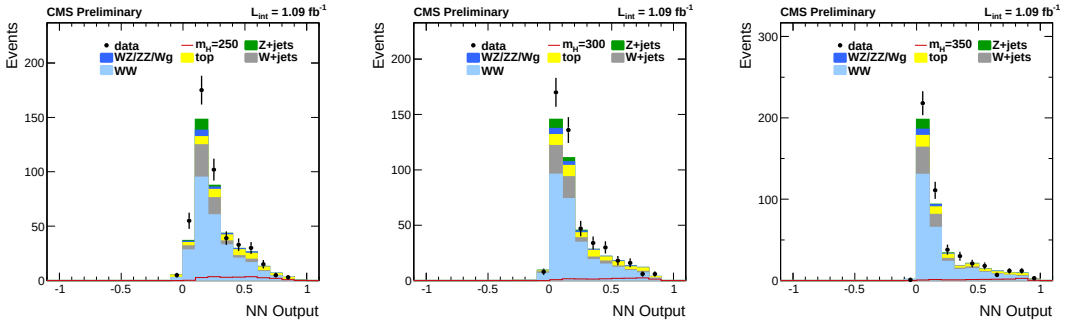
(a) (b) (c)

Figure D.2: Distribution of the best Neural Network output for $m_H = 150, 160, 170$ GeV in the 0 jet bin.



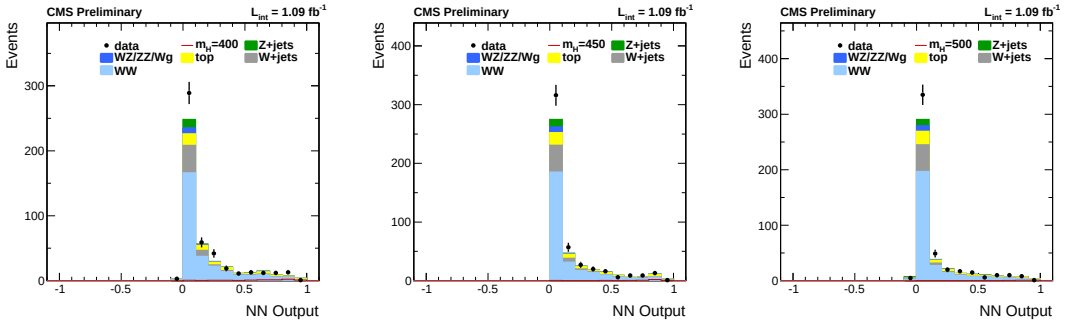
(a) (b) (c)

Figure D.3: Distribution of the best Neural Network output for $m_H = 180, 190, 200$ GeV in the 0 jet bin.



(a) (b) (c)

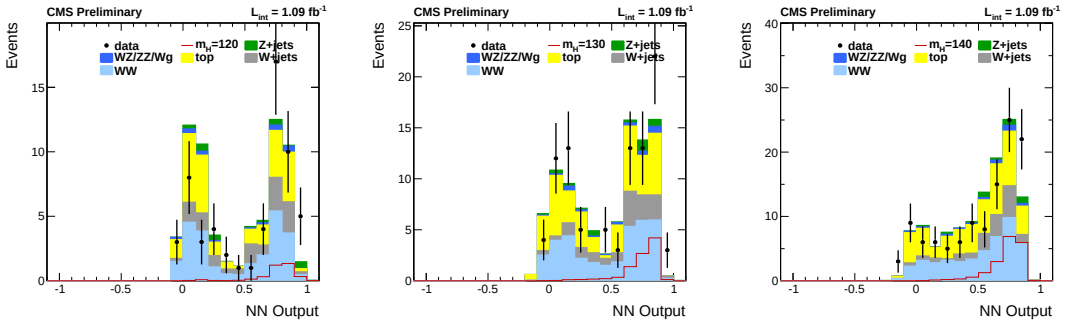
Figure D.4: Distribution of the best Neural Network output for $m_H = 250, 300, 350$ GeV in the 0 jet bin.



