



Master Thesis

Daniele Zanzi

Search for the Standard Model Higgs Boson in the ATLAS Experiment

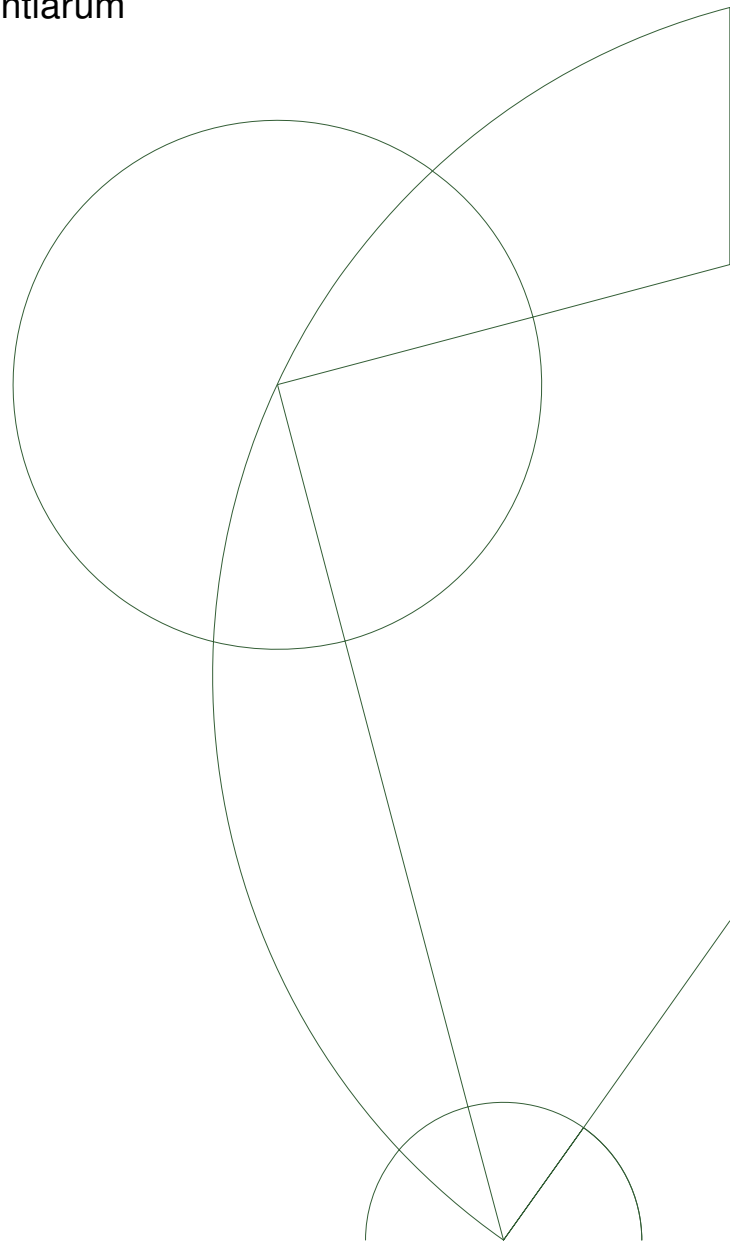
Submitted for the degree of candidatus scientiarum

CERN-THESIS-2011-002
04/02/2011



Supervised by Stefania Xella

February 4, 2011



Contents

Introduction	1
1 The Standard Model	3
1.1 Symmetries	5
1.2 The Electro-Weak Interaction	6
1.3 Local Gauge Invariance	8
1.4 The Higgs Mechanism	10
1.5 The Electro-Weak Symmetry Breaking	14
1.6 The Fermionic Masses	16
2 The SM Higgs Boson Quest	19
2.1 Theoretical Limits	19
2.1.1 Unitarity	19
2.1.2 Triviality	20
2.1.3 Vacuum Stability	20
2.2 Electro-Weak Radiative Corrections	21
2.3 Direct Searches	25
2.3.1 Production Processes	25
2.3.2 Branching Ratios and Total Width	27
2.4 Direct Searches at LEP-II	29
2.5 Direct Searches at Tevatron	32
3 LHC and the ATLAS Experiment	35
3.1 The Large Hadron Collider	35
3.2 The ATLAS Experiment	40
3.3 MonteCarlo event generation and ATLAS simulation	45
3.4 Particles and event properties reconstruction in ATLAS	48
4 SM Higgs Boson Searches in ATLAS	51
4.1 The VBF analysis	53
4.2 An auxiliary analysis	53
5 The $\text{pp}(\text{gg}+\text{VBF}) \rightarrow \text{H}+\text{jets} \rightarrow \tau^+\tau^- + \text{jets}$	57
5.1 Signal	57
5.2 Backgrounds	59
5.3 Trigger	60
5.4 Objects and Event selection	63
5.5 Cut Description	66
5.5.1 Initial Cuts	66
5.5.2 B Jet Veto	66

5.5.3	Eta Tagging Jet	68
5.5.4	Missing Energy, E_T^{miss}	68
5.5.5	Collinear Mass Approximation	69
5.5.6	Visible Mass	75
5.5.7	Invariant Mass $M(h, j)$	75
5.5.8	Factorization	75
5.5.9	Weights and Errors	78
5.6	Results	80
5.7	Combined Results	84
5.8	gg and VBF Comparison	87
5.9	Pile Up Effects	87
Conclusions		94
Bibliografy		94
A Detailed Cut flows		99
B VBF Cut flow		105
C Tau ID Factorization		107

Introduction

The Standard Model (SM) of Particle Physics is a theory established almost 40 years ago and it is corroborated by a long series of experimental results. It cannot be considered completely tested until all its theoretical foundations have been proven. Indeed the pivotal part of this theory, the ElectroWeak Symmetry Breaking, could be explained by the presence of a doublet of scalar fields, but so far we don't have evidence of their existence in Nature. These fields should be unveiled by the detection of one scalar neutral particle, the *Higgs boson*.

Intensive searches for the Higgs boson are ongoing since more than 10 years. The Higgs boson is searched for through its direct production and its virtual effects on electroweak observables, but so far physicists have only been able to set limits on the mass of this particle since no signal has been observed. This quest is very important and the future of particle physics will be driven by either the discovery or the exclusion of the SM Higgs boson. In case of discovery, the Higgs boson mass can give hints on which scale new physics occurs; in case of exclusion, a deep theoretical work will be needed to find an alternative explanation of the experimental successes of the SM.

In 2009 the *Large Hadron Collider* (LHC) started to provide proton proton collision data at the highest energy ever reached. This machine has been designed to provide the ultimate answer about the Higgs boson existence because it will be able to explore the whole mass range, from 114.4 GeV up to 1 TeV, where the Higgs boson is expected to be.

Thanks to the good performance of the accelerator, already in 2010 the LHC experiments, ATLAS, CMS, ALICE and LHCb, started recording events produced at the center of mass energy of $\sqrt{s} = 7$ TeV. At the end of 2011 there should be enough integrated luminosity to exclude most of the low mass range where the Higgs boson is likely to be found, although the very low mass region below 120 GeV will be very tough to explore. Of course, the LHC search will continue after the long shutdown, expected in 2012 or 2013, which will bring an energy and luminosity increase. The only other experiments looking for the Higgs boson are CDF and D0 at the Tevatron accelerator in USA and they use proton antiproton collisions at a center of mass energy lower than LHC. After almost two decades of operation, with the last data taken at the end of 2011, they might be able to reach the sensitivity to exclude the SM Higgs boson in some mass range. But the LHC data will be definitively the data with the final word on the SM Higgs boson.

In this thesis I present a study on the Higgs boson search in $\tau^+\tau^-$ final states that can improve the ATLAS sensitivity at $m_H = 120$ GeV, exactly in the mass range where more efforts are needed. The aim of the study is, indeed, to prove that additional production modes and decay channels, not considered so far, can be used.

I design a new event selection sensitive to both the gluon fusion and the Vector Boson Fusion (VBF) production modes and I include all the possible tau lepton decays, considering therefore fully leptonic, semileptonic and fully hadronic tau decays. The study is an event counting experiment performed on simulated Monte Carlo events produced at the center of

mass energy of 10 TeV. The discovery sensitivity obtained combining all the channels for an integrated luminosity of 30 fb^{-1} is more than 3σ , which corresponds to more than 4σ at 14 TeV and more than 2σ at 7 TeV. The sensitivity is comparable to the one achieved by the existing $H \rightarrow \tau^+\tau^-$ analysis which includes only the VBF as production mode and only the leptonic and semi-leptonic channels. However, since the event selection proposed in this thesis is almost orthogonal to the one used in the baseline $H \rightarrow \tau^+\tau^-$ analysis, the two sensitivities can be combined. Consequently this study proves that it is possible to increase the sensitivity of ATLAS to the search for the Higgs boson. Moreover a significant part of the combined sensitivity of this study comes from the hadronic channel which has not been included so far. So a remarkable improvement for the SM Higgs boson search can be achieved.

In order to provide the necessary preliminary notions, in the first chapter an introduction to the Standard Model and the SM ElectroWeak Symmetry Breaking is given. Then the second chapter is devoted to summarize our present knowledge about the SM Higgs boson and to describe the latest results from the direct and the indirect searches. The third chapter presents the main features of the LHC accelerator and the ATLAS detector, with a specific focus on the present and future technical challenges. In addition a description of how Monte Carlo simulations of proton proton collisions are performed is provided.

The analysis, subject of this thesis, is described in the fourth and fifth chapter. A summary of the included signals and backgrounds and of the event selection is given. Then the sensitivity of the search strategy proposed is summarized, and the results are compared to the current baseline analysis for SM Higgs boson search at low mass, the VBF analysis. In the conclusion a list of improvements on the search strategy and a brief outline of the steps needed to apply the search to 2011 data is presented.

Acknowledgments

This work was only possible thanks to the support of my supervisor Stefania Xella and the Niels Bohr Institute. The year I've spent at NBI has been extremely helpful for my growth as a physicist. I really benefited from the informal atmosphere and the experience of each member of the ATLAS Group. I'll always be thankful for the huge amount of time that Stefania spent on me and I hope that my work will be useful for the activities of the group.

Chapter 1

The Standard Model

The Standard Model (SM) of Particle Physics is the best theory at the moment to explain the phenomena of the microcosm. Nature can be described at the scale of 1 fm (10^{-15} meters) in terms of fundamental particles and interactions between them. This general framework has been built up in the second half of the XX century by great theoretical and experimental achievements. Fig. 1.1 illustrates the fundamental particles observed in Nature and described within the SM theory.

In the SM particles can be *bosons* or *fermions*. The bosons, by definition, have integer spins and all the fundamental particles of this type observed so far are *vectors*, i.e. they have spin one. These bosons are the carriers of the fundamental interactions which are the *electromagnetic* force, the *strong* force, the *weak* force and the *gravitational* force. Photons and gluons mediate the electromagnetic and the strong force, respectively, and they are massless. The W and Z bosons, instead, are involved in the weak interactions and, as we will see, their masses were correctly predicted by the SM [2]. Gravitation should also be mediated by a boson, the graviton, even if with spin 2 and not 1 as the others, but there is no evidence of its existence. The gravitational force is not included in the SM because there is no a coherent approach to describe this interaction in the same ways as the others. Moreover in the microscopic world gravity is so weak that it can be neglected. The last boson which appears in Fig. 1.1 is the Higgs boson. This particle has never been observed and its search is the topic of this thesis. Here we only mention that, on the contrary of the other bosons, it is not a mediator of a force and it is expected by SM predictions to be scalar, i.e. spin 0, and neutral.

The other particles included in the SM are fermions. These are the bricks of the ordinary matter. Fermions, by definition, have semi-integer spins and the fundamental fermions in Nature and therefore in SM have all spin $1/2$. They can be divided in quarks and leptons. The first ones interact via all the three SM forces, while the second ones are not strongly interacting. All fermions can be arranged in three symmetric sets of particles called *families*. Each of them includes two quarks and two leptons.

Looking at Fig. 1.1, the huge spread in masses between particles makes Nature look rather asymmetric. Nonetheless, at the basis of the SM stands the groundbreaking concept of *symmetry*. It can be surprising but the dynamics and the masses of the fundamental particles can spring from the fact that physical laws should obey certain symmetries or, in other word, should be invariant under specific transformations. This approach has been so successful that today most of the particle physicists believe that all the laws of the microcosm might be a manifestation of the same unique symmetry.

In this chapter we will introduce first how fundamental interactions stem from principles of invariance and then we will take the reader to the core of the SM: the Electro-Weak Unification through the Spontaneous Symmetry Breaking. The books used to write this

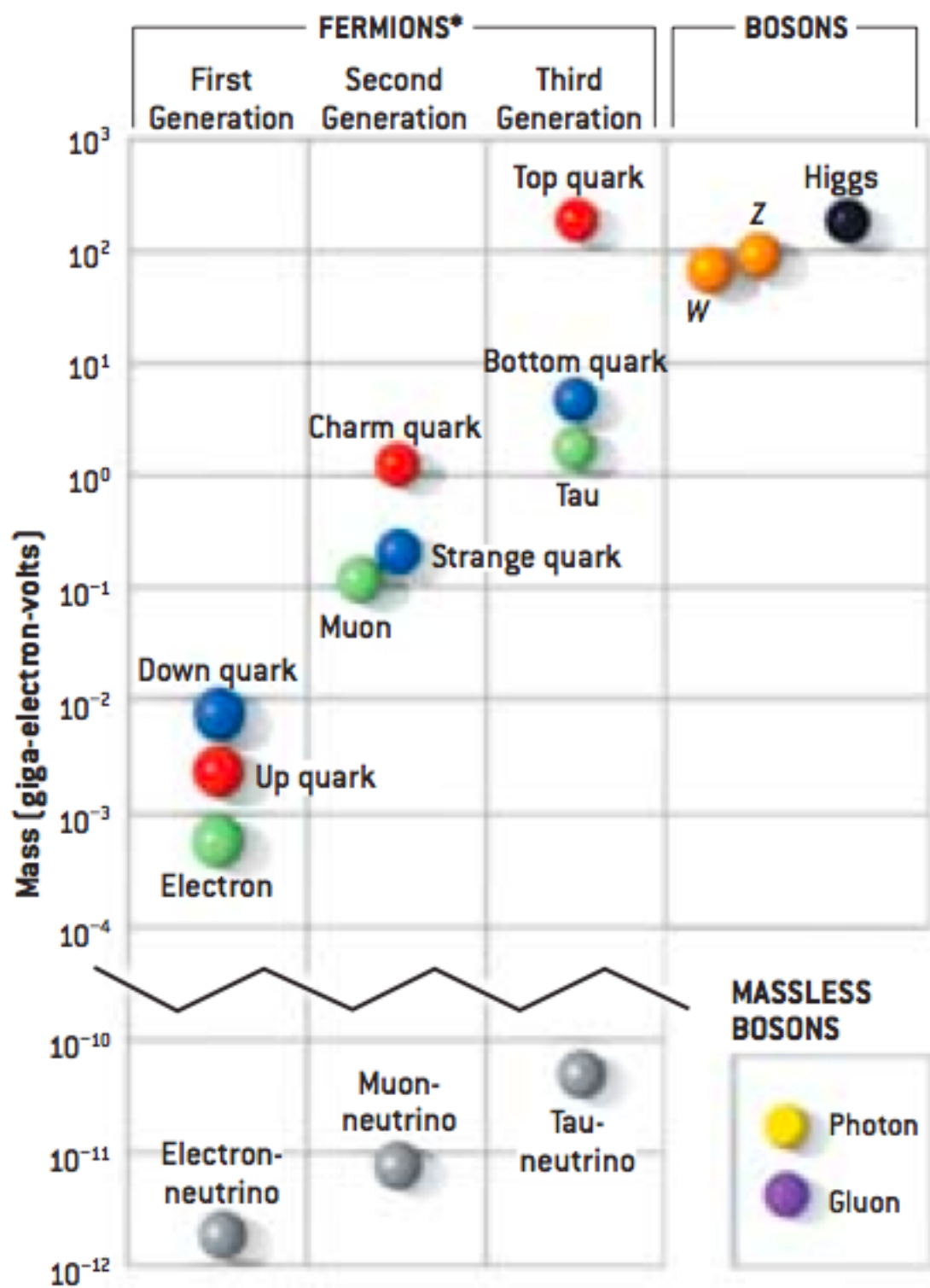


Figure 1.1: Schematic representation of the fundamental particles in the SM arranged according to their masses [1].

chapter are [3, 4, 5, 6].

1.1 Symmetries

When we talk about symmetries we refer to symmetries in the *Lagrangian* of a physical system. The Lagrangian is defined, both in classical and quantum mechanics, as the difference between the kinetic and the potential energy

$$L = T - V \quad (1.1)$$

and it contains all the dynamical informations, like the degrees of freedom, of the system. Once the Lagrangian and the initial conditions are given, then the dynamical evolution is completely determined. So a system is symmetric or invariant if the Lagrangian remains the same after the application of a specific transformation. According to the parameters and the properties of the transformations, symmetries can be, for instance, finite or continuous, local or global, Abelian or non Abelian. This means that the transformations that implement the symmetry can be in a finite or infinite set, that they can be dependent or not on the space-temporal coordinates, and that they commute or not with each other. In the following we will see that the most important symmetries in the SM are the *local gauge* ones, which can be interpreted as principles of invariance under local phase transformations. Today the common belief is that gauge symmetries are at the origin of all the interactions between particles.

In order to understand how the symmetries of a physical system can lead to dynamical consequences, it is necessary to emphasize that each principle of invariance is intimately linked to the conservation of a physical quantity, as proved by the Noether's theorem. Classical examples are the conservation of momentum and angular momentum in case of invariance under spatial translations and rotations respectively. In the quantum world, we deal not only with "external" properties like momenta, but also with "internal" quantities such as quantum numbers and consequently *internal* symmetries as well. The conserved quantities associated to these are, for instance, the electric charge and the color, the strong charge, which are quantum numbers that define each particle.

It is possible to understand how we can get a dynamical description of physical systems from these mathematical properties if we think about a principle of invariance where the transformations depend on spatial and temporal coordinates, i.e. a local symmetry. In this case, the system is unchanged by the transformations undergone in a given region if in the other regions counterbalancing transformations can be performed. This means that the information needed to implement these transformations must be *carried* from one point to another of the system. And this leads to the idea of interactions mediated by carriers. This illustration, despite its simplicity, can give a clue about how the fundamental interactions can stem from symmetries and why they are described by the exchange of particles.

The Standard Model description of particles and forces in Nature is based on the mathematical language of the quantum theory of fields, where particles are excitations of fundamental fields which are functions of, or extend in space and time. This description of particles and forces is based on both special relativity and quantum mechanics. In this theoretical framework, the three fundamental interactions can be described by the product of three mathematical groups: $SU(2)_L \times U(1)_Y \times SU(3)$. This structure is composed by two "independent" blocks, $SU(3)$ which represents the strong force and $SU(2)_L \times U(1)_Y$ that represents the weak and the electromagnetic forces.

In the following part of the chapter, we will focus on the latter block of the Standard Model in order to introduce the theoretical foundations of this thesis.

1.2 The Electro-Weak Interaction

The unification of the weak and the electromagnetic interactions was formalized in the 1960s through the work of Glashow, Weinberg and Salam. Before the introduction of these ideas, the two interactions were described in an independent way and the weak force was depicted by the Fermi Theory as an effective force mediated by massive carriers. Although this was a seminal model, it was not satisfactory both on the experimental and the theoretical point of view.

The first attempt to describe the weak interactions in the framework of symmetries was with the $SU(2)_L$ group. This idea was based on the observation that weak *Charged Currents* (CC) associate pairs of particles like $\nu_e \leftrightarrow e$, as in the beta decay $p \rightarrow n e^+ \nu_e$ or in the neutrino scatterings $\nu_\mu N \rightarrow \mu^- X$ [6]. So, having in mind that these currents might transform particles within the same multiplet and that this doublet structure is present also among quarks, the $SU(2)$ group came straightforward. The ‘L’ means that the CC couple only to *left-handed* particles, meaning that in such particles spin and momentum are opposite in direction. Left-handedness is only one of the two intrinsic state of *chirality* that particles can have. So far there is no experimental evidence of right-handed weak couplings.

As a consequence of the $SU(2)_L$ hypothesis, the two CC should form a triplet together with a third neutral current. Schwinger (1957) suggested that this current could be associated with the photon in order to achieve the electro-weak unification. The problem of this model is that *ad-hoc* couplings are needed to give mass to the weak vector bosons leaving the photon massless. Because of the lack of predictive power this idea was replaced by the hypothesis that the third current was responsible for the weak *Neutral Current* (NC) acting e.g. in neutrino scattering $\nu N \rightarrow \nu X$ or in $\bar{\nu}_e e^- \rightarrow \bar{\nu}_e e^-$ which cannot be mediated by a charged current (Bludman, 1958 [7]). Even this proposal was not well founded because it didn’t account for some experimental observations. For instance, since these three currents are in the same multiplet, they should have the same properties like the ‘V-A’ structure¹ which makes the CC interacting only with left-handed particles. But experiments proved that NC can couple also with right-handed particles.

The turning point was reached when Glashow [8] proposed in 1961 to combine the $SU(2)_L$ with a $U(1)$ group. The quantum numbers related to these symmetries were called *weak isospin* and *weak hypercharge* following the analogy with hadronic isospin and electrical charge. Glashow’s idea now is at the basis of the electro-weak unification and of the Standard Model. Weinberg [9] and Salam [10] improved later this description introducing a spontaneous symmetry breaking that can account for the difference in mass between vector bosons.

We try now to express what just said in the mathematical formalism of quantum field theory. A mathematical illustration of the Charged Currents in terms of the particle fields is:

$$\begin{cases} J_\mu^+ = \bar{\nu} \gamma_\mu \frac{1}{2} (1 - \gamma^5) e = \bar{\nu}_L \gamma_\mu e_L \\ J_\mu^- = \bar{e} \gamma_\mu \frac{1}{2} (1 - \gamma^5) \nu = \bar{e}_L \gamma_\mu \nu_L \end{cases} \quad (1.2)$$

In these formulas, ν and e are the Dirac spinor fields of a neutrino and an electron, γ_μ is a Dirac matrix defined by the Lorentz index μ , and $\frac{1}{2}(1 - \gamma^5)$ is the chiral projector which

¹The notation ‘V-A’ means ‘vectorial - axial’ and it refers to the properties of vectors under parity transformations, i.e. under the inversion of the spatial axes. Axial vectors are invariant, while vectorial vectors invert their directions. Usually physics observables are invariant under parity transformations since we expect that the world seen through a mirror has the same laws as the one that we experience. But this is not true for the weak interactions which involve physics operators of the form $V \pm A$ which are not parity invariant. The sign ‘-’ means that CC couple only one to chiral state, which is called left-handed.

manifests that the weak interactions violate parity and couple only to left-handed particles.

We can write a more compact formulation introducing the left-handed doublet and the right-handed singlet

$$\chi_L = \begin{pmatrix} \nu \\ e^- \end{pmatrix}_L \quad \chi_R = e_R \quad (1.3)$$

The absence of a right-handed neutrino is due the fact that, assuming that neutrinos are massless, they can be only in chiral eigenstates and measurements showed only left-handed neutrinos. So far, there is no experimental evidence of right-handed neutrinos, although neutrino masses appear to be non-vanishing. We will use only the first generation of leptons in order to have a simple illustration, but the same considerations apply to all the generations of leptons and quarks.

Using the chiral doublet and the operators $\sigma_{\pm} = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$ where σ_i are the Pauli matrices, CC becomes

$$J_{\mu}^{\pm} = \bar{\chi}_L \gamma_{\mu} \sigma_{\pm} \chi_L \quad (1.4)$$

The Pauli matrices are 2×2 matrices and they are the fundamental representation of the $SU(2)$ symmetry group. This basically means that applying one of these matrices it is possible to transform one state to another within the $SU(2)$ doublet. Specifically σ_+ transforms the bottom state of the doublet in the top state and the σ_- makes the opposite. The physical meaning of (1.4) is that CC represents the interaction between the up(down) part of the doublet, which is χ_L , with the down(up) part, which is $\sigma_{-(+)}\chi_L$. It is possible to define an orthogonal neutral current

$$J_{\mu}^3 = \bar{\chi}_L \gamma_{\mu} \frac{1}{2} \sigma_3 \chi_L = \frac{1}{2} \bar{\nu}_L \gamma_{\mu} \nu_L - \frac{1}{2} \bar{e}_L \gamma_{\mu} e_L \quad (1.5)$$

So, thanks to the doublet χ_L we can build up a triplet of currents which leads to the $SU(2)$ symmetry group. But the neutral current proposed in this way is not the one observed in the experiments since the former couples only to left-handed particles, while the NC observed experimentally have also a right-handed component.

This is the point where Glashow's idea and the electromagnetic force come into play. Indeed, the electromagnetic interaction couples in the same way both to the left- and the right-handed particles since it doesn't violate parity. So, a proper combination of J_{μ}^3 and the electromagnetic current, neutral as well, might provide a description of the Neutral Current observed in experiments. A naive idea could be to simply add the $U(1)_{em}$ symmetry to $SU(2)_L$, but this doesn't work since the electromagnetic current doesn't have defined properties under $SU(2)_L$ transformations. Indeed, if we look at

$$J_{\mu}^{em} = -\bar{e} \gamma_{\mu} e = -\bar{e}_R \gamma_{\mu} e_R - \bar{e}_L \gamma_{\mu} e_L = \bar{e} \gamma_{\mu} Q e \quad (1.6)$$

where Q is the charge operator and the generator of $U(1)_{em}$, this current couples only to one member of the chiral doublet. The solution is to find another neutral current that is independent from $SU(2)_L$. That means that it should be a singlet under $SU(2)_L$ transformations and that it should produce the electrical current when it is combined to the neutral current of the J_{μ}^i triplet. This can be achieved introducing another $U(1)$ symmetry group with the neutral current

$$J_{\mu}^Y = \bar{\psi} \gamma_{\mu} Y \psi \quad (1.7)$$

where Y is the *hypercharge* and the generator of the new symmetry and ψ represents the particle field. We define the relation between the generators of $SU(2)_L$ and $U(1)_Y$ as in the Gell-Mann-Nishijima relation

$$Q = T^3 + \frac{Y}{2} \quad (1.8)$$

and $T^3 = \frac{1}{2}\sigma_3$ where σ_3 is the third Pauli matrix. The outcome is J_μ^Y neutral and invariant under $SU(2)_L$

$$J_\mu^Y = -\bar{e}_R \gamma_\mu \mathbf{2} e_R - \bar{e}_L \gamma_\mu \mathbf{1} e_L - \bar{\nu}_L \gamma_\mu \mathbf{1} \nu_L \quad (1.9)$$

where the hypercharge eigenvalues are expressed in bold characters.

In this way we have built the $SU(2)_L \times U(1)_Y$ symmetry group, where the two CC of $SU(2)_L$ are associated with the $W^\pm$ bosons and the remaining two neutral currents can be combined in a such a way to describe the electromagnetic current

$$\frac{1}{2}J_\mu^Y + J_\mu^3 = J_\mu^{em} \quad (1.10)$$

Looking at Table 1.1 we can see explicitly why the $U(1)_{em}$ current was not suitable to be combined with $SU(2)_L$. Indeed, the eigenvalues, which are related to conserved quantities like the weak isospin T and the electrical charge Q , have to be the same within each multiplet. This requirement, in fact, is fulfilled by $SU(2)_L$, which is defined by T eigenvalues, and $U(1)_Y$, which is defined by Y eigenvalues, but not by $U(1)_{em}$ with the Q eigenvalues.

Table 1.1: $SU(2)_L \times U(1)_Y$ quantum numbers of leptons.

	T	T^3	Q	Y
ν_L	1/2	+1/2	0	-1
e_L	1/2	-1/2	-1	-1
e_R	0	0	-1	-2

What we have seen so far is that with the $SU(2)_L \times U(1)_Y$ structure, we get four currents. Two of them are charged and have the V-A structure of the CC. The other two are neutral. One violates the parity conservation as the CC and the second doesn't. The combination of these two produces a neutral current with the same properties as the electromagnetic current. It can be proved that the orthogonal current couples to left- and right-handed particles and can be used to describe the NC. So, the group structure proposed by Glashow can account for the experimental observations on CC and NC and provides a good framework for the electro-weak unification.

1.3 Local Gauge Invariance

In order to get deeper in the electro-weak unification, now we need to understand how the forces stem from the local gauge principles of invariance. In the following we will consider a $U(1)$ symmetry, but it is possible to generalize for any symmetry group. In this example a gauge transformation is equivalent to a local phase transformation of a field $\psi(x)$ where the parameter α depends on the spatial and temporal coordinates

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}\psi(x) \quad (1.11)$$

Before going further it is necessary to say few words about the physical meaning of the Lagrangian in the SM. The formalism of quantum field theory contains complex mathematical objects and operators. However, from a phenomenological point of view it is possible to associate symbols and terms of the Lagrangian with physical particles, interactions, kinematic and mass terms.

So, let's consider a Lagrangian² of a fermion with spin 1/2:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (1.12)$$

Eq. 1.12 describes a freely moving fermion ψ and it includes the kinetic term $i\bar{\psi}\gamma^\mu\partial_\mu\psi$ and the mass term $m\bar{\psi}\psi$. If we want this Lagrangian to be invariant under (1.11) then it is necessary to introduce the *covariant derivative* $\mathcal{D}_\mu$ and a *vector field* $\mathcal{A}_\mu$, called *gauge boson*, with specific transformation properties

$$\mathcal{D}_\mu = \partial_\mu - ie\mathcal{A}_\mu \quad (1.13)$$

$$\mathcal{A}_\mu \rightarrow \mathcal{A}_\mu + \frac{1}{e}\partial_\mu\alpha \quad (1.14)$$

where e is the coupling constant associated to the $U(1)$, i.e. the characteristic parameter of this group, which determines the strength of the interactions that that $U(1)$ induces. The name 'constant' can be misleading since this value is not properly a constant, but it can 'run' because of virtual effects and the vacuum polarization. The reason why $\mathcal{A}_\mu$ has to be necessary a vector is because it needs to have the same properties as the relativistic derivative ∂_μ . So, the lagrangian which is invariant under (1.11) is

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\mathcal{D}_\mu\psi - m\bar{\psi}\psi = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi + e\bar{\psi}\gamma^\mu\psi\mathcal{A}_\mu \quad (1.15)$$

We can see that now there is a new term $e\bar{\psi}\gamma^\mu\psi\mathcal{A}_\mu$ which generates an interaction between the field current $\bar{\psi}\gamma^\mu\psi$ and vector boson $\mathcal{A}_\mu$. The strength of this interaction is determined by the value of the constant e .

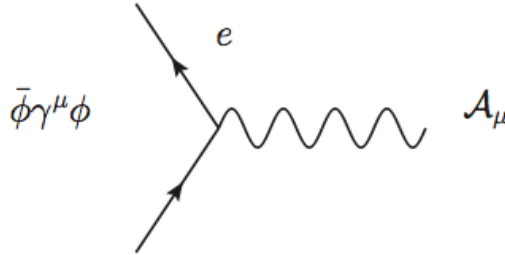


Figure 1.2: Feynman diagram of the interactions $e\bar{\psi}\gamma^\mu\psi\mathcal{A}_\mu$.

Interaction terms like this are represented pictorially by *Feynman diagrams* like in Fig. 1.2. These are more than an illustration since each branch of the diagram is related to a part of the interaction term in the Lagrangian. The solid line is the field current, the curly line is the gauge boson and the vertex is associated to the coupling constant. This is a simple diagram, but this schematization is very powerful since it allows to translate the Lagrangian in a sets of diagrams and to derive *Feynman rules* to compute the amplitude of each diagram. In this way manipulations on the Lagrangian or on the Feynman diagrams are totally equivalent and using the diagrams it is possible to visualize better the physical processes.

All the observables in the Relativistic Quantum Field Theories are treated in the perturbative approach. This means that they are computed as expansions in powers of the coupling constant. This means that the most simple Feynman diagram of the given process

²In the following we will not use the Lagrangian L to describe the system, but the Lagrangian density $\mathcal{L}$ which is related to L by $L = \int \mathcal{L} d\vec{x}$. However for simplicity we will call $\mathcal{L}$ Lagrangian as well.

is combined to higher order diagrams with with real and virtual corrections. This expansion has to be performed properly since divergences may occur. A *Renormalizable Theory*, such as the SM, has the property that all these divergences can be eliminated by redefining quantities such as the mass and the charge of the electron that appear in the Lagrangian. Indeed, the “bare” value of these variables are not accessible experimentally because there are always screened by virtual effects.

Going back to what we obtained in (1.15), in order to have a complete Lagrangian we should include also a kinetic and a mass term for the new field. Even if it is possible to find a kinetic term which doesn't break the local phase invariance, the same cannot be done for the mass term. The reason is that $m^2 \mathcal{A}_\mu \mathcal{A}^\mu$ breaks explicitly the invariance (1.11).

So the most general Lagrangian for a fermion with a $U(1)$ symmetry is

$$\mathcal{L} = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi + e \bar{\psi} \gamma^\mu \psi \mathcal{A}_\mu - \frac{1}{4} \mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu} \quad (1.16)$$

where $\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$ and the gauge field has to be *massless*. It is possible to get a clue about the reason why gauge bosons cannot have a mass through a simple consideration. Since there is no limitation in the size of the region where the local gauge invariance is valid, then the interaction needs to have an infinite range, which means that the mediator has to be massless. This is a problem since we want to describe the weak interaction which is mediated by massive gauge bosons. A choice could be to put a mass term ‘by-hand’ in the Lagrangian breaking the gauge invariance, but this symmetry is really crucial since it prevents, for instance, from the occurrence of divergences in the theory.

The solution can be a mechanism of *spontaneous symmetry breaking* which leaves the Lagrangian still invariant, but with a non symmetric appearance. This is what Weinberg [9] and Salam [10] proposed in 1967 and 1968.

1.4 The Higgs Mechanism

Weinberg and Salam implemented the method of Spontaneous Symmetry Breaking that Higgs [11, 12], Brout and Englert [13] and Kibble, Guralnik and Hagen [14] developed independently in 1964. Even though this mechanism is commonly called the *Higgs mechanism*, the proper name should be the *BEHHGK mechanism*. The decisive idea is to add in the Lagrangian a new field, the *Higgs field*, which has a *Vacuum Expectation Value* (VEV) different from zero. This means that the state with lowest energy for this field, the one realized in Nature, is reached not when the field is null, but when it permeates all the space.

A pedagogic example to understand how this field can generate the masses of the gauge bosons is to consider the Lagrangian of a complex scalar field with a *Abelian $U(1)$ global gauge symmetry*. This physical system can be described as

$$\mathcal{L} = T - V = (\partial_\mu \phi^*) (\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2 \quad (1.17)$$

where $\lambda > 0$. The potential is in the most general renormalizable form invariant under $U(1)$. The same Lagrangian can be written parametrizing the complex field with two real ones like $\phi = (\phi_1 + i\phi_2)/2$

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi_1)^2 + \frac{1}{2} (\partial_\mu \phi_2)^2 - \frac{1}{2} \mu^2 (\phi_1^2 + \phi_2^2) - \frac{\lambda}{4} (\phi_1^2 + \phi_2^2)^2 \quad (1.18)$$

If $\mu^2 > 0$, the potential has a minimum at $\phi = 0$ and the result is a system with two massive scalar fields degenerated in mass with $m = \mu$. This result is not useful for our purpose because we are looking for a way in which one field gives mass to another dynamically.

If $\mu^2 < 0$, then the ground state is no more at $\phi = 0$, but in the continuum of states $\phi_1^2 + \phi_2^2 = v^2 = -\mu^2/\lambda$. So, in order to have a description in terms of particles, i.e. fluctuations or quanta of the fields, it is necessary to make an expansion around a stable state. This means that we need to *choose* the VEV for the ground state. One possibility is to take a real VEV

$$\begin{cases} \langle 0|\phi_1|0\rangle = v \\ \langle 0|\phi_2|0\rangle = 0 \end{cases} \quad (1.19)$$

and perform the translation to the ground state

$$\phi(x) = \sqrt{\frac{1}{2}} (v + \eta(x) + i\xi(x)) \quad (1.20)$$

The Lagrangian becomes

$$\mathcal{L}' = \frac{1}{2} (\partial_\mu \xi)^2 + \frac{1}{2} (\partial_\mu \eta)^2 + \mu^2 \eta^2 + \dots \quad (1.21)$$

One of the two fields, η , has acquired a mass $m_\eta = \sqrt{-2\mu^2}$, while ξ is massless. The latter field is called the *Goldstone boson*. This results can be easily understood having a look at the shape of the potential, which is often called the ‘Mexican Hat’, as in Fig. 1.3. There are two

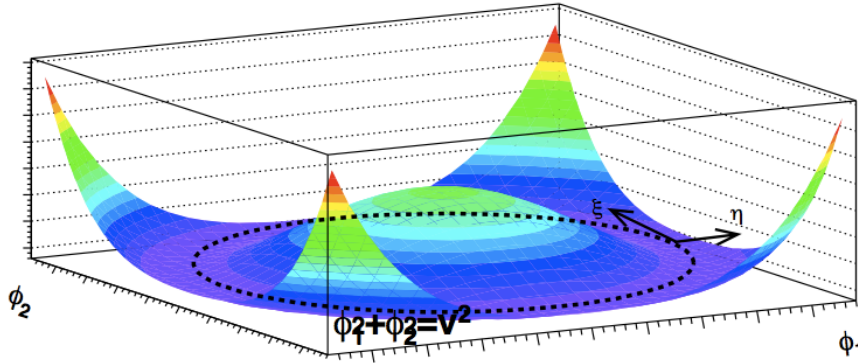


Figure 1.3: Illustration of the Higgs potential, the dashed line represents the continuum of states at the lowest energy and the two arrows are the independent excitations around the ground state.

independent fluctuations around the ground state, one orthogonal to the circle $\phi_1^2 + \phi_2^2 = v^2$, which requires energy and which is related to the massive field η , and the second tangent to the circle. The latter excitation connects states at the same energy so it is associated to massless excitations, ξ . The *Goldstone's Theorem* generalizes this exercise saying that a Goldstone boson appears whenever a continuous symmetry is not apparent in the ground state.

So far we have found a way to give mass to a field dynamically, i.e. without including a mass term in the Lagrangian, but we still have massless particles. The next step is to consider no more a global, but a *local gauge* $U(1)$ symmetry. As we have seen in the previous section, the local gauge symmetry requires the introduction of the covariant derivative together with a vector field, so the Lagrangian this time is

$$\mathcal{L} = (\partial^\mu + ie\mathcal{A}^\mu) \phi^* (\partial_\mu - ie\mathcal{A}_\mu) \phi - \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2 - \frac{1}{4} \mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu} \quad (1.22)$$

We consider again the case where $\lambda > 0$ and $\mu^2 < 0$. The same choice of the VEV and the relative translation (1.20) lead to the new Lagrangian

$$\mathcal{L}' = \frac{1}{2} (\partial_\mu \xi)^2 + \frac{1}{2} (\partial_\mu \eta)^2 - v^2 \lambda \eta^2 + \frac{1}{2} e^2 v^2 \mathcal{A}_\mu \mathcal{A}^\mu - ev \mathcal{A}_\mu \partial^\mu \xi - \frac{1}{4} \mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu} + \dots \quad (1.23)$$

There are three particles in this system:

- | | | |
|---|-----------------------------|-----------------------|
| 1. a massless scalar (Goldstone) boson ξ | $m_\xi = 0$ | d.o.f. ³ 1 |
| 2. a massive scalar boson η | $m_\eta = \sqrt{2\lambda}v$ | d.o.f. 1 |
| 3. a massive vector gauge boson $\mathcal{A}_\mu$ | $m_{\mathcal{A}} = ev$ | d.o.f. 3 |

Our aim to have a massive gauge boson seems to be achieved, but there is still a problem since not all these particles correspond to physical states. A quick counting of the initial and the final degrees of freedom makes this clear. Indeed, since the physics of the system is not affected by the translation (1.20), $\mathcal{L}$ and $\mathcal{L}'$ are equivalent. So, the number of degrees of freedom has to be the same, but in $\mathcal{L}$ there are two scalar real fields and a massless vector field with a total $n_{dof} = 4$, while in $\mathcal{L}'$ $n_{dof} = 5$. The additional degree of freedom is not physical, but it corresponds to the freedom of performing a gauge transformation. So it is necessary to find a proper gauge transformation that leaves the Lagrangian with only the physical particles. This can be done reformulating the previous translation

$$\phi(x) = \sqrt{\frac{1}{2}} (v + \eta(x) + i\xi(x)) \simeq \sqrt{\frac{1}{2}} (v + \eta(x)) e^{i\xi(x)/v} \quad (1.24)$$

If we think of this transformation as a combination of a translation plus a local phase rotation, then we need to transform accordingly also the gauge field

$$\left\{ \begin{array}{l} \phi(x) = \sqrt{\frac{1}{2}} (v + h(x)) e^{i\theta(x)/v} \\ \mathcal{A}_\mu \rightarrow \mathcal{A}_\mu + \frac{1}{ev} \partial_\mu \theta \end{array} \right. \quad (1.25)$$

so $\mathcal{L}$ changes into

$$\mathcal{L}'' = \frac{1}{2} (\partial_\mu h)^2 - \lambda v^2 h^2 + \frac{1}{2} e^2 v^2 \mathcal{A}_\mu^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4 + \frac{1}{2} e^2 \mathcal{A}_\mu^2 h^2 + v e^2 \mathcal{A}_\mu^2 h - \frac{1}{4} \mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu} \quad (1.26)$$

The disappearance of the field $\theta(x)$ can be easily shown applying (1.25) to (1.22). Indeed it vanishes in the Higgs potential since it enters as a phase and the potential is phase invariant. Moreover it cannot survive in the kinetic term of the gauge field since this is invariant under gauge transformation. So we only need to check the kinetic term of the scalar field. Let's look at the term

$$\begin{aligned} (\partial^\mu + ie\mathcal{A}'^\mu) \phi'^* &= (\partial^\mu + ie\mathcal{A}^\mu + \frac{i}{v} \partial^\mu \theta) \sqrt{\frac{1}{2}} (v + h) e^{-i\theta/v} \\ &= \sqrt{\frac{1}{2}} \left(-i\partial^\mu \theta - \frac{ih}{v} \partial^\mu \theta + ie\mathcal{A}^\mu (v + h) + i\partial^\mu \theta + \frac{ih}{v} \partial^\mu \theta \right) e^{-i\theta/v} \\ &= \sqrt{\frac{1}{2}} ie\mathcal{A}^\mu (v + h) e^{-i\theta/v} \end{aligned} \quad (1.27)$$

The product of this times its complex conjugate makes θ disappear. So the transformed Lagrangian doesn't depend on θ at all.

³degrees of freedom

Finally we have reached our goal: starting from two scalar real fields and a massless vector field, we obtained a system of two interacting massive particles, the gauge boson $\mathcal{A}_\mu$ with $m_{\mathcal{A}} = ev$, and scalar particle h with $m_h = \sqrt{2}\lambda v^2$. The total amount of degrees of freedom is 4, so there is no unphysical state as before. This is the *Higgs mechanism* and we can reinterpret the result saying that one of the scalar field has been ‘eaten’ by the vector field acquiring mass.

In order to apply this mechanism to the $SU(2)_L \times U(1)_Y$, it is still necessary to understand what happens to a $SU(2)$ local gauge symmetry. In this case the scalar field is a complex doublet

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.28)$$

and the Lagrangian is

$$\mathcal{L} = (\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \quad (1.29)$$

The $SU(2)$ symmetry group has three generators, the Pauli matrices $\frac{1}{2}\sigma_i$ where $i = 1, 2, 3$. Consequently the phase transformation is defined by three parameters and the covariant derivative has three gauge bosons.

$$\phi \rightarrow \phi' = e^{i\alpha_i \sigma_i / 2} \phi \quad (1.30)$$

$$\mathcal{D}_\mu = \partial_\mu + ig \frac{\sigma_i}{2} W_\mu^i \quad (1.31)$$

$$W_\mu^i \rightarrow W_\mu^i - \frac{1}{g} \partial_\mu \alpha^i - (\vec{\alpha} \times \vec{W}_\mu)^i \quad (1.32)$$

g is a new coupling associated to the $SU(2)$ symmetry group.

