



Thèse préparée à l'Universitè Pierre et Marie Curie

École doctorale des Sciences de la Terre et de l'environnement et Physique de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE)

Étude des propriétés du boson de Higgs dans les désintégrations en deux quark b dans l'experience ATLAS au LHC

Par Carlo Enrico PANDINI

Thèse de doctorat de physique des particules

Dirigée par Giovanni CALDERINI

Presentée et soutenue publiquement le 21 octobre 2016

Devant un jury composé de:

M.	SIROIS	Yves	*Rapporteur*
Mme.	GRAY	Heather	*Rapporteur*
M.	HAMEL DE MONCHENAULT	Gautier	
M.	FORTE	Stefano	
M.	KADO	Marumi	
M.	BERNARDI	Gregorio	
M.	CALDERINI	Giovanni	

Study of the properties of the Higgs boson in the decay channel to a b-quark pair with the ATLAS experiment at the LHC

Abstract

This thesis focuses on two separate research axes: first, the search for the Standard Model (SM) Higgs boson H in the associated production mode (VH) decaying to b-quark pairs, with data collected at $\sqrt{s}$ =7 TeV and 8 TeV during the LHC Run-1, and separately at $\sqrt{s}$ =13 TeV during the LHC Run-2 (until July 2016). Furthermore, the search for a high-mass CP-odd pseudoscalar A boson decaying to Zh pairs, with the h being a SM Higgs-like boson decaying to bottom-quark pairs, with data collected at $\sqrt{s}$ =13 TeV during the LHC Run-2 (until December 2015).

The event selection, the analysis techniques, the main backgrounds and their modeling, the signal properties, and the statistical discrimination between the signal and background in data are discussed for all analyses presented in this document, as well as the interpretation of the results.

In the SM $VH(b\bar{b})$ channel no clear excess over the background prediction is observed in the analyses of the LHC Run-1 and Run-2 datasets. From the Run-1 data analysis the observed (expected) significance under the hypothesis of a SM Higgs boson can be quoted at $\sigma_{obs(exp)} = 1.4(2.6)$, with a ratio of the measured signal yield to the Standard Model expectation (signal strength) of $\mu = 0.51 \pm 0.31(\text{stat.}) \pm 0.24(\text{syst.})$. The significance of the Run-2 data analysis is $\sigma_{obs(exp)} = 0.42(1.94)$, with a signal strength of $\mu = 0.21^{+0.36}_{-0.35}(\text{stat.}) \pm 0.36(\text{syst.})$, corresponding to a 95% observed upper limit on the cross section times branching ratio $pp \to (W/Z)(H \to b\bar{b})$ of 1.2 times the SM prediction, compared to the limit of $1.0^{+0.4}_{-0.3}$ expected in the absence of a SM Higgs signal.

In the search for a CP-odd pseudoscalar A boson two upward deviations from the background-only hypothesis are observed, for masses of the A boson of $m_A = 260$ GeV and 440 GeV, with a local significance of approximately 2σ. No other significant excess over the background prediction is found, therefore 95% upper limits are set on the production cross sections times branching fractions, $\sigma(pp \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$, in the range of [4.0, 0.017] pb ([6.9, 0.026] pb) for $m_A = [220, 2000]$ GeV assuming gluon-fusion (b-quark-associated) exclusive production, which are furthermore interpreted in the context of 2HMDs as constraints in the $\tan(\beta)$ and $\cos(\beta - \alpha)$ parameter plane.

Keywords

LHC, ATLAS experiment, Higgs boson, VH associated production, bottom-quark, 2HDM

Contents

1 | Introduction

The Large Hadron Collider (LHC) [1] at CERN accelerates and collides proton bunches at unprecedented energies, with the highest instantaneous luminosities achieved in hadron colliders so far. The proton-proton collisions provide access to an extended kinematic regime, which is probed by two multi-purpose experiments, ATLAS [2] and CMS [3], able to test Standard Model (SM) predictions for elementary particles and their interactions, and to explore new physics scenarios beyond the SM (BSM).

In the Standard Model of particle physics the Higgs mechanism is required to introduce mass terms for all massive elementary particles, without spoiling the fundamental properties of the model, and it predicts the existence of a new scalar massive state: the Higgs particle H. The theory of the mass-generating mechanism was independently hypothesized by Brout, Englert and Higgs in 1964 [4, 5], and it has been experimentally confirmed with the discovery of the predicted Higgs boson, announced by the ATLAS and CMS experiments on July 4, 2012 [6, 7]. At the time of this announcement only Higgs couplings to bosons could directly be proven.

To probe the structure of the Higgs sector of the SM it is critical to investigate its coupling to fermions as well: the $H \to b\bar{b}$ channel allows to directly test the Higgs coupling to bottom-type fermions, and it provides the largest branching fraction of 58% for a Higgs boson with mass of 125 GeV, making it an important actor in the study of the total Higgs decay width under general assumptions [8].
The overwhelming background from QCD jet production makes the inclusive search for $H \to b\bar{b}$ decays extremely challenging at hadron colliders: the experimental signature of the different Higgs production mechanisms is the best handle to disentangle the signal from the backgrounds. The associated production mechanism (VH) provides the highest sensitivity thanks to the leptonic decays of the vector boson, while other modes, such as vector boson fusion (VBF) and top-quark pair associated production ($t\bar{t}H$), have been studied as well by the ATLAS and CMS Collaborations.

The first evidence of the Higgs decay to bottom-quark pairs have been reported by the D0 and CDF Collaborations [9], with the full dataset collected during the Run-2 of the Tevatron proton-antiproton collider, corresponding to a maximum local (global) significance of 3.3 (3.1) standard deviations for an Higgs mass of 135 GeV, and an observed significance of 2.8 standard deviations for an Higgs mass of 125 GeV.

The investigation of this Higgs process at the LHC aims at providing a definite and solid answer about the nature of the Higgs coupling to bottom-quarks.
The CMS Collaboration reported a moderate excess of events from the analysis of the LHC Run-1 data, with an observed significance of 2.1 standard deviations [10], while no result from the LHC Run-2 data has been presented yet. A combination of the ATLAS and CMS

Run-1 analyses has been performed [11], leading to an excess in the $H \to b\bar{b}$ decay channel of 2.6 standard deviations (from both VH and $t\bar{t}H$ production modes combined).
In this thesis I describe the search for the $VH(b\bar{b})$ process with data collected by the ATLAS detector during the LHC Run-1 and Run-2: while the analysis of the LHC Run-1 data represents the 'legacy result' for this search from the ATLAS Collaboration, the Run-2 analysis is the first investigation of Higgs to bottom-quark pair decays in associated production mode performed at a center of mass energy of $\sqrt{s} = 13$ TeV.

The $VH(b\bar{b})$ final state is also a powerful probe in the quest for BSM physics. At the center of mass energies reached by the LHC, massive resonances can be produced by proton-proton collisions, and several BSM scenarios [12] predict the existence of CP-odd neutral bosons A with masses above $\sim$ 200 GeV. The Zh decay channel, where h is a SM-like Higgs boson decaying to bottom-quark pairs, has a good sensitivity to this type of signal, thanks to the large $h \to b\bar{b}$ branching fraction coupled with the experimental signature from leptonic V boson decays.
The $A \to Zh \to l^+l^-b\bar{b}$ channel has already been investigated by the ATLAS and CMS Collaborations with data collected during the LHC Run-1 [13, 14]: in this thesis I present the results of this search, extended to include $Z \to \nu\bar{\nu}$ decay modes, and improved with the use of jet substructure techniques, performed with data collected at $\sqrt{s} = 13$ TeV during the LHC Run-2. The search for a high-mass resonance strongly benefits from the increase in center of mass energy, from 8 to 13 TeV, and it represents one of the top candidates for the analysis of the LHC Run-2 data.

Structure of the thesis and personal contributions The thesis is divided into nine Chapters, including this Introduction. Chapter 2 is an overview of the fundamental theory of particle physics, with emphasis on the construction of the Higgs scalar sector and its main features. This Chapter contains a short outline of the BSM physics framework used for the interpretation of the search for a CP-odd A boson documented in the thesis.
Chapter 3 provides a general description of the LHC, introducing some fundamental concepts of accelerator physics, and contains a detailed description of the ATLAS detector with its several sub-systems. A short outline of the upgrades introduced between the LHC Run-1 and Run-2 is provided for the main components of the ATLAS detector.
In Chapter 4 I describe the main statistical methodologies developed for the analysis of LHC data, and applied to obtain the results presented in this thesis. The Chapter is divided into two main parts: first I cover the fundamental principles of the statistical approach, while in the second part I discuss some specific techniques applied in the analyses described in the following Chapters. The results presented in this thesis heavily rely on this statistical methods, which are a core aspect of the investigation of LHC data.
Chapter 5 contains the details of the object identification and reconstruction with the ATLAS detector, for all the main objects used in the analyses. The final part of the Chapter covers the specific techniques used to improve the event reconstruction for $VH(b\bar{b})$ final states. When relevant, the differences between the strategies adopted for the identification or the reconstruction of objects and events in the analysis of LHC Run-1 and Run-2 data are outlined.
In Chapter 6 the search for $VH(b\bar{b})$ processes with data collected during the LHC Run-1 is described. I have been involved in this search since December 2013, mainly focusing on the analysis of the $ZH(\nu\bar{\nu}b\bar{b})$ channel. The most relevant contributions I gave to this results include the estimate of modeling systematic uncertainties (specifically for the diboson background), the definition of the statistical analysis and the extraction of the results

for the diboson $VZ(b\bar{b})$ cross-check measurement, and the combination of the analyses of $\sqrt{s} = 7$ and 8 TeV datasets.
Chapter 7 describes the search of a CP-odd A boson decaying to $Zh(\rightarrow b\bar{b})$ final states, performed with the first data collected by the LHC during Run-2, until December 2015. My focus in this analysis remained on the $Zh(\nu\bar{\nu}b\bar{b})$ channel, from the definition of the selection cuts until the production of the final distributions used to extract the results. I studied the optimal strategy to include the analysis of high-transverse-momentum final states (where jet substructure techniques are used), combining it with the search at lower transverse momenta. I took care of the estimate of modeling systematic uncertainties for the main backgrounds ($t\bar{t}$ and V + jets backgrounds) and I was editor of the supporting documentation for signal and background modeling studies [15]. Finally, I worked on the definition of the statistical analysis and the extraction of the final results.
In Chapter 8 I present the analysis of the LHC Run-2 data collected until July 2016, for the search of $VH(b\bar{b})$ processes. While I kept contributing to the analysis of the $ZH(\nu\bar{\nu}b\bar{b})$ channel, I studied the modeling systematics for the main backgrounds (with focus on the V + jets and diboson processes) and for the Higgs signal, as well as their implementation in the statistical analysis.
The studies on the modeling of the VH signal allowed me to be involved in the LHC Higgs Cross Section Working Group, profitting from the useful and stimulating discussions taking place in the $VH + VBF$ sub-group, and to participate to the ATLAS publication on Higgs signal Monte Carlo modeling [16]. The work on the diboson background prediction allowed me to contribute to the ATLAS publication on the Monte Carlo modeling of multi-boson processes [17].
Finally, Chapter 9 outlines the general conclusions of the work presented in this thesis, with an outlook on the future prospects for these analyses at the LHC.

2 | The Theoretical Framework

In this Chapter I discuss the theoretical foundations that motivate the research work presented in this thesis, and provide the basis for its interpretation.
I start with an overview of the Standard Model (SM) of particle physics as fundamental framework and I present the Higgs boson as natural consequence of the introduction of a new scalar sector of the theory. I then discuss the phenomenology of this new boson, from the main production modes to the decay channels relevant at hadronic colliders such as the Large Hadron Collider (LHC), with particular emphasis on the processes studied in this work. I finally present some of the theoretical issues and open points that arise from the formulation of the SM with the addition of this new scalar sector, and the attempts that have been made to develop the theoretical framework to address them.
In the second part of the Chapter I focus on a specific set of 'beyond the Standard Model' (BSM) theories which target the SM issues by extending its scalar sector to include additional Higgs bosons, known as Two Higgs Doublet Models (2HDMs).

2.1 The Standard Model of Particle Physics

The SM is a Quantum Field Theory (QFT) developed in the 1970s that unifies three of the four fundamental interactions known today, and gives an almost complete description of the elementary particles physics. In the last two decades the SM has been tested by several experiments at LEP, Tevatron and LHC and other accelerators, and so far all the precision measurements are almost completely reproduced by its theoretical predictions.

A detailed description of the Standard Model is beyond the scope of this thesis, nevertheless I discuss the core concepts on which the SM has been built and that characterize it. The structure and the content of this Chapter are inspired by the complete texts covering the theoretical foundations and the details of the SM of particle physics [18, 19, 20].
First I introduce from a phenomenological point of view the fundamental blocks used to built the SM as a Quantum Field Theory, the elementary particles' fields that describe the components of matter. I'll discuss the structure of interactions between these particles, and how it's implemented within a Quantum Field Theory without spoiling its most appealing properties.
Finally I illustrate the details of the Higgs mechanism, which allowed us to build a Quantum Field Theory of massive fields that satisfy the criteria of gauge invariance and renormalizability. The most important theoretical aspects of the SM are highlighted, explaining why it's such a strong and durable theory, which incorporates the fundamental principles required from an elementary particle theory.
At the end, I briefly outline the main issues and unsolved problems of the SM. It is important to point out that, if today the SM doesn't provide a complete description of particle physics, it is expected to be, at worst, incomplete rather than wrong. That is, the Standard Model is at worst a subset of a more advanced theory of particle physics.

The Standard Model of particle physics is built with the language of Quantum Field Theories: it is important to remind the basic reasons and motivations behind this choice. The idea to reduce particles and interactions to fundamental objects was predominant in the XIX century, moved by the reduction of thermodynamics to statistical mechanics and the unification of magnetism and electricity by Maxwell [21], whose intentions were to create a mechanical model of electromagnetism.
Nevertheless the first cracks of this project appeared precisely in the field of electromagnetism, with the inclusion of special relativity [22].
The goal became then to build the most general theory that included both relativistic kinematics and quantum dynamics of particles, therefore following two main guidelines:

1. *covariant theory*: the Lagrangian is invariant under Lorentz transformation, and all the observables are covariant objects. The theory's states are built as irreducible representations of Lorentz or Poincaré group.

2. *quantum theory*: the states of the theory are described by the rules of quantum mechanics, characterized by the quantum numbers of the symmetry group of the theory.

The procedure to build a theory with such properties is simple: one has to consider all the fundamental objects (that is the most fundamental particles observed by experiments) and find the largest symmetry group that allows to classify these objects as irreducible representations, making sure that the theory respects the two points listed before. The symmetry structure of the group defines then the interactions between the described particles.
This is the simple and general idea behind the construction of Quantum Field Theories, which led to the crisis of mechanical reductionism.

2.1.1 Fundamental Particles of Matter

In this paragraph I describe the elementary matter blocks, the *fermions*. This group of particles includes all the fundamental objects of quantum number of spin $s = \frac{1}{2}$ observed by experiments, and it is further divided in two sub-groups: *leptons* and *quarks*. Leptons only interact through *electroweak* force, while quarks have an additional quantum number, the *color*, related to *strong* force.
Fermions are also divided in three generations of increasing mass, both for leptons and quarks, as shown in Table2.1. The main difference between leptons and quarks is that

	I family	II family	III family	Q/e	Interactions
Quarks	u	c	t	$+2/3$	electroweak
	d	s	b	$-1/3$	strong
Leptons	e	μ	τ	-1	electroweak
	ν_e	ν_μ	ν_τ	0	

Table 2.1: Fermions families with spin and electromagnetic charge quantum numbers

leptons can be experimentally observed as single objects, while quarks are always bounded together to form colourless particles with integer charge, called *hadrons*.
Hadrons composed of three quarks are classified as half-integer spin *baryons*, while hadrons

composed of a quark anti-quark pair are integer spin particles called *mesons*.

Field description of fermions Fermions can be described by three main fields: Weyl, Majorana and Dirac fields. The SM has to describe massive fermions, therefore they cannot be introduced as Weyl fields (which are defined in two-dimensional representations of the Lorentz group, and represent massless particles with positive or negative helicity). In order to write a fermionic mass term in the SM Lagrangian, four-dimensional fields are required: Majorana fields are built from one Weyl field and its hermitian conjugate, while Dirac fields are composed of two different Weyl fields.
It can be easily shown that the free Lagrangian for massive Majorana fields is not invariant under $U(1)$ transformations, therefore electromagnetic charge is not conserved for Majorana particles: as a consequence most of the SM particles cannot be described by such fields. The SM Lagrangian is written using Dirac fields for all the included fermions.
Majorana fields however are not useless and can be used to describe massive neutrinos in SM extensions where neutrinos have only left handed component [23, 24, 25]. In these models the presence of Majorana mass terms should lead to violation of the lepton number as a consequence of the non invariance under global $U(1)$ symmetries.

2.1.2 Standard Model Lagrangian and Symmetry Group

After defining the fundamental objects of the theory, the fermionic fields, the next step is to describe them as irreducible representations of the most large symmetry group that can describe their properties.
This is a fundamental step in building a Quantum Field Theory: writing the Lagrangian as a locally invariant object under the defined symmetry group provides us not only the free theory of fermions, but also the structure of interactions between these free fields. The interaction's terms of the Lagrangian are then identified as additional particles (physical observables), the boson force carriers of spin 1 that are described in Sections 2.1.3 and 2.1.4.

The symmetry structure of the SM has been worked out taking into account several years of experiments and is the result of a challenging theoretical work. The global symmetry group of the SM is:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$$

We can divide the global symmetry group in two main sub-groups:

$\mathbf{SU(2)_L \otimes U(1)_Y}$ is the symmetry structure of the electroweak theory, where $U(1)_Y$ is the hypercharge symmetry group and $SU(2)_L$ describes the weak interaction symmetries.
From this group it's possible to extract a sub-group $U(1)_{em}$ that describes the electromagnetic interactions and the electrical charge conservation. This group allows to describe in a unified way both the weak interaction, at first included as an effective Fermi theory, and the electromagnetic force.

$\mathbf{SU(3)_C}$ describes the colour symmetry of quarks and it's related to strong interaction. From an historical point of view this symmetry group has been introduced to provide a theoretical description of the confinement of quark particles in hadrons, and received validation by predicting successfully all the hadronic structures observed by experiments.

The SM Lagrangian is written as an invariant object under the global symmetry group now described.
In order to describe not only the free theory, but also the interactions, the SM is built as a gauge theory, requiring that the Lagrangian's terms are *locally* invariant under $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$.
The gauge formulation of the theory is achieved by introducing additional terms to the Lagrangian that provide us the interaction fields for the fundamental forces associated with the symmetry group. The complete Lagrangian can be written as the sum of multiple terms:

$$\mathbf{L_{SM} = L_{SU_3} + L_{SU_2 \otimes SU_1}}$$

2.1.3 QCD: $SU(3)$ Lagrangian

The strong interaction (QCD) part of the Lagrangian can be written as:

$$_{SU_3} = -\tfrac{1}{4} F^i_{\mu\nu} F^{i\mu\nu} + \textstyle\sum_r \bar{q}_{r\alpha} i \not{D}^\alpha_\beta \, q^\beta_r$$

where $F^i_{\mu\nu}$ is the field strength tensor for the gluon fields G^i_μ:

$$F^i_{\mu\nu} = \partial_\mu G^i_\nu - \partial_\nu G^i_\mu - g_s f_{ijk} \, G^j_\mu \, G^k_\nu$$

which depends on the QCD coupling constant g_s, with f_{ijk} $(i, j, k = 1, \cdots, 8)$ the structure constants defined by the non-abelian $SU(3)$ algebra.
In order to obtain gauge invariance for the strong interaction term of the Lagrangian, the covariant derivative is defined as: $D^\alpha_{\mu\beta} = (D_\mu)_{\alpha\beta} = \partial_\mu \delta_{\alpha\beta} + i g_s \, G^i_\mu \, L^i_{\alpha\beta}$
where α runs over color quantum numbers for the gluon G^i_μ and $L^i = \frac{\lambda^i}{2}$ are the representation matrices of $SU(3)$.

The strong interactions alone are diagonal in the flavour indices (so there is no mixing between flavours), but they can change the quark colors and are purely vector interactions. In this $SU(3)$ Lagrangian there are no bare mass terms for the quarks: these are allowed by QCD alone but would spoil the chiral symmetry of the electroweak part of the Lagrangian. The quark masses are generated by the scalar sector of the theory, described later.

2.1.4 Electroweak: $SU(2) \otimes SU(1)$ Lagrangian

This second part of the Lagrangian can be further divided in several sub-terms, in order to simplify its structure:

$$L_{SU_2 \otimes U_1} = L_{gauge} + L_f + L_\Phi + L_{Yukawa}$$

where L_{gauge} is the free Lagrangian (or *Maxwell Lagrangian*) for the gauge boson fields that are introduced to preserve gauge invariance:

$$_{\text{gauge}} = -\tfrac{1}{4} F^i_{\mu\nu} F^{\mu\nu i} - \tfrac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

with field strength tensors that reflects the abelian nature of $U(1)_Y$ and the non-abelian one of $SU(2)_L$:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

$$F_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk} W^j_\mu W^k_\nu$$

where g is the gauge coupling constant for the $SU(2)_L$ group, and W^i_μ, $i = 1,\ 2,\ 3$ and B_μ are respectively the SU_2 and U_1 gauge fields.

The fermion part of the Lagrangian, L_f, is simply a Dirac-like Lagrangian for the fermionic fields written with the covariant derivative of $SU(2)_L \otimes U(1)_Y$ to obtain gauge invariance:

$$_F = \sum_{m=1}^{F} \left(\bar{q}^0_{mL} i \not{D} q^0_{mL} + \bar{l}^0_{mL} i \not{D} l^0_{mL} + \bar{u}^0_{mR} i \not{D} u^0_{mR} + \bar{d}^0_{mR} i \not{D} d^0_{mR} + \bar{e}^0_{mR} i \not{D} e^0_{mR} \right)$$

and

$$D_\mu \varphi = \left(\partial_\mu + ig \frac{\tau^i}{2} W^i_\mu + \frac{ig'}{2} B_\mu \right) \varphi$$

m is the family index and F is the number of families (for the SM, F = 3). The fermionic fields are included as left and right chiral projectors defined at the beginning of Section 2.1.1. As for the QCD Lagrangian bare mass terms are forbidden by the requirement of chiral invariance.

The Lagrangian's terms introduced so far describe the free propagation and the interaction not only of the fermion's fields, but also for the gauge fields introduced to build the Lagrangian itself. Nevertheless we are neglecting the masses of all the particles, in order to preserve local gauge invariance (for gauge fields) and chiral symmetry (for fermions).
To overcome this issue a scalar sector is added to the SM, described by two additional Lagrangian terms: L_Φ to describe the new scalar field with its characteristic potential and L_{Yukawa} to include its interactions with fermions. These additional terms are described in full detail, together with the Higgs mechanism that justifies their introduction, in Section 2.2.

2.2 Scalar Sector of the Standard Model: Higgs Mechanism

Before introducing the details of this theoretical model, I would like to highlight the need to avoid bare masses for both the fermions and the gauge bosons in the SM Lagrangian, as stated in the previous section.
The main motivation behind the great theoretical effort represented by the construction of the SM is the will to describe the elementary particles with a *gauge invariant* theory. The gauge invariance, in fact, grants us that the model satisfys several important properties, first of all renormalizability, and that has the correct behaviour in high-energy regimes.
Including bare masses in the Lagrangian would spoil the gauge invariance in two different ways:

gauge fields: adding a brute-force mass term in the gauge Lagrangian for any of the gauge fields would lead to explicit violation of gauge invariance.
In order to obtain the gauge invariance of the fermion fields, that locally transform like:

$$\Phi(x) \to e^{-ie\alpha(x)} \Phi(x)$$

The gauge fields have to transform according to the following expression (for example B_μ):

$$B_\mu \to B_\mu + \partial_\mu \alpha(x)$$

Therefore, a mass term like $\mu^2 B^\mu B_\mu$ is clearly not invariant under B_μ transformations.

fermions: considering a Dirac-mass for the fermions: $m(\bar{\Phi}_L \Phi_R + \bar{\Phi}_R \Phi_L)$, it is clear that, since the electroweak coupling is different between the right-handed and left-handed part of the field, such a term is not invariant under a chiral transformation of $SU(2)_L$.

Properties of gauge theories Including naive mass terms for fermions and bosons would spoil gauge invariance: which are the reasons that make a gauge invariant theory so appealing, and why it's widely believed that the fundamental theory of nature has to be a gauge theory? Gauge theories show some important features which allow to avoid unphysical behaviours and make the SM a robust and solid theory of elementary particles.

1. ***renormalizability*** [26]: the computation of probability amplitudes for processes with one or more loops leads to divergent results; the theory is called renormalizable if there is a finite number of divergences at all perturbative orders (the number of divergences must not increase with the perturbative order). As a consequence it is possible to measure a finite number of physics observables to absorb the divergences at every perturbative order.
   This divergent behaviour may seem suspicious, and one may think that it hints at some unsolved issue of the theory: on the contrary renormalization confirms the correctness of the SM as a Quantum Field Theory. It reflects the fact that the Lagrangian is built using parameters (coupling constants, particles' masses, ...) that have to be experimentally measured in order to be well defined. The divergences appear trying to express physics observables (as cross sections or decay widths) in terms of the conventional parameters written in the Lagrangian.

   Renormalizable theories show adimensional coupling constants, while non renormalizable ones have dimensional couplings. A non renormalizable theory is an effective theory: it is still predictive if we measure enough observables to absorb all the divergences, but increasing the perturbative order the number of measurements becomes higher.
   The idea that the fundamental theory of nature has to be a renormalizable theory is a theoretical preconception, supported by the success of the SM: the fundamental laws of nature could be very well described by different non renormalizable theories, each valid at its own energy scale. In this scenario there wouldn't be any larger renormalizable theory which generates the effective theories we observe at lower energy scale throughout decoupling.

2. ***high-energy behaviour*** [27, 28]: probability amplitudes must satisfy a unitarity requirement, which can be spoiled if the energy behaviour of the amplitude is not correct. In detail, if the amplitude increases linearly with the center of mass energy, unitarity is violated. Interaction terms coming from the gauge structure of the theory (for example trilinear electroweak boson couplings) cancel the amplitudes that show an incorrect energy behaviour, thus avoiding unitarity violation at high energy scale

(see for example the diagrams in Figure 2.1).
The gauge symmetry of the theory sets not only the structure of the coupling, but also grants the universality of the coupling constants: this feature grants the cancellation of problematic terms.

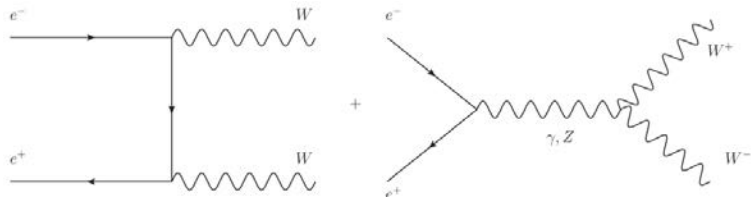


Figure 2.1: Feynman diagrams for the $e^+e^- \to W^+W^-$ process

In order to overcome this *mass issue* we can include an additional scalar sector to the model, introducing a mechanism that provides mass to all the massive fermions and to three of the electroweak bosons (Z^0, $W^\pm$) leaving the photon massless.

Symmetries in a Quantum Theory The following description of the scalar sector and especially of the Higgs mechanism are based on a simple but fundamental idea: a symmetry transformation has different definitions for classic and quantum systems.

classic system: if the Lagrangian L is invariant under symmetry transformation, all the states of the system are invariant.

quantum system: the system is symmetric if the Lagrangian L and all the quantum states of the system are invariant under the transformation, but states invariance is not a consequence of the Lagrangian invariance.
Since the states are built from the vacuum state $|0>$ with ladder operators, state invariance is achieved if the vacuum is invariant under symmetry transformations.

2.2.1 Construction of the Scalar Sector

The scalar sector of the SM is surely the less constrained sector: there are several experimental results about the fermionic sector (coupling and mass measurements), while we have only less direct and indirect information about the scalar one.
Keeping this in mind we want to build a SM scalar sector introducing the minimum number of required degrees of freedom, following a principle of simplicity.

The minimal model that can give mass to three of the four bosons of $SU(2)_L \otimes U(1)_Y$ is built with three scalar fields, providing one degree of freedom for each massive boson. This choice however is not the one realized in nature: the radiative corrections computed within this model are not in good agreement with the SM precision tests.

The second minimal model has one additional degree of freedom, and shows interesting symmetry properties that grant us much smaller radiative corrections and a better agreement with experimental data. This model is built using a complex doublet of scalar fields, defined as:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi_1 + i\phi_2 \end{pmatrix}$$

This field has to satisfy the gauge invariance properties of L under $SU(2)_L \otimes U(1)_Y$ transformations ($SU(3)_C$ does not contribute, since the gluons are massless and don't couple with the scalar sector):

$$\begin{aligned} SU(2)_L \qquad & \Phi \to e^{[i\frac{g}{2}\tau_j \alpha_j(x)]}\Phi \\ U(1)_Y \qquad & \Phi \to e^{[i\frac{g'}{2}Y\beta(x)]}\Phi \end{aligned}$$

where g is the $SU(2)_L$ coupling constant, Y is the ipercharge quantum number and g' the $U(1)_Y$ coupling constant. As for the electroweak sector, gauge invariance can be achieved introducing the adequate covariant derivative for the scalar field:

$$D_\mu \Phi = \left(\partial_\mu + ig\frac{\tau^i}{2}W^i_\mu + \frac{ig'}{2}B_\mu\right)\Phi$$

Finally one can write the Lagrangian term for the scalar sector as:

$$_\Phi = (D^\mu \Phi)^\dagger D_\mu \Phi - V(\Phi) = (D^\mu \Phi)^\dagger D_\mu \Phi - \mu^2 \Phi^\dagger \Phi - \lambda(\Phi^\dagger \Phi)^2$$

L_Φ symmetry group The Lagrangian L_Φ for a doublet of complex fields is explicitly invariant under O(4) transformations. It can be easily shown, writing L_Φ in terms of a complex matrix H_{2x2}, that it is invariant under $SU(2)_L \otimes SU(2)_R$ transformations as well (O(4) and $SU(2)_L \otimes SU(2)_R$ are isomorphic groups)

In addition to this standard construction of the scalar sector, I can define some useful quantities related to the scalar field:

- $|0>$ is *vacuum state* of the system, the lowest energy state
- $v = <0|\Phi|0>$ is the *vacuum expectation value* (*vev*) of the scalar field Φ

The core concept behind the Higgs mechanism is that the vacuum expectation value of the new scalar field introduced does not respect gauge symmetry.

Since the vacuum state $|0>$ is the lowest energy state of the system, it's important to study the potential (or the energy) for the scalar sector now defined, and therefore the vacuum state itself.

2.2.2 Scalar Field Potential

The potential for Φ is:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda(\Phi^\dagger \Phi)^2$$

Since the potential depends on Φ, field configuration, we cannot find explicitly the solution for the vacuum state from $V(\Phi)$, but we can redefine the potential as a function of $v = <0|\Phi|0>$ and minimize $V(v)$, interpreting v as the solution of the classical equation of motion.
With few mathematical steps it can be shown that:

$$V(\Phi) \to V(v) = \tfrac{1}{2}\mu^2|v|^2 + \tfrac{1}{4}\lambda v^4$$

There are two possible solutions:

> $\boldsymbol{\mu^2 > 0}$, there is a single minimum at $v = 0$, therefore the vacuum is empty and cannot break the symmetry (this case is not of physical interest).

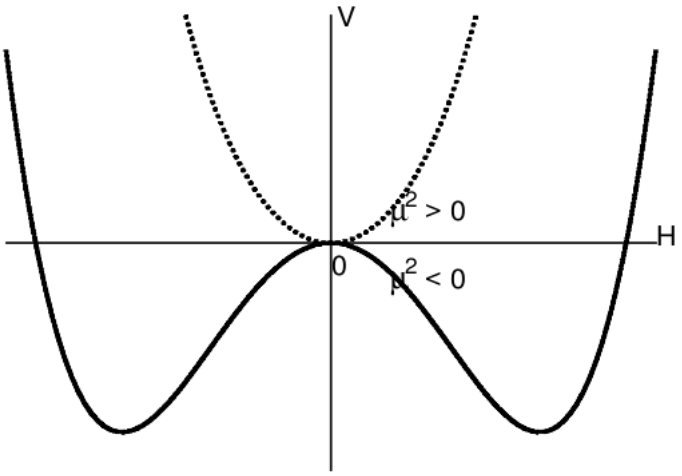


Figure 2.2: The Higgs potential $V(v)$ for $\mu^2 > 0$ (dashed line) and $\mu^2 < 0$ (solid line).

$\mu^2 < 0$, the solution is $|v| = \left(-\mu^2/\lambda\right)^{1/2}$ and is not unique (for $\mu^2 = 0$ the solution is similar to $\mu^2 < 0$).

Considering the $\mu^2 < 0$ case, it's clear that the choice of the minimum is not trivial, since the point of minimum is not unique. To further discuss this point the introduction of a new concept is required.

Spontaneous Symmetry Breaking (SSB) A spontaneous symmetry breaking is a symmetry breaking related to the choice of a particular state of the system. This kind of symmetry breaking is possible also for classic systems, but plays a much more relevant role in quantum theories when the choice of the vacuum state is not trivial, and is connected with the appearance of a bosonic degree of freedom.

Goldstone Theorem The Goldstone theorem is built on the concept of SSB and provides a fundamental result for the Higgs mechanism [29].

Hp: Given a Lagrangian invariant under a group of continuous transformations with N generators, and assuming that M of N generators are spontaneously broken: the vacuum state is not any more invariant under the M transformations related to the broken generators.

Th: In the particles' spectrum of the theory, developed around the vacuum expectation value v, there are M massless particles called Goldstone bosons (one boson for each spontaneously broken generator).

It's important to remind that the Goldstone bosons are not affected by radiative corrections, therefore they are always massless.
The Goldstone bosons are massless scalar fields of the theory, that cannot be interpreted as physics observables (real particles): they appear in unitary gauge as the longitudinal degree of freedom of the massive spin 1 particles that acquire mass, while in arbitrary gauge they can be expressed as non-physical particles of the theory.

2.2.3 Higgs Mechanism

The core of the Higgs mechanism [4, 5] is the application of the concepts described so far to the potential introduced in Section 2.2.2.
The *vev* is not unique, therefore the symmetry group can be spontaneously broken on the vacuum state by choosing a specific value for the *vev* itself. The field Φ can then be re-written as an expansion around the *vev* chosen: in the resulting Lagrangian the required mass terms for the gauge bosons of the theory appear as a direct consequence.

Choice of the vacuum expectation value Generally there isn't a prescription to choose the *vev* to obtain a specific result. However, after the SSB, the resulting theory must still include the QED with its un-broken $U(1)_{em}$, since the photon must not acquire mass. The *vev* chosen during the SSB must be therefore invariant under $U(1)_{em}$.
Given the general form of the vacuum expectation value:

$$vev = <0|\Phi|0> = \Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

there are two choices that preserve $U(1)_{em}$ (conventionally the first one is used):

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}; \; \Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix}$$

Now the scalar field Φ can be written as an expansion around the *vev* (on the right in unitary gauge for a simpler expression):

$$\Phi = \begin{pmatrix} \phi^+ \\ v + \sigma + i\eta \end{pmatrix} \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \sigma \end{pmatrix}$$

Mass of the vector bosons Z^0, $W^\pm$

The scalar Lagrangian L_Φ can now be written with the explicit form of the scalar complex doublet Φ and the covariant derivative D_μ, that is combining the following equations:

$$_\Phi = (D^\mu \Phi)^\dagger D_\mu \Phi - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2$$

$$D_\mu \Phi = \left(\partial_\mu + ig \frac{\tau^i}{2} W_\mu^i + \frac{ig'}{2} B_\mu \right) \Phi$$

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \sigma \end{pmatrix}$$

The computation is long but straightforward and the result shows interesting features, and can be divided in two parts:

- charged component: W^1, W^2
- neutral component: B, W^3

Writing the $W^{(1,2)}$ fields as linear combination, $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2)$, we recognize immediately the mass eigenstates for the $W^\pm$ bosons, and their mass terms within the Lagrangian.
The neutral component contains mixed terms of B and W^3 fields, therefore the physical fields can be retrieved by diagonalizing the mass matrix with a rotation of the Weinberg angle.

$$A_\mu = \cos(\theta_W)B_\mu + \sin(\theta_W)W^3_\mu$$
$$Z^0_\mu = \cos(\theta_W)W^3_\mu - \sin(\theta_W)B_\mu$$

With the constraining relation on the value of θ_W:

$$g\sin(\theta_W) = g'\cos(\theta_W) = e \Rightarrow \tan(\theta_W) = \frac{g'}{g}$$

Once the fields are written as mass eigenstates the gauge Lagrangian of $SU(2)_L \otimes U(1)_Y$ together with the scalar Lagrangian L_Φ can be conveniently expressed as:

$$\begin{aligned} L_\Phi + L_{gauge} = \Big(& \frac{1}{2}(\partial_\mu\sigma)(\partial^\mu\sigma) + \mu^2\sigma^2 - \lambda v\sigma^3 - \frac{\lambda}{4}\sigma^4 + \\ & + (\frac{gv}{2})^2 W^+_\mu W^\mu_- + \frac{1}{2}(\frac{g^2+g'^2}{4})v^2 Z_\mu Z^\mu + \\ & + \frac{1}{2}g^2 v W^+_\mu W^\mu_- \sigma + \frac{g^2 v}{2}\frac{1}{2\cos(\theta_W)^2} + Z_\mu Z^\mu\sigma + \\ & + \frac{g^2}{4}W^+_\mu W^\mu_- \sigma^2 + \frac{g^2}{4}\frac{1}{2\cos(\theta_W)} Z_\mu Z^\mu \sigma^2 \Big) \end{aligned} \tag{2.1}$$

The first line describes the self-couplings of the σ scalar field, the second line contains the physical fields' mass terms, the third and fourth lines contain the couplings between the field σ and the gauge bosons.
The mass of the W and Z bosons can be explicitly written as:

- $m_W = \frac{gv}{2} = 81$ GeV
- $m_Z = \frac{v}{2}\sqrt{g^2+g'^2} = 91$ GeV

The vacuum expectation value v can be obtained from the measurement of the muon decay width [30, 31, 32], and corresponds to: $vev = v = 246,22$ GeV. Using this measure m_W and m_Z can be computed, obtaining the value quoted before, in very good agreement with the direct mass measurements [33, 34, 35].

Coupling between the gauge bosons and the Higgs From the final form of the Lagrangian we can see the coupling between the W and Z bosons with the Higgs field σ. This couplings can be re-written in order to make a very interesting property explicit:

- $\frac{1}{2}g^2 v W^+_\mu W^\mu_- \sigma = g\,\mathbf{m_W}\,W^+_\mu W^\mu_- \sigma$
- $\frac{g^2 v}{2}\frac{1}{2\cos(\theta_W)^2} Z_\mu Z^\mu \sigma = g\,\mathbf{m_Z}\,\frac{1}{2\cos(\theta_W)} Z_\mu Z^\mu \sigma$

The coupling of the Higgs boson σ with a second particle is always proportional to the mass of the particle itself (this result is valid for all bosonic and fermionic fields).

Mass for the fermionic sector

The scalar sector introduced in the SM Lagrangian to provide mass to the gauge bosons can be used to define the fermionic mass in a gauge invariant way.
We can write a candidate Lagrangian as:

$$L_{Yukawa} = -Y_u(\bar{Q}_L\tilde{\Phi}u_R + \bar{u}_R\tilde{\Phi}^\dagger Q_L) - Y_d(\bar{Q}_L\Phi d_R + \bar{d}_R\Phi^\dagger Q_L)$$

where $Y_{(u,d)}$ are called *Yukawa coupling constant* and are specific for each fermion, and the $Q_{(L,R)}$ terms are doublet of up and down fermionic fields: $Q_{(L,R)} = \begin{pmatrix} u_{(L,R)} \\ d_{(L,R)} \end{pmatrix}$. Therefore the explicit expression of the Yukawa Lagrangian contains:

- $\bar{Q}_L \tilde{\Phi} u_R = (\frac{\bar{u}_L \, v \, u_R}{\sqrt{2}})(\frac{\bar{u}_L \, \sigma \, u_R}{\sqrt{2}})$
- $\bar{Q}_L \Phi d_R = (\frac{\bar{d}_L \, v \, d_R}{\sqrt{2}})(\frac{\bar{d}_L \, \sigma \, d_R}{\sqrt{2}})$

where the first term provides the mass to the fermion field and the second one shows the coupling between the fermion and the Higgs boson σ. Both the mass and the coupling terms are explicitly invariant under $SU(2)_L \otimes U(1)_Y$ transformations.
The fermionic mass can be then written as:

$$m_u = \frac{Y_u \, v}{\sqrt{2}}$$
$$m_d = \frac{Y_d \, v}{\sqrt{2}}$$

Mass eigenstates and properties The fermionic fields used to write the Yukawa Lagrangian are flavour eigenstates, so they are not the physical observables detected in the experiments. Expressing the Yukawa coupling constant as mass matrices and writing the Lagrangian for the physical mass eigenstates we can look at some very interesting features: $Y_{(u,d)} \to M_{ij}^{(up,down)}$

$$L_{Yukawa} = -M_{ij}^{up}(\bar{Q}_{iL}\tilde{\Phi}u_{jR} + h.c.) - M_{ij}^{down}(\bar{Q}_{iL}\Phi d_{jR} + h.c.)$$

In can be shown that the mass matrices $M_{ij}^{(up,down)}$ can be diagonalized with a pair of unitary transformations, that act as basis transformations on the fermionic fields. It is therefore possible to write the Lagrangian for the mass eigenstates exploiting these unitary transformations, different for up and down fields:

$$\bar{u}_L = U_L \bar{u}_L^0 \qquad \bar{u}_R = U_R \bar{u}_R^0$$
$$\bar{d}_L = V_L \bar{d}_L^0 \qquad \bar{d}_R = V_R \bar{d}_R^0$$

where u^0 and d^0 are mass eigenstates.
This result becomes very interesting if we look at the neutral and charged components of the electroweak Lagrangian, writing them for the mass eigenstates.

- $L_{NC} = \bar{u}_L \gamma_\mu u_L + \ldots \Longrightarrow \bar{u}_L^0 \gamma_\mu u_L^0 + \ldots$
- $L_{CC} = \frac{g}{\sqrt{2}} \bar{u}_L \gamma_\mu \frac{1-\gamma_5}{2} d_L + \ldots \Longrightarrow \frac{g}{\sqrt{2}} \bar{u}_L^0 \gamma_\mu \frac{1-\gamma_5}{2} d_L^0 [U_L^\dagger V_L] + \ldots$

Therefore the neutral electroweak interactions are diagonal for mass eigenstates. This is a fundamental result for the SM: from a phenomenological point of view flavour changing neutral couplings (FCNC) are extremely suppressed, and every theoretical model that predict them at tree level is strongly disfavored.
Considering the charged term we can write the $[U_L^\dagger V_L]$ matrix product as a single unitary matrix which summarizes the electroweak mixing between quark flavours: the CKM matrix V_{CKM} [36, 37].

2.2.4 Main Standard Model features involving the Higgs scalar sector

In this Section I would like to point out some of the main properties that make the model just described so solid and predictive for an elementary theory of particles.

Free parameters of the SM

The gauge sector of the SM can be written using four constant parameters:

- g, g' electroweak coupling constants
- v vacuum expectation value
- λ quartic coupling constant of the SM scalar sector

The fermionic sector gives twelve Yukawa coupling constants, which can be directly measured from fermionic masses. It is possible to write the four gauge constants in terms of physics observables, in order to measure them, choosing quantities that can be measured with extremely high precision:

$$\alpha_{em} = \frac{e^2}{4\pi}\frac{(g\sin\theta_W)^2}{4\pi}$$
$$\frac{G_\mu}{\sqrt{2}} = \frac{g^2}{8m_W^2} = \frac{1}{2v^2}$$
$$m_Z = \frac{v}{2}\sqrt{g^2+g'^2}$$

Therefore it's possible to obtain the numerical value for v, g and g', while λ remains a free parameter: this implies that the mass of the Higgs boson is not fixed by the model but can assume different values. The mass m_H can be written explicitly as:

$$m_H = \sqrt{2\lambda v^2}$$

The scale of the Higgs mass is thus controlled by the vacuum expectation value of the field, but its precise theoretical value is unknown because of the new coupling constant λ.
This results were quite important from a phenomenological point of view, because the theoretical model didn't provide a direct information about the mass range where the Higgs boson should have been found.
The observation of the Higgs boson, announced on July 4th 2012 by the ATLAS and CMS collaborations [6, 7], has now been well established and a measurement of the Higgs boson mass has been performed by the ATLAS and CMS Collaborations using data collected during the Run-1 of the LHC [38], resulting in:

$$m_H = 125.09 \pm 0.21(stat) \pm 0.11(syst)\ \text{GeV} = 125.09 \pm 0.24\ \text{GeV}$$

Stability of the Higgs potential

The beta function describing the running of the quartic Higgs coupling λ introduced in the scalar potential in Section 2.2.2 is given by:

$$\beta(\lambda) = \frac{1}{16\pi^2}(24\lambda^2 - 6Y_t^4 + \frac{9}{8}g^4 + \frac{3}{8}g'^4)$$

where Y_t is the top Yukawa coupling (neglecting the Yukawa for lighter fermions). The β function is thus positive or negative whether the negative contribution from the large top Yukawa coupling is compensated by the gauge couplings and the Higgs quartic coupling. We can thus observe that the running of the Higgs quartic coupling is determined by the measured values of m_t and m_H at the electroweak scale.
Had we measured a larger value of m_H we could have implied a non-perturbative behavior of the Higgs self-coupling above some scale Λ, but from the measured value of m_H (as well as the other EW SM parameters) we can conclude that the Higgs quartic coupling remains

perturbative up to the Planck scale [39, 40].
However, for the same measured value of m_H we find that the $\beta(\lambda)$ function is negative and the λ coupling decreases becoming negative at energies $\Lambda = O(10^{10} - 10^{12})$ GeV, as shown in Figure 2.3, although it has a broader range when considering a 3σ fluctuation in the top quark mass value. For negative values of λ the SM vacuum state becomes unstable: this behavior may call for new physics at an intermediate scale before the instability (hence below the Plank scale), or alternatively the electroweak vacuum may remain in a metastable state [41, 42, 43].
This question, within the SM, is related to the lifetime of the metastable vacuum, determined by the rate of quantum tunneling from this vacuum into the true vacuum state of the theory.

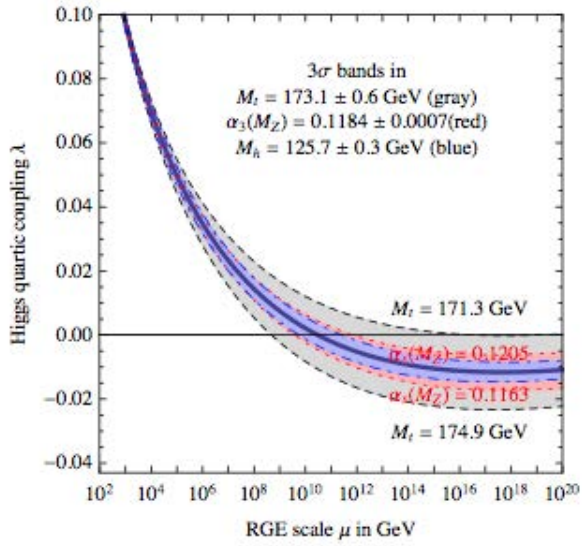


Figure 2.3: Renormalization group evolution of the Higgs quartic coupling λ, for $m_H = 125.7$ GeV, $m_t = 173.4$ GeV and $\alpha_S(M_Z) = 0.1184$, with variations of $\pm 3\sigma$ for these three parameters separately [41].

Higgs boson width

As already derived explicitly the Higgs boson couplings to fundamental fermions and bosons are proportional to the masses of coupled particles. In detail we can write the SM Higgs boson couplings as:

$$g_{Hf\bar{f}} = \frac{m_f}{v}, \quad g_{HVV} = \frac{2m_V^2}{v}, \quad g_{HHVV} = \frac{2m_V^2}{v^2}, \quad g_{HHH} = \frac{3m_H^2}{v}, \quad g_{HHHH} = \frac{3m_H^2}{v^2}$$

where V = $W^{\pm}, Z$. Since all Higgs couplings to fermions and bosons are specified knowing the mass of the Higgs itself and of its decay products, in the SM the natural width of the Higgs boson can be precisely calculated by summing the partial width of each decay mode. The total width can thus be described as a function of the Higgs mass alone, as shown in Figure 2.4. The natural width of a SM Higgs boson with a mass of 125 GeV is about 4.1 MeV (smaller than the experimental mass resolution obtained by the ATLAS and CMS Collaborations, which is of the order of 1-2% of the Higgs mass), with a theoretical uncertainty of approximately 4%: if the Higgs width were experimentally measured to be larger than the SM prediction, it may indicate the presence of new particles to which the

Higgs boson could decay. For this reason Higgs width measurements can be used to probe the existence of new physics coupled to the Higgs sector.
The direct upper limit on the total width Γ_H is currently quoted by the CMS Collaboration at 1.7 GeV [44]. More stringent limits can be obtained from the measurement of the off-shell coupling strength of the Higgs boson decaying to WW and ZZ pairs, with the stronger assumptions that the SM Higgs couplings are unchanged and no new physics intervenes at energy scales larger than m_H (in the probed off-shell regime): observed limits on Γ_H are quoted at 22.7(26) MeV by the ATLAS(CMS) Collaboration [45, 46].

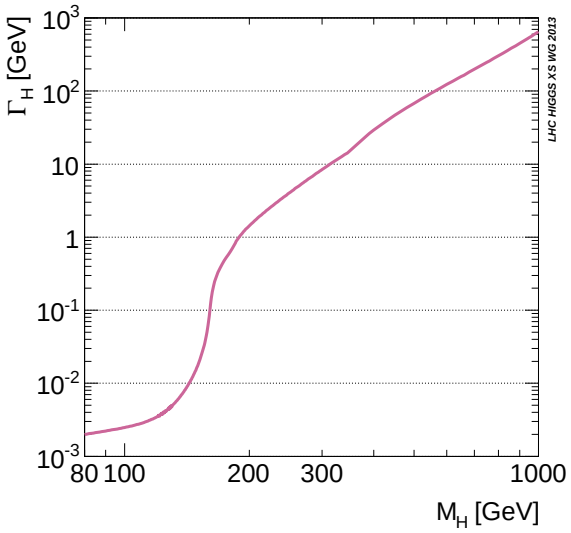


Figure 2.4: Theoretical prediction of the Higgs total width as a function of the mass of the Higgs booson.

SM custodial symmetry

In addition to its $SU(2)_L \otimes U(1)_Y$ gauge symmetry, the SM Higgs Lagrangian also respect an additional approximate global symmetry, which is verified in the limit of $g' \to 0$ and $h_f \to 0$ as a $SU(2)_R$ symmetry. Under this approximation the Higgs Lagrangian is therefore globally invariant under $SU(2)_R \times SU(2)_L$, and this symmetry remains preserved for non vanishing Yukawa couplings Y_f as long as $Y_u = Y_d$.
After the EWSB mechanism the $SU(2)_R \times SU(2)_L$ global symmetry is spontaneously broken to a global $SU(2)_C$ called custodial symmetry group [47, 48]:

$$SU(2)_R \times SU(2)_L \to SU(2)_C$$

In the custodial symmetry limit ($g' \to 0$) we can indeed see that the $W^\pm$ and Z gauge bosons have identical mass and form an $SU(2)_C$ triplet. We can characterize the custodial symmetry by introducing a parameter ρ, which is expressed at tree level as:

$$\rho = \frac{M_W^2}{M_Z^2 cos^2\theta_W} = 1 \text{ or } \frac{M_W^2}{M_Z^2} = \frac{g^2}{g'^2 + g^2} = cos^2\theta_W$$

These relations always hold at tree level, and clearly show that in the $g' \to 0$ limit we obtain the mass degeneracy $M_W = M_Z$. The ρ parameter is a key quantity in the context of precision electroweak measurements, since it's measured experimentally with very high

accuracy and may be especially affected by new physics. Since the custodial symmetry protects the relation between the W and Z masses under radiative corrections (therefore $\rho = 1$ beyond tree level), all corrections to the value of ρ are proportional to terms that break the custodial symmetry.

To give one example, if we had built the SM scalar sector from a triplet of scalar fields (rather than the Higgs doublet) we would have found very large radiative corrections to ρ which allow us to discard such theoretical model.
An other example comes from the heavy quarks' masses: since $m_t \neq m_b$, the $Y_u = Y_d$ approximation is not valid and the custodial symmetry is broken, hence we see relevant radiative corrections generated by massive fermions, proportional to $m_t^2 + m_b^2 - 2(m_t^2 m_b^2) log(m_t^2/m_b^2)/(m_t^2 - m_b^2)$ [49]. Indeed, before the direct discovery of the top quark at Tevatron, the ρ parameter has provided the main indirect information on the top mass m_t, which was determined with approximately 30% accuracy. At one loop order the ρ corrections from top-bottom loop is of the order of 1% and can be expressed as:

$$\rho = 1 + \frac{3Y_t^2}{4\pi^2} = 1 + \frac{3G_F m_t^2}{8\sqrt{2}\pi^2}$$

Conversely, since the Higgs potential respects exactly the $SU(2)_R \times SU(2)_L$ custodial symmetry, we cannot find any one-loop correction to ρ proportional to the Higgs quartic coupling λ which would lead to corrections proportional to m_H^2. To find this type of corrections we need to consider two-loop level contributions which explicitly break the custodial symmetry, and are too small to put significant constraints on m_H [50].

2.3 SM Higgs Boson Phenomenology at Hadron Colliders

In Section 2.2 I gave an overview of the foundations of the SM of particle physics, with an additional scalar sector which introduces the Higgs boson as a new particle. I now use this theoretical framework and its predictions to characterize the behavior of the SM Higgs boson particle in the context of a proton-proton hadronic collider such as the LHC, discussing its main modes of production and decay. Particular emphasis is given to the VH associated production mode and the $H \to b\bar{b}$ decay channel, which are the focus of this research work.
Complete and detailed references for the theoretical calculations for the different Higgs production modes and decay channels are given in [51, 52].

2.3.1 Higgs Production Mechanisms

The main production modes of a SM Higgs boson are categorized depending on the different initial and final states that characterize the process. As shown in Figure 2.5 the production mechanisms, ordered from the highest to the lowest cross section at the LHC, are: gluon-fusion (ggH), vector boson fusion (VBF), associated production to vector bosons (VH), associated production to heavy quarks pair ($b\bar{b}H$, $t\bar{t}H$), associated production to single top quark (tH).

Gluon fusion production mode: ggF The dominant production mechanism for a SM Higgs boson at the LHC is via gluon fusion: $pp \to gg \to H$. Gluons do not couple directly to the Higgs boson, hence at the lowest order the production is mediated by a fermion loop, as illustrated in Figure 2.6, and for this reason it may seem suppressed with respect

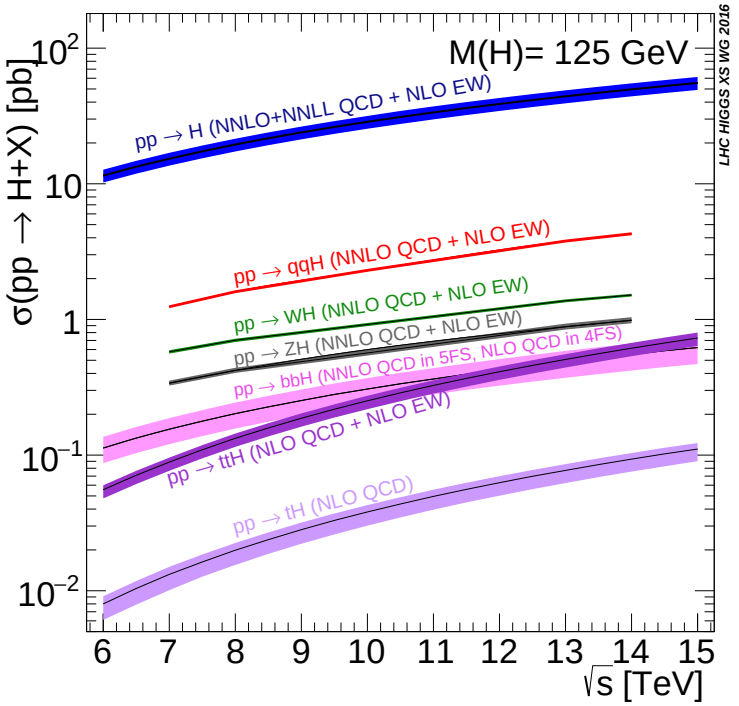


Figure 2.5: Production cross sections for a SM Higgs boson with $m_H = 125$ GeV as a function of the center of mass energy ($\sqrt{s}$ [TeV]) for the different Higgs production modes [52].

to other processes at tree level. Nevertheless the gluon density within colliding protons is highly dominant with respect to other partons, as clearly shown by Figure 2.7 where the PDFs for gluon and quarks are shown (the typical range of the average momentum fraction carried by partons in Higgs production at LHC is $0.001 \leq< x >\leq 0.01$), therefore the Higgs production cross section trough gluon fusion is dominant at hadron colliders. The intermediate fermionic loop has to contain colored particles (to allow coupling with gluons), and because the Higgs coupling is directly proportional to the mass of the interacting particles, the most important contributions come from heavy quark loops (top, bottom) with large Yukawa coupling.

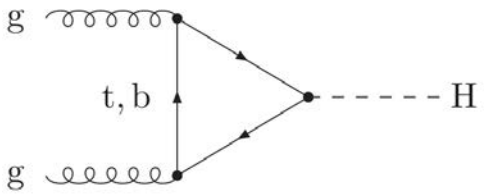


Figure 2.6: Feynman diagram of gluon fusion Higgs production at lowest order.

The cross section for this process has been recently determined at next-to-next-to-next-to-leading-order (NNNLO or N^{3}LO) in the effective theory as an expansion around threshold, with a theoretical uncertainty of the order of 5-6% and PDF$\oplus\alpha_S$ uncertainties of the order of 3% [54].

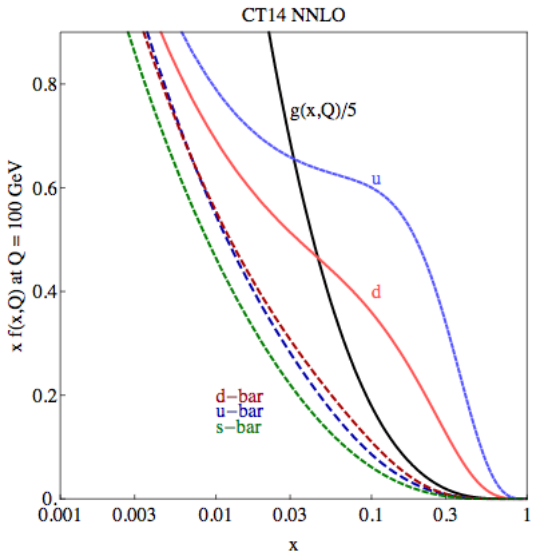


Figure 2.7: The CT14 NNLO parton distribution functions at $Q = 100$ GeV for $u, \bar{u}, d, \bar{d}, s = \bar{s}$ and g as a function of the transferred momentum fraction x [53].

Vector boson fusion production mode: VBF The second production mode in order of decreasing rate is the vector boson fusion (VBF) mechanism: $pp \to qq \to qqV^*V^* \to qqH$. This process has a very distinct signature in the detector, characterized by the presence of two hard jets in the forward and backward regions of the detector.
The two main contributions to VBF Higgs production come from the t- and u-channel diagrams shown in Figure 2.8, which define the *genuine* VBF channel: the two hard jets have a strong tendency to be forward-backward, in direct contrast to other jet-production mechanisms, providing a good background rejection through selection criteria on transverse momentum (p_T) and pseudorapidity (η) of the jets, jets η gap and central jet activity (for a precise definition of the pseudorapidity η see Section 3.2.1). The s-channel diagrams and the interference terms are strongly suppressed by these standard VBF cuts, hence the cross section can be properly approximated by the contributions of squared t- and u- amplitudes without interferences. The QCD corrections to LO diagrams are thereby reduced to vertex corrections to the weak boson quark coupling, while explicit NLO QCD calculations in this approximation show that these corrections are small (about $5 - 10\%$). The NLO EW corrections amount to approximately 5% [55, 56]. A first calculation of NNLO QCD corrections to the total and differential inclusive cross section for VBF has been obtained using the structure-function approach [57, 58].

Associated production to a vector boson: VH In the associated production of the Higgs boson (or *Higgsstrahlung*) along with a vector boson $W^\pm$ or Z, the leading-order diagrams are quark-initiated ($pp \to q\bar{q} \to V^* \to VH$): as a result this is the third mechanism, as far as production rate is concerned, at proton-proton colliders as the LHC, while it ranked second (after gluon-fusion production) at the proton-antiproton Tevatron collider. Furthermore, it was the most important production mode exploited at LEP for Higgs searches, allowing to avoid Yukawian suppression due to the direct coupling of the Higgs boson to electrons.
The associated production of Higgs and gauge bosons is one of the simplest production mechanisms at the LHC: the final state VH is not affected by strong interactions, which

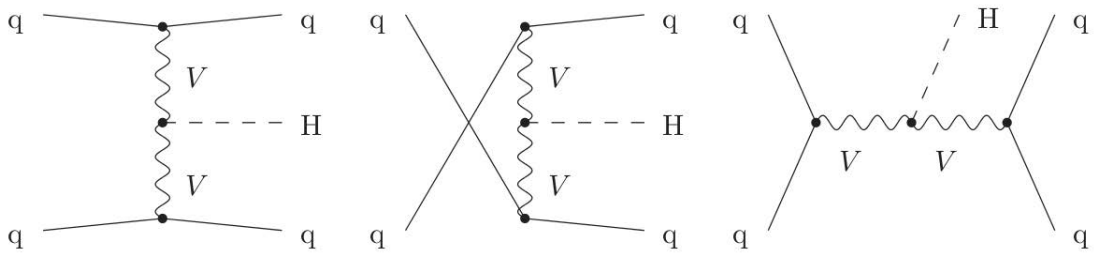


Figure 2.8: Feynman diagram of vector boson fusion Higgs production $qq \to qqH$ at lowest order, in t-, u- and s-channels, where V denotes the W or Z bosons and q denotes any quark or antiquark.

affect only the quark-antiquark initial state. At the lowest perturbation order in fact, this process can be viewed simply as the Drell-Yan production of a virtual $W^{\pm}$ or Z boson, which then splits into a real vector boson and a Higgs particle. The partonic cross section can therefore be expressed as a function of the momentum of the virtual gauge boson k as:

$$\frac{d\hat{\sigma}}{dk^2}(q\bar{q} \to VH) = \sigma(q\bar{q} \to V^*) \times \frac{d\Gamma}{dk^2}(V^* \to VH)$$

corresponding to the leading-order (LO) Feynman diagram shown in Figure 2.9.

Next-to-leading-order (NLO) QCD corrections to Drell-Yan processes only affect the $q\bar{q}$

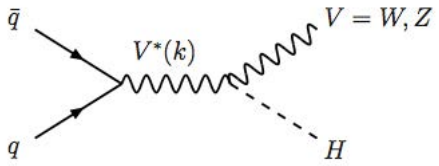


Figure 2.9: Leading-order Feynman diagram for associated Higgs and vector boson production in hadronic collisions.

initial state, consisting in virtual corrections with gluon exchange on the $q\bar{q}$ vertex, quark self-energy correction and emission of an additional gluon, as shown in Figure 2.10. When these corrections are included the LO cross section is modified as

$$\sigma_{NLO} = \sigma_{LO} + \Delta\sigma_{q\bar{q}} + \Delta\sigma_{qg}$$

For an appropriate choice of renormalization and factorization scales ($\mu_R = \mu_F = M_{VH}$) the NLO corrections increase the LO cross section by O(30%) [59].

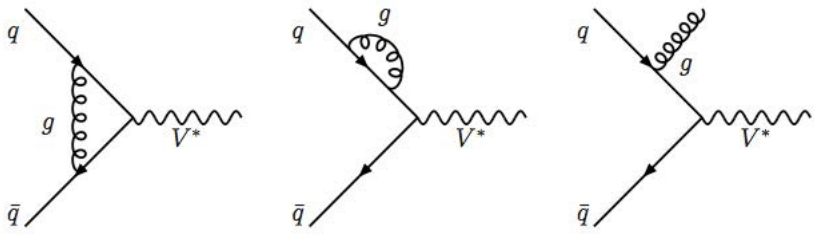


Figure 2.10: NLO QCD corrections to the vector boson-quark-antiquark vertex.

The NNLO QCD corrections to the VH associated production [59, 60, 61] (i.e. the $O(\alpha_S^2)$ contributions to the $pp \to VH$ process) can be categorized in separate sets of diagrams:

1. NNLO QCD correction to the Drell-Yan $pp \to V^*$ process, consisting in two-loop corrections to $q\bar{q} \to V^*$, one-loop corrections to the $qg \to qV^*$ and $q\bar{q} \to gV^*$ processes and tree-level contributions from $q\bar{q}, qq, qg, gg \to V^* + 2$ partons in all possible ways. These corrections affect both $W^{\pm}H$ and ZH production processes (see Figure 2.11(a)), and they increase the NLO production cross section by approximately 1-3%.

2. two-loop $O(\alpha_S^2)$ box+triangle diagrams in $q\bar{q} \to Z^*$, and one-loop vertex diagrams for $q\bar{q} \to gZ^*$ and $qg \to qZ^*$ processes. These sets of diagrams are only allowed for neutral Z bosons, and cannot be drawn for $W^{\pm}$ bosons thanks to the conservation of the electromagnetic charge (see Figure 2.11(b)). In the SM only the top and bottom quarks contribute to these topologies (since the Z boson couples only axially to internal quarks). These contributions are extremely small (at the level of less than 1%) and can therefore be safely neglected.

3. $O(\alpha_S^2)$ diagrams containing closed fermion loops, with the Higgs boson coupling directly to an heavy quark (predominantly top quark), allowed both for $W^{\pm}H$ and ZH production (*t-loop* contributions of the order of 1% on the total cross section, see Figure 2.11(c)).

4. gluon-initiated mechanism $gg \to ZH$ (not contributing for $W^{\pm}H$ production) consisting of triangle and box diagrams (in both cases only heavy quarks loop - bottom and top - give relevant contributions). It's important to note that the contribution of this kind of diagrams can be substantial at the LHC, since the $O(\alpha_S^2)$ suppression of the cross section is in part compensated by the increased gluon luminosity (see Figure 2.11(d)).

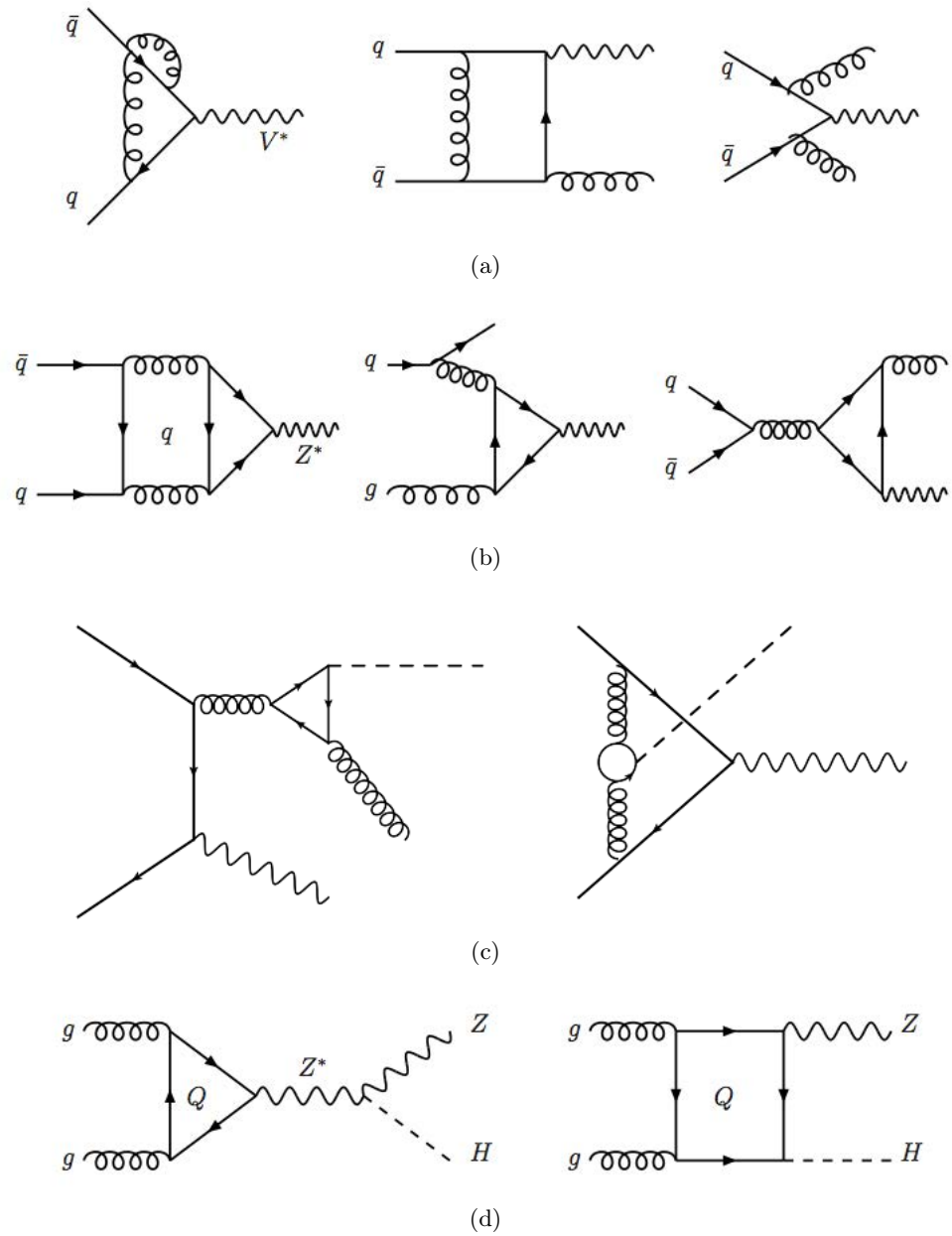


Figure 2.11: NNLO corrections at $O(\alpha_S^2)$ for VH associated production processes.

Loop-initiated $gg \to ZH$ processes at the LHC This particular subset of ZH associated production processes deserves special attention for Higgs boson searches at the LHC: in this Section I discuss some of its main features and I motivate why it can play an important role for the analyses discussed in this research work.

The $gg \to ZH$ production can be considered as a standalone process whose LO QCD contributions start at $O(\alpha_S^2)$, and correspond to the set of diagrams of Figure 2.11(d). Calculations for the inclusive cross section at LO QCD are available and are relatively straightforward; the main feature of the LO QCD cross section is the destructive interference between the two main types of diagrams shown in Figure 2.11(d): box diagrams for $gg \to ZH$ and triangle diagrams for $gg \to Z^* \to ZH$. The two initial-state gluons lead to a rather strong renormalization and factorization scale dependence of about 30%, thus increasing the theoretical uncertainty of the ZH relative to the WH process, where the gg channel does not exist (at this order). Experience from the gluon-fusion process $gg \to H$ (same initial state and color structure of $gg \to ZH$) shows, however, that the LO scale uncertainty drastically underestimates the actual size of the higher-order corrections (as a consequence of the initial state color charge $C_A/C_F = 9/4$, larger than in quark-initiated Drell-Yan type of diagrams).

The NLO QCD $O(\alpha_S^3)$ corrections to this process are beyond the technology currently available for these calculations, because of the presence of massive multi-scale double-box integrals. However a NLO perturbative correction factor ($k_{NLO} = \sigma_{NLO}/\sigma_{LO}$) can be calculated in the limit of infinite top quark mass and vanishing bottom quark mass (known as the 'effective field theory' approach) [62], and used to rescale the LO cross section to an approximate NLO prediction:

$$\begin{aligned} \sigma^{NLO}_{\text{approx}}(m_\text{t}, m_\text{b}) &= \sigma^{LO}(m_\text{t}, m_\text{b}) \frac{\sigma^{NLO}(m_\text{t} \to \infty, m_\text{b} = 0)}{\sigma^{LO}(m_\text{t} \to \infty, m_\text{b} = 0)} \\ &= \sigma^{LO}(m_\text{t}, m_\text{b}) k(m_\text{t} \to \infty, m_\text{b} = 0) \end{aligned} \tag{2.2}$$

This approximation holds well for $m_H = 125$ GeV at the center of mass energies considered at the LHC (from 7 to 13 TeV), but of course worsens for larger $\sqrt{\hat{s}}$ (above the $2m_\text{t}$ threshold, which is not reproduced in the $m_\text{t} \to \infty$ limit) and in specific kinematic regimes (for instance in the boosted Higgs regime).
The validity of this approach has been tested by comparing the LO partonic cross section obtained with the exact top-mass dependence, and with the effective field theory method: the agreement is of the order of 2% (25%) for $\sqrt{s} = 8$ TeV (14 TeV) at $m_H = 125$ GeV, and it deteriorates to the level of 74% (143%) for $\sqrt{s} = 8$ TeV (14 TeV) at $m_H = 200$ GeV. It is worth noting that the kinematic constraint $\sqrt{\hat{s}} > m_H + m_Z$, in the boosted Higgs regime ($p_T^H > 200$ GeV) is sufficient to impose $\sqrt{\hat{s}}$ above the $2m_\text{t}$ threshold, already for $m_H = 100$ GeV, hence the effective field theory approach cannot be used for the partonic cross section calculation in boosted topologies. However the correction factor defined above depends only weakly on the top-quark mass, even outside the convergence region of the large m_t limit [63, 64, 65], therefore this method is applied for the calculation of $k(m_\text{t} \to \infty, m_\text{b} = 0)$.
The result is a $k_{NLO} \sim 2$ factor for $m_H = 125$ GeV at $\sqrt{s} = 14$ TeV, approximately 5% larger at 8 TeV, and indeed not covered by the size of the LO scale uncertainty. As we could expect the uncertainty from renormalization and factorization scale variation on the NLO result is decreased approximately by a factor of two, compared to the LO case. The behavior of this correction factor is very similar to the one observed for gluon fusion

single-Higgs production or Higgs pair production.

Finally, the impact of a threshold resummed cross section for $gg \to ZH$ at NLL has been considered [66], matched to the NLO result: the central value of the inclusive σ^{ggZH} cross section increases by 15(18)% at $\sqrt{s} = 8(13)$ TeV, while the uncertainty from scales variation decreases by a factor of three to four. This translates in an increase of 1.5-2% for the total inclusive σ^{ZH} cross section, with a reduction of scale variation uncertainties of approximately a factor of two.

Based on this description I want to highlight two points relevant for these processes:

- The NLO+NLL σ^{ggZH} contribution to the total σ^{ZH} cross section is of the order of 8-9% at $\sqrt{s} = 7 - 8$ TeV, and it increases to approximately 14% at $\sqrt{s} = 13$. This already sizable contribution becomes even more relevant when considering Higgs production with relatively large transverse momenta. This is a consequence of the threshold effect from the presence of top quark loops which makes the $gg \to ZH$ process especially important for $p_T^H \gtrsim m_H$, with a transverse momentum spectrum fundamentally different from the dominant quark-initiated contribution, as shown in Figure 2.12.

- The gluon-initiated associated production brings a peculiar sensitivity to new physics: through modified Higgs coupling to SM states, through new heavy colored states participating in the loops or through new s-channel pseudoscalars predicted for instance in the SM extensions discussed in detail in Section 2.5 (motivating the $gg \to A \to Zh$ search described in Chapter 7 of this thesis) [67].
  Focusing in detail on modified couplings, the LO $gg \to ZH$ process introduces a dependence on the couplings of the Higgs particle to top quarks, which doesn't appear from the LO quark-induced VH processes. Due to the destructive nature of the interference between triangle and box $gg \to ZH$ diagrams, this process is furthermore sensitive to the relative sign of the Higgs couplings to top quark (k_t) and to the Z vector boson (k_Z). Given the different structure of the p_T^H spectra for gluon-initiated and quark-initiated contributions, a parametrized study of the couplings' dependence becomes therefore very interesting. Note that the p_T^H spectrum is also sensitive to the presence of additional heavy colored states in the loops which would cause new mass-threshold effects in the distribution.

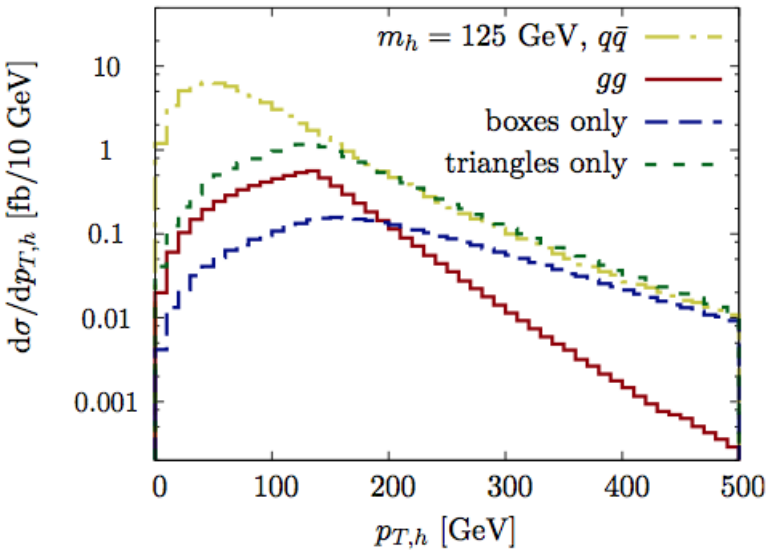


Figure 2.12: Transverse momentum p_T^H spectra for $pp \to ZH$ production at $\sqrt{s} = 14$ TeV [67], showing quark-initiated and gluon-initiated contributions. The dotted blue and green lines represent the triangle and box diagrams contributing to the $gg \to ZH$ process, which interfere destructively to give the red line gluon-induced differential cross section. The top quark mass-threshold effect is clearly visible at $p_T^H \sim m_H$.

The total production cross sections for $pp \to q\bar{q} \to V^* \to VH$ at NNLO QCD and NLO EW accuracy are reported in Tables 2.2 and 2.3 separately for W^+H, W^-H and ZH production modes, at 7,8 and 13TeV, for $m_H = 125.09$ GeV.
The NNLO QCD calculation is performed with `VH@NNLO` [68], Renormalization and factorization scales are set to $\mu = \mu_R = \mu_F = M_{VH}$ and PDFs are taken from the set PDF4LHC15_nnlo_mc PDFs [69]. The NLO EW calculation is performed with `HAWK` [70], with $\mu = \mu_R = \mu_F = M_V + M_H$, using NNPDF2.3QED PDFs [71], which include EW corrections.
The total cross sections are calculated as:

$$\sigma^{WH} = \sigma^{WH,DY}_{NNLOQCD}(1 + \delta_{EW}) + \sigma_{t-loop}$$

$$\sigma^{ZH} = \sigma^{ZH,DY}_{NNLOQCD}(1 + \delta_{EW}) + \sigma_{t-loop} + \sigma^{ggZH}$$

with the σ^{ggZH} contribution also quoted separately in Table 2.3, calculated at NLO QCD accuracy with `VH@NNLO`, including next-to-leading-log (NLL) effects. Photon-induced diagrams are not included in this calculation, and they contribute to O(1%) or less.
The reported uncertainties include independent scales variations in the range $1/3 < \mu_{(R,F)}/M < 3$ and combined PDF$\oplus\alpha_S$ uncertainties at 68% CL from the PDF4LHC15_nnlo_mc PDFs set.

$\sqrt{s}$ [TeV]	σ [pb]	Δ_{scale} [%]	$\Delta_{PDF\oplus\alpha_S}$	W^+H [pb]	W^-H [pb]
7	0.5756	$^{+0.7}_{-0.8}$	$\pm$2.1	0.3685	0.2070
8	0.7009	$^{+0.7}_{-0.8}$	$\pm$2.0	0.4443	0.2566
13	1.373	$^{+0.5}_{-0.7}$	$\pm$1.9	0.8380	0.5313

Table 2.2: Total $W^\pm H$ cross sections at 7,8 and 13 TeV, with Δ_{scale} and $\Delta_{PDF\oplus\alpha_S}$ uncertainties, and separate contributions for W^+H and W^-H production modes.

$\sqrt{s}$ [TeV]	σ [pb]	Δ_{scale} [%]	$\Delta_{PDF\oplus\alpha_S}$	σ^{ggZH} [pb]
7	0.3384	$^{+2.6}_{-2.3}$	$\pm$1.7	0.02792
8	0.4199	$^{+2.8}_{-2.4}$	$\pm$1.7	0.03942
13	0.8824	$^{+3.8}_{-3.0}$	$\pm$1.6	0.1227

Table 2.3: Total ZH cross sections at 7,8 and 13 TeV, with Δ_{scale} and $\Delta_{PDF\oplus\alpha_S}$ uncertainties, and separate contributions for the σ^{ggZH} contribution alone.

From Table 2.2 we can see the charge asymmetry for the $W^\pm H$ cross section at a proton-proton collider as the LHC, due to the different parton distribution functions for quarks and antiquarks in protons. The total cross sections for both $W^\pm H$ and ZH production modes increase by approximately a factor of two when comparing the LHC Run-1 (7 and 8 TeV) and Run-2 (13 TeV) center of mass energies (a larger increase is observed for ZH production due to the presence of gluon-induced processes). Scale uncertainties for the ZH production mode are significantly larger compared to the WH process mainly thanks to the contribution of $gg \to ZH$.

Production in association with heavy quarks: $b\bar{b}H$, $t\bar{t}H$, tH This production mechanisms have the lowest cross sections and rates at the LHC, and are particularly challenging from an experimental point of view, but they present the opportunity to study the direct coupling of the Higgs boson to fermions. In particular the $t\bar{t}H$ and tH production modes provide us the only *direct* probe of the nature of the top-Higgs Yukawa coupling (otherwise accessible only at loop level in production or decay).
It is particularly interesting to note that the $t\bar{t}H$ production cross section increases more rapidly (by approximately a factor of two) than the other Higgs production modes when comparing LHC Run-1 and Run-2 center of mass energies, therefore this channel becomes more and more interesting for searches with $\sqrt{s} = 13$ TeV data from the LHC (although the cross section of the main background, SM top-antitop pair production, is also substantially increasing for higher $\sqrt{s}$).

2.3.2 Higgs Decay Modes

The Higgs boson has no appreciable lifetime, and decays immediately into final state fermions or bosons. The decay into bosons includes two, three and four body bosonic decays, photons pair decay, gluons pair decay or $Z\gamma$ associated decay through loop vertices. Decays to fermions can be divided in leptonic decay modes ($\tau^+\tau^-$, $\mu^+\mu^-$ and e^+e^- pairs) and hadronic decay modes (mainly $b\bar{b}$ and $c\bar{c}$).
The branching ratio to any single mode is expressed as the ratio of the partial width to the total width, where the total width is the sum of all possible partial widths:

$$BR(H \to XX) = \frac{\Gamma(H \to XX)}{\sum_i \Gamma(H \to X_i X_i)}$$

The branching ratios for the different decay modes for m_H around 125 GeV are shown in Figure 2.13. For an Higgs particle with $m_H = 125$ GeV the dominant branching fraction is to bottom quarks pairs at $\sim 58\%$, covering more than half of the Higgs decay width. The searches described in this document focus indeed on this specific decay mode, hence more detail on its features are given in Section 2.5.

The second largest decay rate comes from W boson pairs production, with one of the bosons off-shell. This decay channel is studied at the LHC considering all Higgs production modes, and presents different challenges: the hadronic W decay channels show similar issues to the $b\bar{b}$ case, with large hadronic activity in the detector, while for the leptonic W decay channels we have to treat the presence of neutrinos in the final state.

The decay to gluon pairs takes the third largest branching fraction but it's not distinguishable from the SM background, and it's not studied at the LHC. The following decay mode is τ leptons pair production: similarly to the WW case all production modes are exploited for this channel, and the main difficulties come from the reconstruction of a final state with undetectable neutrinos from the τ lepton decays, and the discrimination between hadronically decaying τ leptons and light-quark jets.

The charm quarks pair production has a very low branching ratio and overwhelming background contributions from QCD SM processes (dominated by gluon-splitting $g \to c\bar{c}$ production), with the additional experimental challenge of tagging hadronic jets from charm quarks. For these reasons this channel hasn't yet been exploited at the LHC.

With a similar rate to the $c\bar{c}$ channel, the $ZZ^\star$ production has a much cleaner experimental

signature from the leptonic decays of the Z bosons: even if the total $H \to ZZ^\star \to 4l$ production rate is extremely low, this channel has almost no background contributions and presents a distinct opportunity to precisely study the Higgs boson's properties (for this reasons it is known as the 'golden channel').

The $\gamma\gamma$ decay channel has a lower decay rate but it's a relatively clean process. Since the Higgs boson has no direct coupling to photons, the $H \to \gamma\gamma$ production is always a loop-induced decay with main contribution from top quark and W boson loops. In this channel the main background contribution arises from SM photon pair production, therefore this search is characterized by a signal peaking over a smoothly decreasing $m_{\gamma\gamma}$ invariant mass spectrum. A similar process, with even lower branching fraction, is the $H \to Z\gamma$ production.

Finally the $\mu\mu$ decay channel presents a very clean experimental signature: a di-muon pair with invariant mass peaking at m_H. The extremely suppressed branching fraction makes it a challenging channel for Higgs boson searches, but it's a prime candidate to probe the nature of the Higgs boson couplings with fermions of the second family.

The decay of the Higgs boson to invisible particles, or the decay channels to Higgs boson pairs are not treated here.

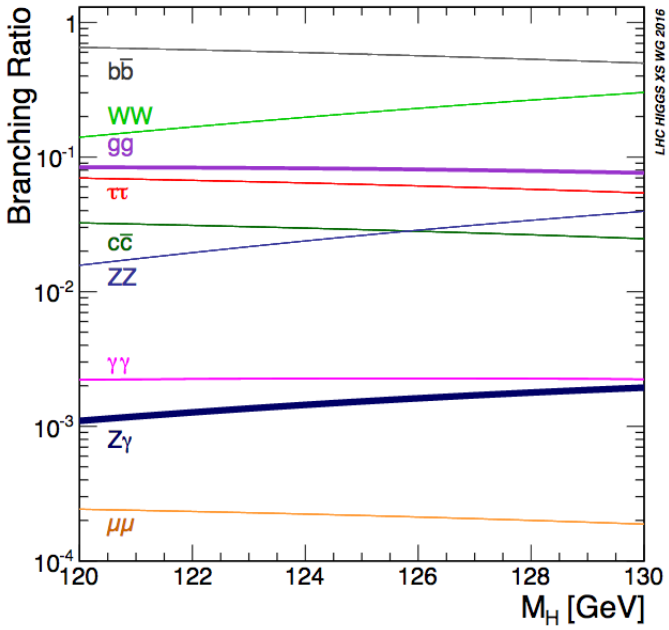


Figure 2.13: Branching ratios for various decay modes of a SM Higg boson, as a function of the Higgs mass m_H [52].

2.3.3 Phenomenology of $pp \to VH, H \to b\bar{b}$+leptons processes

The research work documented in this thesis focuses on the search for a Standard Model Higgs boson with mass of 125 GeV, produced in association with a vector boson V ($W^{\pm}$ and Z) and decaying to a pair of $b\bar{b}$ quarks: $pp \to VH, H \to b\bar{b}$.
In this Section I discuss the phenomenological features that characterize this set of processes: most of the aspects treated here are also valid for the BSM search for a CP-odd A boson via $pp \to A \to Zh, h \to b\bar{b}$ discussed in Chapter 7 (where h is the SM Higgs boson candidate with $m_h = 125$ GeV), however limited to the production in association with Z bosons. Some key aspects of these processes are nonetheless different from the SM Higgs phenomenology, thus they are discussed in detail in Section 2.5.

As described in Section 2.3.2 the $b\bar{b}$ decay channel is particularly appealing since it allows us to consider more than half of the Higgs total with, with a branching ratio of

$$\mathrm{BR}(H \to b\bar{b}) = 0.5809$$

with an uncertainty of 0.65% from missing higher order corrections, and parametric uncertainties of $O(0.75\%)$ of α_S, m_c, m_b and m_t, and a partial width of $\Gamma_{H\to b\bar{b}} = 2.407 \cdot 10^{-3}$ GeV, for the measured value of $m_H = 125.09$ GeV. In addition this decay channel is a unique probe to study the direct coupling of the Higgs boson to down-type quark.

The main production mechanism at the LHC, gluon fusion Higgs production, cannot be exploited since the SM $b\bar{b}$ pair production constitutes an irreducible and overwhelming background with cross section several orders of magnitude larger than the $gg \to H \to b\bar{b}$ one, as can be seen from Figure 2.15. We have to consider other production modes, with lower cross section but cleaner experimental signatures which allow for triggering, identifying and discriminating signal events with more peculiar features. VBF, VH and $t\bar{t}H$ mechanisms have been studied at the LHC in conjunction with $H \to b\bar{b}$ decays: the most significant results and the ones considered in this work are obtained from the VH associated production mode, exploiting the leptonic decays of the V boson to achieve good triggering conditions and strong background rejection.

$VH(b\bar{b})$ signal: main experimental signature The $pp \to VH \to b\bar{b}$ processes, henceforth referred to as $VH(b\bar{b})$, have three possible final states according to the leptonic decays of the vector bosons: the '1-lepton' channel for $W \to l\nu$ decays, the '2-lepton' channel for $Z \to l^+l^-$ and the '0-lepton' channel for $Z \to \nu\bar{\nu}$, represented as LO diagrams in Figure 2.14.
The signature of these processes consists of a pair of hadronic jets originated from the bottom quarks (b-jets) from the Higgs decay, produced in association with charged lepton(s) (electrons, muons and possibly leptonically decaying τ's) and/or large missing transverse energy from undetected neutrino(s) in the detector.

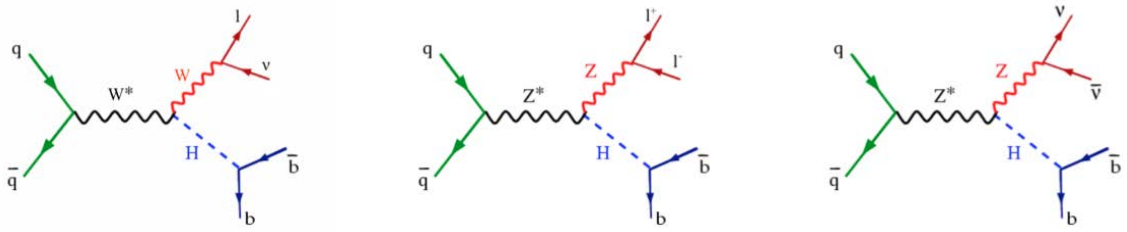


Figure 2.14: Feynman diagrams for the three leptonic decay channels of $VH(b\bar{b})$.

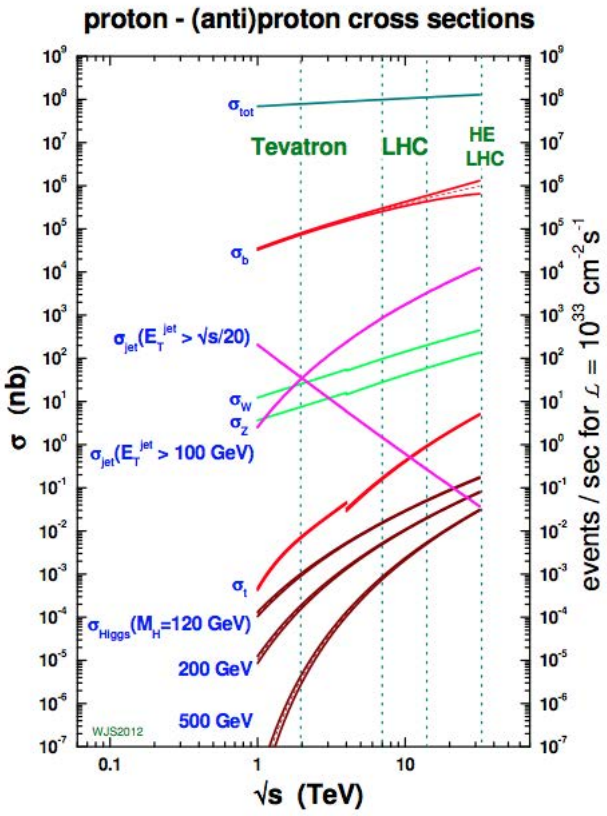


Figure 2.15: Production rates for signal and background processes at hadron colliders as a function of the center of mass energy of the collisions [72].

