



Search for the Standard Model Higgs boson
in the $H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}$ final state
with the ATLAS experiment and
study of its spin and parity quantum numbers

Manuela Venturi



Dissertation
zur Erlangung des Doktorgrades der
Fakultät für Mathematik und Physik der
Albert-Ludwigs-Universität
Freiburg im Breisgau

DISSERTATION

**Search for the Standard Model Higgs boson
in the $H \rightarrow W^+W^- \rightarrow \ell^+\nu \ell^-\bar{\nu}$ final state
with the ATLAS experiment and study
of its spin and parity quantum numbers**

Manuela Venturi

**Search for the Standard Model Higgs boson
in the $H \rightarrow W^+W^- \rightarrow \ell^+\nu \ell^-\bar{\nu}$ final state
with the ATLAS experiment
and study of its spin
and parity quantum numbers**

DISSERTATION

zur Erlangung des Doktorgrades der
Fakultät für Mathematik und Physik der

ALBERT-LUDWIGS-UNIVERSITÄT
Freiburg im Breisgau

vorgelegt von
Manuela Venturi

January 2014

Dekan: Prof. Dr. Michael Růžička
Betreuer der Arbeit: Prof. Dr. Karl Jakobs
Referent: Prof. Dr. Karl Jakobs
Prüfer: Prof. Dr. Gregor Herten
Prof. Dr. Harald Ita
Prof. Dr. Karl Jakobs

Datum der mündlichen Prüfung:
26. Februar 2014

Considerate la vostra semenza:
fatti non foste a viver come bruti,
ma per seguire virtute e canoscenza.

Dante Alighieri
Inferno - Canto XXVI

Contents

Introduction	ix
1 Electroweak Symmetry Breaking in the Standard Model	1
1.1 Matter fields and interactions	1
1.1.1 Quantum electrodynamics	3
1.1.2 Quantum Chromodynamics	5
1.1.3 Electroweak interactions	7
1.1.4 Massless Standard Model lagrangian	10
1.2 The Higgs mechanism	11
1.2.1 The Goldstone theorem	11
1.2.2 Spontaneous symmetry breaking in the Standard Model	12
1.2.3 Yukawa terms	13
1.2.4 Free parameters of the theory	13
1.3 Constraints on the Higgs boson mass	14
1.3.1 Theoretical constraints	14
1.3.2 Experimental constraints before the LHC	17
1.4 Phenomenology of hadron collisions	20
1.4.1 Factorization	20
1.4.2 Parton distribution functions	23
1.4.3 Monte Carlo generators	25
2 Higgs Boson Physics	27
2.1 Production modes	27
2.1.1 The gluon-fusion production mechanism	29
2.1.2 The vector-boson fusion production mechanism	31
2.1.3 The associate production mechanism	32
2.2 Decay modes	33
2.2.1 The $H \rightarrow f\bar{f}$ channel	34
2.2.2 The $H \rightarrow \gamma\gamma$ channel	35
2.2.3 The $H \rightarrow V^{(*)}V^{(*)}$ channel	36
2.3 Spin and CP quantum numbers	37
2.3.1 The case of a spin-0 resonance	37
2.3.2 The case of a spin-1 resonance	41

2.3.3	The case of a spin-2 resonance	41
3	LHC, ATLAS and Performance	49
3.1	The Large Hadron Collider	49
3.1.1	Past and future operations	52
3.2	The ATLAS detector	52
3.2.1	Inner detector	54
3.2.2	Calorimeters	58
3.2.3	Muon spectrometer	61
3.2.4	Trigger system	62
3.3	Performance	63
3.3.1	Track reconstruction	63
3.3.2	Vertex reconstruction	64
3.3.3	Electrons	65
3.3.4	Muons	69
3.3.5	Jets	71
3.3.6	b -Jets	73
3.3.7	Missing transverse energy	74
4	Search for the $H \rightarrow WW^{(*)}$ Decay Using a Multivariate Analysis	77
4.1	Signal topology and analysis categories	78
4.2	Background topologies	80
4.2.1	The W^+W^- background	80
4.2.2	The top-quark backgrounds	82
4.2.3	The Z + jets background	83
4.2.4	The W + jets background	85
4.2.5	The diboson ($WZ, ZZ, W\gamma$) backgrounds	85
4.3	Monte Carlo samples	86
4.4	Object selection	88
4.4.1	Data-taking conditions	88
4.4.2	Event cleaning	88
4.4.3	Trigger requirements	89
4.4.4	Electron definition	89
4.4.5	Muon definition	90
4.4.6	Jet definition	90
4.5	Event preselection	91
4.6	The Boosted Decision Tree technique	96
4.6.1	BDT settings	98
4.6.2	BDT input variables	99
4.7	BDT response	104
4.7.1	Signal region for the cut-based analysis	104
4.7.2	Overtraining checks	105
4.7.3	Signal region in 0-jet	106
4.7.4	Signal region in 1-jet	109

4.7.5	WW control region in 0-jet	112
4.7.6	WW control region in 1-jet	114
4.7.7	Top-quark control region in 1-jet	115
4.8	Background estimation	117
4.8.1	Top-quark backgrounds in 0-jet	117
4.8.2	Z/γ^* +jets background	117
4.8.3	W +jets background	117
4.9	Systematic uncertainties	119
4.9.1	Experimental uncertainties	120
4.9.2	Theoretical uncertainties	123
4.9.3	Uncertainties from data-driven methods and control regions	125
4.10	Statistical model and fit procedure	126
4.10.1	Remapping	127
4.10.2	Construction of the likelihood function	130
4.10.3	Normalization, shape and statistical uncertainties	131
4.10.4	Test statistic and CL_s procedure	132
4.11	Results	134
4.11.1	CL_s exclusion limits	135
4.11.2	Fitted signal strength	143
4.11.3	Local p_0 -values	144
5	Study of the Spin of the Higgs Boson in the $H \rightarrow WW^*$ Channel	147
5.1	Monte Carlo samples	149
5.2	Object selection	151
5.2.1	Data-taking conditions	151
5.2.2	Trigger requirements	151
5.2.3	Electron definition	151
5.2.4	Muon definition	152
5.2.5	Jet definition	153
5.3	Event preselection	153
5.4	BDT training	156
5.4.1	Input variables	157
5.4.2	BDT settings	160
5.5	BDT Output	161
5.5.1	BDT output in the signal region	161
5.5.2	WW control region	165
5.5.3	$Z \rightarrow \tau\tau$ control region	168
5.5.4	Other backgrounds	170
5.6	Systematic uncertainties	171
5.6.1	Experimental uncertainties	171
5.6.2	Theoretical uncertainties	174
5.6.3	Uncertainties from data-driven methods and control regions	174
5.6.4	Theoretical uncertainty on the WW background shape	176
5.6.5	Higgs boson transverse momentum	177

5.6.6	Impact on the analysis	177
5.7	Statistical model and fit procedure	179
5.7.1	Likelihood and test statistic	179
5.7.2	Choice of the binning	182
5.8	Results for the spin-2 analysis	183
5.8.1	Post-fit yields	183
5.8.2	Post-fit BDT distributions	185
5.8.3	Pulls in nuisance parameters	185
5.8.4	Signal strength and compatibility with the rate measurement	188
5.8.5	Sensitivities	189
5.9	Results for the spin-1 analysis	190
5.9.1	Post-fit yields	190
5.9.2	Post-fit BDT distributions	192
5.9.3	Pulls in nuisance parameters	192
5.9.4	Signal strengths	194
5.9.5	Sensitivities	196
5.10	Spin Combination	197
5.10.1	The $H \rightarrow \gamma\gamma$ analysis	198
5.10.2	The $H \rightarrow ZZ^* \rightarrow 4\ell$ analysis	201
5.10.3	Combined results	203
A	Comparison with other multivariate techniques	209
B	Choice of the input variables	211
C	Input variables for the individual lepton flavours	213
D	BDT with two variables	217
E	BDT with the WW background as signal	220
F	Topology of signal-like and background-like events	222
G	BDT response for the cut-based candidates	227
H	Additional details on limits	228
I	Post-fit nuisance parameters	236
J	Choice of the training configuration for the spin analysis	238
K	The theoretical systematic uncertainty on the shape of the WW background	242

L	Additional studies on the shape of the WW background	249
L.1	Compatibility of the MC@NLO and POWHEG predictions with the data	249
L.2	Constraint to UEPS from the signal-like region only	253
M	Training on backgrounds	255
M.1	WW background	255
M.2	$t\bar{t}$ background	255
M.3	$Z/\gamma^* + \text{jets}$ background	256
N	Definition of the test statistic: comparison to the PLR method	258
O	Results for all the $J^P = 2^+$ fractions	262
	Conclusions	269
	Bibliography	273
	Acknowledgements	291

Introduction

The Standard Model of particle physics represents a major success in the attempt of describing the nature of elementary particles and of their interactions. It is based on two cornerstones, the quantum chromodynamics and the electroweak theory.

Quantum chromodynamics [1–5] acts among quarks and gluons. The latter are eight spin-1 massless bosons, which are responsible for mediating the strong interaction. Quantum chromodynamics is responsible for the confinement of quarks into hadrons and hence for the stability of ordinary matter, made up of protons and neutrons.

The electroweak theory has been formulated by Glashow, Salam and Weinberg [6–9] in order to unify the electromagnetic and weak interactions. The former acts between electrically charged particles and is mediated by a massless spin-1 boson, the photon, while the latter acts on all known fermions, has a shorter range and a weaker strength and is mediated by massive gauge particles, the $W^\pm$ and Z bosons. The electroweak theory recognizes both interactions as manifestations of the same fundamental force, related to the symmetry group $SU(2)_L \times U(1)_Y$, where L states the fact that this interaction only acts between left-handed fermions (and right-handed antifermions).

The principle of local gauge invariance, which is a fundamental concept of the Standard Model, would prevent the presence of massive bosons and fermions, if an additional mechanism would not come into play, i.e. the Brout-Englert-Higgs mechanism of spontaneous symmetry breaking [10–14]. The latter generates the masses of all known particles in a unified and elegant way, preserving the requirements of renormalizability [15] and unitarity [16, 17] of the theory. To this extent, a doublet of complex scalar fields is introduced, with a neutral component which develops a non-zero vacuum expectation value, spontaneously breaking the $SU(2)_L \times U(1)_Y$ symmetry into the electromagnetic $U(1)_{\text{QED}}$ symmetry. The same mechanism explains the masses of the W and Z bosons and of the fermions, via the interaction with the newly postulated Higgs field.

The search for the Higgs boson, whose mass is a free parameter of the theory, has represented a challenge for particle physics over the last four decades, being regarded, as time went by, as the last missing ingredient of the Standard Model. Before the LHC experiments started their operations, direct and indirect constraints on its mass existed: the direct constraints (in particular from the LEP [18] and the Tevatron [19] experiments) had ruled out most of the high-mass range, with a lower threshold at $m_H > 114.4$ GeV at 95% CL, while the indirect constraints [20] had excluded the range above ~ 144 GeV at 95% CL.

One of the main physics motivations for the design and the construction of the Large Hadron Collider (LHC) [21–23], a circular proton accelerator located at CERN (Geneva), has been to prove or conclusively rule out the existence of the Higgs boson. With an energy in the centre-of-mass ranging from $\sqrt{s} = 7$ TeV in 2011 to $\sqrt{s} = 14$ TeV during the future run starting in 2015, the LHC meets all the requirements to produce a significant number of Higgs bosons over the whole allowed mass range. The ATLAS experiment [24–26], in particular, is well equipped to detect leptons and jets with very high accuracy, over a vast range of momenta, besides ensuring, thanks to its hermiticity, a precise determination of the missing energy, in the plane transverse to the beam axis. It follows that the ATLAS experiment can reconstruct all characteristic decays of a Standard Model Higgs boson which involve leptons or photons in the final state. Among them, the $H \rightarrow WW^{(*)}$ final state has the largest branching ratio for masses above $m_H \gtrsim 135$ GeV.

This thesis presents a study of the $H \rightarrow WW^{(*)} \rightarrow \ell^+ \nu \ell^- \bar{\nu}$ final state, with ℓ denoting electrons and muons, using data collected by the ATLAS experiment, corresponding to an integrated luminosity of 4.7 fb^{-1} at $\sqrt{s} = 7$ TeV in 2011, and of 20.7 fb^{-1} at $\sqrt{s} = 8$ TeV in 2012. With the first part of the data, at $\sqrt{s} = 7$ TeV, a multivariate search for the Standard Model Higgs boson in the range $110 \text{ GeV} < m_H \leq 600 \text{ GeV}$ is performed. It employs the Boosted Decision Tree (BDT) technique, which is based on the standard cut-based analysis in the same channel, but is expected to outperform the latter since it takes into account the correlations among the variables characterizing the final-state leptons. Through a series of binary splittings, the events are classified as signal-like or background-like, so that an increased significance can be extracted from the signal-like part of the phase space. The work summarized in this thesis constitutes a leading contribution to the conception, to the development and to the validation of the $H \rightarrow WW^{(*)}$ BDT analysis, up to the final results, which have been published in Ref. [27]. This analysis confirms and improves the results of the corresponding cut-based analysis [28],

leading to an exclusion for a Standard Model Higgs boson, with a luminosity corresponding to 4.7 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$, in the range $130 \text{ GeV} \leq m_H \leq 281 \text{ GeV}$ at 95% CL, to be compared with an expectation of $127 \text{ GeV} \leq m_H \leq 254 \text{ GeV}$.

In July 2012, after having analysed an integrated luminosity corresponding to $\sim 5.9 \text{ fb}^{-1}$ at $\sqrt{s} = 8 \text{ TeV}$, a first observation was made by the ATLAS [29, 30] and CMS [31] collaborations, of an excess of events at a mass close to 125 GeV, compatible with the sought Standard Model Higgs boson, in the $\gamma\gamma$ and $ZZ^* \rightarrow 4\ell$ final states, with a significance of 4.5 and 3.4 standard deviations, respectively. Shortly after, the same excess was observed by the ATLAS collaboration also in the $WW^{(*)} \rightarrow \ell^+\nu\ell^-\bar{\nu}$ final state [32], with a significance of 2.8 standard deviations using an integrated luminosity of 13.0 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$. The BDT analysis, run on the same dataset, was proved to reproduce the signal strength measured by the cut-based analysis, namely $\mu = 1.5 \pm 0.6$ for $m_H = 125 \text{ GeV}$, in units of the Standard Model cross section. While a cut-based approach seemed good enough to increase the accuracy on μ and the significance of the observation, the multivariate analysis developed in 2011 was the best suited for studying the quantum numbers of the new resonance, and in particular its spin and parity properties.

The aim is to test the Standard Model hypothesis, $J^P = 0^+$, against alternative assignments, namely $1^+, 1^-, 2^+$. The $J^P = 2^+$ hypothesis follows the model for Kaluza-Klein-like gravitons [33–36], and is investigated in five possible values for the couplings of the Higgs-like resonance to quarks and gluons (from pure gg production to pure $q\bar{q}$). For what concerns the $J^P = 1^\pm$ hypothesis, it is worthwhile to test it in the $H \rightarrow WW^*$ final state, even if the Landau-Yang theorem [37, 38] forbids the decay of a $J^P = 1^\pm$ boson into two on-shell photons, in contradiction with the evidence achieved by the ATLAS collaboration in the $\gamma\gamma$ channel [39]. In fact, as predicted in various extensions of the Standard Model, there could be two resonances, close in mass, one decaying into photons, the other into vector bosons.

Most of the fundamental aspects of the $H \rightarrow WW^*$ spin analysis have been developed in the context of this thesis, from the very beginning to the final results, published in Ref. [40], which have later served as a leading contribution to the ATLAS spin combination [41]. The latter paper represents the first evidence for the spin-0 nature of the new particle, which is needed in order to identify it with the Standard Model Higgs boson. For this reason, it has been cited in the official motivation for the award of the 2013 Nobel Prize in Physics to Peter Higgs and François Englert [42].

This thesis is organised as follows. In Chapter 1, the Standard Model is briefly

introduced, with an overview of the general characteristics of elementary interactions, the Goldstone theorem and the Higgs mechanism. The pre-LHC constraints on the Higgs boson mass are then summarized, both from the theoretical and the experimental points of view. Finally, the salient features of proton-proton collisions at the LHC are reviewed, including the role of parton distribution functions and of Monte Carlo generators. Chapter 2 focuses on the physics of the Standard Model Higgs boson at the LHC, describing its production and decay modes and discussing the status of the theoretical computations. The second part of the chapter presents an overview of the Higgs boson properties, and in particular of the kinematical and angular distributions which are expected to characterize the final-state particles in the various spin/CP hypotheses. In particular, the decay into vector bosons of the new resonance, $X \rightarrow VV^{(*)}$ with $V = W, Z$, is discussed, for X being either a spin-0, spin-1 or spin-2 boson, with special emphasis on the models which are later used in Chapter 5. Chapter 4 presents a multivariate search in the $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell^-\bar{\nu}$ channel, carried out with an integrated luminosity of 4.7 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$ in 2011, describing the analysis method and the final results. Several additional details and cross-checks are given in Appendices A – I. Chapter 5 illustrates the spin analysis, in the same channel, for which an integrated luminosity of 20.7 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$ has been used. Also in this case, only the baseline results are shown in the main body of the text, while additional details are reported in Appendices J – O. Finally, the statistical combination of the three bosonic channels ($H \rightarrow WW^* \rightarrow \ell^+\nu\ell^-\bar{\nu}$, $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$) is described and the final results are presented.

Chapter 1

Electroweak Symmetry Breaking in the Standard Model

In this Chapter, the Standard Model (SM) of particle physics is briefly introduced, with an overview of the general characteristics of elementary particles and their interactions, the Goldstone theorem and the Higgs mechanism. Afterwards, the pre-LHC constraints on the Higgs boson mass are summarized, both from the theoretical point of view (unitarity violation, convergence issues for perturbation theory, triviality bound, vacuum stability) and from the experimental side (global fit to the electroweak parameters, searches at LEP and at the Tevatron). Finally, the salient features of proton-proton collisions at the LHC are discussed, including the role of parton distribution functions and Monte Carlo generators.

If not stated otherwise, the description follows Refs. [43, 44]; Ref. [45] is used for the QCD section, while Ref. [46] is used for the summary of the theoretical and experimental constraints on the Higgs boson mass.

Throughout this work, natural units are used, based on universal constants: in particular, the reduced Planck constant $\hbar$ and the speed of light c are assumed to be $\hbar = c = 1$. In this system, energy, mass and momentum are all measured in electron volts (eV).

1.1 Matter fields and interactions

In the Standard Model, elementary particles are classified into two main categories, according to their spin: half-integer for fermions (which can be divided further into quarks and leptons), integer for bosons (category to which photons, gluons, W and Z bosons belong – alike the Higgs boson).

The fundamental forces acting between elementary particles – the strong, weak and electromagnetic interactions – are all expressed through gauge theories, with the only exception being the gravitational force. In a gauge theory, fields are described by representations of a symmetry group; requiring the Lagrangian to be invariant under local phase transformations of the fields (the so-called principle of local gauge invariance), the interactions between the corresponding particles are generated, mediated through gauge bosons.

The Standard Model of elementary particles is described by a Lagrangian which is invariant under the symmetry group:

$$G_{\text{SM}} = SU(3)_C \times SU(2)_L \times U(1)_Y, \quad (1.1)$$

where $SU(3)_C$ is the colour group of strong interactions, while $SU(2)_L$ and $U(1)_Y$ are the electroweak groups of isospin and hypercharge, respectively. The G_{SM} group is spontaneously broken to $SU(3)_C \times U(1)_{\text{QED}}$, as discussed in the following. The gauge bosons corresponding to the unbroken groups are massless: eight gluons for $SU(3)_C$ and one photon for $U(1)_{\text{QED}}$, respectively. The mediators of the weak force are, the $W^\pm$ and Z bosons, are, on the other hand, massive.

The fermionic content of the SM is organized into three generations, which have identical properties, but increasing masses and different flavour quantum numbers:

$$\begin{aligned} & \left(\begin{smallmatrix} \nu_e \\ e^- \end{smallmatrix} \right)_L, e_R^-, \left(\begin{smallmatrix} u \\ d \end{smallmatrix} \right)_L, u_R, d_R \\ & \left(\begin{smallmatrix} \nu_\mu \\ \mu^- \end{smallmatrix} \right)_L, \mu_R^-, \left(\begin{smallmatrix} c \\ s \end{smallmatrix} \right)_L, c_R, s_R \\ & \left(\begin{smallmatrix} \nu_\tau \\ \tau^- \end{smallmatrix} \right)_L, \tau_R^-, \left(\begin{smallmatrix} t \\ b \end{smallmatrix} \right)_L, t_R, b_R \end{aligned} \quad (1.2)$$

with left-handed particles belonging to $SU(2)_L$ doublets, while right-handed particles are singlets (left- and right-handed particles are identified by the subscripts L and R, respectively). Furthermore, antiparticles are defined as having the same mass as their corresponding particles, but opposite additive quantum numbers. Quarks exist in six flavours (u , d , s , c , b and t) and three colours (red, green and blue). Stable matter is made up of fermions belonging to the first generation, while the electrically charged fermions of the second and third generations are unstable and decay into lighter particles. Table 1.1 summarizes masses, charges and spin quantum numbers for the SM particles [47]. In the following, additional details are given concerning the electromagnetic, electroweak and strong interactions.