The Lagrangian is

$$\mathcal{L} = \left(\partial_\mu \phi + ig \frac{\vec{\sigma}}{2} \cdot \vec{W}_\mu \phi \right)^\dagger \left(\partial^\mu \phi + ig \frac{\vec{\sigma}}{2} \cdot \vec{W}^\mu \phi \right) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 - \frac{1}{4} \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} \quad (1.33)$$

where

$$\vec{W}_{\mu\nu} = \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu - g \vec{W}_\mu \times \vec{W}_\nu \quad (1.34)$$

It is important to highlight that the vector products in (1.32) and (1.34) arise from the non-Abelian structure of the symmetry group, which means that the generators don’t commute with each other. This feature will produce interactions between the gauge bosons.

As in the previous examples, the lowest energy states are at $\phi^\dagger \phi = v^2/2$. We choose a VEV that breaks the $SU(2)$ symmetry and that is real

$$\langle 0 | \phi | 0 \rangle = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.35)$$

and we perform the transformation

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} e^{i\vec{\sigma} \cdot \vec{\theta}(x)/v} \quad (1.36)$$

The result is equivalent to what we got in the $U(1)$ example: the Goldstone bosons θ_i disappears, eaten by the three gauge bosons W_μ^i which gain a mass $m_W = \frac{1}{2}gv$.

1.5 The Electro-Weak Symmetry Breaking

Now we can apply the Higgs mechanism to $SU(2)_L \times U(1)_Y$. As Weinberg and Salam suggested at the end of 1960s, the aim is to build a theory with one massless boson, the photon, three massive gauge bosons, the $W^\pm$ and the Z , one scalar massive particle, the Higgs boson, and possibly no massless Goldstone bosons.

Let's start defining $T_i = \frac{\sigma_i}{2}$ and Y as the generators for $SU(2)_L$ and $U(1)_Y$, respectively. They satisfy the relation

$$Q = T^3 + \frac{1}{2}Y \quad (1.37)$$

$SU(2)_L$ is characterized by a coupling constant g and three gauge bosons W_μ^i and $U(1)_Y$ by g' and B_μ . We will consider only the Lagrangian for leptons, as we did before, since the generalization that includes quarks is straightforward. The only difference is that, since quarks are massive, there will be two right-handed singlets, u_R and d_R , instead of one.

The covariant derivative is

$$\mathcal{D}_\mu = \partial_\mu - ig\vec{T} \cdot \vec{W}_\mu - i\frac{g'}{2}YB_\mu \quad (1.38)$$

and the Lagrangian is

$$\mathcal{L} = \bar{\chi}_L \gamma_\mu \mathcal{D}^\mu \chi_L + \bar{\chi}_R \gamma_\mu \mathcal{D}^\mu \chi_R - \frac{1}{4} \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} - \frac{1}{4} \mathcal{B}_{\mu\nu} \mathcal{B}^{\mu\nu} \quad (1.39)$$

where $\mathcal{B}_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$. Both fermions and gauge bosons are massless because a gauge boson mass term is not gauge invariant and a fermion mass term is not $SU(2)_L$ invariant. Indeed

$$m\bar{e}e = m\bar{e} \left[\frac{1}{2}(1 - \gamma^5) + \frac{1}{2}(1 + \gamma^5) \right] e = m(\bar{e}_R e_L + \bar{e}_L e_R) \quad (1.40)$$

since e_L and e_R belong to two different multiplets of $SU(2)_L$, then such a mass term cannot be invariant under $SU(2)_L$ transformations.

A possible choice for the Higgs field is a weak-isospin doublet scalar field with hypercharge $+1$, which means that the component with $T^3 = +1/2$ has a positive electrical charge, while the other component is neutral

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.41)$$

The Higgs sector of the Lagrangian has the same formulation as 1.29.

Now, the crucial step is the choice of the ground state since it is necessary to break the $SU(2)_L \times U(1)_Y$ symmetry, saving the $U(1)_{em}$. In order to do that, the VEV has to be neutral since if $Q\phi = 0$, performing a $U(1)_{em}$ transformation

$$\phi \rightarrow \phi' = e^{i\alpha(x)Q}\phi \approx (1 + i\alpha Q)\phi = \phi \quad (1.42)$$

This requirement leads to the following choice

$$\phi_0 = \langle 0|\phi|0 \rangle = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.43)$$

which breaks both $SU(2)_L$ and $U(1)_Y$, but not $U(1)_{em}$. This decision leaves only one *CP-even scalar neutral field*, whose quanta are the *Higgs bosons*, and it determines the mass spectrum of the gauge bosons. In fact, it is possible to calculate the boson masses from

$$(\mathcal{D}_\mu \phi_0)^\dagger (\mathcal{D}^\mu \phi_0) = \left| \left(-ig\frac{\vec{\sigma}}{2} \cdot \vec{W}_\mu - ig'\frac{Y}{2}B_\mu \right) \phi_0 \right|^2 \quad (1.44)$$

The outcome is a couple of W bosons

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad m_W = \frac{1}{2}vg \quad (1.45)$$

and then, diagonalizing the remaining terms in the W_μ^3 and B_μ basis, we get

$$\begin{aligned} A_\mu &= \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}} & m_A &= 0 \\ Z_\mu &= \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} & m_Z &= \frac{1}{2}v\sqrt{g^2 + g'^2} \end{aligned} \quad (1.46)$$

$W^\pm$ will be associated to the CC, A to the electromagnetic currents and Z to the NC. This result is very stringent. Once A_μ is recognized as the photon, then from the interaction terms that come from 1.39 the three couplings are related to each others by

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \quad (1.47)$$

Usually the ratio between g and g' is defined through the *Weinberg's angle* θ_W

$$\tan \theta_W = \frac{g'}{g} \quad (1.48)$$

so

$$e = g \sin \theta_W = g' \cos \theta_W \quad (1.49)$$

It is important to notice that θ_W is a free parameter of the model since it is the ratio of two coupling constants related to independent symmetry groups.

Up to this point, we have presented a quick overview of the Electro-Weak Unification in the SM. It is worth mentioning some of the predictions that support this model and strengthen the belief that this unification is actually performed by a spontaneous symmetry breaking.

We have seen that this model requires a very small number of parameters. Basically they are the two constants g and g' related to the symmetry $SU(2)_L \times U(1)_Y$ and the two parameters of the Higgs potential μ and λ . Usually they are parametrized with the observables α , the fine structure constant, G_F , the Fermi constant, m_Z , the Z boson mass, and m_H , the mass of the Higgs boson

$$\alpha = \frac{g^2 g'^2}{4\pi(g^2 + g'^2)} \quad (1.50)$$

$$m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \quad (1.51)$$

$$G_F = \frac{1}{\sqrt{2}v^2} \quad (1.52)$$

$$m_H = \sqrt{2\lambda}v^2 \quad (1.53)$$

where $v^2 = -\mu^2/\lambda$. G_F is the strength of the weak interaction in the effective and point-like description of weak interactions formulated by Fermi. By measuring G_F in processes like the muon decay we find that $v \approx 246$ GeV. This value set the scale of the Electroweak Symmetry Breaking, but it is not predicted by the SM. The relation between G_F and the VEV v comes from

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} \quad (1.54)$$

which is a comparison between the Fermi theory and the CC in the limit of highly massive gauge bosons.

Once we know the values of α , G_F and m_Z we can predict from (1.45) and (1.46) the mass of the W boson at the lowest order

$$m_W = m_Z \cos \theta_W \approx 80 \text{ GeV} \quad (1.55)$$

which has been confirmed experimentally by the precise measurements of M_W and M_Z at LEP, reported here [6]. Then, if we take into account the relative strength of CC and NC, this can be expressed with a parameter ρ , which at the lowest order should be equal to one according to (1.55)

$$\rho = \frac{m_W^2}{m_Z^2 \cos^2 \theta_W} = 1 \quad (1.56)$$

and this is proven experimentally [15].

These results are a direct consequence of the model and of the choice of a doublet for the Higgs field. Other representations would have lead, for instance, to different values of ρ .

1.6 The Fermionic Masses

Previously we described how the Higgs mechanism can account for the generation of the gauge boson masses without breaking explicitly the gauge invariance, but we still have to deal with massless fermions due to the $SU(2)_L$ symmetry. The impressive power of the Higgs mechanism is that the same doublet of complex scalar fields can generate also mass terms for fermions. Indeed, we can simply add a new sector $SU(2)_L \times U(1)_Y$ gauge invariant in the Lagrangian:

$$- G_e \left[(\bar{\nu}_e, \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^-, \bar{\phi}^0) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right] \quad (1.57)$$

here, again, we write only the part for the first family of leptons. The other families will have the same terms with proper couplings G_μ and G_τ . These terms are referred as the *Yukawa sector* since analogous couplings between a scalar field and a fermionic field of spin 1/2 was proposed by Yukawa to describe the interaction of a nucleon with a pion in the nuclear force.

Then, performing the spontaneous symmetry breaking

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.58)$$

only the neutral field $h(x)$ remains since the other three disappears. (1.57) becomes

$$- \frac{G_e}{\sqrt{2}} v (\bar{e}_L e_R + \bar{e}_R e_L) - \frac{G_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) h(x) \quad (1.59)$$

if $m_e = \frac{G_e v}{\sqrt{2}}$, we get the mass term for the electron and the interaction between the electron and the Higgs field

$$- m_e \bar{e} e - \frac{m_e}{v} \bar{e} e h \quad (1.60)$$

The important feature of this outcome is that the strength of the coupling between the Higgs field and the fermions is proportional to the mass of the fermions. We can compare this with the coupling between the Higgs field and the massive gauge bosons which is proportional to the square of the mass of the gauge bosons (1.26). So the relative intensity of the Higgs couplings to the electron and to the W boson is roughly $m_e/m_W^2 \approx 8 \cdot 10^{-11}$. This comment

will be relevant when we will study the Higgs decay branching ratio, i.e. the probability that the Higgs boson decays in a specific final state.

The generation of masses for quarks is slightly different due to the fact that there are two right-handed singlets u_R and d_R . What we need is the charge conjugate Higgs field

$$-i\tau_2\phi^* = \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} \quad (1.61)$$

An important feature which is a fundamental property of the $SU(2)$ group is that the doublet and the anti-doublet transform exactly in the same way. Consequently we still have a gauge invariant Lagrangian using also the charge conjugate Higgs field

$$-G_d(\bar{u}, \bar{d})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R - G_u(\bar{u}, \bar{d})_L \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} u_R + h.c. \quad (1.62)$$

With the symmetry breaking only the neutral fields survive. So

$$-m_d\bar{d}d - m_u\bar{u}u - \frac{m_d}{v}\bar{d}dh - \frac{m_u}{v}\bar{u}uh \quad (1.63)$$

Here we have neglected that weak currents don't interact with eigenstate of mass and the actual $SU(2)_L$ doublets are defined by the flavor mixing represented by the CKM matrix. However this doesn't change substantially what we have done so far. We only need to be aware that the coupling G_u and G_d are not scalar, but 3×3 matrices.

As a general comment on the Yukawa sector, we can say that this is not as satisfactory as the gauge sector. Indeed, all the couplings, such as G_e , are not predicted by the theory and they should be un-naturally spread in a wide range of intensity. This can be interpreted as a hint that at least this sector is only a model and a more fundamental theory is needed to completely explain the origin of masses of fermions. Nevertheless once the Higgs boson is detected and we know its mass, then we can easily prove whether this sector is correct comparing the expected and the measured branching ratios.

Chapter 2

The SM Higgs Boson Quest

In the previous chapter we have introduced the reader to the Electro-Weak Symmetry Breaking providing some predictions in support of this model. But the ultimate proof would be the discovery of the Higgs boson. A lot of efforts were spent in the last twenty years, after the detection of the remaining missing parts of the SM. Indeed, after the discovery of W and Z bosons (1983 [16]) and of the top quark ([17, 18]) at the predicted masses, then the idea that the renormalizable, spontaneously broken, non-Abelian chiral gauge theory is a fundamental law flourished.

In this chapter we will give a review of the theoretical argumentations that can give us some clues about the Higgs boson mass and the experimental studies performed so far. It is important to remember that these investigations can be carried on because all the features of the Higgs boson (branching ratios, widths, couplings and so on) are completely determined once we choose the value of its mass.

2.1 Theoretical Limits

Upper and lower constraints on the Higgs boson mass can be set from theoretical argumentations [5, 19, 20]. They come from very reasonable considerations, but they cannot provide stringent limits since they depend on the absence of new physics up to a cut-off energy scale. As we will see, this means that we can set a range of masses that is valid as long as virtual effects of new physics enter in the calculation of the Higgs boson mass.

2.1.1 Unitarity

This is the most tight prediction and it comes from the SM calculation of the scattering amplitude of longitudinal gauge bosons $V_L V_L \rightarrow V_L V_L$ where $V = W^\pm, Z$. If we don't include virtual effects of the Higgs boson or new physics, then this amplitude grows proportionally to the center of mass energy of the scattering. This behavior violates the unitarity, which means that at some energy this process has the probability to occur greater than one. This argument leads to two conclusions:

1. In case the Higgs boson exists and its mass is $m_H \ll s$, then the amplitude grows as the square of m_H and this means that $m_H \leq 1$ TeV;
2. If the Higgs boson doesn't exist or its mass is $m_H \gg s$, then there must be a critical energy scale above which new physics appears and this scale should be around 1 or 2 TeV.

In other words, if we can explore energies up to 2 TeV, then we will be able either to discover the Higgs boson either to exclude it and to reach the limit where the SM fails. Luckily, this

is now possible thanks to the collisions produced at the Large Hadron Collider that will be described later.

2.1.2 Triviality

Another argument for the fact that the Higgs boson mass cannot be arbitrarily large comes from the *self-coupling* λ . Indeed, as all the renormalizable coupling constant, λ runs accordingly to

$$\lambda(E) = \frac{\lambda(v)}{1 - \frac{3\lambda(v)}{8\pi^2} \ln(E/v)} \quad (2.1)$$

Similarly to the QED, this coupling is not asymptotically free and for $E \rightarrow \infty$ it approaches the *Landau pole*. Since the Higgs boson mass is $m_H = \sqrt{2\lambda}v$, we can set an upper limit requiring that λ is finite. This means that from (2.1)

$$\frac{1}{\lambda(E)} = \frac{1}{\lambda(v)} - \frac{3}{4\pi^2} \ln(E/v) \quad (2.2)$$

$1/\lambda(\Lambda) > 0$ up to a large scale Λ where new physics appears. So,

$$m_H^2 < \frac{8\pi^2 v^2}{3 \ln(\Lambda^2/v^2)} \quad (2.3)$$

If the cut-off energy is at the Plank scale around 10^{16} GeV, then the Higgs boson mass should be little

$$m_H < 160 \text{ GeV} \quad (2.4)$$

the lower we set this energy, the looser is the upper constraint on the Higgs boson mass. In the Higgs boson mass calculation we cannot neglect the contribution from top and gauge boson loops. If we include these corrections and we require that the theory is perturbative (i.e. λ is finite) below a given energy, then we can set an upper limit on m_H as a function of the top quark mass. For $m_t = 175$ GeV, $m_H < 170$ GeV [21]. Fig. 2.1 shows the upper limits that prevent the self-interaction to become infinite. In the picture there are also lower limits that will be explained next. As we said, this is not a tight constraint. Indeed, it is possible to find a Higgs boson heavier than predicted and this would be a hint that new physics enters at an energy scale below what expected.

2.1.3 Vacuum Stability

Theoretical arguments can provide also a lower bound for the Higgs boson mass due to the condition $V(v) < V(0)$. This is equivalent to the requirement that λ is positive otherwise the potential is unbounded from below and there is no ground state. In this case it is possible to get a lower bound as a function also of the top mass [22]. If we set the cut-off energy at the Planck scale 10^{16} GeV, then

$$m_H(\text{ GeV}) > 130.5 + 2.1(m_t - 174) \quad (2.5)$$

if we take the more conservative assumption that the SM is valid up to the energy scale of 1 TeV, then

$$m_H(\text{ GeV}) > 71 + 0.74(m_t - 174) \quad (2.6)$$

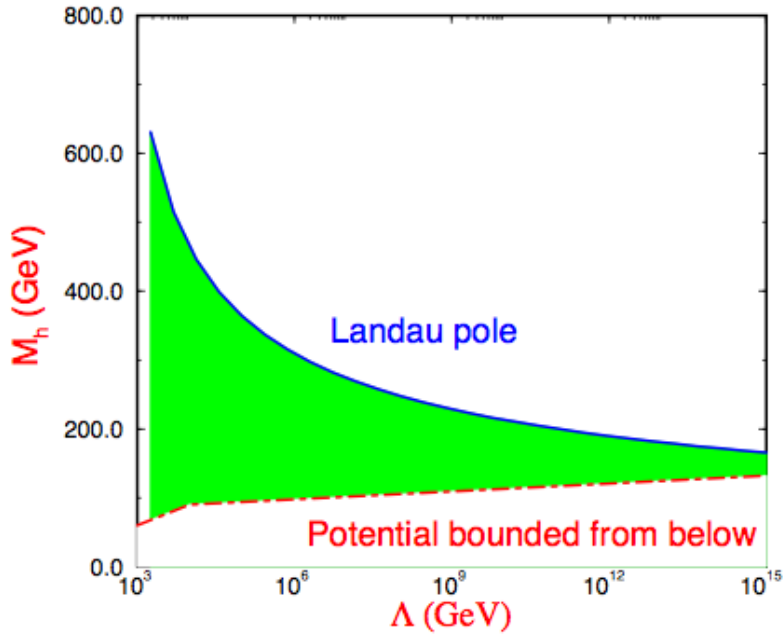


Figure 2.1: Theoretical constraints on the Higgs boson mass as a function of the cut-off energy scale.

2.2 Electro-Weak Radiative Corrections

Before looking for the Higgs boson directly produced in accelerators, physicists tried to detect the virtual effects of this particle in measurements of electro-weak precision observables (EWPO). The drawback of these studies is that usually virtual corrections (Fig. 2.2) depend on the logarithm of the Higgs boson mass, while the top quark gives contributions related to the square of its mass. So the quantum effects of the Higgs boson are pretty weak and it is hard to set tight constraints on its mass.

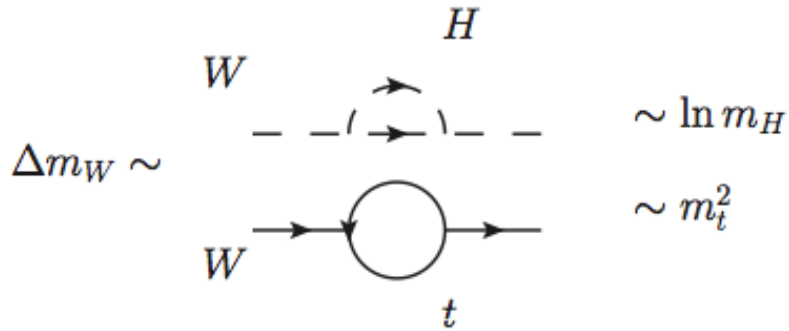


Figure 2.2: Virtual effects on the W mass from a Higgs and a top loop.

It is possible to demonstrate that at one loop all electroweak parameters have at most a logarithmic dependence on m_H . This is due to the so-called *screening theorem* [23]. Since

the form of the electro-weak corrections involving the Higgs boson mass is [19]

$$g^2 \left(\ln \frac{m_H}{m_W} + g^2 \frac{m_H^2}{m_W^2} \right) \quad (2.7)$$

the quadratic term is ‘screened’ by additional two powers of g and so the corrections are dominated by the logarithmic term.

The two observables that are most sensitive to the Higgs boson mass are the W boson mass and the effective leptonic weak mixing angle $\sin^2 \theta_{\text{eff}}$. Even if the precision of m_W is better compared to $\sin^2 \theta_{\text{eff}}$, the latter has a more pronounced dependence on m_H [24]. $\sin^2 \theta_{\text{eff}}$ is a particular renormalization prescription of

$$\sin^2 \theta_W = 1 - \frac{m_W^2}{m_Z^2} \quad (2.8)$$

which can be taken as the tree-level definition of the Weinberg angle. $\sin^2 \theta_{\text{eff}}$ value is mainly determined by asymmetry measurements like the left-right or the b quark forward-backward asymmetry.

Fig. 2.3 [25] shows the comparison between the direct and indirect measurements on m_W and m_t and the relationship with the Higgs boson mass. It is possible to see that bounds are tighter on the top quark mass with respect to the Higgs boson mass.

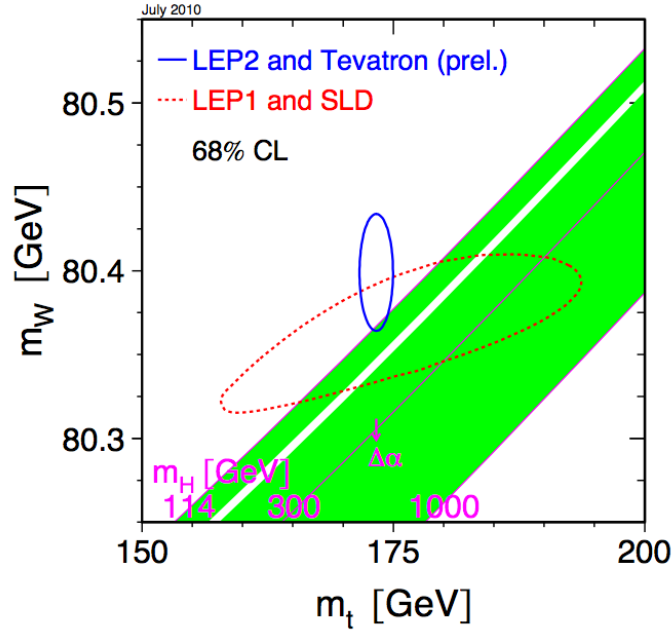


Figure 2.3: Comparison of the indirect constraints on m_W and m_t based on LEP-I/SLD data (dashed contour) and the direct measurement from the LEP-II/Tevatron experiments (solid contour). The green area indicates the relationship for the masses as a function of the SM Higgs boson mass [26].

Fig. 2.4 shows all the measurements of EWPO updated to November 2010 [27]. These values are compared to predictions from the global fit of all these variables. The model parameters that are free to vary in the fit are m_Z , m_c , m_b , m_t , m_H , $\Delta\alpha_{\text{had}}(m_Z^2)^1$, $\alpha_s(m_Z^2)^2$

¹The five-quark hadronic contribution to the running QED coupling constant.

²The strong coupling constant at the Z peak.

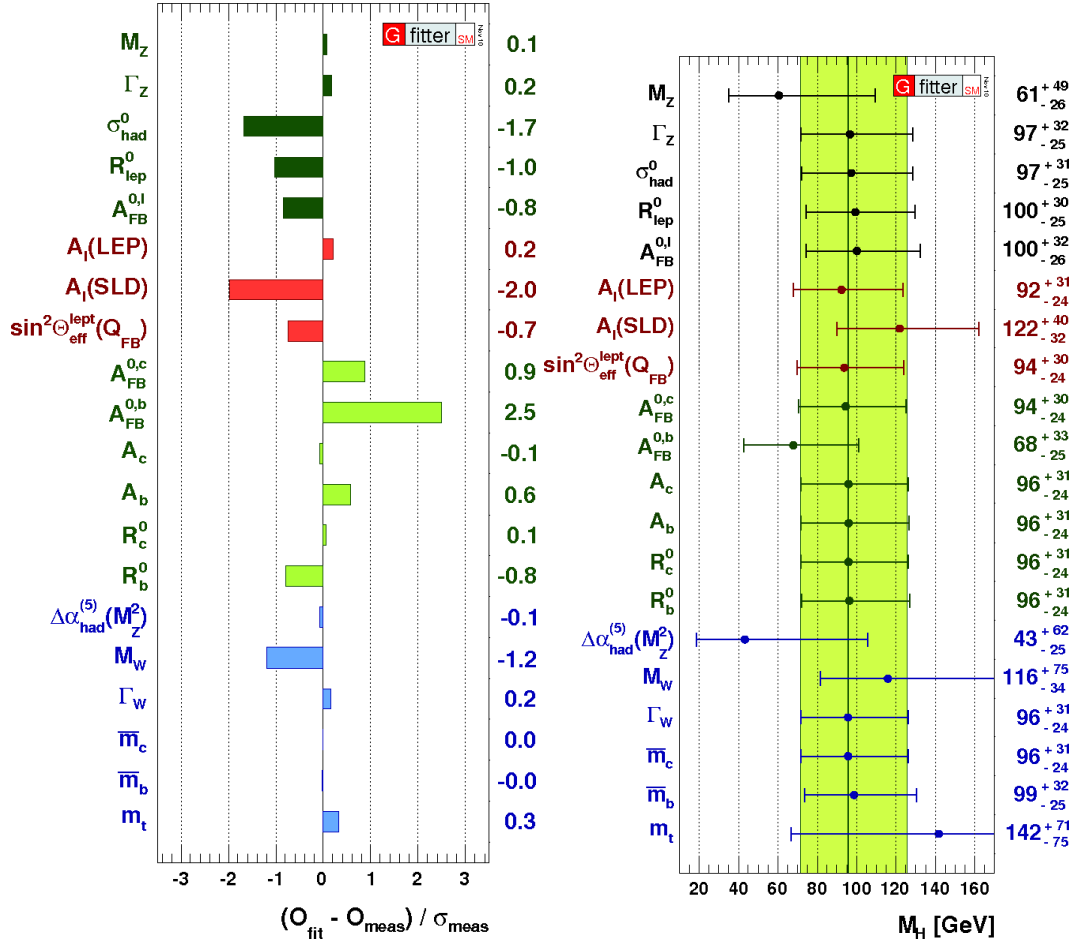


Figure 2.4: Comparison of fit results with direct measurements: pull values for the complete fit (left) and results for m_H from the standard fit excluding the respective measurements (right).

and four theoretical error parameters, accounting for theoretical uncertainties. The main experimental uncertainties come from the top quark mass, while the theoretical uncertainties are related to the prediction for $\sin^2 \theta_{\text{eff}}$. The fit is performed minimizing the statistic test χ^2 which considers the difference between measurements and SM predictions.

All the plots and fits shown in this section indicate a preference for a small Higgs mass. Indeed the most probable value (Fig. 2.5) [27] is around 100 GeV, under the LEP exclusion limit that will be described in the next section.

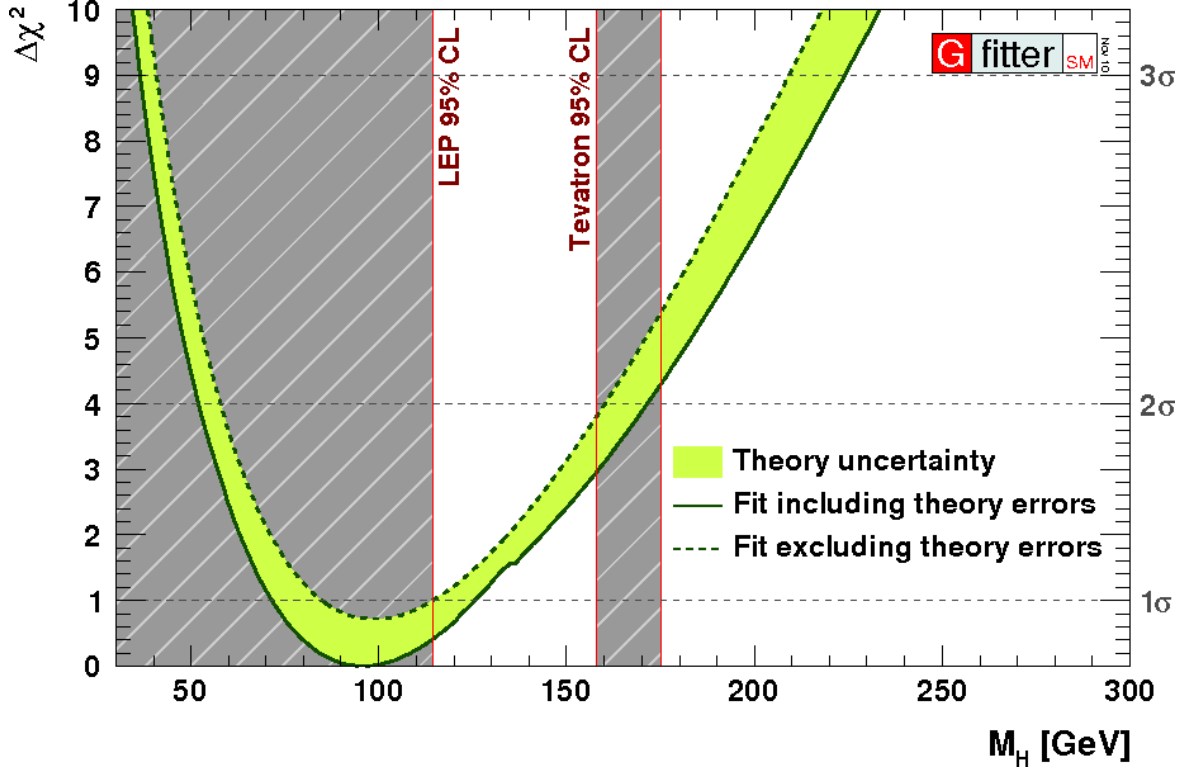


Figure 2.5: χ^2 distribution for the fit as a function of the SM Higgs boson mass.

2.3 Direct Searches

In this section we will present the latest results on the SM Higgs boson searches from LEP-II and Tevatron experiments. But first we give an overview of the Higgs boson phenomenology which makes clear how we can detect this particle.

2.3.1 Production Processes

The main production processes in a proton-proton collider are the *Gluon Fusion* (GF), the *Vector Boson Fusion* (VBF), the *associated production with a vector boson* and the *associated production with $t\bar{t}$* and they are shown in Fig. 2.6 at the tree level.

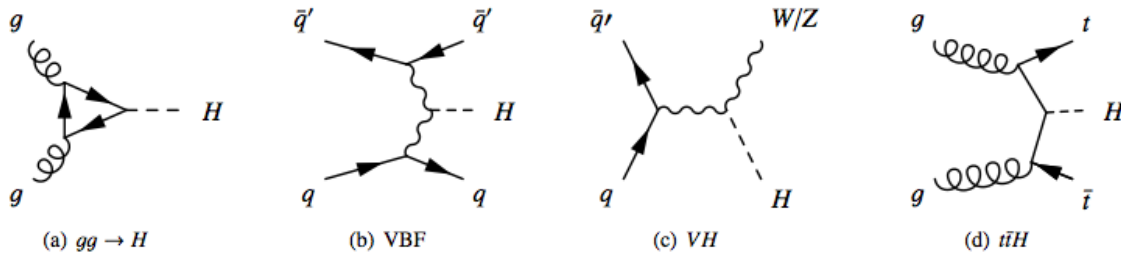


Figure 2.6: Feynman diagrams of the SM Higgs boson production modes in hadron colliders.

Fig. 2.7 [28] shows the inclusive cross sections of these processes at LHC for the center of mass energy of 14 TeV.

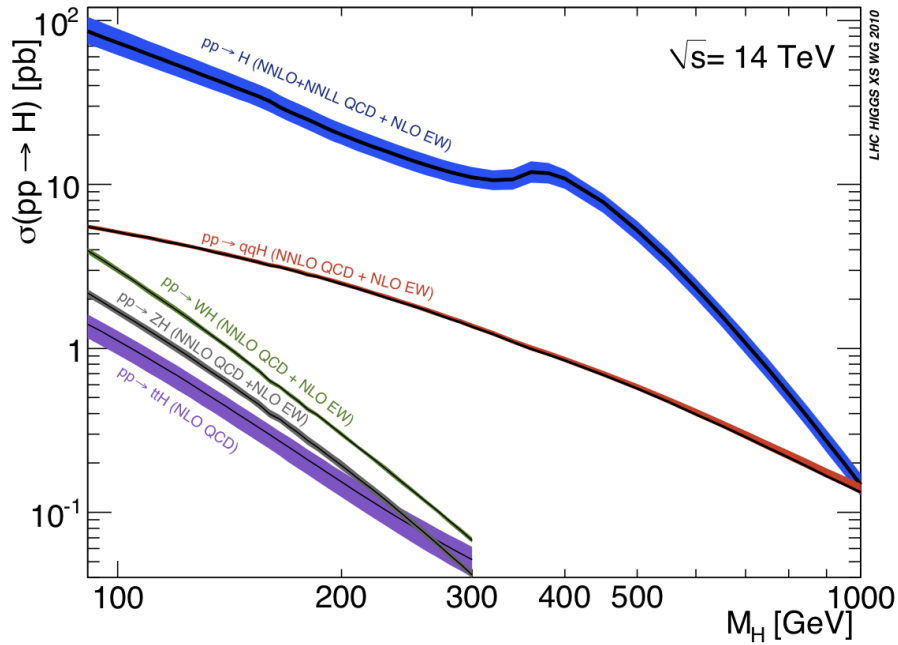


Figure 2.7: SM Higgs boson production cross sections at LHC at the center of mass energy of 14 TeV as a function of the Higgs mass.

The gluon fusion is the primary production process since the Higgs boson coupling to up and down quark is very little. Indeed, as we have seen, the Higgs boson prefers to couple to

fermions with high mass. The gluon fusion has a large cross section compared to the other processes because the production is performed by a loop of a heavy quark. The dominant contribution comes from a top quark loop and a little also from a b quark loop. This production process is strongly enhanced by QCD corrections and the inclusive cross section is slowly converging including higher order corrections. At LHC the NLO correction, i.e. the one virtual loop correction, is about 80-100% of the LO cross section [28]. Often these calculations are made in the *large- m_t* approximation, where the top quark loop is shrunk into an *effective interaction* between gluons and the Higgs boson in the limit $m_t \rightarrow \infty$. Even though this approximation is formally valid if $m_H < m_W$, the agreement with the full top mass dependence is within less than 0.5% between 100 and 300 GeV at NNLO [29]. So, calculations are performed with the exact treatment of the top and bottom mass up to the NLO and higher order corrections are made in the *large- m_t* approximation.

A very important production process is the VBF. In this reaction two quarks radiate two vector bosons which annihilate producing the Higgs boson. As we will see, this is the ‘golden’ production mode in the low mass region since the events generated by this interaction are very peculiar. Indeed, in the final state the two quarks hadronize in two forward jets while the Higgs boson decays in the central part of the detector. Moreover, since this is a pure electroweak process there are no color fields connecting the two quarks. The result is that gluons cannot be emitted in the central part of detector, but they are mostly radiated collinearly to the interacting quarks. So a *Central Jet Veto* is an effective cut that discriminates between signal and backgrounds. Fig. 2.8 shows the complete set of Feynman diagrams that enters in the computation of the LO VBF inclusive cross section. Although all the three channels

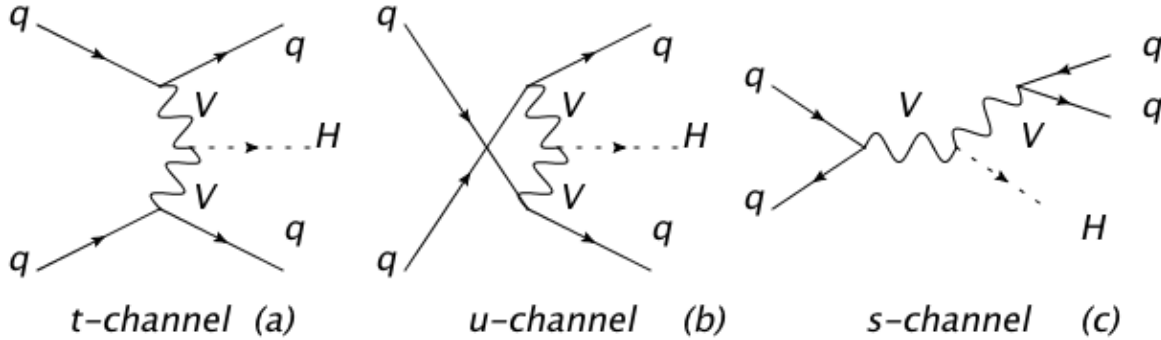


Figure 2.8: Feynman diagrams of the VBF production process.

have to be included in order to save the gauge invariance, only the first two are recognized as proper VBF events. The s-channel, indeed, doesn't have the ideal VBF topology with two forward jets. So its contribution in the VBF selection is very small.

Besides the fact that the gluon fusion and the VBF are important for the high cross sections, they are also complementary because the former is specified by the Yukawa Higgs boson couplings to fermions, while the latter is fixed by the gauge Higgs boson couplings to vector bosons. Since the Yukawa and the gauge sectors are not really connected, it will be important to explore both these production processes in order to understand the role of the Higgs boson in the SM Lagrangian.

The Higgs boson production in association with a vector boson, W or Z , is called *Higgs-strahlung*. As we will see, it is the main channel explored in the direct searches at LEP and Tevatron³, but it can be useful also at LHC since the decay products of the vector boson

³At Tevatron, a $p\bar{p}$ collider, the second main production process is not the VBF as in the pp collisions at LHC, but the Higgs-strahlung.

can be used to ‘tag’ this event. The production in association with $t\bar{t}$ pair will be relevant at LHC in the low mass range and it can provide information on the Yukawa top-Higgs coupling. However it would be pretty hard to separate such events from the $pp \rightarrow t\bar{t}b\bar{b}$ and full numerical analyses of backgrounds at NLO have not been performed yet.

In the e^+e^- collider there is basically only one feasible production mode which is the *Higgs-strahlung* $e^+e^- \rightarrow HV$ where the electron and the positron annihilate producing a virtual vector boson which becomes real emitting the Higgs boson. The direct production $e^+e^- \rightarrow H$ is not practicable since the Yukawa coupling of the Higgs boson to electrons is negligible.

2.3.2 Branching Ratios and Total Width

Besides the production modes, the Branching Ratio (BR), which are the probability of decay in specific final states, is the second main feature that has to be considered in order to find the best channels where the Higgs boson might be detected. However we cannot simply choose the final states with the highest branching ratios, since we need to think if we are actually able to recognize such signal events among backgrounds.

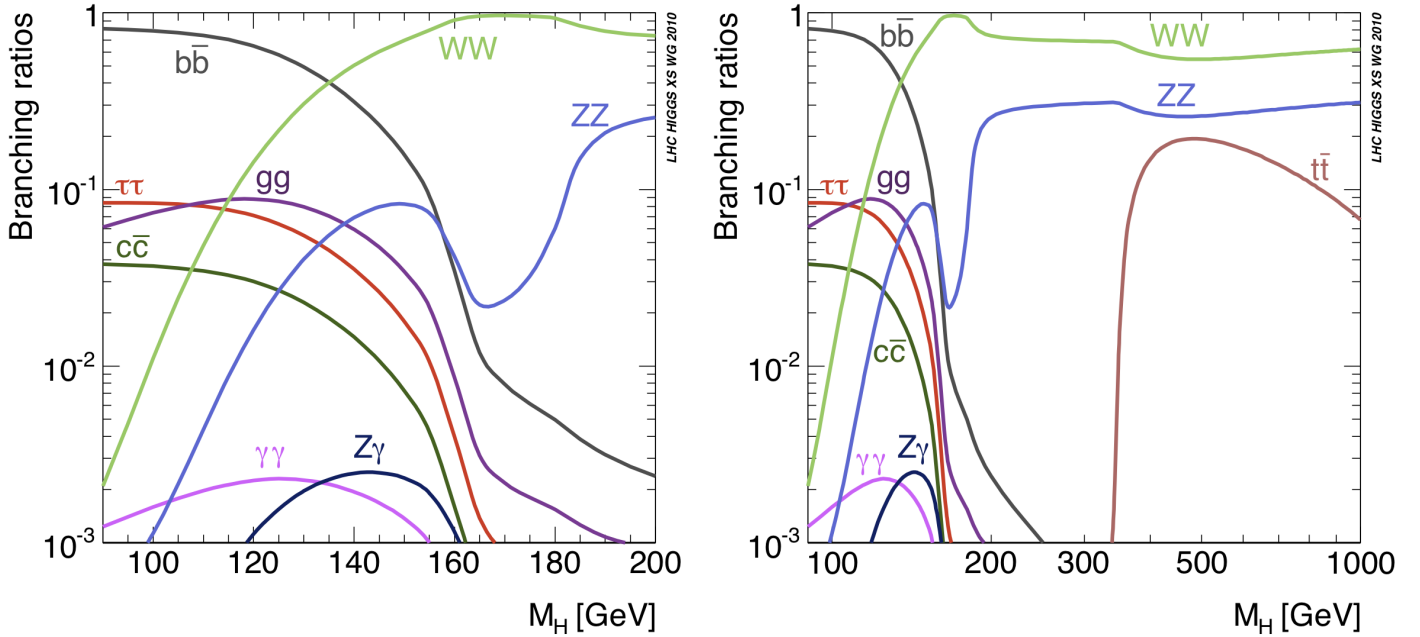


Figure 2.9: SM Higgs boson branching ratios in the low mass range (left) and for the Higgs boson mass up to 1 TeV (right).

In Fig. 2.9 [28] we can see the branching ratios for different values of the Higgs boson mass. These are directly related to the Higgs boson couplings to fermions and bosons. This is the complete list where f is a fermion and V is $W^\pm, Z$:

- $g_{Hf\bar{f}} = \frac{m_f}{v}$
- $g_{HVV} = \frac{2m_V^2}{v}$
- $g_{HHVV} = \frac{2m_V^2}{v^2}$

- $g_{HHH} = \frac{3m_H^2}{v}$
- $g_{HHHH} = \frac{3m_H^2}{v^2}$

It is possible to split the mass range from 100 GeV to 1 TeV in the low mass [100,140] GeV and the high mass region [140, 1000] GeV.

In the low mass region, the main decay mode is in $b\bar{b}$ because the b quarks are the heaviest particle in which the Higgs boson can decay. Then there is a set of final states with branching ratios one order of magnitude smaller than $b\bar{b}$ which are $\tau^+\tau^-$, gg and $c\bar{c}$. Finally the last one with a probability of per mille is $\gamma\gamma$. As already mentioned, the Higgs couplings with gluons and photons are mediated by a virtual loop of top quarks and W s respectively. Among all these final states not all of them can be used. The $b\bar{b}$ decay mode, for instance, can be used only if the QCD background is not overwhelming as at LHC. $c\bar{c}$ and gg are never considered because, firstly, if the hadronic final states can be detected they would be dominated by $b\bar{b}$, secondly, there are no proper ‘tagging’ algorithms for c quarks and gluons such as for b quarks. At LHC the only two decay modes explored for the low mass range are $\tau^+\tau^-$ and $\gamma\gamma$. Even if the probability that the Higgs boson decays in two photons is very little, the backgrounds for these events are well known and so even a small bump in the invariant mass spectrum of the two photons can be significant. In the low mass region the branching ratios for final states with a couple of vector bosons is pretty small since they have to be both virtual. But moving to higher Higgs boson masses, then the probability increases steeply. Indeed, since the Higgs couplings to W and Z bosons are proportional to the square of the mass, while the couplings to fermions are proportional only to the mass, then, as soon as the decays with at least one real vector boson are allowed by kinematics, they become dominant. Around 160 GeV, the threshold for the production of a pair of real W s, all BRs into fermions and even into ZZ drop. The decay mode in WW remains the dominant one in all the Higgs boson high mass range. Even at the ZZ kinematic threshold, the WW final state is still more probable than ZZ because the NC coupling is smaller than the CC one. Finally, above 350 GeV also the decay in $t\bar{t}$ is allowed, but its BR remains smaller than WW and ZZ ones. On the contrary of the WW and ZZ BRs which are almost constant up to $m_H = 1$ TeV, $t\bar{t}$ BR decreases significantly due to polarization effects.

It is also important to mention how the total width of the Higgs boson behaves as a function of m_H (Fig. 2.10) [28].

