



Scuola Dottorale in Scienze Matematiche e Fisiche

XXVII Ciclo di Dottorato

Associated production of the Higgs boson with a W boson in
proton-proton collisions: an explorative analysis of the
three-leptons final state with the ATLAS experiment

Dottorando:
Monica Trovatelli

Docenti Guida:
Prof. Filippo Ceradini

Coordinatore:
Prof. Roberto Raimondi

Prof. Fabrizio Petrucci

Estuvo varios días como hechizado, repitiéndose a sí mismo en voz baja un sartal de asombrosas conjeturas, sin dar crédito a su propio entendimiento. Por fin, un martes de diciembre, a la hora del almuerzo, soltó de un golpe toda la carga de su tormento. Los niños habían de recordar por el resto de su vida la augusta solemnidad con que su padre se sentó a la cabecera de la mesa, temblando de fiebre, devastado por la prolongada vigilia y por el encono de su imaginación, y les reveló su descubrimiento:

”La tierra es redonda como una naranja.”

Úrsula perdió la paciencia: ”Si has de volverte loco, vuélvete tú solo” gritó. ”Pero no trates de inculcar a los niños tus ideas de gitano.”

GABRIEL GARCÍA MÁRQUEZ,
Cien años de soledad (1967)

Acknowledgments

I would like to express my sincere gratitude to my advisor Professor Filippo Ceradini, for his continuous support of my PhD research work; his guidance was for me extremely important and his advices precious.

A special thanks goes to Professor Fabrizio Petrucci, who encouraged me during all these years and taught me how to pursue my own research approach. He was always patient as no one else, calming me down whenever anxious (and it doesn't happen rarely!) and motivating me when things became hard. I sincerely thank him for all the nights he spent to help me with my research work, and for all the enjoyable (and often fruitful) coffee breaks we had together.

Thanks to my friend and colleague Marta, for sharing with me this experience and for joining me in all the shopping afternoons.

Thanks to Cecilia and Valerio ("Mr. simpatia"), for all the nice discussions we had and for those I hope we will have in the future.

Thanks to Daniele, for always helping me when needed (and for having tried to teach me, unfortunately with no success, how to save money!), and thanks to Marco, a sincere friend I am glad to have met.

Thanks to Monica, for having been my dear friend for so many years.

Thanks to Mauro, Michela, Paolo, Giuseppe, Domizia, Toni, Ada, Antonio and to all the friends and colleagues in Rome for having made me feel part of a big family.

Of course thanks to Lorenzo, since he was (and incredibly still is!) able to stay and live with me. He is my confidant and my first supporter, thanks for enjoying with me my successes (but also for comforting me for the defeats).

Last but not least, thanks to all my family, for having always believed this was possible, especially to my grandparents, for those who are still here and for those who are not anymore: thanks for being so proud of me, I am also proud of being your granddaughter!

Contents

Introduction	9
1 Higgs boson discovery at LHC	11
1.1 The Standard Model	11
1.2 The Brout-Englert-Higgs mechanism	14
1.2.1 Bosons masses	16
1.2.2 Fermions masses	16
1.3 The Higgs boson at the LHC	16
1.3.1 Production	17
1.3.2 Decay	19
1.3.3 The Higgs boson discovery	21
1.3.4 Higgs boson properties	23
1.4 The VH process at hadron colliders	28
2 The ATLAS detector	31
2.1 The Large Hadron Collider	31
2.2 The ATLAS experiment	33
2.2.1 ATLAS coordinate system	34
2.2.2 The Inner detector	34
2.2.3 The Calorimeter system	37
2.2.4 The Muon Spectrometer	39
2.2.5 ATLAS Trigger	40
3 Physics objects definition and reconstruction	43
3.1 Track reconstruction	43
3.1.1 Vertex finding algorithm	45
3.2 Leptons	46
3.2.1 Electron reconstruction	46
3.2.2 Muon reconstruction	47
3.3 Jet reconstruction	49
3.3.1 b-tagging algorithm	50
3.4 Tau reconstruction	51
3.5 Missing transverse momentum	54
4 Analysis	56
4.1 Physics process	56
4.1.1 Signal yield	56
4.1.2 Background composition	58
4.2 Data sample and data taking conditions	59
4.2.1 Data preparation	61

4.3	Monte Carlo samples	61
4.4	Analysis strategy	64
4.5	Cut-based analysis	64
4.5.1	Event selection	65
4.5.2	Selection Optimization	68
4.5.3	Blind criteria	75
4.6	Data-MC comparison	76
4.6.1	Control regions	76
4.6.2	NF evaluation	81
4.6.3	Data/MC agreement in CR	82
4.6.4	Data/MC agreement in blinded SR	84
4.6.5	Unblinding: cut-based analysis results	88
4.7	Systematic uncertainties	90
4.7.1	Theoretical uncertainties	91
4.7.2	Experimental uncertainties	91
5	MVA analysis	93
5.1	Event classification in a statistical perspective	93
5.2	Training and classification	94
5.3	BDT	94
5.4	BDT against leading backgrounds	95
5.4.1	WH vs Z+jets	95
5.4.2	WH vs Top	101
5.4.3	WH vs WZ/ γ^*	103
6	Results	106
6.1	Cut-based analysis results	108
6.2	MVA analysis results	109
6.3	Upper limit on VH cross section	112
6.4	Combination with other analysis	112
7	Prospects for LHC Run 2	113
	Conclusions	118
	Appendix	119
A	Data format and reduction	120
B	Tau-Jet fake rate	122
B.1	Fake rate measurement	122
C	Spare Cutflows	124
	Bibliography	127

Introduction

Since from the ancient era, man had the curiosity to know what we are made of, and looked for an explanation of everyday phenomena. Most of the questions found an answer in the early 20th century, when studies carried on by brilliant scientists gave birth to the Standard Model (SM) of particle physics. The SM is an elegant theory which explains, in a remarkable way, the structure of matter and the nature of fundamental interactions. In this theory particles divide in matter constituents, the *fermions*, and particles which mediate the interactions between *fermions*, the *bosons*. Experiments carried on at the particle colliders, such as the LEP at the CERN laboratory or the Tevatron at the Fermilab laboratory, have confirmed all the SM predictions with a high level of accuracy, also considering the wide energy range scanned. However, one important piece was missing until its recent discovery at the Large Hadron Collider (LHC), the *Higgs Boson*, the particle responsible for the mass of the other SM particles, whose existence is postulated in the Brout-Higgs-Englert (BEH) mechanism. On the 4th July 2012 the ATLAS and CMS experiments at the CERN laboratory announced the observation of a new particle, with a mass around 125 GeV having, so far, all the characteristics of the Higgs boson. In 2013, the Nobel Prize in Physics was awarded jointly to Peter Higgs and François Englert for their theoretical discovery confirmed by the ATLAS and CMS experiments. The detailed studies of couplings and properties of this new particle and the comparison with the expectations for the SM Higgs boson are the current focus of the ATLAS Collaboration. This thesis lies in this context, and presents the work done by the author within the ATLAS Collaboration, aiming at studying the Higgs boson production and decay in the channel $WH \rightarrow WWW^* \rightarrow l\nu l\nu\tau\nu$, ($l = e/\mu$), using the proton-proton collision data collected in 2012.

The Higgs boson production in association with a vector boson offers the possibility to measure the coupling of the newly discovered particle with the W and Z bosons. These couplings are predicted by the SM but have not been measured yet; in particular, the channel studied in this thesis, with the Higgs boson further decaying in two W bosons, allows to directly probe the Higgs boson coupling exclusively with W bosons. The analysis here reported is an explorative study of the three leptons final state in the WH channel, one of the leptons being a hadronically decaying tau. The presence of a hadronic tau makes this channel challenging to study at the LHC, because of the high-jet activity at a hadron collider. However, the measure presented is interesting as a feasibility study; the main issues of the measurement are addressed and the foundations for a similar study in LHC Run 2 are set.

This thesis is structured as follows. Chapter 1 introduces the Standard Model and the Brout-Englert-Higgs mechanism, the theory in which this study is embedded. Some recent results obtained by the ATLAS Collaboration about the measurements of spin, mass and couplings of the new particle are also reported. Chapter 2 gives an overview of the ATLAS detector, which is the experimental apparatus used to collect data which this thesis is based on. In chapter 3 a description of the lepton, jet and event identification and reconstruction techniques is presented. Chapter 4 presents the cut-based analysis done in the study of the $WH \rightarrow WWW^* \rightarrow l\nu l\nu\tau\nu$ process. A detailed description of the event selection using MC samples is reported, together with the procedure used to normalize these MC samples to reproduce what observed in data.

Due to the low sensitivity of the analysis, in chapter 5 an alternative approach to the cut-based analysis is presented: the multivariate analysis, aimed at reducing the background processes that mimic the signal topology while keeping as much signal events as possible. Chapter 6 reports and discusses the statistical analysis done to compare the results with the SM expectations. Finally in chapter 7 prospects for this measurement in LHC Run 2, when increased luminosity and centre-of-mass energy will be available, are given.

Chapter 1

Higgs boson discovery at LHC

The recent discovery of the Higgs boson at LHC (2012) was a big step towards the confirmation of the Standard Model as the theory that describes the sub-atomic particles and their interactions. In this chapter, after a brief review of the Standard Model theory, including the formulation of the *simmetry breaking mechanism* and the first attempts to search for the Higgs boson at the CERN electron-positron (LEP) and proton-antiproton (Tevatron) colliders, we will go through the main steps that led to the Higgs boson discovery.

1.1 The Standard Model

The Standard Model (SM) is the theory that accurately describes the elementary particles and their interactions as shown in the experiments. It is a renormalizable field theory that has been developed during the 60's thanks to the work of many people [1][2][3][4][5]. At the moment three of the four forces observed in nature are described by the SM theory: the electromagnetism, the weak and the strong nuclear force. Any attempt to allocate the gravity in the SM has failed so far. In the SM two type of point-like particles exist: *fermions*, spin-1/2 particles that are matter content, further divided in *leptons* and *quarks*, and the *bosons*, integer spin particles which are the mediators of the interaction fields. Fermions interacts by the exchange of spin-1 bosons: eight massless gluons and one massless photon for the strong and the electromagnetic interactions, respectively, and three massive bosons, $W^\pm$ and Z for the weak interaction. Fermions are organized in a three-fold family structure, as shown in figure 1.1 which summarizes the current knowledge of the sub-atomic world.

The SM is a non-Abelian Gauge theory based on the simmetry group $SU(3) \times SU(2) \times U(1)$, where $SU(3)$ is the non-Abelian gauge group of the *Quantum Chromodynamics* (QCD) [6], the theory describing the interaction of quarks and gluons due to colour charge, while the $SU(2) \times U(1)$ is the non-Abelian symmetry group of the combined electromagnetic and weak forces (electroweak). The SM lagrangian can then be written as:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} \quad (1.1)$$

It is invariant under a $SU(3) \times SU(2) \times U(1)$ local gauge symmetry, where "local" means that the transformation depends on the specific space-time point x . Both the QCD and the EW theories are built by requiring a local invariance of the Dirac free lagrangian¹ for the elementary matter fields, quarks in the first case, fermions in the second case. The local gauge invariance leads to

¹The Dirac free lagrangian for a generic matter field Ψ is:

$$\mathcal{L}_{free} = \bar{\Psi}(\gamma^\mu \partial_\mu - m)\Psi, \quad \bar{\Psi} \equiv \Psi^\dagger \gamma^0 \quad (1.2)$$

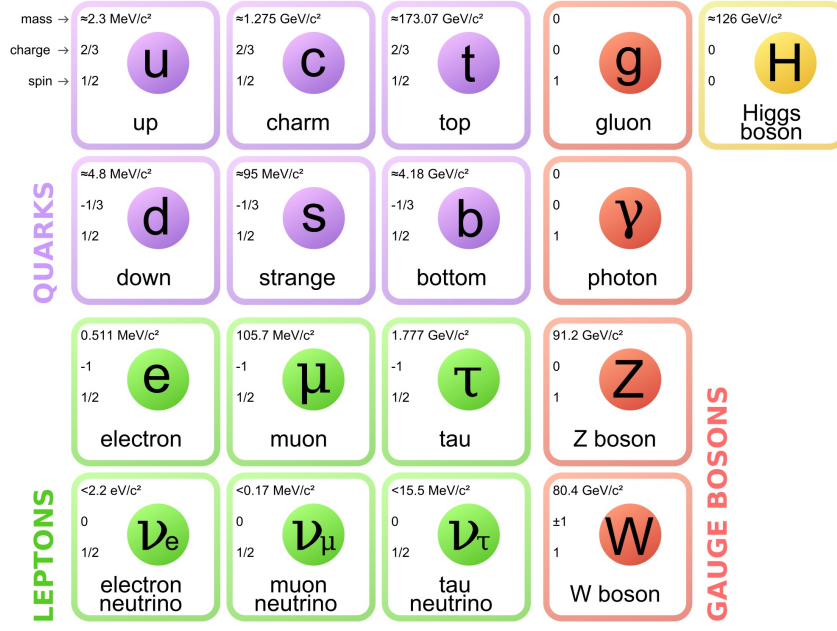


Figure 1.1: Schematic representation of the Standard Model theory.

the addition of new quantum fields in the theory, as in case of the photon in electrodynamics. These quantum fields are responsible for the particles interactions.

Due to the non-Abelian nature of the strong force the gluons are self-interacting, while photons don't interact with each other. This results in the quark asymptotic freedom at high energy and strong interactions at low energy (quark confinement in hadrons).

The $SU(3)_c$ invariant QCD lagrangian, where C stands for colour symmetry, is:

$$\mathcal{L}_{QCD} \equiv -\frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a + i \sum_f \bar{q}_f \gamma^\mu D_{\mu\beta}^\alpha q_f^\beta \quad (1.3)$$

In this formula the $G_a^{\mu\nu}$ field tensor is given by:

$$G_a^{\mu\nu} = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f_{abc} G_\mu^b G_\nu^c \quad (1.4)$$

where G_ν^a corresponds to the gluon fields, with a running from 1 to $N_c^2 - 1 = 8$, with $N_c = 3$ (quarks come in three colours). f_{abc} are the $SU(3)$ structure constants and g_s is the strong coupling constant. In the second term of equation (1.3) q_f is the quark field of flavour f with α and β as colour indices; $D_{\mu\beta}^\alpha$ is the covariant derivative defined as:

$$D_{\mu\beta}^\alpha = \partial_\mu \delta_\beta^\alpha + \frac{i}{2} g_s \sum_a G_\mu^a \lambda_\beta^{a,\alpha} \quad (1.5)$$

where λ^a are the generator matrices of $SU(3)$. The lagrangian in equation (1.3) then describes the quarks q_f interacting by means of gluons; the first term describes the gluon dynamics, including the self-interacting gluon term.

The weak interaction mediates transitions between fermions. The EW lagrangian describes the $SU(2)_L \times U(1)_Y$ group where $SU(2)_L$ group refers to the weak isospin I and the $U(1)_Y$ group refers to the weak hypercharge Y. In this picture the fundamental constituents of matter are arranged in doublets of left-handed and singlets of right-handed fermions:

$$L_i = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \quad \begin{matrix} I_3 = +1/2 \\ I_3 = -1/2 \end{matrix} \quad Y = -1 \quad (1.6)$$

$$Q_i = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b \end{pmatrix}_L \quad I_3 = +1/2 \quad Y = +1/3 \quad (1.7)$$

$$l_{R,i} = e_R^-, \quad \mu_R^-, \quad \tau_R^-, \quad I_3 = 0 \quad Y = -1 \quad (1.8)$$

$$u_{R,i} = u_R, \quad c_R, \quad t_R, \quad I_3 = 0 \quad Y = -4/3 \quad (1.9)$$

$$d_{R,i} = d_R, \quad s_R, \quad b_R, \quad I_3 = 0 \quad Y = -2/3 \quad (1.10)$$

In the above equations I_3 is the third component of the weak isospin. The weak ipercharge is related to the weak isospin through the following equation:

$$Y = 2 \left(\frac{Q}{e} - I_3 \right) \quad (1.11)$$

The request of the local gauge invariance leads to the introduction of four vector bosons: the W^i fields ($i=1,2,3$) for the $SU(2)_L$ group and the field B for the $U(1)_Y$ group. The physical fields A_μ (photon field), Z_μ (the field associated to the neutral boson Z^0) and $W^\pm$ (the fields describing the two charged bosons) can be obtained by a combination of the gauge fields:

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \quad (1.12)$$

$$Z_\mu = W_\mu^3 \cos \theta_W - B_\mu \sin \theta_W \quad (1.13)$$

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad (1.14)$$

In the above equations the angle θ_W , which specifies the mixture of Z_μ and A_μ fields in W_μ^3 and B_μ is known as the mixing angle. The weak mixing angle θ_W also relates the masses of the weak bosons, as shown in section 1.2.

The analytic form of the EW lagrangian is:

$$\mathcal{L}_{EW} = -\frac{1}{4} \sum_G F_G^{\mu\nu} F_{\mu\nu} + i \sum_f \bar{f} D_\mu \gamma^\mu f \quad (1.15)$$

where the index G indicates that the first sum in equation (1.15) is extended to all the vectorial fields, while the index f indicates that the second sum is extended to all the fermionic fields. The first term in equation (1.15) describes the dynamics of the bosons, while the second term the interaction between fermions, interaction that is mediated by the four bosons. The interaction between fermions and bosons can be derived by writing down the definition of the covariant derivative:

$$D_\mu = \partial_\mu - ig_G(\lambda^\alpha G_\alpha)_\mu \quad (1.16)$$

where g_G is the coupling constant to the G field ($G = A, Z, W^\pm$ and λ^α are the generators of the group to which the G field refers ($SU(2)$ or $U(1)$)).

The SM lagrangian as written above is gauge invariant but it doesn't contain any mass term for fermions and bosons. This contradicts the experimental evidence that, apart for the photon, the particles that we observe have a non-zero mass. Any attempt to include ad-hoc mass terms in the lagrangian spoils the gauge invariance and the renormalizability of the theory. In the mid-1960 a couple of different works carried on by several theoreticians ([7][8][9][10]) tried to explain the origin of particle masses; these works showed how the gauge invariance of the lagrangian can be preserved by invoking the *spontaneous EW lagrangian symmetry breaking* (also known as *the Brout-Englert-Higgs (BEH) mechanism*).

1.2 The Brout-Englert-Higgs mechanism

The BEH mechanism is the generalization of the Goldstone model (details can be found in [11]) to the case of a lagrangian invariant for a local phase transformation. In this mechanism the assumption is made that everywhere in space, fluctuations in the vacuum can occur which correspond to the emission or the absorption of a Higgs boson, a spin 0, electrically neutral particle with no colour charge. As a result of their interactions with the Higgs field, the $W^\pm$ and Z^0 bosons and the fermions acquire mass, but gluons and photons remain massless. The choice of a specific vacuum state results in the spontaneous symmetry breaking of the local $SU(2) \times U(1)$ gauge symmetry and gives rise to the spectrum of particles we observe. Spontaneous symmetry breaking is relevant in a field theory only if the ground state is not-unique. In the following the BEH mechanism is briefly derived starting from the Goldstone model; a detailed description can be found in [7][8][9][10].

The simplest example of a field theory exhibiting the spontaneous symmetry breaking is the Goldstone model. The Goldstone model is the model of a real scalar field ϕ with a lagrangian given by

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - V(\phi) \quad (1.17)$$

with

$$V(\phi) = \mu^2\phi^2 + \lambda\phi^4 \quad (1.18)$$

λ and μ^2 are arbitrary real parameters. The first term in equation (1.17) is positive defined and vanishes for constant ϕ . It follows that the minimum of the total energy of the field corresponds to the minimum of the potential $V(\phi)$. To guarantee the existence of a ground state for such a potential $\lambda > 0$ is also requested. For positive values of μ^2 , the minimum of the potential is at $\phi = 0$. However another situation can occur, in case $\mu^2 < 0$ the potential possesses a local minimum at $\phi(x) = 0$ and a whole circle of absolute minimum at

$$\phi(x) = \phi_0 = \left(\frac{-\mu^2}{2\lambda}\right)^{1/2} e^{i\theta} \quad (1.19)$$

The vacuum expectation value is one of those minima ϕ_0 . The angle θ runs from 0 to 2π and

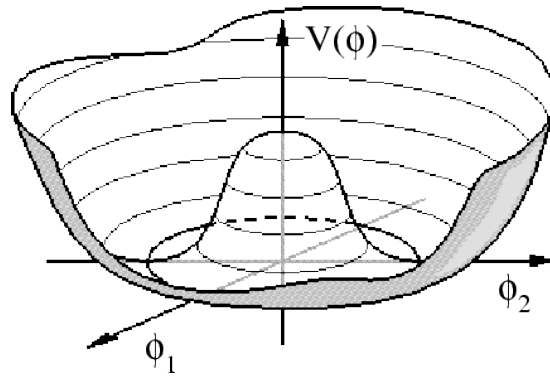


Figure 1.2: Shape of the Higgs potential for $\mu^2 < 0$.

defines a direction in the ϕ -plane. Spontaneous symmetry breaking will occur if we choose one particular direction θ to represent the ground state. Since the lagrangian in equation (1.17) is global invariant we can arbitrary choose the value of θ ; we then take $\theta = 0$ so that:

$$\phi(x) = \phi_0 = \left(\frac{-\mu^2}{2\lambda}\right)^{1/2} = \frac{1}{\sqrt{2}}v \quad (1.20)$$

is real.

This is how the spontaneous symmetry breaking occurs. In the following we will show how this concept can be applied to the SM and how the spontaneous breaking can generate the masses of bosons and fermions.

The simplest way to introduce the spontaneous symmetry breaking in the SM lagrangian is by adding a new $SU(2)_L$ doublet of complex scalar field (called the Higgs field):

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.21)$$

The lagrangian of this scalar doublet is

$$\mathcal{L}_H = (D^\mu \Phi)^\dagger (D_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \quad (1.22)$$

where the covariant derivative is

$$D^\mu = \partial^\mu + \frac{i}{2} g \tau_j W_j^\mu + \frac{1}{2} g' Y B^\mu \quad (1.23)$$

where the sum over the index $j = 1, 2, 3$ is implied, τ_j are the Pauli matrices, g and g' are the coupling constants of fermions to the W^μ and B^μ respectively and Y is the weak hypercharge operator. This lagrangian contains the symmetric potential in figure 1.2, which has again a vacuum expectation value different from zero, which can be chosen to be

$$\Phi_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \quad (1.24)$$

where $v = (-\mu^2/\lambda)^{1/2}$. Equation (1.24) states that the ground state of the $V(\Phi)$ potential occurs for a non-vanishing value of the Φ field. The ground state is not symmetric under $SU(2)_L \times U(1)_Y$ transformation since there is a preferred direction, and the symmetry is spontaneously broken. To understand the physical content of this mechanism we expand the lagrangian perturbatively around its ground state. In general we can express the Φ field around the ground state as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \sigma(x) \end{pmatrix} \quad (1.25)$$

The gauge in which the Higgs field has the above form is called the *unitary gauge*. In this gauge the imaginary part of the complex Φ field can be eliminated through a local transformation of the field. The Φ field has then become real. By substituting the expanded expression for the Higgs field in the lagrangian in equation (1.22) we find

$$\mathcal{L} = \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma + 2\mu^2 \sigma^2 - 2\sqrt{-2\lambda\mu^2} \sigma^3 - \lambda \sigma^4 + \text{const.} \quad (1.26)$$

The real field $\sigma(x)$ measures the deviation of the field $\Phi(x)$ from the equilibrium ground state configuration $\Phi(x) = \Phi_0$. Equation (1.26) can be interpreted as the lagrangian of a scalar field $\sigma(x)$ with mass $\sqrt{2\lambda}v$. The equation includes a cubic term that breaks the symmetry of the potential in picture 1.2 (the potential is anymore invariant under the transformation $x \rightarrow -x$). The Higgs field $\Phi(x)$ describes a scalar neutral particle, the Higgs boson, of mass

$$m_H = \sqrt{2}\mu = \sqrt{2\lambda}v \quad (1.27)$$

The value of m_H depends on μ and it is a free parameter of the SM.

1.2.1 Bosons masses

Bosons masses originate from the interaction of the $SU(2) \times U(1)$ gauge fields (W^μ and Z^μ) with the Higgs field. This interaction takes place through the covariant derivative D^μ in equation (1.23). Substituting the covariant derivative in the scalar lagrangian of equation (1.22) one gets for the kinetic term:

$$(D_\mu \Phi)^\dagger D^\mu \Phi \rightarrow \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma + (v + \sigma)^2 \left(\frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right) \quad (1.28)$$

Both the $W^\pm$ and the Z^0 bosons have acquired mass, since they appear in the previous formula as quadratic terms. The mass of the $W^\pm$ and the Z^0 bosons are related through the equation:

$$M_Z \cos \theta_W = M_W = \frac{1}{2} v g \quad (1.29)$$

and given that

$$\tan \theta = \frac{g'}{g} \quad (1.30)$$

we can also write

$$M_W = \frac{v g}{2} \quad (1.31)$$

$$M_Z = \frac{v}{2} \sqrt{g^2 + g'^2} \quad (1.32)$$

1.2.2 Fermions masses

Fermions masses are generated by coupling the Higgs doublet and the fermions. The additional Yukawa term to add at the SM lagrangian has the form

$$\mathcal{L} = -g_\psi (\bar{\psi}^L \Phi \psi^R) + h.c. \quad (1.33)$$

where g_ψ is the coupling constant of the fermionic field ψ to the Higgs field. ψ^L and ψ^R are the left- and right-handed fermion fields respectively. Expanding again equation (1.33) around the ground state of the Higgs field we can derive the fermion mass term:

$$m_\psi = g_\psi v / \sqrt{2} \quad (1.34)$$

From equations (1.29) and (1.34) it is possible to note that bosons and fermions masses strongly depend on the value of the parameter v , as well as on the m_H . It can be shown [12] that the parameter v is related to the Fermi constant G_F through

$$v = (\sqrt{2} G_F)^{-1/2} \approx 246 \text{ GeV} \quad (1.35)$$

This in the past allowed to predict the mass of the $W^\pm$ bosons and the mass of the Z^0 boson, before they were discovered at the UA1 and UA2 experiments. In the SM framework the Higgs boson self-coupling parameter λ is then the only free parameter of the theory; because of the equation (1.27) also the Higgs boson mass value was unknown until its discovery at LHC. As shown in the next section, the Higgs boson production and decay modes depend on its mass; for this reason some decades ago people started to look for the Higgs in the various production and decay channels, since each mode was in principle possible and accessible at a given energy.

1.3 The Higgs boson at the LHC

Decades after the Higgs theory was established, the Higgs boson was discovered in 2012 by the ATLAS and CMS collaborations [13][14]. In this section production and decays modes of the Higgs boson at a hadron collider are discussed.

1.3.1 Production

As shown in the previous section, the Higgs boson couples to bosons and fermions with different couplings; in particular the coupling with fermions is proportional to the mass of the fermion, while the coupling with the bosons is proportional to the square mass of the boson. The Higgs-fermions coupling $g_{Hf\bar{f}} = \frac{m_f}{v}$ is of the order of m_f/m_W , and it is weak for $m_f \ll m_W$; this condition is satisfied for neutrinos, electrons, muons and the light quarks (u,d,s). Hence, although is in principle possible for the Higgs boson to be produced by these particles, the production cross sections are very small. The value of the Higgs-boson coupling $g_{HVV} = \frac{2m_V^2}{v}$, being proportional to the square mass of the boson V which the Higgs boson couples to, makes this production mode more probable, thanks to the fact that these bosons are very massive with respect to fermions. The four main Higgs production modes are summarized in figure 1.3. Their relative importance depends on the centre-of-mass energy; at LHC running with a centre-of-mass energy of 8 TeV they are, from the most to the least probable:

Gluon-gluon fusion (ggF): it is the dominant Higgs boson production mode. Two gluons from the colliding protons couple to the Higgs boson through a fermion loop; because of their mass, top quarks are the most probable to contribute to the loop. This production channel has no distinctive experimental signature; it can be detected only with a clear identification of the Higgs boson decay products.

Vector boson fusion (VBF): it is the second Higgs boson production mode at the LHC, with a cross section a factor 10 smaller than the ggF. Two quarks from the colliding protons emit two virtual bosons V , which in turn fuse to produce the Higgs boson. The process is characterized by the emission of two *high* $-p_T$ jets, called *tagging jets*, directed predominantly in the forward region.

Associated production with a vector boson (VH, $V=W/Z$): it is known as the *Higgsstrahlung* process. In fact it consists in a quark-antiquark annihilation to create a vector boson V , with $V=W/Z$, which then radiates an Higgs boson. Although the cross section for this process is low compared to the ggF and VBF processes, it gives the chance to directly test the coupling of the Higgs boson with a vector boson, and so to test the SM predictions.

A study of the VH associated production in final states with three leptons, one being a tau, is the subject of this thesis. For this reason the last paragraph of this chapter will contain a brief review of the VH process at hadron colliders.

Associated production with heavy quarks ($t\bar{t}H$): it is the least probable Higgs boson production mode at the LHC. The initial state gluons exchange a top quark from which a Higgs boson is produced. This process offers the possibility to measure the top Yukawa coupling. Even though the latter is large, thanks to the high mass of the top quark, the heavy $t\bar{t}H$ final state is kinematically suppressed. The cross section results to be a factor 100 smaller than ggF cross section.

In figure 1.4 the cross section values for the production modes listed above are shown as a function of the Higgs boson mass. The exact values, for a centre-of-mass energy of 8 TeV and for the measured value of the Higgs boson mass $m_H = 125$ GeV, are reported in table 1.1 and compared with the 13 TeV values, 13 TeV being the centre-of-mass energy expected in the second data-taking run of LHC (Run 2). Run 2 starting date is scheduled for March 2015.

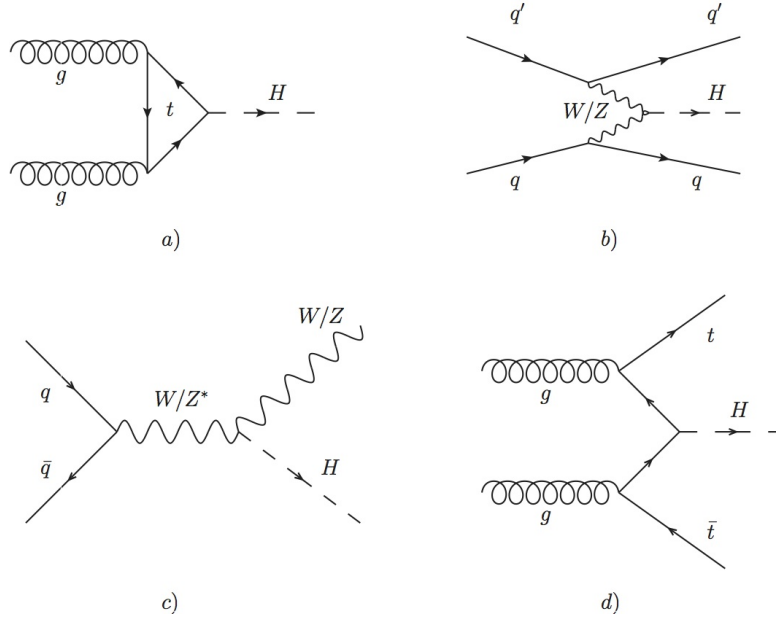


Figure 1.3: Tree-level Feynman diagrams for the Higgs production modes: (a) the gluon-gluon fusion, (b) the Vector Boson Fusion, (c) the associated production with a $W^\pm$ or a Z boson, (d) the $t\bar{t}$ associated production.

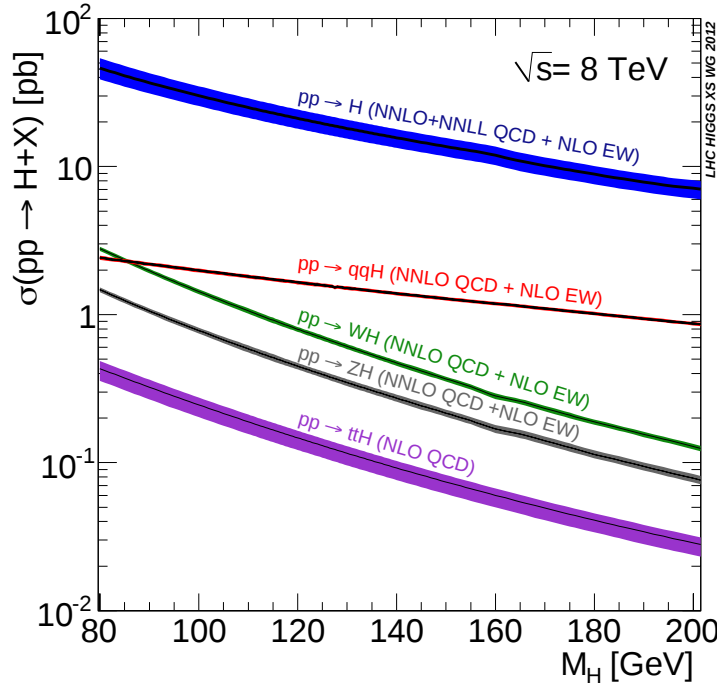


Figure 1.4: Cross section values for the Higgs boson production modes in proton-proton collisions at a centre-of-mass energy of 8 TeV, as a function of the Higgs boson mass [15]. Higgs boson masses up to 200 GeV are shown. The bands include theoretical uncertainties.

Higgs production cross sections in pb at LHC for $m_H = 125$ GeV					
$\sqrt{s}$	ggF	VBF	WH	ZH	$t\bar{t}H$
8 TeV	$19.3^{+10\%}_{-10\%}$	$1.6^{+3\%}_{-3\%}$	$0.7^{+3\%}_{-3\%}$	$0.4^{+4\%}_{-4\%}$	$0.1^{+9\%}_{-12\%}$
13 TeV	$44.0^{+10\%}_{-10\%}$	$3.8^{+5\%}_{-5\%}$	$1.4^{+3\%}_{-4\%}$	$0.9^{+4\%}_{-4\%}$	$0.5^{+11\%}_{-13\%}$

Table 1.1: Cross section values (in pb) for the main Higgs production modes at LHC, for a centre-of-mass energy of 8 TeV and 13 TeV. All the cross section values, except for the $t\bar{t}H$, are computed at NNLO in perturbation theory for the QCD corrections [16][17][18], and at NLO for the EW corrections. The $t\bar{t}H$ cross-section which is computed at NLO in QCD. The quoted uncertainty has been computed by adding in quadrature the error obtained by varying the QCD scale and that obtained by varying the PDF set.

1.3.2 Decay

Since the couplings of the Higgs boson are proportional to masses, as m_H increases the Higgs particle becomes strongly coupled. This reflects in the sharp rise of the Higgs boson total width, shown in figure 1.5 as a function of the Higgs mass. In figure 1.6 the Higgs boson branching ratios are shown; the values for $m_H = 125$ GeV are reported in table 1.2.

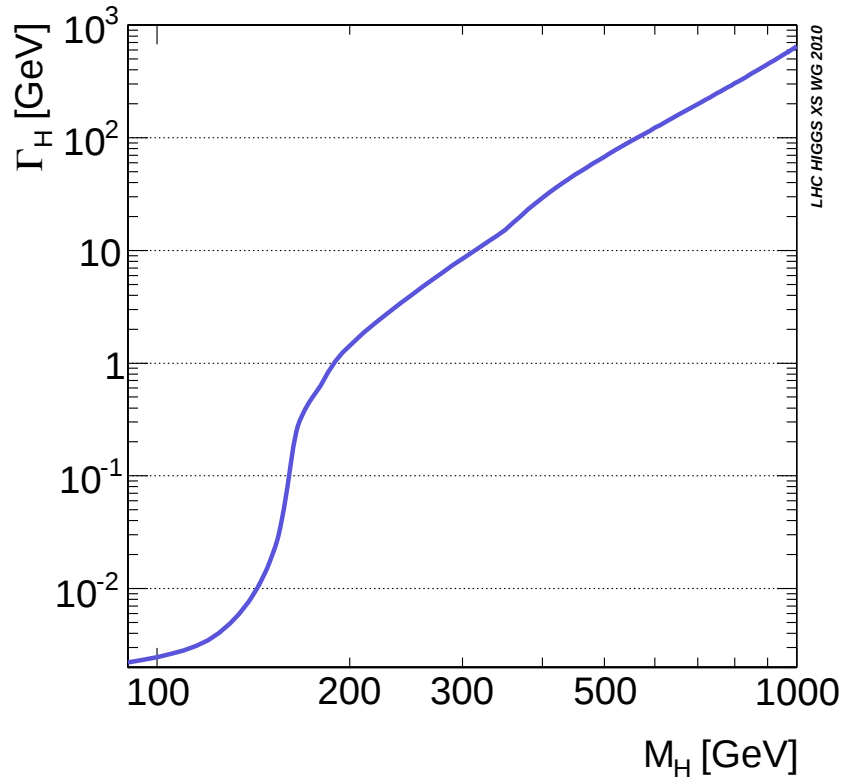


Figure 1.5: Higgs total width as a function of its mass [16].

Two Higgs boson mass region can be distinguished, the low mass region for $m_H \lesssim 135$ GeV, and the high mass region, for $m_H > 135$ GeV. In the low mass region the dominant decay mode is the $H \rightarrow b\bar{b}$, which is hard to be identified at LHC because of the high multi-jets background. The $H \rightarrow \gamma\gamma$ proceeds via a W loop; despite the branching ratio of this process is very low compared to the $b\bar{b}$ mode, the clean signature of two photons made this channel one of the

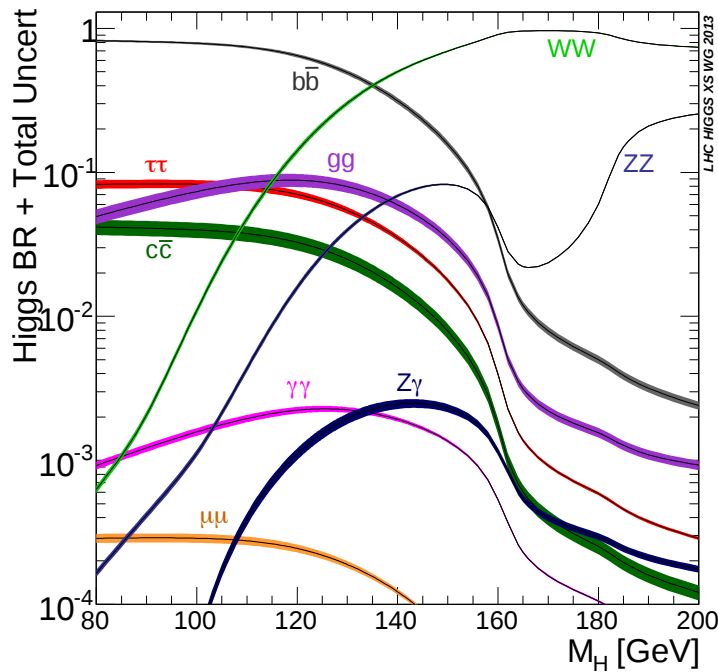


Figure 1.6: Higgs branching ratios as a function of its mass [18]. The bands represent the theoretical uncertainties.

Higgs branching ratios for $m_H = 125$ GeV				
$H \rightarrow b\bar{b}$	$H \rightarrow \tau\tau$	$H \rightarrow \gamma\gamma$	$H \rightarrow WW$	$H \rightarrow ZZ$
0.577	0.064	0.002	0.216	0.027

Table 1.2: Main SM Higgs boson branching ratio (BR) at LHC, for $m_H = 125$ GeV.

preferred for its discovery at LHC. The width sharply increases as soon as the WW threshold is approached. Below this threshold the decays into off-shell V particles is important, for example the $H \rightarrow WW^*$ decay. The dip of the ZZ branching ratio just below the ZZ threshold, in particular, is due to the fact that the W boson is lighter than the Z boson, and the opening of its threshold depletes all the other branching ratios. For $m_H > 160$ GeV both the $H \rightarrow WW$ and the $H \rightarrow ZZ$ modes are possible. These two channels had a leading role in the Higgs boson discovery (see section 1.3.3); leptonic vector boson decays were selected, which allowed to decrease the background contribution while keeping high the acceptance on the signal events.

1.3.3 The Higgs boson discovery

The Higgs boson has been the subject of many physics searches in the last decades at particle colliders. Before the advent of LHC, direct searches for the Higgs boson were carried on at the Large Electron-Positron collider (LEP), first, and then at the Tevatron proton-antiproton collider. A lower bound on the Higgs boson mass of 114.4 GeV at 95% CL has been set with LEP [19] data, while Tevatron studies reported an excess of events around $m_H = 125$ GeV with a significance of 3.0σ [20], mainly from searches in the $VH \rightarrow Vb\bar{b}$ channel.

The search for the Higgs boson culminated in its discovery on the 4th July 2012, when both ATLAS and CMS Collaborations reported an excess of events in the region 124-126 GeV, compatible with the existence of a SM Higgs boson of that mass. The significance of the excess obtained by combining the 8 TeV result with the previous 7 TeV result, was 4.9 and 5.0 standard deviations respectively, which in both cases is enough to claim the discovery of an Higgs-boson like particle [13][14]. In this section only ATLAS published results on the Higgs search and discovery are discussed.

The observation of the SM Higgs boson was possible thanks to the combination of the individual searches, carried on in the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^{(*)} \rightarrow 4l$ and $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ channels. In november 2013 also the evidence for the decays into fermions was obtained in the $H \rightarrow \tau\tau$ and $H \rightarrow b\bar{b}$ channels [21][22].

Four leptons: The $H \rightarrow ZZ^{(*)} \rightarrow 4l$ channel, where $l = e, \mu$, is also called the "golden-channel" for the Higgs boson discovery at LHC, since it has a clean signature and very small background contribution, even if it is characterized by a tiny cross section. It provides good sensitivity over a wide mass range (110-600 GeV), largely due to the excellent momentum resolution of the ATLAS detector (see chapter 2). The selection of four charged leptons in the final state allows to fully reconstruct the Higgs boson invariant mass. Data are compared with the expected distribution of the four leptons invariant mass m_{4l} for the background and for a Higgs boson signal with $m_H = 125$ GeV [23]; the result is shown in figure 1.7.

Two photons: Thanks to the excellent di-photon invariant mass resolution, the $H \rightarrow \gamma\gamma$ channel was one of the most important channels for the Higgs discovery at LHC. In fact it was possible to distinguish the peak due to the tiny expected signal over the huge diphoton background with a smooth distribution. The result obtained with the combination of the $\sqrt{s} = 7$ TeV and the $\sqrt{s} = 8$ TeV data is shown in figure 1.8. An excess of events around $m_H = 126.5$ GeV was observed [23].