	Name	Mass	Charge (e)	Spin
e	Electron	0.511 MeV	-1	1/2
ν_e	Electron Neutrino	< 2 eV (95% CL)	0	1/2
μ	Muon	105.66 MeV	-1	1/2
ν_μ	Muon Neutrino	< 0.19 MeV (90% CL)	0	1/2
τ	Tau	1.777 GeV	-1	1/2
ν_τ	Tau Neutrino	< 18.2 MeV (95% CL)	0	1/2
u	Up	$2.3^{+0.7}_{-0.5}$ MeV	2/3	1/2
d	Down	$4.8^{+0.5}_{-0.3}$ MeV	-1/3	1/2
c	Charm	1.275 ± 0.025 GeV	2/3	1/2
s	Strange	96 ± 5 MeV	-1/3	1/2
t	Top	173.07 ± 0.89 GeV	2/3	1/2
b	Bottom	4.18 ± 0.03 GeV	-1/3	1/2
γ	Photon	0	0	1
$W^\pm$	W Boson	80.385 ± 0.15 GeV	± 1	1
Z	Z Boson	91.1876 ± 0.0021 GeV	0	1
g	Gluon	0	0	1
H	Higgs Boson	—	0	0

Table 1.1: Summary of Standard Model particles, with the indication of their mass, charge and spin quantum numbers. Electric charges and spins correspond to the SM prediction. All masses are given according to the world average in Ref. [47] (apart from photons and gluons, which are predicted to be massless): for light quarks, they are estimated at a scale $\mu \simeq 2$ GeV, m_c and m_b are the running masses in the $\overline{\text{MS}}$ scheme (cf. Section 2.2.1), while m_t is measured by the combined Tevatron experiments. The mass of the Higgs boson is discussed later, hence omitted here.

1.1.1 Quantum electrodynamics

A fermion of mass m and spin 1/2 is described by the free Dirac Lagrangian:

$$\mathcal{L}_0 = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) \quad (1.3)$$

where $\psi(x)$ is the fermionic spinor and $\bar{\psi}(x) = \psi^\dagger(x)\gamma^0$ its adjoint. The partial derivative is defined as $\partial_\mu = \partial/\partial x_\mu$ and a sum is assumed when indices are repeated. The Lagrangian in Eq. 1.3 is not invariant under local phase transformations such as $\psi(x) \rightarrow e^{iQ\theta(x)}\psi(x)$, with Q being a real constant, because this would introduce an additional term proportional to $iQ\partial_\mu\theta(x)$. The local gauge invariance is restored if, in Eq. 1.3, the standard derivative is substituted by the covariant one:

$$D_\mu\psi(x) = [\partial_\mu + ieQA_\mu(x)]\psi(x) \quad (1.4)$$

which contains a new spin-1 field, the electromagnetic field $A_\mu(x)$. The covariant derivative applied to the field transforms like the fermionic field itself, $D_\mu\psi(x) \rightarrow e^{iQ\theta(x)}(D_\mu\psi(x))$, provided that $A_\mu(x)$ transforms as $A_\mu(x) \rightarrow A_\mu(x) - (1/e)\partial_\mu\theta$. The

Dirac Lagrangian contains now a vectorial gauge field interacting with the fermionic field, with a coupling given by the electric charge eQ :

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) = \mathcal{L}_0 - eQA_\mu(x)\bar{\psi}(x)\gamma^\mu\psi(x). \quad (1.5)$$

Introducing a kinetic term for the gauge field, the latter becomes a propagating field: the full Lagrangian of Quantum Electrodynamics (QED) is then given by:

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_0 - eQA_\mu(x)\bar{\psi}(x)\gamma^\mu\psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) \quad (1.6)$$

where $F_{\mu\nu}(x) = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength. By means of the Euler-Lagrange equation, applied to the QED Lagrangian, Maxwell equations are recovered, $\partial_\mu F^{\mu\nu}(x) = eJ^\nu = eQ\bar{\psi}\gamma^\nu\psi$.

It should be noticed that the inclusion of a mass term for the electromagnetic field, proportional to $m^2 A_\mu A^\mu$, would break the local gauge invariance and is, therefore, forbidden, resulting in the prediction of massless photons.

The elementary QED process is $e^- \rightarrow e^- + \gamma$, depicted in Fig. 1.1 (a), and proceeds through an electron (or any other charged lepton, or a quark) emitting (or absorbing) a photon. Modifying the time arrow, i.e. changing the topological configuration of the decay, the other QED elementary processes, e.g. $e^- + e^+ \rightarrow \gamma$, are constructed; combining more of these vertices, all QED processes are obtained.

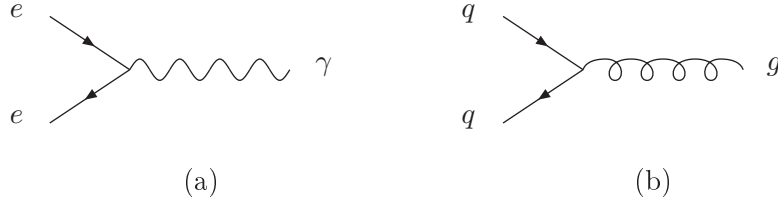


Figure 1.1: Elementary vertices: (a) QED vertex for $e \rightarrow e\gamma$, (b) QCD vertex for $q \rightarrow qg$. Time flows vertically downwards in these diagrams.

The QED vertex in Fig. 1.1 (a) is at lowest order in the QED coupling constant, $\alpha_{\text{QED}} = e^2/(4\pi)$. Adding virtual corrections introduces ultraviolet divergences; for instance, the photon self-energy contribution, depicted in Fig. 1.2 for the e^-e^- scattering, consists of a finite part and of a divergent one, proportional to $\ln(-q^2/\mu^2)$, where q is the photon momentum and μ is some reference energy scale, which diverges as $q \rightarrow \infty$. The infinity can be reabsorbed into a redefinition of masses, charges and coupling constants; namely, the scattering amplitude, which at lowest order depends

on α/q^2 (omitting the QED subscript), becomes:

$$\frac{\alpha_R(\mu^2)}{q^2} (1 - \Pi_R(q^2/\mu^2)) \quad (1.7)$$

where $\alpha_R(\mu^2)$ is the *renormalized running* coupling at the scale μ^2 , while $\Pi_R(q^2/\mu^2)$ is the renormalized self-energy correction. Both terms in Eq. 1.7 depend on the arbitrary scale μ and on the chosen renormalization scheme, but the physical scattering amplitude does not. All ultraviolet divergent contributions to all scattering processes are in fact eliminated through the redefinition of the running coupling, a fact that ensures the renormalizability of QED. Because of its running nature, the value of the fine structure constant, usually defined at the electron scale, $\alpha(\mu^2 = m_e^2) = 1/137$, changes at larger scales; its evolution is determined by the β function:

$$\mu \frac{d\alpha}{d\mu} = \alpha \beta(x) = \alpha \left(\beta_1 \frac{\alpha}{\pi} + \beta_2 \left(\frac{\alpha}{\pi} \right)^2 + \dots \right) \quad (1.8)$$

whose solution is, at one-loop level and assuming $Q^2 = \mu^2$:

$$\alpha(Q^2) = \frac{\alpha(Q_0^2)}{1 - \frac{\beta_1 \alpha(Q_0^2)}{2\pi} \ln(Q^2/Q_0^2)}. \quad (1.9)$$

Since $\beta_1 = 2/3 > 0$, $\alpha(Q^2)$ increases for increasing energies, which also means that the effect of the electromagnetic charge decreases at larger distances from the charged particle.

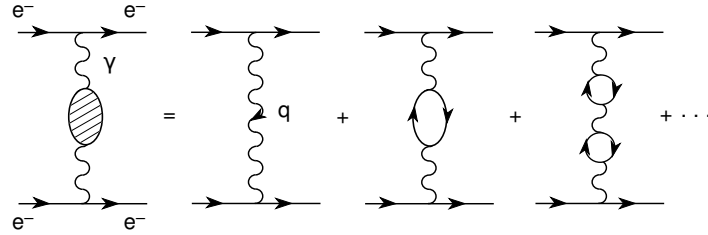


Figure 1.2: Photon self-energy contribution to e^-e^- scattering [45].

1.1.2 Quantum Chromodynamics

The same idea of local gauge invariance, that leads to QED in case of a $U(1)$ symmetry, results in Quantum Chromodynamics (QCD) when considering $SU(3)_C$ as a symmetry group. C labels the colour charge: its introduction as an additional degree of freedom [3, 48, 49] is needed in order to explain the observation of baryonic states such as Δ^{++} , which otherwise would be constituted by three u quarks in an

identical quantum state and would thus violate the Pauli exclusion principle. The observation that only colour singlets exist as asymptotic states can then be justified assuming that quarks are confined into hadrons, either in the form of baryons, $B \propto \epsilon^{\alpha\beta\gamma}|q_\alpha q_\beta q_\gamma\rangle$, or of mesons, $M \propto \delta^{\alpha\beta}|q_\alpha \bar{q}_\beta\rangle$.

The elementary QCD interaction is $q \rightarrow qg$, depicted in Fig. 1.1 (b). The quark colour may change in this interaction and, since colour is conserved, the difference is carried away by the gluon. The free QCD Lagrangian describing a quark field q_f of flavour f is given by:

$$\mathcal{L}_0 = \sum_f \bar{q}_f (i\gamma^\mu \partial_\mu - m_f) q_f. \quad (1.10)$$

This is similar to Eq. 1.3, but now q_f is a vector in colour space, q_f^α ; under global $SU(3)_C$ transformations ($q_f^\alpha \rightarrow U^\alpha_\beta q_f^\beta$), the free Lagrangian in Eq. 1.10 is invariant, provided that U is unitary ($UU^\dagger = U^\dagger U = 1$ and $\det U = 1$). The matrix U can be parametrized as $U = \exp(i\lambda^a \theta_a/2)$, in terms of the eight generators of the $SU(3)$ algebra, $\lambda^a = 2T^a$, which are traceless and Hermitian, and satisfy the commutation relation $[T^a, T^b] = if^{abc}T^c$, with f^{abc} being the totally antisymmetric structure constants.

As in the QED case, requiring the Lagrangian to be invariant under arbitrary local transformations, $\theta_a = \theta_a(x)$, the partial derivative in Eq. 1.10 has to be replaced by its covariant form:

$$D_\mu q_f = \left[\partial_\mu + ig_s \frac{\lambda^a}{2} G_a^\mu(x) \right] q_f. \quad (1.11)$$

Analogously to the QED derivative in Eq. 1.4, the requirement for covariance has introduced the interaction of the quark fields with the eight gluonic fields, G_a^μ . The strong coupling g_s is unique for all of them. As stated in Eq. 1.11, the quark-gluon interaction does not affect the quark flavour; moreover, since leptons are not colour-charged, they do not participate in the strong interaction. The gluon field strength is given by:

$$G^{\mu\nu}(x) = -\frac{i}{g_s} [D^\mu, D^\nu] = \frac{\lambda^a}{2} G_a^{\mu\nu}(x) = \frac{\lambda^a}{2} \left(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f^{abc} G_b^\mu G_c^\nu \right). \quad (1.12)$$

By means of it and the quark fields, the QCD Lagrangian can be written as:

$$\begin{aligned}
\mathcal{L}_{\text{QCD}} = & \sum_f \bar{q}_f (i\gamma^\mu \partial_\mu - m_f) q_f - \frac{1}{4} (\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) \\
& - g_s G_a^\mu \sum_f \bar{q}_f \gamma_\mu \frac{\lambda^a}{2} q_f \\
& + \frac{g_s}{2} f^{abc} (\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) G_\mu^b G_\nu^c - \frac{g_s^2}{4} f^{abc} f_{ade} G_b^\mu G_c^\nu G_\mu^d G_\nu^e.
\end{aligned} \tag{1.13}$$

The terms in the first line are the kinetic ones for the free quark and gluon fields, respectively, while the second line contains their mutual interaction. Finally, as a consequence of the non-Abelian nature of the colour group, the terms grouped in the last line describe the self-interactions among the gluon fields: the cubic term, proportional to g_s , corresponding to three-gluon vertices, and the quartic term, proportional to g_s^2 , corresponding to four-gluon vertices. There is no mass term for the gluon fields, because $\propto m_G^2 G_a^\mu G_\mu^a$ would break the covariance of the theory. So gluons are predicted to be massless spin-1 particles.

Gluon self-interactions are responsible for another crucial feature of QCD: the strong running coupling, $\alpha_s(Q^2)$, decreases with decreasing distances, unlike the QED one. Renormalizing the gluon self-energy, as discussed for the photon self-energy before, owes a value for the first β coefficient of $\beta_1 = (2N_f - 11N_c)/6$. While the positive contribution comes from the quark loop corrections to the gluon internal line, the negative one derives from the gluon self-interactions. For $N_c = 3$ colours and $N_f = 6$ flavours, β_1 is negative, so that

$$\lim_{Q^2 \rightarrow \infty} \alpha_s(Q^2) = 0 \tag{1.14}$$

which is known as the *asymptotic freedom* property of QCD. On the other hand, at lower energies, $Q^2 \rightarrow 0$, which is the confinement regime, the running coupling blows up, so that perturbation theory can not be applied anymore.

1.1.3 Electroweak interactions

The unification of electromagnetic and weak interactions has been first formulated by Glashow, Salam and Weinberg [6–9]: the electroweak theory recognizes both interactions as manifestations of the same fundamental force, related to the symmetry group $SU(2)_L \times U(1)_Y$. The label L refers to the fact that only left-handed fermions (and right-handed antifermions) participate in this interaction, as already shown in Eq. 1.2. Referring to quarks of the first generation (the same holds for the other

generations and for the leptonic sector), the free massless Lagrangian is given by:

$$\mathcal{L}_0 = \sum_{j=1}^3 i \bar{\psi}_j(x) \gamma^\mu \partial_\mu \psi_j(x) \quad (1.15)$$

where ψ_1 is a $SU(2)_L$ doublet, while $\psi_{2,3}$ are singlets. The former transforms in flavour space as:

$$\psi_1(x) \rightarrow e^{iy_1\beta} U_L \psi_1(x), \quad (1.16)$$

with $U_L = \exp(i\sigma_i \alpha^i/2)$ being a non-Abelian matrix, written in terms of the Pauli matrices σ_i , and β an arbitrary constant phase. On the other hand, the singlets acquire only a phase in the transformation, $\psi_{2,3} \rightarrow \exp(iy_{2,3}\beta)\psi_{2,3}$. The factors y_i are the hypercharges.

The spinor field can be explicitly broken down into its left- and right-handed components, introducing the chirality projectors $P_{L,R}$, in terms of the matrix $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$:

$$\psi = (P_L + P_R)\psi = \left(\frac{1 - \gamma^5}{2} + \frac{1 + \gamma^5}{2} \right) \psi = \psi_L + \psi_R \quad (1.17)$$

so that the free Lagrangian can be expressed as:

$$\mathcal{L}_0 = i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + i\bar{\psi}_R \gamma^\mu \partial_\mu \psi_R. \quad (1.18)$$

A mass term, proportional to $m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L)$, would break the gauge invariance of the theory, since it would mix the left- and right-handed components.

Requiring the transformation in Eq. 1.16 to be local, the derivative is substituted by the covariant one, namely:

$$D_\mu \psi_1(x) = [\partial_\mu + ig\sigma_i W_\mu^i(x)/2 + ig'y_1 B_\mu(x)] \psi_1(x) \quad (1.19)$$

$$D_\mu \psi_{2,3}(x) = [\partial_\mu + ig'y_{2,3} B_\mu(x)] \psi_{2,3}(x)$$

where four gauge bosons have been introduced, and two couplings, g and g' . The corresponding field strengths are given by $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and $W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk} W_\mu^j W_\nu^k$; the total electroweak Lagrangian can then be written as:

$$\mathcal{L}_{\text{EW}} = \sum_{j=1}^3 i \bar{\psi}_j(x) \gamma^\mu D_\mu \psi_j(x) - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu}. \quad (1.20)$$

Cubic and quartic interactions among the gauge bosons are present, due to the quadratic term in $W_{\mu\nu}^i$. Like the fermions, also the gauge bosons are predicted to

be massless.

The Lagrangian in Eq. 1.20 contains both charged and neutral interactions. The former ones, which derive from the term proportional to $g\bar{\psi}_1\gamma^\mu\sigma_i W_\mu^i\psi_1$, have the form, referring only to quarks and leptons of the first generations:

$$\mathcal{L}_{\text{CC}} = -\frac{g}{2\sqrt{2}} \left\{ W_\mu^\dagger [\bar{u}\gamma^\mu(1 - \gamma_5)d + \bar{\nu}_e\gamma^\mu(1 - \gamma_5)e] + \text{h.c.} \right\} \quad (1.21)$$

where $W_\mu = (W_\mu^1 + iW_\mu^2)/\sqrt{2}$ and its complex conjugate $W_\mu^\dagger = (W_\mu^1 - iW_\mu^2)/\sqrt{2}$ are the physical $W^\pm$ bosons. The vertex contains a factor $\gamma^\mu(1 - \gamma_5)$, therefore weak charged currents are of the V-A type, where V labels the vectorial current (γ^μ) and A the axial one ($\gamma^\mu\gamma_5$). Such an interaction violates both parity ($\mathcal{P}$: $\vec{x} \leftrightarrow -\vec{x}$) and charge conjugation ($\mathcal{C}$: particle $\leftrightarrow$ antiparticle); their combination is, however, still a good symmetry for V-A interactions.

For what concerns neutral interactions, on the other hand, they are expressed in Eq. 1.20 in terms of the fields W_μ^3 and B_μ , which do not correspond to the observed physical states. The latter are rather obtained via a rotation:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad (1.22)$$

where θ_W is the weak mixing angle, so that the neutral-current Lagrangian can be written as:

$$\mathcal{L}_{\text{NC}} = -\sum_j \bar{\psi}_j \gamma^\mu \left\{ A_\mu \left[g \frac{\sigma_3}{2} \sin\theta_W + g' y_j \cos\theta_W \right] + Z_\mu \left[g \frac{\sigma_3}{2} \cos\theta_W - g' y_j \sin\theta_W \right] \right\} \psi_j. \quad (1.23)$$

The correct quantum numbers for QED are obtained imposing $e = g \sin\theta_W = g' \cos\theta_W$, which correlates the electromagnetic and weak couplings. Another equation links the charges: $Y = 2(Q - I^3)$, where Y is the weak hypercharge and I^3 is the third component of the weak isospin. They are generated by the $U(1)_Y$ and $SU(2)_L$ groups, respectively; their values for the SM fermions are summarized in Table 1.2. Right-handed neutrinos are not considered here because they would not interact electroweakly ($Q = I^3 = Y = 0$) and thus would be sterile.

The couplings of Z bosons to fermions are described by the last term in Eq. 1.23 and can be recast into the form:

$$\mathcal{L}_{\text{NC}}^Z = -\frac{e}{2 \sin\theta_W \cos\theta_W} Z_\mu \sum_f \bar{f} \gamma^\mu (c_V^f - c_A^f \gamma_5) f \quad (1.24)$$

where the vectorial and axial neutral couplings have been introduced: $c_V^f = 2I_f^3 -$

$4Q_f \sin^2 \theta_W$ and $c_A^f = 2I_f^3$, respectively.

	$Q(e)$	I	I^3	Y
e_L^-, μ_L^-, τ_L^-	-1	1/2	-1/2	-1
$\nu_{e,L}, \nu_{\mu,L}, \nu_{\tau,L}$	0	1/2	+1/2	-1
e_R^-, μ_R^-, τ_R^-	-1	0	0	-2
u_L, c_L, t_L	2/3	1/2	+1/2	1/3
d_L, s_L, b_L	-1/3	1/2	-1/2	1/3
u_R, c_R, t_R	2/3	0	0	4/3
d_R, s_R, b_R	-1/3	0	0	-2/3

Table 1.2: Overview of the electroweak quantum numbers for the Standard Model fermions: the electromagnetic charge (in units of e), the weak isospin, its third component and the resulting electroweak hypercharge, according to the relation $Y = 2(Q - I^3)$. All particles are taken to be interaction eigenstates.

The quarks which take part in the electroweak interaction (labelled as d', s', b' in Eq. 1.25) are superpositions of the eigenstates of the strong interaction. They can be obtained from the latter via a rotation, as described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [50, 51]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.25)$$

The CKM matrix is unitary and has $(N-1)^2$ free parameters, with $N = 3$ being the number of observed generations: three rotation angles and one complex phase, the latter able to accomodate the amount of CP violation observed in weak interactions, e.g. in the neutral kaon system.

1.1.4 Massless Standard Model lagrangian

Putting together the various terms discussed so far, the complete SM Lagrangian can be written as:

$$\begin{aligned} \mathcal{L}_{\text{SM}} = & -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} + \bar{L}_i iD_\mu \gamma^\mu L_i + \\ & \bar{e}_{Ri} iD_\mu \gamma^\mu e_{Ri} + \bar{Q}_i iD_\mu \gamma^\mu Q_i + \bar{u}_{Ri} iD_\mu \gamma^\mu u_{Ri} + \bar{d}_{Ri} iD_\mu \gamma^\mu d_{Ri} \end{aligned} \quad (1.26)$$

with the index i accounting for the various generations. This Lagrangian is explicitly invariant under local gauge transformations of the G_{SM} group, as in Eq. 1.1. However, gauge bosons and fermions are still massless, since the addition of mass terms for either the fermions or the bosons would break the covariance of the theory, as previously discussed.

The most natural solution to this problem is provided by the Brout-Englert-Higgs mechanism of spontaneous symmetry breaking – or the Higgs mechanism for short [10–14].

1.2 The Higgs mechanism

1.2.1 The Goldstone theorem

The Goldstone theorem [52–54] describes the mechanism of spontaneous symmetry breaking (SSB), according to which a non-symmetric ground state arises from a symmetric Lagrangian, implying the existence of additional massless degrees of freedom.

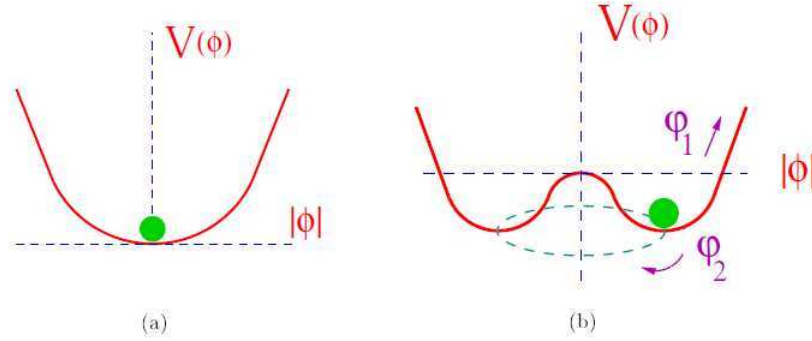


Figure 1.3: Representation of the potential V of the scalar field ϕ , $V = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$, for $\mu^2 > 0$ (a) and $\mu^2 < 0$ (b). Both illustrations are taken from Ref. [43].