This observable is related to the number of allowed decay modes and the phase space available for each of them. We can see also in this plot a noticeable division in the low and high mass range. Indeed, below the WW kinematic threshold, the width is smaller than the experimental resolution and a ‘peak’ can be detected in the invariant mass of the decay products. As soon as the vector boson pair final states are accessible, then the total width becomes bigger than the experimental resolution. The resonance is no longer visible in the invariant mass plots since it is too spread. So the Higgs boson signal cannot be detected in a invariant mass fit, but with an event counting experiment. In this case the life time of the Higgs boson is so short that it would be improper to call it ‘particle’. It would be better to look at the Higgs boson as an intermediate state that enhances the cross sections or opens new final state channels in the scatterings of two particles.

Both the branching ratios and the total and partial widths can be precisely predicted from the choice of the Higgs boson mass value. So, it will be extremely important, once the Higgs boson is observed, to measure these variables in order to discriminate between the SM Higgs boson and other methods of EWBS.

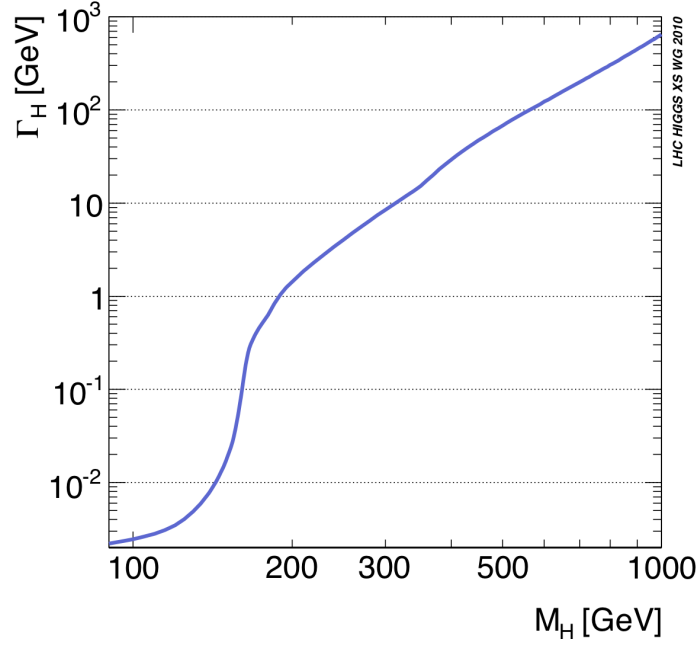


Figure 2.10: SM Higgs boson total width as a function of its mass.

2.4 Direct Searches at LEP-II

The first direct search was carried out at the *Large Electron-Positron Collider* (LEP). This was an e^+e^- collider built at CERN about 100 m underground with a circumference of 27 Km. It was the biggest accelerator ever built and it made its first collision in August 1989. It collected data at the center of mass energy of $\sqrt{s} = 90$ GeV until 1995. With these events produced at the Z peak an impressive amount of physics results were published. In the second phase of the LEP lifetime (LEP-II), from 1996 to the end of 2000, the energy increased up to 209 GeV. Reaching such a high energy was a great challenge since the design center of mass energy was 200 GeV. After the upgrade in 1996, LEP restarted at $\sqrt{s} = 189$ GeV in order to reach the threshold to produce W boson pairs. Then the machine was pushed to its limits with the hope to find the SM Higgs boson. The need to reach the highest energy was due to the production mode to which the LEP experiments were sensitive. Indeed, as we have already mentioned, since the direct production $e^+e^- \rightarrow H$ has a too little cross section, the main process is the *Higgs-strahlung* $e^+e^- \rightarrow Z^* \rightarrow ZH$. So, the Higgs boson mass range accessible was roughly $m_H < \sqrt{s} - m_Z \approx 118$ GeV. At that moment the prediction from EWPO was $m_H = 81^{+52}_{-33}$ GeV [30] and engineers tried to increase the LEP center of mass energy to raise the kinematic threshold as high as possible.

The decay modes available for the SM Higgs boson in this low mass range are $b\bar{b}$ (BR 74%), $\tau^+\tau^-$ (BR $\approx 7\%$), WW^* (BR $\approx 7\%$), gg (BR $\approx 7\%$) and $c\bar{c}$ (BR $\approx 4\%$). The final states were split in four categories:

- *four-jet*: $(H \rightarrow b\bar{b})(Z \rightarrow q\bar{q})$;
- *missing energy*: $(H \rightarrow b\bar{b})(Z \rightarrow \nu\bar{\nu})$;
- *leptons*: $(H \rightarrow b\bar{b})(Z \rightarrow l\bar{l})$ where $l = e, \mu$;
- *tau leptons*: $(H \rightarrow b\bar{b})(Z \rightarrow \tau^+\tau^-)$ and $(H \rightarrow \tau^+\tau^-)(Z \rightarrow b\bar{b})$.

Not all possible combinations of decay products have been explored in order to reduce backgrounds from $Z\gamma$, fermion pairs, WW and ZZ . The $H \rightarrow b\bar{b}$ decay mode was a useful channel since at LEP it was not overwhelmed by the huge level background of QCD dijets present in hadron machines like Tevatron or LHC.

In 2003 the combination of the results from all the four experiments installed along the LEP ring, ALEPH [31], DELPHI [32], L3 [33] and OPAL [34], was published [30]. The total integrated luminosity, which is proportional to amount of data collected⁴, was 2461 pb^{-1} and the likelihood test for the consistency of data with the hypotheses of background or signal-plus-background was performed. Only ALEPH observed an excess beyond 95% CL that could be compatible with the signal of a SM Higgs boson with $m_H = 115 \text{ GeV}$. However this observation didn't occur in the other experiments and DELPHI even reported a deficit for the same signal mass. So the combined sensitivity was not enough significant. The four experiments could only set a lower bound at $m_H > 114.4 \text{ GeV}$ at the 95% confidence level. Fig. 2.11 shows the ratio $\text{CL}_s = \text{CL}_{s+b}/\text{CL}_b$ which is used to set the lower bound.

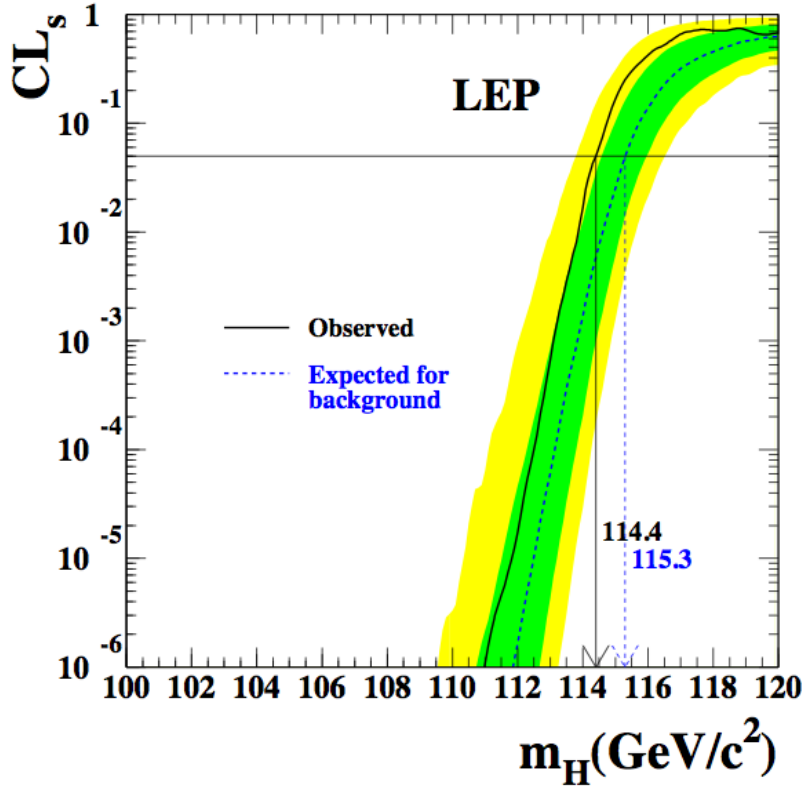


Figure 2.11: Ratio $\text{CL}_s = \text{CL}_{s+b}/\text{CL}_b$ for the signal-plus-background hypothesis as a function of the m_H . The solid line is the observed confidence level and the dashed line comes from the median background estimation. The intersection of the horizontal line for $\text{CL}_s = 0.05$ with the observed curved is used to define the 95% confidence level lower bound on the mass of the SM Higgs boson.

The difference between the expected and the observed limit is due to the excess observed by ALEPH. This deviation, which has a 3σ significance if only ALEPH data is considered, is mainly produced by three candidates selected in the four-jet analysis at center of mass energies greater than 206 GeV [31].

⁴For a proper definition see Section 3.1.

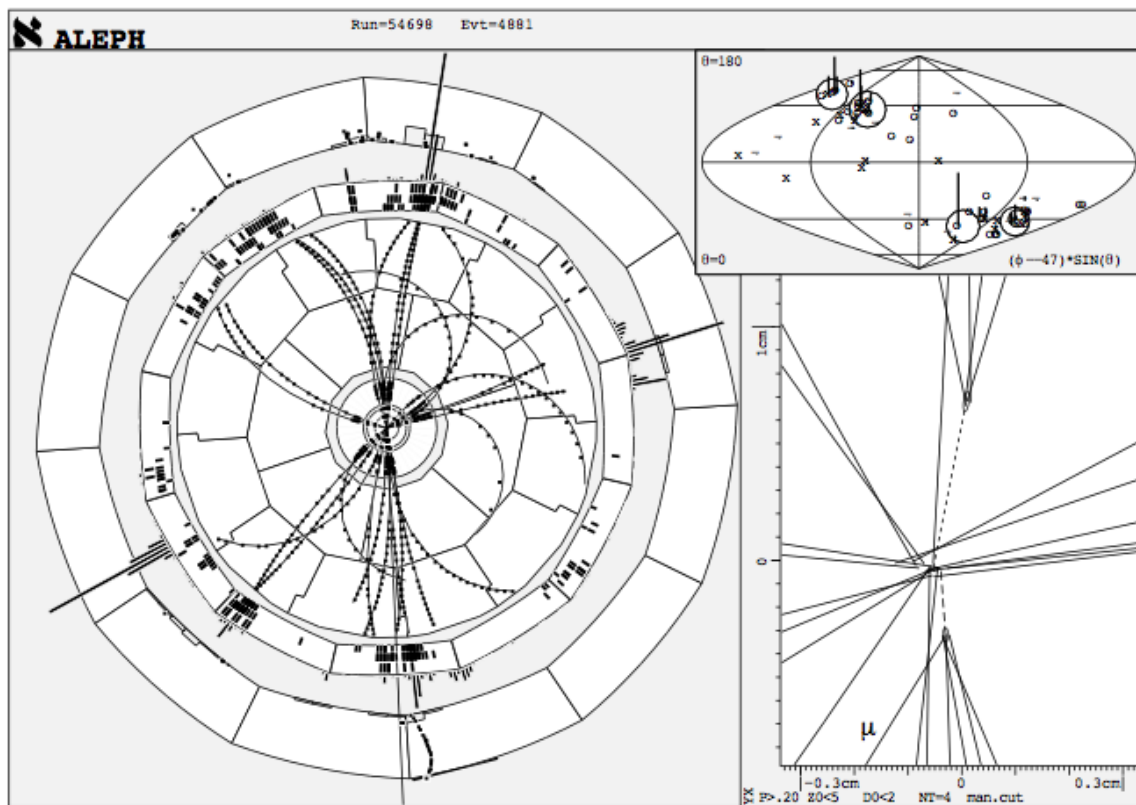


Figure 2.12: Four-jet Higgs boson candidate with a reconstructed mass of 114.3 GeV detected in the ALEPH experiment.

Fig. 2.12 displays one of these three events. The Higgs boson candidate has the reconstructed mass of 114.3 GeV and it decays in two well tagged b jets. The invariant mass of the two other not b tagged jets is 92.1 GeV. The difference between these two values makes the hypothesis of a ZZ event unlikely. The background most compatible with this events is $b\bar{b}gg$ even if the energies of the non b tagged jets at 43.5 and 49.0 GeV are typical of the Z boson decay.

2.5 Direct Searches at Tevatron

Another accelerator where direct searches are ongoing is Tevatron at Fermi National Accelerator Laboratory. It is a proton-antiproton collider which is running at $\sqrt{s} = 1.96$ TeV. It started smashing particles in the 1983 and it will be shut down in September 2011. Several analyses are carried on in order to span all the mass range even if the best sensitivity necessary for setting new constraints is reached only around 160 GeV. The analyses are performed with *multivariate methods* like Neural Networks, Matrix Element probabilities and Boosted Decision Trees. On the contrary of cut-based analyses, these take advantage of the correlations between different variables.

In the low mass range, the production mode considered is the *Higgs-strahlung* and the Higgs decay mode is in $b\bar{b}$. The gluon fusion production process cannot be considered since the multijet background is too large. In the associated production, instead, the detection of the vector boson helps suppressing backgrounds. The final states analyzed for the vector bosons are: $W \rightarrow l^\pm \nu$, $Z \rightarrow \nu\nu$ and $Z \rightarrow l^- l^+$. The missing energy and the leptons coming from the W and Z bosons are used to tag the events and a resonance is expected to appear in the invariant mass of the two b jets. The b tagging algorithms are effective in suppressing backgrounds from $W + j$ and $Z + j$, as Fig. 2.13 shows [35]. Nevertheless the selection

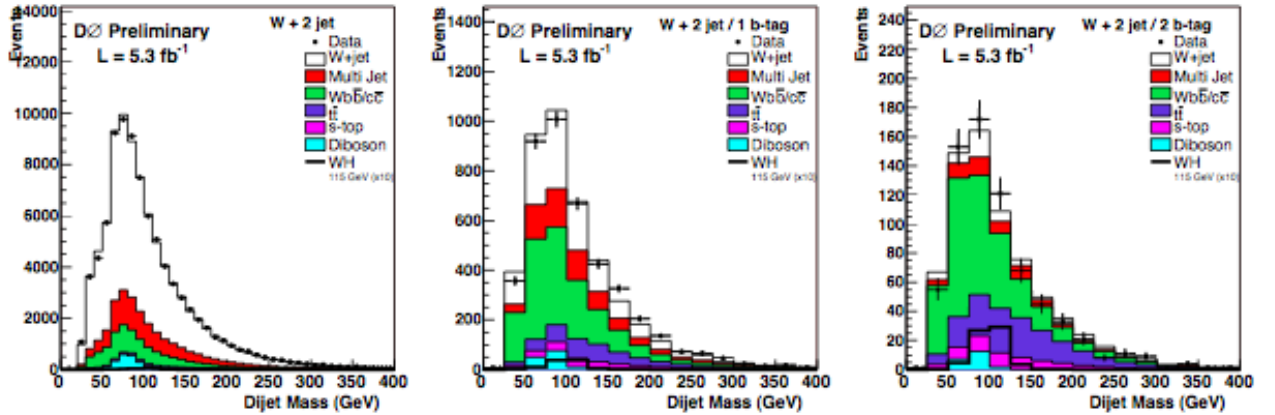


Figure 2.13: Dijets invariant mass of $WH \rightarrow l\nu b\bar{b}$ with no b tags, one b tag and two b tags (from left to right).

with no b tags is an important *control region* to test the background estimations and the kinematic modeling of the relevant variables (Fig. 2.13, left). In the high mass region, studies are focused on the $gg \rightarrow H \rightarrow WW \rightarrow l^\pm \nu l^\mp \nu$ channel ($l = e, \mu, \tau$) [36]. Since in this mass range the dominant decay mode is in a pair of W , it is possible to use the main production process which is the gluon fusion. The high production cross section and the large branching ratio makes this analysis the most sensitive to the SM Higgs boson in the mass range $m_H > 140$ GeV. Analyses are split in several sub-channels according to leptonic

flavors or jet multiplicities, but all the signatures of high missing transverse momentum, E_T^{miss} , and two oppositely charged leptons are considered. The angle between the two W 's provides a good separation against the $Z \rightarrow WW$ background thanks to the difference in the spins of H and Z . Since the full reconstructed mass cannot be used to detect the signal because of the presence of two neutrinos, it is necessary to estimate the mass from the transverse invariant mass. Recently new channels have been explored like $H \rightarrow WW \rightarrow l\nu jj$ and $H + W/Z \rightarrow WW + W/Z \rightarrow l^\pm l^\mp(l) + X$. The latter channel is based on the associated production mode.

The sensitivity reached combining the results from all the analyses carried out by the two experiments at Tevatron, D0 and CDF, enabled to set new limits on the SM Higgs boson mass in 2010. Analyzing an integrated luminosity of 5.9 (CDF) and 5.4-6.7 (D0) fb^{-1} at $\sqrt{s} = 1.96$ TeV, the mass range excluded with 95% CL is 158-175 GeV[36].

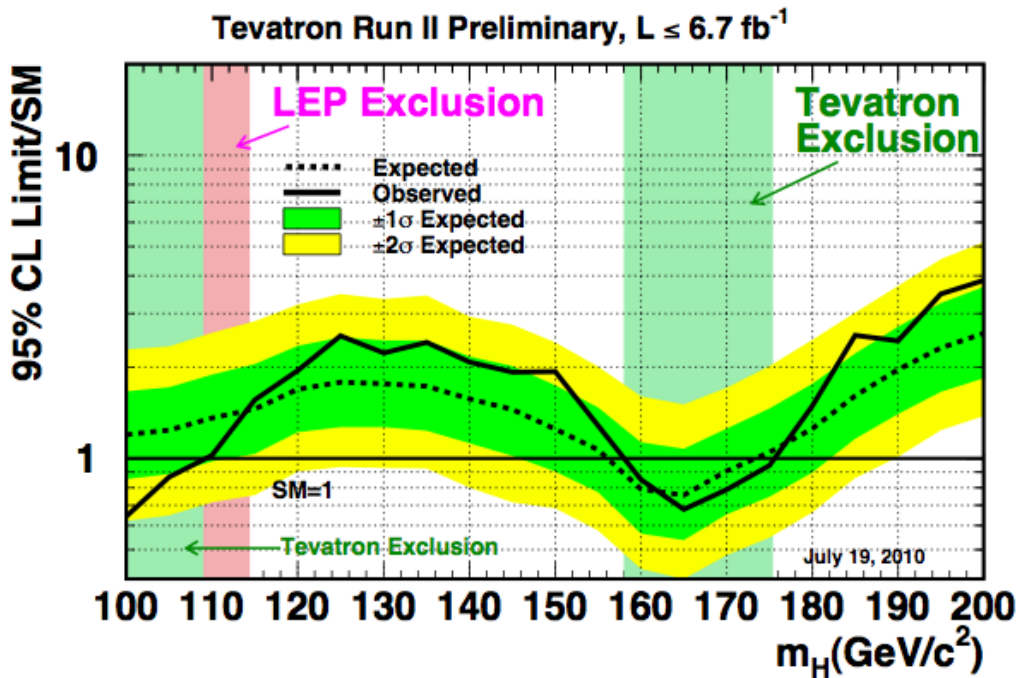


Figure 2.14: Observed and expected 95% CL upper limits on the ratio to the SM cross section as a function of the SM Higgs boson mass. A value of the limit ratio less than one indicates that that mass region is excluded at the 95% CL. The bands indicate the 68% and 95% probability regions where the limits can fluctuate in the absence of signals.

Fig. 2.14 shows the combined 95% exclusion limits. It is clear that despite the high sensitivity reached in the high mass region, the low mass range is most troublesome for the hadron collider. D0 and CDF haven't reached yet the LEP lower bound and they can exclude at the 95% CL a signal with 1.56 times the SM cross section at $m_H = 115$ GeV.

Both experiments will collect data until September 2011 and before the winter shut down in 2010 the total integrated luminosity analyzed was less than 7 fb^{-1} . At the end of 2011, when about 10 fb^{-1} are expected to be collected, the estimated sensitivity at $m_H = 115$ GeV should be 3σ and at least 2.4σ in $100 < m_H < 185$ GeV [35]. This is not the necessary for a discovery, which usually requires 5σ , but enough for a 'strong' observation at least below 120 GeV.

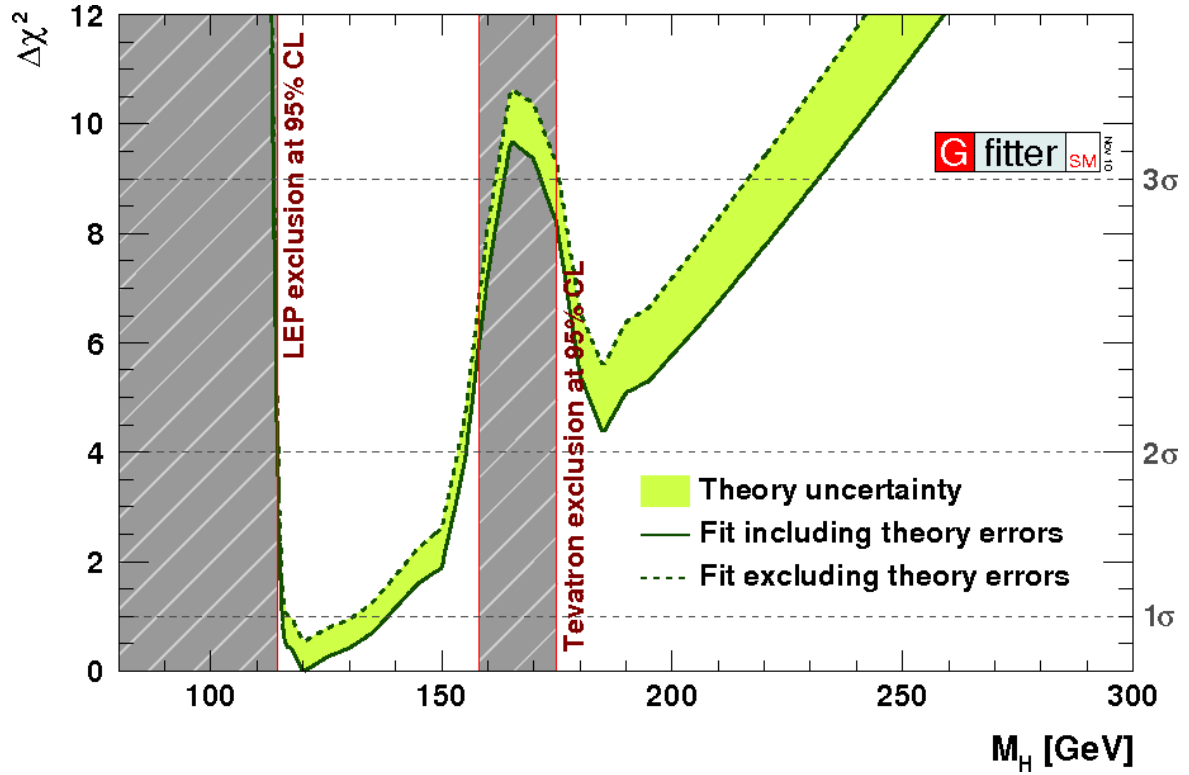


Figure 2.15: $\Delta\chi^2$ as a function of the SM Higgs boson mass for the fit described in section 2.2 including also the LEP and Tevatron exclusion limits.

If we consider both the direct and indirect searches carried on so far, we find out that the most probable value for the SM Higgs boson is around 120 GeV, as Fig. 2.15 shows [27].

Chapter 3

LHC and the ATLAS Experiment

3.1 The Large Hadron Collider

Direct searches for the SM Higgs boson are also performed by the experiments located around the *Large Hadron Collider* (LHC) [37] at CERN. LHC is the accelerator that replaced LEP in the 27 Km ring which crosses the French-Swiss border 100 m underground. This is a proton-proton collider which should reach the center of mass energy of 14 TeV and the instantaneous luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ at design performance.

The instantaneous luminosity L is one of the crucial parameters of High Energy Physics colliders and is defined by $\frac{dN}{dt} = \sigma L$ where dN/dt is the number of events which are observed per unit of time and σ is the cross section, i.e. the probability that each single event occurs. The integrated luminosity is the integral in time of the instantaneous luminosity and it is measured in *barns*, $1\text{b} = 10^{-24} \text{ cm}^{-2}$. The relation between the instantaneous luminosity and the amount of data delivered by the accelerator is given by

$$L = fn \frac{N_1 N_2}{A}$$

where f is the revolution frequency and n is the number of bunches in each beam, N_i is the number of particles in each bunch and A is the cross section of the beams, i.e. the ‘size’ of the beams in the transverse plane.

One of the main reasons for the construction of LHC is to explore a new energy scale, to understand the electro-weak spontaneous symmetry breaking and to discover the physics beyond the Standard Model. This means that both the beam energy and the luminosity must be pushed to their technological limits. The choice of proton beams was the only possibility available. In fact electrons and positrons cannot be accelerated so hard because most of energy would be lost in radiation. Antiprotons cannot be used as well because it is not possible to produce an antiproton beam with high luminosity.

In order to reach such a high energy protons have to pass through a complex system of accelerators where they are accelerated and stored, as Fig. 3.1 illustrates. The injection in the LHC ring is at 450 GeV and then the energy rumps up to 7 TeV per beam. Protons circulate in vacuum chambers in the middle of superconducting magnets which provide the strong magnetic field needed to bend these particles. To operate these 9300 magnets installed along the ring, 10 thousands tonnes of liquid nitrogen and 60 tonnes of superfluid helium are required to bring the temperature down to -271.3°C (1.9 K).

LHC is a record machine. Besides being the biggest machine ever built, it is one of the emptiest places in the Solar System because of the vacuum which protons pass through. Moreover it is at the same time the hottest and coldest place in the galaxy thanks to the high energy reached in the collisions (100 000 times the temperature of the heart of the Sun)

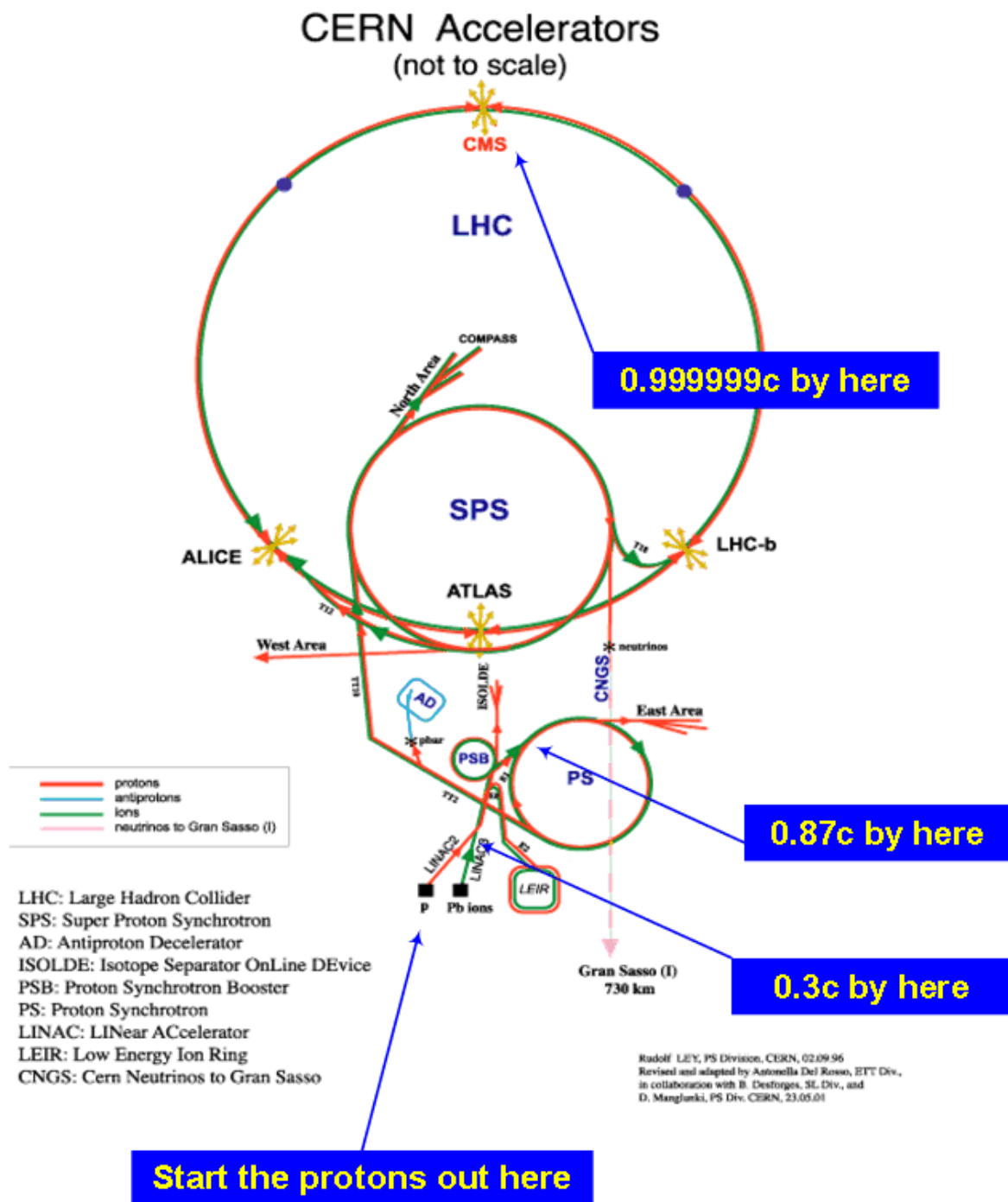


Figure 3.1: LHC system of accelerators. Protons receive the first acceleration in the LINAC and then they circulate in the other rings increasing energy and being collected in bunches. The injection in LHC is at 450 GeV.

and the cryogenic system that makes the magnets even colder than the outer space. If this is not enough, we should also mention that data recorded in LHC collisions are handled by the largest distributed computing network in the world, the Grid.

After a false start of LHC on the 10th of September 2008, which was interrupted by an incident ten days later, particles have been injected again in the LHC ring at the end of October 2009. Since that time the machine has performed surprisingly well and already after a month, on the 30th of November 2009, LHC became the world's highest energy particle accelerator hitting 1.18 TeV per beam. On the 30th of March 2010 the center of mass energy of 7 TeV was reached. Few days later the instantaneous luminosity raised to $10^{28} \text{ cm}^{-2} \text{ s}^{-1}$ and the run length was extended to 30 hours. Then big efforts were made in raising the luminosity by increasing the number of bunches in the beam and the number of protons in each bunch. The target luminosity for 2010 at $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ was achieved in the middle of October. In this regime physicists have already started facing one of the challenges that high luminosity brings into data analysis. Indeed, with such a high instantaneous luminosity the average number of interactions per bunch crossing was 4 (Fig.3.2).

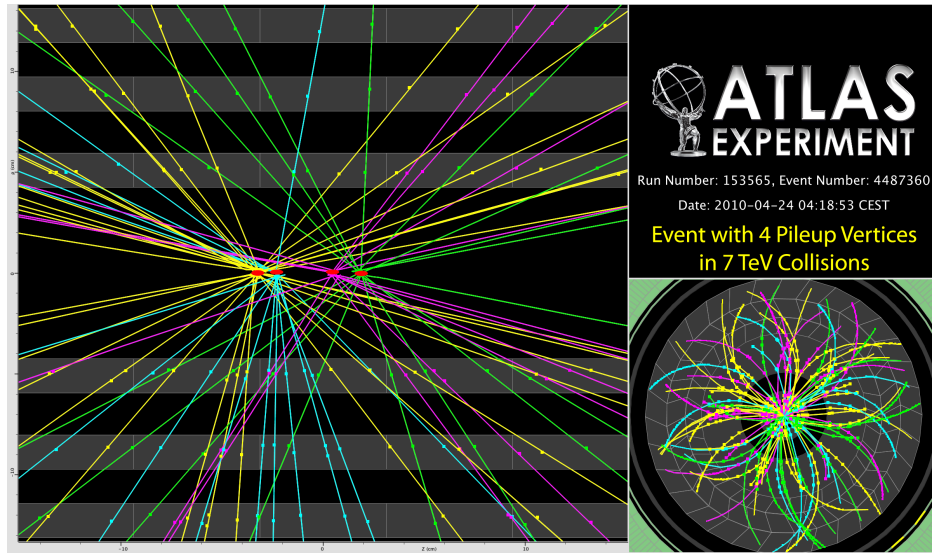


Figure 3.2: ATLAS event display of a proton-proton collision with 4 primary reconstructed vertices [38].

This means that on top of the events we are interested in, there are on average three minimum bias events. This sort of event is produced by a pp scattering at a large impact parameter or small energy transfer. The final state is therefore characterized by low-pt objects, mainly jets from light quark hadronizations. Such events are the most common in the pp collisions, but they don't contain interesting physics, which is rather produced in high momentum transfer scatterings. The presence of multiple minimum bias events produced in the same bunch crossing is called *in-time pile up*. At the designed luminosity more than 20 interactions are expected to occur in each bunch crossing. It will be crucial to develop adequate reconstruction algorithms since these overlapping events can affect the energy resolution, especially for the transverse missing energy, E_T^{miss} , and the identification of particles like taus and b quarks. In addition to the *in-time*, there is also the possibility of the *out-of-time* pile up, which is produced by events in successive bunch crossings. In 2010 data this is not relevant since bunches are separated by 150 ns, but it will become a significant source of pile up at the design bunch spacing of 25 ns.

The 2010 run was terminated on the 4th of November when the Heavy Ion Programme

started. Indeed, LHC is able to accelerate not only protons but also lead nuclei for specific physics studies.

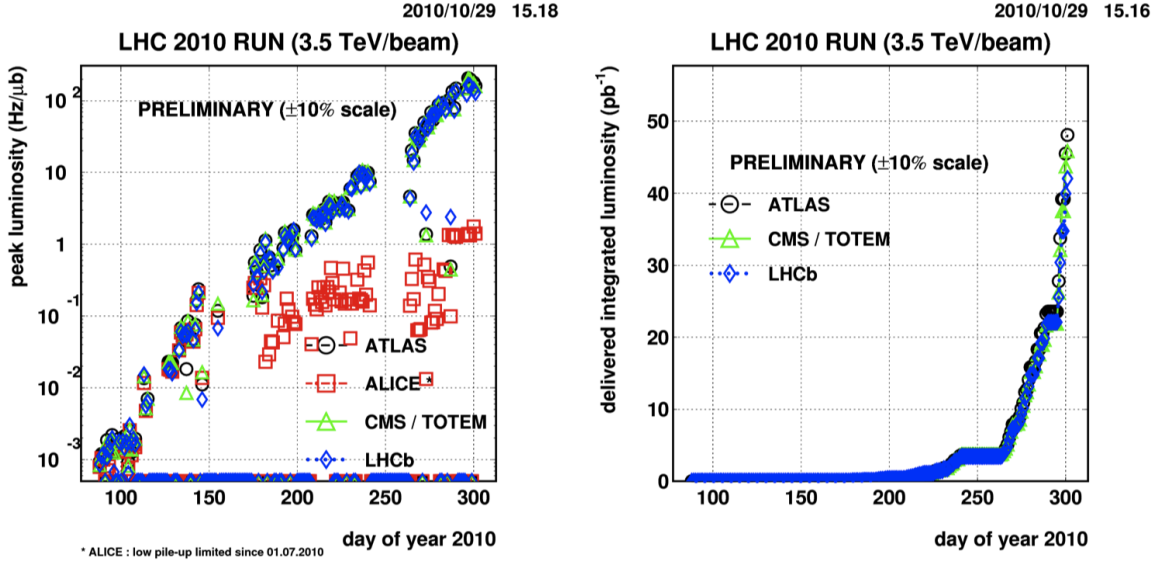


Figure 3.3: LHC instantaneous luminosity (left) and integrated luminosity (right) delivered in the four main LHC experiments in the 7 TeV proton run in 2010 [39].

In the 7 TeV proton run, LHC delivered roughly 50pb^{-1} (Fig. 3.3) and the Standard Model Physics has been rediscovered. Indeed, even the top quark, that have been observed so far only at Tevatron, was detected also at LHC. As Fig. 3.4 shows, the next step will be the exploration of new regimes where new particles like the SM Higgs boson or even new physics beyond the SM occurs.

The impressive achievements of LHC in this short time are reflected, for instance, by the number of $t\bar{t}$ events collected so far by ATLAS, one of the LHC experiments, which is only 8 times less than what CDF or D0 recorded so far in 17 years of operation. Moreover the exploration of the TeV scale has already started with the observation of a dijet event with an invariant mass of 3.7 TeV (ATLAS)[38].

At the beginning of 2011 the future operation plans will be decided. At the end of 2011 LHC might be shut down for one year for the upgrade needed to reach the 14 TeV and the luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. It is necessary to choose if LHC will operate until that time at the same energy as in 2010 with the instantaneous luminosity at $10^{33}\text{cm}^{-2}\text{s}^{-1}$ or at 8 TeV. Moreover the number of bunches in each beam has to be decided.

These are critical decisions since the competition with Tevatron on important discoveries such as the SM Higgs boson is tough. At the moment the baseline is to run at 7 TeV until 1fb^{-1} is collected, but if the center of mass is pushed up to 8 TeV then the same sensitivity can be reached with about 20% less integrated luminosity [41]. If then each beam is filled with 450 bunches, at the end of 2011, 3.66fb^{-1} might be delivered. Doubling the number of bunches to 900 the integrated luminosity could be 5.49fb^{-1} . The sensitivity for the SM Higgs boson discovery at $m_H = 120\text{GeV}$ would be above 3σ in the first case and 5σ in the second. The mass range most difficult for LHC which is at $m_H = 115\text{GeV}$ can be explored with 5σ only with at least 7fb^{-1} at 8 TeV considering only data from ATLAS.

As we mentioned in the previous chapter, at the end of 2011 Tevatron is expected to reach the 3σ sensitivity for discovery at $m_H = 115\text{GeV}$ and at least 2.4σ in the whole range.

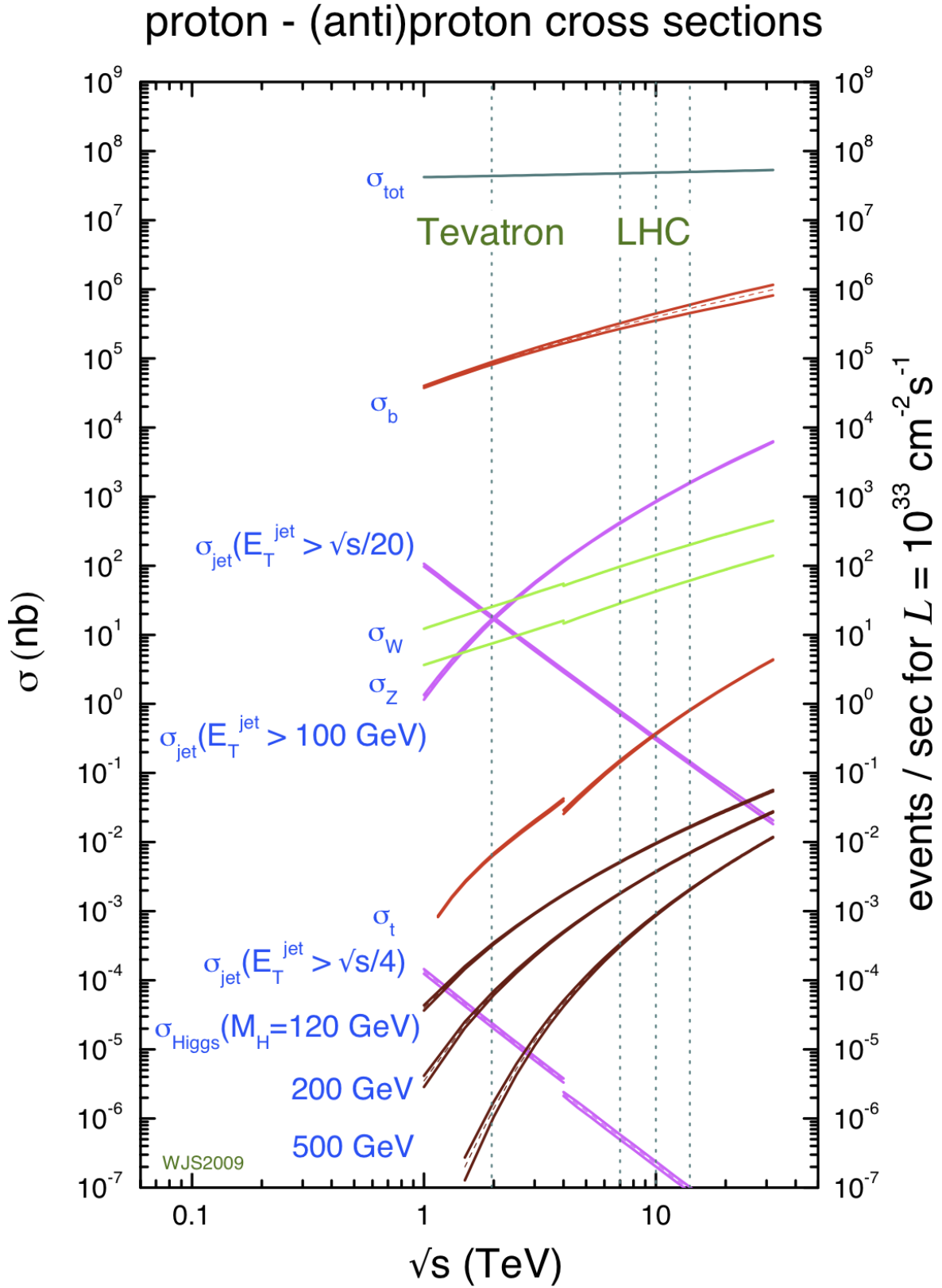


Figure 3.4: Proton-proton cross sections for SM processes as a function of the center of mass energy. The dashed lines correspond to the Tevatron (1.96 TeV) and LHC (7 and 14 TeV) collision energies [40].

The integrated luminosity needed at LHC to reach similar results is 2.5 fb^{-1} at 8 TeV.

The LHC management will have to take the tough decision to try to push the accelerator a bit harder before the long shut down in order to compete with Tevatron on the discovery which has the highest media coverage and which is probably the last one achievable by the american accelerator.

3.2 The ATLAS Experiment

The *A Toroidal LHC ApparatuS* (ATLAS) [42] is one of the four main experiments installed along LHC ring. It is a multi purpose detector as the *Compact Muon Solenoid* (CMS). These are the two biggest experiments in the size both of the detector and of the collaboration at LHC and they will crosscheck one another in probing the SM and exploring the new physics that might be revealed in the TeV energy scale. The other two experiments are ALICE, which is specialized in heavy ion physics, and LHCb, which is, on the contrary of the others, a fixed target experiment and it will focus on b-flavor physics.

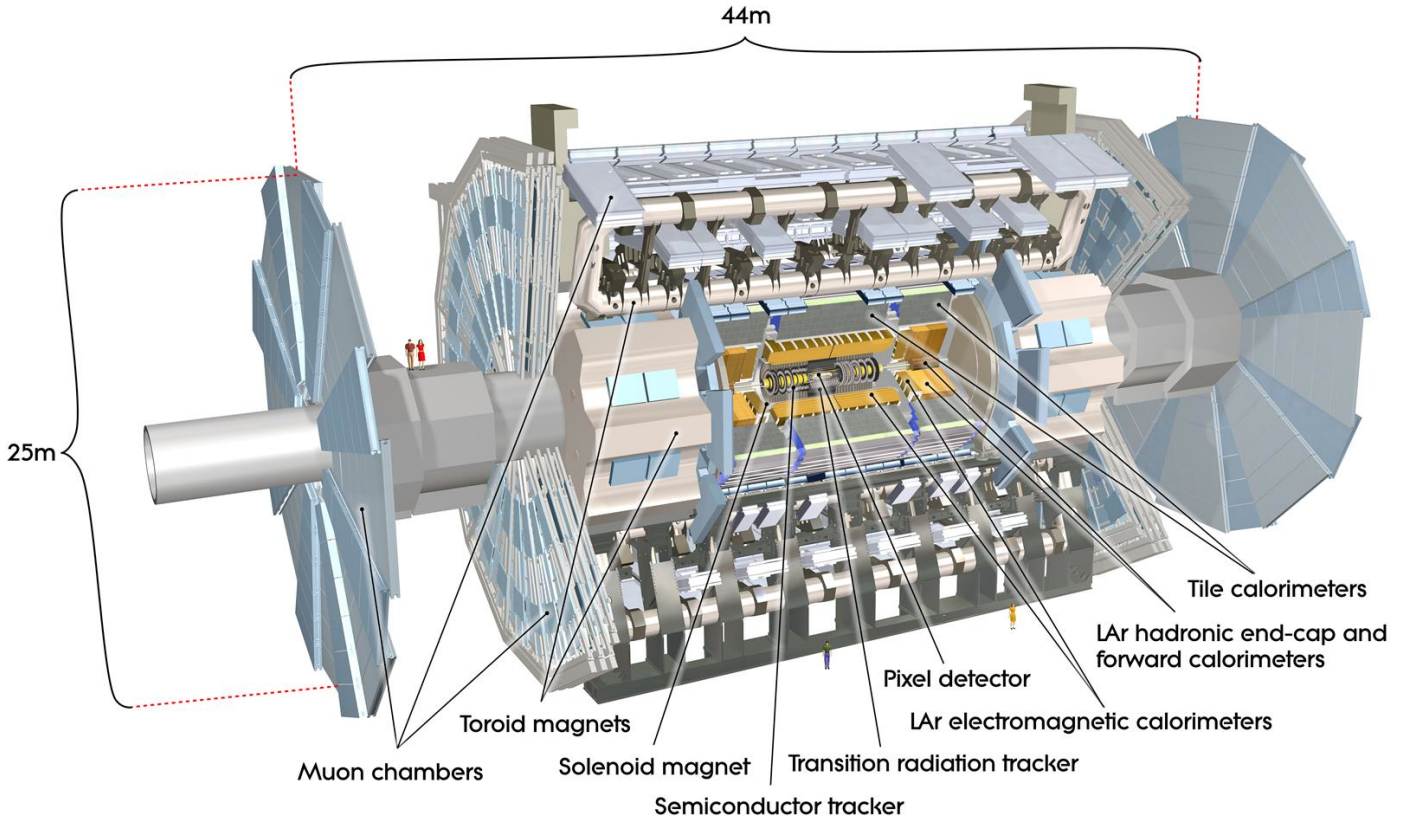


Figure 3.5: Schematic view of the ATLAS detector and its components [38].