WW channel: The process $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ is highly sensitive to a SM Higgs boson in the mass range around the WW threshold of 160 GeV. The signature for this channel is two opposite charge leptons with large transverse momentum and a large momentum imbalance in the event due to the escaping neutrinos. The presence of neutrinos in the final state doesn't allow the reconstruction of the Higgs invariant mass. For this reason

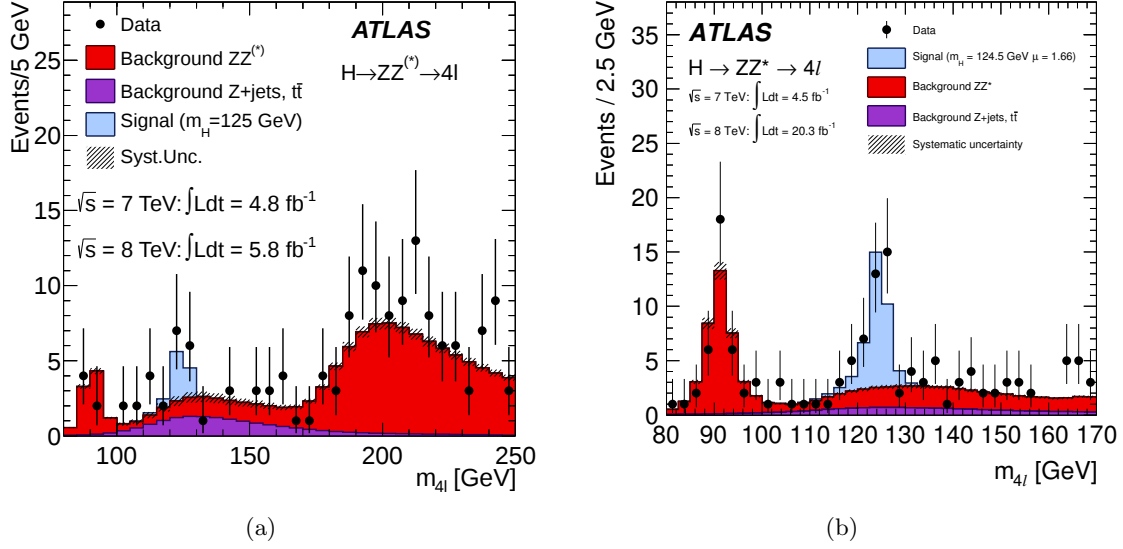


Figure 1.7: The distribution of the four-lepton invariant mass, $m_{4\ell}$, for the selected candidates, compared to the background expectation in the 80–250 GeV (left) 80–170 GeV (right) mass range, for the combination of the $\sqrt{s} = 7$ TeV and the $\sqrt{s} = 8$ TeV data. The signal expectation for a SM Higgs with $m_H = 125$ GeV is also shown. Figure (a) shows the distribution obtained in July 2012, when only a fraction of the 8 TeV data were analyzed, while figure (b) shows the results obtained with the full 8 TeV statistics [23].

the transverse mass, defined as $m_T = \sqrt{(\mathbf{E}_T^{\text{ll}} + \mathbf{E}_T^{\text{miss}})^2 - |\mathbf{p}_T^{\text{ll}} + \mathbf{E}_T^{\text{miss}}|^2}$, where $\mathbf{E}_T^{\text{ll}}$ and $\mathbf{p}_T^{\text{ll}}$ are respectively the transverse energy and the transverse momentum of the dilepton system, has been used for the data comparison with background expectations [24] (see figure 1.9).

Statistical interpretation of the excess

An excess of events was observed around $m_H = 126$ GeV in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels, both of which provided fully reconstructed candidates with high resolution in invariant mass. To quantify the significance of the excess, an hypothesis test has been performed. Hypothesis testing [25] is a tool used for decision making and for drawing conclusions based on an acquired set of measurements. In the search for the Higgs boson, the test aimed to understand if the observed excess of events was compatible with the Higgs boson predicted by the SM. More details about the statistical hypotheses can be found in section 6. The result of the test was given in terms of the p_0 value. The p_0 value is a measurement of the observed level of significance. It quantifies how often, if the measurements were repeated many times, one would obtain data as far away (or more) from the so-called *null hypothesis* as the observed data, assuming the *null hypothesis* to be true. The *null hypothesis* was the background-only one. In this way the p_0 value shown in figure 1.10 is the probability that a fluctuation in background events would reproduce, in absence of any signal from a Higgs boson, a number of events at least as large as the observed one. The observed local p_0 value from the combination of channels is shown in figure 1.10 as a function of m_H . The largest local significance for the combination of the 7 and 8 TeV data is found for a SM Higgs boson mass hypothesis of $m_H = 126.5$ GeV, where it reaches 6.0σ .

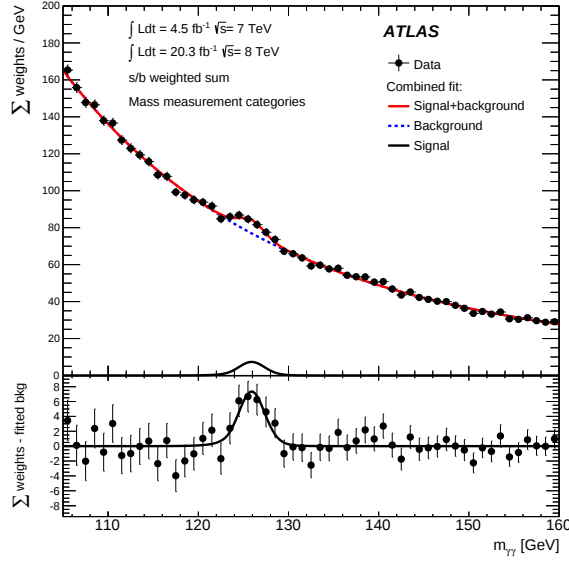


Figure 1.8: Distribution of the invariant mass of diphoton candidates for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data sample [23]. The result of a fit to the data of the sum of a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order polynomial is superimposed.

1.3.4 Higgs boson properties

After the Higgs boson discovery, the experimental challenge became the comparison of its properties with the SM predictions. In this section the ATLAS recent results on the Higgs boson mass and spin measurements, together with its production modes and couplings are discussed.

Mass

The ATLAS Collaboration used a model-independent approach to measure the Higgs boson mass based on fitting the mass spectra of the two decay modes $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$ [23]. In these two channels the invariant mass distribution produces a narrow peak over a smooth background, from which the mass can be extracted without assumptions on the signal production and decay yields. From fits to the mass spectra the two values are obtained:

$$m_H(\gamma\gamma) = 126.8 \pm 0.2(stat) \pm 0.7(syst) \text{ GeV} \quad (1.36)$$

$$m_H(ZZ^*) = 124.3^{+0.6}_{-0.5}(stat)^{+0.5}_{-0.3}(syst) \text{ GeV} \quad (1.37)$$

For the combined mass result a profile likelihood ratio $\Lambda(m_H)$ has been used:

$$\Lambda(m_H) = \frac{\mathcal{L}\left(m_H, \hat{\mu}_{\gamma\gamma}(m_H), \hat{\mu}_{4l}(m_H), \hat{\theta}(m_H)\right)}{\mathcal{L}\left(\hat{m}_H, \hat{\mu}_{\gamma\gamma}, \hat{\mu}_{4l}, \hat{\theta}\right)} \quad (1.38)$$

where the individual signal strengths $\mu_{\gamma\gamma}$ and μ_{4l} are treated as independent nuisance parameters in order to allow for the possibility of different deviations from the SM expectation in the two channels. The likelihood functions in the numerator and denominator of equation (1.38) are built using sums of signal and background probability density functions in the $\gamma\gamma$ and $4l$ mass spectra. The combined mass measurement is:

$$m_H = 125.36 \pm 0.37(stat) \pm 0.18(syst) \text{ GeV} \quad (1.39)$$

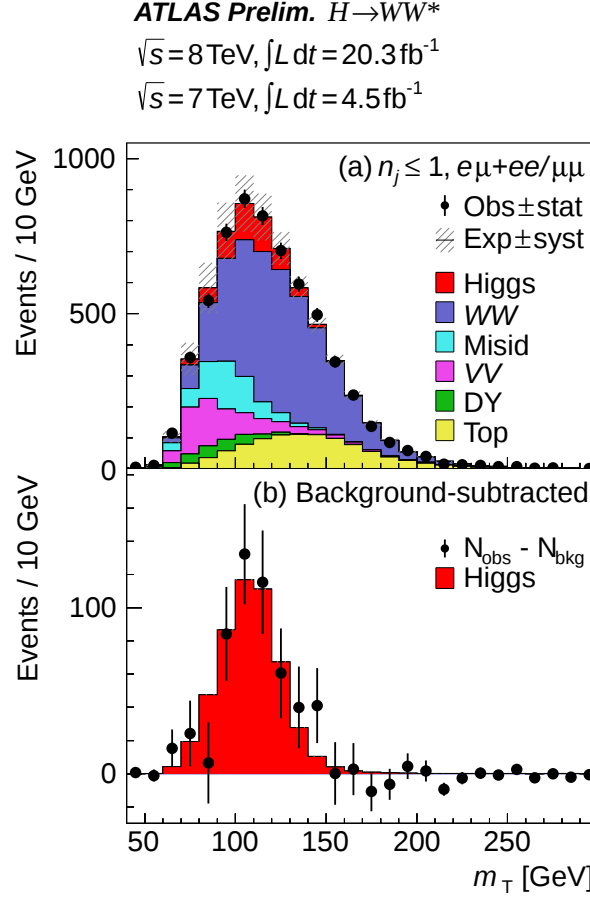


Figure 1.9: Combined transverse mass distributions of $n_{jets} \leq 1$ for all lepton-flavour samples in 7 and 8 TeV data [24]. The plot in (b) shows the residuals of the data with respect to the estimated background compared to the expected distribution for a SM Higgs boson with $m_H = 125$ GeV; the uncertainties on the data are statistical, i.e., $\sqrt{N_{\text{obs}}}$, and the uncertainty on the background (not shown) is up to about 25 events per m_T bin and partially correlated between bins. In both plots, background processes are scaled by post-fit normalization factors and the signal processes by the observed signal strength μ from the likelihood fit to all regions.

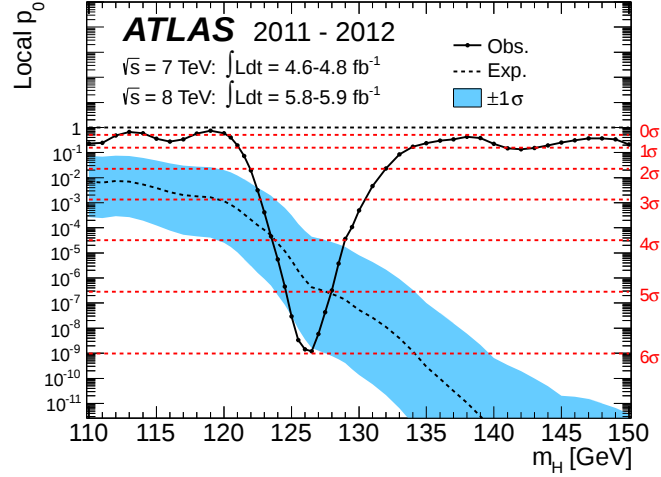


Figure 1.10: The observed (solid) local p_0 as a function of m_H in the low mass range [13]. The dashed curve shows the expected local p_0 under the hypothesis of a SM Higgs boson signal at that mass with its plus/minus one sigma band. The horizontal dashed lines indicate the p-values corresponding to significances of 1 to 6 sigma.

The $-2 \ln \Lambda$ value as a function of m_H for the individual $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$ channels and their combination is shown in figure 1.11. In order to assess the compatibility of the mass measurements from the two channels a dedicated test statistic has been used. From the value of $-2 \ln \Lambda$ at $\Delta m_H = 0$, a compatibility of 4.8%, equivalent to 1.98σ , is estimated.

Spin

In the SM, the Higgs boson is a spin-0 and CP-even particle ($J^P = 0^+$). The spin-parity of the observed Higgs boson has been evaluated independently in the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$ channels; the results have then been combined and the spin obtained. The analysis in each channel relies on discriminant observables chosen to be sensitive to the spin and the parity of the signal. Several spin-parity hypothesis have been tested: $J^P = 0^+, 0^-, 1^+, 1^-, 2^+$. A likelihood function $\mathcal{L}(J^P, \mu, \theta)$ is built for each spin-parity assumption and a test statistic q used to distinguish between two signal spin-parity hypothesis at a time is performed, based on a ratio of likelihoods:

$$q = \log \frac{\mathcal{L}(J^P = 0^+, \hat{\mu}_{0^+}, \hat{\theta}_{0^+})}{\mathcal{L}(J_{alt}^P, \hat{\mu}_{J_{alt}^P}, \hat{\theta}_{J_{alt}^P})} \quad (1.40)$$

where J_{alt}^P is the alternative hypothesis to be tested. Variables sensitive to the Higgs spin-parity are for example the dilepton invariant mass m_{ll} and the azimuthal separation of the two leptons, $\Delta\phi_{ll}$, in the $H \rightarrow WW^* \rightarrow l\nu l\nu$ event topologies. The data favour the SM quantum numbers of $J^P = 0^+$ [27], while the other hypotheses are rejected. Results are shown in figure 1.12.

Production and couplings

The Higgs boson production strength, the parameter μ , has been determined from a fit to data using the profile likelihood ratio $\Lambda(\mu)$ for a fixed mass hypothesis corresponding to the measured value. The overall signal production strength is measured to be:

$$\mu = 1.33 \pm 0.14(stat) \pm 0.15(sys) \quad (1.41)$$

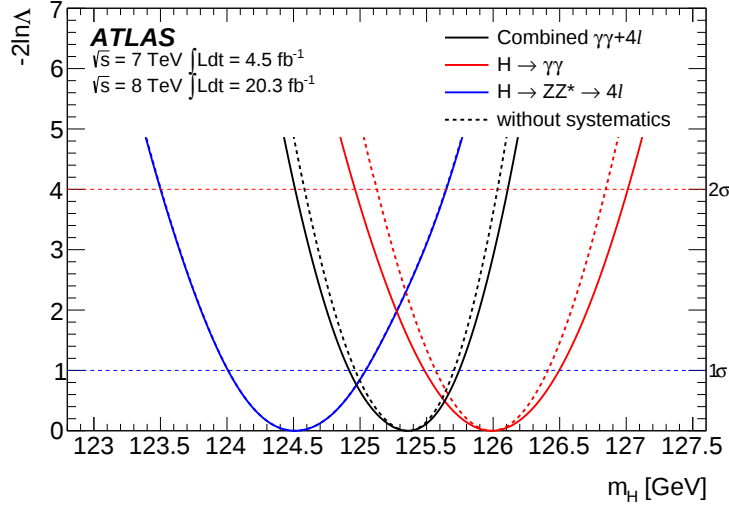


Figure 1.11: Value of $-2 \ln \Lambda$ as a function of m_H for the individual $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$ channels and their combination, where the signal strengths $\mu_{\gamma\gamma}$ and μ_{4l} are allowed to vary independently [23]. The dashed lines show the statistical component of the mass measurements. For the $H \rightarrow ZZ^* \rightarrow 4l$ channel, this is indistinguishable from the solid line that includes the systematic uncertainties.

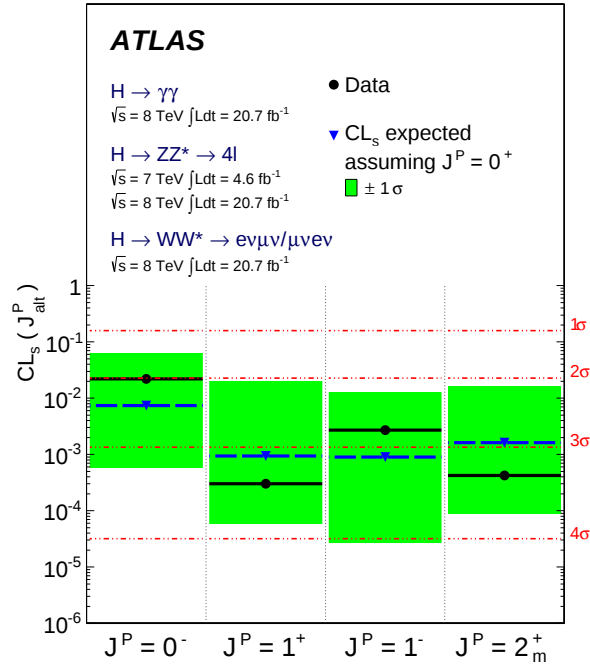


Figure 1.12: Expected (blue triangles/dashed lines) and observed (black circles/solid lines) confidence level CLs for alternative spin-parity hypotheses assuming a 0^+ signal [27]. The green band represents the 68% CLs (J^P_{alt}) expected exclusion range for a signal with assumed 0^+ . On the right y-axis, the corresponding numbers of Gaussian standard deviations are given, using the one-sided convention.

The consistency between this measurement and the SM expectation ($\mu = 1$) is about 7%. This measurement does not give any information about the relative contributions of the different production mechanisms. For this reason the data are fitted separately for vector-boson-mediated processes, VBF and VH, and for gluon-mediated processes, ggF and ttH; two signal strength parameters are used, $\mu_{ggF+ttH} = \mu_{ggF} = \mu_{ttH}$ and $\mu_{VBF+VH} = \mu_{VBF} = \mu_{VH}$. The results are shown in figure 1.13. To test the sensitivity to VBF production alone, the data are also fitted

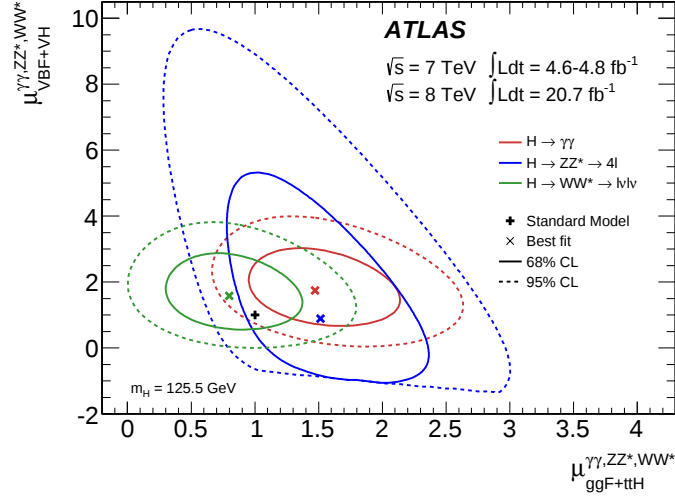


Figure 1.13: Likelihood contours for the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$ channels in the $(\mu_{ggF+ttH} \times B/BSM, \mu_{VBF+VH} \times B/BSM)$ plane for a Higgs boson mass $m_H = 125.5$ GeV [26]. The branching-ratio scale factors B/BSM can a priori be different for the different final states. The sharp lower edge of the $H \rightarrow ZZ^* \rightarrow 4l$ contours is due to the small number of events in this channel and the requirement of a positive pdf. The best fits to the data ($\times$) and the 68% (full) and 95% (dashed) CL contours are indicated, as well as the SM expectation (+).

with the ratio $\mu_{VBF}/\mu_{ggF+ttH}$. A value

$$\mu_{VBF}/\mu_{ggF+ttH} = 1.4^{+0.4}_{-0.3}(stat)^{+0.6}_{-0.4}(sys) \quad (1.42)$$

is obtained combining the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$ channels. This result provides evidence at the 3.3σ level that a fraction of Higgs boson production occurs through VBF. At the moment no evidence has been found for production in the VH and ttH channels.

The coupling of the Higgs boson to fermions and bosons is parametrized in term of scale factors k_i [26]. If the cross section for a given process is

$$\sigma \times BR(ii \rightarrow H \rightarrow ff) = \frac{\sigma_{ii} \cdot \Gamma_{ff}}{\Gamma_H} \quad (1.43)$$

possible deviations from SM predictions can be parametrized through scaling factors for couplings such that

$$\Gamma_{ff} = k_f^2 \Gamma_{ff}^{SM}; \quad \Gamma_H = k_H^2 \Gamma_H^{SM}; \quad \sigma_i = k_i^2 \sigma_i^{SM} \quad (1.44)$$

These scale factors are extracted from fits to the data using the profile likelihood ratio $\Lambda(k)$, where the k_i couplings are treated either as parameters of interest or as nuisance parameters, depending on the measurement. Figure 1.14 shows the results of the fit for the three channels ($\gamma\gamma$, ZZ^* and WW^*) and their combination.

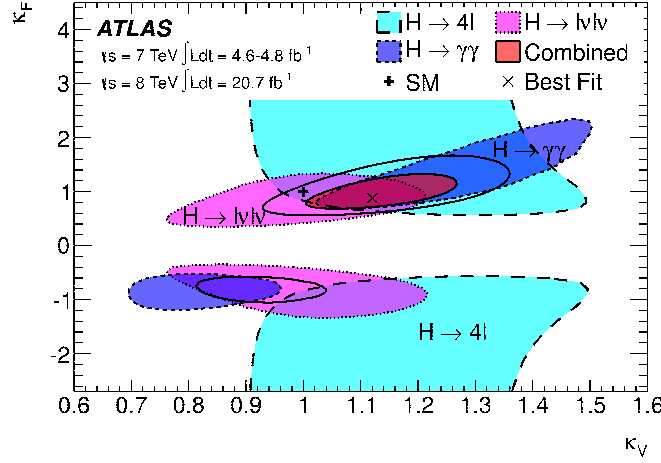


Figure 1.14: Likelihood contours (68% CL) of the coupling scale factors k_F and k_v for fermions and bosons, as obtained from fits to the three individual channels and their combination (for the latter, the 95% CL contour is also shown) [26]. The best-fit result ($\times$) and the SM expectation (+) are also indicated.

1.4 The VH process at hadron colliders

The associated production of the Higgs boson with a vector boson (V) at the LHC suffers from a lower cross section compared to the dominant ggF and VBF processes. The value of the cross section can be found in table 1.1 separately for WH and ZH processes. These refer to the NNLO calculation. In fact at hadron colliders the lowest-order (LO) cross sections are affected by large uncertainties arising from higher-order QCD corrections. The cross section variation with the renormalization and factorization scales can be stabilized by evaluating the NLO and the NNLO radiative corrections.

The associated production cross section is one of the simplest production mechanism at hadron colliders; the final state does not feel strong interactions, and can be viewed simply as the production of a virtual W or Z boson, which then splits in a real boson and a Higgs boson (see figure 1.3 (c)). Strong interactions affect only the quark-antiquark initial state, so the total $pp \rightarrow VH$ cross section can be factorized as the product of the partons distribution functions and the cross section for the subprocess $q\bar{q}' \rightarrow VH$:

$$\sigma_{LO}(pp \rightarrow VH) = \int_{\tau_0}^1 d\tau \sum_{ij} \frac{d\mathcal{L}^{ij}}{d\tau} \hat{\sigma}_{LO}(\hat{s} = \tau s) \quad (1.45)$$

where $\tau_0 = (M_V)^2/s$, s being the total centre-of-mass energy of the proton-proton collision, and the luminosity is defined in terms of the parton densities². At NLO the QCD corrections to the process $pp \rightarrow V^* + X$ consist of virtual corrections due to gluon exchange in the $q\bar{q}$ vertex, quark self-energy corrections and the emission of a gluon in the initial state, as shown in the Feynman diagrams in figure 1.15. Since these QCD corrections to the sub-process $pp \rightarrow V^* + X$, for a given mass of the virtual boson, are the same as the QCD corrections to the Drell-Yan

²The differential parton-parton luminosity is defined as [16]

$$\frac{\tau}{\hat{s}} \frac{d\mathcal{L}^{ij}}{d\tau} \equiv \frac{\tau/\hat{s}}{1 + \delta_{ij}} \int_{\tau}^1 dx \left[f_i^{(a)}(x) f_j^{(b)}(\tau/x) + f_j^{(a)}(x) f_i^{(b)}(\tau/x) \right] / x \quad (1.46)$$

It is a useful quantity to introduce, since when multiplied by the dimensionless cross section $\hat{\sigma}$ for a process, provide an estimate of the size of an event cross section at the LHC.

process, the latter have been used as reference in [29]. NLO electroweak (EW) corrections

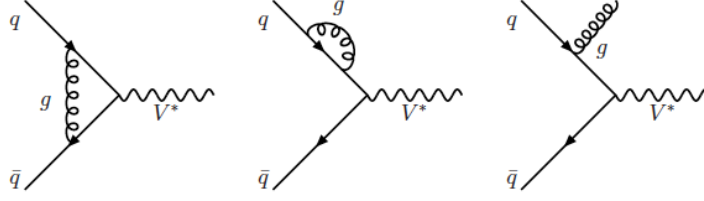


Figure 1.15: NLO corrections to the vector boson quark-antiquark vertex.

have been evaluated in [30]. In contrast to the NLO QCD corrections, EW corrections do not respect a factorization into Drell–Yan-like production and decay, since there are irreducible (box) corrections to $q\bar{q}' \rightarrow VH$ already at one loop.

The NNLO corrections are the $O(\alpha_s^2)$ contributions to the annihilation process. They consist in the two-loop vertex corrections and the emission of two gluons in the initial state of the $q\bar{q}$ interaction. A detailed calculation can be found here [31] and here [16] for a more recent result. Since the cross section is evaluated at a fixed order in perturbation theory, the cross section depends on the unphysical renormalization and factorization scales μ_R and μ_F , which nominal value is chosen as the mass of the intermediate gauge boson. The variation of these scale together with that of the parton density functions define the uncertainty of the cross section. The cross section results for WH production at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 14$ TeV is shown in figure 1.16.

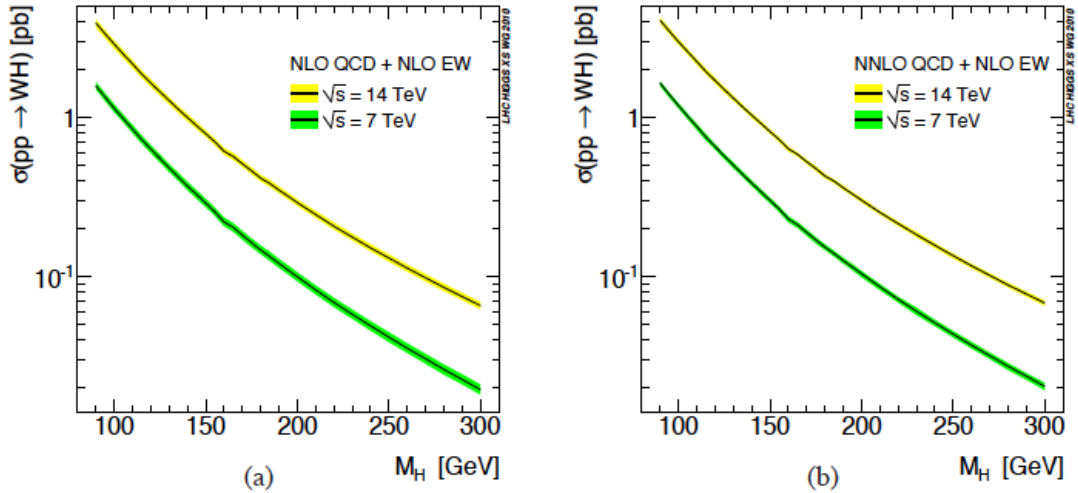


Figure 1.16: Cross section for the WH production for 7 TeV and 14 TeV at (a) NLO and (b) NNLO QCD, including NLO EW effects in both cases [16].

The impact of higher order (HO) corrections to the tree-level diagram, is quantified in term of the so-called K-factor, defined as:

$$K_{HO} = \frac{\sigma_{HO}}{\sigma_{LO}} \quad (1.47)$$

The K-factor at NLO and NNLO for the Higgs boson associated production at the LHC is shown in figure 1.17 as a function of the Higgs boson mass. The value found is around 1.2.

As will be shown in chapter 4 with the luminosity collected by ATLAS during Run 1, the Higgs boson associated production with a vector boson will not be observed. However, a detailed study of the process in a Higgs boson decay channel with high BR will result useful in preparing the analysis for Run 2.

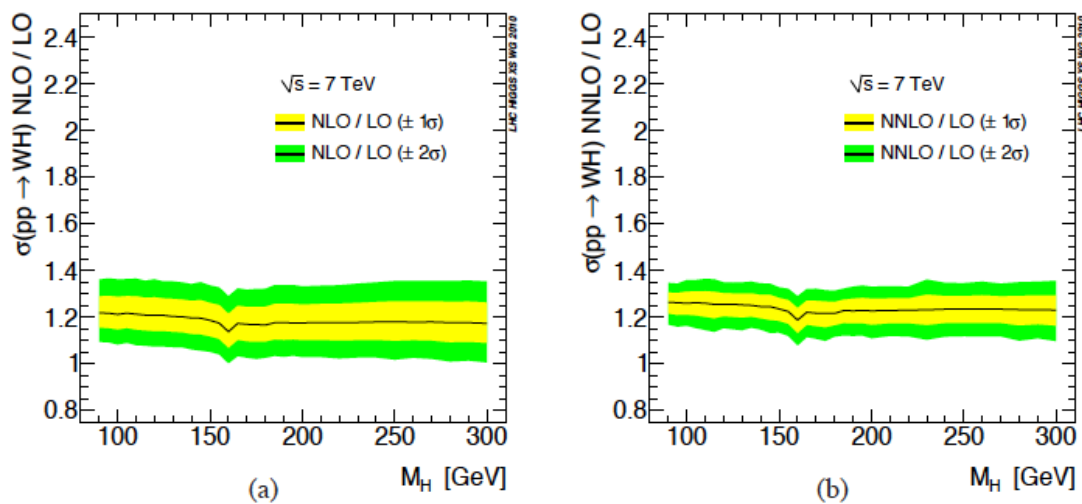


Figure 1.17: The K-factors (ratio to LO prediction) for $pp \rightarrow WH$ at the LHC 7 TeV as a function of m_H for the NLO and NNLO cross sections [16]. The little kinks at around 160 GeV and, somewhat smaller, 180 GeV, are due to the WW and ZZ thresholds that occur in the EW radiative corrections.

Chapter 2

The ATLAS detector

In this chapter the experimental apparatus used to measure the VH process explained in the previous chapter is described. ATLAS is a general-purpose experiment recording the collisions provided by the Large Hadron Collider (LHC), the proton-proton collider which is operational at CERN since 2009. The ATLAS detector is going to be explained in all its sub-components in the next sections.

2.1 The Large Hadron Collider

The Large Hadron Collider [28] is the 27 km long particle collider hosted in the same tunnel where the LEP collider was, about 100 meters underneath the Swiss-French national border near Geneva. It is the most energetic particle collider ever built. In the LHC ring two proton bunches at a time collide, each bunch having approximately 10^{11} protons, with a centre-of-mass energy equals to $\sqrt{s}= 8$ TeV during 2012 data taking period (2012 was part of the so-called *Run 1*). The proton-proton (pp) collider is currently in a shutdown phase, to allow for the replacement and upgrade of some machine and detectors components; it will start again with protons collisions in spring 2015, with a centre-of-mass energy of 13 TeV.

Before particles are injected into the LHC they go through several acceleration stages, shown in figure 2.1. After their production, the proton beams are accelerated up to 50 MeV by the LINAC2 machine. The protons are then injected into the Proton Synchrotron Booster (PSB) which accelerate the particles to 1.4 GeV. After that, the particles are injected first into the Proton Synchrotron (PS) where they are accelerated to 26 GeV, and then into the Super Proton Synchrotron (SPS) which increases their energy to 450 GeV. The particles enter the LHC into two parallel rings and after ramping up to the desired energy the beams are squeezed and directed to collisions in the dedicated LHC experiments. Beams focusing and acceleration is obtained with magnets and radiofrequency (RF) cavities: LHC is equipped with 1232 superconducting magnets and 6 RF cavities which bend and accelerate the proton beams in the two parallel beam lines in the machine. The magnetic field used to bend such energetic protons is of 8.3 T. For reaching and keeping the superconductivity range of the cold masses, the cooling system provides the magnets with fluid helium at a temperature of 1.9 K.

In studying high energy particle collisions one of the crucial parameter of the collider is the *instantaneous luminosity*, since it is proportional to the production rate:

$$\frac{dN}{dt} = \mathcal{L} \times \sigma \quad (2.1)$$

where σ is the cross section of the considered process. The instantaneous luminosity depends

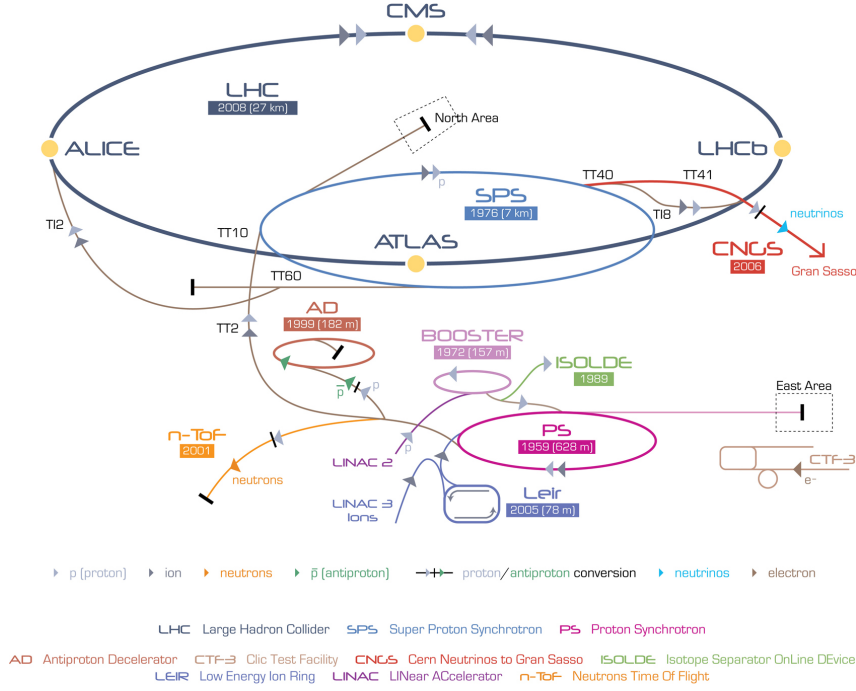


Figure 2.1: The CERN particles accelerator complex. The injection chain together with the four interaction points are visible.

on the intrinsic properties of the machine:

$$\mathcal{L} = \frac{f N_p^2 k}{4\pi\sigma_x\sigma_y} \quad (2.2)$$

where N_p is the number of particles in a bunch, f is the revolution frequency of the proton in the accelerating ring, k is the number of bunches circulating in the beam, σ_x and σ_y are the gaussian beam profiles in the transverse plane with respect to the beam direction. A good knowledge of the expected beam properties give access to a good expectation value for the luminosity. The luminosity can then be used in two ways from equation (2.1): either predict the number of expected background or signal events from a prior cross section value, or measure a cross section from a number of observed events in data. To reduce the uncertainty on the luminosity and hence on the measured cross sections, its value is monitored regularly during the data-taking period, for example with the Van Der Meer scans method [32]. Figure 2.2 shows the peak luminosity delivered to ATLAS in the various data-taking period of Run1, while the 2012 integrated luminosity used for the scope of this thesis, corresponds to 20.3 fb^{-1} (see picture 4.4). Although the high intensity of the beams allows to probe for rare process, it also gives rise to some disadvantages, as the possibility to have more than one interaction per bunch crossing. This phenomenon is called *pile-up*. We distinguish between *in time pile-up*, when many interactions arise from the same bunches collision, and *out-of-time pile-up*, when detector signals occurred in a bunch crossing before the event of interest but are recorded later because of the latency time of some detectors. In both cases the pile-up produces a high particle multiplicity in the detector which makes harder the reconstruction of the event of interest. The average number of pile-up interactions in 2012 collisions was $\langle \mu \rangle = 20.7$.

Table 2.1 shows the LHC parameters in 2012. The LHC provides collisions in four collision points along its circumference where detector experiments are hosted: ALICE (*A Large Ion Collider Experiment*) [33], ATLAS (*A Toroidal Lhc ApparatuS*) [34], CMS (*Compact Muon Solenoid*) [35][36] and LHCb (*Large Hadron Collider beauty*) [37]. ATLAS and CMS are multi-

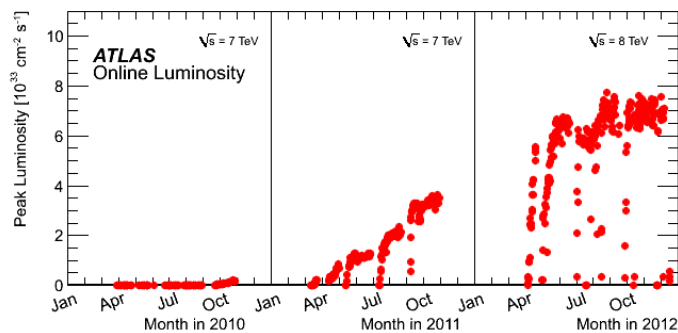


Figure 2.2: The peak instantaneous luminosity delivered to ATLAS per day versus time during the pp runs of 2010, 2011 and 2012.

Parameter	Value
Beam energy [TeV]	4
Bunch spacing [ns]	50
Number of bunches	2748
Protons per bunch	$\approx 10^{11}$
Mean number of interactions per bunch crossing	20.7
Mean luminosity [$cm^{-2}s^{-1}$]	7.7×10^{33}
Bunch transverse dimension [μm]	≈ 30

Table 2.1: LHC operational parameters during 2012 data-taking period.

purpose detectors, while ALICE and LHCb focused respectively on the quark-gluon plasma studies and on the study of CP violation processes occurring in b and c hadron decays. The ATLAS detector is described in the next section.

2.2 The ATLAS experiment

ATLAS is a general-purpose experiment designed to maximize the physics discovery potential offered by the LHC accelerator. It is called general-purpose since it is capable of detecting and measuring every particle in collisions allowing to reconstruct the whole final state. The detector development and operation as well as the data analysis is performed within the ATLAS Collaboration, an international organization of ~ 3000 people from 173 institutes in 37 countries.

The ATLAS detector is 20 meters tall and 45 meters long and weights ~ 7000 tons. It has a cylindrical shape centered at the interaction point with its axis along the beam line, and it is composed of several concentric subdetectors which measure different features of the particles generated in the pp collisions as they move from the center of the detector to the outer part. From the innermost to the outermost layer, the ATLAS experiment is composed of (see figure 2.3):

- an inner tracking system, designed to detect charged particles and measure their momentum and direction;
- a superconducting solenoid providing a uniform magnetic field along the beam axis;

- a calorimeter system, with an electromagnetic calorimeter to measure the energy deposited by electrons and photons, followed by hadronic calorimeter;
- a muon spectrometer, to reconstruct the muon tracks and to measure their momentum, in a system of an air-cored toroidal magnets.

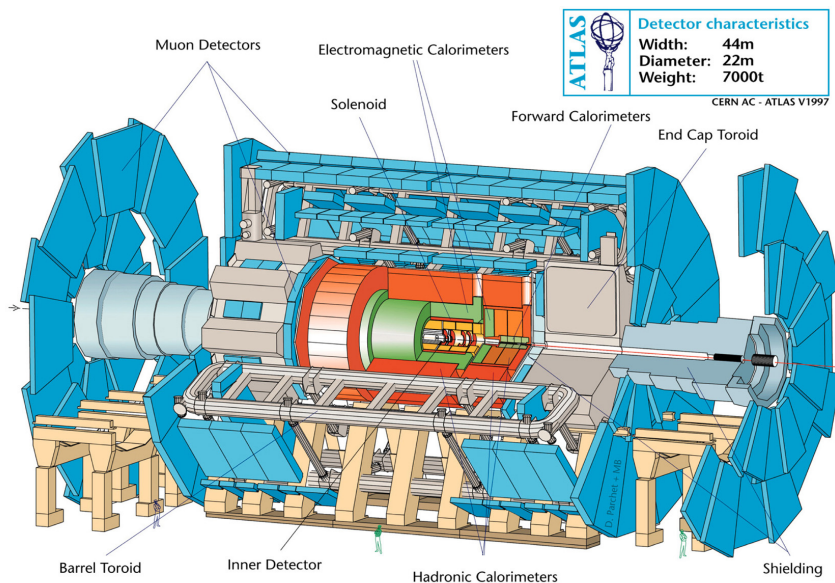


Figure 2.3: Schematic design of the ATLAS detector.

2.2.1 ATLAS coordinate system

The ATLAS reference system is shown in figure 2.4: the origin of the system is in the nominal interaction point, the z axis is along the beam line, while the $x-y$ plane is the plane perpendicular with respect to beam line. The positive x -axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y -axis is defined as pointing upwards. The azimuthal angle ϕ is defined around the beam axis, while the polar angle θ is the angle from the z axis in the $y-z$ plane. The θ variable is not invariant under boosts along the z axis, and so the *rapidity* y is used:

$$y = \frac{1}{2} \ln \frac{E + p \cos \theta}{E - p \cos \theta} \quad (2.3)$$

where E and p are respectively the energy and the momentum of the particle. In the ultra-relativistic limit the pseudorapidity η is a very good approximation of y :

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (2.4)$$

With the relation in equation (2.4) to smaller values of η correspond higher θ angles and vice versa.

2.2.2 The Inner detector

The ATLAS Inner detector (ID) has a fully coverage in ϕ and covers the pseudorapidity range $|\eta| < 2.5$. It consists of a silicon Pixel detector (Pixel), silicon strip detector (SCT) and for $|\eta| < 2$ a Transition Radiation Tracker (TRT). This set of detectors covers the radial distance

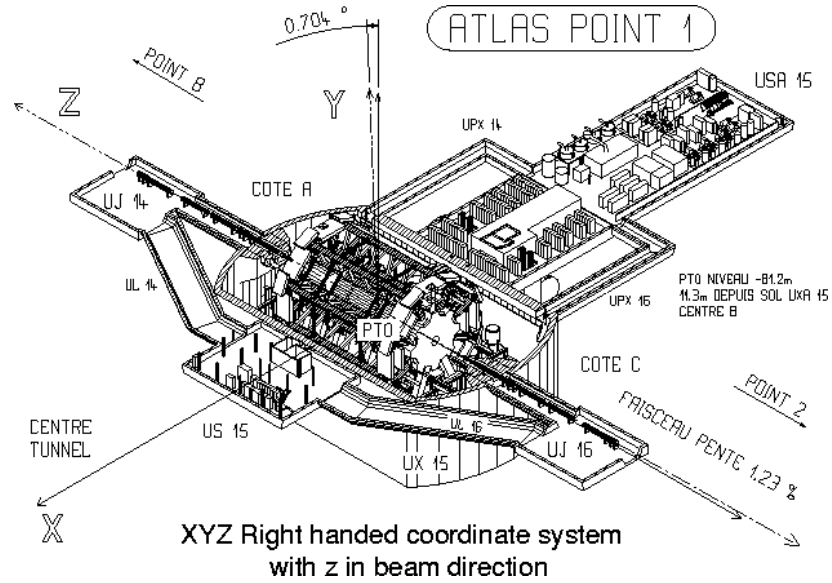


Figure 2.4: ATLAS experiment coordinate system.

of 50.5 mm from the interaction point up to 1066 mm. A detailed conceptual layout of the ID can be seen in figure 2.5 and a brief description is given below:

- **Pixel Detector:** is the innermost tracking device and that with the highest granularity. It is composed of three layers of silicon pixels. It provides high-precision track measurement, since the spatial resolution on the single hit is $\approx 10 \mu\text{m}$ in the ϕ coordinate and $\approx 115 \mu\text{m}$ along the z coordinate.
- **Semiconductor Tracker (SCT):** is designed to provide additional precision space point measurements per track. It is composed of eight layers of silicon strips with a spatial resolution on the single hit of $17 \mu\text{m}$ in ϕ and $580 \mu\text{m}$ along z . The Pixel Detector and the Semiconductor Tracker together provide on average eight high-precision hits per track.
- **Transition Radiation Tracker (TRT):** is a combined straw tube tracker and transition radiation detector. The resolution of such a detector is lower than the previous one ($\approx$

130 μm per straw), but it is compensated by the high number of points per track (36 on average) that it can provide.