Considering the case of a complex scalar field $\phi(x)$, and of its Lagrangian $\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - (\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2)$, the symmetry under which $\mathcal{L}$ is invariant is a global $U(1)$ transformation: $\phi(x) \rightarrow \exp(i\theta)\phi(x)$. The scalar potential has two free parameters, μ^2 and λ : while λ is required to be positive in order to have a ground state, μ^2 can be either positive or negative. The first case, $\mu^2 > 0$, is the trivial one, describing a potential with only one minimum at $|\phi| = 0$ (see Fig. 1.3 (a)); the case $\mu^2 < 0$, on the other hand, corresponds to an infinite set of degenerate minima, as shown in Fig. 1.3 (b), satisfying the condition:

$$|\phi_0| = \sqrt{\frac{-\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}. \quad (1.27)$$

Infinite states $\phi_0 = (v/\sqrt{2})e^{i\theta}$ are at the minimum energy; the symmetry is broken by making a particular choice for the ground state, e.g. $\theta \equiv 0$ in the *unitary gauge*. The complex field ϕ can be expanded around the vacuum state introducing the field

excitations $\phi_{1,2}$:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \phi_1(x) + i\phi_2(x)) \quad (1.28)$$

in terms of which the potential takes the form:

$$V(\phi) = V(\phi_0) - \mu^2 \phi_1^2 + \lambda v \phi_1(\phi_1^2 + \phi_2^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2. \quad (1.29)$$

Out of the two field excitations, ϕ_1 is massive, $m_{\phi_1}^2 = -2\mu^2$, while ϕ_2 results to be massless. This is in fact a general result: according to the Goldstone theorem, whenever a Lagrangian is invariant under a group G but its vacuum state is not, there exist as many massless spin-0 bosons as broken generators.

1.2.2 Spontaneous symmetry breaking in the Standard Model

The same mechanism can be expanded to describe the SM case: the goal is the generation of masses for the W and Z bosons, leaving the photon massless (indeed QED shall remain unbroken).

This can be achieved introducing a doublet of complex scalar fields, with quantum numbers $Y_\phi = I_\phi = 1/2$:

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} \quad (1.30)$$

described by the Lagrangian:

$$\mathcal{L}_S = (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \quad (1.31)$$

which is invariant under a local $SU(2)_L \times U(1)_Y$ transformation. As before, for $\lambda > 0$ and $\mu^2 < 0$ the potential in Eq. 1.31 possesses an infinite number of equivalent minima, satisfying a condition analogous to Eq. 1.27, $|\langle 0 | \phi^0 | 0 \rangle| = v/\sqrt{2}$. This holds only for the neutral component of the doublet, since electric charge has to be conserved. The choice of one particular minimum breaks $SU(2)_L \times U(1)_Y$ into $U(1)_{\text{QED}}$, generating three Goldstone bosons. Expanding ϕ at first order around its minimum v , one gets:

$$\phi(x) = \exp\left(i\frac{\sigma_i}{2}\theta^i(x)\right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.32)$$

which contains four real fields, $\theta^i(x)$ and $H(x)$. Thanks to the $SU(2)$ invariance of the model, the phase in Eq. 1.32 is arbitrary and can be rotated away choosing, as

before, $\theta^i(x) \equiv 0$. In this gauge, the covariant derivative in Eq. 1.31 takes the form:

$$(D_\mu \Phi)^\dagger D^\mu \Phi = \frac{1}{2} \partial_\mu H \partial^\mu H + (v + H)^2 \left\{ \frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right\}. \quad (1.33)$$

The presence of v , the vacuum expectation value, generates a mass for the $W^\pm$ and Z bosons, $m_W = m_Z \cos \theta_W = gv/2$, while the photon remains massless. In other words, by spontaneously breaking $SU(2)_L \times U(1)_Y$ into $U(1)_{QED}$, three massless Goldstone bosons, θ^i , emerged, which have been subsequently absorbed by the $W^\pm$ and Z bosons to form their longitudinal degrees of freedom, and thus their masses.

1.2.3 Yukawa terms

Also fermion masses can be explained through the introduction of the scalar doublet in Eq. 1.30. The following Yukawa-type term results to be gauge-invariant:

$$\mathcal{L}_Y = -c_1(\bar{u}, \bar{d})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R - c_2(\bar{u}, \bar{d})_L \begin{pmatrix} \phi^{0*} \\ -\phi^- \end{pmatrix} u_R - c_3(\bar{\nu}_e, \bar{e})_L \begin{pmatrix} \phi^+ \\ -\phi^0 \end{pmatrix} e_R + \text{h.c.} \quad (1.34)$$

Imposing the unitary gauge again, $\mathcal{L}_Y$ takes the simpler form:

$$\mathcal{L}_Y = -\frac{1}{\sqrt{2}}(v + H)\{c_1 \bar{d}d + c_2 \bar{u}u + c_3 \bar{e}e\} \quad (1.35)$$

which defines the masses of the light fermions to be:

$$(m_d, m_u, m_e) = (c_1, c_2, c_3) \frac{v}{\sqrt{2}} \quad (1.36)$$

while neutrinos stay massless. The couplings c_i are not fixed by the theory, so the masses are free parameters that need to be determined experimentally.

1.2.4 Free parameters of the theory

The scalar and gauge sectors described so far contain only four free parameters: the couplings to the gauge bosons, g and g' (Eq. 1.19), and the two independent terms in the Higgs boson potential, μ and v (Eq. 1.29). One can trade them by G_F , α^{-1} , m_Z and m_H (the first two being respectively the Fermi constant and the fine structure constant), with the advantage of using the three most precisely measured quantities in the SM. The weak mixing angle θ_W can be used to relate couplings and masses:

$$\sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}, \quad \sin^2 \theta_W = 1 - \frac{m_W^2}{m_Z^2} \quad (1.37)$$

while from the measurement of G_F a determination of the vacuum expectation value can be obtained:

$$v = \left(\sqrt{2} G_F \right)^{(-1/2)} = 246 \text{ GeV}. \quad (1.38)$$

The scalar Lagrangian introduced in Eq. 1.31 can be rewritten in terms of the physical fields (in the unitary gauge) as $\mathcal{L}_S = \frac{1}{4}\lambda v^4 + \mathcal{L}_H + \mathcal{L}_{HG^2}$, where:

$$\mathcal{L}_H = \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 H^2 - \frac{m_H^2}{2v} H^3 - \frac{m_H^2}{8v^2} H^4 \quad (1.39)$$

contains the triple and quartic Higgs boson self-interaction terms, and:

$$\mathcal{L}_{HG^2} = m_W^2 W_\mu^\dagger W^\mu \left\{ 1 + \frac{2}{v}H + \frac{H^2}{v^2} \right\} + \frac{1}{2}m_Z^2 Z_\mu Z^\mu \left\{ 1 + \frac{2}{v}H + \frac{H^2}{v^2} \right\} \quad (1.40)$$

the Higgs boson couplings to the gauge bosons, whose strength is proportional to the square of the gauge boson mass. Moreover, all Higgs boson couplings are determined by m_H , m_W , m_Z and v . Fig. 1.4 shows a summary of the Higgs boson couplings to fermions, to massive gauge bosons and to itself.

1.3 Constraints on the Higgs boson mass

1.3.1 Theoretical constraints

Despite being a free parameter in the Standard Model, the Higgs boson mass can be used to derive limits on the stability of the Standard Model itself, and on its range of validity.

The first limit comes from considerations of unitarity violation in the WW scattering processes. Historically, in fact, massive vector bosons were introduced to explain weak interactions, since the simpler Fermi model (contact coupling), postulating a cross section scaling with energy like $\sigma \sim G_F^2 s$, would have violated unitarity above $\sqrt{s} \sim G_F \sim 300 \text{ GeV}$. The same argument holds, at much higher energies, for the scattering processes involving the longitudinal components of the vector bosons, W_L and Z_L , which proceed, among the other contributions, through the exchange of a Higgs boson. In the high-energy limit $s \gg m_W^2, m_H^2$, taking as an example the $W^+W^- \rightarrow W^+W^-$ scattering, the amplitude scales like $a_0 \sim -m_H^2/(8\pi v^2)$, which is unitary only if $m_H < 870 \text{ GeV}$. Hence, the Higgs boson cannot be heavier than that, unless New Physics appears in the TeV range to restore unitarity.

The second limit comes from perturbation theory, which also requires the Higgs boson not to be too heavy. The decay width into longitudinally polarized vector bosons is proportional to $\Gamma_H \propto m_H^3$; when higher-order radiative corrections, involv-

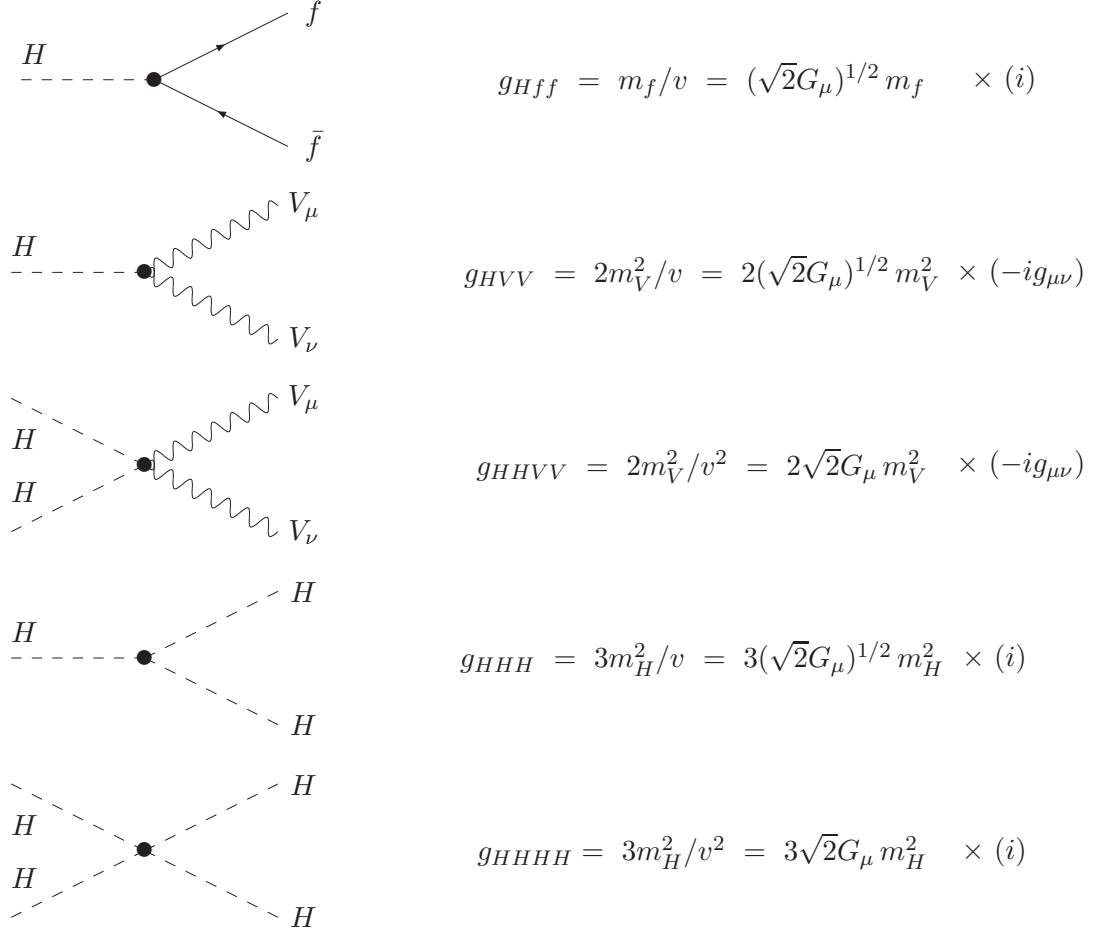


Figure 1.4: The Higgs boson couplings in the SM and their strengths.

ing the exchange of virtual Higgs bosons, are taken into account, the Born width becomes:

$$\Gamma_{\text{tot}} \simeq \Gamma_{\text{Born}} \left[1 + 3\hat{\lambda} + 62\hat{\lambda}^2 + \mathcal{O}(\hat{\lambda}^3) \right] \quad (1.41)$$

with λ being the quartic coupling, and $\hat{\lambda} = \lambda/(16\pi^2)$. For m_H above $\mathcal{O}(10 \text{ TeV})$, this series is not convergent ($3\hat{\lambda} \sim 1$), but already for $m_H \sim 1 \text{ TeV}$ the two-loop term becomes as important as the one-loop term ($3\hat{\lambda} \sim 62\hat{\lambda}^2$). Therefore, perturbation theory holds only if $m_H < 1 \text{ TeV}$. In addition, it should be noticed that, already at the TeV scale, the total width and the mass become comparable, $\Gamma_H \sim m_H$, so that, in this regime, the Higgs boson cannot be considered as a true resonance anymore.

The third limit involves again the quartic coupling λ and its one-loop corrections. The running of λ with the energy scale Q is described by, choosing v as the natural

reference energy:

$$\lambda(Q^2) = \lambda(v^2) \left[1 - \frac{3}{4\pi^2} \lambda(v^2) \log \frac{Q^2}{v^2} \right]^{-1} \quad (1.42)$$

which vanishes for $Q^2 \ll v^2$, at which point the theory becomes trivial, i.e. non interacting. On the other hand, in the opposite limit $Q^2 \gg v^2$, λ diverges, becoming infinite at the Landau pole Λ_C , implying that the Standard Model would not be valid for higher energies. Since $\Lambda_C = v \exp(4\pi^2 v^2 / m_H^2)$, m_H needs to be small to push Λ_C at energies large enough; for instance, $\Lambda_C = \Lambda_{\text{GUT}} \sim 10^{16}$ GeV (which is the scale of Grand Unification) would imply $m_H \lesssim 200$ GeV. Setting the cut-off at m_H itself, computations on the lattice (able to take into account also the non perturbative effects) yield the triviality bound $m_H < 640$ GeV.

The fourth limit is a lower bound, which can be set when considering that not only the Higgs boson self-couplings contribute to the running of λ , but also the couplings to gauge bosons and to fermions. Since, as shown already in Fig. 1.4, the couplings depend on the mass, the main contributions come from the top-quark loop and loops of massive gauge bosons, but they appear with opposite sign: in particular, since the top-quark loop has a negative sign, it can drive $\lambda(Q^2)$ to negative values if $\lambda(v^2)$ is too small, hence $V(v)$ is not the absolute minimum anymore and the theory is not stable. To avoid this, i.e. to keep $\lambda(Q^2) > 0$, a boundary on m_H is needed, which depends on the value of Λ_C : a low cut-off ($\Lambda_C \sim 10^3$ GeV) would imply $m_H \gtrsim 70$ GeV, a large one ($\Lambda_C \sim 10^{16}$ GeV) would bring it up to $m_H \gtrsim 130$ GeV. It should be noticed, anyhow, that these bounds become weaker in case the vacuum is metastable, rather than stable, with a life longer than the observed lifetime of the Universe.

Putting together the triviality upper bound with the stability lower one, the constraints represented in Fig. 1.5 are obtained, which include all corrections up to the two-loop level, and use as inputs the physical masses of the top-quark and of the gauge bosons, and their couplings [55]. The allowed range for m_H as a function of the cut-off Λ_C is shown; in particular, if New Physics appears at the TeV scale, the following limit can be derived:

$$50 \text{ GeV} \lesssim m_H \lesssim 800 \text{ GeV} \quad \text{for } \Lambda_C \sim 10^3 \text{ GeV} \quad (1.43)$$

while, if the Standard Model is valid up to the GUT scale, this range shrinks significantly:

$$130 \text{ GeV} \lesssim m_H \lesssim 180 \text{ GeV} \quad \text{for } \Lambda_C \sim 10^{16} \text{ GeV} . \quad (1.44)$$

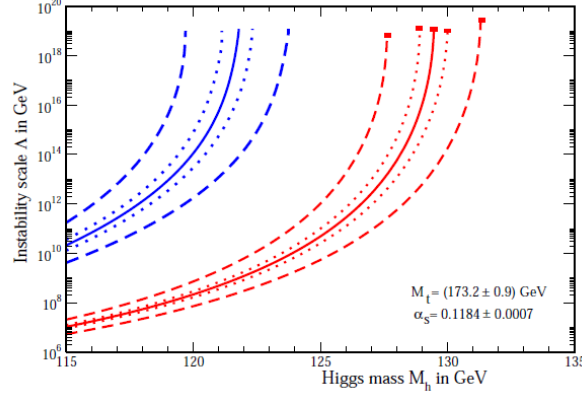


Figure 1.5: Bounds on the Higgs boson mass as a function of the New Physics scale Λ , combining triviality and vacuum stability arguments. The plain red line is obtained for the central value of m_t and α_s , the dashed red one for $\pm 2\sigma$ variations of m_t , the dotted red one for $\pm 2\sigma$ variations of α_s . The same in blue for the metastable scenario [55].

1.3.2 Experimental constraints before the LHC

Indirect searches

The Higgs boson mass is tightly bounded to the other SM parameters, as previously discussed; moreover, radiative corrections to the electroweak observables are sensitive to m_H . For these reasons, a global fit to the SM observables allows to set stringent limits on m_H . To this extent, all quantities accurately predicted by the SM and precisely measured in data are used, namely: $\alpha(m_Z)$, G_μ , m_Z , observables from Z analyses at LEP1, from W analyses at LEP2 and at the Tevatron, the left-right and forward-backward asymmetries for heavy quark decays, $A_{LR,FB}^{b,c}$, as measured at LEP1 and at the SLC, and finally a few low-energy measurements. The result of the global fit, $\Delta\chi^2$, is shown in Fig. 1.6 (a) as a function of m_H ; taking into account the lower bound from direct searches at LEP (discussed in the following), the allowed range at 95% C.L. is:

$$114.4 \text{ GeV} < m_H < 144 \text{ GeV} \quad (1.45)$$

It should be noticed that such a range is obtained from the average of the 3σ -distant measurements of $\sin^2 \theta_W$ from leptonic data (A_{LR}) and b -quark data (A_{FB}^b).

Combining this result with the theoretical limit in Eq. 1.44, a preference for the $\sim (130 - 140)$ GeV range is obtained, in case the Standard Model is valid up to the GUT scale.

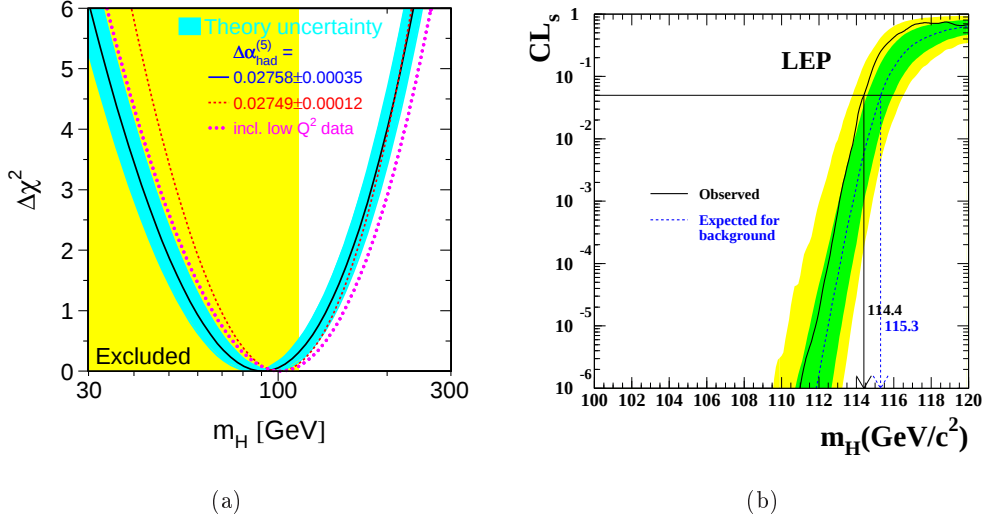


Figure 1.6: (a) Distribution of $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ as a function of m_H , resulting from the global electroweak fit; the shaded area marks the region excluded by direct searches [20]. (b) Final limit on m_H from LEP2 [18]. The ratio $\text{CL}_s = \text{CL}_{s+b}/\text{CL}_b$ for the signal plus background hypothesis, as a function of the test mass m_H , is shown. The solid line refers to observation, the dashed line is the median background expectation for the background-only hypothesis.

Direct searches at LEP

Before the era of the TeV colliders, the highest centre-of-mass energy used to be provided by the LEP (Large Electron Positron) collider, which operated between 1989 and 2000 at CERN, with $\sqrt{s}$ ranging from m_Z (LEP1) to 209 GeV (LEP2). The accelerator complex included four experiments: ALEPH, DELPHI, OPAL and L3. At LEP2 energies, the main production mechanism for a SM Higgs boson is the associate production $e^+e^- \rightarrow Z^* \rightarrow ZH$, or Higgs-strahlung; its cross section scales as $1/s$, vanishing around $m_H^{\max} \sim \sqrt{s} - m_Z \sim 117$ GeV. Considering the total integrated luminosity collected by LEP2, $\int L dt = 0.1 \text{ fb}^{-1}$, this mechanism would have provided around 30 Higgs bosons for $m_H = 110$ GeV. The most abundant decay channels are those involving heavy particles in the final state, i.e. b -quarks or τ leptons.

Results from the four LEP2 collaborations have been combined, resulting in the exclusion limit $m_H > 114.4$ GeV [18], as depicted in Fig. 1.6 (b), to be compared with 115.3 GeV expected in absence of signal. This difference is caused by a 1.7σ excess around $m_H = 116$ GeV (the corresponding p -value for the background-only hypothesis is 9%).