ATLAS [43] has the standard structure of the modern high energy physics detector, as Fig. 3.5 shows. It was designed to be able to reconstruct and identify effectively all the final states that can be a hint of new physics in the proton-proton collisions in the LHC regime of energy and luminosity. These signatures are mainly missing energy, secondary vertices and high-pt leptons and hard jets. One of the biggest challenges is the reconstruction of the interesting physics process in a very ‘dirty’ environment dominated by minimum-bias

events. Indeed, with the designed bunch crossing of 25 ns it is estimated that interactions will occur at the rate of 10^9 Hz and the number of pile up events will be 23 on average. This is the only way to access very rare physics, but it represents a technological challenge since a perfect resolution and timing have to be achieved by the instrumentation. Moreover, the huge detector output cannot be stored on tape, so an effective trigger has to be implemented in order to reduce the rate of recorded events down to about 200 Hz. The basic features necessary to ATLAS to be a *discovery* experiment are:

- Fast and hard-radiation sensors and front-end electronics;
- High granularity and high acceptance;
- Good momentum resolution and high reconstruction efficiency in the tracking system;
- Good identification of secondary vertices;
- Accurate electromagnetic and hadronic calorimeters for electron, photon, jet, tau and E_T^{miss} identification and energy measurements;
- Good muon momentum resolution on a wide range of transverse momentum.

ATLAS and CMS share the same physics goals, but they have different geometries and detectors [44]. Indeed, both experiments were designed to achieve the optimal performance

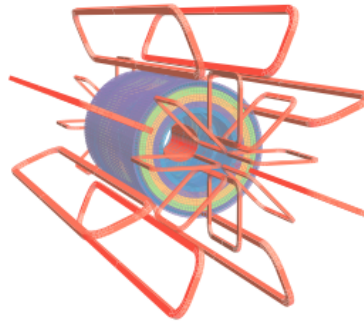


Figure 3.6: Illustration of ATLAS magnetic field system [42].

needed to detect the process $H \rightarrow ZZ \rightarrow 4\mu$, which is considered the benchmark process for the discovery of the Higgs boson. The key feature is therefore the momentum resolution of muons and this is proportional to:

$$\frac{\Delta p}{p} \sim \frac{1}{BL^2} \quad (3.1)$$

where B is the average magnetic field and L is the arm length used to measure the sagitta. A good resolution can be achieved with a *intense* and *compact* magnetic field or with a *large* volume immersed in a moderate magnetic field. CMS geometry is designed for the first option, while ATLAS for the second. Indeed, the inner part of the ATLAS detector is contained in a 2 T solenoidal magnetic field and there is a large volume in the outer part of the detector with a 1 T toroidal magnetic field. This structure (Fig. 3.6) defines the geometry and the size of the ATLAS detector, as Fig. 3.5 shows.

The 3-D coordinate system has z along the beam tunnel and y pointing upwards and x orthogonal to these two and pointing towards the center of the LHC circle. Paths of particles crossing the ATLAS detector are usually given in the polar angle coordinates, where the

azimuthal angle ϕ goes from $-\pi$ to π , with $\phi = 0$ at the positive direction of x axis, and the polar angle θ goes from 0 to π , with $\theta = 0$ along the positive direction of the z axis. The pseudo rapidity η is defined as [42]

$$\eta = -\ln \tan(\theta/2) \quad (3.2)$$

The core of ATLAS is the *Inner Tracking System* or *Inner Detector* (ID). It is composed by pixel and silicon microstrip (SCT) trackers close to the interaction point and the Transition Radiation Tracker (TRT). The first two detectors cover the region $|\eta| < 2.5$ and they are designed for pattern recognition, excellent momentum resolution for tracks with $p_T > 0.5$ GeV and vertexing. The pixel tracker has the highest granularity with more than 80 millions read-out channels placed on three layers in order to resolve primary and secondary vertices. The SCT, instead, uses strips and not pixels and has a ‘coarser’ granularity with 6 millions read-out channels. Each charged track is defined by four space points. The information from the pixel detector mainly is important for secondary vertexing and consequentially for b-jet tagging and tau lepton identification. The TRT is placed around the SCT covering the region in $|\eta| < 2.0$. It consists of gaseous straw tubes immersed in the transition radiation material. Each track leaves on average 36 hits on a long arm, so this detector improves significantly the momentum resolution of the ID. Moreover, thanks to the transition radiation photons, it can identify electrons on a wide range of energies.

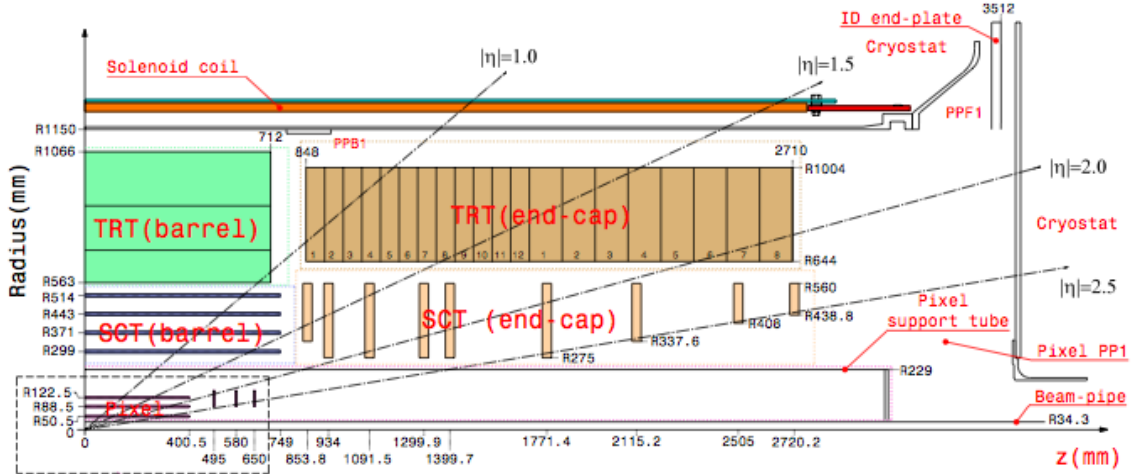


Figure 3.7: Section of the Inner Detector [42].

The ATLAS electromagnetic calorimeter is based on sampling detectors with liquid argon (LAr) as active medium and lead as passive medium. The ATLAS hadronic calorimeter is a scintillator tile calorimeter in the central region and a continuation of the electromagnetic LAr calorimeter in the forward regions. Both electromagnetic and hadronic calorimeters cover the region $|\eta| < 4.9$, but the precision electromagnetic calorimeters with high granularity are only in $|\eta| < 3.2$. This is needed to have good measurements in energy and position for electrons or tau leptons in the central high- p_T range matching the ID precision. The hadronic calorimeters have a poorer granularity which is enough for jet reconstruction and E_T^{miss} measurement. The total thickness of the calorimetry system at $|\eta| = 0$ is 11 radiation lengths which are adequate for avoiding the punch-through into the muon system and providing good E_T^{miss} resolution. The calorimetry is one of the biggest differences between ATLAS and CMS. In the former experiment, calorimeters are placed outside the solenoid, while in the latter they are inside and therefore they have limitations in size. That is why

ATLAS calorimeters are larger and not compact like the CMS electromagnetic lead-tungstate homogeneous calorimeter.

Another important component of ATLAS is the *Muon Spectrometer*. It provides momentum and charge measurements of particles exiting the calorimeters in $|\eta| < 2.7$. It is composed of three stations of Monitored Drift Tubes (MDT) or Cathode-Strip Chambers (CSC) which can measure the sagitta of charged tracks, mainly muons, and consequently the momentum. The multiple scattering which is detrimental for this measurement is minimized by the choice of the toroidal air-core magnet. In addition to these detectors there are also Resistive Plate Chambers (RPC) or Thin Gap Chambers (TGC) which are used for triggering on muons thanks to the excellent time resolution.

Fig. 3.8 illustrates a nice event display of a $t\bar{t}$ candidate and shows all the main ATLAS components and their responses.

The ATLAS detector system is completed by three *Forward Detectors*. LUCID is a Cherenkov detector which monitors online the relative luminosity through the detection of inelastic proton-proton scatterings in the forward region. Zero-Degree Calorimeter is designed to detect forward neutrons in heavy-ion collisions which are correlated to the centrality of the interaction. Then, at the end of 2010 the ALPHA detector has been placed 240 m far from the interaction point and it will measure the absolute luminosity detecting pp elastic scatterings with scintillating-fibre trackers.

The *Trigger* and *Data Acquisition* (DAQ) systems are important elements of modern particle detectors, but this is exceptionally true for LHC experiments. As we have already mentioned, due to the overwhelming high event rate at 1 GHz, an excellent trigger performance is needed to reduce this rate at least of 6 order of magnitudes and then an outstanding data acquisition system has to handle such a data stream. The trigger system has been split in three levels. At the first level L1, distinctive signatures such as high- p_T muons, electrons, photons, jets, tau-jets, E_T^{miss} and high total transverse energy are searched. The informations used for L1 decision are reduced-granularity data from subsets of the calorimeters and the muon spectrometer. The L1 trigger find Regions-of-Interest (RoIs), that are regions of the detector which survive selection criteria based typically on the amount of transverse energy deposited in these regions. Isolation criteria can be applied too, based on the energy collected around a core region of interest. The L1 RoIs are then inputs for the L2 trigger which uses the full granularity detector data in the RoI. The final stage is the event filter (EF) whose selection is based on data from the whole detector and uses algorithms close to the ones that are implemented in the offline analysis. The DAQ is composed of a complex system of read-out modules, buffers, buses and event-building systems which takes the output from each detector channel, holds them during the trigger latencies, combines and then sends them to the storage elements preserving the event integrity. In order to have $2.5 \mu\text{s}$ latency, the L1 trigger is implemented by custom hardware, while L2 and the EF are software based and use commercial computers and networks. A frantic activity is on going to define the trigger *menu*¹ for the 2011 run at the luminosity of $10^{33} \text{cm}^{-2} \text{s}^{-1}$. Indeed, the choice of the selection criteria are crucial for the offline physics analyses. It is important to define trigger objects that keep the signal efficiency high and at the same time provide good background rejections. These performances should be stable in p_T and η and the rate cannot exceed the limited bandwidth assigned. Moreover, although the background rejection should be as high as possible, it is necessary to select backgrounds in specific phase space for data driven background estimations. It is pretty hard to take in account all these demands, especially for physics analyses based on complex objects like tau leptons decaying hadronically. As we will see, a lot of options are available and a thorough work is needed to find the proper

¹Trigger menus are tables which list thresholds and selection criteria performed at all the three levels of triggers.

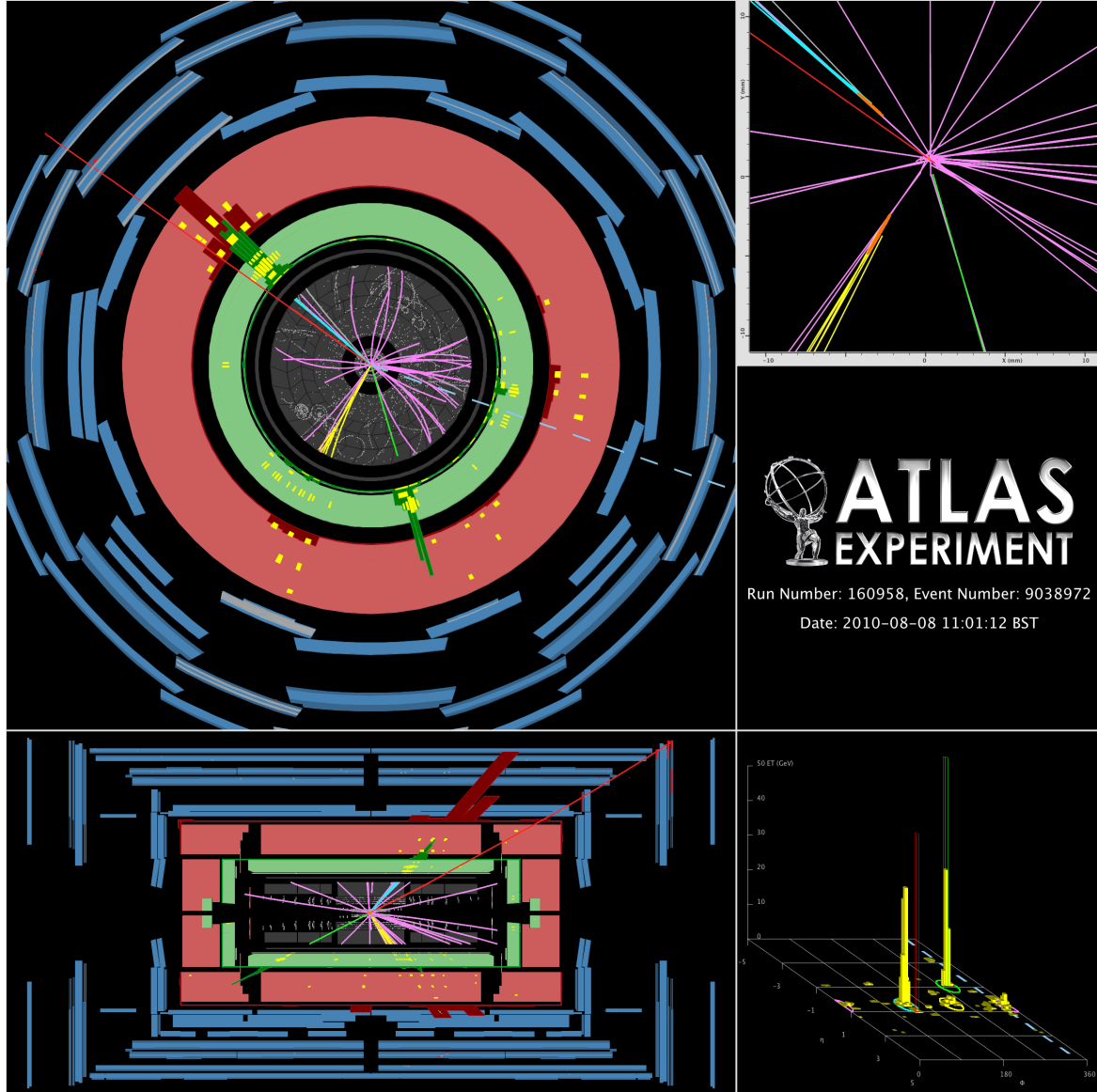


Figure 3.8: ATLAS event display of a $t\bar{t} \rightarrow e\mu + X$ candidate with two b-tagged jets. The electron is the green line, the muon is the red line, E_T^{miss} is the light blue dashed line and the b jets secondary vertices are indicated by the orange ellipses in the top right zoom. The detectors displayed are the electromagnetic (green) and the hadronic (red) calorimeters and the muon spectrometer (blue) [38].

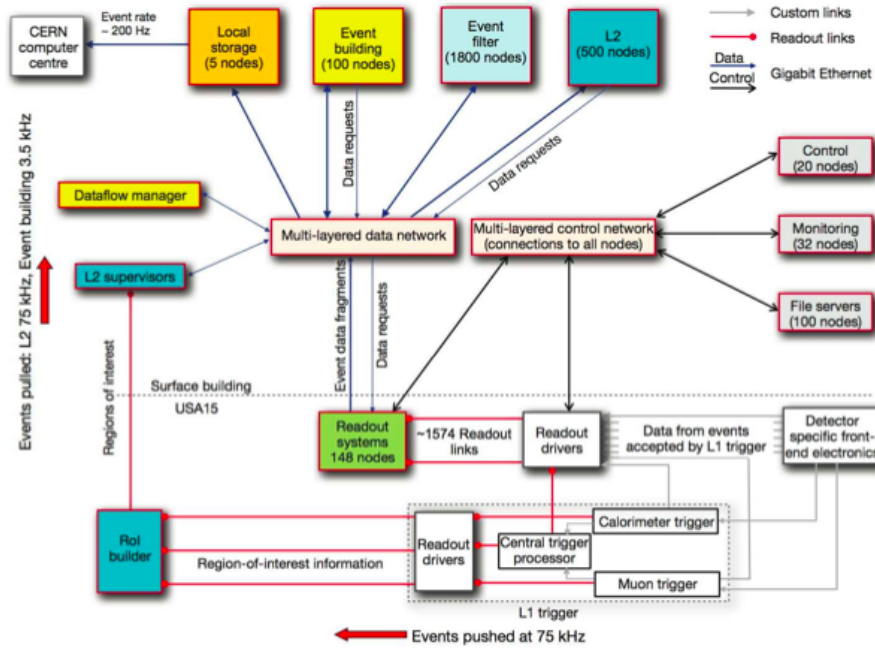


Figure 3.9: Block diagram of ATLAS trigger and data acquisition systems [42].

combinations of them.

3.3 MonteCarlo event generation and ATLAS simulation

The simulation of the events that are expected to be recorded in ATLAS is essential. It allows to evaluate the expected performance of the detectors and the reconstruction algorithms and to compare our understanding of the physical phenomena with what comes out of the collisions. A simulation is performed in two steps:

1. the MonteCarlo event generation
2. the detector response simulation.

The tools used in the first step are called ‘Monte Carlo’ event generators and they simulate the final particle states of the proton proton interactions at the LHC center of mass energies. In the following we will try to illustrate all the processes that are involved in the high energy interactions and which make the simulation so difficult. However it is important to bear in mind that these tools are the only link between theory and observations.

A high energy interaction can be imagined as the combination of four different processes² which are illustrated in Fig.3.10 [46]. The ‘real’ interaction in which new and interesting physics might occur is between two partons of the colliding protons. This reaction is called *Hard Process* (HP) and it is characterized by a high- Q^2 scattering, i.e a scattering where a high momentum is exchanged. This interaction can be calculated using quantum field theory and the uncertainty is only in the numerical approximation coming from the procedure used to include all possible higher order corrections to the basic tree level Feynman diagram for the interaction. Then all the particles fly away from the interacting point entering a new stage of the event which is the *Parton Shower* (PS). In this step all the partons radiate

²This section is based a lecture given by Fabio Maltoni at NBI (slides are available at [45]) and on [44].

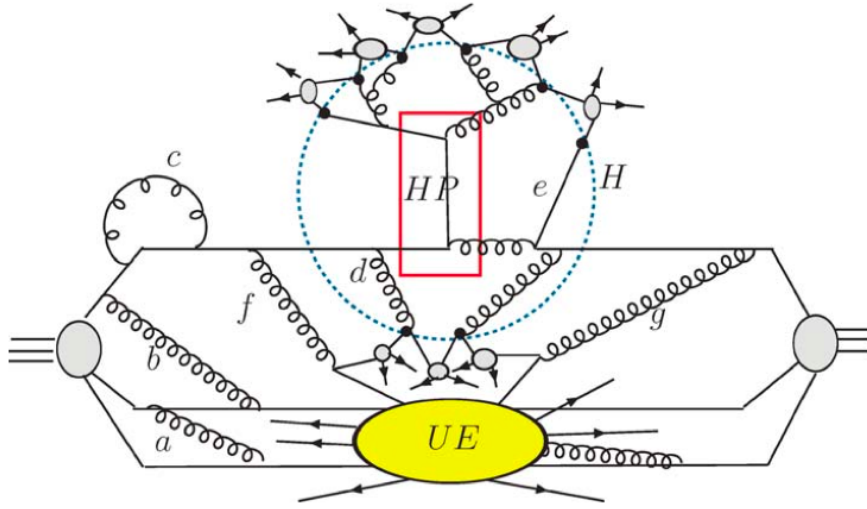


Figure 3.10: Illustration of the structure of a high energy proton proton collision. The *Hard Process* occurs in the red box, the *Parton Shower* in the blue dashed circle and the *Underlying Event* in the yellow spot.

gluons and $q\bar{q}$ couples which combine themselves in color singlet clusters. Indeed no colored final states are observed, this is called infrared slavery in QCD theory³. So this stage has a deep impact and complicates significantly the picture of the HP produced in the event. The PS algorithm is not dependent on the specific HP. The third stage is the *Hadronization* and it occurs when all partons have degraded their energies and they start gathering in clusters which then become hadrons. This step is characterized by QCD interaction at low- Q^2 . In this regime QCD cannot be treated perturbatively and it is necessary to use a model dependent description. As the PS, this algorithm is universal, no matter which was the original HP, and this feature is used to tune phenomenological models. The last part of a high energy interaction is the *Underlying Event* (UE). This part is also dominated by low- Q^2 QCD, so it cannot be treated perturbatively as the hadronization. But the description is even more complex since it has memory of the HP. Indeed, the main characters in this part of the high energy interaction are the remnant partons inside the protons that are not involved in the HP. As we can understand looking at the Fig. 3.10, these particles not only produce a series of complex scatterings, but they also interact with the HP, sharing $q\bar{q}$ pairs. So the tuning of the UE is not only dependent on the energy scale of the collision as the PS and the hadronization, but also on the HP. This scenario is made even more intricate from the experimental point of view by the simultaneous interactions of other pairs of protons that produce the pile up events.

This qualitative description is intended to give a picture of the complexity of the proton proton collisions at LHC and the difficulty in simulating them. As we have seen, only the core of the interaction can be describe exactly by theory, but the following evolution towards the final observable states deeply relies on model dependent tuning. Most LHC discoveries of new physics in proton proton collisions rely deeply on these simulations. They are needed

³Since the QCD QFT is not Abelian the vacuum polarization doesn't screen the charges like in the QED, but produces an 'anti-screening'. This means that the strength of the interaction increases with the increase of the distance between two colored partons. Moreover, since the masses of the u and d quarks are very little, as soon as two partons are separated by a distance comparable with the QCD characteristic scale of 1 fm, then the color potential has enough energy to materialize in a pair of light quarks. This process goes on until there are only color singlet clusters that are detected as hadrons.

to estimate the performance of the detector and they are the only available tool to go from data back to theory.

Once particles and momenta of the final state are simulated, then it is necessary to translate them into data with the detector simulation. This process can be split in three steps. The first one is the proper *Simulation* where the event generated by MC is passed through a model of the detector. This first step is done using the Geant4 software package [47]. The interaction between particles and detectors is simulated and the hits, i.e. the energy deposits, are computed. The second step is the *Digitization* where the hits are translated in time and voltage measurements reproducing the detector response. The output of this stage is exactly the same as the raw data recorded by the experiment with the real collisions. The last step is the *Reconstruction* where digits are transformed in physical objects like tracks, momenta and particles using the reconstruction algorithms. At the end of this chain the simulated events have the same layout as data, so they can be directly compared. If there are discrepancies between the recorded data and the MC events, the cause have to be traced back in the simulation chain. It can be a wrong simulation of the detector or a wrong tuning of MC generators or finally a new occurring physics process.

In the analysis performed in this thesis we used MC events fully simulating the ATLAS detector. However, since the production of such events is highly time and CPU consuming not all physics processes can be simulated in this way. For example QCD di-jets, which are one of the toughest backgrounds to the search for the Higgs boson at low mass values, produced by hard parton scatterings with p_T scale in the range [35-70] GeV have a cross section of about a tenth of millibarn. This means that more than 100 billion events would be needed to prove that this background can be suppressed enough to allow a clear observation of new physics. Such process is simulated with a simplified model of the detector and produced in a single step. This ‘fast’ simulation is performed in ATLAS by the software called *Atlfast* [48].

The events analyzed in this thesis are produced by several MC generators. Here we give a brief description of their features:

MC@NLO[49]: This is a MC generator with Next-to-Leading-Order (NLO)⁴ calculations.

Since the NLO requires a big number of Feynman diagrams with respect to the LO, only specific physics processes are implemented and the multiplicity in the final state is limited. The real corrections included in the NLO calculation can produce events where a hard gluon is radiated at the initial state or in the final state.

AlpGen[50]: This is a LO order MC generator. It is used especially for the simulation of events with a high multiplicity of jets. As we will see later, at the LHC energy events with several high-pt and well separated jets are important, especially for the quest of rare processes such as the Higgs boson or SUSY particles. Indeed, because of the small but not negligible fake rate of such jets, these events have to be properly simulated. However, the need of producing events with multiple jets in the HP makes computationally impossible to reach the NLO level of description. Nevertheless, it is important to mention that the difference with NLO generators is not so big. Indeed, if we compare events like $gg \rightarrow H + 1j$ at LO and $gg \rightarrow H + 1j$ at NLO where the additional jet comes from a real NLO correction, then the level of accuracy at which this jet is described is at LO in both cases.

Herwig[51] and Pythia[52]: They are LO MC generators as well, but they are specialized in the PS simulation. Indeed they can be used to produce the HP and the PS, even if

⁴The NLO is an improvement of the Leading-Order (LO) calculation. It includes the effects of virtual and real corrections on the tree level Feynman diagram which is considered at the LO.

for only a small set of processes like $2 \rightarrow 2$ or $2 \rightarrow 1$, or to make the PS when interfaced to other MC generators like MC@NLO or Alpgen which simulate only the HP. The PS is performed in the *soft* and *collinear* approximation that means that only soft gluons can be radiated at high angles with respect to the mother particle, while hard gluons are produced in a narrow cone. This makes clear why the correct simulation of events with multiple hard and well separated jets has to be done at the HP and not in the PS.

The *matching* between the HP and PS is a delicate issue. Indeed it is necessary to avoid the *double counting* of some Feynman diagrams that can lead to a wrong prediction. The PS algorithm basically adds gluon radiation to any parton in the HP as we can see in the first row of Fig. 3.11.

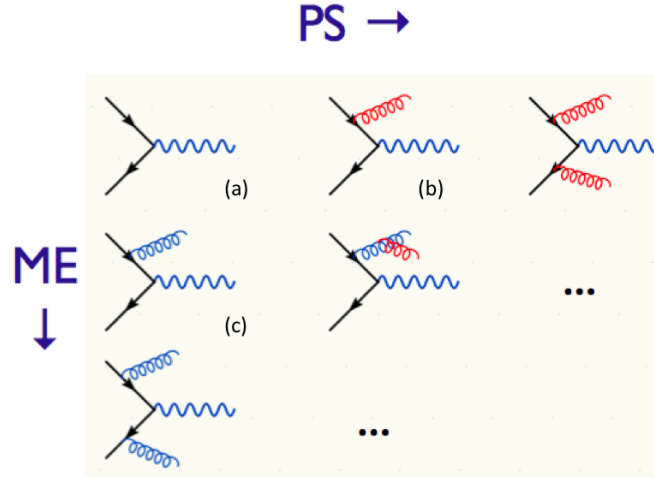


Figure 3.11: Illustration of the matching of the HP, here called ME, and the PS. Moving from the top to the bottom the level of accuracy of the HP increases from LO to NLO and NNLO, while if we move from the left to the right side the PS radiations are sequentially included[45].

However if the HP is at the NLO, then the diagrams of (a) and (c) in Fig. 3.11 are already included. So, if the PS add a gluon like in (b), then the contribution of (c) would be double counted. Several algorithms are implemented in the PS tools like Herwig and Pythia in order to avoid this issue.

3.4 Particles and event properties reconstruction in ATLAS

ATLAS is described as a multi-purpose detector because different fundamental particles interact differently in the various parts of the detector and by this mean are reconstructed and identified as electrons, muons, etc... Some fundamental particles have short lifetime, like tau leptons or b quarks, and they can be reconstructed and identified only through their stable decay products. It is important at this point to briefly describe how tracks and energy deposits, like the ones illustrated in Fig. 3.12, can be translated in four-vector momenta of fundamental particles produced in high energy proton proton collisions.

Electron and Photon: Electrons and photons leave very similar signatures in the detector, the only difference is that electrons leave also a signal in the inner tracking detectors

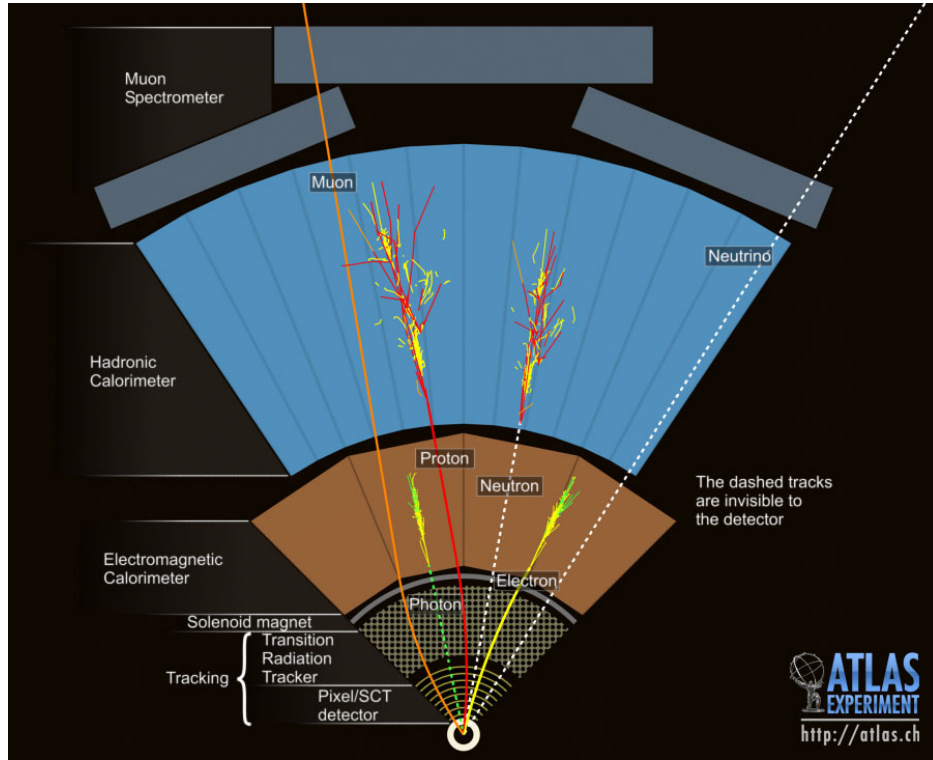


Figure 3.12: Illustration of the signals that each kind of particles leaves in the ATLAS detector and which are used in the reconstruction process.

and photons don't. Indeed the reconstruction of both these objects starts from the detection of a cluster of energy deposits in the electromagnetic calorimeters. This cluster is then the seed for the matching with tracks in the inner detector. From the selected tracks it is then possible to distinguish between electrons and photon conversions. So, basically, an electron is defined by a cluster in an EM calorimeter with an associated track which doesn't come from a photon conversion into electron positron pair, while a photon is an energy deposit in the EM calorimeter without any track or with a reconstructed conversion associated. More precise measurements both in the calorimeter, like the shower shape and the isolation, and in the inner detector, like the track quality, improve the identification of these objects. In many analyses a tight identification is required, which provides good rejection for instance against conversions (high number of hits in the vertexing layers) and charged hadrons (ratio of the high-threshold hits to the number of hits in the TRT).

Muon: Muons typically leave a very clean signature and consequently are used in several physical analyses. Indeed a muon leaves a tracks in the inner detector because it is charged, but it does not leave large energy deposits in the calorimeters as an electron. This is due to the fact that the probability for electrons and muons of radiating photons is proportional to $1/m^2$. This means that an energy loss from photon radiation is 40000 times less probable for a muon with respect to an electron. So typically a muon can fly through all the detector and it can hardly be misidentified. Two independent measurements of the muon tracks are performed, one in the inner detector and one in the muon spectrometer. Then the matching between these two tracks improves the momentum resolution.

Jets and Missing Energy: The reconstruction of jets and missing energy heavily relies on the hadronic calorimeters. There is not a unique strategy of jet finding as well as there is not a unique definition of jets. A jet can be in general defined as energy deposits in calorimeters and tracking devices contained within a certain angular cone or region and originating from particles stemming from the same parton (quark or gluon). The two main jet finding algorithms are based on fixed size cones and sequential recombinations. The inputs are signals measured in the calorimeters cells combined in towers or topological clusters. The fixed size cone finder starts building the jet from any input above a given threshold and it sums the objects in the cells within the cone. The direction of the selected cluster is computed and the cone re-centered. This algorithm is performed iteratively summing the energies in the new cone and updating the direction of the cluster until the resulting four-vector is stable. The sequential recombination jet finder, which is the one used by default in ATLAS, computes weighted distances for any couple of objects and any objects themselves from a seed. Then a jet is associated to the objects or the couple of objects with the minimum weight. If this is a couple, then the two objects are removed and only the combination of them is kept. If the minimum weight is associated to a single input then it is replaced by a jet. This procedure is repeated on the new set of weights until all objects are removed.

The missing energy is an important event signature and it reveals the presence of both neutrinos and new stable neutral particles. It is considered in the object selection of several physics analyses. Its measurement can be performed simply adding the energy deposits in the calorimeters and the reconstructed muons or summing the energy of the reconstructed and classified final state particles. Consequently a global calibration or specific weights need to be applied. In both cases, the real challenge is the evaluation of the contributions from noise and dead regions.

Tau: When we refer to a tau usually we consider tau leptons decaying hadronically since it is impossible to distinguish an electron or muon from a tau lepton decay from one to other particles decays. There are several features that make a hadronic tau different from a parton initiated jet, even if this identification is still very tough. The main hadronic tau decay modes are [6]:

- “1-prong” $\tau^- \rightarrow \nu_\tau h^- \geq 0 \text{ neutrals} \geq 0 K^0$ (77%) like $\tau^- \rightarrow \pi^-(\pi^0)\nu_\tau$
- “3-prong” $\tau^- \rightarrow \nu_\tau h^- h^- h^+ \geq 0 \text{ neutrals}$ (23%) like $\tau^- \rightarrow \pi^-\pi^-\pi^+(\pi^0)\nu_\tau$.

A tau jet can be identified by the track multiplicity (1 or 3), by the shape and the isolation of the energy deposits in the calorimeter and by the reconstruction of the secondary vertex in case of “3-prong” decay. Tau leptons travel about 100 μm before decaying and this flight distance can be resolved. There are two reconstruction algorithms, one calorimetry-based which is seeded by clusters in the calorimeters and the other track-based which starts from a high-quality and high p_T track. Variables useful for tau hadronic decay identification are typically combined in a likelihood method, to improve the performance with respect to a simple cut based method.

b Jets: b jets are also very important for physics analyses because they are a signature of top quark decay and they are important for the Higgs and beyond SM searches. They have a mean path longer than tau leptons and it is hard to distinguish them from other particles that generate a secondary vertex as strange hadrons like K_S^0 or Λ or electrons from conversions. There are several b tagger algorithms and they rely mainly on the impact parameter significance of tracks associated to b jets and/or the secondary vertex found within the jet acceptance.

Chapter 4

SM Higgs Boson Searches in ATLAS

The SM Higgs boson search is one of the original benchmarks for the design and performance of the ATLAS experiment and $H \rightarrow ZZ \rightarrow \mu\mu\mu\mu$ has been promoted to be the *golden channel* driving the experimental design, especially for the muon system. There are other channels which are used in order to explore the SM Higgs boson mass range from the LEP exclusion limit at 114.4 GeV up to 1 TeV. In this section we will describe the main analyses developed in the ATLAS experiment and then we will focus on the Higgs search in the $\tau^+\tau^-$ final states which is relevant for this thesis.

In the low mass region ($m_H < 140$ GeV) the SM Higgs boson is searched in $\gamma\gamma$ and $\tau^+\tau^-$ final states. These are the only feasible channels since the decay modes in $b\bar{b}$, $c\bar{c}$ and gg can hardly be discriminated from QCD backgrounds. Even if the $H \rightarrow \gamma\gamma$ branching ratio is extremely small (Fig. 2.9), this channel has the advantage with respect to the di-tau final state, that the Higgs mass can be precisely reconstructed (Fig. 4.1). In the high mass region

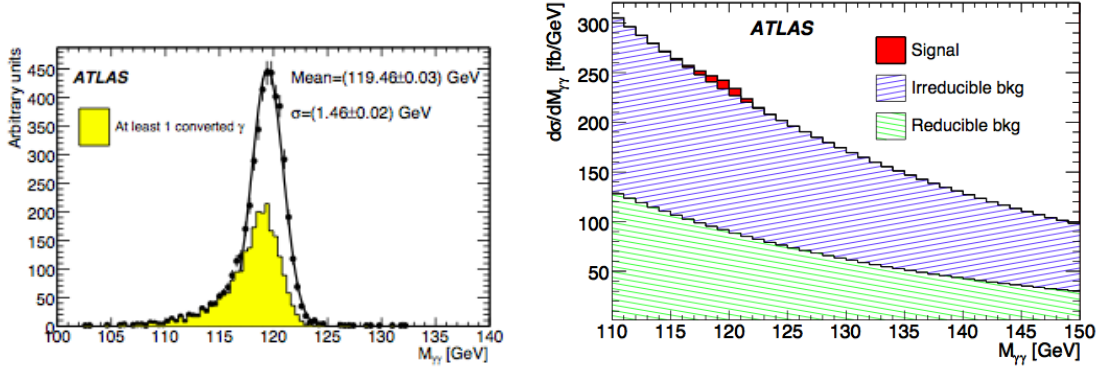


Figure 4.1: Left: Invariant mass for photons pairs from Higgs boson with $m_H = 120$ GeV after trigger and identification cuts. Right: Diphoton Invariant mass spectrum after the application of cuts of the inclusive analysis. The ‘irreducible bkg’ is $\gamma\gamma$ events and the ‘reducible bkg’ is γj and $j j$ events [43].

($m_H > 140$ GeV) the most sensitive search is $H \rightarrow WW$ thanks to the high branching ratio (Fig. 2.9). The final state in pair of leptons ($ee, \mu\mu, e\mu$) has a high sensitivity because these events are characterized by large missing energy and a small angle between the leptons due to spin correlations. Indeed, as Fig. 4.2 shows, since the Higgs boson is a scalar particle (spin 0), in the final state leptons and neutrinos are produced in opposite directions. This

channel will be used to set the first exclusion limits but it is not optimal to determine the mass of Higgs boson since only the transverse mass can be estimated.

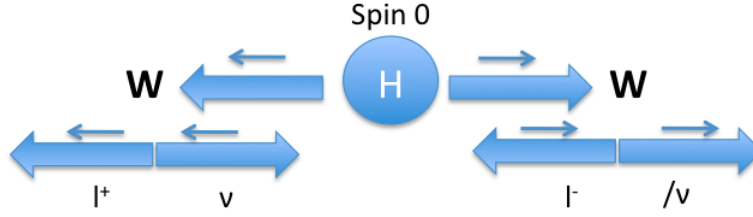


Figure 4.2: SM Higgs boson decay in a pair of W bosons. Thick arrows illustrate momenta and thin arrows are spins.

In order to measure all the features of the Higgs boson it is necessary to use the golden channel $H \rightarrow ZZ^{(*)} \rightarrow 4l(e, \mu)$. This channel has an excellent energy resolution since there are no neutrinos in the final state and it is possible to get a narrow peak in the invariant mass spectrum (Fig. 4.3). Moreover each couple of leptons can be constrained to have an invariant mass compatible with the Z mass, strongly reducing the background events.

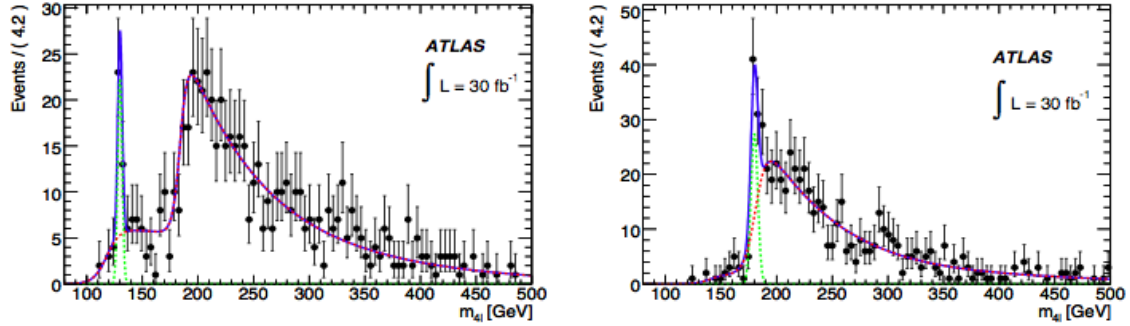


Figure 4.3: Invariant mass of the four leptons in a pseudo-experiment corresponding to an integrated luminosity of 30 fb^{-1} at 14 TeV for $m_H = 130 \text{ GeV}$ (left) and 180 GeV (right). The solid line is the fit of signal and background [43].

The channels that we have described, apart from $H \rightarrow \tau^+\tau^-$ that will be illustrated in the next section, are the most sensitive, but the complete list of SM Higgs boson searches in ATLAS includes also:

- $pp \rightarrow H \rightarrow WW \rightarrow l\nu, qq$
- $pp \rightarrow t\bar{t}H \rightarrow t\bar{t}b\bar{b}$
- $pp \rightarrow t\bar{t}H \rightarrow t\bar{t}WW$
- $pp \rightarrow ZH \rightarrow llWW$

The sensitivity achieved combining all the analyses is shown in Fig. 4.4 [53]. These plots illustrate the value of the cross section, multiple of the SM Higgs boson cross section, which can be excluded with the 95% CL with an integrated luminosity of 1 fb^{-1} at the center of mass energy of 7 TeV. The region where the dashed lines, which are the combined limits, is

below one excludes at the 95% CL a SM Higgs boson in that mass region. It is possible to see the contribution from the single channels. Above $m_H > 200$ GeV the relevant search is $H \rightarrow ZZ$ while from 130 to 200 GeV the important channel is $H \rightarrow WW$.

It is clear that the mass range below 120 GeV is the most tough for the ATLAS experiment. The sensitivity in this range relies basically on the $H \rightarrow \gamma\gamma$ channel. Comparing the ATLAS (Fig. 4.4) and the Tevatron (Fig. 2.14) sensitivities, we can see that the shapes in the low mass region are different and the most difficult range for the Tevatron experiments is not below 120 GeV, but around 130 GeV. Since the most probable SM Higgs boson mass value, as we know so far, is at 120 GeV (Fig. 2.15), then ATLAS needs to improve its analysis in this mass region in order to set limits before Tevatron.