(a) (b) (c)

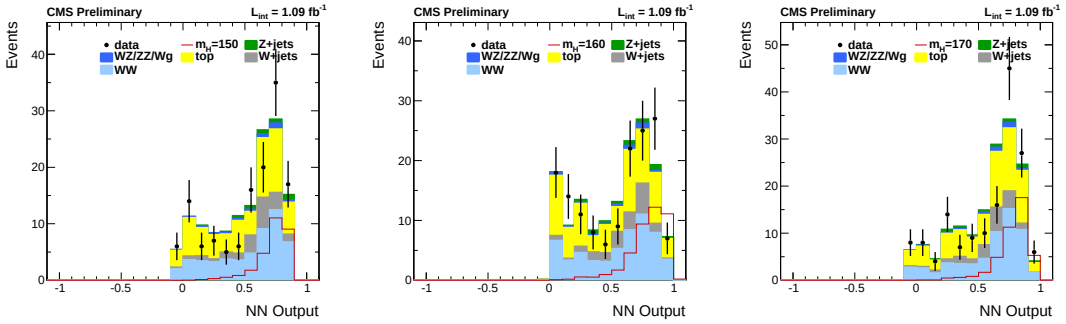
Figure D.5: Distribution of the best Neural Network output for $m_H = 400, 450, 500$ GeV in the 0 jet bin.

D.2 1 jet bin



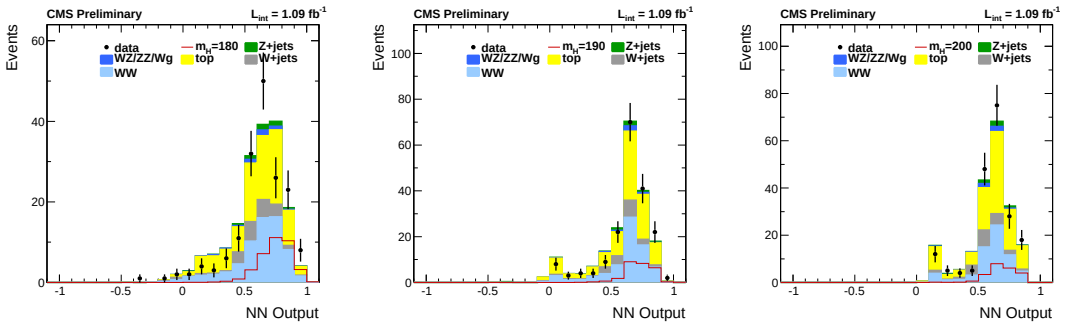
(a) (b) (c)

Figure D.6: Distribution of the best Neural Network output for $m_H = 120, 130, 140$ GeV in the 1 jet bin.



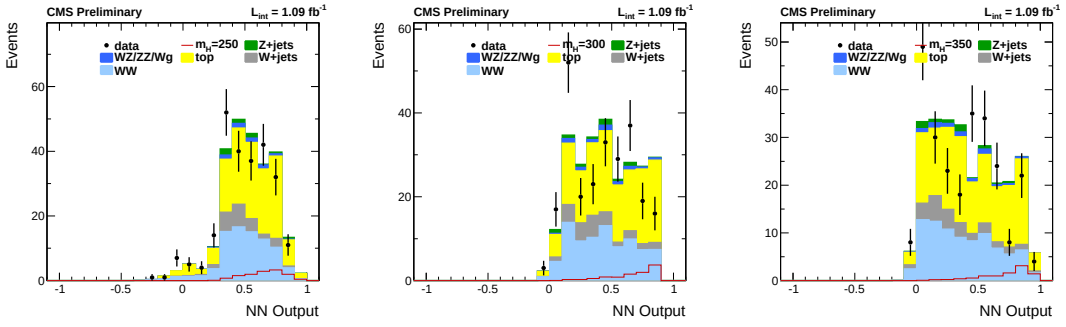
(a) (b) (c)

Figure D.7: Distribution of the best Neural Network output for $m_H = 150, 160, 170$ GeV in the 1 jet bin.



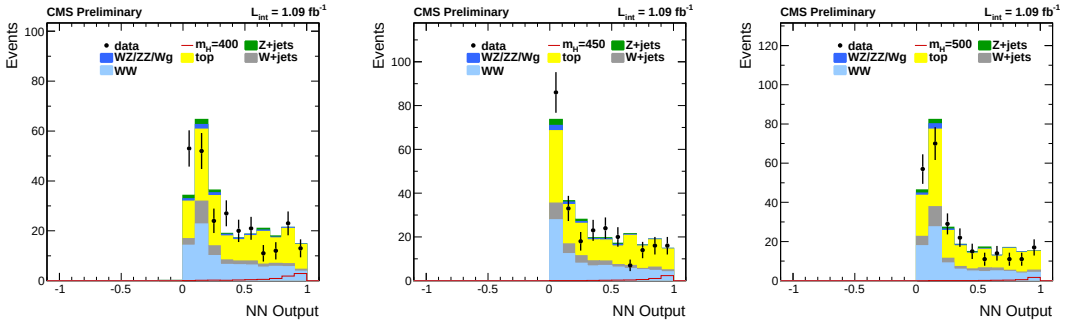
(a) (b) (c)

Figure D.8: Distribution of the best Neural Network output for $m_H = 180, 190, 200$ GeV in the 1 jet bin.