Direct searches at the Tevatron

The Tevatron, a 6.28 km long proton-antiproton collider, has been operative at Fermilab (US) between 1987 and 2011, with an energy in the centre-of-mass up to $\sqrt{s} = 1.96$ TeV. Its two main experiments, CDF and DØ, collected and analyzed a total luminosity of 10.0 fb^{-1} [19]. Similarly to the LHC, the most important production processes for a SM Higgs boson at the Tevatron are gluon-fusion ($gg \rightarrow H$) and associate production ($W^\pm H$ and ZH), the latter having a smaller cross section but a distinctive signature, with two hard jets in the final state. Concerning decay modes, $b\bar{b}$ dominates below $m_H \approx 135$ GeV, WW otherwise.

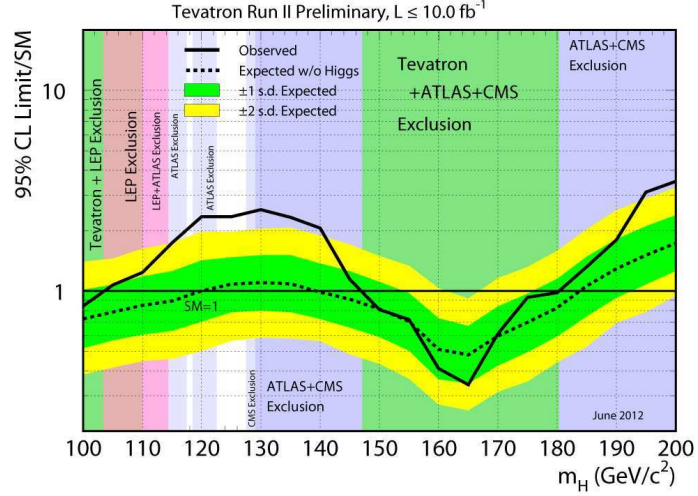


Figure 1.7: Observed and expected 95% C.L. upper limits on the ratios to the SM cross section, as a function of the Higgs boson mass, for the combined CDF and DØ analyses [19].

Despite being the most abundant process at low masses, the $gg \rightarrow H \rightarrow b\bar{b}$ channel is overwhelmed by the inclusive production $p\bar{p} \rightarrow b\bar{b} + X$, which proceeds via the strong interaction; as a consequence, the search for the associate production mode VH , with isolated leptons from the $W^\pm/Z$ decays, is experimentally more convenient. In particular, searches have been performed in the $W \rightarrow \ell\nu$ channel, in the $Z \rightarrow \nu\bar{\nu}, \ell^+\ell^-$ channels, and at CDF also in the hadronic decay modes.

The second most powerful channel at low masses is VH with $H \rightarrow \tau^+\tau^-$, where the hard jets are used to clean up the dominant $Z \rightarrow \tau^+\tau^-$ background. The $H \rightarrow \gamma\gamma$ decay mode, on the other hand, is suppressed by its low signal-to-background ratio, so that this channel does not contribute more than 5% to the overall combined sensitivity [56].

For masses above $m_H \approx 135$ GeV, the leading channel is $H \rightarrow W^+W^- \rightarrow$

$\ell^+\nu\ell^-\bar{\nu}$. For this channel, the $D\emptyset$ analysis uses the Boosted Decision Tree technique (cf. Section 4.6). A total of 14 BDTs are trained for the various phase space regions, lepton flavours and jet bins, resulting in the exclusion of a SM Higgs boson in the range $159 < m_H < 176$ GeV at 95% C.L. [57].

The CDF analysis, on the other hand, makes advantage of the Neural Network (NN) technique to maximize its sensitivity. After a common preselection, the phase space is splitted according to the jet bin, to the lepton flavour and to the dilepton invariant mass, training a NN with specific variables for each configuration. In the most sensitive bin, the $e\mu$ 0-jet one, the output of a Matrix Element computation at leading order is used as additional input. To extract final exclusion limits, a binned likelihood is constructed using 12 NN templates, leading to an exclusion at 95% C.L. in the range $148 \text{ GeV} < m_H < 175 \text{ GeV}$ [58].

Combining all the channels, above the whole mass range (100-200 GeV), for both experiments, an exclusion for the regions $100 \text{ GeV} < m_H < 120 \text{ GeV}$ and $139 \text{ GeV} < m_H < 184 \text{ GeV}$ is expected in the absence of signal, while the observed exclusion is found to be limited to $100 \text{ GeV} < m_H < 103 \text{ GeV}$ and $147 \text{ GeV} < m_H < 180 \text{ GeV}$ [19]. The reason for this is the presence of an excess of data events over the background expectation between 115 GeV and 140 GeV. The significance of such an excess is 3.0σ locally at $m_H = 120$ GeV; the corresponding global significance, taking into account the look-elsewhere effect, is 2.5σ . Most of the sensitivity comes actually from the $H \rightarrow b\bar{b}$ channel (with most of the signal candidates coming from the CDF $ZH \rightarrow \ell^+\ell^-b\bar{b}$ analysis).

1.4 Phenomenology of hadron collisions

The following Section briefly describes the fundamental aspects of phenomenology at hadron colliders; the description follows Ref. [59] if not stated otherwise.

1.4.1 Factorization

Collisions at hadron colliders can be divided into either *hard* or *soft*. Hard collisions, characterized by large momentum transfer among the partons and thus capable of significantly redirecting the energy flow, are the rarest; they can be described by means of perturbation theory, allowing for precise predictions of their rates and properties. Soft scattering, on the other hand, involves small momentum transfer; perturbation theory cannot be applied in this case, since, due to the running nature of the strong coupling constant, $\alpha_S(Q^2)$ increases with smaller energies, preventing the convergence of the perturbative expansion.

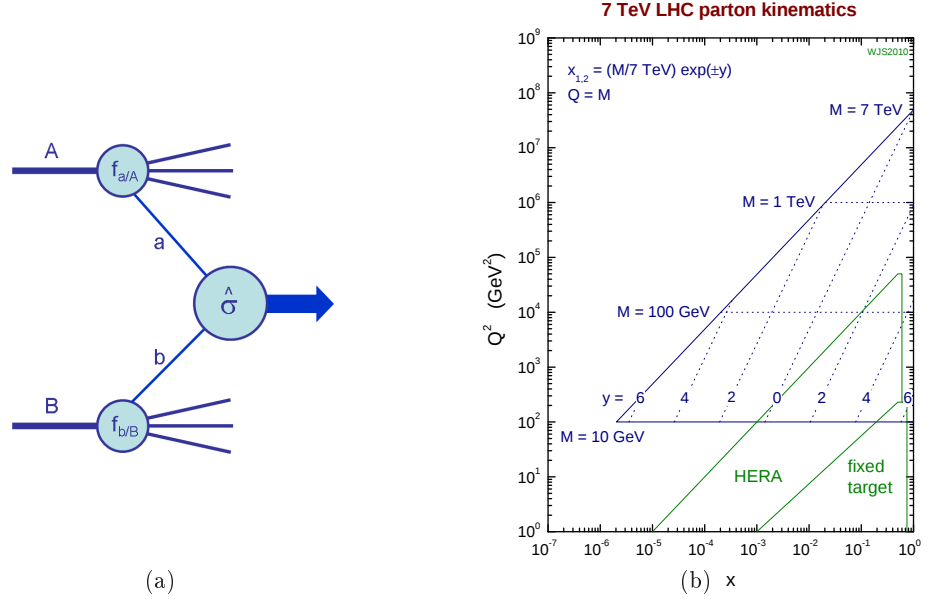


Figure 1.8: (a) Schematic representation of a proton-proton interaction, showing the partonic collision corresponding to the reduced cross section $\hat{\sigma}_{ab}$ [59]. (b) Kinematical reach of the LHC collider, in the (Q^2, x) phase space, at $\sqrt{s} = 7$ TeV, compared to HERA and previous fixed-target experiments [60].

Proton-proton collisions can be schematically represented as in Fig. 1.8 (a): when two protons, A and B , collide, the interaction actually happens between two of their constituents, a and b , with reduced cross section $\hat{\sigma}$. The proton constituents, i.e. quarks and gluons, are generically referred to as *partons*. The proton remnants can not exist as asymptotic states, due to confinement, and thus hadronise into colour-neutral states, in the so-called *underlying event*. Moreover, the incoming partons can radiate additional partons, due to their colour charge, resulting in additional jets which might be reconstructed in the event; they constitute the *initial state radiation* (ISR), while, if the final objects radiate bremsstrahlung radiation, this takes the name of *final state radiation* (FSR).

The centre-of-mass energy for the partonic system is reduced with respect to $\sqrt{s}$, the hadronic one, namely $\hat{s} = x_a x_b s$, where $x_{a,b}$ label the fractions of proton momenta carried by the two colliding partons. In the reference frame of the two hadrons, with z being the beam axis, the four-momenta of the colliding partons are given by (neglecting their masses):

$$p_a^\mu = \frac{\sqrt{s}}{2}(x_a, 0, 0, x_a) \quad , \quad p_b^\mu = \frac{\sqrt{s}}{2}(x_b, 0, 0, -x_b). \quad (1.46)$$

Two kinematical variables, the rapidity y and the pseudorapidity η , are widely used in collider phenomenology, since distances in y or η are invariant under a Lorentz

boost in the z direction:

$$y = \frac{1}{2} \ln((E + p_z)/(E - p_z)) \quad , \quad \eta = -\ln(\tan(\theta/2)) . \quad (1.47)$$

These two quantities are identical for massless particles. In a partonic collision like $ab \rightarrow X$, where a state X is produced, its mass M and rapidity y are linked to the fractional momenta of the incoming partons by the following relation:

$$x_a = \frac{M}{\sqrt{s}} e^y \quad , \quad x_b = \frac{M}{\sqrt{s}} e^{-y} . \quad (1.48)$$

This is shown in Fig. 1.8 (b), where the kinematical reach of the LHC experiments (for $\sqrt{s} = 7$ TeV) is compared with that of the previous fixed-target experiments and of HERA: with respect to them, the LHC reach extends to harder scales, hence to heavier objects. For a given rapidity y , two lines are drawn, corresponding to x_a and x_b ; the vertical axis corresponds to Q^2 , the characteristic scale of the process ($Q^2 = M^2$).

The hadronic cross section σ_{AB} can be obtained from the partonic one, σ_{ab} , convoluting it with the *parton distribution functions* (PDFs) $f_{a/A}(x_a, Q^2)$ and $f_{b/B}(x_b, Q^2)$ [61]:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab \rightarrow X} . \quad (1.49)$$

Here, Q^2 is again the momentum transferred in the process; this equation is valid in the *asymptotic scaling limit*, $s \rightarrow \infty$ and $\tau = m_X^2/s$ fixed. When real and virtual *collinear* gluon emissions are considered, they give rise to large logarithms, which apparently spoil the convergence of the series. In fact, a general result, known as *factorization theorem*, shows that collinear emissions are process-independent and hence can be reabsorbed at all orders into a redefinition of the PDFs themselves [62]. Only finite corrections are left behind, which are on the other hand not universal and constitute the perturbative $\mathcal{O}(\alpha_S^n)$ corrections, so that Eq. 1.49 can be rewritten as:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma}_0 + \alpha_S(\mu_R^2) \hat{\sigma}_1 + \dots]_{ab \rightarrow X} . \quad (1.50)$$

The strong coupling α_S is a function of the *renormalization scale* μ_R , while μ_F is the *factorization scale*, which separates small and large distances, i.e. the perturbative regime from the soft one. Formally, once all α_S orders have been resummed, the cross section in Eq. 1.50 does not depend on any of the scales. On the other hand, the dependence on μ_F and μ_R can be used to quantify the uncertainty on σ_{AB} given by missing higher orders. Typically, to avoid the presence of large logarithms, values

close to the energy scale of the process under consideration are chosen for μ_F and μ_R , e.g. $\mu_F = \mu_R = m_H$.

1.4.2 Parton distribution functions

Scaling violations, i.e. the dependence of the parton distribution functions on Q^2 , are described by the DGLAP equations [63–66]:

$$\begin{aligned}\frac{\partial q_i(x, Q^2)}{\partial \log Q^2} &= \frac{\alpha_S}{2\pi} \int_x^1 \frac{dz}{z} \{P_{q_i q_j}(z, \alpha_S) q_j(\frac{x}{z}, Q^2) + P_{q_i g}(z, \alpha_S) g(\frac{x}{z}, Q^2)\} \\ \frac{\partial g(x, Q^2)}{\partial \log Q^2} &= \frac{\alpha_S}{2\pi} \int_x^1 \frac{dz}{z} \{P_{g q_j}(z, \alpha_S) q_j(\frac{x}{z}, Q^2) + P_{g g}(z, \alpha_S) g(\frac{x}{z}, Q^2)\}.\end{aligned}\tag{1.51}$$

The coefficients P_{ab} are the *splitting functions*, describing how parton densities change due to gluon and quark emissions. They can be expanded perturbatively [67, 68]:

$$P_{ab}(x, \alpha_S) = P_{ab}^{(0)}(x) + \frac{\alpha_S}{2\pi} P_{ab}^{(1)}(x) + \frac{\alpha_S^2}{2\pi} P_{ab}^{(2)}(x) + \dots\tag{1.52}$$

It should be noticed that the DGLAP evolution only applies to multiple parton emissions where partons are strongly ordered in transverse momentum, $k_{T1}^2 \ll \dots \ll k_{Tn}^2 \ll Q^2$, since it resums only the leading logarithms in the expansion of $(\alpha_S \log Q^2)^n$, and may therefore not be accurate at low x , where $\log(1/x) > \log(Q)$.

For what concerns the x dependence of parton densities, it cannot be obtained perturbatively and has to be determined by a fit on data. The result is illustrated in Fig. 1.9 (a) for a representative set [69]: the gluon density dominates for $x \lesssim 0.01$, while the up-quark distribution peaks around $x \approx 0.2$ ⁽¹⁾.

The fit is performed by several independent groups, among them the CTEQ [71], MSTW [72], NNPDF [73] and HERA [74] collaborations. The data used in the fits cover a wide range in x and Q^2 , and include HERA data (H1 and ZEUS, mainly at low x), Deep Inelastic Scattering and Drell-Yan data (at higher x), and finally collider data. In particular, the LHC experiments have already provided valuable data for the precision determination of the PDF distributions, in both the jet and electroweak sectors. An example is shown in Fig. 1.9 (b), where the ATLAS measurement of the $W \rightarrow \ell \nu$ fiducial cross section is compared with the latest PDF predictions, allowing for their discrimination. Ellipses are obtained through the *Hessian* method, which is the standard procedure to compute PDF uncertainties and their impact on collider observables.

The Hessian method quantifies uncertainties by means of a complete set of eigen-

¹For this reason, since to produce a state of mass ≈ 100 GeV at intermediate rapidities, $y = 2$, two partons of $x_a \approx 0.05$ and $x_b \approx 0.001$ are needed, partonic interactions at the LHC are most likely gluon-gluon collisions in the low and intermediate mass range.

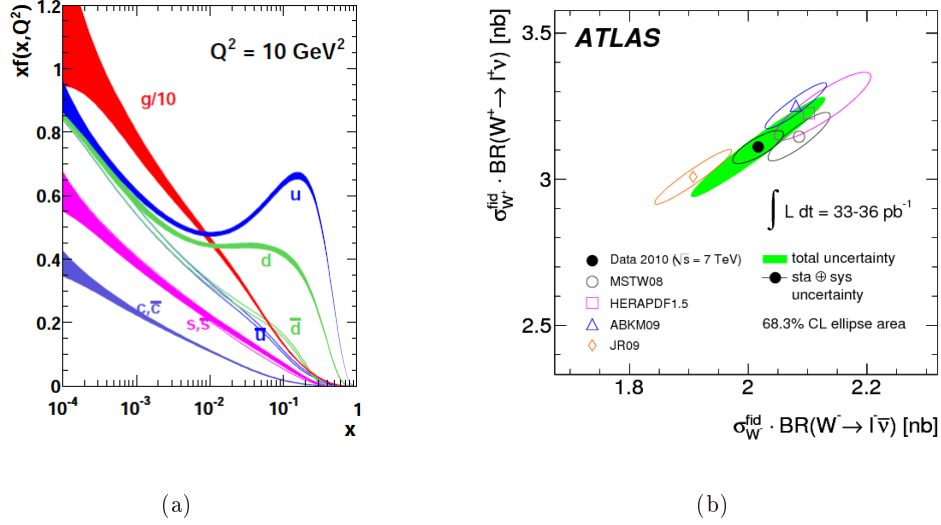


Figure 1.9: (a) PDF distributions for the MSTW 2008 NLO set at $Q^2 = 19 \text{ GeV}^2$ [69]. The filled areas correspond to the 68 % C.L. uncertainties. (b) Correlation between the $W^+ \rightarrow \ell^+ \nu$ and $W^- \rightarrow \ell^- \bar{\nu}$ cross sections as measured by the ATLAS collaboration [70]. The theoretical predictions are shown for various PDF sets, compared with the result observed with an integrated luminosity corresponding to 33 pb^{-1} at $\sqrt{s} = 7 \text{ TeV}$. The ellipses span the 68 % C.L. coverage for total uncertainties (green) and excluding the luminosity uncertainty (black).

vectors, obtained from the diagonalization of the $N \times N$ Hessian matrix⁽²⁾, where N is the number of free parameters in the fit. Their eigenvalues span many orders of magnitude, with the largest ones being related to the best-determined directions. For each eigenvector there are two error sets, “+” and “-”, obtained via a $\Delta\chi^2$ excursion from the global minimum⁽³⁾. The positive and negative variations induced on an observable can be asymmetric, especially for the eigenvectors which are poorly determined. This is taken into account when computing the PDF envelope for a variable X via the *Master Formula* [75, 76]:

$$\Delta X^+ = \sqrt{\sum_{i=1}^N [\max(X_i^+ - X_0, X_i^- - X_0, 0)]^2} \quad (1.53)$$

$$\Delta X^- = \sqrt{\sum_{i=1}^N [\max(X_0 - X_i^+, X_0 - X_i^-, 0)]^2}.$$

²For CTEQ latest sets, N goes from 20 (CTEQ6.1) to 26 (CT10 NLO), while for MSTW2008 $N = 20$.

³The choice of the $\Delta\chi^2$ criterium is different for the CTEQ ($\Delta\chi^2 = 100$, corresponding to 90% C.L.) and the MSTW ($\Delta\chi^2 = 50$) collaborations, reflecting the fact that a coherent and unique statistical treatment is not possible, given the large variance of experimental data to be included.

An additional uncertainty to be considered when computing theoretical uncertainties related to PDFs on physical observables, is the one deriving from the running strong coupling constant, since this value is used in the global PDF fits.

1.4.3 Monte Carlo generators

Monte Carlo (MC) generators are an essential tool for particle physics: they are used to simulate collision events in a probabilistic way, which is intrinsic to the quantistic nature of elementary particles. According to the theoretical cross sections of physics processes and to their expected kinematical and angular distributions, MC generators produce a large number of events, providing the information about the four-momenta of the particles involved. The name “Monte Carlo” derives from the fact that, in order to simulate the probabilistic outcome, random numbers are used. Fig. 1.10 shows a schematic break down of the various steps involved in the

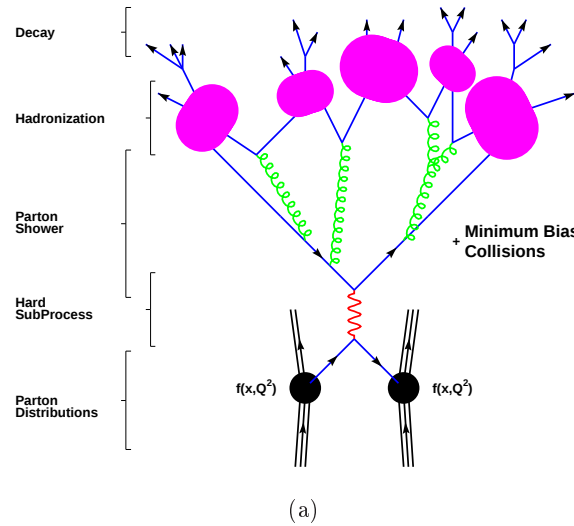


Figure 1.10: Schematic representation of the various phases of a collision event, as implemented in Monte Carlo generators [77]. Time flows from the bottom to the top of the figure.

generation of a MC event, with the time flowing from the bottom to the top of the figure [77]. The starting point is the hard collision between two of the incoming partons, based on the matrix element of the interaction. Most MC generators only allow the presence of a fixed number of partons in the final state. Subsequently, higher order effects in perturbation theory are added by evolving the event thanks to an iterative process, known as the *parton shower* method. Starting from the hard interaction, an algorithm evolves the event backward and forward in time, by means of the DGLAP equations, to add initial and final state radiation, respectively.

With this procedure, soft and collinear partons are added to the event, so that at the end of the series of splittings (which is determined by an energy cut-off, typically $\mathcal{O}(1 \text{ GeV})$), a number of particles moving coherently with the incoming partons is produced. **PYTHIA** [78] and **HERWIG** [79] are two of the most commonly used generators and are able to simulate both the hard collision and the following parton shower; the generators which do not possess this latter feature, on the other hand, are commonly interfaced to either **PYTHIA** or **HERWIG**.

After the shower is complete, the produced particles are generally unstable and are forced to decay, until stable final particles are obtained; moreover, there might be particles with non-zero colour charge, for which the *hadronization* step takes place. In this case, partons are grouped into hadrons using a phenomenological model, since perturbation theory cannot describe long-range effects. Treating the hadronisation step as completely decoupled from the perturbative part of the event, allows the tuning of the phenomenological model to data, using minimum bias events⁽⁴⁾ [80–82].

The last step in the MC event generation consists in adding the details of the underlying and pile-up events. Also in this case, the parameters of the model have to be tuned to data, as measured by the ATLAS experiment at $\sqrt{s} = 900 \text{ GeV}$ and $\sqrt{s} = 7 \text{ TeV}$ [83]. While **PYTHIA** uses its own implementation for the underlying event and multiple particle interactions, **HERWIG** is interfaced with **JIMMY** [84].