SM backgrounds: main experimental signature These signal processes have large background from numerous SM processes in large part common, but contributing in different proportions, across the three lepton channels. In this Section I give a general overview of these backgrounds, while a more detailed description of their modeling is left to the following Chapters. The dominant SM backgrounds arise from vector boson production in association with hadronic jets (in particular jets originated by heavy flavor quarks) and top-antitop ($t\bar{t}$) pair production. Albeit with lower rates, also semileptonic diboson production VV (with one boson decaying leptonically and the other decaying hadronically), single-top production and QCD multi-jet, give significant background contributions.

- **V+jets background**: the production of a vector boson decaying to leptons in association with two b-jets ($Wb\bar{b}$, $Zb\bar{b}$) is an irreducible background. For quark-induced initial states the main contribution comes from V boson production in association with gluon-splitting to $b\bar{b}$, while gluon-induced processes (allowed only for $Z + jets$ production) contributes as bbZ associated production: both sets of processes are shown schematically in Figure 2.16.
  Other combinations of quark flavours in the final state can become a background for the $VH(b\bar{b})$ signal due to the misidentification of jets from charm or lighter quarks (c-jets and light-jets respectively) as b-jets.We note that given the proton-proton nature of the LHC hadronic collision, the flavour composition of $Z + jets$ and $W + jets$ events is not identical and it is driven by the composition of the proton's PDFs: for

instance W production in association with a single b-jet plus additional c- or light(l)-jets ($W + bl$ or $W + bc$) has a lower rate than the corresponding $Z + bl$ or $Z + bc$ production.

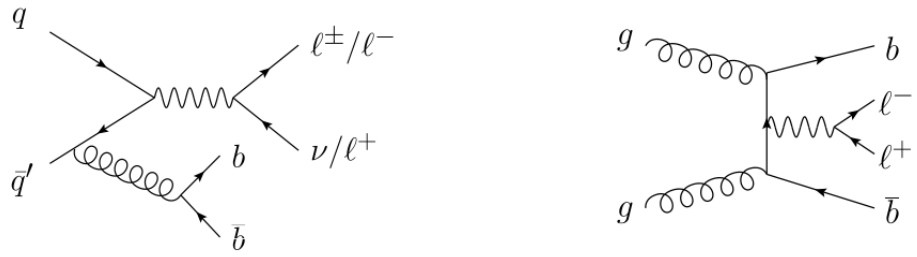


Figure 2.16: Lowest order Feynman diagrams for quark-induced $W+b\bar{b}$ or $Z+b\bar{b}$ production (left) and gluon-induced $Zb\bar{b}$ production.

- **top-antitop pair production**: as can be noted from Figure 2.15 the production of $t\bar{t}$ pairs has a quite large rate, however it is a *reducible* background which can be suppressed with specific analysis cuts (described in full details in the next Chapters). The top quark decays to Wb as shown by the diagram in Figure 2.17, thus the final state signature of a $t\bar{t}$ decay can vary considerably.
  In di-leptonic decays both W bosons decay leptonically to lv pairs, resulting in a final state with two b-jets, large missing transverse energy and two leptons with opposite charge.
  In semi-leptonic decays one W boson decays to leptons, while the second to hadrons, providing a final state with a single charged lepton, significant missing transverse energy and four hadronic jets, two of which are b-jets.
  In fully-hadronic decays both W bosons decay hadronically, producing a final state with six jets, two of which are b-jets.

  For the $VH(b\bar{b})$ signal the di-leptonic and semi-leptonic contributions are the main backgrounds, and can be rejected by vetoing events with multiple additional jets in addition to the two b-jets from the Higgs decay. Since the b-jets are produced from recoiling top quarks, they tend to be less collimated, and the p_T spectrum of the selected b-jets pair is softer than in the $VH(b\bar{b})$ case (in fact, for high $p_T(b\bar{b})$ the main $t\bar{t}$ contribution comes from events with a misidentified c-jet from the W decay, close-by to the b-jet from the same top decay).

- **semileptonic VV production**: diboson processes with semileptonic decays have much smaller rates than $V + jets$ or $t\bar{t}$, but constitute a very important background for the $VH(b\bar{b})$ search as they can mimic very closely the signal signature, as shown by the diagrams in Figure 2.18.
  The main contributions come from $ZZ \to l^+l^-b\bar{b}$, $ZZ \to \nu\bar{\nu}b\bar{b}$ and $WZ \to l\nu b\bar{b}$ processes in which the Z boson plays the role of the Higgs decaying to a $b\bar{b}$ pair with invariant mass lower than the Higgs one ($m_Z < m_H$), but with a cross section approximately 5 times larger than the $VH(b\bar{b})$ one. Considering the relatively poor m_H resolution for $H \to b\bar{b}$ events, this background can bring significant contamination in the region of the Higgs invariant mass peak.

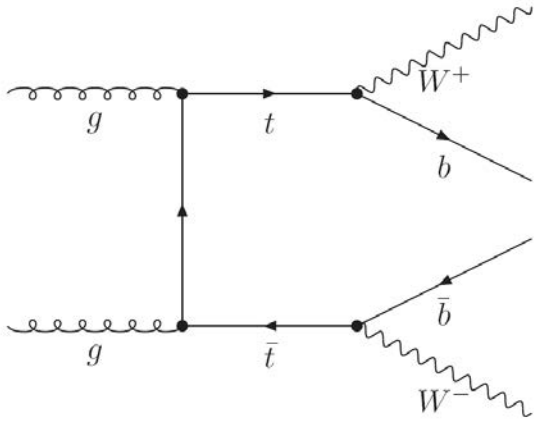


Figure 2.17: Lowest order Feynman diagram for top-antitop pair production and decay.

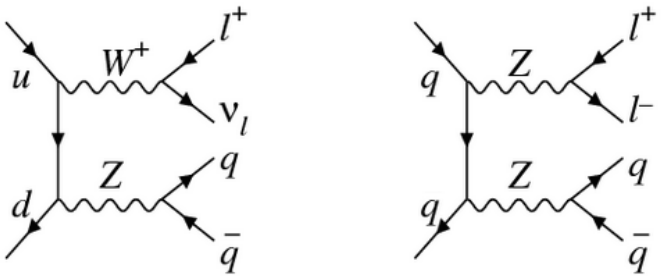


Figure 2.18: LO Feynman diagrams for semileptonic diboson production in WZ and ZZ channels.

- **single-top production**: three main sets of single-top processes contribute as background in these searches: *s-channel* production, *t-channel* production and *Wt-channel* production, shown schematically in Figure 2.19. In all cases the final state contains at least one W boson and at least one b-jet, and can thus reproduce the signal experimental signature. Similarly to the $t\bar{t}$ production case the angular separation between b-jets can be a good quantity to reject the background.

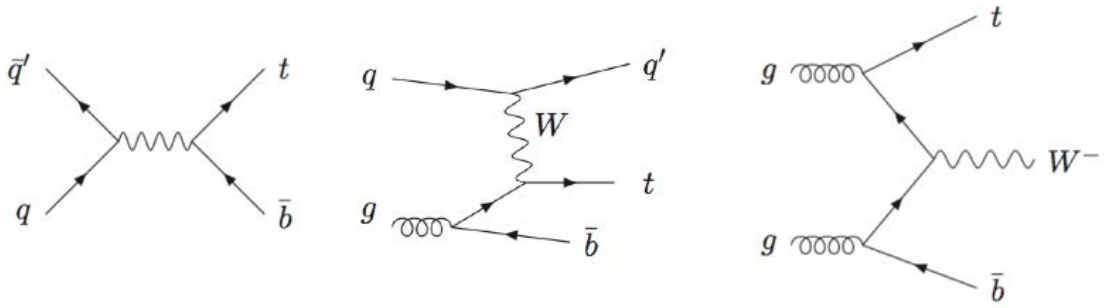


Figure 2.19: Feynman diagrams for single-top production: (a) LO s-channel diagram, (b) NLO t-channel diagram and (c) NLO Wt-channel diagram.

- **QCD multi-jet production**: these processes are strongly suppressed when selecting leptonic final states, but they can still produce a small background due to the extremely large production rate at the LHC.
  Their main contribution arises from the misidentification of hadronic jets as charged leptons, and the genuine production of b-jets (possibly containing semileptonic decays to muons, and resulting in missing transverse energy). Isolation criteria for the selected leptons are crucial to reject this background, which is mostly relevant for the $WH \to l\nu b\bar{b}$ decay mode for low transverse momenta of the W boson.
  This background can be sizable in the $ZH \to \nu\bar{\nu}b\bar{b}$ channel as well, mostly due to multi-jet events with poorly reconstructed hadronic jets faking a genuine missing transverse energy contribution from neutrinos. In this case the background is suppressed with cuts on angular variables, defined from the reconstructed jets, the Higgs candidate dijet-pair, and the missing transverse energy.

2.4 Standard Model Issues and Open Points

In spite of the many features that qualify the SM as a solid and robust theory of elementary particles, there are some critical points that show the limits of its theoretical structure and hint at the fact that new physics may be required in order to describe consistently the known phenomena of particle physics.
Some of this issues come from experimental observations that are not predicted or described by the SM, other rely on theoretical preconceptions about the nature of a fundamental theory of particles: in this Section I give a brief overview of some of these problems.

2.4.1 Higgs/Hierarchy problem

As shown in Section 2.2 it's possible to generate masses for the electroweak bosons and the fermions thanks to the SM scalar sector and, in detail, the introduction of an elementary Higgs field. For the consistency of the model the Higgs mass cannot be too different from the W boson mass, and the experimental discovery of the Higgs boson at LHC in the last year shows that $m_H \sim 125$ GeV.
However there is a complication: the tree-level Higgs mass receives quadratically divergent radiative corrections from interactions with gauge bosons, fermions and self interactions, as show in Figure 2.20 [73, 74].
In detail we can write the corrections as:

$$m_H^2 = (m_H^2)_{\text{bare}} + O(\lambda, g^2, h^2)\Lambda^2$$

where Λ is the next higher scale of the theory (after electroweak scale). If there were no higher scale Λ could be interpreted simply as an ultraviolet cutoff, assuming that m_H is a measured parameter and that $(m_H)_{bare}$ is not a physics observable. However the general assumption is that the SM should be embedded in a different theory valid at a larger energy scale, therefore the integral has a cut off at this new higher scale Λ.
For example the larger theory that includes the SM could be a theory of gravity, valid at the Planck scale $M_P \sim 10^{19}$ GeV, or a grand unified theory (GUT), so that $\Lambda \sim 10^{14}$ GeV. In any cases the natural scale for m_H is $O(\Lambda)$, a much larger value than the measured (and expected) one.
As a consequence there should be a fine tuned cancellation, of the order of $\frac{m_H}{\Lambda^2} \sim \frac{10^2}{(10^{19})^2}$ for the Planck scale, between the bare value of the mass and its correction: this seems highly unnatural and is a strong argument against the validity of the SM alone, that hints at the presence of new physics at higher energy scales.
A viable solution to this issue is the possibility that the W and Z bosons are composite objects: however in this case we would be losing the beautiful (and successful) $SU(2) \otimes U(1)$ gauge symmetry of the SM. There are other models that address this problem, like Technicolor and composite Higgs models [75, 76, 77, 78] (which replace the Higgs mechanism with dynamical mechanisms where the Higgs field is replaced by bound states of fermions) or Supersymmetry [79, 80, 81].

2.4.2 Gravity problem

The SM doesn't describe gravity, while unifying all the other known interactions, even if it's possible to use it together with classical general relativity by hand. This comes from the fact that it is not clear yet how to implement general relativity as a quantum theory and doesn't seem possible to generate the interactions from the SM itself [82].
In addition to the non-unification and non-quantization of gravity there is a different issue

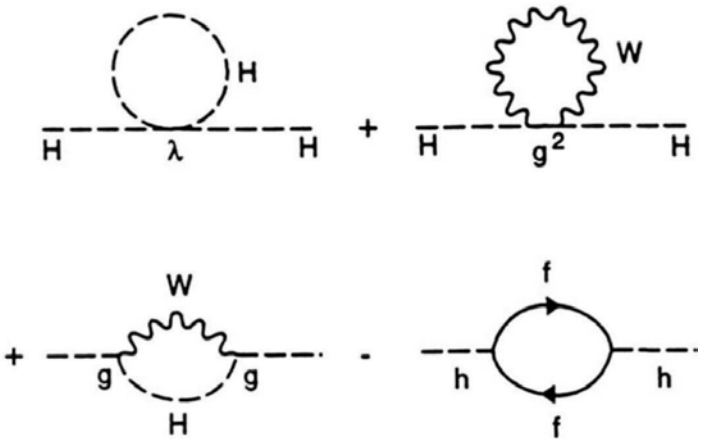


Figure 2.20: Divergent loop corrections to m_H

related to the cosmological constant, recalling that it can be thought of as energy of the vacuum. The energy density induced by SSB in the SM is about 50 orders of magnitude larger than the observational upper limit of the cosmological constant [83]. Again this would imply the presence of high fine tuning between the bare and the SSB value of the cosmological constant (while in principle there is no reason to think that they are related):

$$\Lambda_{\text{cosm}} = \Lambda_{\text{bare}} + \Lambda_{\text{SSB}}$$
$$|\Lambda_{\text{SSB}}| \sim 10^{50} |\Lambda_{\text{obs}}|$$

Superstring theories attempt to unify gravity and may yield finite theories of quantum gravity and other interactions (but it's not clear whether or not they may solve the cosmological constant problem), but the lack of experimental predictions makes impossible to validate these theories so far [84]. This limits the SM range of validity to 10^{19} GeV

2.4.3 Dark matter and dark energy

The SM can describe the behaviour of ordinary matter, but we already know from cosmological measurements and gravitational effects that this only accounts for 4% of the content of the universe [85, 86]. The remaining part is divided between dark matter (22%) which has no electromagnetic interaction and can be detected only through gravitational force, and dark energy (74%) which appears to be associated with the vacuum in space. The dark energy is distributed evenly throughout the universe (in space and in time) so it doesn't have any local gravitational effects and leads to a repulsive force which tends to accelerate the expansion of the universe.
The SM cannot offer any explanation for this different kind of matter, and while some of its extensions (as Supersymmetry) provide good dark matter candidates, these have not yet been observed in experiments.

2.4.4 Matter anti-matter asymmetry

In our universe there is a large predominance of matter over anti-matter: from a cosmological point of view this may cause troubles, since it's difficult to describe the evolution of the universe from a situation of balance, at the Big Bang, to the very asymmetric condition observed today [85].
The SM provide a source of CP violation that can take into account part of this asymmetry,

coming from the presence of a complex phase in the CKM matrix [37] that describe the mixing of different quark flavours, within electroweak interactions. CP violation can break the balance between matter and anti-matter, throughout asymmetry between creation of particles and antiparticles.
Nonetheless the CP violation provided by the CKM matrix of the SM is not enough to justify the observed matter anti-matter asymmetry, therefore there should be other mechanisms which may explain it.

2.4.5 Strong CP problem

This is an additional fine-tuning problem. Considering the QCD Lagrangian introduced previously, it is possible to include an additional term which breaks P, T and CP symmetry [87, 88]. This term would induce an electric dipole moment d_N for the neutron, in disagreement with the stringent experimental limits ($\theta < 10^{-10}$). Such a small value of θ seems again unnatural, and radiative correction due to CP violation in electroweak interactions of the order of 10^{-3} make impossible to set θ directly to zero, therefore it should have an extremely small, but non zero, value. The fine tuning is required in order to cancel this correction against the bare value.

2.4.6 Unification of coupling constants

The coupling constants of the SM interactions are *running coupling constants*: their value depends on the specific energy scale considered, as a result of the renormalization. Looking at the running coupling constants of the fundamental interactions (α_{QED}, α_S, α_{EW}) for increasing energy scales, we can see an interesting feature (as shown in Figure 2.21): the coupling constants seem to converge at high energy, but the intersection is not the same for all of them. In detail:

- α_S and α_{QED} converge at $\Lambda \sim 10^{14}$ GeV
- α_{QED} and α_{ew} converge at $\Lambda \sim 10^{15}$ GeV
- α_S and α_{ew} converge at $\Lambda \sim 10^{12} - 10^{14}$ GeV

This result seems, one more time, unnatural and not consistent with the idea of the SM as an approximation at low energy of a larger theory, valid at much higher scales. If this idea would be correct, the SM coupling constants should be obtained from the decoupling at low energy scales of a single unified interaction. This issue can be solved in different extensions of the SM, as Supersymmetry of Grand Unified Theories [89, 90].

2.4.7 Electromagnetic charge quantization

In the SM the electromagnetic charge is the quantum number of an abelian $U(1)_{em}$ symmetry group, hence there is no motivation that justify the absence in nature of particles with arbitrary charge values: the charge quantization of quarks as $\frac{1}{3}$ of leptons' charges is accidental.
This particular feature observed in nature is therefore not described properly by the SM, while it's possible to derive it from more complex models, as Grand Unified Theories, where the electromagnetic charge originates as the quantum number of a non abelian symmetry group and, as a consequence, has a discrete spectrum of values [91].

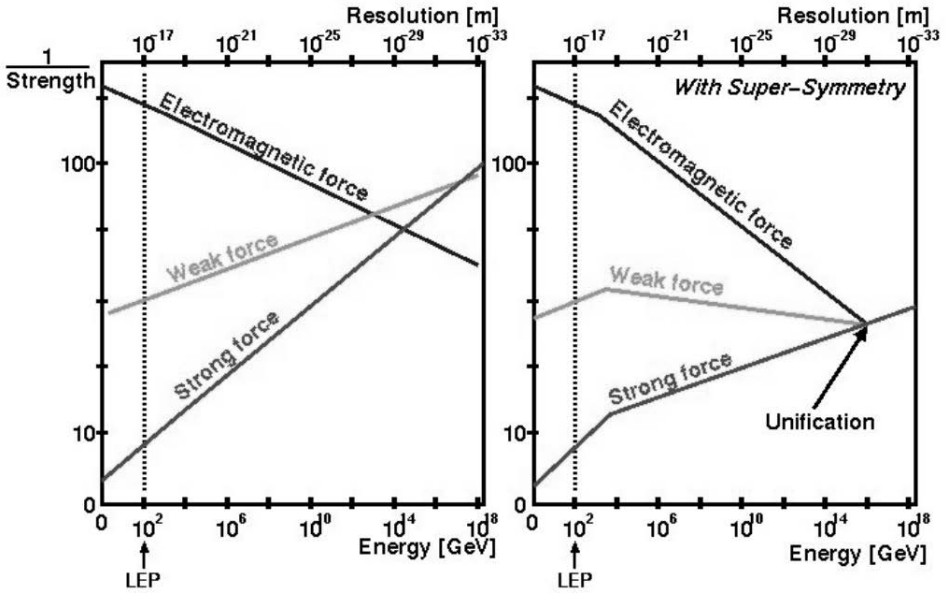


Figure 2.21: Running of the coupling constants in the SM (on the left) and in a Supersymmetric Model (on the right)

2.4.8 Hierarchy of fermionic families

The hierarchy of quark-lepton masses in the charged fermion spectrum is a striking feature of the SM: the fermionic masses range from the top quark mass, of the order of the electroweak scale, down to the electron's mass of about $\frac{1}{2}$ MeV. Not only the SM doesn't give any motivation for the existence of the heavier families, that can be observed only in laboratory and are not found in the ordinary matter, but there is no explanation or prediction of the fermion masses, which vary over six order of magnitudes. There are many possible models of new physics which generate fermionic masses at the loop level. Despite all of these ideas there is no compelling model and none of these gives detailed predictions.

2.4.9 Neutrino masses

In the SM neutrinos are described as massless left handed particles, with no right handed component (following a minimality assumption: since right handed neutrinos don't interact in the SM they are simply omitted). However experimental observations of neutrino mass oscillations show that neutrinos have a small mass [92]. The approximation that $m_\nu << m_f$, where f is any fermion in the SM, is extremely good for most applications, but still it is interesting to consider how mass terms for neutrinos can be consistently introduced in the SM Lagrangian.
There are two possible mass terms for neutrinos, that can be written in the Lagrangian: a Dirac mass (as for the other fermions) and a Majorana mass.
In the first case the existence of a right handed neutrino is assumed, and it becomes possible to build a four-dimensional Dirac field for neutrinos, using left and right handed components. This RH neutrino is not yet observed because it doesn't interact with any other particles of the SM, and therefore it's called *sterile*.
Since the neutrino masses, in this scenario, are introduced as for the other fermions, they must be generated throughout the Higgs mechanism, but this implies that the yukawa coupling constants of neutrinos should be at least 12 orders of magnitude weaker than

that of the top quark. This is an arbitrary assumption that seems quite unnatural and it's widely refused.
The Majorana model is built by using only left handed neutrino fields and their complex conjugates, interpreting the neutrinos as their own antiparticles: right handed neutrinos are still introduced as charge conjugate of left handed neutrinos, which are chargeless. Using these fields it's possible to introduce a Majorana mass term in the Lagrangian, without exploiting the Higgs mechanism.
The Majorana mass can assume arbitrarily large values since no extra symmetry is recovered in the massless limit. Moreover, since Majorana masses violate lepton number conservation, we must assume that the introduced Majorana mass is large enough that lepton number violation effects are compatible with observations.
Writing the mass term for these Majorana neutrinos and diagonalizing the mass matrix to get the mass eigenstates we obtain two eigenvalues m_1 and m_2. The first mass m_1 is extremely heavy (related to the large Majorana mass introduced for right handed neutrinos) and has no observable effects on physics at the electroweak scale; the second one m_2 is suppressed with respect to the typical fermion masses m my a factor $\frac{m}{m_1}$. In this way, light neutrinos arise without the need of assuming unnaturally small values of the Yukawa couplings. On the other hand a new heavy particle has been introduced, (that cannot be observed at the electroweak scale) and allows to validate or disprove this model, called *see-saw* mechanism [93].

Most of the SM issues presented are related to the problem of fine-tuning.
While considering the impact of fine-tuning in different models it is important to remember that it is always possible to fix the fine-tuning problem of a theory by adding to the theory itself an arbitrarily large contribution of new fields, particles or free parameters. However in this way the naturalness of the theory is not improved. While building a new model we must be very careful and find a way to give the correct estimate of the order of magnitude of the fine-tuning, to check if we are proceeding in the right direction.

2.5 Beyond the Standard Model Extensions: Two Higgs Doublet Models

In this Section I discuss the theoretical and phenomenological aspects of the two-Higgs-doublet extensions of the Standard Model (2HDM) which provides an interpretation framework for the search for a CP-odd A boson via $pp \to A \to Zh, h \to b\bar{b}$ discussed in Chapter 7. The outline given in this Chapter is mostly inspired by the comprehensive review from [12].
A comprehensive theoretical overview of 2HDM is beyond the scope of this document, therefore only the key points of the models explored by the research work presented in this thesis are treated in this Section. These models are 2HDM of Type I and Type II with CP-conserving Higgs potential.
Two-Higgs doublet models are one of the simplest possible extensions of the SM which does not entail large radiative corrections to the ρ parameter introduced in Section 2.2.4: these models are built by extending the SM scalar sector from a single scalar $SU(2)$ doublet to two scalar doublets (hence the '2HDM' denomination).
There are several reasons that can motivate the extension to 2HDM, among which the main ones are:

- **supersymmetry**: a single Higgs doublet is unable to generate mass terms simultaneously for charge 2/3 and charge $-1/3$ quarks, since the scalars and their complex conjugates belong to multiplets of opposite chirality and cannot therefore couple together in the Lagrangian. In addition, since scalars and spin-1/2 fields belong to the same chiral multiplets, additional Higgs doublets are needed for the cancellation of anomalies. For these reasons the Minimal Supersymmetric Standard Model (MSSM) contains two scalar doublets.

- **axion models**: in some of these models with an additional singlet at high energy scales, the presence of a second Higgs doublet arises as effective low-energy approximation and allows to eliminate CP-violating terms in the QCD Lagrangian (strongly suppressed phenomenologically) by imposing a global $U(1)$ symmetry.

- **baryogenesis**: 2HDM can provide a mechanisme to generate baryon asymmetry in the sufficient size thanks to a flexible scalar mass spectrum and the presence of additional sources of CP-violation.

2.5.1 Scalar potential in 2HDM

In order to write the general form of the potential of the scalar sector in 2HDM we can rely on some simplifying assumptions: CP is conserved in the Higgs sectors (this allows to distinguish between scalars and pseudoscalars), CP is not spontaneously broken, and discrete symmetries are present, which eliminate from the potential all quartic terms odd in one of the doublets (all possible real quadratic coefficients are usually considered, with a term which softly breaks these discrete symmetries). For two scalar doublets ϕ_1, ϕ_2 with hypercharge +1 the general scalar potential can be expressed as:

$$\begin{aligned} V = \Big(& m_{11}^2 \phi_1^\dagger \phi_1 + m_{22}^2 \phi_2^\dagger \phi_2 \\ & - m_{12}^2 (\phi_1^\dagger \phi_2 + \phi_2^\dagger \phi_1) + \frac{\lambda_1}{2} (\phi_1^\dagger \phi_1)^2 + \frac{\lambda_2}{2} (\phi_2^\dagger \phi_1)^2 \\ & + \lambda_3 \phi_1^\dagger \phi_1 \phi_2^\dagger \phi_2 + \lambda_4 \phi_1^\dagger \phi_2 \phi_2^\dagger \phi_1 + \frac{\lambda_5}{2} [(\phi_1^\dagger \phi_2)^2 + (\phi_2^\dagger \phi_1)^2] \Big) \end{aligned} \tag{2.3}$$

Following the same procedure introduced in Section 2.2.3 we can write the vacuum expectation values for the two doubles and express the two Higgs doublets as an expansion around them. Two complex $SU(2)$ doublets correspond to eight fields:

$$\phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, \; a = 1, 2$$

Of these eight degrees of freedom three are used to generate masses for the $W^\pm$ and Z boson, as in the SM case, while the remaining five can be expressed as physical scalar Higgs fields: one charged scalar $H^\pm$, two neutral scalars h, H^0, and one pseudoscalar A Using these fields we can write from the scalar potential the mass terms for the charged and neutral scalars and pseudoscalars degrees of freedom (not quoted here).
Two of the crucial parameter in the study of 2HDM are the rotation angle β, which diagonalizes the mass-squared matrices of the charged scalars and pseudoscalars, which we can write as:

$$\tan\beta = \frac{v_2}{v_1}$$

(where v_2 and v_1 are the vacuum expectation values of the ϕ_1 and ϕ_2 fields) and the rotation angle α which diagonalizes the mass-squared matrix of the scalar fields. These two parameters (α, β) determine the interactions of the various physical scalar fields with the vector bosons and with the fermions (for given fermions' masses): while in the SM the Higgs branching fractions are determined by the mass of the particles involved in the decay, in 2HDM an additional dependence on α and β comes into play.

2.5.2 Flavour Changing Neutral Currents (FCNC) in 2HDM

An important feature of general 2HDM is the presence of tree-level FCNC, strongly suppressed by experimental measurements. These FCNC arise at tree-level since the Yukawa interactions for same charge quarks are not automatically diagonalized in a simultaneous way by a diagonalization of the mass matrix (contrary to the SM) for the two scalar degrees of freedom $\phi_{1,2}$. For this reason neutral Higgs scalars can mediate FCNC.
The FCNC in 2HDM can be avoided by imposing discrete symmetries on the models: based on the Paschos-Glashow-Weinberg theorem, all right-handed quarks of a given charge coupling to a single Higgs multiplet is necessary and sufficient condition for the absence of tree-level FCNC. This can be realized in two different ways:

1. **Type I** 2HDM: all quarks couple to just one of the Higgs doublet (chosen to be ϕ_2). This can be enforced with a discrete $\phi_1 \to -\phi_1$ symmetry.

2. **Type II** 2HDM: the $Q = 2/3$ right-handed (RH) quarks couple to one Higgs doublet (ϕ_2) and $Q = -1/3$ RH quarks couple to the other (ϕ_1). This can be enforced with a discrete $\phi_1 \to -\phi_1$, $d_R^i \to -d_R^i$ symmetry.

(Supersymmetric models give the same Yukawa couplings as Type II 2HDM, but using continuous symmetries).
Type I and Type II 2HDM are the only ones explored in the search described in Chapter 7, therefore additional models allowing some degree of FCNC (Type III 2HDM) are not treated here.

Discussing Type I and II 2HDM with no CP violation in the vacuum expectation values of the scalar doublets ($v_{1,2}$, which can thus be assumed to be both real and non-negative

without loss of generality, and can be written as $v_1 = v\cos\beta$ and $v_2 = v\sin\beta$.
The five physical degrees of freedom introduced in Section 2.5.1 can be expressed as linear combination of Goldstone bosons:

$$A = \eta_1 \sin\beta - \eta_2 \cos\beta$$
$$h = \rho_1 \sin\alpha - \rho_2 \cos\alpha$$
$$H^0 = -\rho_1 \cos\alpha - \rho_2 \sin\alpha$$
$$H^\pm = -\phi_1^\pm \sin\beta + \phi_2^\pm \cos\beta$$

2.5.3 Relevant phenomenology for Type I and II 2HDM

The results for the search of a massive CP-odd pseudoscalar A boson obtained in Chapter 7 are interpreted in the framework of Type I and II 2HDM, with CP-conserving Higgs potential, under the assumption that the lightest Higgs field h corresponds to the Higgs boson with mass of 125 GeV observed by the ATLAS and CMS Collaboration at the LHC.

The CP-odd nature of the A boson makes its decay to Zh possible (given the CP-even nature of the h Higgs boson, and the assumption of a CP-conserving scalar potential), while $H^0 \to Zh$ decays are forbidden. Clearly in the search for a *neutral* A boson, only Zh decay modes are considered, while charged $W^\pm h$ channels are not included (in principle $H^\pm$ charged Higgs decays to $W^\pm h$ are possible, but with very low branching ratio, subdominant with respect to fermionic decay channels).

Two main different mechanisms are considered for the production of the pseudoscalar A boson:

- **gluon fusion**: the mechanism is similar to the gluon fusion production of a single SM Higgs boson described in Section 2.3.1. In both Type I and Type II 2HDM the coupling of the A boson to top quarks is proportional to $\tan^{-1}\beta$, therefore the coupling is suppressed for large values of $\tan\beta$.
- $b\bar{b}A$ **associated production**: in this production mode the A boson is radiated directly from a bottom quark line. The coupling to bottom quark is proportional to $\tan^{-1}\beta$ in Type I models and to $\tan\beta$ in Type II models, therefore this mechanism is particularly sensitive to the large $\tan\beta$ phase space region for Type II 2HDM.

As a consequence of the different dominant production mode with respect to the SM $VH(b\bar{b})$ phenomenology discussed in Section 2.3.3 we can observed a larger hadronic jets activity for $pp \to A \to Zh, h \to b\bar{b}$ final states (because of the larger initial state color charge for gluon-initiated processes). In addition, when considering $b\bar{b}A$ modes, we expect final states with larger b-jets multiplicity, originated not only from the Higgs decay but also from the production mechanism itself.

The phenomenology of the Zh decay is not substantially different from the SM case discussed in Section 2.3.3: the main differences arise when considering very large values of m_A (above ~ 1 TeV), for which the decay to Zh is considerably boosted, producing Z and h bosons with very large transverse momenta. The treatment of this peculiar final states from an experimental point of view is discussed in detail in Chapter 7.

3 | ATLAS and LHC

The research work documented in this thesis is based on data collected from proton-proton (pp) collision at a center of mass energy ($\sqrt{s}$) of 7, 8 and 13 TeV by the ATLAS experiment [94, 95, 2] at the Large Hadron Collider (LHC) [1]. Data at $\sqrt{s} = 7 - 8$ TeV have been collected during the LHC Run-1 (2011-2012), while data at $\sqrt{s} = 13$ TeV have been collected during the LHC Run-2 (2015-2016).

3.1 Large Hadron Collider LHC

The LHC is a two-ring superconducting hadron accelerator and collider, designed for protons and lead ions collisions, which has been installed in the existing 26.7 km tunnel constructed between 1984 and 1989 for the CERN LEP machine [96], on the French-Swiss border at the European Organisation for Nuclear Research, known as *Conseil Européen pour la Recherche Nucléeaire* (CERN).
The initial LHC project was approved by the CERN Council in December 1994, planning to build the machine in two stages (with a center of mass energy of 10 TeV, to be upgraded later to 14 TeV), and was then modified in 1995 to approve the conctruction of a 14 TeV machine in a single stage. CERN was the perfect site for this accelerator, thanks to the possibility to re-use the LEP tunnel and its injection chain, which guaranteed a considerable cost saving.

The LHC design depends on some basic principles taking advantage from the latest and most advanced available technology. In order to build a particle-particle collider two rings with counter rotating beams are required, unlike particle-antiparticle colliders (as Tevatron) where the beams can share the same space in a single ring.
The tunnel geometry has been designed for the LEP machine and includes eight interaction points followed by long straight sections for RF cavities, as outlined in the schematic view in Figure 3.1. LHC is a proton machine, hence it doesn't suffer from synchrotron radiation losses as much as LEP, and would ideally have longer arcs and shorter straight sections, but using the LEP tunnel without heavy changes was the most cost-effective solution. Only four of the eight crossing points (IPs) are used for interactions, and host the four main LHC experiments: ATLAS, CMS [3], LHCb [97] and ALICE [98]. Three smaller experiments are instead located at or nearby the main interaction points: LHCf [99], MoEDAL [100] and TOTEM [101].

Upstream the LHC there is a series of less powerful accelerators, referred to as the LHC injector chain outlined in Figure 3.2 [102], whose function is to accelerate the protons to an energy of 450 GeV each, at which the beams are injected into the LHC itself. Protons are originally extracted by stripping away electrons from a cylinder of hydrogen gas passing through an electric field, and are first accelerated to an energy of 50 MeV in the linear accelerator LINAC2. They are then injected first into the Proton Synchrotron Booster

(PSB) which accelerates them up to an energy of 1.4 GeV, and then into the proton synchrotron (PS) up to 25 GeV. The beams pass to the Super Proton Synchrotron (SPS) through the Transfer Tunnels (TTs) TT2 and TT10 to finally reach the energy of 450 GeV before being injected into the LHC ring through the Injection Tunnels (TIs) TI12 and TI8. In the LHC the proton beams are accelerated with radio frequency (RF) cavities up to the collision energy. The LHC ramp to nominal energy per beam takes around 20 minutes. The two beams circulating in the LHC collide in the IPs under stable conditions for up to

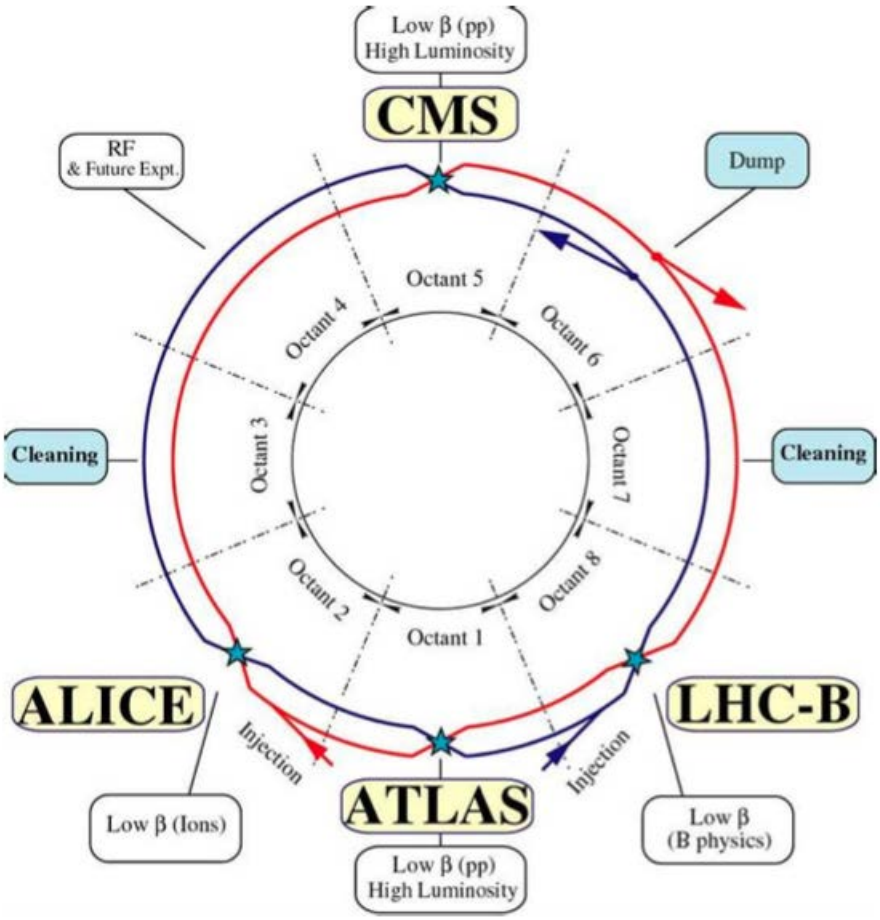


Figure 3.1: Layout of the LHC ring showing the positions of each experiment at the collision points as well as injection, beam dump and beam cleaning regions [103].

approximately 10 hours (depending on the conditions of the run), and are finally extracted from the LHC to inject new beams. The whole process of injection, acceleration, collision and extraction is referred to as a fill.
The protons circulating in the LHC are arranged in *bunches*, and each beam is able to contain 2808 bunches with approximately 10^{11} protons per bunch. Proton bunches fit into *bunch trains*, where the separation between each bunch is referred to as *bunch spacing*. During 2011 and 2012 the bunch spacing was set to 50 ns, while during 2015 and 2016 the bunch spacing has been reduced to 25 ns (except for a short period at the beginning of the 2015 data taking).

LHC magnetic system Opposite magnetic dipole fields are required in each ring to accelerate counter-rotating proton beams: the LHC has been designed as a collider with

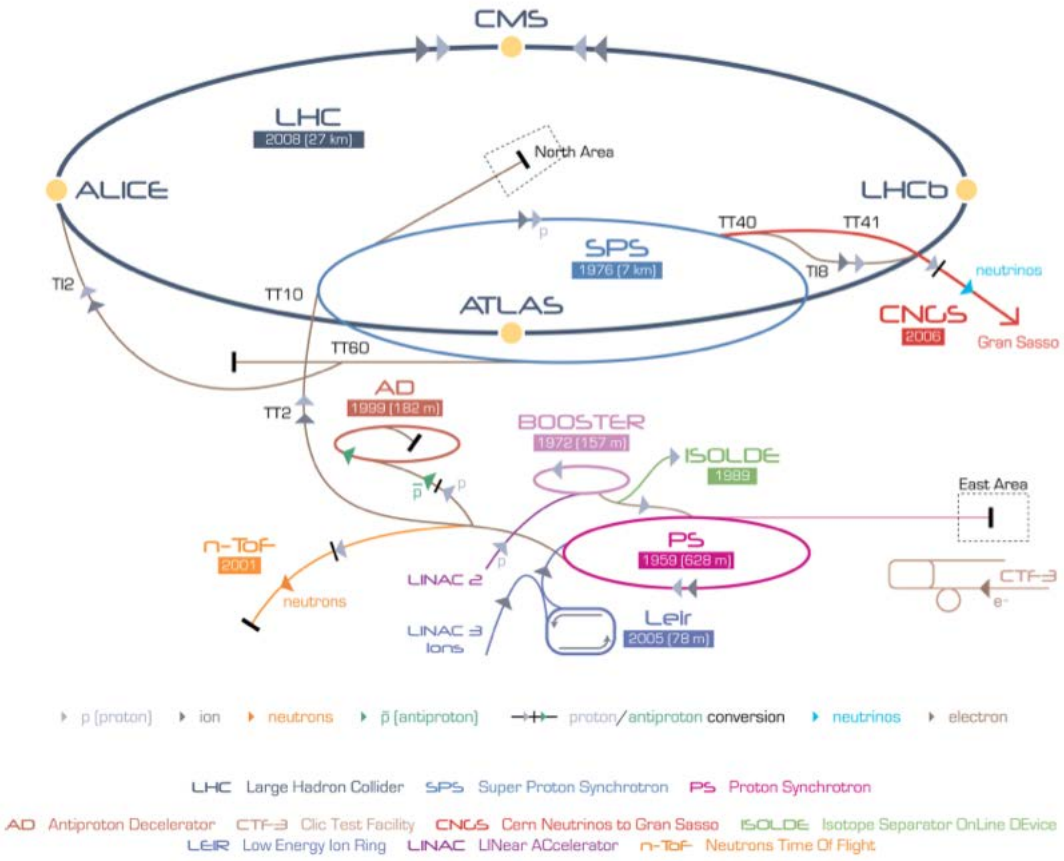
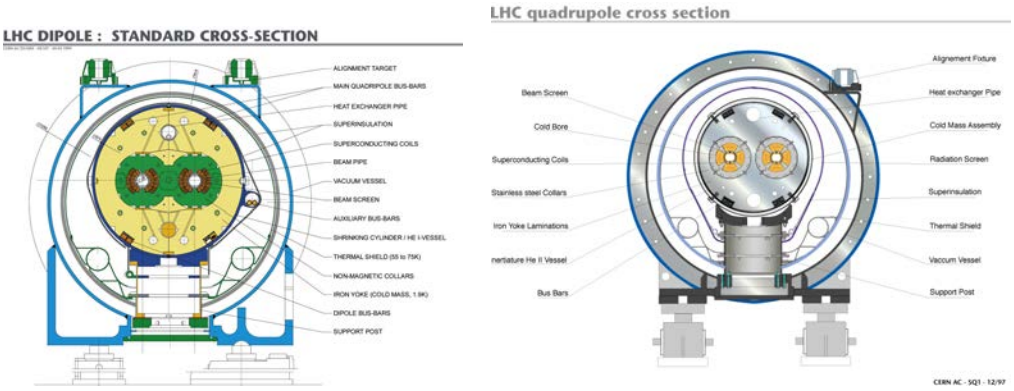


Figure 3.2: Overview of the CERN accelerator complex including the LHC.

separate magnetic fields and vacuum chambers in the main arcs and with common sections only at the intersection regions, where the experimens are situated.
The LHC relies on 1232 superconducting niobium-titanium dipole magnets at the edge of present technology to bend the path of proton beams travelling through the ring: the dipole magnets are cooled to a temperature of 1.9 K using superfluid helium, and operates at fields of 8.36 T generated by an electric current of 11700 A. Additional 392 quadrupole magnets complement the dipoles system by stabilizing and focusing the beams.
Superconducting technology is essential in order to obtain such high magnetic fields and to avoid big resistive losses. Space limitations in the tunnel have led to a two-in-one or twin-bore design for almost all of the LHC superconducting magnets, which accommodates the two beam channels in a common cold mass and cryostat, with magnetic flux circulating in the opposite sense through the two channels, show schematically in Figure 3.3. This makes the magnet structure complicated, especially for the dipoles, for which the separation of the two beam channels is small enough that they are coupled both magnetically and mechanically.

Luminosity In a particle-particle collider the processes under study usually involve a two body initial state which leads to an unknown final state: $|1,2>\rightarrow|f>$. The cross-section for the process $\sigma_{int|1,2>\rightarrow|f>}$ is related to the probability of an initial state interaction leading to the final state $|f>$. In order to write this quantity in term of physical observables a fundamental quantity is the average number of interaction per bunch crossing is:



(a) (b)

Figure 3.3: Cross section of the superconducting dipole 3.3(a) and quadrupole 3.3(b) magnets used in the LHC ring.

$$< \Delta N^{crossing}_{|1,2>\to|f>} >= \sigma_{int|1,2>\to|f>} \frac{N_1 N_2}{\sigma_{xy}}$$

where σ_{xy} is the transverse section of the bunches, N_1 and N_2 are the number of particles in each bunch. Therefore the rate of interaction, with a bunch crossing frequency f, is given by:

$$\frac{dN_{|1,2>\to|f>}}{dt} = \sigma_{int|1,2>\to|f>} \mathcal{L}_0$$

Where $\mathcal{L}_0 \propto \frac{f N_1 N_2}{\sigma_{xy}}$ is a new quantity called instantaneous luminosity, usually quoted in units of $(cm^{-2}s^{-1})$. A detailed definition of $\mathcal{L}_0$ is given as a function of the beam parameters:

$$\mathcal{L}_0 = \frac{N_b^2 n_b f_{ref} \gamma_r}{4\pi\epsilon_n \beta^*} \left[1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right]^{-1/2}$$

where the definition of the parameters and their values for typical 8 and 13 TeV collisions are shown in Table 3.1.

The number of total collisions in a given period of time is given by the integral over time of the interaction rate, that is the product of cross section and integrated luminosity $\mathcal{L}_{int}$:

$$\mathcal{L}_{int} = \int \mathcal{L}_0 dt$$
$$N_{evts} = \sigma_{|1,2>\to|f>} \mathcal{L}_{int}$$

In order to obtain a significant number of events for very rare processes (hence very low cross sections) an extremely high integrated luminosity is required.

The luminosity in the LHC is not constant over a physics run, but decays due to the degradation of intensities and emittances of the circulating beams. The main cause of the

Parameter	Design	2011	2012	2015	June 2016	Definition
Beam Energy [TeV]	7.0	3.5	4.0	6.5	6.5	
$N_b[10^{11}]$	1.15	1.3	1.5	1.1	1.1	Protons per bunch
n_b	2808	1380	1380	2244	2076	Max bunches per beam
f_{rev}			11.245			Revolution frequency [s^{-1}]
γ_r			4260			Relativistic γ factor
ϵ_n [μm]	3.75	2.4	2.4	3.5	3.4	Transverse emittance
β^*	0.55	1.0-1.5	0.6	0.8	0.4	β function at IP
$\theta_c/2$ [μrad]	143	120	146	145	185	Crossing angle at IP
σ_z [cm]	7.55		9.4	10		RMS bunch length
σ^*	16.7	23	19	16		RMS transverse beam size
$\mathcal{L}_0^{max}$ [10^{34} $cm^{-2}s^{-1}$]	1.0	0.35	0.77	0.51	1.01	Maximum peak luminosity
$\mathcal{L}_{int}^{tot}$ [fb^{-1}]	-	5.5	22.8	4.2	15.1	Total integrated luminosity

Table 3.1: Beam parameters for typical 8 and 13 TeV LHC runs.

luminosity decay during nominal LHC operation is the beam intensity loss from collisions. The $\mathcal{L}_{int}$ is given as a function of the characteristic decay time τ_L and length of the run T_{run}:

$$\mathcal{L}_{int} = \mathcal{L}_0 \tau_L \left[1 - exp\left(-\frac{T_{run}}{\tau_L}\right)\right]$$

The LHC has been designed to collide proton beams with a center of mass energy $\sqrt{s}$ of 14 TeV and a luminosity of $10^{34} cm^{-2} s^{-1}$; it can also collide heavy ions (Pb) with an energy per nucleon of 2.8 TeV and a peak luminosity of $10^{27} cm^{-2} s^{-1}$. The nominal design number of protons per bunch is $15 \cdot 10^{11}$, and the time between successive bunch collisions is designed to be 25 ns.
In 2010 the ATLAS detector recorded an integrated luminosity of 45 pb^{-1} at a 7 TeV center of mass energy, which increased in 2011 data taking period to 5.1 fb^{-1}. During the 2012 data taking ATLAS recorded a total integrated luminosity of 21.3 fb^{-1} at $\sqrt{s} = 8$ TeV in condition of stable beams. During 2015 ATLAS collected an integrated luminosity of 3.9 fb^{-1}, while the total integrated luminosity recorded by ATLAS at $\sqrt{s} = 13$ TeV until July 2016 amounts to 13.7 fb^{-1}.
Figure 3.4 shows the total luminosity delivered by LHC and the integrated one collected by ATLAS during the 2011, 2012 and 2015+2016 LHC runs.

Pile-up An important parameter related to the instantaneous luminosity is the mean number of inelastic interactions per bunch crossing, called pile-up events. Pile-up events are mainly soft interactions which constitute a background to the hard interaction targeted by the analysis.
Pile-up events also affect the measurement of the physics objects used in the analysis, worsening the resolution with which we can reconstruct hard-scattering events.
This quantity can be expressed in relation to the instantaneous luminosity:

$$\mathcal{L}_0 = \frac{\mu n_b f}{\sigma_{inelastic}}$$

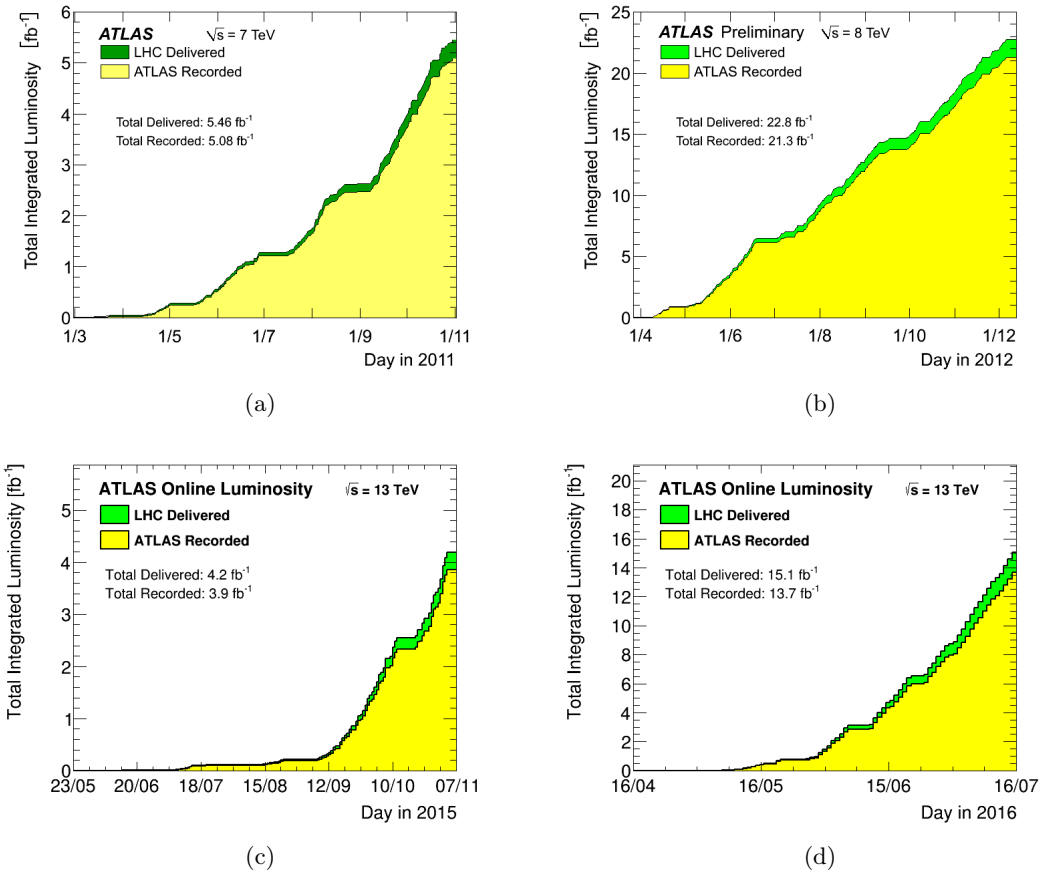
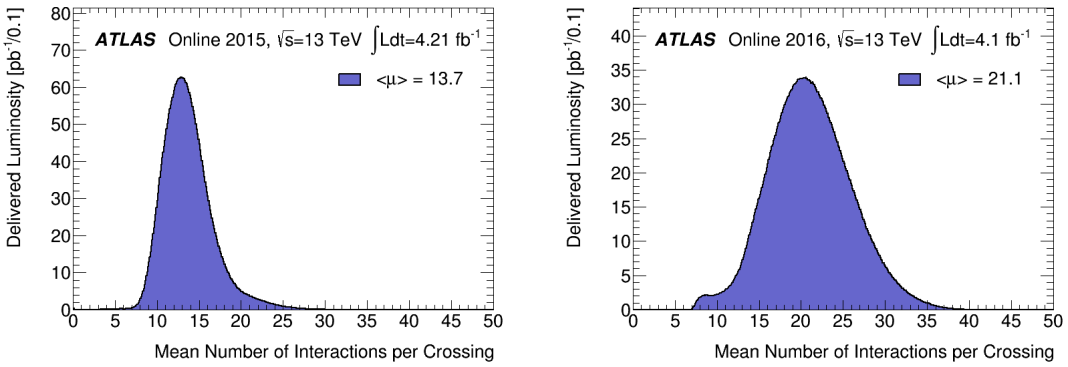


Figure 3.4: Cumulative luminosity versus time delivered to (green) and recorded by ATLAS (yellow) during stable beams for *pp* collisions at 7,8 and 13 TeV centre-of-mass energy in 2011 3.4(a) , 2012 3.4(b) , 2015 3.4(c) and 2016 (a) 3.4(d) [104].

where μ is the number of inelastic interactions per bunch crossing, n_b is the number of colliding bunches and f is the bunch crossing frequency, $\sigma_{inelastic}$ is the total inelastic cross section for *pp* collisions (for instance, $\sigma_{inelastic} = 73$ mb for 8TeV collisions). The number of pile-up interactions per bunch crossing, μ, is therefore proportional to $\frac{\mathcal{L}_0}{f}$ and increases with the peak luminosity. During the 2011 data taking the mean number of inelastic interactions per bunch crossing $< \mu >$ was of approximately 10, increasing to $< \mu > \sim 20$ during the 2012 run, decreasing to $< \mu > \sim 13$ for data collected in 2015 (thanks to the reduced bunch spacing) and reaching $< \mu > \sim 21$ during the 2016 data taking. Figure 3.5 shows the mean number of interactions per crossing for 2015 and 2016 *pp* collisions.

Interactions at hadronic colliders At hadronic colliders, protons (and possibly anti-protons) are accelerated and collide, therefore the objects that interact are not elementary particles, but instead have the composite structure of partons (quarks and gluons). To highlight the main consequences I note a few important remarks:

1. *minimum bias events (pile-up)*: inelastic *pp* collisions are dominated by soft interactions where perturbative QCD calculations are not valid. Most of the time colliding protons produce interactions with very low transverse momentum, generating a large number of soft particles with a uniform distribution in rapidity. Occasionally there is a hard scattering among the parton constituents producing outgoing particles and



(a) (b)

Figure 3.5: Luminosity-weighted distribution of the mean number of interactions per crossing for 2015 3.5(a) and 2016 3.5(b) *pp* collision, showing all data delivered to ATLAS during stable beam conditions. The mean number of interactions per crossing is the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch [104].

jets with high transverse momentum. We are in general more interested to hard-scattering events, with large transverse momentum transfer which makes the QCD perturbative regime valid, however we must have a good understanding and modeling of the multiple soft interactions to assess their impact on the high-p_T physics. To do so we use minumum-bias (MB) data: inelastic hadron-hadron events selected with an experiment's minimum-bias trigger, largely dominated by soft interaction events.

2. *underlying event*: protons are composite objects, therefore any hard scattering process between partons has an underlying event associated, characterized by many objects at low transverse momentum and small angle, which come from the rest of the colliding protons: interactions at hardon colliders are never clean.

3. *partonic center of mass energy* $\sqrt{\hat{s}}$: we are interested in collisions between protons which are able to probe the partonic structure of the hadrons, thus involving a momentum transfer that is large compared to the proton mass. Each parton inside the protons however carries an unknown fraction of the proton energy, hence the center of mass energy of the single partonic interaction $\sqrt{\hat{s}}$ is unknown (but of course related to the center of mass energy of the proton collision $\sqrt{s}$). The measurement of hard scattering processes, with energy scales of more than a few GeV, is the main method by which physics is explored at the LHC.

3.1.1 LHC Experiments

There are 7 main experiments ongoing at the LHC: ATLAS, CMS, ALICE, LHCb, TOTEM, LHCf and MoEDAL. ATLAS and CMS are installed at the high luminosity intersection points, where the highest peak luminosity is achieved, while ALICE and LHCb are low luminosity experiments.
In order to operate successfully in the LHC standard conditions a detector must satisfy some particular requirements regarding geometry, data acquisition and general features. The following points focus mostly on the main requirements for the ATLAS and CMS detectors, but are widely shared by the other experiments:

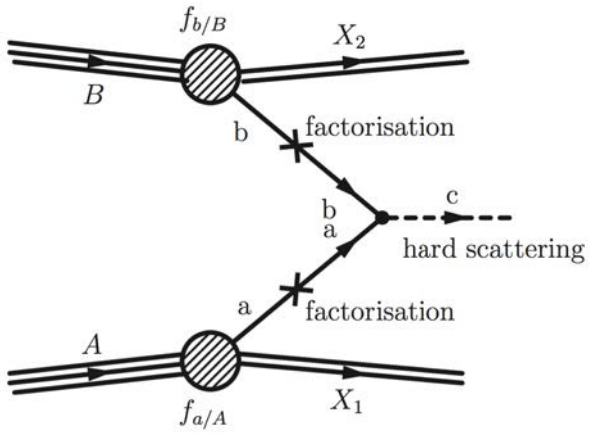


Figure 3.6: Diagram of proton-proton interaction: the partons a and b contribute to the hard scattering, while the rest of the protons X_1 and X_2 generate the underlying event. $f_{a/A}$ and $f_{b/B}$ are the parton distribution functions PDF for the partons a and b.

- **speed of response**: this characteristic is essential to collect the largest fraction of delivered luminosity, given the extremely high LHC collision rate.

- **resistance to radiation**: proton-proton interactions inside the detector cause large amounts of radiation to hit the machine's components, in order to obtain good performances over a long period of time both read-out electronics and sensitive materials must have a strong resistance to radiation.

- **high granularity**: to separate the interesting hard scattering events from pile-up and underlying events it is very important to achieve a fine granularity that provides the good spatial discrimination. Good resolution and reconstruction efficiency for charged particles in the tracking detectors.

- **impact parameter**: precise determination of the interaction vertices and the impact parameter are keys to identify long lived particles as b-jets or τ leptons.

- **energy measurement**: efficiency and accurate energy measurement in the electromagnetic and hadronic calorimeters, with maximal geometric coverage is crucial to fully reconstruct events within the detector.

- **muon detection**: muon identification and reconstruction at high resolution and efficiency have to be achieved for the widest possible range of momenta.

- **trigger system**: the detector must have a fast, efficient and flexible trigger system which allows to reject background and pile-up events and discriminate interesting physics processes.

- **data acquisition system (DAQ)**: in order to save and successfully store all the information coming from the detector the data acquisition system must be able to handle large amount of data.

Keeping in mind these fundamental features that must be the core of the detectors, the 6 experiments are designed to accomplish different goals, performing specific analyses and searches:

ATLAS and **CMS** are the two main and largest LHC experiments, designed as general purpose and high luminosity detectors. They are both conducting studies of SM and BSM physics, and both contributed to the discovery of the Higgs boson particles. ATLAS and CMS have similar structures (which involve however different technologies) and similar expected performances: the ATLAS detector is described in detail in Section 3.2.

LHCb has been built to study the B mesons physics and to measure CP-violation parameters with very high precision. Its main feature is the particular geometry: the detector covers only the forward pseudorapidity region $2 < \eta < 5$ where the probability to observe B mesons' decays is maximal. LHCb is much smaller than ATLAS and CMS, it is installed in a low luminosity intersection since the cross section for B-physics processes is very high.

ALICE has been optimized for heavy ions collisions, in order to study the properties of high energy interactions between nuclei, and properties of the quark-gluon plasma.

TOTEM is designed to accurately monitor the luminosity of LHC, and to study forward particles to focus on physics that is not accessible to the general-purpose experiments: it is able to detect objects produced very close to the LHC beams.

LHCf exploits particles produced in the forward region by collisions in the Large Hadron Collider as a source to simulate cosmic rays in laboratory conditions. The study of the cascades of particles will help to interpret and calibrate large-scale cosmic-ray experiments.

MoEDAL is an experiment designed to search for highly ionizing avatars of new physics such as magnetic monopoles or massive (pseudo-)stable charged particles. It is installed in the same cavern of LHCb.

3.2 ATLAS Detector

The ATLAS detector is designed to satisfy all the requirements listed in the Section 3.1.1 and to act as a *general-purpose* experiment: the physics program of the ATLAS Collaboration includes both precise measurements of SM processes and searches for BSM physics, therefore the detector is optimized for a broad variety of goals. Since the main features of the detector are set by the physics processes under study at the TeV energy scale, it's worth to list some of the main goals of the ATLAS experiment:

- high precision tests of QCD, electroweak interaction and flavor physics
- high rate top quark production, study of top coupling properties and spin
- study of the Higgs boson properties within the discovery decay channels, and search in the remaining channels
- search for BSM boson(s) (Higgs, W', Z', ...)
- search and study of supersymmetric particles in different supersymmetric models
- observation and study of nucleus-nucleus collisions, study of strong interaction properties at high energy densities where the transition to a non-confined color state is possible (*quark-gluon plasma*)

To achieve these goals the detector is designed to be able to identify and reconstruct a wide variety of physical objects, in order to resolve as well as possible the different decay modes and final states of these processes. Figure 3.7 shows a schematic view of the interaction of different types of particles within the ATLAS subdetectors. The interaction with the detector can be briefly outlined depending on the properties of the detected objects and the measured quantities:

- electrically charged particles are detected by bending their trajectories with a magnetic field and exploiting the Lorentz force to measure electrical charge and momentum. Their paths are measured in silicon or gaseous detectors from ionisation effects: this kind of signal is generally referred to as 'hit' and defines the spatial coordinate of the particle interaction. A charged energetic particle can leave several hits in the tracking detectors, from which it's possible to reconstruct a trajectory. Accurate spatial resolution is needed to reconstruct tracks in high luminosity environments as at the LHC, in tracking detectors which are the closest to the interaction points. These detectors are also used to reconstruct the primary collision vertex (PV) and possible secondary interaction vertices (SV) from long lived particles or pile-up events.

- to obtain an accurate energy measurement for charged particles, we exploit the electromagnetic interactions and bremsstrahlung radiation processes. For electrons, and photons above the $2m_e$ energy threshold, this process can produce electromagnetic showers (measured in units of radiation length X_0) which are detected in electromagnetic calorimeters with scintillating material, designed to contain showers from highly energetic photons and electrons.

- strongly interacting particles can also produce hadronic showers (characterized in units of interaction length λ) in heavy materials by inelastic nuclear processes, which are collected in dedicated hadronic calorimeters designed to contain the most part of the hadronic showers. The energy deposits in the calorimeters are the input seeds used to build hadronic jets.

- muons can only be detected through electroweak interactions, and release minimum ionisation energy in the calorimeters, which they can cross retaining most of their energy. They can be detected with specific muon tracking chambers outside of the calorimeters, whose information is combined with that from the inner tracking detectors.

3.2.1 Detector structure

Coordinate system The nominal interaction point inside the detector is defined as the origin of the coordinate system, the z axis is defined as the beam direction, the x-y plane is the transverse plane with respect to the beam direction (where the x axis goes towards the center of the LHC ring, and the y axis goes upward), as shown in Figure 3.8. The detector is divided into two parts: A-side for positive values of z, and C-side for negative value of z; it is also symmetric along the z axis with respect to the interaction point.
The azimuthal angle ϕ is measured around the beam axis, while the polar angle θ is defined starting from the beam axis.
A new angular variable is introduced, the pseudorapidity η, defined as: $\eta = -ln(\tan\frac{\theta}{2})$.
The pseudorapidity is quite useful to describe the geometry of the detector since it transforms nicely under Lorentz boost. Useful pseudorapidity values are:

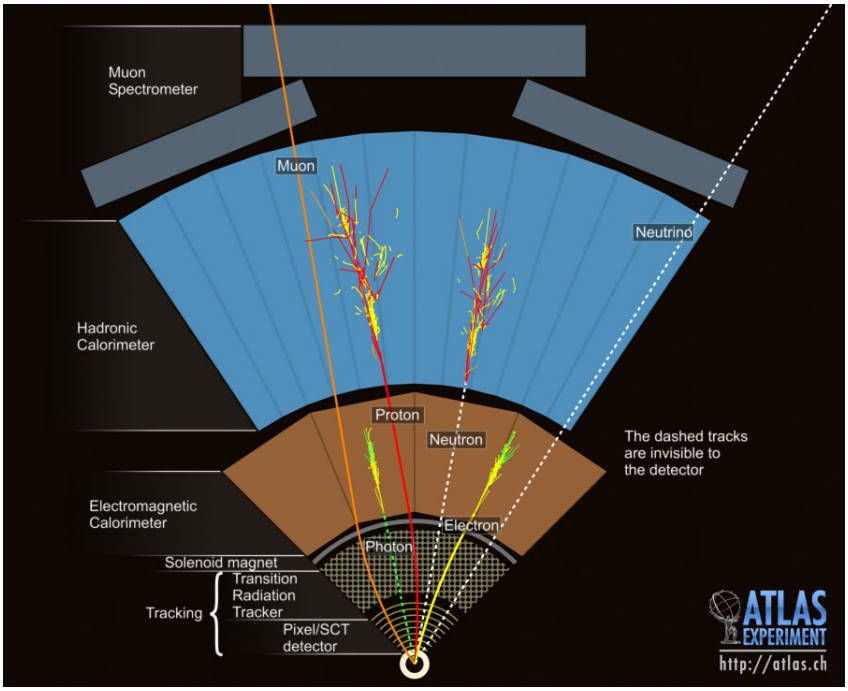


Figure 3.7: Interaction of the different particles with the ATLAS detector active volumes [105].

$\eta = 0$ for a particle in the transverse plane of the detector ($\theta = \frac{\pi}{2}$)

$\eta = \pm\infty$ for a particle with $\theta = 0, \pi$

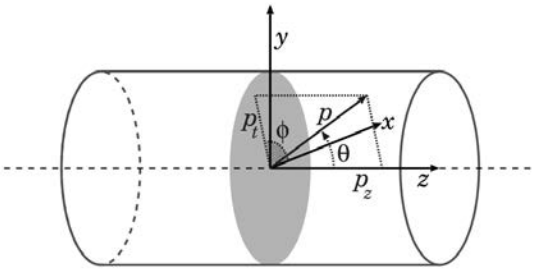


Figure 3.8: Diagram of standard coordinates for a cilindrical detector

The angular distance in the (η, ϕ) plane is defined by $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.

Detector components The ATLAS detector is a very large apparatus, 45 meters long and 25 meters tall, with a cylindrical symmetry, and it is composed of several sub-detector systems designed to accomplish specific tasks and provide particular features, shown in its entirety in Figure 3.9. ATLAS is composed of three main sub-detectors, each optimized for a specific physics goal (and a magnet system [3.2.2] which provides the required magnetic fields):

- *Inner Detector (ID)*[3.2.3]: it takes care of the precise measurement of the charged particles tracks and theirs momentum, allowing to reconstruct the production and decay vertices. It is immersed in a 2T solenoidal magnetic field.

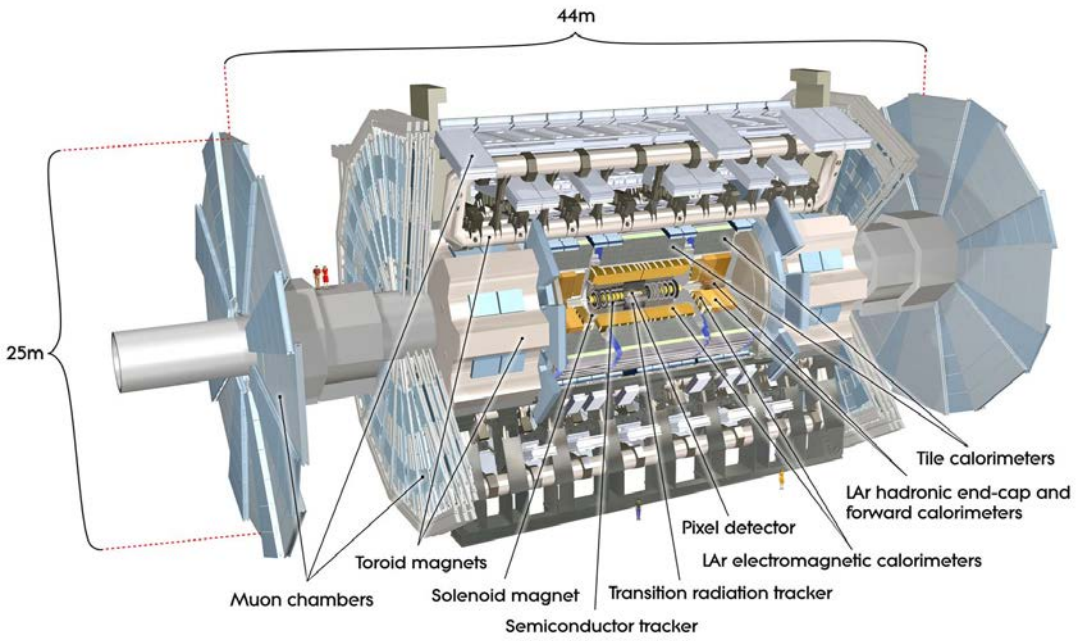


Figure 3.9: Overview of the ATLAS detector with its sub-components

- *Calorimeter System*[3.2.4]: it includes two different calorimeters to measure hadronic and electromagnetic energy losses. The electromagnetic component is a Liquid Argon sampling calorimeter with high granularity and good coverage; in the barrel section the hadronic calorimeter is a scintillator-tile calorimeter, while in the end-caps the hadronic calorimeter uses the Liquid Argon technology too.

- *Muon Spectrometer*[3.2.5]: this is the outermost component of the detector and surrounds the calorimeters. It uses three layers of high precision tracking chambers to measure the muons momentum and reconstruct their tracks within a toroidal magnetic field.

In the next sections I will describe in detail not only these three fundamental sub-components of the ATLAS detector, but also the magnet system that provides the different magnetic fields required by the various parts of the apparatus. Table 3.2 show a summary of the design performance of the ATLAS systems, which will be discussed in detail in the next Sections.

3.2.2 Magnet System

The main feature of the ATLAS magnet system is the presence of two magnetic subsystems, one for the Inner Detector and one for the Muon Spectrometer. A solenoid [106] is installed between the electromagnetic calorimeter and the Inner Detector, and generates the axial magnetic field for the ID. Three sets of toroidal magnets [107, 108] (one for the barrel and two for the endcaps) provide the magnetic field necessary for the muon spectrometer. Therefore the magnet system is composed of four main magnets, three toroids and one solenoid.
This is one of the most remarkable differences with respect to the CMS design, which includes a single solenoid providing a magnetic field of 4 T: while both detectors achieve

Detector system	Design resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
EM calorimeter	$\sigma_E/\mathrm{E} = 10\%/\sqrt{\mathrm{E}} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimeter			
barrel and endcap	$\sigma_E/\mathrm{E} = 50\%\sqrt{\mathrm{E}} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/\mathrm{E} = 100\%\sqrt{\mathrm{E}} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 3.2: Design performance parameters of the ATLAS detector. E and p_T are expressed in units of GeV.

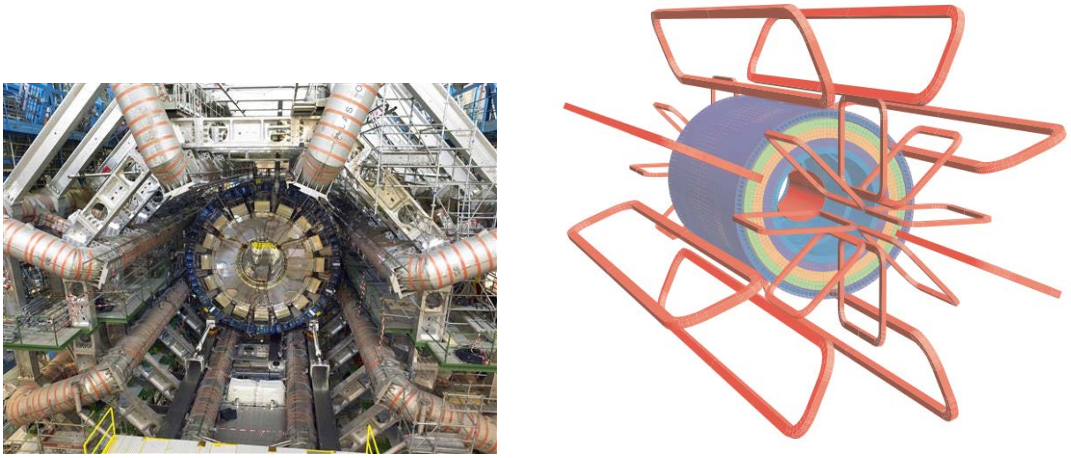
similar muon momentum resolution (ATLAS benefits from the larger radial extension, while CMS from the stronger magnetic field), the ATLAS Muon Spectrometer can provide a standalone muon momentum measurement.

Solenoid The solenoidal magnet is a coil of superconducting material with a 8kA electric current, which provides a magnetic field of 2T for the inner detector, and it is cooled to a temperature of 4.5 K to maintain superconducting properties. It is 5.8 m long and has an inner radius of 1.23 m and an outer radius of 1.28 m (the material it adds in front of the calorimeters amounts to 0.66 radiation lengths to avoid disrupting the calorimeter performances)

Toroids A toroidal magnet has the same structure of a solenoid whose extremities are connected together. The ATLAS toroidal system includes a barrel toroid (composed of 8 separate coils with an electric current of 20 kA) and two endcaps toroids, which together provide a 0.5 T field in the central region and a 1T field in the end-caps.
The toroids are installed just outside the haronic calorimeter (hence outside the calorimeter system) with an inner radius of 4.7 m and an outer radius of 10.05 m and they are 25.3 m long.
The endcaps toroids generate a magnetic field near the proton beam, in order to deflect particles with small polar angles θ. Together with the barrel toroid they provide an almost complete geometric coverage of the magnetic field. Figure 3.10 shows an overview of the toroidal magnets, with the calorimeter system installed within.

3.2.3 The Inner Detector

The Inner Detector [109] is the central part of the detector, immersed in an axial mangetic field of 2 T from the solenoid magnet. The ID must provide high granularity to measure with very good precision tracks and transverse momentum of charged particles down to 0.1 GeV. It also takes care of reconstructing the particles' primary and secondary interaction vertices. The ID also identify electrons in a pseudorapidity range $|\eta| < 2$ from 0.5 to 150 GeV. To achieve this it exploits a discrete combination of high resolution silicon pixels and micro-strip detectors in the innermost part of the tracking volume, and straw tubes in the outermost part, combining a smaller number of high precision measurements close to



(a) (b)

Figure 3.10: ATLAS detector after the installation of the 8 toroidal coils 3.2.2, geometry of the magnet system and the hadronic calorimeter steel 3.2.2 showing the eight barrel toroid coils and the endcap coils interleaved. The solenoid is inside the calorimeter volume.

the interaction point with a higher number of measurements in the outer zone. Table 3.3 outline the ID geometry with all its components, which will be discussed in detail in the next Sections.

The important parameters that can be measured in a tracker and characterize the helix motion of a charged particle in a uniform field are:

- transverse momentum p_T
- azimuthal angle ϕ defined from the momentum as $\tan\phi = p_y/p_x$
- polar angle θ, defined from the momentum as $\cot\phi = p_z/p_T$
- transverse impact parameter d_0: distance in the xy plane between the track closest point to the z axis and the z axis itself
- longitudinal impact parameter z_0: z coordinate of the closest approach point to the z axis

The amount of material in the ID necessary to achieve high granularity and including readout and cooling system is quite sizable: a map of its impact in the ATLAS simulation is shown in Figure 3.11 (excluding the IBL, the additional inner layer inserted after the LHC Run-1). The overall ID momentum resolution (achieved in Run-1 before the IBL insertion) is:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\% \cdot p_T \oplus 1\%$$

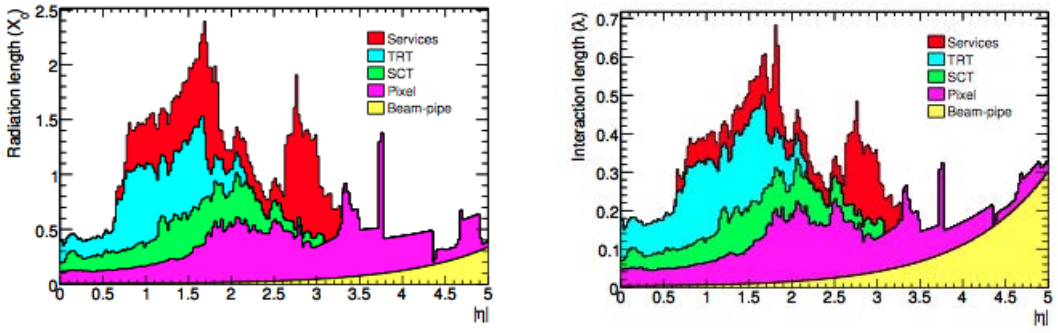
Pixel The pixel detector [110] is the nearest detector system to the interaction region, where the particle flux is the highest, hence it must have a good resistance to radiation. It has a very high spatial resolution and covers a pseudorapidity region $|\eta| < 2.5$.
The pixels have different structures in the barrel and end-cap regions: in the barrel they are

System/Subsystem		Radial extension (mm)	Length (mm)
Overall ID envelope		$0 < R < 1150$	$0 < \|z\| < 3512$
Beam pipe		$29(25) < R < 36(29)$	-
IBL	Overall envelope	$31.0 < R < 40$	-
	Sensitive barrel	$R > 25.7$	$0 < \|z\| < 232$
Pixel	Overall envelope	$45.5 < R < 241$	$0 < \|z\| < 3092$
3 cylindrical layers	Sensitive barrel	$50.5 < R < 122.5$	$0 < \|z\| < 400.5$
2 x 3 disks	Sensitive endcap	$88.8 < R < 149.6$	$495 < \|z\| < 650$
SCT	Overall envelope	$255 < R < 549$ (barrel)	$0 < \|z\| < 805$
		$251 < R < 610$ (endcap)	$810 < \|z\| < 2797$
4 cylindrical layers	Sensitive barrel	$299 < R < 514$	$0 < \|z\| < 749$
2 x 9 disks	Sensitive endcap	$275 < R < 560$	$839 < \|z\| < 2735$
TRT	Overall envelope	$554< R < 1082$ (barrel)	$0 < \|z\| < 780$
		$617 < R < 1106$ (endcap)	$827 < \|z\| < 2744$
73 straw planes	Sensitive barrel	$563 < R < 1066$	$0 < \|z\| < 712$
160 straw planes	Sensitive endcap	$644 < R < 1004$	$848 < \|z\| < 2710$

Table 3.3: Geometry parameters of the Inner Detector - values between parenthesis refer to the updated geometry after the insertion of the IBL.

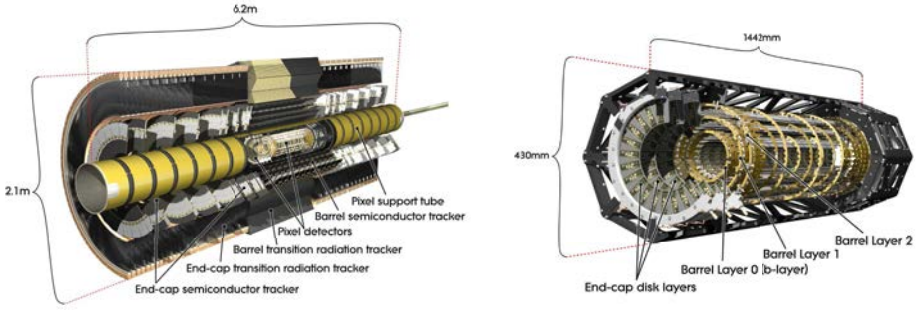
arranged in three cylindrical layers at a distance of 5.05, 8.85, 12.25 cm from the center of the beam pipe. In the end-caps the pixels are divided into three disks: the detector design is optimized to reconstruct high resolution tracks using three points within the pixels.
All the pixels are segmented in R - ϕ and z, with a minimum active size of 50μm in R - ϕ by 400 μm in z, achieving an accuracy of the order of 12 μm in R - ϕ and 115 μm in z.
The active region of the pixel detector is shown in Figure 3.12 with the three cylindrical layers and two identical end-cap regions composed of three disks. It is worth noting that the detector modules used to build the cylindrical part and the disks are identical. The total number of pixels is of about $67 \cdot 10^6$ for the barrel and $13 \cdot 10^6$ for the end-caps, to cover a global area of 1.7 m^2.

Semi-Conductor Tracker SCT The SCT [111, 112] follows the same principles of the pixel detector, applied to a larger area and using micro-strip detectors instead of pixels. As the pixel detector also the SCT is divided into barrel and end-caps regions and covers the pseudorapidity range $|\eta| < 2.5$. The SCT is designed to provide 4 additional track points in the barrel to contribute to the momentum measurement (and the tracks reconstruction), therefore it is organized in 4 layers approximately at 30, 37, 44, 51 cm from the interaction point, each of ones is composed of 2 microstrip sensors. These two strip layers are assembled together with a relative angle of 40 mrad in order to obtain a stereoscopic measurement in the R - ϕ plane. The end-cap parts of the detector are organized in 2 $\times$ 9 disks of micro-strips.
Each microstrip is 6.4 cm long with a pitch of 80 μm, the global area covered is of 61 m^2, with a spatial resolution of 16 μm in R - ϕ and 580 μm along the z direction.



(a) (b)

Figure 3.11: Material distribution of the inner detector including services and beampipe, expressed in terms of radiation length X_0 3.2.3 and interaction length λ 3.2.3



(a) (b)

Figure 3.12: Cut view of the entire ID 3.12(a) and schematic overview of the pixel detector 3.12(b).

Transition Radiation Tracker The TRT [113, 114] covers the pseudorapidity region $|\eta| < 2$ and provides information only in the R - ϕ plane with a resolution of 130μm.
The TRT is the outermost part of the ID and has the worse spatial resolution among the ID sub-detectors, but provides up to 36 additional track points which are fundamental for tracks reconstruction, particle identification and momentum measurement. The TRT is designed as a drift chamber which measures the charge drift time and also the transition radiation, identifying electrons thanks to transition radiation photons.
The drift tubes are filled with Xenon gas: there are about 50000 barrel tubes, which are parallel to the beamline, 150 cm long, and with a diameter of 4 mm can work at very high rates; each barrel tube is divided into two sections with two different outputs. In the end-cap regions there are about 320000 tubes disposed radially in wheels.

Inner B-Layer (IBL) The Insertable B-Layer [115] (IBL) is a fourth layer added to the present Pixel detector between a new smaller beam pipe and the current inner Pixel layer (B-layer), which has been inserted in the ATLAS detector in the period between the LHC Run-1 and Run-2 data taking (known as Long Shutdown 1 or LS1).
The IBL consists of 14 staves arranged around the beam pipe with a tilt of 14°, allowing for an azimuthal overlap of the staves of $\Delta\phi = 0.18°$ (as shown in Figure 3.14 and 3.15)

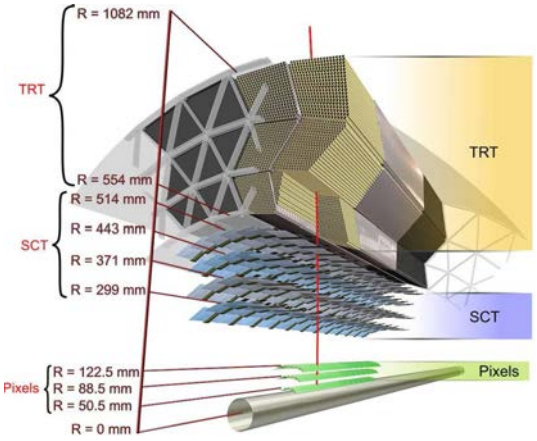


Figure 3.13: Overview of the general ID structure.

which ensures complete coverage in ϕ and to compensate for the Lorentz angle in the 2T solenoidal magnetic field of the ATLAS detector. Since there is no overlap in the z direction, the design tries to minimize the gap between sensor areas, which is of ~ 250 μm.
The IBL is the innermost layer of the pixel detector, with an average distance of the staves from the center of the beam pipe of 33.25 mm, and has been designed to achieve very good spatial resolution with special care to be resistant to high radiation. Each stave is

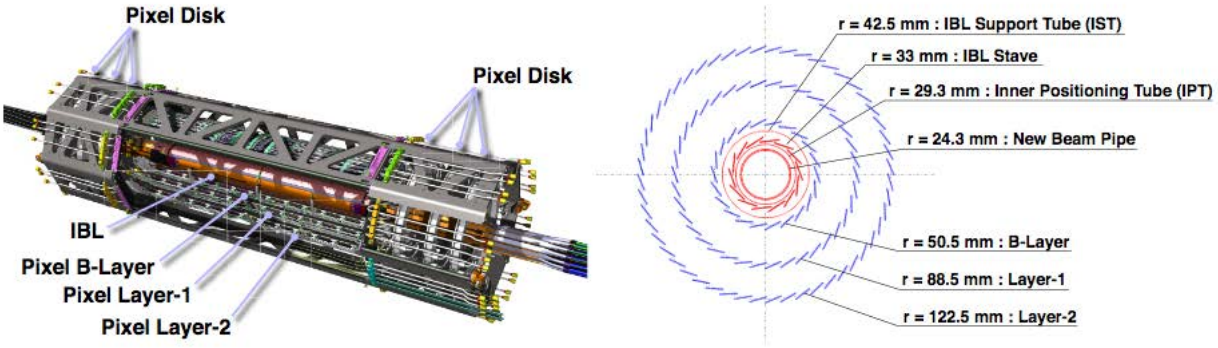


Figure 3.14: Schematic view of the ATLAS 4-Layer pixel Detector for Run-2.

64 cm long with a coverage of $|\eta| < 2.9$, and consists of 12 planar double-chip modules in the central part and two times 4 3D sensor single-chip modules. The planar sensors use a technology similar to the rest of the pixel detector, with some changes to improve radiation hardness given the very short distance from the beam line. These sensor are built in n-in-in technology and are 200 μm thick with a 4×2 cm^2 area. The 3D sensor use n-in-p silicon technology, and are 230 μm thick with a 2×2 cm^2 area. To allow lower bias voltage with respect to the planar sensor, and improve the radiation hardness of the 3D modules, electrodes passing through the bulk are used. The pixel size is reduced with respect to the outermost pixel layer from 50×400 μm^2 to 50×250 μm^2.
The IBL is a very light detector with a material budget of only 1.9% of a radiation length (60% of the Pixel B-layer). The installation of the IBL required additional radial space between the B-Layer and the beam pipe: a new beryllium beam pipe with reduced inner diameter (47 mm) has replaced the previous one to fit inside the IBL.

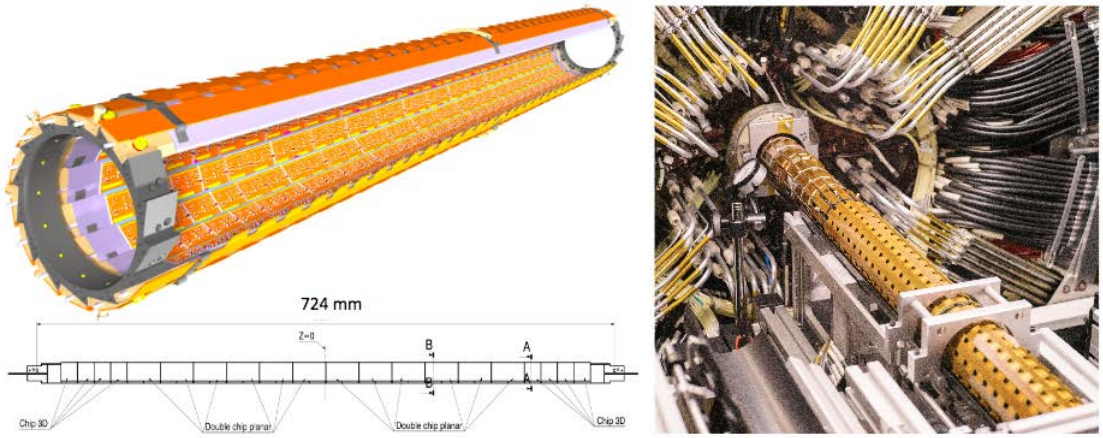


Figure 3.15: (left) 3D rendering of the IBL layer with its 14 staves, (right) photo of the IBL during its installation inside the Pixel Detector in the ATLAS cavern.

The main gains from the insertion of the IBL can be summarized:

- tracking robustness: loss of data in the Pixel B-layer (from failures, radiation damage or inefficiencies) strongly deteriorates the impact parameter resolution, and thus the b-tagging. With the IBL the full b-tagging efficiency can be restored even in case of complete B-layer failure.

- luminosity effects: the Run-1 Pixel detector was designed for a peak luminosity of $1 \times 10^{34}\text{cm}^2\text{s}^{-1}$ which has been exceeded during the LHC Run-2 (the peak luminosity recorded by ATLAS up to October 2016 corresponds to $13.2 \times 10^{33}\text{cm}^2\text{s}^{-1}$ [116]), leading to high occupancy from pileup events which will affect the B-layer in particular limiting the b-tagging efficiency. The IBL, with comparably low occupancy, helps in preserving good tracking performance despite luminosity effects.

- tracking precision: the very close distance to the IP allows to improve the accuracy of track impact parameter reconstruction, vertexing and hence b-tagging performance.

It's worth noting how the IBL design targets the improvement of the b-tagging performance for the LHC Run-2, a crucial aspect for the analyses described in this document.

Inner detector: Upgrades for LHC Run-2 After the LHC Run-1 several upgrades have affected the ATLAS Inner Detector, driven by the insertion of an additional inner layer (IBL) which made the extraction of the Pixel detector necessary. Both the Pixel and the SCT detectors have been made more robust with new service panels (Pixel), additional readout drivers and faster electronics with improved optical transmitters (SCT).
To cope with gas leaks in the TRT (abundant during Run-1 due to corrosion of outflow pipes) it has been decided to use an Argon-based gas mixture (less expensive but with reduced absorption of transition-radiation) in places where the Xenon losses were too high and not affordable, adding remote gas regulation.

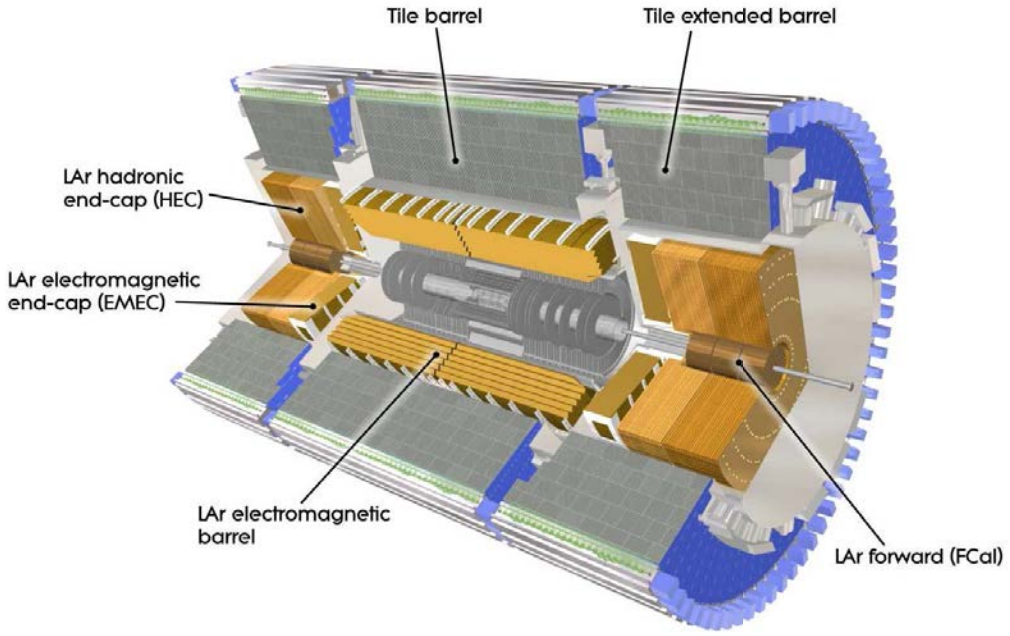


Figure 3.16: Cut view of the ATLAS electromagnetic and hadronic calorimeters

3.2.4 The Calorimeter System

The ATLAS calorimeter system consists of two separate subsystems: the electromagnetic and the hadronic calorimeters. The system is designed to meet precise physics requirements, mainly to provide the best possible resolution for high-energy photons, electrons, jets and missing transverse energy.
To obtain a good missing transverse energy resolution, the main requirement for a calorimeter is hermeticity. Good containment for jets and missing transverse energy is also important, and it is achieved because of the calorimeter thickness, allowed by the large radius toroidal magnet muon system. The total thickness of the electromagnetic calorimeter is larger than $22X_0$ in the barrel region and larger than $24X_0$ in the endcap region. The total interaction length of the entire system is approximately 10λ, with a peak value of 11λ at $\eta = 0$.
The ATLAS calorimeters are designed to limit the punch-through of hadrons into the muon system, enhancing the performance of the muon chambers, and to make sure that electromagnetic and hadronic showers are contained within the respective calorimeter system.
In Figure 3.17 are shown the calorimeter thickness in radiation lengths and interaction lengths. Thanks to these features the electromagnetic and hadronic showers coming from particles produced with energies of the TeV scale can be resolved and contained in the calorimeters, for jets with a momentum of up to ~ 1.5 TeV.
The full calorimetric system covers a pseudorapidity range of $|\eta| < 4.9$: the larger η coverage of the calorimeter system compared to the Inner Detector is due to the importance of geometric coverage for the missing transverse momentum reconstruction.