(a) (b) (c)

Figure D.9: Distribution of the best Neural Network output for $m_H = 250, 300, 350$ GeV in the 1 jet bin.



(a) (b) (c)

Figure D.10: Distribution of the best Neural Network output for $m_H = 400, 450, 500$ GeV in the 1 jet bin.

Appendix E

Method used to estimate Parton-distribution functions uncertainties

E.1 Introduction

In order to generate events according to hadron collision processes, MC generators need to randomly produce events with parton momentum fraction and energy taken from some distributions. These distributions are called parton distribution functions (PDFs) and they are flavor specific. Unfortunately PDFs are empyrical functions, they are obtained as a result of a QCD fit using data collected dependent from many different experiments and therefore are subject to uncertainties coming from these QCD analyses. Consequently, a consistent approach for parametrization has to be defined in order to be properly able to propagate these systematics to MC cross section (or other physical variables) calculations. Indeed it is crucial to point out how important is the knowledge of PDFs in order to correctly predict the cross section of a given process, which can be written as:

$$\sigma_{had} = \sum_{m,n} \int f_m(x_1, Q) f_n(x_2, Q) dx_1 dx_2 \sigma_{part} \tag{E.1}$$

The variables $x_{1,2}$ are respectively the momentum fractions of the incoming partons entering the hard interaction and Q is the event scale of the process. σ_{part} is the partonic cross section and $f_{m,n}(x_{1,2}, Q)$ are the PDFs of the incoming partons at a given momentum fraction and energy scale. The indexes m, n denote the flavor of the partons involved. In the follwing sections we will first discuss the framework and the method which is necessary to compute the propagation of PDFs uncertainties and finally we will present some results concerning the H→WW and its main backgrounds.

E.2 PDF errors

PDFs are parametrized in a polynomial form. As an example, the PDF parametrization for the gluon by the CTEQ group is

$$f(x, Q_0) = A_0 x^{A_1} (1 - x)^{A_2} (1 + A_3 x^{A_4}) \tag{E.2}$$

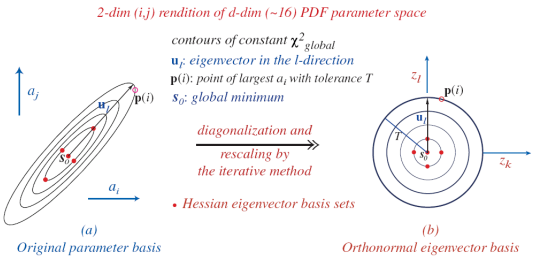


Figure E.1: Iso-contours of χ^2_{global} before and after diagonalization of the Hessian matrix

where the A_i are the free parameters of the fit. Depending on the analysis, the group defines a minimal set of parameters necessary to describe the QCD data. Then a global fit is performed leading to a minimal χ^2_{global} estimation which defines the optimal description of the different experiments that were used. The optimal set of parameters defines the "best" PDFs but does not give any information about the uncertainties. The CTEQ group has developed a method in order to provide the error associated with the estimation of each of the free parameters. Additional information can be extracted from the χ^2 derivatives, in particular the Hessian matrix can give an estimation of the uncertainties coming from the error on the parameters estimation. In order to have independent errors the Hessian matrix is diagonalized and a new set of parameters is computed in this new orthogonal basis. A visual explanation is given in Figure E.1. The up-down variations of the new parameters are then computed by indepedently shifting them in this new basis and calculating the relative $\Delta\chi^2$ variation. The main PDF groups work on this basis to provide the central value and 2d parameters variation (there are d parameters and therefore 2d +/- variations)
.

E.3 PDFs uncertainty propagation

E.3.1 The reweighting method

As we saw in the previous section, PDF sets are provided with a central value $f_0(x,Q)$, which is used for calculations, such as cross sections or efficiencies, and a set of 2d +/- variations with respect to the central value, $f_{k^\pm}(x,Q)$. The propagation of PDF uncertainties to physical variables is therefore straightforward. If X is the physical quantity of interest (i.e cross section, efficiency, PDF itself), we need to compute X not only for the central value but also for the 2d "deviations" due to PDF uncertainty and endup with 2d calculations of X which we label as $X_{k^\pm}$ (see FigureE.2).