The generated events are finally propagated through the full ATLAS detector, simulated by the **GEANT4** toolkit [85]. It includes an exact description of the material distribution in the detector, of the interactions undergone by the particles during their propagation, of the energy measurement and resolution and the read-out electronics. For a few of the backgrounds used in the $H \rightarrow WW^{(*)}$ analysis documented in this thesis, the full simulation of the ATLAS detector has been substituted by a faster (but less precise) one, provided by the **ATLFAST** software [86]. The compatibility of the two simulations, within the respective uncertainties, has been verified in each case.

Finally, for some of the theoretical studies (like those aiming to quantify the impact of PDFs on some observable), no detector simulation is required: in these cases, MC samples *at generator level* are used.

⁴The term minimum bias refers to soft, inelastic collisions, with low transverse momentum and low particle multiplicity.

Chapter 2

Higgs Boson Physics

This Chapter offers an overview of the most important production and decay channels for the Standard Model Higgs boson at the LHC. The gluon-fusion, vector-boson fusion and associate production modes are briefly described, with a focus on their potential at the LHC and on the status of the theoretical computations. With regards to the decay channels, the mechanisms for the Higgs boson decaying into two fermions, two photons or two massive vector bosons are introduced. The description follows Refs. [46, 87] if not stated otherwise. The second part of the Chapter focuses on the Higgs boson properties, and in particular on its spin-CP assignments in the Standard Model and on their consequences for the final-state particles. The kinematical and angular distributions expected for leptons arising from $X \rightarrow VV^*$ decays are discussed, with X being either a spin-0, spin-1 or spin-2 resonance. In the latter case, the Randall-Sundrum model for Kaluza-Klein gravitons is assumed.

2.1 Production modes

As discussed in Chapter 1, the Higgs boson couples preferentially to the heaviest particle allowed by the phase space constraints. This property, together with the hypothesis on the Higgs boson mass (that in the following is assumed to be $m_H = 125$ GeV), defines completely the production and decay channels, and their relative abundances. Fig. 2.1 shows the production modes at the LHC (for (a) $\sqrt{s} = 7$ TeV and (b) $\sqrt{s} = 8$ TeV), in the intermediate mass range, stating also the precision achieved by the theoretical computations.

The most important production channels are the gluon-fusion mechanism ($gg \rightarrow H$), the vector-boson fusion process ($qq \rightarrow V^*V^* \rightarrow qq + H$), the associated production with W or Z bosons ($q\bar{q} \rightarrow VH$) and the associated production with top- or bottom-quarks ($gg, q\bar{q} \rightarrow Q\bar{Q} + H$). The corresponding cross sections are summarized in Ta-

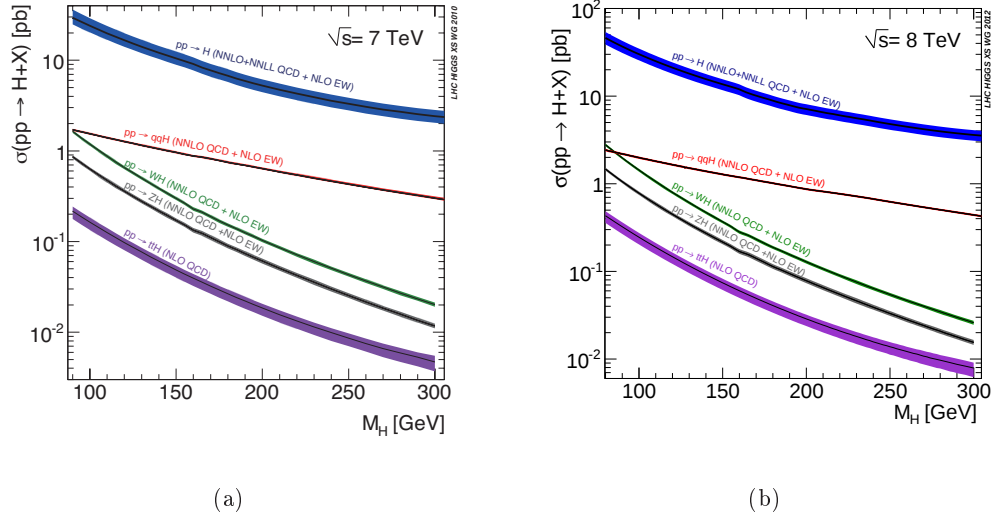


Figure 2.1: Cross sections for Higgs boson production in the Standard Model, for (a) $\sqrt{s} = 7$ TeV and (b) $\sqrt{s} = 8$ TeV [88].

ble 2.1, for two different center-of-mass energies, $\sqrt{s} = 7$ and 8 TeV, corresponding to the 2011 and 2012 LHC runs, respectively. The theoretical uncertainties derive from the scale variations ($\mu_R, \mu_F = m_H/2, 2m_H$), which account for the higher orders missing in the computation, and from the PDFs (and the corresponding uncertainty on α_s), both introduced in Section 1.4.

	Process	Cross section (pb)	+ Error (%)	- Error (%)	$\pm$ Scale (%)	$\pm$ (PDF, α_s) (%)
$\sqrt{s} = 7$ TeV	gg	15.32	+14.7	-14.9	7.8	7.6
	VBF	1.222	+2.8	-2.4	+0.3	2.5
	WH	0.5729	+3.7	-4.3	0.8	3.5
	ZH	0.3158	+4.9	-5.1	1.6	3.5
	ttH	0.0863	+11.8	-17.8	9.3	8.5
$\sqrt{s} = 8$ TeV	gg	19.52	+14.7	-14.7	7.8	7.5
	VBF	1.578	+2.8	-3.0	0.2	2.8
	WH	0.6966	+3.7	-4.1	0.6	3.5
	ZH	0.3943	+5.1	-5.0	+1.6	3.5
	ttH	0.1302	+11.6	-17.1	9.3	7.8

Table 2.1: Summary of the LHC reference values for the production cross sections of a SM Higgs boson with $m_H = 125$ GeV, at $\sqrt{s} = 7$ and 8 TeV [87]. The total theoretical uncertainties are broken down into the scale variation term and the PDF term (in the latter case, considering also the associated uncertainty on α_s). If the uncertainties are asymmetric, the largest value is given.

2.1.1 The gluon-fusion production mechanism

The gluon-fusion process, depicted in Fig. 2.2 at leading order (LO), proceeds through a quark loop, mainly constituted by t - and b -quarks; this is due to the fact that the Higgs boson coupling to fermions, g_{Hff} , is proportional to m_f (cf. Fig. 1.4), which also means that the b -loop is suppressed with respect to the t -loop by a factor $m_t/m_b \sim 35$.

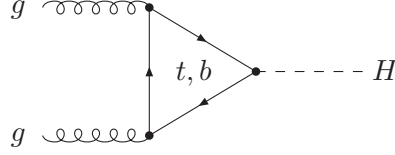


Figure 2.2: Feynman diagram contributing to $gg \rightarrow H$ at leading order.

The partonic cross section at leading order can be written as [89]:

$$\sigma_0^H(gg \rightarrow H) = \frac{G_\mu \alpha_s^2(\mu_R^2)}{288\sqrt{2}\pi} \left| \frac{3}{4} \sum_q A_{1/2}^H(\tau_Q) \right|^2 \quad \text{with } \tau_Q = m_H^2/4m_Q^2 \quad (2.1)$$

where Q labels heavy quarks and $A_{1/2}^H(\tau_Q)$ is the form factor for spin-1/2 particles. Below the threshold for producing $Q\bar{Q}$ pairs, the real part of the amplitude dominates, going from $A_{1/2}^H = 4/3$ for small values of τ_Q , to $A_{1/2}^H = 2$ for $\tau_Q = 1$, and decreasing linearly afterwards, modulo logarithmic corrections (the amplitude vanishes in the chiral limit $m_Q \rightarrow 0$). Eq. 2.1 is obtained in the narrow-width approximation (NWA).

Convoluting $\sigma_0^H(gg \rightarrow H)$ with the gluonic PDFs, one can obtain the proton-proton cross section at LO for the gluon-fusion process:

$$\sigma_{\text{LO}}(pp \rightarrow H) = \sigma_0^H \tau \frac{d\mathcal{L}^{gg}}{d\tau} \quad (2.2)$$

with $\frac{d\mathcal{L}^{gg}}{d\tau} = \int_\tau^1 \frac{dx}{x} g(x, \mu_F^2) g(\tau/x, \mu_F^2)$ and $\tau = m_H^2/s$.

At next-to-leading order (NLO), virtual and real QCD corrections have to be taken into account, namely $gg \rightarrow H(g)$ (Fig. 2.3 (a),(c)), $q\bar{q} \rightarrow Hg$ (b) and $gq \rightarrow Hq$ (d), all contributing at the same order in α_s . When all QCD corrections are added up⁽¹⁾, ultraviolet and infrared divergences cancel, while the remaining collinear divergences are absorbed into a redefinition of the parton densities [90–92]. The cross section

⁽¹⁾Electroweak radiative corrections are not discussed here, since they are negligible with respect to the sizeable QCD ones (less than 5% on σ_{LO} for $m_H = 125$ GeV).

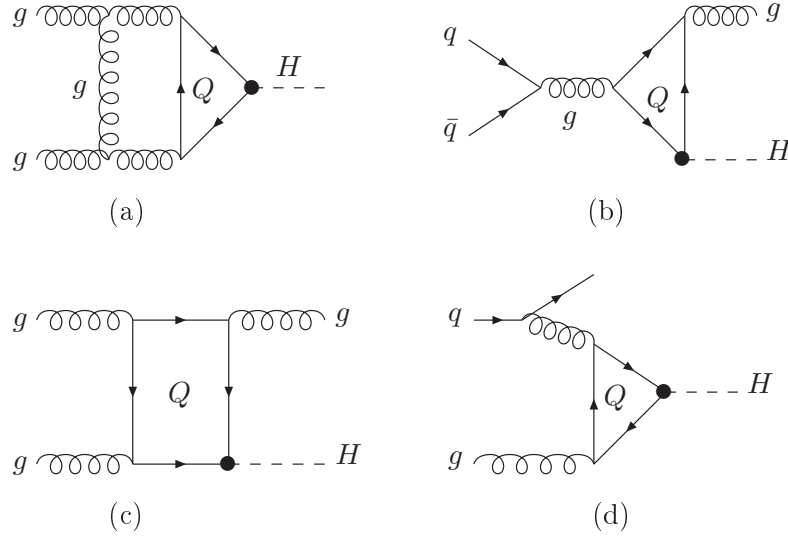


Figure 2.3: Typical diagrams for the virtual and real QCD corrections to $gg \rightarrow H$ at next-to-leading order: (a) virtual gluon correction to $gg \rightarrow H$, (b) $q\bar{q} \rightarrow Hg$, (c) real gluon correction to $gg \rightarrow H$, (d) $gq \rightarrow Hq$.

becomes then [93,94]:

$$\sigma(pp \rightarrow H + X) = \sigma_0^H \left[1 + C_H \frac{\alpha_s}{\pi} \right] \tau_H \frac{d\mathcal{L}^{gg}}{d\tau_H} + \Delta\sigma_{gg}^H + \Delta\sigma_{gq}^H + \Delta\sigma_{q\bar{q}}^H \quad (2.3)$$

where the C_H coefficient, denoting the two-loop quark corrections after regularization, is equal to $11/2$ in the limit of large quark masses, $\tau_Q = m_H^2/4m_Q^2 \ll 1$ (which is the case if $m_H = 125$ GeV). In this limit, also the remaining real contributions $\Delta\sigma$ take simpler expressions, omitted here. The K -factors, defined as $K = \sigma_{\text{NLO}}/\sigma_{\text{LO}}$, are tiny for $K_{q\bar{q}}$ and K_{gg} , the latter being negative, while K_{gq} and K_{virt} are rather large, of the order of 50% each. This result, which holds for all masses at the LHC, brings K_{tot} to increase the LO cross section of a factor 60 – 90%, both in the limit $\tau_Q \ll 1$ and when the full dependence on m_t is taken into account, the difference between the two cases being only a few percent.

On the other hand, the inclusion of next-to-next-to-leading order (NNLO) corrections has been performed only in the large m_t limit, leading to an increase with respect to σ_{NLO} of about 25%, and of an additional 5% when resumming also the soft-gluon terms up to NNLL (next-to-next-to-leading logarithms) [95–97]. Recently, also approximate NNNLO corrections have been computed, showing that the convergence of the perturbative series is slower than expected ($k_{\text{NNNLO}} = 17\%$ for $m_H = 125$ GeV and $\sqrt{s} = 8$ TeV [98]), yielding a central value of $\sigma_{\text{NNNLO}} = 22.58$ pb.

The reference value for LHC experiments is on the other hand $19.52 \text{ pb} \pm 14.7\%$ for $m_H = 125 \text{ GeV}$ (cf. Table 2.1), computed at NNLL QCD and NLO EW, with the error given by the linear combination of scale, PDF and α_s uncertainties [87].

The dependence on μ_R and μ_F (varied in the range $\mu_R, \mu_F = m_H/2, 2m_H$), shown in Fig. 2.4, is reduced when going from LO to NNNLO, being $\pm 7\%$ in the latter case.

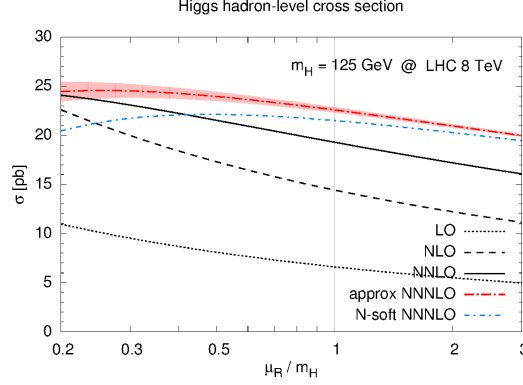


Figure 2.4: Dependence of the gluon-fusion cross section on the renormalization scale μ_R , for all orders up to NNNLO [98].

2.1.2 The vector-boson fusion production mechanism

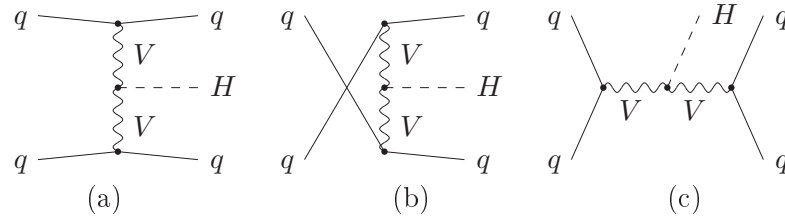


Figure 2.5: Feynman diagrams for the (a) t -, (b) u -, and (c) s -channel contributions to $qq \rightarrow qqH$ at LO, where $V = W, Z$ and q denotes any quark or antiquark.

The vector-boson fusion mechanism (VBF) $qq \rightarrow qqH$ (where q denotes any quark or antiquark), depicted in Fig. 2.5 at LO in the t -, u - and s -channels, is characterized by the presence of two hard jets in the final state, separated by a large pseudorapidity gap. The intermediate vector bosons tend to carry only a small fraction of the energy of the incoming partons; the rest is released in terms of jets, with large energies and small transverse momenta ($p_T \sim m_V$), and hence at small angles from the beam axis. This is in contrast with the large QCD background, which normally exhibits substantial central activity; these features are used to tag the VBF jets and suppress QCD and other backgrounds.

The theoretical uncertainty given by the NLO QCD and NLO EW computation [99] is reduced from 5 – 10% to 1 – 2% when including approximate NNLO QCD corrections, which makes the VBF channel the most theoretically accurate Higgs boson production channel at hadron colliders [100]. The reference value for LHC experiments is $\sigma = (1.578 \text{ pb})_{-3.0\%}^{+2.8\%}$ at $\sqrt{s} = 8 \text{ TeV}$ for $m_H = 125 \text{ GeV}$ (cf. Table 2.1) [87].

This mechanism proceeds through radiation of vector bosons from a $q\bar{q}$ pair, with $W^*W^* \rightarrow H$ being three times more likely than $Z^*Z^* \rightarrow H$ at the LHC, due to the fact that W -boson couplings to fermions are larger than those of the Z boson. While at the LHC the VBF mechanism is the second most abundant production process, at the Tevatron it does not exceed 0.1 pb, even for low m_H values; this is because the main contribution to VBF originates from longitudinal vector-bosons exchange, whose cross section grows with s , as seen in Section 1.3.1.

The NLO K -factor increases the LO cross section by about 5% and, at the LHC, is practically constant over the entire mass range [101, 102]. For what concerns the dependence on scales and PDFs, it is of the order of 2% at NLO [103].

2.1.3 The associate production mechanism

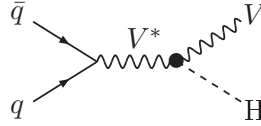


Figure 2.6: The VH production mechanism at leading order.

The third most abundant production process at hadron colliders is the associate production mode, $q\bar{q} \rightarrow V^* \rightarrow VH$, depicted in Fig. 2.6. It can be factorized into the Drell-Yan production of a virtual vector boson, followed by its decay into a VH pair, so that the partonic cross section at any order can be written as [46]:

$$\sigma(q\bar{q} \rightarrow HV) = \sigma(q\bar{q} \rightarrow V^*) \times \frac{d\Gamma}{dk^2}(V^* \rightarrow HV) \quad (2.4)$$

where k is the momentum of V^* ($k^2 \neq m_V^2$). With this factorization, it is only the $\sigma(q\bar{q} \rightarrow V^*)$ term that has to be modified to include NLO corrections, which increase the LO cross section by approximately 30%, setting $\mu_R = \mu_F = m_V$ [104–106]; adding also the NNLO QCD terms [95] and the NLO EW ones [107–109], both of the order of a few percent, yields the reference values of $\sigma_{WH} = (0.697 \text{ pb})_{-4.1\%}^{+3.7\%}$ and $\sigma_{ZH} = (0.394 \text{ pb})_{-5.0\%}^{+5.1\%}$ for $m_H = 125 \text{ GeV}$ [87], at $\sqrt{s} = 8 \text{ TeV}$ (cf. Table 2.1).

σ_{WH} is a factor two larger than σ_{ZH} at both the Tevatron and the LHC, over the whole mass range; in addition, given the fact that $\text{BR}(W^\pm \rightarrow \ell^\pm \nu) \simeq 20\%$ while $\text{BR}(Z \rightarrow \ell^+ \ell^-) \simeq 6\%$, the WH channel results to be much more abundant than ZH .

2.2 Decay modes

Fig. 2.7 (a) shows the decay branching ratios for a SM Higgs boson over the whole mass range, up to 1 TeV. In (b), the branching ratios are multiplied by the corresponding production cross sections at $\sqrt{s} = 8$ TeV.

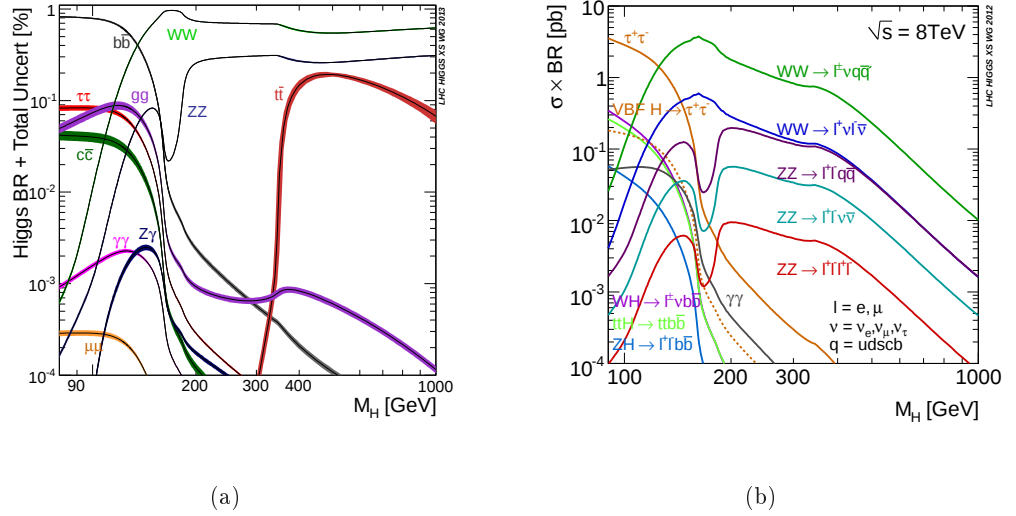


Figure 2.7: Branching ratios for Higgs boson decay in the Standard Model (a), multiplied by the corresponding production cross sections at $\sqrt{s} = 8$ TeV (b) [88].

The most abundant decay channel for a light Higgs boson ($110 \text{ GeV} \lesssim m_H \lesssim 130 \text{ GeV}$) is $b\bar{b}$, with a branching ratio between 75% and 50%, followed by WW^* , which increases up to 30% in this mass range, and gg (around 10%, mediated by loops of heavy quarks). The remaining channels ($\tau^+\tau^-$, $c\bar{c}$, $\gamma\gamma$, ZZ^* , γZ) are under the percent level. The actual values, used as a reference for LHC experiments, can be seen in Table 2.2 for $m_H = 125 \text{ GeV}$; the corresponding total width is $\Gamma_H = 4.07 \text{ MeV}$, with an uncertainty of $\pm 4\%$ [87].

In the intermediate mass range ($130 \text{ GeV} \lesssim m_H \lesssim 180 \text{ GeV}$), the $WW^{(*)}$ and $ZZ^{(*)}$ channels are fully accessible (one or two on-shell bosons for $m_H < 2m_V$ and $m_H > 2m_V$, respectively) and dominate the total width, with a small contribution from $b\bar{b}$, dropping from 50% at $m_H = 130 \text{ GeV}$ as mentioned above, to a few percent

at $m_H = 180$ GeV. For $2m_W \lesssim m_H \lesssim 2m_Z$, given the virtuality of one of the Z bosons, the WW channel is close to a branching ratio of 100%.