4.1 The VBF analysis

Besides $\gamma\gamma$, the other final state relevant for the Higgs boson with a mass around 120 GeV is $\tau^+\tau^-$ [43]. The ATLAS search in such a final state is optimized for the *Vector Boson Fusion* (VBF) production mode (Fig. 2.8). The reason of using this and not the process with the highest cross section (gluon fusion) is because VBF events have a peculiar topology that makes them nicely distinguishable from backgrounds. Indeed the two quarks produced together with the Higgs boson typically hadronize in two jets which are observed in the forward region of the detector. In addition, since there is no strong interaction between these partons, they are not correlated by color fields. This means that gluons can be radiated only collinearly to the quark directions and so the jet activity in the central part of the detector is expected to be very low. The final state are therefore characterized by two forward jets and a couple of taus produced between them. This topology can be implemented in a set of cuts that provide a nice discrimination against backgrounds.

Fig. 4.5 shows an illustration of how a VBF event ideally looks like. The event selection requires two ‘tagging jets’ emitted with high rapidity and widely separated. Then it looks for a couple of tau leptons between the two jets. Finally it applies the *Central Jet Veto* since no other jets are expected to be radiated in the central part of the detector. This analysis is used in the current ATLAS public limits and sensitivity predictions, including only the leptonic and the semi-leptonic final state, which means that the event is selected only if at least one tau decays leptonically.

4.2 An auxiliary analysis

The aim of this thesis is to prove that the VBF analysis doesn’t explore the full potential of the $\tau^+\tau^-$ final states in the search for the Higgs boson. Indeed, it is possible to design another event selection that can accept also SM Higgs boson produced by the gluon fusion and includes the double hadronic tau decay channel, therefore significantly improving the combined sensitivity.

The idea that the gluon fusion signal might be considered in the $H \rightarrow \tau^+\tau^-$ quest comes from two previous studies. The first one [55] is a theoretical study at parton level, which means that it doesn’t consider the detector simulation. The authors claimed that the SM Higgs boson could be observed with few fb^{-1} of integrated luminosity in pp collisions at the center of mass energy of 14 TeV in the channel $pp(gg + gq) \rightarrow H + j \rightarrow \tau^+\tau^- + j$. In this channel the Higgs boson is produced in association with a hard jet. The second study has been presented within the ATLAS collaboration [56]. It considered only the leptonic final state and events with missing energy and a high-pt jet.

The analysis performed in this thesis takes some ideas from these two papers and goes in details and extends these studies, proving that the gluon fusion process indeed can be

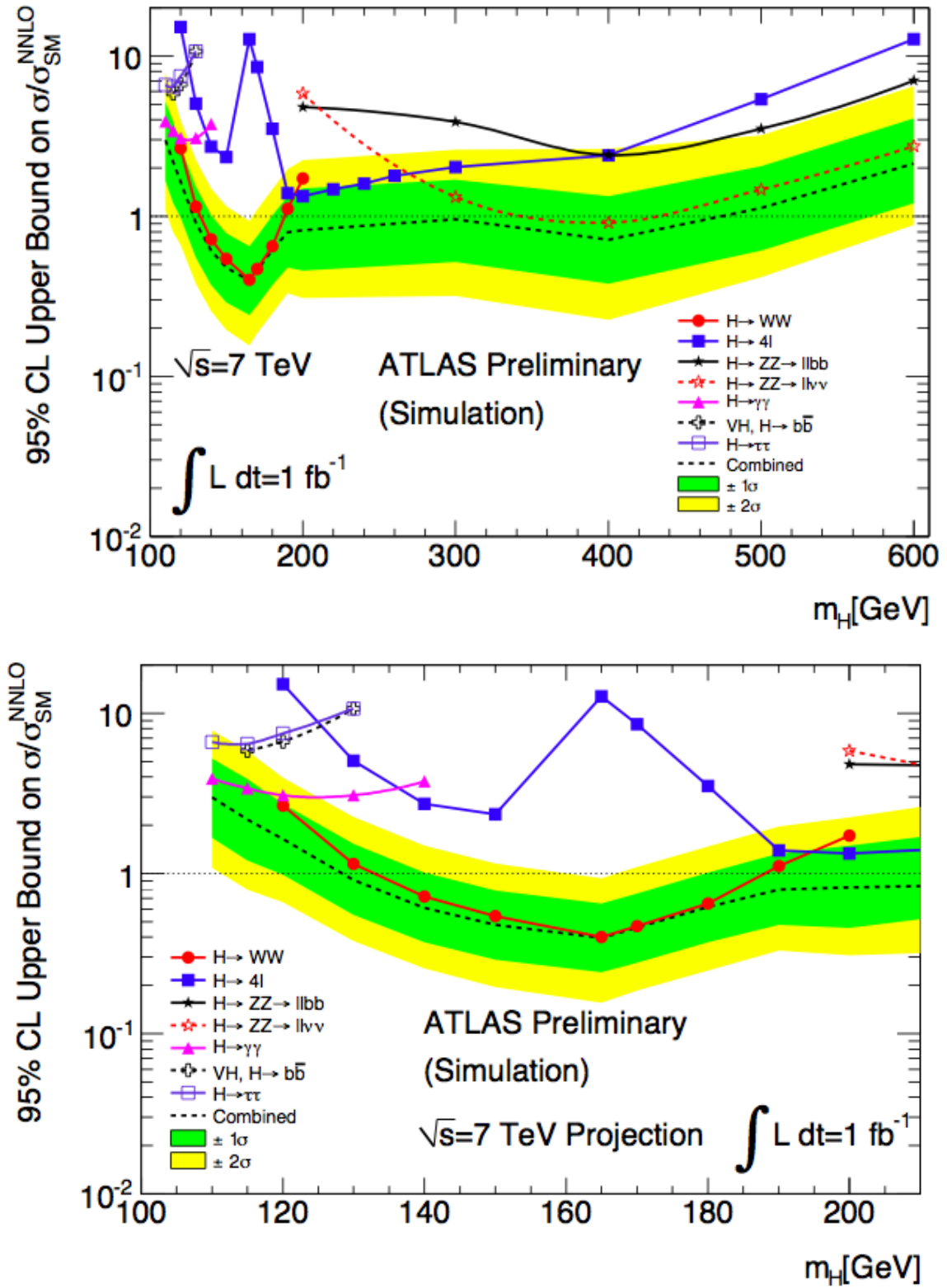


Figure 4.4: Multiple of the cross section of the SM Higgs boson that can be excluded with the 95% CL with an integrated luminosity of 1fb^{-1} in proton-proton collision at the center of mass energy of 7 TeV.

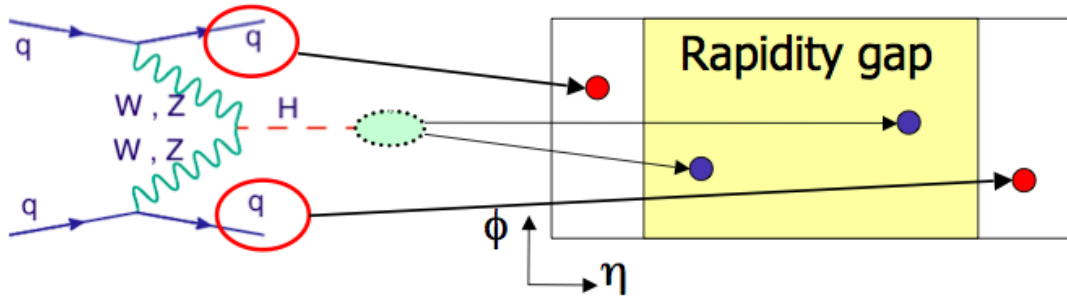


Figure 4.5: Illustration of the selection strategy for VBF events. Two ‘tagging jet’ are selected in the forward region with a wide angular separation and a pair of tau leptons with opposite charge is detected in the central region between the tagging jets [54].

used as well to search for the Higgs boson and that all the possible decays for tau leptons should be and can be included to increase the sensitivity. The study is performed using a full simulation of the ATLAS detector and effects of pileup are evaluated.

Chapter 5

The $pp(gg+VBF) \rightarrow H+jets \rightarrow \tau^+\tau^- + jets$

The analysis presented in this thesis is a SM Higgs boson search in $\tau^+\tau^-$ final states in simulated events of proton-proton collisions at the center of mass energy of 10 TeV. Signal events are produced with an Higgs boson mass of 120 GeV. All the tau decays, full leptonic, semi-leptonic and full hadronic, are included. Given the probability of a leptonic decay of 35%, the branching ratios are:

- leptonic channel $H \rightarrow \tau^+\tau^- \rightarrow ll$ 12%
- semi-leptonic channel $H \rightarrow \tau^+\tau^- \rightarrow lh$ 45%
- hadronic channel $H \rightarrow \tau^+\tau^- \rightarrow hh$ 43%

5.1 Signal

The Higgs production processes examined are the *Gluon Fusion* (GF) and the *Vector Boson Fusion* (VBF). The Feynman diagrams of these processes are illustrated in Figg. 5.1 and 5.2.

In Fig. 5.1 (a) the gluon fusion is at the leading order and the other three diagrams are next-to-leading order processes where there is also a gluon in the final state. As explained later, in this analysis we select events where there is at least one hard jet in addition to the Higgs boson in the final state. So the greatest yield of gluon fusion signal comes from the diagrams (b),(c) and (d). It is important to mention that the loss of the LO signal is not so dramatic since the NLO corrections are of the same size of the LO cross section. Indeed the perturbative expansion of the inclusive cross section is slowly converging [57]:

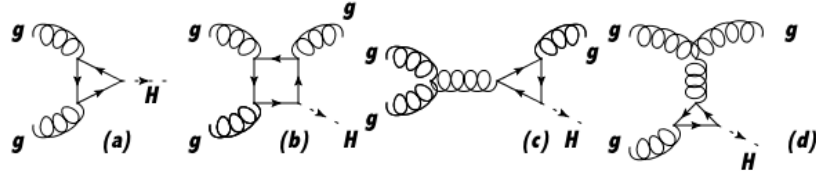


Figure 5.1: Feynman diagrams of H production in the gluon fusion process. The loops includes all quarks, even if the main role is played by the top quark. (a) describes the production at the leading order and (b),(c) and (d) are some of the next to leading order diagrams where the Higgs boson is produced together with a gluon.

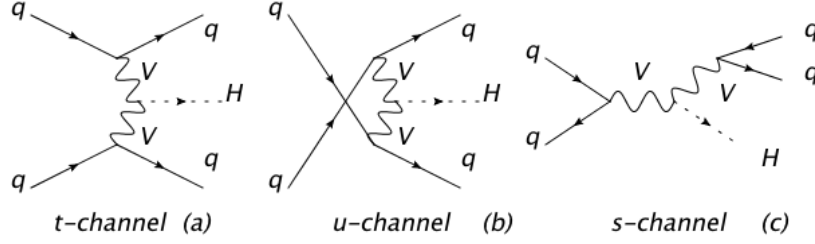


Figure 5.2: Feynman diagrams of the Vector Boson Fusion production mode. ‘V’ represents all weak bosons, W^+ , W^- and Z .

$$\sigma = \sigma_{LO} + \sigma_{NLO} + \sigma_{NNLO} + \dots = \sigma_{LO} \cdot (1 + 0.8 + 0.3 + 0.1 + \dots) \quad (5.1)$$

Furthermore the final state with only a Higgs decay into tau leptons is very hard to distinguish experimentally from QCD dijets backgrounds.

The VBF diagrams are shown in Fig. 5.2. In order to preserve the gauge invariance all the three processes have to be included in the computation of the inclusive cross section, even if usually only the t- and u-channels are considered as ‘true’ VBF processes. The s-channel is typically regarded as a Higgs radiation, so it is called ‘Higgs-strahlung’. The reason for this distinction is mainly due to the fact that the main signature of the VBF events is a couple of forward jets. Indeed, all the VBF searches pick up events with such a topology and this selection makes the s-channel almost negligible. Due to this, most of the MC generators used in ATLAS include only the t- and u- channels in the VBF process and this is the case also for the sample used in this analysis. This is a drawback for this study where the s-channel can give a significant help in the signal final yield. Indeed, looking at the Table 5.1 where the VBF cross sections with and without the s-channel are reported, we can see that the difference is close to 50%. Since half or more of the signal events accepted in this analysis are produced through the VBF, the inclusion of the s-channel in the search strategy outlined here promises to confirm even more the di tau channel as a crucial final state for SM Higgs search at low masses¹.

Table 5.1: Total inclusive cross sections for Higgs production in proton-proton collisions.

$m_H = 120 \text{ GeV}$	$\sqrt{s}=7 \text{ TeV}$	$\sqrt{s}=10 \text{ TeV}$	$\sqrt{s}=14 \text{ TeV}$
GF: $\sigma_{NNLO}(gg \rightarrow H)$	16.69 pb	31.03 pb	53.50 pb
VBF: $\sigma_{NLO}(gg \rightarrow ggH) \text{ (t+u)}$	1.24 pb	2.40 pb	4.26 pb
VBF: $\sigma_{NLO}(gg \rightarrow ggH) \text{ (t+u+s)}$	1.84 pb	3.41 pb	5.84 pb

In Table 5.1 it is also possible to compare gluon fusion and VBF cross sections at different center of mass energies. These are the latest calculations available and they are reviewed in [57].

Regarding the MC samples used for the simulation of gluon fusion events, these are made with MC@NLO[58]. The process generated is $H_1 H_2 \rightarrow H + X$, where H_1 and H_2 are generic partons and X stands for all possible final states. This process includes diagrams like $gg \rightarrow H$ and $gg \rightarrow Hg$, that are illustrated in Fig. 5.1, and also $qg \rightarrow Hq$. VBF is not considered since at LO it is a pure electro-weak process. At the moment the total cross section used to normalize gluon fusion events is the one indicated in Table 5.1 and the qg

¹This issue will be examined in the future with a proper simulation of the s channel.

process is not included. An estimation with MadGraph [59] showed that the qg contribution is only few percents of the gg cross section. The VBF events are produced with Herwig[51], a LO event generator since the VBF process doesn't have QCD corrections at NLO.

The center of mass energy used in all the MC samples, both signal and backgrounds, is 10 TeV because when this study started, the default energy used in the MC simulations available was 10 TeV, and not 7 TeV as in real collisions in 2010. It turns out, given the current rumors of pushing energy to 8 or even 9 TeV in 2011, that these simulated samples are very useful to estimate the possible reach in 2011. Only the MC sample for the gluon fusion signal in the hadronic final state is simulated at 7 TeV because it is the only one available. A comparison between gluon fusion signals in the semileptonic channel at 7 and 10 TeV has shown that the main variables used in this analysis are not so different.

5.2 Backgrounds

In this section we report a list of the sources of background events to be considered in this analysis.

$W \rightarrow l\nu_l + j$ This kind of events might contaminate the signal region especially if the W boson is produced with 2 or 3 hard jets. In the semi-leptonic and hadronic channels, the probability that a QCD jet fakes a tau jet is small, but still significant, and therefore $W + j$ events have to be considered. Indeed even if the tau jet fake rate is less than a per mille, the probability of such events is about 4 orders of magnitude bigger than the chance to detect a signal event. In this analysis we examine all the leptonic decays of the W boson with a jet multiplicity up to 5. The MC samples are made with AlpGen [50] and the total cross sections are scaled by a k-factor of 1.22 from LO to NLO.

$Z \rightarrow \ell^+\ell^- + j$ This is the most troublesome background especially because it is impossible to distinguish a pair of taus coming from a Z or a Higgs boson². This kind of events is called *irreducible background*.

More specifically, in the leptonic channel the selected pair of leptons can come both from $Z \rightarrow e^+e^-, \mu^+\mu^-$ and from $Z \rightarrow \tau^+\tau^- \rightarrow \ell^+\ell^-$ events. In the semi-leptonic channel it is still possible to accept Z decaying in electrons or muons because of the imperfect performance of electron and muon vetoes, but in the hadronic channel only $Z \rightarrow \tau^+\tau^-$ events might be selected because the probability of identifying two muons or electrons as hadronic tau decays is too small. In this study all the decay modes in leptons in all the three channels with a jet multiplicity up to 5 are considered, even if only events with 1, 2 or 3 jets are relevant due to the veto for high jet multiplicities. In the MC samples produced by AlpGen³, processes where the Z boson is replaced by a virtual photon are implemented as well. The interference between intermediate Z and γ^* is also included. In order to increase the efficiency in the event generation, cuts have been applied on the di-lepton invariant mass $60 < M(\ell\ell) < 200$ GeV and on the lepton transverse momentum $p_T > 15$ GeV. Also in this case a k-factor of 1.22 is applied to scale the total cross sections from LO to NLO.

single top and $t\bar{t}$ $t\bar{t}$ events cannot be neglected since they are characterized by chains of decays where E_T^{miss} and jets that can fake taus are produced. Indeed, the top quark decays almost all the times in a b quark and a W boson. So leptons can be produced

²The difference in the spins of H and Z affects the polarization of tau leptons differently. Future improvements of this analysis will take these effects into account.

³The $Z \rightarrow \tau^+\tau^- + 1j$ sample used in the leptonic and semileptonic channels has been privately simulated with Atfast since the one officially produced has a too little statistics for this analysis.

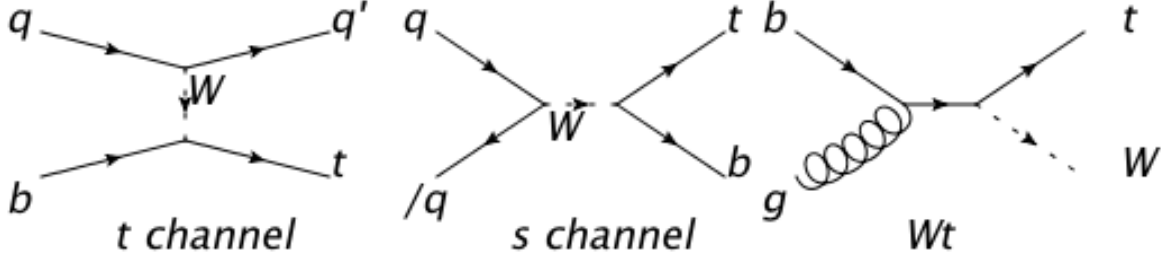


Figure 5.3: Feynman diagrams of single top production processes.

in the W decay and fake tau jets can be generated by the decay of the b or the second W (in $t\bar{t}$ bar events) in light quarks. MC samples are produced with MC@NLO and they include full and not full hadronic $t\bar{t}$ and single top in the t -, s - and Wt channel, from left to right in Fig. 5.3 respectively.

di-jets QCD is the background most difficult to handle. Indeed, because of the huge cross section it is necessary to simulate a rejection power at least up to 10^{10} . This is not feasible. Some compromises have to be accepted. These events are simulated with a ‘fast’ algorithm called AtI Fast [48]. The samples used in this analysis contain di-jets events produced with Pythia [52]. With these approximations we are able to run on samples with about 40 million events. When enough data are collected the QCD background will be estimated on data rather than on MC.

The MC datasets are split in p_T ranges of the hard jets, so we will use the following labels: J0 [8-17] GeV, J1 [17-35] GeV, J2 [35-70] GeV, J3 [70-140] GeV and J4 [140-280] GeV.

5.3 Trigger

Since the main backgrounds made with AtI fastII don’t have any trigger bits and the trigger menu in the signal sample is not feasible for high luminosities such as $10^{33}\text{cm}^{-2}\text{s}^{-1}$, we decided to perform this analysis without any trigger requirements to start with. However we choose the offline selection thresholds on transverse momenta high enough to match the plateau region for trigger efficiency for the triggers expected in $10^{33}\text{cm}^{-2}\text{s}^{-1}$. These efficiencies will be applied on top of the final sensitivity. In this section we will present some studies on which kind of triggers could be suitable for this analysis.

As a general statement, it is better to keep the p_T thresholds for taus and leptons as low as possible otherwise the signal statistics is strongly affected. This can be seen in Table 5.2. So, multiple-objects triggers are preferred to single-object ones that necessarily have high thresholds to keep rates low. Regarding the leptonic channel, a plausible choice might be two leptons with trigger p_T thresholds around 15 GeV. For this reason in the offline selection we consider electrons and muons with $p_T > 20$ GeV to account for trigger resolution effects. In the semileptonic channel combined triggers like tau hadronic plus electron or muon, with a trigger threshold of 15 GeV on electron or muon and 16 GeV on the hadronic tau are a reasonable choice for $10^{33}\text{cm}^{-2}\text{s}^{-1}$. The offline cuts are 20 GeV for leptons and 30 GeV for taus. Since the turn-on curve in tau triggers is slower than in the lepton ones, it is necessary to have a wider difference between trigger and offline thresholds in order to be in the plateau efficiency region.

For the hadronic channel, it is possible to choose between double tau or tau+missing energy triggers. Figg. 5.4 and 5.5 show a rough estimation of the efficiencies of tau16_loose+xe35,

Table 5.2: Estimation of offline cut efficiencies as a function of the p_T of tau and muon for events with 1 tau (LlhMedium, 1 or 3 tracks) and 1 isolated muon with opposite charges. The sample is gluon fusion Higgs with semileptonic decay.

μp_T [GeV]	τp_T [GeV]							
	15	20	25	30	35	40	45	50
15	1.00	0.93	0.81	0.68	0.54	0.41	0.29	0.20
20	0.76	0.70	0.61	0.51	0.41	0.31	0.22	0.15
25	0.55	0.51	0.45	0.37	0.30	0.22	0.16	0.11
30	0.39	0.36	0.32	0.27	0.21	0.16	0.11	0.07
35	0.27	0.25	0.22	0.18	0.14	0.10	0.07	0.05

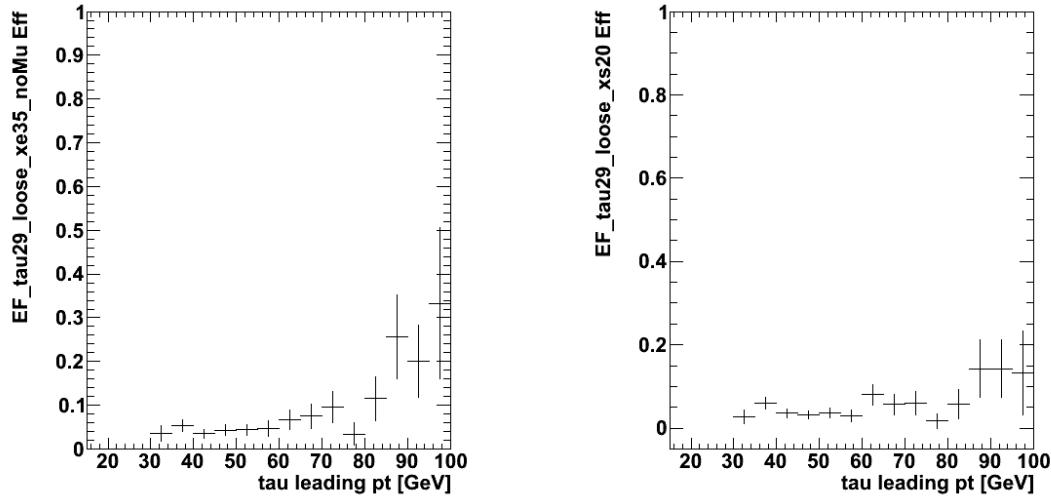


Figure 5.4: Estimation of EF_tau29_loose_xe35 and EF_tau29_loose_xs20 trigger on events selected offline with 2 taus with $p_T > 30$ GeV, $E_T^{\text{miss}} > 15$ GeV and at least one jet with $p_T > 20$ GeV. The sample is gluon fusion Higgs with double hadronic decay.

tau16_loose+xs20⁴, tau29_loose_tau20_loose and 2tau29_loose triggers. These efficiencies are computed for signal events selected offline with 2 taus with $p_T > 30$ GeV and $E_T^{\text{miss}} > 15$ GeV.

By applying a cut on E_T^{miss} , the signal efficiency drops very quickly. The lowest threshold on E_T^{miss} for a tau+xe trigger affordable at the luminosity of $10^{33}\text{cm}^{-2}\text{s}^{-1}$ might be 35 GeV and offline one does not need to require more than 20 GeV. As Fig. 5.4 shows, this choice would be detrimental for the signal. It is better to not require any E_T^{miss} already in the trigger. Regarding double tau triggers, two different thresholds would be advisable. Indeed Table 5.3 (right) shows that the p_T distributions of the two taus are sensibly different. The tau29_loose_tau20_loose is the best choice for this analysis. If the rate is too high, it is possible to put a tighter cut on the number of tracks. This will not have any impact on the offline analysis since this requirement is also present in the offline tau selection.

At the moment these thoughts are very preliminary since big efforts are ongoing especially in the improvement of the tau identification at trigger level. So estimations of efficiencies and

⁴‘xs’ stands for *Met significance* and it is defined as the ratio of transverse missing energy over the square root of SumEt, the scalar sum of E_T^{miss} along the x and y axes.

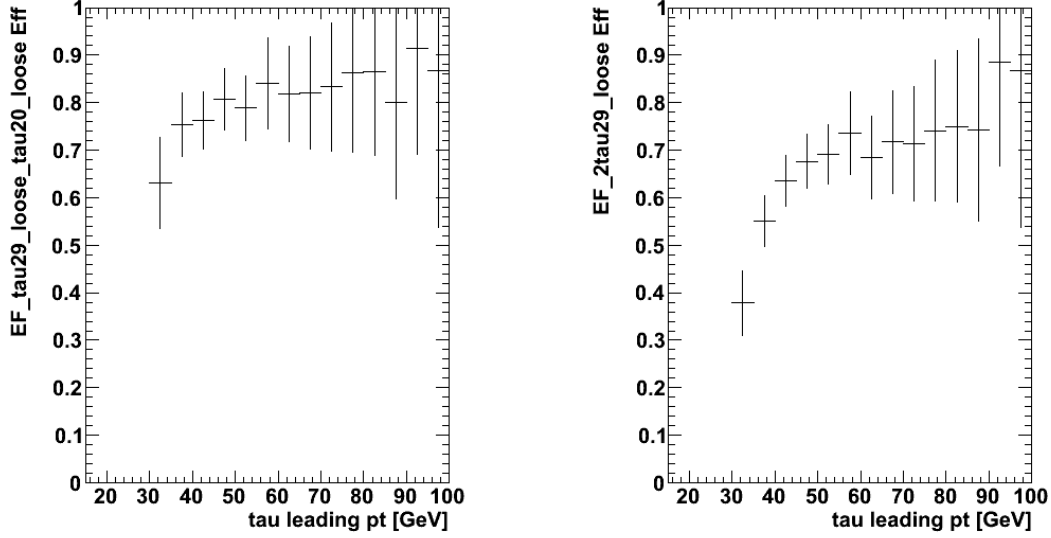


Figure 5.5: Estimation of tau29_loose_tau20_loose (left) and tau29_loose_tau29_loose triggers (right) at EF on events selected offline with 2 taus with $p_T > 30$ GeV, $E_T^{\text{miss}} > 15$ GeV and at least one jet with $p_T > 20$ GeV. The sample is gluon fusion Higgs with double hadronic decay.

Table 5.3: Estimation of offline cut efficiencies as a function of E_T^{miss} and p_T of the leading tau (left) and as a function of the p_T of the two taus (right). The selected events have 2 taus (LlhMedium) with opposite charges and 1 or 3 tracks, $E_T^{\text{miss}} > 15$ GeV and at least 1 jet with $p_T > 20$ GeV. The sample is gluon fusion Higgs with double hadronic decay.

E_T^{miss} [GeV]	leading τp_T [GeV]					subleading τp_T [GeV]	leading τp_T [GeV]				
	30	35	40	45	50		30	35	40	45	50
15	1.00	0.95	0.81	0.65	0.49	30	1.00	0.95	0.81	0.65	0.49
20	0.69	0.65	0.56	0.44	0.33	35	0.56	0.56	0.52	0.43	0.33
25	0.47	0.44	0.37	0.29	0.23	40	0.29	0.29	0.29	0.26	0.22
30	0.32	0.30	0.26	0.20	0.16	45	0.15	0.15	0.15	0.15	0.13
35	0.23	0.21	0.19	0.15	0.12	50	0.08	0.08	0.08	0.08	0.08

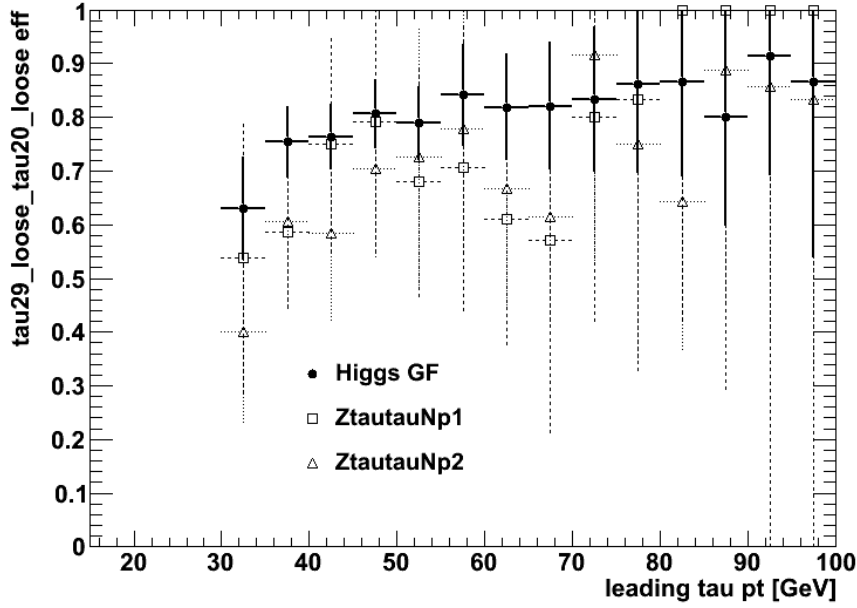


Figure 5.6: Comparison of trigger efficiencies for gluon fusion signal, $Z + 1j$ and $Z + 2j$.

rates are difficult. Despite this, $\text{tau29_loose_tau20_loose}$ can be taken as a baseline trigger and we will present the final sensitivity taking an efficiency on signal of 80% (Fig. 5.5, left). Fig. 5.6 shows that if the same efficiency is taken for both signal and $Z + j$ background, then the estimation of the sensitivity will be conservative.

5.4 Objects and Event selection

In this section we list the requirements that objects in each event have to fulfill in order to be selected and used to compute event variables such as the mass of the Higgs candidate.

Muons Only muons with transverse momentum higher than $p_T > 20$ GeV in the pseudo-rapidity range $|\eta| < 2.5$ are accepted. The limits on η correspond to the coverage of the tracking system of the inner detector. In order to select high-quality muons, the selection requires that the reconstructed muon has two tracks, one in the muon spectrometer and one in the inner detector. Furthermore the matching of these two tracks performed by the ‘Staco’ algorithm needs to be with a $\chi^2 < 100$.

Electrons Only reconstructed electrons identified as ‘Tight’ are accepted and they need also to be emitted with $p_T > 20$ GeV in $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$. The limit at 2.47 is due to the EM calorimeter acceptance, while the intermediate range between $1.37 < |\eta| < 1.52$, called ‘crack’ region, is excluded since it corresponds to the transition between the barrel and the endcap calorimeters.

Jets All jets reconstructed in the ‘AntiKt4H1Topo’ collection are selected. These are produced by the sequential recombination jet finder ‘AntiKt’ which is seeded by topological clusters. The rapidity range is $|\eta| < 4.4$ and it is determined by the acceptance of the forward hadronic calorimeter.

Tau Only tau jets identified by a likelihood medium selection are considered (efficiency of 50%) and have to be in $|\eta| < 2.5$ with at least $p_T > 30$ GeV. An electron and muon

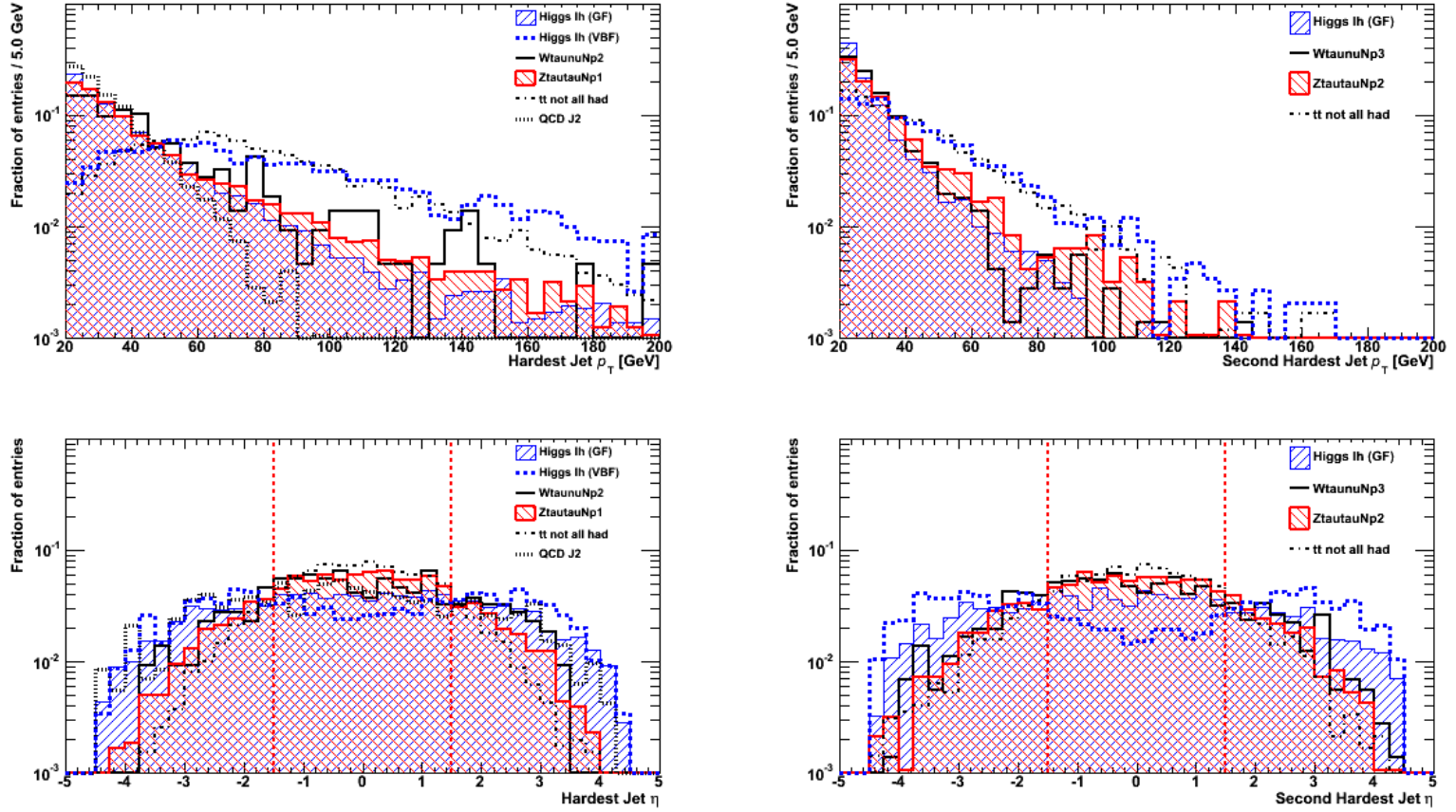


Figure 5.7: Distributions of the hardest and second hardest jets.

veto is also applied.

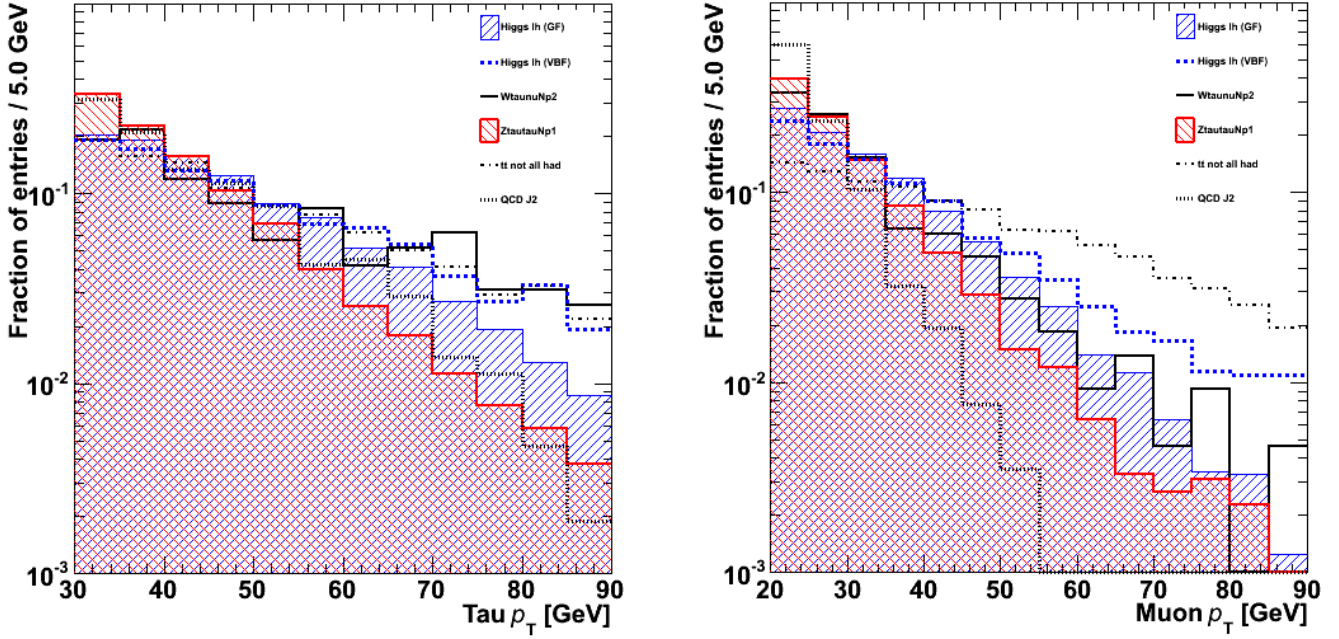
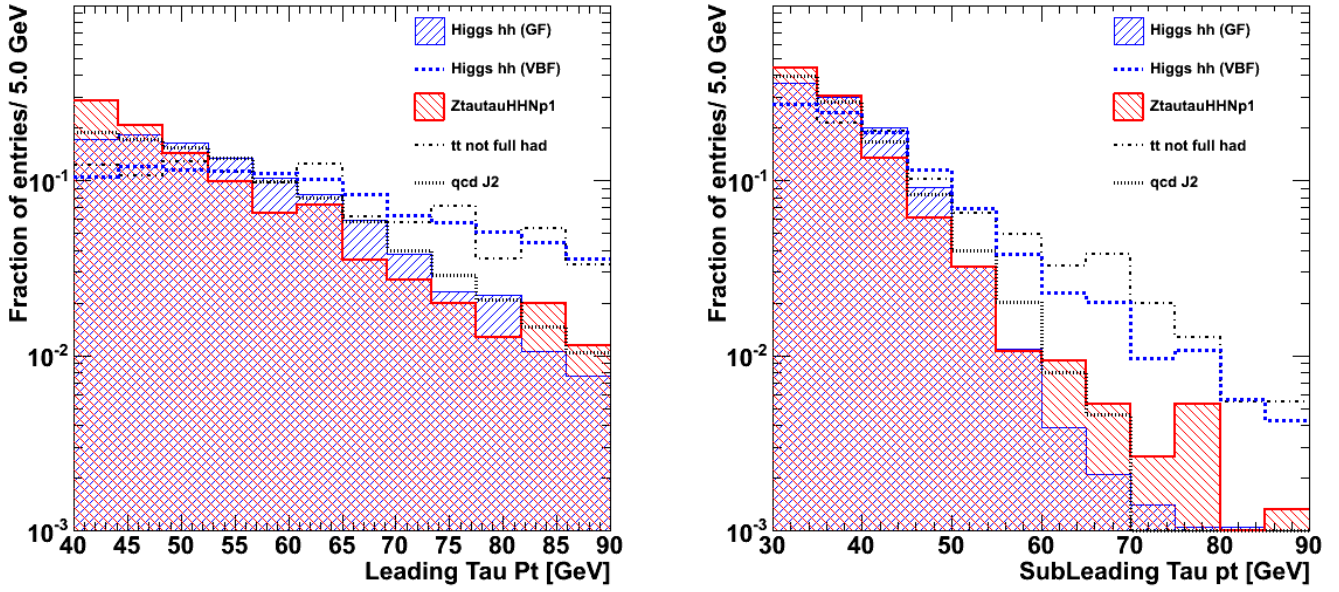
The selected objects in the final state are requested not to overlap with each others. This means that if the angular distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.2$ for a given couple of objects, then only one of the two is kept and the other is rejected according to these priorities: for lh-channel $\mu \rightarrow e \rightarrow \tau \rightarrow j$, for ll-channel $\mu \rightarrow e \rightarrow j$ and for hh-channel $\tau \rightarrow j$.

As it will be explained later it is necessary to select events where the Higgs candidate is boosted, so we require the presence of at least one hard jet with $p_T > 20$ GeV in $|\eta| < 4.4$. Moreover, since the Higgs production mode considered in this analysis hardly have more than two jets, we reject events with more than two jets with $p_T > 20$ GeV in $|\eta| < 4.4$. This veto helps reducing the background from events with high jet multiplicity like $t\bar{t}$ or $W/Z + 3, 4, 5j$. The hardest jet will be called *tagging jet*.

Finally each event is assigned to one of the three channels according to the number of tau jets and leptons. If there are exactly two leptons, the event is selected for the leptonic channel and if there is one lepton and one tau jet it is selected for the semileptonic channel. The events accepted in the hadronic channel need to have exactly two hadronic taus, at least one of which has a transverse momentum $p_T > 40$ GeV.

Fig. 5.8 and 5.9 show the p_T distributions of taus and muons in the μh and hh channels. These are exemplifying plots to illustrate the effects of applying thresholds on the transverse momentum of taus and leptons.

In order to have enough statistics after the full selection to prove the level of background left, in the QCD backgrounds the τ identification efficiency has been factorized with the

Figure 5.8: Distributions of tau and muon p_T in the μh channel.Figure 5.9: Distributions of tau p_T in the hh channel.

efficiency of the rest of the selection. This has been done using all the events where there is at least one τ with all the features described before except for the tau medium identification. Then two sets of 2D histograms in $(\eta - p_T)$ have been plotted for only identified and all τ . Finally the weights were determined dividing the first set by the second set of plots. These histograms are shown in Appendix C. The method to define the weight for each event will be described in section 5.5.9.

5.5 Cut Description

Once the event and its reconstructed objects are accepted for one of the channels, then a further event selection is applied. In this section all the cuts implemented are illustrated even if some of them are not applied in all the channels. All the plots shown in the following, if not differently specified, are drawn at the ‘TauTrack’ step for the semi-leptonic and hadronic channels and at the ‘Charge’ step for the leptonic channel.

5.5.1 Initial Cuts

The first set of cuts is a further cleaning of the selected leptons and tau jets:

Charge The charges of the leptons or tau jets selected need to be opposite;

TauTrack Tau jets are required to be 1- or 3-prong, which means that they need to have only 1 or 3 tracks.