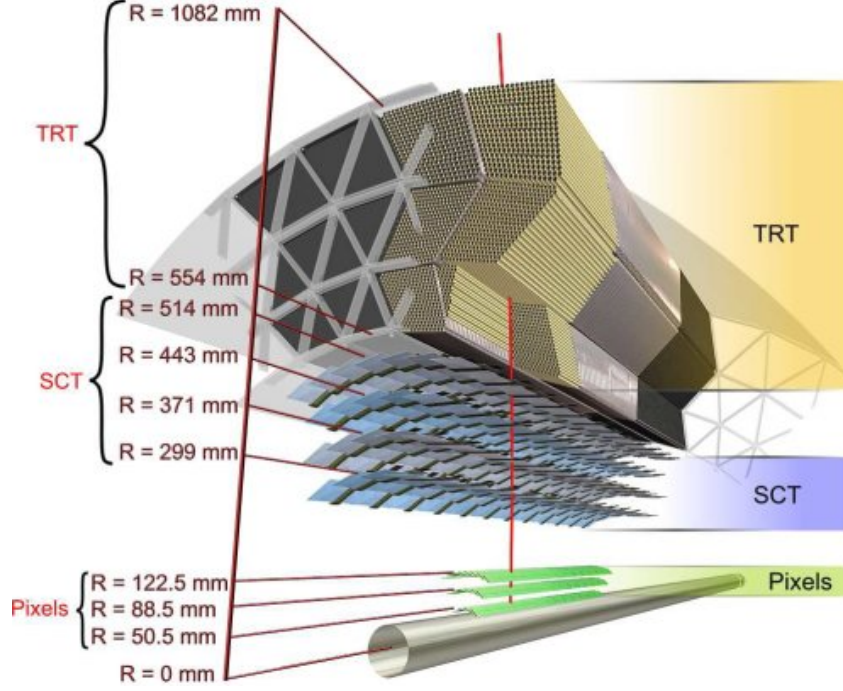


Figure 2.5: ATLAS inner detector. Each sub-component is shown, together with its radial dimensions.

A superconducting solenoid provides a uniform magnetic field of approximately 2 T to the inner tracking system. The solenoid, shown in figure 2.6, covers the central region of the detector; it is located between the inner detector and the electromagnetic calorimeter and its dimensions have been optimized in order to minimize the amount of dead material (only 0.83 radiation lengths) in front of the calorimeter system. The resulting magnetic field is along the z axis, bending particles tracks in the transverse plane. The aim of the ATLAS ID is indeed to measure the tracks of the charged particles produced in the pp collision and all their related features: p_T , η , ϕ , and eventually the secondary vertices due to long-lived particles. To estimate the resolution on the momentum the *sagitta* method can be used: the magnetic field bends the trajectory of the charged particles because of the Lorentz's force:

$$\vec{F}_L = q\vec{v} \times \vec{B} \quad (2.5)$$

where q is the charge of the particle, $\vec{v}$ is its velocity and $\vec{B}$ is the magnetic field. The resolution of the momentum measurement depends on many detector-related parameters:

$$\frac{\delta p}{p^2} = \frac{8}{0.3 \cdot B \cdot L^2} \Delta s \quad (2.6)$$

where B is the magnetic field expressed in Tesla, L is the length of the reconstructed track expressed in meters, while Δs is:

$$\Delta s = \frac{\epsilon}{8} \sqrt{\frac{720}{N+4}} \quad (2.7)$$

where N is the number of measured points on the track and ϵ is the resolution on the point measurement. s is called *sagitta*, and it is the divergence of the track due to the magnetic field

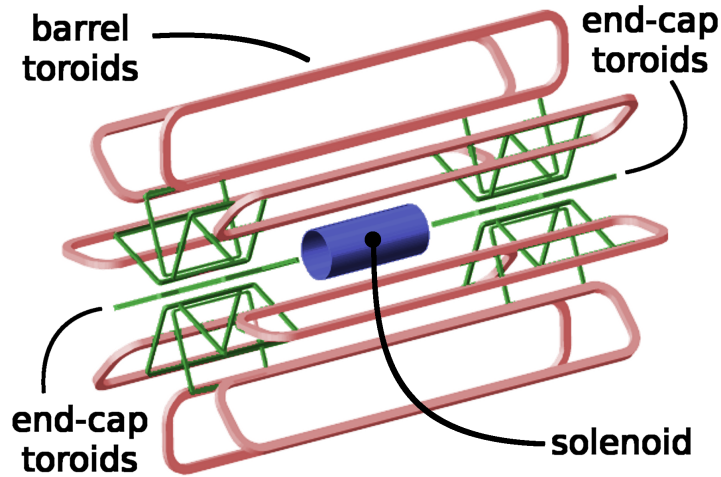


Figure 2.6: The magnetic system of the ATLAS detector. The superconducting solenoid (the inner cylinder) providing the magnetic field to the ID is shown, together with the external coils of the toroid that provide the magnetic field to the muon system.

and multiple scattering in the detector material. From equations (2.6) and (2.7) it is possible to note how is crucial to have a strong magnetic field, an high number of points per track and a good spatial resolution on these points to have a good resolution in the track p_T .

The performance of the ID are shown in the next chapter, where also the tracking algorithms are explained.

2.2.3 The Calorimeter system

The calorimeter system measures the energy of photons, electrons, hadrons and jets, together with the missing transverse momentum (due to undetected particles like neutrinos). It is hermetic in ϕ , and the coverage in pseudo-rapidity is up to $|\eta| < 4.9$. The system is composed by an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL) (figure 2.7).

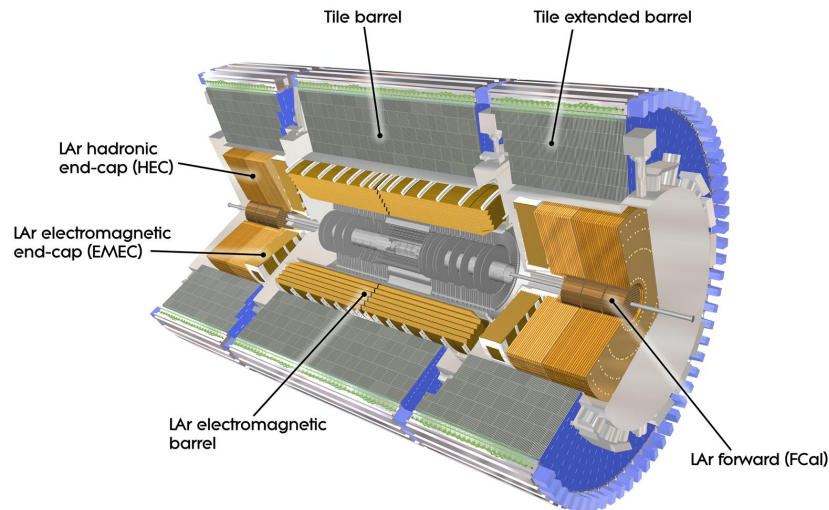


Figure 2.7: ATLAS calorimeter system.

The Electromagnetic Calorimeter

The electromagnetic calorimeter is made up of alternated layers of lead (Pb), which is the absorber, and Liquid Argon (LAr), the active material. It covers the region up to $|\eta| < 3.2$. For the absorption material and for the electrodes it has been chosen an accordion structure (see figure 2.8), able to cover all the azimuth angle; this structure allows for a fast signal extraction. The calorimeter is segmented in cells of variable dimensions as a function of η as well as its

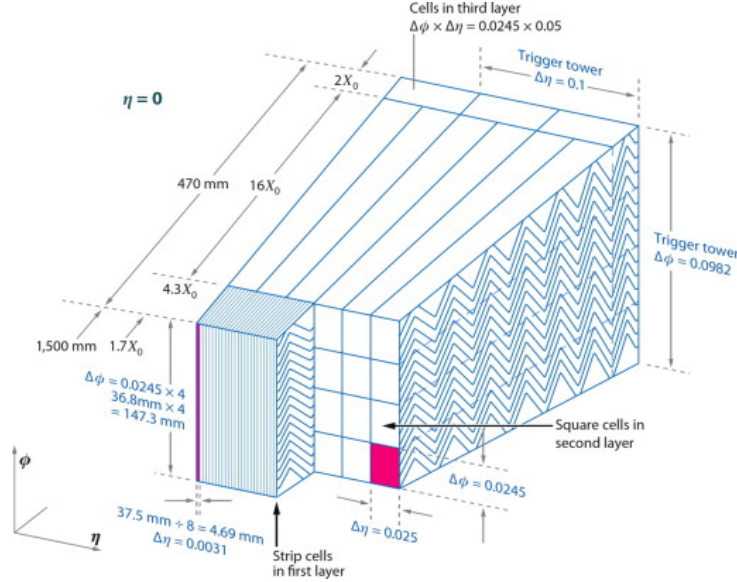


Figure 2.8: The accordion structure of the electromagnetic calorimeter and its radial segmentation.

thickness (> 24 radiation length in the central region and > 26 radiation length in the forward region). In the central region the segmentation is $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$.

The energy resolution for electrons/photons is parametrized as:

$$\frac{\Delta E}{E} = \frac{11.5\%}{\sqrt{E[\text{GeV}]}} \oplus 0.5\% \quad (2.8)$$

where 11.5% is the sampling term and 0.5% is the constant (intercalibration) term.

The Hadronic Calorimeter

The Hadronic Calorimeter covers the region $|\eta| < 4.5$ and it is made by several different technologies depending on η . The Tile Calorimeter (TileCal) covers $|\eta| < 1.7$; it is made of alternating layers of iron (used as absorbers) and scintillating tiles as active materials, and its thickness offers about 10 interaction lengths at $\eta = 0$. It is segmented in $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ pseudo-projective towers pointing to the interaction point.

The HEC (End-Cap) region, $1.7 < |\eta| < 3.1$, is placed after the end-cap of the electromagnetic calorimeter and it is equipped with liquid Argon and lead. The forward region, $3.1 < |\eta| < 4.51$, is equipped again with liquid Argon, but the accordion structure is replaced by concentric rods and tubes made of copper. This variety of materials and structures is due to the different radiation hardness required in the different parts of the detector.

2.2.4 The Muon Spectrometer

The Muon Spectrometer (MS) is the outermost component of the ATLAS detector. It is designed to detect minimum ionizing particles (muons) exiting the calorimeter system and to measure their momenta in the pseudorapidity range $|\eta| < 2.7$. The MS is instrumented with both trigger and high-precision chambers immersed in the magnetic field provided by air-core toroidal magnets (figure 2.6) which bends the particles along the η coordinate (being $\int B \cdot dl$ between 2 and 6 T · m). A sketch of the MS is displayed in figure 2.9. The MS chambers devoted to

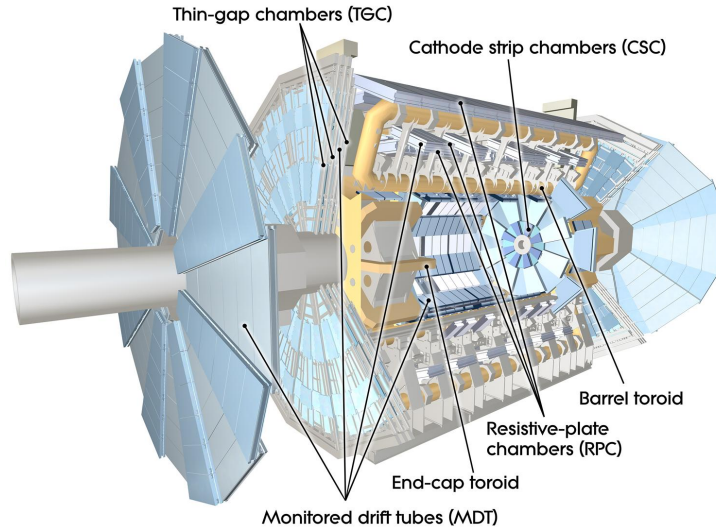


Figure 2.9: The ATLAS muon spectrometer.

the precision tracking are the *Monitored Drift Tubes* (MDT) and the *Cathode Strip Chambers* (CSC), while for the trigger measurement the *Resistive Plate Chambers* (RPC) and the *Thin Gap Chambers* (TGC) are respectively in the barrel and in the endcap:

- **Monitored Drift Tubes:** they are used in the central region ($|\eta| < 2$) of the detector. The MDT chambers are composed of aluminium tubes of 30 mm diameter and 400 μm thickness, with a 50 μm diameter central wire. The tubes are filled with a mixture of Argon and CO_2 at high pressure (3 bars), and each tube has a spatial resolution of 80 μm .
- **Cathode Strip Chambers:** they are used at higher pseudo-rapidity ($2 < |\eta| < 2.7$) with respect to the MDT. CSC chambers are multiwire proportional chambers in which the readout is performed using strips forming a grid on the cathode plane in both orthogonal and parallel direction with respect to the wire. The spatial resolution of the CSC is about 60 μm .
- **Resistive Plate Chambers:** the RPC produce the trigger signal in the barrel. They are also capable to measure the transverse coordinate and are therefore complementary with the MDT. 544 chambers are located in three concentric layers connected to the MDT. Every chamber has 2-layers of gas gap filled with a gas mixture of 94.74% $\text{C}_2\text{H}_2\text{F}_4$ + 5% $\text{isoC}_4\text{H}_{10}$ + 0.3% SF_6 , where the last one is added to limit the charge avalanches in the chamber. The chambers are made with bakelite plates of 2 mm and readout strips with pitches of about 3 cm. The RPC work at 9.8 kV and have a time resolution of 1.5 ns.
- **Thin Gap Chambers:** the TGC are multiwire proportional chambers dedicated to the trigger system on the endcap part of the ATLAS detector. The TGC, like the RPC,

provide also a measurement of the muon track coordinate orthogonal to the one provided by the precision tracking chambers. The nominal spatial resolution for the TGC is 3.7 mm in the $R - \phi$ plane. The gas mixture used for these chambers is 55% $CO_2 + 45\%$ nC5H12 and they work at 2.9 kV. The time resolution is about 4 ns.

As shown in figure 2.9, in the central region the MS is arranged on a three layer (or stations) cylindrical structure which radii are 5, 7.5 and 10 meters; while in the forward region the detectors are arranged vertically, forming four disks at 7, 10, 14 and 21 – 23 meters from the interaction point (see figure 2.10). The resolution of the MS is dominated by the energy

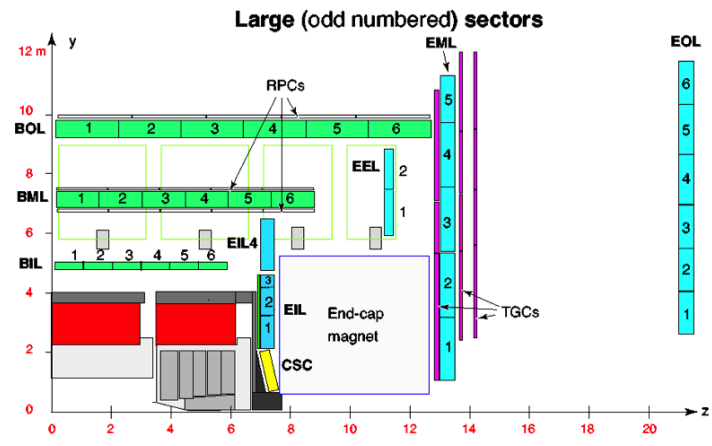


Figure 2.10: The cross-section of the muon system in a plane containing the beam axis (bending plane).

loss of the muons in the calorimeter for low momenta, $p_T < 30$ GeV, by multiple scattering effects for $30 \text{ GeV} < p_T < 100$ GeV while, above 100 GeV, calibration and alignment of the spectrometer itself become the most significant factors in momentum resolution. The relative resolution $\sigma p_T / p_T$ that can be obtained is better than 3% over a wide p_T range.

2.2.5 ATLAS Trigger

The proton-proton interaction rate, at the LHC designed luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and with 25 ns of bunch spacing, is 40 MHz, while the event data recording limit is about 200 Hz. This means that a sophisticated trigger system that selects only the relevant events for the interesting physics processes at LHC, processes that occur at a lower rate as shown in figure 2.11, is needed. The ATLAS trigger system is organized in three levels as shown in figure 2.12: the Level-1, the Level-2 and the Level-3, also called the *Event Filter*. Each step refines the decision taken at the previous step, using new algorithms and combining the information from different subdetectors.

- **Level 1 (LVL1):** it is a hardware-based trigger system, which uses low granularity informations from a given subset of detectors. The LVL1 trigger looks for high- p_T muons candidates or calorimetric objects (electrons/photons or jets) using the fast and rough informations from the muons trigger detectors (RPC and TGC) and from the calorimeter system (for the clusters). The LVL1 is designed to reduce the rate below 75 kHz, which is the maximum acceptance that the detector readout can handle. The LVL1 decision must reach the front-end electronics within $2.5 \mu\text{s}$ after the bunch-crossing and the decision is given in terms of *Regions of Interest* (RoI), $\eta - \phi$ regions of the detector in which interesting activity has been detected. The list of RoI is the input of the Level-2.

Event rate and decision stages

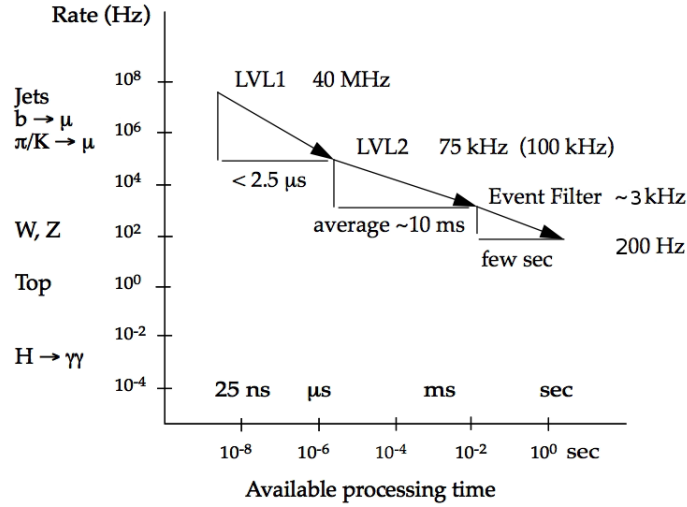


Figure 2.11: The event rate at which interesting physics occur (referred to LHC design parameters) and the processing time of each trigger level.

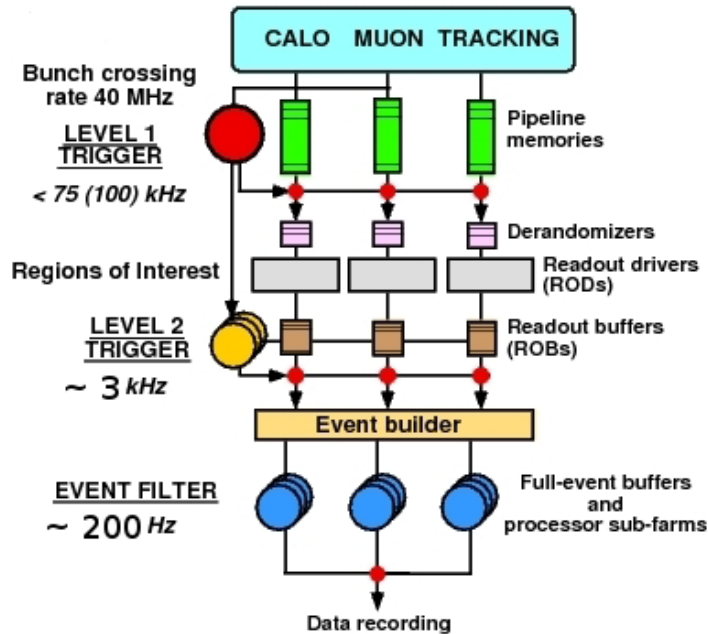


Figure 2.12: The ATLAS trigger system.

- **Level 2 (LVL2):** it is a software-based trigger which refines the decision taken from LVL1 to further reduce the trigger rate to about 3 kHz. To this purpose the data of the precision chambers are used in the Muon Spectrometer (MDT, CSC) as well as the data from the ID, while the measurement of the calorimetric objects is refined using higher level algorithms. Moreover the data of the different subdetectors are combined together in order to obtain better object reconstruction/identification. The decision is taken in ≈ 10 ms.
- **Level 3 (LVL3):** the level 3 or Event Filter (EF) is again a software-based level and forms, together with the LVL2, the so called *High Level Trigger* (HLT). At this stage

the reconstruction algorithms, also used during the offline event reconstruction, are used, and a full reconstruction is performed. The output rate of the LVL3 is of the order of ≈ 100 Hz. All the events that have been selected by the LVL3 trigger are then written to mass storage (disks or tapes), and are then used in the analysis.

Chapter 3

Physics objects definition and reconstruction

The analysis described in this thesis relies on the off-line reconstruction of electrons, muons, taus and jets, together with the determination of the missing transverse momentum. This chapter describes the main algorithms used for objects reconstruction in ATLAS during Run1 data taking period, while analysis specific object selection criteria are described in detail in chapter 4.

3.1 Track reconstruction

Hits recorded in the individual ID layers are used to reconstruct the trajectories of charged particles inside the tracker and to estimate their kinematic parameters. Particles trajectories are parametrized with a five parameter vector [38]:

$$\tau = (d_0, z_0, \phi_0, \theta, q/p) \quad (3.1)$$

where d_0 is the transverse impact parameter defined as the distance of closest approach in the transverse plane of the track to the primary vertex; z_0 is the longitudinal impact parameter, ϕ_0 is the azimuthal angle of the track and θ is its polar angle (figure 3.1). q/p is the inverse of the particle momentum multiplied by its charge. Track reconstruction in ATLAS is a two-stages

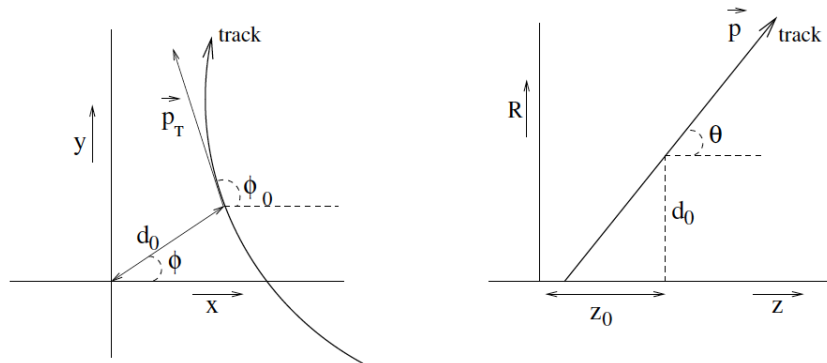


Figure 3.1: Illustration of the track parameters in the transverse plane (left) and RZ-plane (right), as defined in the ATLAS tracking frame.

procedure:

Track finder: Assignment of the hits left in the detector by charged particles traversing active detector elements to the track candidate.

Track fitter: The hits are used to reconstruct the trajectories by performing a fit to the track kinematic parameters. The track fitting is based on the minimization of the track-hit residuals.

In the above procedure a sequence of algorithms is used [39]. The default, named *inside-out* algorithm, starts with combining hits from the three pixel detector layers and the first SCT layer, in the so-called *track seed*. This seed is then extended to all the SCT candidates to form a track candidate that is then extended into the TRT and refitted using the full ID information. If some TRT hit worsening the fit quality is found, this is not included in the final fit and is labelled as *outlier* and kept for off-line studies. A track in the barrel region of the ID has typically 3 Pixel hits, 8 SCT hits and approximately 30 TRT hits.

The accuracy of the track reconstruction can be limited by a combination of: the resolution of the detector elements, the detailed map of the magnetic field, the misalignment of the detector elements and the multiple scattering and energy losses due to the material in the detector. The track reconstruction efficiency, defined as the fraction of primary particles (particles with a mean lifetime not greater than 3×10^{-11} s) matched to a reconstructed track, as a function of track p_T is shown in figure 3.2. A deterioration of the track reconstruction performance can also come

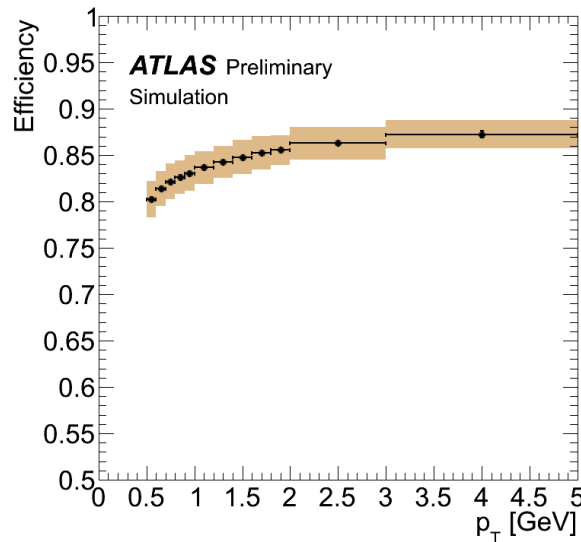


Figure 3.2: Tracking reconstruction efficiency vs transverse momentum. Tracking reconstruction efficiency is defined as the ratio between “number of matched tracks” and “number of generated charged particles”. The colour band shows the effects of selection cuts and detector material on the efficiency.

from the increasing detector occupancy with pile-up, which results in nearby hits from other particles confusing the pattern recognition algorithm.

Track reconstruction performances have been evaluated first using cosmic rays and then directly with data from pp collisions [40]. Figure 3.3 shows the distribution of pixel and SCT number of hits per track, together with data-MC comparison as a function of the pseudorapidity and the transverse impact parameter d_0 . Good data-MC agreement is found.

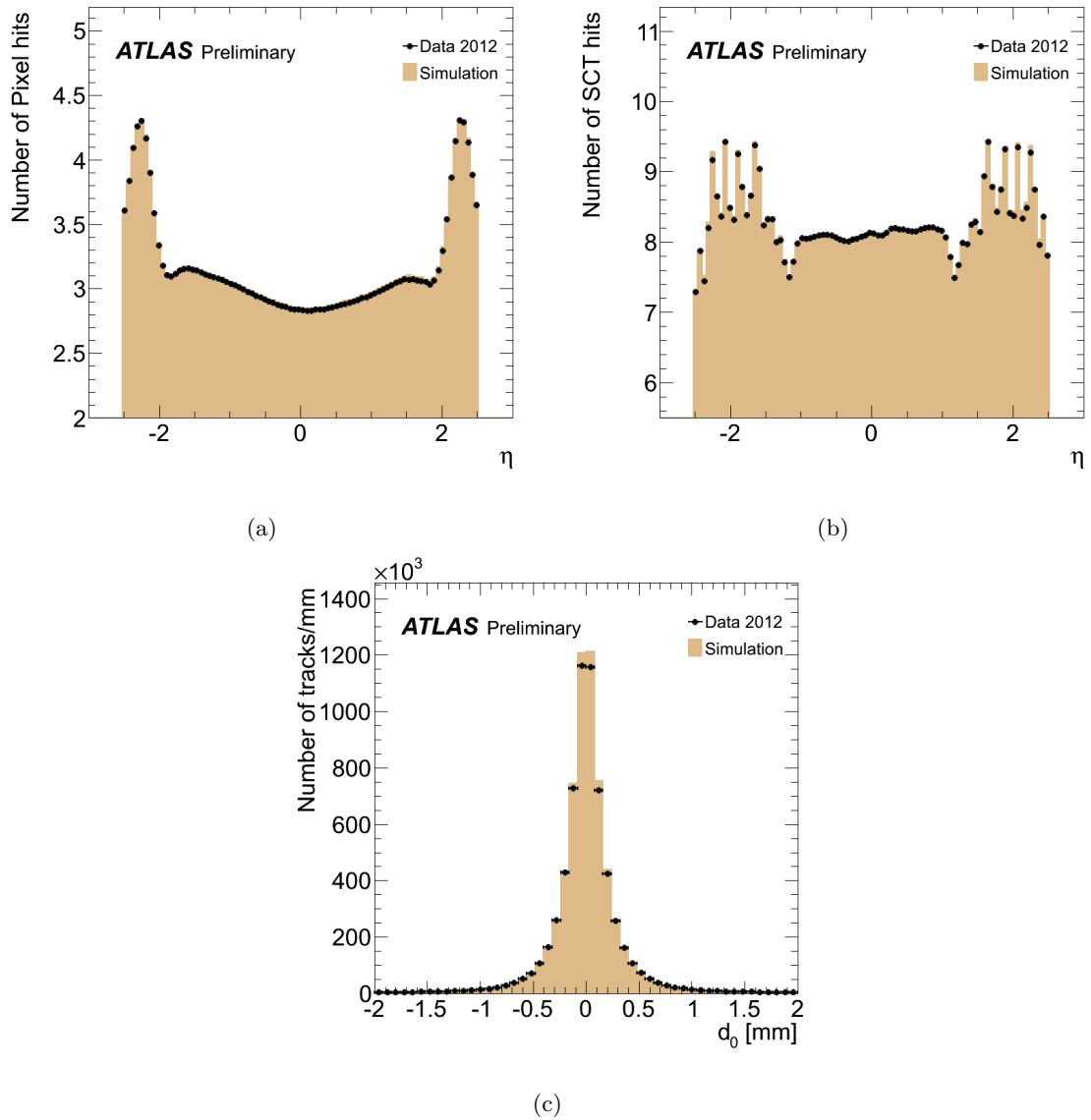


Figure 3.3: 2012 data and simulation comparison of Pixel hits vs eta (a) SCT hits vs eta (b), transverse impact parameter d_0 (c).

3.1.1 Vertex finding algorithm

Two types of vertex can be distinguished in an ATLAS event:

- **Primary vertex (PV):** This vertex is characterized by a high particle multiplicity. Vertices due to pile-up interactions may cause the PV reconstruction efficiency to decrease, both for the shadowing (when the nearby interaction prevents the reconstruction of the vertex) and for the closeness of two vertices that cannot be resolved.
- **Secondary vertex:** Vertex corresponding to the decay of short-lived particles, such as b-hadrons. These vertices are usually characterized by a lower track multiplicity with respect to PV.

PV are reconstructed using an iterative vertex finding algorithm [41]. Vertex seeds are obtained from the z-position at the beamline of the reconstructed tracks. An iterative χ^2 fit is

made by using seed and nearby tracks. Each track carries a weight that is a measure of its compatibility with the fitted vertex. Tracks displaced by more than 7σ from the vertex are used to seed a new vertex. The procedure goes on until no additional vertices are found. At least two charged particles with $|\eta| < 2.5$ and $p_T > 400$ MeV are required to define the interaction.

Figure 3.4 shows the number of reconstructed vertices per event in 2012 data, as a function of the number of pile-up interactions. Data are compared with MC expectations.

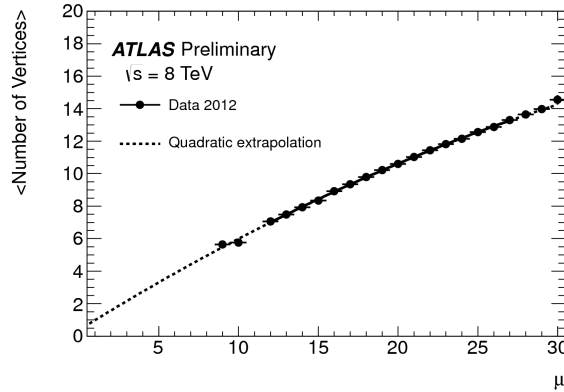


Figure 3.4: Average number of reconstructed primary vertices per event as a function of average number of pp interactions per bunch crossing measured for the data of 2012. Data are collected using a minimum bias trigger. A second order polynomial fit is performed in the upper range of μ . For the lower values of μ the result of extrapolation is shown.

3.2 Leptons

Leptons represent the final state of the process studied in this thesis. Their signature and the procedure adopted in ATLAS for their reconstruction and identification is the subject of this section. Tau lepton reconstruction is postponed after the jet reconstruction; the reason for this will become clear in going further with the chapter.

3.2.1 Electron reconstruction

The characteristic signature of electrons consists in a track in the ID together with a narrow shower in the EM calorimeter. Electrons reconstruction uses different algorithms depending on where the electron is found: we refer to *central electron*, if the electron is in $|\eta| < 2.5$, and to *forward electron*, if it is in $2.5 < |\eta| < 4.9$ [42]. An electron in the central region is reconstructed with an algorithm that combines the energy deposit in the EM with the track in the ID. First the EM cluster is built using the so-called sliding window algorithm, which adds together energy deposits with window size 3×5 in units of 0.025×0.05 in (η, ϕ) space. Then the reconstructed track is extrapolated from the ID to the EM calorimeter to look for a matching with the cluster. A track and a cluster are considered to be successfully matched if the distance between the track impact point and the EM cluster barycentre is $|\Delta\eta| < 0.05$. An electron candidate is considered to be reconstructed if at least one track is matched to the seed cluster. The four-momentum of central electrons is computed using information from both the final cluster and the best track matched to the seed cluster.

Forward electrons are instead reconstructed using only the information from the EMEC and the forward calorimeter. No distinction is then possible between electrons and photons. A topological cluster algorithm is used: cells with deposited energy significantly above the noise

level are grouped in three dimensions in an iterative procedure, starting from seed cells. Due to the reduced detector information forward electrons are used in physics analysis only for $E_T > 20 \text{ GeV}$.

Once the reconstructed candidate electron is found an identification criterion is applied, which relies on the shape of electromagnetic shower in the calorimeter as well as on tracking and track-to-cluster matching quantities. The identification can be based either on independent cuts on these quantities or on a single cut on the output of a likelihood function taking as input these quantities. Three reference sets of cuts have been defined with increasing background rejection power: *loose*, *medium* and *tight*. In general, each cut adds to the previous some additional requirements. From these categories, three more categories are derived, named "++". They were added in order to accomplish the trigger bandwidth restrictions for high luminosity, and have been used in most of 2012 ATLAS data analysis:

- **Loose++**: Is based on calorimeter informations only. Shower shape variables of the EM calorimeter middle layer and hadronic leakage variables are used. The hadronic leakage (R_{had}) allows to distinguish electron shower from hadrons shower, thanks to the different width of the two (usually electron showers have narrower width in η than those created by jets). It is defined as the ratio of the transverse energy reconstructed in the first layer of the hadronic calorimeter to the transverse energy reconstructed in the EM calorimeter. This category has a high identification efficiency ($\approx 97\%$) but poor background rejection.
- **Medium++**: With respect to the previous category this adds information from the first layer of the EM calorimeter, aimed to improve the $e - \pi$ separation. In addition further qualities on the ID tracks are requested. A background rejection of a factor 3-4 higher than the loose++ category is achieved, while the identification efficiency is reduced by about 10%.
- **Tight++**: This category adds many additional selections, as the particle identification using the TRT, the discrimination against photon conversions via a b-layer hit requirement and information about reconstructed conversion vertices. Background rejection is $\approx 10^5$ while the identification efficiency is $\approx 80\%$.

For the scope of this thesis the identification based on a likelihood output has been used. The electron likelihood makes use of signal and background probability density functions (PDF) of the discriminating variables. Based on these PDF, an overall probability is calculated for the object to be signal or background. The signal and background probabilities are combined into a discriminant on which a cut is applied. The electron reconstruction efficiency measured in 2011 and 2012 data is shown in figure 3.5, while figure 3.6 shows the electron identification efficiency for the Loose likelihood selection (LooseLLH) and the Tight++ selection (VeryTightLLH) in 2012 data [43].

3.2.2 Muon reconstruction

Muon reconstruction uses the available information of MS and ID systems, and partially the information of the calorimeter. Several reconstruction criteria are used, leading to different types of muons [44]:

- **Stand-alone (SA) muon**: only the hits in the MS stations are used to reconstruct the muon track. The hits are combined to form segments that are then used in a fit for the track reconstruction. The track is then extrapolated back to the point of closest approach to the beam line, taking into account the energy loss in the calorimeter. SA muons are used to extend the acceptance to the range $2.5 < |\eta| < 2.7$ where there is not coverage of the ID.

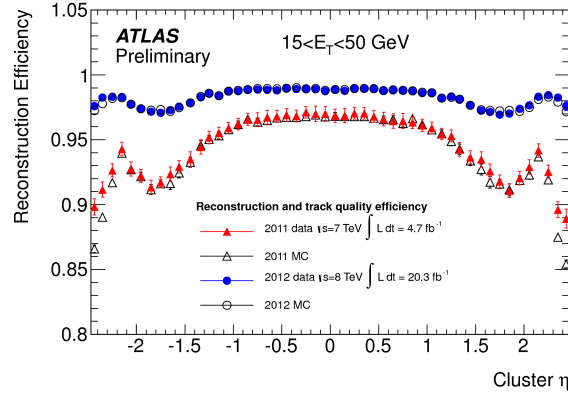


Figure 3.5: The efficiency of electron reconstruction including criteria on track quality (namely that there is at least one hit in the pixel detector and at least seven hits in the pixel and SCT detectors together) is shown as a function of the pseudorapidity for electrons with transverse energy E_T between 15 and 50 GeV for data (filled markers) and MC (open markers) for the 2011 (triangles) and the 2012 (circles) data samples. The total (statistical and systematic) uncertainty is displayed. The combined track reconstruction and the track-cluster matching performance is measured using a tag and probe method utilizing $Z \rightarrow e^+e^-$ events.

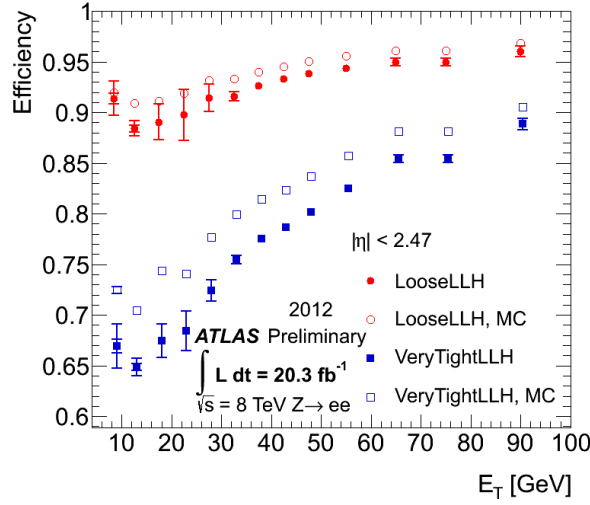


Figure 3.6: The identification efficiency of electrons from the $Z \rightarrow e^+e^-$ decay for the Loose, VeryTight Likelihood is shown as a function of E_T for $|\eta| < 2.47$. Both data and MC efficiencies are shown.

- Combined (CB) muon: a combination of the local MS track with the local ID track is used. CB muons are the most commonly used muons in physics analysis since they have the highest purity together with the best resolution on the muon kinematical properties.
- Segment-tagged (ST) muon: this muon is reconstructed using an *inside-out* algorithm. The ID track is extrapolated to the MS and then associated with at least one local track segment in the MDT chambers or CSC chambers. This type of muon is used to increase the acceptance on muons in all those cases in which the muon crosses only one layer of MS chamber (for example for low p_T muons).

- Calorimeter-tagged (CaloTag) muon: also this muon is reconstructed using an *inside-out algorithm*, combining the ID track with a calorimeter energy deposit compatible with a minimum ionizing particle. CaloTag muons have low purity but are used to recover acceptance in the uninstrumented regions of the MS.

To add redundancy muons have been reconstructed using two different software packages, named *Chains*. The first chain performs a statistical combination of the track parameters of the SA and ID tracks using the corresponding covariance matrices. The second chain performs a global refit of the muon tracks using the hits in both the MS and the ID systems. A unified reconstruction programme, *Chain 3*, has been developed and used in 2012 data taking; it includes the best features of both *Chain 1* and *Chain 2*.

The reconstruction efficiency for CB muons, evaluated with a tag-and-probe method on selected $Z \rightarrow \mu\mu$ events, is shown in figure 3.7.

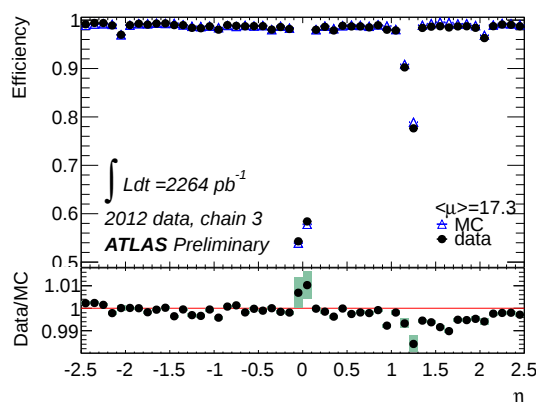


Figure 3.7: Reconstruction efficiency for combined muons versus the pseudorapidity for muon tracks with hits in at least two muon chambers in the MS. The efficiency drop at $\eta = 1.2$ is due to muon chambers not yet installed in the MS. The efficiency drop around $\eta = 0$ is instead due to the partial equipment of this region with muon chambers, because of the presence of services.

3.3 Jet reconstruction

Jets are produced in pp collisions from the hadronization of quarks and gluons. In this section a brief review of the jets reconstruction is given.

Jets reconstruction uses the informations given by the calorimeter system. Jet finding algorithm can use as input different type of seeds, which distinguish *calorimeter jets* and *track jets*. The input to *calorimeter jets* are *topological clusters* or *calorimeter towers* [45] [46].

Topological cluster: three-dimensional groups of calorimeter cells designed to follow the shower development taking advantage of the fine segmentation of the calorimeter (see chapter 2). The cluster is seeded by a calorimeter cell with a total energy above the noise ($|E_{cell}| > 4\sigma_{cell}$). σ_{cell} corresponds to the RMS of the energy distribution measured in events with a random trigger. Cells neighbouring the seed with $|E_{cell}| > 2\sigma_{cell}$ are then added to the cluster, increasing its size. Finally, all calorimeter cells neighbouring the formed topo-cluster are added. The topo-cluster has an energy equal to the energy sum of all the included calorimeter cells, zero mass and a direction calculated by averaging the pseudorapidities and the azimuthal angles of all the constituent cells.

Calorimeter tower: calorimetric cells that contain the energy deposit are projected into a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ grid. The energy in each tower is computed using only cells belonging to topological clusters.

Track jets are built from charged particle tracks originating from the PV; they are almost insensitive to pile-up and so constitute a stable reference for the jet reconstruction. The track jets must have at least two tracks each with $p_T > 0.5$ GeV and $|\eta| < 2.5$.

Once clusters and towers have been built, the jet reconstruction takes place. Jets are reconstructed using the *anti- k_t* [48] algorithm with a radius parameter $R = 0.4$ or $R = 0.6$. In this thesis jets with $R = 0.4$ seeded from a topological cluster and having $p_T > 25$ GeV and $|\eta| < 4.9$ are used. The *anti- k_t* algorithm is a *cluster algorithm* that builds jets by clustering them in an iterative procedure. The method can be described as follows:

1. Compute $d = \min(d_{ij}, d_{iB})$, where i, j are particles or jets, d_{ij} is the distance ΔR between them and d_{iB} is the distance between the entity i and the beam. The smallest distance is i.e. identified.
2. If $d = d_{ij}$ then merge i and j and goto 1)
3. If $d = d_{iB}$ then i is a single complete jet, remove it from the list

This procedure is repeated until no entities are left in the list. The quantities d_{ij} and d_{iB} are defined as follows:

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.2)$$

$$d_{iB} = p_{Ti}^{-2} \quad (3.3)$$

where

$$\Delta R_{ij}^2 = \Delta\phi_{ij}^2 + \Delta\eta_{ij}^2 \quad (3.4)$$

and R is the reference radius parameter. In this way soft particles tend to cluster with hard particles long before they cluster among themselves and the shape of the jet is unaffected by soft radiation.