In the high-mass range ($m_H \gtrsim 2m_Z$), the Higgs boson decays exclusively into massive gauge bosons, with a branching ratio of $\sim 2/3$ for WW and $\sim 1/3$ for ZZ final states. For $m_H \gtrsim 350$ GeV, also the $t\bar{t}$ channel opens up, but the pattern just mentioned does not change significantly, given that the $t\bar{t}$ branching ratio stays below 20%.

	$b\bar{b}$	WW^*	gg	$\tau\tau$	$c\bar{c}$	ZZ^*	$\gamma\gamma$	$Z\gamma$	$\mu\mu$
Branching ratio	0.577	0.215	0.0857	0.0632	0.0291	0.0264	0.0228	0.00154	0.00022
$\pm$ Relative error (%)	3.3	4.3	10.2	5.7	12.2	4.3	5.0	9.0	6.0

Table 2.2: Summary of the LHC reference values for the decay branching ratios of a SM Higgs boson with $m_H = 125$ GeV, with their relative theoretical uncertainties [87]. If the uncertainties are asymmetric, the largest value is given.

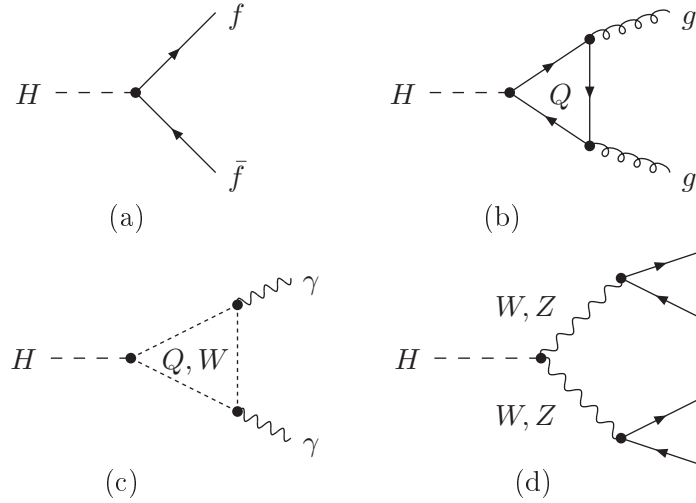


Figure 2.8: Leading order decay diagrams, with Q denoting any heavy quark: (a) $H \rightarrow f\bar{f}$, (b) $H \rightarrow gg$, (c) $H \rightarrow \gamma\gamma$, (d) $H \rightarrow VV$.

2.2.1 The $H \rightarrow f\bar{f}$ channel

The $H \rightarrow f\bar{f}$ decay channel is depicted in Fig. 2.8 (a) at leading order. The partial decay width can be cast into the form [110,111]:

$$\Gamma_{\text{LO}}(H \rightarrow f\bar{f}) = \frac{G_\mu N_c}{4\sqrt{2}\pi} m_H m_f^2 \beta_f^3 \quad (2.5)$$

where N_c is the number of colours ($N_c = 1$ for leptons, $N_c = 3$ for quarks). The partial width is proportional to m_f^2 , signaling the preference for heavy fermions,

and to $\beta_f^3 = (1 - 4m_f^2/m_H^2)^{3/2}$, β_f being the relativistic velocity of the final-state fermions. Due to this latter factor, $\Gamma_{\text{LO}}(H \rightarrow f\bar{f})$ is strongly suppressed near the mass threshold⁽²⁾, $\Gamma_{\text{LO}}(H \rightarrow f\bar{f}) \propto \beta_f^3 \rightarrow 0$ for $m_H \simeq 2m_f$.

For decays into light quarks ($m_q \ll m_H < m_t$, with $q \neq t$, so that $\beta_q \sim 1$), large logarithms appear in the decay width at one loop. The Feynman diagrams to consider are those involving the emission of a real or virtual gluon by one of the quark lines, yielding the following expression for the NLO partial decay width [112–115]:

$$\Gamma_{\text{NLO}}(H \rightarrow q\bar{q}) \simeq \frac{3G_\mu}{4\sqrt{2}\pi} m_H m_q^2 \left[1 + \frac{4}{3} \frac{\alpha_s}{\pi} \left(\frac{9}{4} + \frac{3}{2} \log \frac{m_q^2}{m_H^2} \right) \right] \quad (2.6)$$

proportional to $\log(m_q^2/m_H^2)$, which is negative for $m_q < m_H$. This can be avoided via a redefinition of the quark masses; in particular, using the $\overline{\text{MS}}$ factorization scheme, these logarithms are resummed to all orders and absorbed into $\overline{m}_q(m_H^2)$ and $\overline{\alpha}_s(m_H^2)$.

2.2.2 The $H \rightarrow \gamma\gamma$ channel

As shown in Fig. 2.8 (c), due to their massless nature, the Higgs boson decay into photons or gluons proceeds at lowest order via loops involving massive particles. In particular, the $H \rightarrow \gamma\gamma$ and $H \rightarrow Z\gamma$ decays are mediated by loops constituted by W bosons and charged fermions (essentially top-quarks, since the other contributions can be neglected), while for Hgg only quark loops come into play. Due to the presence of additional electroweak or strong coupling constants, these loop-decays become negligible when the WW and ZZ kinematical thresholds open up.

At LO, the partial decay width into photons is given by [110, 116–118]:

$$\Gamma_{\text{LO}}(H \rightarrow \gamma\gamma) = \frac{G_\mu \alpha m_H^3}{128\sqrt{2}\pi^3} \left| \sum_f N_c Q_f^2 A_{1/2}^H(\tau_f) + A_1^H(\tau_W) \right|^2 \quad (2.7)$$

where $A_{1/2}^H$ and A_1^H correspond to the form factors for spin-1/2 and spin-1 particles, respectively (cf. Eq. 2.1), which are functions of the parameter $\tau_{f,W} = m_H^2/(4M_{f,W}^2)$. In particular, for $m_H = 125$ GeV, $A_{1/2}^H(\tau_t) \simeq 4/3$ and $A_1^H(\tau_W) \simeq -8$, so that the W -loop dominates. The decay width into photons is just a few keV in the low mass range, but, thanks to its clear experimental signature, the $H \rightarrow \gamma\gamma$ channel has been the first, together with $H \rightarrow ZZ^* \rightarrow 4\ell$ [29], where an excess of events compatible with the SM Higgs boson has been detected by the ATLAS experiment

²This does not happen for a pseudoscalar Higgs boson A , for which the partial decay width would have been suppressed by a factor β_f only [46].

in 2012 [30, 119].

Concerning the NLO QCD corrections to the quark loop, they are only virtual, and the following parametrization can be used [94, 120–125]:

$$A_{1/2}^H(\tau_t) = (A_{1/2}^H(\tau_t))_{\text{LO}} \left[1 + \frac{\alpha_s}{\pi} C_H(\tau_t) \right]. \quad (2.8)$$

The perturbative factor $C_H(\tau_t)$ can be approximated as -1 in the region of interest (or, more in general, in the $m_t \rightarrow \infty$ limit), while it gets much larger above the $t\bar{t}$ threshold, in particular above 650 GeV, where the W - and t -loop amplitudes exhibit destructive interference.

2.2.3 The $H \rightarrow V^{(*)}V^{(*)}$ channel

The Higgs boson decay into real and virtual gauge bosons is depicted in Fig. 2.9 at leading order.

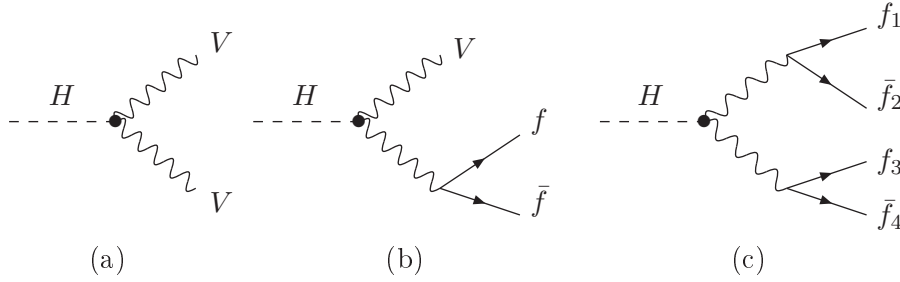


Figure 2.9: Higgs boson decays into real and virtual gauge bosons: (a) both V s are on-shell, (b) one V is on-shell, the other is off-shell, (c) both V s are off-shell.

For $m_H = 125$ GeV, the decay into two real vector bosons is not possible, so one of them has to be off-shell (the probability for both V s being off-shell is smaller than 1%). In this range, the decay into WW^* dominates over the decay into $b\bar{b}$, notwithstanding the suppression of the former by an additional power of the electroweak coupling and the virtuality of the W^* boson, due to the fact that $g_{HVV} \gg g_{Hff}$ (the first involves $2m_V^2$, the second m_b). The width for the three-body decay is obtained assuming massless fermions in the final state, and yields [126, 127], [128]:

$$\Gamma(H \rightarrow VV^* \rightarrow V f \bar{f}) = \frac{3G_\mu^2 m_V^4}{16\pi^3} m_H \delta'_V R(x) \quad (2.9)$$

with $\delta'_W = 1$ and $\delta'_Z = \frac{7}{12} - \frac{10}{9} \sin^2 \theta_W + \frac{40}{9} \sin^4 \theta_W$, where all the dependence on $x = m_V^2/m_H^2$ is contained in the $R(x)$ factor. The partial width is proportional to m_H , rather than to m_H^3 as in the case of a two-body decay.

The higher order corrections to $H \rightarrow V^{(*)}V^{(*)}$ have first been computed for on-shell gauge bosons, in the narrow-width approximation [129–131], and later in the general case [132].

At the LHC, the following branching ratios are used, for $m_H = 125$ GeV: $\Gamma_{H \rightarrow WW} = 0.215^{+4.3\%}_{-4.2\%}$ and $\Gamma_{H \rightarrow ZZ} = 0.0264^{+4.3\%}_{-4.2\%}$ (cf. Table 2.2) [87]. They are computed with **HDecay** [133], which takes into account all corrections up to two loops in the heavy-Higgs boson-mass hypothesis, and folded with the result of **Prophecy4f** [132], which includes also the interference effects between WW and ZZ , namely $\Gamma_{4f}^{\text{Prophecy}} = \Gamma_{H \rightarrow W^*W^* \rightarrow 4f} + \Gamma_{H \rightarrow Z^*Z^* \rightarrow 4f} + \Gamma_{WW/ZZ}$. Missing higher orders account for a residual uncertainty $\lesssim 0.5\%$.

2.3 Spin and CP quantum numbers

2.3.1 The case of a spin-0 resonance

The production and decay modes discussed so far, have been presented under the assumption of a SM Higgs boson, corresponding to the $J^{CP} = 0^{++}$ assignment for its spin and CP quantum numbers. As a consequence, the couplings to the vector bosons have the form:

$$\mathcal{L}(HVV) = \left(\sqrt{2}G_\mu\right)^{1/2} m_V^2 H V^\mu V_\mu \quad (2.10)$$

as in Fig. 1.4, which are even under parity and charge conjugation.

On the other hand, an hypothetical CP-odd Higgs boson A , predicted by many extensions of the SM (e.g. in two-doublet Higgs boson models [134], 2HDM, as the Minimal Supersymmetric Standard Model [135], or in axion models [136, 137]), would couple to the vector bosons with an effective term of the form [46]:

$$\mathcal{L}(AVV) = \frac{1}{4}\eta \left(\sqrt{2}G_\mu\right)^{1/2} m_V^2 A V^{\mu\nu} \tilde{V}_{\mu\nu} \quad (2.11)$$

with $\tilde{V}_{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} V_{\rho\sigma}$ being the field conjugate to $V_{\mu\nu}$, $\epsilon^{\mu\nu\rho\sigma}$ the totally antisymmetric tensor and η a dimensionless constant. If CP is conserved at tree-level, $\mathcal{L}(AVV)$ is induced only at higher orders, which limits its possible size. Finally, it should be noticed that a CP-odd coupling has to be a P-wave (total angular momentum equal to one), while the SM CP-even coupling is a S-wave; the P-wave reduces in the laboratory frame to $(\epsilon_1 \times \epsilon_2) \cdot (p_1 - p_2)$, rather than to $\epsilon_1 \cdot \epsilon_2$, which can be rephrased saying that A does not couple to longitudinal gauge bosons.

The SM Higgs boson coupling to fermions has the form $g_{Hff} = (\sqrt{2}G_\mu)^{1/2} m_f$; the corresponding g_{Aff} contains an additional γ_5 factor, resulting in a pseudoscalar

quantity [46]:

$$g_{Aff} = (\sqrt{2}G_\mu)^{1/2} m_f \gamma_5. \quad (2.12)$$

Such a coupling is purely axial, thus manifestly parity-violating.

The angular and kinematical distributions of the final-state fermions in the $H/A \rightarrow VV^* \rightarrow (f_1\bar{f}_2)(f_3\bar{f}_4)$ decay depend on the CP properties of the intermediate boson (H or A). The reference frame is defined in Fig. 2.10.

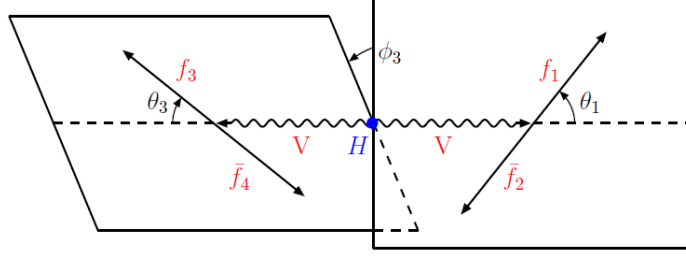


Figure 2.10: Definition of the decay angles for the final state $H \rightarrow VV \rightarrow (f_1\bar{f}_2)(f_3\bar{f}_4)$ in the Higgs boson rest frame [138].

Five kinematical quantities – the Cabibbo-Maksymowicz variables [139] – are in fact sufficient to describe this final state: the polar and azimuthal angles of f_1 and f_3 in the rest frame of the vector bosons, $(\theta_1, 0)$ and (θ_3, ϕ_3) , respectively, and the masses of the two vector bosons. In this frame, the differential decay rate can be written as [46, 140]:

$$\begin{aligned} \frac{d\Gamma(H \rightarrow VV)}{dc_{\theta_1} dc_{\theta_3} d\phi_3} \sim & s_{\theta_1}^2 s_{\theta_3}^2 + \frac{1}{2\gamma_1 \gamma_3 (1 + \beta_1 \beta_3)} s_{2\theta_1} s_{2\theta_3} c_{\phi_3} \\ & + \frac{1}{2\gamma_1^2 \gamma_3^2 (1 + \beta_1 \beta_3)^2} [(1 + c_{\theta_1}^2)(1 + c_{\theta_3}^2) + s_{\theta_1}^2 s_{\theta_3}^2 c_{2\phi_3}] \\ & - \frac{4A_{f_1} A_{f_3}}{\gamma_1 \gamma_3 (1 + \beta_1 \beta_3)} \left[s_{\theta_1} s_{\theta_3} c_{\phi_3} + \frac{1}{\gamma_1 \gamma_3 (1 + \beta_1 \beta_3)} c_{\theta_1} c_{\theta_3} \right]. \end{aligned} \quad (2.13)$$

Here, c_θ and s_θ are shortcuts for $\cos \theta$ and $\sin \theta$, while β and $\gamma = (1 - \beta^2)^{-1/2}$ refer to the velocities of the vector bosons; for large Higgs boson masses, the vector bosons are boosted, thus $\beta \sim 1$ and $\gamma \rightarrow \infty$, so that all terms depending on $1/\gamma$ in Eq. 2.13 become negligible, and the dependence on the azimuthal angle disappears, as expected for longitudinally polarized vector bosons. The factors A_{f_i} correspond to combinations of the $Vf\bar{f}$ vectorial and axial couplings, c_V^f and c_A^f , defined in Eq. 1.24.

Integrating out the polar angles, the azimuthal distribution becomes [140]:

$$\frac{d\Gamma(H \rightarrow VV)}{d\phi_3} \sim 1 + a_1 c_{\phi_3} + a_2 c_{2\phi_3}. \quad (2.14)$$

On the other hand, for a CP-odd Higgs boson, the differential decay rate is given by [140]:

$$\frac{d\Gamma(A \rightarrow VV)}{dc_{\theta_1}dc_{\theta_3}d\phi_3} \sim 1 + c_{\theta_1}^2 c_{\theta_3}^2 - \frac{1}{2} s_{\theta_1}^2 s_{\theta_3}^2 - \frac{1}{2} s_{\theta_1}^2 s_{\theta_3}^2 c_{2\phi_3} - 2A_{f_1} A_{f_3} c_{\theta_1} c_{\theta_3} \quad (2.15)$$

which, after integration over the polar angles, reduces to:

$$\frac{d\Gamma(A \rightarrow VV)}{d\phi_3} \sim 1 - \frac{1}{4} c_{2\phi_3}. \quad (2.16)$$

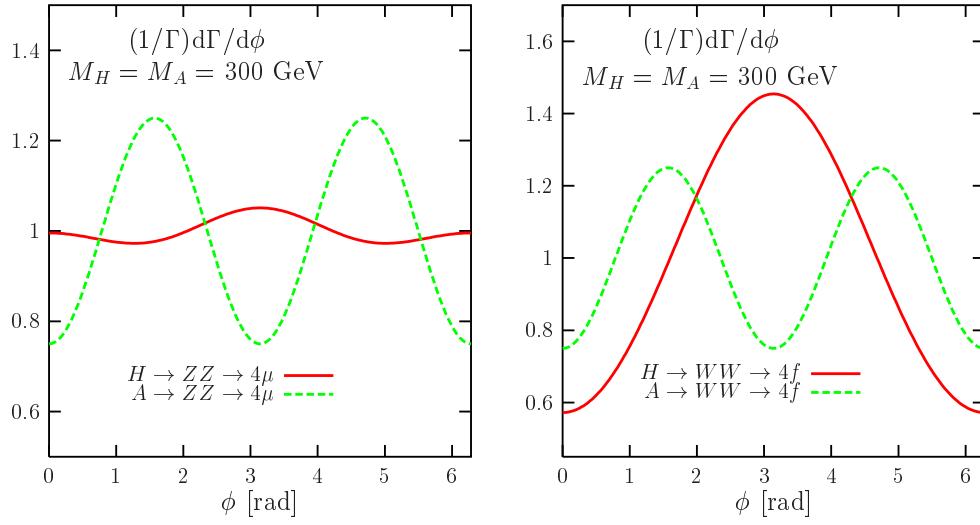


Figure 2.11: Azimuthal distributions of final-state fermions in the processes $H/A \rightarrow ZZ \rightarrow 4\mu$ (left) and $H/A \rightarrow WW \rightarrow 4f$ (right), comparing the CP-even and CP-odd hypotheses. The angle ϕ corresponds to ϕ_3 in Fig. 2.10. The mass is assumed to be $m_H = m_A = 300$ GeV [46].

The difference between the two CP hypotheses is illustrated in Fig. 2.11, which shows the azimuthal distribution of the final-state fermions for $H/A \rightarrow ZZ \rightarrow 4\mu$ on the left and for $H/A \rightarrow WW \rightarrow 4f$ on the right, with $m_H = 300$ GeV in both cases. In the $H \rightarrow ZZ$ case, the modulation with ϕ is small, since the term $\cos \phi_3$ is suppressed by the coefficient a_1 , with $a_1 \propto v_\ell^2 \ll 1$.

The partial width for the three-body decay in the CP-odd case, to be compared with Eq. 2.9 for the SM case, exhibits a different dependence on the mass [140]:

$$\Gamma(A \rightarrow VV^* \rightarrow Vf\bar{f}) = \frac{3G_\mu^2 m_V^6}{8\pi^3 m_A} \delta'_V \eta^2 R_A(x). \quad (2.17)$$

where $x = M_V^2/M_A^2$ and η accounts for the cross section normalization.

The distinction between CP-even and CP-odd states can be studied also in the CP-

mixing case, where the discovered resonance would be a superimposition of even and odd states, so close in mass to be observed as degenerate. According to Refs. [35,36], the general scattering amplitude for such a mixed X boson can be written as:

$$A(X \rightarrow V_1 V_2) = v^{-1} \left(g_1^{(0)} m_V^2 \epsilon_1^* \epsilon_2^* + g_2^{(0)} F_{\mu\nu}^{*(1)} F^{*(2),\mu\nu} + g_3^{(0)} F^{*(1),\mu\nu} F_{\mu\alpha}^{*(2)} \frac{q_\nu q^\alpha}{\Lambda^2} + g_4^{(0)} F_{\mu\nu}^{*(1)} \tilde{F}^{*(2),\mu\nu} \right) \quad (2.18)$$

where $F_{\mu\nu}$ is the field tensor for a generic gauge boson, $F_{\mu\nu} = \epsilon_\mu q_\nu - \epsilon_\nu q_\mu$, while $\tilde{F}_{\mu\nu}$ is the conjugate field tensor (cf. Eq. 2.11). The $g_i^{(0)}$ form factors are in general complex and momentum-dependent, with (0) labelling the spin quantum number. The first term in Eq. 2.18 is the usual SM CP-even contribution, while the fourth is the pure CP-odd term. $g_2^{(0)}$ multiplies a dimension-five operator ($\mathcal{L} \sim \frac{g_2^{(0)}}{v} X Z_{\mu\nu} Z^{\mu\nu}$) which accounts for the loop-induced Higgs boson decays into massless bosons, such as $\gamma\gamma$ or gg . Finally, the term introduced by $g_3^{(0)}$ is a dimension-seven operator, suppressed by the New Physics scale Λ^2 .

The SM decays $H \rightarrow WW, ZZ$ are described with $g_1^{(0)} = 1$ and $g_{2,3,4}^{(0)} = 0$. On the contrary, a pure pseudoscalar Higgs boson would have only $g_4^{(0)} \neq 0$.