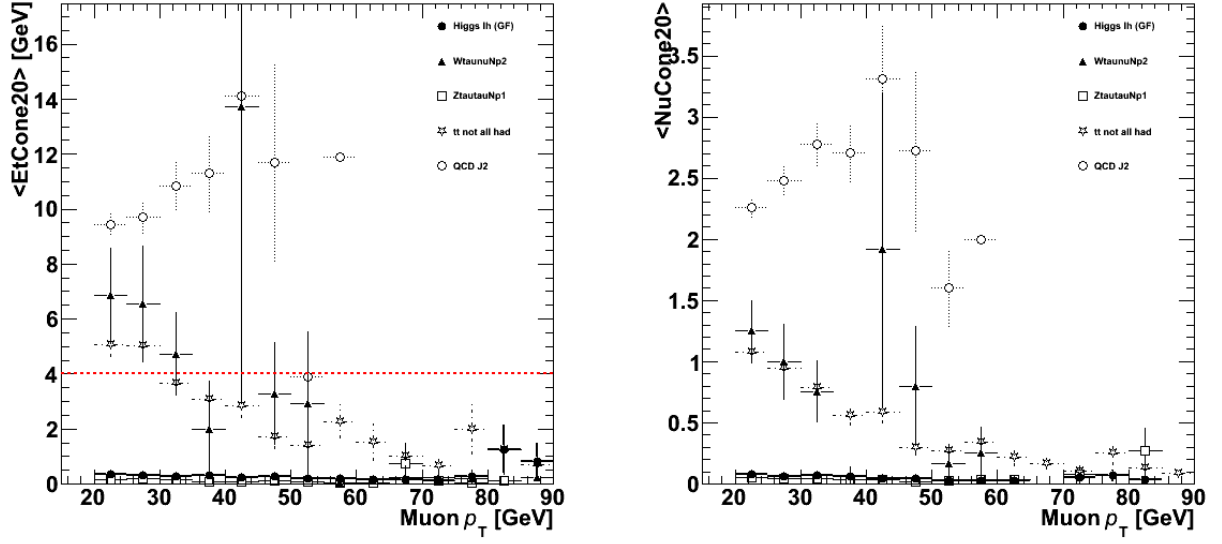
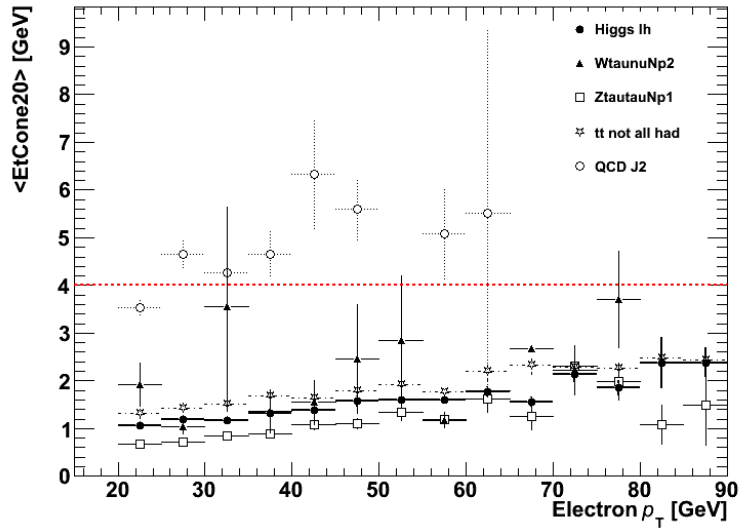
Isolation In order to select only isolated muons and electrons, a cut on the energy deposited in the calorimeter is applied. The energy measured in a cone of radius $\Delta R < 0.2$ around the track associated to the lepton has to be less than 4 GeV (this variable is called EtCone20). This energy doesn’t include the energy of the lepton itself. In case of muons, it is also required that the number of additional tracks within the same cone (NuCone20) is zero.

The distributions of the isolation variables are plotted in Figg. 5.10 and 5.11. The isolation cuts are more tight on muons because of the requirement of no additional tracks in the cone (right plot of Fig. 5.10). The same cut is not applied on electrons since this variable is not officially recommended within ATLAS. The isolation requirement is effective especially against QCD background. For muons the most effective cut is on the number of tracks and a lower threshold on Etcone20 for muons is useless since at the end of the cut flows all muons are well isolated.

5.5.2 B Jet Veto

$t\bar{t}$ events are characterized by two main features: high multiplicity of jets and the presence of b quarks in the decay products of top quarks. Since it is preferable to not apply any veto on hard jets additional to the tagging one in order to save the VBF signal, events with b quarks need to be excluded. There are several algorithms that provide the probability that a jet is generated by the hadronization of a b quark and in this study the one called IP3D+SV1 is used. This method is a combination of two algorithms that provide a weight for each jet depending on the impact parameter and the secondary vertex respectively⁵. We decide to exclude events where there are one or more jets with a weight greater than 4.0. It has been

⁵The IP3D tagger is based on the 3D impact parameter significance, i.e. the ratio of the impact parameter to its error, and the SV1 tagger is based on properties of the secondary vertex such as its mass, the fraction of charged energy and the number of tracks.

Figure 5.10: Muon isolation in μh channel.Figure 5.11: Electron isolation in eh channel.

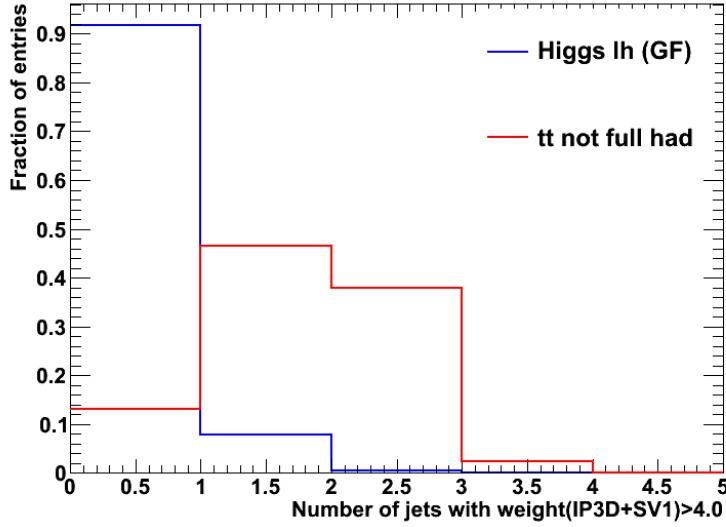


Figure 5.12: Number of jets in each event with a bjet weight greater than 4.0.

simulated on a $t\bar{t}$ sample that the selection of jets with a weight higher than 4.0 has a b-jet efficiency around 50%.

Since this weight is not computed for events produced by AtIfastII, this cut is *not* applied in QCD samples and in the privately produced $Z \rightarrow \tau^+\tau^- + 1j$ sample⁶. Fig. 5.12 illustrates the effectiveness of this cut for the μh channel.

5.5.3 Eta Tagging Jet

The η distribution (bottom plots of Fig. 5.7) of the two hard jets is significantly different between signal and backgrounds. For VBF events, this comes from the peculiar topology of the final state with two forward jets, but for gluon fusion events the reason is not the same clear. This feature is real, since the hardest jet should be generated at the partonic level and not in the parton shower, therefore the distribution is correct at the Leading Order and it is not affected by the approximations included in the parton shower. A possible explanation is that the processes that generate this jet in gluon fusion signal and backgrounds are different. Probably, in gluon fusion, the interacting gluons can mainly go through an Initial State Radiation (ISR) and the generated parton hadronizes in an hard jet. The emission in this case would be not very central. Instead, in background events, this jet might come really from the interaction of colliding partons, so its direction is closer to the plane transverse to the beam.

In order to reduce the background from $Z + 1, 2j$ and $W + 2, 3j$, we require that the rapidity of these hard jets has to be in the range $[1.5, 4.4]$. So events with very central jets are rejected.

5.5.4 Missing Energy, E_T^{miss}

In this analysis signal events are marked by the unbalance of the total energy in the transverse plane due to the double tau decay. This signature is stronger in case of leptonic decay where there are two neutrinos from the tau decay and they carry a more relevant fraction of tau energy with respect to the hadronic decay where there is only one neutrino.

⁶A study on the official full reco sample showed that the efficiency for this cut is high.

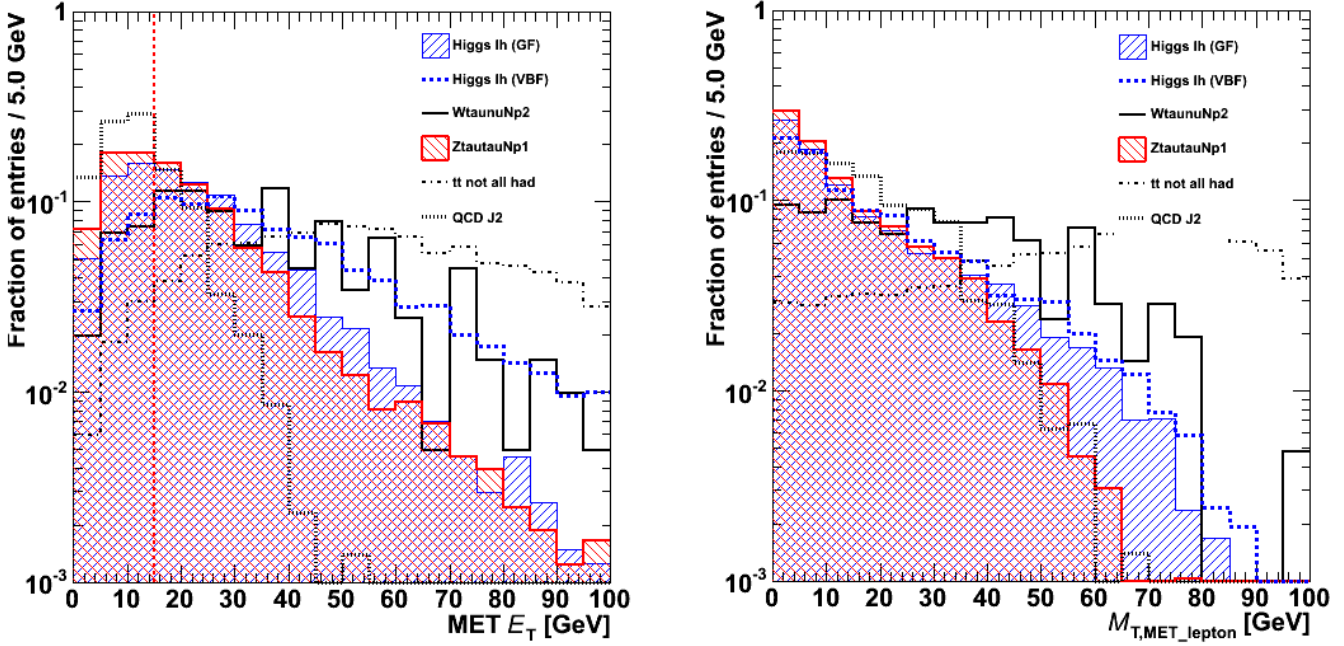


Figure 5.13: Distributions and thresholds on E_T^{miss} and $M_T(E_T^{\text{miss}}, \text{lep})$ in the μh channel.

Since the relevant backgrounds in each channel are not the same and they have different E_T^{miss} spectra, it is not possible to set the same threshold in all the channels. Nevertheless a common feature is that these thresholds should be as low as possible. Indeed, only QCD J2 and $Z \rightarrow e^+e^-, \mu^+\mu^-$ have softer E_T^{miss} distributions with respect to the gluon fusion signal. These considerations are also relevant for the choice of the trigger, especially in the hadronic channel.

The threshold applied is 15 GeV in the semileptonic channel (Fig. 5.13), 20 GeV in the leptonic channel in order to suppress $Z \rightarrow e^+e^-, \mu^+\mu^-$ events (Fig. 5.14) and 20 GeV in the hadronic channel as well in order to have a safe cut against QCD dijets J2 (Fig. 5.15).

5.5.5 Collinear Mass Approximation

The decay mode in a couple of tau leptons produces necessarily events with missing energy due to the presence of tau neutrinos. This means that it is not possible to reconstruct the mass of the Higgs boson directly from the visible objects and an approximation is needed. The standard method is called *Collinear Mass Approximation* and is based on the assumption that particles generated in the tau decay are emitted in a very narrow cone because of the momentum conservation. This hypothesis can be assumed safely since the tau system is heavily boosted. Knowing the four momentum of the two visible particles (e , μ or tau jet) and the projections of E_T^{miss} along x and y axes, it is possible to build a system of two equations which gives an estimation of the energies of the two tau leptons. So, the Higgs boson mass can be evaluated.

The drawback of this method is that the visible objects must not be emitted back-to-back. This requirement is needed otherwise the system doesn't have a unique solution. Let's take the case where both taus decay in leptons, even if the method is valid for all the decays. Each tau can be associated to the sum of the lepton and the neutrinos four-vectors. In the following we do not distinguish between ν_τ and ν_l since the purpose is to estimate the

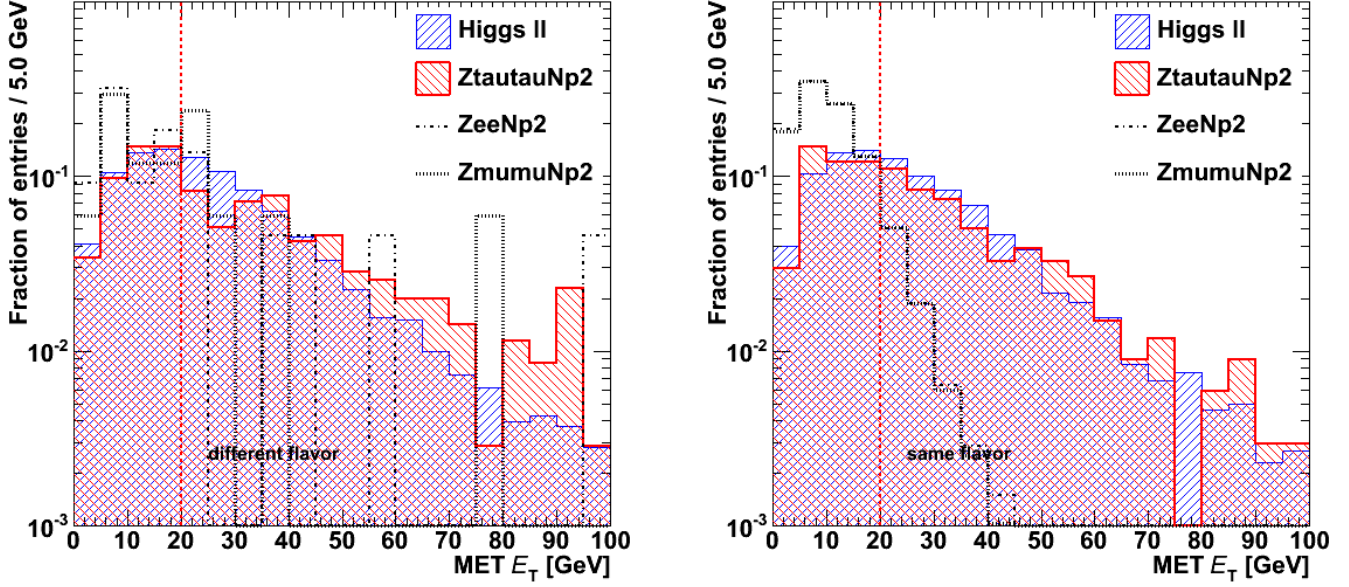


Figure 5.14: Distribution and threshold of E_T^{miss} in the ll channel.

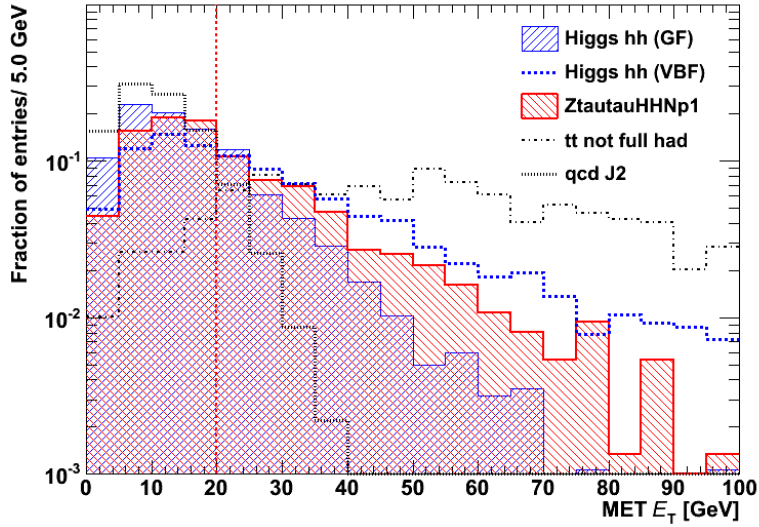


Figure 5.15: Distributions and thresholds on E_T^{miss} in the hh channel.

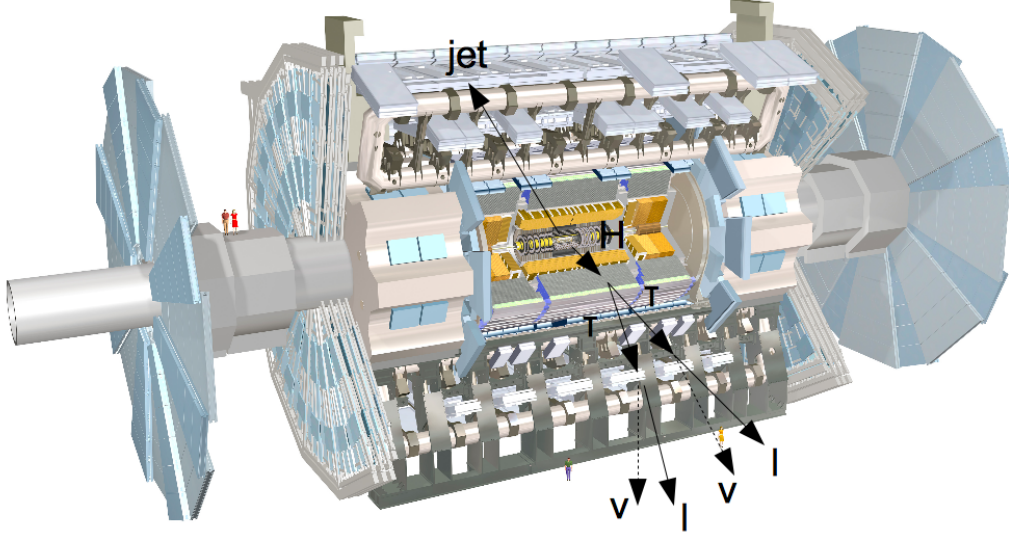


Figure 5.16: Illustration of a signal event where the Higgs mass can be reconstructed by the collinear mass approximation. The jet emitted in the Higgs opposite direction is the tagging jet selected in this analysis and reduces the probability that taus are produced back-to-back. The dashed arrows represent the sum of the tau and the leptonic neutrinos.

component of the total E_T^{miss} coming from each tau. So

$$p_{\tau,i} = p_{l,i} + p_{\nu,i} \quad (5.2)$$

$$p_{l,i} = (E_{l,i}, \vec{p}_{l,i}) \quad (5.3)$$

$$p_{\nu,i} = \left(E_{\nu,i}, E_{\nu,i} \frac{\vec{p}_{l,i}}{|\vec{p}_{l,i}|} \right) \quad (5.4)$$

where $i = 1, 2$ is the index for the two taus and $l = e, \mu$ according to the decay mode. Eq. (5.4) embodies the assumption that neutrinos fly in the direction of the leptons. For each event it is possible to measure the missing energy only along the x- and y-axis. Indeed, it is important to bear in mind that in hadron colliders, such as LHC, it is not possible to know the total momentum along the beam before the collision. Only the initial transverse momentum can be assumed to be zero. So, because of the momentum conservation, the total p_T in the final state must be zero as well. This means that we can estimate only two unknowns from the measurements of E_T^{miss} in the x and y axes and these unknowns are the energies carried by neutrinos in the two tau decays:

$$E_x^{\text{miss}} = E_{\nu,1,x} + E_{\nu,2,x} = E_{\nu,1} \frac{p_{l,1,x}}{|\vec{p}_{l,1}|} + E_{\nu,2} \frac{p_{l,2,x}}{|\vec{p}_{l,2}|} \quad (5.5)$$

$$E_y^{\text{miss}} = E_{\nu,1,y} + E_{\nu,2,y} = E_{\nu,1} \frac{p_{l,1,y}}{|\vec{p}_{l,1}|} + E_{\nu,2} \frac{p_{l,2,y}}{|\vec{p}_{l,2}|} \quad (5.6)$$

This system can be inverted in

$$E_{\nu,1} = |\vec{p}_{l,1}| \frac{+E_x^{\text{miss}} p_{l,2,y} - E_y^{\text{miss}} p_{l,2,x}}{p_{l,2,y} p_{l,1,x} - p_{l,2,x} p_{l,1,y}} \quad (5.7)$$

$$E_{\nu,2} = |\vec{p}_{l,2}| \frac{-E_x^{\text{miss}} p_{l,1,y} + E_y^{\text{miss}} p_{l,1,x}}{p_{l,2,y} p_{l,1,x} - p_{l,2,x} p_{l,1,y}} \quad (5.8)$$

As we mentioned before, in order to get two independent equations, taus cannot be back-to-back. It is straightforward to check this taking

$$\frac{p_{l,1,x}}{|p_{l,1}|} = -\frac{p_{l,2,x}}{|p_{l,2}|} \quad (5.9)$$

$$\frac{p_{l,1,y}}{|p_{l,1}|} = -\frac{p_{l,2,y}}{|p_{l,2}|} \quad (5.10)$$

The result is a couple of identical equations. So, a cut in the angle between the two leptons is needed to reject events where taus are emitted in opposite directions:

$$\cos(\Delta\phi_{ll}) > -0.9$$

This cut is illustrated in Fig. 5.17.

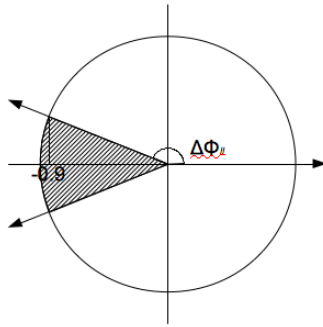


Figure 5.17: Cut on the angle between leptons. Events where a lepton is emitted in the dashed region are rejected.

Once the energies of the neutrinos are known, it is possible to estimate the four-vectors of each tau. We use the approximation that $m_\tau \approx 0$:

$$p_{\tau,i} = \left(E_{\nu,i} + E_{l,i}, (E_{\nu,i} + E_{l,i}) \frac{p_{l,i}}{|p_{l,i}|} \right) \quad (5.11)$$

It is necessary to check if each event is physically meaningful, i.e. that the energies carried by leptons are positive and less than the energies of the mother tau. This can be done taking x_i , which is the ratio of the energy of the visible objects to the estimated energy of the tau, to be less than one:

$$x_i = \frac{E_{l,i}}{E_{\tau,i}} = \frac{E_{l,i}}{E_{l,i} + E_{\nu,i}} < 1.0 \quad (5.12)$$

In order to reject background events like $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$ (Fig. 5.19) where there are no neutrinos, in leptonic decay x_i is required to be even smaller. So the thresholds in case of leptonic and hadronic decays are:

$$0 < x_l < 0.7 \quad (5.13)$$

$$0 < x_h < 1.0 \quad (5.14)$$

Finally, the invariant mass of the di-tau system which represents the Higgs candidate is:

$$m_{\tau^+\tau^-}^2 = 2E_{\tau,1}E_{\tau,2}(1 - \cos(\Delta\phi_{\tau^+\tau^-})) = 2(E_{\nu,1} + E_{l,1})(E_{\nu,2} + E_{l,2})(1 - \cos(\Delta\phi_{ll})) \quad (5.15)$$

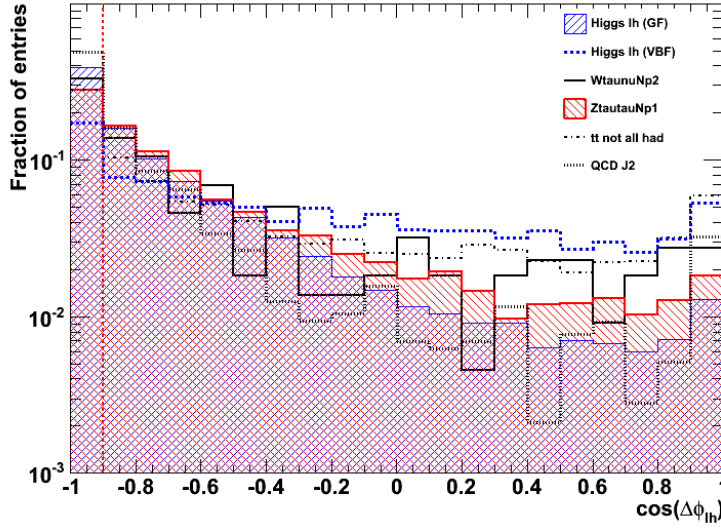


Figure 5.18: Distribution of $\cos(\Delta\phi_{lh})$ which is the angle between the muon and the tau jet in the μh channel.

This mass can also be expressed in terms of x_i as

$$m_{\tau^+\tau^-} = \frac{m_{ll}}{\sqrt{x_1 x_2}}$$

where m_{ll} is the invariant visible mass.

Usually the rejection of events with back-to-back taus affects significantly the signal statistics when the Higgs boson is produced at rest. But the selection of events with at least one jet in the final state used in this analysis reduces the probability of such topology. However the signal reconstruction efficiency is higher for the VBF events than for the gluon fusion. Even if there is one jet, the Higgs boson produced by gluon fusion is less boosted on average with respect to the VBF and the probability of a back-to-back decay is still high, as Fig. 5.18 shows.

This difference causes a 20% less efficiency in the gluon fusion events with respect to the VBF ones. The cuts on x_l and x_h affect the signals almost equally in the leptonic and semi-leptonic channels. Only in the hadronic channel there is a significant difference. Fig. 5.20 shows the distributions of the fractions of the tau energies carried by the leading and the subleading tau jets.

Two other mass reconstruction methods with higher efficiencies have been tested. The first one is the effective mass $m_{eff} = m_{vis} + E_T^{miss}$ and the second is the likely mass, which is the invariant mass of the di- τ system plus the E_T^{miss} 4-vector ($met_e, met_x, met_y, p_{z,\tau1} + p_{z,\tau2}$) where $met_e = E_T^{miss} \frac{E_{\tau1} + E_{\tau2}}{|p_{\tau1} + p_{\tau2}|}$. Even if the effective mass saves a lot of statistics (Fig. 5.21), it also accepts a lot of background events.

The mass reconstruction of the Higgs boson candidate in a pair of tau leptons will be improved in the future steps of this analysis thanks to new ideas that have been recently proposed as [60]. This new method has the great advantage of being able to reconstruct the mass also for back-to-back Higgs decay. This means that the gluon fusion signal can be enhanced and a higher signal yield can be obtained in the event selection.

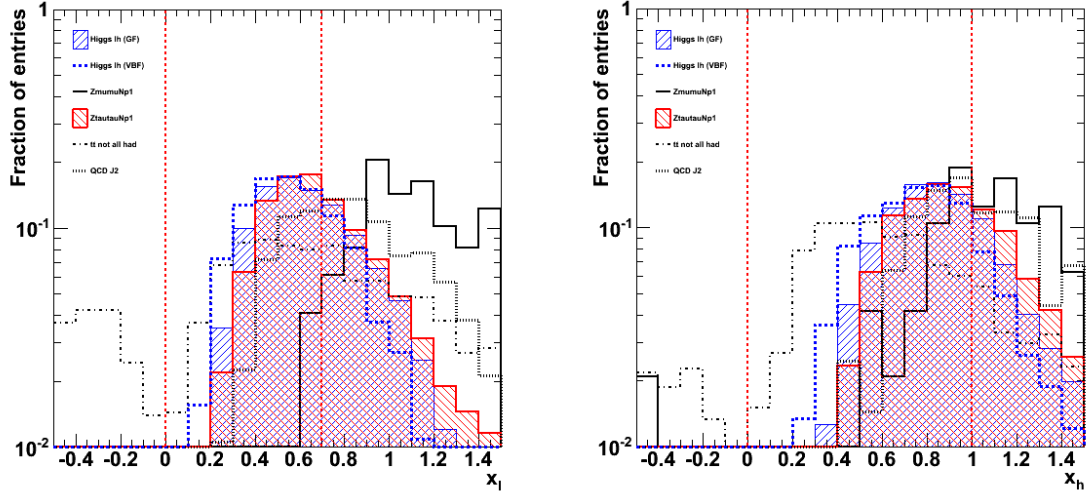


Figure 5.19: Distributions of the fractions of tau energy carried by leptons (left) and hadrons (right) in the μh channel.

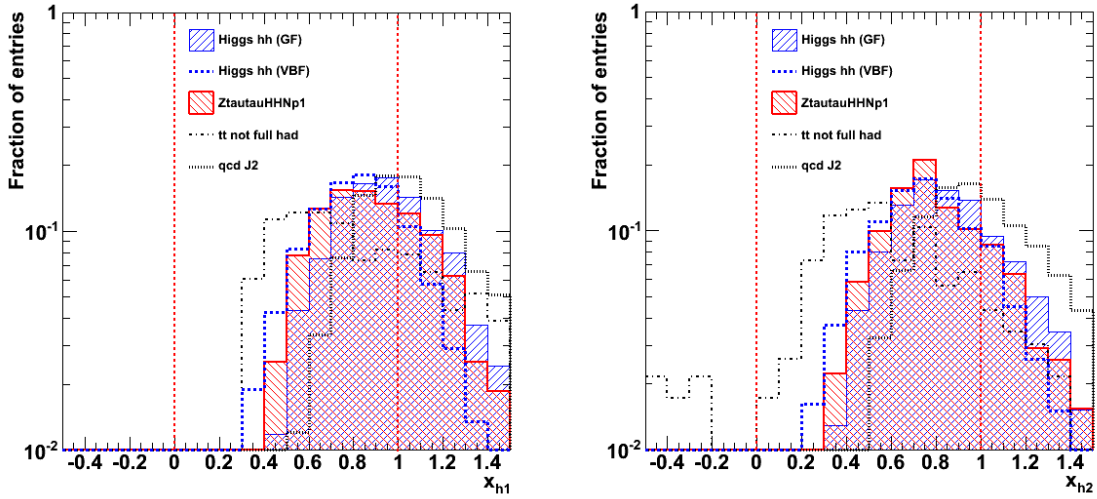


Figure 5.20: Distributions of the fractions of tau energy carried by the leading (left) and the subleading (right) tau jets in the hadronic channel. There is a significance difference between gluon fusion and VBF in the plot on the left.

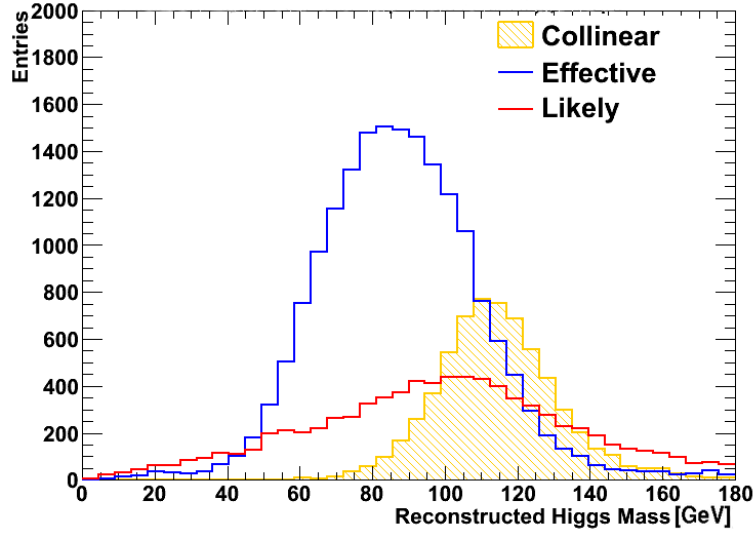


Figure 5.21: Different mass reconstruction methods for gluon fusion signal events in μh channel.

5.5.6 Visible Mass

The window applied on the visible mass is intended to be a further cleaning from events with a mass of the reconstructed Higgs candidate decay products in background events not coming from a resonance, like QCD and $W+j$ events. This cut additionally helps against $Z \rightarrow \tau^+\tau^-$ because of the softer spectrum of taus from Z with respect to H . In the semi-leptonic and in the hadronic channels the window is set in the range $[50,100]$ GeV (Fig. 5.22 and 5.23). In the leptonic channel, the window is at $[40,80]$ GeV in order to cut away the Z peak from $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ events (Fig. 5.24).

5.5.7 Invariant Mass $M(h, j)$

The final cut is on the invariant mass of the Higgs candidate and the tagging jet. The spectrum of this invariant mass is significantly different in signal and background events. As Fig. 5.25 shows, especially $Z+j$ and QCD events have a softer distribution. The threshold has been set to 600 GeV. Only in the eh channel, we prefer to keep the threshold lower at 500 GeV (Fig. 5.26).

5.5.8 Factorization

Tables 5.4 summarizes the selections applied in each channel.

Since the MC samples have very different luminosities, especially between signal and backgrounds, it is hard to reach the end of the cut flow with enough statistics for all of them. In order to be confident on the background estimation it is necessary to get some events in the end at least for the most important backgrounds. To do this, a factorization of the effects of subsets of cuts can be applied. We remind that the tau identification is already factorized for the QCD samples in the object selection and this approach with respect to applying tau identification directly on the sample has been studied and proven to be giving consistent results. The cut factorization is performed splitting the cut flow in three parts (Table 5.5) and it is applied on all backgrounds, but not on signal samples.

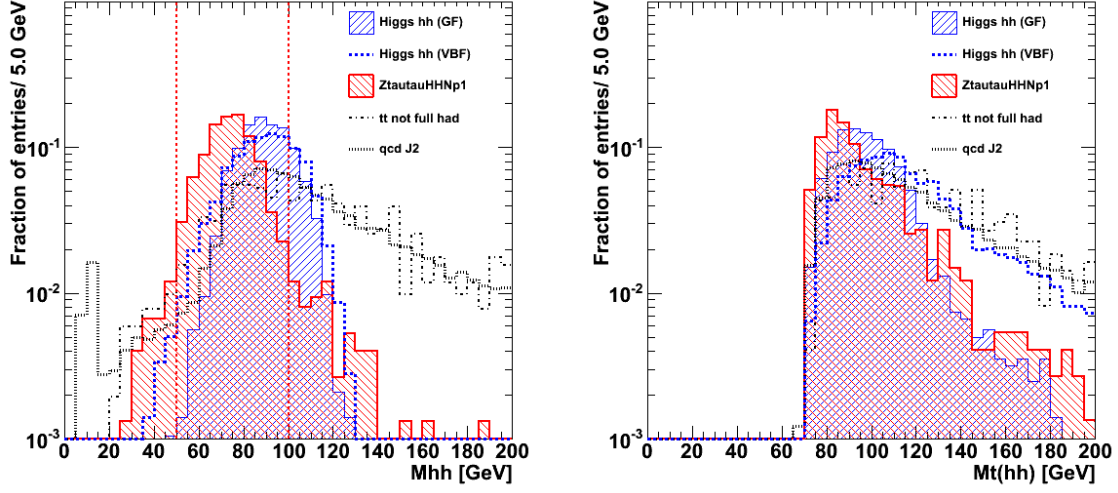


Figure 5.22: Invariant visible mass and transverse visible mass in the hadronic channel. The suppression of $Mt(hh)$ under 70 GeV is due to the p_T cut on the selected tau jets.

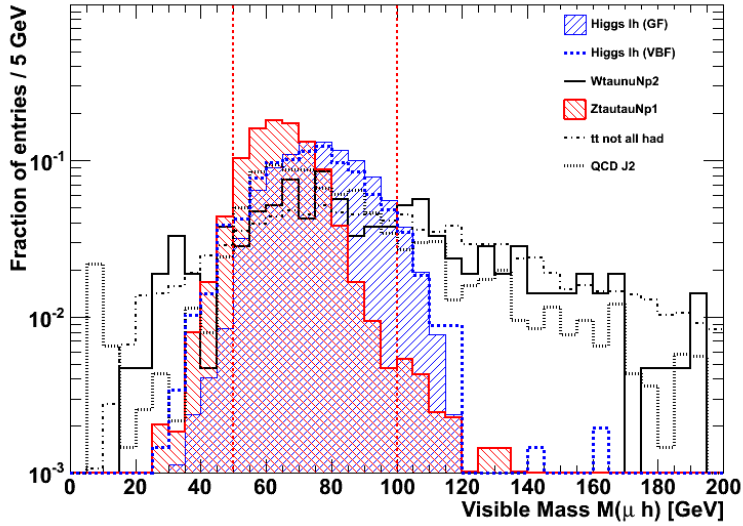
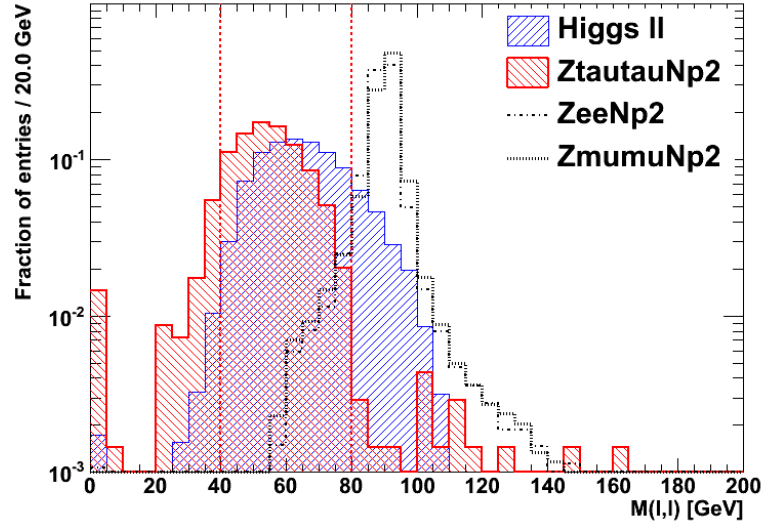
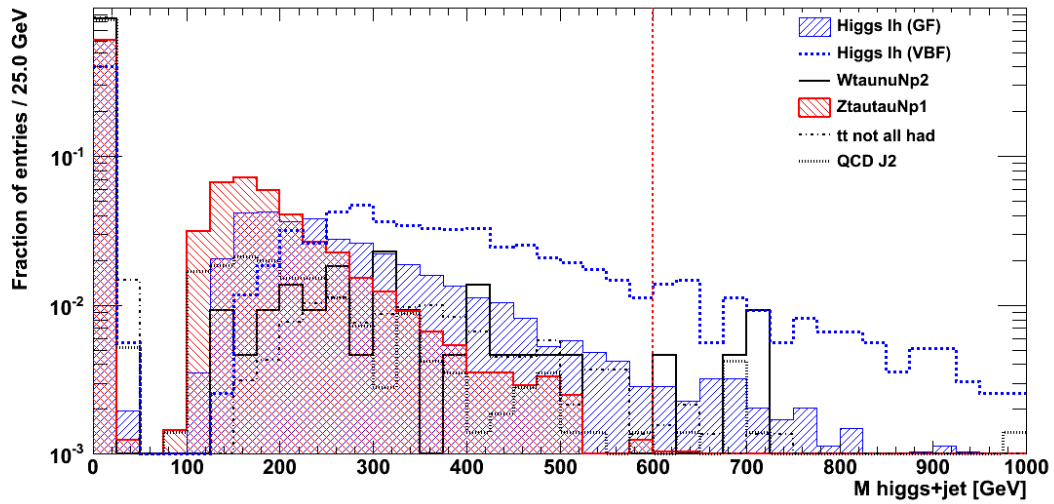


Figure 5.23: Invariant visible mass in the μh channel.

Figure 5.24: Invariant visible mass in the ll channel.Figure 5.25: Invariant mass distribution of the tagging jet plus the Higgs candidate in the μh analysis.

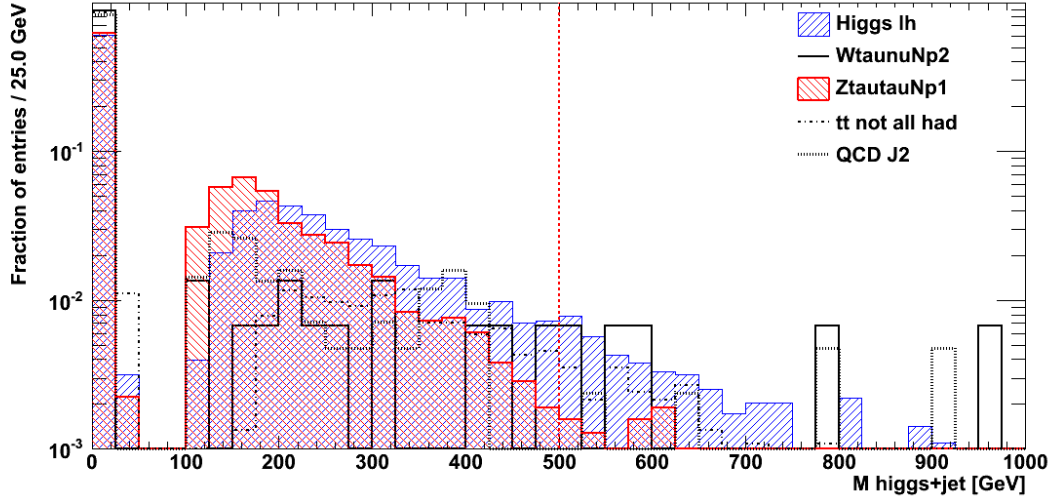


Figure 5.26: Invariant mass distribution of the tagging jet plus the Higgs candidate in the eh analysis.

We compute the selection efficiency for the two first sets with respect to the number of events that are accepted in the object selection. Finally we apply the third set weighting each event with the efficiency previously estimated. The error associated to this procedure is only statistical, meaning that the precision is $1/\sqrt{N_{accepted}}$. It should be noted that this factorization approach can only overestimate the backgrounds and is therefore conservative, due to the fact that the set of “Final cuts” includes jets with lower η (hence higher p_T), which make the events more likely to pass the selection for background events like QCD dijets. The consistency of this method has been tested with the $Z \rightarrow \tau^+\tau^- + 2j$ sample, where the yields at each step of the event selection without factorization are compatible within errors with the yields with the factorization.

5.5.9 Weights and Errors

Each event is weighted by the efficiency of the factorized cuts. So errors are the combination of $\sqrt{N_{ev}}$ and the error on the efficiency factor, which is only statistical as well. They are scaled by the total cross section and divided by the initial number of events in the MC sample.

For the QCD samples, events have an additional weighting factor due to the fact that tau ID is also factorized. Indeed, there can be multiple combinations of taus that fulfill a set of cuts. Therefore it is necessary to consider the possibility of picking up one good tau-lepton or tau-tau pair among all the combinations. So, in the semileptonic analysis, the weight of each event is computed as

$$\omega_{ev} = \frac{1}{n_{\text{all comb}}} \sum_{\text{good comb}} \omega_{\tau} \quad (5.16)$$

where $n_{\text{all comb}}$ is the total number of tau-lepton pairs among the selected objects and $\sum_{\text{good comb}}$ is the sum of weights ω_{τ} defined by the histograms in Appendix C. The order of magnitude of these weights is the permille.

Cut Name	Description
Charge	The leptons and the τ -jets selected are required to have opposite charges
Tau Tracks	The number of tracks associated to the τ -jet has to be 1 or 3
Isolation	The energy deposited in the calorimeter in a cone of radius $\Delta R < 0.2$ has to be less than 4 GeV. In case a muon is selected, then the additional requirement of no other tracks in a cone of radius $\Delta R < 0.2$ is applied
BJet	There must be no jets with a IP3D+SV1 weight greater than 4.0
EtaJet	The hardest and eventually the second hardest jet ($p_T > 20$ GeV) have to be in the range of rapidity $1.5 < \eta < 4.4$
MET	$E_T^{\text{miss}} > 15$ GeV (lh), 20 GeV (ll and hh)
Coll	$\Delta\phi > -0.9$ and the fractions of energies carried by visible objects have to be in the ranges $0 < x_h < 1.0$ for hadronic taus and $0 < x_l < 0.75$ for leptonic taus
Visible	The visible mass has to be in the window [50,100] GeV (lh, hh) and [40,80] GeV (ll)
Mhj	The invariant mass of the Higgs candidate plus the tagging jet is required to be greater than 600 GeV (500 GeV in the eh channel)
Mass Window	[110,140] GeV

Table 5.4: Summary of the selection applied in the ll , lh and hh channels.

Correspondingly, in the hadronic channel,

$$\omega_{ev} = \frac{1}{n_{\text{all comb}}} \sum_{\text{good comb}} \omega_{\tau,1} \omega_{\tau,2} \quad (5.17)$$

in this case $n_{\text{all comb}}$ can be considered as the binomial factor $\binom{n_\tau}{2}$. In the hadronic channel $n_{\text{all comb}}$ is not the total number of combinations among the selected tau jets, but the number of couples of taus with 1 or 3 tracks and opposite charges⁷. This means that $n_{\text{all comb}}$ is smaller than what it should be and consequently the weights, and therefore the QCD backgrounds, are overestimated.