Different quality selections for the jet reconstruction exist [47]: *Loose*, *Looser*, *Medium*, *Tight*. The *Looser* selection was designed to provide an efficiency above 99.8% with an high fake jet rejection as possible (fake jets deriving principally from calorimeter electronics noise, cosmic rays and no-collision backgrounds). The *Tight* selection was designed to provide a much higher fake jet rejection, above 99%, with a consequent lower identification efficiency.

Calorimeter jets are reconstructed from calorimeter energy deposits measured at the EM scale, the baseline signal scale for the energy deposited by EM showers in the calorimeter. Their transverse momentum is evaluated at the EM energy scale. For hadrons this leads to a jet energy measurement that is typically 15 – 55% lower than true energy. ATLAS has developed several jet calibration schemes aimed at determining the jet energy scale (JES). The jet calibration is performed by applying corrections derived from test beam and MC simulations. The JES strategy adopted for the calibration of jets used in the analysis is called *Local Cluster Weighting (LCW) calibration*. The LCW calibration uses properties of clusters to calibrate them individually. A more detailed description can be found in [45].

3.3.1 b-tagging algorithm

B-tagging is used in this analysis to put a veto on backgrounds involving high- p_T b-jets in the final state, such as the top background. B-jet tagging relies on b-hadrons properties, such as:

- A displaced secondary vertex due to the b-hadron long life time ($\approx 1.6 ps$). The b-hadrons typically travel a few mm in the ID. A large impact parameter d_0 is then measured
- High track multiplicity
- High p_T of the decay products

Various b-tagging algorithm have been developed [49] with the aim of achieving an high b-jet tagging efficiency for real b-jets whilst keeping the misidentification efficiency for c-jets and light flavour jets at low levels. The key ingredient for each b-tagging algorithm are the calorimeter jets, the tracks reconstructed in the ID and the selected primary vertex. B-tagging algorithms distinguish in impact parameter based algorithm, as the *IP3D* and the *JetFitter*, which use the transverse and longitudinal IP significance as input, secondary vertex based algorithm, as the algorithm *SV1*, which attempts to reconstruct the bottom-quark decay vertex, and a more complex algorithm, the *MV1* algorithm, which is a neural network based tagger that combines the information of the previous algorithms together with some other input variables. The *MV1* algorithm has been used in this analysis. A tag weight distribution is obtained and three working points for the b-tagging efficiency are defined, corresponding to efficiencies of 60%, 70% and 80% [50]. Figure 3.8 shows the performance of the *MV1* tagging algorithm evaluated in $t\bar{t}$ MC samples for jets satisfying $p_T > 20$ GeV and $|\eta| < 2.5$.

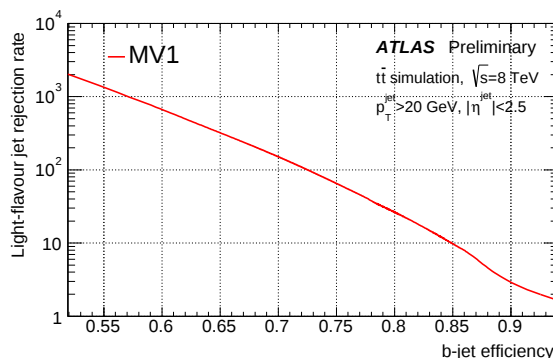


Figure 3.8: Performance (light-flavour rejection, defined as the inverse of the mistag rate, versus b-jet efficiency) of the MV1 tagging algorithm, as evaluated for jets with $p_T > 20$ GeV and $|\eta| < 2.5$, as evaluated on a sample of simulated $t\bar{t}$ events.

3.4 Tau reconstruction

Due to their mass taus are the only leptons that can also decay to hadrons. Figure 3.9 shows a diagram summarizing all the possible final states in tau decays. Leptonic tau decays cannot be distinguished from prompt electrons and muons, while for hadronically tau decays an algorithm for both the reconstruction and the identification is implemented in the reconstruction framework. Hadronic taus decays are present in the final state of the channel studied in this thesis; from now on we will call them simply taus (τ).

Taus are jets whose signature consists in a narrow shower respect to quark- and gluon-initiated jets and in the characteristic number of charged tracks of one (*1-prong* tau) or three (*3-prongs* tau) (see figure 3.10). To correctly discriminate taus from jets and other leptons, due to the similarity of their signature, the combined informations from many sub-detectors are used.

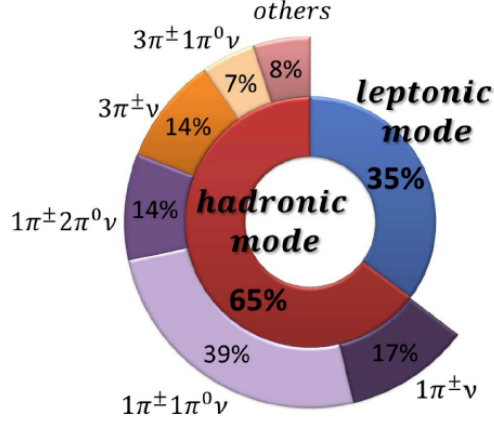


Figure 3.9: Tau lepton decay modes and corresponding branching fractions.



Figure 3.10: Typical signature of a tau (left) and quark/gluon-initiated (right) jets.

Tau reconstruction [68] starts with the individuation of the so-called tau-seed. Each jet reconstructed with the *anti- k_t* algorithm with a radius parameter $R = 0.4$ and having $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$ is potentially a good candidate to be a tau. Tracks are then associated to the tau candidate only if inside the *core-cone*, defined as the region within $\Delta R < 0.2$ of the jet axis. Tracks outside the *core-cone*, in the so-called *isolation annulus*, $0.2 < \Delta R < 0.4$, are also used for kinematical variables calculation and used in the identification stage. After the tau has been reconstructed, a multivariate technique (BDT or projective likelihood) is used for its discrimination against jets. A multivariate technique and a cut-based approach are also used to distinguish taus from electrons and muons respectively. Several calorimetric and tracking variables are computed and given as input to the BDT (or to the likelihood). One of the most discriminating variable is the *core energy fraction* defined as:

$$f_{core} = \frac{\sum_{i \in all}^{\Delta R_i < 0.1} E_{T,i}^{EM}}{\sum_{j \in all}^{\Delta R_j < 0.2} E_{T,j}^{EM}} \quad (3.5)$$

where $E_{T,i}(E_{T,j})$ is the transverse energy, calibrated at the EM energy scale, deposited in cell $i(j)$, and i runs over the cells in all layers associated with the τ candidate within $\Delta R < 0.1$ of the axis, while j runs over all cells in all layers within $\Delta R < 0.2$. This variable measures the energy fraction in the innermost part of the τ core cone and so accounts for differences of the τ shape with respect to jet shape. An important track variable is instead the *track radius* defined as:

$$R_{track} = \frac{\sum_i^{\Delta R_i \leq 0.4} p_{T,i} \Delta R_i}{\sum_i^{\Delta R_i \leq 0.4} p_{T,i}} \quad (3.6)$$

where i runs over all core and isolation tracks of the τ candidate, within $\Delta R \leq 0.4$. ΔR_i is defined relative to the τ axis and $p_{T,i}$ is the track transverse momentum. Figure 3.11 compares

the distribution of these variables in simulated τ events with the corresponding distribution in a jet sample obtained in 2012 data. Tau identification efficiency is defined as the number of truth-

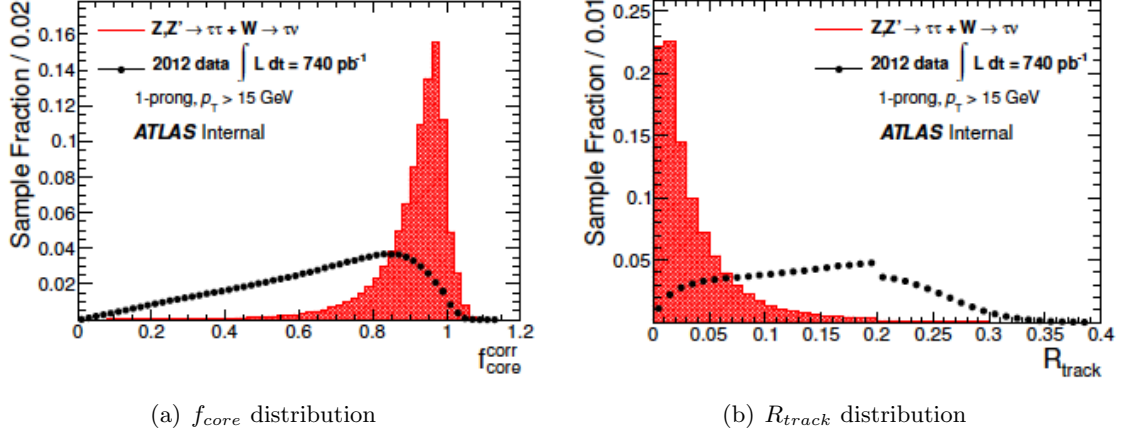


Figure 3.11: Distributions of a selection of jet discriminating variables for simulated $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ signal samples and a jet background sample selected from 2012 data.

matched τ candidates passing identification criteria divided by all true hadronic tau decays. Three working points have been defined, namely *Loose*, *Medium* and *Tight* and corresponding to efficiencies reported in table 3.1. In figure 3.12 the tau identification efficiency with BDT is

	Loose	Medium	Tight
1-prong τ	70%	60%	40%
3-prongs τ	65%	55%	35%

Table 3.1: Tau identification efficiency working points.

shown for both 1-prong and 3-prongs τ as a function of the number of reconstructed primary vertices in the event. The BDT is approximately independent of the pile-up conditions.

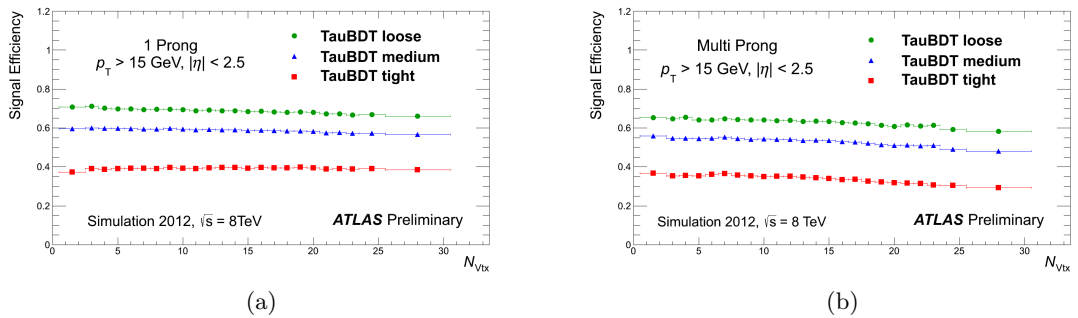


Figure 3.12: Signal efficiency for (a) 1-prong and (b) 3-prongs τ candidates for the three working points of the BDT tau ID as a function of the number of reconstructed vertices. The efficiency were obtained using $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ simulated samples.

1-prong tau has a characteristic signature quite similar to that produced by an electron. For this reason after the identification step a further discrimination with respect to electrons

is needed in order to reduce the probability of fake taus. The electron veto BDT is then built by using some calorimeter informations, such as the presence of transition radiation which is typically emitted by an electron, or the longitudinal and lateral extension of the shower in the hadronic calorimeter. Usually the shower produced by the tau decay products is longer and wider with respect to the one created by an electron. Even in this case three working points have been defined corresponding to the e-veto efficiency reported in table 3.2.

	Loose	Medium	Tight
e-veto efficiency	95%	85%	75%

Table 3.2: Electron veto efficiency working points for τ with $p_T > 20$ GeV and satisfying the BDT *Loose* ID.

A small probability of mis-identifying a muon as a tau exists in all those cases in which an energetic cluster in the EM calorimeter is associated to a muon track. By simply removing geometrical overlaps between muons and taus most of these fake taus disappear. However it can still be that the muon has not been reconstructed because passed in an inefficient region of the MS. For this reason a cut-based discriminant is trained and used, with the effect of reducing the muon fakes of about 40%.

3.5 Missing transverse momentum

The missing transverse momentum, E_T^{miss} or MET, is defined as the momentum imbalance in the plane transverse to the beam axis.

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad (3.7)$$

This imbalance can derive from undetected particles, such as neutrinos produced in the pp interaction. The E_T^{miss} is defined as the negative vectorial sum of the momenta associated to all the reconstructed objects. In particular it is reconstructed mainly from the calorimeter energy deposits that are associated with reconstructed and identified objects, the so-called *calorimeter term* $E_T^{miss,calo}$, but it takes also into account muons reconstructed from the ID to recover muons in regions not covered by the MS, the so-called *muon term* $E_T^{miss,\mu}$ [52]. Then

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,calo} + E_{x(y)}^{miss,\mu} \quad (3.8)$$

where

$$E_x^{miss,calo} = - \sum_{i=1}^{N_{cell}} E_i \sin\theta_i \cos\phi_i \quad (3.9)$$

$$E_y^{miss,calo} = - \sum_{i=1}^{N_{cell}} E_i \sin\theta_i \sin\phi_i \quad (3.10)$$

E_i , θ_i and ϕ_i are the energy, the polar angle and the azimuthal angle respectively. The $E_T^{miss,calo}$ term is calculated using calorimeter cells calibrated according to the reconstructed object to which they are associated: electrons, photons, hadronically decaying τ , jets and muons. Cells not associated to any reconstructed objects are also taken into account to improve the E_T^{miss} resolution, in the so-called $E_T^{miss,CellOut}$ term. In both cases only cells belonging to three-dimensional topological clusters are included in the calculation, in order to suppress noise contribution. The

$E_T^{miss,\mu}$ term is calculated from the momenta of muon tracks reconstructed with $|\eta| < 2.7$. Whenever the momenta of physics objects in the analysis are smeared or rescaled the effect is propagated to the E_T^{miss} which is recomputed.

A large deterioration of the E_T^{miss} resolution was observed in 2011 data when the average number of pile-up interactions per event increased. Methods to suppress pile-up are therefore needed, which can restore the E_T^{miss} resolution to values more similar to those observed in the absence of pile-up [53]. To this purpose in this analysis two modified definitions of E_T^{miss} are used: the MET_STVF and the $MET_TrackClj$. The MET_STVF is a pile-up robust variable computed by scaling down the energy in SoftTerm (SoftTerm is the E_T^{miss} term accounting for calibrated jets with $7 < p_T < 25$ GeV together with the unassociated cells) by the fraction of the momenta of the SoftTerm-matched tracks that are also associated with the hard scattering vertex. The method used to correct the E_T^{miss} for pile-up effects is called Soft Term Vertex Fraction method [53]. MET_STVF is defined as:

$$E_T^{miss} \times STVF \quad (3.11)$$

where

$$STVF = \frac{\sum_{tracksSoftTerm,PV} p_T}{\sum_{tracksSoftTerm} p_T} \quad (3.12)$$

where the sums are taken over the tracks unmatched to physics objects and PV denotes the tracks associated to the PV. Figure 3.13 shows the E_T^{miss} distribution before and after the rescaling for the STVF correction.

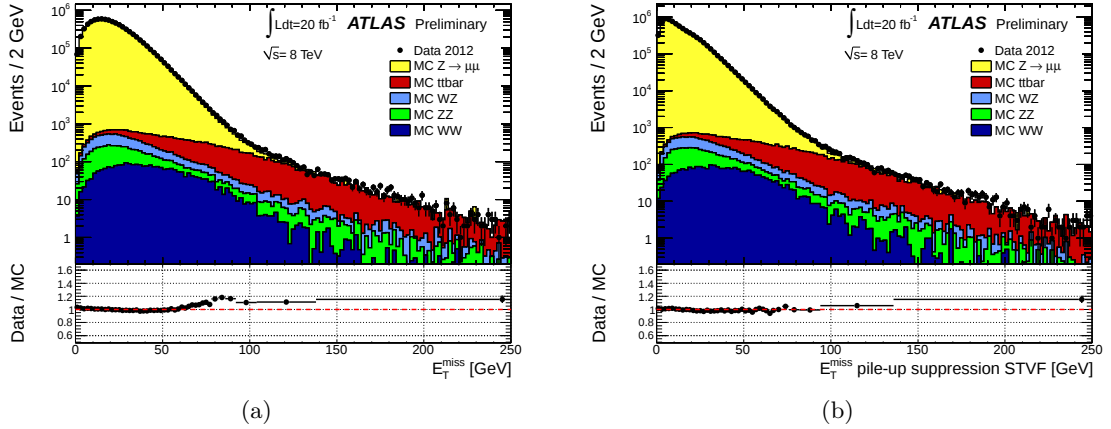


Figure 3.13: Distribution of E_T^{miss} as measured in a data sample of $Z \rightarrow \mu\mu$ events before (a) and after (b) pile-up suppression. The expectation from Monte Carlo simulation is superimposed and normalized to data, after each MC sample is weighted with its corresponding cross-section. The lower parts of the figures show the ratio of data over MC.

The $MET_TrackClj$ is the E_T^{miss} calculated using the missing transverse momentum reconstructed from tracks instead of calorimeter objects, and is expected to be less dependent on pileup conditions. In addition the track momenta are replaced by the calorimeter energy when an electron or a jet contribute to the track-based E_T^{miss} variable.

Chapter 4

Analysis

This chapter describes the data analysis for the study of the associated production of the Higgs boson with a W boson. Goal of the work is to explore the WH production at $\sqrt{s} = 8$ TeV by looking at those events with the Higgs boson decaying in two W bosons. In particular the study will focus on final states with exactly two charged leptons ($e^\pm$ or $\mu^\pm$) plus a tau from the three W decays. The channel studied is then the following:

$$WH \rightarrow WWW^* \rightarrow l\nu l\nu\tau_{had}\nu \quad (4.1)$$

with $l = e, \mu$. As mentioned in the introduction, the associated production of the Higgs boson is the only production mode in which the Higgs boson has not been observed yet (including the associated $t\bar{t}H$ mode) and so it is an important channel for the completeness of the Higgs boson physics program. The WH process suffers from a lower cross section with respect to the other two dominant production mechanisms [31], the gluon-gluon fusion and the Vector Boson Fusion, but the clear signature of three high momentum leptons in the final state makes this channel accessible. Furthermore this process probes the coupling of the Higgs boson with a gauge boson, specifically the W, whose coupling is predicted by the Standard Model to be $g_{HVV} = \frac{M_V^2}{v}$, where M_V^2 is the squared mass of the gauge boson and $v = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV is the vacuum expectation value of the Higgs boson fixed by the Fermi coupling G_F . The measurement of this coupling and its comparison with the Standard Model expectations would again confirm (or not) the validity of the Standard Model predictions leading to a more complete understanding of the Higgs mechanism.

4.1 Physics process

The signature of the process in equation (4.1) is three charged leptons plus appreciable missing transverse momentum due to the neutrinos from the W decays. Since the Higgs boson is neutral the event total charge will be equal to ± 1 , depending on the charge of the W boson associated with the Higgs boson. In the decay of the spin-0 Higgs boson to W bosons the angular momentum is conserved, and the W's are emitted with opposite spins. In the weak decay of the W, due to the V-A nature of the interaction, the positively charged lepton is likely to be emitted in the direction of the W^+ spin and the negative lepton in the opposite direction of the W^- spin (see figure 4.1). This results in the two charged leptons being emitted close to each other: their di-lepton invariant mass together with their angular distance will be small.

4.1.1 Signal yield

Before starting with the analysis description it is useful to estimate the event yield of this channel, which means how many WH events, called *signal events* from now on, should we

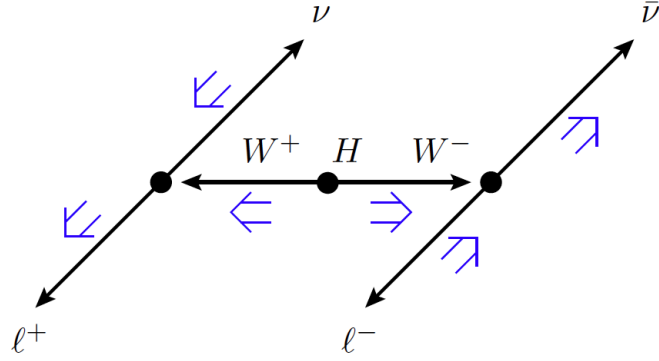


Figure 4.1: Illustration of the $H \rightarrow WW^*$ decay. The small arrows indicate the particles directions of motion and the large double arrows indicate their spin projections. The spin-0 Higgs boson decays to W bosons with opposite spins, and the spin-1 W bosons decay into leptons with aligned spins. The H and W decays are shown in the decaying particle's rest frame. Because of the V-A decay of the W bosons, the charged leptons have a small opening angle in the laboratory frame.

expect in proton-proton collisions at a centre-of-mass energy of 8 TeV. We can evaluate the expected number of WH events by considering the following formula:

$$N_{events} = \sigma_{WH} \times \mathcal{L} \times \text{branching fraction} \times \text{acceptance} \times \text{efficiency} \quad (4.2)$$

where referring to figure 1.4 the WH cross section, as predicted by the Standard Model and calculated at NNLO, is $\sigma \approx 0.7 \text{ pb}$ (table 1.1), while the integrated luminosity is 20.3 fb^{-1} . In order to get the event yield in equation (4.1) we have to select only the $H \rightarrow WW^*$ decays and ask for all the W bosons to decay leptonically. The branching ratio for the $H \rightarrow WW^*$ decay is $\approx 20\%$ while the W bosons decay 11% of the time in each lepton flavour. The flavour combinations for the 3-lepton final state given three possible lepton flavour (electron, muon or tau) are $(3)^3 = 27$, 12 of which have exactly one τ , 6 two τ , and 1 three τ , regardless if the tau decays hadronically or not. By asking for exactly one tau and having in mind that the tau branching ratio to hadrons is $\approx \frac{2}{3}$ (figure 3.9) we can compute the signal event yield:

$$N_{WH \rightarrow WWW^* \rightarrow l\nu l\nu \tau_{had}\nu} \approx 0.7 \text{ pb} \times 20.3 \cdot 10^3 \text{ pb} \times 0.2 \times 0.3^3 \times f_c \sim 30 \text{ events} \quad (4.3)$$

where f_c are all the flavour combinations with two leptons plus one tau:

$$f_c = \frac{12}{27} \times \frac{2}{3} + \frac{6}{27} \times \frac{2}{3} \times \frac{1}{3} \times 2 + \frac{1}{27} \times \frac{2}{3} \times \frac{1^2}{3} \times 3 \approx 40\% \quad (4.4)$$

To get a realistic value of the expected number of signal events we have to take into account the detector response and so consider the leptons selection and reconstruction efficiencies. By considering an average efficiency of 90% for muons, 80% for electrons and 55% for taus with the "medium" selection (chapter 3) we then get:

$$N_{WH \rightarrow WWW^* \rightarrow l\nu l\nu \tau_{had}\nu} \approx 10 \text{ events} \quad (4.5)$$

The number of expected signal events in equation (4.5) is an upper limit of the number we could be able to observe in the analysis since we didn't consider the analysis acceptance (p_T cuts, η and ϕ restrictions), which will reduce this number. To understand if we would have the chance or not to see the signal, with such a low event yield, we should roughly also estimate the expected number of background events. To this purpose a first idea can be obtained comparing signal and main backgrounds cross section at a centre-of-mass energy of 8 TeV. This comparison can be found in the next paragraph.

4.1.2 Background composition

Background topologies that mimic our signal events could be classified in two main categories:

- 3 real leptons backgrounds
- real + fake leptons backgrounds

The first category will include all those processes with exactly three real leptons, the diboson Standard Model processes $pp \rightarrow WZ$ and $pp \rightarrow W\gamma^*$ (WZ/γ^* in the following), as well as the tri-boson background $pp \rightarrow VVV$ ($V = Z, W, \gamma$), together with the $pp \rightarrow ZZ^*$ production with an undetected lepton. In the second category will fall those processes with one or more fake leptons, as the $pp \rightarrow Z + jets$, the $pp \rightarrow WW$ and the $pp \rightarrow t\bar{t}$ processes. Fake leptons originate from hard jets (jets with high p_T) that leave a track in the ID and some calorimeter deposits and can then be misidentified as electron or tau (the muon fake probability from jets is lower). We will also refer to fake leptons in case of real leptons from light flavour, beauty and charm decays. The distinguishing features of these backgrounds, discussed in detail in section 4.4, motivate the definition of event categories based on the flavour of leptons.

Table 4.1 lists signal and main backgrounds cross sections. With the exception of WH cross section the others cross sections listed correspond to the experimental value measured by ATLAS in Run1, as can be seen from figure 4.2.

Process	Cross section (pb)
WH	$0.70 \pm 0.01 \pm 0.02$ (theory)
$t\bar{t}$	$242.4 \pm 1.7 \pm 10.2$ (data)
$Z + jets$	$(27.9 \pm 0.2 \pm 1.1) \cdot 10^3$ (data)
WZ	$20.3^{+0.8+1.4}_{-0.7-1.3}$ (data)
WW	$71.4^{+1.2+5.5}_{-1.2-4.9}$ (data)

Table 4.1: Inclusive cross section for signal and main backgrounds. WH cross section was evaluated in [16]; the 1% error is due to the uncertainty on the QCD scale, while the 2.3% error is due to the uncertainty on the PDF. $Z + jets$ and WZ cross sections have been measured with 7 TeV data, while the value reported for $t\bar{t}$ and WW cross section has been extracted from 8 TeV data.

The efficiency for the 3-lepton selection is process-dependent, which means that the requirement of three leptons in the final state, one of which must be a tau, implies an inhomogeneous reduction of the various background processes. In table 4.2 the *effective* backgrounds cross-section after the 3-lepton selection is reported; for each background process the expected fraction of events with respect to the total background is also shown. The *effective* cross-section is the actual cross section seen in the analysis, thus scaled for accounting for three reconstructed objects in the event. The *effective* cross section is defined as:

$$\sigma_{eff} = \frac{N_{3lep\ events}}{\mathcal{L}} \quad (4.6)$$

where $N_{3lep\ events}$ is the number of events having two leptons plus one tau. Table 4.2 has to be compared with table 4.1. Because of the 3-lepton detection efficiency more backgrounds processes with respect to those in the previous table are here reported. Indeed their *effective*

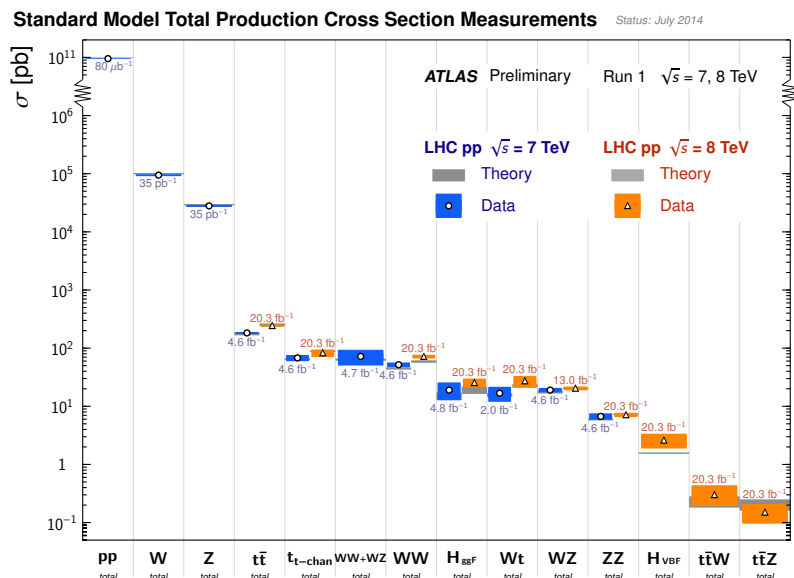


Figure 4.2: Summary of several Standard Model total production cross section measurements, corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher. The luminosity used for each measurement is indicated close to the data point.

cross section becomes comparable or even higher than those of main background samples listed in table 4.1. It is worth to note that the main background is still the Z +jets process. The origin of such background is examined in appendix B.

Similarly to what done for the background, we can calculate the *effective* cross section for the signal, since this is the number to compare with those in table 4.2 to get an idea of the difficulty of the measurement: the value found is ≈ 0.3 fb, smaller than that of all the background processes. The strategy followed to reduce the huge background is described in detail in section 4.4.

4.2 Data sample and data taking conditions

The data sample consists on the full integrated luminosity collected in 2012 during the 8 TeV proton-proton run, corresponding to 20.3 fb^{-1} . During the data taking period the status of the several detector components was monitored by the Data Quality (DQ) group. The role of the DQ was to monitor each subsystem, to record any inefficiency and to transfer the information to analysts. DQ compiled a list called *Good Runs List* (GRL) to be used, including only good runs. For physics analysis it is crucial to know about degraded conditions in order to be able to exclude data from periods where detector problems would affect measurements. In figure 4.3 the luminosity weighted relative fraction of good quality data delivered by the various subsystems during proton-proton collisions between April 4th and December 6th is shown. In figure 4.4 the total LHC delivered luminosity versus time (green distribution), recorded (yellow distribution), and certified to be good quality data (blue distribution) during stable beams for proton-proton collisions at 8 TeV centre-of-mass energy in 2012 are shown.

Process	σ_{eff} (fb)	Total bkg fraction
Top	200	$\sim 2\%$
$Z + jets$	$8 \cdot 10^3$	$\sim 89\%$
WZ/γ^*	40	$\sim 0.4\%$
WW^*	15	$\sim 0.2\%$
VVV	0.5	$\sim 0.001\%$
ZZ^*	30	$\sim 0.3\%$
$Z\gamma$	230	$\sim 3\%$
$W + jets$	470	$\sim 5\%$
$W\gamma$	10	$\sim 0.1\%$
$ggF/VBF/t\bar{t}H$	1	$\sim 0.01\%$

Table 4.2: Backgrounds event fraction with respect to the total background amount after the three leptons selection. Numbers here reported are obtained using MC simulation only. The corresponding *effective* cross section obtained is also reported. The simulation of the ATLAS detector and its response is also included. Leptons are reconstructed and identified using the common techniques discussed in chapter 3.

ATLAS p-p run: April-December 2012										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.9	99.1	99.8	99.1	99.6	99.6	99.8	100.	99.6	99.8	99.5
All good for physics: 95.5%										
Luminosity weighted relative detector uptime and good quality data delivery during 2012 stable beams in pp collisions at $\sqrt{s}=8$ TeV between April 4 th and December 6 th (in %) – corresponding to 21.3 fb ⁻¹ of recorded data.										

Figure 4.3: Luminosity weighted relative fraction of good quality data delivery by the various ATLAS subsystems during LHC fills with stable beams in pp collisions at $\sqrt{s} = 8$ TeV, and after switching the tracking detectors on. Runs between April 4th and December 6th, corresponding to a recorded integrated luminosity of 21.3 fb⁻¹, are accounted.

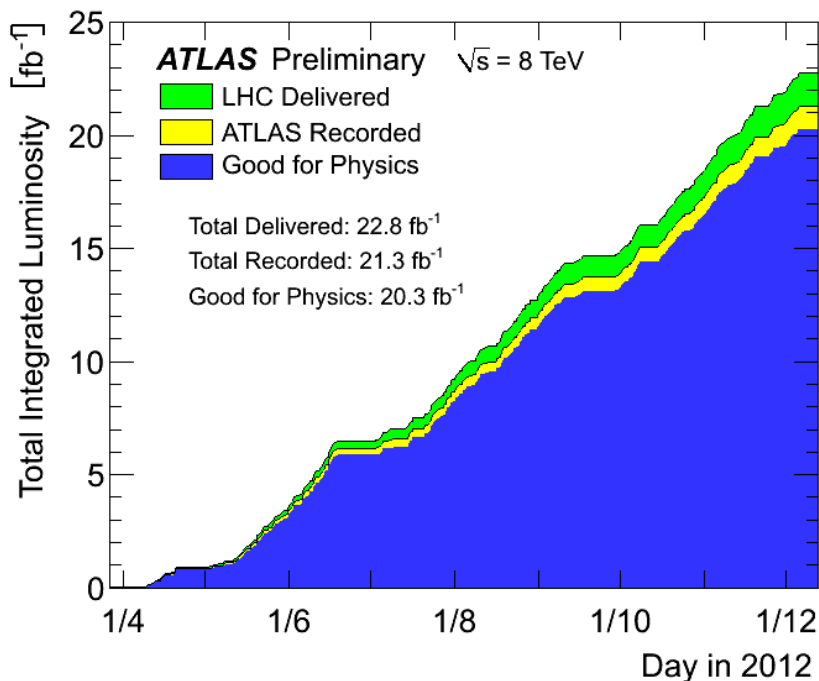


Figure 4.4: Total integrated luminosity and Data Quality in 2012. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity reflects the DAQ inefficiency.

4.2.1 Data preparation

Data are recorded and grouped in periods named from A to M. A data period is a selection of runs with a coherent configuration of the detector and the trigger. Thus, events belonging to the same period have been processed and reconstructed with the same software configuration (trigger and detector conditions). Each period contains several "ntuples" in which data are stored. The analysis here described uses a recent ntuple format called *COMMON*, since it contains lots of informations in common to several analyses. To limit the running time on this huge amount of data we accessed the ntuples through the ATLAS grid infrastructure and reduced their size by performing a preliminary three leptons selection. This reduction allowed the analysis to take place locally at Roma Tre University. More details about the data format and the reduction process used to this thesis scope can be found in appendix A.

4.3 Monte Carlo samples

Signal and background processes are modelled with Monte Carlo samples prepared to include the simulation of particles interaction with detector material. Simulations are largely used in particle physics both to get the analysis signal and background estimation.

Signal

The associated production of the Higgs boson is modelled with `PYTHIA8` MC generator [54] and the cross section has then been normalized to the next-to-next-to-leading order (NNLO)

calculation. The Higgs boson branching ratios are calculated with PROPHECY4f [55]¹. The MC generators used to simulate all those processes including a Higgs boson with a mass of $m_H = 125 \text{ GeV}$ are reported in table 4.3, together with the cross section assumed in the analysis.

Process	Higgs decay	Generator	Cross section @ 8 TeV(pb)
$WH + ZH$	$H \rightarrow WW^*$	PYTHIA v8.165	0.24
ggF	$H \rightarrow WW^*$	POWHEG-BOX v1.0 + PYTHIA v8.165	0.44
VBF	$H \rightarrow WW^*$	POWHEG-BOX v1.0 + PYTHIA v8.165	0.035
$t\bar{t}H$	$H \rightarrow WW^*$	PYTHIA v8.165	0.028

Table 4.3: MC generators used to model Higgs boson production and decay. On the WH and the $t\bar{t}H$ a 3-leptons filter is applied at the event generation level, while on the ZH , ggF and VBF generation a 2-leptons filter is applied.

Backgrounds

In modelling background processes several MC generators have been used. The simulation of QCD processes is generally divided in two steps: the first step is the event generation, which includes the simulation of the particular process from the proton-proton collision. This part is simulated using POWHEG [56], ACERMC [62], MADGRAPH [61], depending on the particular process. The second step consists in parton showering and hadronization, usually modelled by PYTHIA (version 6 or 8), HERWIG [58] or JIMMY [59]. Electroweak processes are instead modelled with ALPGEN [57] or SHERPA [60], even if including jets in the final state. Except for the top samples produced at NNLO and those samples produced with POWHEG, the background samples are normalized to the NLO cross section. Acceptances and efficiencies are obtained for most processes from a full simulation [64] of the detector using GEANT4 [63]. In few MC samples, as the WZ/γ^* sample, the detector response has been obtained through a fast simulation software. A realistic treatment of the pile-up conditions is also included. Minimum bias events are reweighted so that the distribution of the average number of interactions matches what observed in data. The CT10 [65] parton distribution function (PDF) is used for the POWHEG samples, while the CTEQ6L1 [66] is used for the ALPGEN, MADGRAPH and PYTHIA samples.

The complete list of MC samples and generators used in the analysis can be found in [67]. In table 4.4, background processes are grouped to form the main categories already shown in table 4.2. For each subprocess the number of fake leptons needed to enter the tri-lepton selection is specified.

To account for differences in the integrated luminosity of the MC samples with respect to the integrated luminosity recorded in data, (the former being much higher to reduce the statistical uncertainty on the templates), a scale factor has to be applied to MC samples. The scale factor is evaluated as the ratio of the recorded luminosity to the Monte Carlo luminosity:

$$SF_{lumi} = \frac{\mathcal{L}_{data}}{\mathcal{L}_{MC}} = \frac{\mathcal{L}_{data}}{N} \times \sigma \quad (4.7)$$

where N is the number of generated events for a given MC sample. This SF is then modified to include the effect of pile-up on MC events, the so-called MC *pile-up reweighting*. To model

¹PROPHECY4f is a MC event generator for precise simulations of the Higgs boson decay $H \rightarrow WW/ZZ \rightarrow 4 \text{ fermions}$. In calculating the partial width of the $H \rightarrow 4 \text{ fermions}$, the interference between the $H \rightarrow WW \rightarrow l\nu l\nu$ and the $H \rightarrow ZZ \rightarrow ll\nu\nu$ diagrams has to be taken into account. However, to the purposes of this thesis, the events generation for the signal process is done by including all the possible final states in the W decays, not only the leptonic ones. This means that the $\text{BR}(H \rightarrow WW)$ is used in the calculation instead of the specific $\text{BR}(H \rightarrow WW \rightarrow l\nu l\nu)$ and so the interference term is neglected.

Name	Process	Features
Top		
$t\bar{t}$	$t\bar{t} \rightarrow WbW\bar{b}$	one fake lepton
t	$t\bar{b}, tq\bar{b}$ tW, tZ	two fake leptons
Misidentified leptons		
Z+jets	$pp \rightarrow Z \rightarrow l^+l^- + jets$	fake lepton from jet
W+jets	$pp \rightarrow W + jets, W^\pm \rightarrow l^\pm\nu$	two fake leptons
WW	$pp \rightarrow WW, W^\pm \rightarrow l^\pm\nu$	one fake lepton
$W\gamma$	$pp \rightarrow W\gamma, W^\pm \rightarrow l^\pm\nu$	γ misidentified as e + one fake lepton
$Z\gamma$	$pp \rightarrow Z\gamma, Z \rightarrow l^+l^-$	γ misidentified as e
Other dibosons		
ZZ^*	$pp \rightarrow ZZ^*, Z \rightarrow l^+l^-$	unidentified lepton
WZ/γ^*	$pp \rightarrow WZ/\gamma^* \rightarrow l\nu l^+l^-$	three real leptons
Drell-Yan (DY)		
Z/γ^*	$pp \rightarrow Z/\gamma^* \rightarrow ee, \mu\mu, \tau\tau$	one fake lepton
Tribosons		
	WWW^*	three real leptons
VVV	ZWW^*, WWg^*	one undetected lepton
	$ZZZ^*, Z \rightarrow l^+l^-$	three undetected leptons

Table 4.4: Background topologies to the $WH \rightarrow WWW^*$ analysis, in the final state with three charged leptons plus neutrinos.

the effects of pile-up, the Monte Carlo samples for the above processes were simulated with a fixed distribution of additional minimum-bias interactions. This is subsequently reweighted to the distribution observed in the data, taking into account the mean number of interaction per bunch crossing in both data and Monte Carlo as a function of the data-taking period. This gives a weight, weight_{PU} , that has to be applied to the Monte Carlo events in order to reproduce the distributions measured in data. The last, but not least, correction to be taken into account is the differences measured in leptons reconstruction between MC and data. This correction, named SF_{lep} in the following, is evaluated from dedicated studies carried on by each performance group and has to be applied to each event, depending on the flavour of the leptons in the event. In this framework, once a template is obtained applying the analysis selection and scaling for the SF_{lumi} , it has also to be scaled for an event-by-event correction:

$$\text{MC_event_weight} = \text{weight}_{PU} \times \text{SF}_{lep0} \times \text{SF}_{lep1} \times \text{SF}_{lep2} \quad (4.8)$$

4.4 Analysis strategy

As previously discussed, the associated production process at the LHC has a small cross section with respect to many other physics processes; for this reason the small statistics predicted for WH events makes the measurement not possible. Only an upper limit on the observed cross section can be obtained. In this sense the work presented in this thesis is an explorative study of the WH sub-process in (4.1), in view of a future measurement with higher statistics and of a combination with the other VH analysis ².

The analysis described here has many analogies with the purely leptonic analysis, without any hadronic tau in the final state, which is documented in [67]. The two analyses share many definitions and some analysis tools, together with most of MC samples used, making it possible to combine the results and put a more stringent upper limit on the $WH \rightarrow WWW^*$ cross section. However the tau request was addressed in this thesis for the first time: new background processes with tau-fakes arised, as the Z+jets and the WW backgrounds. This implied that new analysis techniques and tools had to be specifically developed.

The analysis was divided in two stages: first a *cut-based* strategy has been followed in order to reduce the background contamination of signal topology events (section 4.5). Topological and kinematical selections were applied to the events sample previously categorized in signal regions depending on the leptons flavour and charge. Then a multivariate approach was used to better discriminate the signal from the background minimizing the loss of signal events (chapter 5).

4.5 Cut-based analysis

This section describes how events with characteristics resembling those of the desired signal have been selected, while rejecting as many non-signal events as possible. A cut-based analysis uses a sequence of cuts to select events of interest, with the aim of reducing the background-like events and maximizing the significance, defined as:

$$s = \sqrt{2[(S + B) \cdot \ln(1 + (S/B)) - S]} \quad (4.9)$$

where S and B are respectively the number of signal and backgrounds events. When the number of signal events is very small compared to the number of background events the equation (4.9) can be replaced by the simpler $S/\sqrt{B}$.

²The analysis is optimized for the 3-leptons WH process, but it has a significative acceptance also for the ZH process, and a small contribution from $H \rightarrow \tau\tau$ channel, the latter rejected by the event selection (cfr table 4.9). For this reason, in the following we will refer to the VH process as the signal we are looking for.

Before describing the *cut-based* analysis let's first observe that many background processes listed in table 4.2 contain two same flavour and opposite charge leptons. This is the case of those processes including a Z boson. For this reason it will be useful to split our 3-leptons final states according to the flavour and charge composition:

- Same flavour, opposite charge (SFOC): $e^+e^-\tau$, $\mu^+\mu^-\tau$; 4 combinations
- Same flavour, same charge (SFSC): $e^\pm e^\pm \tau^\mp$, $\mu^\pm \mu^\pm \tau^\mp$; 4 combinations
- Different flavour, opposite charge (DFOC): $e^\pm \mu^\mp \tau$; 2 combinations
- Different flavour, same charge (DFSC): $e^\pm \mu^\pm \tau^\mp$; 2 combinations

The flavour splitting always refers to leptons different from the tau. This event classification does partially reflect the background composition of our signal region. Indeed the SFOC signal events will be mainly contaminated by background topologies including a Z boson.

4.5.1 Event selection

First of all, events are required to have two leptons plus one tau. Quality cuts and trigger requirements for the leptons are included. Then, some kinematical cuts are applied to refine the events selection. Here a complete list of the applied selections:

- Preselection
- Lepton flavour splitting
- Number of jets
- Top veto
- Z veto
- MET cut
- ΔR_{01} cut

Preselection: Leptons are required to have a transverse momentum greater than 15 GeV and at least one out of the three leptons must have fired a trigger. Events are accepted if there is at least one successful match between a reconstructed lepton and an object responsible for an inclusive electron or muon trigger. For a successful match, muons and electrons reconstructed within a cone of $\Delta R = 0.2$ around a lepton trigger object are required to have $p_T > 24$ GeV for both electrons and muons. Auxiliary triggers for high p_T ($p_T > 60$ GeV for electrons, $p_T > 35$ GeV for muons) single leptons without isolation requirement are also used to recover efficiency.