The electromagnetic calorimeter exploits liquid argon (LAr) as active material and lead as absorber; the hadronic calorimeter uses LAr technology with a copper absorber in the forward region (HEC) and active scintillating tiles with steel as absorbing material in the barrel (TileCal). The forward calorimeter (FCAL) uses LAr and copper for the elec-

tromagnetic interactions and tungsten for the hadronic part. The general scheme of the calorimeter system is outlined in Figure 3.16.

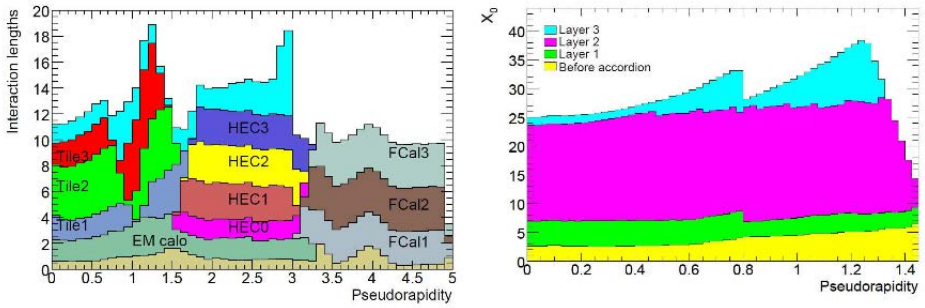


Figure 3.17: Cumulative amounts of material in the ATLAS calorimeters in interaction length (left), and radiation length (right)

Electromagnetic Calorimeter The ATLAS electromagnetic calorimeter [117, 118] is composed of two parts, a barrel section that covers $|\eta| < 1.4$ and an endcap section where $1.4 < |\eta| < 3.2$. It is a sampling calorimeter which uses Liquid Argon (LAr) as active material and lead as absorber, disposed in an accordion geometry with full coverage in ϕ and no crack regions: this particular geometry (shown in Figure 3.18) provides uniform performances through the detector, a good energy and spatial resolution and low electronic noise with high resistance to radiation. The disadvantage of this layout is the long time of response (about 400 ns), longer than the delay between two interactions. The ATLAS EM

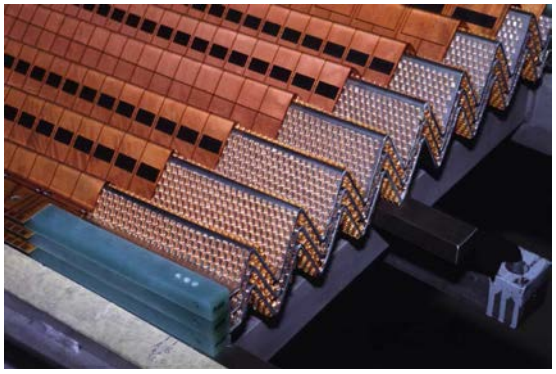


Figure 3.18: Photo of one component of the LAr calorimeter

calorimeter follows the general principles of electromagnetic calorimeters: when a charged particle enters the calorimeter it produces a shower of electrons and photons, the LAr calorimeter then measures the shower energy collecting the charge at electrodes. Due to the presence of lead absorber some energy is lost, but the global charge collected is proportional to the energy of the particle that initiated the shower, therefore, by calibrating the calorimeter, it is still possible to measure this energy. In the LAr calorimeter the absorber is composed of lead plates with a thickness of 1.13 or 1.53 mm, glued to stainless steel sheet 0.2 mm thick, to provide mechanical strength.
The barrel region is divided into two parts, for positive and negative z values, each 6.4 m long with inner radius of 1.15 m and outer radius of 2.25 m. In order to reduce the

material budget in front of the calorimeter, the central solenoid and the EM calorimeter barrel share a cryostat and are cooled together.
There are four calorimeter longitudinal layers with different features, which provide discrimination power between electrons, photons and pions, and provide useful informations regarding the longitudinal shape of the shower: presampler, strips, middle and back.

- The **presampler** is a very thin module, 1.1 cm thick, situated inside the solenoid magnet and before the cryostat. It doesn't contain any absorber material and is built with very high granularity in order to measure the energy lost by the particles before entering the calorimeter.
- The second layer is composed of $\Delta\eta = 0.0031$ **strips** with high granularity that allow for charged and neutral pion separation, and to resolve adjacent energy deposits. Together with the presampler it adds up to 6 X_0.
- The **middle** part is the most thick, extending up to 22 X_0 and segmented in towers of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, and usually contains most of the showers initiated by electrons and photons with up to 50 GeV of energy.
- Finally the **back** module provides 2 additional X_0 of $\Delta\eta \times \Delta\phi = 0.050 \times 0.025$ towers, specifically for electrons and photons with E> 50 GeV producing large showers, and has the role of presampler for the hadronic calorimeter.

The full barrel calorimeter covers a pseudorapidity region of $|\eta| < 1.4$.
The EM calorimeter has a linear response within 1% for clusters down to 10 GeV. The

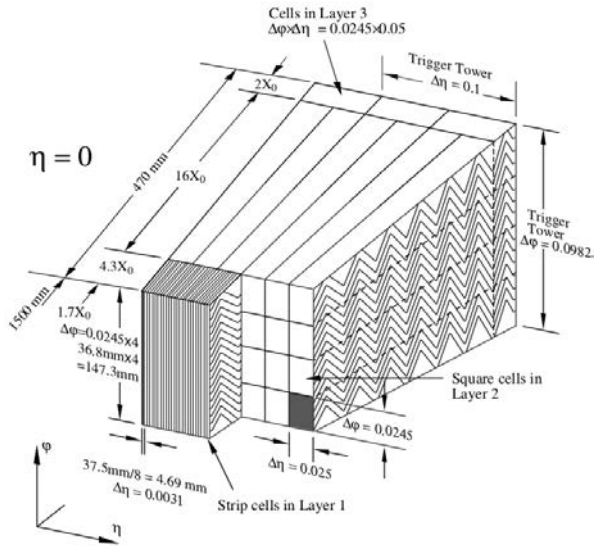


Figure 3.19: Schematic view of one of the EM barrel modules

lead thickness in the barrel is optimized as a function of η providing the best possible resolution in energy measurement:

$$\frac{\sigma_E}{\mathrm{E}} = \frac{10\%}{\sqrt{\mathrm{E[GeV]}}} \oplus 0.3\%$$

with a pseudorapidity resolution of:

$$\Delta\eta = \frac{40 \text{ mrad}}{\sqrt{\text{E[GeV]}}}$$

The electromagnetic calorimeter has a depth of about 2 interaction lengths (λ_0), therefore a significant fraction of jet energy can be deposited in the ECAL: this calorimeter can contribute significantly to jet performance.

Hadronic Calorimeter The barrel region of the hadronic calorimeter [119, 120] is divided into a central part ($|\eta| < 1$) and two extended barrels ($1 < |\eta| < 1.7$). Scintillators are used to recover part of the energy loss in the gap between the two barrel regions, corresponding to the readout of the EM calorimeter. The hadronic endcap (HEC) covers the pseudorapidity region of $1.5 < |\eta| < 3.2$ and it is contained in the same cryostat of the electromagnetic endcap calorimeter (EMEC) and the forward calorimeter (FCAL). The hadronic endcaps are composed of two independent wheels each, with approximately 2 m of outer radius, segmented longitudinally in two sectors: HEC modules are composed of 25 copper plates in the front wheels, 25 mm thick, and 15 copper plates 50 mm thick in the rear wheels.
The tile modules of the tile calorimeter are built with a structure that alternates 14 mm iron absorber plates and 3 mm scintillator polystyrene tiles with a ratio of 4.7:1 (a schematic view of this structure in Figure 3.20).

The TileCal is built with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, while the HEC has a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ or 0.2×0.2 (notably worse than the EM calorimeter). The overall energy resolution of the TileCal and HEC system is:

$$\frac{\sigma_E}{\text{E}} = \frac{50\%}{\sqrt{\text{E[GeV]}}} \oplus 3\%$$

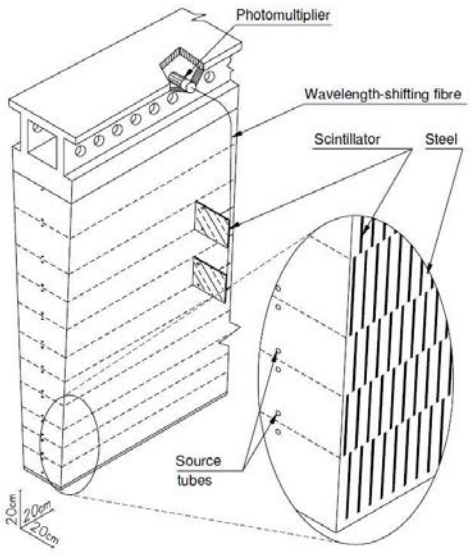


Figure 3.20: Tile from the ATLAS tile hadronic calorimeter.

Some loss of performance in the scintillation material is expected in the tile calorimeter due to radiation damage, but the energy resolution should degrade by only $\sim 10\%$ after about ten years of nominal operations (while no significant damage is expected for the Liquid Argon calorimeters).

FCAL The forward calorimeter [121] is designed using an intrinsic radiation-hard technology because of the high amount of radiation it is exposed to. It is placed in the same cryostat of the EMEC and HEC, with its front face at 4.7 m from the interaction point. The FCAL consist of three parts: the electromagnetic part is made with copper, while the hadronic parts are made with tungsten as absorber material. It has a metallic matrix structure with regularly spaced longitudinal channels filled with concentric rods (at positive high voltage) and tubes (grounded). The sensitive medium is LAr which fills the gap between.
The FCAL has a compact design to fit in the available longitudinal space, and to avoid energy leakage to its neighbors. The energy resolution is:

$$\frac{\sigma_E}{\mathrm{E}} = \frac{100\%}{\sqrt{\mathrm{E}}} \oplus 10\%$$

Calorimeter system: Upgrades for LHC Run-2 Between the LHC Run-1 and Run-2 both ATLAS calorimeters underwent large consolidations: the common strategy for both systems has been to increase reliability of the detectors, with new on-detector power supplies, consolidated electronics and improved data and power transmission.

3.2.5 The Muon Spectrometer

Muons require a dedicated detector since they are the only detectable particles that can consistently cross the ID and the calorimeter system without being stopped. The Muon Spectrometer [122] (MS) is the outermost part of ATLAS and has its own trigger system and tracking chambers. The MS covers out to $|\eta| < 2.7$ and can trigger on particles within $|\eta| < 2.4$
The magnetic field generated by the toroids is usually perpendicular to the muons tracks and allows us to measure the muon momentum with high precision (up to a few TeV) in the $R - z$ projection, in a direction parallel to the bending direction of the solenoidal magnetic field in the inner detector.
The muon chambers of the barrel region are designed in three concentric shells at radius of 5, 7.5, 10 m from the beamline. In the endcaps the muon chambers are arranged in large wheels at 7.4, 10.8, 14, 21.5 m from the interaction point along the z axis. The muon system structure is shown in Figure 3.21.
The muon spectrometer is composed of two types of subdetectors: chambers for the precision measurement of the particle momentum, chambers with coarser resolution and fast response for online triggering. The first type of subdetectors are:

- MDTs or Monitored Drift Tube chambers, which take care of the precise measurement of muon momenta providing high accuracy and good mechanical stability. MDTs cover the region $|\eta| < 2.7$ except in the innermost endcap layer where the coverage ends at $|\eta| = 2.0$; each chamber of the MDTs is built with 3-8 layers of drift tubes with an average resolution of 80 μm per tube and 35 μm per chamber

- Cathode-strip Chambers (CSCs), used in the innermost layer of the forward region due to their ability to tolerate higher rates; CSCs are multiwire proportional chambers optimized to have the wires oriented in the radial direction. This design provide a good resolution of 60 μm in the bending plane and 4 mm in the transverse plane.

The second type of subdetectors integrates the MDTs measuring the ϕ coordinates and provide online event selection by identifying the bunch crossings containing hard-scattering events with 99% accuracy (therefore they need to have very fast response).

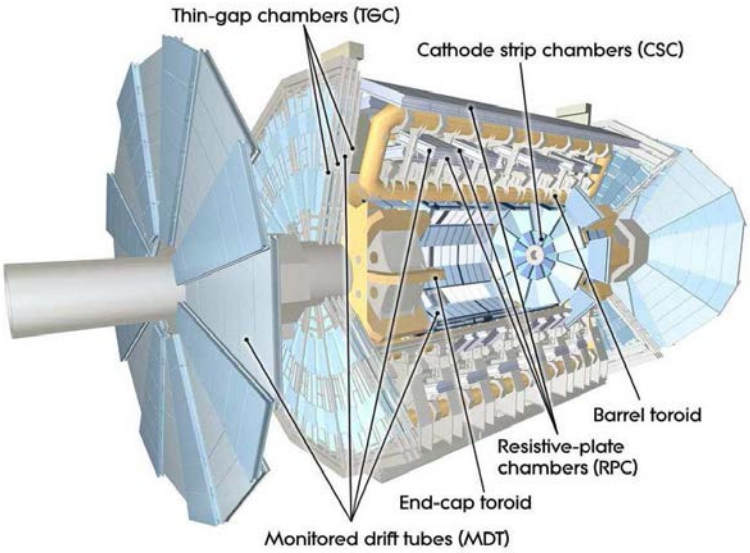


Figure 3.21: Cut view of the muon system structure.

- RPCs or Resistive Plate Chambers provide the fast trigger in the barrel region ($|\eta| < 1.05$): RPCs are gaseous parallel electrode-plate detectors. RPCs are arranged in planes in three cylindrical layers around the z axis, each plane composed of two layers which allow to measure the η and ϕ coordinates.
- In the end-cap regions ($1.05 < |\eta| < 2.4$) there are Thin Gap Chambers (TGCs) which take care of the trigger: TGCs are multiwire proportional chambers with high granularity and a very good rate of tolerance compared to RPCs. TGCs are segmented radially into the endcap and forward region, measuring the bending plane coordinate.

The accuracy of the measured momentum for a high-p_T muon track depends mainly on the resolution of the sagitta (the deviation in the $R - z$ plane from a straight line): for a track of $p_T \sim 1$ TeV the typical sagitta is approximately 500 μm. The MDTs provide a momentum measurement with $\sigma_{p_T}/p_T \sim 10\%$ resolution for 1 TeV muons, and 2-3% for lower momenta. In case of low-p_T muons the muon system information is combined with the inner detector to achieve a good efficiency and resolution.

Muon system: Upgrades for LHC Run-2 The muon spectrometer upgrade has been driven by the need to increase muon acceptance and the reduction of the Level-1 trigger rate in the endcap (dominated during Run-1 by slow particles not coming from the interaction region).
Muon acceptance have been improved with new or repaired chambers: ~ 30 TGCs and a few CSC repaired or replaced, new RPCs in the barrel holes due to the toroid feed and the access elevators (+2.8% acceptance) and extra endcap chambers. Improved muon purity at trigger level is achieved by adding coincidences with inner TGC layers and external TileCal layers.

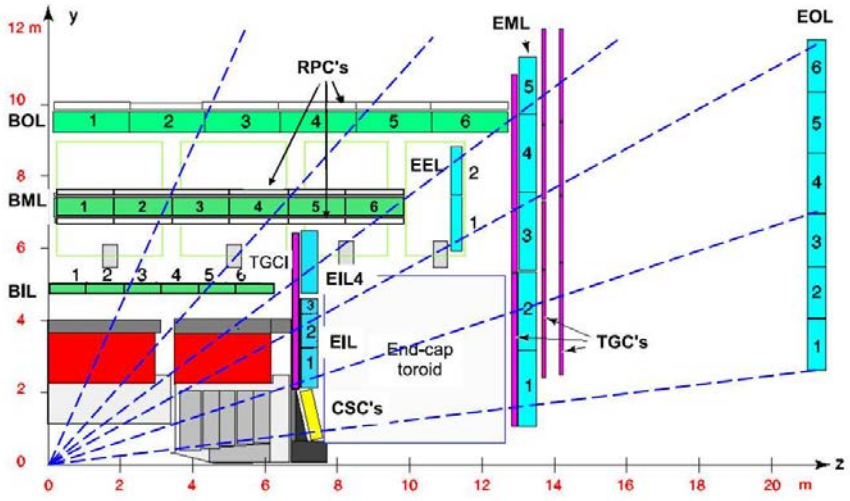


Figure 3.22: Schematic cutview of the ATLAS muon system, transverse to the beam line (before the introduction of upgrades for the LHC Run-2).

3.2.6 The ATLAS Forward Detectors

The ATLAS forward detectors [123] are a set of smaller detectors designed to provide coverage in the very forward region, in order to study inelastic *pp* scattering at small angles. From the closest to the most distant from the IP: LUCID [123] (Luminosity measurement Using Čerenkov Integrating Detectors), ZDC [124] (Zero-Degree Calorimeter) and ALFA (Absolute Luminosity For ATLAS). Their position along the beam line is outlined in Figure 3.23.

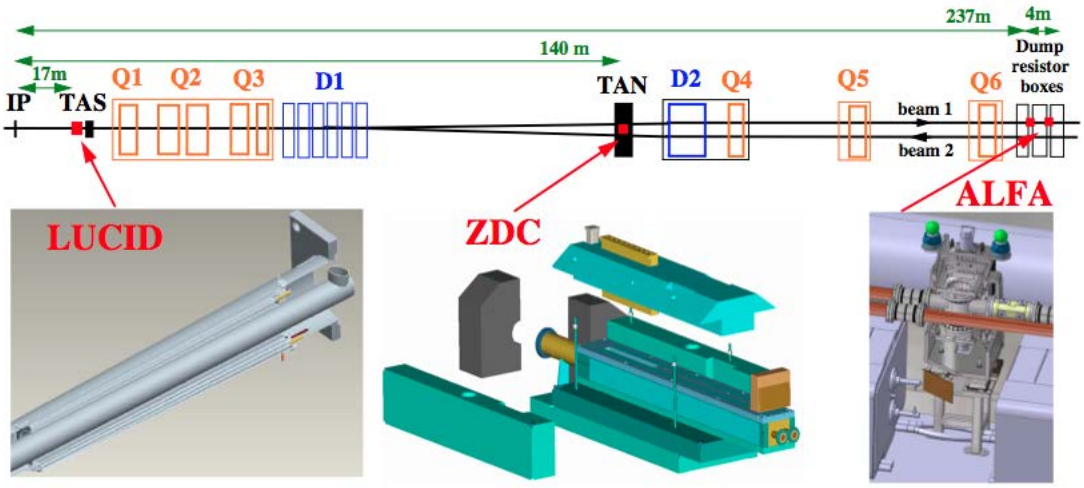


Figure 3.23: Placement of the forward detectors along the beam line with respect to the ATLAS interaction point.

LUCID It is situated at ±17 m from the IP, close to the Target Absorber Secondaries (TAS) collimator, with the main purpose of online luminosity monitoring (integrated and instantaneous luminosity and monitoring of beam conditions) via the measurement of *pp* inelastic collisions. Its design goal is an absolute luminosity measurement with uncertainty lower or of the order of 5%.
LUCID's main features and requirements are: resistance to a large radiation dose, good acceptance for MB events, possibility of counting charged particles and pointing capabilities to suppress tracks not coming from the IP.
It is designed as an array of Čerenkov tubes to count charged particles: 20 tubes of aluminum, 1.5 m long with a diameter of 15 mm, installed around the beam pipe and filled with C_4F_{10} at constant pressure of 1.2-1.4 bar. The Čherenkov threshold is 2.8 GeV for pions and 10 MeV for electrons.
Its two components are installed at $\sim$ 10 cm radial distance from the beam pipe, at $|\eta| \sim 5.8$. The read out system is composed of photomultipliers (PMT) with quartz windows (which suffer from the high radiation environment of LUCID).

ZDC It is placed at ±140 m from the IP, where the beam pipe is divided into two separate pipes, and it is embedded in the Target Absorber Neutral (TAN), between the two pipes after the split. ZDC is designed to detect forward neutrons from heavy-ion collisions ($|\eta| > 8.3$ to measure the centrality of heavy-ion collisions versus the number of forward spectator neutrons. By requiring a tight coincidence between the two ZDC detectors it has the role to reduce beam-halo and beam-gas effects.

Each arm is composed of four modules: one EM module of $\sim 29X_0$ (built from eleven tungsten plates normal to the beam axis, each vertically extended by 290 mm steel plates - quartz rods of 1 mm diameter enter the plates parallel to the z direction to form a 8×12 matrix) and three hadronic modules of $\sim 1.4\lambda$.
The time resolution is $\sim$ 100 ps, which allows to determine the IP with a 3 cm precision in z.

ALFA It is the furthest detector from the IP, encased into 'Roman pots' at $\pm$240 m, built with scintillating fiber trackers. Its goal is to detect elastic scattering at small angles (of the order of 3 μrad), related to the total cross section by the optical theorem. Since such scattering angles are typically smaller than the beam divergence, special beam conditions are required (high-β^* optics with small beam emittance), which result in low collected instantaneous luminosity (as a consequence non-radiation-hard technologies can be used). Using Roman pots the detector can be placed as close as possible to the beam pipe ($\sim$ 1 mm), to achieve the design spatial resolution of $\sim 25 \pm 3$ μm.

3.2.7 The ATLAS Trigger and Data Acquisition System

The collision rate provided by the LHC corresponds to 20 MHz during Run-1 and 40 MHz during Run-2: such rates are far exceeding the read-out and storage capabilities allowed by the current ATLAS technology. A dedicated trigger system is used to rapidly decide whether each event is stored for offline analyses, reducing the data-flow rate to a more manageable value of about 1 kHz, and providing a first discrimination between hard-scattering events (interesting for SM measurements and BSM physics searches) and soft-physics events.
The trigger system consists of two main levels: the first stage or Level-1 [125] (L1) finds regions of interest (RoIs) in the calorimeters and MS by using fast, dedicated hardware, reducing the data-flow rate to approximately 70 kHz (100 kHz for Run-2 collisions) with about 2 μs total latency.
Events retained by the L1 trigger are passed to the high-level trigger [126] (HLT), which during the Run-1 data-taking is further divided into two parts, the Level-2 (L2) and the Event Filter (EF) triggers, combined together during the Run-2. The HLT runs a more complete event reconstruction on a read-out of the whole detector, and results in a trigger acceptance rate up to 700 Hz (1 kHz during Run-2). For the Run-1 data-taking, the L2 trigger considers the RoIs provided by the L1 including full granularity information from the Inner Detector, reducing the data-flow rate to about 3.5 kHz, while the EF runs the event reconstruction. For the Run-2 data-taking, the HLT tracking algorithms have also been prepared for the inclusion of a new, fast hardware based tracking system (FTK [127]) which will become fully operational during 2017.
It is worth noting that the objects reconstructed at trigger level ('online') can be remarkably different from the corresponding objects passing the offline reconstruction and identification detailed in Section 5, and they do not benefit from dedicated calibration. Events used in the analyses documented in this thesis are collected by single-lepton, di-lepton and E_T^{miss} triggers, whose details are described in the corresponding analysis Chapters.
Finally, a trigger selection (consisting in the full chain from L1 to HLT) can be 'prescaled' or 'un-prescaled': prescaled triggers only retain a fraction of the events that would pass the trigger selection, to further reduce the data-flow rate and avoid over-burdening the data taking system. All triggers used in these analyses are un-prescaled.

4 | Statistical Analysis: The Methodology

This Chapter contains a detailed description of the statistical procedures implemented in the different analyses described in this thesis [128] [129] in order to extract the final results. The main statistical model used to perform the analyses is the binned Likelihood, built by taking into account the scaling factor for the expected signal rate (*signal strength* μ) as the parameter of interest under study, and all the systematic uncertainties introduced in the analysis as nuisance parameters (NPs). In the searches for the SM Higgs boson presented in Chapters 6 and 8 the signal strength is defined as the ratio of the measured Higgs rate to its SM prediction:

$$\mu = \frac{\sigma \cdot BR}{\sigma_{SM} \cdot BR_{SM}}$$

4.1 Physics search as a statistical test

In this scenario the statistical significance of an experiment can be quoted by measuring a p-value or its equivalent Gaussian significance. Under the hypothesis of signal absence, it is also particularly useful to characterize the sensitivity of an experiment by means of the expected statistical significance.

The general procedure for the search of a new physics process, in the framework of frequentist statistical tests, is often configured as the test of a first hypothesis (in this case predicted by a physics theory, but not yet observed by experiments) against an alternative hypothesis. For the purposes of setting limits on the presence of new phenomena the first hypothesis (H0) corresponds to the presence of signal, which is tested against the background-only alternative hypothesis (H1). When testing the discovery of a new signal process, the role of the two hypotheses is inverted.
The result of the physics search can be presented in terms of *p-value*, a quantity which estimates the level of agreement between a specific hypothesis H and the observed data, formally defined as the probability, under the assumption of the hypothesis H, of finding data of equal or larger incompatibility with the predictions of H. The tested hypothesis can be considered excluded if the correspondent p-value is below a specific threshold.
The p-value is usually expressed in terms of the 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, as shown from the example in Figure 4.1. In a more formal way the significance Z is the quantile Φ^{-1} of the standard Gaussian, computed for (1-p):

$$Z = \Phi^{-1}(1 - p)$$

The choice of the threshold for the p-value which corresponds to the rejection of the hypothesis H is not fixed and is in principle arbitrary. In the context of high energy physics

the rejection of the background hypothesis with a significance of at least $Z = 5$ (hence p $= 2.87 \times 10^{-7}$) is regarded as an appropriate level to quote a discovery. On the opposite, to exclude a signal hypothesis and set an upper limit on the presence of a new process, the threshold for the p-value is set at 0.05 (corresponding to a confidence level of 95% and a significance $Z = 1.64$).
It is important to point out that the rejection of the background-only hypothesis from statistical tests is only one of the steps required to claim the discovery of a new phenomenon. The degree of belief for a new process depends also, for example, on the plausibility of the new signal hypothesis and the level of agreement between its prediction and the observed data.
While the significance is the primary instrument to reject or accept the tested hypotheses, when performing an experiment to test the presence of new physics it is meaningful to quantify the sensitivity of the analysis, measuring the expected significance that would be observed in the measurement under the assumption of different hypotheses.

In this section the statistical procedures required to discard the tested hypothesis are outlined in a very formal and general way. In order to implement these tests on the analyses described in this thesis a specific method is applied: the frequentist significance test of the Profile Likelihood Ratio (PLR). The core idea of this method is the use of a Likelihood function as basic element to build a more complex test statistic, which is finally used to compute the p-value. The strength of this concept relies on the great flexibility of the parametric model that defines the Likelihood function, that makes possible to account not only for the parameters of interest of the analysis (rate and cross section of the signal process, the signal strength), but allows to introduce additional terms to fully describe signal and background models. These additional parameters are called *nuisance parameters* (NP) and are fitted from the data when deriving the p-value from the PLR: the range and the flexibility of the fit for each of these additional terms can be used to model systematic uncertainties, or to provide additional constraints on the background estimate. The additional degrees of freedom, introduced with the nuisance parameters to model systematic effects and uncertainties result as it is expected in a loss of sensitivity.

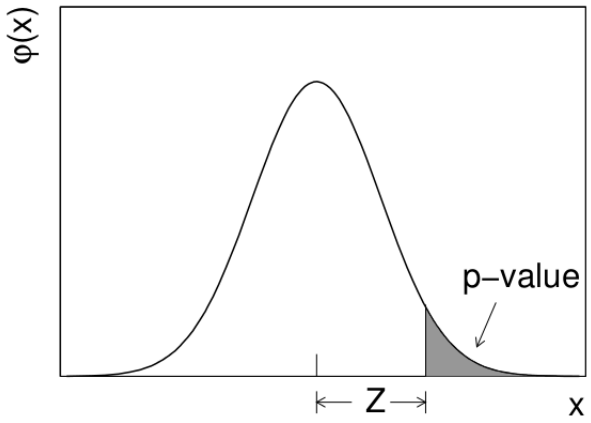


Figure 4.1: Significance Z and p-value for a normal distribution.

Likelihood function The Likelihood function built for these searches is a complicated object, which has to account for several systematic effects and includes a wide range of nuisance parameters: a detailed definition of the Likelihood for the different searches is given in the corresponding chapters. The general formula for a binned Likelihood function is thereby given as:

$$L(\mu, \vec{\beta}_{samp}, \vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global}) = Pois(n|\mu_T) \cdot Pois(n_{samp}|\beta_{samp}) \cdot (\vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global}),$$

where:

- n = number of events measured in the signal region.
- $\vec{\beta}_{samp}$ = statistical uncertainties quoted on the MC or data-driven control sample for the backgrounds, using the initial event numbers (n_{samp}) before scaling to the cross section.
- $\vec{\theta}_{s,b}$ = nuisance parameters associated to signal and background estimate, they account for systematic uncertainties on the analysis model.
- $\vec{\theta}_{global}$ = nuisance parameters shared amongst different decay channels, as the luminosity uncertainty (it is separated from the others to provide a good model for the combination with other channels)
- μ_T = total number of expected events, defined as:

  $$\mu_T = \sum_l \mu \mathcal{L} \sigma_l(m_X) f_s(\vec{\theta}_s) f_g(\vec{\theta}_{global}) + \sum_j \mathcal{L} \beta_j f_b(\vec{\theta}_b) f_g(\vec{\theta}_{global}) \ ,$$

  where the sum runs over the different signal production modes and/or contributions, and the different parameters are:

  → μ = parameter of interest (POI) for the statistical analysis, it is the scaling factor for the expected signal rate, called *signal strength*.

  → $\mathcal{L}$ = nominal value of the integrated luminosity used for the analysis.

  → $\sigma_l(m_X)$ = nominal effective cross section for signal events, where the signal is the production of a new particle X with mass m_X

  → $f_s(\vec{\theta}_s)$, $f_b(\vec{\theta}_b)$, $f_g(\vec{\theta}_{global})$ = functions of the analysis nuisance parameters that account for the dependence of the number of signal, background and total events on these NPs.

  → β_j = nominal effective cross section for the background process j

In these analyses the shape information of the discriminating variable used in the Profile Likelihood Ratio fit is extracted from template histograms, therefore the Likelihood is built as a binned function of this variable. This means that the complete Likelihood function is obtained by the product of one binned Likelihood term (as the one shown above) for each bin of the discriminating distributions.

As explained before, the presence of the NPs characterizes the shapes of probability density functions for signal and background, encoding the information of systematic uncertainties and providing the required handles to adjust the expected predictions during the Likelihood fit, and thus resulting in a loss of sensitivity: this is represented by the broadening of the profiled Likelihood function caused by the set of θ terms that enters the maximization

procedure. The θ NPs represent additional degrees of freedom introduced in the Likelihood fit, therefore when introducing them it is crucial to make sure that the information available from the fitted data is sufficient to control these degrees of freedom, and that the correlation model among them is properly assessed and implemented.

The NPs θ included in the Likelihood function can be generally divided into three classes: (i) unconstrained parameters controlling the normalization of the backgrounds (which are formally identical to the signal strength μ); (ii) normalization NPs controlling the yield of signal and background across the different regions of the phase space, included in the Likelihood with a log-normal probability density function (pdf) which prevents the normalisation factors from becoming negative in the fit; (iii) shape NPs controlling the variations in the templates of the fitted distributions, parametrized with a Gaussian pdf.
Parameters from (ii) and (iii) are in general associated with systematic uncertainties, and the prior constraint of their pdfs is added as a penalty term to the Likelihood, related to the magnitude of the corresponding uncertainty.
When building the Likelihood the pdfs parametrizing these NPs are in general known for the nominal case and for two extreme values of the NPs corresponding to the $\pm 1\sigma$ systematic variations: the pdfs are thus interpolated between these points to obtain a smooth pdf, using a quadratic interpolation with linear extrapolation for the normalization NPs (ii), and a piecewise exponential extrapolation with 6th order polynomial interpolation for shape NPs (iii).

Profile Likelihood Ratio (PLR) and test statistic The PLR is built from the Likelihood function here defined, as the ratio between the Likelihood maximized with respect to the full set of NPs, with the parameter of interest set as an arbitrary value to be tested, and the Likelihood maximized with respect to both the NPs and the parameter of interest:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})},$$

where:

- $\hat{\hat{\theta}}$ is the conditional maximum-Likelihood (ML) estimator of θ, the value of θ that maximizes the Likelihood for a specified μ (hence it is a function of μ itself).
- $\hat{\theta}$ and $\hat{\mu}$ are the ML estimator of the unconditional Likelihood at the denominator, for the nuisance parameters θ and the parameter of interest μ respectively.

From the Profile Likelihood Ratio it is possible to build different test statistics, which are the fundamental objects to compute the p-value of a measurement. From the given definition of $\lambda(\mu)$ it is clear that $0 \leq \lambda \leq 1$, where $\lambda \lesssim 1$ corresponds to a good agreement between data and the tested value of μ in the conditional Likelihood. For this reason the test statistic in usually defined as $t_\mu \sim -2 \log \lambda(\mu)$, so that higher values of t_μ imply an increasing incompatibility between the data and μ.
It is important to note that when quoting the results for the 'fitted signal strength' $\hat{\mu}$ or the fitted values of the NPs, a maximum Likelihood fit is performed rather than the complete PLR fit, since only the $\hat{\mu}$ and $\hat{\theta}$ maximum Likelihood estimators are needed. To obtain the uncertainty on $\hat{\mu}$ however, the full PLR is used, considering variations of $2 \log \lambda(\mu)$ by one unit (without any constraint on $\hat{\mu}$).
From the PLR $\lambda(\mu)$ several test statistics can be obtained, here I quote only the most interesting choices for the analyses discussed in this document.

Test statistics for the discovery of a positive signal q_0 This test statistics is introduced to test the background-only hypothesis $\mu = 0$ against alternative hypotheses where μ is assumed to be positive or zero. Rejecting the background-only hypothesis leads thus to the the discovery of a new positive signal.

$$q_0 = \begin{cases} -2\log\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} = \begin{cases} -2\log\dfrac{L(0,\hat{\hat{\theta}}(0))}{L(\hat{\mu},\hat{\theta})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}.$$

The constraint on $\hat{\mu} \geq 0$ means that data are considered in disagreement with the background-only hypothesis only if an upward signal strength fluctuation is observed: a negative $\hat{\mu}$ may indeed highlights some evidence against the background-only model but does not show that the measured data contain signal events. If the measured event yield is larger than the expectation ($\hat{\mu} > 0$) the test statistics becomes larger, indicating higher incompatibility between the data and the background-only hypothesis. This incompatibility can be expressed with a p-value, in this case named p_0, defined as:

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0|0)dq_0,$$

where $q_{0,obs}$ is the value of the test statistic measured from the observed data, and $f(q_0|0)$ the probability distribution function of the test statistic itself, under background-only hypothesis.

Test statistics for upper limits setting $\tilde{q}_\mu$ for a positive signal This test statistics is useful with the purpose of setting exclusion limits on the parameter of interest μ, when $\mu \geq 0$:

$$\tilde{q}_\mu = \begin{cases} -2\log\tilde{\lambda}(\mu) & \tilde{\mu} \leq \mu \\ 0 & \tilde{\mu} > \mu \end{cases} = \begin{cases} -2\log\dfrac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2\log\dfrac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}.$$

This quantity has some interesting properties: when $\hat{\mu} > \mu$ the test statistic is set to zero, because, while determining an upper limit, data with $\hat{\mu} > \mu$ are not regarded as less compatible with μ than the data obtained, hence they are not part of the rejection region of the test. When $\hat{\mu} < \mu$ two different scenarios are considered: if $\hat{\mu}$ is negative, it is replaced arbitrarily by zero to account for the fact that $\mu \geq 0$, hence a test of negative value wouldn't be significant. When $0 \leq \hat{\mu} \leq \mu$ the test statistic is written in the standard form.

As introduced before, higher value of the test statistic correspond to increasing disagreement between data and tested hypothesis, therefore the level of disagreement can be quantified by means of the test statistic by defining the p-value for the tested μ as:

$$p_\mu = \int_{\tilde{q}_{\mu,obs}}^{\infty} f(\tilde{q}_\mu|\mu)dq_\mu,$$

where $q_{\mu,obs}$ is the value of the test statistic measured from the observed data, and $f(\tilde{q_\mu}|\mu)$ the probability distribution function of the test statistic itself, under the hypothesis of

signal strength μ.

The $\tilde{q}_\mu$ test statistics is used in these analyses to set exclusion limits derived with the CL_S method [128], while the q_0 test statistics is used to assess the probability p_0 of the background-only hypothesis.

Asymptotic formulae For both the test statistics introduced in this Section, several useful approximations are available to make their calculation more straightforward and less taxing from the computation point of view. In detail, exploiting the Wald approximation [129] [130], the test statistics q_0 and $\tilde{q}_\mu$ can be written respectively as:

$$q_0 = \begin{cases} \frac{\mu^2}{\sigma^2} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad ; \qquad \tilde{q}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0 \\ \frac{(\mu-\hat{\mu})^2}{\sigma^2} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}$$

where $\hat{\mu}$ follows a Gaussian distribution with mean value corresponding to the signal strength according to whom the observed data are distributed (μ') and standard deviation σ (obtained from the covariance matrix of the estimators for all the parameters, POI and nuisance parameters) and μ is the parameter of interest under test. With few straightforward steps it can be shown that the corresponding significance Z_0 or Z_μ respectively is:

$$Z_0 = \Phi^{-1}(1-p_0) = \sqrt{q_0} \quad ; \qquad Z_\mu = \begin{cases} \sqrt{\tilde{q}_\mu} & 0 < \tilde{q}_\mu \leq \frac{\mu^2}{\sigma^2} \\ \frac{\tilde{q}_\mu + \frac{\mu^2}{\sigma^2}}{\frac{2\mu}{\sigma}} & \tilde{q}_\mu > \frac{\mu^2}{\sigma^2} \end{cases}$$

In order to set an upper limit, the p_μ-value is computed for several tested signal strengths μ, until the value corresponding to the fixed threshold α is found (for example $\alpha = 0.05$ to exclude the signal with CL=95%). The upper limit is the largest μ with $p_\mu \leq \alpha$, while the μ parameters with p_μ above the threshold are excluded with a confidence level CL = 1 - α.
Recalling that:

$$p_\mu = 1 - \Phi(\tilde{q}_\mu|\mu),$$

where Φ is the cumulative function of the Gaussian distribution, it is clear that the upper limit on the signal strength μ for a CL = 1 - α is obtained by setting $p_\mu = \alpha$, as:

$$\mu_{up} = \hat{\mu} + \sigma\Phi^{-1}(1-\alpha)$$

Experimental sensitivity and the Asimov [131] dataset The sensitivity of an experiment is a direct and quantitative estimate of the rejection power of the experiment itself against different hypotheses (hence different values of the parameter of interest). Although this sensitivity does not depend on the observed data, but can be estimated from simulations, it is still a very important estimator of the experiment performance and it is often used as discriminating variable to tune and optimize the analysis.

The sensitivity is characterized as the expected (median) significance with which it is possible to reject different values of μ. For the case of exclusion limit it is computed assuming data generated under the background-only ($\mu = 0$) hypothesis, and testing the median significance with which nonzero value of μ can be rejected (specifically the nominal $\mu = 1$). The choice of the median value as expected significance is not accidental: in general the expected value of Z would not satisfy the relation with the p-value ($Z = \Phi^{-1}(1-p)$), but since it is monotonic the medians of Z and p respect the same relation. Note that since the p-value is a monotonic function of the test statistic itself, establishing the p-value correspondent to the median value of the test statistic under the assumption of μ' (and thus the sensitivity, or median Z) is equivalent to compute the median p-value, assuming μ'.
In order to easily obtain the median value of the significance Z the so called *Asimov dataset* is introduced. It is an artificial dataset, defined so that when it is used to evaluate the ML estimators for the Likelihood parameters, the results are the input values of the parameters themselves. These true values are, in practice, the ones that would be estimated from Monte Carlo simulations in the limit of very large statistics. The Asimov dataset condition is easily found by computing the ML estimators from the Likelihood function, and imposing that they are equal to their expectation values.
This special dataset can be used to obtain the median value for the test statistic in a very straightforward way, leading to a simple expression for the corresponding median significance. As show before Z is a monotonic function of the test statistic, thus the median significance is obtained evaluating Z as a function of the median test statistic, which is approximated by the Asimov value.
When treating the discovery of a positive signal with q_0 the median discovery significance assumes a signal strength parameter μ' tested, while in the case of exclusion limits setting with $\tilde{q}_\mu$, the median exclusion significance is evaluated assuming $\mu' = 0$, obtaining:

$$med[Z_0|\mu'] = \sqrt{q_{0,A}},$$
$$med[Z_\mu|0] = \sqrt{\tilde{q}_{\mu,A}},$$

where the test statistics are obtained from the Asimov dataset.

Expected statistical variations It is important to emphasize that, even if the value μ' used to compute the median significance of rejecting the μ under test is correct, real data will be affected by statistical fluctuations, hence the observed significance may not coincide to the median. Knowing the expected variation of the significance, given the statistical fluctuation that are expected in data, is fundamental to obtain a reliable result.
In particular, since the fitted $\hat{\mu}$ follows a Gaussian distribution, it is easy to compute error bands for the sensitivity corresponding to the $\pm N\sigma$ variation of $\hat{\mu}$, because they are conveniently the quantiles of the Gaussian function.
For the discovery test, the significance values corresponding to $\mu' \pm N\sigma$ are:

$$Z_0(\mu' + N\sigma) = med[Z|\mu'] + N$$
$$Z_0(\mu' - N\sigma) = max[med[Z|\mu'] - N, 0]$$

While for the exclusion limits setting, recalling the expression found for μ_{up}:

$$\mu_{up} = \hat{\mu} + \sigma\Phi^{-1}(1-\alpha),$$

the median upper limit under the μ' hypothesis can by computed by replacing $\hat{\mu} = \mu'$:

$$med[\mu_{up}|\mu'] = \mu' + \sigma\Phi^{-1}(1-\alpha),$$

and the $\pm N$ error bands on the sensitivity are given by:

$$band_{N\sigma} = \mu' + \sigma(\Phi^{-1}(1-\alpha) \pm N).$$

4.2 Treatment of the nuisance parameters in the Likelihood Fit

In Section 4.1 I introduced the more formal principles guiding the definition of the statistical tests used to perform the analyses. In this Section I focus instead on the treatment of the NPs in the Likelihood fit, discussing how to deal with statistical fluctuations and how to limit the fit complexity by removing negligible NPs. I then discuss some of the most critical tests to assess the stability of the Likelihood fit, looking at the behavior of the NPs during the fit and their impact on the result, and I introduce some fundamental ideas that guide the determination of the correlation model among NPs.

4.2.1 Smoothing and pruning of systematic uncertainties

The effect on the fit of possible statistical fluctuations affecting the estimate of systematic uncertainties on the shape of the distributions has to be tested. Shape uncertainties are implemented in the Likelihood fit as alternative templates for the discriminating variable relative to the nominal prediction, therefore they can suffer from statistical fluctuation in the simulation. If the systematic uncertainty is obtained by reweighting the nominal template distribution (e.g. uncertainties on the b-tagging efficiencies), no statistical fluctuation is expected beyond those included in the nominal template. When the systematic variations can change the event selection (e.g. jet energy scale variations), additional statistical fluctuations may be introduced and affect the systematic shape. If multiple systematic uncertainties are derived from the same template distribution (e.g. multiple components of the jet energy scale uncertainty), the statistical error on the template should not be included in all of them. To mitigate this effect a smoothing procedure is applied to all systematic variations across all regions, merging bins of the fitted distribution based on the constraints that the statistical uncertainty in each bin must be below 5% and the shape of the systematic templates remains physical (monotonous for the MVA discriminant, with at most one local extreme for invariant mass spectra).
Given the very complex analyses and the extended model of nuisance parameters, the fit may contain a very large number of systematic-variation template pairs ($+1\sigma$ and -1σ): to reduce this number in order to obtain a more efficient and solid fit model, a pruning procedure is applied to remove systematic uncertainties that have a negligible impact on the final result. Normalization (shape) uncertainties are dropped if the variation of the corresponding template is below 0.5% (below 0.5% in all bins). Additional pruning criteria are applied in all regions where the signal contribution is less than 2% of the total background and the systematic variations impact the total background prediction by less than 0.5%. Furthermore, shape uncertainties are removed if the up- and down-varied shapes are more similar to each other than to the nominal shape (only applied for shape systematic where opposite-sign variations are expected).

4.2.2 Stability of the Likelihood fit: post-fit NPs and rankings

The stability and the behavior of the combined fit is tested with several checks and studies: the most notable tests include the comparison of the central value and uncertainty of the nuisance parameters before (pre-) and after the fit (post-fit), which can be respectively pulled or constrained by the Likelihood fit, as well as the study of correlations developed between uncertainties during the fit. To assess these effects the expected modeling of the NPs is compared to the pulls, the constraints and the correlations observed in the fit to data: when striking differences are found their source is investigated and this may lead to further uncorrelate systematic uncertainties by introducing additional NPs, providing the the fit with the degrees of freedom required to adjust the MC expectation to the observed data.

A second critical test allows to order the systematic uncertainties according to their impact on the final results (the fitted value of the signal strength $\hat{\mu}$): this information is extremely useful to individuate the most important NPs in the Likelihood fit, and to isolate the aspect of the analysis which could most benefit from improvements or additional studies.
The NP ranking is obtained by performing one iteration of the fit for each systematic uncertainty, fixing the corresponding NP to the fitted value $\hat{\theta}$ from the baseline fit, shifted up or down by its fitted uncertainty, and allowing all remaining parameters to vary, in order to properly account for correlations between the systematic uncertainties. The magnitude of the shift in the fitted signal strength $\hat{\mu}$ compared to the baseline Likelihood fit gives a measure of the impact of the considered NP. Similarly the expected impact of each uncertainty is obtained from a fit in which the corresponding NP is set to its pre-fit value and shifted by the associated uncertainty.

4.2.3 Nuisance parameters correlation

For the nuisance parameters related to the background and signal modeling the specific correlation scheme is introduced case by case in the analysis Chapters. The choice to uncorrelate NPs related to different systematic effects (for instance the p_T^V shape uncertainty and the jet flavor fraction for a given background) appears clear when considering their role in the fit: to correlate two NPs is equivalent to the assumption that information on one of them can affect the other. This is a strong statement as one of the systematic effects may be strongly constrained by the Likelihood fit (for instance the p_T^V spectrum), and propagate this strong constraint on the second NP (in this example controlling the jet flavor fractions of the background). Furthermore, special care is needed when the correlation model between uncertainties is obtained from MC-to-MC comparisons: if the correlation is not reproduced in data, the NPs will be pulled during the fit to recover data-to-MC discrepancy, causing potential bias in the final result.
Keeping the NPs uncorrelated may represent a more conservative approach, since it increases the number of degrees of freedom in the fit: this is adopted for large backgrounds whose modeling can be controlled with good accuracy from the fit to data. For smaller backgrounds, such as diboson or single-top production, less information is available from data and it is less likely to cause biases in the result due to mismodeled correlations, therefore NPs can be correlated more extensively across regions and/or systematic effects. It is however important to highlight that the correlation model implemented for the NPs has to be studied case by case, considering our best knowledge of the modeling for the process in consideration.

5 | Event and Object Reconstruction

5.1 Physics objects reconstruction

In this Section I describe the details of the identification and reconstruction of the main objects used in the analyses documented in this thesis. Since these analyses have been performed with LHC data collected at different center of mass energies, during different Runs, the features of the object reconstruction may be different due to changes in the identification and reconstruction algorithms, upgrades of the ATLAS sub-detectors (for instance, the insertion of the IBL in the Pixel detector) and different detector conditions. In this cases I give full details of the algorithms an techniques used for data collected at $\sqrt{s} = 8$ and 13 TeV, while I shortly summarize possible differences in the analysis of $\sqrt{s} = 7$ TeV data in Section 5.3. These objects include tracks from charged particles and interaction vertexes (Section 5.1.1), charged leptons (Section 5.1.2), hadronic jets - possibly originated by heavy-flavor quarks - (Section 5.1.3, 5.1.4) and the missing transverse energy, a quantity related to the reconstruction of the full event (Section 5.1.5).

5.1.1 Tracks and vertexes

Charged particle tracks are reconstructed by fitting track hypotheses to the set of clustered hits left by charged particles passing through the inner detector. The tracking algorithm uses clusters collected from the pixel layers of the ID as seeds, and is able to resolve ambiguities in the hit assignments arising from high multiplicity environment. The p_T-threshold for charged particle tracks is at 400 MeV.
In the analysis of Run-1(Run-2) data, interaction vertices are reconstructed from at least three(two) selected tracks in the inner detector, and the fit to the candidate vertex is performed including the beam spot as an additional constraint. The 'primary vertex' is selected among the candidate vertices as the one with the largest sum of squared transverse momenta of associated-tracks $\sum p_T^2$.

5.1.2 Leptons

Leptons are critical objects for all the analyses discussed in this document: not only the main search channels are defined from the leptonic decays of the vector boson V produced in association with the Higgs boson, but isolated leptons are also used to reject backgrounds containing misidentified leptons, or leptons originated by semileptonic decays of heavy quarks (as multijet processes).
Three categories of leptons are defined in all analyses presented in this thesis, corresponding to three different purity criteria: loose, medium, tight leptons. In this section these three definitions are presented, highlighting the main differences for the reconstruction strategies used for data collected during the LHC Run-1 and Run-2.

Electrons

Electrons are reconstructed as noise-suppressed clusters of energy in the calorimeter using a sliding window algorithm with a matching charged track from the inner detector [132, 133, 134]. Electron energy calibration is driven by data from reference processes such as $Z \to e^+e^-$ [135].

- **Loose electrons** are required to have $E_T > 7$ GeV and $|\eta| < 2.47$, they have to fulfill loose likelihood identification criteria, which encodes information on the electron shower-shape, track-quality criteria, quality of the matching between the track and its associated energy cluster in the calorimeter (direction and momentum/energy), TRT information, and an identification criterion for electrons originated by photon conversions. In the Run-1 analysis the lepton must pass an isolation criteria, requiring that the transverse momentum carried by tracks within $\Delta R = 0.2$ around the lepton, excluding its own track, is less than 10% of the lepton energy. In the Run-2 analyses a similar isolation criteria is applied with 99% efficiency for real electrons; furthermore the lepton track has to satisfy impact parameter criteria to reject tracks from pile-up events, based on the transverse impact parameter significance and the longitudinal impact parameter difference between the track and the primary vertex.

- **Medium electrons** are required to satisfy the loose criteria with a tighter selection $E_T > 25$ GeV.

- **Tight electrons** are required to satisfy the medium criteria with tighter likelihood identification. The Run-1 track-based isolation criteria is tightened to 4%; an additional isolation cut is applied by requiring the sum of the calorimeter energy deposits within $\Delta R = 0.3$ around the lepton, excluding clusters associated to the lepton itself, to be less than 4% of the lepton energy. In the Run-2 analyses similar calorimeter-based isolation and track-based isolation criteria are combined, with an efficiency of 95% for real electrons.

Muons

Muon leptons are reconstructed from tracks in the inner detector and the muon spectrometer [136, 137].

- **Loose muons** are required to have $E_T > 7$ GeV. Different types of muons are included in the analyses: almost all muons are reconstructed in both the MS and the ID, with $|\eta| < 2.5$ coverage; to overcome the limited muon chambers coverage muons identified in the calorimeter with $p_T > 20$ GeV, associated to an ID track with $|\eta| < 0.1$ are retained; muons reconstructed in the forward region ($2.5 < |\eta| < 2.7$) of the MS are considered even if not matching full ID tracks (to cope with the limited inner-detector coverage). When muons are associated to ID tracks, additional requirements on their impact parameter are applied (in both Run-1 and Run-2 analyses), and in all cases isolation criteria are included to reject non-prompt muons, as done for electrons.

- **Medium muons** have to satisfy the loose criteria, with $E_T > 25$ GeV and $|\eta| < 2.5$ (in the analysis of Run-1 data, only muons reconstructed in both MS and ID are retained at this step).

- **Tight muons** have to satisfy the medium criteria, with tighter identification and isolation requirements (as done for electrons) to further reduce the multi-jet background.

Taus

In the analysis of Run-1 data, tau leptons are not treated specifically: taus decaying hadronically (τ_{had}) are treated as light-jets, while tau decaying leptonically are not explicitly selected (but electrons or muons from their decay are included, if reconstructed as isolated leptons according to Section 5.1.2).
On the contrary in the analyses of Run-2 data, hadronically decaying taus are used in the physics selection: they are reconstructed [138, 139] using the anti-k_T algorithm [140] with a radius parameter R of 0.4. τ_{had} leptons are required to have exactly 1 or 3 matching charged tracks within a cone around the jet axis of $R = 0.2$ (1-prong or 3-prong τ candidates). To reject fake τ-leptons from hadronic jets a multivariate discriminator is built with boosted decision trees, combining information from calorimeters and tracking detectors.
Dedicated energy corrections are applied to calibrate the hadronically decaying tau lepton candidates [141].
Hadronically decaying τ leptons are required to have $p_T > 20$ GeV and $|\eta| < 2.5$, excluding the transition region between the barrel and forward calorimeters corresponding to $1.37 < |\eta| < 1.52$. The τ_{had} reconstruction and identification efficiency is about 55% in the 1-prong case and about 40% in the 3-prong case.

5.1.3 Hadronic Jets

Elementary particle physics processes are often described in terms of quarks and gluons, fundamental objects of Quantum Chromodynamic. These objects, however, are not observed in a physics experiment, due to the structure and the features of strong interaction: quarks and gluons fragment and hadronise immediatly after production, therefore the only observable object in the detector is a spray of hadrons, which is called jet [142].
Jets are particularly important because by measuring their properties it is possible to obtain information about the partons that originated them. Even if usually a jet is initiated from a single parton, this is not always the case: first, partons are ill-defined objects especially when it comes to physics measurements, and also two or multiple partons may be highly collimated and result combined in a single jet, with a specific structure. Therefore in physics analysis useful informations come from both the number of jets and their substructure, originated by multiple partons within a single jet.
Jets need a precise definition in order to reconstruct and use them. A *jet definition* is a set of rules, organized as an algorithm, that allows to group particles into jets and to compute the momentum of the resulting objects. In the analyses presented in this thesis different jet definitions are used, depending on what physics information is required from the initial partonic state which originate the jet. In this section I describe all jet definitions used in the following chapters, highlighting their main features and the motivations of these choices.
All jets used in the analyses are built using the anti-k_T algorithm [140], therefore the main differences among the jet algorithms considered come from the parameters of this algorithm (e.g. the radius R), and from the fundamental objects used as inputs for the jet algorithms, which may be energy clusters in the calorimeter or inner detector tracks.

Calorimeter jets Calorimeter jets are reconstructed from noise-suppressed topological clusters of energy in the calorimeters [143]. In the search for the SM $VH(b\bar{b})$ (Chapters 6 and 8) only jets with radius parameter $R = 0.4$ are considered. In the search for a CP-odd A boson presented in Chapter 7 also jets with larger radius $R = 1.0$ are used. Both jet

collections are here described in detail.
Calorimeter jets are simply referred to as 'jets'; when both jet collections are used in parallel, the $R = 0.4$ jets are named 'small-R' jets, while the $R = 1.0$ jets are named 'large-R' jets.

- calorimeter jets with $R = 0.4$, or **small-R jets**: the energies of reconstructed jets is calibrated using p_T- and η-dependent correction factors (JES) derived from simulation, with residual corrections from in-situ measurements applied to data [144, 145, 146]. The jet energy is furthermore corrected to account for the contribution of pile-up interactions with a jet-area-based technique [147, 148]. Further adjustments are made based on jet internal properties, which improve the energy resolution without changing the average calibration (global sequential calibration [144, 146]).
  Jet cleaning criteria are applied to identify jet candidates likely to be originated from non-collision sources and noise: any event containing such jets is removed [149].
  Different criteria are in place to reduce the contamination by jets from pile-up interactions: in the analysis of Run-1 data, the scalar sum of the p_T of tracks matched to the jet, and originating from the primary vertex, must be at least 50% of the scalar sum of the p_T of all tracks matched to the jet (this requirement is applied only to jets with $p_T < 50$ GeV and $|\eta| < 2.4$, and it is named 'jet vertex fraction' or JVF selection), jets without matched tracks are retained. In the analysis of Run-2 data this selection is based on a likelihood discriminant which combines information about the primary vertex and charged tracks associated with the jet [150], applied for jets with $p_T < 50$ GeV and $|\eta| < 2.4$.

  In the SM $VH(b\bar{b})$ analysis presented in Chapter 6 jets with $p_T > 20$ GeV and $|\eta| < 4.5$ are considered; while in the searches presented in Chapter 7 and 8, jets with $p_T > 20$ GeV and $|\eta| < 2.5$ ('central jets') and jets with $p_T > 30$ GeV and $2.5 < |\eta| < 4.5$ ('forward jets') are included.

- calorimeter jets with $R = 1.0$, or **large-R jets**: this jet collection is only used in the search presented in Chapter 7. A dedicated procedure named trimming [151] is applied to the large-R jets to remove the energy of clusters originating from initial state radiation, pile-up interactions or underlying event. The trimming is applied by reclustering the initial jet constituents into sub-jets with radius $R_{sub} = 0.2$, using the k_T algorithm [152, 153], and removing subjets with p_T below 5% of the p_T of the parent large-R jet [154].
  The energy of large-R jets is corrected for losses in passive material of the detector and for the non-compensating response of the calorimeter.
  Only large-R jets with $p_T > 250$ GeV and $|\eta| < 2.0$ are considered in the analysis.

Track jets Track jets are built from ID tracks using the anti-k_T algorithm with radius parameter $R = 0.2$, and are required to have $p_T > 10$ GeV and $|\eta| < 2.5$, and to be formed by at least two tracks.
These jets are used in the search presented in Chapter 7, and are only considered if associated to a large-R jet, to extract information on its sub-structure and composition. Track jets are associated to large-R jets via *ghost-association* [155]: the large-R jet clustering algorithm is run including the track jets among the seeds, setting their p_T to an infinitesimal value, and if they are found to be included in the large-R jet constituents they are defined as *ghost-associated* to it.

5.1.4 B-Jets

In these analyses 'b-jets' are hadronic jets reconstructed as small-R or track jets which pass some identification criteria based on dedicated b-tagging algorithms developed by the ATLAS Collaboration. These algorithms are built to identify jets originated from the fragmentation of b-quarks by using information from the long b-quark lifetime.
The b-tagging algorithm has evolved between the analysis of LHC Run-1 and Run-2 data, to cope with the different conditions of pp collisions and to fully exploit the upgrades of the ATLAS detector (such as the insertion of the IBL). In this section I describe the b-tagging algorithms considered, for different LHC runs and types of jets, as well as some details on their calibration and the associated systematic uncertainties. In all cases the b-tagging algorithms provide as output a weight w for each tested jet, related to the likelihood of it being a b-jet: several operating points, corresponding to different b-tagging efficiencies and calibration scale factors, are provided for each algorithm.

b-tagging for the analysis of Run-1 data Jets originated from b-quarks are selected with the 'MV1c' b-tagging algorithm [156, 157, 158]: MV1c is a neural network multivariate regression function which combines together various b-tagging algorithms built to exploit the information of the track impact-parameter significance, and explicit reconstruction of b- and c-hadron decay vertices and $b \to c$ hadron decay chain. The MV1c algorithm is an improved version of the MV1 function [159, 160, 161] which achieves a better rejection for jets originated by c-quarks, at the cost of slightly worse light jet rejection. This feature is especially important to suppress the $t\bar{t}$ background in the high-p_T region of the analysis, where the flavor composition of this process has a strong contribution from events with jets originated by c-quarks, as stated in Section 2.3.3.
In the analysis described in Chapter 6 four different MV1c operating points are used, with dedicated calibrations, corresponding to average b-tagging efficiencies of 80%, 70%, 60% and 50% for b-jets with $p_T > 20$ GeV (measured in a dedicated sample of $t\bar{t}$ events), denoted respectively as loose, medium and tight. The improved c-jets rejection factor obtained with MV1c corresponds to 26 (5) for the tight (medium) operating point, while the light-jets rejection factor corresponds to 1400 (136).
The efficiencies for b-, c- and light-jets are corrected by measuring them in data and simulation in a specific $t\bar{t}$ control region (for b-jets), a $D^\star$-enriched region (for c-jets) or in multi-jet samples (for light-jets), extracting scale factors (SFs) from the data-to-MC comparison for each exclusive region between operating points, as a function of the jet kinematic (p_T and - for light-jets - $|\eta|$).
To keep into account the dependency of these SFs on the MC generator used for the simulation, an additional set of 'MC-to-MC' SFs is derived by comparing the efficiencies across different generators. The MV1c b-tagging algorithm is developed for jets with $p_T > 20$ GeV and $|\eta| < 2.5$, thus only this kinematic regions is considered when applying b-tagging criteria in the analysis.

b-tagging for the analysis of Run-2 data In the analysis of LHC Run-2 data, two types of jets are considered for b-tagging: small-R calorimeter jets and track jets. The b-tagging algorithm 'MV2c' is a multivariate function similar to MV1c, including information from the improved tracking capabilities achieved by ATLAS in particular with the insertion of the additional pixel layer IBL. The MV2c algorithm has been deployed in two separate versions, corresponding to a first optimization MV2c20 with the 2015 dataset [162, 163, 164], and a second instance MV2c10 for the combined 2015+2016 dataset [165]. The former is used in the search presented in Chapter 7 (it is worth noting that the same b-tagging

algorithm is used for both small-R and track jets in this analysis, at the same operating point), while the latter is applied in the analysis described in Chapter 8.

- MV2c20 (2015 dataset): this algorithm is used with a single 70% efficiency working point, providing a c-jets rejection factor of 8 (5.6), and a light-jets rejection factor of 400 (180) for small-R jets (track jets). The algorithm is applied to jets in the region of $|\eta| < 2.5$, and only to track jets which are ghost-associated to a selected large-R jet.

- MV2c10 (2015+2016 dataset): this algorithm is only applied to small-R jets, as track jets are not used for b-tagging criteria in the analysis presented in Chapter 8. While the b-tagging efficiency working point used is again 70%, the c-jets rejection factor is improved to 12, with a small reduction in the light-jets rejection factor which corresponds to 380. Given the strong impact of backgrounds characterized by the presence of c-jets in the final states considered (especially from top-quark pair production), this feature makes this improved algorithm a better candidate than the MV2c20 used for the preliminary analysis of the 2015 dataset. The algorithm is applied to jets with $|\eta| < 2.5$.

For both instances of the MV2c algorithm the tagging efficiencies are corrected by data measurements in specific control regions, similarly to what described for the b-tagging algorithm used for $\sqrt{s} = 8$ TeV data. No MC-to-MC SFs are needed for the MV2c algorithms, since the MC simulated samples produced for the analysis of LHC Run-2 data have been generated with consistent setup for the heavy-flavor simulation and the b-quark decays through the `EvtGen1.2.0` software [166].

5.1.5 Missing Transverse Energy

The missing transverse momentum (E_T^{miss}) is a crucial quantity to identify final states with undetectable neutrinos, whose presence can be inferred by an apparent momentum imbalance in the transverse plane, and to suppress the background contribution in signal topologies where all the particles in the final state can be identified in the detector (as the 2-lepton $ZH \to l^+l^-b\bar{b}$ process).
Since the E_T^{miss} quantify the transverse momentum imbalance, it is really an 'event quantity' whose definition relies on all other objects reconstructed for a given event.

E_T^{miss} for the analysis of Run-1 data The E_T^{miss} [167, 168] is defined as the negative vector sum of the transverse momenta associated with energy clusters in the calorimeters with $|\eta| < 4.9$. When clusters are associated with reconstructed objects (electrons, τ leptons, photons and jets) the appropriate calibration for these objects is used. Muons are treated separately by adding their transverse momenta to the E_T^{miss} calculation, properly removing the energy deposited in the calorimeters to avoid double-counting. The components not associated to any reconstructed objects are part of the E_T^{miss}(SoftTerm), or E_T^{miss}(ST), which encodes the contribution of all energy deposits that do not satisfy any criteria to be identified as physics objects (thus can be associated to mismeasured or misidentified leptons or jets due to limited detector coverage and resolution, or to noise sources inside and outside the detector as cosmic-ray and beam-halo muons).

E_T^{miss} for the analysis of Run-2 data The approach used in the analyses described in Chapters 7 and 8 differs from the reconstruction strategy adopted for Run-1 data in the treatment of the E_T^{miss}(ST): this component is based on the well-reconstructed tracks

originating from the primary vertex, which are not already included in any of the physics objects [169]. Building the E_T^{miss}(ST) from tracks rather than from energy deposits in the calorimeters, makes this quantity more robust with respect to the pile-up contribution.

In addition to the E_T^{miss}, a track-based missing transverse momentum vector p_T^{miss} is built from the negative vector sum of the transverse momenta of all well-reconstructed tracks associated with the primary vertex. This quantity is mainly used to suppress non-collision and multi-jet backgrounds.

5.2 Event reconstruction

In this Section I describe the techniques, methods and corrections dedicated to the optimization of the reconstruction of full physics events, using the objects introduced in the previous Section. This includes specific b-jet energy corrections for the improvement of the Higgs invariant mass resolution (Section 5.2.1), a kinematic fit applied in the $l^+l^-b\bar{b}$ channel - where the event kinematic can be fully reconstructed and constrained (Section 5.2.2), the overlap removal techniques applied to avoid double counting of objects (Section 5.2.3), and finally the truth-tagging method applied to improve the statistical power of MC simulation when applying b-tagging requirements (Section 5.2.4).

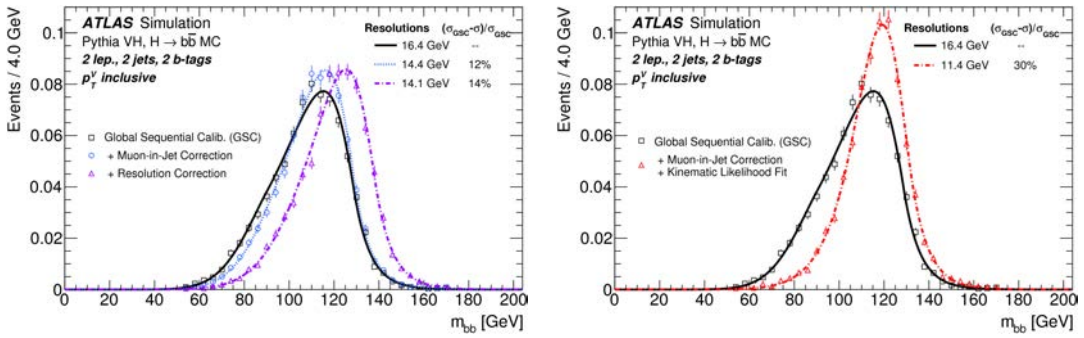
5.2.1 Dijet mass resolution: b-jets energy corrections

In all the analyses discussed in this document the invariant mass of the Higgs candidate is reconstructed from its decay to a b-quark pair, and it is either used as final discriminating variable or contributes to its definition. The resolution of this invariant mass is therefore a crucial aspect of the analysis performance, and plays a central role in the rejection of backgrounds which are either non-resonant (such as $t\bar{t}$ or V+jets production) or have a different invariant mass (as diboson VZ production).
Since the Higgs invariant mass is built with hadronic jets originating from b-quarks, objects whose reconstruction is very challenging, special care is needed to to maximize its resolution. In this section I present a series of corrections applied to the b-jets used to define the Higgs invariant mass: although these corrections are applied to the specific objects, their target is the improvement of the invariant mass resolution, thus they are considered 'event-related'.
The following corrections are derived for the analysis of Run-1 and Run-2 LHC data:

- 'muon-in-jet': this correction targets semileptonic decays of b-quarks which produce muons inside the jet cone. The 4-vector of muons passing selection criteria and reconstructed within the b-jet cone, is added to the jet 4-vector, after subtracting the muon's energy deposit in the calorimeter (if more than one muon is found, the closer to the jet is used). Systematic uncertainties on this kind of correction are found to be negligible.

- 'p_T-reco': this correction (applied after the muon-in-jet one) compensates for biases in the jet response due to resolution effects, out-of-cone energy and undetectable neutrinos. It is derived from the p_T spectrum of jets from the decay of a Higgs boson with $m_H = 125$ GeV in simulated VH events. The correction is obtained by comparing the energy of calibrated jets to the energy of corresponding 'truth jets' (reconstructed from simulated MC events before detector simulation - including all stable hadrons from the hadronization of the generated partons and any muons or neutrinos from semileptonic jet decays).

Both corrections are applied to the selected b-jets used to build the Higgs candidate, leading to an improvement of 14% in the dijet mass resolution (which is typically of 11%). The impact of both corrections applied in the analysis of Run-1 LHC data is shown in Figure 5.1(a). In the search presented in Chapter 7 the muon-in-jet correction is applied to large-R jets as well, while the p_T-reco correction is derived and used only for small-R jets. Figure 5.2 shows the impact of muon-in-jet and p_T-reco corrections for small-R jets in the analysis of LHC Run-2 data presented in Chapter 7, compared to the standard jet calibration applied to hadronic jets.



(a) (b)

Figure 5.1: Dijet-invariant-mass distribution for the decay products of a Higgs boson with $m_H = 125$ GeV in the 2-lepton MVA selection. The distributions are shown (a) using jets after global sequential calibration (GSC, solid), and after adding muons inside jets (dotted) and after correcting for resolution effects specific to the kinematics of the decay of a Higgs boson with $m_H = 125$ GeV (dash-dotted); (b) using jets after global sequential calibration (GSC, solid), and after adding muons inside jets and applying the kinematic fit (dash-dotted). The distributions are fit to the Bukin function [170] and the parameter representing the width of the core of the distribution is shown in the figures, as well as the relative improvement in the resolution with respect to jets after the global sequential calibration.

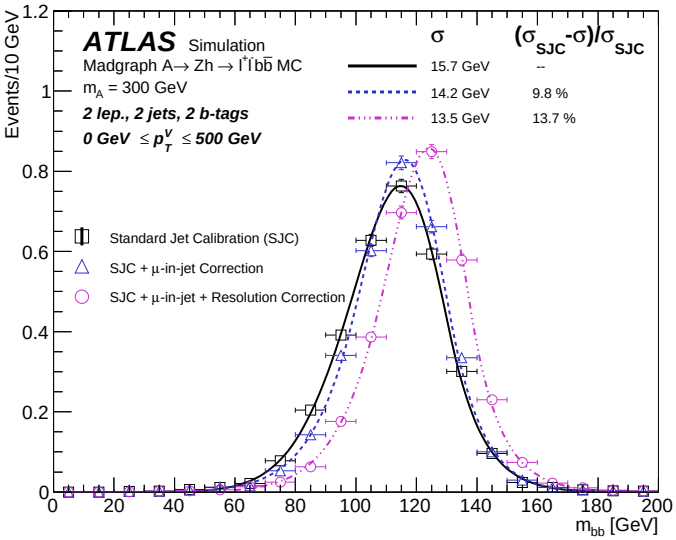


Figure 5.2: Dijet invariant mass distribution for the decay products of a Higgs boson with $m_h = 125$ GeV in the 2-lepton low-p_T^Z 2-tag category in simulated $gg \to A \to Zh \to l^+l^-b\bar{b}$ events, from the analysis presented in Chapter 7. The distributions are shown using jets after the standard jet calibration, after the muon-in-jet correction, and finally after the p_T-reco (resolution) correction. The distributions are fit using the Bukin function and the parameter representing the width f the core of the distribution is shown above, as well as the relative improvement with respect to the jets after the standard jet calibration.

5.2.2 Kinematic fit

As outlined in Section 2.3.3 the $ZH(b\bar{b})$ process is studied in the 2-lepton channel, where the Z boson decays leptonically to a pair of electrons or muons. In this topology there is no 'true' E_T^{miss} (originated from undetectable particles in the detector) other than the small contribution from neutrinos coming from semileptonic heavy flavor decays, therefore the full $ZH(l^+l^-b\bar{b})$ event kinematic can be reconstructed.
Since the lepton resolution is much higher than the jet resolution, the jet energy can be corrected in order to fully balance the transverse momentum of the event, improving its resolution and, as a consequence, the resolution of the invariant mass of the Higgs boson candidate.
This is achieved with a kinematic likelihood fit (KF) which combines a Breit-Wigner constraint on the dilepton invariant mass, a Gaussian constraint on each component of the transverse momentum of the $(l^+l^-b\bar{b})$ system (centered at zero, with width of 9 GeV determined from ZH MC simulated events), dedicated transfer functions relating the true jet transverse momenta to their reconstructed values (after applying muon-in-jet correction, without p_T-reco resolution correction), and finally a prior constraint obtained from the true jet p_T spectrum in simulated ZH events (which plays a role similar to the p_T-reco correction).
This procedure is applied in the 2-lepton channel, in both the analysis of Run-1 and Run-2 data with the same kinematic function, leading to an improvement of $O(20\%)$ of the Higgs candidate mass resolution, which is shown in Figure 5.1(b) compared to the mass resolution obtained applying only the GSC correction for the analysis of Run-1 data.

5.2.3 Overlap removal

The 'overlap removal' procedure takes into account all reconstructed objects and applies specific criteria to avoid double-counting, treating objects which are reconstructed from the same detector signature (for instance an electron, which is also reconstructed as fake-jet). Only small-R jets are used for the overlap removal. Differences between the techniques used in the analyses of Run-1 and Run-2 are highlighted in the following description of the overlap removal steps: when two cut values are quoted, the one in parentheses refers to the Run-2 analyses.

1. if an electron and a muon share the same matched ID track, the electron is removed

2. if a jet and an electron are separated by $\Delta R < 0.4(0.2)$ the jet is discarded

3. (Run-2 analyses only) if a surviving jet and an electron are separated by $\Delta R < 0.4$, the electron is removed

4. if a jet and a muon are separated by $\Delta R < 0.4(0.2)$ the jet is discarded if it has three or fewer matched tracks since it is likely to originate from a muon showering in the calorimeter, otherwise the muon is discarded. The discarded muons are used in the E_T^{miss} computation and in the jet energy corrections described in Section 5.1.4.

5. (Run-2 analyses only) if a surviving jet and an muon are separated by $\Delta R < 0.4$, the muon is removed

6. (Run-2 analyses only) if a surviving jet and an hadronic τ candidate are separated by $\Delta R < 0.2$, the jet is removed

The more complicated treatment of the overlap between muons and jets is introduced to avoid loosing signal-like events where muons from the V boson decay are close to hadronic jets. The selection criteria on the number of tracks for the muons (which are associated to the primary vertex) also helps in reducing the likelihood of considering muons from semileptonic b (or c) quark decays, and improves the rejection of pile-up jets.
Jets satisfying the b-tagging requirement are treated separately in order to properly treat semileptonic decays of the b quarks seeding them (as described in Section 5.1.4), hence they are not considered for the overlap removal procedure.

5.2.4 Truth-tagging

The b-tagging algorithms presented in Section 5.1.4 have high rejection power against c- and light-jets: it is difficult to produce MC samples with large enough statistics after the b-tagging requirements for processes dominated by c- and light-jets. To overcome this issue a dedicated procedure named 'truth-tagging' (or 'tag rate function') is applied to obtain reasonable statistics. The core principle of this method consist in retaining all MC events, whether they contain jets which are indeed satisfying the b-tagging criteria or not, assigning an event-by-event weight which depends on the b-tagging efficiency.
The truth-tagging procedure is applied for the $V+cc$, $V+cl$, $V+l$, and WW samples in the 2-tag categories to produce the template histogram input of the Likelihood fit. Furthermore the training of the BDTs strongly benefits from high statistics samples, therefore the truth-tagging is applied to all events used for the MVA training. This procedure is only applied in the SM $VH(b\bar{b})$ analyses presented in Chapters 6 and 8.
The b-tagging efficiency ϵ_{jet} is a function of the jet kinematic (p_T and η) and its true flavor label (based on the flavor of the MC generated hadrons with $p_T > 5$ GeV within a cone of size $\Delta R = 0.3$ around the jet axis), related to the likelihood of b-tagging the jet.
For each event all possible combinations of b-tagged and non b-tagged jets satisfying the analysis selection are considered: for each combination a single weight is defined as the product of the b-tagging efficiency for each b-tagged jet, times the complement of the b-tagging efficiency for each non b-tagged jet. The truth-tagging weight is finally defined as the sum of the weights for all these possible combinations.
For events containing exactly 2 selected jets (j_1, j_2) in a region where 2 b-tagged jets are required, the truth-tagging weight is defined by the single term:

$$w_{\text{tot}} = \epsilon_{j_1} \cdot \epsilon_{j_2}$$

while for events containing exactly 3-jets, in the same region, the truth-tagging weight is written as the sum:

$$\begin{aligned} w_{\text{tot}} = &(\epsilon_{j_1} \cdot \epsilon_{j_2} \cdot (1 - \epsilon_{j_3}))+ \\ &(\epsilon_{j_2} \cdot \epsilon_{j_3} \cdot (1 - \epsilon_{j_1}))+ \\ &(\epsilon_{j_1} \cdot \epsilon_{j_3} \cdot (1 - \epsilon_{j_2})) \end{aligned} \tag{5.1}$$

and similarly for events with higher jet multiplicities (adding terms to the sum) or different b-tagging requirements.

In some cases it is required to define precisely which are the b-tagged jets among the selected ones, in order for instance to apply specific analysis selections. When this is necessary, a random combination of b-tagged and non-b-tagged jets is chosen among the possible ones with probability given by the single term of the w_{tot} sum corresponding to the specific combination.

An additional complication is introduced to the truth-tagging method when the specific value of the b-tagging algorithm output is used in the analysis, for instance to define multiple exclusive b-tagging categories as described in Section 6.3. In this scenario, in fact, a jet is not only defined as b-tagged or non-b-tagged, but it can satisfy different b-tagging purity criteria (e.g. loose, medium, tight).
A random value of the b-tagging algorithm output weight is drawn for each jet in the event, extracted from the cumulative distribution of the output weights (generated from the the b-tagging efficiency) above or below the b-tagging selection criteria, depending on the b-tagging assignment of the jet in the truth-tagging procedure. The b-tagging algorithm output values are described by discrete distribution comprising one value for each b-tagging operating point (hence these information is called 'pseudo-continuous' b-tagging).

One of the disadvantage of the truth-tagging method comes from ignoring possible correlation between the b-tagging efficiency of different jets in an event. When applying the truth-tagging to the MC samples used for the Run-1 $VH(b\bar{b})$ analysis (Chapter 6), a bias in the tagging efficiency for c-jets is observed in events with two c-jets, as a function of $\Delta R(c,c)$, by comparing the direct- and truth-tagging approaches. From this comparison a reweighting correction function of $\Delta R(c,c)$ is extracted and applied in the analysis, with a corresponding systematic uncertainty quoted as half the size of the correction.
This $\Delta R(c,c)$ correction is not found to be necessary for the Run-2 $VH(b\bar{b})$ search presented in Chapter 8

5.3 Objects and event reconstruction in the analysis of $\sqrt{s} = 7$ TeV data

The main features of the object and event reconstruction in $\sqrt{s} = 7$ TeV data are very close to what described in the previous sections concerning the analysis of $\sqrt{s} = 8$ TeV data, with a few small differences, which are only relevant for inclusion of the $\sqrt{s} = 7$ TeV dataset in the measurement presented in Chapter 6, detailed in Section 6.7.

- electrons: loose electrons are defined with a transverse energy cut at 10 GeV, rather than 7 GeV. Tight leptons are required to pass a looser calorimetric isolation, down to 7% (from 4%).
- hadronic jets: the global sequential calibration is not applied. The cut on the scalar sum of the p_T of the tracks matched to the jet, originating from the primary vertex, is raised from 50% to 75%.
- b-tagging: the b-tagging algorithm employed is MV1 [171, 172], predecessor of the improved version used in the analysis of $\sqrt{s} = 8$ TeV data (MV1c). A single operating point is available for this tagger with an efficiency of 70%.
- overlap removal: double-counting between reconstructed muons and hadronic jets does not include any requirement on the number of tracks associated with the jet.

6 | Search for Standard Model $VH(b\bar{b})$: Run-1 data

6.1 Introduction: The Analysis Strategy

In this Chapter I describe the search for the $b\bar{b}$ decay of the Standard Model Higgs boson in $(W/Z)H$ associated production, performed with data collected by the ATLAS detector during the Run-1 of the LHC [173] at the center of mass energies of $\sqrt{s} = 7$ and 8 TeV. An analysis of the 7 TeV dataset had already been published by the ATLAS collaboration [174], and a preliminary result for the full Run-1 dataset (including data collected at $\sqrt{s} = 7$ and 8 TeV) had been presented in [175]: the analysis discussed in this thesis not only benefits from a larger dataset with respect to [174], but also includes several improvements in the analysis strategy and the object reconstruction not included in [175].
Given the low sensitivity of the 7 TeV data, and the changes in the ATLAS object reconstruction strategies between 7 and 8 TeV datasets, most of these improvements target only the analysis of 8 TeV data. The two datasets are therefore analyzed separately, with minimal changes for the analysis of 7 TeV data with respect to [174].

This chapter focuses on discussing the improved analysis of the 8 TeV dataset, while the main differences in the 7 TeV data analysis are outlined in Section 6.7. The datasets used in this analysis as well as the simulated samples employed for the signal and background predictions are outlined in Section 6.2.

As outlined in Section 2.3.3 this analysis targets separate $VH(b\bar{b})$ final states depending on the leptonic decay mode of the vector boson V, including events containing zero ($Z \to \nu\bar{\nu}$), one ($W \to l\nu$) or two ($Z \to ll$) charged leptons (electrons or muons), named respectively as 0-lepton, 1-lepton and 2-lepton channels. To optimize the analysis sensitivity events are further categorized according to the reconstructed transverse momentum of the vector boson (p_T^V), the number of jets and the number of b-tagged jets. The object reconstruction is described in Chapter 5, and it is not repeated in the analysis chapters, while the event selection and categorization is described in Section 6.3.

One of the defining points of this analysis is the choice of adopting a multivariate (MV) approach to perform the data analysis for the 8 TeV dataset, including different variables in the definition of the final discriminant. The features and needs of a multivariate analysis (MVA) have a strong impact on the definition of the event selection described in 6.2. For the 7 TeV dataset the analysis is performed consistently with [174] using the invariant mass of the Higgs candidate (dijet-mass) as discriminating variable. The dijet-mass analysis is also performed for the 8 TeV dataset in order to gain confidence with the results of the MV approach and provide a separate cross-check of the results. The details of the MVA are described in Section 6.4.

The main backgrounds are described in Section 2.3.3: except for the multi-jet QCD pro-

duction which is estimated with data driven methods, all other background predictions come from Monte Carlo (MC) simulations. The MC modeling of the background is thus one of the main components of this analysis, and special care is needed in the treatment of background processes which present mismodelings in the simulation, with data driven corrections and extensive modeling studies, documented in Section 6.5.

The analysis results are obtained from a binned maximum Likelihood fit which determines the signal yield and the background normalization, following the approach described in Chapter 4. The Likelihood fit includes experimental and theoretical systematic uncertainties as nuisance parameters (NP), which models the deviation from the nominal prediction in the fit: the uncertainties are reflected in penalty terms which constrain the NP in the fit. The details of the Likelihood fit for this analysis are summarized in Section 6.8, along with the treatment of all systematic uncertainties which affect the analysis.

As a validation procedure for the analysis strategy, both for the MVA and the dijet-mass approaches, a measurement of the diboson $VZ(b\bar{b})$ yield is performed in the same final states and with the same event selection: this SM process is remarkably similar to the $VH(b\bar{b})$ signal (replacing the $H \to b\bar{b}$ decay with the $Z \to b\bar{b}$ one) with a larger cross section, and provides a good check of the validity of the analysis (details on the implementation of this separate measurement are presented in Section 6.8.4).

Finally the results of the $VH(b\bar{b})$ analysis for the 7 and 8 TeV datasets, from the MVA and the dijet-mass approaches are presented in Section 6.9, along with the outcome of the validation $VZ(b\bar{b})$ measurement.

6.2 Data and Simulated Samples

The datasets used in this analysis include only *pp* collision data recorded with the ATLAS detector in stable beam conditions and with all relevant sub-detectors providing high-quality data, during the Run-1 of the LHC. Two separate datasets for $\sqrt{s} = 7$ TeV (2011) and for $\sqrt{s} = 8$ TeV (2012) data are analyzed simultaneously, corresponding to integrated luminosities of 4.7 and 20.3 fb^{-1} respectively.

Data quality and event cleaning Only data recorded with stable beam and in optimal functional conditions of the ATLAS detector are used in the analysis, ensuring all essential elements of the ATLAS detector were operational with good efficiency. These are summarized in the ATLAS Good Run Lists (GRL) of runs and luminosity blocks whose high quality has been assessed by the Data Quality ATLAS group, taking into account the status and performance of each sub-detector, beam conditions and objects reconstruction performance.
On top of the GRL request, a series of standard cleaning cuts are applied to avoid sporadic event problems in the detector, during reconstruction, or due to activity in the detector from non-collision background:

- remove incomplete events which do not contain information from all relevant subdetectors
- remove events which triggered error flags from the LAr or TileCal calorimeters
- remove events not containing a primary vertex associated to at least three tracks
- E_T^{miss} cleaning: remove events containing poorly reconstructed jets with $p_T > 20$ GeV and $|\eta| < 4.5$, after overlap removal and before JVF cut, to avoid jets not associated to real energy deposits in the calorimeters.

Trigger selection In the 0-lepton channel events are triggered by the magnitude of the E_T^{miss} vector: the E_T^{miss} trigger configuration evolved during the Run-1 data taking to cope with the increasing luminosity. The E_T^{miss} trigger used for the 7 TeV data taking has a lower threshold at 70 GeV and an efficiency above 50% above 120 GeV, exceeding 99% above 170 GeV. The trigger used for 8 TeV data has an increased E_T^{miss} threshold at 80 GeV.
The dependence of this trigger on the offline E_T^{miss} reconstruction is measured from $W \to \mu\nu$ and $Z \to \mu\mu + jets$ events selected from single-muon triggers, in which the E_T^{miss} is reconstructed including the muon contribution. The luminosity of the 7 TeV dataset is reduced to 4.6 fb^{-1} as a consequence of a brief data taking period in which the E_T^{miss} triggers were not available for the first bunch crossings of two bunch trains.

The 1-lepton channel relies on single-lepton (electron or muon) triggers to select events. The single lepton trigger threshold changed during the Run-1 data taking: the E_T threshold for the single-electron trigger increased from 20 to 22 GeV, while the p_T threshold for the single-muon trigger was set at 18 GeV throughout the 7 TeV data taking period. For the 8 TeV data two sets of single-lepton triggers are considered: a first set with lower thresholds ($E_T > 24$ GeV for single-electron, $p_T > 24$ GeV for single-muon) and isolation criteria, which were not included for 7 TeV triggers; a second set with higher energy threshold ($E_T > 60$ GeV for single-electron, $p_T > 36$ GeV for single-muon) but no isolation requirements. Single-lepton trigger efficiencies are extracted using a tag-and-probe method

from $Z \to ee$ and $\mu\mu$ events.
The 1-muon sub-channel also includes events collected by the E_T^{miss} triggers, to compensate for the limited muon trigger-chamber coverage (since muons are not included in the E_T^{miss} calculation at trigger level, events triggered by significant E_T^{miss} can contribute to the 1-lepton muon sub-channel after the full offline reconstruction).

In the 2-lepton channel events are selected from single-lepton (with same trigger threshold as in the 1-lepton channel) and di-lepton triggers, with thresholds at 12 GeV and 13 GeV for the electron and muon channels respectively.

Monte Carlo Samples Monte Carlo simulated samples are produced for signal and background processes using the `ATLFAST-II` simulation [176], including a complete simulation of the ATLAS detector based on the `GEANT4` program [177], except for the response of the calorimeters for which a parameterised simulation is used.
The $VH(b\bar{b})$ signal from qq-initiated WH and ZH processes is simulated with Pythia 8 [178] with the CTEQ6L1 [179] parton distribution functions (PDFs), the AU2 ATLAS tune [180, 181] for parton shower, hadronisation and multiple parton interactions. For the analysis of 8 TeV data, the Powheg generator [182, 183, 184] has been used with the MiNLO approach [185] (with CT10 PDFs [186], interfaced to Pythia 8 via the AU2 tune), to produce samples for the evaluation of theoretical systematic uncertainties, and also to generate $gg \to ZH$ events at LO(QCD) (for the 7 TeV data analysis this contribution is included only in the total cross section for the normalization of the signal contribution, but not as a separate MC sample), validating its prediction with an independent calculation [67]. Signal samples are simulated for Higgs boson masses from 100 to 150 GeV in steps of 5 GeV, simulating all charged-lepton flavours from W and Z decays, and only the $b\bar{b}$ mode for the Higgs decay.
The qq-initiated VH processes are normalized to the total production cross section calculated at NNLO in QCD [60, 61, 59] with EW corrections at NLO accuracy [187], while the gg-initiated production cross section is computed at LO with NLO corrections [62]. The Higgs boson decay branching ratios are calculated with hdecay [188].

For the main background processes introduced in Section 2.3.3 the following MC generators are used: version 1.4.1 of the Sherpa generator [189] with CT10 PDFs for (W/Z) + jets events at LO in QCD, with massive c- and b-quarks. In order to obtain a good statistical size for the (W/Z) + jets MC samples, even in regions with heavy flavor production or boosted vector bosons, dedicated filters are employed for the generation of this simulation: one filter allows to select events containing b-, c- or light-flavored hadrons (to enhance the statistics of V production in association with heavy hadrons), a second filter selects events according to the vector boson transverse momentum. The (W/Z)+jets samples are thus 'sliced' in several exclusive sub-sample according to the p_T of the vector boson, with boundaries defined as follow:

p_T^V bins = [0, 40, 70, 140, 280, 500) GeV (the last bin contains events with $p_T^V > 500$ GeV)

Top-pair production events are simulated with the Powheg generator with CT10 PDFs interfaced with the Pythia 6 parton shower [190] (for which the CTEQ6L1 PDFs and the ATLAS Perugia2011C tune [180, 181] are used). The $t\bar{t}$ samples include a filter to require at least one W boson to decay leptonically.
The total yield for these backgrounds is determined from data in the Likelihood fit, however they are normalized to theoretical cross sections for the analysis optimization. The

(W/Z) + jets cross sections are computed at NNLO QCD [191], while the $t\bar{t}$ cross section is given at NNLO QCD with next-to-next-to-leading logarithmic (NNLL) resummation of soft gluon terms [192].
V + jets events obtained from SHERPA are labelled according to the flavor of the reconstructed jets with the following procedure: for each jet all generated hadrons with $p_T > 5$ GeV within a cone of size $\Delta R = 0.4$ around the jet axis are considered:

- if a b-hadron is found the jet is labelled as b-jet
- if not, and if a c-hadron is found the jet is labelled as c-jet
- if none of the above the jet is labelled as light-jet

The flavor of the two reconstructed jets used to build the Higgs candidate is used to label each (W/Z) + jets event in

- V+(heavy flavour) (or V+hf), if at least one of the two jets is labelled as b-jet, or both jets are labelled as c-jets
- $V + cl$, if one jet is labelled as c-jet and the other as light-jet
- V+light (or $V + l$), is both jets are labelled as light-jets

where the $V + cl$ subsample is kept separate from the V+hf one, since the main production process is $gs \to Wc$ rather than gluon splitting.