Moreover, the QCD estimates for the semi-leptonic and hadronic channels are conservative since with this procedure we don't require that the other tau candidates would not pass the tau identification. We only consider the probability that the chosen tau can be identified as a good one. On the contrary, in all other MC samples we take events where there is exactly one (or two) identified tau(s). Due to this, the QCD background is overestimated. An example is the following: if in a QCD event there are one lepton and two tau jets selected, then the weight is computed as

$$\omega_{ev} = \frac{1 \cdot \omega_{\tau,1} + 1 \cdot \omega_{\tau,2}}{2} > \frac{1 \cdot \omega_{\tau,1} \cdot (1 - \omega_{\tau,2}) + 1 \cdot \omega_{\tau,2} \cdot (1 - \omega_{\tau,1})}{2} \quad (5.18)$$

The second term of the disequation takes into account the probability that exactly one tau jet is identified in the event and it gives weights smaller than the ones used in this

⁷This choice reduces the computational time.

Cut sets	Description
Initial cuts	Charge
	Tau Track
	Isolation
	BJet Veto
Eta Jets	Eta jets
Final cuts	E_T^{miss}
	Collinear Mass
	Visible Mass
	$M(h, j)$
	Mass Window

Table 5.5: Cut flow factorization in three sets of cuts.

analysis. The errors on the weights associated to each tau jets in the QCD events are of the order of 2-3%. They have been neglected in the computation of the final results since the corresponding errors on the event weights are of the same order, or even smaller in case of multiple combinations, and they are tiny compared to the statistic uncertainties.

5.6 Results

In this section we report the summary of the signal and background yields in each channel (Tables 5.6, 5.7, 5.8 and 5.9). More detailed informations are provided in Appendix A. In the last two columns of each table we estimate the sensitivity for an integrated luminosity of 30 fb^{-1} at the center of mass energy of 10 TeV with and without systematic uncertainties on the background. In the second column from the right we simply report the ratio $s/\sqrt{b}$, which is a rough estimation of the signal yield in units of background fluctuations. Indeed, if the background is poisson distributed, then the variance is the background yield itself. In order to take into account systematic uncertainties on the background estimation, in the first column from the right we evaluate the significance using the formula implemented in RooStat [61], which is a statistical tool developed within ROOT. This formula computes the significance in 1-sided gaussian standard deviation in a number counting experiment for the signal-plus-background hypothesis against the only-background hypothesis. Its code is based on [62, 63, 64] and it allows to take into account a relative uncertainty on the background estimate. We set this uncertainty to 0.1 which is the same value used in the $H \rightarrow \tau^+\tau^-$ baseline analysis [43].

As explained in the captions of Tables 5.6-5.9, the factorized cuts are labelled with a ‘*’.

Cut	Signal (with VBF) [fb]	W + jets [fb]	Z + jets [fb]	top [fb]	QCD [fb]	Total Bg [fb]	$s/\sqrt{b}$ for 30 fb^{-1}	σ for 30 fb^{-1} (10% sys on b)
Start	$1\text{E}+03\pm 1.4$	$4.8\text{E}+07\pm 2\text{E}+04$	$4.4\text{E}+06\pm 4\text{E}+03$	$4.2\text{E}+05\pm 2\text{E}+02$	$5.9\text{E}+10\pm 9\text{E}+06$	$5.9\text{E}+10\pm 0.08\text{E}+08$	0.023	0.065
Selection	22 ± 0.2	$1.3\text{E}+04\pm 0.2\text{E}+03$	$3.9\text{E}+03\pm 0.1\text{E}+3$	$1\text{E}+03\pm 10$	$7.2\text{E}+04\pm 6\text{E}+02$	$8.9\text{E}+04\pm 0.7\text{E}+04$	0.41	0.064
Charge*	20 ± 0.2	$5.6\text{E}+03\pm 1\text{E}+02$	$3\text{E}+03\pm 60$	$6.7\text{E}+02\pm 9$	$1.8\text{E}+04\pm 3\text{E}+02$	$2.7\text{E}+04\pm 0.2\text{E}+04$	0.68	0.059
TauTrack*	20 ± 0.2	$5.3\text{E}+03\pm 0.1\text{E}+03$	$3\text{E}+03\pm 60$	$6.5\text{E}+02\pm 9$	$1.1\text{E}+04\pm 2\text{E}+02$	$2\text{E}+04\pm 0.2\text{E}+04$	0.8	0.056
Isolation*	19 ± 0.2	$4.9\text{E}+03\pm 0.1\text{E}+03$	$2.8\text{E}+03\pm 60$	$5.5\text{E}+02\pm 8$	$4.4\text{E}+02\pm 60$	$8.7\text{E}+03\pm 0.2\text{E}+03$	1.1	0.045
BJet*	19 ± 0.2	$4.8\text{E}+03\pm 100$	$2.8\text{E}+03\pm 60$	$1.7\text{E}+02\pm 5$	$4.4\text{E}+02\pm 60$	$8.2\text{E}+03\pm 0.2\text{E}+03$	1.1	0.044
EtaJet*	9.1 ± 0.1	$1.6\text{E}+03\pm 50$	$8.3\text{E}+02\pm 40$	26 ± 2	$1.4\text{E}+02\pm 20$	$2.6\text{E}+03\pm 80$	0.97	0.032
Met	6.2 ± 0.1	$1.5\text{E}+03\pm 54$	$4.1\text{E}+02\pm 20$	24 ± 2	47 ± 6	$2\text{E}+03\pm 60$	0.77	0.035
Coll	2.6 ± 0.07	48 ± 3	$1.2\text{E}+02\pm 5$	2.2 ± 0.2	6.7 ± 0.9	$1.7\text{E}+02\pm 6$	1.1	0.084
mlh	2.5 ± 0.07	24 ± 2	85 ± 4	1 ± 0.1	4.4 ± 0.6	$1.1\text{E}+02\pm 5$	1.3	0.15
Mhj	0.45 ± 0.03	2 ± 0.5	1.8 ± 0.3	0.1 ± 0.03	0.21 ± 0.07	4.1 ± 0.6	1.2	0.72
Window	0.33 ± 0.02	0.2 ± 0.2	0.21 ± 0.12	0.005 ± 0.003	0.08 ± 0.05	0.5 ± 0.2	2.5	2.1

Table 5.6: μh -channel summary cut flow. (*) factorized cuts and errors are only statistical. The yields in the last row include additional contributions estimated in Section 5.6.

Cut	Signal (with VBF) [fb]	W + jets [fb]	Z + jets [fb]	top [fb]	QCD [fb]	Total Bg [fb]	$s/\sqrt{b}$ for 30 fb^{-1}	σ for 30 fb^{-1} (10% sys on b)
Start	$1\text{E}+03\pm 1.4$	$4.8\text{E}+07\pm 2\text{E}+04$	$4.4\text{E}+06\pm 4\text{E}+03$	$4.2\text{E}+05\pm 3\text{E}+02$	$5.9\text{E}+10\pm 9\text{E}+06$	$5.9\text{E}+10\pm 7\text{E}+08$	0.023	0.065
Selection	16 ± 0.2	$8.9\text{E}+03\pm 0.1\text{E}+03$	$1\text{E}+04\pm 1\text{E}+02$	$6.9\text{E}+02\pm 9$	$1.3\text{E}+04\pm 3\text{E}+02$	$3.3\text{E}+04\pm 0.2\text{E}+04$	0.49	0.062
Charge*	15 ± 0.2	$4.2\text{E}+03\pm 90$	$6.9\text{E}+03\pm 0.1\text{E}+03$	$4.8\text{E}+02\pm 8$	$3.6\text{E}+03\pm 0.2\text{E}+03$	$1.5\text{E}+04\pm 0.8\text{E}+04$	0.66	0.057
TauTrack*	15 ± 0.2	$3.9\text{E}+03\pm 90$	$6.9\text{E}+03\pm 0.1\text{E}+03$	$4.6\text{E}+02\pm 7$	$2.3\text{E}+03\pm 0.1\text{E}+03$	$1.4\text{E}+04\pm 0.6\text{E}+04$	0.69	0.056
Isolation*	14 ± 0.2	$3.7\text{E}+03\pm 80$	$6.7\text{E}+03\pm 0.1\text{E}+03$	$4.3\text{E}+02\pm 7$	$1.1\text{E}+03\pm 0.1\text{E}+03$	$1.2\text{E}+04\pm 0.3\text{E}+04$	0.72	0.055
BJet*	14 ± 0.2	$3.6\text{E}+03\pm 80$	$6.6\text{E}+03\pm 0.1\text{E}+03$	$1.4\text{E}+02\pm 4$	$1.1\text{E}+03\pm 0.1\text{E}+03$	$1.1\text{E}+04\pm 0.3\text{E}+04$	0.72	0.054
EtaJet*	6.8 ± 0.1	$1.2\text{E}+03\pm 40$	$2.2\text{E}+03\pm 90$	19 ± 1	$3.5\text{E}+02\pm 30$	$3.8\text{E}+03\pm 0.1\text{E}+03$	0.61	0.049
Met	4.9 ± 0.1	$1.1\text{E}+03\pm 50$	$5.5\text{E}+02\pm 30$	18 ± 1	$1.1\text{E}+02\pm 10$	$1.8\text{E}+03\pm 70$	0.63	0.039
Coll	2.1 ± 0.06	37 ± 3	85 ± 4	1.6 ± 0.1	17 ± 2	$1.4\text{E}+02\pm 8$	0.98	0.082
Visible	2 ± 0.06	18 ± 2	62 ± 4	0.67 ± 0.06	11 ± 2	92 ± 6	1.1	0.15
Mhj	0.57 ± 0.03	2.1 ± 0.6	1.8 ± 0.4	0.1 ± 0.02	1.6 ± 0.6	5.6 ± 1.2	1.3	0.72
Window	0.43 ± 0.03	0.5 ± 0.3	0.5 ± 0.2	0.004 ± 0.003	0.34 ± 0.25	1.3 ± 0.5	2.1	1.6

Table 5.7: eh -channel summary cut flow. (*) factorized cuts and errors are only statistical.

Cut	Signal (with VBF) [fb]	W + jets [fb]	Z + jets [fb]	top [fb]	QCD [fb]	Total Bg [fb]	$s/\sqrt{b}$ for $30 fb^{-1}$	σ for $30 fb^{-1}$ (10% sys on b)
Start	$1.6E+02 \pm 0.3$	$4.8E+07 \pm 2E+04$	$4.4E+06 \pm 4E+03$	$4.1E+05 \pm 2E+02$	$5.9E+10 \pm 9E+06$	$5.9E+10 \pm 7E+08$	0.0035	0.074
Selection	9.5 ± 0.08	$4.6E+03 \pm 0.2E+03$	$2.3E+05 \pm 9E+02$	$4.9E+03 \pm 20$	$1E+05 \pm 9E+03$	$3.4E+05 \pm 0.3E+05$	0.09	0.066
Charge*	9.5 ± 0.08	$2.5E+03 \pm 0.1E+03$	$2.3E+05 \pm 9E+02$	$4.3E+03 \pm 20$	$7.4E+04 \pm 0.8E+04$	$3.1E+05 \pm 0.2E+05$	0.094	0.066
Isolation*	8.7 ± 0.07	$1.2E+03 \pm 90$	$2.1E+05 \pm 8E+02$	$3.1E+03 \pm 20$	$7E+03 \pm 3E+03$	$2.2E+05 \pm 5E+03$	0.1	0.066
Bjet*	8.5 ± 0.07	$1.1E+03 \pm 90$	$2E+05 \pm 8E+02$	$8.9E+02 \pm 9$	$7E+03 \pm 3E+03$	$2.1E+05 \pm 5E+03$	0.1	0.066
EtaJet*	4.1 ± 0.05	$4.2E+02 \pm 60$	$6.2E+04 \pm 0.6E+04$	$1.2E+02 \pm 2$	$3E+03 \pm 2E+03$	$6.5E+04 \pm 2E+04$	0.088	0.066
Met	2.6 ± 0.04	$3.8E+02 \pm 60$	$3.3E+03 \pm 60$	$1.1E+02 \pm 2$	$2.3E+02 \pm 0.2E+03$	$4E+03 \pm 0.2E+03$	0.22	0.06
Coll	1.1 ± 0.03	1.1 ± 0.3	$1.2E+02 \pm 8$	7.3 ± 0.2	4 ± 5	$1.3E+02 \pm 10$	0.54	0.019
Visible	1 ± 0.02	0.48 ± 0.17	52 ± 4	2.8 ± 0.1	2 ± 2.7	58 ± 5	0.73	0.1
Mhj	0.17 ± 0.01	0.13 ± 0.08	0.82 ± 0.25	0.28 ± 0.03	0 ± 0	1.2 ± 0.3	0.83	0.61
Window	0.13 ± 0.01	0 ± 0	0.1 ± 0.1	0.016 ± 0.008	0 ± 0	0.11 ± 0.09	2.1	1.7

Table 5.8: leptonic channel summary cut flow. (*) factorized cuts and errors are only statistical. The yields in the last row include additional contributions estimated in Section 5.6.

Cut	Signal (with VBF) [fb]	W + jets [fb]	Z + jets [fb]	top [fb]	QCD [fb]	Total Bg [fb]	$s/\sqrt{b}$ for $30 fb^{-1}$	σ for $30 fb^{-1}$ (10% sys on b)
Start	$7.4E+02 \pm 2$	$4.8E+07 \pm 2E+04$	$3E+06 \pm 3E+03$	$4.2E+05 \pm 2E+02$	$5.9E+10 \pm 3E+07$	$5.9E+10 \pm 0.1E+010$	0.017	0.062
Selection	28 ± 0.4	$2.2E+03 \pm 60$	$1.8E+03 \pm 40$	$1.5E+02 \pm 4$	$7.6E+08 \pm 3E+06$	$7.6E+08 \pm 0.5E+08$	0.0056	0.067
Charge*	25 ± 0.4	$9.1E+02 \pm 40$	$1.4E+03 \pm 30$	76 ± 3	$1.9E+03 \pm 10$	$4.2E+03 \pm 0.1E+03$	2.1	0.0085
TauTrack*	24 ± 0.4	$8.1E+02 \pm 40$	$1.3E+03 \pm 30$	70 ± 3	$1.9E+03 \pm 10$	$4.1E+03 \pm 0.1E+03$	2.1	0.0071
BJet*	24 ± 0.4	$7.8E+02 \pm 30$	$1.3E+03 \pm 30$	23 ± 2	$1.9E+03 \pm 10$	$4E+03 \pm 0.1E+03$	2.1	0.0068
EtaJet*	12 ± 0.3	$2.6E+02 \pm 20$	$3.5E+02 \pm 20$	3 ± 0.4	$3.2E+02 \pm 4$	$9.3E+02 \pm 50$	2.1	0.058
Met	3.9 ± 0.14	$2E+02 \pm 15$	$1.2E+02 \pm 7$	2.7 ± 0.3	13 ± 0.22	$3.4E+02 \pm 20$	1.2	0.05
Coll	1.8 ± 0.09	17 ± 2	54 ± 4	0.41 ± 0.06	13 ± 0.6	85 ± 5	1.1	0.14
Mhj	0.48 ± 0.04	3.4 ± 0.7	3.7 ± 0.6	0.05 ± 0.01	1.5 ± 0.2	9 ± 1	0.9	0.4
Visible	0.45 ± 0.04	0.7 ± 0.3	3 ± 0.5	0.007 ± 0.004	0.3 ± 0.08	4 ± 0.6	1.2	0.75
Window	0.35 ± 0.03	0 ± 0	0.4 ± 0.2	0.003 ± 0.002	0.14 ± 0.05	0.5 ± 0.2	2.6	2.1

Table 5.9: hh -channel summary cut flow. (*) are factorized cuts and errors are only statistical. In the QCD estimation the cuts ‘Charge*’ and ‘TauTrack*’ are applied at the same time.

In order to be confident about the background estimations, we tried to loosen the event selection. Indeed it is necessary to give an estimations of the backgrounds that cannot reach the end of the cut flow because of the small statistics. All the following considerations are documented in the tables of Appendix A.

In the eh channel, thanks to the softer cut on $M(h, j)$, all the relevant backgrounds can survive at the event selection with some events. Moreover, lowering the threshold on $M(h, j)$ from 500 to 400 GeV, the sensitivity of this channel doesn't change. So the background estimation in the eh channel is appropriate. In the leptonic channel we tried to lower the threshold on $M(h, j)$ as well. This time from 600 to 500 GeV. The result is that the sensitivity goes from 1.8 to 1.6. The background from $W + j$ is negligible, but a contribution of 0.03 fb from $Z \rightarrow \tau^+\tau^- + 1, 3j$ has to be added on the total $Z + j$ background. This quantity has been estimated applying to the yields with the $M(h, j)$ threshold at 500 GeV the efficiency of raising this threshold to 600 GeV. In the hadronic channel we think that the background estimation is reasonable without any additional contribution since the sensitivity doesn't change if we loose the event selection asking $M(h, j)$ above 500 GeV, and even if the threshold is at 400 GeV the sensitivity only goes from 2.1 to 1.9. In the μh channel we tried to loosen the event selection as we did for the other channels moving the lower limit on $M(h, j)$ from 600 to 500 GeV. For the $Z + j$ background we can estimate an additional contribution of 0.1 fb from $Z \rightarrow \tau^+\tau^- + 1, 3j$ exactly in the same way as for the leptonic channel. The background from $W + j$, specifically $W \rightarrow \mu\nu, \tau\nu + 2, 3j$, is more difficult to evaluate. Since the statistics is very low, we tried to loosen the object selection requiring events with one tau identified as 'Loose' and not 'Medium'. This doubles the statistics at the beginning of the event selection, but in the end we get only one event from $W \rightarrow \mu\nu + 3j$ as in the baseline selection. So we loosened also the $M(h, j)$ threshold from 600 to 500 GeV and we computed the efficiencies (see tables in Appendix. A). The outcome is that in the selection with 'Loose' taus we would need to add 0.2 fb, which corresponds to about 0.1 fb in the baseline event selection with 'Medium' taus. We think that this estimation is fair and if in the future, with a better knowledge on this background maybe from data, it turns out that this yield is too high it is still possible to gain a factor of ~ 4 using 'Tight' taus without a significant loss in the signal.

More detailed comments are needed on the signal yield in the hadronic channel. Looking at the tables in Appendix A, we can see that the gluon fusion signal is very small compared to the VBF. The main differences in the efficiencies are in the E_T^{miss} , $M(h, j)$ and in the collinear mass cuts. The higher threshold on the missing energy, although helpful to kill the QCD background, affects greatly the gluon fusion signal. Moreover the collinear approximation, which strongly suppresses the gluon fusion signal in all the channels, in the hadronic final states has the worst performance. It is hard to save a higher fraction of gluon fusion signal. If the E_T^{miss} threshold is set at 15 GeV, the gluon fusion is still one third of the VBF signal and the sensitivity drops from 2.1 to 1.7 because of the increase of QCD backgrounds. Probably a cut-based selection is not the proper strategy since all the hard cuts on tau p_T , E_T^{miss} and $M(h, j)$ are more efficient for the VBF events. A multivariate technique can for sure select more gluon fusion events. Moreover, as we mention before, the collinear approximation works better for the VBF than for the gluon fusion events. This issue has to be studied further, but an improvement can be gained trying to use a different mass reconstruction recently suggested in [60]. This new method is based on a likelihood minimization and it doesn't have the limitations of the collinear approximation. Indeed it can reconstruct also back-to-back decays which are the majority of the gluon fusion events. The hadronic channel should have the best performances with this technique compared to the other channels. Moreover, the authors claim that the invariant mass distribution doesn't have a long tail as for the collinear approximation and this is also an advantage for the gluon

Energy	Signal [fb]	Total Bg [fb]	$s/\sqrt{b}$ $30 fb^{-1}$	σ (10% syst on bg) $30 fb^{-1}$
10 TeV	1.20 ± 0.05	2.6 ± 0.7	4.6	3.6
14 TeV	2.20 ± 0.10	4.4 ± 1.1	6.0	4.6
7 TeV	0.67 ± 0.03	1.5 ± 0.5	3.3	2.6

Table 5.10: Combined analysis at 7, 10 and 14 TeV.

channel	μh	$e h$	$l l$	$h h$
trigger	tau16_loose_mu15	tau16_loose_e15	double lepton e/mu10	tau29_loose_tau20_loose
signal	90%	90%	100%	80%
qcd	60%	60%	100%	40%
sensitivity	2.1→2.0	1.6→1.5	1.8→1.8	2.1→2.0

Table 5.11: Effects of trigger efficiencies.

fusion signal.

5.7 Combined Results

In this section we report the combined results of all the channels. In Table 5.10 we sum all the signal and background yields at the bottom of each cut flow. The final sensitivity is the sum in quadrature of the sensitivity for each channel. We want to mention again that the signal is likely to be underestimated because of the absence of one channel in the VBF production (Section 5.1). At the same time the QCD background is overestimated since we don't apply the bjet veto and the weights are computed with some approximations that make them higher (Section 5.5.9).

Since the 10 TeV run is no more scheduled⁸, we report in Table 5.10 results for the center of mass energies of 7, 10 and 14 TeV. The extrapolations at 7 and 14 TeV are made through a rescaling of the total cross sections of each sample. Different scenarios of integrated luminosity are shown in Fig. 5.27.

These values are not affected by the application of trigger efficiency. It is possible to give an educated guess applying the efficiencies listed in Table 5.11. This estimation is conservative since the same values are taken for both signal and W/Z backgrounds. From the table it is clear that the only effect is a decrease of the expected number of events.

It is interesting to show also the collinear mass distributions of the di-tau system. Since the number of events left at the end of the cut flows is very poor, it is necessary to factorize the shape of the backgrounds at an earlier stage of the selection. Specifically, when indicated in the Figs. 5.28, 5.29, 5.30 and 5.31, $Z+j$, $W+j$ shapes are estimated applying only the cuts 'Charge', 'TauTrack', 'Met' and 'Coll'. For the QCD shape we even don't apply the 'Charge' cut. We don't expect that the other cuts which are not applied can change significantly the shape of the background because they involve variables linked to the Higgs candidate and the tagging jet and not the two taus separately. Once the shape is determined, then it is scaled to the actual yield in the mass window.

⁸The LHC plans changed during 2010 and when we started working on this analysis the MC samples at 10 TeV were the ones with the higher center of mass energy available.

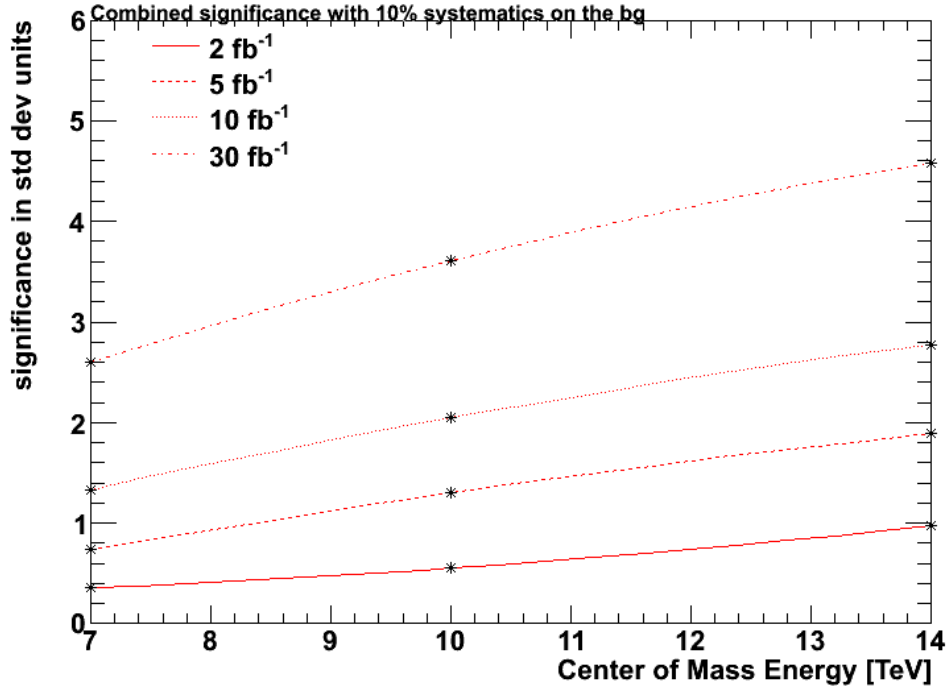


Figure 5.27: Combined significance at the center of mass energy of 7, 10 and 14 TeV for different integrated luminosities.

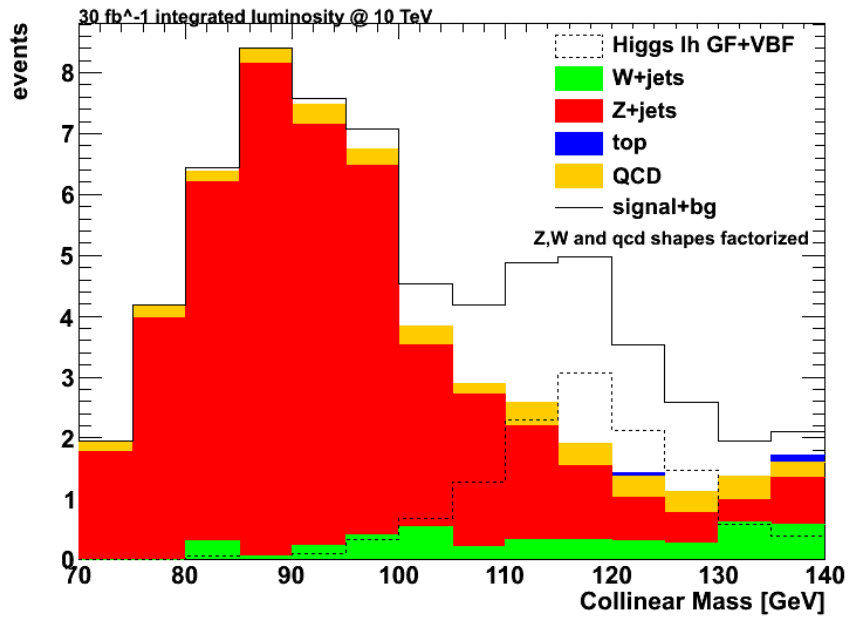


Figure 5.28: Collinear mass distribution at the end of the cut flow in the μh channel.

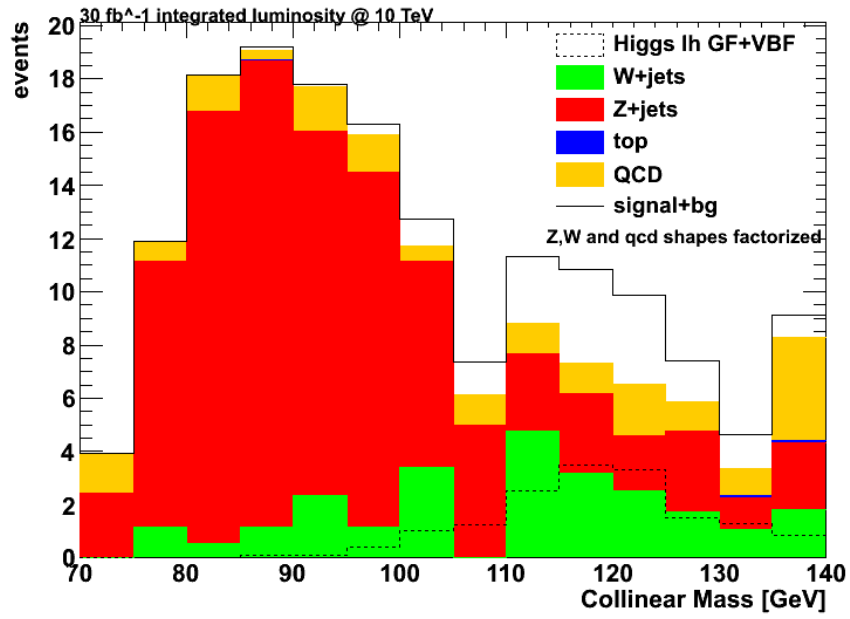
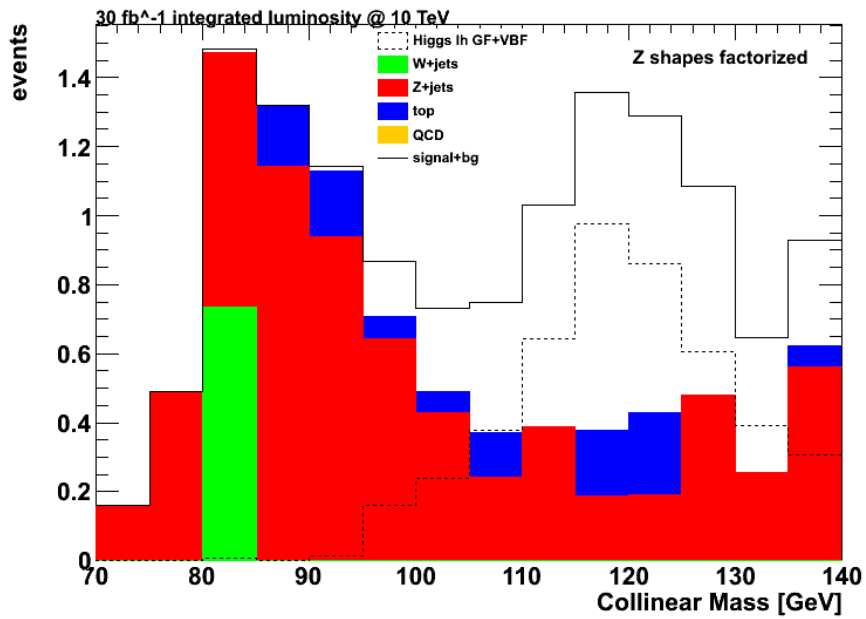
Figure 5.29: Collinear mass distribution at the end of the cut flow in the eh channel.

Figure 5.30: Collinear mass distribution at the end of the cut flow in the leptonic channel.

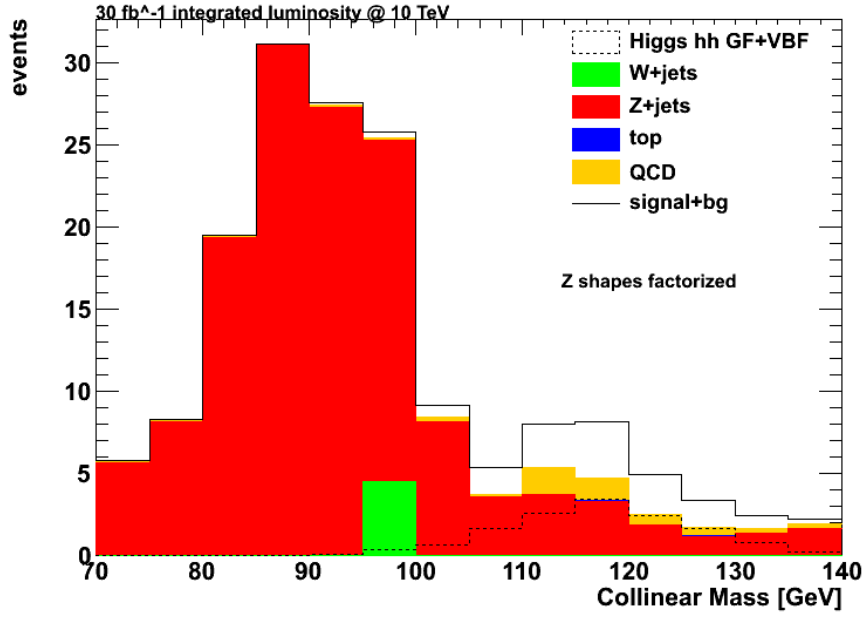


Figure 5.31: Collinear mass distribution at the end of the cut flow in the hadronic channel.

5.8 gg and VBF Comparison

The main aim of this study is to develop a new analysis that can be combined with the baseline analysis based only on the VBF signal in order to increase the sensitivity for the SM Higgs boson in the low mass range. Since we include both the gluon fusion and the VBF as signal production processes, it is necessary to know which is the fraction of events that might be accepted by both analyses.

Even if the gluon fusion has a production cross section dominant with respect to the VBF and we tried to optimize the cuts for the gluon fusion signal, in the final yield more than half of the signal is produced by VBF. This is clear, looking at Tables A.1, A.4, A.6 and A.10. So, we try to estimate how many events selected in this analysis would be accepted also in the VBF study. To do this we implement some of the standard VBF cuts. The main features of the VBF analysis is the requirement of two ‘tagging’ forward jets j_1 , j_2 . They have to be in opposite hemispheres with the Higgs boson emitted between them. Moreover there is a veto in case of any other hard jet is produced in the central region. In Appendix B there is the list of the cuts for each channel.

Table 5.12 reports the fractions of event accepted in this analysis, both in the gluon fusion and in the VBF samples, that pass the VBF cuts and consequently that might be selected also by the VBF analysis.

The overlap between the two analyses is small. In Table 5.13 we report the impact of each of the VBF cuts on the events accepted by our selection for the μh -channel. It seems that the requirements on the two tagging jets make the two analyses very different and this might be a support for the combination of these studies in the near future.

5.9 Pile Up Effects

Since next year LHC will already run at the luminosity of $10^{33} \text{cm}^{-2} \text{s}^{-1}$, the effects of pile up cannot be neglected any longer. This issue has to be studied carefully since a lot of the

Table 5.12: Fractions of events selected in this study that would be selected also in the VBF analysis.

channel	gluon fusion sample	VBF sample
μh	2.4%	13%
eh	3.6%	14%
ll	6.5%	16%
hh	8.3%	0.0%

Table 5.13: Cut flow of the basic cuts used in the VBF analysis applied on the events accepted in this analysis in the μh channel. The thresholds are reported in App. B

cut	GF sample	VBF sample
Start	1.0	1.0
Met	0.89	1.0
Mt	0.75	0.72
≥ 2 jets	0.14	0.59
$\eta_1 \times \eta_2$	0.05	0.25
$\Delta\eta_{jj}$	0.04	0.14
Central Jet Veto	0.04	0.14
M_{jj}	0.02	0.14
$\eta_{min} < \eta_\tau < \eta_{max}$	0.02	0.13
Mass Window [105,135]	0.02	0.13

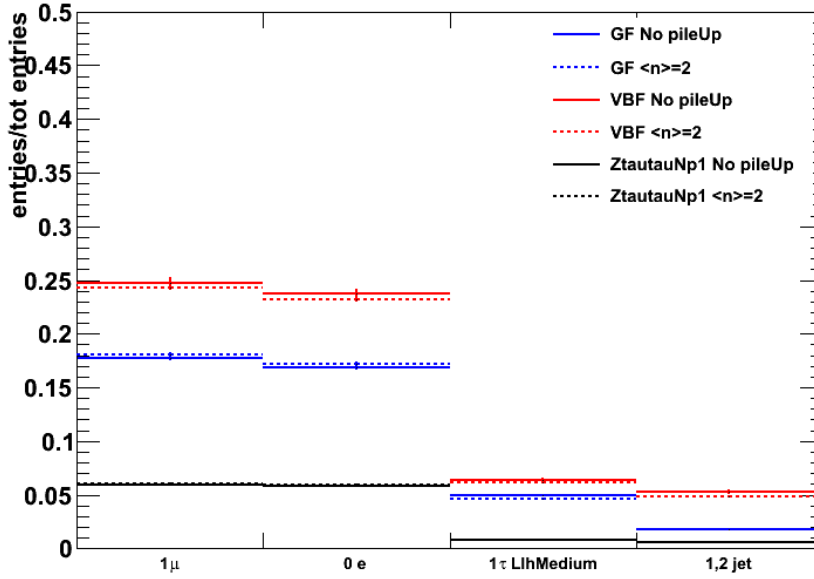


Figure 5.32: Object selection efficiencies in the μh channel with and without pile up. From left to right the efficiency in each bin is for the selection of exactly 1 muon, 0 electrons, 1 identified tau jet and 1 or 2 hard jets. The red line is for the VBF signal, the blue line for the gluon fusion signal and the black line for the $Z \rightarrow \tau^+ \tau^- + 1j$ background.

reconstruction performances can be significantly affected. Indeed the presence of additional tracks and energy clusters in the detector can for instance worsen the tau identification and the estimation of the missing energy.

At the moment the MC samples with pile up available have on average only two events of Minimum Bias (MB) that simulate the *in-time* pile up. In order to examine also a scenario with a higher number of primary vertices, we have reweighted the events in the pile up samples obtaining an average of 4 MB events. The reweighting has been performed dividing the normalized distributions of the number of vertices of type ‘3’⁹ in a sample with 2 pile-up events and in a sample with 4. All the MC used in this section are produced at a center of mass energy of 7 TeV.

Since these MC samples have a smaller statistics compared to the ones used in the analysis, we cannot compare the selection efficiencies through all the cut flow. Nevertheless we can estimate the impact of the pile up at least in the object selections. Figg. 5.32 and 5.33 illustrate the object selection efficiencies in the μh and hh channels. The values for 4 MB events is not shown since it would be overlapped with the values for 2 MB events. There are not significant differences.

The tau identification for true taus is only slightly affected as it is possible to see from the likelihood distribution in Fig. 5.34 on the left. It is encouraging that the pile up seems to have a bigger impact on fake tau jets (Fig. 5.34, right).

Another important feature that has to be checked in an environment with pile up is the lepton isolation. Fig 5.35 shows the isolation variables used in the event selection of the μh channel. It is clear, particularly for the average energy deposit in the cone of radius $\Delta R < 0.2$, that the values increases as the number of pile up events. However this should

⁹Vertexes with type ‘3’ are reconstructed vertices that are not identified as the primary vertex nor as secondary vertices.

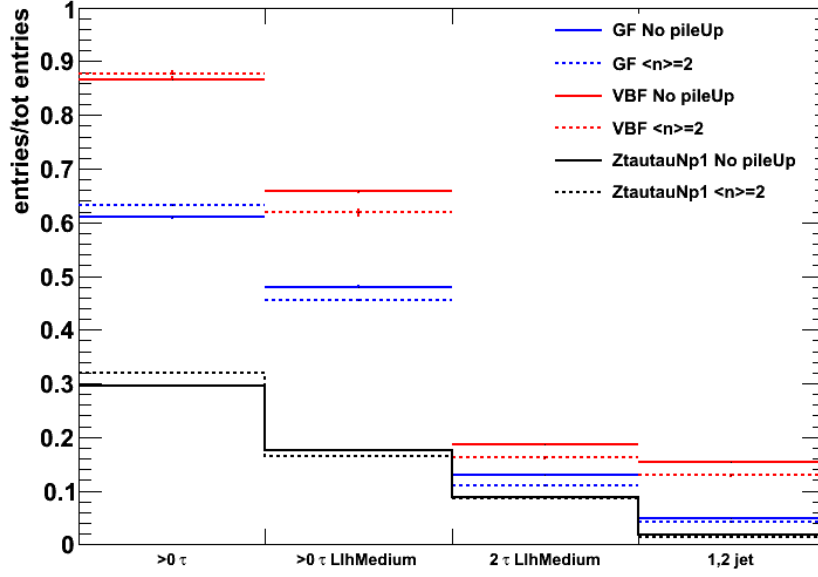


Figure 5.33: Object selection efficiencies in the hh channel with and without pile up. From left to right the efficiency in each bin is for the selection of at least 1 tau, at least 1 identified tau, exactly 2 identified taus and 1 or 2 hard jets. The red line is for the VBF signal, the blue line for the gluon fusion signal and the black line for the $Z \rightarrow \tau^+\tau^- + 1j$ background

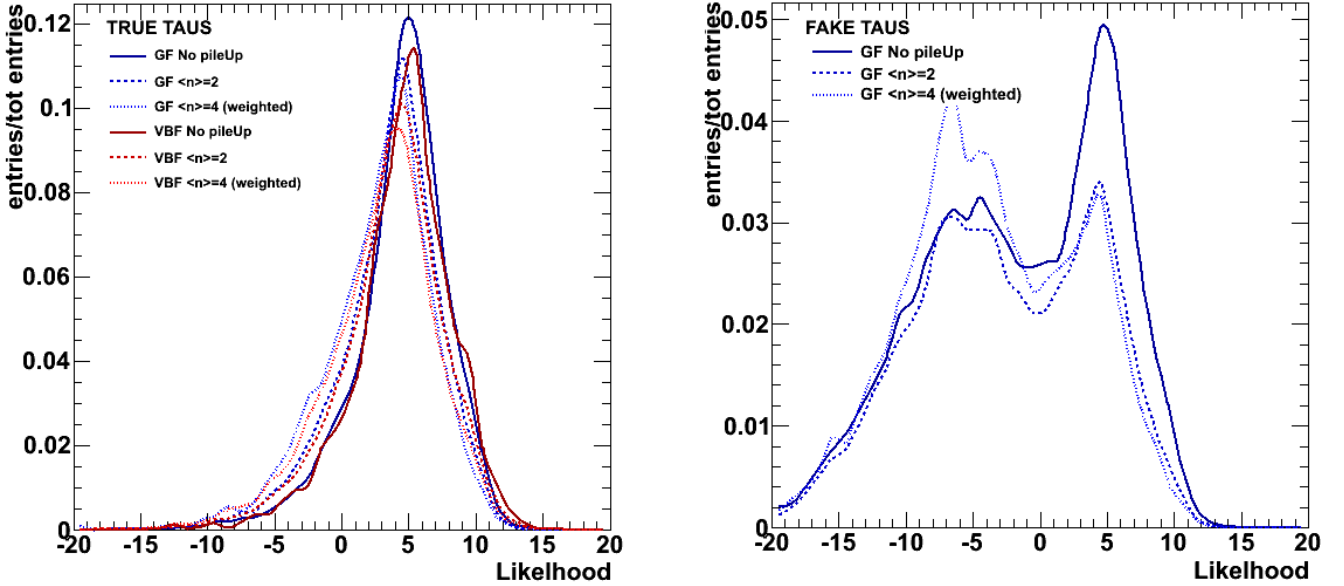


Figure 5.34: Tau likelihood distributions with and without pile up for taus jets matched with true taus in $\Delta R < 0.2$ (left) and for fake tau jets which don't match any truth taus in $\Delta R < 0.4$ (right). The threshold for a tau 'LlhMedium' is at 0.0 and for 'LlhTight' is at 4.5. The different normalizations are due to underflows.

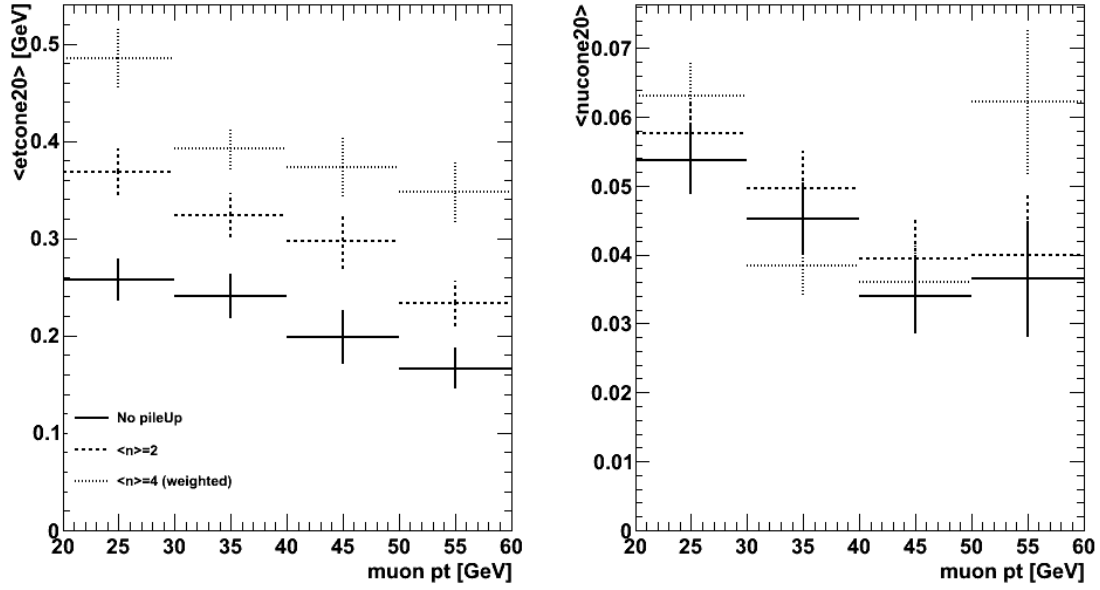


Figure 5.35: Muon isolation in the gluon fusion samples with and without pileup at 7 TeV.

not affect the event selection used in this analysis since the cut on this variable is very loose.