Moreover, leptons are required to be isolated as well as far away from other reconstructed objects, like jets or leptons. Lepton isolation is both track-based and calorimeter-based, requiring respectively that the scalar sum of the transverse momenta of tracks from the primary vertex, in a cone around the candidate lepton track, and the scalar sum of the calorimeter cells transverse energy in a cone around the lepton track, is small compared to the lepton transverse momentum. The p_T dependent lepton isolation criteria is specified in table 4.5 for electrons and muons.

When the lepton is found in overlap with another reconstructed object, the undesired object is removed and only one object is kept in the analysis. The overlap removal condition is defined by a threshold on the two objects angular distance. The criteria adopted are explained in table 4.6.

Cone size	p_T range	Calorimetric isolation	Tracking isolation
$\Delta R = 0.2$	$> 20 \text{ GeV}$	$E_T \text{Cone}/p_T < 0.10$	$P_T \text{Cone}/p_T < 0.04$
	$< 20 \text{ GeV}$	$E_T \text{Cone}/p_T < 0.07$	$P_T \text{Cone}/p_T < 0.04$

Table 4.5: Lepton isolation criteria. A cone of radius $\Delta R = 0.2$ around the lepton is defined. The scalar sum for transverse momenta of tracks in this cone, $P_T \text{Cone}$, and the scalar sum for calorimeter trasverse energy, $E_T \text{Cone}$, around the lepton track are evaluated. For the lepton to be isolated these quantities should be a small fraction of the lepton transverse momentum.

$e - e$	$e - \mu$	$e - \tau$	$\mu - \tau$	$e - jet$	$\mu - jet$	$\tau - jet$
$\Delta R < 0.1$	$\Delta R < 0.1$	$\Delta R < 0.1$	$\Delta R < 0.1$	$\Delta R < 0.3$	$\Delta R < 0.3$	$\Delta R < 0.3$
remove e with lower p_T	remove e	remove τ	remove τ	remove jet	remove μ	remove jet

Table 4.6: Overlaps removal criteria adopted in the analysis. For each lepton the distance from the other reconstructed objects is checked.

Leptons are ordered according to the following criteria: $lep0$ is the lepton with a unique charge, $lep1$ and $lep2$ have charge opposite to $lep0$. $lep1$ is the lepton with minimum angular distance from $lep0$, $lep2$ is the remaining lepton. The angular distance between leptons is defined as follows:

$$\Delta R_{ll} = \sqrt{\Delta \eta^2 + \Delta \phi^2} \quad (4.10)$$

As explained in section 4.1, leptons from the Higgs boson decay are expected to be close to each other; this reflects in a ΔR_{01} distribution populated at small values (figure 4.5³).

Lepton flavour splitting: Events are split in two main regions: the Z-enriched region, characterized by SFOC leptons events, and the Z-depleted region, containing the remaining events. This events classification reflects the composition of our background processes. To enhance the significance, the Z-depleted region is further split in two more sub-regions, the SC region and the OC region. The SC region will then include both the DFSC and the SFSC topologies, while the OC region will include the DFOC topology. From now on we will refer to Signal Region 1 (SR1) for the Z-enriched region, Signal Region 2 (SR2) for the Z-depleted OC region and Signal Region 3 (SR3) for the Z-depleted SC region. The SR definitions are reported in table 4.7.

Number of jets: Signal events are characterized by a low jet multiplicity. The number of jets distribution for jets with transverse momentum greater than 25 GeV in $|\eta| < 4.5$ is shown in figure 4.6 for signal and background processes. Events are accepted if the number of jets (Njets) is not larger than 3.

Top veto: To reduce the top background (single top and $t\bar{t}$ production) we accept jets only if not b-tagged (figure 4.7). The b-tagging algorithm used in the analysis is called MV1; it has a nominal efficiency of 80% in identifying jets from the hadronization of b-quarks (par. 3.3.1).

³In this thesis, plots display MC contributions as stacked histograms, while the signal is shown as a non-stacked red line. The signal refers to SM expectation for a Higgs boson with $m_H = 125 \text{ GeV}$; it is also multiplied by a factor 100 to make it visibile. Only main backgrounds are drawn, other contributions go under the label "OtherBkgs". Errors are shown as a dotted line on the total SM blue line. When numerical suffixes are drawn on the x-axis label, as in plotting the $\Delta R_{n_1 n_2}$ variable, the variable is computed for leptons $lep n1$ and $lep n2$. When the data/MC ratio is shown, the yellow band represents the $\pm 1\sigma$ uncertainty due to the statistical error on MC.

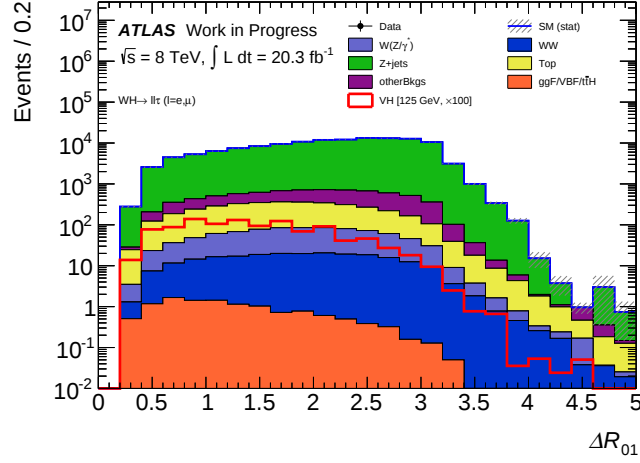


Figure 4.5: ΔR_{01} MC distribution for two charged leptons plus one tau events after the pre-selection. Expectations for Standard Model Higgs boson associated production with mass $m_H = 125 \text{ GeV}$ are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown. The lower cut at $\Delta R_{01} = 0.2$ is due to leptons OR.

Signal Region	Description
SR1	two same flavour opposite charge leptons plus a tau $(e^+e^-\tau, \mu^+\mu^-\tau)$
SR2	two opposite charge and different flavour leptons plus a tau $(e^\pm\mu^\mp\tau)$
SR3	two same charge leptons plus a tau $(e^\pm e^\pm\tau^\mp, \mu^\pm\mu^\pm\tau^\mp, e^\pm\mu^\pm\tau^\mp)$

Table 4.7: Signal Regions definition and events flavour splitting.

MET cut: Signal events include three neutrinos in the final state, which reflect in a appreciable missing transverse momentum (MET). However the presence of multiple neutrinos may imply a partial balance of their transverse momentum. A tight MET selection is applied in SR1 to reduce the Z+jets background. In this analysis two different MET definition have been used: the MET_STVF and the MET_TrackClj (chapter 3).

Z-veto: To reduce the contribution from backgrounds including a Z boson a cut on the SFOC dilepton invariant mass is applied. Events are rejected if have $|m_{01(02)} - m_Z| < 10 \text{ GeV}$.

ΔR_{01} : Leptons from the Higgs boson decay tend to be close to each other. To distinguish signal events from background events we set an upper cut on the ΔR_{01} variable.

Figure 4.8 shows the signal and background effective cross section before and after the selections listed above. Main backgrounds are $t\bar{t}$, $Z+jets$, WZ/γ^* and WW , whose contribution can be significantly reduced by the application of the selections described, as can be seen from the right part of the figure. As usual, the "OtherBkgs" category includes all the minor background processes listed in table 4.2.

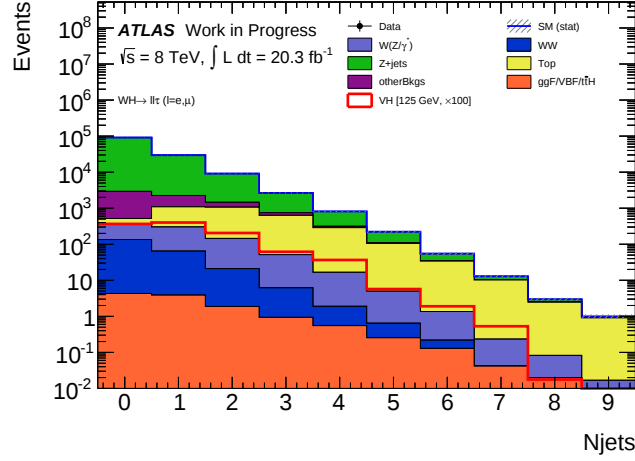


Figure 4.6: MC distribution of the jet multiplicity, N_{jets} . Expectations for Standard Model Higgs boson associated production with mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown. Only events with $N_{\text{jets}} \leq 3$ are accepted in the analysis.

4.5.2 Selection Optimization

The selections listed in the previous section have been optimized independently in each SR in order to obtain the highest significance. The optimization consists in finding the best value for each variable that maximizes the signal region significance. For each variable an interval of variation is defined (Var_{min}, Var_{max}) together with the number of steps in the interval. The optimization proceeds scanning the first variable defined, in our case the tau p_T . The number of steps for each variable has been fixed at 50. At the end of this variable scan we fix it at its best value and proceed with the second variable. When all the variables are considered, the first iteration is concluded. The whole procedure is repeated for at least three times, that are enough to obtain the best set of selections, as shown in figure 4.9. Indeed the procedure converges after 2 or 3 iterations, which means that increasing the number of iterations would not lead to a different set of best values nor to a higher significance. In figure 4.10 an example of such variable scan is shown for ΔR_{01} in SR3. If two variables are strictly correlated, as in the case of MET_STVF and MET_TrackClj, a bidimensional significance scan is performed. An example is shown in figure 4.11 for SR3. In table 4.8 the optimized variables selections in each SR are reported.

As a general rule the optimal value for each variable is chosen in a region of constant significance. A local maximum in the significance should not be taken as best value; a step in the significance distribution with respect to a cut variation can be due to a statistical fluctuation (few events with large weight that have been cut away). In the region of constant significance, the choice was to use a cut in such a way that the loss of signal events was minimized.

Once the optimization is completed, the optimized selections are applied in each SR and the results are shown in tables 4.9 and 4.10. Only main backgrounds are here explicitly reported; the breakdown of the "OtherBkgs" category can be found in Appendix C.

	Variable	Optimal value
SR1	$p_T(\tau)$	$> 15 \text{ GeV}$
	Njets	≤ 3
	MET_STVF	$> 50 \text{ GeV}$
	MET_TrackClj	$> 50 \text{ GeV}$
	Z-veto	$> 10 \text{ GeV}$
	ΔR_{01}	< 1.1
SR2	$p_T(\tau)$	$> 24 \text{ GeV}$
	Njets	≤ 3
	MET_STVF	$> 25 \text{ GeV}$
	MET_TrackClj	$> 30 \text{ GeV}$
	ΔR_{01}	< 1.4
SR3	$p_T(\tau)$	$> 15 \text{ GeV}$
	Njets	≤ 3
	MET_STVF	$> 35 \text{ GeV}$
	MET_TrackClj	$> 30 \text{ GeV}$
	ΔR_{01}	< 1.0

Table 4.8: Selection optimization result. Z-veto selection is applied only in SR1 since this is the only region with abundance of Z-like events.

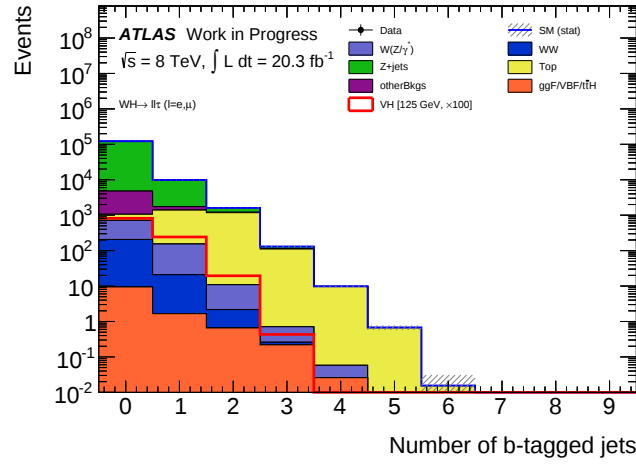


Figure 4.7: MC distribution of the b-tagged jets multiplicity. Expectations for Standard Model Higgs boson associated production with mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown. Only events with no b-tagged jets are accepted in the analysis (first bin).

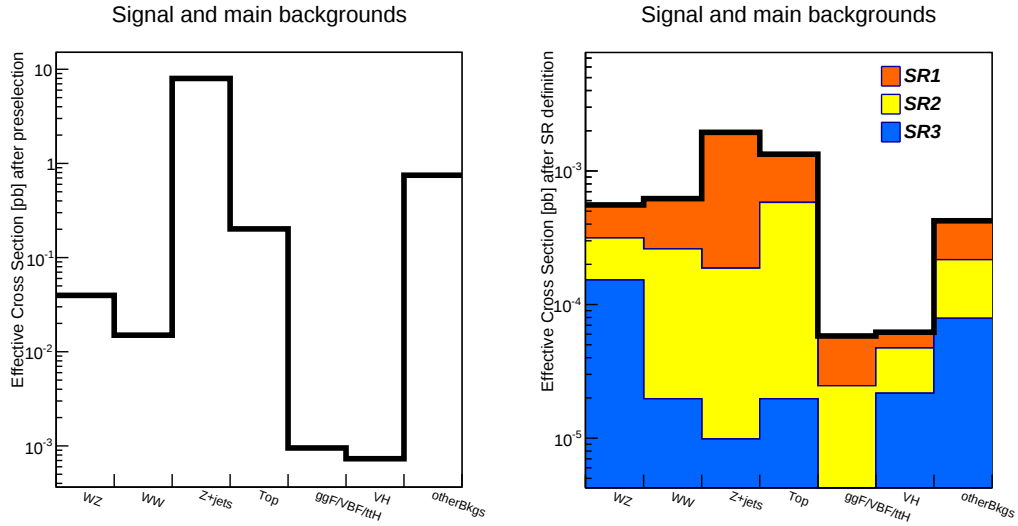


Figure 4.8: Effective cross section for signal and backgrounds after preselection (left) and after the SR definition (right). In the right figure the different contributions in each SR are shown as stacked histograms. Note that the scale on the Y-axis is different in the two plots, due to the reduction in the effective cross section obtained with the SR definition. In the right plot, the $ggF/VBF/ttH$ contribution is present in SR3 although not visible since it is smaller than the lower value on the Y axis.

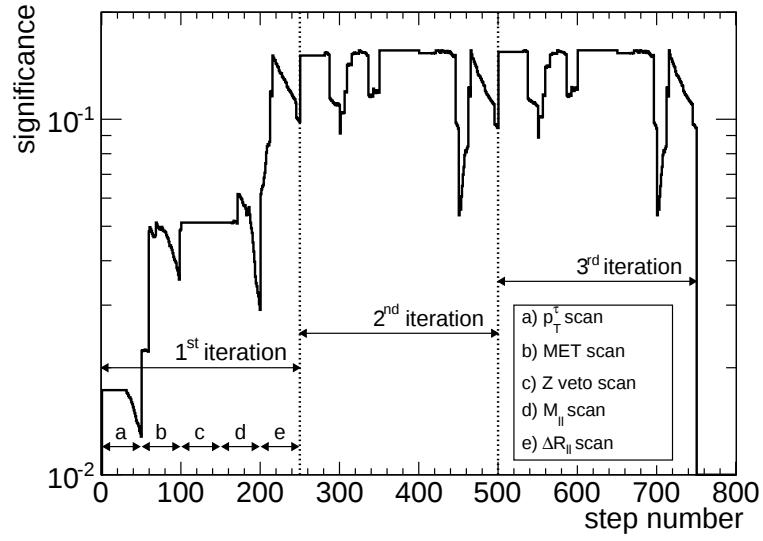


Figure 4.9: Full optimization scan in SR3. The optimization procedure scan five variables ($p_T(\tau)$, MET variables, $Z - veto$, m_{01} and ΔR_{01}), and is repeated for three times. In each iteration, the significance variation for a given variable scan is visible, and is indicated with a letter, from a) to e).

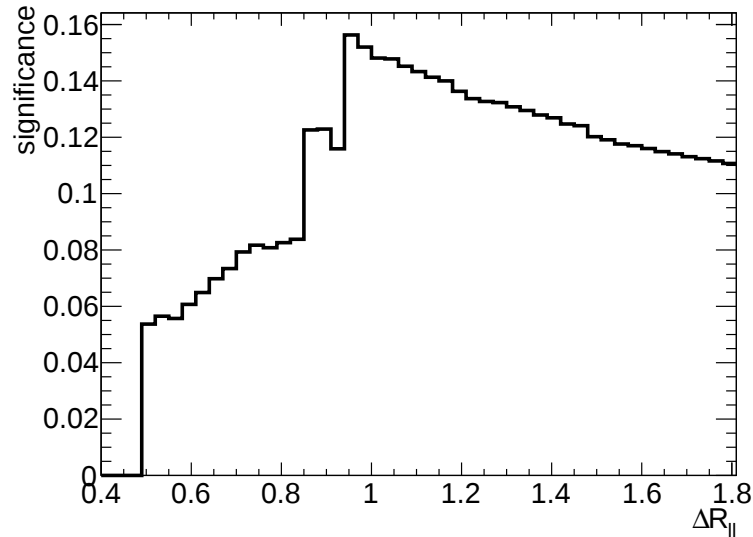


Figure 4.10: The plot shows the significance variation as a function of ΔR_{01} scanning in the interval $\Delta R \in (0.4, 1.8)$.

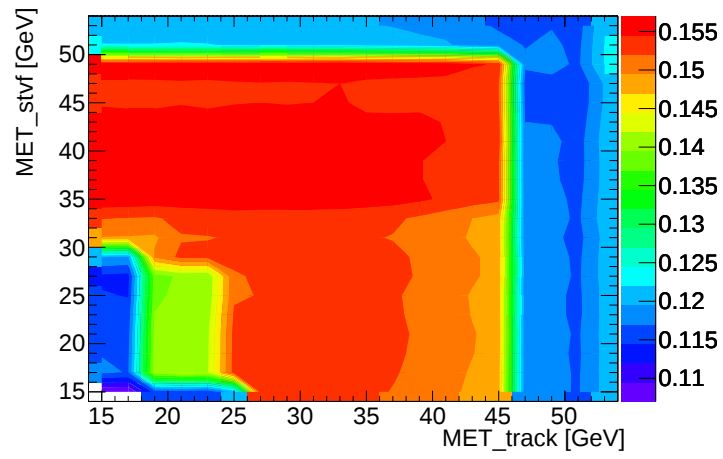


Figure 4.11: The plot shows the significance variation in SR3 while scanning on both MET_STVF and MET_TrackClj variables.

Signal Region	Selection	$WH(WW)$	$ZH(WW)$	$VH(\tau\tau)$	Total Background	$S/\sqrt{B}$
SR1	3 leptons	5.9 ± 0.5	8.1 ± 0.2	0.9 ± 0.0	$(182.5 \pm 0.7)10^3$	0.03 ± 0.00
	Preselection	4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	$(133.2 \pm 0.6)10^3$	0.03 ± 0.00
	$p_T(\tau)$	4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	$(133.2 \pm 0.6)10^3$	0.03 ± 0.00
	Flavour splitting	1.9 ± 0.3	4.6 ± 0.2	0.3 ± 0.0	$(131.1 \pm 0.6)10^3$	0.02 ± 0.00
	Njets	1.7 ± 0.3	4.4 ± 0.2	0.3 ± 0.0	$(130.2 \pm 0.6)10^3$	0.02 ± 0.00
	Top veto	1.4 ± 0.3	3.1 ± 0.1	0.2 ± 0.0	$(119.7 \pm 0.6)10^3$	0.01 ± 0.00
	Z-veto	1.1 ± 0.2	$(6.8 \pm 0.6)10^{-1}$	0.2 ± 0.0	$(19.0 \pm 0.2)10^3$	0.01 ± 0.00
	MET	0.3 ± 0.0	$(1.6 \pm 0.3)10^{-1}$	0.0 ± 0.0	246 $\pm$ 15	0.03 ± 0.00
	ΔR_{01}	0.2 ± 0.0	$(8 \pm 2)10^{-2}$	0.0 ± 0.0	68 $\pm$ 7	0.04 ± 0.00
SR2	$p_T(\tau)$	2.0 ± 0.2	3.1 ± 0.1	0.3 ± 0.0	$(36.9 \pm 0.3)10^3$	0.03 ± 0.00
	Flavour splitting	$(6.4 \pm 0.8)10^{-1}$	$(4.3 \pm 0.5)10^{-1}$	0.1 ± 0.0	635 $\pm$ 7	0.05 ± 0.00
	Njets	$(6.4 \pm 0.8)10^{-1}$	$(4.1 \pm 0.5)10^{-1}$	0.1 ± 0.0	574 $\pm$ 7	0.05 ± 0.01
	Top veto	$(5.1 \pm 0.8)10^{-1}$	$(2.9 \pm 0.4)10^{-1}$	0.1 ± 0.0	166 $\pm$ 5	0.07 ± 0.01
	MET	$(4.2 \pm 0.8)10^{-1}$	$(2.0 \pm 0.3)10^{-1}$	0.0 ± 0.0	81 $\pm$ 3	0.07 ± 0.01
	ΔR_{01}	$(3.6 \pm 0.8)10^{-1}$	$(1.3 \pm 0.3)10^{-1}$	0.0 ± 0.0	27 $\pm$ 1	0.10 ± 0.02
SR3	$p_T(\tau)$	4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	$(133.2 \pm 0.6)10^3$	0.03 ± 0.00
	Flavour splitting	1.4 ± 0.3	$(5.6 \pm 0.5)10^{-1}$	0.0 ± 0.0	$(2.7 \pm 0.2)10^2$	0.12 ± 0.02
	Njets	1.4 ± 0.3	$(5.6 \pm 0.5)10^{-1}$	0.0 ± 0.0	$(2.7 \pm 0.2)10^2$	0.12 ± 0.02
	Top veto	1.1 ± 0.2	$(4.2 \pm 0.5)10^{-1}$	0.0 ± 0.0	$(2.3 \pm 0.2)10^2$	0.10 ± 0.02
	MET	0.7 ± 0.2	$(2.4 \pm 0.3)10^{-1}$	0.0 ± 0.0	45 $\pm$ 4	0.14 ± 0.03
	ΔR_{01}	0.4 ± 0.1	$(9 \pm 2)10^{-2}$	0.0 ± 0.0	5.6 $\pm$ 0.5	0.18 ± 0.05

Table 4.9: Signal regions cut-flow obtained with MC signal and background samples. Each row shows the number of signal events and total background events after the selection is applied; the significance is also evaluated. The 5th column shows the analysis acceptance for the process $VH \rightarrow \tau\tau$. The number of events is reported with its statistical uncertainty. Except for the column showing the significance, numbers smaller than 0.1 are rounded as 0 in this table.

Signal Region	Selection	T_{op}	$Z + jets$	WZ/γ^*	WW	$ggF/VBF/t\bar{t}H$	OtherBkgs
SR1	3 leptons	4093 $\pm$ 8	(162.0 $\pm$ 0.7)10 ³	806 $\pm$ 4	304 $\pm$ 3	19.3 $\pm$ 0.4	(15.2 $\pm$ 0.2)10 ³
	Preselection	2866 $\pm$ 6	(125.4 $\pm$ 0.6)10 ³	651 $\pm$ 3	218 $\pm$ 2	12.0 $\pm$ 0.3	(41.1 $\pm$ 0.2)10 ²
	$p_T(\tau)$	2866 $\pm$ 6	(125.4 $\pm$ 0.6)10 ³	651 $\pm$ 3	218 $\pm$ 2	12.0 $\pm$ 0.3	(41.1 $\pm$ 0.2)10 ²
	Flavour splitting	1432 $\pm$ 5	(124.9 $\pm$ 0.6)10 ³	572 $\pm$ 3	109 $\pm$ 2	6.2 $\pm$ 0.2	(40.3 $\pm$ 0.2)10 ²
	Njets	1432 $\pm$ 5	(124.3 $\pm$ 0.6)10 ³	553 $\pm$ 3	100 $\pm$ 0	5.8 $\pm$ 0.2	(40.0 $\pm$ 0.2)10 ²
	Top veto	150 $\pm$ 2	(115.4 $\pm$ 0.6)10 ³	423 $\pm$ 3	98 $\pm$ 2	4.8 $\pm$ 0.2	(36.0 $\pm$ 0.2)10 ²
	Z-veto	127 $\pm$ 1	(16.3 $\pm$ 0.2)10 ³	74 $\pm$ 1	83 $\pm$ 1	4.2 $\pm$ 0.2	(24.2 $\pm$ 0.2)10 ²
	MET	72 $\pm$ 1	(1.1 $\pm$ 0.1)10 ²	16.1 $\pm$ 0.4	29.5 $\pm$ 0.8	1.1 $\pm$ 0.1	18 $\pm$ 2
	ΔR_{01}	15.3 $\pm$ 0.5	36 $\pm$ 7	4.9 $\pm$ 0.3	7.3 $\pm$ 0.4	(68 $\pm$ 8)10 ⁻²	4.2 $\pm$ 1.0
SR2	$p_T(\tau)$	984 $\pm$ 4	(34.1 $\pm$ 0.3)10 ³	354 $\pm$ 2	53 $\pm$ 1	3.7 $\pm$ 0.2	(10.6 $\pm$ 0.1)10 ²
	Flavour splitting	484 $\pm$ 3	93 $\pm$ 6	16.1 $\pm$ 0.4	24.4 $\pm$ 0.8	1.5 $\pm$ 0.1	16 $\pm$ 2
	Njets	425 $\pm$ 3	92 $\pm$ 6	15.8 $\pm$ 0.4	24.1 $\pm$ 0.7	1.3 $\pm$ 0.1	16 $\pm$ 2
	Top veto	41.3 $\pm$ 0.8	76 $\pm$ 4	12.7 $\pm$ 0.4	21.0 $\pm$ 0.7	1.0 $\pm$ 0.1	14 $\pm$ 2
	MET	33.3 $\pm$ 0.7	15 $\pm$ 2	8.5 $\pm$ 0.3	15.3 $\pm$ 0.6	0.8 $\pm$ 0.1	8 $\pm$ 2
	ΔR_{01}	11.4 $\pm$ 0.4	3.6 $\pm$ 0.7	3.3 $\pm$ 0.3	4.9 $\pm$ 0.3	0.5 $\pm$ 0.1	2.8 $\pm$ 0.6
SR3	$p_T(\tau)$	2866 $\pm$ 6	(125.4 $\pm$ 0.6)10 ³	651 $\pm$ 3	218 $\pm$ 2	12.0 $\pm$ 0.3	(41.1 $\pm$ 0.2)10 ²
	Flavour splitting	18.4 $\pm$ 0.6	(1.6 $\pm$ 0.2)10 ²	49.3 $\pm$ 0.8	3.8 $\pm$ 0.2	(5.0 $\pm$ 0.2)10 ⁻²	38 $\pm$ 3
	Njets	15.7 $\pm$ 0.5	(1.6 $\pm$ 0.2)10 ²	48.4 $\pm$ 0.8	3.5 $\pm$ 0.2	(2.8 $\pm$ 0.2)10 ⁻¹	38 $\pm$ 3
	Top veto	3.7 $\pm$ 0.3	(1.5 $\pm$ 0.2)10 ²	38.9 $\pm$ 0.7	2.7 $\pm$ 0.2	(1.0 $\pm$ 0.1)10 ⁻¹	34 $\pm$ 3
	MET	2.7 $\pm$ 0.2	7 $\pm$ 4	21.2 $\pm$ 0.4	1.9 $\pm$ 0.2	(3 $\pm$ 1)10 ⁻²	12 $\pm$ 2
	ΔR_{01}	0.4 $\pm$ 0.1	0.2 $\pm$ 0.2	3.1 $\pm$ 0.2	0.4 $\pm$ 0.1	(1 $\pm$ 0)10 ⁻²	1.6 $\pm$ 0.4

Table 4.10: Signal regions cut-flow obtained with MC background samples. At each stage of the cut-flow the number of background events for the main background processes is reported. The "Other Bkgs" category includes all the minor processes listed in table 4.4 which contribution is not explicitly reported in the previous columns. The number of events is reported with its statistical uncertainty. Numbers smaller than 0.1 are rounded as 0 in this table.

4.5.3 Blind criteria

In this thesis a blind data analysis technique was adopted. A blind analysis consists in the temporarily exclusion, from the available data sample, of all those events that fall into the defined SR. This technique is commonly used in high energy physics whenever small signals have to be measured; in fact hiding the SR is an optimal way to reduce or eliminate the experimenter bias. In a cut-based analysis the bias can origin from a fine tuning of the selections. If these selections are applied to the SR knowing their effect on data the danger is that the cuts are chosen to remove few backgrounds events while increasing the significance.

Several blind analysis techniques have been used in particle physics in recent years. We decided to use the *hidden signal box* technique [69] which consists in hiding the SR until the analysis is completed. In this way the consistency of the measurement and the stability of the selections applied are verified using MC samples and only data events outside the defined SR. Once the analysis is completed and everything is under the experimenter control the box is opened and the result is obtained.

The blind region was defined starting from the SR definition in section 4.5.1 and slightly relaxing the cuts by looking at how the significance varies around the optimal values in table 4.8. The hidden box is the multidimensional region defined in table 4.11.

SR	Blinding criteria
Blind SR1	$N_{\text{jets}} \leq 3$ Top veto Z-veto $\text{MET_STVF}(\text{TrackClj}) > 50 \text{ GeV}$ $\Delta R_{01} < 1.2$
Blind SR2	$N_{\text{jets}} \leq 3$ Top veto $\text{MET_STVF}(\text{TrackClj}) > 9 \text{ GeV}$ $\Delta R_{01} < 1.6$
Blind SR3	$N_{\text{jets}} \leq 3$ Top veto $\text{MET_STVF}(\text{TrackClj}) > 15 \text{ GeV}$ $\Delta R_{01} < 1.6$

Table 4.11: Blinded region definition. For each SR all the events satisfying the listed selection are excluded from the blinded analysis.

4.6 Data-MC comparison

In this section a method used to verify the modelling of the MC samples for the background is described. To be trustworthy, MC predictions should include the best knowledge of physics processes and detector response. However, even if these conditions are satisfied, it can happen that the theoretical calculation and the simulation of the detector response, on which MC is based, do not have the desired accuracy to reproduce data. This could lead to an incorrect representation of what is found in the data. It is then extremely important to evaluate MC samples reliability, especially when the number of expected signal events is very low.

The scope of this section is to compare MC predictions with data and eventually to evaluate a normalization factor (NF) that can take into account any disagreement. In this analysis, background properties are estimated from the so-called *Control Regions* (CR). These are kinematical regions different from our Signal Regions so that what we want to measure does not lie in these regions. In principle CR should only contain the known physics processes. In the next sections we describe our CR definitions and the NF evaluation for the main backgrounds.

4.6.1 Control regions

Backgrounds contributing to final states with three isolated leptons are dominated by $t\bar{t}$, $Z+jets$ WZ/γ^* and WW production. In order to verify the modelling (normalization and shape description) of these components, four CR have been defined with selections aimed at ensuring high purity for the background under study and orthogonality between the CR and SR. This means that no events present at the last selection of the SR must be present in the CR. The orthogonality between CR and SR is twofold: first it allows to compare data with MC, not possible in SR due to the blinding condition, second it avoids calculation problems that could originate when using the same event twice in a fit. CR for the above backgrounds have been defined starting from the SR definition by changing just one analysis cut. This allows to define CR adjacent to SR, where we are confident that the shape of the various kinematical distributions are quite similar. In principle the modelling could indeed be pretty different in a phase space region disjoint from the SR and so the error made in extrapolating and applying the NF to the SR could be significant. For this reason CR and SR share most of the selections.

In table 4.12 CR definition for main backgrounds is reported. The WW CR is missing. In fact any attempt to define a CR with high statistics failed. For this reason we do not apply any NF to the WW background (i.e. $NF(WW) = 1$).

Top CR

Top CR has been defined with the same criteria adopted in SR2 except for the inversion of the Top veto selection. Events are selected requiring at least one $p_T > 20$ GeV jet satisfying the b-tagging selection. Indeed we expect one b-jet in each top decay. Figure 4.12 shows the b-jet multiplicity for events in SR2, before applying the request of at least one b-jet. Reverting the Top veto not only enhances the top background in the CR but also allows to reduce other backgrounds as $Z+jets$ and WZ/γ^* making Top CR 97% pure in top events.

Z+jets CR

$Z+jets$ CR is defined starting from SR1, the *Z-enriched* region, since is in this region that we have the abundance of Z-like backgrounds, due to the presence of SFOC leptons. $Z+jets$ events are selected with the same criteria of events in SR1 with the exception of the MET cut. Referring to the cut-flow in table 4.10 we observe that is the tight MET selection that mostly reduces the $Z+jets$ contribution in SR1, even more than the Z-veto cut. Indeed the Z-veto cut reduce $Z+jets$

CR	Defined in	Selection	Reverted cut with respect to SR
Top CR	SR2	jet multiplicity MET cut ΔR_{01}	Top veto
Z+jets CR	SR1	jet multiplicity Top veto Z-veto ΔR_{01}	MET cut
WZ/γ^* CR1	SR1	jet multiplicity Top veto MET cut ΔR_{01}	Z-veto
WZ/γ^* CR2	SR3	jet multiplicity Top veto MET cut	—

Table 4.12: Control Regions definition.

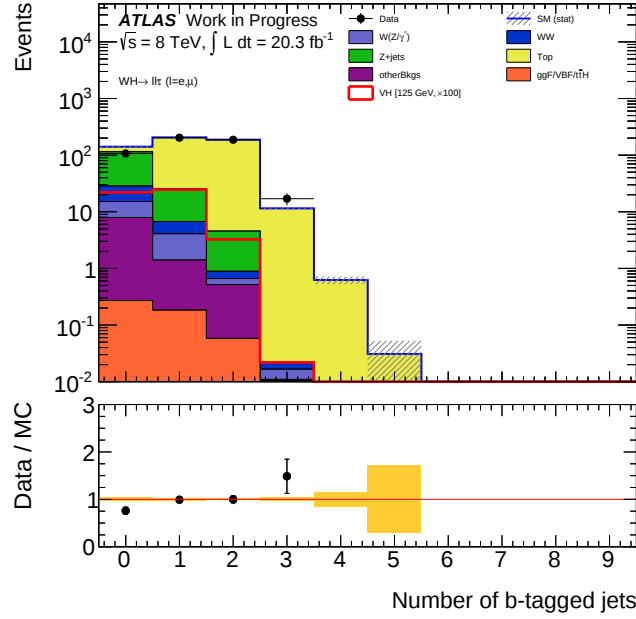


Figure 4.12: Number of b-tagged jets distribution in Top CR before requiring at least one b-tagged jet. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

background of $\approx 85\%$, while MET cut allows a reduction of $\approx 95\%$. It was then natural to define Z+jets CR starting from the SR1 definition and reverting the MET cut, from $\text{MET} > 50$ GeV to $\text{MET} < 50$ GeV. In fact no significant MET is expected in Z+jets events. In figure 4.13 the MET_STVF and MET_TrackClj distributions are shown, before the inverted MET selection.

WZ/γ^* CR

For WZ/γ^* CR two different definitions could work. WZ/γ^* is indeed a not negligible fraction of the total background both in SR1 and in SR3 (cfr table 4.10). The only difference is that SR1 is highly contaminated by other background sources like Z+jets and $t\bar{t}$ and could then be difficult to select a pure WZ/γ^* CR. On the contrary in SR3 WZ/γ^* is the main background, see cut-flow in table 4.10, corresponding to $\approx 60\%$ of the whole background in this CR; this feature should help in defining WZ/γ^* CR as pure as possible. On the other hand, in SR3 the overall WZ/γ^* statistics is smaller and the calculation at the NF could be less accurate. We then tested both these two WZ/γ^* CR, named respectively CR1 and CR2.

WZ/γ^* CR1 has been defined with the same criteria adopted in SR1 apart from the Z-veto selection that has been inverted, by asking for leptons to have an invariant mass $|m_{01(02)} - m_Z| < 10$ GeV. In figure 4.14 the invariant mass of opposite charge leptons, m_{ll} , is shown before the application of the inverted Z-veto cut.

WZ/γ^* CR2 has been defined with the same criteria adopted for SR3 definition. By using the SR3 selections WZ/γ^* CR2 is orthogonal to all the others CR already defined, since it is the only CR to select same charge leptons. Unfortunately this definition doesn't grant the orthogonality against SR3, by definition. We are of course aware of this and we know that this could be a non-negligible problem when the blinding of the SR will be released. By the

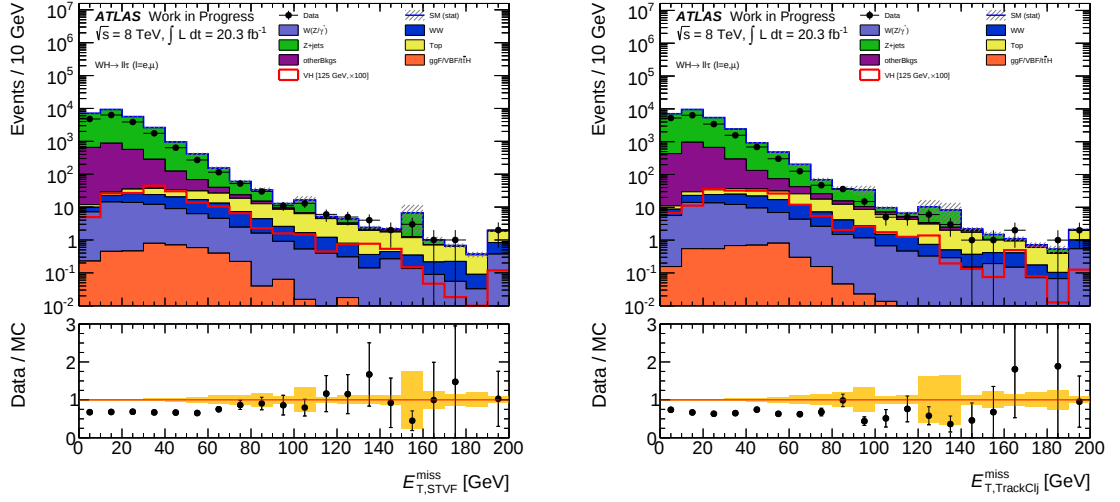


Figure 4.13: Distributions of (a) MET_STVF and (b) MET_TrackClj in Z+jets CR before cutting on the MET variables. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

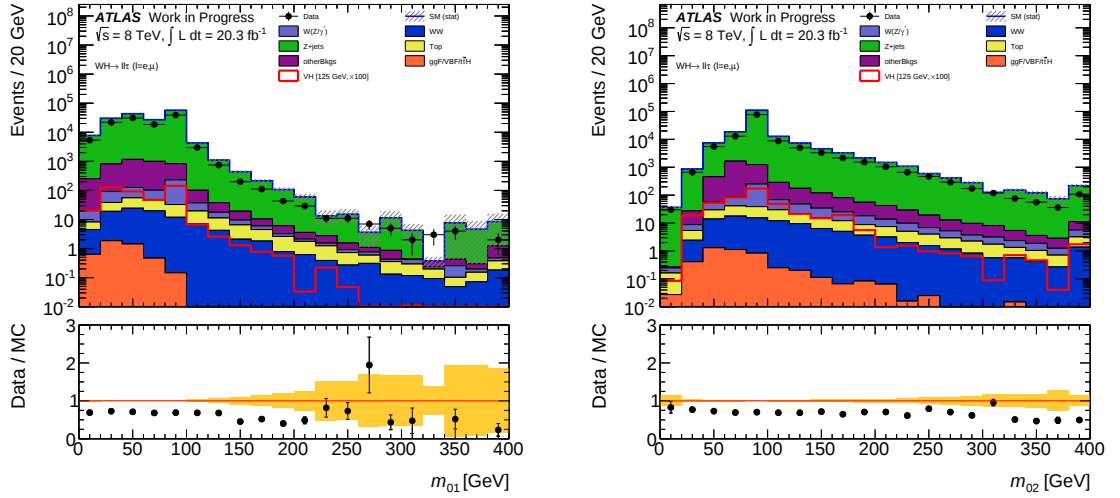


Figure 4.14: Distributions of (a) m_{01} and (b) m_{02} in WZ/γ^* CR1 before the application of the inverted Z-veto cut. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown. Both variables show clearly the Z boson mass peak.

way, compared with Z+jets and $t\bar{t}$ backgrounds, WZ/γ^* has low statistics that will be further reduced if we add a new selection to the CR. To avoid the statistics reduction we decided not to care for the moment of the overlap between WZ/γ^* CR2 and SR3 while to carry on with the evaluation of the NF. Of course we will be back on this issue later. In figure 4.15 the three leptons invariant mass distribution in WZ/γ^* CR2 after the Top veto selection is shown.

In table 4.13 the MC statistics breakdown in each control region together with the purity

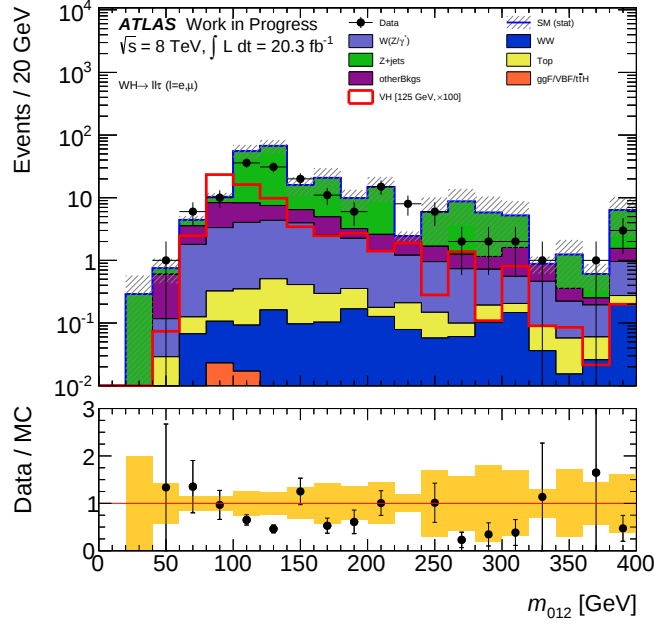


Figure 4.15: Three leptons invariant mass in WZ/γ^* CR2 after the Top veto selection. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

obtained in the main background process is reported. The table also shows the data/MC ratio without any NF applied.

CR	Data/MC	Total background	Top	Z+jets	WZ/γ^*	Purity
Top	0.94 ± 0.09	107.6 ± 1.4	104.3 ± 1.2	1.1 ± 0.5	0.9 ± 0.1	0.97 ± 0.00
Z+jets	0.67 ± 0.02	$(5.7 \pm 0.1)10^3$	12.2 ± 0.4	$(5.2 \pm 0.1)10^3$	13.6 ± 0.5	0.92 ± 0.02
WZ/γ^* CR1	0.75 ± 0.14	$(1.4 \pm 0.2)10^2$	1.7 ± 0.2	$(1.2 \pm 0.2)10^2$	15.4 ± 0.4	0.11 ± 0.02
WZ/γ^* CR2	0.82 ± 0.19	$(0.5 \pm 0.1)10^2$	1.7 ± 0.2	20 ± 8	16.2 ± 0.4	0.33 ± 0.06

Table 4.13: Control regions: main backgrounds events breakdown and purity. The "Total background" column includes the contribution of all the non-leading backgrounds.