Semileptonic diboson events are modeled with the POWHEG generator with CT10 PDFs interfaced to the PYTHIA 8 parton shower with the AU2 tune (for the analysis of 7 TeV data the HERWIG generator [193] is used, with CTEQ6L1 PDFs and the AUET2 tune [180, 181]). The diboson POWHEG prediction is normalized to the NLO cross section provided by the MC generator itself, cross-checked with the MCFM [194] software, while the HERWIG prediction is scaled to the NLO cross section obtained from MCFM directly (with MSTW2008NLO PDFs [195]). Note that WW processes actually include all possible decays, including fully leptonic final states.
The single-top background from s- and Wt- channels is simulated with the same POWHEG setup used for $t\bar{t}$ events, while t-channel events are modeled by the ACERMC generator [196] interfaced to PYTHIA 6 with CTEQ6L1 PDFs and the Perugia2011C tune, with cross sections taken from [197, 198, 199].

Minimum-bias interaction events are simulated with the PYTHIA 8 generator with MSTW2008LO PDFs [200] and the A2 ATLAS tune [180, 181]. These events are included by overlaying them to signal and background simulated events according to the luminosity profile of the recorded data: these 'pile-up' events are simulated both within the same bunch crossing of the hard-scattering process and in the neighboring bunch crossings.
A complete and schematic list of the MC generators used in this analysis for all signals and backgrounds is reported in Table 6.1.

Process	Generator	$\sigma \times BR$
Signal		
$q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}/l^+l^-b\bar{b}$	Pythia 8	59.1 fb
$gg \to ZH \to \nu\bar{\nu}b\bar{b}/l^+l^-b\bar{b}$	Powheg +Pythia 8	5.1 fb
$q\bar{q} \to WH \to l\nu b\bar{b}$	Pythia 8	131.7 fb
Vector boson + jets		
$W \to l\nu$	Sherpa 1.4.1	12.07 nb
$Z/\gamma^\star \to l^+l^-$ ($m_{ll} > 40$ GeV)	Sherpa 1.4.1	1.24 nb
$Z \to \nu\bar{\nu}$ ($m_{\nu\bar{\nu}} > 5$ GeV)	Sherpa 1.4.1	6.71 nb
Top-quark		
$t\bar{t}$	Powheg +Pythia 6	262.89 pb
t-channel	AcerMC +Pythia 6	87.76 pb
s-channel	Powheg +Pythia 6	5.61 pb
Wt-channel	Powheg +Pythia 6	22.37 pb
Diboson(*)		
WW	Powheg +Pythia 8	52.44 pb
WZ ($m_{ll} > 40$ GeV)	Powheg +Pythia 8	9.241 pb
ZZ ($m_{ll} > 40$ GeV)	Powheg +Pythia 8	3.171 pb

Table 6.1: The generators used for the simulation of the signal and background processes, along with the cross sections times branching ratios (corrected for filter efficiencies when relevant) used to normalize the MC samples. (*) For the analysis of the 7TeV data, no MC simulation is used for $gg \to ZH$ signal (which enters in the analysis only as a normalization factor on the total VH cross section, and Herwig for the simulation of diboson processes.

6.3 Event selection

The detailed definition of the objects (leptons, jets, E_T^{miss}) used in the analysis to reconstruct the event properties and kinematics is given in Chapter 5, leaving to this Section the outline of the selection cuts and the event categories implemented for the analysis of 8 TeV data (while differences in the 7 TeV selection are summarized in Section 6.7).
The analysis has been optimized for the search of a $VH(b\bar{b})$ signal of a SM Higgs with mass $m_H = 125$ GeV, whose phenomenological features are given in Section 2.3 and 2.3.3.

Lepton channels Events are divided into lepton channels by counting the number of reconstructed leptons, to target the different vector boson leptonic decay modes:

- 0-lepton channel: events with exactly zero loose leptons
- 1-lepton channel: events with exactly one tight lepton and no additional loose leptons
- 2-lepton channel: events with exactly one medium lepton and one additional loose lepton of the same flavor (electrons or muons), and no other loose lepton

For 1-lepton and 2-lepton events the objects that satisfied the trigger requirements have to be associated with the selected leptons, for at least one of the lepton triggers considered.

Jet categories Reconstructed jets used in this analysis are described in Section 5.1.3: 'selected jets' are defined as central jets with $p_T > 20$ GeV and $|\eta| < 2.5$ (corresponding to the η validity range for the application of the b-tagging algorithm).
Only events with exactly two or three selected jets are retained, rejecting events containing additional jets with $p_T > 30$ GeV and $|\eta| > 2.5$ to suppress the background from top-pair production, resulting in two jet-categories or jet-bins (2-jet and 3-jet). From this point on, only selected jets are considered to define analysis variables. All selected jets are considered for the b-tagging algorithm, in order to define three exclusive double b-tag (2-tag) categories and one single b-tag (1-tag) category. In all cases at least one of the b-tagged jets must have $p_T > 45$ GeV.
The 2-tag categories are defined by events with exactly two selected jets passing the loose b-tagging requirement, discarding events with 3 selected jets in which the lowest-p_T one is loosely b-tagged. These events are furthermore divided into three exclusive sub-categories:

- 2-tag TT (or Tight) category: events in which the two b-tagged jets satisfy the tight (50% efficiency) b-tagging requirement
- 2-tag MM (or Medium) category: events which don't belong to the TT category, in which both b-tagged jets satisfy the medium (70% efficiency) b-tagging requirement
- 2-tag LL (or Loose) category: events which don't belong to any of the above, in which both b-tagged jets satisfy the loose (80% efficiency) b-tagging requirement

The MM and TT 2-tag categories bring the largest contribution to the analysis sensitivity, while the LL category can provide constraints on the background contributions without two real b-jets.
In the dijet-mass analysis these three 2-tag subcategories are kept separated, to exploit

the information from the different b-tagging purity; in the MVA the b-tagging discriminant variable MV1c is included directly in the BDT definition, therefore the MM and TT categories are merged together for the 0-lepton and 2-lepton channels.
The 1-tag category contains events with exactly one loosely b-tagged jet. Finally an additional zero b-tagged jet (0-tag) category is defined requiring no loosely b-tagged jet, although it is not included in the extraction of the final results but only used for modeling studies. Figure 6.1 shows a scheme describing the events categorization according to the number of b-tagged jets and their purity.
The Higgs candidate in 2-jet events is built from the selected dijet pair, while in 3-jet categories it is formed by the two b-tagged jets in the 2-tag regions, by the b-tagged jet and the highest-p_T non b-tagged jet in the 1-tag region, and by the leading (highest p_T) jets pair in the 0-tag region.

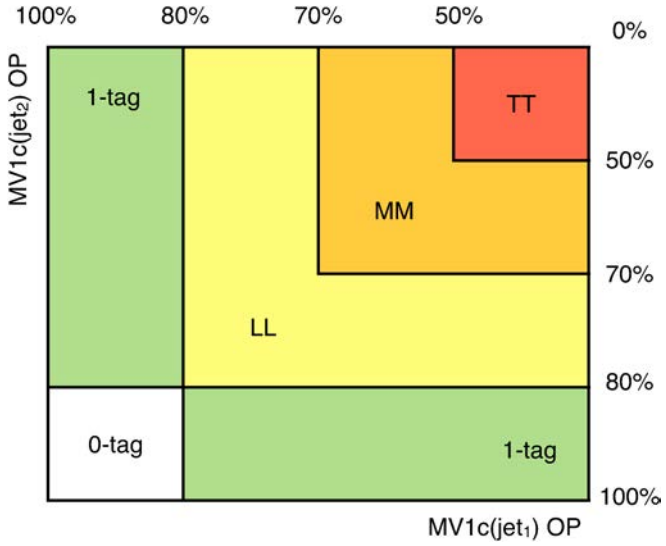


Figure 6.1: Schematic view of the event categorization in b-tagging regions according to the output of the MV1c b-tagging algorithm applied to the selected jets. The bin boundaries corresponds to the MV1c algorithm operating points discussed in Section 5.1.4 for efficiencies of 100%, 80% (loose), 70%(medium) and 50% (tight).

Vector boson transverse momentum p_T^V categories Events selected in the lepton channels and jet categories described above are also separated in multiple regions depending on the magnitude of the transverse momentum of the vector boson (p_T^V), to improve the analysis sensitivity by isolating high-p_T^V events (where the signal over background ratio is larger). In the 0-lepton channel the transverse momentum of the vector boson is defined as the E_T^{miss}, since we can assume its main contribution comes from the neutrinos from the Z decay. In the 1-lepton channel p_T^W equals to the vector sum of the selected lepton transverse momentum and the E_T^{miss} contribution. Finally for 2-lepton events p_T^Z is the vector sum of the transverse momenta of the two selected leptons.

The p_T^V categorization is different in the MVA approach and the dijet-mass analysis. For the multivariate analysis two categories are defined, separating events with $p_T^V \lesseqgtr 120$ GeV; for the dijet-mass analysis 5 separate regions are defined with boundaries at 0, 90, 120, 160 and 200 GeV.
The main reason behind this difference comes from the nature of the MVA: the p_T^V is

included among the input variables used by the MV algorithm to define the final discriminant, which thus includes information from the shape of the p_T^V distribution, making a finer transverse momentum categorization not needed. It is however still useful to separate low-p_T^V and high-p_T^V events for the MVA: when training the MV algorithm over the inclusive p_T^V spectrum in fact, the low-p_T^V region is dominating thanks to the higher statistics, thus decreasing the performance in the most sensitive high-p_T^V category.

The p_T^V categorization is slightly different for 0-lepton events, because of the lower threshold of the E_T^{miss} trigger which makes selecting very low-p_T^V events not possible. In the 0-lepton channel a cut on the scalar sum of the jets transverse momenta $\sum p_T^{jet_i}$ at 120(150) GeV for 2-jet(3-jet) events is applied to reject events for which the E_T^{miss} trigger efficiency is not well modeled, due to its non-trivial dependence on the jet activity. After this cut the E_T^{miss} trigger is 100% efficient above 160 GeV, 97% efficient above 120 GeV and 80% efficient above 100 GeV, with rapidly decreasing efficiency below this threshold. For this reason, only four p_T^V regions are included in the dijet-mass analysis for the 0-lepton channel, with the first p_T^V boundary at 100 GeV rather than 90 GeV, while in the MVA approach the lower boundary of the first p_T^V bin is increased from 0 to 100 GeV.
For similar reasons in the 1-lepton muon sub-channel events collected by the E_T^{miss} trigger are included only for $p_T^W > 120$ GeV.
The definition of three lepton channels, two jet-bins, four b-tag categories (1-tag, 2-tag LL, MM and TT) and multiple p_T^V regions results in a quite complex split of the analysis selection consisting of 44 categories in the MVA approach and 108 categories in the dijet-mass analysis.

Topological and kinematic selection Additional kinematic cuts are applied to further suppress the background contribution, following two main guidelines: the selections cuts are different among lepton channels when they target the specific signature from the leptonic decay of the vector boson, and they are different between MVA and dijet-mass analysis. In the multivariate approach the selection cuts are looser to retain the highest possible statistics for the MVA training, relying on the multivariate algorithm itself to use the input variables information in order to separate the signal from the background.

- A **common selection** cut can be applied on the $\Delta R(\text{jet}_1, \text{jet}_2)$ separation between the two selected jets which form the Higgs candidate, within a $[\Delta R_{min}, \Delta R_{max}]$.
  The ΔR_{min} boundary suppresses the $V+$jets background (which receives large contribution from gluon-splitting $g \to b\bar{b}$).
  The upper ΔR_{max} boundary reject top-pair production background, and can be tightened for higher p_T^V regions since boosted $VH(b\bar{b})$ events tends to produce very collimated b-jets, contrary to b-jets from $t\bar{t}$ events. For the same reason the lower cut is removed in the highest p_T^V bin to increase signal acceptance.
  In the dijet-mass analysis all these cuts are applied, while in the MVA approach only the lower cut is considered and only for $p_T^V < 200$ GeV (but the $\Delta R(\text{jet}_1, \text{jet}_2)$ variable is included in the MVA training).

- In the **0-lepton channel** the main set of cuts is applied to suppress QCD multijet (MJ) and non-collision background, selecting events with large magnitude of the track-based missing transverse momentum vector (p_T^{miss}), small azimuthal angle between E_T^{miss} and p_T^{miss} ($\Delta\phi(E_T^{miss}, p_T^{miss})$), large azimuthal angle between E_T^{miss} and the closest jet ($\min[\Delta\phi(E_T^{miss}, \text{jet})]$) to target signatures with genuine missing transverse energy (instead of E_T^{miss} produced by misidentified or misreconstructed

jets), and large azimuthal angle between E_T^{miss} and dijet system ($\Delta\phi(E_T^{miss}, \text{dijet})$) to target the back-to-back signature typical of signal events. In addition a cut on the $\sum p_T^{jet_i}$ is applied to improve the trigger efficiency modeling as already stated.

An additional set of cuts is applied only for low-p_T^V events ($100 < p_T^V < 120$ GeV), where the MJ background starts to rise because of the looser E_T^{miss} cut: only events with exactly two jets are considered (removing the 3-jet category), with small azimuthal angle between selected jets $\Delta\phi(\text{jet}_1, \text{jet}_2) < 2.7$, large E_T^{miss} significance $\mathcal{S} > 7$ (where $\mathcal{S}$ is the ratio of the E_T^{miss} to the root square of $\sum p_T^{jet_i}$)) and large values of a likelihood ratio built specifically to discriminate against the MJ background $\mathcal{L} > 0.5$.

- In the **1-lepton channel** a cut on the transverse mass m_T^W, defined from the selected lepton and the E_T^{miss} as

$$m_T^W = \sqrt{2p_T^l E_T^{miss}(1 - \cos(\phi^l - \phi^{miss}))},$$

is applied in the dijet-mass analysis to suppress the ttbar background (while m_T^W enters directly in the BDT discriminant, therefore no cut is used in the MVA).
To reject the MJ background, only events with large H_T (scalar sum of E_T^{miss}, p_T of the two leading jets and the selected leptons) or large E_T^{miss} are retained, respectively for the low-p_T^W and high-p_T^W categories. Nevertheless the contribution of the MJ background is significant in this channel, and it is especially sizable for electron events in the low-p_T^W region: given the difficulties in the MJ background modeling and the relatively lower sensitivity of this region, only muon events are considered for low-p_T^W (< 120 GeV).

- In the **2-lepton channel** the only requirements applied are on the invariant mass of the dilepton system (m_{ll}), which has to be consistent with the Z boson mass, and on the E_T^{miss} for the dijet-mass analysis (while the E_T^{miss} is included in the BDT definition for the MVA). The kinematic fit detailed in Section 5.2.2 is applied to 2-lepton events selected by this analysis.

These selection cuts are outlined schematically in Table 6.2. After event selection, the energy calibration of the b-tagged jets is enhanced as detailed in Section 5.2.1, leading to an improvement of the signal dijet-mass resolution of the order of 14%, and a mass resolution of typically 11%. In the 2-lepton channel the kinematic likelihood fit described in Section 5.2.2 is applied, resulting in an improvement of 30% for the dijet-mass resolution.

The goal of this event selection is to reject as much background as possible, and split the phase space to improve the analysis sensitivity, while maintaining the highest signal acceptance. After the cuts, the signal acceptance for the VH signal, with leptonically decaying V boson, is shown in Table 6.3 for the MVA and the dijet-mass based analyses, along with the total signal cross sections (the acceptance for other signal production and decay modes is negligible). The $W(\to l\nu)H$ signal shows significant contamination in the 0-lepton channel (with 7% of the total WH acceptance), and the $Z(\to ll)H$ signal in the 1-lepton channel (with 10% of the total $Z(\to ll)H$ acceptance).

Variable	Dijet-mass analysis					Multivariate analysis	
Common selection							
p_{T}^{V} [GeV]	0–90	$90^{(*)}$–120	120–160	160–200	> 200	0–120	> 120
$\Delta R(\mathrm{jet}_1, \mathrm{jet}_2)$	0.7–3.4	0.7–3.0	0.7–2.3	0.7–1.8	< 1.4	> 0.7 (p_{T}^{V}<200 GeV)	
0-lepton selection							
$p_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	NU	> 30	> 30			NU	> 30
$\Delta\phi(\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}, \boldsymbol{p}_{\mathbf{T}}^{\mathbf{miss}})$		$< \pi/2$	$< \pi/2$				$< \pi/2$
$\min[\Delta\phi(\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}, \mathrm{jet})]$		–	> 1.5				> 1.5
$\Delta\phi(\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}, \mathrm{dijet})$		> 2.2	> 2.8				–
$\sum_{i=1}^{N_{\mathrm{jet}}=2(3)} p_{\mathrm{T}}^{\mathrm{jet}_i}$ [GeV]		> 120 (NU)	> 120 (150)				> 120 (150)
1-lepton selection							
m_{T}^{W} [GeV]	< 120					–	
H_{T} [GeV]	> 180		–			> 180	–
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	–		> 20		> 50	–	> 20
2-lepton selection							
$m_{\ell\ell}$ [GeV]	83-99					71-121	
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	< 60					–	

Figure 6.2: event selection for the different lepton channels and analysis strategies (MVA, dijet-mass). $(*)$ In the 0-lepton selection the first p_T^V interval starts at 100 GeV rather than 90 GeV; in the 1-lepton selection only the muon channel is included for $p_T^V < 120$ GeV (NU = 'Not Used').

$m_H = 125\,\mathrm{GeV}$ at $\sqrt{s} = 8\mathrm{TeV}$				
Process	Cross section × BR [fb]	Acceptance [%]		
		0-lepton	1-lepton	2-lepton
$q\bar{q} \to (Z \to \ell\ell)(H \to b\bar{b})$	14.9	–	1.3 (1.1)	13.4 (10.9)
$gg \to (Z \to \ell\ell)(H \to b\bar{b})$	1.3	–	0.9 (0.7)	10.5 (8.1)
$q\bar{q} \to (W \to \ell\nu)(H \to b\bar{b})$	131.7	0.3 (0.3)	4.2 (3.7)	–
$q\bar{q} \to (Z \to \nu\nu)(H \to b\bar{b})$	44.2	4.0 (3.8)	–	–
$gg \to (Z \to \nu\nu)(H \to b\bar{b})$	3.8	5.5 (5.0)	–	–

Figure 6.3: Cross section times branching ratio and acceptance for the three channels at $\sqrt{s} = 8$ TeV, showing separately qq- and gg-initiated processes for ZH. The branching ratio for $Z \to ll$ includes only decays to electrons and muons, while all three lepton flavors enter in the $W \to l\nu$ and $Z \to \nu\bar{\nu}$ processes. The acceptance is computed considering all events passing the MVA (or dijet-mass analysis) selection in the 2-tag categories combined.

6.4 Multivariate Analysis: Boosted Decision Trees

This analysis is aimed to the search of the SM Higgs boson decaying to a $b\bar{b}$ final state, therefore the invariant mass of the dijet pair selected as the Higgs boson candidate is the primary discriminating variable to separate signal from background. However given the small signal-over-background ratio in the regions considered, the nature of the background processes which can very well mimic the signal signature, and the relatively coarse resolution of the dijet invariant mass, including the discriminating power of additional variables can be extremely beneficial and lead to large sensitivity improvements. The main analysis of the 8 TeV dataset is thus performed following a multivariate approach, defining boosted decision tree (BDTs) [201, 202] selectors which combine the discriminating power of several input variables.
A multivariate decision tree is a binary tree classifier which recursively partitions the parameter space into multiple regions where signal or background purities are enhanced, using multiple discriminating variables and properly accounting for the correlation between them. Boosting is a method which improves the performance and the stability of the decision trees and involves the combination (as weighted average) of many trees into a single final discriminant. The use of a multiple tree classifiers is a crucial step to improve the stability with respect to statistical fluctuations in the training sample.

Dedicated BDTs are built, trained and evaluated for each separate region of the following list: the 2-tag region (merging together LL, MM and TT categories) of the 0-lepton, 1-lepton, and 2-lepton channels, for 2-jet and 3-jet events. In the 0-lepton channel the BDTs are defined only for events with $p_T^V > 120$ GeV, while in the low-p_T^V bin ($100 - 120$ GeV) the dijet-mass approach is used. In the 1-lepton and 2-lepton channels separate BDTs are considered for $p_T^V \leqslant 120$, with no separation between electron and muon channels (since no BDT variable depends specifically on the lepton flavor). The BDTs are defined and trained to separate the $VH(b\bar{b})$ signal from SM background processes. The output of the BDT multivariate algorithm is here defined as BDT_{VH}.

Input variables The BDT input variables are defined as the smallest set of variables which provides the best discriminating power for the final BDT output, avoiding the inclusion of variables which show large mismodelling in the MC simulation with respect to data, or that do not bring significant improvement to the final discriminant. The first variable considered is the dijet invariant mass of the Higgs candidate (m_{bb}) which provides the best sensitivity. Several additional variables are tested by adding them one at a time to the m_{bb}, retaining the one which provides the largest signal versus background separation. This procedure is iterated (by adding each tested variable to the set defined in the previous step) until the inclusion of new variables doesn't bring any significant improvement. The complete list of variables for each lepton channel is outlined in Table 6.4 (the small differences among lepton channels stem from the inclusion of quantities which are specific to the kinematic of the separate channels).
Common variables across lepton channels include: the transverse momentum of the vector boson p_T^V (defined as the E_T^{miss} in the 0-lepton channel), the transverse momentum of the two b-tagged jets ($p_T^{b_{(1,2)}}$), the missing transverse energy E_T^{miss}, the $\Delta R(b_1, b_2)$ separation between the two b-tagged jets, the $\Delta\phi(V, bb)$ azimuthal angle between the vector boson and the dijet system and finally the output of the MV1c b-tagging neural network for the two b-tagged jets. For 3-jet events only, two more variables are included: the transverse momentum of the third jet $p_T^{jet_3}$, and the invariant mass of the 3-jet system m_{bbj} (increas-

ing the sensitivity of the 3-jet category by roughly 30%).
Additional variables are included for specific lepton channels: the pseudorapidity separation $|\Delta\eta(b_1, b_2)|$ between the b-tagged jets (0-lepton, 2-lepton), the pseudorapidity separation between the vector boson and the dijet system $|\Delta\eta(V, bb)|$ (2-lepton), the scalar sum H_T of the transverse momenta of all jets and the E_T^{miss} (0-lepton), the minimum azimuthal separation between the selected lepton and the two b-tagged jets $min[\Delta\phi(l, b)]$ and the transverse mass m_T^W (1-lepton), the invariant mass of the dilepton pair m_{ll} (2-lepton).

Variable	0-Lepton	1-Lepton	2-Lepton
p_{T}^{V}		×	×
$E_{\mathrm{T}}^{\mathrm{miss}}$	×	×	×
$p_{\mathrm{T}}^{b_1}$	×	×	×
$p_{\mathrm{T}}^{b_2}$	×	×	×
m_{bb}	×	×	×
$\Delta R(b_1, b_2)$	×	×	×
$\lvert\Delta\eta(b_1, b_2)\rvert$	×		×
$\Delta\phi(V, bb)$	×	×	×
$\lvert\Delta\eta(V, bb)\rvert$			×
H_{T}	×		
$\min[\Delta\phi(\ell, b)]$		×	
m_{T}^{W}		×	
$m_{\ell\ell}$			×
$MV1c(b_1)$	×	×	×
$MV1c(b_2)$	×	×	×
	Only in 3-jet events		
$p_{\mathrm{T}}^{\mathrm{jet}_3}$	×	×	×
m_{bbj}	×	×	×

Figure 6.4: Variables used in the MVA BDT for the 0-lepton, 1-lepton and 2-lepton channels.

The quality of the modeling of the input variables for the signal and the expected backgrounds (being them estimated from MC simulation or data-driven techniques) is crucial: mismodelings of the distributions used in the BDT training could lead to bias in the output of the BDT discriminant and lower performances of the MVA when applied to data. It is therefore important to study the agreement of the signal and the expected background with the data, not only for the final BDT output, but for each one of the input variables. From this comparison a good agreement between data and expectation is found within the estimated uncertainties. Distributions for a selected set of variables for the three lepton channels, in 2-jet and 3-jet events, are shown in Figure 6.5, 6.6, 6.7, 6.8. In these distributions the multijet background is estimated as described in Section 6.5.1, the corrections to the MC simulation defined in Section 6.5.2 are applied and the background prediction is adjusted in shape and normalization by the combined Likelihood fit described in Section 6.8 through the action of all the included nuisance parameters and normalization factors. The 2-lepton channel expectation is furthermore corrected by the kinematic fit described in Section 5.2.2.
Information on the correlation among the input variables is also used by the BDT to build the final discriminant, thus its modeling has to be studied comparing the expectation to the data: as it is shown in Figure 6.9 a good agreement for the correlation between input

variables is observed between data and MC simulation (shown here for the most powerful variables - m_{bb} and $\Delta R(b_1, b_2)$ - but tested for all pairs of variables).
The correlation between the input variables and the BDT output itself is equally important and can provide information on the impact and the features of individual variables on the final discriminant: this correlation is shown in Figure 6.10 for the dominating variable, the dijet invariant mass m_{bb}, which appears in good agreement between data and MC simulation.

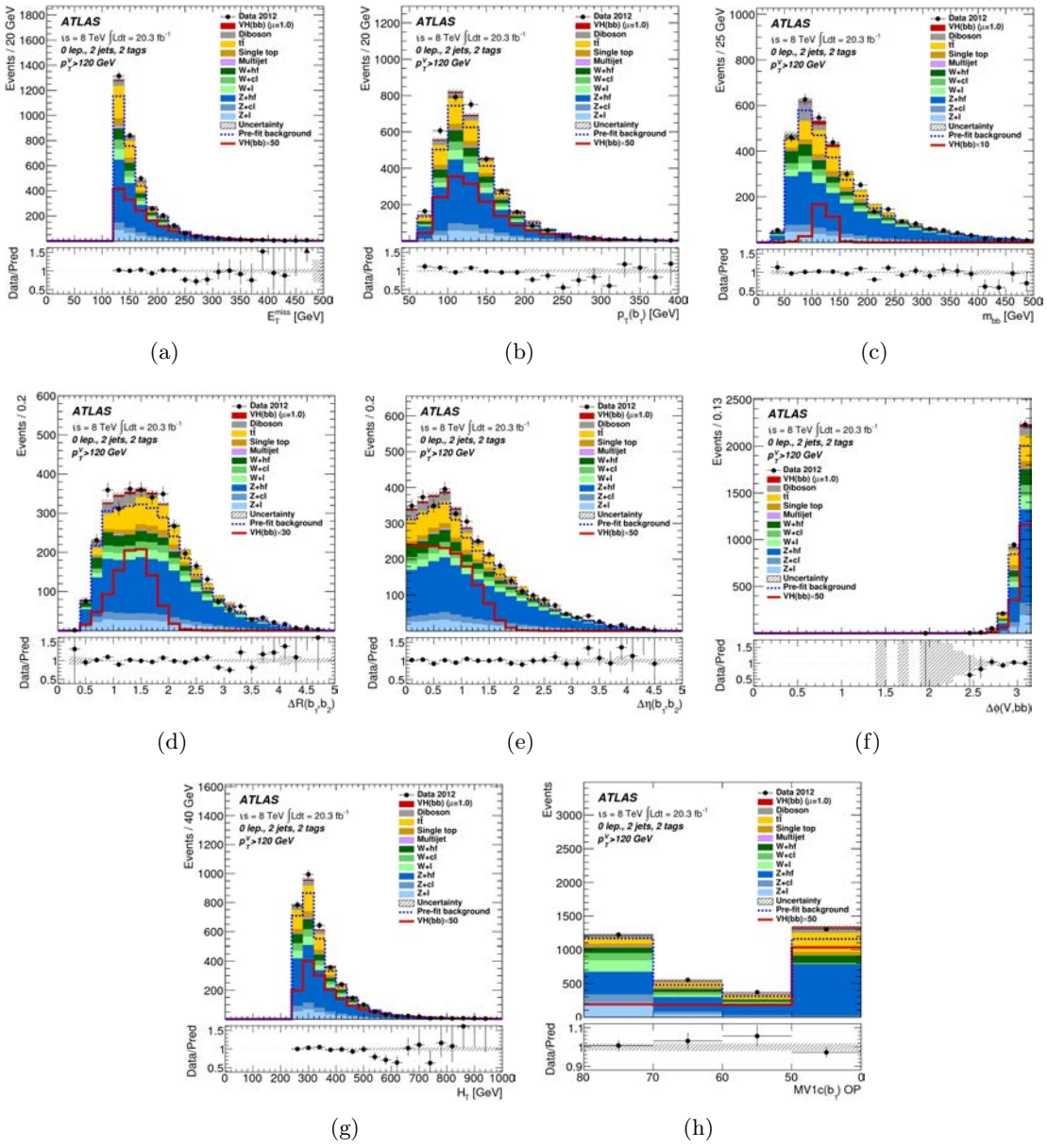


Figure 6.5: BDT input variables in the 0-lepton channel for the 2-jet 2-tag category (LL, MM and TT combined) and $p_T^V > 120$ GeV: (a) E_T^{miss}, (b) $p_T(b_1)$, (c) m_{bb}, (d) $\Delta R(b_1, b_2)$, (e) $|\Delta\eta(b_1, b_2)|$, (f) $|\Delta\phi(V, bb)|$, (g) H_T. In (h) the output of the MV1c b-tagging algorithm is shown for the first identified b-jet, where the left bin boundaries denote the operating points (MV1c(b_1) OP) of the MV1c b-tagging algorithm, corresponding to b-tagging efficiencies of 80%, 70%, 60%, 50%. Background contributions after adjustments explained in the text are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as μ=1.0), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

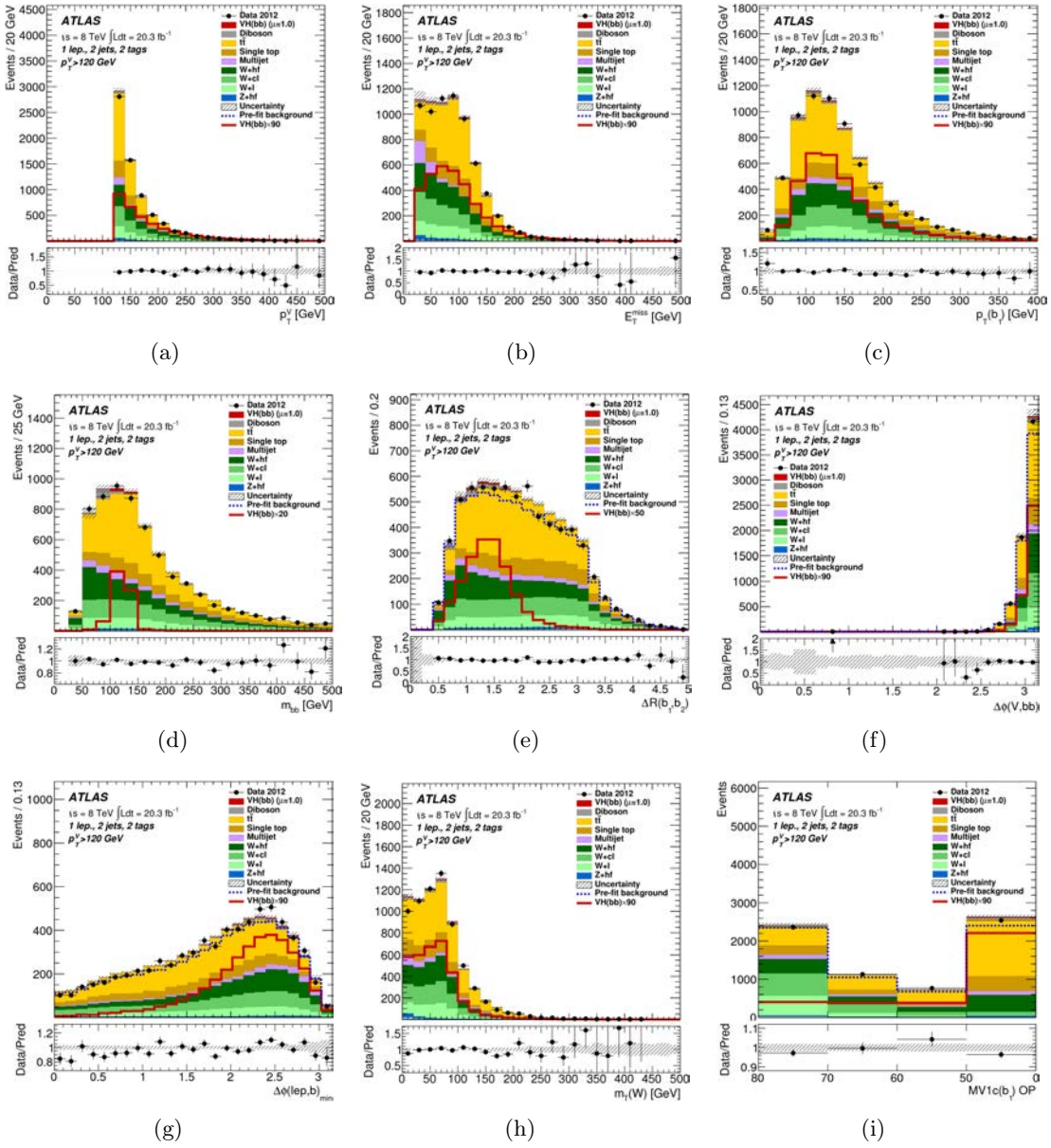
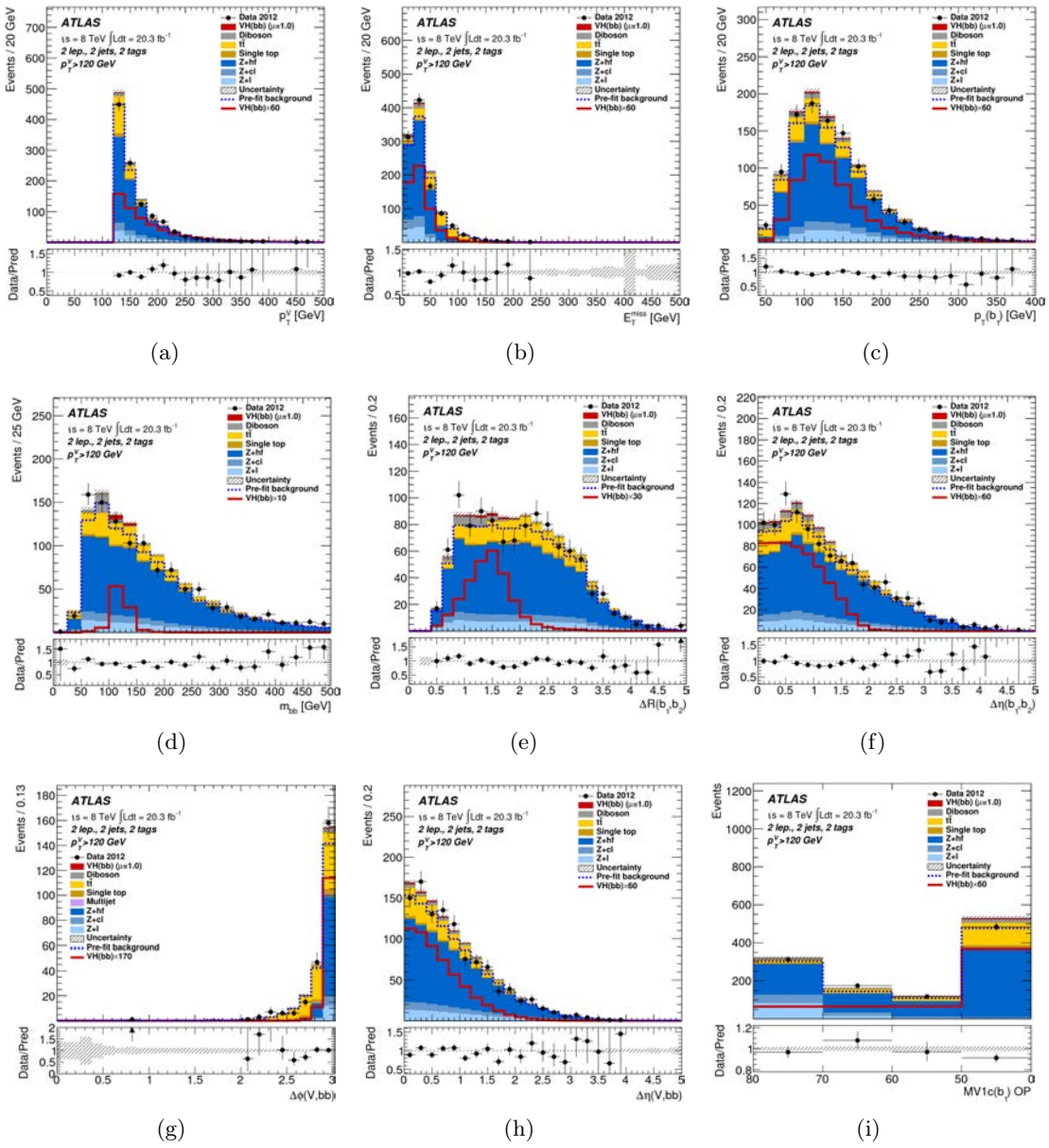


Figure 6.6: BDT input variables in the 1-lepton channel for the 2-jet 2-tag category (LL, MM and TT combined) and $p_T^V > 120$ GeV: (a) p_T^V, (b) E_T^{miss}, (c) $p_T(b_1)$, (d) m_{bb}, (e) $\Delta R(b_1, b_2)$, (f) $|\Delta\phi(V, bb)|$, (g) $\min[\Delta\phi(l, b)]$, (h) m_T^W. In (i) the output of the MV1c b-tagging algorithm is shown for the first identified b-jet, where the left bin boundaries denote the operating points (MV1c(b_1) OP) of the MV1c b-tagging algorithm, corresponding to b-tagging efficiencies of 80%, 70%, 60%, 50%. Background contributions after adjustments explained in the text are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as μ=1.0), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

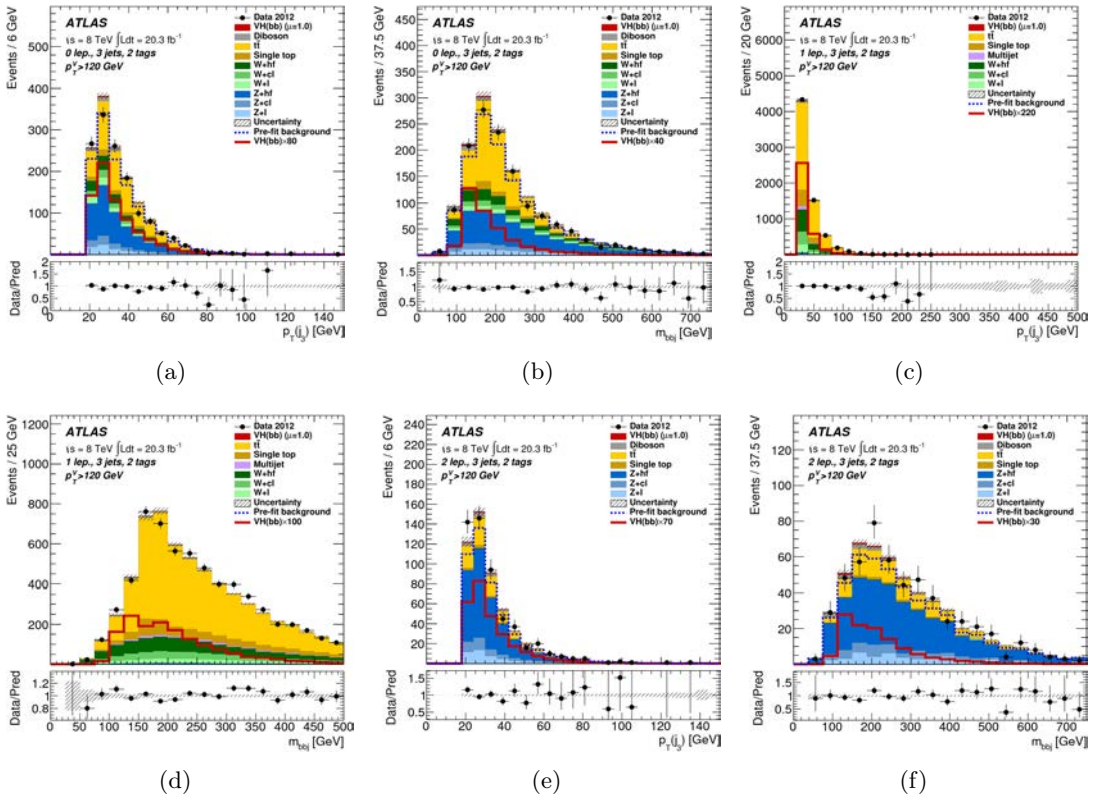


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 6.7: BDT input variables in the 2-lepton channel for the 2-jet 2-tag category (LL, MM and TT combined) and $p_T^V > 120$ GeV: (a) p_T^V, (b) E_T^{miss}, (c) $p_T(b_1)$, (d) m_{bb}, (e) $\Delta R(b_1, b_2)$, (f) $|\Delta\eta(b_1, b_2)|$, (g) $|\Delta\phi(V, bb)|$, (h) $|\Delta\eta(V, bb)|$. In (i) the output of the MV1c b-tagging algorithm is shown for the first identified b-jet, where the left bin boundaries denote the operating points (MV1c(b_1) OP) of the MV1c b-tagging algorithm, corresponding to b-tagging efficiencies of 80%, 70%, 60%, 50%. Background contributions after adjustments explained in the text are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as μ=1.0), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.



(a) (b) (c)

(d) (e) (f)

Figure 6.8: Variables input to the BDT specific for 3-jet events in 2-tag category (LL, MM and TT combined) and $p_T^V > 120$ GeV: (a) 0-lepton, $p_T^{\text{jet}_3}$; (b) 0-lepton, m_{bbj}; (c) 1-lepton, $p_T^{\text{jet}_3}$; (d) 1-lepton, m_{bbj}; (e) 2-lepton, $p_T^{\text{jet}_3}$; (f) 2-lepton, m_{bbj}. The distributions for the 2-lepton channel in (e) and (f) are shown after having applied the kinematic fit. The background contributions after the adjustments explained in the text are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as μ =1.0), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

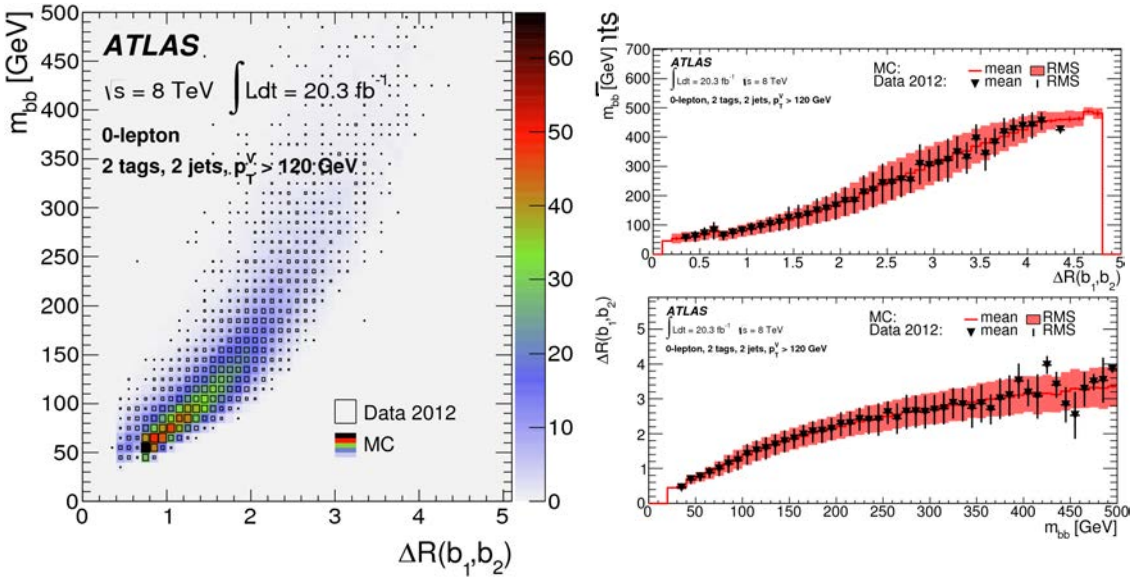


Figure 6.9: Correlation plots for the two most discriminating variables included in the BDT: the dijet-mass m_{bb} and the separation $\Delta R(b_1, b_2)$, shown in the 2-tag 2-jet high-p_T^V signal region (merging LL, MM and TT categories) of the 0-lepton channel. In (a) the 2D scatter plot of m_{bb} versus $\Delta R(b_1, b_2)$ for the total expected background (the shading indicates the number of events) and data (the size of the open boxes is proportional to the number of events); in (b) and (c) the mean values and the RMS of the projections onto the $\Delta R(b_1, b_2)$ and m_{bb} respectively, for the total expected background after the combined Likelihood fit to data.

6.4.1 Boosted Decision Trees

A decision tree is a binary tree structured classifier which performs a sequence of binary splits (yes/no decisions) on a single variable at a time from the full set of input variables until a stop criterion (which limits the complexity of the tree) is fulfilled, splitting the phase space into several regions which can be finally classified as signal-like or background-like according to the majority of training events in the final node. The root node of the tree is the entire training sample consisting of signal (s) and background (b) events: the split is determined at each node of the tree by finding the variable and corresponding cut value that provides the best separation between signal and background. The main principle of decision trees is not dissimilar to a cut-based approach, but it is able to split the phase space in a large number of hypercubes, each labeled as signal or background-like, while a cut-based analysis can only select a single hypercube of phase space for each region considered.

Boosted decision trees extend these concepts from one to multiple trees, all built from the same training ensemble by reweighting the original events, combined in a single MVA classifier given by the weighted average of the individual decision trees. The boosting step stabilizes the response of the decision trees with respect to fluctuations in the training sample. Note however that the advantage of a relatively straightforward interpretation of the decision tree (as consequence of its similarities with the cut-based approach) is lost when combining multiple trees in a single BDT, but the performance improvement and the possibility to strongly reduce the sensitivity of the classifier to fluctuations in the training

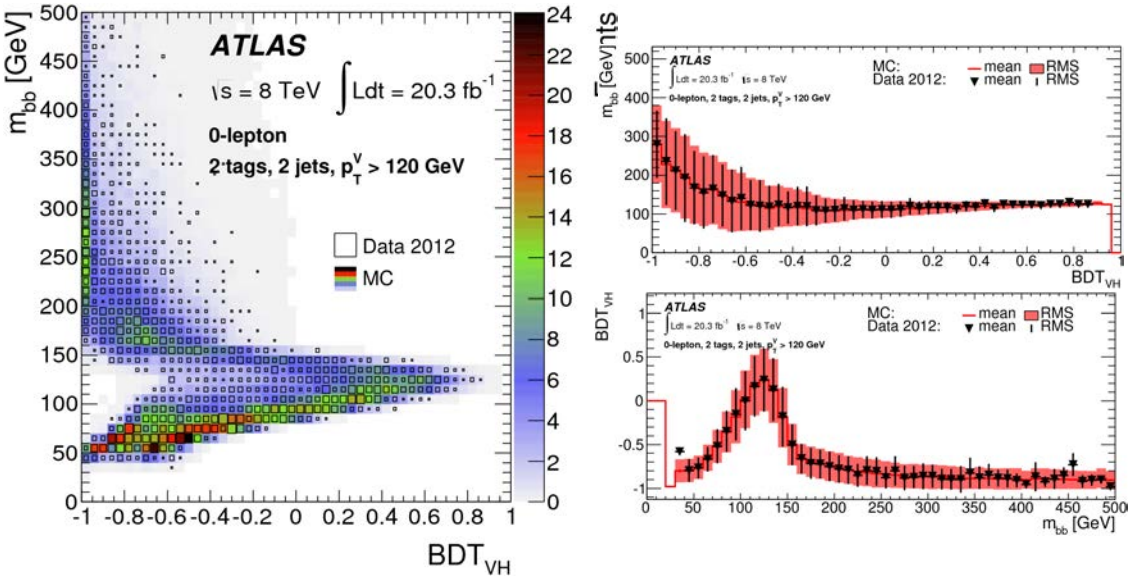


Figure 6.10: Correlations between the dijet-mass m_{bb} and the BDT output discriminant BDT_{VH} in the 2-tag 2-jet high-p_T^V signal region (merging LL, MM and TT categories) of the 0-lepton channel. In (a) the 2D scatter plot of m_{bb} versus BDT_{VH} for the total expected background (the shading indicates the number of events) and data (the size of the open boxes is proportional to the number of events); in (b) and (c) the mean values and the RMS of the proejctions onto the axes of the BDT_{VH} discriminant and dijet-mass respectively, for the total expected background after the combined Likelihood fit to data.

samples (known as *overtraining*) makes this the best option.

The BDTs are still tested to evaluate possible *overtraining*: fluctuations in the training sample can be due to limited statistics or an excessive number of nodes in the trees, and lead to a different response of the BDT between the training sample and an independent sample: a simple *overtraining* test is therefore used to compare the BDT performance between independent training and test samples.
In order to exploit the full available statistics of signal and background simulations without introducing biases, the events are split into two equal size samples (A and B): each BDT is trained on the entirety of A+B events, but the BDTs trained with events from A(B) are tested on the B(A) subsample, to avoid using the same events for training and evaluation of the BDTs. Half of the data are analysed with BDTs trained on sample A and the other half with BDTs trained on sample B. The output distributions of BDTs trained on samples A and B are merged for both simulated and data events.
In order to perform the training of the BDTs with the largest statistics available for the background samples, the truth-tagging procedure described in Section 5.2.4 is applied to all background processes. In the plots presented in this Section, and in the rest of the analysis, the truth-tagging method is only used for $V + cl$, $V + l$ and WW samples in 2-tag categories, in order to reduce statistical uncertainties on these processes which would suffer from low statistics after the double b-tagging requirement.

The BDTs used for this analysis are trained and tested using the Toolkit for Multivariate Data Analysis, TMVA [203]. This tool provides access to several parameters of the BDTs which can be fine tuned (as the number of tree nodes, the training stopping criteria, ...)

whose values have been determined in order to achieve the best expected sensitivity while limiting the complexity of the BDTs. The main parameters with their corresponding values are summarized in Table 6.2: the parameter names refers to the naming convention from [203].

TMVA parameter	Value	Description
`NTrees`	200	number of trees in the forest
`MaxDepth`	4	maximum depth of the decision tree allowed
`nCuts`	100	number of equally spaced cuts tested per variable per node
`nEventsMin`	100	minumum number of training events in a leaf node
`BoostType`	AdaBoost	Boosting type for the trees in the forest

Table 6.2: Main parameters of the BDTs with corresponding values from the TMVA [203] software.

Transformation of the output distribution The BDT output is provided by the MVA algorithm as a distribution between -1 (background-like events) and +1 (signal-like events) stored in histograms with 1000 bins. This histogram is rebinned in order to obtain a smoother distribution in the background-dominated regions (where a finer binning doesn't bring sizable improvements, and less bins provide stability to the Likelihood fit) and a fine binning in the signal-dominated regions, to achieve the best sensitivity while retaining a sufficiently large number of background events in each bin. This approach is viable since the original BDT output distribution does not have a well-defined physical interpretation. The binning transformation is optimized considering the expected sensitivity from the Likelihood fit and the total number of bins.
The core steps of this rebinning are the following: starting from the highest (rightmost) bin of the BDT_{VH} distribution, the bins to the left are merged one by one until the value of a dedicated likelihood-based function Z, which encodes the signal over background information, exceeds 1. This step is then repeated starting from the highest bin which has not been merged in the previous step, until the full distribution has been scanned. Furthermore to limit the number of bins and reduce the impact of statistical fluctuations the expected statistical uncertainty from the total background has to be smaller than 10% in each merged bin. The Z function is defined as:

$$Z = \sqrt{z_b n_b / N_b} + \sqrt{z_s \mathcal{L}_s}$$
$$\mathcal{L}_s = \sum_i s_i \times log(1 + s_i/b_i)$$

where s_i and b_i are the number of signal and background events in the bin i, $N_{b(s)}$ is the total number of background events in the histogram, $n_{b(s)}$ the number of background events in the merged bins, and $z_{s(b)}$ are free parameters optimized to achieve the best rebinning, both set to $z_s = z_b = 4.5$.

A similar strategy is adopted for the dijet-mass analysis as well, to rebin the m_{bb} distribution. The main differences are in the definition of the Z likelihood function which is now written as

$$Z = z_s n_s / N_s + z_b n_b / N_b$$

and in the optimization of the $z_{s(b)}$ parameters which is performed independently for different lepton channels (ranging from 2 to 6).

The m_{bb} distribution is binned more coarsely in the tail of the mass spectrum where the background is dominating, and finer in the central region where the signal to background ratio is largest. A similar effect appears for the BDT_{VH} distribution in which the signal is dominating on the high side of the spectrum. Figure 6.11 shows the impact of the rebinning transformation on the BDT_{VH} distribution and the dijet-mass distribution for one of the 1-lepton signal region, with the caveat that both distributions are mapped to histograms with fixed-size bins, required for the technical implementation of the Likelihood fit.

Alternative MVAs The BDTs are built and optimized to provide separation power between the Higgs signal $VH(b\bar{b})$ with a mass of 125 GeV and the SM backgrounds. In order to perform the analysis across a wider m_H range the same BDTs are re-trained (without changing the parameters optimization) with alternative signal samples corresponding to the different mass points under study, since using the BDTs trained at $m_H = 125$ GeV across the whole mass range would lead to a degradation of the expected limits of approximately 20% for $m_H \neq 125$ GeV.
In a similar way the same BDTs are used to perform the diboson cross-check measurement described in Section 6.8.4, re-training, testing and evaluating the BDTs using the semileptonic diboson processes as signal versus the remaining SM backgrounds (which include in this case the Higgs processes as well). The BDTs binning parameters for the diboson measurement have also been re-tuned, moving the z_s value from 4.5 to 2.5 to further reduce the number of bins with no loss in expected sensitivity.

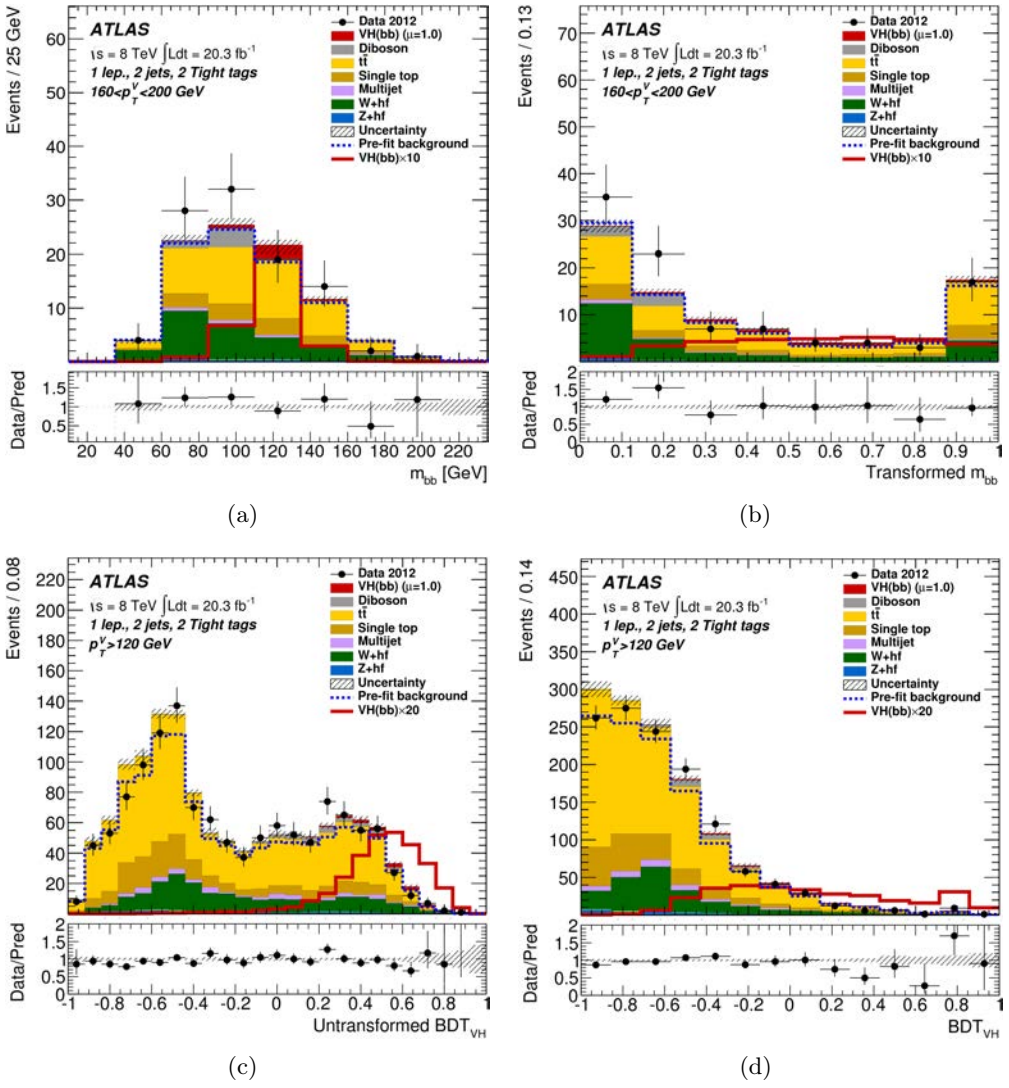


Figure 6.11: Top: dijet-mass m_{bb} distribution for the expected background and signal contributions in the 2-tag 2-jet TT category of the 1-lepton channel with 160 GeV $< p_T^W \leq$ 200 GeV (a) before and (b) after applying the binning transformation. Bottom: BDT output distribution BDT_{VH} for the expected background and signal contributions in the 2-tag 2-jet TT category of the 1-lepton channel with $p_T^W >$ 120 GeV (c) before and (d) after applying the binning transformation. The background expectation corrected by the combined Likelihood fit is shown as filled histogram, while the total expected background contribution before applying any correction from the Likelihood fit is shown as dashed blue histogram. The Higgs signal for $m_H = 125$ GeV is shown as filled histogram in red stacked on the fitted background for a SM expectation corresponding to a signal strength μ =1, and as simple solid red line scaled by a large factor to make it more visible, unstacked to the background prediction. All overflow entries are included in the last bin. The combined statistical and systematic uncertainties on the signal plus fitted background expectation is plotted as hatched band, while the lower panel shows the ratio of the data to the sum of signal and fitted background.

6.5 Background Estimate and Modeling

The quality and the accuracy of the SM background expectation is a crucial part of this analysis: several backgrounds have strong contributions in the most sensitive signal regions and their modeling has to be carefully considered, not only for the discriminating variables used in the Likelihood fit (the BDT_{VH} classifier, the m_{bb} spectrum and the distribution of the b-tagging algorithm output weights for the selected dijet pair MV1c($b_{1(2)}$)), but for all the variables entering in the analysis selection and in the construction of the final BDT discriminant.
The estimate of the QCD multijet background which is used as input in the Likelihood fit is obtained from a data-driven technique, detailed in this Section. For all remaining backgrounds the expected distributions used in the Likelihood fit rely on MC simulations, which are studied in detail considering both comparisons to alternative MC generators and data-to-MC comparisons to assess potential issues and derive the corrections described in this Section.

The Likelihood fit is able to adjust the shape and the normalization of the backgrounds (within certain constraints corresponding to the systematic uncertainties documented in Section 6.6) to exploit information from the data. The normalization of the dominant V + jets and $t\bar{t}$ backgrounds is left free to float in the fit and it is therefore determined from the data.
Given the strong role of the Likelihood fit to data in the determination of the final background modeling, all the distributions shown in this Section, unless otherwise specified, show the backgrounds after the Likelihood fit corrections in shape and normalization.

6.5.1 Estimate of the multijet background

As described in Section 2.3.3 multijet production from strong interactions has an extremely large cross-section an can thus lead to significant background in the analysis, even with effective MJ-suppression selection cuts. Two main types of MJ processes affect the $VH(b\bar{b})$ final states: the first coming from the misidentification of jets or photon conversions as electrons or from semileptonic decays of heavy flavour jets, which contributes mostly to the 1-lepton and 2-lepton channels. The second arising from strong fluctuations in the jet energy measurement in the calorimeters, which create 'fake' contributions to the E_T^{miss}, and affects mostly the 0-lepton channel. The lepton channel which suffer the most from this background is the 1-lepton, since the high E_T^{miss} requirement in the 0-lepton and the signature from two well reconstructed leptons in the 2-lepton allow to strongly suppress the MJ contribution.
For these MJ processes no reliable MC simulation is yet available, therefore their estimate has to rely on data. Different data-driven techniques are adopted for the three lepton channels, given the different nature of the background contributions.
Systematic uncertainties on the estimate of the MJ background for all the lepton channels have been studied and are documented in Section 6.6.2.

Multijet in 0-lepton channel The technique adopted in this channel is based on the definition of three background-enriched regions (named B,C and D) and one signal region corresponding to the baseline 0-lepton selection (named region A). The background estimate in the signal region is taken from the multijet contribution in the regions B, C and D with the following method.
Two of the variables described in Section 6.3 are considered: the min$[\Delta\phi(E_T^{miss}, jet)]$ (if

the E_T^{miss} arises from mismeasured jets, its direction is expected to be close their direction) and the $\Delta\phi(E_T^{miss}, p_T^{miss})$ separation (in events without 'fake' E_T^{miss} the calorimeter-based and the track-based missing transverse momenta should have similar directions).
The signal(VH)-region (A) is defined by the nominal selection cuts introduced in Section 6.3, $\min[\Delta\phi(E_T^{miss}, jet)] > 1.5$ and $\Delta\phi(E_T^{miss}, p_T^{miss}) < \pi/2$. Three background-enriched regions are defined by reverting only the $\Delta\phi(E_T^{miss}, p_T^{miss})$ cut (region C), or requiring $\min[\Delta\phi(E_T^{miss}, jet) < 0.4$ (region B and D - with $\Delta\phi(E_T^{miss}, p_T^{miss})$ cut as in region A and C respectively). The correlation between this two variables has been shown to be weak (both in data and MC simulation), thus they are good candidates for this technique.
The statistics of the four regions, after subtracting the EW background estimate, is fairly low, therefore the b-tagging requirement is removed from regions B,C and D, and a normalization correction factor R is applied to the final estimate as the fraction of 2-tag events in region D. The MJ background template in region A is obtained from data events in region B, C and D after subtracting the remaining EW backgrounds:

$$N_{QCD}(A) = \frac{N(B)}{N(D)} \times N(C) \times R$$

The MJ background contribution in the 0-lepton channel is estimated at approximately 1% of the total background.

Multijet in 1-lepton channel The MJ estimate in this channel rely on the definition of a background-enriched region by loosening some of the lepton identification criteria to increase the first type of MJ contribution described above: the MJ template is obtained from this region by subtracting the remaining EW background contribution (adjusted by a preliminary fit to data). The estimate is performed separately in the electron and muon channels and for each signal and control region.
The background-enriched region is defined from the nominal selection using medium leptons instead of tight ones, and looser track- and calorimeter-based isolation criteria: the track-based isolation is widened to intervals of 5%-12%(7%-50%) for electrons(muons) instead of $< 4\%$, while the calorimeter-based isolation is loosened from $< 4\%$ to $< 7\%$.
As in the 0-lepton case the 2-tag requirement strongly reduces the available statistics in the background-enriched region. To increase the sample size, 1-tag events are used to enrich the 2-tag MJ template, since their kinematic properties are similar in the background-region, by assigning to the untagged jet an emulated MV1c value drawn from the MV1c distribution observed in the corresponding 2-tag MJ template. The MV1c distribution chosen from the 2-tag sample depends on the rank (leading or sub-leading) of the untagged jet and on the MV1c value of the tagged jet in the 1-tag events.
Some residual differences are observed between real 2-tag MJ events and pseudo-2-tag MJ events in data, therefore a reweighting is applied on the MV1c variable of the tagged jet and, only in the electron sub-channel, on the $\Delta R(j_1, j_2)$ and p_T^W variables. The 2-tag background-enriched region is thus defined by this procedure, separately for 2-jet, 3-jet, LL, MM and TT categories.

The normalization of the MJ template is determined by fitting the E_T^{miss} distribution separately in 2-jet, 3-jet, 1-tag and 2-tag (merging LL, MM and TT categories) background-enriched regions leaving the normalization of the other EW background processes free to float.
The MJ background is dominant for low p_T^W since it can hardly mimic the signature from

a boosted W decay: in the 2-tag 2-jet $p_T^W < 120$ GeV category it ranges from 11% of the total background in the LL region to 6% in the TT region. Moreover for $p_T^W < 120$ the MJ contribution is twice as large in the electron sub-channel than in the muon sub-channel. The low-p_T^W region is included mainly to provide constraints and the modeling of the major backgrounds (V + jets and $t\bar{t}$) in the Likelihood fit, with relatively small signal-to-background ratio. For these reasons the electron sub-channel is only included for $p_T^W > 120$ GeV, while the muon sub-channel is considered across the whole p_T^W range to provide constraints on the EW backgrounds.
The removal of the low-p_T^W electron sub-channel from the analysis affects also the search performed with 7 TeV data, and it is one of the remarkable differences in the analysis of the 7 TeV dataset with respect to the results presented in [174].

Multijet in 2-lepton channel The strategy adopted follows closely the 1-lepton channel, by loosening identification and isolation requirements for the selected leptons. The MJ template is however determined from a fit of the dilepton-mass distribution m_{ll} in which only the Z + jets and the MJ components are freely floating, while the remaining EW backgrounds come from the MC simulation. The fitted normalization factors are consistent among 0-tag, 1-tag and 2-tag regions. The size of the MJ data sample is strongly reduced in the 2-tag region due to the b-tagging selection, therefore events without any b-tagging requirements are included. Their normalization is scaled by the fraction of 2-tag events, and MV1c values are randomly assigned to the selected jets, according to the fraction of observed events in the 2-tag region with the corresponding MV1c values.
Comparing data and MC prediction in the sidebands of the Z mass peak, the contribution of the MJ background is found to be negligible in the muon sub-channel; the total estimated amount of MJ background is $< 1\%$ of the total background in the 2-lepton channel.

6.5.2 Modeling of the EW backgrounds: corrections and reweightings

In this section I will focus on the description of the quality and the features of the modeling of the main backgrounds of this analysis: V+jets and $t\bar{t}$ production. In both cases extensive studies have been conducted to assess the accuracy of the MC description by comparing it to the data, and corrections have been derived to improve the MC modeling. Backgrounds whose MC simulation does not require specific corrections are studied in full detail to assess possible systematic uncertainties related to their modeling, which are described in Section 6.6.3

V+jets modeling and corrections This background is studied separately for Z+jets and W + jets events. The available statistics is not large enough to allow conclusive modeling studies directly in the 1-tag and 2-tag categories, so instead the 0-tag region is considered to investigate the details of the MC prediction compared to data.
The crucial variable for these studies, encoding prime information on the kinematic of the system recoiling from the Higgs candidate, is the transverse momentum of the vector boson p_T^V. The analysis categorization is indeed defined by splitting the p_T^V spectrum, with the boosted categories providing the highest sensitivity.

For W+jets events in the 0-tag region of the 1-lepton channel sizable discrepancies between the MC description and data are observed: in particular the p_T^W spectrum is found to be

softer in data, as shown in Figure 6.13(a) in the muon sub-channel. The mismodeling of the W transverse momentum is shown to be correlated to a mismodeling of the $\Delta\phi(j_1, j_2)$ distribution, shown in Figure 6.13(b) again in the muon sub-channel. Studies conducted in [175] showed that this kind of discrepancies persists when applying harsh cuts to reject the MJ background, and are consistent for the electron and muon sub-channel, thus they are unlikely to be related to the modeling of MJ background.
To improve the modeling of the MC simulation a set of reweighting corrections is extracted for the W+l and W+cl background components which are dominant in the 0-tag region. Four corrections are computed as function of the $\Delta\phi(j_1, j_2)$ variable, separately for 2-jet and 3-jet events, and for $p_T^W \lessgtr 120$ GeV. The choice of two separate sets of functions for the low- and high-p_T^W regimes is motivated by the fact that a single set of functions inclusive in p_T^W leads to an overestimate of the correction in the high-p_T^W regime. To correct the relative normalization of the $p_T^W \lessgtr 120$ GeV regions with respect to each other between data and MC additional rescaling factors are introduced to increase(reduce) the background yield by 0.7%(5.6%) in the $p_T^W < (>)120$ GeV.
The result of these reweighting corrections is shown in Figure 6.13(c) and 6.13(d), where the modeling of the p_T^W spectrum (and of course of the reweighted $\Delta\phi(j_1, j_2)$ variable) is considerably improved (other variables as the dijet-mass are improved by the reweighting as well). The reweighting functions are applied to the W+l and W+cl backgrounds in all regions and lepton channels. No correction is extracted for the W+cc or the W+hf components, since their contribution to the 0-tag and 1-tag categories of the studied control regions is too low to allow conclusive studies.

For Z + jets events in the 2-lepton channel a similar approach is adopted, although it is easier to define a category with high purity in Z($\to ll$)+jets background: a 0-tag control region is defined from the nominal MVA 2-lepton selection applying the tighter cuts of the dijet-mass analysis on the E_T^{miss} and m_{ll} variables. In this region it is found that a $\Delta\phi(j_1, j_2)$ reweighting analogous to the W + jets one improves the modeling of the p_T^Z: this correction is therefore applied to the Z+l background component in all regions and lepton channels.
A second control region is obtained applying the 2-tag MVA selection limiting the dijet-mass spectrum to $m_{jj} < 100$ GeV and $m_{jj} > 150$ GeV, to obtain a signal-depleted sample: in this control region there is no evidence of a need for the $\Delta\phi(j_1, j_2)$ reweighting but the p_T^Z distribution is mismodeled nonetheless. A dedicated p_T^Z reweighting is therefore extracted and applied to the Z+hf and Z+cc background components in all regions and lepton channels, and it is shown to improve the modeling in the 1-tag categories as well.

$t\bar{t}$ modeling and corrections For the $t\bar{t}$ background it has been shown that the POWHEG generator interfaced to the PYTHIA 6 parton shower describes a p_T spectrum of the top quarks which is too hard compared to the unfolded measurement from ATLAS [204]. To account for this discrepancy a correction, function of the generated top quarks p_T, has been derived, and it is applied in this analysis for the $t\bar{t}$ production process in all regions and lepton channels (a recent higher-order differential cross-section calculations [205] - not available when this analysis has been performed - shows that the fixed order NNLO (QCD) results would improve the description of the top quarks p_T spectrum, correcting at least in part this mismodeling).

6.5.3 Modeling of VH signal: NLO EW differential corrections

As described in Section 6.2 the signal prediction is normalized to the NNLO(QCD)+NLO(EW) cross section, therefore the effect of the NLO EW contributions is already included on the total normalization. However it has been shown [206] that the NLO EW corrections have a strong dependence on the vector boson transverse momentum p_T^V, therefore the HAWK MC program has been used to derive differential cross section corrections Δ_{NLO} as a function of this variable (reconstructed from the MC before detector simulation), along with the associated uncertainty, with special care not to include NLO EW inclusive effects, already accounted for in the signal normalization. Fluctuations due to limited MC statistics are smoothed by fitting the correction above 200 GeV with a first-order polynomial. The correction is derived and applied only for qq-initiated VH production (the effect on gg-initiated production has not been calculated yet).
The uncertainty on the NLO EW correction is dominated by next-order (NNLO) corrections from additional EW vertices, which are expected to be of the order of the squared NLO corrections: to ensure that the uncertainty does not vanish, it is applied as $\max(2\%, \Delta^2_{NLO})$. The size of the Δ_{NLO} EW p_T^V reweighting, with the associated uncertainty, is shown in Figure 6.12 separately for the three VH signals ($llb\bar{b}$, $\nu\bar{\nu}b\bar{b}$ and $l\nu b\bar{b}$).

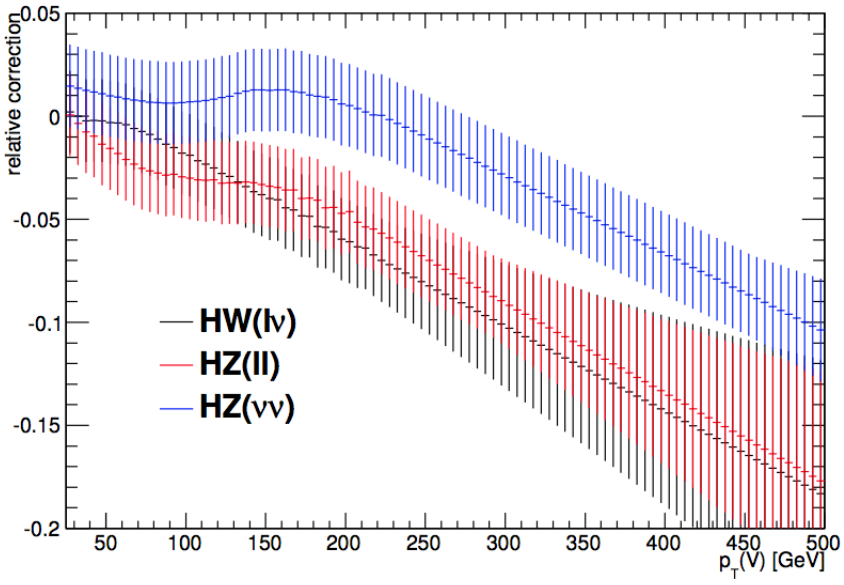


Figure 6.12: Relative NLO electroweak correction to the signal cross sections as a function of p_T^V, and size of the associated uncertainties, calculated from the HAWK MC generator.

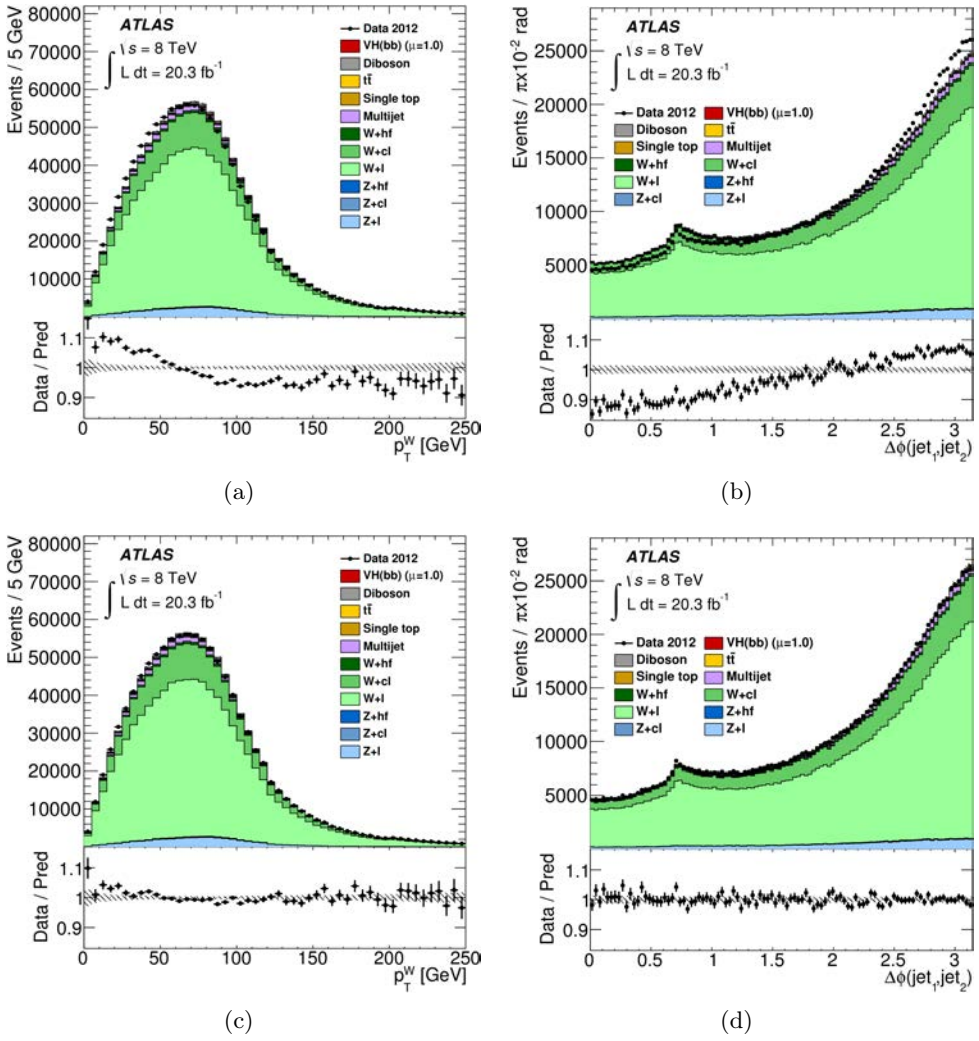


Figure 6.13: The (a) p_T^W and (b) $\Delta\phi(j_1, j_2)$ distributions observed in data (points with error bars) and expected (histograms) in the 0-tag 2-jet control region for the 1-lepton muon sub-channel (MVA selection) before and (c,d - respectively) after the reweighting $\Delta\phi(j_1, j_2)$ function is applied. The multijet and simulated background normalization are provided by the multijet fits. The size of the statistical uncertainty is indicated by the shaded band. The data-to-background ratio is shown in the lower panel.

6.5.4 Background composition in the $VH(b\bar{b})$ analysis

Having introduced the details of the SM background expectation derived from data and MC, as well as the definition of the analysis categorization and selection for both the dijet-mass approach and the MVA, I show in this section the background composition across the different signal and control regions for the discriminating variables used to extract the final results. The details of the background composition are slightly different for the dijet-mass analysis and the MVA and are thus outlined separately.

Dijet-mass analysis Figure 6.14, 6.15 and 6.16 show distributions of the p_T^V and dijet-mass variables for a selection of 2-tag signal regions from the dijet-mass analysis.
The most sensitive categories are the 2-jet 2-tag signal regions with tighter b-tagging requirements: for the higher jet multiplicity the $t\bar{t}$ and V + jets backgrounds increase, while the signal has larger acceptance in the 2-jet region (as a result of the main production modes being qq-initiated rather than gluon-fusion initiated). The background composition is very different across lepton channels, as well as for different p_T^V intervals, jet multiplicities and b-tagging categories:

- **2-lepton channel**: the dominant background is $Z + bb$ across all categories. Top-quark pair production has a significant contribution in the lower-p_T^Z intervals, and the relative semileptonic diboson contribution increases with p_T^Z.
- **1-lepton channel**: in 2-jet events the W+bb and $t\bar{t}$ backgrounds are clearly dominant in the MM and TT categories, while the relative contribution of W+bb and diboson is largest in tighter b-tagging categories and increases with p_T^W. For the $t\bar{t}$ background the flavors of the two selected jets show a dependence on the p_T^W interval: at high-p_T^W, where the top decay products (W, b-quark) are collimated, the bc contribution is large, with the c-quark coming from the $W \to cs$ decay. Significant contribution from single-top-quark production is observed.
  In 3-jet events the $t\bar{t}$ background is the dominant one, with a sizable contribution from single-top-quark production, especially from Wt-channel, and from W+bb (increasing in relative contribution with p_T^W as for 2-jet events).
  The MJ background is non negligible in the 1-lepton channel and can be clearly observed in the lowest p_T^W intervals of the 2-jet category.
- **0-lepton channel**: the dominant processes are Z+bb and $t\bar{t}$, with sizable background from W+bb. The relative $t\bar{t}$ contribution is largest in the lowest p_T^V intervals and in the 3-jet category (as expected)

Multivariate analysis Figures 6.17, 6.18 and 6.19 show distributions of the BDT_{VH} variable for 2-tag signal regions in 2-jet and 3-jet categories, across the three lepton channels.
Since the BDTs are trained to separate signal-like events from background-like events, it can be observed that background processes dominated by light-jets and (to a lesser extent) c-jets are pushed to the lower values of the BDT_{VH} distributions, mainly thanks to the inclusion of the MV1c b-tagging information among the input variables of the MVA. At higher values of the MVA discriminant the background composition is fairly similar to what is observed in the 2-tag signal regions of the dijet-mass analysis, with largest contribution from $V + bb$ and $t\bar{t}$. The top-quark-pair production has however the largest contribution in this analysis due to the looser selection on $\Delta R(j_1, j_2)$.
Figure 6.20 shows the distributions of the output of the MV1c b-tagging algorithm for the

b-tagged jet in the 1-tag categories of the MVA, for 2-jet categories with $p_T^V > 120$ GeV. The four bins considered in these distributions correspond to the four b-tagging operating points and are ordered from left to right in increasing b-jet purity. These plots show clearly how the inclusion of the MV1c variable in the combined fit is able to provide strong constraints on $V + cl$ and V+light backgrounds.

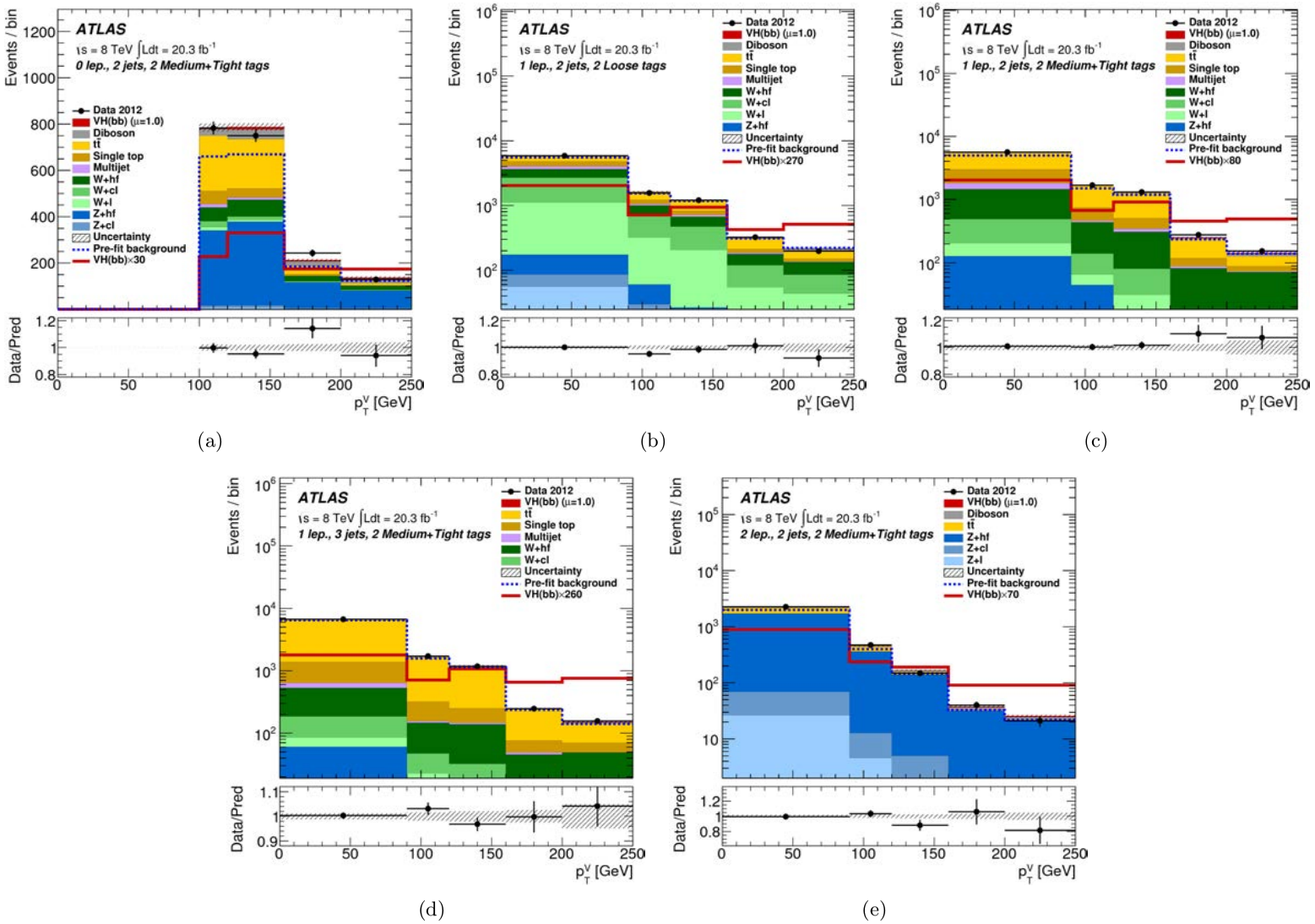
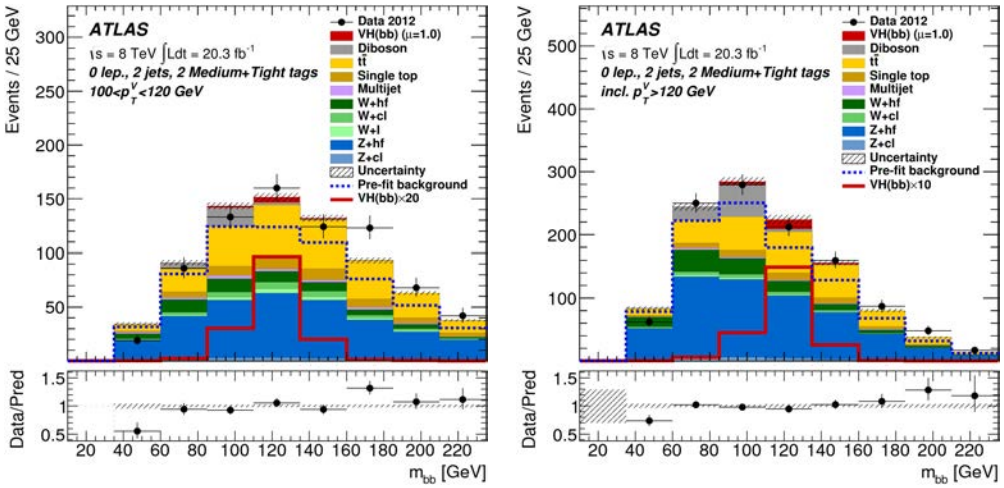
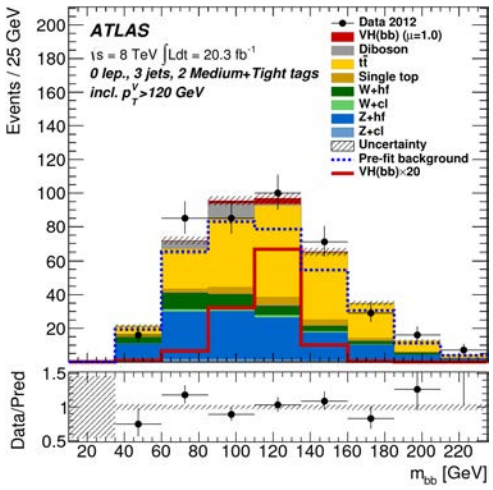


Figure 6.14: The p_T^V distribution observed in data (points with error bars) and expected (histograms) for (a) the 2-jet signal regions of the 0-lepton channel with the Medium and Tight b-tagging categories (also referred to as MM and TT in the text) combined, (b) the 2-jet signal regions of the 1-lepton channel for the LL category, (c) the 2-jet signal regions of the 1-lepton channel with the MM and TT categories combined, (d) the 3-jet signal regions of the 1-lepton channel with the MM and TT categories combined, and (e) the 2-jet signal regions of the 2-lepton channel with the MM and TT categories combined. The background contributions after the global fit of the dijet-mass analysis are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. Overflow entries are included in the last bin. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.



(a)

(b)



(c)

Figure 6.15: The dijet-mass distribution observed in data (points with error bars) and expected (histograms) in the 0-lepton channel with the Medium and Tight b-tagging categories (also referred to as MM and TT in the text) combined for (a) the 2-jet signal region in the $100 < p_T^V < 120$ GeV interval, (b) the 2-jet signal regions in the three intervals with $p_T^V > 120$ GeV combined, and (c) the 3-jet signal regions in the three intervals with $p_T^V > 120$ GeV combined. The background contributions after the global fit of the dijet-mass analysis are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

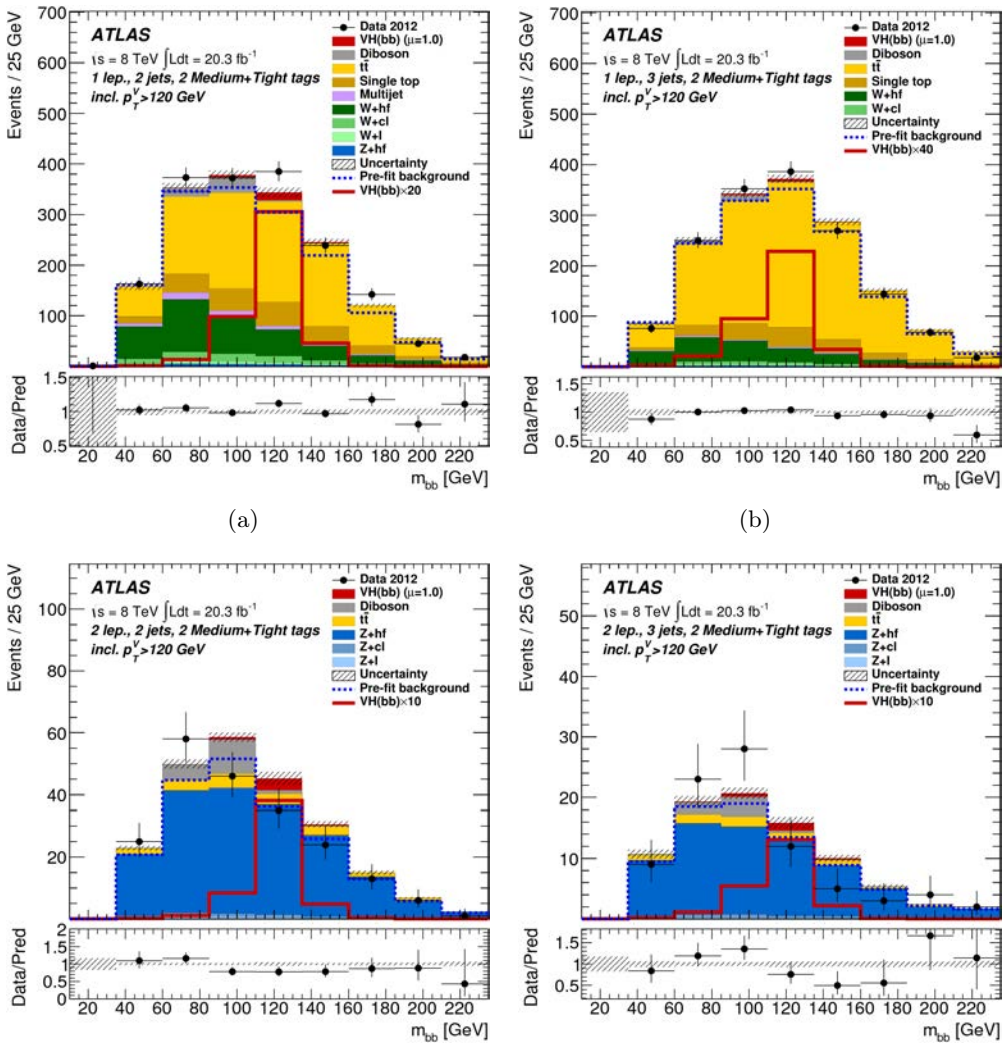
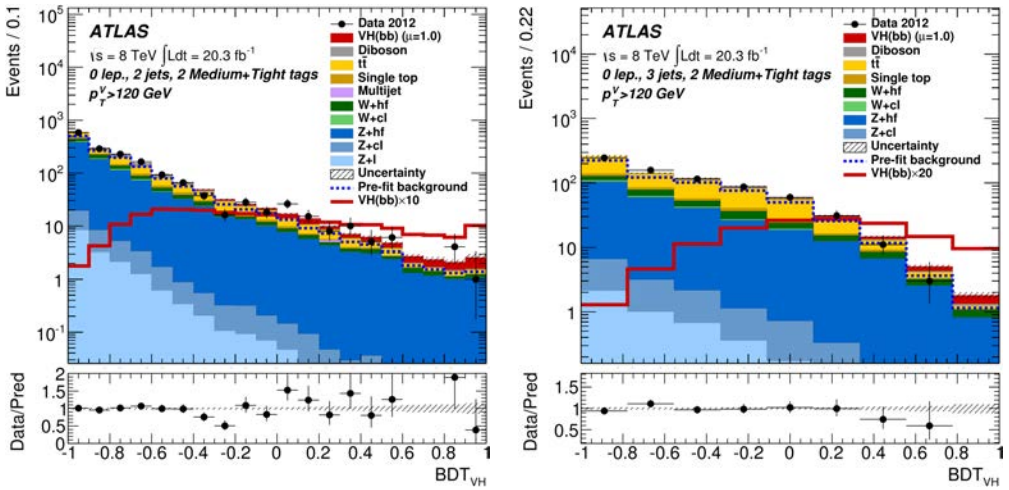


Figure 6.16: The dijet-mass distribution observed in data (points with error bars) and expected (histograms) with the Medium and Tight b-tagging categories (also referred to as MM and TT in the text) combined and the three intervals with $p_T^V > 120$ GeV combined for (a) the 2-jet signal regions of the 1-lepton channel, (b) the 3-jet signal regions of the 1-lepton channel, (c) the 2-jet signal regions of the 2-lepton channel, and (d) the 3-jet signal regions of the 2-lepton channel. The background contributions after the global fit of the dijet-mass analysis are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.