There are two methods to calculate these deviations. The most intuitive is the "brute force" production of 2d+1 MC samples, corresponding to the central value and the 2d deviations. This method is indeed correct but time consuming. The other approach one can have is to simply produce one MC sample, corresponding to the central value, and assign to each event 2d-couple of weights defined by:

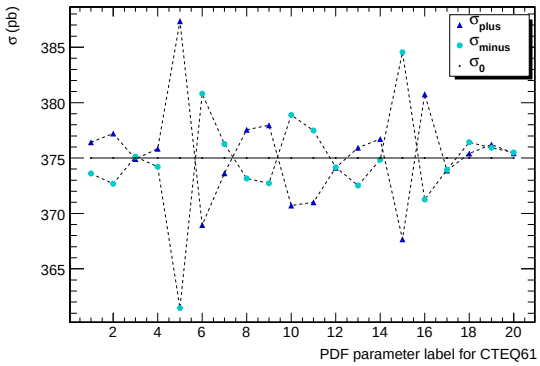


Figure E.2: Calculation of the $t\bar{t}$ cross-section and its d=20 +/- parameters deviations with the use of CTEQ6.1 tables

$$w^i_{k^\pm} = \frac{f_{k^\pm}(x^i_1, f^i_1, Q^i) f_{k^\pm}(x^i_2, f^i_2, Q^i)}{f_0(x^i_1, f^i_1, Q^i) f_0(x^i_2, f^i_2, Q^i)} \tag{E.3}$$

for $1 \leq i \leq N_{total}$ and $1 \leq k \leq d$ and where $x_{1,2}$ and Q are the momentum fractions and scale of the event i. This method works provided that two conditions are fulfilled. The first is that events simulated with a deviated PDF exhibit the same topology as the ones simulated with the central value. This seems as fairly good assumption at leading order since the $f_{k^\pm}$ are only small deviations of $f_0(x, Q)$ and cover the same x,Q range with approximately the same integral. The second condition is that enough events need to be generated in order to have a good coverage of the momentum fraction spectrum. Once the weights are stored for each event, the calculation of $X_{k^\pm}$ is straightforward. For instance if X is a cross section, $X_{k^\pm}$ would then be:

$$\sigma^\pm_k = \sum_{i=1}^{N_{events}} w^i_{k^\pm} \tag{E.4}$$

or an efficiency:

$$\varepsilon^\pm_k = \frac{\sum_{i=1}^{N_{selected}} w^i_{k^\pm}}{\sum_{i=1}^{N_{total}} w^i_{k^\pm}} \tag{E.5}$$

and if the physical observable would be the PDF itself then clearly one has $X_{k^\pm} = f_{k^\pm}$.

E.3.2 The Master Formula

The merit of the Hessian approach is producing uncorrelated parameters. In other words, the total uncertainty can be computed by adding in quadrature the "d" independent contributions. Nevertheless, a convention needs to be introduced for specific cases where the +/- contribution falls on the same side with respect to the central value, as it happens for parameter 19 in Figure E.2. Therefore a general formula has to be introduced. It will

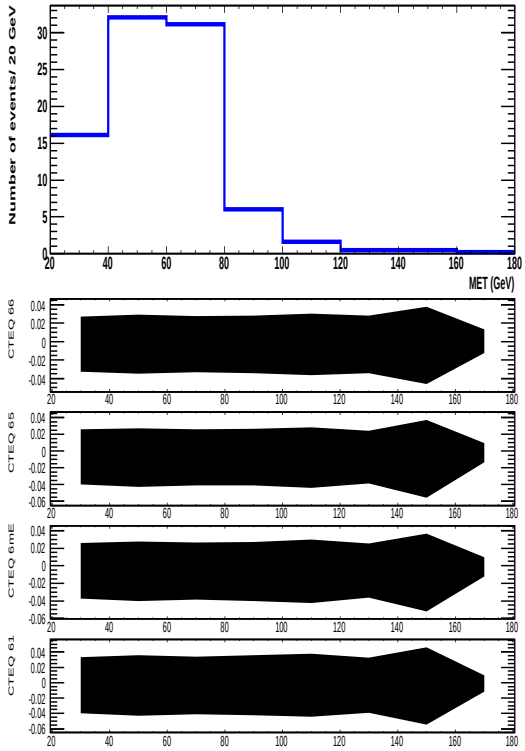


Figure E.3: MET distribution for signal with m_H=160 GeV and its relative PDF uncertainty per bin using different CTEQ tables