The last column is the purity of each CR defined as:

$$\text{Purity} = \frac{\text{Number of events for the leading background}}{\text{Total background events}} \quad (4.11)$$

where the "leading background" is the background process for which the NF has to be evaluated from the given CR. As shown in table 4.13 both Top CR and Z+jets CR have a high statistics and a high purity, while for WZ/γ^* CR things are more difficult. In the case of WZ/γ^* CR1 indeed the Z-veto cut inversion not only allows the selection of WZ/γ^* events, but also the selection of Z+jets, lowering the purity of the region. For WZ/γ^* CR2 instead the purity is higher with respect to CR1 but the statistics is quite low. The data-MC agreement, with the

exception of Top CR, is far away from being equal to one. This confirms the assumption at the beginning of this chapter, that MC samples have to be partially corrected to reproduce data. We will discuss again the data-MC agreement after having evaluated and applied the NF. The complete CR cut-flow can be found in appendix C .

4.6.2 NF evaluation

This paragraph describes the method used for the NF evaluation. The NF is defined as the factor for which the given MC process has to be scaled in order to correctly describe data. Ideally the NF definition is the following:

$$\text{NF} = \frac{\text{Number of simulated events in CR}}{\text{Number of data events in CR}} \quad (4.12)$$

The equation (4.12) would work only with 100% purity CR, which is never our case. Indeed each CR defined in the previous paragraph is contaminated by several background processes and all of them should be taken into account when considering the data/MC agreement. It is worth noting that the NF as described in equation (4.12) acts as a scale factor for a given background process, which means that it cannot correct the simulation for any mismodelling in the variables distribution. The NF does only correct the integral of the given distribution to reproduce what observed in data.

To evaluate a NF for each of the background processes a simultaneous fit to all the CR has been set up. NF are computed by a χ^2 minimization built with the difference between data and MC events in each CR:

$$\chi^2 = \sum_{\text{all CR, all samples}} \frac{(N_{\text{Data}} - N_{\text{MC}})^2}{\sigma^2} \quad (4.13)$$

where the sum runs over all the CR and all the MC background processes, and

$$N_{\text{MC}} = N_{\text{Top}} \cdot \text{NF}_{\text{Top}} + N_{\text{Z+jets}} \cdot \text{NF}_{\text{Z+jets}} + N_{\text{WZ}/\gamma^*} \cdot \text{NF}_{\text{WZ}/\gamma^*} + N_{\text{othersBkgs}} \quad (4.14)$$

$$\sigma = \sqrt{(\sigma_{\text{Data}})^2 + (\sigma_{\text{MC}})^2} \quad (4.15)$$

In this way the NF for a given background is evaluated considering the background contribution in each CR defined. Results for Top, Z+jets and WZ/γ^* samples are shown in table 4.14. For the WZ/γ^* sample two different values are quoted, depending on which WZ/γ^* CR is included in the χ^2 formula. Indeed the sum in equation (4.13) runs over three CR at a time, and just one out of two WZ/γ^* CR is considered. Top and Z+jets NF are stable and they do not

	NF
Top	1.02 ± 0.06
Z+jets	0.68 ± 0.01
WZ/γ^* CR1	2.07 ± 0.43
WZ/γ^* CR2	1.05 ± 0.40

Table 4.14: The table shows NF for the main analysis backgrounds as obtained with the chi-square minimization in equation (4.13).

depend on which WZ/γ^* CR is used in the χ^2 . The stability of Top and Z+jets NF is granted

by the high statistics and purity of the relative CR. Indeed with high statistics a better precision on the NF is obtained, which means that statistical fluctuations of few events in the CR cannot distort the fit result. On the contrary for WZ/γ^* NF two different values with a relative large uncertainty are obtained. The problem in this case is due to the low statistics that we have in both CR1 and CR2, which then turn out to be more sensitive to statistical fluctuations. Moreover different selections are applied in CR1 with respect to CR2, each of them affecting the shape of the distribution we fit. The total number of events discarded by the application of a given selection could be quite different in the MC sample with respect to the data sample, and this happens every time the simulation does not reproduce the shape of a variable distribution. Once more shape mismodelling becomes a real issue when the statistics is low, as in WZ/γ^* CR1 and CR2.

To solve this problem we implemented a simultaneous *shape fit*, i.e. a fit that takes into account the shape of the fitted distribution. The formula in equation (4.13) is modified in order to compare data and MC events in each bin of a given variable distribution:

$$\chi^2 = \sum_{\text{all CR, all samples, all bins}} \frac{(N_{\text{Data}} - N_{\text{MC}})^2}{\sigma^2} \quad (4.16)$$

This fit methodology consists again in finding a unique factor to normalize the MC distribution to data, but this time the NF is such to minimize the bin per bin differences in data-MC number of events. Each bin in the chi-square formula acts as a single CR. The shape fit and the previous fit on the integral of a distribution are supposed to give the same results in case no shape mismodelling is observed in MC. As a general rule the fit should never be performed on a variable that has already been used for a CR definition. By applying a selection on a variable the corresponding effect on data and MC distribution could be very different. For this reason we choose to test the *shape fit* on the three leptons invariant mass m_{lll} variable. m_{lll} distribution is shown in figure 4.16 (a) for WZ/γ^* CR1 and in figure 4.16 (b) for WZ/γ^* CR2. Table 4.15 shows the NF derived with the *shape fit*.

	m_{lll}
Top	1.02 ± 0.06
Z+jets	0.68 ± 0.01
WZ/γ^* CR1	1.33 ± 0.45
WZ/γ^* CR2	3.17 ± 0.41

Table 4.15: The table shows the NF for the main backgrounds as obtained with the shape fit described in equation (4.16).

Again Top and Z+jets NF are stable, while for WZ/γ^* NF it is not possible to get a single and stable value. Both values in table 4.15 are potentially acceptable, even if the mean value is quite high. There is no reason to prefer either of them. This means that the *shape fit* does not solve the problem with the WZ/γ^* sample.

4.6.3 Data/MC agreement in CR

NF as reported in table 4.14 are then applied to each CR for the normalization of both Z+jets and Top samples; the data/MC agreement is then verified by comparing some variable distributions. As explained in the previous paragraph, due to the instability of WZ/γ^* NF, we decided to not apply any NF to the WZ/γ^* background (i.e. $\text{NF}(WZ/\gamma^*) = 1$). Results are shown in

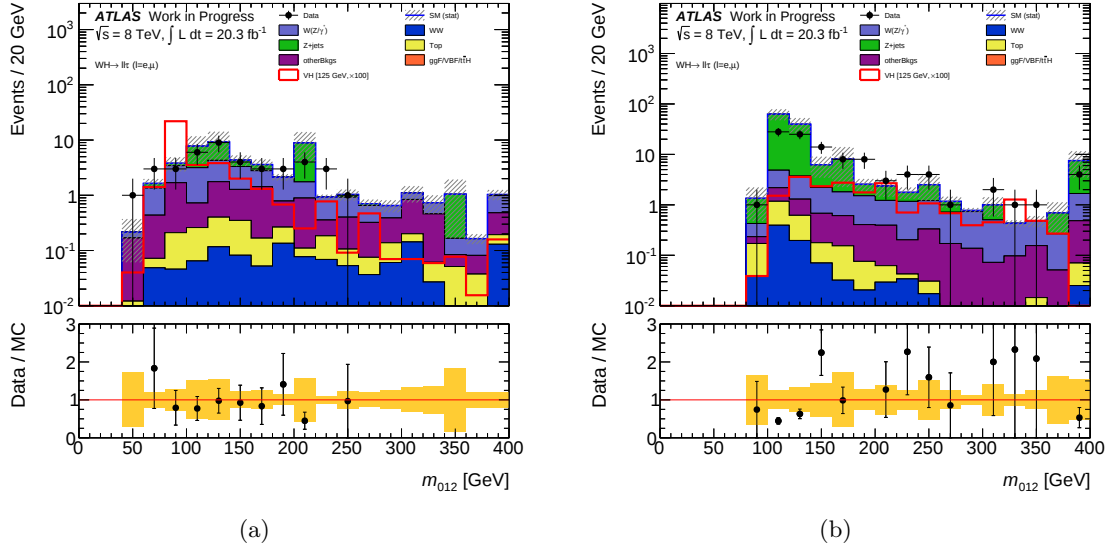


Figure 4.16: Distributions of normalized m_{ll} in (a) WZ/γ^* CR1 and (b) WZ/γ^* CR2. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

figures 4.17-4.18. In figure 4.17 the MET distributions in Top CR are shown. In figure 4.18 the ΔR_{01} distribution is shown for the Z+jets CR.

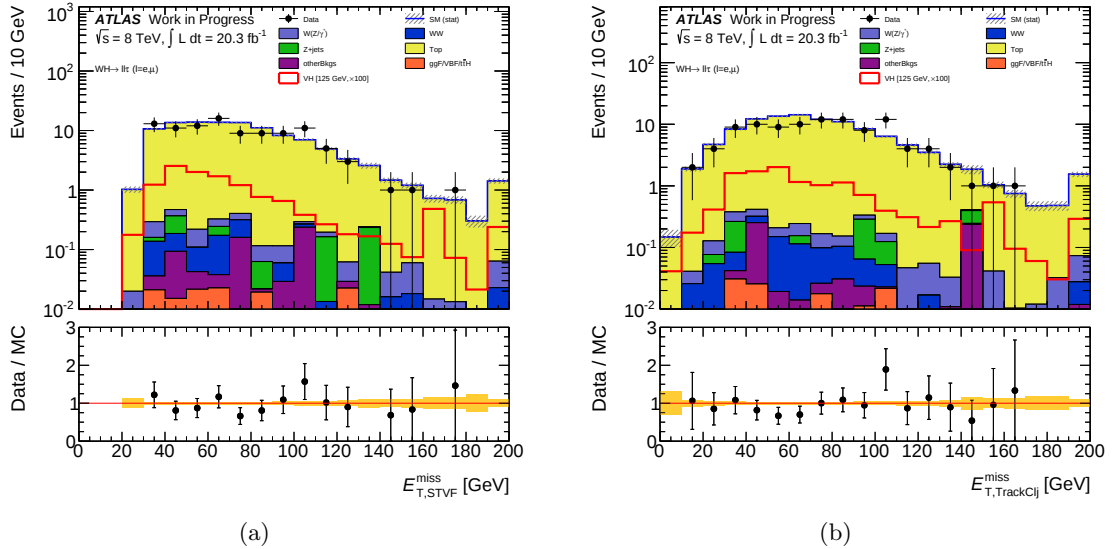


Figure 4.17: Distributions of normalized (a) MET_STVF and (b) MET_TrackClj in Top CR. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

Except for the distributions in the Z+jets CR that are characterized by a high statistics, all the other distributions show some statistical fluctuations. In both WZ/γ^* CR1 and WZ/γ^* CR2

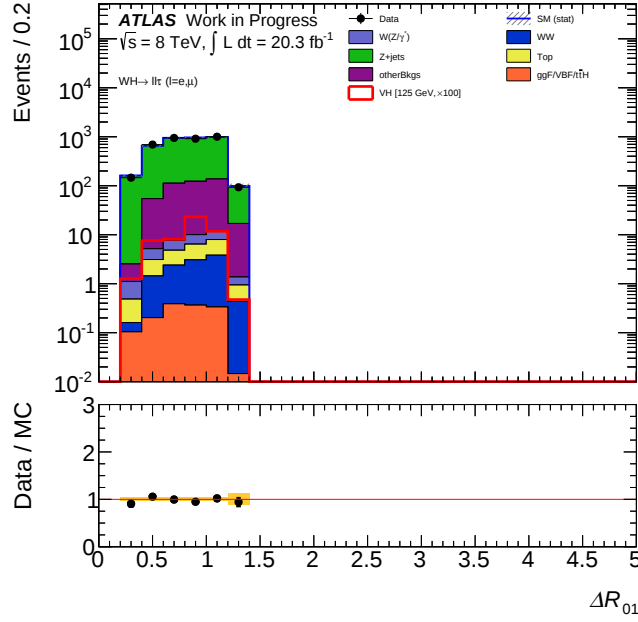


Figure 4.18: Distribution of normalized ΔR_{01} in Z+jets CR. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown. The distribution has an upper limit equal to 1.4 due to the blinding criteria in SR1.

the data and MC statistics is quite low and the distributions are dominated by the statistical fluctuations (figure 4.16). Generally speaking, the data/MC ratio is around one and inside the yellow error band in all the CR considered; MC samples are adequate to reproduce the shape of the distributions.

4.6.4 Data/MC agreement in blinded SR

Once NF have been evaluated we want to apply them to the SR. Due to the blinding in the SR the comparison at the end of the cut-flow is not possible. At any rate a preliminary useful information can be derived by comparing the data/MC ratio, together with some distributions, after the first selections in each SR. Checking data/MC agreement in the blinded SR is mandatory before proceeding with the unblinding, since any incostincency could be the proof of something unclear. Let's recall that CR have been used in the fit to evaluate NF, the (good) agreement we found in the previous paragraph is then expected. SR are instead kinematical regions disjoint from CR, in principle the modelling could be quite different and it has to be checked. In fact, by excluding all those events falling in the SR, we could end to define the CR with few left events, that fall in the tails of variable distributions. Due to the lack of statistics in the tail of the distribution the shape for that variable could not be well modelled by MC.

Plots showing data/MC agreement in blinded SR at different cut-flow levels are shown in figures 4.19-4.21, while table 4.16 reports the SR cut-flow before the MET selection. A good agreement is evident in all the regions, with a data/MC ratio in general compatible with 1, within the statistical uncertainty. Due to the lack of statistics in both SR2 and SR3 some local (bin-per-bin) disagreements are visible, which could not be recovered by the application of a different NF to the leading backgrounds. For this reason a different analysis approach, using

a multivariate technique, has been carried on and is reported in the next chapter. Chapter 6 compares the two methods results.

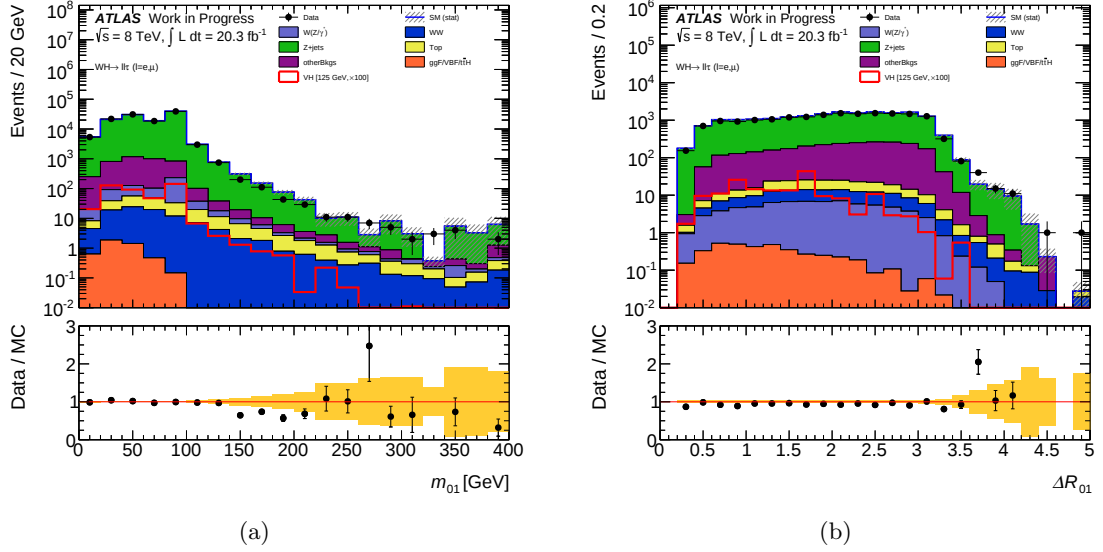


Figure 4.19: MC and data distributions in blinded SR1: (a) the m_{01} dilepton invariant mass after the Top veto, (b) the ΔR_{01} distribution after the Z-veto cut. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

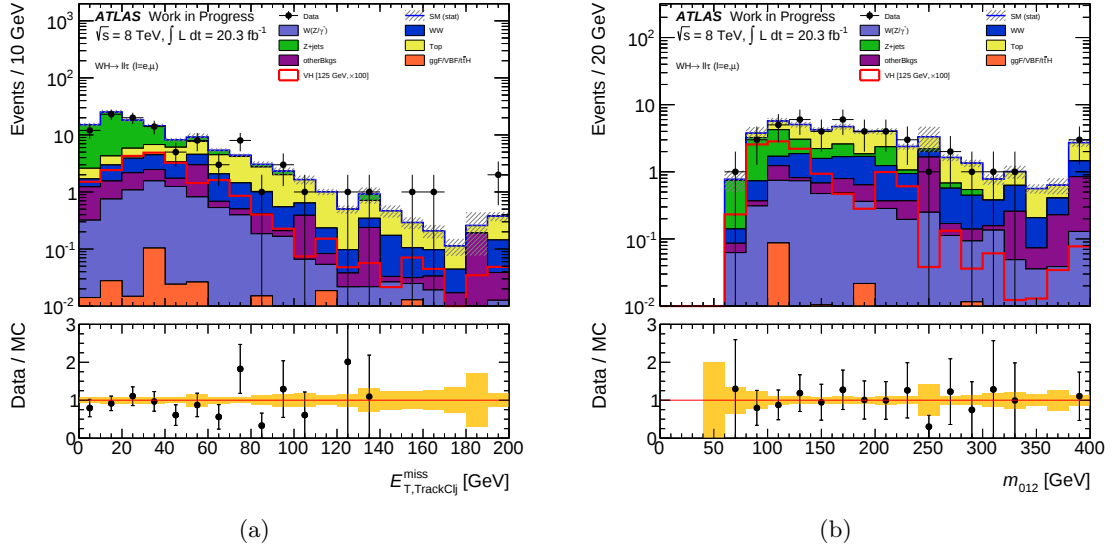


Figure 4.20: MC and data distributions in blinded SR2: (a) the $MET_TrackClj$ distribution after the Top veto, (b) the m_{012} 3-leptons invariant mass after the MET cut. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

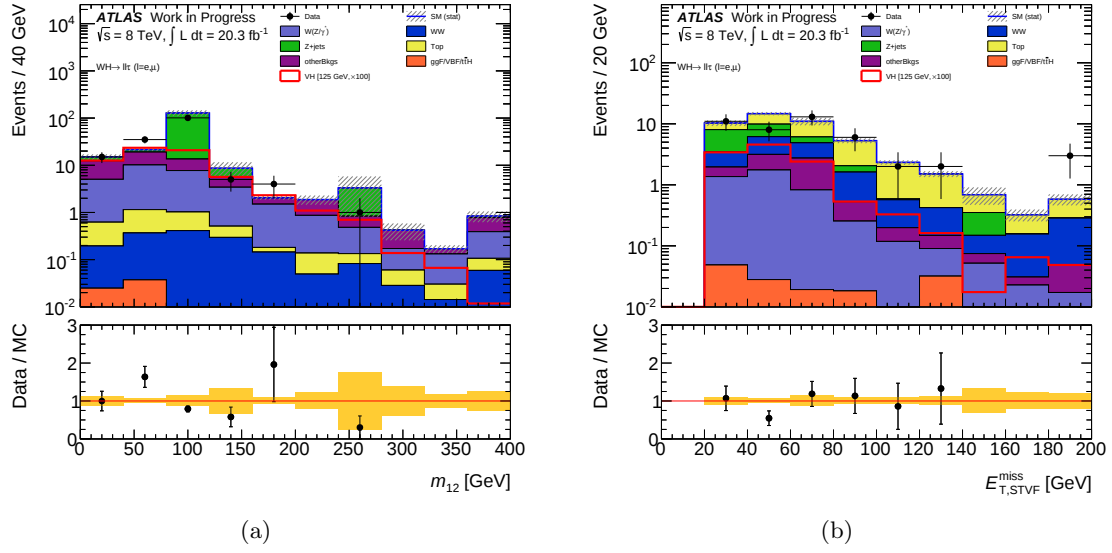


Figure 4.21: MC and data distributions in blinded SR3: (a) the M_{12} same-charge dilepton invariant mass after the Top veto, (b) the MET_STVF distribution after the MET cut. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

Signal Region	Selection	$WH(WW)$	$ZH(WW)$	Observed	Total Background	Data/MC
SR1	3 leptons	4.4 ± 0.5	7.4 ± 0.2	187972	$(182.2 \pm 0.7)10^3$	1.03 ± 0.00
	Preselection	2.9 ± 0.4	5.2 ± 0.2	136945	$(133.0 \pm 0.6)10^3$	1.03 ± 0.01
	$p_T(\tau)$	2.9 ± 0.4	5.2 ± 0.2	136945	$(133.0 \pm 0.6)10^3$	1.03 ± 0.01
	Flavour splitting	1.7 ± 0.3	4.5 ± 0.2	135033	$(131.1 \pm 0.6)10^3$	1.03 ± 0.01
	Njets	1.5 ± 0.3	4.3 ± 0.2	134061	$(130.2 \pm 0.6)10^3$	1.03 ± 0.01
SR2	Top veto	1.2 ± 0.3	3.1 ± 0.1	120055	$(119.7 \pm 0.6)10^3$	1.00 ± 0.01
	Z-veto	1.0 ± 0.2	$(6.1 \pm 0.6)10^{-1}$	17891	$(189.8 \pm 0.2)10^3$	0.94 ± 0.01
	$p_T(\tau)$	1.2 ± 0.2	2.7 ± 0.1	29020	$(36.8 \pm 0.3)10^3$	0.79 ± 0.01
	Flavour splitting	0.2 ± 0.0	$(2.3 \pm 0.3)10^{-1}$	560	580 ± 6	0.97 ± 0.04
	Njets	0.2 ± 0.0	$(2.1 \pm 0.3)10^{-1}$	514	519 ± 6	0.99 ± 0.05
SR3	Top veto	$(0.7 \pm 0.0)10^{-1}$	$(0.9 \pm 0.2)10^{-1}$	104	111 ± 4	0.94 ± 0.10
	$p_T(\tau)$	2.9 ± 0.4	5.2 ± 0.2	136945	$(133.0 \pm 0.6)10^3$	1.03 ± 0.11
	Flavour splitting	0.9 ± 0.2	$(2.8 \pm 0.4)10^{-1}$	219	$(2.3 \pm 0.2)10^2$	0.96 ± 0.11
	Njets	0.9 ± 0.2	$(2.7 \pm 0.4)10^{-1}$	215	$(2.2 \pm 0.2)10^2$	0.96 ± 0.11
	Top veto	0.5 ± 0.2	$(1.5 \pm 0.3)10^{-2}$	161	$(1.8 \pm 0.2)10^2$	0.90 ± 0.10

Table 4.16: Blinded signal regions cut-flow. Each row shows the number of signal events and total background events after the selection is applied; the number of observed data events is also shown, and the Data/MC can be found in the last column. The number of events is reported with its statistical uncertainty. Numbers smaller than 0.1 are rounded as 0 in this table.

4.6.5 Unblinding: cut-based analysis results

This section describes the results obtained with the cut-based analysis approach, once the blinding veto has been removed from data. MC expectations for the Higgs boson associated production with $m_H = 125$ GeV are compared with data. The statistical interpretation of the results is given in chapter 6, as a conclusion of the analysis. In table 4.17 the unblinded cut-flow is shown. In this table the column reporting the number of events observed in each SR together with that displaying the data/MC ratio have been added. NF as in table 4.14 are applied to Top and Z+jets backgrounds. A good data/MC agreement is obtained, with a ratio compatible with 1 in all the SR. Some fluctuations in the data/MC ratio are visible especially in the *Z-depleted* regions, SR2 and SR3, dominated by a low statistics. The only disagreement is indeed found in SR3, where the ratio increases from 1.0 to 2.1 by cutting on the ΔR_{01} variable. The ΔR_{01} distribution before the last selection in SR3 is shown in figure 4.22. By selecting $\Delta R_{01} < 1$ only the first four non-empty bins are considered, three of which have both data and MC points (bins 0.4 - 1). In these three bins data points are systematically over the total background, but the statistics is too low, as reflects in the statistical error which comes with these points and with the data/MC ratio in this region, to draw any conclusion. Moreover within two standard deviations data are compatible with MC, and the disagreement is then not worrisome.

Signal Region	Selection	$WH(WW)$	$ZH(WW)$	$VH(\tau\tau)$	Observed	Total Background	Data/MC
SR1	3 leptons Preselection	5.9 ± 0.5	8.1 ± 0.2	0.9 ± 0.0	188217	$(182.5 \pm 0.7)10^3$	1.03 ± 0.00
		4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	137172	$(133.2 \pm 0.6)10^3$	1.03 ± 0.01
	$p_T(\tau)$ Flavour splitting Njets Top veto Z-veto MET ΔR_{01}	4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	137172	$(133.2 \pm 0.6)10^3$	1.03 ± 0.01
		1.9 ± 0.3	4.6 ± 0.2	0.3 ± 0.0	135062	$(131.1 \pm 0.6)10^3$	1.03 ± 0.01
		1.7 ± 0.3	4.4 ± 0.2	0.3 ± 0.0	134090	$(130.2 \pm 0.6)10^3$	1.03 ± 0.01
		1.4 ± 0.3	3.1 ± 0.1	0.2 ± 0.0	120084	$(119.7 \pm 0.6)10^3$	1.00 ± 0.01
		1.1 ± 0.2	$(6.8 \pm 0.6)10^{-1}$	0.2 ± 0.0	17920	$(19.0 \pm 0.2)10^3$	0.94 ± 0.01
		$(2.9 \pm 0.0)10^{-1}$	$(1.6 \pm 0.3)10^{-1}$	0.0 ± 0.0	218	246 $\pm$ 15	0.89 ± 0.08
		$(2.1 \pm 0.0)10^{-1}$	$(8 \pm 2)10^{-2}$	0.0 ± 0.0	63	68 $\pm$ 7	0.93 ± 0.15
		2.0 ± 0.2	3.1 ± 0.1	0.3 ± 0.0	29103	$(36.9 \pm 0.3)10^3$	0.79 ± 0.01
		$(6.4 \pm 0.8)10^{-1}$	$(4.3 \pm 0.5)10^{-1}$	0.1 ± 0.0	613	635 $\pm$ 7	0.97 ± 0.04
		$(6.4 \pm 0.8)10^{-1}$	$(4.1 \pm 0.5)10^{-1}$	0.1 ± 0.0	567	574 $\pm$ 7	0.99 ± 0.04
SR2	$p_T(\tau)$ Flavour splitting Njets Top veto MET ΔR_{01}	$(5.1 \pm 0.8)10^{-1}$	$(2.9 \pm 0.4)10^{-1}$	0.1 ± 0.0	157	166 $\pm$ 5	0.94 ± 0.08
		$(4.2 \pm 0.8)10^{-1}$	$(2.0 \pm 0.3)10^{-1}$	0.0 ± 0.0	81	81 $\pm$ 3	1.00 ± 0.11
		$(3.6 \pm 0.8)10^{-1}$	$(1.3 \pm 0.3)10^{-1}$	0.0 ± 0.0	32	27 $\pm$ 1	1.21 ± 0.22
		4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	137172	$(133.2 \pm 0.6)10^3$	1.03 ± 0.01
		1.4 ± 0.3	$(5.6 \pm 0.5)10^{-1}$	0.0 ± 0.0	256	$(2.7 \pm 0.2)10^2$	0.94 ± 0.09
		1.4 ± 0.3	$(5.6 \pm 0.5)10^{-1}$	0.0 ± 0.0	252	$(2.7 \pm 0.2)10^2$	0.94 ± 0.10
		1.1 ± 0.2	$(4.2 \pm 0.5)10^{-1}$	0.0 ± 0.0	198	$(2.3 \pm 0.2)10^2$	0.88 ± 0.10
		$(6.5 \pm 1.6)10^{-1}$	$(2.4 \pm 0.3)10^{-1}$	0.0 ± 0.0	44	45 $\pm$ 4	0.98 ± 0.17
		$(3.5 \pm 1.1)10^{-1}$	$(9 \pm 2)10^{-2}$	0.0 ± 0.0	12	5.6 $\pm$ 0.5	2.13 ± 0.64
SR3	$p_T(\tau)$ Flavour splitting Njets Top veto MET ΔR_{01}						

Table 4.17: Unblinded signal regions cut-flow. Each row shows the number of signal events and total background events after the selection is applied; the number of observed data events is also shown, and the Data/MC can be found in the last column. The number of events is reported with its statistical uncertainty. Numbers smaller than 0.1 are rounded as 0 in this table.

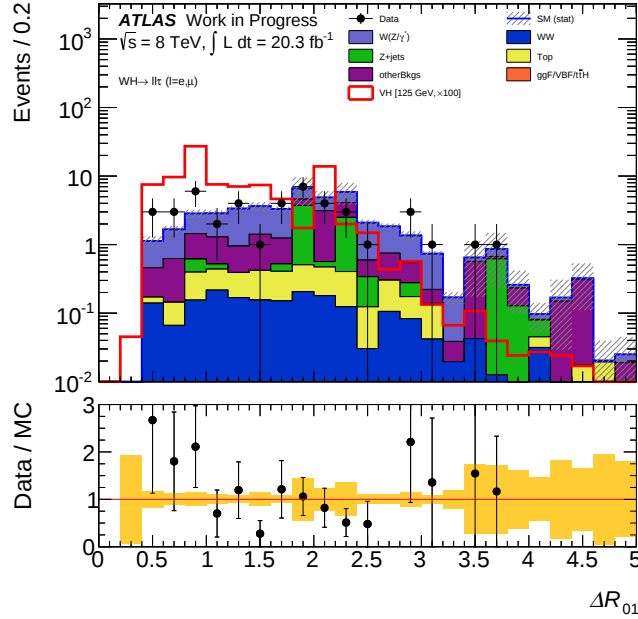


Figure 4.22: Unblinded ΔR_{01} distribution in SR3 before the last selection. Data (dots) are compared to expectations from the simulation of the background components (stacked filled histograms). Expectations for a Standard Model Higgs boson of mass $m_H = 125$ GeV are multiplied by a factor 100 and are presented as a non-stacked histogram (red line). Only statistical errors are shown.

4.7 Systematic uncertainties

Results shown in the previous sections were obtained by considering only the statistical uncertainty. Statistical uncertainties are the result of stochastic fluctuations arising from the fact that a measurement is based on a finite set of observations. Repeated measurements of the same phenomenon will therefore result in a set of observations that will differ, and the statistical uncertainty is a measure of the range of this variation. In any measurement, systematic uncertainties must be also taken into account.

Systematic uncertainties arise from uncertainties associated to the nature of the measurements apparatus (e.g. uncertainties from the calibration of the detector), to assumptions made by the experimenter to obtain the result, or to the model used to make inferences based on the observed data [77]. Unlike the statistical uncertainties, systematic errors are generally correlated from one measurement to the next, and cannot be reduced by increasing the statistics of the data sample. In this section, the main sources of systematic errors and their impact on the measurement are discussed. A detailed study of systematic uncertainties was not performed so far. Most of the systematics are in common with the purely leptonic analysis [67] and we will exploit the systematic uncertainties already computed in that context whenever possible. A careful study of systematics will be certainly needed for the future but with the data set available at the moment, the statistical error is so large that the impact of systematic uncertainties does not affect the analysis result.

Systematic uncertainties to be considered can be divided in two main groups:

- Theoretical uncertainties

- Experimental uncertainties

4.7.1 Theoretical uncertainties

In predicting a process rate, e.g. the WH production at the LHC, some assumptions are needed: the energy scale to which the process occurs, Q^2 (for the WH system Q^2 is the invariant mass of the WH system) is fixed and a specific choice for the PDF of the interacting partons is done. The cross section obtained with such calculations is therefore affected by uncertainties; these uncertainties are evaluated, for the scale, by varying the renormalization and the factorization scale independently within the interval $[Q/3, 3Q]$, and for the PDF by using several set of PDF (e.g. CT10 [65], MSTW2008 [78],...).

For the Higgs boson cross-section, the uncertainties are evaluated following the recommendation of the LHC Higgs cross-section working group [16][17][18]; these include also the uncertainty on the $H \rightarrow WW^*$ branching ratio that arises from two main sources, missing high order corrections (theoretical uncertainties) and experimental errors on the SM input parameters, such as quark masses (parametric uncertainties). For the signal at $m_H = 125$ GeV, the size of the scale uncertainty and the PDF uncertainties are evaluated as 1.0% and 2.3%, respectively, while the uncertainty on the $H \rightarrow WW^*$ branching ratio has been evaluated as $\sim 4\%$ [16].

For non-Higgs processes which constitute the main backgrounds in each of the SR defined (Top, Z+jets, WZ/ γ^* and WW), the uncertainties used for the purely leptonic analysis have been applied.

The impact of these uncertainties on each MC process was evaluated by varying each source by $\pm 1\sigma$, σ being the theoretical error, and then quantifying the change in the event yield obtained. For non-Higgs processes theoretical uncertainties are negligible with respect to the experimental uncertainties. For Higgs processes, the impact of these systematics on the event yield is found to be $\sim 4\%$ for the Higgs branching ratio, $\sim 1.2\%$ for the PDF variation and $\sim 1.5\%$ for QCD scale.

4.7.2 Experimental uncertainties

The measurement described in this thesis relies on the knowledge of the detector performance. These quantities, such as reconstruction and identification efficiency for leptons and jets, the energy scales and the energy resolution of the different sub-detectors, were measured by the performance groups. Systematic uncertainties on these measurements are also provided and should be taken into account. The systematic uncertainties divide in two main categories: 1) uncertainties common to all the samples which result in a change of the event weight and 2) uncertainties affecting objects' kinematics whose impact depends on the particular event considered. The first category includes the uncertainty on the integrated luminosity, evaluated as the 2.8% in 2012 data. Moreover, the statistical error relative to the NF, obtained when normalizing Top and Z+jets background processes to data in the CR (cfr table 4.14), is also taken into account and treated as a systematic error. The overall impact of the remaining event-dependent uncertainties on the analysis is instead evaluated with the following procedure:

- The systematic source of interest is varied by 1σ ;
- All the MC sample are re-reconstructed with this change and the analysis repeated without changing anything else;
- The impact of the systematic source is evaluated as a variation on the event yield.

As already explained, the evaluation of the impact of the systematic uncertainties is derived from the purely leptonic analysis, as documented in [67]. The re-processing of the entire MC samples

has not been done directly for this thesis; the impact of an experimental systematic on a given sample has been increased by 20% respect to that found in purely leptonic analysis. This was done to use a conservative estimation of the errors, since most of the experimental systematics are analysis dependent and their impact should in principle be evaluated on the phase space relative to the SR defined in this analysis. Moreover, the request for a tau decaying in hadrons, and the probability for it to be a fake tau (see Appendix B), increase the uncertainty on the total number of events with respect to that in the purely leptonic analysis. The main sources of experimental uncertainty are the modelling of the pile-up conditions, the lepton identification efficiency (whose impact is $\sim 3\%$ on the signal, $\sim 4\%$ on the Z+jets background) and the jet energy scale and resolution (whose impact is $\sim 3\%$ on the signal, $\sim 16\%$ on the Top background), the latter affecting also the estimation of the missing transverse momentum.

In table 4.18 the relative impact of both the statistical and the systematic uncertainties on signal and background event yields is given. The uncertainties have then been used in the final statistical fit in chapter 6. The relative contribution of each background process to the total background in the three SR is also reported.

		VH	$ggF/VBF/t\bar{t}H$	Top	WW	WZ/γ^*	$Z + jets$	$OtherBkgs$	Total Bkg
SR1	Stat error	3.3%	14%	3.3%	5.5%	6%	25%	25%	9%
	Syst error teo	5.2%	11%	-	-	-	-	-	0%
	Syst error exp	4%	16%	24%	17%	2.4%	32.4%	24%	17%
	% total bkg	0%	0%	23%	10%	7%	51%	9%	100%
SR2	Stat error	20%	20%	3.4%	6.1%	9%	10%	21.4%	9.6%
	Syst error teo	5.2%	11%	-	-	-	-	-	0%
	Syst error exp	4.6%	17%	18%	10%	7%	13%	24%	8.6%
	% total bkg	0%	2%	42%	19%	13%	14%	12%	100%
SR3	Stat error	20%	0%	25%	25%	6.5%	100%	25%	9.3%
	Syst error teo	5.2%	11%	-	-	-	-	-	0%
	Syst error exp	4.6%	17%	18%	10%	7%	13%	24%	8%
	% total bkg	0%	0%	6%	7%	55%	4%	29%	100%

Table 4.18: Summary of the impact of statistical and systematic uncertainties in each SR for signal and background processes. Systematic uncertainty is divide in theoretical (teo) contribution and experimental (exp) contribution. Where not present, the error was considered negligible. The last row in each SR section shows the relative contribution of a given background process to the total background.

Even if differences are observed among the various SR, the systematic uncertainties have the same impact on the background event yield as the statistical uncertainty. On the signal, on the contrary, the statistical uncertainty is dominant with respect to the systematic uncertainty. However, the statistical uncertainty in our measurement is dominated by the statistical error on data; therefore we do expect, and will be proved in chapter 6, that systematic uncertainties will have a small impact on our result.

Chapter 5

MVA analysis

A multivariate analysis (MVA) [25] is a statistical technique that examines the relations among multiple variables at the same time to classify an event as signal or background. Rather than assigning an event to a definite class membership, a probability for it to belong to a certain class is defined. With respect to cut-based analysis a MVA does not act on the single variables but combine several variables into a more powerful variable, called the *classifier*, taking into account also the variables correlations. The analyzer can then use this variable to decide if an event can be selected as signal or rejected as background-like event, depending if it passes a threshold or not. Alternatively, the variable distribution can be used in a fit with the signal and the background components in order to obtain their relative amount in the data sample.

MVA analysis is commonly used in particle physics especially when the expected number of signal events is very low. In this case applying sequentially several selections to reduce the background, as in a cut-based analysis, is not convenient, since with the background reduction usually comes also the signal reduction. In addition, a signal event that might look background-like in only a single observable will inevitably be mis-classified as background in a cut-based analysis. However, it might be correctly classified with a multivariate classification approach that is able to compensate for this one background-like feature by exploiting all the other observables that might look very signal-like.

In this thesis two different MVA based on Boosted Decision Tree (BDT) were adopted. First a MVA was applied in all the SR in conclusion of the cut-based analysis to enhance the sensitivity. Then another MVA was used as an alternative approach to the cut-based analysis: a combined fit on the BDT results from all the SR has been performed. Before describing in detail the analysis strategy a preliminary introduction about the multivariate classification technique and the BDT method is given.

5.1 Event classification in a statistical perspective

Hypothesis testing is a commonly used method for decision making and for drawing conclusions on an acquired set of measurement. This test consists in formulating a null hypothesis (H_0) and in determining if a dataset is consistent or not with H_0 . In event classification, where we want to select the signal events we are interested in, H_0 is the hypothesis that the event originates only from SM background processes. In MVA the null hypothesis is rejected or not depending on the value of the MVA variable (the *classifier*).

In most cases the probability densities functions (PDF) of the observables for signal and background events overlap; this means that there is a region in phase space where one can find both signal and background events. For this reason some background events will be misclassified as signal (leading to a *type I error*), or one fails to classify signal events and assign them to the

background category (*type II error*). These errors come with a probability, denoted α for the type I error, and β for the type II errors. The quantities $(1 - \alpha)$ and $(1 - \beta)$ are then called respectively *background rejection* and *signal efficiency*. In each individual classification problem one has to find the best balance between type I and type II errors.

To display the performance of a classification algorithm the *Receiver-Operating-Characteristics (ROC)* curve is usually drawn. It shows the background rejection as a function of the signal efficiency that can be obtained by varying the cut on the MVA classifier.

5.2 Training and classification

A MVA consists in two main steps:

Training: It is the process that defines the splitting criteria for each node (par. 5.3). A sample of events for which membership is known (Monte Carlo) is used to train the multivariate analysis method in separating signal events from background events. A classifier is built whose performances are then evaluated and tested using a different known data sample. During the training the classifier learns how the discriminating variables have to be chosen in order to get an optimal signal/background separation.

The most commonly used separation criteria is through the *Gini index*, defined as $p \cdot (1 - p)$, where p is the purity of the node, defined as the ratio of signal events to all events in that node. Pure background nodes have zero purity.

Classification: Uses the result from the training step to classify unknown (data) events into signal or background.

5.3 BDT

A decision tree is a tree-structured classifier that consists of a series of binary splits, as displayed in figure 5.1. The tree starts from a root node and then it is built up of repeating splits and nodes down to the final leaf nodes. The set of nodes and splits leading to a leaf node is called branch. Thanks to these repeated yes/no decisions the phase space is split into several sub-regions that are classified as signal or background depending on the majority of events that end up in the final leaf node. The split criteria are simple cuts on individual observables.

Unfortunately decision trees appear to be quite unstable with respect to statistical fluctuations in the training sample from which the tree structure is derived. For example if two variables exhibit a similar signal/background separation power, then a fluctuation in the sample may cause the tree growing algorithm to decide to split on one variable, while without the fluctuation the other variable would have been selected. To overcome this problem in the early 1990s a new technique was developed aimed to increase the decision tree performance: the *Boosting* [70] [71]. Applying this technique to a decision tree we then get a *Boosted Decision Tree* (BDT) that combines many different decision trees to form what is called a *forest*. In a BDT first a classifier is trained using the training data sample; then a new training iteration starts with a modified training sample in which the previously misclassified events are given a larger weight. The procedure is iterated many times and finally the result of all the different classifiers obtained is averaged. The BDT response will result in a distribution peaked at -1 for background-like events and peaked at 1 for signal-like events.

To the scope of this thesis the TMVA [72] toolkit was used. TMVA is an analysis framework integrated in ROOT which hosts a large variety of multivariate classification algorithms, including the BDT.

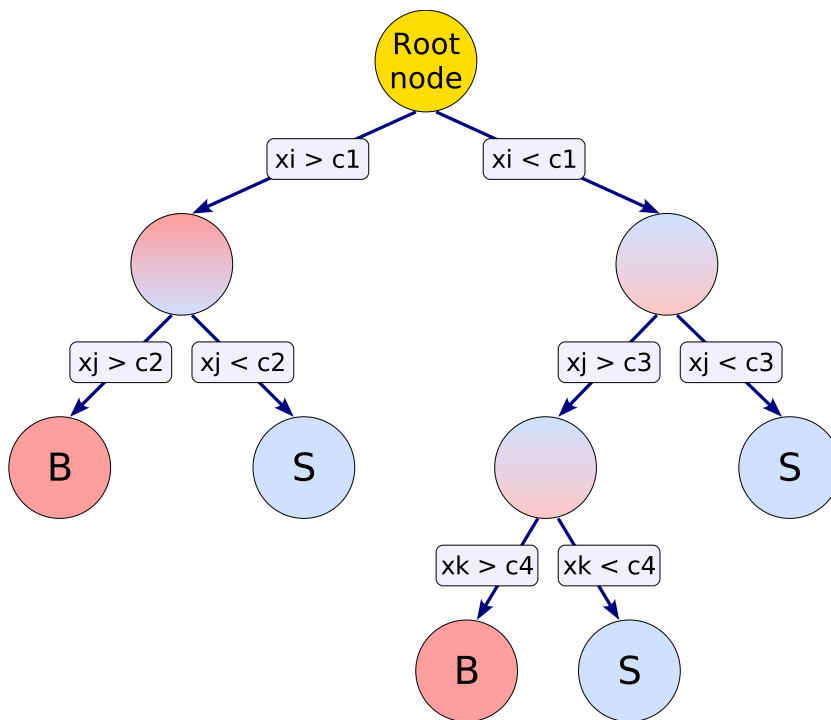


Figure 5.1: Schematic view of a decision tree.