(a) (b)

Figure 6.17: The BDT_{VH} discriminant distribution observed in data (points with error bars) and expected (histograms) for the 0-lepton channel combining the 2-tag Medium and Tight b-tagging categories (also referred to as MM and TT in the text) for $p_T^V > 120$ GeV for (a) 2-jet events and (b) 3-jet events. The background contributions after the global fit of the MVA are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

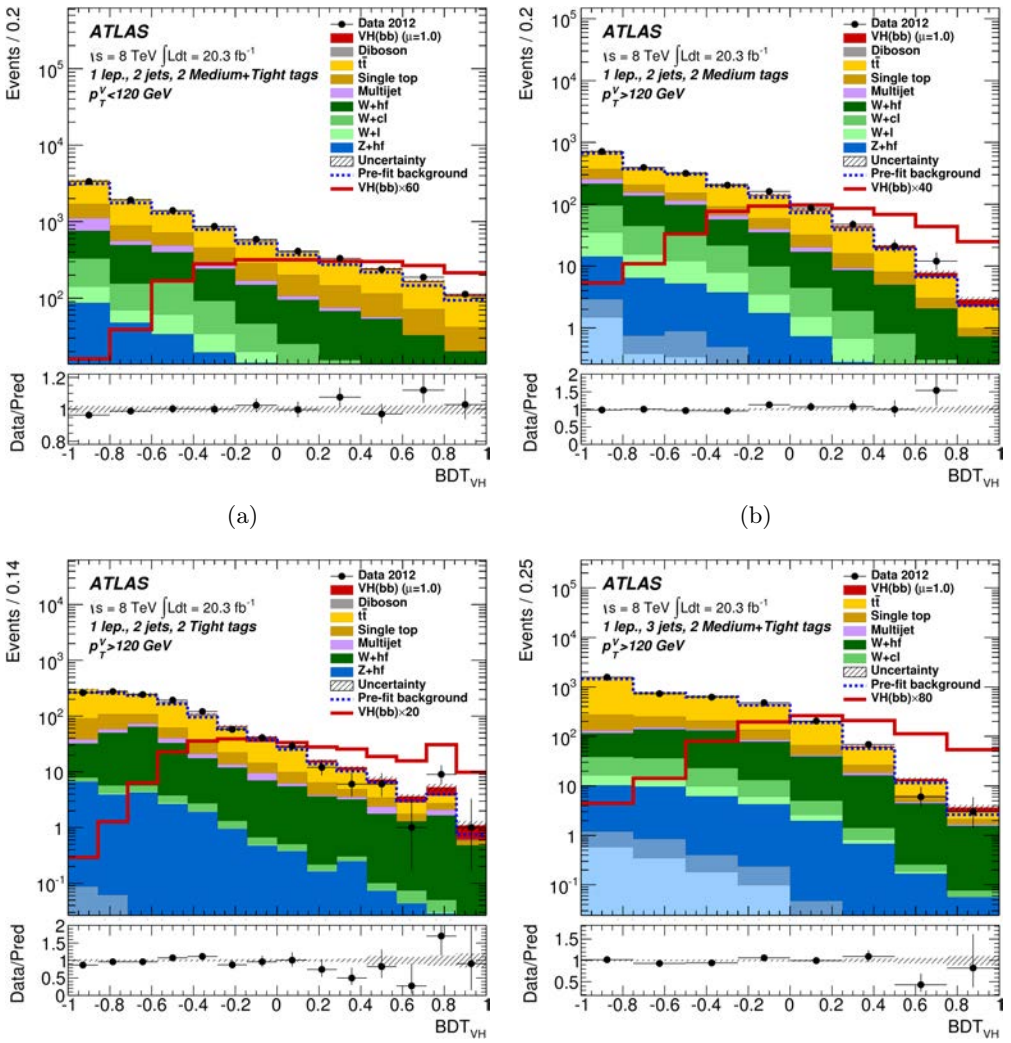


Figure 6.18: The BDT_{VH} discriminant distribution observed in data (points with error bars) and expected (histograms) for the 2-tag signal regions of the 1-lepton channel for (a) 2-jet events with the Medium and Tight b-tagging categories (also referred to as MM and TT in the text) combined and with $p_T^W \leq 120$ GeV, (b) MM 2-jet events with $p_T^W > 120$ GeV, (c) TT 2-jet events with $p_T^W > 120$ GeV, and (d) MM and TT combined 3-jet events with $p_T^W > 120$ GeV. The background contributions after the global fit of the MVA are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

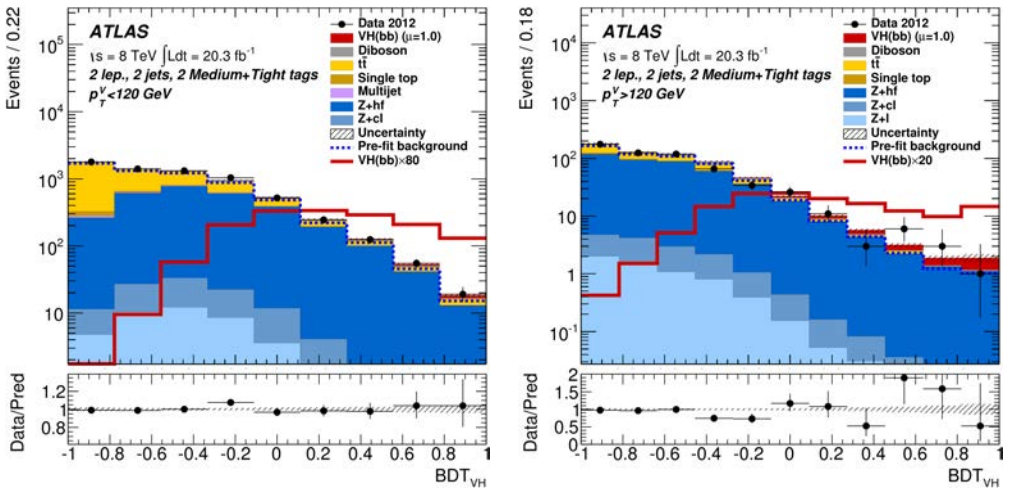


Figure 6.19: The BDT_{VH} discriminant distribution observed in data (points with error bars) and expected (histograms) for the 2-lepton channel combining the 2-tag Medium and Tight b-tagging categories (also referred to as MM and TT in the text) for (a) 2-jet events with $p_T^Z \leq 120$ GeV, (b) 2-jet events with $p_T^Z > 120$ GeV, and (c) 3-jet events with $p_T^Z > 120$ GeV. The background contributions after the global fit of the MVA are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

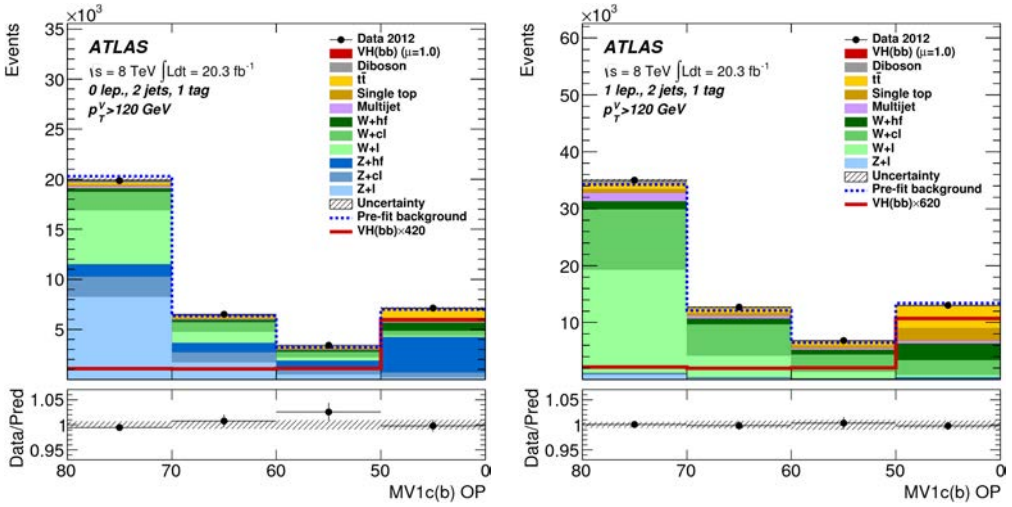


Figure 6.20: Distribution of the output of the MV1c b-tagging algorithm observed in data (points with error bars) and expected (histograms) for the 1-tag control regions of the MVA in the 2-jet category with $p_T^V > 120$ GeV (a) in the 0-lepton channel, (b) in the 1-lepton channel, and (c) in the 2-lepton channel. The left bin boundaries denote the operating points (MV1c(b) OP) as defined in Section 5.1.4, corresponding to b-tagging efficiencies of 80%, 70%, 60%, 50%, i.e., the b-jet purity increases from left to right. The background contributions after the global fit of the MVA are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

6.6 Systematic Uncertainties

The estimate of systematic uncertainties and their contribution to the analysis is a fundamental step of every research work, which needs to be well understood and carefully validated: quoting the result of a measurement without assigning systematic errors is misleading and untrustworthy. In this chapter I describe in detail the numerous sources of systematic uncertainties which affect this analysis, explaining how they are derived and included and what is their impact on the measurement: the experimental systematics, related to the performance of the detector and the reconstruction of measured objects and events; the systematics on the data-driven estimate of the MJ background; the systematics on the MC modeling of the EW backgrounds and the $VH(b\bar{b})$ signal.
In understanding what role these systematic uncertainties play in the final results it is crucial to remember that the hypothesis tested in this search is studied by building and fitting a Profile Likelihood Ratio function, where each systematic uncertainty enters as a nuisance parameter which characterizes the shape of the signal and background probability density functions, with a constraint related to the estimated magnitude of the uncertainty.

6.6.1 Experimental systematic uncertainties

This type of systematics mostly stems from the uncertainty on the corrections to the reconstruction and the calibration of simulated objects in the detector. The $VH(b\bar{b})$ final state in the three lepton channels contains almost all possible physics objects that can be identified in the ATLAS detector, therefore several sources of experimental systematics can play an important role in different regions of the analysis, depending on the targeted final state.

- **integrated luminosity and pile-up**: the uncertainty on the integrated luminosity corresponding to 2.8% is applied to the signal and backgrounds estimated from MC simulation (except the data-driven MJ background), and it is derived following the methodology detailed in [207], from a preliminary calibration of the luminosity scale obtained from beam-separation scans performed in November 2012. An uncertainty of 4% on the modeling of additional interactions per bunch crossing is applied.

- **leptons**: for both electrons and muons, uncertainties are considered on the trigger, reconstruction, identification and isolation efficiencies and on corrections on the energy and resolution. For the 0-lepton channel an additional uncertainty on the lepton veto efficiency is estimated using the 2-lepton selection. The impact of lepton systematic uncertainties in the analysis is fairly small at the order of $O(1\%)$.

- **$\mathbf{E_T^{miss}}$ trigger**: the efficiency of the E_T^{miss} trigger is corrected for the MC simulations with scale factors derived from $W(\mu\nu)$+jets and $Z(\mu\mu)$+jets events (separately). The correction amounts to 4.5% in the low-E_T^{miss} regime (100 GeV $< E_T^{miss} <$ 120 GeV) and 1% in the high-E_T^{miss} regime (above 120 GeV). Systematic uncertainties on this correction are applied, derived from the statistical uncertainty on the correction factor and the difference between corrections from $W(\mu\nu)$+jets and $Z(\mu\mu)$+jets events: the uncertainty is smaller than 1% for high-E_T^{miss} and $O(3\%)$ for low-E_T^{miss} events.

- **jet energy scale (JES) and resolution (JER)**: the calibration of the energy scale of hadronic jets involves several steps, therefore multiple sources of systematic uncertainties are considered, related to in situ calibration analyses, pile-up dependent corrections and flavor composition of jets in different event classes. These uncertainties are decomposed into uncorrelated components and treated as independent

sources in the analysis. The total relative systematic uncertainties on the JES ranges from 3% for central jets with p_T of 20 GeV to 1% for central jets with p_T of 1 TeV. An additional uncertainty of 1-2% is introduced specifically for the energy calibration of b-jets (including uncertainties from MC response differences, lepton energies in semi-leptonic b hadron decays and different composition and response of light quarks and gluons within the jet).
Similarly, a total uncertainty on the JER has been applied, estimated at 10% to 20% depending on the η range (for jets with p_T of 20 GeV) and less than 5% for jets with $p_T > 200$ GeV. This uncertainty already includes a specific contribution derived for b-jets [208]. From in-situ analyses the JER appears well described by the MC simulation: the JER uncertainty includes both the data-to-MC differences observed in these analyses, and the uncertainties on the method used for the JER measurement itself, therefore it is expected to be slightly overestimated and constraints in the Likelihood fit can be expected.

- $\mathbf{E_T^{miss}}$: both the JES uncertainty and the uncertainties on the lepton energy and momentum calibration are propagated to the E_T^{miss} calculation. Additional uncertainties on the energy calibration (8%) and resolution (2.5%) of calorimeter energy clusters not associated with any reconstructed object (the *soft-term* entering in the E_T^{miss} definition) are included.

- **flavor-tagging**: several sources of uncertainties related to the flavour-tagging of hadronic jets are considered. Jets are tagged according to the value of the MV1c discriminant: specific **MC-to-data scale factors** (SFs) are derived by measuring the tagging efficiency in data (with dedicated samples) and simulation, within intervals between operating points and depending on the jet kinematics (p_T and η), as detailed in Section 5.1.4. Uncertainties on these SFs are considered, encoding experimental (e.g. JES propagation), theoretical (e.g. p_T spectrum of $t\bar{t}$ events) and statistical errors. Following the approach adopted for the JES, the uncertainties are decomposed into uncorrelated components, and only the most significant are applied in the analysis, resulting in 10 components for b-jets, 15 components for c-jets and 10 components for light-jets. The impact of the remaining uncertainty components has been checked to be negligible in the analysis. For b-jets the MC-to-data SFs are close to unity, with uncertainties of the order of 2-3% across the jet p_T range, reaching 5% for $p_T = 20$ GeV and 8% above 200 GeV (the accuracy of the b-jet SFs is driven by the analysis of $t\bar{t}$ events in 2-lepton final states).
  An additional set of **MC-to-MC SFs** is derived for b- and c-jets (described in Section 5.1.4), to account for the differences between the MC simulation used in the analysis and the one used for the MC-to-data SFs measurement. Half of the value of the SFs is applied as a systematic uncertainty.
  When performing the **parametrised tagging** (or *truth-tagging*) a correction is applied on the c-jets in $V + cc$ samples in the low-ΔR region, as described in Section 5.2.4. Half of the correction is included as systematic uncertainty.

6.6.2 Uncertainties on the modeling of the multi-jet background

The data-driven estimate of the MJ background is described in detail in Section 6.5.1. In the 0-lepton and 2-lepton channels, where the MJ contribution is estimated at the level of 1%, a systematic uncertainty of 100% is applied to the normalization of this background. This uncertainty includes both systematic and statistical effects in the methods used to determine the MJ background, and it is considered uncorrelated across jet bins and b-tag

categories in the 0-lepton mode.
In the 1-lepton channel the MJ background has a more significant contribution, hence the systematic uncertainties cover the different systematic effects in full detail: they can be divided into uncertainties affecting the shape, the normalization, or derived from the reweighting applied to the MJ template.
Uncertainties on the normalization of the MJ background include both the statistical uncertainty from the MJ fits and the EW background subtraction. Furthermore they include the difference between the MJ fits performed separately in LL, MM and TT categories, or inclusively in the 2-tag region. These uncertainties in the 2-tag 2-jet region amount to 11%, 14% and 22% (in LL, MM and TT categories respectively) for the electron channel, and are three times larger in the muon channel due to the lower statistics of the MJ sample.
Uncertainties on the shape of the MJ template are estimated comparing the nominal MJ background to alternative templates obtained with different isolation requirements: in the electron channel the track-based isolation is varied in the range of 12% to 50% and the calorimeter based isolation is varied from 0% to 4%. In the muon channel alternative templates are derived tightening the track based isolation (in the range of 7%-9.5%) and loosening it (in the range of 9.5%-50%).
Finally a reweighting in $\Delta R(j_1, j_2)$ and p_T^W is applied to the 1-lepton MJ estimate in the electron channel, as described in Section 6.5.1: half of the reweighting is applied as systematic uncertainty for both variables.

6.6.3 Uncertainties on the MC modeling of signal and backgrounds

As already stated this analysis relies strongly on MC simulation to obtain an accurate estimate of the several sources of SM background. The estimate of the background however is also driven by the data sample analyzed, in the sense that all backgrounds enter in a combined Profile Likelihood Fit, and their shape and normalization can be adjusted by the fit to properly model the data distributions.
The degree to which the fit can modify the background estimate obtained from the MC simulation is determined by the features of the nuisance parameters affecting the background itself: these nuisance parameters can be either freely floating normalization parameters (included only for the main backgrounds: V+h.f. and $t\bar{t}$) or constrained parameters on the shape and normalization of the background prediction, where the constraint corresponds to the estimate of the systematic uncertainty on such quantities.
In this section the uncertainties related to the MC modeling of the different backgrounds are described, with the premise that the full model of systematic uncertainties has not only to reflect our theoretical knowledge of the background prediction, but also provide the Likelihood fit with a complete set of handles to obtain an accurate background estimate when fitting the MC expectation to data.
These uncertainties are derived considering, when possible, all the input variables included in the BDTs, or at least the most important ones, starting from the m_{bb} and the p_T^V. They can be divided into two main groups: systematic uncertainties obtained from the comparison of the MC simulation to data, in specific control regions, and uncertainties derived from the comparison of different MC generators among themselves.
The same uncertainty can affect different regions of the analysis: whether to correlate or not the variation across categories depends if the constraint in the Likelihood fit on one region should be propagated to the other regions. This is determined case by case and it is outlined in the description of the fit model in Section 6.8.

Signal systematic uncertainties

Theoretical systematic uncertainty on the signal prediction can be categorized in: uncertainties on the total cross section, uncertainties on the signal acceptance across different analysis categories, and uncertainties affecting the shape of the main discriminating variables.

The inclusive VH cross sections used for the signal normalization have been introduced in Section 6.2: normalization uncertainties on this estimate are obtained from [209] and cover missing higher-order corrections via renormalization (μ_R) and factorization (μ_F) scale variations and PDFs uncertainties. For WH processes, the scale uncertainties amount to 1%, while PDFs uncertainties are quoted at 2.4%. For ZH processes, the uncertainties from [209] are quoted inclusively for qq- and gg-initiated processes: since the two classes of signal processes are included as separate MC simulations in the analysis, we derive separate uncertainties as follow. The $qq \to ZH$ scale and PDFs uncertainties are taken identical to the WH ones (assuming that they are comparable given the similar initial state), while the $gg \to ZH$ values are taken such that the quadrature sum of the uncertainties for qq- and gg-initiated processes gives the estimate obtained from [209]. The scale uncertainties on the gg-initiated ZH contribution are furthermore inflated to account for differences in the number of additional jets between the LO and the NLO (QCD) prediction. This procedure results in scale uncertainties of 50% and PDFs uncertainties of 17% for $gg \to ZH$, while the values for $qq \to ZH$ are identical to the WH case.
The relative uncertainty on the Higgs boson branching ratio to $b\bar{b}$ is 3.3% for $m_H = 125$ GeV [188]. The contribution of decays to final states other than $b\bar{b}$ is verified to amount to less than 1% after selection.

The effect of missing higher-order corrections on the acceptance and the shape of the MC prediction is estimated by generating alternative signal samples with Powheg interfaced to the Pythia 8 parton-shower, with μ_R and μ_F varied independently between 0.5 and 2.0 in all possible combinations.
The *Stewart-Tackmann* procedure [210] is applied to estimate the impact of scale variations to the signal acceptance after kinematic selections applied at generator level. This procedure allows to properly account for the correlation of scale variations between different jet bins (exclusive among themselves), and results in an acceptance uncertainty on the combined 2-jet and 3-jet categories (3.0%, 3.4% and 1.5% for WH, $qq \to ZH$ and $gg \to ZH$) and an anti-correlated acceptance uncertainty on the 3-jet category (4.2%, 3.6% and 3.3% for WH, $qq \to ZH$ and $gg \to ZH$) which is related to the vetoing of 3-jet events applied to define the 2-jet exclusive bin.
These scale variations have a sizable effect on the shape of the p_T^V distribution, therefore an additional uncertainty is derived as function of this variable. For qq-initiated processes this uncertainty starts from a positive effect of 1% at $p_T^V = 50$ GeV and goes to a negative impact of 3% at $p_T^V = 200$ GeV. For gg-initiated processes the same shape effect is observed, with variations from a positive 2% to a negative 8%, respectively. The p_T^V shape uncertainty is obtained from a linear fit performed to envelope the most discrepant scale variations.

PDFs uncertainties on the signal acceptance and shape are obtained in a similar way, following the PDF4LHC prescription: alternative Powheg +Pythia 8 samples are generated with the central values of the MSTW2008NLO and NNPDF23NLO PDF sets, and compared to the Powheg +Pythia 8 samples generated with CT10NLO, after kinematic

selections applied at generator level. Acceptance uncertainties are derived separately for 2-jet and 3-jet events and different signal processes, but they are considered correlated across jet categories, and range from 2% to 5%. No evidence for any shape uncertainties is found from PDF variations.

Uncertainties on the modeling of underlying event, multiple parton interactions and details of the parton-shower tuning are studied comparing different tunes of the PYTHIA 8 parton-shower using the same POWHEG +PYTHIA 8 MC samples employed for the estimate of scale and PDFs effects, and found to be negligible for this analysis.
An uncertainty on the parton-shower and hadronization model used for the signal prediction is derived by comparing POWHEG signal samples interfaced to PYTHIA 8 and HERWIG parton-showers respectively: this comparisons results in an acceptance variation of 6.5% for events with $p_T^V > 120$ GeV, while below 120 GeV the variation is of 8% and 13% for 2-jet and 3-jet events respectively.

Finally the uncertainty on the NLO EW differential correction on the shape of the p_T^V spectrum, already described in Section 6.5.3 is included.
The full model of theoretical systematic uncertainties for the signal prediction is outlined in Table 6.22 along with the SM backgrounds.

Top-pair production systematic uncertainties

The nominal prediction for the $t\bar{t}$ background is obtained from the POWHEG +PYTHIA 6 generator, detailed in Section 6.2. In the global Likelihood fit the normalization of this background is left free to float, separately for the different lepton channels, therefore no systematic uncertainty on the inclusive $t\bar{t}$ cross section, used to normalize the MC samples, is applied.
Theoretical uncertainties are assessed by comparing this nominal estimate with alternative MC simulation encoding several effects:

- PDFs variations: POWHEG +PYTHIA 6 with HERAPDF [211] sample
- parton-shower and hadronization scheme: POWHEG +HERWIG sample
- matrix-element NLO calculation and matching: MC@NLO [212]+HERWIG sample
- varied initial- or final-state radiation (ISR/FSR): ACERMC +PYTHIA 6 sample
- higher-order tree-level matrix elements: ALPGEN [213]+PYTHIA 6 sample

The comparison focuses on the 1-lepton channel selection, where the background contribution is the largest, and shows that the dominant variation is obtained from the ALPGEN +PYTHIA 6 sample, which is thus used to assess most of the $t\bar{t}$ systematic uncertainties described in this section.
From the comparison of the POWHEG +PYTHIA 6 and ALPGEN +PYTHIA 6 simulations an uncertainty of 20% on the ratio of 3-jet to 2-jet events is obtained. This uncertainty is treated as correlated across 0- and 1-lepton channels, while uncorrelated to the 2-lepton channel.
Variations in the m_{bb} and the p_T^V spectra are studied with the same comparison. Uncertainties on the shape of the m_{bb} distribution are derived separately for 2-jet, 3-jet, low-p_T^V and high-p_T^V events, correlated in the global fit across these categories. The variation is larger for higher p_T^V categories: in 2-jet events a positive impact of 3% at $m_{bb} = 50$ GeV

changes to a negative impact of 1% at 200 GeV; the effect is similar but opposite in sign for 3-jet events.
From the study of the p_T^V distribution an uncertainty of 7.5% on the relative acceptance variation between high- and low-p_T^V categories is obtained.
In the 1-lepton channel only, an additional shape uncertainty on the E_T^{miss} distribution is found to be significant, derived separately for low- and high-p_T^V, although applied as correlated between the two regions.

Finally a systematic uncertainty related to the top quark p_T reweighting described in Section 6.5.2 is applied, amounting to half the correction, and correlated across lepton channels.

$Z+$jets systematic uncertainties

In the combined Likelihood fit the $Z+$jets background is divided into three main components, according to the flavor labels of the two selected jets used to reconstruct the Higgs candidate: Z+light, $Z+cl$ and Z+hf (the last including $Z+bb$, $Z+bc$, $Z+bl$ and $Z+cc$). For each one of the three background components two nuisance parameters are included in the fit to control the total normalization and the ratio between 3-jet and 2-jet events (*3-to-2-jet* ratio). For Z+light events the uncertainty on both the normalization and the 3-to-2-jet ratio is derived from data selected in the 0-tag region of the 2-lepton channel, and correspond to 5%.
For $Z+cl$ and $Z+bb$ events the overall normalization is left free to float in the fit, without any constraints or uncertainties, while the 3-to-2-jet ratio is determined from the comparison of the nominal SHERPA MC prediction to the alternative ALPGEN generator, in the 2-tag 2-lepton category, and amount to 26% and 20% respectively.
Furthermore, from the same comparison, uncertainties on the relative acceptance variations between the flavor components of the Z+hf sub-sample (with respect to the $Z+bb$ one, whose normalization is free to float) are estimated at 12% for each bc/bb, cc/bb and bl/bb fraction (uncorrelated between 2-jet and 3-jet events for the bl/bb one).
From a data-to-MC comparison in the 2-tag region of the 2-lepton channel (excluding the region in the 100-150 GeV range) an additional systematic effect is observed, on the shape of the m_{bb} spectrum. This results in a systematic uncertainty on the shape of this variable with a positive effect of 3% at $m_{bb} = 50$ GeV decreasing to a negative effect of 5% at 200 GeV. The m_{bb} shape systematic is treated as correlated between $Z+cl$ and Z+hf components, but uncorrelated for the Z+light one, and it is shown to cover the MC-to-MC comparison between SHERPA and ALPGEN.
The last set of systematic uncertainties is derived from the corrections applied to the Z+jets MC prediction, described in Section 6.5.2. A systematic variation on the $\Delta\phi(j_1, j_2)$ variable is quoted as half the magnitude of the corresponding correction for the Z+light background component, and the full size of the correction for the Z+hf and $Z+cl$ components. This uncertainty is uncorrelated for the Z+light component (but correlated across $Z+cl$ and Z+hf), and between 2-jet and 3-jet events.
A systematic uncertainty is assigned to the p_T^V correction, corresponding to half the size of the reweighting, uncorrelated for Z+light events (but correlated across $Z+cl$ and Z+hf).

$W+$jets systematic uncertainties

The treatment of the $W+$jets modeling is similar to the $Z+$jets one. In the combined Likelihood fit the $W+$jets background is divided into three main components, according to

the flavor labels of the two selected jets used to reconstruct the Higgs candidate: W+light, $W + cl$ and W+hf (the last including $W + bb$, $W + bc$, $W + bl$ and $W + cc$).
For each one of the three background components, two nuisance parameters are included in the fit to control the total normalization and the ratio between 3-jet and 2-jet events (*3-to-2-jet* ratio). For W+light events the uncertainty on both the normalization and the 3-to-2-jet ratio is derived from data selected in the 0-tag region of the 2-lepton channel, and correspond to 5%. The normalization of $W + cl$ and W+hf events is determined from the fit as an unconstrained floating parameter, while for $W + cl$ events an uncertainty on the 3-to-2-jet is quoted at 10% from the same data-to-MC comparison used for the W+light sample.
Since for the W+hf background there are no dedicated control regions to study the accuracy of the MC simulation compared to data, alternative MC samples are used to assess the modeling systematic uncertainties on this background, including several systematic sources: the nominal SHERPA prediction is compared to alternative POWHEG +PYTHIA 8, aMC@NLO +HERWIG ++ (including independent μ_R and μ_F variations between 0.5 and 2.0, and different PDF sets - CT10NLO, MSTW2008NLO, NNPDF2.3) and ALPGEN +HERWIG simulations. The comparison is performed after kinematic selections on generator level quantities which mirror the cuts applied in the analysis, and the largest source of variation is taken as systematic uncertainty.
This procedure leads to an uncertainty on the 3-to-2-jet ratio of 10% from the comparison between SHERPA and the aMC@NLO +HERWIG ++ nominal samples, which is applied to the whole W+hf component. From the same generator comparison an uncertainty on the shape of the p_T^V spectrum is derived, with a positive impact of 9% at $p_T^V = 50$ GeV decreasing to a negative impact of 23% at 200 GeV, treated as correlated across the W+hf sub-samples and uncorrelated between 2-jet and 3-jet events.
From the comparison between SHERPA and ALPGEN +HERWIG a sizable effect on the m_{bb} distribution is observed, and it is applied as systematic uncertainty on the whole W + jets sample, uncorrelated among W+light, $W + cl$, $W + (bb, cc)$ and $W + (bl, bc)$ components, and furthermore uncorrelated between low- and high-p_T^V events for the $W + (bb, cc)$ component. This uncertainty has a a positive impact of 23% for $m_{bb} = 50$ GeV decreasing to a negative impact of 28% at 200 GeV.
The uncertainties on the relative acceptance variations between the flavor components of the W+hf sub-sample (with respect to the $W + bb$ one, whose normalization is freely floating) are estimated from a comparison between SHERPA and ALPGEN +HERWIG after detector simulation and full event reconstruction, applying the nominal 1-lepton channel analysis selection: this comprises effects on the prediction of flavor composition from both the matrix-element calculation and the parton-shower model. The systematic uncertainties are quoted at 12% for bc/bb and cc/bb, and at 35% for the bl/bb fraction (which is uncorrelated among low- and high-p_T^V events).
Similarly to the Z + jets case, the last set of uncertainties is obtained from the $\Delta\phi(j_1, j_2)$ correction applied to the MC simulation for W+light and W+cl components. A systematic uncertainty corresponding to half the magnitude of the correction is applied, uncorrelated between W+light and $W + cl$, and 2-jet and 3-jet events. No correction is applied for W+hf events, but a systematic uncertainty is included corresponding to the full correction for $W + l$ and $W + cl$ events, uncorrelated between 2-jet and 3-jet events.

Diboson systematic uncertainties

The diboson background includes the diboson production processes which originate a final state with two vector boson (WW, WZ, ZZ). Several diboson processes give significant

contribution in the analysis signal regions: $Z \to \nu\bar{\nu}$ and $Z \to b\bar{b}$ for the 0-lepton channel, $Z \to ll$ and $Z \to b\bar{b}$ for the 2-lepton channel; $W \to l\nu$ and $Z \to b\bar{b}$ for the 1-lepton channel. Additional processes can give smaller contributions, thanks to the failed identification of a jet from the W decay or to the failed reconstruction of one of the leptons: $Z \to \nu\bar{\nu}$ and $W \to q\bar{q}$, $Z \to ll$ and $W \to q\bar{q}$, WW (inclusive) production.
The POWHEG generator provides a NLO (QCD) estimate of the diboson background, relying on the CT10NLO PDF set, interfaced to the PYTHIA 8 parton-shower and hadronization model.
The NLO (QCD) cross sections are provided by the POWHEG generator, and are used to normalize the diboson MC simulation. For WW and ZZ processes the POWHEG cross-section does not include the contributions of gg-initiated processes ($gg \to VV$), which is thus computed and added back to the cross-section value using the MCFM software, setting the main parameters of the calculation (renormalization and factorization scales, PDF) consistently with the POWHEG simulation. For the processes with one leptonically decaying Z boson a cut on $m_{ll} > 20$ GeV is applied for consistency with the POWHEG samples, where m_{ll} is the invariant mass of the lepton pair produced by the Z boson. As a result the diboson cross-section is increased by 2.5% for the WW sample and by 6% for the ZZ samples.

Different sources of systematic uncertainties on the modeling of the diboson background are considered: perturbative uncertainties on the fixed-order NLO cross-section computation, uncertainties on the knowledge of PDF and α_S coupling constant, parton-shower and hadronization model. While the first two are estimated using fixed order calculation with varied parameters from the MCFM software, the last one is obtained comparing the nominal POWHEG +PYTHIA 8 sample to an alternative HERWIG MC simulation.

Perturbative uncertainties: missing higher-order corrections Uncertainties due to unknown higher-order perturbative corrections to the total and differential cross-sections are usually estimated through scales variations. Applying this method on a phase space divided into exclusive jet bins, as it is the case in this analysis, however, leads to an underestimation of the uncertainty on exclusive 2-jet fixed-order cross-section, due to cancellations occurring between the large corrections to the total cross section and the perturbative corrections from logarithmic dependence on the third jet veto. To obtain a robust estimate of these uncertainties, treating properly these cancellations, the Stewart-Tackmann method [210] is applied to derive systematic errors in the 2-jet and 3-jet categories.
These uncertainties are obtained from the MCFM fixed order cross-section calculation performed with renormalization and factorization scales varied independently between 0.5 and 2.0, considering all possible combinations, defining the 2- and 3-jet categories by cutting on the transverse momentum and the pseudorapidity of the final state partons in the NLO calculation. Since MCFM does not provide the cross-section computation at higher order than NLO, the 3-jet bin corresponds to the inclusive cross-section $\sigma_{\geq 3}$ for VV+1 jet (while the 2-jet bin corresponds to the exclusive cross-section σ_2 for VV+0 jets).

According to [210] the covariance matrix for the perturbative uncertainties on the 2-jet and 3-jet categories is built as:

$$\mathrm{C}(\sigma_2,\, \sigma_{\geq 3}) = \begin{pmatrix} \Delta^2_{\geq 2} + \Delta^2_{\geq 3} & -\Delta^2_{\geq 3} \\ -\Delta^2_{\geq 3} & \Delta^2_{\geq 3} \end{pmatrix},$$

where $\Delta_{\geq 2}$ and $\Delta_{\geq 3}$ are the absolute uncertainties on the inclusive 2-jet cross-section ($\sigma_{\geq 2}$)

and the inclusive 3-jet cross-section, derived varying the μ_R and μ_F down to half and up to double the initial value. The covariance matrix C is decomposed in the sum of two matrices:

$$\mathrm{C}(\sigma_2,\, \sigma_{\geq 3}) = \begin{pmatrix} \Delta^2_{\geq 2} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \Delta^2_{\geq 3} & -\Delta^2_{\geq 3} \\ -\Delta^2_{\geq 3} & \Delta^2_{\geq 3} \end{pmatrix} = \mathrm{C}_1 + \mathrm{C}_2.$$

The first term C_1 results in an absolute systematic error $\Delta_{\geq 2}$ on the 2jet-exclusive cross section σ_2, while the second term C_2 has one singular eigenvector (hence no systematic variation) and one non-singular eigenvector which provides an anticorrelated $\Delta_{\geq 3}$ absolute systematics on σ_2 and $\sigma_{\geq 3}$. This procedure is applied for each of the diboson processes, obtaining two orthogonal relative uncertainties on σ_2 and one relative systematic uncertainty on $\sigma_{\geq 3}$.

The scale variations are derived separately in the different p_T^V regions used in the dijet-mass analysis: the variations show a non negligible dependence with respect to p_T^V, therefore they are applied as uncertainties affecting both the normalization of the MC samples and the shape of its p_T^V distribution. Figure 6.21 shows the two orthogonal shape systematics for the $W \to l\nu$ and $Z \to q\bar{q}$ process obtained with this procedure on the 2-jet bin cross-section σ_2. Table 6.3, 6.4 and 6.5 summarize the values of these uncertainties across the 5 p_T^V bins for the 6 different diboson processes considered.
These uncertainties are fully correlated across diboson processes and categories (with the proper correlation model across jet bins implemented via the Stewart-Tackmann procedure).

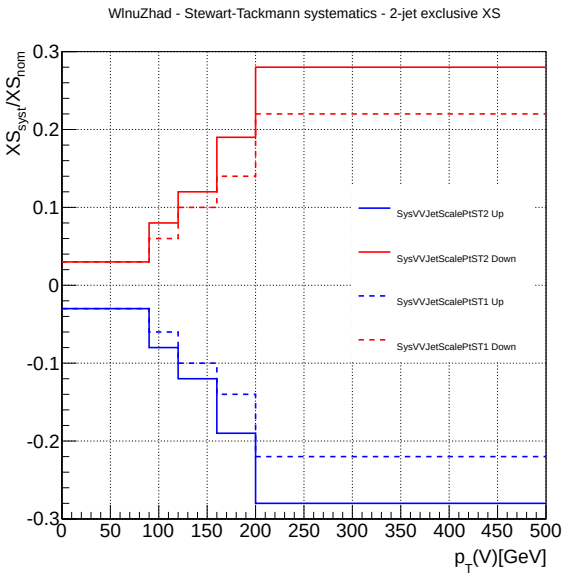


Figure 6.21: Shape perturbative systematics on σ_2 for the $W \to l\nu$ and $Z \to q\bar{q}$ process obtained applying scale variations according to the Stewart-Tackmann method. The solid line corresponds to the errors derived from the C_1 term and the dotted line from the C_2 term of the covariance matrix C.

PDF and α_S uncertainties The uncertainties on PDFs and α_S are evaluated according to the PDF4LHC prescription: the global PDF+α_S uncertainty is derived taking the envelope of the error bands provided by the CT10NLO and MSTW2008NLO PDF sets,

p_T^V bins [GeV]	0-90	90-120	120-160	160-200	> 200
$Z \to \nu\bar{\nu}$ $Z \to q\bar{q}$	3%	5%	7%	10%	14%
$Z \to ll$ $Z \to q\bar{q}$	3%	5%	7%	10%	13%
$W \to l\nu$ $Z \to q\bar{q}$	3%	8%	12%	19%	28%
$Z \to \nu\bar{\nu}$ $W \to q\bar{q}$	3%	7%	12%	18%	28%
$Z \to ll$ $W \to q\bar{q}$	3%	8%	13%	19%	29%
WW	3%	6%	9%	13 %	19 %

Table 6.3: Perturbative systematics on σ_2 from the C_1 term of the covariance matrix across the different p_T^V bins.

p_T^V bins [GeV]	0-90	90-120	120-160	160-200	> 200
$Z \to \nu\bar{\nu}$ $Z \to q\bar{q}$	2%	4%	5%	7%	11%
$Z \to ll$ $Z \to q\bar{q}$	2%	4%	6%	8%	11%
$W \to l\nu$ $Z \to q\bar{q}$	3%	6%	10%	14%	22%
$Z \to \nu\bar{\nu}$ $W \to q\bar{q}$	2%	5%	8%	12%	20%
$Z \to ll$ $W \to q\bar{q}$	2%	6%	9%	13%	21%
WW	2%	4%	6%	9%	13%

Table 6.4: Perturbative systematics on σ_2 from the C_2 term of the covariance matrix across the different p_T^V bins.

evaluated for two different α_S central values of 0.117 and 0.119. This set of uncertainties does not show a significant dependence with respect to p_T^V, therefore the errors are considered as normalization systematics on the cross-sections of the diboson backgrounds. The systematic errors are derived separately for the 2-jet and 3-jet categories, and do not depend on the decay of the vector boson but only on the V boson pair produced. The PDF+α_S uncertainties range from 2% for WW 3-jet events to 4% for WZ 2-jet events. These uncertainties are fully correlated across diboson processes and analysis categories.

Parton-shower and hadronization model uncertainties The shape of the reconstructed $Z \to b\bar{b}$ lineshape in VZ production is affected by the parton-shower and hadronisation model. Two different shower models are considered to study this effect: the nominal Powheg +Pythia 8 prediction and the LO Herwig MC simulation with LL parton-shower. Since Herwig does not provide a separate estimate of the different decay modes for the boson pair, the comparison is performed inclusively for WW, WZ and ZZ events. The effect of higher-order corrections is already included via scale variations, therefore the normalization difference between the LO Herwig calculation and the NLO Powheg prediction is neglected normalizing the samples to the same cross section. A systematic uncertainty is derived from the comparison of the dijet-mass spectrum applying the 0-lepton selection without any b-tagging requirement, fitting the MC-to-MC ratio with a parametric function to reduce the impact of statistical fluctuation. The variation has a negative impact of 10%(20%) for $m_{jj} = 50$ GeV increasing to a positive impact of 20%(40%) above 200 GeV for ZZ(WZ) events. The resulting uncertainties are applied for all lepton channels,

p_T^V bins [GeV]	0-90	90-120	120-160	160-200	> 200
$Z \to \nu\bar{\nu}$ $Z \to q\bar{q}$	-11%	-12%	-13%	-15%	-16%
$Z \to ll$ $Z \to q\bar{q}$	-10%	-12%	-14%	-15%	-16%
$W \to l\nu$ $Z \to q\bar{q}$	-12%	-13%	-15%	-16%	-17%
$Z \to \nu\bar{\nu}$ $W \to q\bar{q}$	-11%	-12%	-14%	-15%	-17%
$Z \to ll$ $W \to q\bar{q}$	-12%	-13%	-14%	-16%	-18%
WW	-11%	-12%	-14%	-15%	-17%

Table 6.5: Perturbative systematics on $\sigma_{\geq 3}$ from the C_2 term of the covariance matrix across the different p_T^V bins.

uncorrelated for WW, WZ and ZZ events.

Single-top systematic uncertainties

The single-top MC prediction is normalized to the inclusive cross sections quoted in [197, 198, 199], with uncertainties of 4%, 4% and 7% respectively for the s-channel, t-channel and Wt-channel production from [197].
To estimate systematic effects on the modeling of the MC simulation the nominal prediction is compared to several alternative MC generators. The s-channel POWHEG +PYTHIA 6 nominal sample is compared to alternative ACERMC and MC@NLO predictions; for the t-channel differences between the nominal ACERMC +PYTHIA 6 and aMC@NLO +HERWIG ++ are considered; finally the Wt-channel POWHEG +PYTHIA 6 prediction is compared to ACERMC, POWHEG +HERWIG and MC@NLO simulations. Furthermore, different approaches to the interference between $t\bar{t}$ and Wt-channel are compared, considering both the Diagram Removal (DR) adopted for the nominal MC sample, and the alternative Diagram Subtraction (DS) scheme [214]. For all three channels the ACERMC samples are used to estimate the effect of varied ISR and FSR. The comparisons are performed applying the 1-lepton event selections, where the single-top background is largest.

Acceptance uncertainties for each production channel are taken as the largest deviation among the different MC-to-MC comparisons, separately for low- and high-p_T^V, 2-jet and 3-jet events. These uncertainties range from the largest 52% for 2-jet low-p_T^V t-channel events, to the lowest 5% for the Wt-channel (which increases to 15% in the 3-jet high-p_T^V region), and values around 20% for the s-channel.

Variations in the shape of the BDT input variables are considered for all the above comparisons and the different single-top channels, leading to include three shape uncertainties for the Wt production mode: the m_{bb} spectrum for 2-jet high-p_T^V events is varied from a positive 20% at $m_{bb} = 50$ GeV to a negative 40% at 200 GeV; a second uncertainty is quoted on the same variable in the 3-jet category, with corresponding range between positive 25% and negative 20%; finally, a third uncertainty affects the p_T spectrum of the leading jet for 2-jet low-p_T^V events.

Signal	
Cross section (scale)	1% ($q\bar{q}$), 50% (gg)
Cross section (PDF)	2.4% ($q\bar{q}$), 17% (gg)
Branching ratio	3.3 %
Acceptance (scale)	1.5%–3.3%
3-jet acceptance (scale)	3.3%–4.2%
p_{T}^{V} shape (scale)	S
Acceptance (PDF)	2%–5%
p_{T}^{V} shape (NLO EW correction)	S
Acceptance (parton shower)	8%–13%
Z+jets	
Zl normalisation, 3/2-jet ratio	5%
Zcl 3/2-jet ratio	26%
Z+hf 3/2-jet ratio	20%
Z+hf/Zbb ratio	12%
$\Delta\phi(\mathrm{jet}_1,\mathrm{jet}_2)$, p_{T}^{V}, m_{bb}	S
W+jets	
Wl normalisation, 3/2-jet ratio	10%
Wcl, W+hf 3/2-jet ratio	10%
Wbl/Wbb ratio	35%
Wbc/Wbb, Wcc/Wbb ratio	12%
$\Delta\phi(\mathrm{jet}_1,\mathrm{jet}_2)$, p_{T}^{V}, m_{bb}	S
$t\bar{t}$	
3/2-jet ratio	20%
High/low-p_{T}^{V} ratio	7.5%
Top-quark p_{T}, m_{bb}, $E_{\mathrm{T}}^{\mathrm{miss}}$	S
Single top	
Cross section	4% (s-,t-channel), 7% (Wt)
Acceptance (generator)	3%–52%
m_{bb}, $p_{\mathrm{T}}^{b_1}$	S
Diboson	
Cross section and acceptance (scale)	3%–29%
Cross section and acceptance (PDF)	2%–4%
m_{bb}	S
Multijet	
0-, 2-lepton channels normalisation	100%
1-lepton channel normalisation	2%–60%
Template variations, reweighting	S

Figure 6.22: Summary of the systematic uncertainties on the signal and background modelling. An 'S' symbol is used when only a shape uncertainty is assessed.

6.7 Analysis of the $\sqrt{s} = 7$ TeV dataset

A standalone analysis of the 7 TeV dataset has been published by the ATLAS collaboration [174] in November 2012, and it is included with very small changes in the final results presented in this chapter, combined with the MVA of the 8 TeV dataset. Since the analysis of the 7 TeV data is mostly unchanged from the previous publication, and its impact on the combined sensitivity is relatively small, the analysis will not be covered in full detail: In this section I will briefly discuss the main differences between the analysis of the two datasets, and the few modifications introduced to the 7 TeV analysis with respect to [174].

6.7.1 Main features of the 7 TeV analysis

This analysis is performed following a dijet-mass approach similar to the one adopted as alternative cross-check for the analysis of 8 TeV data, no MVA has been considered. The main differences between the two dijet-mass analysis arise from changes in the object and event reconstruction (which are outlined in the dedicated Section 5.3) and some changes in the event selection and analysis strategy motivated by the different sensitivity of the 7 and 8 TeV datasets.

The selection cuts mirror the ones introduced in Section 6.3 for the 8 TeV dijet-mass analysis, with the following exceptions:

- the analysis is divided into a different number of p_T^V regions, with slightly different boundaries: in the 1- and 2-lepton channels five p_T^V bins are included [0,50,100,150,200,>200] GeV, while in the 0-lepton channel three p_T^V bins are included [120,160,200,>200] GeV.
- only one b-tagging operating point is available, thus a single 2-tag category is defined (rather than three separate LL, MM and TT sub-categories)
- in the 0-lepton channel events with 100 GeV $< p_T^V <$ 120 GeV are not included, and no cut on $\sum p_T^{jet}$ is applied
- in the 1-lepton muon channel events from the E_T^{miss} trigger are retained only in the 2-tag 2-jet category, for $p_T^W > 160$ GeV, and they define a separate category. Furthermore in both the electron and muon channels the $m_T^W > 40$ GeV cut is applied for $p_T^W < 160$ GeV, and the H_T cut is replaced by an $E_T^{miss} > 25$ GeV selection for events with $p_T^W < 200$ GeV.
- in the 2-lepton channel no kinematic fit is applied, and an additional control region is included in the Likelihood fit to constrain the $t\bar{t}$ background from data. This 'top-$e\mu$' CR is defined selecting different lepton flavor events with 2 b-tagged jets, with no veto on the number of selected jets.

The background modeling has some differences from the outline of Section 6.5. The data-driven MJ background in the 1-lepton channel is derived by fitting m_T^W and E_T^{miss} templates obtained from the inversion of lepton track-isolation criteria, in the electron and muon channel respectively. The MC simulations used for the estimate of the EW backgrounds are consistent with Section 6.5 with the following exceptions: a mixture of POWHEG +PYTHIA and ALPGEN +HERWIG is used for the W + jets background (respectively for the modeling of W produced in association with heavy flavors or with light- and c-jets); ALPGEN +HERWIG is also used for the Z+light jets backgrounds; top-quark pair and single-top

productions are obtained from MC@NLO +Herwig and finally Herwig is used for the semileptonic diboson simulation. The VH signal modeling is consistent with Section 6.5 and 6.5.3, but no separate $gg \to ZH$ sample is available, therefore the contribution of gg-induced signals is included by rescaling the inclusive signal cross section to account for it.
A reweighting correction is applied to the (W/Z) + jets MC simulation as function of $\Delta\phi(j_1, j_2)$, derived from 0-tag events in the 1- and 2-lepton channels, inclusively in p_T^V, to account for the same mismodeling observed in the 8 TeV dataset, while no correction to the $t\bar{t}$ simulation is considered.
Figure 6.23 shows a few selected distributions of the discriminating variable, the dijet-mass m_{bb}, in the signal regions of this analysis for the three lepton channels.

The experimental systematic uncertainties are mostly different with respect to the ones described in Section 6.6: either different datasets (7 TeV data rather than 8 TeV data) are used to extract the uncertainties, or the uncertainties pertain to different identification and/or reconstruction algorithms, or finally they are strictly related to the data-taking period (as the uncertainty on the integrated luminosity, quoted at 1.8% for the 7 TeV dataset).
Systematic uncertainties on the MC modeling of signal and backgrounds are derived with strategies similar to the ones adopted for the 8 TeV analysis, comparing the baseline MC simulation to alternative generators and/or fixed order calculations. However the baseline MC predictions themselves are not always consistent among the analyses of the two datasets, therefore it is not trivial to identify the same sources of modeling uncertainties between the two analysis.
The dominant sources of systematics for the 7 TeV analysis are quite similar to the ones affecting the analysis of 8 TeV data: the leading instrumental uncertainties are related to the b-tagging efficiency and the jet energy scale, while the most important modeling uncertainties come from the $t\bar{t}$, Z+jets and W+jets background estimate.
The statistical analysis follows a binned Likelihood fit approach, using the m_{bb} dijet-mass distribution as discriminating variable for the analysis signal regions. The rebinning transformation described in Section 6.4 is not applied. The signal regions include all 2-tag categories, consisting of 13 p_T^V bins across the three lepton channels, separately for 2-jet and 3-jet events. The control regions are included in the Likelihood fit exploiting only the event yields, and consisting of the top-$e\mu$ CRs for 2-tag 2-lepton events divided into five p_T^V bins, and all 1-tag categories with 13 p_T^V bins across the three lepton channels, separately for 2-jet and 3-jet events.
The floating normalizations introduced in the Likelihood fit are consistent with the 8 TeV analysis, with the difference that a single normalization parameter is included for the $t\bar{t}$ background, fully correlated across lepton channels. Furthermore the MJ background normalization is also free to float in all regions of the 1-lepton channel, except in the 2-tag 3-jet regions where the statistical power of the data is not sufficient to provide a reliable constraint. In these regions the MJ normalization is constrained by a 30% uncertainty.

Updates to the analysis of the 7 TeV dataset Two main changes have been introduced to this analysis compared to the results published in [174]:

1. VH signal normalization: the contribution of gg-induced and top-loop-induced ZH signal processes was not included in [174]. A dedicated MC sample for this signal is not available for $\sqrt{s} = 7$ TeV, but its contribution is taken into account by modifying the inclusive cross section used to normalize the signal prediction to include it, leading

to an increase in the total cross section of 6% for $m_H = 125$ GeV.

2. 1-lepton low-p_T^V electron channel: in the MVA of the 8 TeV dataset, events from the 1-lepton electron channel in the $p_T^V < 120$ GeV region are discarded, since their contribution to the combined sensitivity is quite low and the estimate of the MJ background very challenging. To be consistent in the combination of the 7 and 8 TeV results, this region is removed from the 7 TeV analysis as well.

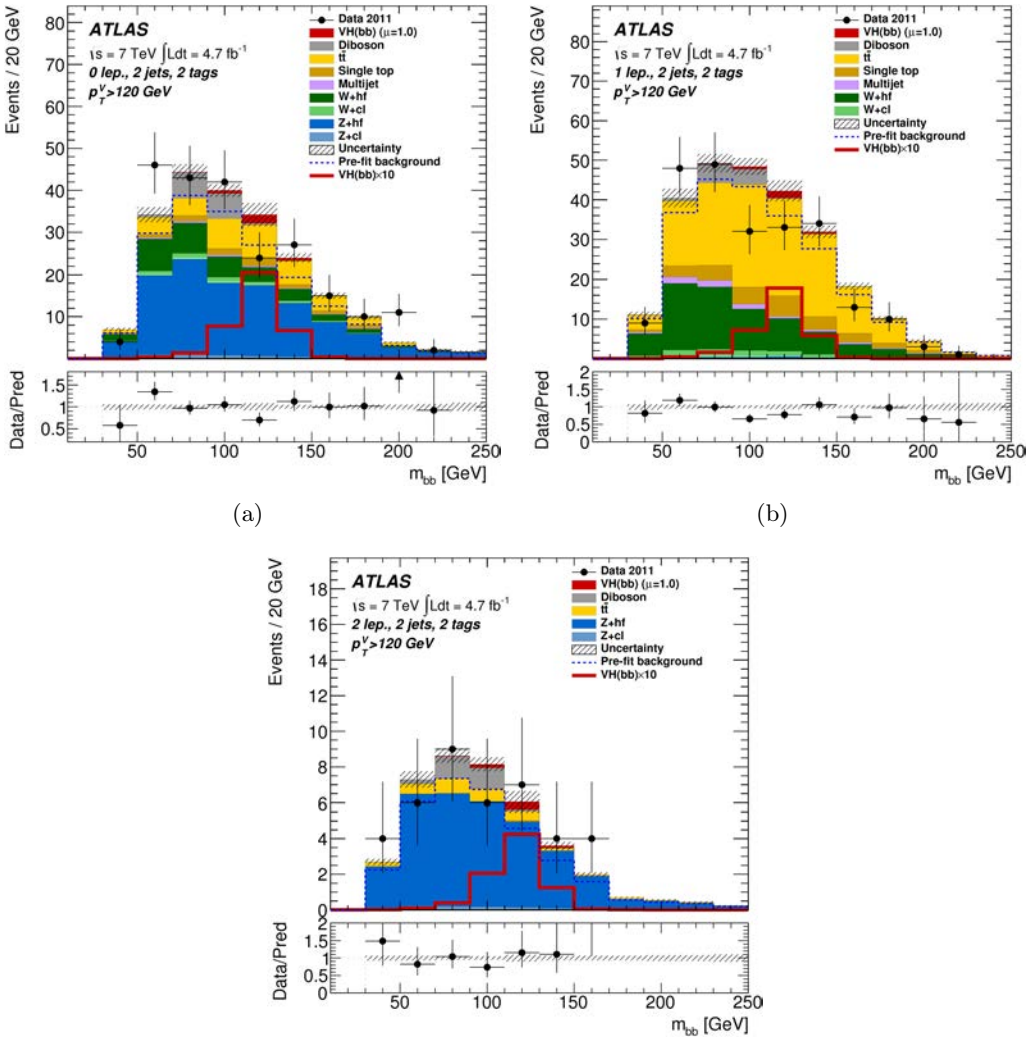


Figure 6.23: The dijet-mass distributions observed in the 7 TeV data (points with error bars) and expected (histograms) for the 2-jet 2-tag signal regions with the $p_T^V > 120$ GeV intervals combined: (a) 0-lepton channel, (b) 1-lepton channel, and (c) 2-lepton channel. The background contributions after the global fit of the dijet-mass analysis are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds, as expected from the SM (indicated as $\mu = 1.0$), and, unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total background as expected from the pre-fit MC simulation. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

6.8 Statistical Analysis: the Likelihood fit

In this section I describe the statistical analysis performed to extract the final combined results of the $VH(b\bar{b})$ analysis of Run-1 LHC data. The core aspects of the Likelihood fit adopted in this search are introduced with more details in Chapter 4, therefore in this section the focus is on the aspect of the statistical procedure which are specific to this analysis.
First I describe the details of the Likelihood function used in the global fit, outlining the features of the different categories included, along with the fitted discriminating variables. I discuss the treatment of systematic uncertainties in terms of nuisance parameters in the fit, and then present the impact of the fit to data on these uncertainties, giving full detail on the many studies performed to gain confidence with the solidity and the performances of the fit.
The main purpose of this section is twofold: first it describes in details the fit model, with the main technical details of its implementation. Furthermore it motivates the specific choices which defined the main building blocks of this fit model, from the different regions included to the treatment of their correlations. Full understanding of these details is crucial to present the result of this search in Section 6.9.

6.8.1 The Likelihood function: categories and variables

The test statistics adopted to extract the final results is built from a Profile Likelihood Ratio, where the binned Likelihood function follows the description given in Chapter 4, and it is built as the product of Poisson probability terms over the bins of the input distributions involving the numbers of data events and the expected signal and background yields. The signal strength is the parameter μ that multiplies the SM Higgs boson production cross section times branching ratio into $b\bar{b}$. The impact of systematic uncertainties on the signal and background expectations is described by nuisance parameters (NPs), θ.
The results of the statistical analysis are presented in terms of: the 95% confidence level (CL) upper limit on the signal strength; the probability p_0 of the background-only hypothesis; the best-fit signal strength $\hat{\mu}$ with its associated uncertainty σ_μ (defined in Chapter 4). Expected results are obtained in the same way as the observed ones by replacing the data in each input bin by the corresponding Asimov data prediction.
The regions entering the Likelihood fit are summarized in Table 6.24 and correspond to the event selection outlined in detail in Section 6.3.

Signal regions of the dijet-mass analysis In the dijet-mass approach the signal regions include all 2-tag categories (separately for LL, MM and TT purities), across 5 p_T^V intervals in the 1- and 2-lepton channels, 3 p_T^V intervals for 0-lepton 3-jet events and 4 p_T^V intervals for 0-lepton 2-jet events, separately for 2-jet and 3-jet events, for a total number of 81 categories. In all these regions the m_{bb} distribution after rebinning transformation is used as input for the global fit.

Signal regions of the MVA In the MVA approach 24 2-tag signal regions are considered, including two $p_T^V \lesseqgtr 120$ GeV intervals for the 1- and 2-lepton channels, and a single $p_T^V > 120$ GeV interval for the 0-lepton channel, separately for 2-jet and 3-jet events. In the 1-lepton channel the b-tagging categories are separated in LL, MM and TT according to the jets' purity, while in the 0- and 2-lepton channel the MM and TT category are merged in a single (MM+TT) region, while keeping the LL one separated. In all these regions the BDT_{VH} is the fitted discriminating variable, after the rebinning transformation.

Furthermore in the 0-lepton channel three additional regions are included (raising the number of signal regions to 27), corresponding to the 100 GeV $< p_T^V <$ 120 GeV interval, for 2-jet events only, separately for LL, MM and TT purities. For these additional categories the m_{bb} distribution is used in the Likelihood fit.

Control regions Control regions are included in the fit to provide constraints on the modeling of the backgrounds. In both the dijet-mass and the MVA approach the main control regions include the 11 1-tag categories, across two $p_T^V \lesseqgtr 120$ GeV intervals for all lepton channels, with lower p_T^V boundary at 100 GeV in the 0-lepton channel, separately for 2-jet and 3-jet events. In all these regions the distribution of the MV1c output of the b-tagging algorithm is used in the fit, to provide further constraints on the flavor composition of the backgrounds.
Note that the MVA p_T^V categorization is used in the 1-tag category for the dijet-analysis as well. The finer p_T^V binning used in the 2-tag regions would require a very accurate study of the modeling of the MV1c distribution as a function of the transverse momentum, in order to properly assess systematic uncertainties and control the relative normalization across p_T^V bins. The use of two p_T^V bins still provides information on the p_T^V spectrum for the different jet flavor components.
When the Likelihood fit is performed in the 2-lepton channel alone, without the contribution of the other lepton channels, an additional control region is included to constrain the $t\bar{t}$ background from data. This 'top-$e\mu$' CR is defined from different lepton flavor and opposite sign events, with no veto on additional selected jets. The top-$e\mu$ CR is not included in the global fit since sufficient information on the $t\bar{t}$ modeling can be extracted from the 1-lepton regions in the dijet-mass analysis, and the 2-lepton low-BDT_{VH} categories in the MVA.

$VH(b\bar{b})$ categories: impact on the Likelihood fit The full set of regions included in the Likelihood definition covers an extremely large phase space with many different SM backgrounds having larger yields than the expected VH signal. Two general considerations are therefore necessary: first, what is the power of the different regions included in the fit, and could some of them be eliminated? Second, how does the Likelihood function reflect the best knowledge of the background and signal predictions across these several regions?
To answer the first question the main sets of regions can be considered one by one, starting from the consideration that, from the topology and kinematic of the $VH(b\bar{b})$ signal, the most natural signal regions for the three lepton channels are the 2-jet 2-tag categories with boosted Higgs candidate (and thus high transverse momentum p_T^V of the recoiling vector boson):

- **1-tag regions**: these categories have very low sensitivity by themselves, since the signal versus background ratio is quite small. However they are crucial to provide information on the modeling of the relative flavor fraction in $V+$jets events ($V+l$, cl, cc, bl, bc, bb), whose theoretical modeling in the MC simulation is extremely challenging. Removing these regions from the combined Likelihood fit degrades the sensitivity by $O(5\%)$ due to degeneracies in the shape of different $V+$jets flavor production in the MC simulation. For these reasons the fitted variable for the 1-tag category is indeed the output distribution of the b-tagging algorithm MV1c, which provides accurate control on the jet flavor composition of the events.

- **3-jet regions**: in the 1-lepton and (to a smaller extent) in the 0-lepton channels the 2-tag 3-jet regions are enriched of $t\bar{t}$ events, providing useful constraints on the

modeling of this major background. In the 2-lepton channel, 3-jet events show a very high purity in Z+hf background, whose modeling is thus controlled by fitting these categories. Removing the 3-jet regions from the combined fit degrades the analysis sensitivity by $O(8\%)$.

- **low-p_T^V regions**: as described in Section 6.5 the transverse momentum spectrum of top-quark pair and vector boson production show mismodeling when comparing the MC simulation to the data distributions. The inclusion of low-p_T^V events not only brings additional sensitivity to the analysis (of the order of 10%) but allows the fit to fix the background normalization from the low-p_T^V categories (with higher statistics) and propagate this information to the boosted regions. The details of this correlation have to be carefully implemented in the fit model to avoid bias in the background modeling at high-p_T^V: this is done through the modeling systematic uncertainties outlined in Section 6.5, which control the relative normalization of the backgrounds across the p_T^V spectrum with normalization and shape parameters.

<table>
<tr><td></td><td colspan="3">Dijet-mass analysis</td><td colspan="3">MVA</td></tr>
<tr><td>Channel</td><td>0-lepton</td><td>1-lepton</td><td>2-lepton</td><td>0-lepton</td><td>1-lepton</td><td>2-lepton</td></tr>
<tr><td>1-tag</td><td colspan="3">$MV1c$</td><td colspan="3">$MV1c$</td></tr>
<tr><td rowspan="3">LL
MM 2-tag
TT</td><td colspan="3">m_{bb}</td><td>BDT$^{(*)}$</td><td colspan="2">BDT</td></tr>
<tr><td colspan="3">m_{bb}</td><td rowspan="2">BDT$^{(*)}$</td><td>BDT</td><td rowspan="2">BDT</td></tr>
<tr><td colspan="3">m_{bb}</td><td>BDT</td></tr>
</table>

Figure 6.24: The distributions used in each region by the likelihood fit in the dijet-mass analysis and in the MVA applied to the 8 TeV data. Here, 'BDT' stands for 'BDT_{VH} discriminant'. For each entry listed, there are additional divisions into p_T^V intervals; five in the dijet-mass analysis and two in the MVA. These distributions are input to the fit for the 2-jet and 3-jet categories separately, except in the low-p_T^V interval ($100 - 120$ GeV) of the 0-lepton channel where only the 2-jet category is used. In the 0- and 2-lepton channels, the MM and TT 2-tag categories are combined in the MVA. $(*)$ In the low-p_T^V interval of the 0-lepton channel, the MVA uses the m_{bb} distributions in the LL, MM and TT 2-tag categories as well as the MV1c distribution in the 1-tag category.

6.8.2 Nuisance parameters in the Likelihood fit

The implementation in the fit of the nuisance parameters controlling the signal and background modeling, and related to systematic uncertainties, reflects the confidence with which these parameters and their behavior are known.
The prior penalty terms constraining the parametrized NPs reflect the degree of knowledge as to where their true value lies, while the correlation of the NPs across analysis categories expresses the statement that their effects have the same source in different regions, and are therefore correlated.
Statistical fluctuations affecting the estimate of systematic uncertainties are dealt with applying the smoothing techniques described in Section 4.2.1, and the number of systematic-variation template pairs ($+1\sigma$ and -1σ) is reduced by a factor of two, from the initial number of twenty-thousands, thanks to the pruning procedure also described in Section 4.2.1, reducing significantly the fit complexity.

Background normalization: floating scale factors The collected data have sufficient statistical power to control the dominant backgrounds, $t\bar{t}$, W+hf, W+cl, Z+hf and Z+cl, whose normalizations are thus left free to float in the fit in the 2-jet categories.
As explained in Section 6.5 the $t\bar{t}$ normalization is uncorrelated among lepton channels. The probed $t\bar{t}$ phase space is largely different for different lepton selections: 2-lepton $t\bar{t}$ events come mostly from fully leptonic decays, while in the 0- and 1-lepton channels there is a sizable contribution from semileptonic decays and fully leptonic decays with mis-identification or mis-reconstruction of one or more leptons. Furthermore, the p_T^V spectrum considered is different in the 0-lepton channel and the 1- and 2-lepton channels. The scale factors extracted from the MVA fit to the 8 TeV data for these backgrounds are summarized in Table 6.25.

Understanding the $VH(b\bar{b})$ Likelihood fit model The correlation scheme for nuisance parameters related to the modeling of signal and background has been introduced case by case in Section 6.5, and it is the result of careful studies based on the core considerations detailed in Chapter 4
The stability and the quality of the fit model are studied performing the different tests outlined in Section 4.2.2, considering the impact of the Likelihood fit on the NPs in terms of central values, uncertainties and correlation arising during the fit itself. Furthermore the ranking of the NPs according to their impact on the measured signal strength $\hat{\mu}$ is considered as a figure of merit to determine the most important NPs in the analysis. The computation of the raking can become a lengthy procedure since the number of NPs in the fit model is very large: systematic uncertainties which are shown to have negligible effect on the expected fitted uncertainty on $\hat{\mu}$ are not included in the study (NPs associated to the muon momentum scale and resolution, to the electron energy resolution, one NP related to a single JES component, NPs associated with the quark-gluon composition of the backgrounds).
The stability of the results is also tested by performing fits for each lepton channel independently, which can also help to identify from which region each constraint originates.

Figure 6.26 shows the ranking of the systematic uncertainties obtained from the MVA fit to 8 TeV data, with NPs ordered by decreasing observed impact on $\hat{\mu}$. The first five systematic uncertainties in the ranking are (from the largest impact in descending order):

- on the dijet-mass shape in W+bb and W+cc backgrounds, for $p_T^W > 120$ GeV
- on the W+bl/W+bb relative acceptance variation, for $p_T^W > 120$ GeV
- on the W+bb background normalization (free to float in the 2-jet categories)
- on the p_T^W shape of the W+hf background in the 3-jet category
- on the signal acceptance from parton-shower modeling effects

It is clear how the modeling of the W+hf in the most sensitive signal region plays a crucial role in the determination of the final results: this is not unexpected since the W + jets background plays a dominant role in the analysis, but the MC simulation of heavy flavor production is still extremely challenging and not fully under control from a theoretical point of view, and there isn't any pure control region where its modeling can be studied accurately.
The NPs which appear lower in the ranking are not unexpected as well: from Figure 6.26 NPs related to the modeling and normalization of the main backgrounds ($t\bar{t}$ and Z+hf) are

observed, along with JES-related NPs (which can impact the relative normalization of the 2- and 3-jet bins) and b-tagging-related NPs (which can affect the relative normalization of the different b-tagging categories).

Pulls and the constraints are shown as the black points plotted in Figure 6.26 for the top ranked parametrized NPs (whose expectation is at 0.0 with a ± 1.0 pre-fit uncertainty), and as the red points for the top ranked unconstrained normalization parameters (whose expectation is at 1.0, without any prior uncertainty). Most of the top ranked NPs have pulls which are compatible with the expected value within the fitted uncertainties. The pull on the NP related to the p_T^W shape uncertainty for W+hf 3-jet events (4th in the ranking) could be explained with a residual mismodeling of the MC simulation of the W+hf background: a mismodeling of this variable is found from the data-to-MC comparison in 0-tag W + jets events (dominated by W+light), but in absence of a pure data control sample the same study cannot be performed for the W+hf background. It is not unlikely that, if such mismodeling is present in the W+hf MC simulation, the p_T^W shape uncertainty, derived from MC-to-MC comparisons in the 2-tag regions, may be pulled to adjust the MC prediction to the observed data.
It can be noted how the normalization parameters free to float in the fit are well constrained for backgrounds which benefit from pure control regions ($t\bar{t}$ and Z + jets), while the W+hf floating normalization shows a significant post-fit uncertainty (the numerical value of the floating normalization with associated uncertainties are reported in Table 6.25 as well). The remaining top ranked parameters are well behaved in the combined Likelihood fit, and the solidity of the fit is reflected in the excellent data-to-MC agreement show in the post-fit plots for the fitted discriminating variables (add ref.).

Process	Scale factor
$t\bar{t}$ 0-lepton	1.36 ± 0.14
$t\bar{t}$ 1-lepton	1.12 ± 0.09
$t\bar{t}$ 2-lepton	0.99 ± 0.04
Wbb	0.83 ± 0.15
Wcl	1.14 ± 0.10
Zbb	1.09 ± 0.05
Zcl	0.88 ± 0.12

Figure 6.25: Factors applied to the nominal normalizations of the $t\bar{t}$, W+hf, W+cl, Z+hf, and Z+cl backgrounds, as obtained from the global MVA fit to the 8 TeV data. The $t\bar{t}$ background is normalized in the 2-jet category independently in each of the lepton channels. The errors include the statistical and systematic uncertainties.

Likelihood fit for alternative m_H hypothesis The analysis has been optimized for the search of a Higgs boson with $m_H = 125$ GeV, but the Likelihood fit is performed for alternative masses as well, to scan the range between 100 and 150 GeV. For the dijet-mass analysis no change is introduced when testing different Higgs masses, except the binning of the transformed m_{bb} distribution. The MVA however is trained with signal samples corresponding to the alternative m_H, as explained in Section 6.4, to avoid degrading the performance for masses away from the baseline $m_H = 125$ GeV. The fit model described in this section and the results, when not specified, refer to the analysis performed for a

Higgs boson with mass of 125 GeV.

6.8.3 Combination of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV analyses

The analyses of the 7 TeV and the 8 TeV datasets are performed separately, and finally combined when building the Likelihood function and performing the Likelihood fit. In this section I describe how the combined Likelihood function is defined in order to include both measurements (this combination is tested and performed only for the MVA of the 8 TeV data).
A combined Likelihood fit including both analyses is simply defined by the product of the Likelihood function defined for the analysis of the 8 TeV dataset with the additional Poisson probability terms corresponding to the categories of the 7 TeV search. The signal strengths which control the normalization of the signal prediction for all Likelihood terms are correlated among themselves.
The features of the combined measurement are thus defined by the level of correlation between the NPs which control the Likelihood terms related to the 7 TeV or the 8 TeV searches: to assess the possible correlation for each NP, the source of the corresponding systematic uncertainty and the methodology to estimate it are considered.
The object reconstruction is substantially different between the two searches: jet calibration, lepton identification and b-tagging algorithms are different, as well as the calculation of the E_T^{miss} soft-term. All experimental systematic uncertainties are therefore kept uncorrelated between the 7 and 8 TeV measurements, in the combined Likelihood.
The background and signal modeling is in many cases relying on different MC generators; furthermore the categories of the 8 TeV analysis have been extended to improve the sensitivity and include larger statistics for the MVA approach, thus it is difficult to assess accurate correlations between modeling uncertainties estimated for the two searches, which are finally kept uncorrelated in the fit.
The theoretical uncertainties on signal cross section (including missing higher order corrections and PDF variations) and $H \to b\bar{b}$ branching ratio are instead computed consistently at 7 TeV and 8 TeV, therefore they are fully correlated between the two analyses. For the 7 TeV analysis the missing higher order corrections and PDF variations for ZH signals are computed only for the $qq \to ZH$ production, and applied to the full ZH signal (rescaled to include the gg-induced contribution as described in Section 6.7), therefore the first are fully correlated to the corresponding uncertainty for both the qq- and gg-induced ZH signal of the 8 TeV analysis, while the second are correlated between qq-initiated signal only.

NPs in the combined fit Since the combined fit of 7 and 8 TeV dataset is performed keeping both measurements almost completely uncorrelated, the results are expected to be consistent with the separate Likelihood fit of the 7 TeV and 8 TeV analyses.
No correlation between NPs belonging to the two analyses arises after the combined fit; the NPs pulls are completely consistent with the behavior observed in standalone Likelihood fit of 7 and 8 TeV dataset and finally the NPs ranking of the combined analysis is identical to the one shown in Figure 6.26: this is a clear indication that the impact of the 7 TeV measurement on the final result is small (as expected, given the lower analysis sensitivity) and no effect on the NPs controlling the 8 TeV measurement is introduced when performing the combined fit. These checks are repeated both on the observed data and on the Asimov dataset built from the 7 and 8 TeV analyses.

Compatibility between the 7 TeV analysis and the expected 8 TeV result A consistency test between the 7 TeV measurement and the expected results of the 8 TeV

analysis has been considered: two minimum Likelihood fits have been performed (building a combined Likelihood including both 7 TeV and 8 TeV analyses), correlating or uncorrelating the Higgs signal strength μ between the two analyses, and the difference between the Likelihood minima has been studied.
The test has been repeated using several Asimov datasets from the 8 TeV measurement, generated with expected values of the signal strength ranging from 0.0 to 3.0, and real data for the 7 TeV measurement. The difference between Likelihood minima for these multiple fits is described by a χ^2 distribution, therefore the consistency between the two measurements can be quoted as the probability of obtaining a result (from the 8 TeV measurement contribution) which is more or less compatible than the one observed from the combined fit with the tested 8 TeV signal strength. For an expected 8 TeV signal strength of 1.0 the consistency is of the level of 5%, increasing to $O(10\%)$ for an 8 TeV signal strength of 0.5. The low consistency between the 7 TeV measured signal strength and the 8 TeV expected one is not surprising, since the 7 TeV signal strength is expected to assume negative values from [174].

6.8.4 Measurement of the diboson $VZ(b\bar{b})$ signal strength

As outlined in the Introduction 6.1 the measurement of the signal strength for semileptonic diboson production processes $VZ(b\bar{b})$ can be used as a validation of the analysis procedure, since their signature is remarkably similar to the $VH(b\bar{b})$ one, with softer p_T^{bb} spectrum and m_{bb} peaking at lower values. This cross-check is performed both for the dijet-mass and the MVA.
In the dijet-mass approach, the binning transformation of the m_{bb} distribution is re-optimized, while for the MVA the BDTs are re-trained to discriminate the diboson signal from the SM backgrounds (including the $VH(b\bar{b})$ and WW processes), as described in Section 6.4. The Likelihood fit is then performed including a signal strength parameter μ_{VZ} controlling the overall normalization of the diboson WZ and ZZ processes, treating the WW contribution as background. The VH signals are included in the fit as background contribution, with a nuisance parameter on the overall normalization constrained by an uncertainty of 50%. The fitted distributions of the BDT_{VZ} discriminant are shown in Figure 6.27 for the 2-jet 2-tag $p_T^V > 120$ GeV signal regions of the three lepton channels.

Compatibility between the diboson and the Higgs measurements Similarly to the consistency check described in Section 6.8.3, the compatibility of the diboson cross-check with the baseline Higgs result is tested by performing a Likelihood fit in which both the VH and the VZ signal strength, μ and μ_{VZ}, are unconstrained. This allows to study the correlation between the two parameters arising during the fit to data. The template distributions of the discriminating variables (m_{bb} or BDT_{VZ}) use the binning optimized for the Higgs search, and in the MVA approach the BDTs are trained to select the Higgs signals.

6.9 Results

The results of this analysis are obtained from the Likelihood fit (detailed in Section 6.8) of the final discriminating variable in the 2-tag signal regions and the MV1c distributions of the b-tagged jet in the 1-tag control regions, with nuisance parameters free to float or constrained by priors. The final discriminant used in the 2-tag regions are the rebinned m_{bb} distribution for the dijet-mass analysis, and the transformed BDT_{VH} multivariate

discriminator in the MVA approach.
Separate results for the VH search are obtained from the dijet-mass and the multivariate analysis, with the latter being used as nominal result given the higher expected sensitivity to the Higgs boson signal targeted in this study. The dijet-mass analysis is maintained as a cross-check, and its results are described in Section 6.9.2. The analysis of the 7 TeV dataset is only performed following the dijet-mass approach, and it is combined to the nominal MVA result obtained from the 8 TeV dataset, as described in Section 6.8.3. If not specified otherwise all results are presented for a Higgs boson mass of 125 GeV. In the following the fitted signal strength parameters are denoted as μ and μ_{VZ}, rather than $\hat{\mu}$ and $\mu\hat{_{VZ}}$.
In Section 6.9.3 the results of the diboson VZ Profile Likelihood Fit described in Section 6.8.4 are reported.

6.9.1 Nominal results for the $VH(b\bar{b})$ search

These results are obtained from the combined Likelihood fit of the dijet-mass analysis of 7 TeV data and the MVA of 8 TeV data. Distributions of the BDT_{VH} discriminant and of MV1c, with background normalization and nuisance parameters adjusted by the global fit to the 8 TeV data were already presented in Section 6.5. Dijet-mass distributions in the 7 TeV data analysis were shown in Section 6.7. Agreement between data and estimated background is observed within the systematic and statistical uncertainties shown by the hatched bands.
Figure 6.28 shows the 95% CL upper limits on the cross section times branching ratio for $pp \to (W/Z)(H \to b\bar{b})$ in the Higgs boson mass range 110-140 GeV obtained from the Profile Likelihood Fit. The observed limit for $m_H = 125$ GeV is 1.2 times the SM value, to be compared to an expected limit, in the absence of signal, of 0.8. For the 8 TeV (7 TeV) data only, the observed and expected limits are 1.4 (2.3) and 0.8 (3.2), respectively.
The probability p_0 of obtaining from background alone a result at least as signal-like as the observed one is 8% for a tested Higgs boson mass of 125 GeV; in the presence of a Higgs boson with that mass and the SM signal strength, the expected p_0 value is 0.5%. This corresponds to an excess observed with a significance of 1.4σ, to be compared to an expectation of 2.6σ. For the 8 TeV data alone the observed and expected levels of significance are 1.7σ and 2.5σ, respectively. For the 7 TeV data alone the expected significance is 0.7σ and there is a deficit rather than an excess in data, as can be seen in Figure 6.23. Figure 6.29 shows the p_0 values in the mass range 110-140 GeV as obtained from the 7 TeV and 8 TeV combined dataset.
The fitted μ values for $m_H = 125$ GeV from the maximum Likelihood fit are shown in Figure 6.30 for the 7 TeV, 8 TeV and combined dataset. With all lepton channels and data-taking periods combined, the fitted value of the signal strength parameter is $\mu = 0.51 \pm 0.31(\text{stat.}) \pm 0.24(\text{syst.})$. For the 8 TeV data the fitted value of the signal strength parameter is $\mu = 0.65 \pm 0.32(\text{stat.}) \pm 0.26(\text{syst.})$. For the 7 TeV data it is $\mu = -1.6 \pm 1.2(\text{stat.}) \pm 0.9(\text{syst.})$.
For a Higgs boson with mass of $m_H = 125.36$ GeV, as was measured by ATLAS [215], the signal strength parameter is $\mu = 0.52 \pm 0.32(\text{stat.}) \pm 0.24(\text{syst.})$.

Additional PLR fits are performed floating independently the signal strengths for (i) the WH and ZH production processes, or (ii) the three lepton channels (note that these are not independent fits across production modes or lepton channels, but combined fits with multiple signal strengths). The results of these fits are shown in Figure 6.31 and 6.32. The consistency of the fitted signal strengths in the WH and ZH processes, assuming the

SM hypothesis, is at the level of 20%. For the lepton channels, the consistency under the SM hypothesis between the three fitted signal strengths is at the level of 72% for the 7 TeV data, and of 8% for the 8 TeV data. These consistency checks are performed with the same methodology outlined in Section 6.8.3. The lower compatibility of the fitted signal strengths for the ZH processes and in the 0-lepton channel are associated with the data deficit observed in the most sensitive bins of the BDT_{VH} discriminant in the 0-lepton channel, shown in Figure 6.17(a).

Figure 6.33 shows the data, background and signal yields, where the bins of the final discriminant distributions from all signal regions are combined into bins of log(S/B), separately for the 7 and 8 TeV datasets, where the rightmost bins show thus the most sensitive analysis categories. To build this plot, S is the expected signal yield and B is the fitted background yield. Details of the fitted values of the signal and of the various background components are provided in Table 6.36.

6.9.2 Dijet-mass analysis as cross-check of the MVA approach

The distributions of m_{bb} in the dijet-mass analysis, with background normalisations and nuisance parameters adjusted by the global fit to the 8 TeV data were already presented in Section 6.5. Agreement between data and estimated background is observed within the uncertainties shown by the hatched bands.
In the dijet-mass analysis a μ value of $1.23 \pm 0.44(\text{stat.}) \pm 0.41(\text{syst.})$ is obtained for the 8 TeV dataset. The consistency for the results of the three lepton channels is at the level of 8%.
The compatibility of the dijet-mass analysis and the MVA results is estimated with a 'bootstrap' method [216] from the 8 TeV dataset results, since the same data sample is used for both analyses. A large number of event samples are randomly extracted from the simulated samples, with the Higgs signal strength μ set to the SM expectation at 1.0. Each of the random extractions is representative of the integrated luminosity used for the data analysis in terms of expected yields as well as of associated Poisson fluctuations. Each of the event samples is used to perform both the dijet-mass analysis and the MVA, thus allowing the two fitted μ values to be compared and their statistical correlation to be extracted (simulating the effect of repeated measurement). At the same time, the expected distributions of μ and of its uncertainty are determined for both analyses. From this procedure the results obtained for the 8 TeV data with the dijet-mass analysis and the MVA are expected to be 67% correlated, and the observed results are found to be statistically consistent at the level of 8%.
The observed significance in the dijet-mass analysis is 2.2σ, while the expected significance is 1.9σ, to be compared to 2.5σ expected for the MVA, which is thus the baseline analysis. Figure 6.34 shows the m_{bb} distribution in data after subtraction of all backgrounds except for diboson production for the 7 TeV and 8 TeV data separately, as obtained with the dijet-mass analysis. In this figure the contributions of all 2-tag signal regions from all lepton channels is combined, weighted by their respective ratios of expected Higgs boson signal to fitted background (S/B). The VZ contribution is clearly seen, located at the expected Z mass. The Higgs boson signal contribution is shown as expected for the SM cross section.

6.9.3 Diboson $VZ(b\bar{b})$ measurement

The analysis has been validated by measuring the diboson VZ signal strength parameter μ_{VZ}, as detailed in Section 6.8.4. The measured signal strength for the 8 TeV dataset

with the MVA is $\mu_{VZ} = 0.77 \pm 0.10(\text{stat.}) \pm 0.15(\text{syst.})$. This result is consistent with the observation already made from Figure 6.34. The signal strengths obtained for the combined fit of three lepton channels are consistent at the 85% level (assessed with the same methodology used for the Higgs measurement). In the dijet-mass analysis at 8 TeV the observed signal strength is $\mu_{VZ} = 0.79 \pm 0.11(\text{stat.}) \pm 0.16(\text{syst.})$. The correlation of the systematic uncertainties on mu_{VZ} and μ is 35% in the MVA and 67% in the dijet-mass approach.
The compatibility between the diboson cross-check measurement and the main Higgs results is obtained as outlined in Section 6.8.4, by performing the PLR fit with the same discriminating variables used for the baseline Higgs analysis of the 8 TeV dataset, with both μ and μ_{VZ} free to float. The fitted Higgs signal strength μ is unchanged from the nominal result presented in Section 6.9.1, and the statistical correlation between the two signal strength parameters is quoted at -3% in the MVA and 9% in the dijet-mass analysis. The main cause of this low correlation can be found in the different shape of the p_T^V distributions for VZ and VH events (softer in the case of diboson processes): since this variable is used both in the MVA training and in the definition of the categories for both dijet-mass and MVA approaches, the final results are remarkably sensitive to its features. The yields quoted in Tables 6.37, 6.38 and 6.39 indeed show that the ratio of the diboson to the Higgs contribution is smaller in the higher p_T^V interval than in the lower one. The additional input variables for the BDT provide further separation in the MVA, leading to a very small diboson contribution in the most significant bins of the BDT_{VH} discriminant, as it appears from Table 6.36.
The diboson signal strength obtained from fit of the 7 TeV dataset corresponds to $\mu_{VZ} = 0.50 \pm 0.30(\text{stat.}) \pm 0.38(\text{syst.})$. The signal strength measured from the combined 7 and 8 TeV datasets is $\mu_{VZ} = 0.74 \pm 0.09(\text{stat.}) \pm 0.14(\text{syst.})$. The diboson VZ signal is thus observed with a significance of 4.9σ, to be compared to the expected significance of 6.3σ. The fitted μ_{VZ} values are summarized in Figure 6.35 for the 7 TeV, 8 TeV and combined dataset, and for the combined fit of the 0-, 1- and 2-lepton channels (in all cases the MVA results are considered for the analysis of 8 TeV data).

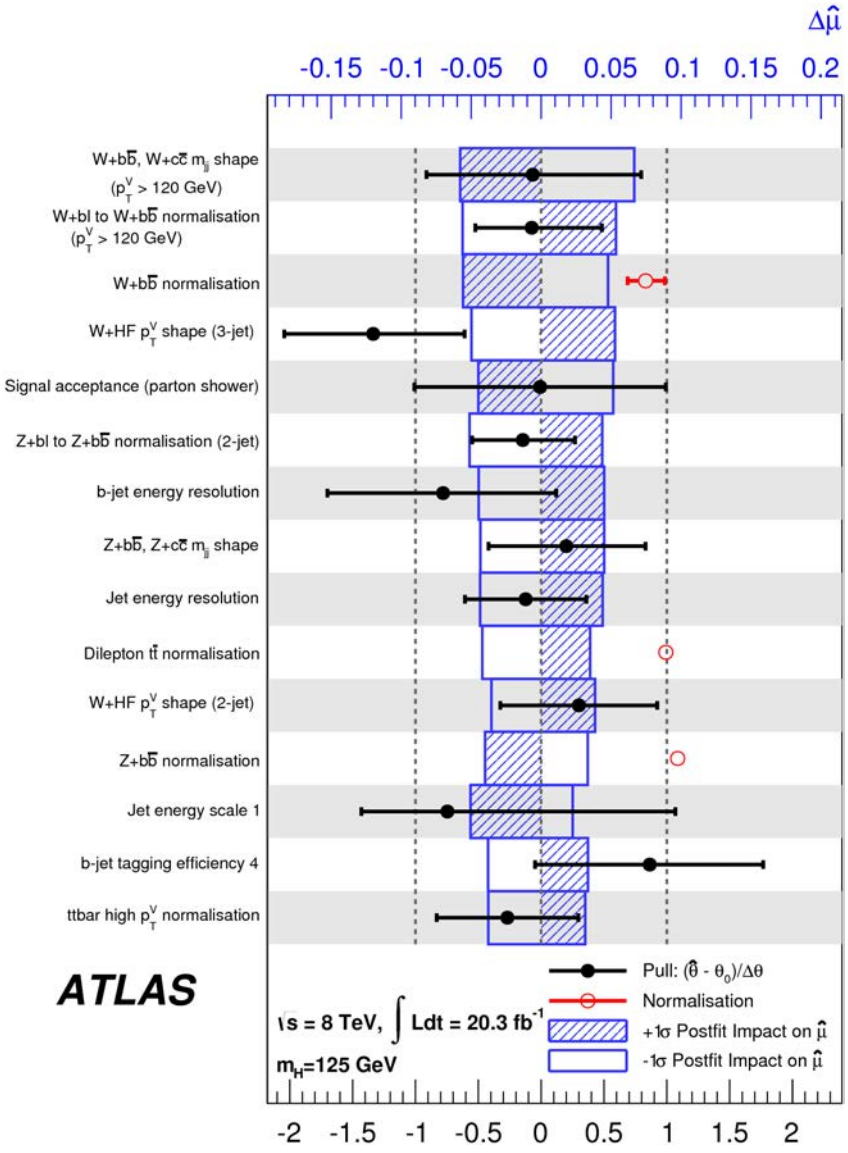
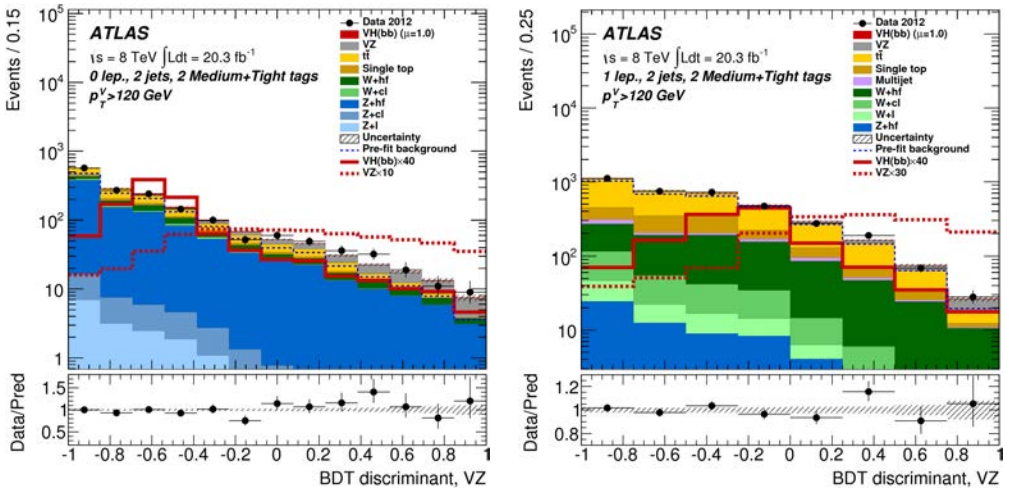
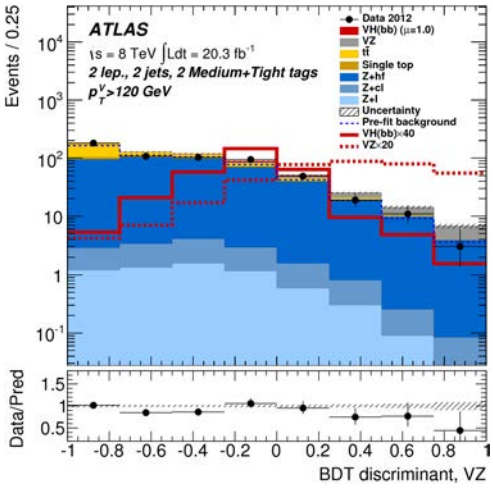


Figure 6.26: Impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$ for the MVA applied to the 8 TeV data. The systematic uncertainties are listed in decreasing order of their impact on $\hat{\mu}$ on the y-axis. The boxes show the variations of $\hat{\mu}$, referring to the top x-axis, when fixing the corresponding individual nuisance parameter θ to its post-fit value $\hat{\theta}$ modified upwards or downwards by its post-fit uncertainty, and repeating the fit as explained in the text. The hatched and open areas correspond to the upwards and downwards variations, respectively. The filled circles, referring to the bottom x-axis, show the deviations of the fitted nuisance parameters $\hat{\theta}$ from their nominal values θ_0, expressed in terms of standard deviations with respect to their nominal uncertainties $\Delta\theta$. The associated error bars show the post-fit uncertainties of the nuisance parameters, relative to their nominal uncertainties. The open circles with their error bars, also referring to the bottom x-axis, show the fitted values and uncertainties of the normalisation parameters that are freely floating in the fit. The normalisation parameters have a pre-fit value of one. As explained in Section 6.6, the jet energy scale and b-tagging uncertainties are decomposed into uncorrelated components; the labels 1 and 4 refer to such components.