The study of the impact of the pile up needs to be improved since the likelihood in this analysis will not be the one used for the 2011 data and the average number of MB events is little with respect to what is expected for the future runs. Nevertheless, these preliminary plots make confident that the effects of the pile up on the signal efficiency can be contained and that the backgrounds shouldn't raise too much. The veto on events with more than two hard jets implemented in this analysis, for instance, should avoid a great increase of the QCD backgrounds.

Conclusions

In this thesis we have presented a new analysis for the $H \rightarrow \tau^+\tau^-$ search, which is sensitive to both the gluon fusion and the VBF production processes. All the possible final states, full leptonic, semi-leptonic and full hadronic, have been examined and the combined sensitivity has been computed for proton-proton collisions at the center of mass energy of 10 TeV. In order to make a comparison with the baseline search proposed for the SM $H \rightarrow \tau^+\tau^-$ at the ATLAS experiment, the VBF analysis, we have rescaled both signal and backgrounds to $\sqrt{s} = 14$ TeV. The resulting sensitivity computed for an integrated luminosity of 30 fb^{-1} is above 4σ taking a systematic uncertainty of 10% on the background. The corresponding sensitivity in the VBF analysis combining the leptonic and semi-leptonic channels is 5.32σ .

The two analyses have very little overlap, as shown in section 5.8. They are therefore almost orthogonal and the combination of them increases the sensitivity of the $H \rightarrow \tau^+\tau^-$ search in the mass region which is the hardest for the ATLAS experiment. More specifically, this thesis proves that it is possible to design an event selection which is sensitive to a signal production process to which the VBF analysis is almost blind and to use a new channel, where both the tau leptons decay hadronically, which has not been considered so far in the sensitivity estimations.

It is important to mention that this analysis is a simple cut-based study with plenty of room for improvements. For instance, a signal enhancement can be achieved including some missing production Feynman diagrams and from the implementation of the event selection in a multivariate analysis. Indeed the hard cuts applied in sequence are more effective for the VBF signal selection rather than for the gluon fusion one. Moreover better simulations can be provided by new MC productions with higher statistics and proper center of mass energies. Improvements in the signal to background ratio can come from the adoption of new mass reconstruction techniques. Indeed a pair of methods recently proposed can reconstruct better the mass of the Higgs candidate without the cuts imposed by the collinear mass approximation. These techniques can rescue a lot of the gluon fusion signal and they perform better in the channel where the collinear approximation has more troubles. In addition they should enhance the sensitivity because they consider spin correlations between taus. All these features are very promising and they need to be studied further. They couldn't be used in this analysis because they still have to be refined and the spin correlations were not properly simulated in the MC samples available. Another important outcome of this study is that there is a missing part of the VBF signal that can be used and which is rejected by the VBF analysis. This means that the $H \rightarrow \tau^+\tau^-$ search can be designed better, maybe splitting the analysis in jet multiplicity bins and even in 1- or 3- prong tau selections. The adoption of one the new mass reconstruction methods mentioned above can allow to include also the 0-jet channel because also the back-to-back events can be reconstructed.

The attempt to find an effective strategy to select the gluon fusion signal is important primarily because, on the contrary of the VBF, this signal is sensitive to the couplings of the SM Higgs boson to the fermions. Moreover it can be also valuable for the study of other features of the Higgs boson like the CP properties. Several papers on this topic, like [65],

are based on the gluon fusion in association with jets as production process for the signal events.

This analysis is based only on MC simulations and the next step is to prepare for the search on real data. Indeed the ultimate search for the Higgs boson will be based on data-driven background estimations as much as possible in order to reduce systematics. Already with the 2010 data it is possible to study and understand the background from $W + j$ and multijets events and large sets of $Z + j$ and $t\bar{t}$ will be available soon after the LHC restart in 2011. Verifying in real data the size and the shapes of backgrounds is very important especially for the new channel proposed here, the hadronic channel. Recent results in ATLAS on multi jets [66] show that the MC expectations are reproducing the data reasonably well for 3 and 4 jets final states (the biggest background to the hadronic channel) but of course a detailed analysis in the context of this study needs to be done. Additionally, Z and W backgrounds in this channel need to be studied.

This analysis is important for two reasons: its timing and its results. At the end of 2011 probably all the low mass range, where the SM Higgs boson is likely to be, can be excluded, so it is necessary to prepare an optimized $H \rightarrow \tau^+\tau^-$ search before the end of this year. Then, even if the results of this study give only a glimpse on the actual sensitivity, they prove that the contribution to the ATLAS combined sensitivity in the mass region $m_H < 120$ GeV which comes from the $H \rightarrow \tau^+\tau^-$ is underestimated. At the moment the relevant channel in that region is $H \rightarrow \gamma\gamma$, but the search in the $\tau^+\tau^-$ final state can be improved. It is important to bear in mind that this is the mass range where the SM Higgs boson has the highest probability to be found and where the ATLAS sensitivity is worst. Any improvements are therefore of the utmost importance.

Bibliography

- [1] G. Kane, *The Dawn of Physics beyond the Standard Model*, January 2006. Scientific American **15** (January 2006) 4–11.
- [2] C. Rubbia, *Experimental observation of the intermediate vector bosons W^+ , W^- , and Z^0* , Reviews of Modern Physics **57** (July, 1985) 699–722.
- [3] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*. J. Wiley & Sons, 1984.
- [4] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [5] I. J. Aitchison and A. J. Hey, *Gauge Theories in Particle Physics, Vol II: QCD and the Electroweak Theory*. CRC Pr I Llc, 2003.
- [6] Particle Data Group Collaboration, K. Nakamura et al., *Review of particle physics*, J. Phys. **G37** (2010) 075021.
- [7] S. A. Bludman, 1958. Nuovo Cimento G (1958) 433.
- [8] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.
- [9] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.
- [10] A. Salam, in *Elementary Particle Theory*, p. 367. Almqvist and Wiksell, Stockholm, 1968.
- [11] P. W. Higgs Phys. Rev. Lett. **12** (1964) 132.
- [12] P. W. Higgs Phys. Rev. **145** (1966) 1156.
- [13] F. Englert and R. Brout Phys. Rev. Lett. **13** (1964) 321.
- [14] G. S. Guralnik, C. R. Hagen, and T. W. Kibble Phys. Rev. Lett. **13** (1964) 585.
- [15] J. E. Kim, P. Langacker, M. Levine, and H. H. Williams, *A theoretical and experimental review of the weak neutral current: a determination of its structure and limits on deviations from the minimal $SU(2)_L \times U(1)$ electroweak theory*, Rev. Mod. Phys. **53** (Apr, 1981) 211–252.
- [16] C. Rubbia, *Experimental observation of the intermediate vector bosons W^+ , W^- , and Z^0* , Rev. Mod. Phys. **57** (Jul, 1985) 699–722.
- [17] B. Jacobsen, *Top mass from electroweak measurements*, arXiv:hep-ex/9407002 [hep-ex].

- [18] D0 Collaboration, S. Abachi et al., *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, [arXiv:hep-ex/9503003](#).
- [19] S. Dawson, *Introduction to electroweak symmetry breaking*, [arXiv:hep-ph/9901280](#).
- [20] L. Reina, *TASI 2004 lecture notes on Higgs boson physics*, [arXiv:hep-ph/0512377](#).
- [21] N. Cabibbo, L. Maiani, G. Parisi, and R. Petronzio, *Bounds on the Fermions and Higgs Boson Masses in Grand Unified Theories*, Nucl. Phys. **B158** (1979) 295–305.
- [22] M. Sher, *Precise vacuum stability bound in the standard model*, Phys. Lett. **B317** (1993) 159–163, [arXiv:hep-ph/9307342](#).
- [23] M. J. G. Veltman, *Second Threshold in Weak Interactions*, Acta Phys. Polon. **B8** (1977) 475.
- [24] The Snowmass Working Group on Precision Electroweak Measurements Collaboration, U. Baur et al., *Present and future electroweak precision measurements and the indirect determination of the mass of the Higgs boson*, [arXiv:hep-ph/0202001](#).
- [25] J. Alcaraz, *Precision Electroweak Measurements and Constraints on the Standard Model*, [arXiv:0911.2604 \[hep-ex\]](#).
- [26] ALEPH Collaboration Collaboration, et al., *Precision Electroweak Measurements and Constraints on the Standard Model*, [arXiv:1012.2367 \[hep-ex\]](#). * Temporary entry *.
- [27] H. Flacher et al., *Gfitter - Revisiting the Global Electroweak Fit of the Standard Model and Beyond*, Eur. Phys. J. **C60** (2009) 543–583, [arXiv:0811.0009 \[hep-ph\]](#).
- [28] *LHC Higgs Cross Section Working Group*, Available at: <https://twiki.cern.ch/twiki/bin/view/lhcphysics/crosssections>.
- [29] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, JHEP **11** (2009) 088, [arXiv:0909.3420 \[hep-ph\]](#).
- [30] LEP Working Group for Higgs boson searches Collaboration, R. Barate et al., *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565** (2003) 61–75, [arXiv:hep-ex/0306033](#).
- [31] ALEPH Collaboration, R. Barate et al., *Observation of an excess in the search for the standard model Higgs boson at ALEPH*, Phys. Lett. **B495** (2000) 1–17, [arXiv:hep-ex/0011045](#).
- [32] DELPHI Collaboration, P. Abreu et al., *Search for the standard model Higgs boson at LEP in the year 2000*, Phys. Lett. **B499** (2001) 23–37, [arXiv:hep-ex/0102036](#).
- [33] L3 Collaboration, M. Acciarri et al., *Higgs candidates in e^+e^- interactions at $\sqrt{s} = 206.6\text{-GeV}$* , Phys. Lett. **B495** (2000) 18–25, [arXiv:hep-ex/0011043](#).
- [34] OPAL Collaboration, G. Abbiendi et al., *Search for the standard model Higgs boson in e^+e^- collisions at $s^{*(1/2)} = 192\text{-GeV} - 209\text{-GeV}$* , Phys. Lett. **B499** (2001) 38–52, [arXiv:hep-ex/0101014](#).
- [35] B. Kilminster, *Higgs boson searches at the Tevatron*, [arXiv:1012.2172 \[hep-ex\]](#).

- [36] B. Penning, *High Mass Higgs Boson Searches at the Tevatron*, [arXiv:1011.6307 \[hep-ex\]](#).
- [37] *Cern Faq: The LHC Guide*, Available at <http://cdsmedia.cern.ch/img/cern-brochure-2009-003-eng.pdf>, February, 2009.
- [38] Available from the atlas collaboration web page <http://atlas.ch/>.
- [39] L. Evans, *The large Hadron Collider at CERN*, in *29th LHC Symposium*. Lisbon, November, 2010.
- [40] J. Stirling. Available at <http://projects.hepforge.org/mstwpdf/plots/plots.html>.
- [41] F. Gianotti, *ATLAS: highlights from the first run and prospects for 2011*, in *LHC Beam Operation workshop*. Evian, December, 2010.
- [42] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [43] The ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, [arXiv:0901.0512 \[hep-ex\]](#).
- [44] M. Spiropulu and S. Stapnes, *LHC's ATLAS and CMS Detectors*, in *Perspective on LHC Physics*, G. Kane and A. Pierce, eds. World Scientific, 2008.
- [45] *LHC Phenomenology with Madgraph - Introduction to MC's for the LHC*, Available at <http://cp3wks05.fynu.ucl.ac.be/twiki/bin/view/physics/copenhagen2010>.
- [46] M. L. Mangano and T. J. Stelzer, *Tools for the Simulation of Hard Hadronic Collisions*, *Annual Review of Nuclear and Particle Science* **55** (Dec., 2005) 555–588.
- [47] S. Agostinelli et al., *G4—a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) no. 3, 250 – 303. <http://www.sciencedirect.com/science/article/B6TJM-48TJFY8-5/2/23ea98096ce11c1be446850c04cfa498>.
- [48] D. Adams, S. Asai, D. Cavalli, M. Dührssen, K. Edmonds, S. Elles, M. Fehling, U. Felzmann, L. Gladilin, L. Helary, M. Hohlfeld, S. Horvat, K. Jakobs, M. Kaneda, G. Kirsch, S. Kuehn, J. F. Marchand, C. Pizio, X. Portell, D. Rebuzzi, E. Schmidt, A. Shibata, I. Vivarelli, S. Winkelmann, and S. Yamamoto, *The ATLFast-II performance in release 14 -particle signatures and selected benchmark processes-*, Tech. Rep. ATL-PHYS-INT-2009-110, CERN, Geneva, Dec, 2009.
- [49] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **0206** (2002) 029, [arXiv:hep-ph/0204244 \[hep-ph\]](#).
- [50] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **0307** (2003) 001, [arXiv:hep-ph/0206293 \[hep-ph\]](#).
- [51] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6.5 release note*, [arXiv:hep-ph/0210213 \[hep-ph\]](#).
- [52] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput.Phys.Commun.* **178** (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](#).

- [53] C. ATLAS, *ATLAS Sensitivity Prospects for 1 Higgs Boson Production at the LHC Running at 7, 8 or 9 TeV*, Tech. Rep. ATL-PHYS-PUB-2010-015, CERN, Geneva, Nov, 2010.
- [54] S. Xella, *VBF SM Higgs boson search with ATLAS*, .
- [55] A. Belyaev, R. Guedes, S. Moretti, and R. Santos, *Higgs boson phenomenology in $\tau^+\tau^-$ final states at the LHC*, JHEP **07** (2010) 051, [arXiv:0912.2620 \[hep-ph\]](#).
- [56] B. Mellado, W. Quayle, and W. S. Lan, *Search for Higgs Boson Decay $H \rightarrow \tau^+\tau^- \rightarrow l^+l^-p_t$ Associated with One High P_T Jet at the LHC*, Tech. Rep. ATL-COM-PHYS-2004-023, CERN, Geneva, Apr, 2004.
- [57] N. Andari, K. Assamagan, A. Bourgaux, M. Campanelli, G. Carrillo-Montoya, M. Escalier, M. Flechl, J. Huston, S. Muanza, B. Murray, B. Mellado, A. Nisati, J. Qian, D. Rebutti, M. Schram, M. Schumacher, R. Tanaka, T. Vickey, M. Warsinsky, and H. Zhang, *Higgs Production Cross Sections and Decay Branching Ratios*, Tech. Rep. ATL-PHYS-INT-2010-030, CERN, Geneva, Mar, 2010.
- [58] J. M. Campbell, R. Ellis, and G. Zanderighi, *Next-to-Leading order Higgs + 2 jet production via gluon fusion*, JHEP **0610** (2006) 028, [arXiv:hep-ph/0608194 \[hep-ph\]](#).
- [59] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, et al., *MadGraph/MadEvent v4: The New Web Generation*, JHEP **0709** (2007) 028, [arXiv:0706.2334 \[hep-ph\]](#).
- [60] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, [arXiv:1012.4686 \[hep-ex\]](#).
- [61] Available at http://root.cern.ch/root/html/roostats__numbercountingutils.html.
- [62] R. D. Cousins, J. T. Linnemann, and J. Tucker, *Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, ArXiv Physics e-prints (Feb., 2007) , [arXiv:physics/0702156](#).
- [63] K. Cranmer, *Statistical Challenges for Searches for New Physics at the LHC*, in *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, L. Lyons & M. Karagöz Ünel, ed., pp. 112–+. 2006. [arXiv:physics/0511028](#).
- [64] J. Linnemann, *Measures of Significance in HEP and Astrophysics*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, L. Lyons, R. Mount, & R. Reitmeyer, ed., pp. 35–+. 2003. [arXiv:physics/0312059](#).
- [65] J. R. Andersen, *CP-properties of the Higgs-boson couplings from $H +$ dijets through gluon fusion*, [arXiv:1006.2563 \[hep-ph\]](#).
- [66] *Measurements of multijet production cross sections in proton-proton collisions at 7 TeV center-of-mass energy with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-084, CERN, Geneva, Oct, 2010.

Appendix A

Detailed Cut flows

Cut	Higgs lh gg [fb]	Higgs lh VBF [fb]	$W\mu\nu + 2j$ [fb]	$W\mu\nu + 3j$ [fb]	$W\tau\nu + 2j$ [fb]	$W\tau\nu + 3j$ [fb]
Start	9.6E+02±1.4	80±0.36	8.3E+05±9.4E+02	2.5E+05±5.4E+02	8.2E+05±9.2E+02	2.5E+05±5.2E+02
Selection	18±0.19	3.8±0.078	7.3E+03±88	2.9E+03±58	4.9E+02±23	2.2E+02±16
Charge*	17±0.18	3.3±0.072	3.4E+03±60	1.3E+03±38	2.4E+02±16	1.1E+02±11
TauTrack*	17±0.18	3.3±0.072	3.2E+03±58	1.2E+03±37	2.3E+02±15	1.1E+02±11
Isolation*	16±0.18	3.1±0.07	3E+03±57	1.1E+03±36	1.7E+02±13	73±9
BJet*	16±0.17	3±0.069	3E+03±56	1.1E+03±36	1.7E+02±13	73±9
EtaJet*	7.5±0.12	1.6±0.05	1.1E+03±34	2.5E+02±14	65±7.6	16±3.4
Met	4.9±0.097	1.3±0.046	1.1E+03±36	2.3E+02±15	56±7.6	14±3.3
Coll	1.9±0.06	0.76±0.035	27±2.3	11±1.3	6.7±1.3	2±0.59
Visible	1.8±0.059	0.67±0.033	13±1.5	4.7±0.74	4.1±0.9	1.1±0.38
Mhj	0.23±0.021	0.22±0.019	0.67±0.33	0.9±0.31	0±0	0.41±0.21
Window	0.16±0.018	0.17±0.016	0±0	0.1±0.1	0±0	0±0

Table A.1: μh -channel cut flow. (*) factorized cuts, errors are only statistical.

Cut	$W\mu\nu + 2j$ [fb]	$W\mu\nu + 3j$ [fb]	$W\tau\nu + 2j$ [fb]	$W\tau\nu + 3j$ [fb]
Start	$8.3E+05 \pm 9.4E+02$	$2.5E+05 \pm 5.4E+02$	$8.2E+05 \pm 9.2E+02$	$2.5E+05 \pm 5.2E+02$
Selection	$1.6E+04 \pm 1.3E+02$	$6.6E+03 \pm 87$	$1E+03 \pm 33$	$4.6E+02 \pm 23$
Charge*	$6.6E+03 \pm 84$	$2.6E+03 \pm 55$	$4.3E+02 \pm 21$	$2.1E+02 \pm 15$
TauTrack*	$5.7E+03 \pm 78$	$2.2E+03 \pm 50$	$3.7E+02 \pm 20$	$1.8E+02 \pm 14$
Isolation*	$5.4E+03 \pm 76$	$2.1E+03 \pm 49$	$2.9E+02 \pm 17$	$1.3E+02 \pm 12$
BJet*	$5.3E+03 \pm 75$	$2E+03 \pm 48$	$2.9E+02 \pm 17$	$1.3E+02 \pm 12$
EtaJet*	$2E+03 \pm 42$	$4.6E+02 \pm 18$	$1.1E+02 \pm 9.2$	32 ± 4.6
Met	$1.8E+03 \pm 44$	$4.3E+02 \pm 19$	94 ± 9.1	28 ± 4.5
Coll	49 ± 2.7	21 ± 1.6	9.5 ± 1.3	4.5 ± 0.9
mlh	22 ± 1.8	8 ± 0.88	5.7 ± 0.94	2.4 ± 0.57
Mhj	1.6 ± 0.45^a	1.2 ± 0.31^b	0.87 ± 0.32^c	0.46 ± 0.2^d
Window	0 ± 0	0.083 ± 0.083	0 ± 0	0 ± 0

Table A.2: μh -channel cut flow with ‘**Loose**’ tau selection. (*) factorized cuts, errors are only statistical. The yields in the footnotes are obtained with the $M(h, j)$ threshold at 500 instead of 600 GeV.

^aMhj $3.7 \pm 0.7 \rightarrow$ Window 0.39 ± 0.23

^bMhj $1.6 \pm 0.37 \rightarrow$ Window 0.33 ± 0.17

^cMhj $1.1 \pm 0.36 \rightarrow$ Window 0.0 ± 0.0

^dMhj $0.61 \pm 0.23 \rightarrow$ Window 0.076 ± 0.077

Cut	$Z\mu\mu + 1j$ [fb]	$Z\mu\mu + 2j$ [fb]	$Z\tau\tau + 1j$ [fb]	$Z\tau\tau + 2j$ [fb]	$Z\tau\tau + 3j$
Start	$2.5E+05 \pm 1E+03$	$8.5E+04 \pm 1.9E+02$	$2.6E+05 \pm 2.5E+02$	$8.6E+04 \pm 1.9E+02$	$2.6E+04 \pm 1E+02$
Selection	$4.6E+02 \pm 43$	$6.2E+02 \pm 16$	$1.4E+03 \pm 18$	$6.4E+02 \pm 16$	98 ± 6.3
Charge*	$4.1E+02 \pm 41$	$2.8E+02 \pm 11$	$1.2E+03 \pm 17$	$5.3E+02 \pm 15$	75 ± 5.5
TauTrack*	$4.1E+02 \pm 41$	$2.8E+02 \pm 11$	$1.2E+03 \pm 17$	$5.3E+02 \pm 15$	75 ± 5.5
Isolation*	$3.9E+02 \pm 40$	$2.6E+02 \pm 10$	$1.1E+03 \pm 17$	$4.9E+02 \pm 14$	68 ± 5.3
BJet*	$3.8E+02 \pm 39$	$2.5E+02 \pm 10$	$1.1E+03 \pm 17$	$4.8E+02 \pm 14$	66 ± 5.2
EtaJet*	$1.1E+02 \pm 25$	82 ± 5.4	$3.7E+02 \pm 11$	94 ± 6.5	10 ± 2
Met	31 ± 9.2	39 ± 3.1	$2.1E+02 \pm 8.2$	63 ± 5	7.7 ± 1.7
Coll	0.99 ± 1	1.7 ± 0.33	81 ± 3.6	26 ± 2.3	3.2 ± 0.75
Visible	0.99 ± 1	0.86 ± 0.22	61 ± 2.9	17 ± 1.6	2 ± 0.49
Mhj	0 ± 0	0 ± 0	0.6 ± 0.2^a	1 ± 0.26	0.17 ± 0.092^b
MassWindow	0 ± 0	0 ± 0	0 ± 0	0.18 ± 0.11	0 ± 0

Table A.3: μh -channel cut flow. (*) factorized cuts, errors are only statistical. The yields in the footnotes are obtained with the $M(h, j)$ threshold at 500 instead of 600 GeV.

^aMhj $1.3 \pm 0.29 \rightarrow$ Window 0.13 ± 0.094

^bMhj $0.39 \pm 0.15 \rightarrow$ Window 0.086 ± 0.063

Cut	Higgs lh gg [fb]	Higgs lh VBF [fb]	$We\nu$ +2j [fb]	$We\nu$ +3j [fb]	$W\tau\nu$ +2j [fb]	$W\tau\nu$ +3j [fb]
Start	9.6E+02±1.4	80±0.36	8.2E+05±9.4E+02	2.5E+05±5.2E+02	8.2E+05±9.2E+02	2.5E+05±5.2E+02
Selection	13±0.16	2.7±0.066	5.3E+03±75	2.1E+03±48	3E+02±18	1.5E+02±13
Charge*	12±0.15	2.4±0.062	2.4E+03±51	1E+03±33	1.6E+02±13	68±8.7
TauTrack*	12±0.15	2.4±0.061	2.3E+03±49	9.3E+02±32	1.5E+02±13	61±8.3
Isolation*	12±0.15	2.3±0.061	2.2E+03±48	8.7E+02±31	1.4E+02±12	58±8
BJet*	12±0.15	2.3±0.06	2.1E+03±48	8.5E+02±31	1.4E+02±12	57±7.9
EtaJet*	5.6±0.1	1.2±0.043	8E+02±28	2.2E+02±13	51±7.3	11±2.8
Met	3.9±0.086	0.98±0.04	7.4E+02±30	2E+02±14	44±7.3	9.5±2.7
Coll	1.5±0.054	0.61±0.031	21±2	8.8±1.2	4.6±1.2	1.1±0.42
Visible	1.5±0.053	0.56±0.03	10±1.4	4.3±0.75	2.3±0.74	0.5±0.24
Mhj	0.3±0.024	0.27±0.021	0.97±0.4 ^a	0.45±0.23 ^b	0.36±0.26 ^c	0±0 ^d
Window	0.21±0.02	0.22±0.019	0.32±0.23	0±0	0.18±0.18	0±0

Table A.4: eh -channel cut flow. (*) factorized cuts, errors are only statistical. The yields in the footnotes are obtained with the $M(h, j)$ threshold at 400 instead of 500 GeV.

^aMhj $2.4 \pm 0.63 \rightarrow$ Window 0.32 ± 0.23

^bMhj $1.5 \pm 0.42 \rightarrow$ Window 0.22 ± 0.16

^cMhj $0.54 \pm 0.32 \rightarrow$ Window 0.36 ± 0.26

^dMhj $0.083 \pm 0.086 \rightarrow$ Window 0.0 ± 0.0

Cut	Zee +1j [fb]	Zee +2j [fb]	$Z\tau\tau$ +1j [fb]	$Z\tau\tau$ +2j [fb]	$Z\tau\tau$ +3j
Start	2.5E+05±1E+03	8.8E+04±1.9E+02	2.6E+05±2.5E+02	8.6E+04±1.9E+02	2.6E+04±1E+02
Selection	4.9E+03±1.4E+02	2.1E+03±29	8.9E+02±15	4.4E+02±13	71±5.4
Charge*	3.3E+03±1.2E+02	1.3E+03±23	7.8E+02±14	3.7E+02±12	53±4.6
TauTrack*	3.3E+03±1.2E+02	1.3E+03±23	7.8E+02±14	3.7E+02±12	53±4.6
Isolation*	3.2E+03±1.1E+02	1.2E+03±22	7.7E+02±14	3.6E+02±12	49±4.5
BJet*	3.1E+03±1.1E+02	1.2E+03±22	7.7E+02±14	3.5E+02±12	46±4.3
EtaJet*	1.2E+03±77	2.8E+02±10	2.5E+02±9.5	77±6.1	7.6±1.7
Met	1.9E+02±19	66±3.2	1.4E+02±6.6	51±4.7	5.7±1.4
Coll	5.8±2.4	2.8±0.41	51±2.8	21±2.2	2.5±0.66
Visible	4.8±2.2	2.5±0.38	38±2.3	15±1.6	1.4±0.41
Mhj	0±0 ^a	0±0 ^b	0.55±0.2 ^c	1.2±0.31 ^d	0.043±0.045 ^e
Window	0±0	0±0	0.069±0.069	0.42±0.18	0±0

Table A.5: eh -channel cut flow. (*) factorized cuts, errors are only statistical. The yields in the footnotes are obtained with the $M(h, j)$ threshold at 400 instead of 500 GeV.

^aMhj $0.0 \pm 0.0 \rightarrow$ Window 0.0 ± 0.0

^bMhj $0.055 \pm 0.055 \rightarrow$ Window 0.055 ± 0.055

^cMhj $2 \pm 0.38 \rightarrow$ Window 0.27 ± 0.14

^dMhj $2 \pm 0.41 \rightarrow$ Window 0.63 ± 0.22

^eMhj $0.13 \pm 0.081 \rightarrow$ Window 0.0 ± 0.0

Cut	Higgs ll gg [fb]	Higgs ll VBF [fb]	$W_{e\nu} + 2j$ [fb]	$W_{e\nu} + 3j$ [fb]
Start	$1.4E+02 \pm 0.32$	12 ± 0.055	$8.2E+05 \pm 9.4E+02$	$2.5E+05 \pm 5.2E+02$
Selection	8 ± 0.076	1.6 ± 0.02	$1.1E+03 \pm 33$	$2.7E+02 \pm 17$
Charge*	7.9 ± 0.076	1.6 ± 0.02	$6.1E+02 \pm 25$	$1.7E+02 \pm 14$
Isolation*	7.3 ± 0.073	1.4 ± 0.019	$2.7E+02 \pm 17$	46 ± 7.2
Bjet*	7.2 ± 0.072	1.4 ± 0.018	$2.6E+02 \pm 16$	45 ± 7.1
EtaJet*	3.4 ± 0.049	0.74 ± 0.013	60 ± 5.8	8.6 ± 1.9
Met	2 ± 0.038	0.59 ± 0.012	51 ± 5.5	7.2 ± 1.7
Coll	0.79 ± 0.024	0.33 ± 0.0089	0.36 ± 0.15	0.1 ± 0.065
Visible	0.74 ± 0.023	0.27 ± 0.0082	0.12 ± 0.086	0.07 ± 0.052
Mhj	0.076 ± 0.0074	0.091 ± 0.0047	0 ± 0	0.035 ± 0.036
Window	0.053 ± 0.0062	0.073 ± 0.0042	0 ± 0	0 ± 0

Table A.6: ll -channel cut flow. (*) factorized cuts, errors are only statistical.

Cut	$W_{\mu\nu} + 2j$ [fb]	$W_{\mu\nu} + 3j$ [fb]	$W_{\tau\nu} + 2j$ [fb]	$W_{\tau\nu} + 3j$ [fb]
Start	$8.3E+05 \pm 9.4E+02$	$2.5E+05 \pm 5.4E+02$	$8.2E+05 \pm 9.2E+02$	$2.5E+05 \pm 5.2E+02$
Selection	$4.7E+02 \pm 22$	$2.4E+02 \pm 17$	91 ± 9.7	37 ± 6.4
Charge*	$3.6E+02 \pm 19$	$1.8E+02 \pm 15$	66 ± 8.3	30 ± 5.8
Isolation*	30 ± 5.6	9.3 ± 3.3	18 ± 4.3	3.3 ± 1.9
Bjet*	29 ± 5.5	9.3 ± 3.3	17 ± 4.1	3.3 ± 1.9
EtaJet*	10 ± 2.2	1.9 ± 0.73	5.8 ± 1.9	0.81 ± 0.57
Met	8.5 ± 1.9	1.6 ± 0.65	4.9 ± 1.8	0.69 ± 0.51
Coll	0.23 ± 0.09	0.054 ± 0.031	0.2 ± 0.13	0.12 ± 0.1
Visible	0.07 ± 0.043	0.009 ± 0.0097	0.13 ± 0.1	0.074 ± 0.068
Mhj	0.023 ± 0.024	0 ± 0	0 ± 0	0.074 ± 0.068
Window	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Table A.7: ll -channel cut flow. (*) factorized cuts, errors are only statistical.

Cut	$Zee + 1j$ [fb]	$Zee + 2j$ [fb]	$Z\mu\mu + 1j$ [fb]	$Z\mu\mu + 2j$ [fb]
Start	$2.5E+05 \pm 1E+03$	$8.8E+04 \pm 1.9E+02$	$2.5E+05 \pm 1E+03$	$8.5E+04 \pm 1.9E+02$
Selection	$4.8E+04 \pm 4.4E+02$	$1.9E+04 \pm 88$	$8.5E+04 \pm 5.9E+02$	$3.2E+04 \pm 1.1E+02$
Charge*	$4.8E+04 \pm 4.4E+02$	$1.9E+04 \pm 87$	$8.5E+04 \pm 5.9E+02$	$3.2E+04 \pm 1.1E+02$
Isolation*	$4.4E+04 \pm 4.2E+02$	$1.7E+04 \pm 83$	$7.7E+04 \pm 5.6E+02$	$2.8E+04 \pm 1.1E+02$
Bjet*	$4.4E+04 \pm 4.2E+02$	$1.6E+04 \pm 82$	$7.6E+04 \pm 5.5E+02$	$2.7E+04 \pm 1.1E+02$
EtaJet*	$1.4E+04 \pm 2.9E+02$	$3.3E+03 \pm 40$	$2.5E+04 \pm 3.9E+02$	$5.5E+03 \pm 52$
Met	$5.7E+02 \pm 29$	$2.6E+02 \pm 5.5$	$1.1E+03 \pm 40$	$4.4E+02 \pm 7.2$
Coll	11 ± 3.6	5 ± 0.59	26 ± 5.5	8.5 ± 0.78
Visible	2.3 ± 1.6	0.28 ± 0.14	3.5 ± 2	0.99 ± 0.26
Mhj	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Window	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Table A.8: ll -channel cut flow. (*) factorized cuts, errors are only statistical.

Cut	$Z\tau\tau + 1j$ [fb]	$Z\tau\tau + 2j$ [fb]	$Z\tau\tau + 3j$ [fb]
Start	$2.6E+05 \pm 2.5E+02$	$8.6E+04 \pm 1.9E+02$	$2.6E+04 \pm 1E+02$
Selection	$6.1E+02 \pm 12$	$2.9E+02 \pm 11$	39 ± 4
Charge*	$5.9E+02 \pm 12$	$2.8E+02 \pm 11$	37 ± 3.9
Isolation*	$5.6E+02 \pm 12$	$2.5E+02 \pm 10$	29 ± 3.4
Bjet*	$5.6E+02 \pm 12$	$2.5E+02 \pm 10$	28 ± 3.4
EtaJet*	$1.9E+02 \pm 8.5$	48 ± 4.9	5.2 ± 1.5
Met	88 ± 5	28 ± 3.3	3.6 ± 1.2
Coll	37 ± 2.5	13 ± 1.7	1.8 ± 0.63
Visible	30 ± 2.1	9.5 ± 1.3	1.2 ± 0.43
Mhj	0.15 ± 0.11^a	0.61 ± 0.21	0.053 ± 0.056^b
Window	0 ± 0	0.068 ± 0.068	0 ± 0

Table A.9: ll -channel cut flow. (*) factorized cuts, errors are only statistical. The yields in the footnotes are obtained with the thresholds on $M(h, j)$ at 500 instead of 600 GeV.

^aMhj $0.91 \pm 0.27 \rightarrow$ Window 0.076 ± 0.076

^bMhj $0.16 \pm 0.1 \rightarrow$ Window 0.053 ± 0.056

Cut	Higgs hh gg [fb]	Higgs hh VBF [fb]	$W\tau\nu + 2j$ [fb]	$W\tau\nu + 3j$ [fb]
Start	$6.9E+02 \pm 2.2$	53 ± 0.27	$8.2E+05 \pm 9.2E+02$	$2.5E+05 \pm 5.2E+02$
Selection	22 ± 0.39	5.7 ± 0.087	$1.1E+03 \pm 34$	$4.4E+02 \pm 22$
Charge*	20 ± 0.37	4.8 ± 0.08	$4.9E+02 \pm 22$	$1.9E+02 \pm 15$
TauTrack*	20 ± 0.37	4.8 ± 0.08	$4.4E+02 \pm 21$	$1.8E+02 \pm 14$
BJet*	19 ± 0.36	4.7 ± 0.079	$4.3E+02 \pm 21$	$1.7E+02 \pm 14$
EtaJet*	9.3 ± 0.25	2.4 ± 0.056	$1.6E+02 \pm 12$	39 ± 5.5
Met	2.6 ± 0.13	1.3 ± 0.042	$1.2E+02 \pm 11$	32 ± 5.1
Coll	0.92 ± 0.079	0.88 ± 0.034	10 ± 1.5	4.5 ± 0.95
Mhj	0.13 ± 0.03	0.35 ± 0.022	1.9 ± 0.56	1.1 ± 0.37
Visible	0.13 ± 0.03	0.32 ± 0.021	0.3 ± 0.21^a	0.4 ± 0.21^b
Window	0.083 ± 0.024	0.26 ± 0.019	0 ± 0	0 ± 0

Table A.10: hh -channel cut flow. (*) factorized cuts, errors are only statistical. The yields in the footnotes are obtained with the thresholds on $M(h, j)$ at 500 instead of 600 GeV.

^aVisible $0.6 \pm 0.3 \rightarrow$ Window 0.0 ± 0.0

^bVisible $0.7 \pm 0.28 \rightarrow$ Window 0.1 ± 0.1

Cut	$Z\tau\tau + 1j$ [fb]	$Z\tau\tau + 2j$ [fb]	$Z\tau\tau + 3j$ [fb]	qcd J2 [fb]	qcd J3
Start	$3.2E+04 \pm 1.4E+02$	$1.6E+04 \pm 60$	$4.9E+03 \pm 19$	$5.6E+10 \pm 2.7E+07$	$3.3E+09 \pm 1.2E+06$
Selection	$5.8E+02 \pm 19$	$6.5E+02 \pm 12$	85 ± 2.4	$5.4E+08 \pm 2.6E+06$	$2.1E+08 \pm 3E+05$
Charge*	$4.8E+02 \pm 18$	$5.4E+02 \pm 11$	64 ± 2.1	$1.5E+03 \pm 11$	$3.2E+02 \pm 0.63$
TauTrack*	$4.8E+02 \pm 18$	$5.3E+02 \pm 11$	63 ± 2.1	$1.5E+03 \pm 11$	$3.2E+02 \pm 0.63$
BJet*	$4.7E+02 \pm 17$	$5.2E+02 \pm 11$	61 ± 2.1	$1.5E+03 \pm 11$	$3.2E+02 \pm 0.63$
EtaJet*	$1.5E+02 \pm 11$	89 ± 4.8	11 ± 0.87	$2.8E+02 \pm 4.3$	35 ± 0.18
Met	65 ± 6.4	35 ± 2.3	4.9 ± 0.48	10 ± 0.22	3.2 ± 0.02
Coll	32 ± 3.5	19 ± 1.3	2.6 ± 0.27	10 ± 0.61	3.2 ± 0.039
Mhj	1.4 ± 0.49	1.9 ± 0.26	0.33 ± 0.061	0.86 ± 0.18	0.52 ± 0.015
Visible	1.3 ± 0.46	1.4 ± 0.23	0.23 ± 0.049	0.2 ± 0.084	0.086 ± 0.0059
Window	0.16 ± 0.16	0.22 ± 0.084	0.018 ± 0.012	0.11 ± 0.054	0.032 ± 0.0037

Table A.11: hh -channel cut flow. (*) factorized cuts, errors are only statistical.

Appendix B

VBF Cut flow

Table B.1 lists the cuts and thresholds applied to estimate the overlap between the analysis presented in this analysis and the VFB one.

Channel	Selection
lh channel	$E_T^{\text{miss}} > 20 \text{ GeV}$ $M_T(l, E_T^{\text{miss}}) < 30 \text{ GeV}$ $N_{jets} \geq 2$ with the p_T of the leading jet above 20 GeV $\eta_{j,1} \times \eta_{j,2} < 0$ $\min(\eta_{j,1}, \eta_{j,2}) < \eta_l, \eta_\tau < \max(\eta_{j,1}, \eta_{j,2})$ $\Delta\eta_{jj} > 3.6$ $M_{jj} > 500 \text{ GeV}$ no other jets with $p_T > 20 \text{ GeV}$ in $ \eta < 3.2$
ll channel	$E_T^{\text{miss}} > 20 \text{ GeV}$ $N_{jets} \geq 2$ with the p_T of the leading jet above 20 GeV $\eta_{j,1} \times \eta_{j,2} < 0$ $\min(\eta_{j,1}, \eta_{j,2}) < \eta_l, \eta_\tau < \max(\eta_{j,1}, \eta_{j,2})$ $\Delta\eta_{jj} > 3.6$ $M_{jj} > 500 \text{ GeV}$ no other jets with $p_T > 20 \text{ GeV}$ in $ \eta < 3.2$
hh channel	$E_T^{\text{miss}} > 40 \text{ GeV}$ $M_T(hh) < 80 \text{ GeV}$ $N_{jets} \geq 2$ with the p_T of the leading jet above 40 GeV $\eta_{j,1} \times \eta_{j,2} < 0$ $\min(\eta_{j,1}, \eta_{j,2}) < \eta_l, \eta_\tau < \max(\eta_{j,1}, \eta_{j,2})$ $\Delta\eta_{jj} > 4.0$ $M_{jj} > 700 \text{ GeV}$ no other jets with $p_T > 20 \text{ GeV}$ in $ \eta < 3.2$ $p_T(j_1 + j_2 + \tau_1 + \tau_2 + E_T^{\text{miss}}) < 60 \text{ GeV}$

Table B.1: Summary of the selection applied to implement the VBF analysis.

Appendix C

Tau ID Factorization

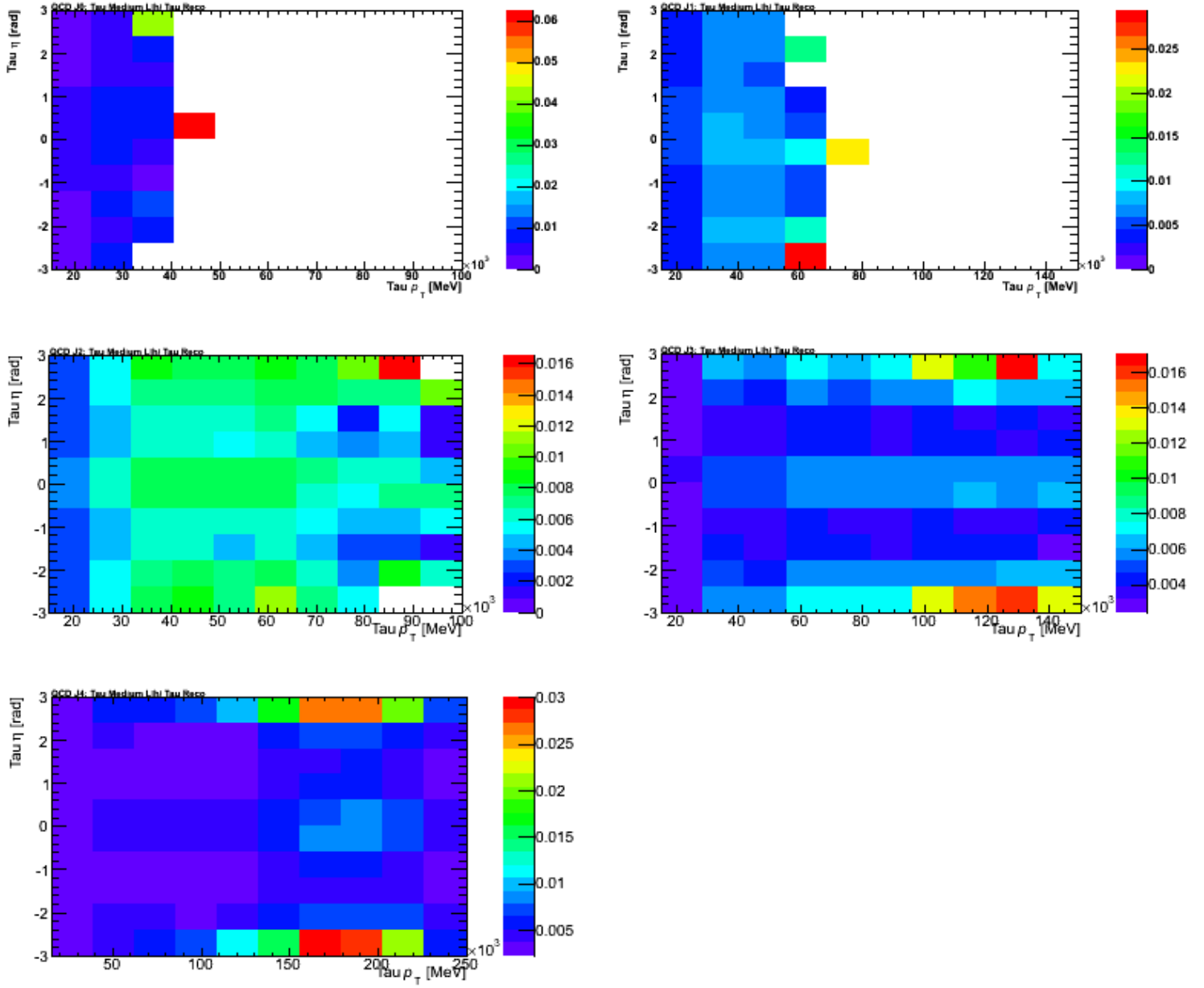


Figure C.1: Tau identification efficiencies in QCD dijets for Medium Likelihood. Top left J0, top right J1, center left J2, center right J3, bottom J4.