5.4 BDT against leading backgrounds

The cut-based analysis explained in section 4.5 was not so effective in reducing the background contaminating the SR, as can be seen by looking at the significance obtained (table 4.9). With the aim of enhancing the sensitivity in our SR, a BDT technique has been applied in SR1 to reduce the dominant Z+jets background, in SR2 to reduce the Top background and in SR3 to reduce WZ/ γ^* background. All these BDT will result effective also in reducing non-leading backgrounds.

In the next sections the MVA analysis procedure is explained and discussed in detail, as an example, for the SR1, and briefly, for the other two SR.

5.4.1 WH vs Z+jets

Although specific selections have been applied to reduce the Z+jets contribution in SR1, as a tight MET cut and the Z-veto requirement, it still remains the main background among the SFOC topology events. In the following a detailed description of the multivariate technique used for the reduction of such background is given. Some general aspects, as the training variables optimization, are in common between the various MVA here described; unless otherwise indicated the same results are applied.

Training

The BDT was built with the parameters in table 5.1, which are the default values when using the BDT method in ATLAS [72].

Some preliminary selections are applied on the input events on which the training is performed. This allows to filter out those events that certainly do not belong to the signal category.

BDT parameter	Value used
NTrees	1000
nEventsMin	≈ 1600
MaxDepth	3

Table 5.1: BDT parameters used in the analysis. NTrees is the number of trees in the forest, nEventsMin is the minimum number of events in a final leaf node and MaxDepth is the number of layer in a tree. In particular nEventsMin determines when stop the node splitting: when the number of events in the node reach the nEventsMin, the node is classified as signal or background according to the class the majority of events belongs to.

The list of the selections applied on the input events is:

- Three $p_T > 15$ GeV isolated leptons, one of which is a tau
- Total charge = ± 1
- Trigger matching for one lepton candidate (e or μ)
- Lepton flavour splitting: SFOC events required
- Jet veto: Njets ≤ 3
- Top veto: no b-tagged jets with $p_T > 20$ GeV
- MET selection: MET_STVF, MET_TrackClj > 15 GeV

Table 5.2 reports the number of events used in the training step, after the previous selections were applied.

Signal events	≈ 12000
Background events	≈ 53000

Table 5.2: Number of signal (WH) events and background (Z+jets) events used for the training of the BDT.

BDT optimization

The first step in MVA, after setting the BDT parameters, consists in choosing the best performing set of variables to be used in the training of the BDT. Variable performance is established in terms of the signal/background separation defined as:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(z) - \hat{y}_B(z))^2}{\hat{y}_S(z) + \hat{y}_B(z)} dz \quad (5.1)$$

where $\hat{y}_S$ and $\hat{y}_B$ are the signal and background PDF of the variable z , respectively. Of course when the signal and background distributions do not overlap the separation is maximum and equals to 1, while is zero for identical shapes. To understand if the chosen set of variables is the optimal and if contains or not meaningless variables, the BDT result stability in terms of statistical fluctuations should be checked. The procedure adopted is:

- Start with a set of variables and order them on the base of their signal/background separation, from the most to the least discriminating variable;
- Train the BDT and check its performance;
- Remove variables one by one, from the one with the worst to the one with the best separation, every time train the BDT again with the new set of variables. Check the BDT performance;
- Repeat the operation until the list will contain just one variable.

This procedure ignores if a different set of variables would have performed better if used together; nevertheless it is a good way to understand if some useless variables are in the set. In table 5.3 the ranking of the variables used is reported. According to the criteria described the first variable

Rank	Variable	Separation
1	MET_TrackClj	0.455
2	MET_STVF	0.453
3	ΔR_{01}	0.382
4	m_{01}	0.303
5	lepPtSum	0.268
6	m_{02}	0.223
7	m_{012}	0.135
8	lepPt1	0.065
9	lepPt2	0.059
10	lepPt0	0.057
11	m_{12}	0.019

Table 5.3: Variables ranking based on the signal/background separation. The variable lepPtSum is the vector sum of the lepton transverse momenta, while m_{012} is the three lepton invariant mass.

removed from the set was then m_{12} . Figure 5.2 shows the distributions of the input training variables for signal (in blu) and background (in red), while figure 5.3 shows variables correlation matrix for both signal and background. The two MET variables are highly-correlated. By removing one of these MET variables the BDT performance get slightly worse, in term of the overall signal/background separation obtained. For this reason both variables are kept for the training. For simplicity, the previous figures contain only those variables chosen with the optimization procedure and effectively used in the training procedure.

BDT performances are evaluated considering the following two aspects

1. Overtraining
2. ROC curve area

Overtraining occurs when the BDT has too few degrees of freedom, because too many model parameters of an algorithm were adjusted to few data points. An overtrained classifier does not

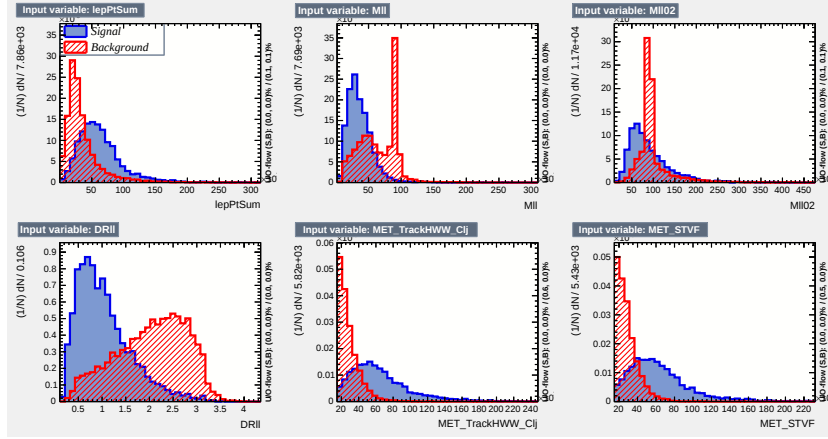


Figure 5.2: Distributions of the input training variables for Z+jets background (red curve) and WH signal (blue curve). Even if not specified, the M_{ll} variable is the dilepton invariant mass m_{01} .

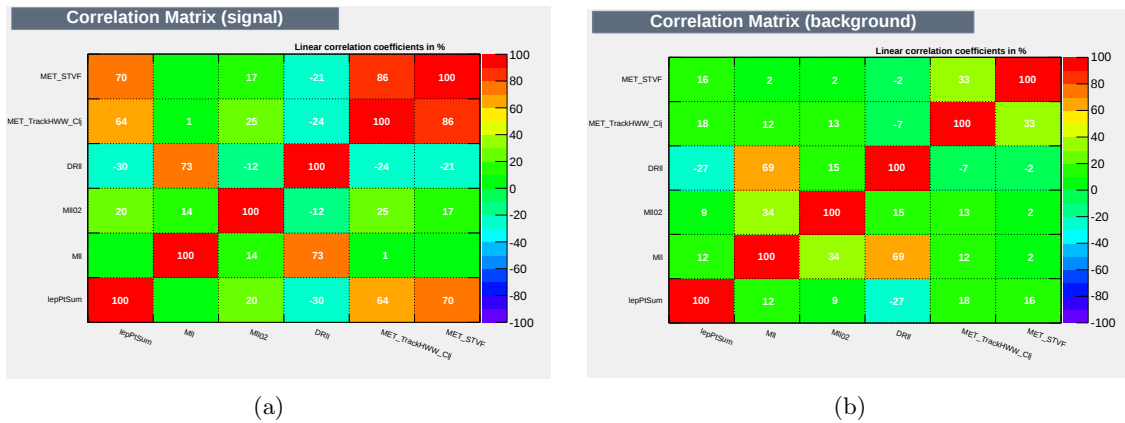


Figure 5.3: Correlation matrices among the training input variables for (a) the signal and (b) the Z+jets background.

capture the general features of the underlying PDF and so it is not flexible enough to adapt to a different set of events. To check if the method is overtrained or not the sample of events is divided in two sub-samples, the *training sample* and the *test sample*, and the BDT is evaluated twice, once for each sample. The results from the training and test samples are compared with the Kolmogorov-Smirnov (KS) test [72]. The KS test [73] is a statistical test that can be used to compare two samples. It measures the probability that a chosen dataset is drawn from the same parent population as a second dataset. The difference between the distribution functions of the two datasets is evaluated. The KS probability quantifies the compatibility between the two results; a small value of the KS parameter is an indication of overtraining. The KS result for the training against the Z+jets background can be found in figure 5.4.

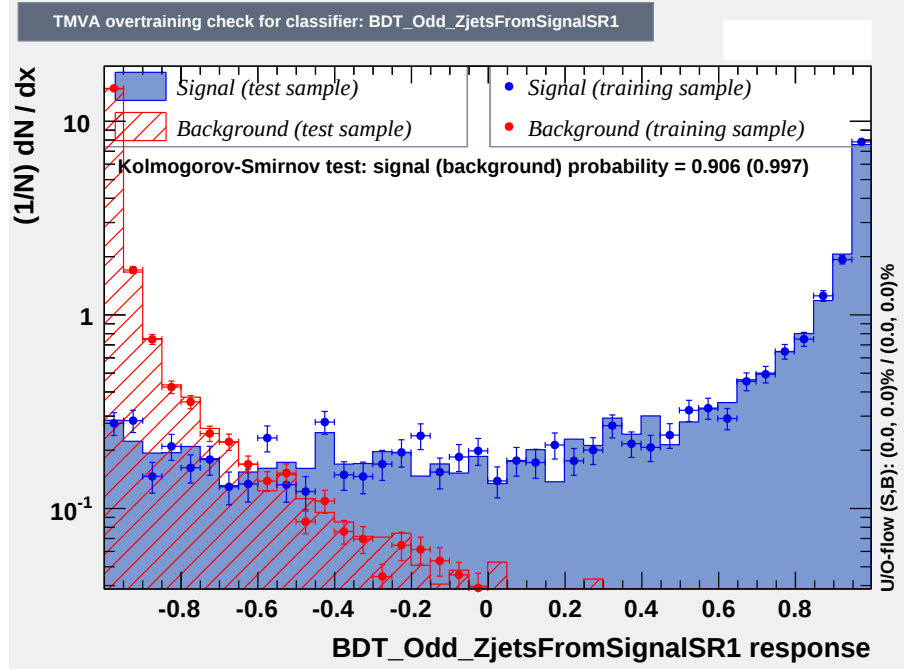


Figure 5.4: Kolmogorov-Smirnov test result for the signal/Z+jets discrimination. The training is performed on the sample of odd-events and the result is applied on the sample of even-events. The high value of the KS parameter (KS=0.906) is an indication of the robustness of the training procedure.

The overall performance for all possible cut values for a given classification algorithm is easily visualised using the ROC curve (figure 5.5): the larger the area underneath the curve the better the algorithm performance. In figure 5.6 the ROC curve area as a function of the number of variables used for the training is displayed.

From the figure a small dependence on the number of variables in the set is evinced. The area does not change by adding more variables to the set of 5/6 variables, which means that the BDT performance remains stable.

Application and results

Once the training is complete, a recipe for discriminating the signal, in this case the WH, from the background, the Z+jets, is obtained. This recipe is then used to discriminate signal events from background events (this time all the topologies contributing to the total background in SR1 are considered), in the so-called "application" stage. During the application the BDT

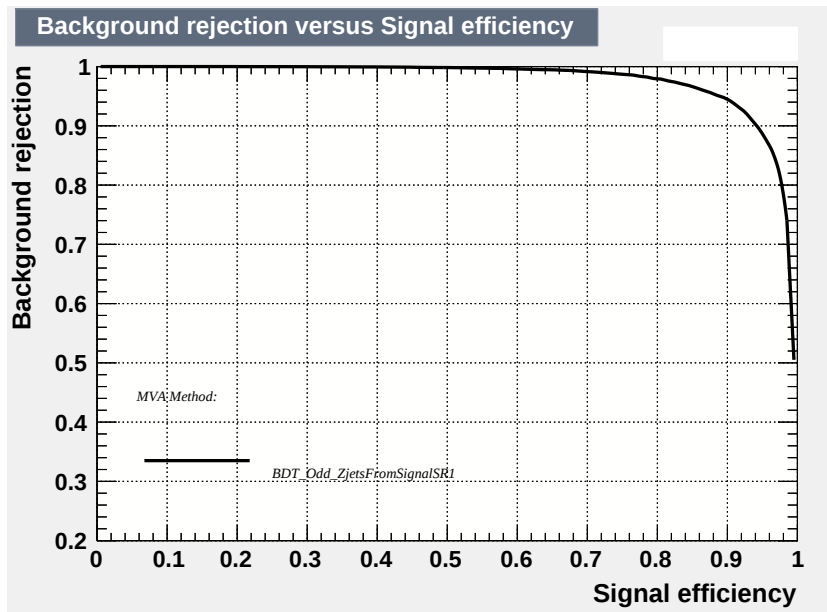


Figure 5.5: ROC curve as a result of the BDT for the signal/ Z +jets discrimination.

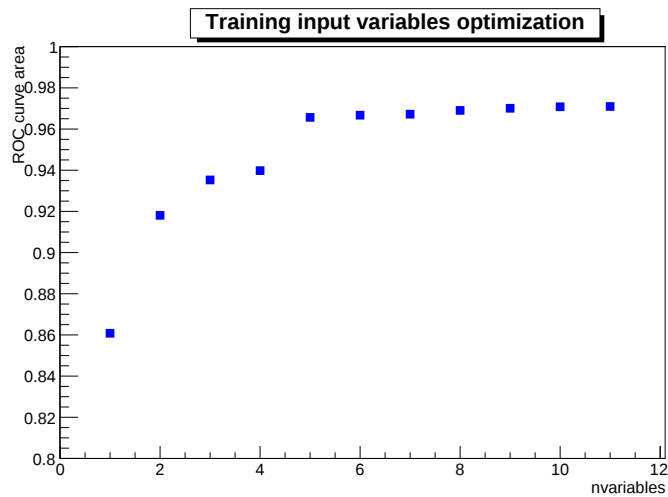


Figure 5.6: ROC curve area value as a function of the number of training variables.

doesn't know if a given event belongs to the signal category or not; in fact the lesson learnt from the training is used to classify data events in signal or background events. The BDT output distribution obtained for SR1 can be found in figure 5.7. As can be seen (figure 5.7(a)) the Z+jets

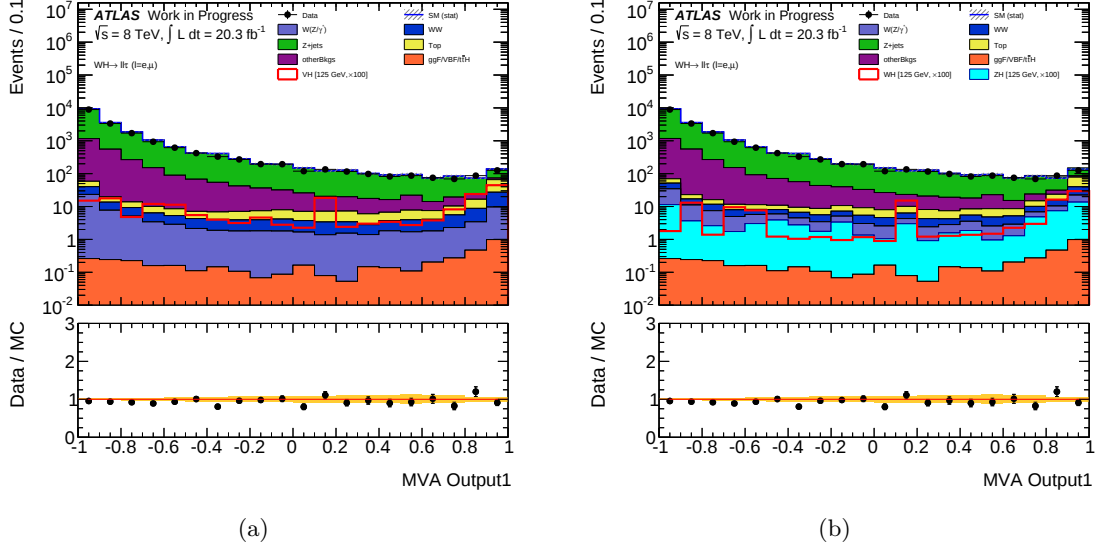


Figure 5.7: BDT output distribution in SR1 with (a) total VH signal shown as red non-stacked histogram (b) WH process shown as red non-stacked histogram, ZH process as cyan stacked histogram. Both WH and ZH contributions have been multiplied by a factor of 100. Only statistical uncertainties are drawn.

is peaked around -1 and then decreases, while the VH, although its shape is less pronounced, tends to be peaked around +1. It should be noticed that the training was performed with the aim of classifying the WH events, while here the BDT output distribution is drawn for WH+ZH events. To demonstrate that the BDT distribution for the ZH process tends to distort the whole signal (VH) distribution, figure 5.7(b) reports the WH process as a red non-stacked line, while the ZH process is drawn as a stacked cyan histogram. The peaks in the WH distribution are only due to few MC events with a big weight (cfr equation 4.8).

It can be also noticed that not only the Z+jets background appears to have a different shape with respect to the signal, other backgrounds behave as the Z+jets, peaking at low values of the BDT output. This means that cutting on the BDT output to reject as much background events as possible while keeping most of the signal statistics, we could end with a significant gain in sensitivity. However in this thesis a different approach has been followed; the BDT output distribution obtained in each SR has been used in a combined shape fit, so that the significance of each bin is combined with that of the others, and an improvement in terms of the sensitivity is obtained without losing any event. This approach is discussed in the next chapter, in the context of the analysis results.

5.4.2 WH vs Top

Top is the main background in SR2. A multivariate technique can be used also in this case to catch the main differences between signal and background. A BDT is trained in SR2 and the result is shown in figure 5.8. The separation power of the BDT is clearly visible in this case, with the signal peaked around +1 while background distribution appears constant.

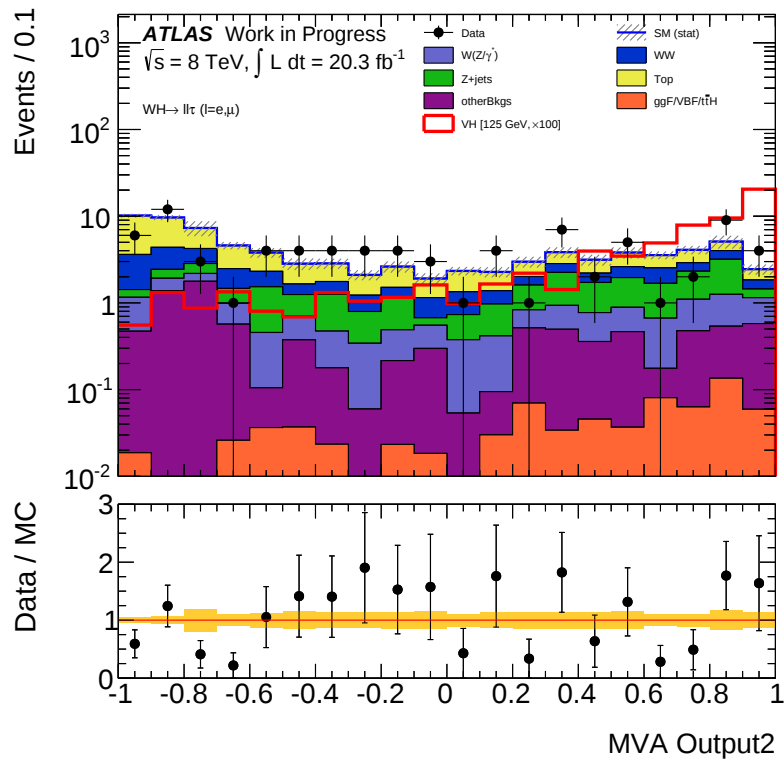


Figure 5.8: BDT output distribution in SR2. Signal is shown as a red non-stacked line, and its contribution as been multiplied by a factor of 100. Only statistical uncertainties are drawn.

5.4.3 WH vs WZ/ γ^*

WZ/ γ^* is the main background process in SR3; this signal region is also the region with the highest significance we have (see table 4.9). A BDT has then been used to distinguish the WH events from this background. The variables used for the training are in figure 5.9. With respect to the other BDT trained for this thesis, the variables used in this region have a poor separation power, with the exception of the ΔR_{01} and the m_{01} variables, which are also highly correlated. No more variables have been found to be better performant than those in figure 5.9. As a

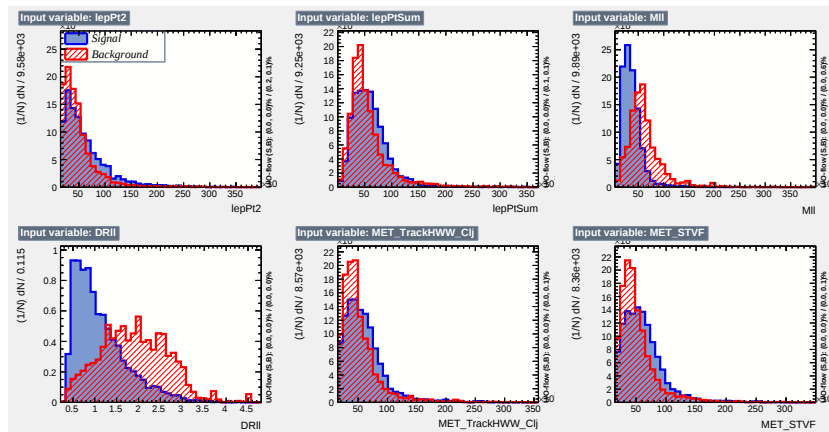


Figure 5.9: Shapes of the input training variables for WZ/ γ^* background (red curve) and WH signal (blue curve).

result, the BDT output distribution for signal and backgrounds overlap over a wide region in the interval $[-1,1]$ (figure 5.10), although the former is again peaked towards positive BDT values while the latter is peaked towards negative BDT values. Of course here the main problem is the lack of MC statistics, which is the main ingredient for finding a general recipe to discriminate signal from background. Moreover, the WZ/ γ^* is the background most similar to our signal, as results from the training variable distributions, and so the hardest to reduce. Nevertheless, the BDT variable has a considerable rejection power against the dominant WZ/ γ^* ; the effect of a cut applied on the BDT output in SR3, required to be greater than 0.0, 0.2 or 0.4, is reported in table 5.4. Comparing the significance obtained at the end of the cut-flow (0.18 ± 0.05) with that obtained with the selection $BDT > 0.4$ (0.21 ± 0.06), the significance gain is around the 20%, even if the statistical error tends to increase. However, as in the previous cases, the BDT output distribution will be used in the final combined shape fit, described in the next chapter, and a better gain in significance is obtained.

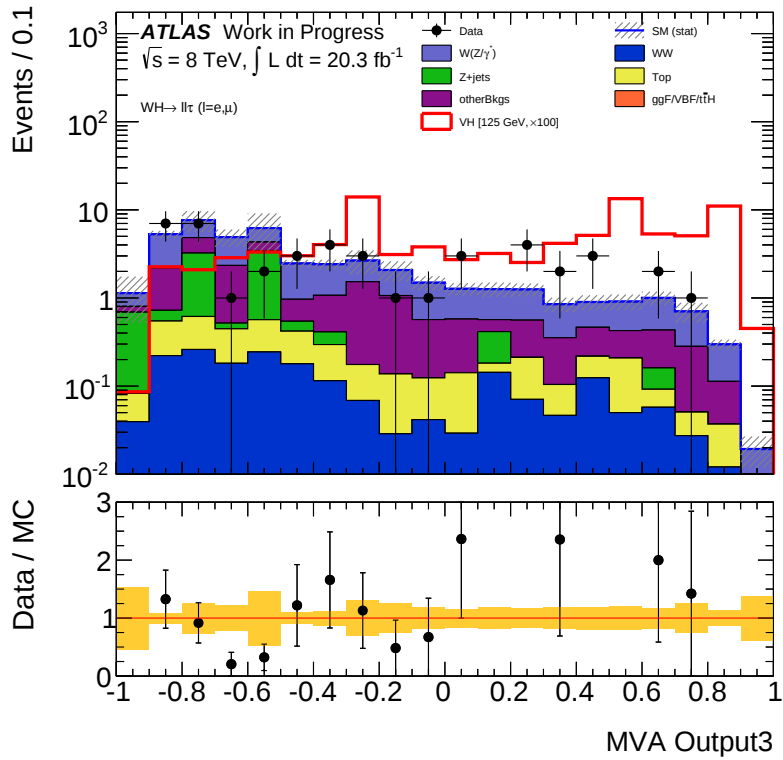


Figure 5.10: BDT output distribution in SR3. Signal is shown as a red non-stacked line, and its contribution has been multiplied by a factor of 100. Only statistical uncertainties are drawn. The peaks observed in the signal are due to few MC events with a big normalization factor.

Signal Region	Selection	$WH(WW)$	$ZH(WW)$	$VH(\tau\tau)$	Total Background	$S/\sqrt{B}$
	3 leptons	5.9 ± 0.5	8.1 ± 0.2	0.9 ± 0.0	$(182.5 \pm 0.7)10^3$	0.03 ± 0.00
	Preselection	4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	$(133.2 \pm 0.6)10^3$	0.03 ± 0.00
SR3	$p_T(\tau)$	4.3 ± 0.4	5.9 ± 0.2	0.6 ± 0.0	$(133.2 \pm 0.6)10^3$	0.03 ± 0.00
	Flavour splitting	1.4 ± 0.3	$(5.6 \pm 0.5)10^{-1}$	0.0 ± 0.0	$(2.7 \pm 0.2)10^2$	0.12 ± 0.02
	Njets	1.4 ± 0.3	$(5.6 \pm 0.5)10^{-1}$	0.0 ± 0.0	$(2.7 \pm 0.2)10^2$	0.12 ± 0.02
	Top veto	1.1 ± 0.2	$(4.2 \pm 0.5)10^{-1}$	0.0 ± 0.0	$(2.3 \pm 0.2)10^2$	0.10 ± 0.02
	MET	0.7 ± 0.2	$(2.4 \pm 0.3)10^{-1}$	0.0 ± 0.0	45 ± 4	0.14 ± 0.03
	ΔR_{01}	0.4 ± 0.1	$(9 \pm 2)10^{-2}$	0.0 ± 0.0	5.6 ± 0.5	0.18 ± 0.05
	$BDT > 0.0$	$(3.5 \pm 1.1)10^{-1}$	$(8 \pm 2)10^{-2}$	0.0 ± 0.0	5.3 ± 0.5	0.19 ± 0.05
	$BDT > 0.2$	$(3.4 \pm 1.1)10^{-1}$	$(5 \pm 2)10^{-2}$	0.0 ± 0.0	4.4 ± 0.5	0.19 ± 0.06
	$BDT > 0.4$	$(3.2 \pm 1.1)10^{-1}$	$(5 \pm 2)10^{-2}$	0.0 ± 0.0	3.2 ± 0.3	0.21 ± 0.06

Table 5.4: Signal regions cut-flow obtained with MC signal and background samples. Each row shows the number of signal events and total background events after the selection is applied; the significance is also evaluated. The last three rows show the effect of cutting on the BDT output in figure 5.10. The fifth column shows the analysis acceptance for the process $VH \rightarrow \tau\tau$. The number of events is reported with its statistical uncertainty. Here each number is shown with more significant digits with respect to the previous cut-flow tables, to show how the cut on the BDT output affects the number of signal and background events.

Chapter 6

Results

The statistical interpretation of the analysis result is given in this chapter. The idea is to understand if the observed number of events found in the various SR is compatible or not with what expected from SM hypothesis. In addition to the comparison between the expected and the observed significance in the analysis, the signal contribution in the data sample is retrieved. The statistical procedure will also indicate the sensitivity of the measurement, reporting the upper limit on the VH cross section.

The statistical analysis is based on a binned likelihood function [74], written as the product of Poisson probability terms, P_s , obtained from the number of expected signal (S_i) and background (B_i) events and from the observed (N_i) data events, in each i -th signal region:

$$\mathcal{L} = \prod_i^{N_{SR}} P_s(N_i | \mu S_i(\theta) + B_i(\theta)) \quad (6.1)$$

N_{SR} is the number of SR considered, three in our case, θ are the so-called nuisance parameters (NP), that take into account the systematic uncertainties described in section 4.7. The expected signal and background yields in the Poisson terms are allowed to vary within the allowed range of the relevant systematic uncertainties, which impact on a given sample is evaluated through the fit procedure. In the Poisson term for the signal regions the μ parameter, called the *signal strength*, scales the expected signal yield, with $\mu = 0$ corresponding to no signal (*background-only hypothesis*) and $\mu = 1$ corresponding to the SM hypothesis. In this way the signal strength measures the signal contribution relative to the SM expectations. It is also called the *Parameter of Interest* (POI) since it is the relevant information that we want to extract from the statistical procedure. Aim of the procedure is to fit data with the SM expectations, leaving μ as free parameter in the fit, together with the NP. The fit result is obtained by maximizing the likelihood in equation (6.1).

Systematics uncertainties NP are taken into account and constraint by Gaussian probability density functions added to the likelihood in equation (6.1). Each θ represents a different systematic source and since one source can affect multiple signal and background rates in a correlated way, the same θ can be used everywhere to represent it. The correlation is implemented in the fit procedure where it is needed, for example most of the experimental systematics are correlated among different samples. When correlated, a single systematic source affecting more than one sample is treated as a single NP in the fit. The results presented in this chapter have been obtained using the RooStats framework [85], which is a C++ class library based in the ROOT and RooFit packages [84].

The fit is performed twice: once the result of the cut-based analysis in each SR is used as input in the fit and a combined value of significance is retrieved. Then the MVA result described in chapter 5 is used, and a shape fit on the BDT output distribution is performed in each SR

and combined with the others. In case of the shape fit, a new likelihood function is built, this time including the product over the number of BDT bins N_{bins} :

$$\mathcal{L} = \prod_i^{N_{SR}} \prod_j^{N_{bins}} P_s(N_{ij} | \mu S_{ij}(\theta) + B_{ij}(\theta)) \quad (6.2)$$

In this way each bin is treated as a SR itself, and the fit takes advantage from the different signal over background ratio in each bin to get a more stringent evaluation of the significance. Results obtained from the two fit procedures are discussed in section 6.1 and 6.2 respectively.

The modified frequentist method known as CLs [75, 76] is used to compute a 95% upper limit on the signal strength and the p_0 value. A profile likelihood ratio is used for the test statistic:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu))}{\mathcal{L}(0, \hat{\boldsymbol{\theta}}(0))} & \hat{\mu} < 0, \\ -2 \ln \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \quad (6.3)$$

Here $\hat{\boldsymbol{\theta}}^1$ in the numerator denotes the value of $\boldsymbol{\theta}$ that maximizes $\mathcal{L}$ for the given μ , i.e. it is the conditional maximum-likelihood (ML) estimator of $\boldsymbol{\theta}$. The denominator of the second case is the maximized (unconditional) likelihood function, i.e., $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are their ML estimators. The presence of nuisance parameters broadens the profile likelihood as a function of μ relative to what would result if their values were fixed.

The level of agreement between the data and hypothesized μ is quantified in term of the *p-value*. p_μ and p_b values are derived from the probability density functions distributions of $\tilde{q}_\mu$:

$$p_\mu = \int_{\tilde{q}_{\mu, obs}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\boldsymbol{\theta}}_\mu) d\tilde{q}_\mu, \quad p_b = \int_{\infty}^{\tilde{q}_{\mu, obs}} f(\tilde{q}_\mu | 0, \hat{\boldsymbol{\theta}}_0) d\tilde{q}_\mu \quad (6.4)$$

CLs is then constructed as the ratio of *p-values*

$$CLs = \frac{p_\mu}{1 - p_b} \quad (6.5)$$

The 95% upper limit on μ is the solution for $CLs = 0.05$. To compute the statistical significance of an excess over the background, the background-only *p-value* is computed from the test statistic q_0 :

$$p_0 = \int_{q_{0, obs}}^{\infty} f(q_0 | 0, \hat{\boldsymbol{\theta}}_0) dq_0 \quad (6.6)$$

The p_0 can also be converted into an equivalent significance Z , defined such that a Gaussian distributed variable found Z standard deviations above its mean has an upper-tail probability equal to p_0 . That is,

$$Z = \Phi^{-1}(1 - p_0) \quad (6.7)$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian distribution (see figure 6.1). In the Higgs searches rejection of the background-only hypothesis with a significance of at least $Z = 5$ has been considered an appropriate level to claim a discovery.

¹The bold $\boldsymbol{\theta}$ symbol denotes all the nuisance parameters $\boldsymbol{\theta} = (\theta_S, \theta_B)$

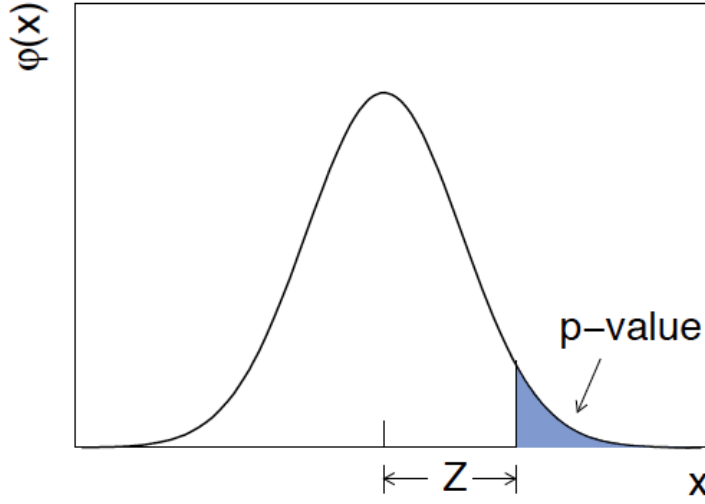


Figure 6.1: The standard normal distribution $\phi(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$ showing the relation between the significance Z and the p -value.

6.1 Cut-based analysis results

In this section the statistical interpretation of the cut-based analysis summarized in table 4.17 is given. The expected and observed event yields are taken as input for the likelihood in equation 6.1. The result of this likelihood maximization is reported in table 6.1 for each signal region separately and for the combination of the three signal regions. The value of the expected significance in each SR is fully compatible with that found with the cut-based analysis (cfr table 4.17), as a proof of the equivalence of the likelihood maximization with the simplest signal over background calculation, in case only the event yields are taken into account for the fit. The significance is indeed the ratio of the signal strength over the uncertainty of the total number of events, and can be approximated by the usual formula $\sigma = S/\sqrt{B}$. If neither an excess nor a underfluctuation in the observed number of events with respect to expectation is obtained, then the expected σ and the observed σ should return similar values. Unfortunately this is not the case, in SR3 an excess of events is observed, as shown also in the combination result. In SR1, on the contrary, a deficit is observed, resulting in a negative value of the observed significance. The negative significance and, correspondingly, the negative μ , are of course unphysical results, since would mean that some events need to be subtracted from the data sample to get the observation. However, the big uncertainty associated to the fitted μ value, stresses the fact that no conclusions can be drawn. Globally, the value of the combined signal strength is quite high, which means that there might be an excess of event. As anticipated in section 4.7, except than in SR1, the error on the signal strength is dominated by the statistical uncertainty. In fact, being μ defined as

$$\mu = \frac{N - B}{S} \quad (6.8)$$

where N is the number of total data events, S is the number of signal events and B is the number of background events, the statistical uncertainty on μ can be approximated by the following formula:

$$\left(\frac{\delta\mu}{\mu}\right)^{stat} = \frac{\sqrt{N}}{S} \quad (6.9)$$

being the statistical error on MC included in the systematic error. The systematic uncertainty is evaluated by the following two formulas:

$$\left(\frac{\delta\mu}{\mu}\right)_{B}^{syst} = -\frac{\delta B}{B} \quad (6.10)$$

$$\left(\frac{\delta\mu}{\mu}\right)_{S}^{syst} = -\frac{N-B}{S^2}\delta S \quad (6.11)$$

being δS and δB the systematic uncertainties on signal and background events respectively. This in SR3, for example, would result in $\mu = 14.5$, $\left(\frac{\delta\mu}{\mu}\right)^{syst} = 1.1$ and $\left(\frac{\delta\mu}{\mu}\right)^{stat} = 7.9$. These approximated results are in agreement with what is shown in table 6.1, where the statistical results are reported together with their statistical and systematic uncertainties.

Signal Region	Expected σ	Observed σ	μ	μ Stat. Error	μ Syst. Error
SR1	0.03	-0.30	$-16.8^{+54.5}_{-54.0}$	+27.5 / -25.2	+47.0 / -47.8
SR2	0.09	0.92	$10.7^{+12.9}_{-11.6}$	+11.6 / -10.3	+5.5 / -5.3
SR3	0.17	2.19	$14.3^{+8.9}_{-7.3}$	+8.6 / -7.0	+2.5 / -1.9
Combination	0.19	2.31	$12.9^{+7.0}_{-6.1}$	+6.4 / -5.6	+2.9 / -2.3

Table 6.1: Unblinded statistical results for the cut-based analysis. The expected and observed significance of the result is reported, together with the signal strength of the measurement. The Stat. Error and Syst. Error columns report respectively the contribution of the statistical uncertainty and the contribution of the systematic uncertainty on the μ error.

6.2 MVA analysis results

The BDT output distributions shown in chapter 5 have been used in the likelihood of equation (6.2). The analysis exploits the different signal and background shapes of the BDT output in the different SR to gain the maximum sensitivity. To this purpose the BDT output distributions are divided in five bins, with the following criteria:

1. Keep the bins with different significance to exploit the shape fit;
2. merge those bins with similar significance;
3. do not have empty bins in the data BDT distribution.

The number of bins used was limited by the available statistics. To extract the results each bin is treated as a single SR and then fitted. With five bins per SR, a total of fifteen SR have been used in the fit procedure.

The three binning intervals used are:

SR1: $\{-1.0, 0.0, 0.5, 0.8, 0.9, 1.0\}$

SR2: $\{-1.0, 0.0, 0.5, 0.8, 0.9, 1.0\}$

SR3: $\{-1.0, 0.0, 0.3, 0.5, 0.7, 1.0\}$

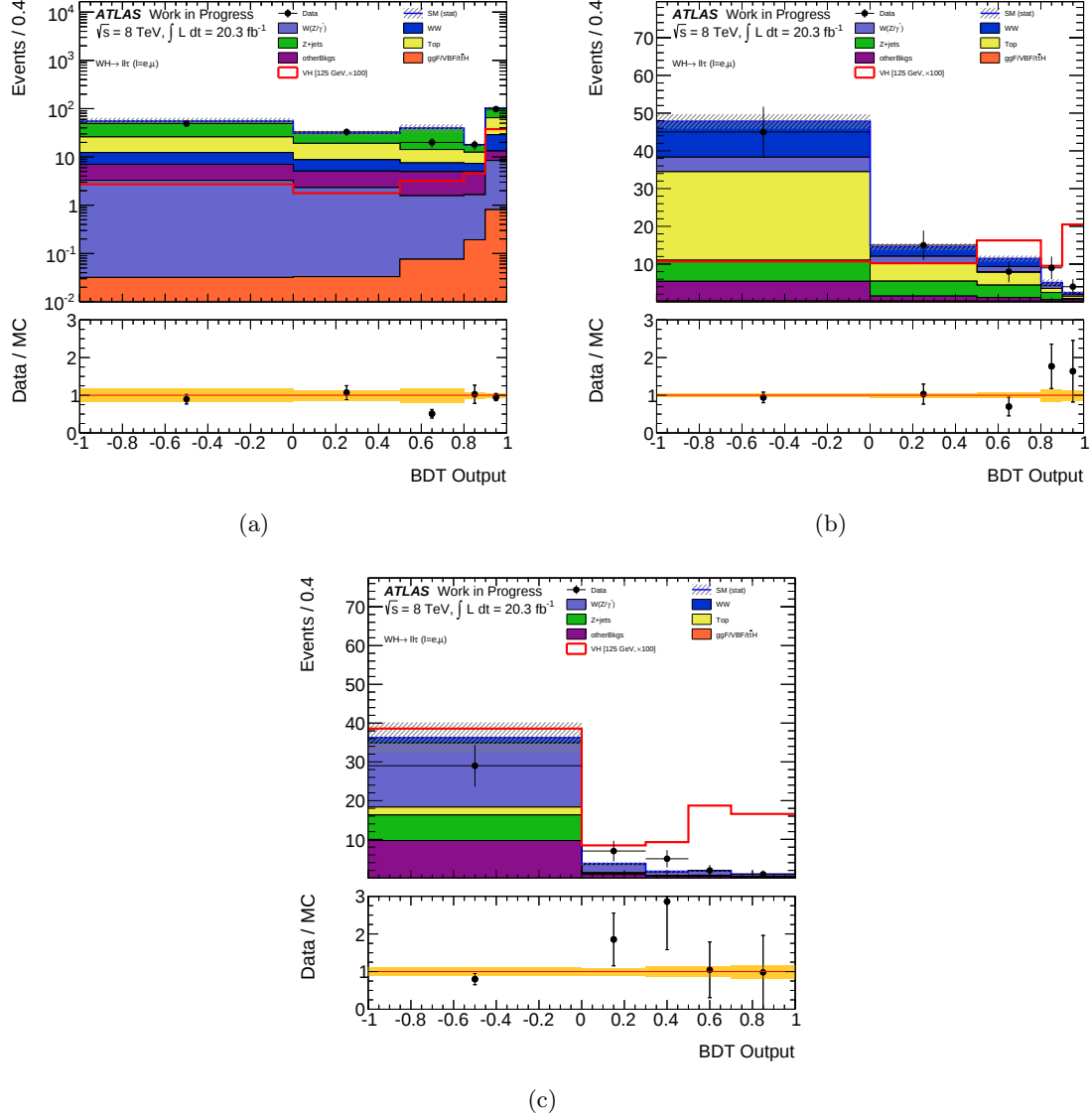


Figure 6.2: Rebinned BDT output distributions for (a) SR1, (b) SR2, (c) SR3 before the ΔR_{01} selection. In SR2 and SR3 the distribution is shown in a linear scale, to make the shape difference between signal and background more visible. In SR1 the logarithmic scale has been chosen due to the greater disparity of signal and background statistics.

Figure 6.2 shows the rebinned BDT output distributions.