(a) (b)



(c)

Figure 6.27: Distribution of the BDT_{VZ} discriminant used in the VZ fit observed in data (points with error bars) and expected (histograms) for the 2-jet 2-tag signal regions with $p_T^V > 120$ GeV in the Medium (MM) and Tight (TT) categories combined of (a) the 0-lepton channel, (b) the 1-lepton channel, (c) the 2-lepton channel. The VZ signal and background contributions are shown as filled histograms after the global VZ fit of the MVA, except for the Higgs boson contribution ($m_H = 125$ GeV). The latter is shown as expected from the SM (indicated as μ=1.0) both as a filled histogram on top of the sum of the fitted VZ signal and other backgrounds, and unstacked as an unfilled histogram scaled by the factor indicated in the legend. The expected VZ contribution is shown in a similar way as an unfilled dotted histogram. The dashed histogram shows the sum of the total VZ signal and background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the fitted VZ signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted VZ signal and background is shown in the lower panel.

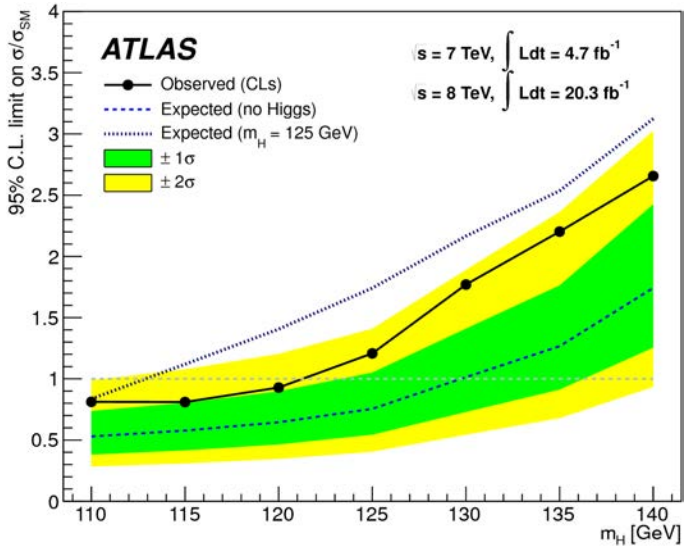


Figure 6.28: Observed (solid) and expected 95% CL cross-section upper limits, normalised to the SM Higgs boson production cross section, as a function of m_H for all channels and data-taking periods combined, as obtained using the dijet-mass analysis for the 7 TeV dataset and BDTs trained at each individual mass for the 8 TeV dataset. The expected upper limit is given for the background-only hypothesis (dashed) and with the injection of a SM Higgs boson signal at a mass of 125 GeV (dotted). The dark and light shaded bands represent the 1σ and 2σ ranges of the expectation in the absence of a signal. For all curves shown, the results obtained at the tested masses are linearly interpolated.

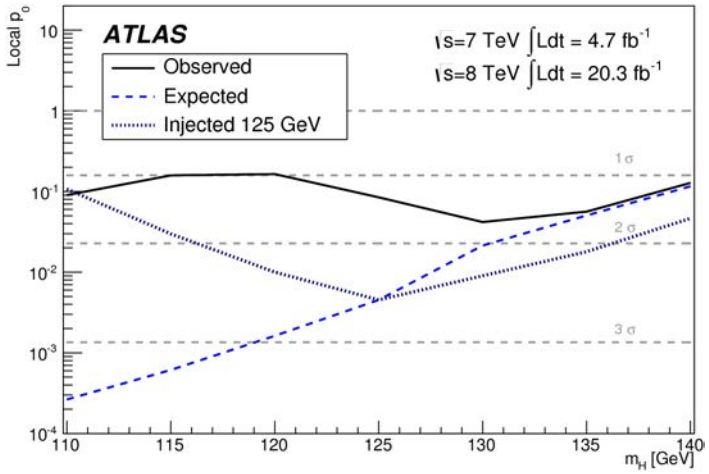


Figure 6.29: Observed (solid) and expected p_0 values as a function of m_H for all channels and data-taking periods combined, as obtained using the dijet-mass analysis for the 7 TeV dataset and BDTs trained at each individual mass for the 8 TeV dataset. The expected p_0 values are given for the background-only hypothesis in the presence of a SM Higgs boson: for the dashed curve the Higgs boson mass corresponds to each tested mass point in turn; for the dotted curve the Higgs boson mass is 125 GeV. For all curves shown, the results obtained at the tested masses are linearly interpolated.

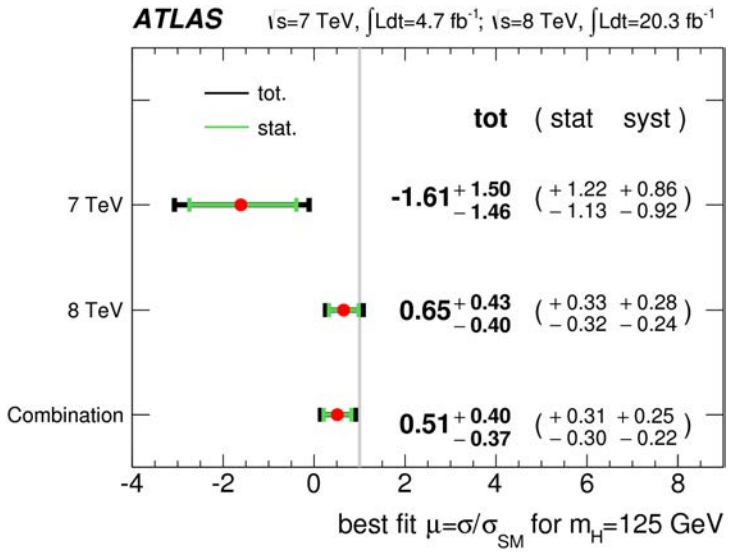


Figure 6.30: The fitted values of the Higgs boson signal-strength parameter μ for $m_H = 125$ GeV for the 7 TeV and 8 TeV datasets and the combination of the 7 TeV and 8 TeV datasets.

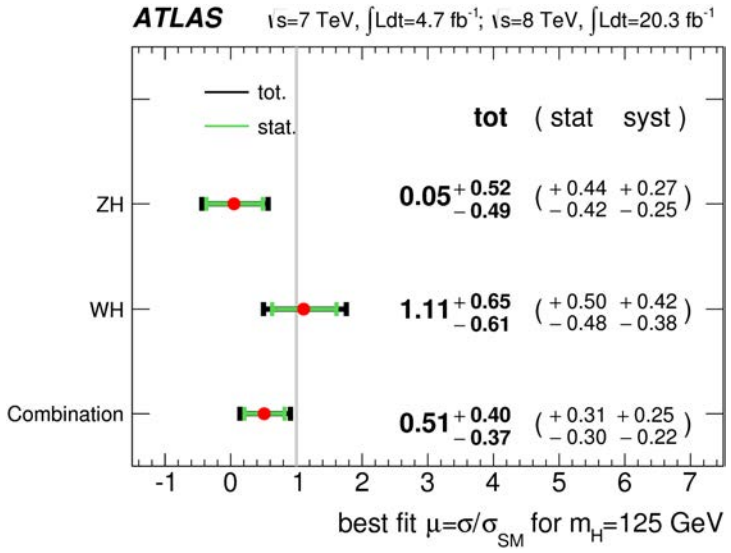


Figure 6.31: The fitted values of the Higgs boson signal-strength parameter μ for $m_H = 125$ GeV for the WH and ZH processes and the combination of the WH and ZH processes, with the 7 and 8 TeV datasets combined. The individual μ values for the $(W/Z)H$ processes are obtained from a simultaneous fit with the signal strength for each of the WH and ZH processes floating independently.

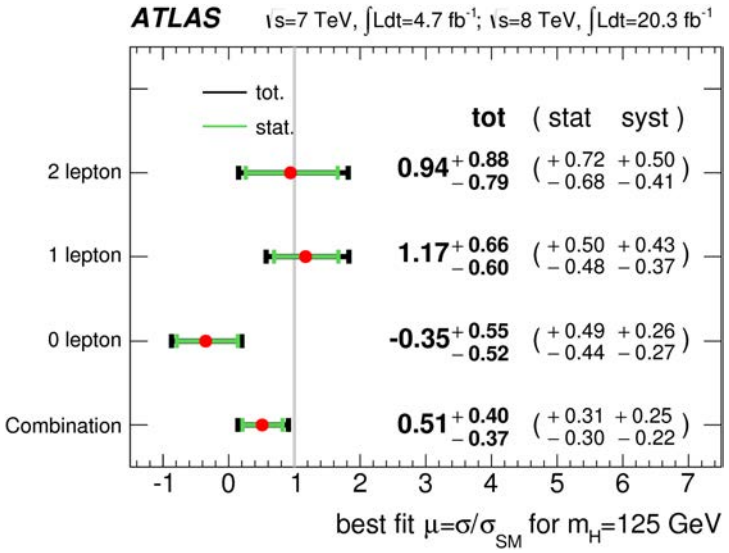
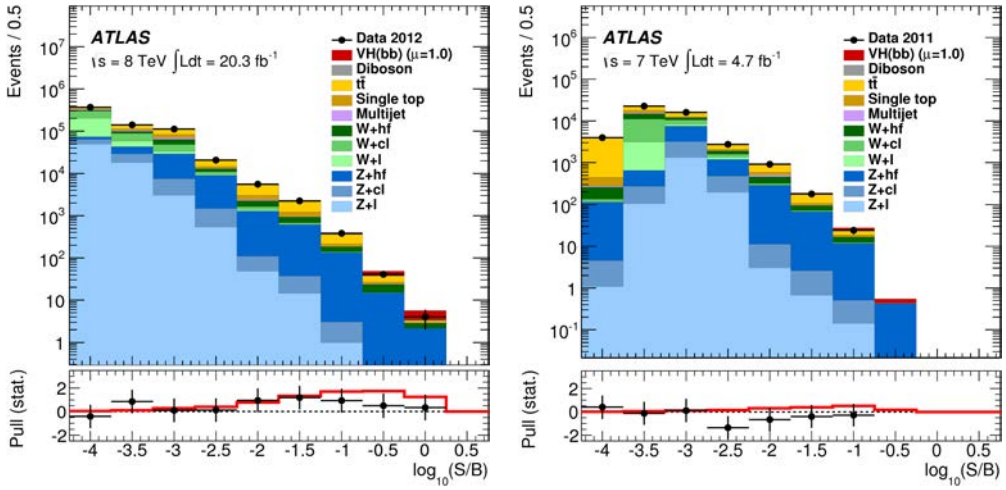
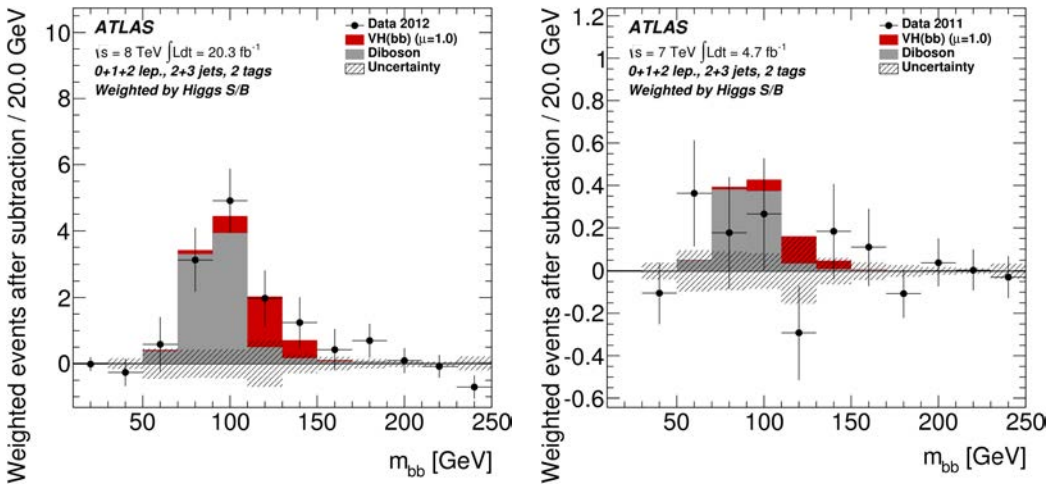


Figure 6.32: The fitted values of the Higgs boson signal-strength paramete μ for $m_H = 125$ GeV for the 0-, 1- and 2-lepton channels and the combination of the three channels, with the 7 and 8 TeV datasets combined. The individual μ values for the lepton channels are obtained from a simultaneous fit with the signal strength for each of the lepton channels floating independently.



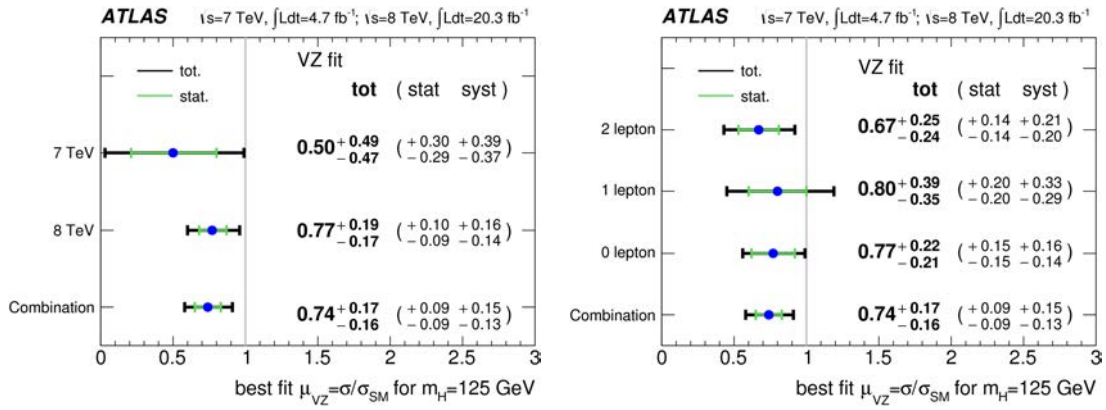
(a) (b)

Figure 6.33: Event yields as a function of log(S/B) for data, background and Higgs boson signal with $m_H = 125$ GeV for the (a) 8 TeV data and (b) 7 TeV data. Final-discriminant bins in all signal regions are combined into bins of log(S/B). The signal S and background B yields are expected and fitted, respectively. The Higgs boson signal contribution is shown as expected for the SM cross section (indicated as μ=1.0). The pull of the data with respect to the background-only prediction is also shown with statistical uncertainties only. The full line indicates the pull of the prediction for signal (μ=1.0) and background with respect to the background-only prediction.



(a) (b)

Figure 6.34: The distribution of m_{bb} in data after subtraction of all backgrounds except for the diboson processes, as obtained with the dijet-mass analysis for the (a) 8 TeV and (b) 7 TeV data. The contributions from all lepton channels, p_T^V intervals, number-of-jets and 2-tag b-tagging categories are summed weighted by their respective values of the ratio of expected Higgs boson signal to fitted background. The contribution of the associated WH and ZH production of a SM Higgs boson with $m_H = 125$ GeV is shown as expected for the SM cross section (indicated as μ=1.0). The size of the combined statistical and systematic uncertainty on the fitted background is indicated by the hatched band.



(a) (b)

Figure 6.35: The fitted values of the diboson signal strength mu_{VZ} for (a) the 7 TeV, 8 TeV and combined datasets, and (b) for the three lepton channels separately and combined, for the combined dataset. The MVA is used for the 8 TeV data. The individual mu_{VZ} values for the lepton channels are obtained from a simultaneous fit with the signal strength for each floating independently.

Process	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Data	368550	141166	111865	20740	5538	2245	382	41	4
Signal	29	43	96	57	58	62	32	10.7	2.3
Background	368802	140846	111831	20722	5467	2189	364	37.9	3.4
S/B	8×10^{-5}	0.0003	0.0009	0.003	0.01	0.03	0.09	0.3	0.7
W+hf	14584	10626	15297	1948	618	250	45	8.2	0.7
Wcl	96282	30184	15227	1286	239	47	4.2	0.2	0.005
Wl	125676	14961	3722	588	107	16	1.3	0.03	0.001
Z+hf	10758	14167	21684	7458	1178	577	130	14.8	2.2
Zcl	13876	11048	4419	941	61	22	2.1	0.1	0.008
Zl	49750	18061	3044	537	48	15	1	0.05	0.004
$t\bar{t}$	30539	24824	26729	5595	2238	922	137	10	0.3
Single top	10356	9492	14279	1494	688	252	31	2.7	0.1
Diboson	4378	1831	1247	474	186	62	9.7	1	0.2
Multijet	12603	5650	6184	400	103	26	3	0.9	0

Figure 6.36: The numbers of expected signal and fitted background events and the observed numbers of events after MVA selection in the bins of Figure 6.33(a). These numbers are for both the 1-tag and 2-tag events in the 8 TeV dataset, corresponding to an integrated luminosity of 20.3 fb^{-1}.

Sample	1-tag		LL-tag		MM-tag	MM+TT-tag	TT-tag
p_{T}^{V}	$100-120\,\mathrm{GeV}$	$>120\,\mathrm{GeV}$	$100-120\,\mathrm{GeV}$	$>120\,\mathrm{GeV}$	$100-120\,\mathrm{GeV}$	$>120\,\mathrm{GeV}$	$100-120\,\mathrm{GeV}$
	2-jet						
VH	7.9	23	2.2	6.9	3.5	23	4.0
WZ, ZZ, WW	235	635	18	49	14.7	81	13.3
$t\bar{t}$	840	1520	114	183	129	332	116
Single top	531	704	40	56	32.6	66	22.8
Wl	5470	7100	159	206	14.4	16	0.2
Wcl	2230	3710	106	159	23.4	27	1.1
W+hf	762	1520	54	124	33.6	128	21.8
Zl	3890	10750	96	272	6.1	17	0.1
Zcl	1590	3990	59	162	9.4	24	0.3
Z+hf	2550	6510	225	607	186	876	151
Total	18340 ± 150	36890 ± 200	886 ± 17	1841 ± 25	458.5 ± 9.8	1599 ± 28	333.2 ± 8.9
Data	18343	36903	887	1860	477	1592	306
	3-jet						
VH	–	8	–	2	–	7	–
WZ, ZZ, WW	–	260	–	17	–	20	–
$t\bar{t}$	–	1670	–	186	–	315	–
Single top	–	318	–	25	–	30	–
Wl	–	2280	–	59	–	4.3	–
Wcl	–	1240	–	53	–	8.9	–
W+hf	–	750	–	60	–	62	–
Zl	–	3190	–	79	–	4.5	–
Zcl	–	1620	–	65	–	9.8	–
Z+hf	–	1890	–	170	–	259	–
Total	–	13310 ± 100	–	718 ± 12	–	719 ± 17	–
Data	–	13344	–	657	–	710	–

Figure 6.37: The expected signal and fitted background yields for each category of the 0-lepton channel after the full selection of the multivariate analysis. The m_{bb} distribution is used in the $100 < p_T^V < 120$ GeV interval. The background yields are normalized by the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties.

Sample	1-tag		LL-tag		MM-tag		TT-tag	
p_{T}^{V}	< 120 GeV	> 120	< 120 GeV	> 120 GeV	< 120 GeV	> 120 GeV	< 120 GeV	> 120 GeV
	2-jet							
VH	39	28	11	9.3	17	14	19	16
WZ, ZZ, WW	1950	927	103.6	62	64	36	52	29.5
$t\bar{t}$	11380	6641	1954	1051	2426	1080	2290	890
Single top	13680	3730	1150	398	975	307	739	219
Wl	65980	23702	1603	697	124	48	3.1	0.9
Wcl	71930	21650	2630	966	465	153	23	6.5
W+hf	16030	6112	1470	648	954	402	506	227
Zl	3940	1223	101	37	7.4	2.8	0.1	0.0
Zcl	1350	333	53	18	10	3.1	0.5	0.2
Z+hf	2080	475	161	45	126	30	85	24
MJ (e)	–	2618	–	162	–	89	–	40
MJ (μ)	10230	164	721	16.0	329	4.8	178	1.3
Total	198540 ± 500	67600 ± 290	9953 ± 91	4106 ± 50	5492 ± 66	2161 ± 33	3889 ± 55	1448 ± 27
Data	198544	67603	9941	4072	5499	2199	3923	1405
	3-jet							
VH	15	14	3.2	3.8	4.8	5.8	5.4	6.5
WZ,ZZ,WW	1100	689	50	39.6	22.6	18	14	14
$t\bar{t}$	18660	10490	3240	1622	4119	1670	4181	1388
Single top	7390	2815	66	318	619	261	503	188
W+l	24980	11320	588	322	42	20	1.1	0.4
W+cl	25900	10080	952	454	164	72	7.7	3.2
W+hf	6530	4740	576	490	353	297	187	168
Z+l	1780	572	43	18.1	2.8	1.4	0.0	0.0
Z+cl	690	193	27	9.8	4.5	1.6	0.2	0.1
Z+hf	1024	272	77	25.9	54	18.8	40	14
MJ (e)	–	1290	–	68.6	–	36	–	15
MJ (μ)	5300	91	227	4.9	117	3.2	58	0.8
Total	93350 ± 320	42570 ± 200	6447 ± 57	3376 ± 43	5501 ± 50	2405 ± 33	4995 ± 55	1796 ± 30
Data	93359	42557	6336	3472	5551	2356	4977	1838

Figure 6.38: The expected signal and fitted background yields for each category of the 1-lepton channel after the full selection of the multivariate analysis. The background yields are normalized by the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties.

Sample	1-tag		LL-tag		MM+TT-tag	
p_{T}^{V}	< 120 GeV	> 120 GeV	< 120 GeV	> 120 GeV	< 120 GeV	> 120 GeV
	2-jet					
VH	20	7.4	5.5	2.0	19	7.0
WZ,ZZ,WW	800	135	73	12.2	129	22
$t\bar{t}$	3490	213	1023	49	3180	137
Single top	385	44	40	5.8	96	12
W+l	245	12	3.7	0.4	0.3	0.0
W+cl	166	18	5.8	0.6	1.0	0.1
W+hf	90	13	7.5	0.7	4.9	0.6
Z+l	29410	4180	648	101	41	6.2
Z+cl	12130	1622	421	64	71	9.7
Z+hf	21090	2853	1916	287.7	2858	414
Total	69360 $\pm$ 280	9123 $\pm$ 86	4271 $\pm$ 50	524.6 $\pm$ 7.9	6473 $\pm$ 76	607 $\pm$ 13
Data	69313	9150	4313	517	6501	570
	3-jet					
VH	7.4	3.9	1.5	0.9	4.9	2.8
WZ,ZZ,WW	347	100	22	6.9	29	8.7
$t\bar{t}$	2170	182	486	29	1476	74
Single top	144	22	18	3.0	41	6.5
W+l	76	7.1	1.4	0.2	0.1	0.0
W+cl	57	8.9	2.3	0.4	0.4	0.1
W+hf	40	6.4	3.1	0.8	0.8	0.4
Z+l	8870	1913	190	46	12	2.7
Z+cl	4650	949	164	38	25	5.7
Z+hf	5790	1223	497	121	766	180
Total	22720 $\pm$ 150	4430 $\pm$ 56	1403 $\pm$ 20	245.2 $\pm$ 5.2	2365 $\pm$ 43	280.9 $\pm$ 8.4
Data	22662	4436	1428	253	2394	283

Figure 6.39: The expected signal and fitted background yields for each category of the 2-lepton channel after the full selection of the multivariate analysis. The background yields are normalized by the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties.

7 | Search for a CP-odd A boson decaying to $Zh(b\bar{b})$ with data collected at $\sqrt{s} = 13$ TeV during the 2015 LHC data-taking

The Higgs boson particle discovered by the ATLAS and CMS Collaborations in 2012 [6, 7] has so far been observed to be consistent with the Standard Model (SM) predictions, and it can be used to probe beyond-the-SM scenarios in which the Higgs SM properties are altered, or additional Higgs bosons are included. In particular many physics models which include an extended Higgs sector also include a CP-odd (pseudoscalar) Higgs boson, denoted A, and a lighter CP-even Higgs boson, denoted h, as described in more detail in Section 2.5. In this analysis we search for the resonant production of a heavy A boson, decaying to a Z vector boson and a h boson with mass of 125 GeV (denoted as $A \to Zh$), targeting leptonic final states in the decays of the Z boson ($Z \to e^+e^-$, $Z \to \mu^+\mu^-$, $Z \to \nu\bar{\nu}$), with the h decaying to a bottom quark pair ($h \to b\bar{b}$).
The search is performed with data collected by the ATLAS detector during the Run-2 of the LHC, for the data-taking period from September until December 2015, corresponding to an integrated luminosity of 3.2 fb^{-1} [207]. The results are interpreted in the context of two-Higgs doublet models (2HDM) of Type I and II, where the Higgs potential is CP-conserving. Previous searches in the same final states have been performed by experiments at LEP, Tevatron, and by both the ATLAS and the CMS experiments using data at $\sqrt{s} = 8$ TeV. These are reported in [217, 218, 13, 14].

7.1 Introduction: The Analysis Strategy

The work presented in this thesis is mainly focused on the search for the SM Higgs boson in the bottom-quark pair decay channel, therefore the inclusion of this search may seem out of place. There are however several reasons which motivate this analysis performed with the first data collected by the ATLAS detector during the LHC Run-2. In this introduction I present the strategy that lead to the publication of these results, and furthermore the main distinguishing features of this analysis which stem from the particular nature of the $A \to Zh$ signal investigated.

From the experience accumulated with the analysis of the LHC Run-1 dataset it was known that with the amount of data collected by the ATLAS detector during the first months of the LHC Run-2 (until the end of December 2015, corresponding to an integrated luminosity of 3.2 fb^{-1}) a rough estimate of the expected sensitivity of the SM $VH(b\bar{b})$ search could reach at most at 1σ (accounting for the change in the signal and background partonic cross sections and in the parton luminosities, between $\sqrt{s} = 8$ TeV and 13 TeV).
However, in order to gain at least partial confidence with the characteristics of the SM $VH(b\bar{b})$ analysis, it is interesting and worthwhile to investigate a different physics scenario

which could lead to promising results with these first data, and at the same time presents a final state very similar to the SM $VH(b\bar{b})$ one. This is the case of the search for a heavy CP-odd A boson decaying to a Zh pair, where the h boson corresponds to a SM-like Higgs boson particle, subsequently decaying to the 0- and 2-lepton final states investigated by the SM analysis.
With the center of mass energy increasing from 8 TeV up to 13 TeV, the production cross section for this type of signals in 2HDMs can be enhanced by a factor of 2-9 (increasing with the mass of the A boson, and depending on the type of 2HDM considered), making this search a suitable candidate for an early investigation of LHC Run-2 data.
To clarify the notation, hereafter in this chapter the SM-like Higgs boson with mass of 125 GeV is named h, consistently with the notation introduced in Section 2.5 for 2HDM particles (while in Chapter 6 and Chapter 8 the SM Higgs boson is indicated with H).

This approach has interesting features: most notably, since the h boson produced in association with the Z by the A boson decay corresponds to a SM-like Higgs boson, a positive result of this search would also lead to an evidence/observation for the SM-like h boson in the bottom-quark pair decay channel. This is not, of course, the primary goal of this analysis, but it represents an attractive byproduct.
There are of course some significant differences in the nature of the $A \to Zh$ signal processes investigated in this Chapter, and the SM $VH(b\bar{b})$ process described in Chapter 6 and 8, which lead to some of the most peculiar aspects of this search.
First, given the neutral electromagnetic charge of the A boson, decays to $W^{\pm}h$ pairs are forbidden, thus excluding the 1-lepton channel from this analysis. This aspect makes this analysis simpler than the SM $VH(b\bar{b})$ search, since the level of accuracy required for the modeling of some of the main 1-lepton channel backgrounds, such as W boson production in association with heavy flavor jets and multi-jet processes, is lower.
Furthermore, the Zh pair is the decay product of a heavy A boson, with mass m_A tested in this analysis from 220 GeV up to 2 TeV. For large values of m_A the Z and h bosons can be strongly boosted, reaching transverse momenta above 500-600 GeV. In this regime the decay products of the h boson tend to be strongly collimated to each other, and the efficiency of reconstructing them as resolved hadronic jets decreases. This key feature motivates the definition of two main regions in which this search is performed: a 'resolved regime', where the h decay products are identified as separated hadronic jets with smaller radius, and a 'merged regime', where the same decay products are reconstructed as a single hadronic jet with large radius.

The structure of this Chapter is the following: in Section 7.2 I present the dataset considered for this search and the MC simulation used to estimate signal and background contributions; in Section 7.3 I describe the details of the selection strategy adopted for the different lepton channels, while the main features of the signal and background MC estimate are given in Section 7.4; in Section 7.5 the experimental and theoretical systematic uncertainties are outlined; Section 7.6 contains the main features of the statistical approach used to extract the final results, which are presented in Section 7.7. The details of the reconstruction of the main physics objects used in this analysis, along with the main event reconstruction techniques, were described in Chapter 5.

7.2 Data and Simulated Samples

Dataset This analysis is performed with data recorded with the ATLAS detector in stable beam conditions from pp collisions at $\sqrt{s} = 13$ TeV during the first part of the LHC Run-2, with all relevant sub-detectors operating in good conditions. The dataset analyzed corresponds to an integrated luminosity of 3.2 fb^{-1}. Data collected in ideal conditions of the ATLAS detector are summarized in GRL, passing high quality controls from the ATLAS Data Quality Group. Cleaning selections analogous to the cuts reported in Section 6.2 are applied, updated to reflect the different beam and detector conditions. For the dataset collected during 2015 considered in this analysis the mean number of pile-up interaction amounts to 13.5.

Trigger selection The trigger strategy is similar to the one adopted for the analysis of Run-1 data described in Section 6.2 (although much simpler since the data-taking period is considerably shorter, therefore data-taking conditions are subject to less changes): in the 0-lepton channel, events are recorded from an un-prescaled E_T^{miss} trigger with online threshold of 70 GeV; 2-lepton events are triggered by at least one of the lowest un-prescaled single-electron or single-muon triggers, with lowest available p_T threshold at 24 GeV for isolated electrons and 20 GeV for isolated muons. In events recorded via the lepton triggers the object that satisfies the trigger requirement has to match an offline-reconstructed lepton.

Monte Carlo samples The signals and all the SM backgrounds considered in this analysis are modeled with MC simulations. All MC samples are processed using the GEANT 4-based ATLAS detector simulation [176], and reconstructed with the standard ATLAS reconstruction software used for collision data. The simulated samples (except those generated with SHERPA) exploit the `EvtGen 1.2.0` program [166] to describe properties of the bottom and charm hadron decays.

The $A \to Zh$ signal samples are generated considering two different production modes, which are tested in this search: both gluon-fusion production $gg \to A$ and b-quark associated production $b\bar{b}A$ processes are simulated with MADGRAPH5_AMC@NLO 2.2.2 [219, 212] with NNPDF2.3LO PDFs, interfaced to the PYTHIA 8 parton shower and hadronization model, using again the NNPDF2.3LO PDFs, configured with the ATLAS A14 tune [220]. The light CP-even h boson in the simulation is always assumed to correspond to the SM-like Higgs boson with $m_h = 125$ GeV and narrow width compared to the resolution of the detector for $h \to b\bar{b}$ decays. The $h \to b\bar{b}$ branching ratio is set to the SM expected value of $58 \pm 2\%$ [51] for the baseline MC prediction, however it is modified by the parameter of the tested 2HDMs when interpreting the results in terms of $\tan\beta$ and $\cos(\beta - \alpha)$. The cross section $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$ for these signal processes is fixed to 0.5 pb (rather than the specific values predicted by one of the different models), and scaled by the Z boson leptonic branching ratios.
The CP-odd A boson has masses m_A ranging from 220 GeV to 2 TeV, and its width is assumed to be narrow when compared to the m_A resolution in the generation of MC events.

The (W/Z) + jets background prediction relies on the SHERPA 2.1.1 [189] generator, with CT10 PDFs and a dedicated SHERPA internal parton shower tune, with massless b- and c-quarks in the ME calculation, while massive quarks can be produced in the parton shower.

In order to provide a good modeling of large jet multiplicities, SHERPA 2.1.1 provides a combination of matrix elements with different parton multiplicities, using the COMIX [221] and OPENLOOPS [222] generators, and merging the MEs with the SHERPA parton shower [223] using the ME+PS@NLO [224] prescription via the CKKW-L [225] merging technique, with a merging scale $Q_{\text{cut}} = 20$ GeV: up to 2 extra partons are included in the NLO ME, and 3 or 4 extra partons are included in the LO ME. The modeling of even higher jet multiplicities relies on the parton shower algorithm.
As for the (W/Z)+jets MC samples described in Section 6.2, also in this case the simulated events are filtered according to the presence of b-, c- or light-flavored hadrons, and sliced in several sub-samples depending on the p_T of the vector boson. This second selection in p_T^V is particularly important since this search probes regions of the phase space at very large transverse momentum: reducing possible statistical fluctuations of the background prediction in the boosted regime is therefore crucial. The boundaries of the p_T^V selection are the following:

$$p_T^V \text{ bins} = [0,\ 40,\ 70,\ 140,\ 280,\ 700,\ 1000,\ 2000)\ \text{GeV}$$
$$(\text{the last bin contains events with } p_T^V > 2000\ \text{GeV})$$

The $(W/Z)+$jets total cross sections are calculated at NNLO (QCD) accuracy [226, 227].
Top-quark pair production is simulated with the POWHEG generator with CT10 PDFs interfaced to the PYTHIA 6 parton shower (with similar setup to what detailed in Section 6.2). The $t\bar{t}$ samples generated with this setup include a filter to require at least one W boson to decay leptonically. The POWHEG resummation damping factor `hdamp`, which is one of the parameters controlling the matching between ME and parton shower and regulating the emission of high-p_T radiation, is set to the mass of the top quark $m_{top} = 172.5$ GeV to obtain a better data-to-MC agreement in the high-p_T region for this process, as show in [228]. The $t\bar{t}$ cross section is calculated at NNLO in QCD including resummation of NNLL soft gluon terms with the `top++2.0` software [229].
The theoretical cross sections for Z+jets and $t\bar{t}$ processes are used to normalize the corresponding MC simulation to perform optimization studies, however the overall event yields of these backgrounds are included as unconstrained parameters in the analysis and are determined from the fit to data described in Section 7.6.
The modeling of single-top events in s-, t- and Wt- production channels is obtained with the POWHEG generator, with CT10 PDFs, interfaced with the PYTHIA 6 parton shower model (with the same settings described in Section 6.2); the single-top cross sections are computed at NLO (QCD) with the `Hator v2.1` software [230, 231] and used to normalize the MC prediction. Single top samples in the t- and s-channels are generated applying a lepton filter to require the leptonic decay of the W boson, while Wt-channel samples are generated inclusively.
Semileptonic diboson production (ZZ, WZ and WW) is simulated with the SHERPA 2.1.1 generator with CT10 PDFs and internal parton shower tuning, normalized to the NLO cross section provided by the MC generator itself. With the same methodology employed for the $(W/Z)+$jets simulation, SHERPA provides a combination of different matrix elements with different parton multiplicities: for WZ and WW, processes with zero extra partons are calculated at NLO in the ME, while 1,2 or 3 additional partons are included at LO, while for ZZ one additional parton is included in the NLO ME. The contribution of gg-initiated ZZ and WW processes is not included in this analysis.
The associated production of a SM Higgs boson Vh is treated as background in this search, and its modeling relies on the PYTHIA 8 generator for qq-initiated processes, while the gg-initiated ZH contribution is simulated with POWHEG +PYTHIA 8 (with similar setup to what detailed in Section 6.2). The Vh signals are normalized to the

NNLO(QCD)+NLO(EW) cross sections [51].

The MC samples used in this analysis include the contribution of multiple pp interactions from the same and from neighboring bunch-crossing (pile-up), overlaying simulated minimum-bias events to each generated event. The simulation of pile-up interactions is obtained from soft QCD processes with PYTHIA 8, with the A2 tune and the MSTW2008LO PDFs.

A complete and schematic list of the MC generators used in this analysis for all signals and backgrounds is reported in Table 7.1.

Process	Generator	$\sigma \times BR$
$A \to Zh$ signal		
$gg \to A \to Zh \to l^+l^-b\bar{b}$	MADGRAPH5_aMC@NLO +PYTHIA 8	0.05 pb
$gg \to A \to Zh \to \nu\bar{\nu}b\bar{b}$	MADGRAPH5_aMC@NLO +PYTHIA 8	0.1 pb
$b\bar{b}A \to Zh \to l^+l^-b\bar{b}$	MADGRAPH5_aMC@NLO +PYTHIA 8	0.05 pb
$b\bar{b}A \to Zh \to \nu\bar{\nu}b\bar{b}$	MADGRAPH5_aMC@NLO +PYTHIA 8	0.1 pb
Vector boson + jets		
$W \to l\nu$	SHERPA 2.1.1	20.080 nb
$Z/\gamma^\star \to l^+l^-$ $(66 < m_{ll} < 116$ GeV)	SHERPA 2.1.1	1.906 nb
$Z \to \nu\bar{\nu}$	SHERPA 2.1.1	1.914 nb
Top-quark		
$t\bar{t}$	POWHEG +PYTHIA 6	831.76 pb
t-channel	POWHEG +PYTHIA 6	69.51 pb
s-channel	POWHEG +PYTHIA 6	3.314 pb
Wt-channel	POWHEG +PYTHIA 6	68.00 pb
Diboson		
WW	SHERPA 2.1.1	51.97 pb
WZ	SHERPA 2.1.1	23.72 pb
ZZ	SHERPA 2.1.1	7.00 pb
SM Vh Higgs (background)		
$q\bar{q} \to ZH \to \nu\bar{\nu}b\bar{b}/l^+l^-b\bar{b}$	PYTHIA 8	132.7 fb
$gg \to ZH \to \nu\bar{\nu}b\bar{b}/l^+l^-b\bar{b}$	POWHEG +PYTHIA 8	18.3 fb
$q\bar{q} \to WH \to l\nu b\bar{b}$	PYTHIA 8	258.0 fb

Table 7.1: The generators used for the simulation of the signal and background processes, along with the cross sections times branching ratios (corrected for filter efficiencies when relevant) used to normalize the MC samples.

7.3 Event selection

This search targets $A \to Zh \to (l^+l^-, \nu\bar{\nu})b\bar{b}$ signals (where $l = e, \mu$), hence one of the main differences with respect to the $Vh(b\bar{b})$ searches described in Chapter 6 and 8 is the lack of

a 1-lepton channel, since charge conservation forbids the $A \to W^{\pm}h$ decay. The analysis is thus optimized categorizing events in two lepton channels, 0-lepton and 2-lepton, according to the number of charged leptons.
Selected events are furthermore divided according to the reconstructed transverse momentum of the vector boson p_T^Z, in low- and high-p_T^Z categories. This is a crucial step since the event reconstruction strategy differs largely between low- and high-p_T^Z events: the low-p_T^Z category corresponds to the 'resolved regime' introduced in Section 7.1, while the high-p_T^Z category defines the 'merged regime'. The p_T^Z is reconstructed as the E_T^{miss} in the 0-lepton channel and as the vector sum of the transverse momenta of the two leptons in the 2-lepton channel. The boundary between low- and high-p_T^Z regions is set at $p_T^Z = 500$ GeV, where this value has been tuned to optimize the expected sensitivity of the analysis.
Specific selection criteria are applied for different p_T^Z regions (common across lepton channels) to reconstruct events in the different regimes:

- **low-p_T^Z**: only events with at least two selected central small-R jets with $p_T > 20$ GeV and $|\eta| < 2.5$ (central) are retained. Events are divided into two b-tagging categories (1-tag, 2-tag) by counting the number of b-tagged central small-R jets: events with more than 2 b-tagged central jets are rejected. At least one of the b-tagged jets must have $p_T > 45$ GeV. In the 2-tag category, the Higgs candidate is built from the two b-tagged jets, while in the 1-tag category the b-tagged jet and the highest-p_T remaining central jet are used.
  The dijet-mass resolution is improved by applying specific b-jet energy corrections described in Section 5.1.4. Events are furthermore categorized according to the value of the reconstructed mass of the Higgs candidate, after applying these corrections: the signal regions are defined by selecting 110 GeV $< m_{\text{dijet}} <$ 140 GeV (consistent with the SM $m_H = 125$ GeV), while the background control regions contain events in the m_{dijet} side-bands. Both signal and control regions are defined separately for 1-tag and 2-tag events. The mass window signal region selection has an efficiency of 70%-80% for signal events [232].

- **high-p_T^Z**: only events with at least one selected large-R jet with $p_T > 250$ GeV and $|\eta| < 2.0$, and at least two track jets with $p_T > 10$ GeV and $|\eta| < 2.0$ associated to it, are retained. Events are divided into two b-tagging categories (1-tag, 2-tag) by applying the b-tagging algorithm to the two leading track jets associated with the leading large-R jet. The Higgs candidate is defined as the highest-p_T large-R jet in both 1- and 2-tag categories, applying to it the jet energy corrections described in Section 5.1.4. Events are divided into signal and control regions with the same approach as in the low-p_T^Z category, however the mass window is widened to 75 GeV $< m_{\text{jet}} <$ 145 GeV to account for the lower mass resolution. The mass window signal region selection has an efficiency of 90% for signal events [232].

Detailed definitions of the hadronic jet collections and b-tagging algorithms used to build these selection criteria are given in Chapter 5. Note that the truth-tagging procedure and the kinematic fit described in the same Chapter are not used in this search.
In both regimes, events with more than 2 b-tagged jets are rejected to suppress the $t\bar{t}$ background contribution. Except for the requirement of at least two central small-R jets (low-p_T^Z) or one large-R jet (high-p_T^Z), no additional categorization in jet-bins is considered and no veto on higher jet multiplicities is applied. This constitutes a major difference with respect to the analysis described in Chapter 6, motivated by the different nature of $A \to Zh$ signal: since the main production mode investigated is through gluon-fusion, signal events

show higher jet multiplicities compared to the SM qq-initiated $Vh(b\bar{b})$ processes. This is due to the different color-factors associated to the emission of real gluons from quarks ($C_F = 4/3$) or from gluons themselves ($C_A = 3$).
Vetoing events with many jets would therefore significantly impact the signal acceptance degrading the analysis performance. Selected small-R jets (in both p_T^Z categories) used in the following selections may be central ($p_T > 20$ GeV and $|\eta| < 2.5$) or forward ($p_T > 30$ GeV and $2.5 < |\eta| < 4.5$).
A schematic representation of all regions included in the analysis, separately for the resolved and merged regimes, is shown in Figure 7.1 (the $t\bar{t}$ control region shown here is introduced in Section 7.4).

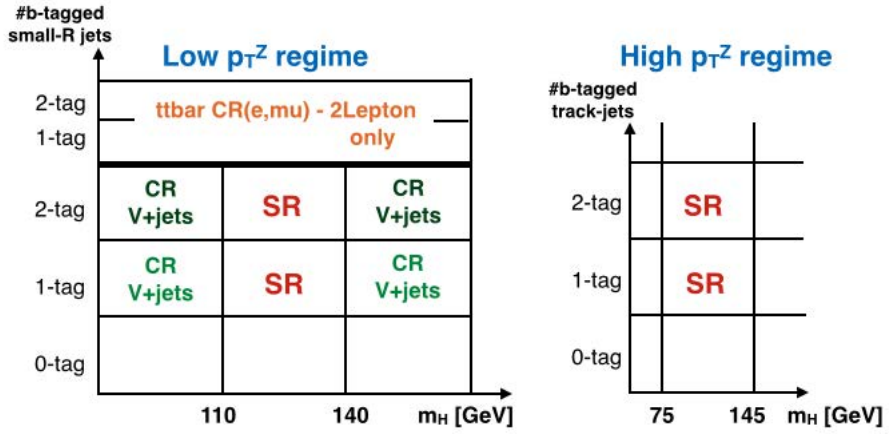


Figure 7.1: Illustration of the signal and control regions in the 0-lepton and 2-lepton low-p_T^Z (left) and high-p_T^Z (right) categories.

After this first categorization, additional topological and kinematic selection criteria are applied for the different lepton channels, summarized in Table 7.2. Most of the selection cuts are taken from the Run-1 $Vh(b\bar{b})$ analysis detailed in Section 6.3 with values adjusted for the different phase space and background composition of this search:

- **0-lepton channel**: only events with exactly zero charged leptons are retained. To select a region where the E_T^{miss} trigger is almost fully efficiency, the reconstructed E_T^{miss} has to be larger than 150 GeV: this ensures a trigger efficiency of 85% at $E_T^{miss} = 150$ GeV, rising to 100% above 200 GeV. In the low-p_T^Z region an additional cut on the scalar sum of the transverse momenta of the selected jets is applied to avoid mismodeling in the turn-on curve of the E_T^{miss} trigger. The $\sum p_T^{jet_i}$ is defined including up to the first three selected central small-R jets ordered in p_T (if only two central small-R jets are selected, the leading forward small-R jet is included if present); the value of the cut is set at 150 GeV for events with more than two selected jets, at 120 GeV for events with exactly two jets (counting both central and forward small-R jets).
  The contribution of the MJ and non-collision background is suppressed by selecting events with: large magnitude of the track-based missing transverse momentum (p_T^{miss}) to reject non-collision background; small azimuthal angle between E_T^{miss} and p_T^{miss} ($\Delta\phi(E_T^{miss}, p_T^{miss})$), large azimuthal angle between E_T^{miss} and the closest jet ($\min[\Delta\phi(E_T^{miss}, \text{small-R jet})]$) calculated using all small-R jets for the high-p_T^Z category and the same jets included in the $\sum p_T^{jet_i}$ definition for the low-p_T^Z category,

to select events with genuine E_T^{miss} contribution, not originated by mismeasured objects; large azimuthal distance between the E_T^{miss} and the h candidate reconstructed as described above for the two p_T^Z regions; small azimuthal distance between the selected small-R jets used to build the h candidate $\Delta\phi(\text{jet}_1, \text{jet}_2)$ (for low-p_T^Z events only), to retain events with the typical $Vh(b\bar{b})$ topology. After these cuts the contribution of MJ background estimated from MC simulation is reduced below 1%.
Furthermore events with at least one hadronically decaying τ leptons, reconstructed as described in Section 5.1.2, are discarded to reduce the contribution of W+jets and $t\bar{t}$ backgrounds. In the high-p_T^Z regions the $t\bar{t}$ background is further suppressed by vetoing events with at least one b-tagged track jets not associated with the selected large-R jet.

- **2-lepton channel**: only events containing exactly two oppositely charged same flavor leptons (electrons or muons) are selected, where at least one lepton is required to have $p_T > 25$ GeV and $|\eta| < 2.5$ (both leptons have to satisfy this requirement in the high-p_T^Z regions). To select events compatible with the leptonic Z boson decay a window selection on the 2-lepton invariant mass is applied, in the range of $70 - 110$ GeV for low-p_T^Z events, widened to $55 - 125$ GeV for high-p_T^Z muon events (due to the worse muon momentum resolution at high transverse momenta). The $t\bar{t}$ background is suppressed by selecting events with $E_T^{miss}/\sqrt{H_T} < 3.5\sqrt{\text{GeV}}$ where H_T is the scalar sum of the p_T of the leptons and the small-R jets in the event.

The discriminating variable used in this search is relatively simpler compared to the complex BDTs used in the SM $VH(b\bar{b})$. In the 0-lepton channel, since fully reconstructing the invariant mass of the Zh system is not possible due to the presence of undetectable neutrinos, the transverse mass m_T^{Zh} is utilized, defined as:

$$m_T^{Zh} = \sqrt{(E_T^{miss} + E_T^h)^2 - (\vec{E_T^{miss}} + \vec{p_T^h})^2}$$

where $\vec{p_T^h}$ and E_T^h are the transverse components of the 3-momentum and the energy of the h candidate respectively, and $\vec{E_T^{miss}}$ is the 3-momentum term of the missing transverse energy 4-vector. When building the m_T^{Zh} variable the 4-momentum of the dijet system is scaled by m_h/m_{dijet}, where $m_h = 125$ GeV, only for low-p_T^Z events, in order to improve the transverse mass resolution.
In the 2-lepton channel the discriminating variable is the reconstructed invariant mass of the Zh system, defined from the h candidate and the two selected leptons. To improve its resolution the 4-momentum of the di-muon system is scaled by $m_Z/m_{\mu\mu}$, where $m_Z = 91.2$ GeV, in both low- and high-p_T^Z regions. Furthermore the 4-momentum of the dijet system is scaled by m_h/m_{dijet}, where $m_h = 125$ GeV, only for low-p_T^Z events, similarly to the 0-lepton channel. The final resolution of the m^{Zh} variable is in the range of 2.5%-3.5%.

Signal acceptances The event selection described in this Section is tailored to achieve the best analysis sensitivity for $A \to Zh$ signal events produced via gluon-fusion mechanism. The results of this search are however also interpreted considering signals from $b\bar{b}A$ associated production (although the analysis is not optimized for such processes), therefore it is useful to quote the signal efficiencies for the two production modes. To give a single figure of merit, all numbers are quoted for a signal with $m_A = 600$ GeV.
The 2-tag signal regions efficiency for gluon-fusion $A \to Zh$ events corresponds to 12.3% for $\nu\bar{\nu}b\bar{b}$ decays and 15.3% for the average of $l^+l^-b\bar{b}$ ($l = e, \mu$) decays. The 1-tag signal regions efficiency corresponds to 7.0% and 7.4% respectively.

The corresponding signal efficiencies for $b\bar{b}A$ produced $A \to Zh$ signal events are in the range of 58%-64% of those found for gluon-fusion production (for the same m_A values). The major drop in efficiency for this production mode corresponds to the veto of events with more than 2 b-tagged jets: removing this veto the $b\bar{b}A$ efficiencies rise to approximately 80% of the gluon-fusion production ones.

Variable	Low-p_{T}^{Z}	High-p_{T}^{Z}
Common selection		
p_{T}^{Z} [GeV]	<500	≥ 500
$N_{b\text{-tag jet}}$	1,2	1,2
$N_{\text{small-}R\text{ jet}}$	≥2	≥0
$N_{\text{large-}R\text{ jet}}$	≥0	≥1
m_{dijet} or m_{jet} [GeV]	110–140	75–145
0-lepton selection		
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 150	–
$\sum_{i=1}^{N_{\text{jet}}=3(2)} p_{\mathrm{T}}^{\text{jet}_i}$ [GeV]	> 150 (120)$^{(*)}$	–
$p_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 30	> 30
$\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})$	< $\pi/2$	< $\pi/2$
$\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, h)$	> $2\pi/3$	> $2\pi/3$
min[$\Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jet)]	> $\pi/9^{(*)}$	> $\pi/9^{(*)}$
$\Delta\phi(j,j)$	< $7\pi/9$	–
Number of hadronic taus	0	0
Number of b-tag track-jets not associated to the leading large-R jet	–	0
2-lepton selection		
m_{ee} [GeV]	70–110	70–110
$m_{\mu\mu}$ [GeV]	70–110	55–125
$E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$ [$\sqrt{\text{GeV}}$]	< 3.5	–

Figure 7.2: Event topological and kinematic selections. (∗) In the low-p_T^Z category of the 0-lepton channel the min[$\Delta\phi(E_T^{miss}$, small-R jet)] and the $\sum_{i=1}^{Njet=3(2)} p_T^{jet_i}$ are calculated using the first three central small-R jets ordered in p_T or, in case there are only two central small-R jets, with these two jets and the first forward small-R jet ordered in p_T.

7.4 Background Estimate and Modeling

As stated in Section 7.2 this analysis relies entirely on MC simulations to obtain the estimate of signal and background processes which is used as input for the Likelihood fit to data. In this section I will briefly highlight some of the key points that differentiate the background estimate from what described in Section 6.5, and characterize this analysis:

1. **MJ background**: since the 1-lepton channel is not included in this search, the impact of the MJ background is largely suppressed. With the selection cuts presented in Section 7.3 its contribution is reduced below 1%: no dedicated estimate of these

processes is therefore required. In order to optimize the cuts and to control its impact on the selected phase space, the PYTHIA 8 MC simulation is used to model the MJ background.

2. $(W/Z)+$**jets background prediction**: these background processes are similar to the ones described in Section 6.5, however the SHERPA MC prediction is improved from the LO estimate of SHERPA 1.4.1 used in Chapter 6, to the higher order prediction of SHERPA 2.1.1. The striking mismodelings found for the V+jets background in the analysis of Run-1 data are not observed in this search, comparing data to MC in the analysis control regions (specifically the 1-tag categories, with higher available data statistics): the variables that were used to extrapolate corrections for the Run-1 $VH(b\bar{b})$ analysis are well described by the simulation, thus no reweighting correction is applied in this search.

3. $(W/Z)+$**jets background flavor labeling**: the $(W/Z)+$jets backgrounds are separated in sub-component according to the true flavor of the two small-R jets forming the h candidate in low-p_T^Z events, or the true flavor of the two leading track jets associated to the selected large-R jet in high-p_T^Z events. The true flavor is determined with the same technique described in Section 6.2. The $(W/Z)+$jets simulated samples are then split into six components based on this association scheme: $W/Z+bb$, $W/Z+bc$, $W/Z+bl$, $W/Z+cc$, $W/Z+cl$ and $W/Z+ll$. For the Z+jets background, $Z+bb$, $Z+bc$ and $Z+cc$ are treated as a single component representing all events with a Z boson and two associated heavy-flavor jets, denoted '$Z+(bb,bc,cc)$', while '$Z+(bl,cl)$' represents $Z+bl$ and $Z+cl$. For the W+jets background, '$W+(bb,bc,cc,bl)$' contains the union of $W+bb$, $W+bc$, $W+cc$ and $W+bl$.

4. $t\bar{t}$ **background**: the $t\bar{t}$ background has sizable contribution in both 0- and 2-lepton channels. It is worth noting that this background shows the largest increase in partonic cross section from $\sqrt{s}=8$ to 13 TeV among all EW backgrounds, and it is furthermore enhanced since no veto on higher jet multiplicities is applied in this analysis. Less control on its estimate is provided by the Likelihood fit to data, since the 1-lepton channel is not included. For this reasons a specific $t\bar{t}$ control region is defined by applying the 2-lepton selection, relaxing the lepton invariant mass cut to $m_{ll}>40$ GeV, removing the $E_T^{miss}/\sqrt{H_T}$ cut and requiring the selected leptons to be of different flavor. This CR is defined for low-p_T^Z events only, separately for 1-tag and 2-tag events.

The remaining backgrounds (semileptonic diboson, single-top production and SM $VH(b\bar{b})$ processes) contribute to the level of a few percents in the analysis regions. When performing this analysis the differential NLO EW reweighting corrections for the SM $VH(b\bar{b})$ estimate (introduced in Section 6.5.3) were not available and therefore are not applied.
The distribution of the discriminating variables m_T^{Zh} and m^{Zh} are shown in Figure 7.3 for the m_{dijet} side-bands control regions, defined in Section 7.3, for the low-p_T^Z categories of the 0- and 2-lepton channels, separately for 1-tag and 2-tag events. The m^{Zh} distribution from events collected in the $t\bar{t}$ specific control regions is shown in Figure 7.4, separately for 1-tag and 2-tag categories.
The distribution of the m_{dijet} variable, key to define signal and control regions and to build the final discriminating variables, is shown in Figure 7.5 and 7.6 for the same set of regions.
A crucial aspect of this search is the possibility to probe two different regimes simultaneously: the low-p_T^Z regime, which shares several features of the background estimate and

modeling with the SM $VH(b\bar{b})$ search, and the high-p_T^Z regime. In the 'resolved regime' the contribution of EW backgrounds is large and their modeling has a sizable impact on the analysis performance, while in the 'merged regime' the number of events is fairly low due to the requirement of $p_T^Z > 500$ GeV, therefore the fine details of the background modeling are less impactful. Figure 7.7 show the p_T^Z distribution for 0- and 2-lepton signal regions, separately for 1- and 2-tag categories, combining together the low- and high-p_T^Z regimes, clearly showing the different magnitude of the background contribution for $p_T^Z \lesseqgtr 500$ GeV (the discontinuity at $p_T^Z = 500$ GeV is due to the different selection cuts in the low- and high-p_T^Z categories).
In all Figures 7.3, 7.4, 7.5, 7.6 and 7.7, the background expectation is shown after the maximum Likelihood fit to the data, described in Section 7.6.

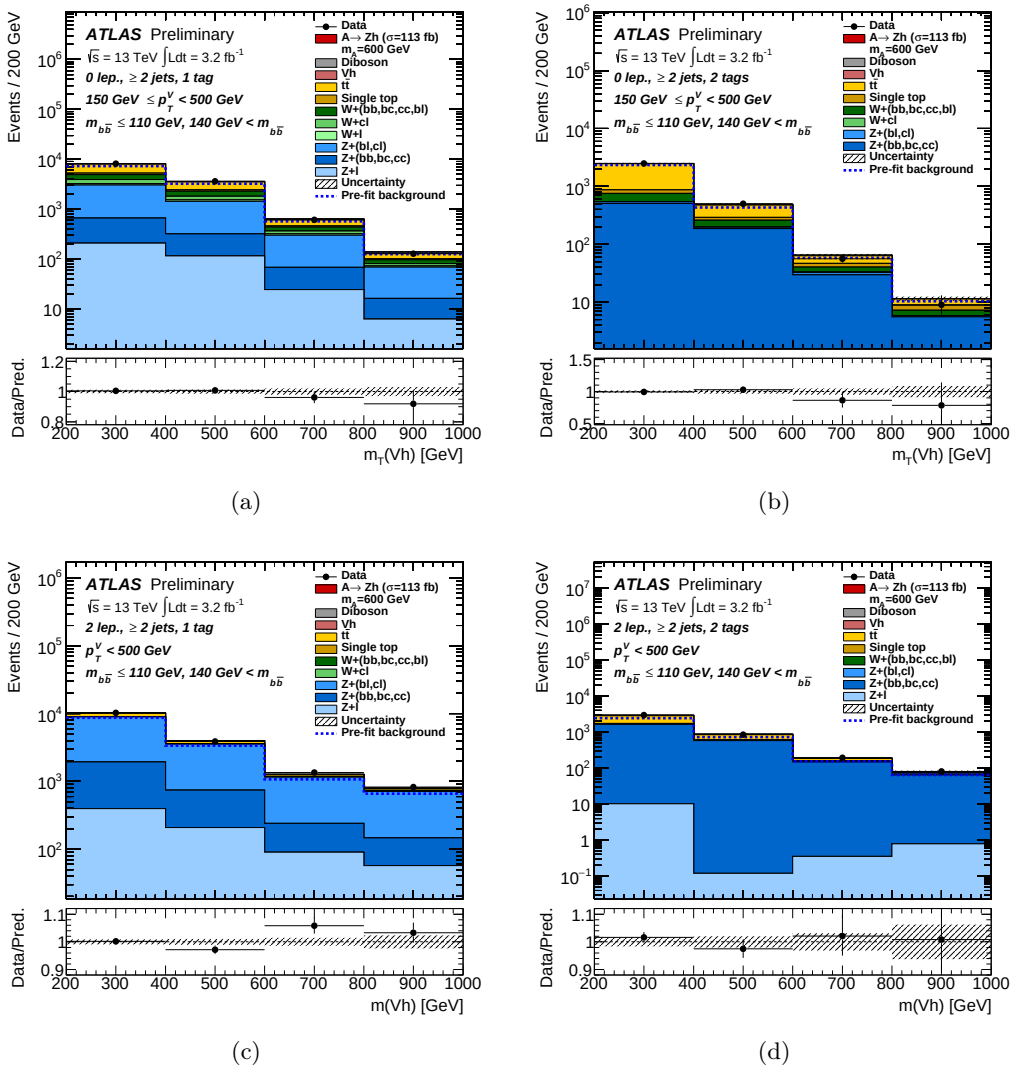
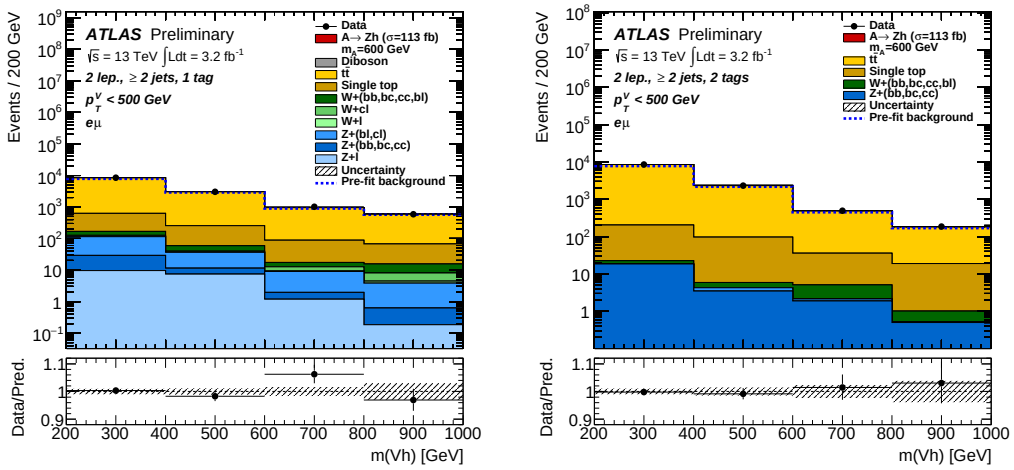


Figure 7.3: Distributions of m_T^{Zh} and m^{Zh} in the low-p_T^Z categories for the 0-lepton (top) and 2-lepton (bottom) channels '(W/Z) + jets' control regions. The left (right) column corresponds to the 1-tag (2-tag) control regions. The background expectation is shown after the maximum Likelihood fit to data; the total background expectation before the fit is denoted by the dotted blue line. The ratio of the data to background prediction is shown below each plot, and contains the total uncertainty band (black diagonal hashing) on the background prediction.



(a) (b)

Figure 7.4: Distributions of m_{Zh} in the $t\bar{t}$-enriched $e\mu$ control region. The left (right) column corresponds to the 1-tag (2-tag) regions. The background expectation is shown after the maximum likelihood fit to the data; the total background expectation before the fit is denoted by the dotted blue line. The ratio of the data to background prediction is shown below each plot, and contains the total uncertainty band (black diagonal hashing) on the background prediction.

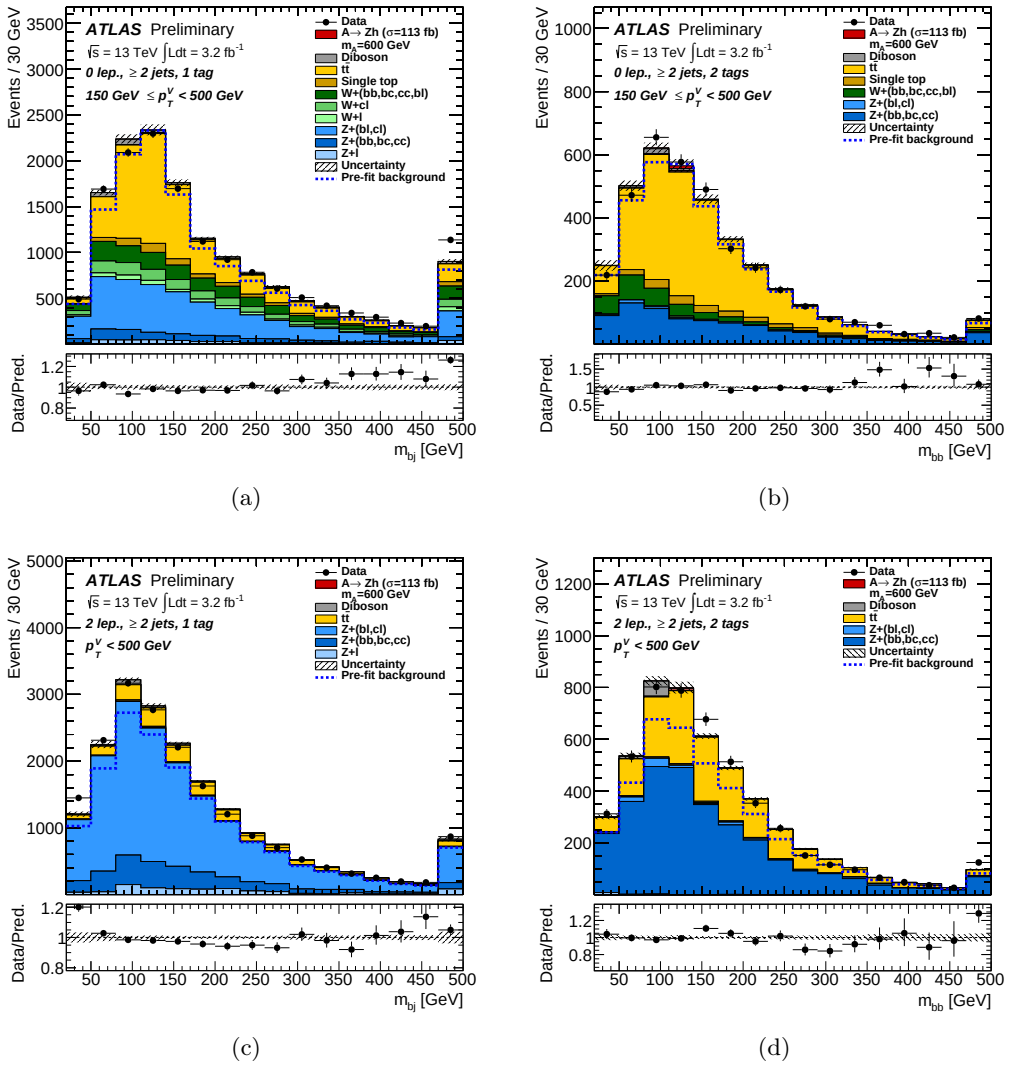
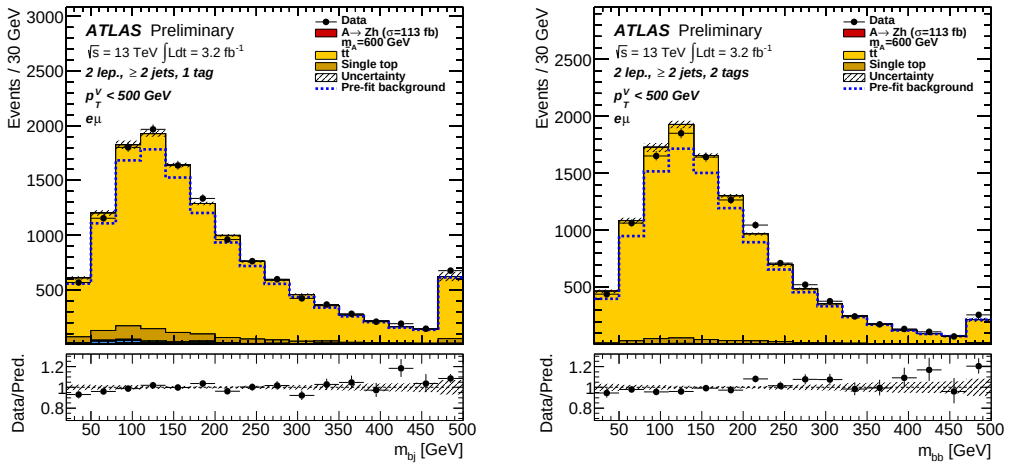


Figure 7.5: Distributions of the dijet-mass in the low-p_T^Z categories for the 0-lepton (top) and 2-lepton (bottom) channels, combining signal regions and '(W/Z) + jets' control regions. The left (right) column corresponds to the 1-tag (2-tag) control regions. The background expectation is shown after the maximum likelihood fit to the data; the total background expectation before the fit is denoted by the dotted blue line. The ratio of the data yield and background prediction is shown below each plot, and contains the total uncertainty band (black diagonal hashing) on the background prediction, excluding the uncertainties specifically derived on the shape of the m_T^{Zh} and m^{Zh} mass distribution for the (W/Z) + jets and the $t\bar{t}$ processes. Signal MC simulation samples are generated using the gluon fusion process. The right-most bin in each plot contains all events in and above that bin.



(a) (b)

Figure 7.6: Distributions of dijet-mass in the $t\bar{t}$-enriched $e\mu$ control region. The left (right) column corresponds to the 1-tag (2-tag) regions. The background expectation is shown after the maximum likelihood fit to the data; the total background expectation before the fit is denoted by the dotted blue line. The ratio of the data yield and background prediction is shown below each plot, and contains the total uncertainty band (black diagonal hashing) on the background prediction, excluding the uncertainties specifically derived on the shape of the m_T^{Zh} and m^{Zh} mass distribution for the $(W/Z)+$jets and the $t\bar{t}$ processes. The right-most bin in each plot contains all events in and above that bin.

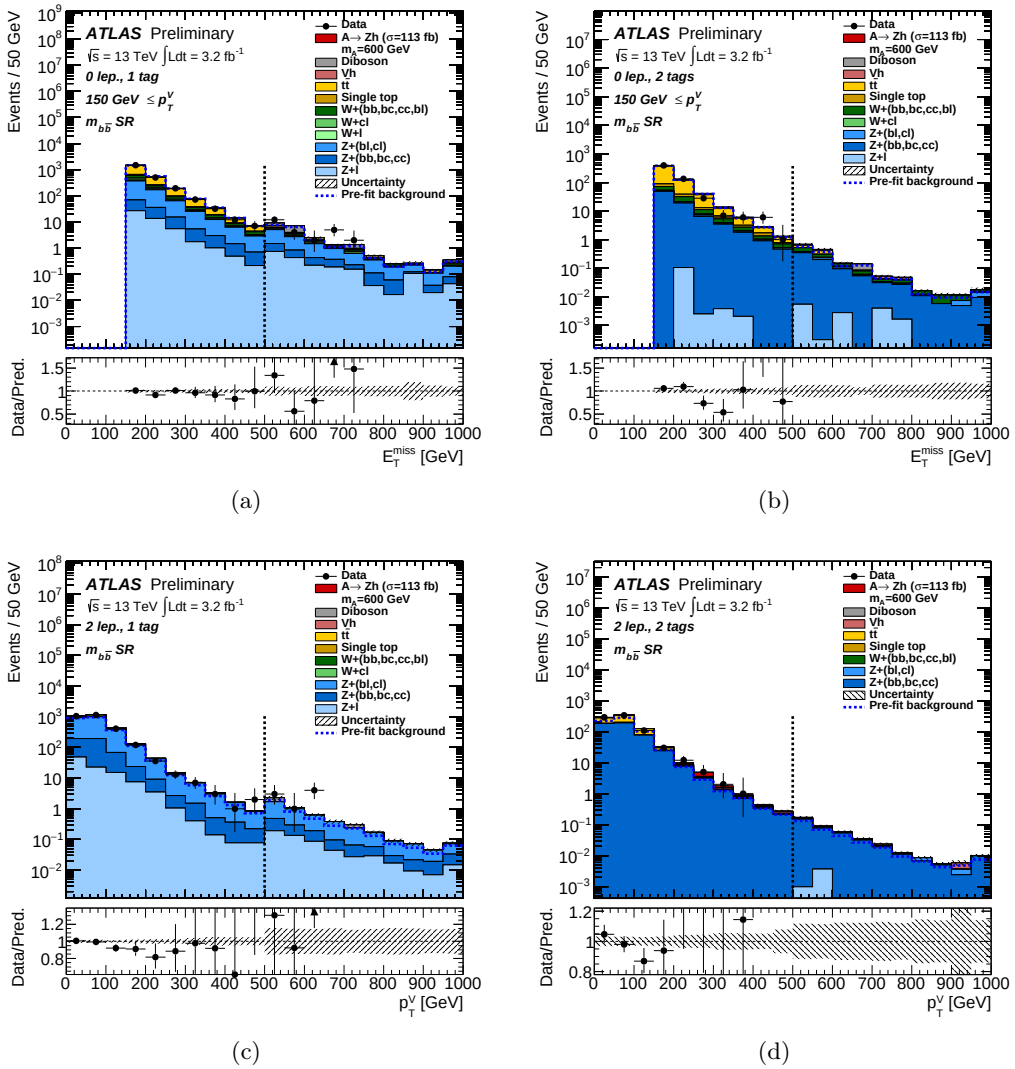


Figure 7.7: Distributions of p_T^Z for the 0-lepton (2-lepton) channels are in the top (bottom) row. The left (right) column corresponds to the 1-tag (2-tag) signal regions. The low-p_T^Z and the high-p_T^Z categories have been merged in a single histogram. The background expectation is shown after the maximum likelihood fit to the data; the total background expectation before the fit is denoted by the dotted blue line. The signal for $pp \to A \to Zh$ with $m_A = 600$ GeV is shown stacked on top of the background and normalised to 113 fb, corresponding to the 95% CL upper limit on $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$. The dotted black line corresponds to the separation between low- and high-p_T^Z regimes at $p_T^Z = 500$ GeV. The ratio of the data and background prediction is shown below each plot, and contains the total uncertainty band - except the uncertainties specifically derived on the shape of the m_T^{Zh} and m_{Zh} mass distribution for the $(W/Z)+$jets and the $t\bar{t}$ processes - (black diagonal hashing) on the background prediction. The right-most bin in each plot contains all events in and above that bin.

7.5 Systematic Uncertainties

In this section I describe the systematic uncertainties affecting this analysis, broadly categorized in experimental systematic uncertainties (related to the performance of the detector and the reconstruction of objects and events) and modeling systematic uncertainties (related to the accuracy of the MC modeling of backgrounds and signal).

7.5.1 Experimental systematic uncertainties

This set of systematics includes uncertainties affecting the trigger selection, the object reconstruction and identification, the object energy, momentum and mass calibrations and resolutions, for the different objects used in this analysis (charged leptons, hadronic jets, missing transverse momentum). The main details of these uncertainties are here summarized:

- **integrated luminosity and pile-up**: the uncertainty on the integrated luminosity amounts to 5%, and it is applied to the MC estimate of signal and backgrounds, determined following the same procedure outlined in [207]. An uncertainty on the modeling of additional interactions per bunch crossing is applied to the MC simulation by varying the average number of additional interactions by $O(10\%)$.

- **leptons**: systematic uncertainties affecting the lepton triggers, lepton reconstruction, identification and isolation efficiencies and the energy and resolution corrections are included, for both electrons and muons, with very low impact on the analysis performance.

- $\mathbf{E_T^{miss}}$ **trigger**: the trigger efficiency is corrected with scale factors (SFs) obtained from data-to-MC comparison for $W(\mu\nu)+$ jets events. Uncertainties on these scale factors account for statistical fluctuations in its determination, and differences in its measurement with alternative physics processes (comparing SFs obtained from $t\bar{t}$ and $Z(\nu\nu)+$ jets events).

- **small-R jets energy scale (JES) and energy resolution (JER)**: the systematic uncertainty on the JES is fully documented in [233] and includes several sources, estimated from data collected at 13 TeV. It is divided into several components covering in-situ analyses, η inter-calibration, high-p_T jets, pile-up effects, flavor composition, flavor response, b-jets specific effects and impact of punch-through jets. These different effects are treated by performing a principal component analysis, to decompose them into uncorrelated components that can be considered as independent sources, and combining the resulting variations in a set of three separated uncertainties, applied in this analysis. The JES uncertainty for central jets ($\eta = 0$) ranges from $O(6\%)$ for jets with $p_T \sim 20$ GeV, to $O(1\%)$ for jets with $p_T > 100$ GeV. A single JER systematic uncertainty is derived from data-to-MC comparisons, ranging from $O(4\%)$ for jets with $p_T \sim 20$ GeV, to $O(1\%)$ for jets with $p_T > 100$ GeV (in the central region $\eta = 0$).

- **large-R jets**: both JES and JER uncertainties are included in the analysis, along with systematic uncertainties on the mass scale and resolution, evaluated by comparing the ratio of calorimeter-based to track-based measurements in multi-jet data and simulation. These systematics have typical sizes of 2-20% for the large-R jet energy/mass scales [154].

- **flavor-tagging**: uncertainties related to the flavor-tagging algorithm applied to small-R and track jets cover different systematic effects. For both jets collections a set of data-to-MC SFs is applied, derived by measuring the tagging efficiency in data, depending on the jet kinematics (p_T and η): for c-jets and light-jets these measurements are performed with data collected during Run-1 (with Run-1 to Run-2 extrapolation uncertainties), while for b-jets the Run-2 data are used with $t\bar{t}$ enriched samples. The uncertainties on these SFs encode several effects from their measurements in data, and result in 40 separate components per flavor (b-, c-, τ- and light-jets). Following the approach used for the JES, this number is reduced via a principal component analysis, resulting in 3 separate uncertainties for b-jets, 4 for c-jets and 5 for light-jets, which are applied in this analysis. Two additional systematic uncertainties are included, related to the extrapolation of the SFs to different jet p_T regimes and to the extrapolation from charm to bottom quark. These uncertainties are treated as fully correlated between small-R and track jets.

- $\mathbf{E_T^{miss}}$: the uncertainties on the jet energy scale are propagated to the E_T^{miss} calculation. Furthermore uncertainties on the determination of the E_T^{miss} *soft-term* are included, covering its calibration and resolution.

7.5.2 MC modeling systematic uncertainties

The dominant background contributions for this search come from $t\bar{t}$ and $(W/Z)+$jets events (mostly $Z+$jets): for these processes dedicated systematic uncertainties on the accuracy of the MC estimate and modeling have been derived. The remaining backgrounds (single-top production, semileptonic diboson processes) have contributions of the order of few percent: for these processes the systematic uncertainties derived for the SM $Vh(b\bar{b})$ search presented in Chapter 6 have been applied without modifications, and their impact on the final result is very small. Finally $Vh(b\bar{b})$ processes are treated as background in this search, applying an overall normalization uncertainty of 50%.
In this section I describe in detail the procedure applied to estimate the MC modeling systematics for $t\bar{t}$ and $(W/Z)+$jets backgrounds and for the $A \to Zh$ signal processes.
The strategy adopted to derive systematic uncertainties for the two main backgrounds is the following: for each process the total normalization is determined by the fit to data, either as a unconstrained nuisance parameter (for the largest background components) or with specific priors. Furthermore systematic uncertainties on the relative variation of the background acceptance across analysis categories are derived by comparing the nominal MC prediction to alternative generators and/or setups. The complete set of nuisance parameters thus derived provides the Likelihood fit with all the necessary degrees of freedom to control the background prediction and adjust its modeling using the information extracted from data. The systematic uncertainties on the normalization of the $t\bar{t}$ and $(W/Z)+$jets backgrounds are summarized in Tables 7.3 and 7.4.

Top-pair production systematic uncertainties The overall normalization of the $t\bar{t}$ background is free to float in the Likelihood fit to data, separately for 0- and 2-lepton channels, in order to fully exploit the constraining power of data from signal and control regions. Additional modeling systematic uncertainties are estimated via the comparison of the nominal POWHEG prediction to the alternative MADGRAPH5_aMC@NLO MC generator, and to POWHEG samples with increased or reduced production of additional radiation. All the alternative samples are generated applying the same requirement on the leptonic decay of the W boson as the nominal one, and undergo the baseline analysis selec-

tion described in Section 7.3. The details of these alternative samples are given in Table 7.2 (where `hdamp` is the POWHEG resummation damping parameter, the transverse momentum scale around which the Sudakov resummation becomes unimportant and can be damped. At this scale the pure NLO matrix element provides a good description, meaning that smaller values of hdamp correspond to a larger suppression of the hardest emission cross section). An uncertainty of 30% is quoted on the relative normalization of the high-p_T^Z and

Generator	Setup Details	Systematic Effect
MADGRAPH5_aMC@NLO	UE-EE-5 tune	hard scatter
+HERWIG ++ [234]	CT10(ME) & CTEQ6L1(PS)	generation and matching
	lepton filter	
POWHEG +PYTHIA 6	nominal setup	*low variation* for additional
	scale variations low ($\mu_R = \mu_F = 2$)	radiation
	`hdamp` $= m_{top}$	
	low radiation PERUGIA2012 tune variation	
POWHEG +PYTHIA 6	nominal setup	*high variation* for additional
	scale variations high ($\mu_R = \mu_F = 0.5$)	radiation
	`hdamp` $= 2 \times m_{top}$	
	high radiation PERUGIA2012 tune variation	

Table 7.2: $t\bar{t}$ MC samples generated for the estimate of systematic uncertainties.

low-p_T^Z regions, and uncertainties in the range of 2.5%-10% on the relative normalization of m_{dijet} signal region to the side-bands control regions: these uncertainties are derived and applied in the analysis separately for 0-lepton and 2-lepton $t\bar{t}$ events, consistently with the overall normalization. Finally an uncertainty of 2.5% is assigned to the relative normalization of the $t\bar{t}$ dedicated control region and the remaining 2-lepton channel. These uncertainties are obtained from the quadrature sum of $t\bar{t}$ acceptance variations between nominal and alternative MC predictions, without including possible differences in the total $t\bar{t}$ cross section. This scheme of normalization uncertainties applied for the $t\bar{t}$ background is reported in Table 7.3.

Additional systematic uncertainties on the shape of the discriminating variables m_T^{Zh} and m^{Zh} are extracted from the same MC comparisons. Separate uncertainties are derived from the comparison of the nominal POWHEG prediction to MADGRAPH5_aMC@NLO, and to the alternative POWHEG samples (treated as correlated), separately for 0- and 2-lepton events, by fitting the ratio between different predictions with a first order polynomial function.
Finally a systematic uncertainty on the shape of the transverse momentum of the top-quark candidates p_T(top) is included: this systematics is taken unmodified from the SM $VH(b\bar{b})$ analysis described in Chapter 6, and it is detailed in Section 6.6.3. Although no striking mismodeling in the p_T(top) distribution has been observed in this search, discrepancies have been found in the analysis of the LHC Run-1 data: since the MC prediction has not changed significantly between the two analyses, it was decided to provide the Likelihood fit with some degrees of freedom to modify the $t\bar{t}$ description according to this variable.

(W/Z) + jets systematic uncertainties Systematic uncertainties on the modeling of these processes are considered separately for the different sub-components: $Z + (bb, bc, cc)$, $Z + (bl, cl)$, $Z + l$, $W + (bb, bc, cc, bl)$, $W + cl$ and $W + l$. The uncertainties are derived by

Lepton channel	Uncertainty	Value
0-lepton 2-lepton	overall normalization	floating
0-lepton 2-lepton	high-to-low p_T^Z	30%
0-lepton 2-lepton	m_{dijet} SR-to-CR	12% 9%
2-lepton	$t\bar{t}$ CR / 2-lepton SR	2.5%

Table 7.3: Summary of the normalization systematic uncertainties for the $t\bar{t}$ background.

comparing the nominal SHERPA MC prediction to alternative MC samples. Two approaches are considered:

1. comparison between the baseline SHERPA MC and alternative MADGRAPH5_AMC@NLO +PYTHIA 8 samples: this comparison is performed applying the full event selection detailed in Section 7.3. The main drawback of this approach is the low statistics of the MADGRAPH5_AMC@NLO samples (generated without hadron-flavor or p_T^V filters) in the analysis signal regions.

2. comparison between the baseline SHERPA MC and alternative SHERPA samples produced with factorization, renormalization and resummazion scales either doubled or halved, or matching scale varied in the range 15 GeV - 30 GeV: these alternative samples are produced at *particle-level*, without any detector simulation, therefore they cannot undergo the baseline event selection from Section 7.3. This comparison is thus performed by implementing an event selection based on particle-level reconstructed quantities through the RIVET [235] software, closely mirroring the baseline selection.

The overall normalization of the largest contributions, $Z + (bb, bc, cc)$ and $Z + (bl, cl)$, is unconstrained in the Likelihood fit to data, while overall normalization uncertainties are assigned to the remaining components to constrain their values in the fit. These normalization uncertainties correspond to 26% for $Z + l$, 30% for $W + (bb, bc, cc, bl)$, 30% for $W + cl$ and 10% for $W + l$ processes, and encode both theoretical uncertainties in the cross section calculations used to normalize these samples [191] and acceptance variations from the particle-level comparisons of the nominal SHERPA prediction to the alternative SHERPA samples produced with varied scales.

Additional uncertainties are quoted on the relative variation of the background contributions across analysis categories and flavor components. An uncertainty of 13% (35%) is assigned to the relative acceptance variation of $Z + cl$ to $Z + bl$ ($W + bl$ to $W + bb$) from the comparison of the alternative SHERPA samples.
Furthermore uncertainties of 10%-20%-30% are quoted on the relative acceptance variation between low- and high-p_T^Z categories, respectively for $Z + (bb, bc, cc)$, $Z + (bl, cl)$ and $Z + l$ components; and corresponding uncertainties of 11%-24%-33%, respectively for $W + (bb, bc, cc, bl)$, $W + cl$ and $W + l$ components. These systematic uncertainties are obtained from the comparison of the baseline SHERPA prediction to the alternative

MADGRAPH5_aMC@NLO samples. From the same comparison, uncertainties of 17%-12%-9% are derived on the relative acceptance variation between 0- and 2-lepton channel for $Z+(bb,bc,cc)$, $Z+(bl,cl)$ and $Z+l$ components (while corresponding uncertainties for the $W+$jets background are not included, since its contribution in the 2-lepton channels is negligible).
All these relative normalization uncertainties include only variation effects on the acceptance of the backgrounds, with no contribution from possible differences in the cross section calculation, and are outlined in Table 7.4.
Finally shape systematic uncertainties for the discriminating variables m_T^{Zh} and m^{Zh} are extracted from the comparison between SHERPA and MADGRAPH5_aMC@NLO: these shape variations are derived separately for all six $(W/Z)+$jets background components, and are estimated and applied separately in the m_{dijet} signal region and the side-bands control regions, in order to cover relative acceptance variations across these categories, as well as shape effects (the impact of these variations is in the range 1%-13%).

$A \to Zh$ signal systematic uncertainties Several systematic effects are considered when studying the modeling of the signal processes, and they are all combined in a single systematic uncertainty on the total signal normalization, since no striking shape effect is observed for the main discriminating variables. The sources of systematic effects include: choice of the PDFs, with an impact of 1%-4% estimated by changing the baseline NNPDF2.3LO PDFs to the MSTW2008LO set; independent variations by a factor of two of the renormalization and factorization scales, with an impact of 1%; choice of the parton shower and underlying event tune, with an impact of 1% to 4%.
These uncertainties are found to be fairly consistent across the m_A mass range considered in this analysis, thus no dependence on the mass of the A boson is introduced.
No modeling uncertainties are considered for the alternative signal samples used to test the $b\bar{b}A$ production hypothesis.

Uncertainty	$Z+(bb,bc,cc)$	$Z+(bl,cl)$	$Z+l$
Overall normalization	floating	floating	26%
high-to-low p_T^Z	10%	20%	30%
0-to-2 lepton	17%	12%	9%
$Z+cl/Z+bl$	-	13%	-
Systematic	$W+(bb,bc,cc,bl)$	$W+cl$	$W+l$
Overall normalization	30%	30%	10%
high-to-low p_T^Z	11%	24%	33%
$W+bl/W+bb$	35%	-	-

Table 7.4: Summary of the normalization systematic uncertainties on the (W/Z)+jets backgrounds

7.6 Statistical Analysis

The statistical approach chosen to extract the final results of this search is presented in its core principles in Chapter 4, and it is consistent with the one adopted for the other analyses presented in this thesis: the measured data are fitted using templates for the

expected signal and background contributions with a binned Likelihood fit approach, which combine simultaneously all the signal and control regions of the analysis. This analysis uses the test statistics specifically defined for upper limits setting $\tilde{q_\mu}$ (rather than the one defined for the discovery of a positive signal q_0), detailed in Section 4.1.
In this section I present an overview of the main features of this fit, summarizing all signal and control regions with their role in the Likelihood fit, while the main results of the search are presented in Section 7.7.

Signal regions Eight different signal regions (SRs), introduced in Section 7.3, are included in the Likelihood fit, separately for 0- and 2-lepton channels, 1-tag and 2-tag categories, low- and high-p_T^Z regimes. The split in 1-tag and 2-tag categories is introduced to provide additional constraints on the background flavor composition, and to recover signal-like events in the high-p_T^Z regime where one of the b-tagged jets is not properly tagged due to the proximity of the selected jets (as a result of the large boost of the h candidate). The efficiency of these regions in collecting signal events is quoted in Section 7.3.
The distribution of the discriminating variables m_T^{Zh} and m^{Zh} in these signal regions is shown in Figure 7.8 and Figure 7.9, respectively for low- and high-p_T^Z events (the background expectation is shown after the binned maximum likelihood fit to the data). It is interesting to note how the high-p_T^Z selection (including the $p_T^Z > 500$ GeV requirement) greatly depletes the number of selected events in data, resulting in zero selected data events in the 2-tag high-p_T^Z signal regions shown in Figure 7.9(b) and 7.9(d).

Control regions Six separate control regions (CRs) are included in the Likelihood fit: four regions are defined as '(W/Z)+jets CRs' by selecting events outside of the m_{dijet} mass window cuts: low-m_{dijet} and high-m_{dijet} events are merged together, while 1-tag and 2-tag categories are kept separate, as well as lepton channels. The background composition of these CRs is clearly shown in the distributions of the m_{dijet} variable shown in Figure 7.5. Two additional regions are defined as '$t\bar{t}$ CRs' as detailed in Section 7.4, from a loosened 2-lepton selection with different flavor leptons, providing strong constraint to the modeling of the $t\bar{t}$ background processes. The distributions of the m_{dijet} variable for these CRs is shown in Figure 7.6.

These six control regions are included only for the low-p_T^Z regime since the background contribution is severely depleted after the high-p_T^Z selection, hence including corresponding CRs in the merged regime does not lead to any improvement in the Likelihood fit. It is therefore crucial to remark that the background modeling and normalization are mostly controlled by the fit of the low-p_T^Z categories (where the background contribution is sizable) and propagated to the high-p_T^Z regions through the overall normalization parameters for the different background processes. To handle this correlation, additional parameters are included for the main backgrounds ($t\bar{t}$ and (W/Z) + jets) controlling the relative normalization between low- and high-p_T^Z categories, as detailed in Section 7.4.

Discriminating variables and binning The variables included in the Likelihood fit for both signal and control regions are the transverse mass m_T^{Zh} and the invariant mass m^{Zh}, respectively for the 0- and 2-lepton channels, introduced in Section 7.3.
Given the different statistical power of low- and high-p_T^Z categories, different binning choices for the discriminating variables are adopted. For high-p_T^Z regions (in both lepton channels) the mass distributions have three equal-size bins from 500 GeV to 2 TeV: the low number of bins is chosen to cope with the very low statistics of these categories while retaining

some information from the shape of the discriminating variable.
For low-p_T^Z signal regions in the 0-lepton channel the m_T^{Zh} distribution contains 8 equal-size bins of 100 GeV, from 200 GeV to 1 TeV; while in the 2-lepton channel variable-width bins are used, with broader bins for larger m^{Zh} masses, in order to reflect the change in background contribution and signal resolution for different masses, exploiting the better mass resolution of this lepton channel. For both lepton channels the template distributions included for the different control regions have equal-size bins of 200 GeV, from 200 GeV to 1 TeV. In all cases events outside the presented range are included in under/over-flow bins.
Statistical uncertainties due to the limited size of simulated background samples are included as nuisance parameters with Poisson parametrization, assigning one NP to each bin of the fitted mass distributions if the relative statistical uncertainty on the total background is above 5%.