be defined in such a way that, for instance, the 19th parameter in Figure E.2 will get no negative contribution to the final uncertainty (but will indeed get a positive one). So if, as before, $X_{k^{\pm}}$ are the varied quantities of the physical observable, the final +/- uncertainty, $\Delta X^{\pm}$, is given by:

$$\Delta X^{\pm} = \sqrt{\sum_{k=1}^{D} \left[max(X_k^{\pm} - X_0, X_k^{\mp} - X_0, 0) \right]^2} \tag{E.6}$$

E.4 Results

E.4.1 Uncertainties

As an example we show the distributions of missing transverse energy (MET) and the angle between the two leptons ($\Delta\phi$) for signal, WW and $t\bar{t}$ samples. Even though limited by small statistics, it is interesting to see that this reweighting method and master formula usage can be used also with kinematic distributions. The Figures E.3-E.4 show the relative uncertainty per bin according to 4 different PDF sets (CTEQ61, CTEQ6mE, CTEQ65, CTEQ66) to be almost stable around 5% for WW and 3% for signal, even though in the last bins there are only few events left.

In Tables E.4.1 and E.4.1 we summarize the total cross section uncertainties. The +/- uncertainties are shown for all signal and background sample and for all CTEQ tables versions. The uncertainties are found to be at the level of 3% for the signal and 5% for backgrounds.

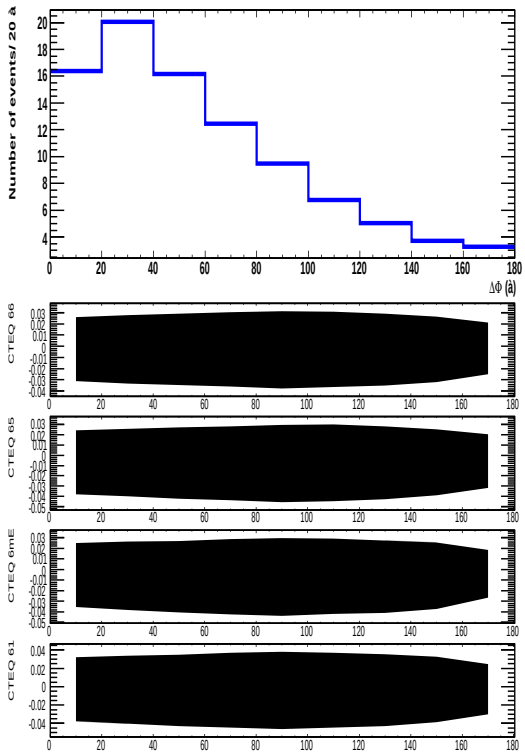


Figure E.4: $\Delta\phi$ distribution for signal with m_H=160 GeV and its relative PDF uncertainty per bin using different CTEQ tables

PDF Uncertainties for H→WW channel					
Mass (GeV)	Cross Section at 14 TeV (pb)	CTEQ 6.1	CTEQ 6mE	CTEQ 6.5	CTEQ 6.6
130	0.91	+ 2.9 % - 3.7 %	+ 2.2 % - 3.3 %	+ 2.3 % - 3.4 %	+ 2.3 % - 2.8 %
150	1.65	+ 2.6 % - 3.2 %	+ 2.0 % - 2.9 %	+ 2.0 % - 3.1 %	+ 2.1 % - 2.5 %
160	1.90	+ 2.5 % - 3.0 %	+ 1.9 % - 2.7 %	+ 1.9 % - 3.0 %	+ 2.0 % - 2.4 %
170	1.81	+ 2.4 % - 2.9 %	+ 1.9 % - 2.7 %	+ 1.9 % - 2.9 %	+ 1.9 % - 2.4 %
200	1.01	+ 2.2 % - 2.5 %	+ 1.7 % - 2.3 %	+ 1.8 % - 2.6 %	+ 1.8 % - 2.1 %

Table E.1: PDF uncertainties for signal samples

PDF Uncertainties for backgrounds					
Sample	Cross Section at 14 TeV (pb)	CTEQ 6.1	CTEQ 6mE	CTEQ 6.5	CTEQ 6.6
$t\bar{t}$	375	+ 5.1 % - 5.0 %	+ 4.9 % - 4.6 %	+ 5.1 % - 4.4 %	+ 4.9 % - 4.6 %
WW	75	+ 3.7 % - 4.7 %	+ 3.6 % - 4.1 %	+ 3.0 % - 3.1 %	+ 3.2 % - 3.2 %
W+jets	40000	+ 4.1 % - 5.0 %	+ 3.5 % - 4.1 %	+ 3.4 % - 3.5 %	+ 3.2 % - 3.3 %

Table E.2: PDF uncertainties for background samples

Bibliography

[1] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl.Phys.*, 22:579–588, 1961.