The fit results are reported in table 6.2. Being ΔR_{01} the variable with the largest discrimination power, the fit should be performed before cutting on it, allowing the BDT to fully exploit the ΔR_{01} distribution. However, the fit has been also done after the ΔR_{01} selection (after the whole set of selections applied in the cut-based analysis), for a cross-check and for a direct comparison with the cut-based result in table 6.1. In terms of the expected significance the two results are compatible, the main difference is found in the observed significance, and correspondingly in the μ value; in fact, as we showed in section 4.6.3, the excess in SR3 is a consequence of the cut on ΔR_{01} variable, which cut enhances the differences in the shapes of data and MC. Comparing

Before ΔR_{01} selection					
Signal Region	Expected σ	Observed σ	μ	μ Stat. Error	μ Syst. Error
SR1	0.03	0.14	$5.8^{+41.7}_{-42.6}$	+25.3 / -23.8	+33.2 / 35.4
SR2	0.14	0.90	$6.7^{+9.3}_{-7.4}$	+8.7 / -6.9	+ 3.5 / -2.7
SR3	0.22	0.78	$4.1^{+6.5}_{-5.1}$	+5.9 / -4.7	+2.6/ -2.0
Combination	0.26	1.16	$5.1^{+5.2}_{-4.4}$	+4.7 / -4.0	+2.3 / -1.8
After ΔR_{01} selection					
Signal Region	Expected σ	Observed σ	μ	μ Stat. Error	μ Syst. Error
SR1	0.04	-0.35	$-19.0^{+53.0}_{-52.7}$	+27.1 / -24.8	+45.6 / -46.6
SR2	0.14	1.35	$10.7^{+10.1}_{-8.2}$	+9.6 / -7.8	+3.4 / -2.6
SR3	0.21	1.26	$7.7^{+8.2}_{-6.4}$	+7.9 / -6.1	+2.4 / -1.8
Combination	0.25	1.78	$8.6^{+6.2}_{-5.2}$	+5.7 / -4.9	+2.3 / -1.8

Table 6.2: Unblinded statistical results for the MVA analysis. The expected and observed significance of the result is reported, together with the signal strength of the measurement. The Stat. Error and Syst. Error columns report respectively the contribution of the statistical uncertainty and the contribution of the systematic uncertainty on the μ error.

table 6.1 with table 6.2 an overall gain of $\approx 40\%$ in significance is obtained with the MVA. The improvement is due to a better signal/background discrimination in both SR2 and SR3, while in SR1 the multivariate analysis seems not to be able to reduce the total background. The expected significance in SR1 remains unchanged with respect to the cut-based analysis result. This is also inferred from the BDT output distribution in figure 6.2(a), noting that signal and background distribution have a very similar shape, especially in the last most sensitive bins.

Together with the significance σ , the fit returns the μ parameter. The combined value found for μ in this and in the previous section and, in particular, its relative error, are a clear indication that the available statistics is too low to draw any conclusion: the mean value of μ is fully compatible with either 0 or 1, within the uncertainty. Once more, the error on μ is dominated by the statistical uncertainty, as expected.

6.3 Upper limit on VH cross section

The 95% upper limit on the VH cross section has been obtained from the test statistics in equation (6.3). The values here reported refer to the MVA analysis combined result (cfr table 6.2). The observed 95% CL upper limit on the VH production cross section is 14.5 times the SM cross section, to be compared with the 9.3 expected. The 1σ band ranges from 6.7 to 13.5 times the SM, while the 2σ ranges from 5.0 to 19.5 times the SM. The observed limit is above the expected value, reflecting the excess of events found in the analysis, but well within the 2σ band.

6.4 Combination with other analysis

This measurement alone doesn't have enough sensitivity to observe the process, but it can be combined with the fully leptonic analysis described in [67]. An estimation of the improvement that can be obtained in the VH search by adding this analysis result, being the improvement defined in terms of the gain in the expected significance, has been performed. The expected significance for the fully leptonic VH analysis is 0.82. The result obtained in this thesis and reported in table 6.2 for the shape case is 0.26. The two analyses results are combined statistically by writing down a likelihood function as that described in equation (6.2), where this time the sum runs not only on the SR (and the bins) defined in the tau analysis, but also on those defined in the fully leptonic analysis (details in [67]). A single POI is fitted, defined as the signal strength parameter μ . The statistical combination of the two analyses results in a significance of 0.87, being the combination possible since the two analyses are completely disjoint (no overlaps exist between events). This means that by adding the tau sub-channel in the VH search an improvement of $\approx 6\%$ in the expected significance can be obtained.

Chapter 7

Prospects for LHC Run 2

In spring 2015 LHC is expected to re-start with protons collisions after two years of pause, time that was necessary to consolidate the magnet interconnections to allow the accelerator to operate at the centre-of-mass energy of 13 TeV. The ATLAS detector was also upgraded, with the insertion in the ID of the new pixel layer, the IBL [81]. The new LHC phase will start ATLAS Run 2 data-taking period, in which $\sim 100 \text{ fb}^{-1}$ of data are expected.

In the long term future, from year 2023 and on, CERN plans to further upgrade the LHC instantaneous luminosity up to reach a peak value of $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ (*High Luminosity LHC*), delivering in this way a total luminosity of about 3000 fb^{-1} . To sustain the luminosity and energy increase, the ATLAS detector will undergo several upgrades: older sections will be replaced by newer and new radiation-hard detector and electronic technologies will be employed in order to improve the performance in tracking and vertex reconstruction, while keeping a reasonable trigger rate. Figure 7.1 shows the LHC upgrade schedule.

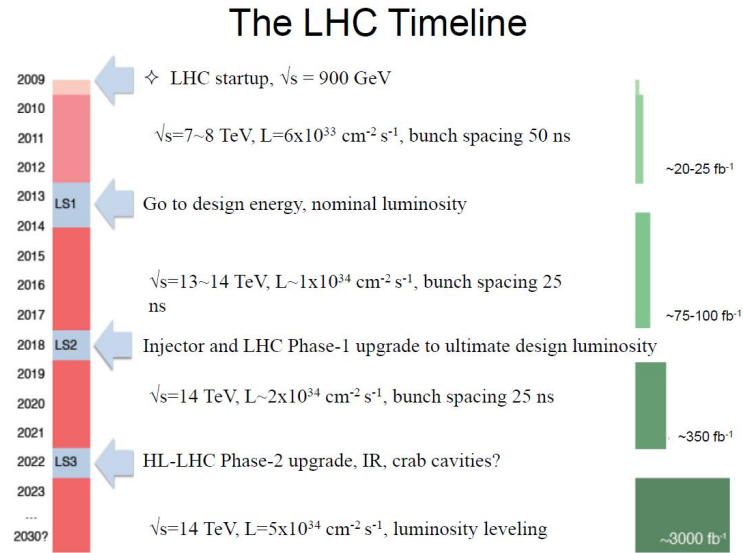


Figure 7.1: LHC upgrade schedule.

Run 2 will give the opportunity to measure more precisely rare Higgs production modes ($t\bar{t}H$, VH and VBF) and improve the measurement of the Higgs mass and couplings, thanks to the reduction of the statistical uncertainty. Aim of this section is to give a first quantitative indication about the expectations for the VH searches in ATLAS Run 2. Figure 7.2 [82] shows the evolution of main processes cross section with the centre-of-mass energy of pp system; in

table 7.1 a comparison with 8 TeV cross section is reported [83].

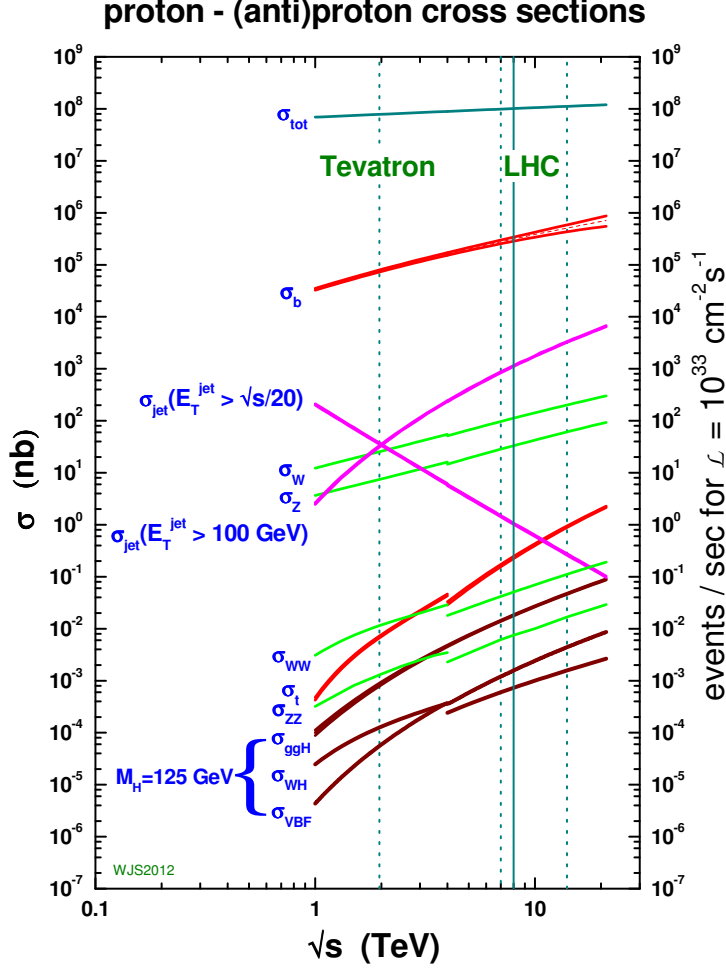


Figure 7.2: Standard Model cross sections as a function of collider energy, with 125 GeV Higgs.

As reported in figure 7.3 [82], cross section of processes dominated by gluon-gluon fusion are expected to grow faster than those processes from quark-antiquark. In fact signal and W/Z backgrounds cross section will double while the $t\bar{t}$ cross section will increase by a factor of 4. Although a detailed study of the VH process at $\sqrt{s} = 13$ TeV goes beyond the purposes of this thesis, it's worth to understand what could be the most challenging issues in Run2 for the analysis with one hadronically decaying tau and how much luminosity is needed to get the observation of the process. To this scope, the cross section of each MC sample used for the 8 TeV analysis was scaled according to the ratio values in table 7.1 and the cut-based analysis was re-run. The NF calculated with 8 TeV data for Top and Z+jets background have been applied to the new samples. The SR definition has been left unchanged. It should be noticed that by simply scaling the 8 TeV cross sections to get the expected 13 TeV values, differences in the analysis acceptance also due to differences in the event-by-event parton density functions, are not taken into account. Moreover, the pile-up is expected to increase from $\langle \mu \rangle = 21$ to $\langle \mu \rangle = 25$ possibly leading to performance degradations (e.g. tracking and vertex reconstruction). With these approximations, the expected number of signal and background events and the

Process	Cross section ratio (14TeV/8TeV) (pb)
WH	2.1
$t\bar{t}$	3.9
$Z/W + jets$	2.3
WZ/γ^*	2.3
WW	2.4
ZZ	2.2
VVV	2.1

Table 7.1: Inclusive cross section 14 TeV/8 TeV ratio for signal (WH) and main backgrounds processes.

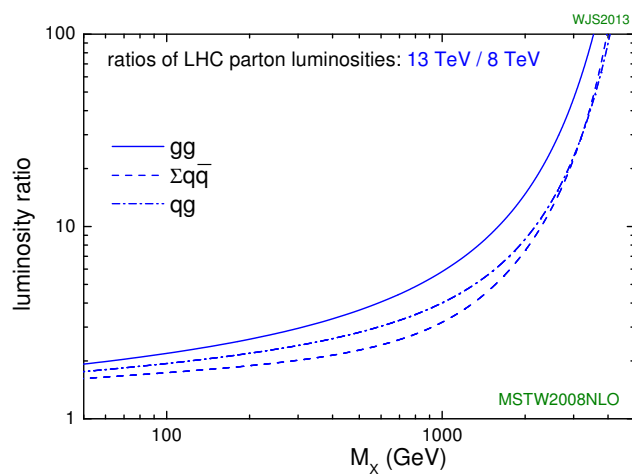


Figure 7.3: 13/8 TeV LHC parton luminosity ratios. These luminosity distributions are calculated using MSTW2008 (NLO) parton distributions [78][82].

corresponding significance obtained for the 13 TeV analysis is reported in table 7.2. The result is shown assuming different integrated luminosity: the same luminosity collected in 2012 run and used in this thesis (20.3 fb^{-1}), the expected at the end of Run2 (100 fb^{-1}), and 300 fb^{-1} .

	Signal Region	VH	Total Background	$S/\sqrt{B}$
$\mathcal{L} = 20.3 \text{ fb}^{-1}$	SR1	0.7 ± 0.0	163 ± 15	0.05 ± 0.01
	SR2	1.1 ± 0.2	77 ± 3	0.13 ± 0.02
	SR3	1.0 ± 0.2	13 ± 1	0.26 ± 0.07
$\mathcal{L} = 100 \text{ fb}^{-1}$	SR1	3.2 ± 0.2	817 ± 73	0.11 ± 0.01
	SR2	5.5 ± 0.9	382 ± 13	0.28 ± 0.05
	SR3	4.7 ± 1.2	64 ± 5	0.58 ± 0.15
$\mathcal{L} = 300 \text{ fb}^{-1}$	SR1	9.6 ± 0.7	$(24.5 \pm 2.2)10^2$	0.19 ± 0.02
	SR2	16 ± 3	$(11.4 \pm 0.4)10^2$	0.48 ± 0.08
	SR3	14 ± 4	193 ± 16	1.00 ± 0.26

Table 7.2: Prospects for 13 TeV VH, $H \rightarrow WW^*$ analysis with two leptons plus one hadronic tau in the final state. The expected cut-flow is reported for the same signal regions definition done in the 8 TeV analysis; different integrated luminosity are assumed. 100 fb^{-1} is the expected luminosity at the end of Run2.

Although the top background increases of a factor of 2 more than the signal, a gain in significance is observed in all the three SR. In fact, the top background contributes to total background for a 35% in SR1 and 56% in SR2, while it is a negligible background in the most sensitive SR3, being only the 10% of the total background. A further improvement could arise by splitting each SR in two sub-regions, one with $N_{\text{jets}} = 0$, the other with $1 \leq N_{\text{jets}} \leq 3$. The reason of the splitting is evident from table 7.3: the top over total background ratio is quite different in the two bins of N_{jets} . The two bin significances are combined in quadrature to give an estimation of the combined significance to be compared with what obtained without splitting in number of jets. No improvements are observed in SR1 with the splitting, while a gain in significance of 15% in SR2 and of 19% in SR3 is obtained.

The preliminary study done for the VH analysis in the 13 TeV scenario shows that the reduction of top background will be the main challenge. The improvement of the b-tagging algorithm performance, hopefully to be obtained thanks to the new IBL detector, could help to this scope. Moreover, an optimization of the analysis strategy, focused on the isolation of the top background with respect to minor backgrounds, e.g. with the splitting of the SR in bins of the N_{jets} variable, will result useful in increasing the significance of the analysis. Last but not least, the usage of multivariate techniques to reduce the leading background should be also considered, since its efficacy was already proved in the 8 TeV analysis. However, even with the integrated luminosity expected by the end of Run2, 100 fb^{-1} , this WH sub-channel cannot be observed alone, a combination with the leptonic WH analysis is mandatory.

	Number of jets	Top/Total Bkg	$S/\sqrt{B}$	$S/\sqrt{B}_{combined}$
SR1	Njets = 0	19%	0.03 ± 0.00	0.05 ± 0.00
	$1 \leq \text{Njets} \leq 3$	42%	0.04 ± 0.00	
SR2	Njets = 0	35%	0.06 ± 0.01	0.15 ± 0.08
	$1 \leq \text{Njets} \leq 3$	64%	0.14 ± 0.04	
SR3	Njets = 0	2.5%	0.11 ± 0.02	0.31 ± 0.24
	$1 \leq \text{Njets} \leq 3$	13%	0.29 ± 0.12	

Table 7.3: Prospects for 13 TeV VH, $H \rightarrow WW^*$ analysis with two leptons plus one hadronic tau in the final state. For each SR, the expected contribution of top background to the total background is quoted. The possibility of splitting the SR in two more sub-regions depending on the number on jets in the event is considered; the significance obtained in this case is also reported. Numbers refer to the 20.3 fb^{-1} scenario.

Conclusions

A study of the Higgs boson associated production in the $WH \rightarrow WWW^* \rightarrow l\nu l\nu\tau\nu$ channel ($l = e/\mu$) has been presented. The dataset used corresponds to the integrated luminosity of 20.3 fb^{-1} from proton-proton collisions at $\sqrt{s} = 8\text{ TeV}$, recorded by the ATLAS experiment during LHC Run1.

The analysis was performed for events containing three charged leptons, one of which being a hadronically decaying tau. Two different strategies have been exploited for the data analysis: in the first, called the "cut-based" analysis, a set of kinematical and topological selections resembling the characteristics of the signal events were applied to the data sample. Among these selections, the ΔR_{01} variable, which exploits the spins correlation between the leptons from the Higgs decay, revealed to be one of the best to discriminate signal from background events. The second analysis approach was the so-called "multivariate" analysis, which incorporates the various kinematic variables used in the cut-based approach to obtain a final single discriminating variable. In both cases, to improve the sensitivity of the measurement the signal region was divided in three more sub-regions, according to flavour and charge of the leptons in the final state. With this splitting the background composition is different in the various signal regions: SR1, for example, including same flavour and opposite charge leptons, is dominated by "Z-like" backgrounds, which can be heavily reduced with the application of some kinematical constraints, as that on the dilepton invariant mass. The analysis main backgrounds are those entering the signal region because of the presence of jet faking a hadronically decaying tau, as the $Z+jets$ and the $t\bar{t}$ background, or because of have three real leptons produced, as the WZ/γ^* background. Dedicated control samples were used to set the normalisation factors for the $Z+jets$ and the $t\bar{t}$ background processes to the data distribution.

A binned maximum likelihood fit was used to extract the signal yield as observed in data, and to compare the results obtained with the two analysis approaches with the Standard Model expectations for a Higgs boson of mass $m_H = 125\text{ GeV}$. The fit was performed twice, using the expected and observed event yields obtained with the cut-based approach, or using different bins of the MVA Boosted Decision Tree output distribution in a combined shape fit. An increase of $\sim 40\%$ in the expected significance was observed. In fact, with the binned fit the significance of the analysis increased of $\sim 40\%$ with respect to the value obtained with the cut-based approach. From the fit the signal strength parameter μ was also retrieved. The value of μ found is $\mu = 5 \pm 5$, with the relative error on μ taking into account both the statistical and the systematic uncertainties. The result is fully consistent with both the signal plus background hypothesis and the background-only hypothesis, being the μ relative error too big to draw any conclusions. In fact the small statistics, together with the low signal over background ratio of the measurement, do not allow neither to measure nor to exclude the WH process from the SM theory. The upper limit obtained on the VH cross section at 95% CL is 14.5 (9.3 expected). In the last pages of the thesis, some considerations for a possible future measurement of the WH production at the LHC were given.

The work done is a first attempt to include hadronically decaying tau in the search for the Higgs boson associated with a W vector boson. Although the available statistics proved not to

be enough to get the observation of the process, the work sets the foundations for performing the analysis in Run 2.

Appendix A

Data format and reduction

The ATLAS experiment records approximately 1 PB of data per year. Data are available for analysis in four formats:

RAW data: Raw data contain the output of the ATLAS detector, produced by real or simulated events after the HLT. It comes in the "bytestream" format as they are delivered from the detector, rather than object-oriented format. The average size of each event is approximately 1.5 MB.

ESD data: The Event Summary Data (ESD) contain the detailed output of the detector reconstruction and are produced from the raw data. They contain sufficient information to allow particle identification, track re-fitting, jet calibration etc. thus allowing for the rapid tuning of reconstruction algorithms and calibrations. An object-oriented format based on ROOT [84] objects is adopted, and the typical event size is 1 MB.

AOD data: The analysis object data (AOD) is a summary of the reconstructed event, and contains sufficient information for common analyses. The AOD is also stored in ROOT format and the nominal event size is of the order of 100 KB.

D3PD data: The derived physics data (D3PD) are a summary of the reconstructed event, and contain sufficient information for common analyses. D3PD contain small subset derived from the AOD / ESD, specific for an analysis or performance group. More than one derivation is possible, in which the data is reduced by removing unnecessary physics blocks (e.g. jets, photons, etc...). User-data can be added in the process, and in the final stage of derivation a flat ROOT ntuple can be produced.

In this thesis a set of D3PD ntuples has been used, called COMMON ntuples (*NTUP.COMMON*). The COMMON ntuples are an ntuple format produced in 2014 with the aim of creating a D3PD that is usable by most ATLAS analysis group. For this reason in a COMMON ntuple more informations with respect to the usual D3PD are included, resulting in a total size of several hundreds of TB for the only 2012 dataset. This huge size required a slimmed procedure to be applied to these ntuples, in order to reduce the data size and allow for the data recording on the Rome disks. In the COMMON ntuples all particle objects are stored in "containers"; this allowed to iterate over all leptons and hadronic taus in a given event, to select and register on the new slimmed ntuple only those events interesting for the WH analysis. The slimming code was indeed developed to select events containing at least one lepton plus one hadronic tau. The lepton and tau identification at this stage is not the final ID used to select the objects in the analysis, but a looser criteria is applied. In this way the slimming allows to filter out all the unwanted events while keeping all those events which are likely to be signal events. The slimming

procedure applied reduced the NTUP_COMMON size up to few TB. The new slimmed ntuples were then trasferred on Roma Tre disks becoming accessible to the purposes of this thesis.

Appendix B

Tau-Jet fake rate

As shown in paragraph 4.1.2, most of the background processes that mimic our signal are due to a jet faking a hadronic tau. This is the case of the Z +jets contribution in which a jet from the hadronization of a quark or a gluon is misidentified as a tau. The purpose of this appendix is to evaluate the tau-jet fake rate in MC and in data, selecting $Z \rightarrow \mu\mu$ events. The tau fake rate is defined as the rate of jets misidentified as hadronically decaying tau leptons by the tau reconstruction algorithms. In contrast with electrons or muons a tau passing through the detector doesn't leave an unambiguous signature, but some energy deposits in the calorimeter and one or more tracks in the inner detector. The challenge when identifying hadronic tau decays is that their signatures in the detector are very similar to quark- or gluon-initiated jets (called *jets* from now on). By the way some topological and kinematical differences between a tau and a jet exist. For example the tau shower shape in the calorimeter is usually narrower than the shower originated from a jet and the number of charged tracks in tau cone is fixed to one or three. All these informations are then combined together by an MVA technique to discriminate taus from jets [68]. Even with a multivariate approach there's still a appreciable probability of such a misidentification.

B.1 Fake rate measurement

The first step in the tau-jet fake rate evaluation was the choice of MC and data sample to be analyzed. To this purpose a pure dilepton sample has to be selected in order to minimize the probability of accounting for lepton reconstruction efficiency or lepton charge flip in the fake rate. A $Z \rightarrow \mu\mu + jets$ sample was then used, since $Z \rightarrow \mu\mu$ events should not have real taus and the two muons in the final state can be detected with high efficiency and purity. The dilepton selection then consists in:

- Two opposite charge muons
- $|m_Z - m_{\mu\mu}| < 25 \text{ GeV}$

In case of MC sample these requirements are partially redundant due to a similar filter applied at the event generation level. On the contrary they are mandatory for the data sample selection and so to compare the fake rate from MC with the rate measured in data. The fake rate is defined as follows:

$$f_r = \frac{\text{Number of } (\mu^+\mu^- + 1\tau) \text{ events}}{\text{Number of } (\mu^+\mu^-) \text{ events}} \quad (\text{B.1.1})$$

where the numerator is the number of events with two opposite charge muons plus one hadronic tau and the denominator is the number of events with two opposite charge muons, regardless if

there are taus in the event and how many they are. Leptons selection is exactly the same as in the main analysis.

Since we expect fake taus to come from reconstructed jets we also expect this ratio to increase with the number of jets in the event. The more jets there are in the event, the higher will be the probability of misidentify them as taus.

The ratio in equation B.1.1 has then been evaluated as a function of the number of jets in the event. Figure B.1 shows the result obtained with the MC $Z \rightarrow \mu\mu + jets$ sample (blue line). The plot shows a rate of approximately 3% in 1-jet events slightly increasing when more jets

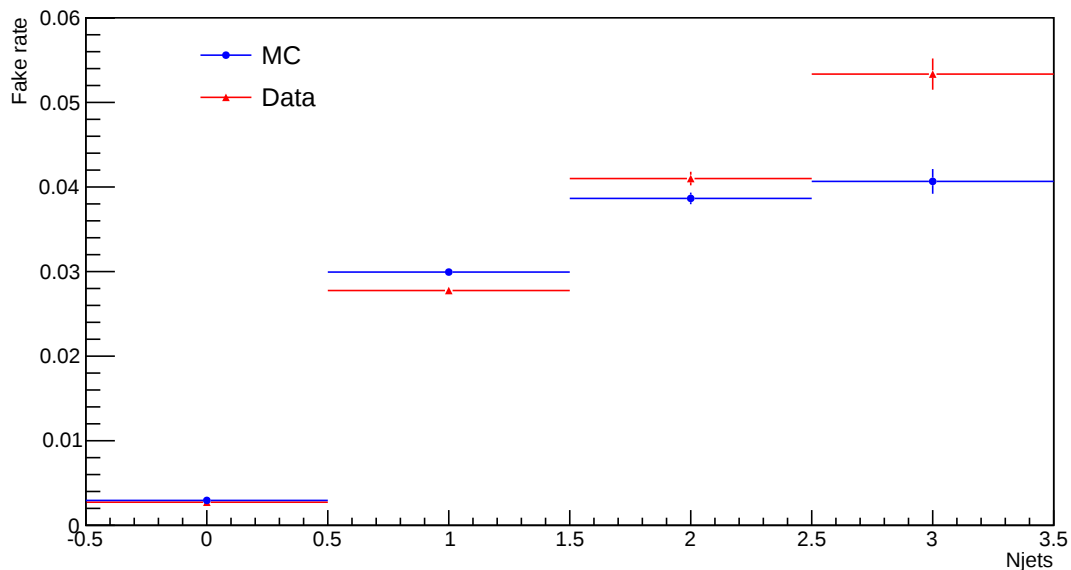


Figure B.1: Tau-jet fake probability in MC (blue line) and data (red line), as a function of the number of reconstructed jets in the event.

are in the event. Only the statistical uncertainty is applied. It should be noticed that, although in events with zero reconstructed jets we would expect a fake rate exactly equals to zero, the first bin in figure B.1 shows a fake rate of approximately 0.3%. This it is not worriesome since the jets acceptance is different from tau acceptance; in the analysis jets are reconstructed with a transverse momentum greater than 25 GeV, while taus with $p_T > 15$ GeV. This means that the first bin shows the fraction of fake taus with $15 < p_T < 25$ GeV.

The same measurement has been performed on the filtered data sample. Result is shown in figure B.1 (red line). Comparing the two distributions a good agreement is found between data and MC.

Appendix C

Spare Cutflows

Control Region	Selection	Data/MC	VH	Total background	T_{op}	$Z + j\text{ets}$	WZ/γ^*	Purity
Top CR	Njets	0.94 ± 0.04	$(5.0 \pm 0.3)10^{-1}$	545 ± 8	400 ± 2	108.8 ± 7.4	10.1 ± 0.3	0.73 ± 0.01
	Top veto	1.01 ± 0.05	$(2.8 \pm 0.3)10^{-1}$	403 ± 6	372.0 ± 2.3	23 ± 6	2.9 ± 0.1	0.92 ± 0.02
	MET	0.99 ± 0.06	$(2.0 \pm 0.2)10^{-1}$	332 ± 3	320 ± 2	5.7 ± 1.3	2.2 ± 0.1	0.96 ± 0.01
	ΔR_{01}	0.94 ± 0.09	$(1.2 \pm 0.1)10^{-1}$	107.6 ± 1.4	104.3 ± 1.2	1.1 ± 0.5	0.9 ± 0.1	0.94 ± 0.09
Z+jets CR	Njets	0.72 ± 0.00	6.1 ± 0.3	$(186.0 \pm 0.9)10^3$	1187 ± 4	$(180.2 \pm 0.9)10^3$	550 ± 3	0.97 ± 0.01
	Top veto	0.70 ± 0.00	4.4 ± 0.1	$(172.0 \pm 0.8)10^3$	134.9 ± 1.4	$(167.3 \pm 0.8)10^3$	419 ± 3	0.98 ± 0.01
	Z-veto	0.68 ± 0.01	1.7 ± 0.2	$(26.3 \pm 0.3)10^3$	112.9 ± 1.3	$(26.3 \pm 0.3)10^3$	70.3 ± 1.1	0.90 ± 0.01
	MET	0.68 ± 0.01	1.4 ± 0.2	$(26.0 \pm 0.3)10^3$	54.2 ± 0.9	$(23.5 \pm 0.3)10^3$	59 ± 1	0.90 ± 0.01
	ΔR_{01}	0.67 ± 0.02	0.5 ± 0.1	$(5.7 \pm 0.1)10^3$	12.2 ± 0.4	$(5.2 \pm 0.1)10^3$	13.6 ± 0.5	0.92 ± 0.02
WZ/γ^* CR1	Njets	0.72 ± 0.00	6.1 ± 0.3	$(185.9 \pm 0.9)10^3$	1187 ± 4	$(180.2 \pm 0.9)10^3$	549 ± 3	0.00 ± 0.00
	Top veto	0.70 ± 0.00	4.4 ± 0.3	$(171.6 \pm 0.8)10^3$	134.9 ± 1.4	$(167.3 \pm 0.8)10^3$	419 ± 3	0.00 ± 0.00
	Z-veto	0.70 ± 0.00	2.7 ± 0.2	$(145.31 \pm 0.8)10^3$	22.0 ± 0.6	$(143.7 \pm 0.8)10^3$	348.6 ± 2.4	0.00 ± 0.00
	MET	0.85 ± 0.07	0.6 ± 0.9	$(6.9 \pm 0.5)10^2$	11.9 ± 0.4	$(5.9 \pm 0.5)10^2$	63.2 ± 0.9	0.09 ± 0.01
	ΔR_{01}	0.75 ± 0.14	$(2.0 \pm 0.3)10^{-1}$	$(1.4 \pm 0.2)10^2$	1.7 ± 0.2	$(1.2 \pm 0.2)10^2$	15.4 ± 0.4	0.11 ± 0.02
WZ/γ^* CR2	Njets	0.75 ± 0.09	1.2 ± 0.2	$(2.9 \pm 0.3)10^2$	14.1 ± 0.5	$(20.3 \pm 0.3)10^2$	35.4 ± 0.7	0.12 ± 0.01
	Top veto	0.68 ± 0.09	0.7 ± 0.2	$(2.4 \pm 0.3)10^2$	2.3 ± 0.2	$(1.8 \pm 0.3)10^2$	25.9 ± 0.6	0.11 ± 0.01
	MET	0.82 ± 0.19	0.37 ± 0.14	$(0.5 \pm 0.1)10^2$	1.7 ± 0.2	20 ± 8	16.2 ± 0.4	0.33 ± 0.06

Table C.1: Control regions cut-flow: main backgrounds events breakdown and purity. The "Total background" column includes the contribution of all the non-leading backgrounds. The number of events is reported with its statistical uncertainty. Numbers smaller than 0.01 are rounded as 0 in this table.

Signal Region	Selection	OtherBkgs	VVV	ZZ*	W + jets	W γ	Z γ
SR1	3 leptons	$(15.2 \pm 0.2)10^3$	11.0 ± 0.2	618 ± 4	$(9.6 \pm 0.2)10^3$	190 ± 6	$(47.8 \pm 0.2)10^2$
	Preselection	$(41.1 \pm 0.2)10^2$	8.6 ± 0.2	478 ± 4	23 ± 7	31.0 ± 2.2	$(35.7 \pm 0.2)10^2$
	$p_T(\tau)$	$(41.1 \pm 0.2)10^2$	8.6 ± 0.2	478 ± 4	23 ± 7	31.0 ± 2.2	$(35.7 \pm 0.2)10^2$
	Flavour splitting	$(40.3 \pm 0.2)10^2$	3.3 ± 0.1	460 ± 4	17 ± 7	8.5 ± 1.1	$(35.5 \pm 0.2)10^2$
	Njets	$(40.0 \pm 0.2)10^2$	3.3 ± 0.1	453 ± 4	17.0 ± 6.7	5.8 ± 0.2	$(35.2 \pm 0.2)10^2$
	Top veto	$(36.0 \pm 0.2)10^2$	2.5 ± 0.1	372 ± 3	16.7 ± 6.8	7.0 ± 1.0	$(32 \pm 0.2)10^2$
	Z-veto	$(24.2 \pm 0.2)10^2$	1.9 ± 0.1	105 ± 2	16.6 ± 6.8	6.0 ± 0.9	$(22.9 \pm 0.2)10^2$
	MET	18 ± 2	1.1 ± 0.1	3.4 ± 0.2	1.0 ± 1.0	1.6 ± 0.5	11 ± 1
	ΔR_{01}	4.2 ± 1.0	0.2 ± 0.0	41.0 ± 0.1	1.0 ± 1.0	0.6 ± 0.3	1.5 ± 0.4
SR2	$p_T(\tau)$	$(10.6 \pm 0.1)10^2$	6.4 ± 0.2	189 ± 2	10 ± 4	13.1 ± 1.4	841 ± 9
	Flavour splitting	16 ± 2	2.1 ± 0.1	3.3 ± 0.2	1.4 ± 1.2	5.0 ± 0.9	4.3 ± 0.9
	Njets	16 ± 2	2.1 ± 0.1	3.3 ± 0.2	1.4 ± 1.2	5.0 ± 0.9	4.3 ± 0.9
	Top veto	14 ± 2	1.7 ± 0.1	3.0 ± 0.2	2 ± 1	4.4 ± 0.8	3.6 ± 0.8
	MET	8 ± 2	1.3 ± 0.1	1.1 ± 0.1	1.6 ± 1.4	3.2 ± 0.7	1.2 ± 0.5
	ΔR_{01}	2.8 ± 0.6	0.4 ± 0.0	0.5 ± 0.1	0.0 ± 0.0	1.4 ± 0.5	0.5 ± 0.3
SR3	$p_T(\tau)$	$(41.1 \pm 0.2)10^2$	8.6 ± 0.2	478 ± 4	23 ± 7	31.0 ± 2.2	$(35.7 \pm 0.2)10^2$
	Flavour splitting	38 ± 3	2.5 ± 0.1	10.8 ± 0.5	4 ± 2	10.6 ± 1.2	10.2 ± 1.2
	Njets	38 ± 3	2.5 ± 0.1	10.7 ± 0.5	4 ± 2	10.6 ± 1.2	10.2 ± 1.2
	Top veto	34 ± 3	2.0 ± 0.1	9.6 ± 0.5	4 ± 2	10 ± 1	9.0 ± 1.2
	MET	12 ± 2	1.5 ± 0.1	2.3 ± 0.1	3.1 ± 1.4	4 ± 1	0.9 ± 0.4
	ΔR_{01}	1.6 ± 0.4	0.3 ± 0.0	0.4 ± 0.0	0.0 ± 0.0	1.0 ± 0.4	0.0 ± 0.0

Table C.2: Signal regions unblinded cutflow displaying the breakdown of the "Other Bkgs" category. At each stage of the cutflow the number of background events for the main background processes is reported. The number of events is reported with its statistical uncertainty. Numbers smaller than 0.1 are rounded as 0 in this table.

Bibliography

- [1] S. L. Glashow, *Partial-symmetry of weak interactions*, Nucl. Phys. 22 no. 4, (1961) 579.
- [2] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak interactions with Lepton-Hadron Symmetry*, Phys. Rev. D2 (1970) 1285.
- [3] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. 19, (1967) 1264.
- [4] A. Salam, *Weak and electromagnetic interactions*, in *Elementary particle theory: relativistic groups and analyticity*, NN. Svartholm, ed p.367, Almqvist & Wiksell, 1968, Proceeding of the eight Nobel Symposium.
- [5] G. 't Hooft and M. Veltman, *Regularization and Renormalization of Gauge fields*, Nucl. Phys. B44, (1972) 189.
- [6] M.H. Seymour, *Quantum Chromodynamics*, arXiv:hep-ph/0505192v2 (2010).
- [7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13 (1964) 321-323**.
- [8] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12 (1964) 132-133**.
- [9] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145 (1966) 1156-1163**.
- [10] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett, **13 (1964) 585-587**.
- [11] Goldstone J. (1961) *Nuovo Cimento* **19 154**.
- [12] F. Mandl, G. Shaw, *Quantum Field Theory*, WILEY (2010).
- [13] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B 716 (2012) 1–29**, [[arXiv:1207.7214](#)].
- [14] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B 716 (2012) 30–61**, [[arXiv:1207.7235](#)].
- [15] <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>.
- [16] S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, arXiv:1101.0593 [hep-ph] and references therein.
- [17] S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka, et al., *Handbook of LHC Higgs 1851 Cross Sections: 2. Differential Distributions*, arXiv:1201.3084 [hep-ph] and references therein.

-
- [18] S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph] and references therein.
- [19] LEP Collaborations, *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B 565** (2003) 61–75, arXiv:hep-ex/0306033.
- [20] CDF and D0 Collaborations, *Higgs Boson Studies at Tevatron*, arXiv:1303.6346v3 [hep-ex] (2013).
- [21] ATLAS Collaboration, *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, ATLAS-CONF-2013-079 (2013), [arXiv:1409.6212].
- [22] ATLAS Collaboration, *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*, ATLAS-CONF-2013-108 (2013).
- [23] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$ channels in pp collisions at a center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. **D. 90**, 052004 (2014).
- [24] ATLAS Collaboration, *Observation and measurement of Higgs boson decays to WW^* with ATLAS at the LHC*, ATLAS-CONF-2014-060 (2014).
- [25] Olaf Behnke, et al., *Data Analysis in High Energy Physics, A Practical Guide to Statistical Methods*, 2013 WILEY-VCH Verlag GmbH & Co. KGaA.
- [26] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector*, Phys. Lett. B 726 (2013), pp. 88–119.
- [27] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. B 726 (2013), pp. 120–144.
- [28] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock, *LHC Design Report*, CERN, Geneva, (2004).
- [29] T. Han and S. Willenbrock, *QCD correction to the $pp \rightarrow WH$ and ZH total cross-sections*, Phys. Lett. **B 273** (1991) 167–172.
- [30] M. L. Ciccolini, S. Dittmaier, and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. **D 68** (2003) 073003.
- [31] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. **B 579** (2004) 149–156, [arXiv:hep-ph/0307206v1].
- [32] S. Van Der Meer, ISR-PO/68-31, KEK 68-64.
- [33] ALICE Collaboration, *ALICE physics performance: Technical Design Report*, CERN-LHCC-2005-030 (2005).
- [34] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report*, CERN-LHCC-99-14/15 (1999).
- [35] CMS Collaboration, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, CERN-LHCC-2006-001 (2006).

- [36] CMS Collaboration, *CMS Physics: Technical Design Report Volume 2: Physics Performance*, CERN-LHCC-2006-021 (2006).
- [37] LHCb Collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007 (2000).
- [38] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its performance in 2012*, ATLAS-CONF-2014-047 (2014).
- [39] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-up LHC Environment*, ATLAS-CONF-2012-042 (2012).
- [40] ATLAS Collaboration, *Performance of the ATLAS Detector using First Collision Data*, JHEP 1009:056 (2010) [arXiv:1005.5254].
- [41] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions $\sqrt{s} = 7\text{ TeV}$ in the ATLAS experiment*, ATLAS-CONF-2010-069 (2010).
- [42] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012), [arXiv:1110.3174].
- [43] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, ATLAS-CONF-2014-032 (2014).
- [44] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, (2014) [arXiv:1407.3935].
- [45] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$* , ATLAS-CONF-2010-053 (2010).
- [46] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$* , Eur. Phys. J. **C73** (2013) 2304.
- [47] ATLAS Collaboration, *Selection of jets produced in proton-proton collision with the ATLAS detector using 2011 data*, ATLAS-CONF-2012-020.
- [48] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, JHEP 0804 (2008) 063.
- [49] ATLAS Collaboration, *Calibration of b-tagging using dilepton top pair events in a combinatorial likelihood approach with the ATLAS experiment*, ATLAS-CONF-2014-004 (2014).
- [50] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046 (2014).
- [51] ATLAS Collaboration, *Identification of Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, ATLAS-CONF-2013-064 (2013).
- [52] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$ with ATLAS*, Eur. Phys. J. **C 72** (2012).
- [53] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at $\sqrt{s} = 8\text{ TeV}$* , ATLAS-CONF-2013-082 (2013).

-
- [54] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852-867.
- [55] A. Bredenstein, A. Denner, S. Dittmaier, A. Mück, and M. M. Weber, *Prophecy4f: A Monte Carlo generator for a proper description of the Higgs decay into 4 fermions*, <http://omnibus.uni-freiburg.de/sd565/programs/prophecy4f/prophecy4f.html>, (2010).
- [56] P. Nason, *Recent Developments in POWHEG*, PoS **RADCOR2009** (2010) 018, arXiv:1001.2747 [hep-ph].
- [57] M. L. Mangano et al., *ALPGEN, a generator for hard multi-parton processes in hadronic collisions*, JHEP **0307** (2003) 001.
- [58] G. Corcella et al., *HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including super-symmetric processes)*, JHEP **0101** (2001) 010.
- [59] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photo-production at HERA*, Z. Phys. **C72** (1996) 637-646.
- [60] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007.
- [61] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5: Going Beyond*, JHEP **1106** (2011) 128.
- [62] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:0405247 [hep-ph].
- [63] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250-303.
- [64] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823-874.
- [65] H. - L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024.
- [66] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004.
- [67] ATLAS Collaboration, *Search for associated production of the Higgs boson in the $WH \rightarrow WWW^{(*)} \rightarrow l\nu l\nu l\nu$ and $ZH \rightarrow ZWW^{(*)} \rightarrow ll\nu l\nu$ channels with the ATLAS detector at the LHC*, ATLAS-CONF-2013-075 (2013).
- [68] ATLAS Collaboration, *Identification of Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, ATLAS-CONF-2013-064.
- [69] A. Roodman, *Blind Analysis in Particle Physics*. PHYSTAT2003, SLAC, Stanford, 1784 California, September 8-11, 2003.
- [70] Schapire, R. (1990) *The strength of weak learnability*. Mach. Learn., 5, 197.
- [71] Freund, Y. (1995) *Boosting a weak learning algorithm by majority*. Inform. Comput., 121, 256.
- [72] A. Hoecker et al., *TMVA 4 - Toolkit for Multivariate Data Analysis with ROOT: User Guide*, 2002 arXiv:0703039v5 [physics].

- [73] Justel, A., Peña, D. and Zamar, R. (1997) *A multivariate Kolmogorov-Smirnov test of goodness of fit*, Statistics & Probability Letters, **35(3)**, **251-259**.
- [74] ATLAS Collaboration, *Measurements of the properties of the Higgs-like boson in the $W^{(*)} \rightarrow l\nu l\nu$ decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2013-030 (2013).
- [75] A. L. Read, *Presentation of search results: the CLs technique*, J. Phys. G **28** (2002) **2963**.
- [76] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) **1554**.
- [77] M. Corradi, *Inclusion of systematic uncertainties in upper limits and hypothesis tests*, CERN-OPEN-2000-213 (2000).
- [78] A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63:189-285** (2009).
- [79] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ using the ATLAS detector at the LHC*, Eur. Phys. J. **C 73** (2013) **2518**, [arXiv:1302.4393 \[hep-ex\]](#).
- [80] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C 73** (2013) **2306**, [arXiv:1210.6210 \[hep-ex\]](#).
- [81] ATLAS Collaboration, *ATLAS Insertable B-Layer technical design report*, CERN-LHCC-2010-013 (2010).
- [82] W.J. Stirling, private communication.
- [83] J.M. Campbell, R. Keith Ellis and C. Williams, *Vector boson pair production at the LHC*, FERMILAB-pub-11-182-T, [arXiv:1105.0020v1 \[hep-ph\]](#) (2011).
- [84] The ROOT Team, *ROOT, An Object-Oriented Data Analysis Framework*, Users Guide 5.26 (2009).
- [85] L. Moneta, K. Belasco, K. Cranmer et al, *The RooStats Project*, proceedings of ACAT, Jaipur, India (2010), [arXiv:1009.1003](#).