Understanding the $AZh(b\bar{b})$ Likelihood fit model The Likelihood fit model and its results are studied and checked by performing the different tests presented in Section 4.2.2, considering the impact of the fit to data on the NPs in terms of central values, uncertainties and correlations arising during the fit itself. In addition, the ranking of the NPs according to their effect on the measured signal strength $\hat{\mu}$ is considered as a figure of merit to determine the most important NPs in the analysis (in this search the signal strength is simply the scaling factor for the expected signal rate, including the branching ratios of the Z and the h bosons). No pathological behavior is found from these tests, showing the stability of the Likelihood fit and the extracted results.

The dominant systematic uncertainties in the Likelihood fit depend on the signal mass hypothesis m_A tested, as fits to lower masses are more sensitive to the low-p_T^Z categories, while fits to larger masses strongly depend on the high-p_T^Z regions. When testing signals with a mass of $m_A = 500$ GeV the highest ranked systematic uncertainties are related to the modeling of $Z + (bb, bc, cc)$ and $Z + (bl, cl)$ processes in the signal regions and to the experimental uncertainties on the b-tagging scale factors, the total integrated luminosity and the JES.
Considering different m_A mass hypothesis it is found that the most important experimental systematic uncertainties are associated with the calibration and resolution of the small- and large-R jet energy, the calibration and resolution of the large-R jet mass (for high-p_T^Z categories) and the determination of the jet b-tagging efficiency and mistag rate. The highest ranked modeling systematic uncertainties are always related to the background estimate for $Z+$jets and $t\bar{t}$ events, and play a stronger role when fitting lower m_A values. Statistical fluctuations on the estimate of systematic uncertainties are treated by applying the smoothing procedure detailed in Section 4.2.1, while the total number of systematic variation template pairs ($+1\sigma$ and -1σ) and normalization NPs is reduced with the pruning methods described in the same Section, in order to eliminate unnecessary parameters and limit the fit complexity.

A crucial message from the Likelihood fit results comes from the total uncertainty on the cross section for the $A \to Zh$ signal processes, after the maximum Likelihood fit to data. For a signal with $m_A = 600$ GeV, after the Likelihood fit, the dominant contribution to this uncertainty comes from the limited number of events in data, which accounts for almost 80% of the cross section uncertainty, while the limited statistics of MC simulated samples, the modeling systematics and the experimental uncertainties play a much smaller

role.

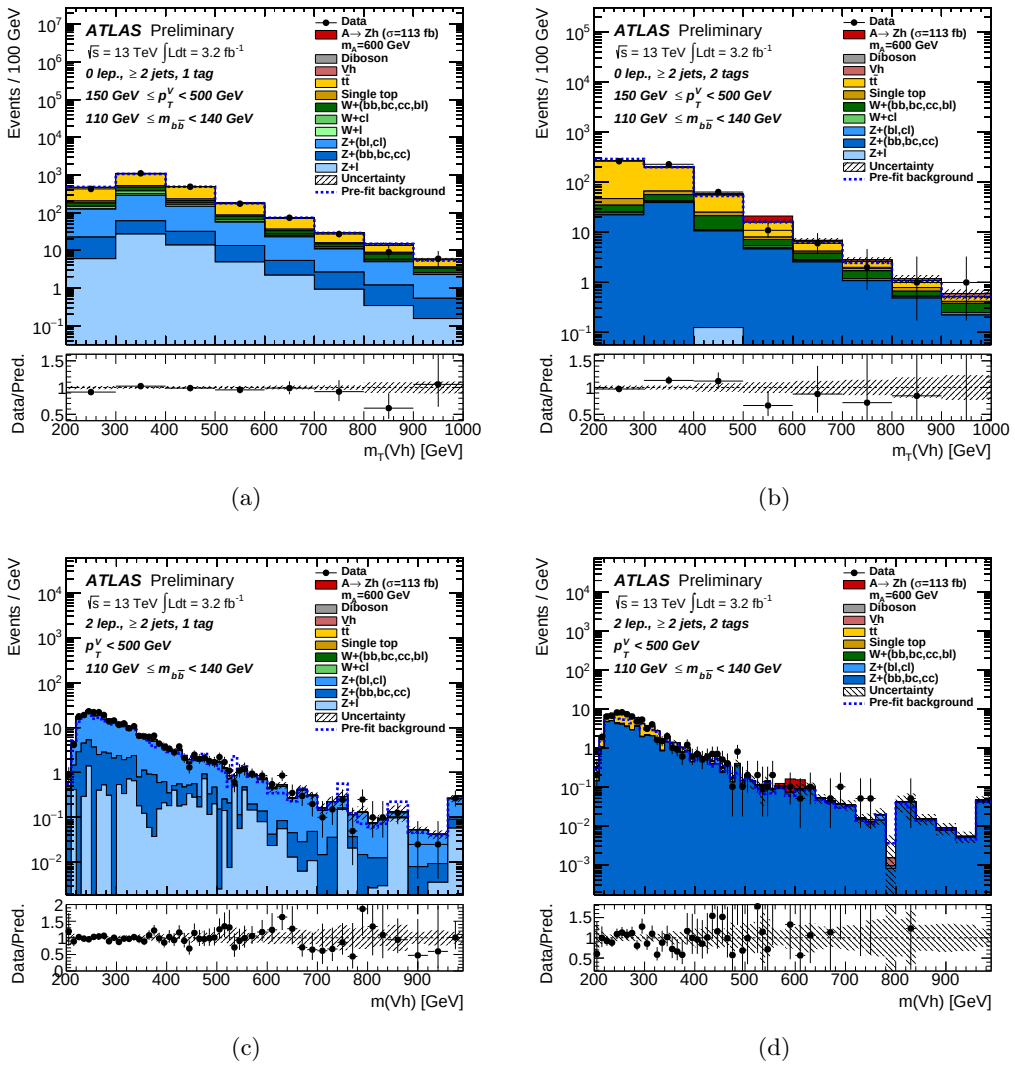


Figure 7.8: Distributions of m_T^{Zh} and m^{Zh} in the low-p_T^Z categories for the 0-lepton (2-lepton) channels are in the top (bottom) row. The left (right) column corresponds to the 1-tag (2-tag) signal regions. The background expectation is shown after the maximum likelihood fit to the data; the total background expectation before the fit is denoted by the dotted blue line. The signal for $gg \to A \to Zh$ with $m_A = 600$ GeV is shown stacked on top of the background and normalized to 113 fb, corresponding to the 95% CL upper limit on $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$. The ratio of the data and background prediction is shown below each plot, and contains the total uncertainty band (black diagonal hashing) on the background prediction. The right-most bin in each plot contains all events in and above that bin. In (c) and (d) the bin widths vary with m_A.

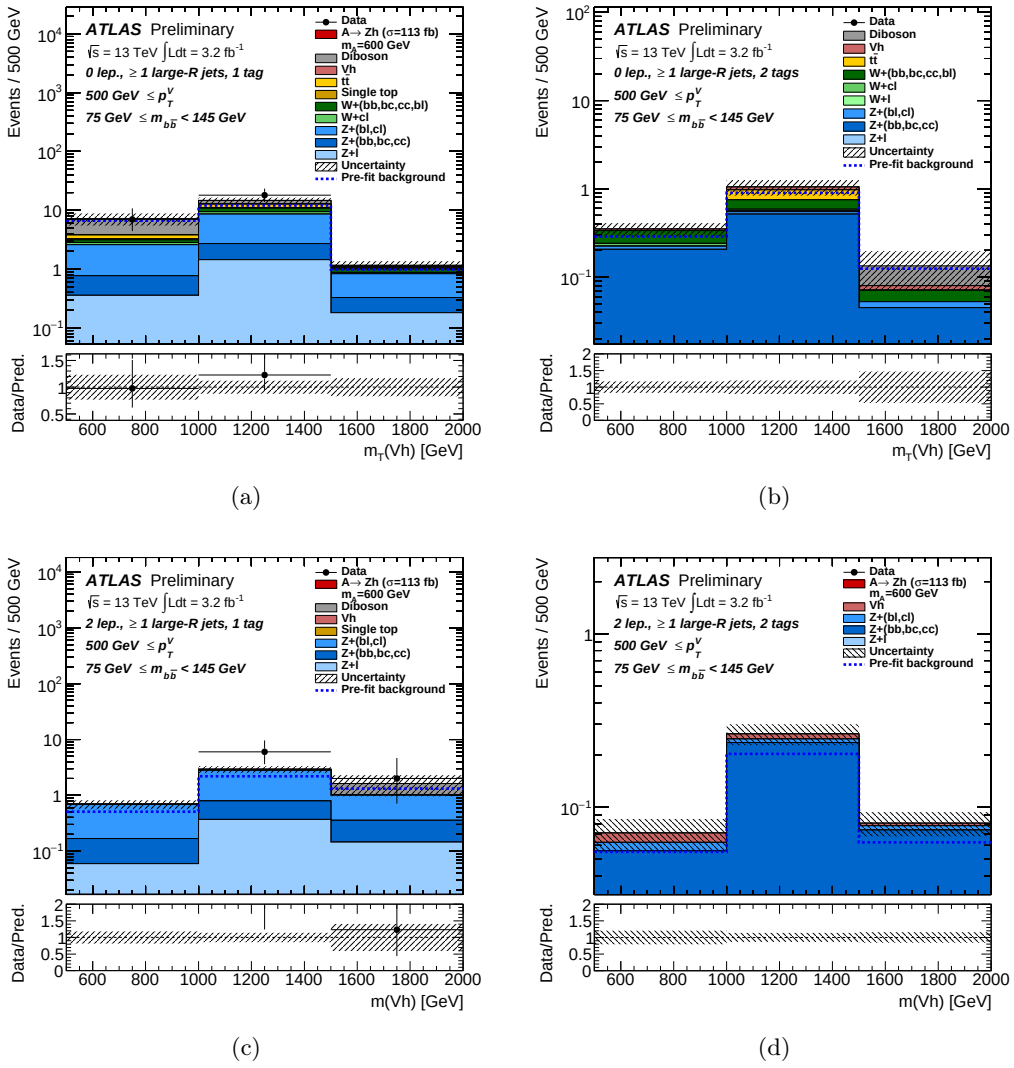


Figure 7.9: Distributions of m_T^{Zh} and m^{Zh} in the high-p_T^Z categories for the 0-lepton (2-lepton) channels are in the top (bottom) row. The left (right) column corresponds to the 1-tag (2-tag) signal regions. The background expectation is shown after the maximum likelihood fit to the data; the total background expectation before the fit is denoted by the dotted blue line. The signal for $gg \to A \to Zh$ with $m_A = 600$ GeV is shown stacked on top of the background and normalized to 113 fb, corresponding to the 95% CL upper limit on $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$. The ratio of the data and background prediction is shown below each plot, and contains the total uncertainty band (black diagonal hashing) on the background prediction. The right-most bin in each plot contains all events in and above that bin.

7.7 Results and Interpretation

The results of this search are obtained from the combined Likelihood fit to data described in Section 7.6 performed for values of the A boson mass in the range 220 GeV - 2 TeV: distributions of the discriminating variables m_T^{Zh} and m^{Zh} for events passing the signal regions selections are shown in Figure 7.8 and Figure 7.9, for the low- and high-p_T^Z categories respectively. The background expectation is shown in these figures after the binned Likelihood fit to data: no significant excess above the expectation from the background prediction is observed in data. Two upward deviations from the background-only hypothesis occur at masses $m_A = 260$ GeV and 440 GeV, each corresponding to a local significance of about 2σ. The expected and observed numbers of events in each signal region (after the Likelihood fit to data) are shown in Table 7.10: the numbers predicted for the $A \to Zh$ signal process assume $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b}) = \sigma(Ab\bar{b}) \times BR(A \to Zh) \times BR(h \to b\bar{b}) = 113$ fb, for $m_A = 600$ GeV. This value of the cross section is chosen because it corresponds to the 95% CL upper limit on the $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$ process at this mass, from the Profile Likelihood Ratio fit.

Since no excess is observed, exclusion limits are quoted using the modified frequentist method named CLs, with the test statistics $\tilde{q_\mu}$ defined in Section 4.1.
Figure 7.11 shows the 95% confidence level (CL) upper limits on the product of production cross section ($pp \to A$) and branching fractions ($A \to Zh$ and $h \to b\bar{b}$) as a function of the resonance mass m_A, for the different production mechanisms tested: pure gluon-fusion and pure b-quark associated production. The same limits are shown super-imposed in Figure 7.12 for a more direct comparison, while the exclusion limits obtained for an equal mixture of gluon-fusion and b-quark-associated production signals are presented in Figure 7.13.
The 95% confidence level upper limits on the triple branching fractions, $\sigma(pp \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$, are set in the range of [4.0, 0.017] pb ([6.9, 0.026] pb) for $m_A = [220, 2000]$ GeV assuming gluon-fusion (b-quark-associated) exclusive production.

In order to show the different features of the sensitivity of the 0- and 2-lepton channels, the expected upper limits for $gg \to A \to Zh \to (l^+l^-, \nu\bar{\nu})b\bar{b}$ are presented separately for the two lepton channels in Figure 7.11(a). For $m_A < 800$ GeV the 2-lepton channel shows the strongest contribution, mainly thanks to the better mass resolution of the m^{Zh} variable, critical in the low-mass region where the contribution of EW backgrounds is sizable. On the contrary the 0-lepton channel is dominant for $m_A > 800$ GeV, due to the largest branching ratio of the Z decay to neutrinos compared to the decay to electron and muon pairs: the larger signal yield has indeed a stronger impact than a finer mass resolution, in a region where the backgrounds are almost completely suppressed.

Finally Figure 7.14 shows the expected upper limits for $gg \to A \to Zh \to (l^+l^-, \nu\bar{\nu})b\bar{b}$ as obtained from the separate standalone fits of the low-p_T^Z and high-p_T^Z categories (combining 0- and 2-lepton channels), compared to the expected combined upper limits. This result clearly shows how the high-p_T^Z categories are dominant for large values of the A boson mass, above 1100 GeV.
For $m_A > 800$ GeV the results presented here in terms of expected limits improve upon the results presented by the ATLAS Collaboration in [13].

Interpretation in the context of the Type-I and Type-II 2HDMs The results presented in this section can be interpreted within the framework of the different 2HDMs introduced in Section 2.5 to impose limits on some combination of parameters of such models. In detail, Type-I and Type-II 2HD models are considered, in order to quote exclusion limits on the $\tan\beta$ and $\cos(\beta-\alpha)$ parameter plane: 95% CL exclusion regions in this plane are shown for $m_A = 600$ GeV in Figure 7.15.
The width of the A boson is assumed to be narrow in the MC simulation, but it is a parameter which depends on the values of $\tan\beta$ and $\cos(\beta-\alpha)$ tested. To properly account for this dependence the A boson width is corrected to the value Γ_A predicted by the tested 2HDMs for each point in the $\tan\beta$ and $\cos(\beta-\alpha)$ parameter space, by using a relativistic Breit-Wigner function. This correction is an approximation which holds for Γ_A width smaller than 5% of the mass m_A, therefore only points of the parameter space where $\Gamma_A/m_A < 5\%$ are considered in this interpretation (points outside this validity region correspond to the gray areas in Figure 7.15, which are thus not tested here).

One b-tag				
	0-leptons low-p_{T}^{Z}	0-leptons high-p_{T}^{Z}	2-leptons low-p_{T}^{Z}	2-leptons high-p_{T}^{Z}
$Z + l$	55 ± 31	2.0 ± 1.0	118 ± 38	0.57 ± 0.28
$Z + (cl, bl)$	518 ± 54	8.2 ± 1.8	1943 ± 65	3.18 ± 0.65
$Z + (bb, bc, cc)$	82 ± 13	1.82 ± 0.35	391 ± 23	0.74 ± 0.13
$W + l$	48 ± 22	–	–	–
$W + cl$	111 ± 42	1.17 ± 0.45	0.67 ± 0.28	–
$W + (bb, bc, cc, bl)$	185 ± 71	1.80 ± 0.63	10.1 ± 4.2	–
$t\bar{t}$	1202 ± 77	2.05 ± 0.72	276 ± 22	–
single top	99 ±10	0.49 ± 0.11	15.3 ± 1.6	–
diboson	27.2 ± 4.6	5.3 ± 1.0	26.3 ± 6.1	0.67 ± 0.32
$(W/Z)h$	3.3 ± 1.6	0.16 ± 0.08	3.7 ± 1.8	0.04 ± 0.02
total background	2332 ± 45	23.0 ± 2.6	2784 ± 47	5.2 ± 0.9
expected $A \to Zh$ (gluon-fusion)	5.00 ± 0.41	–	1.73 ± 0.12	–
expected $A \to Zh$ (b-quark-associated)	3.05 ± 0.26	–	1.01 ± 0.08	–
data	2295	25	2769	8
Two b-tags				
	0-leptons low-p_{T}^{Z}	0-leptons high-p_{T}^{Z}	2-leptons low-p_{T}^{Z}	2-leptons high-p_{T}^{Z}
$Z + l$	0.13 ± 0.44	0.01 ± 0.01	–	–
$Z + (cl, bl)$	4.7 ± 1.8	0.06 ± 0.03	9.6 ± 4.0	0.02 ± 0.01
$Z + (bb, bc, cc)$	81 ± 13	0.75 ± 0.15	490 ± 22	0.36 ± 0.05
$W + l$	–	0.02 ± 0.01	–	–
$W + cl$	3.6 ± 2.1	0.04 ± 0.02	–	–
$W + (bb, bc, cc, bl)$	37 ± 14	0.28 ± 0.09	0.42 ± 0.20	–
$t\bar{t}$	392 ± 24	0.22 ± 0.11	284 ± 22	–
single top	27.8 ± 2.9	–	6.39 ± 0.60	–
diboson	6.1 ± 3.1	0.05 ± 0.03	0.57 ± 0.23	–
$(W/Z)h$	5.4 ± 2.6	0.10 ± 0.05	7.6 ± 3.7	0.03 ± 0.01
total background	557 ± 18	1.53 ± 0.25	799 ± 23	0.42 ± 0.05
expected $A \to Zh$ (gluon-fusion)	9.21 ± 0.76	–	3.65 ± 0.29	–
expected $A \to Zh$ (b-quark-associated)	5.85 ± 0.47	–	2.21 ± 0.17	–
data	577	0	788	0

Figure 7.10: The numbers of expected and observed events for all the signal regions. The expectation is shown after the maximum likelihood fit to the data. The quoted uncertainties are the combined systematic and statistical uncertainties. The numbers predicted for a $A \to Zh$ signal process in gluon-fusion and b-quark-associated production assume that $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b}) = \sigma(pp \to b\bar{b}A) \times BR(A \to Zh) \times BR(h \to b\bar{b}) = 113$ fb, for $m_A = 600$ GeV. This cross section corresponds to the 95% CL upper limit on these effective cross sections at this mass. A dash indicates a negligible yield of events.

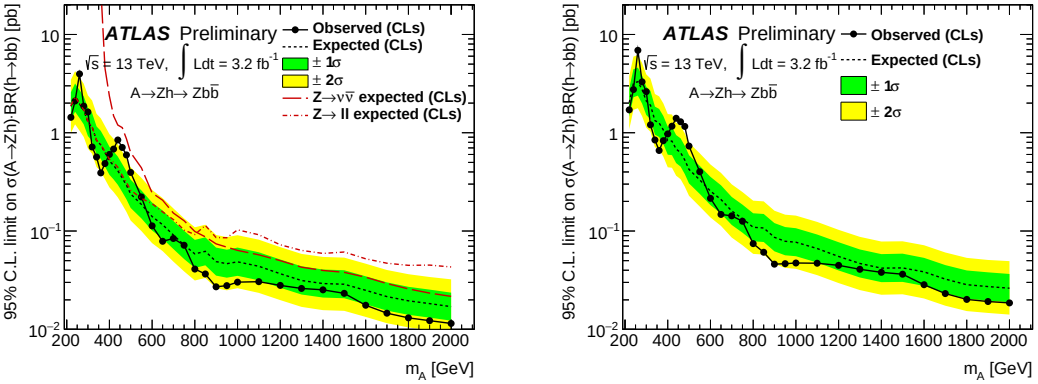


Figure 7.11: Upper limits at the 95% CL on the product of the production cross section for $pp \to A$ and the branching ratios for $A \to Zh$ and $h \to b\bar{b}$ evaluated by combining the 0-lepton and 2-lepton channels. The possible signal components of the data are interpreted using the following production mechanism hypotheses: Assuming (a) pure gluon-fusion production, and (b) pure b-quark-associated production. The green (yellow) band indicates the 1σ (2σ) statistical and systematic uncertainty on the expected exclusion limit (dotted black line). The solid black line is the observed exclusion limit. The long-dashed (fine-dotted) red line indicates the limit for the 0-lepton (2-lepton) analysis alone. The small increase in the expected limit for $A \to Zh \to l^+l^-b\bar{b}$ with m_A around 850 GeV is due to statistical uncertainties on the background estimation, while for m_A around 1 TeV the effect is due to the transition from the low-p_T^Z category to the high-p_T^Z category.

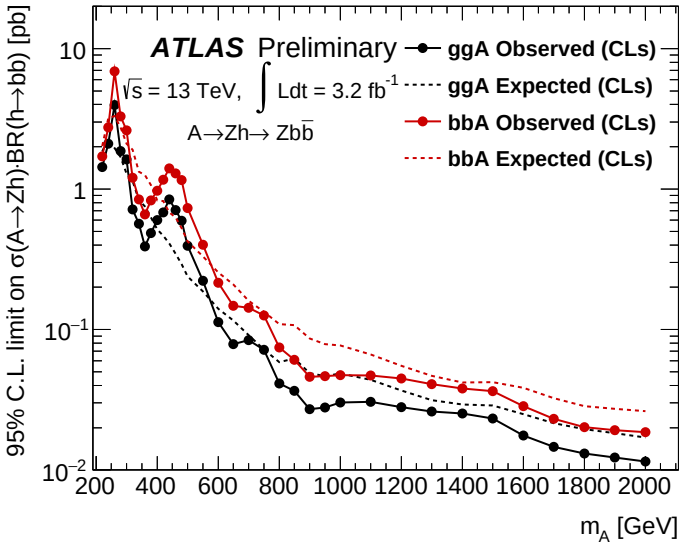


Figure 7.12: Upper limits at the 95% CL on $\sigma(pp \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$ evaluated by combining the 0-lepton and 2-lepton channels. The possible signal components of the data are interpreted using an equal gluon-fusion production (black) and b-quark-associated production (red). The dotted lines indicate the expected exclusion limits. The solid lines are the observed exclusion limits.

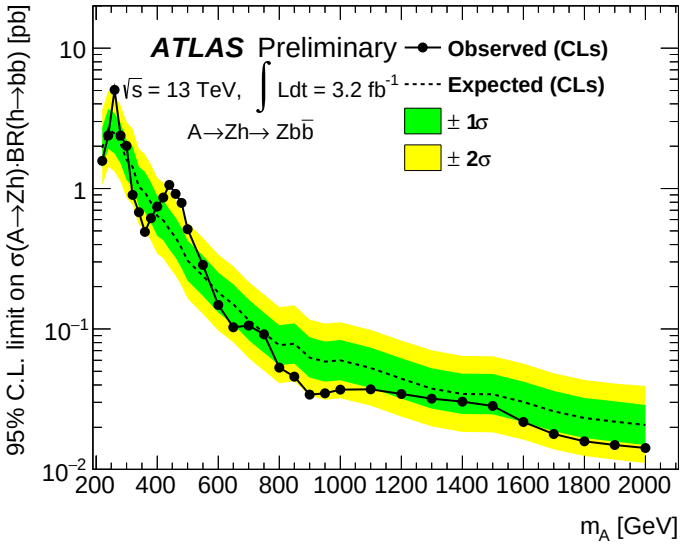


Figure 7.13: Upper limits at the 95% CL on $\sigma(pp \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$ evaluated by combining the 0-lepton and 2-lepton channels. The possible signal components of the data are interpreted using an equal mixture of gluon-fusion and b-quark-associated production. The green (yellow) band indicates the 1σ (2σ) statistical and systematic uncertainty on the expected exclusion limit (dotted black line). The solid black line is the observed exclusion limit.

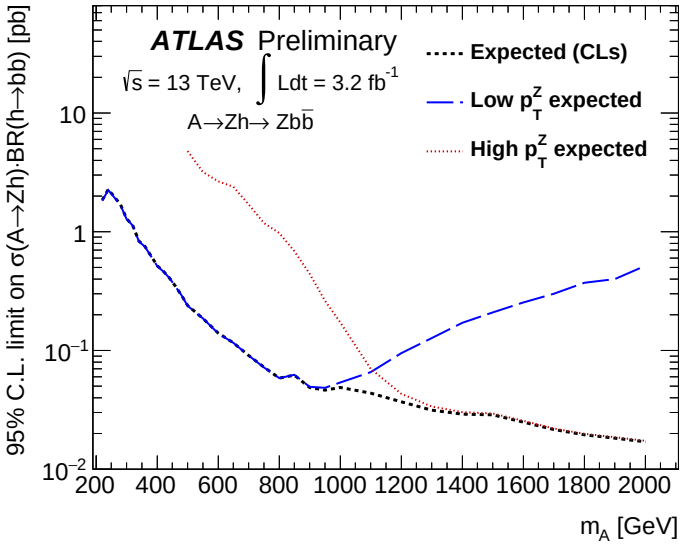
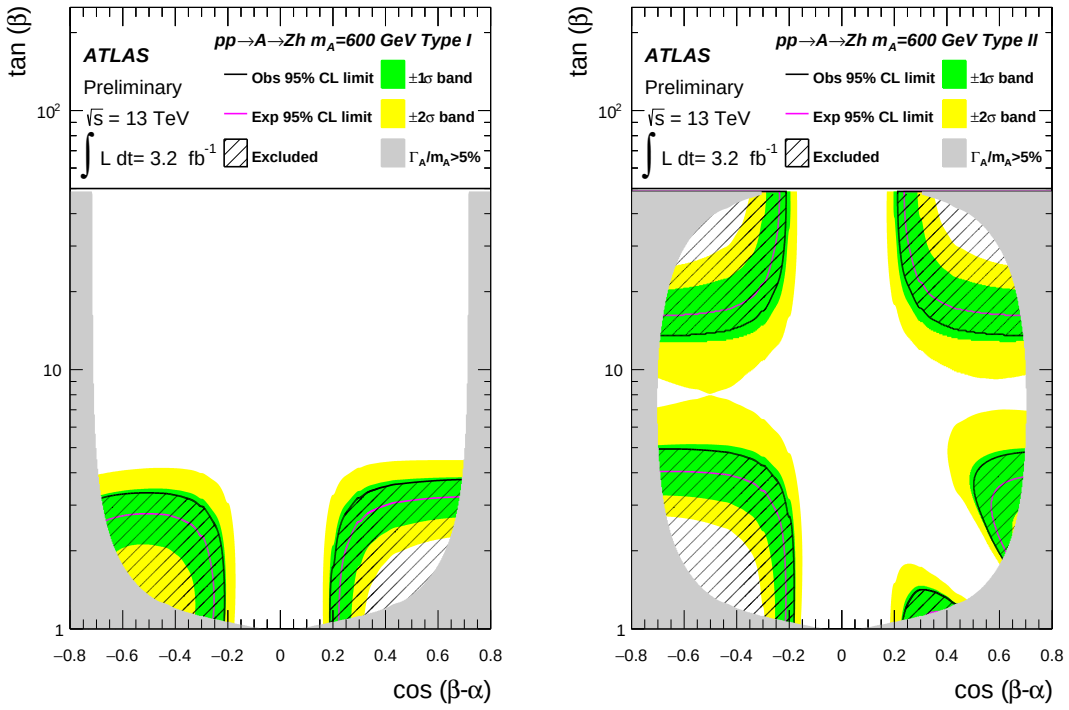


Figure 7.14: Expected upper limits at the 95% CL on $\sigma(gg \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$ evaluated by combining the 0-lepton and 2-lepton channels. The long-dashed line is the upper limit expected from the low-p_T^Z analysis, while the fine-dotted line is that from the high-p_T^Z analysis. The remaining dotted line is the combination of the two analysis channels.



(a) (b)

Figure 7.15: The interpretation of the cross section limits in the context of a Type-I and Type-II 2HDMs as a function of the parameters $\tan\beta$ and $\cos(\beta-\alpha)$ for $m_A = 600$ GeV. Variations of the natural width up to $\Gamma_A/m_A = 5\%$ and different mixtures of gluon-fusion and b-quark-associated production are taken into account. Only points in parameter space where $\Gamma_A/m_A < 5\%$ are considered.

8 | Search for Standard Model $VH(b\bar{b})$: Run-2 data

8.1 Introduction: The Analysis Strategy

In this Chapter I describe the search for the Standard Model Higgs boson in associated production $(W/Z)H$ in the decay channel to a pair of b-quarks, performed with data collected by the ATLAS detector during the LHC Run-2 (until July 2016) at a center of mass energy of $\sqrt{s} = 13$ TeV [236].
This is the first analysis targeting the Higgs $VH(b\bar{b})$ performed with data collected at $\sqrt{s} = 13$ TeV. It starts from the experience gathered from the same analysis performed with the LHC Run-1 data, documented in Chapter 6, and from the search described in Chapter 7 which, although targeting a different type of process, has been critical to begin the understanding of the LHC Run-2 data in the final states here considered.

The main analysis strategy is shaped in its main features on the mold of the $VH(b\bar{b})$ analysis of LHC Run-1 dataset: in this Chapter I give a full description of the analysis, with particular highlights on the updates and changes that have been introduced with respect to Chapter 6.
The dataset and the simulated Monte Carlo samples used for this search are described in Section 8.2. The object and event reconstruction was already detailed in Chapter 5 and thus not repeated here, while the event selection and categorization are described in Section 8.3.
Following the successful experience of the analysis at 8 TeV, the multivariate approach is adopted, by building an MVA whose details are given in Section 8.4. The background composition and modeling, although similar to what observed from Run-1 data, show some important differences. The multijet background is one of the main actors in the SM analysis (contrary to the search described in Chapter 7 where the 1-lepton final state is not included) and its determination from $\sqrt{s} = 13$ TeV data, along with details of the backgrounds and signal estimate, is outlined in Section 8.5.
The experimental systematic uncertainties relevant in this analysis are quite similar to the ones described in Chapter 7, with a few differences. The modeling uncertainties however have been carefully studied and re-derived here. Section 8.6 provides full details on the estimate of the systematic uncertainties included in the analysis.
The results of this search are obtained with a statistical procedure whose main principles are described in the dedicated Chapter 4: in Section 8.7 I provide a detailed discussion of the implementation and the results of this statistical methodology, which leads to the final results presented in Section 8.8.

8.2 Data and Simulated Samples

Dataset This analysis is performed with data collected by the ATLAS detector at a center of mass energy of $\sqrt{s} = 13$ TeV during both the 2015 and the early 2016 LHC running periods, corresponding to integrated luminosities of 3.2 ± 0.1 fb^{-1} and 10 ± 0.4 fb^{-1} respectively [207]. Running conditions changed between 2015 and 2016, and typically events from the combined dataset have a pile-up between 10 and 30. Only data collected in ideal operating conditions of all ATLAS sub-detectors, and passing tight quality controls from the ATLAS Data Quality Group, are used. Cleaning criteria similar to the ones detailed in Section 6.2 are implemented.

Trigger selection The trigger selection follows closely the strategy adopted for the analysis of 2015 data described in Chapter 7: 0-lepton events are collected with E_T^{miss} trigger with online thresholds of 70 GeV (for data recorded in 2015) and 90 GeV (for data recorded in 2016, due to the higher trigger rate).
Events in the 1-lepton electron channel are recorded by at least one of the lowest unprescaled single-electron triggers, with thresholds of 24, 60 and 120 GeV for the 2015 data, and an increased threshold of 140 GeV for the highest threshold in 2016 data; an isolation requirement is included for the lowest threshold trigger, removed for the medium threshold trigger, while the identification requirements are relaxed for the highest threshold trigger. In the 1-lepton muon channel only the the E_T^{miss} trigger is used, as in the 0-lepton channel, providing good efficiency for high p_T muons (the acceptance gain from the inclusion of the lowest unprescaled single-muon trigger has been tested, and it is found to be very small). In the 2-lepton channel events are triggered by at least one of the lowest unprescaled single-electron (the same used for the 1-lepton channel) or single-muon triggers (with p_T threshold of 24 GeV and 40 GeV for the 2015 data, of 24 and 50 GeV for the 2016 data), where one of the offline-reconstructed leptons selected in the analysis has to match the object that satisfied the trigger. The lowest threshold triggers include an isolation requirement which is removed for the higher threshold trigger.

Monte Carlo samples The signals and most of the background processes are predicted with Monte Carlo (MC) simulations, described in this section in detail. The simulation is typically scaled to the best available theoretical prediction for its cross section (even for backgrounds that are left free to float in the Likelihood fit). The MC events are obtained from different generators, and they are processed with a `GEANT4`-based simulation of the ATLAS detector and the standard ATLAS reconstruction software.

The $qq \to VH$ SM signals (both in the W and Z channels) are simulated with the Pythia 8 MC generator at LO(QCD) with leading-log (LL) parton shower, using the A14 tune and NNPDF23LO PDF set. The gluon-induced $gg \to ZH$ production is obtained from the Powheg generator interfaced to the Pythia 8 parton shower, using the AZNLO tune [237] and CT10NLO PDF set. In the simulation the mass of the Higgs boson is fixed at 125 GeV (no mass scan is considered in this analysis), while the Higgs branching ratio is set to BR($H \to b\bar{b}$)=58%. The total cross section used to normalize the Pythia 8 samples for qq-induced processes is computed at NNLO(QCD) and NLO(EW) [60, 61, 59, 52], subtracting the gg-induced contribution to avoid double-counting it. The gg-induced cross section is calculated at NLO+NLL(QCD) [62, 52, 66], and it is used to normalize the Powheg $gg \to ZH$ events.
Top-quark pair, single-top and semileptonic diboson productions are simulated with the

same generators and similar settings as what described in Section 7.2: the only difference concerns the diboson samples for which the clustering algorithm has been improved, bringing the performance of SHERPA 2.1.0 closer to the improved SHERPA 2.2.0 version.
The modeling of (W/Z) + jets processes is provided by the SHERPA 2.2.0 generator: the same generator parameters and filters (both for hadronic flavors and vector boson transverse momentum) as the SHERPA 2.1.0 MC described in Section 7.2 are considered. A theory-based reweighting correction function of the jet multiplicity distribution is applied at event level (where jets are reconstructed from MC simulated events before applying detector reconstruction): this is required to correct the (W/Z) + jets samples generated using a simplified scale setting prescription (to improve the generation speed), to match the prediction obtained from the strict scale prescription. These events are scaled to their NNLO(QCD) cross sections.

The cross sections of all background and signal processes correspond to the numbers already quoted in Table 7.1, with the only exception of the $gg \to ZH$ signal, whose calculation was improved to include NLL(QCD) effects. This change gives an increase of the $gg \to ZH$ cross section of 13%.
All MC samples include the effect of multiple pp interactions from the same and neighboring bunch-crossing by overlaying simulated minimum-bias events on each generated event. These pile-up interactions are obtained from soft QCD processes with PYTHIA 8, using the A2 tune and MSTW2008LO PDFs. The simulated samples (except those generated with SHERPA) exploit the `EvtGen1.2.0` program to describe properties of the bottom and charm hadron decays.

8.3 Event selection

The details of object and event reconstruction were already outlined in Chapter 5, while in this Section I describe the event selection criteria that define this analysis. The events are categorized following the same scheme used for the analysis of 8 TeV data presented in Chapter 6, with the most important differences highlighted in this Section; the selection cuts employed to reject the backgrounds are also based on the experience gathered with the Run-1 analysis, and are optimized to disentangle the signal from $VH(b\bar{b})$ Higgs production (with $m_H = 125$ GeV) from the Standard Model backgrounds.

- **lepton channels**: 0-, 1- and 2-lepton channels are defined requiring respectively exactly 0 loose leptons, 1 tight lepton or 2 loose leptons (of which at least one of medium quality) with same flavor and, in the muon sub-channel, opposite charge. The lepton selection criteria are defined in Section 5.1.2.

- **jets selection and categories**: hadronic jets used in this analysis are the small-R jets defined in Section 5.1.3, and are divided into 'central' jets ($p_T > 20$ GeV and $|\eta| < 2.5$) and 'forward' jets ($p_T > 30$ GeV and $2.5 < |\eta| < 4.5$). Only events with at least two central jets are retained, and exactly two of them must satisfy the b-tagging requirement (these b-tagged dijet pair is used to build the Higgs candidate). The b-tagged jet with the hardest p_T (leading) is required to have $p_T > 45$ GeV. Events are furthermore categorized according to the jet multiplicity (counting both central and forward jets): in the 0- and 1-lepton channels only events with exactly two and exactly three jets are considered, forming the exclusive 2-jet and 3-jet bins. Larger jet multiplicities are vetoed to suppress the $t\bar{t}$ production background. In the 2-lepton channel events with more than three jets are included in the so-called 3-jet bin (which

becomes effectively an inclusive $\geq$ 3-jet bin), leading to an increase in significance of 5-6%. Hereafter, the $\geq$ 3-jet bin of the 2-lepton channel is denoted as '3-jet bin', for consistency with the naming convention adopted for the 0- and 1-lepton channels.

With respect to the $VH(b\bar{b})$ search presented in Chapter 6 it is worth noting some significant differences in the selection and tagging criteria applied to the hadronic jets.
First, no 1-tag category is included in the analysis. This choice is the result of extensive studies that keep into account different factors: no calibration is available for pseudo-continuous distributions of the MV2c b-tagging algorithm output used in this search, thus the b-tagging weight distribution cannot be used as a discriminating variable; the background modeling is estimated to be sufficiently under control without the contribution of 1-tag events in the Likelihood fit (no striking mismodeling for $(W/Z)+$jets processes is observed in $\sqrt{s} = 13$ TeV data); the flavor composition of the backgrounds is more under control thanks to the improved performance of the MV2c algorithm compared to the MV1c algorithm (with a c-jets rejection improved by more than a factor of two, for the same 70% b-tagging efficiency). Finally, the inclusion of the 1-tag category has a relatively small impact of $O(3\%)$ on the expected sensitivity while adding complexity to the fit model.
The second major difference is the inclusion of events with > 3 jets in the 2-lepton channel: while the $t\bar{t}$ background increases significantly from the Run-1 to the Run-2 center of mass energy, it is sufficiently under control (and its suppression also benefits from the improved c-jets rejection) to make this change beneficial in terms analysis sensitivity, improving the expected significance of the 2-lepton channel alone by $O(10\%)$.
Finally, the 2-tag region is not divided into sub-categories according to the b-tagging purity: not having a pseudo-continuous calibration for the MV2c algorithm means that only inclusive b-tagging operating points are calibrated and usable in the analysis, therefore events cannot be categorized in exclusive b-tagging bins within the 2-tag region.

- **Vector boson transverse momentum p_T^V categories**: The p_T^V categorization has the goal of separating high-sensitivity regions (at high p_T^V) from lower-sensitivity categories which are mainly used to constrain the background modeling, thanks to their larger statistical power.
  In the 0-lepton channel the p_T^Z is defined as the reconstructed E_T^{miss}; in 1-lepton events p_T^W is equal to the vector sum of the selected p_T and the E_T^{miss} contribution; in the 2-lepton channel p_T^Z corresponds to the vector sum of the transverse momenta of the two selected leptons.
  In the 0- and 1-lepton channels a single p_T^V region is included, requiring events with vector boson transverse momentum above 150 GeV. Events in the 2-lepton channel are divided into low- and high-p_T^V, with boundary at 150 GeV. A finer split in p_T^V, such as the one adopted in the analysis of Run-1 data, doesn't bring significant improvement given the lower statistics available for this search. While in the 0-lepton channel the p_T^V cut at 150 GeV is justified by the threshold of the E_T^{miss} trigger, in the 1-lepton channel we could in principle include events with lower p_T^W: the low-p_T^W region is not considered in this search to avoid the harsh difficulties related to the modeling and estimate of the multijet background, which contributes significantly in this region, in view of a moderate impact on the analysis sensitivity (of the order of 5-10%).

Topological and kinematic selection After the event categorization just described, an additional set of criteria specific to each lepton channel is applied, in order to reduce the background contribution. A schematic outline of all cuts is given in Table 8.1, along with some details of the previous trigger selection and event categorization. Most of these criteria are inherited from the analysis presented in Chapter 6, with differences arising from the changes in detector conditions, center of mass energy and background composition.

- **0-lepton channel**: a cut on the scalar sum of the p_T of the two (three) selected jets (S_T) is set at 120 (150) GeV to avoid regions with poor modeling of the trigger efficiency. To suppress the multijet background below the $O(1\%)$ level, a set of criteria is applied on angular quantities: events with small azimuthal angle between the E_T^{miss} and the p_T^{miss}, and large azimuthal angle between the E_T^{miss} and the closest jet are retained, since they are likely to contain large genuine missing transverse energy, not originated by object mismeasurements or misidentifications. Furthermore events with large azimuthal angle between the E_T^{miss} and the Higgs candidate, and small azimuthal angle between the b-tagged jets, are selected, to target the typical $VH(b\bar{b})$ signature.

- **1-lepton channel**: the only channel-specific selection is applied in the electron channel, where events with missing transverse energy below 30 GeV are rejected, to reduce the multijet background contribution.

- **2-lepton channel**: the invariant mass of the dilepton system is selected within the window 71 GeV $< m_{ll} <$ 121 GeV, to be consistent with the Z boson mass. The width of this window is relatively large since the m_{ll} variable is also input to the multivariate discriminant, and its asymmetry takes into account the corresponding asymmetry in the multijet background contribution. The kinematic fit technique described in Section 5.2.2 is used in the 2-lepton channel of this analysis.

Selection	0-lepton	1-lepton	2-lepton
Trigger	$E_{\mathrm{T}}^{\mathrm{miss}}$	$E_{\mathrm{T}}^{\mathrm{miss}}$ (μ sub-channel)	
		Lowest unprescaled single lepton	
Leptons	0 loose lepton	1 tight lepton	2 loose leptons ($\geq$ 1 medium lepton)
Lepton pair	-	-	Same flavour opposite-charge for $\mu\mu$
$E_{\mathrm{T}}^{\mathrm{miss}}$	$>$ 150 GeV	$>$ 30 GeV (e sub-channel)	-
m_{ll}	-	-	$71 < m_{ll} < 121$ GeV
S_{T}	$>$ 120 (2 jets), $>$150 GeV (3 jets)	-	-
Jets	Exactly 2 or 3 signal jets		Exactly 2 or $\geq$ 3 signal jets
b-jets	2 b-tagged signal jets		
Leading jet p_{T}	$>$ 45 GeV		
$\min\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet})$	$> 20°$	-	-
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, h)$	$> 120°$	-	-
$\Delta\phi$(jet1,jet2)	$< 140°$	-	-
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, E_{\mathrm{T},trk}^{\mathrm{miss}})$	$< 90°$	-	-
p_{T}^{V} regions	[0, 150] GeV (2-lepton), [150, ∞] GeV		

Figure 8.1: Summary of the event selection in the 0-, 1- and 2-lepton channels.

8.4 Multivariate Analysis: Boosted Decision Trees

As this search follows the path traced by the analysis of data collected during the LHC Run-1, a multivariate approach is adopted to enhance the sensitivity by combining the discriminating power of several variables. The main principles that motivate this choice have been introduced in Section 6.4: the MVA (BDTs) used in this analysis is almost identical to the one described there, with a few distinct aspects that arise from further optimization targeting the search at $\sqrt{s} = 13$ TeV. In this Section I do not repeat a full description of the BDTs, already provided in Section 6.4, but I highlight the differences introduced for this analysis.
Separate BDTs are trained for each lepton channel and in the 2- and 3-jet regions, since the background composition differs considerably across them, and also for the low- and high-p_T^Z categories of the 2-lepton channel, resulting in eight different BDTs trained. Table 8.2 outlines schematically the set of categories included in the analysis with the different BDTs employed in each one of them.
In the training of the BDTs, the truth-tagging method described in Section 5.2.4 is applied to all backgrounds in order to retain the largest available statistics in the 2-tag category.

The input variables combined in the BDTs include all the variables introduced in Section 6.4 (Table 6.4), with two notable differences. Since pseudo-continuous b-tagging calibration is not available, the MV2c output weights cannot be used as discriminating variables in the BDTs, thus they are not considered. Furthermore, two variables are added for the 1-lepton channel, to improve the rejection of the $t\bar{t}$ background: the rapidity difference between the W and the H boson $|\Delta Y(W,H)|$, and the mass of the top quark candidate, reconstructed from the (lepton, neutrino, b-jet) system assuming that each event is obtained from top-quark pair production. In order to build these variables a complete knowledge of the event kinematic (and thus an estimate of the 4-momentum of the neutrino from the W decay) is required. The transverse components of the neutrino 4-momentum are assumed to be equal to the reconstructed E_T^{miss} transverse components, while the longitudinal component p_z^ν is determined, up to a possible two-fold ambiguity, by constraining the mass of the (lepton, neutrino) system to be consistent with the W boson mass. The top quark can be built from the reconstructed W boson and one of the b-jets, choosing p_z^ν and the b-jet so that m_{top} is minimized. The addition of specific variables that improve the $t\bar{t}$ rejection is particularly important since the production cross section for this background shows a larger increase when moving from $\sqrt{s} = 8$ TeV up to 13 TeV, compared to the signal and the remaining background.
The complete list of input variables included in the BDTs is given in Table 8.3, which can be compared with Table 6.4 for an overview of the differences between the setups of the MVAs used for the analysis of Run-1 and Run-2 data.

As introduced in Section 6.4, a proper and well understood modeling of the input variables is critical for the BDTs performances, as well as careful checks of the correlations among the variables themselves and between each of them and the output of the BDTs. In order to gain confidence in the modeling of the different input variables provided by the background prediction, their distributions are compared to data, to spot potential discrepancies. Figures 8.4, 8.5 and 8.6 show the distribution of several key variables for the different lepton channels, across analysis categories. No striking mismodeling is observed, and the prediction from MC simulation is well behaved when compared to data. Correlation tests across variables have been performed, without showing any pathological behavior. Note that in

these plots, and in the rest of the analysis, the truth-tagging procedure is only applied to the $V + cl$, $V + l$ and WW samples, which are the most statistically depleted by the b-tagging requirements.

Details of the BDTs setup inherited from the Run-1 analysis The BDTs definition follows closely the one described in Section 6.4. The following aspects have not been modified: in order to control the overtraining of the MVA the BDTs are 'cross-trained' by splitting the sample of training events in two sub-samples, so that the training on the first sub-sample can be tested on the second sub-sample, and viceversa.
The output distribution of the BDTs weights is transformed to modify its binning in order to maximize the sensitivity of the variable, coping with the available statistics: the same rebinning transformation adopted for the Run-1 analysis is used (with parameters $z_s = z_b = 10$, re-optimized for the different statistics).
Alternative BDTs are trained for the cross-check analysis of semileptonic diboson processes $VZ(b\bar{b})$: no change in the BDTs setup is introduced, and the same output transformation is employed for VH and VZ MVAs. The BDTs output variables are here noted as BDT_{VH} and BDT_{VZ}, when the MVA is trained to discriminate against VH or VZ signals respectively (since no Higgs mass scan is performed in this search, a single set of Higgs BDTs is trained for an Higgs boson mass of $m_H = 125$ GeV).

Channel	Categories					
	2 b-tagged jets					
	$p_{\mathrm{T}}^V < 150$ GeV			$p_{\mathrm{T}}^V > 150$ GeV		
	2 jets	3 jets	≥ 3 jets	2 jets	3 jets	≥ 3 jets
0 lepton	-	-	-	BDT	BDT	-
1 lepton	-	-	-	BDT	BDT	-
2 lepton	BDT	-	BDT	BDT	-	BDT

Figure 8.2: The distributions used in the global likelihood fit for all the categories in each channel.

Variable	0-lepton	1-lepton	2-lepton
p_{T}^{V}		×	×
$E_{\mathrm{T}}^{\mathrm{miss}}$	×	×	×
$p_{\mathrm{T}}^{b_1}$	×	×	×
$p_{\mathrm{T}}^{b_2}$	×	×	×
m_{bb}	×	×	×
$\Delta R(b_1, b_2)$	×	×	×
$\lvert\Delta\eta(b_1, b_2)\rvert$	×		×
$\Delta\phi(V, bb)$	×	×	×
$\lvert\Delta\eta(V, bb)\rvert$			×
H_{T}	×		
$\min[\Delta\phi(\ell, b)]$		×	
m_{T}^{W}		×	
m_{ll}			×
m_{Top}		×	
$\lvert\Delta Y(V, H)\rvert$		×	
	Only in 3-jet events		
$p_{\mathrm{T}}^{\mathrm{jet}_3}$	×	×	×
m_{bbj}	×	×	×

Figure 8.3: Variables used in the multivariate analysis for the 0-, 1- and 2-lepton channels.

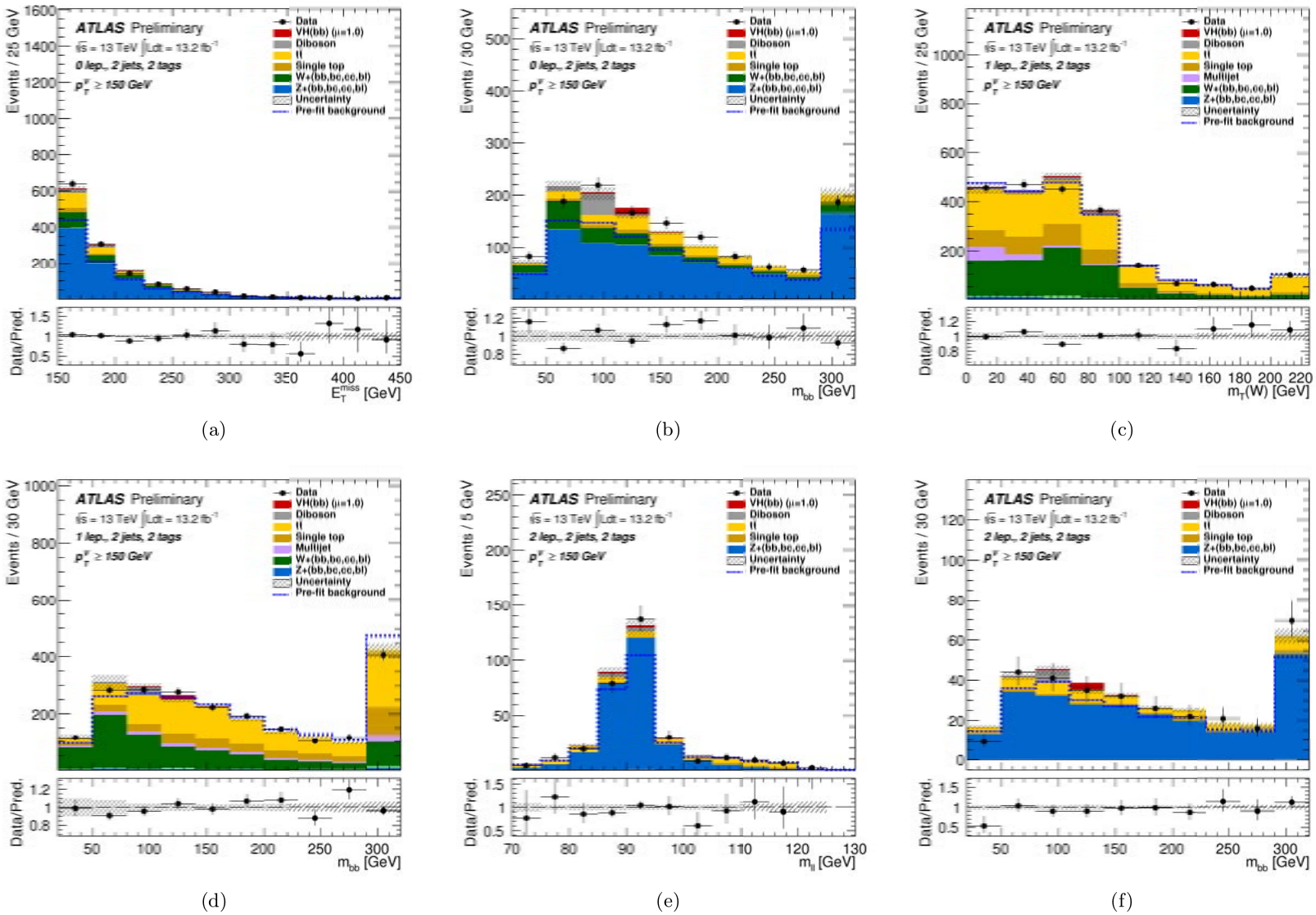


Figure 8.4: The E_T^{miss} (top left), m_T^W (middle left), m_{ll} (bottom left) and m_{bb} (right) post-fit distributions in the 0-lepton (top) 1-lepton (middle) and 2-lepton (bottom) channels for 2-jet, 2 b-tag events in the high p_T^V region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

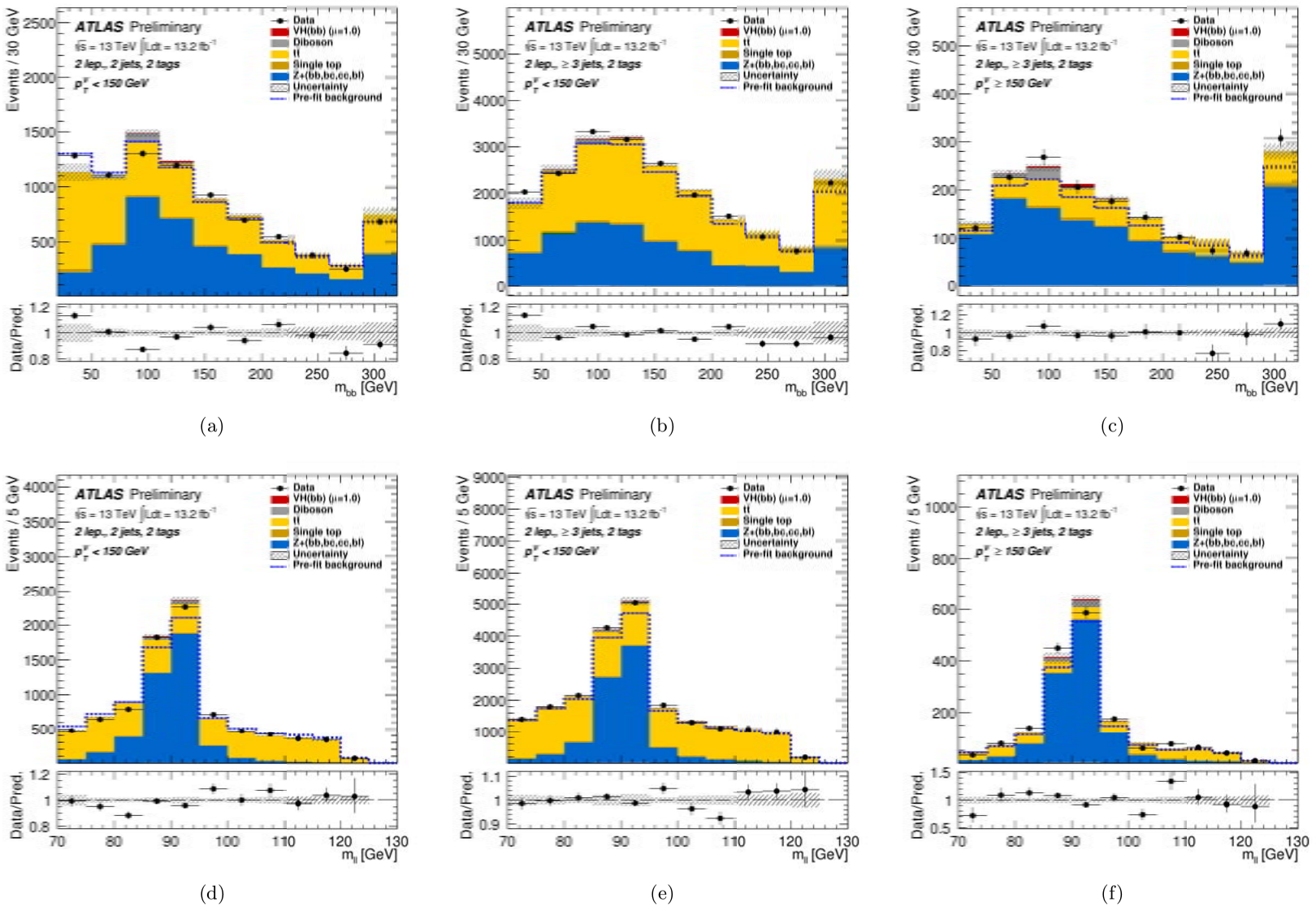


Figure 8.5: The m_{bb} (top) and m_{ll} (bottom) post-fit distributions in the 2-lepton channel for 2-tag events, in the 2-jet low p_T^V (left), $\geq$ 3-jets low p_T^V (center) and $\geq$ 3-jets high p_T^V regions (right). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

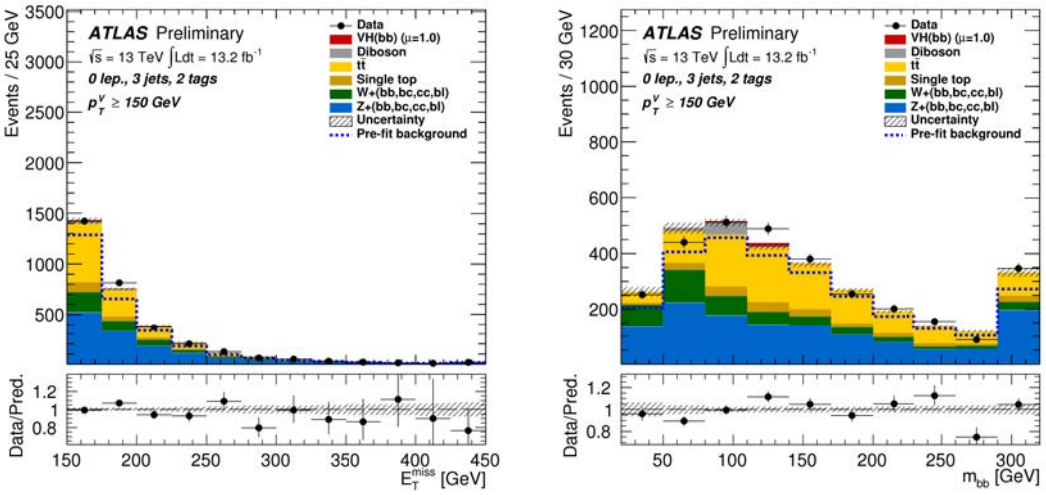


Figure 8.6: The E_T^{miss} (left) and m_{bb} (right) post-fit distributions in the 0-lepton channel for 2-tag events, in the 3-jet high p_T^V region (top). The m_T^W (left) and m_{bb} (right) post-fit distributions in the 1-lepton channel for 2-tag events, in the 3-jet high p_T^V region (bottom). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

8.5 Background Estimate and Modeling

The analysis presented in this Chapter is a very close relative of the search described in Chapter 6, therefore the same considerations introduced in Section 6.5 are in general valid. The description of the backgrounds heavily relies on MC generators: except for the multi-jet processes which are estimated with data-driven techniques, all other backgrounds are taken from the simulation. The normalization of the largest backgrounds is still left free to float in the Likelihood fit, however the correlation scheme introduced to control their yields across the different categories is updated with respect to what was used in Chapter 6.

In the analysis of Run-1 data some striking mismodelings of the V+jets processes have been observed, and dedicated corrections have been derived to adjust the MC prediction to the data distributions. Such mismodelings are not found in the analysis of Run-2 data, therefore no correction for the SM EW backgrounds is introduced in this analysis.
In this Section I describe the data-driven estimate of the multijet background, and the reweighting corrections applied to the VH signals to account for NLO(EW) effects on the differential kinematic distributions.

8.5.1 Estimate of the multijet background

The type of processes which contribute to the MJ background are the same introduced in Section 6.5.1, however the selection considered in this search has some crucial differences from the ones presented in Chapter 6, which affect the background composition and the impact of MJ processes: not only the 1-tag categories are not included in the analysis, but the search is limited in the 0- and 1-lepton channels to the high-p_T^V regions (above 150 GeV), whereas the Run-1 data analysis included 1-lepton events even at low transverse momentum (across the whole p_T^V spectrum in the single-muon channel, above 120 GeV in the single-electron channel). Both these aspects contribute to a reduction of the MJ background in the analysis of Run-2 data, which is therefore relatively simpler to control.

Multijet in 0-lepton channel In the 0-lepton channel the dedicated selection cuts introduced in Section 8.3 are very effective in rejecting MJ and non collision backgrounds. The residual contamination from MJ background has been studied in detail to assess its possible impact in this final state. The MJ estimate for this cross-checks is obtained from the Pythia 8 MC generator with A14 ATLAS tune and NNPDF2.3LO PDFs: the modeling provided by this MC has been validated by applying the analysis selection without b-tagging criteria and in 1-tag regions, in order to benefit from a larger statistics, and has been found to be compatible with the observed data distributions for the relevant kinematic quantities used to reject the MJ background. When possible mismodelings are observed as data-to-MC discrepancies, the MJ MC prediction is inflated to match the data (although this is a conservative approach: given the known difficulties in simulating MJ processes, it is deemed appropriate for this cross-check studies).
After the full analysis selection, in the 2-tag signal regions included in this search, the MJ contribution is not only lower than 1% of the total background, but it is also lower than 10% of the predicted Higgs boson signal yield in the mass window $60 < m_{bb} < 160$ GeV. The MJ background is therefore considered negligible and it is not included in the Likelihood fit for the 0-lepton channel.

Multijet in 1-lepton channel The 1-lepton channel is the most heavily affected by the MJ background, although the p_T^V boundary at 150 GeV has a strong impact on the

rejection of this process. To provide a solid estimate of the background, a data-driven technique similar to the one used for the 0-lepton channel in Chapter 6 is employed, here referred to as 'fake-factor' method.
The typologies of MJ processes relevant for the 1-lepton channel slightly differ in the single-electron and single-muon channel. In the single-muon channel the main contribution comes from multijet production with jets originated by heavy-flavor quarks containing muons from semileptonic decays of the quark itself. In the single-electron channel there is also a sizable contribution from events with jets and/or photon conversions which can mimic a single-lepton signature, therefore the MJ background is larger in this sub-channel. Given these differences, the MJ estimate is performed separately in the single-electron and single-muon sub-channels (as well as separately for 2-jet and 3-jet events), although using the same technique.
The background estimate relies on the definition of a MJ CR enriched in events containing fake leptons: this region is defined by the nominal analysis selection where leptons are identified as failing the tight requirements introduced in Section 5.1.2 but passing the loose ones (loose-not-tight), in order to maintain orthogonality from the signal regions. The MJ background estimate is taken from this region by removing the contribution of all remaining EW backgrounds, and it is corrected by applying a 'fake-factor' that accounts for the different lepton selection requirements between SRs and CRs. The fake-factor is extracted as the ratio of events measured in dijet-enriched data samples with the SR-like lepton selection (at the numerator) and the CR-like lepton selection (at the denominator), thus resulting in the following estimate for the MJ contribution in the analysis signal region:

$$N_{MJ}^{SR} = f \times N_{MJ}^{loose-not-tight} = f \times \frac{N_{dijet-data}^{SR}}{N_{dijet-data}^{loose-not-tight}}$$

under the assumption that f is the same for the baseline analysis selection and the dijet-enriched CRs selection. The dijet-enriched CR is defined by applying the baseline 1-lepton event selection to events with exactly one central jet and one loose lepton. Since the estimate of the fake-factor f is very sensitive to the contribution of EW backgrounds, the normalization of these processes is taken from MC simulation but fitted to data in the dijet-enriched CR, in the range of $150 < E_T^{miss} < 250$ GeV. To account for the fake-factor dependence on the lepton kinematic, f is measured separately in bins of η and p_T of the selected lepton, as well as E_T^{miss} categories for the single-electron channel.
The MJ background estimate naturally includes two components: the normalization and the shape of these processes. The yield of the MJ background is taken from the method described above applied to the 2-tag region, however the shape of the kinematic and topological quantities used in the analysis cannot be reliably extracted from 2-tag categories because of the low statistics. For this reason the template distributions for the MJ background are estimated by applying the fake-factor method to 1-tag categories, where the statistics is larger. Careful cross-checks have been performed to show that the shape of the relevant variables is similar between 1-tag and 2-tag regions for this background: this is mainly due to the fact that the MJ contribution is dominated by non-prompt leptons from heavy flavor jets, while the contribution from light jets faking electrons is smaller. The possible differences between shape distributions in 1-tag and 2-tag categories are however included in the estimate of the systematic uncertainties on the MJ background template.

Multijet in 2-lepton channel The 2-lepton channel shows very small MJ background, thanks to the presence of two isolated leptons in the Z boson invariant mass range. To

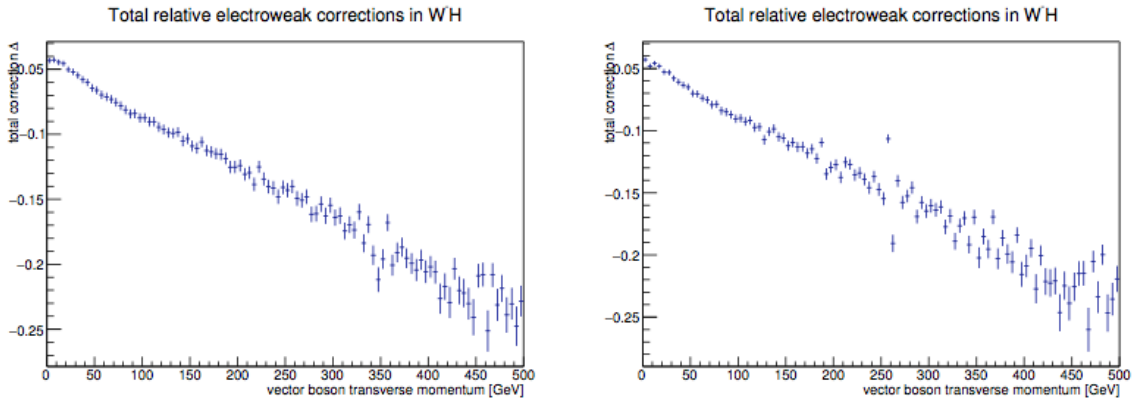
estimate the residual MJ contribution a fit to data is performed by using events from the baseline analysis selection with same-sign (SS) charge requirement for the two selected leptons. The fit includes the contribution of all EW backgrounds from MC simulation, while the MJ background prediction is taken from an exponential model. Assuming that same-sign and opposite-sign events are symmetric for the MJ background, its contribution in the analysis signal region is extracted from this fit, and it is found to be of the order of 0.3% and 1.4% in the muon and electron sub-channels respectively (after applying the dilepton invariant mass window cut of $71 < m_{ll} < 121$ GeV). Given the tiny size of this background in the 2-lepton channel, it is not included in the combined Likelihood fit as in the 0-lepton channel case.

8.5.2 Modeling of VH signal: NLO EW differential corrections

In parallel to the analysis of the LHC Run-2 data, several advancements have been made in the modeling of VH signal processes. However, next-to-leading order EW corrections are not yet included in the currently available generators that can be interfaced to parton shower models, to produce MC events for the analysis. These corrections are included in the normalization of the $qq \to VH$ signal samples, which are scaled to the NNLO(QCD)+NLO(EW) total cross section, however their effect on the differential distributions of signal simulated events is not accounted for.
For this reason, and knowing from the studies conducted for the analysis of Run-1 data that these corrections can be sizable, the VH signal events are reweighted to include differential NLO(EW) corrections as a function of the p_T^V distribution. These corrections are calculated with the HAWK MC software, obtaining a k_{EW}^{NLO}-factor function:

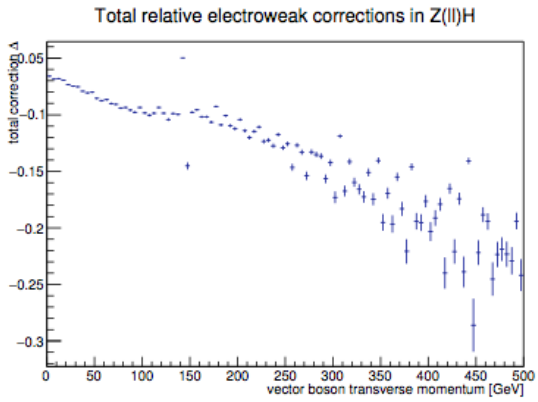
$$k_{EW}^{NLO}(p_T^V) = 1 + \delta_{EW}$$

separately for $Z(\to l^+l^-)H$, $Z(\to \nu\bar{\nu})H$ and $W^\pm(\to l^\pm\nu)H$ events. The NLO(EW) corrections are shown in Figure 8.7, separately for the different processes, as a function of the vector boson transverse momentum.
An uncertainty on these corrections is also estimated: similarly to what introduced in Section 6.5.3 the NNLO(EW) contribution is expected to be of the order of the squared NLO(EW) corrections, thus the uncertainty is quoted as $\max(1\%, \delta_{EW}^2, \Delta_\gamma)$ to ensure that it does not vanish, where 1% is the generic size of the neglected NNLO(EW) higher-order effects and $\Delta_\gamma = \Delta\sigma_\gamma/\sigma_\gamma$ is the relative uncertainty from photon-induced processes, as quoted in [52].

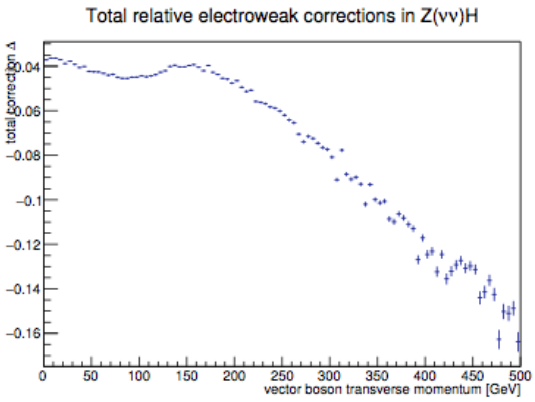


(a) $(W^- \to l^-\bar{\nu})H$

(b) $(W^+ \to l^+\nu)H$



(c) $(Z \to l^+l^-)H$



(d) $(Z \to \nu\bar{\nu})H$

Figure 8.7: Distribution of the NLO(EW) δ_{EW} correction as a function of the vector boson transverse momentum p_T^V for the different signal processes, as obtained from the HAWK MC program.

8.6 Systematic Uncertainties

In this Section I give a description of the different sources of systematic uncertainties included in this analysis, from the experimental systematics related to the performance of the detector and the object identification and reconstruction, to the uncertainties on the data-driven multijet estimate, to the systematics on the MC modeling of EW backgrounds and VH signals.

8.6.1 Experimental systematic uncertainties

The sources of experimental systematics are very similar to the ones described in Section 6.6.1 and 7.5.1, with differences arising from the increased center-of-mass energy, detector conditions, reconstruction and identification algorithms and the event selection applied. Systematics on leptons (lepton triggers, reconstruction, identification and isolation efficiencies and the energy and resolution corrections), E_T^{miss} trigger, jet energy scale (JES) and energy resolution (JER), do not differ substantially from what described in Section 7.5.1 and are therefore not covered in this section, although applied in the analysis.

- **integrated luminosity and pile-up**: the uncertainty on the integrated luminosity is quoted at 2.1% and 3.7% for data collected during 2015 and 2016 respectively, resulting in a global 2.9% uncertainty on the combined dataset, which has been estimated with the methodology introduced in [207], from a preliminary calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016. This uncertainty is applied on both the signal and background prediction. An uncertainty on the modeling of additional interactions per bunch crossing is applied to the MC simulation by varying the average number of additional interactions by $O(10\%)$.

8.6.2 Uncertainties on the modeling of the multijet background

Since this background is neglected in the 0- and 2-lepton channel, based on the studies described in Section 8.5.1, the systematic uncertainties on its estimate only affect the 1-lepton channel, and are divided into three (two) components for the electron (muon) sub-channel.
In the single-electron sub-channel the sources of systematics include the relative amount of jets faking leptons with respect to semileptonic heavy-flavor hadron decays, the impact of the energy scale applied to the 'fake-electrons' on the E_T^{miss} calculation and the amount of non-multijet background in the loose-not-tight isolation region. The latter is considered for the single-muon sub-channel as well, together with the bias from the different missing transverse momentum that isolated and non-isolated muon events may have.

8.6.3 Uncertainties on the MC modeling of signal and backgrounds

As introduced in Section 6.6.3 the analysis of the $VH(b\bar{b})$ heavily relies on MC simulation to estimate the background contributions: the proper assessment of systematic uncertainties on the MC modeling of the backgrounds is therefore crucial as it reflects our degree of confidence in the theoretical prediction for these processes, and the level of accuracy that can be expected from the simulation. The final results of this search are extracted via a Likelihood fit to data, in which the background shape and normalization can be adjusted to data (with or without parametric constraints). The size of the systematic uncertainties on the background prediction, and the correlations among them, are two crucial elements

of the Likelihood fit model that determine how and to what extent the fit to data is able to change and modify the MC estimate of background and signal.
For these reasons extensive studies have been performed to assess modeling systematic uncertainties for the main and the smaller backgrounds after the $VH(b\bar{b})$ analysis selection, as well as for the VH signal processes, which are described in detail in this section. The results of these studies reflect the status of our knowledge of these physics processes in the analysis considered, and quantify the accuracy of their description obtained from the MC simulations.
Since this search is performed by building a MV discriminant, the impact of systematic uncertainties on the modeling is considered for the main variables entering in the multivariate algorithm.

Signal systematic uncertainties Systematics on the signal model may be categorized in uncertainties on the total NNLO(QCD)+NLO(EW) cross section used for the normalization of the signals and the Higgs branching ratio BR($H \to b\bar{b}$), and shape and acceptance uncertainties from the MC simulation.
Uncertainties on the total cross section and the branching ratio are obtained from the prescription of [51] and [52], and affect only the overall normalization of the signal prediction. The former include the effects of renormalization (μ_R) and factorization (μ_F) independent scale variations in a range between 1/3 and 3 times their original values, as well as combined PDFs and α_S uncertainties from the 68% CL interval using the PDF4LHC15_nnlo_mc PDF set. Scale uncertainties are quoted in [52] for ZH and WH processes, without differentiating between qq- and gg-initiated ZH production. Given the different formal order of accuracy of the two production modes, these uncertainties are extracted separately for the two processes by applying the following procedure: scale uncertainties for the qq-initiated ZH production are assumed to be identical to the ones quoted for WH production, while the uncertainties on gg-initiated production are re-derived so that their quadrature sum with the qq-initiated ZH scale uncertainties results in the systematic variations quoted in [52] for the whole ZH production. In a similar way PDF+α_S uncertainties are not quoted separately for qq- and gg-initiated ZH processes: in this case the procedure described above is not suitable, therefore these uncertainties are taken from [51] for the $gg \to ZH$ contribution, while the ones quoted in [52] for the whole ZH production are applied to $qq \to ZH$ processes alone (assuming that the impact of gg-initiated ZH processes is relatively small).
The uncertainty on the Higgs branching ratio is quoted at 1.7% from [52] and it is applied directly in the analysis, accounting for missing higher order effects (QCD and EW) as well as uncertainties on the b-quark mass and the strong coupling constant α_S.
Table 8.1 shows a schematic summary of the normalization systematic uncertainties on the signal theory prediction. Finally a systematic uncertainty related to the reweighting correction function that includes NLO(EW) effects is applied in the analysis, already described in Section 8.5.2.
The second set of systematic uncertainties on the signal prediction covers effects on the acceptance and the shape of the baseline MC simulation after the analysis selection. As introduced in Section 8.2 the VH signals are simulated with the LO(QCD) Pythia 8 generator for qq-initiated processes, and the NLO(QCD) Powheg +Pythia 8 generator for the $gg \to ZH$ contribution.
A set of alternative samples has been produced to assess different sources of systematic uncertainties, relying on the NLO(QCD) Madgraph5_aMC@NLO generator for the hard scattering process, interfaced to the Pythia 8 parton shower, hadronization, underlying

event and multiple-parton-interaction (MPI) simulation, using the NNPDF2.3 5f FFN [186] PDF sets with the A14 tune. These alternative samples only include qq-initiated processes, therefore no specific acceptance or shape systematic is derived for $gg \to ZH$:

- MADGRAPH5_aMC@NLO +PYTHIA 8 nominal prediction, compared to the baseline PYTHIA 8 description to assess differences in the matrix element calculation, and compared to the alternative MADGRAPH5_aMC@NLO samples listed below.
- MADGRAPH5_aMC@NLO +PYTHIA 8 with variations in the A14 tune for the parton shower and the underlying event, covering upwards and downwards variations of the initial-state and final-state radiation, MPI and UE tunes.
- MADGRAPH5_aMC@NLO +PYTHIA 8 with μ_R and μ_F independent scale variations by a factor ranging from 1/2 to 2 times the original value, to assess the uncertainties from missing higher order corrections.
- MADGRAPH5_aMC@NLO +PYTHIA 8 with PDF and α_S variations, generated with the PDF4LHC15 PDFs sets including 30 PDFs variations and 2 α_S variations, to estimate PDF+α_S systematics.

These alternative samples are generated at *particle-level*, thus no detector simulation is considered and therefore they cannot be processed through the baseline event selection described in Section 8.3, since many quantities used to build the selection criteria cannot be defined. To analyze these samples a dedicated analysis selection has been implemented using the RIVET software, reflecting as closely as possible the baseline cuts. All the studies presented in this section using these alternative samples are conducted with this dedicated RIVET $VH(b\bar{b})$ analysis.
Two types of systematic uncertainties are studied: relative acceptance variations across the different analysis categories, and variations in the shape of two of the most discriminating variables used in the BDTs, the vector boson transverse momentum p_T^V and the invariant mass of the Higgs candidate m_{bb}. To extract acceptance variations all samples are normalized to the same total cross section for the different VH processes, in order to avoid double counting of the normalization uncertainties from Table 8.1; quote shape variations all distributions (after selection cuts) are scaled to the same area to avoid double counting of acceptance effects. Relative acceptance variations are studied across jet-bins, while variations across p_T^V categories are provided by the shape systematic on the p_T^V variable.
The following uncertainties are considered:

- **A14 tune variations**: changes in acceptance are calculated separately for 2-jet and 3-jet events and lepton channels, for each A14 tune variation, by taking the maximum among the absolute values of upwards variations, downwards variations and half the difference between the two. Separate tune variations are then summed in quadrature, resulting in acceptance changes very similary between 2-jet and 3-jet events. An overall acceptance variation is assigned inclusively with respect to the jet multiplicity as the maximum between 2-jet and 3-jet variations, and a 3-to-2 jet acceptance variations is quoted as the double-ratio between 3-jet and 2-jet acceptance changes encoding all tune variations as described above. These two acceptance uncertainties are correlated among VH processes, and the variations quoted for $qq \to ZH$ are applied to the $gg \to ZH$ as well: the values of the uncertainties are summarized in Table 8.2.
  Shape variations are assessed by comparing each tune variation to the baseline MADGRAPH5_aMC@NLO +PYTHIA 8 prediction for the p_T^V and m_{bb} variables (the latter

in the range between 0 and 200 GeV), and fitting the ratio with a linear function. For the p_T^V variable there is a single tune variation which shows a non negligible shape effect, for which the maximum of the linear fit between upwards and downwards variations is considered as systematic uncertainty. For the $Z(\to l^+l^-)H$ process the p_T^V shape is derived inclusively for 2-jet and 3-jet events since the effect is very similar, while this shape systematics is not applied for 3-jet $Z(\to \nu\bar{\nu})H$ events, where the shape effect is negligible. For the m_{bb} variable multiple tune variations show sizable discrepancy from the baseline prediction, thus the maxima of the linear fit between upwards and downwards variations are added in quadrature for the separate tune changes, and an uncertainty is derived for all signal processes. For $Z(\to l^+l^-)H$ events the m_{bb} uncertainty is estimated separately in low- and high-p_T^V categories. The p_T^V and m_{bb} shape uncertainties are considered as correlated among VH processes, and the shapes derived for $qq \to ZH$ processes are applied to the $gg \to ZH$ contribution as well.

- **QCD scale variations**: acceptance changes due to μ_R and μ_F scale variations are derived separately in each jet-bin and for each lepton channel, by comparing the baseline MADGRAPH5_aMC@NLO +PYTHIA 8 prediction with the alternative samples with varied scales and taking the envelope of all scale choices. To properly account for the effect of missing higher order corrections and their correlation across the analysis jet-bins, the *Stewart-Tackmann* [210] method was employed, in a similar way to what described in Section 6.6.3. For this reason, an additional variation is calculated for the 0- and 1-lepton channel to account for the effect of vetoing events with jet multiplicity larger than three. In the 2-lepton channel all variations are derived inclusively across the p_T^V spectrum. For $gg \to ZH$ processes the same uncertainties obtained for the $qq \to ZH$ contribution are used, but they are kept uncorrelated among qq- and gg-initiated production mode, as a result of the experience acquired in the analysis of LHC Run-1 data, where scale variation effects are different between these two processes (as described in Section 6.6.3). The acceptance uncertainties from scale variations implemented in the Stewart-Tackmann correlation scheme are given in Table 8.3.
  Shape uncertainties are studied for p_T^V and m_{bb} by considering the ratio between the baseline prediction and each scale variation, parametrized by a linear fit for p_T^V and a second order polynomial for m_{bb}, and considering the scale choice showing the largest discrepancy from the nominal. The p_T^V variations are derived separately for 2-jet and 3-jet events, but treated as correlated among jet-bins. The m_{bb} variations are obtained inclusively across the analysis categories since they are found to be very similar. Shape uncertainties derived for $qq \to ZH$ are applied to $gg \to ZH$ processes as well.

- **PDF+α_S variations**: changes in acceptance from PDFs and α_S variations are considered inclusively across jet-bins since no striking dependence on this categorization is observed, as well as inclusively across the p_T^V spectrum in the 2-lepton channel. The overall acceptance uncertainty is obtained as quadrature sum of all PDF and α_S variations compared to the baseline MADGRAPH5_aMC@NLO +PYTHIA 8 prediction. For ZH processes the uncertainties thus derived are only applied to the $qq \to ZH$ contribution, while for $gg \to ZH$ processes the systematics obtained in Section 6.6.3 for PDFs and α_S variations are conservatively implemented, since their values are significantly larger than the ones obtained for qq-initiated ZH production from this study. Table 8.4 summarizes these acceptance variations.

Shape uncertainties are quoted on p_T^V and m_{bb} with the same methodology used for scale variations: no m_{bb} shape effect is observed, while p_T^V shape variations are derived separately for 2-jet and 3-jet events. The p_T^V shape uncertainties obtained for $qq \to ZH$ are applied to $gg \to ZH$ as well, but treated as uncorrelated between qq- and gg-initiated ZH production modes.

- **ME variations (LO vs NLO)**: the baseline Pythia 8 MC prediction is compared to the alternative Madgraph5_aMC@NLO +Pythia 8 nominal sample, to study the effect of possible differences in acceptance and/or shape between MC simulation at LO and NLO(QCD). The acceptance variation effects observed are small compared to the signal uncertainties introduced so far, and the impact of missing higher order corrections is already encoded in the QCD scale uncertainties quoted in Table 8.3. Furthermore no striking shape effect is observed for the p_T^V and m_{bb} variables. As a result, no additional uncertainty is quoted from the LO-to-NLO comparison between the two generators, which simply constitutes a validation study of the baseline Pythia 8 MC simulation used in the analysis.

Process	Value	Source
$H \to b\bar{b}$ decay	1.7%	missing higher order, m_b and α_S effects
$qq \to ZH$	0.7%	μ_R, μ_F scale uncertainties
$gg \to ZH$	27%	μ_R, μ_F scale uncertainties
$qq \to ZH$	1.9%	PDF+α_S uncertainties
$qq \to WH$	1.6%	PDF+α_S uncertainties
(*correlated*)		
$gg \to ZH$	5%	PDF+α_S uncertainties

Table 8.1: Summary of the normalization systematic uncertainties on the signal theory prediction, showing the sub-processes affected, the relative value of the uncertainty and the source of systematic.

Process	Acceptance variation	
	overall	3-to-2 jet
$Z(\to \nu\bar{\nu})H$	7.5%	5%
$Z(\to l^+l^-)H$	6.5%	4%
$W(\to l\nu)H$	7.5%	6%

Table 8.2: Summary of the acceptance normalization systematics quoted from variations of the A14 ATLAS tune for signal processes.

Process	Acceptance variation			
	2-jet (1st)	2-jet (2nd)	3-jet	4-jet veto
$Z(\to \nu\bar{\nu})H$	3.6%	-1.8%	2.5%	-1.1%
$Z(\to l^+l^-)H$	0.4%	-1.3%	1.4%	-
$W(\to l\nu)H$	4.0%	-3.0%	3.6%	-3.0%
		correlated to 3-jet		
	correlated between WH and qq-ZH			

Table 8.3: Summary of the acceptance normalization systematics from scale variations implemented in the Stewart-Tackmann correlation scheme.

Process	Acceptance variation
$Z(\to \nu\bar{\nu})H$	0.6%
$Z(\to l^+l^-)H$	0.3%
$W(\to l\nu)H$	0.7%

Table 8.4: Summary of the acceptance normalization systematics quoted from variations PDFs and α_S.

Top-pair production systematic uncertainties Systematic uncertainties on the MC modeling of $t\bar{t}$ processes are derived by the comparison of the baseline POWHEG MC prediction to the different generators introduced in Section 7.5.2, with the addition of the following:

- POWHEG +HERWIG ++: this sample is generated using the same setup for POWHEG as for the nominal POWHEG +PYTHIA 6 sample, while parton shower, hadronisation, underlying event and MPI are simulated with HERWIG ++ (version 2.7.1) [234] with the UE-EE-5 [238] tune and the corresponding CTEQ6L1 PDF set. This sample provides the comparison to a different parton shower model (HERWIG ++ vs PYTHIA 6) with respect to the nominal prediction, without changing the hard scattering simulation.

As in Section 7.5.2 the alternative POWHEG samples are compared to the baseline POWHEG prediction, while the MADGRAPH5_AMC@NLO prediction is compared to POWHEG with the same parton shower model (HERWIG ++) to isolate effects arising from the different simulation of the hard-scattering process.
While in this case the $t\bar{t}$ samples are available with detector simulation included, they are analyzed using the same RIVET $VH(b\bar{b})$ analysis introduced for the study of signal systematic uncertainties, to avoid limitations in the statistical size of the samples due to reconstruction effects.
In the 2-tag analysis signal regions the $t\bar{t}$ background is mainly composed of events with two b-tagged jets originated by two b-hadrons (bb), or by one b-hadron and one c-hadron (bc) (a jet is associated to a specific hadronic flavor if the corresponding hadron is found within $\Delta R = 0.3$ from the jet axis among the jet constituents in the MC simulation). For this study of systematic uncertainties only $t\bar{t}$ events labelled as bb or bc, according to this definition, are included, assigning them weights corresponding to the b-tagging efficiencies for the selected jet pair, in order to properly reproduce the flavor composition of 2-tag reconstructed events.

The approach chosen to evaluate the modeling systematic uncertainties on the MC simulation of $t\bar{t}$ production stems from the structure of the Likelihood fit used in this analysis, in which the overall normalization of the main backgrounds ($t\bar{t}$ and (W/Z) + h.f.-jets) is left unconstrained and it is determined from data.
Two types of uncertainties are considered from MC-to-MC comparisons for the $t\bar{t}$ background: relative acceptance variations across different categories of the analysis, and possible shape variations for the most important variables entering in the BDTs (p_T^V and m_{bb}). Relative acceptance variations are derived by summing in quadrature the different MC-to-MC comparisons among acceptance ratios between 3-jet and 2-jet events, separately across lepton channels (and inclusively in p_T^V in the 2-lepton channel). For this evaluation all samples are normalized to the same total cross section, to cover only variation in the acceptance due to the analysis selection. The uncertainties thus derived are very similar for 0- and 1-lepton channel, and a common uncertainty is assigned to these channels. It is worth noting that while $t\bar{t}$ events in the 2-lepton channel are generally fully reconstructed, in the 0- and 1-lepton channels these background events are characterized by the misreconstruction of at least one object. For this reason, and given the different size of the uncertainties obtained for the 2-lepton versus the 0- and 1-lepton channels, the unconstrained overall normalization of the $t\bar{t}$ background is kept uncorrelated among 0+1-lepton and 2-lepton events. A summary of the normalization uncertainties for the $t\bar{t}$ background is given in Table 8.5.
Finally, shape variations for the p_T^V and m_{bb} variables are derived by parametrizing the

largest MC-to-MC variations (after normalizing the template distribution to the same area, to avoid double counting acceptance effects) with a linear or an exponential function. For both variables the largest variation arises from a change in the parton shower model (POWHEG +HERWIG ++ vs POWHEG +PYTHIA 6): the p_T^V shape uncertainty is derived separately for 0+1- and 2-lepton events, while the m_{bb} variations are extracted separately across all three lepton channels (correlating 0- and 1-lepton afterwards).

A cross-check study has been performed by analyzing the alternative $t\bar{t}$ samples with the baseline selection cuts at reconstructed-level, after detector simulation: the results show that no strong correlation between the p_T^V and m_{bb} variables is present, and that the size of the p_T^V shape variations is large enough to cover any shape effect on the transverse momentum spectrum of the leading jet. Both these results give solidity to the estimate and the implementation of the systematic uncertainties derived for the $t\bar{t}$ background.

Lepton channel	Uncertainty	Value
0+1-lepton	overall normalization	floating
2-lepton		
0+1-lepton	3-to-2 jet acceptance	9%
2-lepton		24%

Table 8.5: Summary of the normalization uncertainties for the $t\bar{t}$ background.

(W/Z) + jets systematic uncertainties In a similar way to what is described in Section 7.4 the (W/Z) + jets background is divided into sub-components based on the true hadronic flavor of the jets selected to build the Higgs candidate, determined by considering the flavor of hadrons within $\Delta R = 0.3$ from the jet axis in the MC simulation, as introduced in Section 6.2. The (W/Z) + jets six sub-components are: $W/Z + bb$, $W/Z + bc$, $W/Z + bl$, $W/Z + cc$, $W/Z + cl$ and $W/Z + ll$. The (W/Z) + jets background is then grouped into three sub-samples denoted as '$(W/Z) + (bb, bc, cc, bl)$' (also named (W/Z) + h.f.), '$(W/Z) + (cl)$' and '$(W/Z) + (l)$', which are treated separately in the Likelihood fit, and studied separately for the estimate of systematic uncertainties. In the Likelihood fit the overall normalization of the (W/Z) + h.f. component is left free to float and determined from the fit to data, while the remaining contributions (which amount to less than 1% of the total background in the 2-tag signal regions) are parametrized with a Gaussian constraint, whose prior is discussed in this Section.
The estimate of systematic uncertainties for these backgrounds follow the same methodology employed for $t\bar{t}$ processes: acceptance variations from MC-to-MC comparisons are used to assess priors on the overall normalization of the $(W/Z) + (cl)$ and $(W/Z) + (l)$ backgrounds, as well as relative normalization systematics across the different analysis regions for the (W/Z) + h.f. components (always separately for W and Z events). Furthermore shape systematic uncertainties are extracted, in this case from data-driven studies conducted in dedicated control regions.

The alternative MC samples used for these studies are obtained with the SHERPA generator with setup similar to the baseline simulation introduced in Section 8.2, produced with factorization, renormalization and resummation scales either doubled or halved, or matching scale varied in the range 15 GeV - 30 GeV. The envelope effect of upwards and

downwards scale changes is added in quadrature across the four scale variations. These samples are generated without detector simulation, therefore the studies presented in this Section are also relying on the RIVET analysis introduced for the signal and $t\bar{t}$ systematic studies.