[2] Abdus Salam and John Clive Ward. Gauge theory of elementary interactions. *Phys.Rev.*, 136:B763–B768, 1964.

[3] Steven Weinberg. A Model of Leptons. *Phys.Rev.Lett.*, 19:1264–1266, 1967.

[4] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–322, 1964.

[5] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys.Rev.Lett.*, 13:508–509, 1964.

[6] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble. Global Conservation Laws and Massless Particles. *Phys.Rev.Lett.*, 13:585–587, 1964.

[7] C. Itzykson and J.B. Zuber. Quantum Field Theory. 1980.

[8] Michael E. Peskin and Daniel V. Schroeder. An Introduction to quantum field theory. 1995.

[9] Chen-Ning Yang and Robert L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys.Rev.*, 96:191–195, 1954.

[10] Yoichiro Nambu. Quasiparticles and Gauge Invariance in the Theory of Superconductivity. *Phys.Rev.*, 117:648–663, 1960.

[11] J. Goldstone. Field Theories with Superconductor Solutions. *Nuovo Cim.*, 19:154–164, 1961.

[12] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys.Rev.*, 127:965–970, 1962.

[13] K. Nakamura et al. Review of particle physics. *J.Phys.G*, G37:075021, 2010.

[14] D.J. Gross and Frank Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys.Rev.Lett.*, 30:1343–1346, 1973.

[15] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.

[16] John M. Cornwall, David N. Levin, and George Tiktopoulos. Uniqueness of spontaneously broken gauge theories. *Phys.Rev.Lett.*, 30:1268–1270, 1973.

[17] John M. Cornwall, David N. Levin, and George Tiktopoulos. Derivation of Gauge Invariance from High-Energy Unitarity Bounds on the s Matrix. *Phys.Rev.*, D10:1145, 1974.

[18] Benjamin W. Lee, C. Quigg, and H.B. Thacker. Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass. *Phys.Rev.*, D16:1519, 1977.

[19] J. Bagger, Vernon D. Barger, King-man Cheung, John F. Gunion, Tao Han, et al. The Strongly interacting W W system: Gold plated modes. *Phys.Rev.*, D49:1246–1264, 1994.

[20] Thomas Hambye and Kurt Riesselmann. Matching conditions and Higgs mass upper bounds revisited. *Phys.Rev.*, D55:7255–7262, 1997.

[21] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys.Lett.*, B565:61–75, 2003.

[22] The CDF, D0 Collaborations, the Tevatron New Phenomena, and Higgs Working Group. Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb-1 of Data. *arXiv:1107.5518*.

[23] Henning Flacher, Martin Goebel, Johannes Haller, Andreas Hocker, Klaus Monig, et al. Revisiting the Global Electroweak Fit of the Standard Model and Beyond with Gfitter. *Eur.Phys.J.*, C60:543–583, 2009.

[24] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Ludwig, et al. Updated Status of the Global Electroweak Fit and Constraints on New Physics. *arXiv:1107.0975*, 2011.

[25] Leonard Susskind. Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory. *Phys.Rev.*, D20:2619–2625, 1979.

[26] Steven Weinberg. Implications of Dynamical Symmetry Breaking. *Phys.Rev.*, D13:974–996, 1976.

[27] Howard Georgi and David B. Kaplan. Composite Higgs and Custodial SU(2). *Phys.Lett.*, B145:216, 1984.

[28] Christopher T. Hill, Stefan Pokorski, and Jing Wang. Gauge invariant effective Lagrangian for Kaluza-Klein modes. *Phys.Rev.*, D64:105005, 2001.

[29] Nima Arkani-Hamed, Andrew G. Cohen, and Howard Georgi. (De)constructing dimensions. *Phys.Rev.Lett.*, 86:4757–4761, 2001.

[30] Christopher T. Hill and Elizabeth H. Simmons. Strong dynamics and electroweak symmetry breaking. *Phys.Rept.*, 381:235–402, 2003.