The expression given in Eq. 2.18 is general enough to describe the decay to any on-shell gauge boson, massive or not, including CP-violating terms. This is achieved using both ϵ_i and $F_{\mu\nu}^{(i)}$, which are in fact not independent; reducing the terms from four to three, the amplitude becomes [35]:

$$A(X \rightarrow V_1 V_2) = v^{-1} \epsilon_1^{*\mu} \epsilon_2^{*\nu} \left(a_1 g_{\mu\nu} m_X^2 + a_2 q_\mu q_\nu + a_3 \epsilon_{\mu\nu\alpha\beta} q_1^\alpha q_2^\beta \right) \quad (2.19)$$

where the new $a_{1,2,3}$ coefficients are linked to the $g_{1,2,3,4}^{(0)}$ in the following way:

$$a_1 = g_1^{(0)} \frac{m_V^2}{m_X^2} + \frac{q_1 q_2}{m_X^2} \left(2g_2^{(0)} + g_3^{(0)} \frac{q_1 q_2}{\Lambda^2} \right), \quad a_2 = - \left(2g_2^{(0)} + g_3^{(0)} \frac{q_1 q_2}{\Lambda^2} \right), \quad a_3 = -2g_4^{(0)}. \quad (2.20)$$

Although not present at tree-level in the SM, both a_2 and a_3 may appear through radiative corrections: one loop for a_2 , whose scale is set around $\mathcal{O}(\alpha_{EW}) \sim 10^{-2}$, two or three loops (for decays into WW and ZZ respectively) for a_3 , which is $\mathcal{O} \sim 10^{-10}$ [141].

In summary, since the CP-even, SM-like component arises at tree-level, while possible CP-odd contributions are loop-induced, it would be mostly the former to dominate the angular and kinematical distributions of final-state particles in VV decays at the LHC [35,36,141,142]. In order for a significant deviation from the SM expectation to be observed, the CP-even component should be small, but in this case a clear drop would have been found in the measured cross section, which does

not appear to be currently the case [119].

2.3.2 The case of a spin-1 resonance

The amplitude for the decay of a spin-1 resonance into two identical massive vector bosons ($V = Z, W$) is given by [35, 36]:

$$A(X \rightarrow V_1 V_2) = g_1^{(1)} [(\epsilon_1^* q)(\epsilon_2^* \epsilon_X) + (\epsilon_2^* q)(\epsilon_1^* \epsilon_X)] + g_2^{(1)} \epsilon_{\alpha\mu\nu\beta} \epsilon_X^\alpha \epsilon_1^{*\mu} \epsilon_2^{*\nu} \tilde{q}^\beta \quad (2.21)$$

where $g_1^{(1)}$ and $g_2^{(1)}$ are dimensionless couplings for, respectively, a vector and pseudovector particle (for $J^P = 1^+$, the opposite for $J^P = 1^-$). All terms have been introduced already in Eq. 2.18, apart from $\tilde{q}$, which is the difference in momenta between the two gauge bosons, $\tilde{q} = q_1 - q_2$. In principle, as discussed in the following, the case $V = W$ would require additional degrees of freedom, since the W boson is charged [36], but they are not considered for a first hypothesis testing.

As stated above, the amplitude in Eq. 2.21 is valid only for massive particles. In fact, massive vector bosons cannot decay into two identical, real massless bosons, due to the Bose-Einstein symmetry. This result, known as the Landau-Yang theorem [37, 38], rules out the possibility of a spin-1 Higgs boson decaying into two on-shell photons, which seems to contradict the observation carried out by the ATLAS experiment in the $\gamma\gamma$ channel [39]; however, the presence of a spin-1 resonance decaying into pairs of W or Z bosons can still be tested experimentally.

2.3.3 The case of a spin-2 resonance

Randall-Sundrum gravitons

A spin-2 resonance at the TeV scale, with SM Higgs boson-like couplings, could be the one described by the Randall-Sundrum (RS) model for Kaluza-Klein (KK) gravitons [33, 34].

At variance with previous theories, which predict weakly coupled massive gravitons and require a large number of extra dimensions, e.g. the ADD model [143, 144], the RS model postulates strong couplings and only one extra dimension. This additional fifth dimension is warped, i.e. not factorizable from our standard 4-dimensional subset of the spacetime, which is referred to as a “3-brane”; the non-factorizable metric is obtained by means of an exponential function [33]:

$$ds^2 = e^{-2kr_c\phi} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2 \quad (2.22)$$

where $\eta_{\mu\nu}$ is the usual Minkowski metric $(-1, 1, 1, 1)$, k is a scale of the order of the Planck scale, x^μ are the 4-dimensional coordinates and ϕ is the coordinate of the

extra dimension, which is finite and periodic ($0 \leq \phi \leq \pi$) and whose size is set by the compactification radius r_c . The warp factor $e^{-2kr_c\phi}$ sets the relation between the reduced Planck mass, $\overline{m}_{\text{Pl}} = m_{\text{Pl}}/\sqrt{8\pi}$, and the graviton-SM coupling scale Λ [144]:

$$\Lambda = e^{-kr_c\pi} \overline{M}_{\text{Pl}}. \quad (2.23)$$

By tuning kr_c , it is possible to place Λ at the TeV scale, so that RS gravitons would be accessible at LHC energies with non negligible couplings. On the contrary, the large size of the extra dimensions in the ADD model ($r_c \gg 1/m_{\text{Pl}}$) causes the coupling scale also to be large, $\Lambda = \overline{M}_{\text{Pl}} \sim 10^{18}$ GeV, thus the coupling of gravitons with matter would be heavily suppressed, and they would be detectable in collider experiments only as missing energy [144].

The warp factor sets also the relation between the physical masses m , as experienced in the standard 3-brane, and the fundamental masses m_0 , which are expected to be around the Planck scale:

$$m = e^{-kr_c\pi} m_0. \quad (2.24)$$

In this way, it would be possible to obtain $m_0 \sim 10^{19}$ GeV and $m \sim 10^4$ GeV just setting $e^{kr_c\pi} \sim 10^{15}$, which means $kr_c \sim 50$. This would provide a solution to the hierarchy problem, since all the fundamental quantities would be at the Planck scale, with the exponential form of the warp factor pushing their physical counterparts at the weak scale, i.e. around the TeV [33].

The effective Lagrangian describing the interaction between KK gravitons and the SM fields, in both the RS and ADD models, is given by [144]:

$$\mathcal{L}_{\text{int}} = -\frac{1}{\Lambda} \sum_{\vec{n}} T^{(\vec{n})\mu\nu} \mathcal{T}_{\mu\nu} \quad (2.25)$$

where $\mathcal{T}_{\mu\nu}$ is the SM stress-energy tensor and $T^{(\vec{n})\mu\nu}$ is the $\vec{n}$ -th KK mode. While in the weakly coupled models the mass gap between neighboring gravitons is very small ($\Delta m = r_c^{-1}$) and thus their discrete spectrum can be approximated by a continuum, in the RS model the mass of the n -th excitation is given by $m_n = kx_n e^{-kr_c\pi} = m_1 x_n/x_1$, with x_n being the n -th root of the first order Bessel function, and m_1 the first excitation, possibly in the TeV range. Setting the mass of such a graviton at 125 GeV, in particular, the branching ratios in Table 2.3 are obtained [145]. While a light SM Higgs boson decays most probably in $b\bar{b}$ pairs, as discussed at the beginning of this Chapter, the graviton would decay preferentially into gg pairs. The other couplings branching ratios are a factor ~ 10 smaller than the SM case, and in particular the ratio $(X \rightarrow WW)/(X \rightarrow ZZ)$ is the same in the two hypotheses.

	gg	$b\bar{b}$	$\tau\tau$	WW^*	ZZ^*	$\gamma\gamma$
$m_H = 125 \text{ GeV}$	0.086	0.577	0.063	0.215	0.0264	2.28×10^{-3}
$m_\phi = 125 \text{ GeV}$	0.896	0.0588	9.2×10^{-3}	0.0291	3.6×10^{-3}	0.925×10^{-3}

Table 2.3: Comparison between the predicted branching ratios for the SM Higgs boson (as given in Table 2.2) and for a RS graviton, assuming the same mass of 125 GeV [87, 145].

Lagrangian

A massive spin-2 resonance can be described by the Fierz-Pauli Lagrangian, where it appears as a symmetric, traceless tensor of rank two, satisfying the following properties [146]:

$$G_{\mu\nu} = G_{\nu\mu}, \quad G^\mu{}_\mu = 0, \quad \partial_\mu G^{\mu\nu} = 0 \quad (2.26)$$

where the last equation corresponds to the requirement of positive energy for G . As shown in Ref. [147], G cannot couple to the SM fields through dimension-four operators, due to Lorentz and SM gauge symmetries; furthermore, imposing the flavour and CP symmetries, G has to couple through dimension-five operators, i.e. the same as the SM fields.

The amplitude for the decay of a spin-2 resonance into two identical vector bosons ($V = W, Z, \gamma, g$) is given by [35, 36]:

$$\begin{aligned}
A(X \rightarrow V_1 V_2) = & 2g_1^{(2)} G_{\mu\nu} F^{*(1)\mu\alpha} F^{*(2)\nu\alpha} + 2g_2^{(2)} G_{\mu\nu} \frac{q_\alpha q_\beta}{\Lambda^2} F^{*(1)\mu\alpha} F^{*(2)\nu\beta} \\
& + g_3^{(2)} \frac{\tilde{q}^\beta \tilde{q}^\alpha}{\Lambda^2} G_{\beta\nu} \left(F^{*(1)\mu\nu} F_{\mu\alpha}^{*(2)} + F^{*(2)\mu\nu} F_{\mu\alpha}^{*(1)} \right) + g_4^{(2)} \frac{\tilde{q}^\nu \tilde{q}^\mu}{\Lambda^2} G_{\mu\nu} F^{*(1)\alpha\beta} F_{\alpha\beta}^{*(2)} \\
& + m_V^2 \left(2g_5^{(2)} G_{\mu\nu} \epsilon_1^{*\mu} \epsilon_2^{*\nu} + 2g_6^{(2)} \frac{\tilde{q}^\mu q_\alpha}{\Lambda^2} G_{\mu\nu} (\epsilon_1^{*\nu} \epsilon_2^{*\alpha} - \epsilon_1^{*\alpha} \epsilon_2^{*\nu}) + g_7^{(2)} \frac{\tilde{q}^\mu \tilde{q}^\nu}{\Lambda^2} G_{\mu\nu} \epsilon_1^{*\mu} \epsilon_2^{*\nu} \right) \\
& + g_8^{(2)} \frac{\tilde{q}_\mu \tilde{q}_\nu}{\Lambda^2} G_{\mu\nu} F^{*(1)\alpha\beta} \tilde{F}_{\alpha\beta}^{*(2)} + m_V^2 \left(g_9^{(2)} \frac{G_{\mu\alpha} \tilde{q}^\alpha}{\Lambda^2} \epsilon_{\mu\nu\rho\sigma} \epsilon_1^{*\nu} \epsilon_2^{*\rho} q^\sigma \right. \\
& \left. + \frac{g_{10}^{(2)} G_{\mu\alpha} \tilde{q}^\alpha}{\Lambda^4} \epsilon_{\mu\nu\rho\sigma} q^\rho \tilde{q}^\sigma (\epsilon_1^{*\nu} (q\epsilon_2^*) + \epsilon_2^{*\nu} (q\epsilon_1^*)) \right) \quad (2.27)
\end{aligned}$$

where the terms are the same as those introduced in Eqs. 2.18 and 2.21, and an overall factor Λ^{-1} has been omitted. There are ten dimensionless $g_i^{(2)}$ couplings, eight of which are independent. The first seven terms conserve parity in case of $J^P = 2^+$, while the last three violate it (reversely for $J^P = 2^-$). In principle, the couplings are a function of kinematical invariants, e.g. $g_i^{(2)}(m_X^2, q_1^2, q_2^2)$, but such a dependence can be neglected since the general form of the couplings is anyway unknown. For the same reason, the amplitude in Eq. 2.27 would require more degrees of freedom in case of non identical vector bosons, i.e. $X \rightarrow W^+ W^-$, but it is generally agreed that this form is convenient for a first hypothesis testing, and could eventually be refined in case of anomalies emerging from the data.

The minimal set of couplings includes only dimension-five operators, as discussed above: this is achieved setting all couplings to zero apart from $g_1^{(2)} = g_5^{(2)} \neq 0$. At leading order, such a boson couples to gg and $q\bar{q}$ with 96% and 4% probabilities, respectively [144, 148]. This model is referred to in the following as 2_m^+ , and its amplitude corresponds to [35]:

$$A(X_m \rightarrow V_1 V_2) = \Lambda^{-1} \left[2g_1^{(2)} G_{\mu\nu} F^{*(1)\mu\alpha} F^{*(2)\nu\alpha} + m_V^2 \left(2g_5^{(2)} G_{\mu\nu} \epsilon_1^{*\mu} \epsilon_2^{*\nu} \right) \right] \quad (2.28)$$

which reduces to $\propto G_{\mu\nu} F^{*(1)\mu\alpha} F^{*(2)\nu\alpha}$ in case of massless bosons. It should be noticed that $g_1^{(2)} \neq 0$ allows X_m production through gluon-fusion.

Another interesting model, involving higher order operators, is the so-called 2_h^+ , which is obtained by setting all couplings to zero apart from $g_4^{(2)} \neq 0$. The corresponding pseudo-tensor, 2_h^- , is obtained by setting $g_8^{(2)} \neq 0$ ($F_{\alpha\beta}^{*(2)} \leftrightarrow \tilde{F}_{\alpha\beta}^{*(2)}$).

For what concerns the couplings of X to two fermions, the most general form contains four independent couplings $\rho_i^{(2)}$ [35, 36]:

$$A(X \rightarrow q\bar{q}) = \Lambda^{-1} G^{\mu\nu} \bar{u}_{q_1} \left[\gamma_\mu \tilde{q}_\nu (\rho_1^{(2)} + \rho_2^{(2)} \gamma_5) + \frac{m_{q\bar{q}}}{\Lambda^2} (\rho_3^{(2)} + \rho_4^{(2)} \gamma_5) \right] v_{q_2} . \quad (2.29)$$

In order to have $q\bar{q} \rightarrow X$ production, one needs $\rho_1^{(2)} \neq 0$. Due to momentum composition, fermion-antifermion pairs cannot couple to spin-2 resonances at leading order ($\frac{1}{2} \otimes \frac{1}{2} \neq 2$), unless orbital angular momentum comes into play. For this reason, the observation of substantial, i.e. SM-like, rates for $H \rightarrow \tau\tau$ or $b\bar{b}$ final states, would rule out quite convincingly the spin-2 hypothesis. For the same reason, it is expected that X_2 would be produced mainly through gluon-fusion (with 96% probability at leading order, as stated above), with the vector-boson fusion production being possible but subleading for a mass around 125 GeV. The fractional couplings to gg and $q\bar{q}$ could, however, be modified substantially by higher order corrections: to account for all the possible production modes, the spin analysis documented in Chapter 5 is performed for various coupling configurations, with the $q\bar{q}$ fraction, $f_{q\bar{q}}$, going from 0.00 to 1.00 in steps of 0.25.

Kinematics in the WW final state for a spin-0 or a spin-2 boson

Kinematical and angular distributions of the final-state leptons depend on the spin combinations which give rise to the process $gg \rightarrow X \rightarrow WW^*$, with X being either a spin-0 or spin-2 boson. The discussion hereafter follows Ref. [149].

Quantizing the system along the z axis, a spin state is defined by its quantum

numbers $|J, J_z\rangle$. A spin-2 particle X_2 can be represented in the following basis:

$$\begin{aligned}\epsilon^{s\mu\nu} &= (|2, +2\rangle, |2, +1\rangle, |2, 0\rangle, |2, -1\rangle, |2, -2\rangle) = (\epsilon^{+2\mu\nu}, \epsilon^{+1\mu\nu}, \epsilon^{0\mu\nu}, \epsilon^{-1\mu\nu}, \epsilon^{-2\mu\nu}) \\ &= (\epsilon^{+\mu}\epsilon^{+\nu}, \frac{1}{\sqrt{2}}(\epsilon^{+\mu}\epsilon^{0\nu} + \epsilon^{0\mu}\epsilon^{+\nu}), \frac{1}{\sqrt{6}}(\epsilon^{+\mu}\epsilon^{-\nu} + \epsilon^{-\mu}\epsilon^{+\nu} + 2\epsilon^{0\mu}\epsilon^{0\nu}), \\ &= \frac{1}{\sqrt{2}}(\epsilon^{-\mu}\epsilon^{0\nu} + \epsilon^{0\mu}\epsilon^{-\nu}), \epsilon^{-\mu}\epsilon^{-\nu}) \quad (2.30)\end{aligned}$$

where $\epsilon^\pm$ and ϵ^0 are the transverse and longitudinal polarizations, respectively. Not all combinations can be generated via gluon-fusion. In fact, since gluons are massless spin-1 bosons, they have access with equal probability to two states, $|1, 1\rangle$ and $|1, -1\rangle$, assuming there is no initial-state radiation or transverse momentum. The two-gluon state will then be a composition of $|2, +2\rangle$, $|2, -2\rangle$ and $|2, 0\rangle$ only, with the $|2, 0\rangle$ state having a relative $1/3$ normalization with respect to the others, given by the corresponding Clebsch-Gordan coefficient $\langle 2, 0 | 1, \pm 1 \rangle |1, \mp 1\rangle$. The density matrix is diagonal, i.e. a $|J, J_z\rangle$ two-gluon state decays into a boson having again $|J, J_z\rangle$ quantum numbers. Actually, as computed in Ref. [149], the $|2, 0\rangle$ term does not contribute to the amplitude, thus only transversely, evenly polarized states, $\epsilon_1^+\epsilon_2^+$ and $\epsilon_1^-\epsilon_2^-$, do.

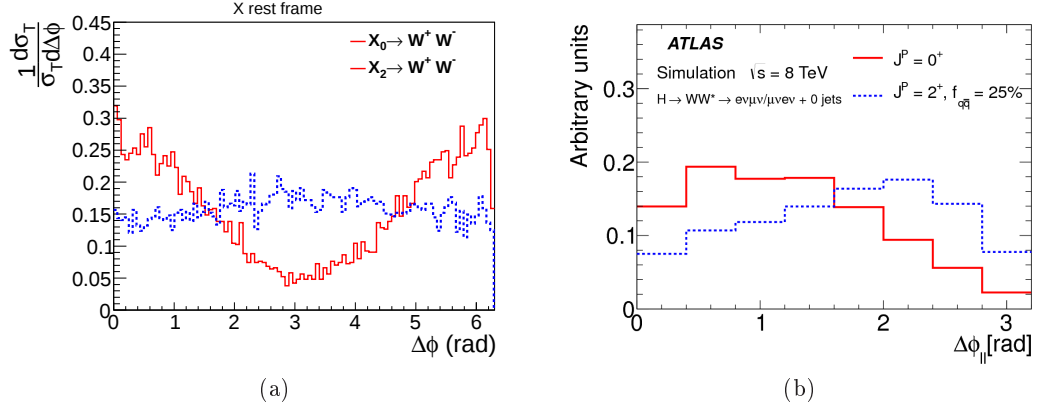


Figure 2.12: Dependence on the dilepton azimuthal angle for final-state leptons, (a) in the X rest frame [150] and (b) in the laboratory frame [41]. The distributions are normalized to unity; in the latter case, the detector simulation and the preselection cuts are applied, following Chapter 5. In both cases, the spin-2 boson is generated in the $f_{q\bar{q}} = 0.25$ hypothesis.

Using the V-A structure for the $W^- \rightarrow \ell^- \bar{\nu}$ decay amplitude, $\mathcal{M}_1 = \bar{u}(p_{\ell^-}) \gamma^\mu \epsilon_{1\mu} (1 - \gamma_5) v(p_{\bar{\nu}})$, and the same for W^+ , the angular dependence of the cross sections for $X_{0,2}$ production and decay is obtained. This is shown in Fig. 2.12, which compares the azimuthal distribution in the X rest frame at generation level (a), and in the laboratory frame after detector simulation and preselection cuts (b). While, in the

former case, the spin-2 distribution exhibits very little dependence on $\Delta\phi_{\ell\ell}$, in the latter case values close to π are preferred by leptons arising from a spin-2 resonance, while for spin-0 smaller angles are chosen in both configurations. In summary, in the azimuthal plane, leptons are mostly emitted close to each other or back-to-back, if produced by a spin-0 or spin-2 resonance, respectively.

For what concerns the polar dependence, the fact that the W bosons need opposite (equal) polarizations in order to produce a spin-0 (spin-2) state, induces a correlation between the polar angles of the outgoing leptons for spin-0, and an anti-correlation for spin-2, as schematically illustrated in Fig. 2.13.

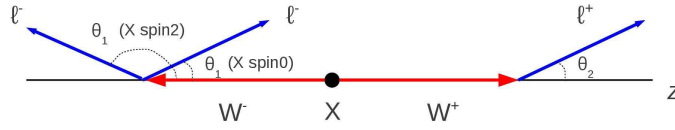


Figure 2.13: Schematic representation of the polar angles in the $X \rightarrow W^+W^- \rightarrow \ell^+\nu_\ell\ell^-\bar{\nu}_\ell$ decay of a spin-0 or spin-2 resonance.

An important consequence of the polar collinearity for a spin-0 resonance, is that also the invariant mass of the dilepton system, $m_{\ell\ell}$, defined as⁽³⁾:

$$m_{\ell\ell} = \sqrt{E_{\ell\ell}^2 - \vec{p}_{T,\ell\ell}^2} \simeq \sqrt{2E_{\ell_1}E_{\ell_2}(1 - \cos^2\theta_{\ell\ell})} \quad (2.31)$$

tends to smaller values than in the spin-2 case. The corresponding invariant mass distributions for leptons arising from a spin-0 or spin-2 resonance are shown in Fig. 2.14 (a). For the same reason, the $p_{T,\ell\ell}$ distribution prefers larger values for spin-0 than for spin-2, as can be seen in Fig. 2.14 (b).

These properties are exploited in Chapters 4 and 5 to distinguish the Higgs boson candidate events from background processes and from the spin and parity alternative hypotheses.