Relative acceptance variations for the (W/Z) + h.f. backgrounds are derived by taking the comparison of each scale variation with respect to the baseline scale choice for the ratios between:

- 3-jet and 2-jet events, separately across lepton channels (and inclusively in p_T^V in the 2-lepton channel)
- 0-lepton and 1-lepton events (for W + h.f processes)
- 0-lepton and 2-lepton events (for Z + h.f processes)

In this way systematic uncertainties are obtained on the relative acceptance variation across jet-bins and across lepton-channels. No systematic is introduced to control the relative normalization of low- and high-p_T^Z events in the 2-lepton channel, since this effect is covered by a dedicated shape uncertainty. Furthermore the relative acceptance variations across the different hadronic flavor components of the (W/Z)+h.f. sample ($(W/Z)+(bb, bc, cc, bl)$) are considered. This results in three additional systematic uncertainties for (W/Z) + h.f. that control the relative flavor composition of the background. Figure 8.8 shows a schematic outline of this set of systematics and the way they control the normalization of the Z + h.f. background, while their values are summarized in Table 8.6 and 8.7 for both (W/Z) + h.f.. Acceptance variations across jet-bins and lepton-channels are estimated inclusively across flavor components (for the whole (W/Z) + h.f. sample), as the same variations are fairly consistent across the bb, bc, cc and bl sub-components. Acceptance variations across flavor components are derived separately across lepton channels and jet-bins, as non negligible differences are observed, but they are treated as fully correlated in the Likelihood fit. All systematics on the 3-to-2 jet ratio are fully correlated across lepton channels, all systematics on the lepton-channel ratios are fully correlated across jet-bins.
For the $(W/Z) + (cl)$ and the $(W/Z) + (l)$ backgrounds a systematic uncertainty on the overall normalization is obtained by adding in quadrature the effect of the four scale variations on the total acceptance within the analysis signal regions, from the envelope of the impact of upwards and downwards scale changes for each variation. This results in acceptance uncertainties of 18% and 23% for $Z + (l)$ and $Z + (cl)$, and 32% and 37% for $W + (l)$ and $W + (cl)$, which are included as prior on the parametrized normalization of these sub-samples in the Likelihood fit.

Shape uncertainties for the (W/Z) + h.f. backgrounds are obtained from data-to-MC comparisons performed in dedicated control regions enriched in these processes, after subtracting all remaining EW backgrounds, and MC-to-MC studies, considering the p_T^V and m_{bb} variables.
The Z + h.f. background is studied applying the 2-lepton selection with a tighter $81 < m_{ll} < 101$ GeV selection and introducing a cut on $E_T^{miss}/\sqrt{H_T}$ below $3.5\sqrt{\text{GeV}}$ (where H_T is the scalar sum of the two selected leptons and all reconstructed jets), to further suppress the $t\bar{t}$ background. The 0-, 1- and 2-tag regions are considered (removing events with $110 < m_{bb} < 140$ GeV for the 2-tag category to avoid the most signal-like region) separately for 2-jet and 3-jet events. The systematic shapes are obtained from a logarithmic fit of p_T^V and a linear fit of m_{bb} in such a way to encompass the data-to-MC differences

observed in the different categories. The resulting systematic uncertainties are quoted as functions of p_T^V and m_{bb}. The p_T^V function ranges from 6% at $p_T^V = 100$ GeV up to 15% at $p_T^V = 250$ GeV, while the m_{bb} function ranges from 0.5% up to 7.5% between 125 GeV and 250 GeV. It is worth noting that the striking mismodelings observed in the analysis of Run-1 data (for instance in the $\Delta\phi(j,j)$ variable) are not found in this study. These shape variations are applied to both 0-lepton and 2-lepton $Z+$ h.f. events.
The $W+$ h.f. background is considered in the 1-lepton categories: in the 1-tag and 2-tag regions the $t\bar{t}$ contribution is too large to allow a data-driven study, therefore the baseline SHERPA prediction is compared to the MADGRAPH5_AMC@NLO samples introduced in Section 7.5.2. The shape uncertainties on p_T^V and m_{bb} are obtained with the same methodology used for the $Z+$ h.f. data-to-MC study. The p_T^V function obtained is identical to the one derived for $Z+$ h.f., while the m_{bb} systematic is larger, ranging from 5% at 125 GeV up to 30% at 250 GeV. These shape variations are applied to both 0-lepton and 1-lepton $W+$ h.f. events.

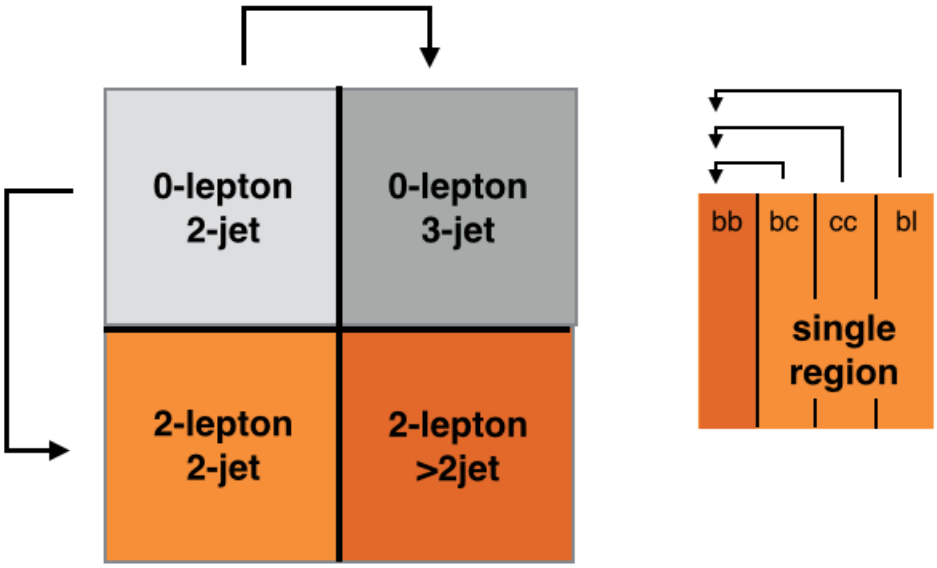


Figure 8.8: Schematic outline of the strategy applied to control the normalization of the $Z+$ h.f. background across analysis categories and flavor components.

Z + h.f. processes	Value	Category
3-to-2 jet acceptance variation	28%	(2-jet, 0-lepton)
	25%	(2-jet, 2-lepton)
0-lepton/2-lepton acceptance variation	26%	0-lepton
$Z + bc/Z + bb$	14%	0-lepton
	27%	(2-jet, 2-lepton)
	16%	(3-jet, 2-lepton)
$Z + cc/Z + bb$	12%	0-lepton
	31%	(2-jet, 2-lepton)
	7%	(3-jet, 2-lepton)
$Z + bl/Z + bb$	17%	0-lepton
	38%	(2-jet, 2-lepton)
	15%	(3-jet, 2-lepton)

Table 8.6: Summary of normalization acceptance uncertainty for the Z + h.f. background. The last column indicates the analysis region in which the corresponding systematic uncertainty is applied.

W + h.f. processes	Value	Category
3-to-2 jet acceptance variation	23%	(2-jet, 0+1-lepton)
0-lepton/1-lepton acceptance variation	17%	0-lepton
$W + bc/W + bb$	43%	0-lepton
	21%	1-lepton
$W + cc/W + bb$	24%	0-lepton
	13%	1-lepton
$W + bl/W + bb$	17%	0-lepton
	31%	1-lepton

Table 8.7: Summary of normalization acceptance uncertainty for the W + h.f. background. The last column indicates the analysis region in which the corresponding systematic uncertainty is applied.

Diboson systematic uncertainties The estimate of modeling uncertainties for semileptonic diboson processes follows very closely the methods applied for the (W/Z)+jets background, with the main differences that the overall normalization of the diboson contribution is always parametrized in the Likelihood fit with a Gaussian constraint, acceptance variations across flavor components are not necessary and no data-driven estimate of shape uncertainties has been performed. Among the three main diboson processes, WZ, ZZ and WW, the latter has a tiny contribution in the analysis selection, thus only an overall normalization uncertainty is considered. For WZ and ZZ processes only the dominant decay modes are considered ($WZ(q\bar{q}l\nu)$, $ZZ(q\bar{q}l^+l^-)$ and $ZZ(q\bar{q}\nu\bar{\nu})$) to extract the systematic uncertainties, which are applied to the sub-dominant modes as well $WZ(q\bar{q}l^+l^-)$ and $WZ(q\bar{q}\nu\nu)$. Relative acceptance variations are obtained from MC-to-MC comparisons, as well as shape uncertainties on the p_T^V and m_{bb} variables.
The nominal MC prediction for these processes is obtained from SHERPA, while alternative SHERPA samples are generated with factorization, renormalization and resummation scales either doubled or halved, or matching scale varied in the range 15 GeV - 30 GeV. Furthermore POWHEG +HERWIG ++ and POWHEG +PYTHIA 8 samples are also available and compared among themselves to assess parton shower model variations. These alternative samples are generated without detector simulation, therefore the studies presented in this Section rely on the RIVET analysis introduced for the previous systematic studies.

Acceptance variations for the diboson background are derived as follow: the envelope of upwards and downwards scale variations (for all the scale variations considered) and the parton shower variation from the comparison of HERWIG ++ and PYTHIA 8 are added in quadrature. Furthermore the full difference between the baseline SHERPA generator and the POWHEG +PYTHIA 8 prediction is considered: if this discrepancy is larger than the former quadrature sum, it is taken as the only acceptance uncertainty (this comparison encodes several effects already covered by parton shower and scale variations, thus including it in the quadrature sum would lead to double counting). The following acceptance ratios are compared for scale, parton shower and generator variations:

- 3-jet and 2-jet events, separately across lepton channels (and inclusively in p_T^V in the 2-lepton channel)
- 0-lepton and 1-lepton events (for $WZ(q\bar{q}l\nu)$ processes)
- 0-lepton and 2-lepton events (for $ZZ(q\bar{q}l^+l^-)$ and $ZZ(q\bar{q}\nu\nu)$ processes)

This results in two set of systematics controlling the relative acceptance variation across jet-bins (correlated among lepton channels) and between 0-lepton events and 1- or 2-lepton channels (correlated across jet-bins). No uncertainty is introduced on the relative acceptance between low-p_T^Z and high-p_T^Z events in the 2-lepton channel: control over this variable is provided by a dedicated shape systematic. Table 8.8 summarizes the values of these systematic uncertainties as well as the regions where they are applied.
Similarly to what is described for the $(W/Z) + (cl)$ and $(W/Z) + (l)$ backgrounds, an uncertainty on the overall normalization of ZZ, WZ and WW processes is derived by simply taking the acceptance variation (computed as described above to encode the different systematic sources) from the full analysis selection. This results in uncertainties of 20%, 26% and 25% on the global normalization of ZZ, WZ and WW processes respectively.
Shape uncertainties on the p_T^V and m_{bb} variables are obtained by comparing the different MC simulations, after normalizing the template distributions to the same area, considering the largest variations among them. Scale variations are always covered by parton shower

effects, therefore for this study only the PYTHIA vs HERWIG ++ and POWHEG vs SHERPA comparisons are included.
The m_{bb} variable shows a distinct feature for both sources of systematic, consisting in a shift of $O(10\text{GeV})$ of the V($q\bar{q}$) invariant mass peak towards higher masses (already documented in [239]), as it is clear by the comparisons shown in Figure 8.9. The variable does not show other striking systematic effects and it is fitted in the region between 25 and 500 GeV from the POWHEG vs SHERPA difference. The fit is performed separately for each diboson process, but inclusively across all analysis categories, resulting in 3 different shape functions which are treated as correlated. The parametrized function used for the fit is the following:

$$f(x) = a + bx + ce^{-x/d} + \frac{Gaus[((x-e)/f)+1]}{Gaus[((x-g)/f)+1}, \text{ where } x = m_{bb}.$$

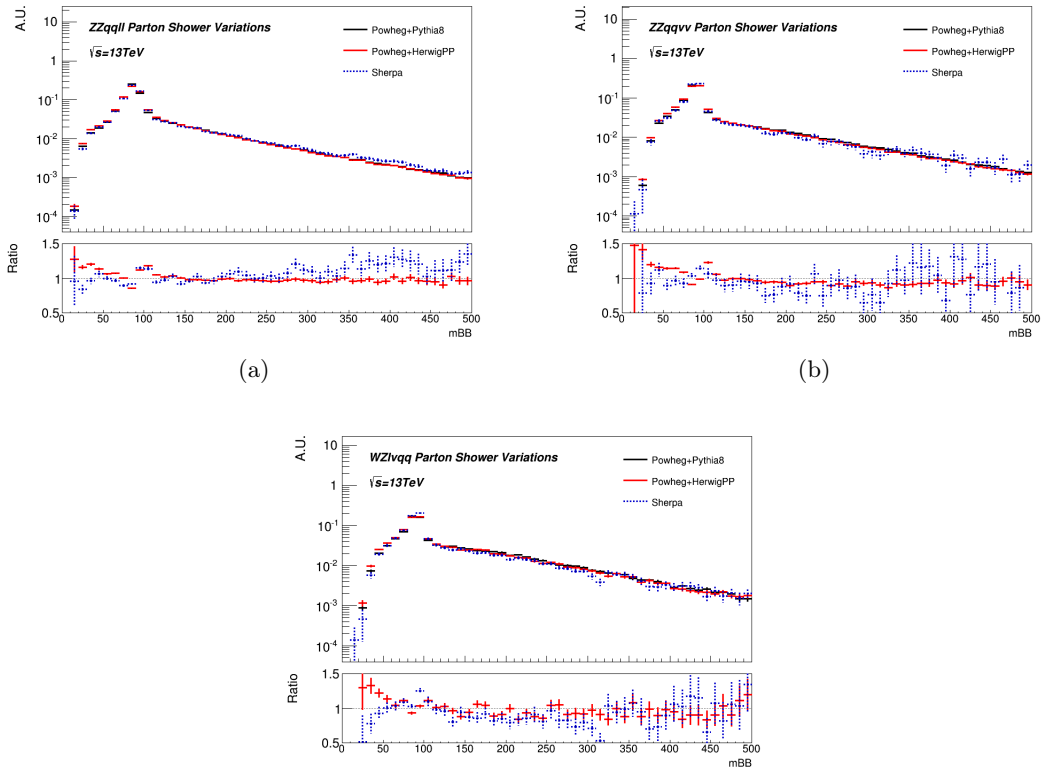
The p_T^V variable shows a mild shape effect for both systematic sources: the POWHEG vs SHERPA difference is fitted for ZZ events (separately for 0-lepton, 2-lepton, 2-jet, 3-jet categories) and WZ events (separately for 2-jet, 3-jet categories and 1-lepton events), deriving 6 shape functions which are treated as correlated across lepton channels, jet bins and diboson processes. The fit is performed in the range of 0-500 GeV for 2-lepton events, and 150-500 GeV for 0-lepton and 2-lepton events. The fitted function for this shape systematic is:

$$f(x) = a + bx + \frac{c}{x+d}, \text{ where } x = p_T^V.$$

It is worth noting that the main contribution from the diboson background in this analysis comes from events with a Z boson decaying to $b\bar{b}$, mimicking the signature of VH signal events, while all uncertainties discussed so far are extracted without any requirement on the flavor composition of the hadronically decaying boson in the VV pair. All acceptance and shape uncertainties have therefore been cross-checked by re-deriving them selecting diboson events with two jets originated by b-hadrons (labelled as bb according to the convention introduced for the $(W/Z)+$jets background), with results well consistent with the ones obtained without any requirement on the jet flavor (which benefit from a larger statistics).

Process	Acceptance variation	Value	Category
$ZZ(q\bar{q}l^+l^- + q\bar{q}\nu\nu)$	3-to-2 jet	19%	2-jet
	0-lepton/2-lepton	30%	0-lepton
$WZ(q\bar{q}l\nu)$	3-to-2 jet	14%	(2-jet, 0-lepton)
	3-to-2 jet	11%	(2-jet, 1-lepton)
	0-lepton/1-lepton	12%	0-lepton

Table 8.8: Summary of the acceptance normalization uncertainties for the semileptonic diboson background.



(a) (b)

(c)

Figure 8.9: Shape variation of the m_{bb} variable comparing POWHEG +PYTHIA 8 versus POWHEG +HERWIG ++ and POWHEG +PYTHIA 8 versus SHERPA, for all considered diboson samples: $ZZ(q\bar{q}l^+l^-)$ (a), $ZZ(q\bar{q}\nu\nu)$ (b) and $WZ(q\bar{q}l\nu)$ (c). All templates are normalised to the same integral to evaluate the effect on the shape of the variable without considering acceptance variation effects.

Single-top systematic uncertainties Systematics on the modeling of single-top processes are derived following the same methodology adopted for diboson processes.
The uncertainties on the overall normalization are provided by the LHC top working group and encodes QCD scale variations, PDFs and α_S effects, summed in quadrature to obtain 4.4%, 4.6% and 6.2% global uncertainty for t-, s- and Wt-single-top processes respectively. Acceptance and shape uncertainties are derived for the t- and Wt-channels by comparing alternative MC samples with the RIVET analysis selection previously introduced: the alternative samples are detailed in Table 8.9 along with the systematic effects considered. All studies are focused on the 1-lepton channel, where this background has a non negligible contribution, and result in acceptance systematic uncertainties for the 2-jet and 3-jet regions (separately) as well as p_T^V and m_{bb} shape uncertainties. The acceptance systematics are derived by comparing the alternative MC samples and taking the quadrature sum of the different systematic sources, resulting in the numbers presented in Table 8.10. Shape uncertainties are obtained by comparing the p_T^V and m_{bb} template distributions normalized to the same area among the different MC generators, and fitting the largest discrepancy with a linear function. The systematics derived in the 1-lepton channel for this background are conservatively applied in the 0- and 2-lepton channels as well.

Generator	Setup Details	Systematic Effect
POWHEG +PYTHIA 6	nominal setup scale variations low ($\mu_R = \mu_F = 2$) low radiation PERUGIA2012 tune variation	*low variation* for additional radiation
POWHEG +PYTHIA 6	nominal setup scale variations high ($\mu_R = \mu_F = 0.5$) high radiation PERUGIA2012 tune variation	*high variation* for additional radiation
POWHEG +PYTHIA 6	Wt-channel nominal setup 'diagram subtraction' scheme setup in the POWHEG ME calculation	alternative ME calculation
POWHEG +HERWIG ++	t-channel nominal setup parton showering with HERWIG ++ CTEQ6L1-UE-EE-5 tune for PS	alternative PS
MADGRAPH5_aMC@NLO +HERWIG ++	t-channel alternative setup ME with MADGRAPH5 5_aMC@NLO CT10f4 PDF in ME	alternative ME

Table 8.9: Single-top MC samples generated for the estimate of systematic uncertainties.

Process	Uncertainty	Value
s-channel	overall normalization	4.6%
t-channel		4.4%
Wt-channel		6.2%
t-channel	2-jet acceptance	16%
Wt-channel		25%
t-channel	3-jet acceptance	19%
Wt-channel		32%

Table 8.10: Summary of the normalization uncertainties for the single-top background.

8.7 Statistical Analysis: the Likelihood fit

The statistical analysis performed to extract the final results adopts the main principles outlined in Chapter 4: in this Section I present the specific details of the combined binned Likelihood fit, from the discriminating variables, to the signal and control regions included in the fit, to the different studies and tests that give solidity to the result. Many aspects of the fit model are shared between the different analyses described in this document, but given the differences in event selection, center of mass energy and detector conditions among the three searches, the model has to be carefully defined and validated to make sure that the results extracted from the Likelihood fit are solid and robust.

As stated in Section 8.6, the normalization of the main backgrounds is left free to float in the fit, in order to be sensitive to the the measured data: these backgrounds are (W/Z)+h.f and $t\bar{t}$ (for the latter the global normalization factor is uncorrelated between 0+1- and 2-lepton channels). The normalization of the remaining backgrounds (diboson, single-top, multijet, $(W/Z)+(cl)$ and $(W/Z)+(l)$) is parametrized in the Likelihood fit with a Gaussian constraint. For all background processes, the relative normalization across the different analysis regions, as well as the shape of the most important variables entering the BDTs, is controlled in the fit by the systematic uncertainties introduced in Section 8.6.3, included in the Likelihood fit as prior penalty factor on the parametrized Gaussian constraints. The normalization of the Higgs VH signal is of course free to float, and constitutes the parameter of interest which is extracted from the fit itself, the Higgs signal strength μ. All systematic uncertainties included in the fit model as nuisance parameters undergo a smoothing and pruning procedure to improve the robustness of the model and reduce its complexity, as described in Section 4.2.1. Statistical uncertainties due to the limited size of the simulated MC samples are taken into account in the fit with nuisance parameters constrained by Poisson distributions in each bin of the distribution, included if the relative statistical uncertainty on the total background is above 5%.

$VH(b\bar{b})$ categories: impact on the Likelihood fit The discriminating variable fitted to the measured data is the output of the multivariate algorithm defined in Section 8.4, BDT_{VH}, after applying the rebinning transformation. The BDT_{VH} output is shown in Figure 8.10 and 8.11: these template distributions are included in the Likelihood fit for all the regions listed in Table 8.2, covering three lepton channels, two jet-bins and two p_T^V regions for the 2-lepton channel only.
The 1-tag regions are not included in the combined fit as their impact is found to be negligible on the final results: the background modeling is already well under control from the fit of 2-tag regions and the effect of additional constraints from the 1-tag categories is not crucial.
This results in a relatively simpler fit model compared to the one detailed in Section 6.8.1: the 0- and 1-lepton channels are divided into just two sub-categories according to the jet multiplicity, where the 3-jet bin provides the strongest constraint on the $t\bar{t}$ background normalization and modeling, while the 2-jet bin has the largest $S/\sqrt{B}$.
The 2-lepton channel is divided into four sub-categories, separating low- and high-p_T^Z events and 2-jet and $\geq$ 3-jet events: the high-p_T^Z regions provide the largest sensitivity, while events at low transverse momentum are crucial to constrain the background modeling, especially for Z + h.f and $t\bar{t}$ processes (while the two backgrounds have similar size in the 2-jet low-p_T^Z bin, top-pair production becomes dominant in the 3-jet region). This aspect makes the proper understanding and implementation of the p_T^V systematic uncertainties on

these backgrounds a crucial aspect of the search, since these uncertainties are the nuisance parameters that control the relative normalization of high- and low-p_T^Z categories in the 2-lepton channel fit model, propagating the information gathered from the high-statistics low-p_T^Z regions to the high-sensitivity boosted categories. Event yields for the signal, backgrounds and data in the regions included in the combined fit are quoted in Table 8.12, after the modifications introduced by the Likelihood fit itself on the signal and background normalization.

Background normalization: floating scale factors During the Likelihood fit the normalization of $t\bar{t}$ and $(W/Z)+$h.f is adjusted to the measured data, without any prior constraint, completely free to float in the fit, with the correlation model described in the previous sections. Table 8.13 shows the scale factors obtained from the fit to data for these backgrounds, separately for $t\bar{t}$ in the 0+1- and 2-lepton channels, $W+$h.f and $Z+$h.f. While the normalization of $t\bar{t}$ and $Z+$h.f is relatively well behaved and compatible with the expectation, the $W+$h.f background requires a large scale factor of 1.59 to adjust the MC modeling to the data distributions. The measure of this SF is driven by the 1-lepton channel, where this background process is dominant compared to the 0- and 2-lepton channels. Two considerations are necessary to comment this results: there is no high-purity control region where the production of W bosons in association with heavy flavor jets can be easily measured and studied, to validate the accuracy of the MC simulation. The $W+$h.f background is a highly non-trivial process, that does not benefit in the fit from any pure control region to constrain its modeling, nor from constraints from low-p_T^V categories, that are not included in the 1-lepton channel. Furthermore it has been observed in previous studies (documented in [240]) that the Sherpa MC simulation has the tendency to underestimate the production of vector bosons in association with heavy flavor: it is possible that this potential mismodeling in the MC simulation results in a large positive scale factor.
This background is clearly a crucial aspect of the $VH(b\bar{b})$ search, and a better understanding of its modeling should be a priority for the next iteration of the analysis: both by improving the accuracy of the MC simulation, and by including additional regions of the 1-lepton channel that could bring precious information on the $W+$h.f features.
Finally it is worth noting that the floating normalization factors only tell part of the story: for these major backgrounds the complex model of normalization and shape systematic uncertainties described in Section 8.6.3 is implemented, providing the fit with several degrees of freedom to act on the background modeling on separate regions and/or variables.

Understanding the $VH(b\bar{b})$ Likelihood fit model Several studies and tests have been performed to assess the solidity of the fit model, following the experience gathered from the analyses presented in Chapter 6 and 7. The NPs included in the fit are investigated by studying their central values, uncertainties and correlations before and after the fit to data. Furthermore the NPs are ranked considering their impact on the measured signal strength $\hat{\mu}$, from the most to the least important, to highlight the crucial parameters in the fit model. These tests do not show striking pathological behaviors in the fit, thus proving the robustness of the results.
Figure 8.14 shows the NPs ranking according the their impact on $\hat{\mu}$, showing their central values and uncertainties after the fit to data as well.
It is interesting to observe how the systematic uncertainties that occupy the highest places in the ranking are all related to either the tagging efficiencies for b- and c-jets, or the MC modeling of the main backgrounds. Tagging efficiencies play a strong role in all three lepton channels (and are correlated to the normalization of the main backgrounds); the

Z + h.f. normalization is in particular dominant in the 0- and 2-lepton channels, while the impact of W + h.f. arises from the 1-lepton channel (and to a less extent it receives contribution from the 0-lepton one as well). It is worth noting that the floating normalization of the Z + h.f. correlates this background between 0- and 2-lepton channels. The m_{bb} shape uncertainty for the Z + h.f. is also driven high in the ranking by the 0- and 2-lepton channels, and gains a slight downwards pull from the 2-lepton channel regions.
The modeling of the Z + h.f. undergoes some adjustment in the fit, reflected in the pulls of the NP central values: both the p_T^V shape systematics (correlated between 0- and 2-lepton channels) and the 0-lepton normalization are pulled towards positive values. This is possibly due to a mismodeling in the p_T^V distribution, which results in a difference between low- and high-p_T^V regions, shown in the pull of the p_T^V Z + h.f. shape uncertainty (which goes in the direction of adjusting the MC simulation to data - note that this uncertainty is derived from data-to-MC comparisons). Since in the 0-lepton channel the low-p_T^V regions are not included, the interplay between Z + h.f. normalization and p_T^V shape from the 2-lepton regions fit may lead to a larger scale factor in the 0-channel alone. The normalization of the same background is also pulled upwards in the 2-jet category (correlated among 0- and 2-lepton channel and across the p_T^V spectrum), where again a possible cause can be the underestimate of Z production in association with heavy flavors from the SHERPA MC generator (common to both lepton channels and more evident when vetoing events with larger jet multiplicities).
The presence of the $t\bar{t}$ normalization in the 2-lepton channel among the top ranking NPs is not surprising as well, as this background has a strong impact in the $\geq$ 3-jets 2-lepton category, and shows some correlation with the Z + h.f. normalization (furthermore is is uncorrelated from the 0+1-lepton channels, thus it does not benefit from their additional constraints). The $t\bar{t}$ normalization in the 3-jet categories appears high in the ranking as well, mostly driven by the 1-lepton channel, and it is anti-correlated with the overall $t\bar{t}$ normalization in the 0+1-lepton channels.

The overall message from this ranking distribution is that, while experimental uncertainties from b- and c-tagging have a strong impact on the results, a proper understanding of the background modeling and the implementation of a complete and robust fit model (in terms of correlations among backgrounds and analysis regions) are critical to gain confidence with the solidity of the final results.

Measurement of the diboson $VZ(b\bar{b})$ signal strength As described in Section 6.8.4, the fit of the signal strength of semileptonic diboson processes μ_{VZ} provides a very useful validation of the $VH(b\bar{b})$ analysis, given the similar signature with the VH signal. To perform this cross-check an alternative BDT training is performed, as outlined in Section 8.4, while the remaining details of the analysis are unchanged. The only difference in the diboson fit model is the implementation of a Gaussian parametrization for the Higgs normalization with a constraint of 50%, while the uncertainties on the overall normalization of the diboson processes introduced in Section 8.6.3 are removed, replaced by an unconstrained diboson signal strength. The BDTs trained against VH or VZ signals provide good separation between diboson and Higgs signal processes, therefore the Higgs and diboson Likelihood fits are only weakly correlated, mainly thanks to the softer p_T^V spectrum and the peak of the m_{bb} distribution at lower mass, for VZ processes. The measured values of the diboson signal strength corresponds to $\mu_{VZ} = 0.91 \pm 0.17(\text{stat.})^{+0.32}_{0.27}(\text{syst.})$, well compatible with one within the uncertainties. The diboson processes are thus observed with a significance of 3.0 standard deviations (σ_{obs}^{VZ}), compared to the expected value of

3.2 (σ^{VZ}_{exp}).

8.8 Results

The main results of the $VH(b\bar{b})$ search with data collected during the LHC Run-2 are extracted from the statistical analysis detailed in Section 8.7 and reported in full detail in this section. From the combined Profile Likelihood Ratio fit including all three lepton channels the observed limit on the ratio of the cross section times branching ratio with respect to the SM expectation for an Higgs mass of $m_H = 125$ GeV is 1.2, compared to an expected limit, in absence of signal, of $1.0^{+0.4}_{-0.3}$. The probability p_0 of obtaining a result at least as signal-like as what is observed in the analyzed data, under the background-only assumption, is 34% for a tested Higgs boson mass of 125 GeV, while the expected p_0 value under the assumption of the signal hypothesis (presence of an Higgs boson with mass and signal strength predicted by the SM) is 3%.

This results in a Higgs boson excess observed with a significance of 0.42 standard deviations (σ^{VH}_{obs}), compared to an expectation of 1.94 standard deviations (σ^{VH}_{exp}). The expected and observed 95% CL limits, p_0 and significances are quoted in Table 8.15 for the single lepton channel fits (performed separately) and the combined global fit.
The Higgs signal strength extracted from the fit to data corresponds to $\mu = 0.21^{+0.36}_{-0.35}(\text{stat.})\pm 0.36(\text{syst.})$. Table 8.11 shows the breakdown of the error on μ in the different sources of uncertainty. It is worth noting that this breakdown is extracted from a combined fit including several categories, on which the effect of the various uncertainties is different: while the impact of systematic uncertainties is close to the statistical error from data, in the most sensitive high-p^V_T categories data statistics is still the dominating factor, therefore they will directly benefit from the data collected by the ATLAS detector until the end of the LHC Run-2. Furthermore, among the systematic uncertainties, the statistical fluctuations in the MC simulation have a strong impact on the result, especially at low p^V_T: this could be directly improved by increasing the size of the simulated samples (in principle a simple task, but often technically demanding). Among the experimental systematics, as expected from the ranking plot of Figure 8.14, the flavor-tagging uncertainties play the stronger role (this type of uncertainties will also largely benefit from the increased statistics that will be collected until the end of the LHC Run-2 data-taking, since they are extracted from data measurements).

Finally Figure 8.17 shows the data, background and signal yields, where the bins of the final discriminant BDT_{VH} in all signal regions are combined into bins of log(S/B). Here, S is the expected signal yield and B is the fitted background yield. This summary plots allows not only to gauge directly at the background composition across the whole $VH(b\bar{b})$ analysis regions, but in particular shows that in the three rightmost, most sensitive, bins of the distribution the observed data are well described by the background-only hypothesis, thus resulting in a fitted signal strength for the $VH(b\bar{b})$ signal lower than one.
The same Likelihood fit is also performed separating the Higgs signal strength in lepton channels, as well as in WH and ZH production modes (still including all regions in the combined fit model), free to float independently from each other in the fit to data, and without changes in the background fit model. The result of these fits with multiple parameters of interest are shown in Figure 8.16: it can be observed how all three lepton channels give a signal strength lower than the Higgs SM expectation (this behavior is evident in particular from the 2-lepton channel, where the fitted signal strength assumes negative

value).

Uncertainty	± Impact on the error on μ
Data Stat.	+ 0.361 - 0.346
Full Syst.	+ 0.358 - 0.360
MC Stat.	+ 0.208 - 0.215
Floating background normalization	+ 0.099 - 0.150
Floating and parametrized normalization	+ 0.126 - 0.171
All except normalization	+ 0.315 - 0.320
Jets and MET	+ 0.050 - 0.046
Flavor-tagging	+ 0.162 - 0.190
Leptons	+ 0.010 - 0.011
Luminosity	+ 0.017 - 0.011
Diboson modeling	+ 0.022 - 0.022
$Z+$jets modeling	+ 0.118 - 0.179
$W+$jets modeling	+ 0.097 - 0.136
$t\bar{t}$ modeling	+ 0.090 - 0.145
Single-top modeling	+ 0.042 - 0.031
Multijet modeling	+ 0.015 - 0.016
Signal modeling	+ 0.081 - 0.027
Total	+ 0.508 - 0.499

Table 8.11: Breakdown of the uncertainties on the fitted signal strength μ for the MVA Higgs fit. The uncertainties are divided into different sources: data statistics ('Data Stat.'), systematic uncertainties together with MC statistical uncertainties ('Full Syst.'), MC statistics alone ('MC Stat.') and different sub-groups of systematic uncertainties.

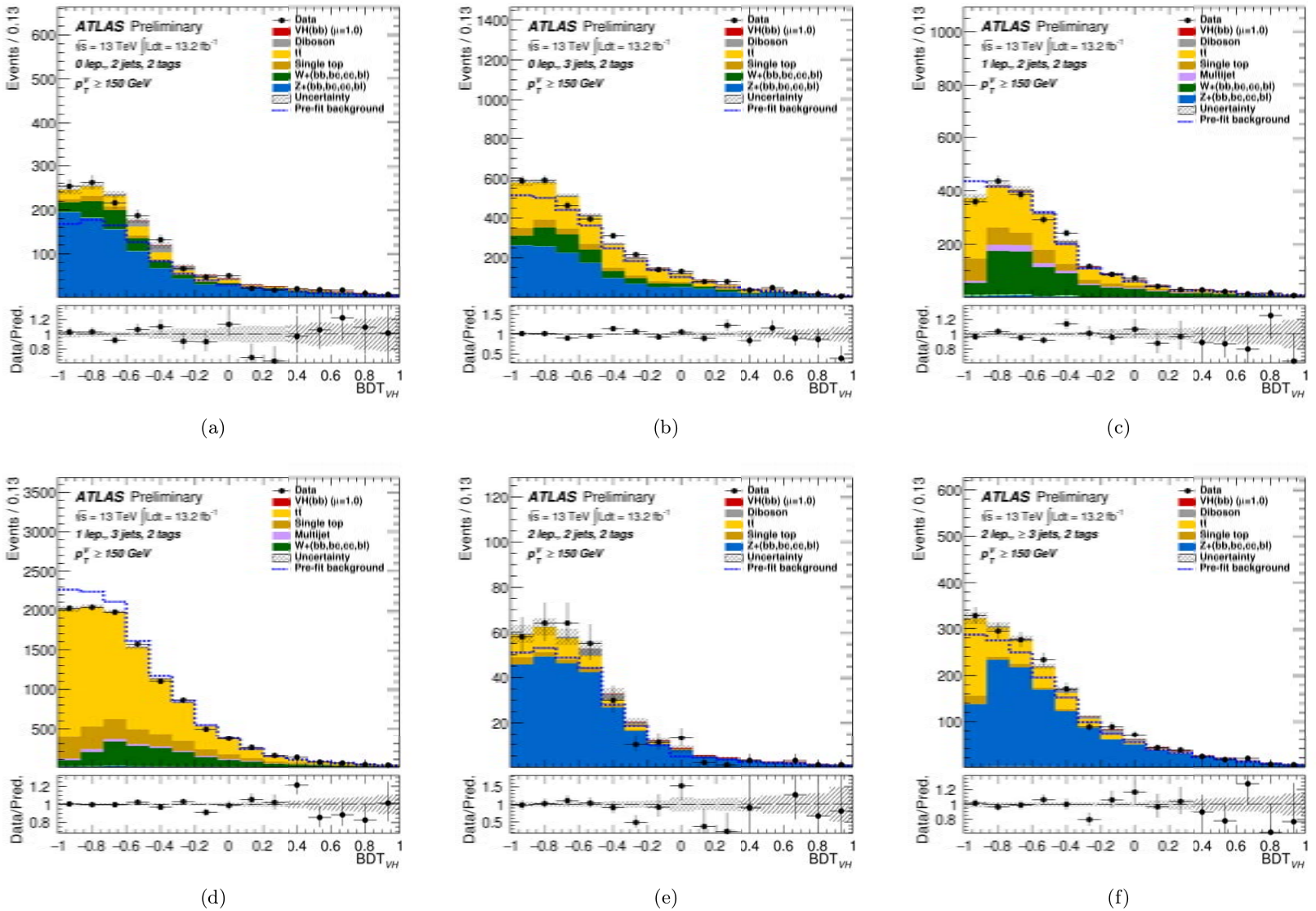
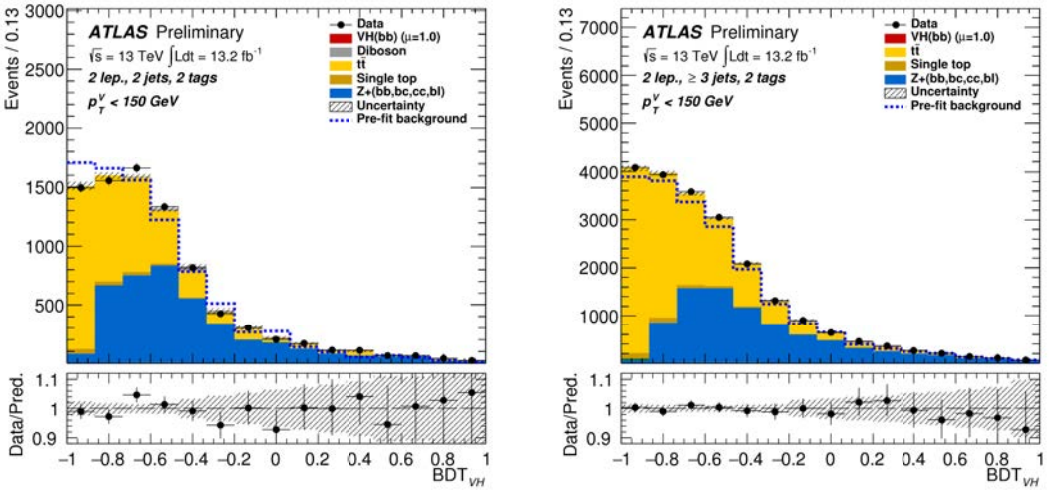


Figure 8.10: The BDT_{VH} post-fit distributions in the 0-lepton (top) 1-lepton (middle) and 2-lepton (bottom) channel for 2-tag events, in the 2-jet (left) and exactly 3 or $\geq$ 3-jets for the 2-lepton case (right) categories in the high p_T^V region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.



(a) (b)

Figure 8.11: The BDT_{VH} post-fit distributions in the 2-lepton channel for 2-tag events, in the 2-jet (left) and $\geq$ 3-jets (right) for the low p_T^V region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.

Sample	0-lepton		1-lepton		2-lepton			
	$p_T^V > 150$ GeV, 2-tag		$p_T^V > 150$ GeV, 2-tag		$p_T^V < 150$ GeV, 2-tag		$p_T^V > 150$ GeV, 2-tag	
	2-jet	3-jet	2-jet	3-jet	2-jet	3-jet	2-jet	3-jet
$Z + l$	1.5±0.1	3.3±2.2	–	–	4.6±0.1	15.4±0.5	0.4±0.0	2.9±0.1
$Z + cl$	4.2±1.8	6.7±2.6	0.9±0.6	–	13.9±5.9	49±21	1.0±0.4	10.0±4.3
Z + HF	864±49	1300±90	29.0±3.0	65.7±3.7	4000±120	8250±300	260±14	1192±49
$W + l$	2.3±1.5	3.8±2.2	4.3±0.1	9.6±0.3	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
$W + cl$	3.7±1.8	7.4±3.5	20±11	33±17	0.0±0.0	0.5±0.0	0.0±0.0	0.0±0.0
W + HF	184±37	440±96	741±114	1610±300	1.2±0.3	42±10	0.4±0.1	1.3±0.3
Single-top	45.5±7.7	204±39	331±55	1590±300	139±39	400±130	10.5±3.0	44±14
Multi-jet	–	–	101±63	210±140	–	–	–	–
$t\bar{t}$	136±14	1081±67	886±82	7520±360	4080±120	12210±340	42.3±4.3	402±36
Diboson	56±17	65±16	39±10	68±16	121±32	190±36	8.3±2.3	46.6±8.4
Total bkg.	1297±35	3110±52	2152±48	11120±110	8358±92	21150±150	322±13	1698±38
$VH(bb)$ (fit)	3.7±8.7	4.3±10.3	4.2±10.1	5.0±12.0	4.2±10.0	6.2±14.9	0.9±2.2	2.5±5.9
Data	1313	3120	2145	11124	8365	21163	316	1700

Figure 8.12: The data, background and signal yields along with the total uncertainty. All the background and signal values are evaluated according to the results of the global fit. The V+HF yield counts includes events from the V+bb, V+bc, V+bl and V+cc categories.

Sample	Scale factor
$t\bar{t}$ 0+1-lepton	0.86 ± 0.13
$t\bar{t}$ 2-lepton	0.94 ± 0.09
W + HF	1.59 ± 0.39
Z + HF	1.04 ± 0.11

Figure 8.13: Factors applied to the nominal normalisations of the $t\bar{t}$, $W+HF$ and $Z+HF$ backgrounds, as obtained from the global likelihood fit. The $t\bar{t}$ background is normalised independently in the 0+1 and 2 lepton channels. The errors include the statistical and systematic uncertainties.

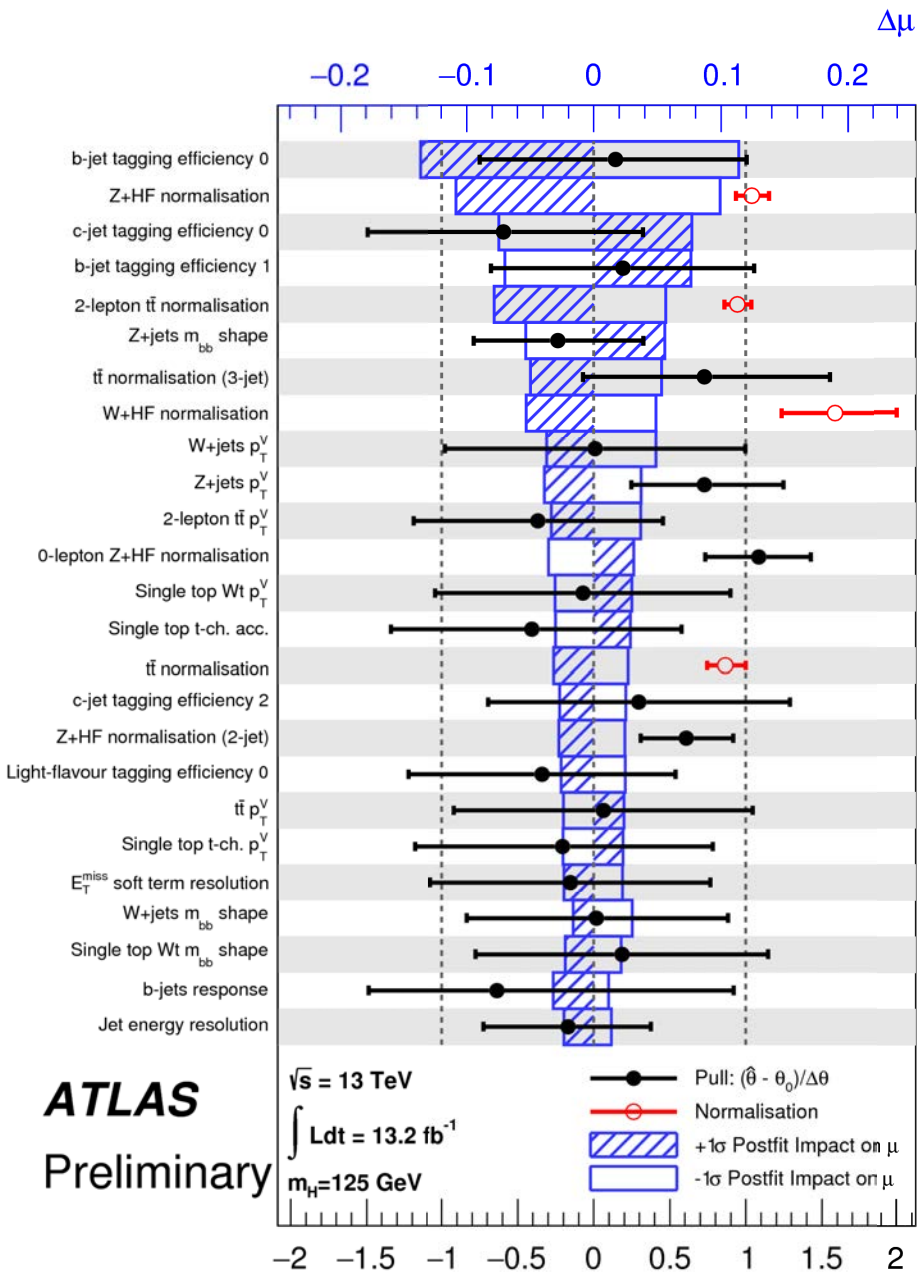
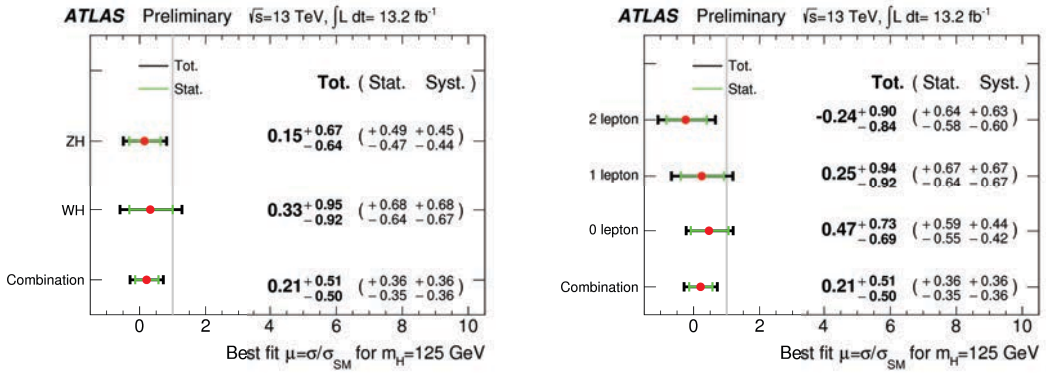


Figure 8.14: Impact of systematic uncertainties on the fitted signal-strength parameter μ. The systematic uncertainties are listed in decreasing order of their impact on $\hat{1}\frac{1}{4}$ on the y-axis. The boxes show the variations of μ, referring to the top x-axis, when fixing the corresponding individual nuisance parameter θ; to its post-fit value $\hat{\theta}$; modified upwards or downwards by its post-fit uncertainty, and repeating the fit as explained in the text. The hatched and open areas correspond to the upwards and downwards variations, respectively. The filled circles, referring to the bottom x-axis, show the deviations of the fitted nuisance parameters $\hat{\theta}$; from their nominal values θ_0 expressed in terms of standard deviations with respect to their nominal uncertainties $\Delta\theta$. The associated error bars show the post-fit uncertainties of the nuisance parameters, relative to their nominal uncertainties. The open circles with their error bars, also referring to the bottom x-axis, show the fitted values and uncertainties of the normalisation parameters that are freely floating in the fit. The normalisation parameters have a pre-fit value of zero. The jet energy scale and b-tagging uncertainties are decomposed into uncorrelated components; the numerical labels refer to such components.

Dataset	Limit		p_0		Significance	
	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
0-lepton	$1.4^{+0.6}_{-0.4}$	2.0	0.07	0.15	1.45	1.02
1-lepton	$2.0^{+0.8}_{-0.6}$	2.1	0.15	0.46	1.04	0.10
2-lepton	$1.8^{+0.7}_{-0.5}$	1.7	0.13	0.57	1.14	−0.17
Combined	$1.0^{+0.4}_{-0.3}$	1.2	0.03	0.34	1.94	0.42

Figure 8.15: The expected and observed 95% CL limits on the ratio of the cross-section times branching ratio with respect to the SM expectation and p_0 and significance values for the individual lepton channels and their combination. The expected limits are evaluated assuming the absence of signal and the expected p_0 and significance assuming a Higgs boson of $m_H = 125$ GeV mass with the SM signal strength.



(a) (b)

Figure 8.16: The fitted values of the Higgs boson signal-strength parameter, μ, for $m_H = 125$ GeV for the WH and ZH processes and their combination (left) and for the 0-, 1- and 2-lepton channels and their combination (right). The individual μ values in either the case of the $(W/Z)H$ processes or individual lepton channels are obtained from a simultaneous fit with the signal strength for each of the processes or lepton channels floated independently.

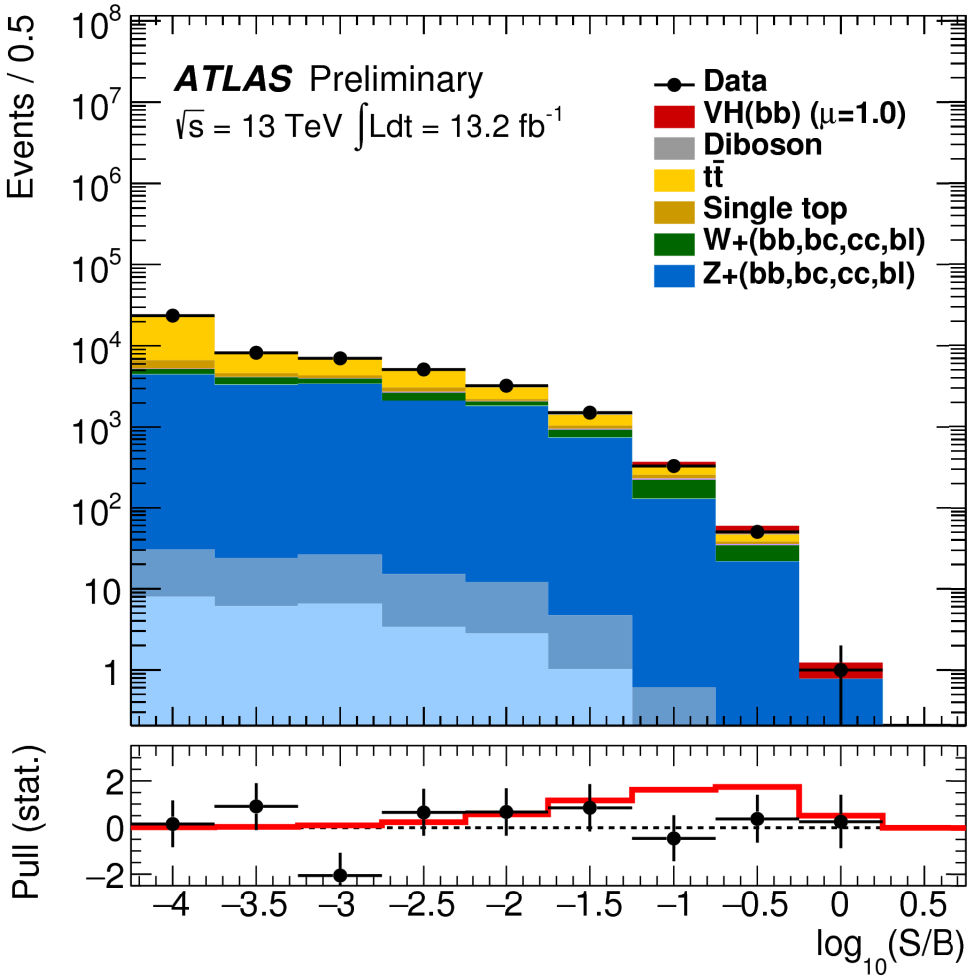


Figure 8.17: Event yields as a function of log(S/B) for data, background and Higgs boson signal with $m_H = 125$ GeV. Final-discriminant bins in all signal regions are combined into bins of log(S/B). The signal S and background B yields are the expected and fitted values, respectively. The Higgs boson signal contribution is shown as expected for the SM cross section (indicated as $\mu = 1.0$). The pull of the data with respect to the background-only prediction is shown without systematic uncertainties. The solid red line indicates the pull of the prediction for signal ($\mu = 1.0$) and background with respect to the background-only prediction.

9 | Conclusions and Outlook

The results of three analyses have been presented in this thesis, corresponding to two separate searches for SM Higgs physics and BSM phenomena.

The search for the SM Higgs boson in the decay channel to b-quark pairs in the associated production mode VH has been performed on the LHC Run-1 dataset, corresponding to an integrated luminosity of 25.0 fb^{-1} at $\sqrt{s} = 7$ and 8 TeV, and the data taken during the LHC Run-2 until July 2016, for a luminosity of 13.2 fb^{-1}, collected with the ATLAS detector. The results of these searches in terms of expected and observed sensitivities, as well as the measured Higgs signal strengths, are shown in Table 9.1. In both cases a signal strength lower than the SM expectation is observed, corresponding to a data deficit with respect to the signal+background hypothesis, which is reflected in an observed sensitivity below the expected value $\sigma_{obs} < \sigma_{exp}$.
The CMS Collaboration reported the results of the $VH(b\bar{b})$ search using the LHC Run-1 dataset [241], obtaining a signal strength of $\mu = 0.89 \pm 0.43$, with an observed significance of $2.09\sigma_{obs}$ and an expected significance of $2.52\sigma_{obs}$.

Data sample	Luminosity	σ_{exp}	σ_{obs}	Signal strength
$\sqrt{s} = 7 + 8$ TeV	25.0 fb^{-1}	2.6	1.4	$\mu = 0.51 \pm 0.31(\text{stat.}) \pm 0.24(\text{syst.})$
$\sqrt{s} = 13$ TeV	13.2 fb^{-1}	1.94	0.42	$\mu = 0.21^{+0.36}_{-0.35}(\text{stat.}) \pm 0.36(\text{syst.})$

Table 9.1: Summary of the results of $VH(b\bar{b})$ searches with Run-1 and Run-2 LHC data, in terms of expected and observed sensitivities and measured value of the Higgs signal strength.

While the signal strength measured from Run-1 data is well compatible with the SM expectation at 1.0, the measurement obtained from the preliminary Run-2 dataset shows a larger disagreement, not fully covered by $\pm 1\sigma$ statistical and systematic uncertainties. It is therefore interesting to quote the 95% CL upper limits on the cross section times branching ratio $pp \to (W/Z)(H \to b\bar{b})$ for this result: the observed limit for $m_H = 125$ GeV corresponds to 1.2 times the SM value, to be compared with the limit of $1.0^{+0.4}_{-0.3}$ expected in the absence of signal. An exclusion of the SM Higgs hypothesis in this search channel is therefore not possible from this analysis.

Although none of the analyses presented in this thesis has been able to reach the level of an observation of the $VH(b\bar{b})$ signal, it is worth noting the very good performance achieved in both cases, quantified by the expected sensitivity. Not only the $2.6\sigma_{exp}$ quoted from the analysis of Run-1 data represents the highest expected sensitivity for $VH(b\bar{b})$ searches among HEP experiments, but considering the Run-1 and Run-2 results together

would likely bring the expected sensitivity above the level of $3\sigma_{exp}$.
No attempt to combine the LHC Run-1 and Run-2 datasets for a comprehensive analysis of the whole available statistics has been made so far, and given the under-fluctuations observed in both searches it would not result in an observation at the level of $3\sigma_{obs}$.

While the analysis of the LHC Run-1 dataset is the ATLAS 'legacy' result for the $VH(b\bar{b})$ search, the $\sqrt{s} = 13$ TeV data will keep increasing throughout the whole LHC Run-2, for an expected integrated luminosity of approximately $30 - 40$ fb^{-1} to be collected by the end of 2016, and approximately 100 fb^{-1} by the end of the run. Under the assumption of the SM Higgs hypothesis, this increase in luminosity should likely be sufficient to bring the analysis of $\sqrt{s} = 13$ TeV data to the level of an observation ($3\sigma_{exp}$) by the end of 2016, and a discovery ($5\sigma_{exp}$) by the end of the LHC Run-2. Of course, the results presented in this thesis show a non negligible deficit in data compared to the SM Higgs hypothesis, which is not reflected in these expected sensitivities. The next months will be therefore critical to determine with better accuracy the nature of the $VH(b\bar{b})$ process, profiting from the increased data statistics.

Additional $\sqrt{s} = 13$ TeV data will deeply impact the results presented for the Run-2 dataset analysis. Although the level of statistical and systematic uncertainties on the measured signal strength is very similar, an increased statistics will allow for a better understanding of the complex background processes competing with the $VH(b\bar{b})$ signal, and a more accurate determination of data-driven experimental uncertainties, thus reducing indirectly the magnitude of the systematic error. It will be critical to improve the analysis design in order to extract as much detail and information as possible from the Run-2 dataset, to pin down with precision the modeling of the most complex backgrounds, such as (W/Z) + heavy flavors and $t\bar{t}$ production.

With the prospect of a rapidly increasing statistics during the LHC Run-2, it becomes very interesting to consider the $VH(b\bar{b})$ channel as a candidate to test possible deviations from the SM. From the analysis of LHC Run-1 and Run-2 data more and more limits on new physics states have been pushed to higher energies, and, in parallel, no significant indication of any deviation of the Higgs boson couplings from those expected in the SM has been observed.
In this scenario, Effective Field Theories (EFTs) in which the new physics is decoupled from the SM at some large energy scale Λ, and its leading effects can be parametrized through higher-dimension effective operators suppressed by Λ, are becoming more and more relevant.
The associated production mechanism VH is a very promising candidate to provide strong constraints on a number of these higher-dimension effective operators, specifically related to the Higgs coupling to vector bosons, competitive (and complementary) to the electroweak precision measurements from LEP [242, 243, 244].
In particular, it is very encouraging to consider possible deviations in differential distributions of quantities highly sensitive to the impact of EFT operators, such as the transverse momentum of the system recoiling away from the Higgs (p_T^V), or the invariant mass (m_{VH}). This impact is enhanced by powers of $\sqrt{\hat{s}}/\Lambda$, playing a strong role in the high tail of the m_{VH} distribution, and in the boosted region at large p_T^V, where the background contribution is extremely small.
With the additional data from the full LHC Run-2 an accurate study of the Higgs boson properties will play a critical role in the quest for new physics: the VH channel will be

an important actor in this precise measurements, with a leading impact from the $VH(b\bar{b})$ analysis, which is already contributing at the level of $O(85\%)$ to the VH expected significance from the Run-1 ATLAS and CMS combined Higgs measurement [11].

The $Zh(b\bar{b})$ final state has been interpreted in the search for a massive CP-odd A boson, decaying to a Zh pair where the h boson corresponds to the SM-like Higgs boson, with data collected at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 3.2 fb^{-1}. No significant excess is found when comparing data to the background expectation, although two upward deviations from the background-only hypothesis are observed, for masses of the A boson of $m_A = 260$ GeV and 440 GeV, with a local significance of approximately 2σ. Limits on the production cross sections times branching fractions, $\sigma(pp \to A) \times BR(A \to Zh) \times BR(h \to b\bar{b})$, in the range of [4.0, 0.017] pb ([6.9, 0.026] pb) for $m_A = [220, 2000]$ GeV assuming gluon-fusion (b-quark-associated) exclusive production are quoted, which are furthermore interpreted in the context of 2HMDs as constraints in the $\tan(\beta)$ and $\cos(\beta - \alpha)$ parameter plane.

The sensitivity of this search is strongly driven by statistics, which is the limiting factor on the obtained results, exceeding by far the impact of systematic uncertainties. The additional data collected by ATLAS at $\sqrt{s} = 13$ TeV will directly affect this analysis, even without fine improvements of the background modeling or the experimental uncertainties.

The analysis presented in this thesis has been optimized for signals produced by gluon-fusion mechanism ($gg \to A \to Zh$), while results are obtained considering b-quark associated production as well ($b\bar{b}A \to Zh$). It is worth noting that the $b\bar{b}A$ production mechanism has an increased sensitivity to the high-$\tan(\beta)$ region of the 2HDMs parameter space, thanks to the direct proportionality of the $b\bar{b}A$ coupling to $\tan(\beta)$ in Type-II 2HDMs. Revisiting the analysis optimization to improve its sensitivity to these types of processes would provide more stringent limits under these scenarios, although it would also open the analysis to background processes with complex features, such as top-quark pair production in association with heavy flavor quarks, whose understanding will be a challenge for the analysis of $Zh(b\bar{b})$ final states.

Acknowledgements

I would like to thank first my thesis director and supervisor, Giovanni Calderini: I first met him when I started the application process at the *Laboratoire de physique nucléaire et de hautes énergies* in Paris, in the summer of 2013, and he has been an outstanding guide during the three years of Ph.D. program. I had the chance, thanks to his support, to pursue many different opportunities and to extend my curiosity across several aspects of particle physics, always within the boundary of a coherent research project and a well designed plan.
The whole LPNHE ATLAS group has proven to be an incredible environment in these past years, source of several extremely interesting discussions, very useful advice, and a critical eye on my work from many areas of expertise. Among them I want to thank in particular the Hbb LPNHE team, including Giovanni Marchiori, Sandro de Cecco and Gregorio Bernardi: thanks to the experience and the effort of this group of people I could face the challenging analyses presented here. I express my appreciation and my thanks to the LAL Hbb group as well: the collaboration with Jean-François Grivaz, Nicolas Morange, Camilla Maiani, David Delgove, Jean-Baptiste de Vivie, Charles Delporte has been precious, and the frequent discussions with them extremely valuable.
I want to express a special thank you to Paolo Francavilla who, among the physicists I had the chance to meet during my Ph.D., left the most impactful mark on my formation as a young researcher. I thoroughly enjoyed the time we spent working on the analyses, as well as discussing in front of a blackboard, and I owe so much of what I learned in the past years to his experience, enthusiasm and especially his will to share, teach, debate and discuss.
Many thanks to all the people of LPNHE, whom I had the chance to meet and with whom I could discuss, chat and laugh (even outside the scope of my work): Sylvestre Pires, Louis Henry, Samuel Coquereau, Lorenzo Caccianiga, Ayan Mitra, Dimitris Varouchas, Yee Yap, Dilia Portilio, Changqiao Li, Stefano Manzoni, Marco Bomben, Francesco Crescioli.
The work presented in this document, and the whole experience of my Ph.D. program, would not have been possible without the contribution and the support of several people from the ATLAS experiment: from the Hbb working group, the main arena where the results described here were shaped, to the FTK group, the Higgs group, the Physics Modeling group, and several others. I had the invaluable chance to work in contact with many theorists, as well as colleagues from the CMS Collaboration: this experience enriched me and widened my view of particle physics outside the boundaries of the ATLAS experiment, and it was mainly possible thanks to the LHC Higgs Cross Section Working group, and the participation to the Les Houches 'Physics at TeV Colliders' workshop. The list of physicists and friends to thank from these various communities would be too long to be quoted, but I am confident I spent these three years surrounded by outstanding people, without whom this work would not have been possible.
Finally, I want to thank my closest friends and my family, who have always been there (even from afar) throughout the past years.

Bibliography

[1] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008.

[2] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[3] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[4] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[5] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[6] The 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.*, B716:1–29, 2012.

[7] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[8] Remi Lafaye, Tilman Plehn, Michael Rauch, Dirk Zerwas, and Michael Duhrssen. Measuring the Higgs Sector. *JHEP*, 08:009, 2009.

[9] T. Aaltonen et al. Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron. *Phys. Rev. Lett.*, 109:071804, 2012.

[10] Serguei Chatrchyan et al. Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks. *Phys. Rev.*, D89(1):012003, 2014.

[11] The ATLAS Collaboration. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 08:045, 2016.

[12] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva. Theory and phenomenology of two-Higgs-doublet models. *Phys. Rept.*, 516:1–102, 2012.

[13] The ATLAS Collaboration. Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys. Lett.*, B744:163–183, 2015.

[14] Vardan Khachatryan et al. Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in $l^+l^-b\bar{b}$ final states. *Phys. Lett.*, B748:221–243, 2015.

[15] Carlo Enrico Pandini and Paul Thompson. Theoretical studies for the 2HDM $A \to Zh \to, h \to b\bar{b}$ search and the SM $Vh \to b\bar{b}$ search
- Supporting Document. Technical Report ATL-COM-PHYS-2015-1474, CERN, Geneva, Dec 2015.

[16] Studies of Monte Carlo generators in Higgs boson production for ATLAS Run 2. Technical Report ATL-PHYS-PUB-2014-022, CERN, Geneva, Dec 2014.

[17] Multi-Boson Simulation for 13 TeV ATLAS Analyses. Technical Report ATL-PHYS-PUB-2016-002, CERN, Geneva, Jan 2016.

[18] Michael E Peskin and Daniel V Schroeder. *An Introduction to Quantum Field Theory; 1995 ed.* Westview, Boulder, CO, 1995. Includes exercises.

[19] Mark Srednicki. *Quantum Field Theory.* Cambridge Univ. Press, Cambridge, 2007.

[20] Richard Keith Ellis, William James Stirling, and Bryan R Webber. *QCD and Collider Physics.* Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge Univ. Press, Cambridge, 2003. Photography by S. Vascotto.

[21] J. Clerk Maxwell. A Dynamical Theory of the Electromagnetic Field. *Philosophical Transactions of the Royal Society of London Series I*, 155:459–512, 1865.

[22] H.A. Lorentz, A. Einstein, H. Minkowski, H. Weyl, and A. Sommerfeld. *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity.* Dover Books on Physics and Mathematical Physics. Dover, 1952.

[23] Ettore Majorana. Teoria simmetrica dell'elettrone e del positrone. *Il Nuovo Cimento (1924-1942)*, 14(4):171–184, 2008.

[24] S. M. Bilenky and S. T. Petcov. Massive neutrinos and neutrino oscillations. *Rev. Mod. Phys.*, 59:671–754, Jul 1987.

[25] S. T. Petcov. The Nature of Massive Neutrinos. *Adv. High Energy Phys.*, 2013:852987, 2013.

[26] G. Peter Lepage. What is renormalization? In *Boulder ASI 1989:483-508*, pages 483–508, 1989.

[27] John M. Cornwall, David N. Levin, and George Tiktopoulos. Derivation of gauge invariance from high-energy unitarity bounds on the s matrix. *Phys. Rev. D*, 10:1145–1167, Aug 1974.

[28] Benjamin W. Lee, C. Quigg, and H. B. Thacker. Weak interactions at very high energies: The role of the higgs-boson mass. *Phys. Rev. D*, 16:1519–1531, Sep 1977.

[29] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken symmetries. *Phys. Rev.*, 127:965–970, Aug 1962.

[30] E. Śanchez. Muon lifetime measurement with the fast detector at psi. *Physics Procedia*, 17:168 – 174, 2011.

[31] M. Steinhauser and T. Seidensticker. Second order corrections to the muon lifetime and the semileptonic B decay. *Phys. Lett.*, B467:271–278, 1999.

[32] D. M. Webber, V. Tishchenko, Q. Peng, S. Battu, R. M. Carey, D. B. Chitwood, J. Crnkovic, P. T. Debevec, S. Dhamija, W. Earle, A. Gafarov, K. Giovanetti, T. P. Gorringe, F. E. Gray, Z. Hartwig, D. W. Hertzog, B. Johnson, P. Kammel, B. Kiburg, S. Kizilgul, J. Kunkle, B. Lauss, I. Logashenko, K. R. Lynch, R. McNabb, J. P. Miller, F. Mulhauser, C. J. G. Onderwater, J. Phillips, S. Rath, B. L. Roberts, P. Winter, and B. Wolfe. Measurement of the positive muon lifetime and determination of the fermi constant to part-per-million precision. *Phys. Rev. Lett.*, 106:041803, Jan 2011.

[33] Timo Antero Aaltonen et al. Combination of CDF and D0 W-Boson Mass Measurements. *Phys. Rev.*, D88(5):052018, 2013.

[34] S. Schael et al. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP. *Phys. Rept.*, 532:119–244, 2013.

[35] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups. Precision Electroweak Measurements on the Z Resonance. *Phys. Rept.*, 427:257, 2006.

[36] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.

[37] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[38] The ATLAS Collaboration. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.

[39] N. Cabibbo, L. Maiani, G. Parisi, and R. Petronzio. Bounds on the Fermions and Higgs Boson Masses in Grand Unified Theories. *Nucl. Phys.*, B158:295–305, 1979.

[40] Pham Quang Hung. Vacuum instability and new constraints on fermion masses. *Phys. Rev. Lett.*, 42:873–876, Apr 1979.

[41] Giuseppe Degrassi, Stefano Di Vita, Joan Elias-Miro, Jose R. Espinosa, Gian F. Giudice, Gino Isidori, and Alessandro Strumia. Higgs mass and vacuum stability in the Standard Model at NNLO. *JHEP*, 08:098, 2012.

[42] Joan Elias-Miro, Jose R. Espinosa, Gian F. Giudice, Hyun Min Lee, and Alessandro Strumia. Stabilization of the Electroweak Vacuum by a Scalar Threshold Effect. *JHEP*, 06:031, 2012.

[43] Gino Isidori, Giovanni Ridolfi, and Alessandro Strumia. On the metastability of the standard model vacuum. *Nuclear Physics B*, 609(3):387 – 409, 2001.

[44] Vardan Khachatryan et al. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. *Eur. Phys. J.*, C75(5):212, 2015.

[45] The ATLAS Collaboration. Constraints on the off-shell Higgs boson signal strength in the high-mass ZZ and WW final states with the ATLAS detector. *Eur. Phys. J.*, C75(7):335, 2015.

[46] Vardan Khachatryan et al. Limits on the Higgs boson lifetime and width from its decay to four charged leptons. *Phys. Rev.*, D92(7):072010, 2015.

[47] P. Sikivie, Leonard Susskind, Mikhail B. Voloshin, and Valentin I. Zakharov. Isospin Breaking in Technicolor Models. *Nucl. Phys.*, B173:189–207, 1980.

[48] H. Georgi. Effective field theory. *Ann. Rev. Nucl. Part. Sci.*, 43:209–252, 1993.

[49] M. J. G. Veltman. Limit on Mass Differences in the Weinberg Model. *Nucl. Phys.*, B123:89–99, 1977.

[50] Riccardo Barbieri. *Ten Lectures on the ElectroWeak Interactions.* 2007.

[51] J R Andersen et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013.

[52] LHC Higgs Cross Section Working Group. *CERN Yellow Report 4.*

[53] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin, Carl Schmidt, Daniel Stump, and C. P. Yuan. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev.*, D93(3):033006, 2016.

[54] Charalampos Anastasiou, Claude Duhr, Falko Dulat, Elisabetta Furlan, Thomas Gehrmann, Franz Herzog, Achilleas Lazopoulos, and Bernhard Mistlberger. High precision determination of the gluon fusion Higgs boson cross-section at the LHC. *JHEP*, 05:058, 2016.

[55] M. Ciccolini, Ansgar Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC. *Phys. Rev. Lett.*, 99:161803, 2007.

[56] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. *Phys. Rev.*, D77:013002, 2008.

[57] Tao Han, G. Valencia, and S. Willenbrock. Structure function approach to vector boson scattering in p p collisions. *Phys. Rev. Lett.*, 69:3274–3277, 1992.

[58] Matteo Cacciari, FrÃ©dÃ©ric A. Dreyer, Alexander Karlberg, Gavin P. Salam, and Giulia Zanderighi. Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order. *Phys. Rev. Lett.*, 115(8):082002, 2015.

[59] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys. Lett.*, B579:149–156, 2004.

[60] J Ohnemus and William James Stirling. Order α_s corrections to the differential cross section for the WH intermediate-mass Higgs signal. *Phys. Rev. D*, 47(DTP-92-74. FERMILAB-PUB-92-278-T):2722–2729. 26 p, Oct 1992.

[61] Howard W Baer, B Bailey, and J F Owens. O(α_s) Monte Carlo approach to W + Higgs associated production at hadron supercolliders. Technical Report FSU-HEP-92-09-15, Florida State Univ. Dept. Phys., Tallahassee, FL, Oct 1992.

[62] Lukas Altenkamp, Stefan Dittmaier, Robert V. Harlander, Heidi Rzehak, and Tom J. E. Zirke. Gluon-induced Higgs-strahlung at next-to-leading order QCD. *JHEP*, 02:078, 2013.

[63] Robert V. Harlander, Hendrik Mantler, Simone Marzani, and Kemal J. Ozeren. Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass. *Eur. Phys. J.*, C66:359–372, 2010.

[64] Alexey Pak, Mikhail Rogal, and Matthias Steinhauser. Finite top quark mass effects in NNLO Higgs boson production at LHC. *JHEP*, 02:025, 2010.

[65] Simone Marzani, Richard D. Ball, Vittorio Del Duca, Stefano Forte, and Alessandro Vicini. Higgs production via gluon-gluon fusion with finite top mass beyond next-to-leading order. *Nucl. Phys.*, B800:127–145, 2008.

[66] Robert V. Harlander, Anna Kulesza, Vincent Theeuwes, and Tom Zirke. Soft gluon resummation for gluon-induced Higgs Strahlung. *JHEP*, 11:082, 2014.

[67] Christoph Englert, Matthew McCullough, and Michael Spannowsky. Gluon-initiated associated production boosts Higgs physics. *Phys. Rev.*, D89(1):013013, 2014.

[68] Oliver Brein, Robert V. Harlander, and Tom J. E. Zirke. vh@nnlo - Higgs Strahlung at hadron colliders. *Comput. Phys. Commun.*, 184:998–1003, 2013.

[69] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys.*, G43:023001, 2016.

[70] Ansgar Denner, Stefan Dittmaier, Stefan Kallweit, and Alexander MÃ$\frac{1}{4}$ck. HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders. *Comput. Phys. Commun.*, 195:161–171, 2015.

[71] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology. *Nucl. Phys.*, B849:296–363, 2011.

[72] "J. Stirling". Cross section as a function of the center of mass energy. http://mstwpdf.hepforge.org/plots/plots.html. Accessed: 2016-09-28.

[73] Kenneth G. Wilson. Renormalization group and strong interactions. *Phys. Rev. D*, 3:1818–1846, Apr 1971.

[74] G.t'Hooft. Recent developments in gauge theories. *Proc. of 1979 Carg'ese Institute*, pages p. 135 Press, New York 1980.

[75] Steven Weinberg. Implications of Dynamical Symmetry Breaking. *Phys. Rev.*, D13:974–996, 1976.

[76] Leonard Susskind. Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory. *Phys. Rev.*, D20:2619–2625, 1979.

[77] David B. Kaplan and Howard Georgi. SU(2) x U(1) Breaking by Vacuum Misalignment. *Phys. Lett.*, B136:183–186, 1984.

[78] Csaba Csaki, Jay Hubisz, and Patrick Meade. TASI lectures on electroweak symmetry breaking from extra dimensions. In *Physics in D >= 4. Proceedings, Theoretical Advanced Study Institute in elementary particle physics, TASI 2004, Boulder, USA, June 6-July 2, 2004*, pages 703–776, 2005.

[79] Stephen P. Martin. A Supersymmetry primer. 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[80] Leonard Susskind. THE GAUGE HIERARCHY PROBLEM, TECHNICOLOR, SUPERSYMMETRY, AND ALL THAT. *Phys. Rept.*, 104:181–193, 1984.

[81] J. Wess and B. Zumino. A Lagrangian Model Invariant Under Supergauge Transformations. *Phys. Lett.*, B49:52, 1974.

[82] John F. Donoghue. The effective field theory treatment of quantum gravity. *AIP Conf. Proc.*, 1483:73–94, 2012.

[83] F. Pisano and N. O. Reis. Cosmological constant and spontaneous gauge symmetry breaking: The Particle physics and cosmology interface charade. 2001.

[84] John H. Schwarz. Superstring Theory. *Phys. Rept.*, 89:223–322, 1982.

[85] Gianfranco Bertone. *Particle Dark Matter: Observations, Models and Searches*. Cambridge University Press, 2010.

[86] D. B. Cline, editor. *Sources of Dark Matter in the Universe: Proceedings, 1st International Symposium, February 16-18, 1994, Bel Air, CA*, Singapore, 1995. World Scientific, World Scientific.

[87] Thomas Mannel. Theory and phenomenology of CP violation. *Nucl. Phys. Proc. Suppl.*, 167:115–119, 2007. [,115(2007)].

[88] H. Banerjee, D. Chatterjee, and P. Mitra. Is there still a strong CP problem? *Phys. Lett.*, B573:109, 2003.

[89] Howard Georgi and S. L. Glashow. Unity of all elementary-particle forces. *Phys. Rev. Lett.*, 32:438–441, Feb 1974.

[90] A.J. Buras, J. Ellis, M.K. Gaillard, and D.V. Nanopoulos. Aspects of the grand unification of strong, weak and electromagnetic interactions. *Nuclear Physics B*, 135(1):66 – 92, 1978.

[91] Robert Foot, H. Lew, and R. R. Volkas. Electric charge quantization. *J. Phys.*, G19:361–372, 1993. [Erratum: J. Phys.G19,1067(1993)].

[92] B. Pontecorvo. Neutrino Experiments and the Problem of Conservation of Leptonic Charge. *Sov. Phys. JETP*, 26:984–988, 1968. [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[93] Peter Minkowski. $\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays? *Phys. Lett.*, B67:421–428, 1977.

[94] *ATLAS: letter of intent for a general-purpose pp experiment at the large hadron collider at CERN*. Letter of Intent. CERN, Geneva, 1992.

[95] W. W. Armstrong et al. ATLAS: Technical proposal for a general-purpose p p experiment at the Large Hadron Collider at CERN. 1994.

[96] S. Myers and E. Picasso. The Design, construction and commissioning of the CERN Large Electron Positron collider. *Contemp. Phys.*, 31:387–403, 1990.

[97] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[98] K Aamodt The ALICE Collaboration et al. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08002, 2008.

[99] The LHCf Collaboration. The lhcf detector at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08006, 2008.

[100] J L Pinfold. The moedal experiment at the lhc - a new light on the terascale frontier. *Journal of Physics: Conference Series*, 631(1):012014, 2015.

[101] G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008.

[102] Michael Benedikt, Paul Collier, V Mertens, John Poole, and Karlheinz Schindl. *LHC Design Report*. CERN, Geneva, 2004.

[103] "LHC outreach". Lhc ring layout. http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/lhc-schematic.jpg. Accessed: 2016-10-06.

[104] "The ATLAS Collaboration". Luminosity public results (run-1). https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults. Accessed: 2016-09-28.

[105] "The ATLAS Collaboration". Particle interactions in the atlas detector. http://www.interactions.org/imagebank/images/CE0155M.jpg. Accessed: 2016-09-28.

[106] ATLAS central solenoid: Technical design report. 1997.

[107] J P Badiou, J Beltramelli, J M Baze, and J Belorgey. *ATLAS barrel toroid: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. Electronic version not available.

[108] *ATLAS end-cap toroids: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. Electronic version not available.

[109] The ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. *Eur. Phys. J.*, C70:787–821, 2010.

[110] Norbert Wermes and G Hallewel. *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.

[111] A. Abdesselam et al. The barrel modules of the ATLAS semiconductor tracker. *Nucl. Instrum. Meth.*, A568:642–671, 2006.

[112] A. Abdesselam et al. The ATLAS semiconductor tracker end-cap module. *Nucl. Instrum. Meth.*, A575:353–389, 2007.

[113] E. Abat et al. The ATLAS TRT barrel detector. *JINST*, 3:P02014, 2008.

[114] E. Abat et al. The ATLAS TRT end-cap detectors. *JINST*, 3:P10003, 2008.

[115] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010.

[116] "The ATLAS Collaboration". Luminosity public results (run-2). https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2016/DataSummary/figs/peakLumiByFill.png. Accessed: 2016-10-07.

[117] ATLAS liquid argon calorimeter: Technical design report. 1996.

[118] The ATLAS Collaboration. Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions. *Eur. Phys. J.*, C70:723–753, 2010.

[119] ATLAS tile calorimeter: Technical design report. 1996.

[120] The ATLAS Collaboration. Readiness of the ATLAS Tile Calorimeter for LHC collisions. *Eur. Phys. J.*, C70:1193–1236, 2010.

[121] A Artamonov, D Bailey, G Belanger, M Cadabeschi, T Y Chen, V Epshteyn, P Gorbounov, K K Joo, M Khakzad, V Khovanskiy, P Krieger, P Loch, J Mayer, E Neuheimer, F G Oakham, M O'Neill, R S Orr, M Qi, J Rutherfoord, A Savine, M Schram, P Shatalov, L Shaver, M Shupe, G Stairs, V Strickland, D Tompkins, I Tsukerman, and K Vincent. The atlas forward calorimeter. *Journal of Instrumentation*, 3(02):P02010, 2008.

[122] *ATLAS muon spectrometer: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[123] Peter Jenni, Markus Nordberg, Marzio Nessi, and Kerstin Jon-And. *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity.* Technical Design Report ATLAS. CERN, Geneva, 2008.

[124] Peter Jenni, Marzio Nessi, and Markus Nordberg. Zero Degree Calorimeters for ATLAS. Technical Report LHCC-I-016. CERN-LHCC-2007-001, CERN, Geneva, Jan 2007.

[125] ATLAS first level trigger: Technical design report. 1998.

[126] Peter Jenni, Marzio Nessi, Markus Nordberg, and Kenway Smith. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 2003.

[127] M Shochet, L Tompkins, V Cavaliere, P Giannetti, A Annovi, and G Volpi. Fast TracKer (FTK) Technical Design Report. Technical Report CERN-LHCC-2013-007. ATLAS-TDR-021, Jun 2013. ATLAS Fast Tracker Technical Design Report.

[128] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys.*, G28:2693–2704, 2002. [,11(2002)].

[129] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[130] Abraham Wald. Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. *Transactions of the American Mathematical Society*, 54(3):426–482, November 1943.

[131] Isaac Asimov. *Franchise*, volume 1. 1990. Broadway Books.

[132] The ATLAS Collaboration. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. *Eur. Phys. J.*, C74(7):2941, 2014.

[133] The ATLAS collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. 2014.

[134] The ATLAS collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. 2016.

[135] The ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J.*, C74(10):3071, 2014.

[136] The ATLAS Collaboration. Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s}$ = 7 TeV in 2010. *Eur. Phys. J.*, C74(9):3034, 2014.

[137] The ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ = 13 TeV. *Eur. Phys. J.*, C76(5):292, 2016.

[138] Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC. Technical Report ATL-PHYS-PUB-2015-045, CERN, Geneva, Nov 2015.

[139] The ATLAS Collaboration. Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment. *Eur. Phys. J.*, C76(5):295, 2016.

[140] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008.

[141] Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC. Technical Report ATL-PHYS-PUB-2015-045, CERN, Geneva, Nov 2015.

[142] Gavin P. Salam. Towards Jetography. *Eur. Phys. J.*, C67:637–686, 2010.

[143] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal. Calorimeter clustering algorithms: Description and performance. 2008.

[144] The ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s}$=7 TeV. *Eur. Phys. J.*, C73(3):2304, 2013.

[145] The ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s}$=7 TeV with the ATLAS detector. *Eur. Phys. J.*, C75:17, 2015.

[146] Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}$=13 TeV. Technical Report ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015.

[147] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett.*, B659:119–126, 2008.

[148] Pile-up subtraction and suppression for jets in ATLAS. Technical Report ATLAS-CONF-2013-083, CERN, Geneva, Aug 2013.

[149] Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical Report ATLAS-CONF-2015-029, CERN, Geneva, Jul 2015.

[150] The ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. 2015.

[151] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet Trimming. *JHEP*, 02:084, 2010.

[152] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187–224, 1993.

[153] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*, D48:3160–3166, 1993.

[154] Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS. Technical Report ATL-PHYS-PUB-2015-033, CERN, Geneva, Aug 2015.

[155] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04:005, 2008.

[156] b-jet tagging calibration on c-jets containing D^{*+} mesons. Technical Report ATLAS-CONF-2012-039, CERN, Geneva, Mar 2012.

[157] Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector. Technical Report ATLAS-CONF-2012-040, CERN, Geneva, Mar 2012.

[158] Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector. Technical Report ATLAS-CONF-2012-043, CERN, Geneva, Mar 2012.

[159] Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. Technical Report ATLAS-CONF-2011-102, CERN, Geneva, Jul 2011.

[160] The ATLAS collaboration. Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment. 2014.

[161] The ATLAS collaboration. Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data. 2014.

[162] The ATLAS Collaboration. Performance of b-Jet Identification in the ATLAS Experiment. *JINST*, 11(04):P04008, 2016.

[163] Expected performance of the ATLAS b-tagging algorithms in Run-2. Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul 2015.

[164] Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data. Technical Report ATL-PHYS-PUB-2015-039, CERN, Geneva, Aug 2015.

[165] Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run. Technical Report ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun 2016.

[166] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[167] The ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS. *Eur. Phys. J.*, C72:1844, 2012.

[168] The ATLAS collaboration. Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV. 2013.

[169] Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-023, CERN, Geneva, Jul 2015.

[170] Bukin A. Fitting function for asymmetric peaks.

[171] Measuring the b-tag efficiency in a top-pair sample with 4.7 fb^{-1} of data from the ATLAS detector. 2012.

[172] Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector. 2012.

[173] The 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. *JHEP*, 01:069, 2015.

[174] The ATLAS Collaboration. Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b-quark pair with the ATLAS detector. *Phys. Lett.*, B718:369–390, 2012.

[175] Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production with the ATLAS detector. Technical Report ATLAS-CONF-2013-079, CERN, Geneva, Jul 2013.

[176] The ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010.

[177] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[178] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[179] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

[180] ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. 2011.

[181] New ATLAS event generator tunes to 2010 data. 2011.

[182] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[183] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[184] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[185] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. $HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013.

[186] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C. P. Yuan. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010.

[187] M. L. Ciccolini, S. Dittmaier, and M. Kramer. Electroweak radiative corrections to associated WH and ZH production at hadron colliders. *Phys. Rev.*, D68:073003, 2003.

[188] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension. *Comput. Phys. Commun.*, 108:56–74, 1998.

[189] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.

[190] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[191] Kirill Melnikov and Frank Petriello. Electroweak gauge boson production at hadron colliders through O(alpha(s)**2). *Phys. Rev.*, D74:114017, 2006.

[192] Michal Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\pm\frac{4}{S})$. *Phys. Rev. Lett.*, 110:252004, 2013.

[193] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.

[194] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205-206:10–15, 2010.

[195] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[196] Borut Paul Kersevan and Elzbieta Richter-Was. The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1. *Comput. Phys. Commun.*, 184:919–985, 2013.

[197] Nikolaos Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Phys. Rev.*, D83:091503, 2011.

[198] Nikolaos Kidonakis. NNLL resummation for s-channel single top quark production. *Phys. Rev.*, D81:054028, 2010.

[199] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys. Rev.*, D82:054018, 2010.

[200] A. Sherstnev and R. S. Thorne. Parton Distributions for LO Generators. *Eur. Phys. J.*, C55:553–575, 2008.

[201] R. Olshen L. Breiman, J. Friedman and C. Stone. *Classification and Regression Trees.*

[202] Y. Freund and R. Schapire. *Experiments with a new boosting algorithm, in Machine Learning.*

[203] Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[204] The ATLAS Collaboration. Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. *Phys. Rev.*, D90(7):072004, 2014.

[205] Michal Czakon, David Heymes, and Alexander Mitov. High-precision differential predictions for top-quark pairs at the LHC. *Phys. Rev. Lett.*, 116(8):082003, 2016.

[206] Ansgar Denner, Stefan Dittmaier, Stefan Kallweit, and Alexander Muck. EW corrections to Higgs strahlung at the Tevatron and the LHC with HAWK. *PoS*, EPS-HEP2011:235, 2011.

[207] The ATLAS Collaboration. Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C73(8):2518, 2013.

[208] The ATLAS Collaboration. Jet energy resolution in proton-proton collisions at $\sqrt{s} =$ 7 TeV recorded in 2010 with the ATLAS detector. *Eur. Phys. J.*, C73(3):2306, 2013.

[209] S. Dittmaier et al. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. 2011.

[210] Iain W. Stewart and Frank J. Tackmann. Theory Uncertainties for Higgs and Other Searches Using Jet Bins. *Phys. Rev.*, D85:034011, 2012.

[211] F. D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive e+-p Scattering Cross Sections at HERA. *JHEP*, 01:109, 2010.

[212] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002.

[213] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[214] Stefano Frixione, Eric Laenen, Patrick Motylinski, Bryan R. Webber, and Chris D. White. Single-top hadroproduction in association with a W boson. *JHEP*, 07:029, 2008.

[215] The ATLAS Collaboration. Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data. *Phys. Rev.*, D90(5):052004, 2014.

[216] G. Bohm and G. Zech. *Introduction to statistics and data analysis for physicts.*

[217] S. Schael et al. Search for neutral MSSM Higgs bosons at LEP. *Eur. Phys. J.*, C47:547–587, 2006.

[218] T. Aaltonen et al. Search for neutral higgs bosons in events with multiple bottom quarks at the Tevatron. *Phys. Rev. D*, 86:091101, Nov 2012.

[219] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[220] ATLAS Run 1 Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014.

[221] Tanju Gleisberg and Stefan Hoeche. Comix, a new matrix element generator. *JHEP*, 12:039, 2008.

[222] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012.

[223] Steffen Schumann and Frank Krauss. A Parton shower algorithm based on Catani-Seymour dipole factorisation. *JHEP*, 03:038, 2008.

[224] Stefan Hoeche, Frank Krauss, Marek Schonherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013.

[225] Leif Lonnblad. Correcting the color dipole cascade model with fixed order matrix elements. *JHEP*, 05:046, 2002.

[226] J Butterworth, E Dobson, U Klein, B Mellado Garcia, T Nunnemann, J Qian, D Rebuzzi, and R Tanaka. Single Boson and Diboson Production Cross Sections in pp Collisions at sqrts=7 TeV. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Aug 2010.

[227] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order. *Comput. Phys. Commun.*, 182:2388–2403, 2011.

[228] Comparison of Monte Carlo generator predictions to ATLAS measurements of top pair production at 7 TeV. Technical Report ATL-PHYS-PUB-2015-002, CERN, Geneva, Jan 2015.

[229] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014.

[230] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann. HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR. *Comput. Phys. Commun.*, 182:1034–1046, 2011.

[231] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. MÃ¶lbitz, P. Rieck, and P. Uwer. HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions. *Comput. Phys. Commun.*, 191:74–89, 2015.

[232] Expected Performance of Boosted Higgs ($\rightarrow b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-035, CERN, Geneva, Aug 2015.

[233] Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015.

[234] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.

[235] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. Rivet user manual. *Comput. Phys. Commun.*, 184:2803–2819, 2013.

[236] Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a $b\bar{b}$ pair in pp collisions at 13 TeV using the ATLAS detector. Technical Report ATLAS-CONF-2016-091, CERN, Geneva, Aug 2016.

[237] The ATLAS Collaboration. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 09:145, 2014.

[238] Michael H. Seymour and Andrzej Siodmok. Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data. *JHEP*, 10:113, 2013.

[239] Multi-Boson Simulation for 13 TeV ATLAS Analyses. Technical Report ATL-PHYS-PUB-2016-002, CERN, Geneva, Jan 2016.

[240] Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2. Technical Report ATL-PHYS-PUB-2016-003, CERN, Geneva, Jan 2016.

[241] S. Chatrchyan et al. Search for the standard model higgs boson produced in association with a w or a z boson and decaying to bottom quarks. *Phys. Rev. D*, 89:012003, Jan 2014.

[242] John Ellis, Veronica Sanz, and Tevong You. Complete Higgs Sector Constraints on Dimension-6 Operators. *JHEP*, 07:036, 2014.

[243] Anke Biekotter, Alexander Knochel, Michael Kramer, Da Liu, and Francesco Riva. Vices and virtues of Higgs effective field theories at large energy. *Phys. Rev.*, D91:055029, 2015.

[244] Celine Degrande, Benjamin Fuks, Kentarou Mawatari, Ken Mimasu, and Veronica Sanz. Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD. 2016.