[31] Nir Polonsky. Supersymmetry: Structure and phenomena. Extensions of the standard model. *Lect.Notes Phys.*, M68:1–169, 2001.

[32] Richard P. Feynman. Very high-energy collisions of hadrons. *Phys.Rev.Lett.*, 23:1415–1417, 1969.

[33] V.N. Gribov and L.N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:438–450, 1972.

[34] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl.Phys.*, B126:298, 1977.

[35] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.

[36] M. R. Whalley, D. Bourilkov, and R. C. Group. The Les Houches Accord PDFs (LHAPDF) and Lhaglue. *arXiv:0508110*, 2005.

[37] John C. Collins and Andreas Metz. Universality of soft and collinear factors in hard-scattering factorization. *Phys.Rev.Lett.*, 93:252001, 2004.

[38] G. Watt. Parton distribution function dependence of benchmark Standard Model total cross sections at the 7 TeV LHC. *arXiv:1106.5788*, 2011.

[39] Charalampos Anastasiou, Stephan Buehler, Elisabetta Furlan, Franz Herzog, and Achilleas Lazopoulos. Higgs production cross-section in a Standard Model with four generations at the LHC. *Phys.Lett.*, B702:224–227, 2011.

[40] J. Baglio, A. Djouadi, S. Ferrag, and R.M. Godbole. The Tevatron Higgs exclusion limits and theoretical uncertainties: A Critical appraisal. *Phys.Lett.*, B699:368–371.

[41] Stefano Catani, Daniel de Florian, and Massimiliano Grazzini. Universality of non-leading logarithmic contributions in transverse momentum distributions. *Nucl.Phys.*, B596:299–312, 2001.

[42] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini. Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC. *arXiv:1109.2109*.

[43] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.

[44] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001.

[45] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys.Rept.*, 97:31–145, 1983.

[46] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002.

[47] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[48] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[49] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[50] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 1106:128, 2011.

[51] Andy Buckley, Jonathan Butterworth, Stefan Gieseke, David Grellscheid, Stefan Hoche, et al. General-purpose event generators for LHC physics. *Phys.Rept.*, 504:145–233, 2011.

[52] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV. *Phys.Lett.*, B122:103–116, 1983.

[53] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider. *Phys.Lett.*, B122:476–485, 1983.

[54] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider. *Phys.Lett.*, B126:398–410, 1983.

[55] P. Bagnaia et al. Evidence for Z0 –> e+ e- at the CERN anti-p p Collider. *Phys.Lett.*, B129:130–140, 1983.

[56] M. Acciarri et al. Determination of the number of light neutrino species from single photon production at LEP. *Phys.Lett.*, B431:199–208, 1998.

[57] A. Alavi-Harati et al. Measurements of direct CP violation, CPT symmetry, and other parameters in the neutral kaon system. *Phys.Rev.*, D67:012005, 2003.

[58] J.R. Batley et al. A Precision measurement of direct CP violation in the decay of neutral kaons into two pions. *Phys.Lett.*, B544:97–112, 2002.

[59] (Ed.) Bruning, Oliver S., (Ed.) Collier, P., (Ed.) Lebrun, P., (Ed.) Myers, S., (Ed.) Ostojic, R., et al. LHC Design Report. 1. The LHC Main Ring. *CERN-2004-003-V-1*, 2004.

[60] (Ed.) Buning, O., (Ed.) Collier, P., (Ed.) Lebrun, P., (Ed.) Myers, S., (Ed.) Ostojic, R., et al. LHC Design Report. 2. The LHC infrastructure and general services. *CERN-2004-003-V-2*, 2004.

[61] (Ed.) Benedikt, M., (Ed.) Collier, P., (Ed.) Mertens, V., (Ed.) Poole, J., and (Ed.) Schindl, K. LHC Design Report. 3. The LHC injector chain. *CERN-2004-003-V-3*, 2004.

[62] ATLAS: Detector and physics performance technical design report. Volume 1. *ATLAS-TDR-014 ; CERN-LHCC-99-014*, 1999.

[63] ATLAS: Detector and physics performance technical design report. Volume 2. *ATLAS-TDR-015 ; CERN-LHCC-99-015*, 1999.

[64] LHCb technical design report: Reoptimized detector design and performance. *CERN-LHCC-2003-030 ; LHCb-TDR-9*, 2003.

[65] G.L. Bayatian et al. CMS technical design report, volume I: Detector Performance and Software. *CERN-LHCC-2006-001 ; CMS-TDR-008-1*, 2006.

[66] G.L. Bayatian et al. CMS technical design report, volume II: Physics performance. *J.Phys.G*, G34:995–1579, 2007.

[67] R. Adolphi et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[68] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[69] Serguei Chatrchyan et al. Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays. *JINST*, 5:T03021, 2010.

[70] CMS, tracker technical design report. *CERN-LHCC-98-006 ; CMS-TDR-005*, 1998. 15 April 1998.

[71] CMS Collaboration. Measurement of Momentum Scale and Resolution using Low-mass Resonances and Cosmic Ray Muons. *CMS Physics Analysis Summary*, 2010.

[72] CMS: The electromagnetic calorimeter. Technical design report. *CERN-LHCC-97-033 ; CMS-TDR-004*, 1997.

[73] CMS: The hadron calorimeter technical design report. *CERN-LHCC-97-031 ; CMS-TDR-002*, 1997.

[74] CMS, the Compact Muon Solenoid. Muon technical design report. *CERN-LHCC-97-032 ; CMS-TDR-003*, 1997.

[75] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[76] John Allison, K. Amako, J. Apostolakis, H. Araujo, P.A. Dubois, et al. Geant4 developments and applications. *IEEE Trans.Nucl.Sci.*, 53:270, 2006.

[77] CMS: The computing project. Technical design report. *CERN-LHCC-2005-023 ; CMS-TDR-007*, 2005.

[78] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl.Instrum.Meth.*, A389:81–86, 1997.

[79] Michael Dittmar and Herbert K. Dreiner. h0 $\rightarrow$ W+ W- $\rightarrow$ l+ l- neutrino neutrino as the dominant SM Higgs search mode at the LHC for M(h0) = 155-GeV - 180-GeV. 1996.

[80] The CMS collaboration. Track reconstruction in the cms tracker. *CMS-TRK-09-001*, 2009.

[81] The CMS collaboration. Tracking and vertexing results from first collisions. *CMS-TRK-10-001*, 2010.

[82] The CMS Collaboration. Vertex Recosntruction at the CMS experiment. *Journal of Physics*, 110:81–86, 2007.

[83] J.Huth et al. Snowmass Accord. *Fermilab-Conf-90/249-E*, 1990.

[84] Gregory Soyez. The SISCone and anti-k(t) jet algorithms. *arXiv:0807.0021*, 2008.

[85] The CMS collaboration. Calorimeter jet quality criteria for the first cms collision data. *CMS-JME-09-008*, 2009.

[86] The CMS collaboration. Particle-flow event reconstruction in cms and performance for jets, taus, and etmiss. *CMS-PFT-09-001*, 2009.

[87] The CMS collaboration. Jet performance in pp collisions at $\sqrt{s}$ =7 tev. *CMS-JME-10-003*, 2010.

[88] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 0804:005, 2008.

[89] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys.Lett.*, B659:119–126, 2008.

[90] The CMS collaboration. Search for the higgs boson decaying to ww in the fully leptonic final state. *CMS-HIG-11-003*, 2011.

[91] The CMS collaboration. Algorithms for b jet identification in cms. *CMS-BTV-09-001*, 2009.

[92] J. M. Campbell, R. K. Ellis. Mcfm for the tevatron and the lhc. *arXiv:1007.3492*, 2010.

[93] R. Gavin, Y. Li, F. Petriello and S. Quackenbush. Fewz 2.0: A code for hadronic z production at the next-to-next-to-leading order. *arXiv:1011.3540v1*, 2010.

[94] T. Binoth, M. Ciccolini, N. Kauer, M. Kramer. Gluon-induced w-boson pair production at the lhc. *JHEP0612:046*, 2006.

[95] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl.Instrum.Meth.*, A362:487–498, 1995.

[96] The CMS collaboration. Measurements of inclusive w and z cross sections in pp collisions at $sqrts = 7$ tev. *CMS-EWK-10-002*, 2010.

[97] The CMS collaboration. Performance of cms muon reconstruction in pp collisions at $sqrts = 7$ tev. *CMS-MUO-10-004*, 2010.

[98] CMS generator group.

[99] John M. Campbell, R.Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 1107:018, 2011.

[100] Jan Therhaag. TMVA Toolkit for multivariate data analysis in ROOT. *PoS*, ICHEP2010:510, 2010.

[101] Alexander L. Read. Modified frequentist analysis of search results (The CL(s) method). *CERN-OPEN-2000-205*, 2000.