³This definition assumes energy and momentum to be the same, given that lepton masses can be neglected.

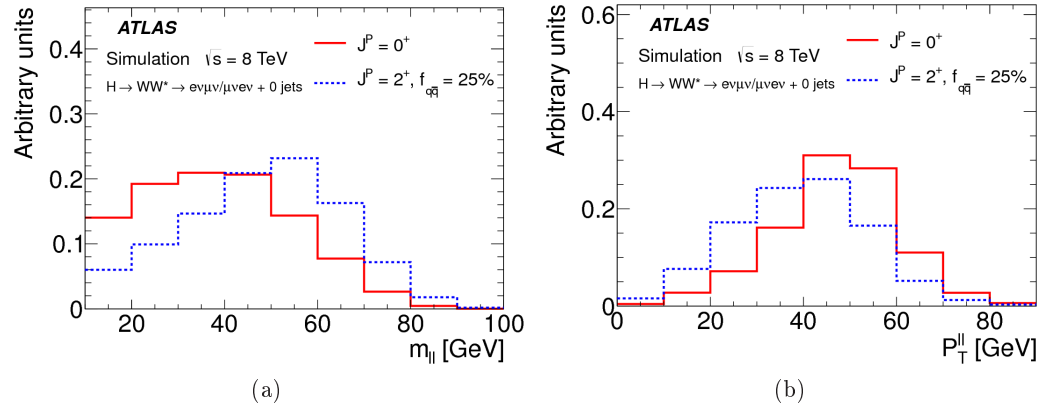


Figure 2.14: Comparison of (a) invariant mass and (b) transverse momentum distributions for the dilepton system arising from a spin-0 (in red) or spin-2 (in blue, in the $f_{q\bar{q}} = 0.25$ hypothesis) resonance [41]. Detector simulation and preselection acceptance cuts are applied, following Chapter 5.

Chapter 3

LHC, ATLAS and Performance

In the first part of this Chapter, the LHC accelerator complex and the ATLAS detector are described. For the latter, a review of the various subdetectors is presented: the inner detector, the electromagnetic and hadronic calorimeters, the muon spectrometer and the trigger system. In the second part of the Chapter, the performance of the ATLAS detector at $\sqrt{s} = 7$ and 8 TeV is discussed, with special emphasis on the objects used in the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ analysis: tracks and vertices, electrons, muons, jets and missing transverse energy.

3.1 The Large Hadron Collider

The Large Hadron Collider [21–23] is the largest and most powerful particle accelerator ever operated. It consists of a 27 km ring, located underground at a mean depth of 100 metres, in the former LEP tunnel, in Geneva (Switzerland), and is part of the CERN⁽¹⁾ accelerator complex, depicted in Fig. 3.1. The LHC is a proton-proton collider, designed to run at a centre-of-mass energy of 14 TeV. It hosts two multi-purpose detectors, ATLAS and CMS [151], and two smaller ones, ALICE and LHCb. The ALICE experiment [152] focuses on Pb-Pb and Pb-p collisions (the Pb beams are also delivered by the LHC, in shorter and dedicated runs) to study heavy-ion physics, in particular the so-called quark-gluon plasma, which represents the “super-dense” state of matter that may have existed in the early phases of the universe. The LHCb experiment [153] looks for CP-violation effects through the study of B -hadron decays.

The protons are extracted from a hydrogen source by means of an electric field which strips off the electrons, and are transferred to a linear accelerator, the Linac2, where

¹CERN stands for Conseil Européen pour la Recherche Nucleaire. Other acronyms used throughout this Section are CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider Beauty).

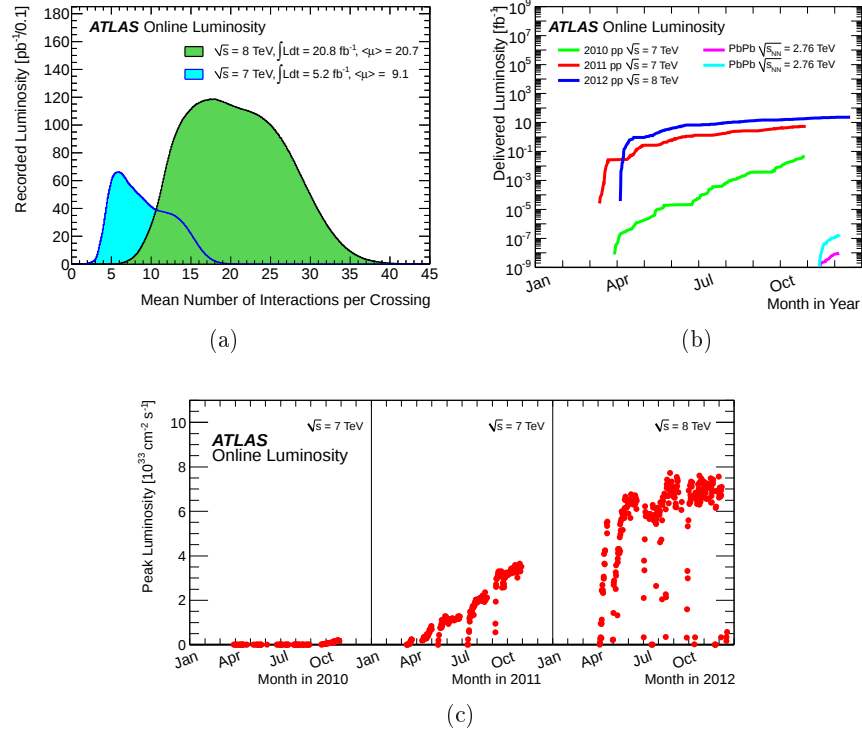


Figure 3.2: (a) Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the 2011 and 2012 data. (b) LHC delivered luminosity versus time for 2010, 2011 and 2012 runs, including both p-p and Pb-Pb data. (c) Instantaneous peak luminosity delivered to the ATLAS experiment per day as a function of time, during the p-p runs of 2010, 2011 and 2012. All plots are taken from Ref. [156].

where N_p is the number of protons per bunch (beam intensity), N_b the number of bunches, f the frequency of bunch crossing and $\sigma_{x,y}$ the beam sizes in the Gaussian approximation. Typical values for N_p and N_b are $1.48 \cdot 10^{11}$ and 1380, for the 2012 run at $\sqrt{s} = 8$ TeV. The resulting peak instantaneous luminosities is illustrated in Fig. 3.2 (c) as a function of time.

Given these extremely large values for N_p and N_b , multiple inelastic p-p interactions can occur in the detector alongside the main hard one. This phenomenon is called *in-time pile-up* when happening between protons of the same bunch, and *out-of-time pile-up* when the protons belong to different bunches. The mean number of interactions per bunch crossing $\langle \mu \rangle$ is shown in Fig. 3.2 (a), for 2011 data ($\langle \mu \rangle = 9.1$) and 2012 ($\langle \mu \rangle = 20.7$).

The bunch spacing has been chosen to be $n_s = 50$ ns in 2012, rather than the design value of 25 ns, in order to minimize electron cloud instabilities and to reduce the total energy stored in the beam, $K_{\text{tot}}^2 \propto n_s^{-1}$. Moving to $n_s = 25$ ns for the Phase I run would ensure a decrease in pile-up of a factor two, given that all other

conditions remain the same [157, 158].

3.1.1 Past and future operations

The operations of the LHC machine and of the ATLAS detector have been carried out successfully in 2011 and 2012. The luminosity delivered from the LHC to the ATLAS experiment is represented in Fig. 3.2 (b) in logarithmic scale; ATLAS performed extraordinarily well, with 95.5% of delivered data being flagged as “AllGood” for physics analyses.

After having accumulated data corresponding to an integrated luminosity of 4.7 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$ and of 20.7 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$, and shorter heavy-ion runs, the LHC has been turned off in February 2013 for the so-called Long Shut Down 1 (LS1), a period of major repairs and substitutions which is foreseen to last until Spring 2015. Afterwards, operations will resume with an energy of 13 TeV in the centre-of-mass, to be progressively ramped up to the nominal value of 14 TeV. Instantaneous luminosity will further increase with respect to 2012, $L \approx 1 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, with a bunch spacing of 25 ns as discussed above. These parameters should allow to collect an integrated luminosity of 100 fb^{-1} until the end of 2017 (Phase I). The whole 2018 will be devoted to the second long shutdown (LS2), which will consist of extensive upgrades to the LHC collimators and injectors (among others, LINAC4 (160 MeV) will replace LINAC2, a new PS Booster (2 GeV) will replace the current one, while PS2 (50 GeV) will replace PS). The instantaneous luminosity will be doubled, $L \approx 2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, with the prospect to accumulate up to 350 fb^{-1} of 14 TeV data until the end of 2022 (Phase II). After that, depending on the unveiled physics scenarios, an ultimate high-luminosity phase (HL-LHC) is foreseen, which will allow the experiments to collect up to 3000 fb^{-1} in ten years, in order to observe rare processes like H self-couplings or $H \rightarrow \mu\mu$, and to measure H properties and branching ratios with a precision up to a few percent, only to mention the Higgs boson sector.

3.2 The ATLAS detector

The ATLAS experiment has been designed to provide precision measurements over a wide range of physics processes, interaction rates and energies, detecting all the known Standard Model processes with excellent precision and being able to extend its reach to a vast range of exotic final states. The main challenges are represented by the unprecedented collision centre-of-mass, instantaneous luminosity, radiation doses and pile-up. The ATLAS detector consists of four main parts, around the interaction point (IP):

1. the inner detector, whose performance is characterized by good momentum resolution and high reconstruction efficiency for charged particles originating from the IP, allowing for detection and reconstruction of primary and secondary vertices;
2. the calorimeter system, including electromagnetic and hadronic calorimeters, for identifying electrons, photons, jets and missing transverse energy with great accuracy;
3. the muon spectrometer, to detect muons over a wide range of momenta;
4. the magnetic system, composed of a central solenoid for the inner detector and toroids for the muon spectrometer.

Important features are also the fast and reliable trigger system, which has to face ever-increasing rates, and the *hermiticity* of the detector, providing full coverage in ϕ and large coverage in η .

The detector has a forward-backward symmetric structure around the beam pipe, and is divided into a central structure (the *barrel*) and two orthogonal structures in the forward directions (the *end-caps*). Its size is of 45 metres in length and 25 metres in diameter, with a total weight of about 7000 tons. It has been designed and built over the last 20 years (Letter of Intent, Oct. 1992 [24]) thanks to the common effort undertaken by the ATLAS collaboration, which involves roughly 3000 physicists at 175 institutions in 38 countries.

In the following, a description of the individual subdetectors, shown schematically in Fig. 3.3, is presented, following Refs. [25,26], if not stated otherwise.

The ATLAS coordinate system

The coordinate system is a cartesian frame having its origin at the nominal interaction point, the z axis along the beam direction (with the positive side set by the anti-clockwise beam, as seen from above), the x and y axes in the transverse plane. The x axis points towards the centre of the LHC ring, so that y is directed upwards. Angles are defined as follows:

$$\theta = \tan^{-1}(\sqrt{x^2 + y^2}/z) \quad (3.2)$$

is the polar angle, measured with respect to the z axis, and:

$$\phi = \tan^{-1}(x/y) \quad (3.3)$$

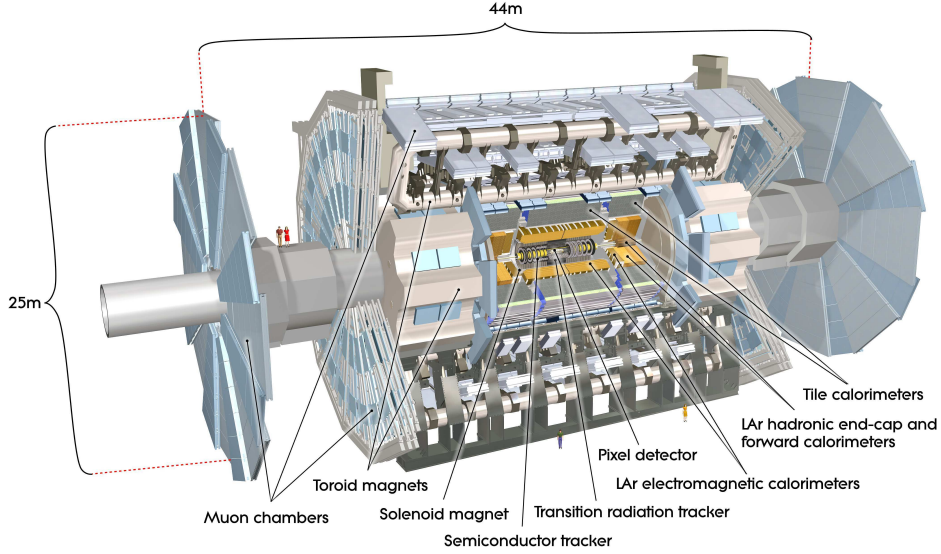


Figure 3.3: Schematic view of the subcomponents of the ATLAS detector. Shown are, in particular, the inner detector, the electromagnetic and hadronic calorimeters, the muon spectrometer, the magnets and the trigger system [25].

is the azimuthal angle in the $x - y$ plane. The polar angle θ is used to define the rapidity y and pseudorapidity η variables, as described in Section 1.4. Pseudorapidity is the massless limit of rapidity, used for light objects; from its definition in Eq. 1.47, it follows that η is maximum in correspondence with the beam axis ($\eta = \infty$ for $\theta = 0$), then decreases when moving away from z , up to $\eta = 0$ for $\theta = \pi/2$. All transverse quantities are defined as $q_T = |\vec{q}| \sin \theta$ in the $x - y$ plane. Finally, a distance in the three-dimensional volume is computed as $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$.

3.2.1 Inner detector

The inner detector (ID) tracks the electrically charged particles originating from the interaction point, measures their directions and transverse momenta with great accuracy and allows for vertex reconstruction. Charged particles inside a magnetic field experience a bending force, with a curvature radius proportional to their momentum; in the ID the bending magnetic field is provided by a central solenoid (5.3 m long, 2.5 m wide in diameter) generating a strength of 2 Tesla. The inner detector consists of three components, at increasing distances from the interaction point: the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT), all depicted in Fig. 3.4. In total, the ID is 7 m long and 1.15 m wide. From the mechanical point of view, the ID is made of three units, a central barrel and two identical end-caps.

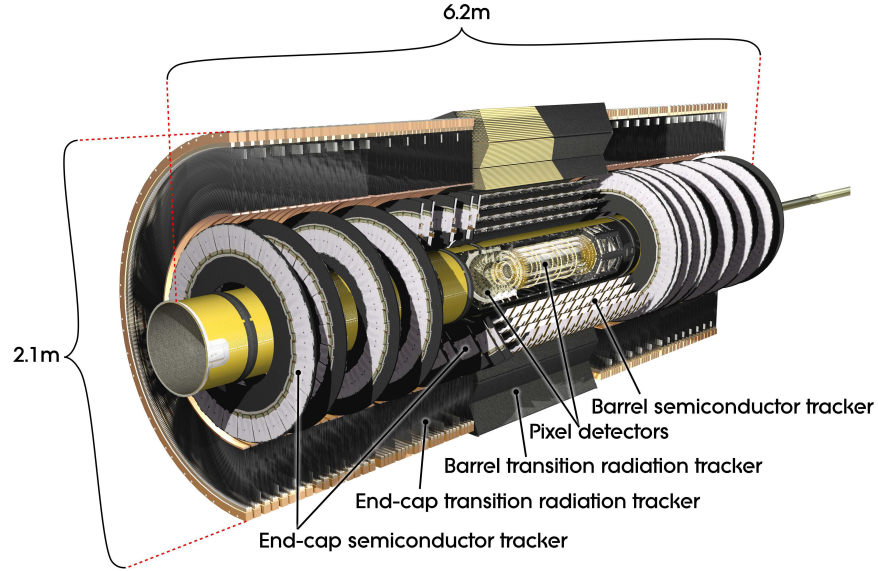


Figure 3.4: The inner detector of the ATLAS experiment. Its three components (Pixel Detector, Semiconductor Tracker, Transition Radiation Tracker) are shown [25].

Given the very large track density in a proton-proton collider and the small distance from the IP, tracking detectors need to have fine granularity, fast response and radiation-hardness; semiconductor trackers, based on silicon, possess all these features, both in the form of pixels and of strips. In the outer region (beyond a radius of 56 cm), precision tracking is substituted by continuous tracking, provided by the TRT, which ensures a high number of hits⁽²⁾ (36 per track on average) at the cost of lower precision. This design provides full coverage (up to $|\eta| < 2.5$) without introducing too much dead material, yielding an average number of three hits in the pixels and eight hits in the strips (four space points).

In the following, the three components of the ID are described.

The pixel detector

The tracking element closest to the beam pipe is the Pixel Detector. Due to its location, this unit is responsible for providing excellent impact parameter resolution and for detecting short-lived particles as τ -leptons and B -hadrons. As illustrated in Fig. 3.5, the system consists of three concentric barrels (at a radius of 5.05, 8.85 and 12.25 cm from the IP respectively) and of three transverse disks in the end-cap on either side (at increasing distance from the IP, up to 6.50 cm).

²A hit is a channel with an output signal above a certain threshold.

Each one of these elements is constituted of identical modules, containing a total of ~ 80 million pixels; there are 1456 modules in the barrel and 288 in the end-caps. Pixels are read out by chips (16 per module), that are provided with buffering electronics in order to store the data while waiting for level-1 trigger decision (see Section 3.2.4).

The most internal layer in the barrel is the so-called B-layer, fundamental for tagging secondary vertices from B -hadron decays. Due to its proximity to the IP, this layer has the shortest lifetime: to maintain robust tracking despite effects arising from increasing luminosity, increasing radiation and hardware lifetime, a new pixel layer will be added during LS1 between the beam pipe and the B-layer, the so-called Insertable B-layer (IBL). For this being possible, the beam pipe itself needs to be changed, reducing its radial size from the actual value of $29 < R < 36$ mm to $25 < R < 29$ mm, allowing the IBL installation at a distance of $31 < R < 40$ mm from the beam axis. Thanks to the IBL presence, resolution for primary vertex reconstruction will increase from the actual $15 \mu\text{m}$ in $x - y$ to $11 \mu\text{m}$, and from $34 \mu\text{m}$ in z to $24 \mu\text{m}$, while the B-layer foreseen lifetime will increase from 300 fb^{-1} to 550 fb^{-1} , at a design luminosity up to $3 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.

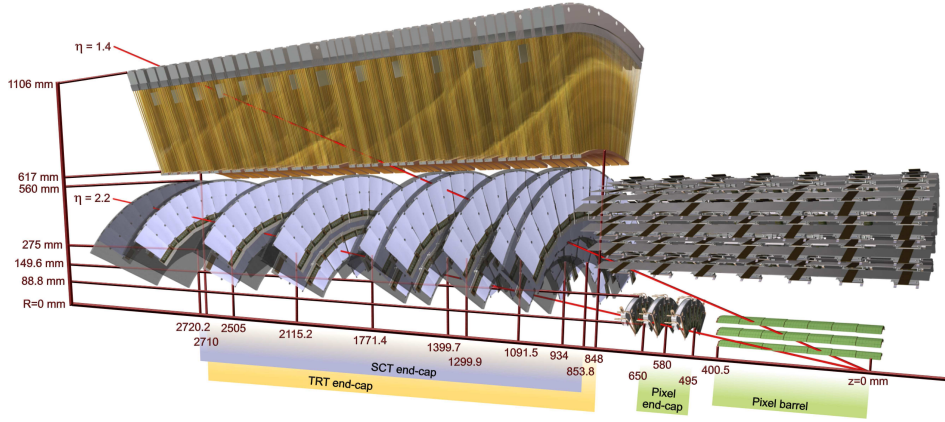


Figure 3.5: Break-down of the internal structure of the ID, showing the different shapes of the disks and of the layers, both in the barrel and in the end-caps [159].

The semiconductor tracker

The second detector crossed by a particle travelling from the IP to the outside, is the Semiconductor Tracker. Unlike the Pixel Detector, it is based on silicium technology, with sensors organized in strips rather than in pixels. The active detector component is constituted by p-on-n micro-strips, glued together in pairs, with a relative rotation of 40 mrad to ensure the measurement of the third dimension. There are four cylindrical layers in the barrel (extending up to $|\eta| < 1.4$) and nine disks in the

end-caps, as shown in Fig. 3.5. The end-cap strips are similar to those in the barrel, but have a tapered shape, extending the coverage up to $|\eta| < 2.5$. The SCT layout is such to provide eight measurements per track, with a spatial resolution of $17\ \mu\text{m}$ in $R - \phi$ and $580\ \mu\text{m}$ in z (in R for the end-caps) per module.

Both the pixel and the SCT systems are placed in a cold environment in order to improve their operational stability and to limit the radiation damage. The read-out chips of the pixels dissipate $15\ \text{kW}$ into the detector volume, while the silicon sensors dissipate more than $25\ \text{kW}$. The silicon temperature has to be maintained $\leq 7^\circ\text{C}$ degrees for the SCT and $\sim 0^\circ\text{C}$ for the pixels, to prevent reverse annealing of the silicon detectors. This implies a coolant temperature of -25°C , to be obtained while introducing a minimal amount of extra material into the detector volume. These features are satisfied by an evaporative cooling system, based on Octafluoropropane (C_3F_8).

The transition radiation tracker

The outermost ID component is the Transition Radiation Tracker. It is able to track charged particles, by means of drift straw tubes, and to identify electrons, which produce detectable X -rays in the interleaved radiators.

Drift tubes are filled with a gas mixture (70% Xe, 27% CO_2 , 3% O_2) and kept at high voltage of negative polarity; at the passage of charged particles, electrons from the medium get free and are collected on the anode wire. There are around 50000 straw tubes in the barrel, organized in three main rings, covering up to $|\eta| = 0.8$, while those in the end-caps, organized in wheels, extend the coverage up to $|\eta| = 2.0$.

The tubes are interleaved with foils of Polypropylene to induce transition radiation. When a charged particles crosses the boundary between two media with different dielectric constants, it emits forward radiation proportionally to its Lorentz factor, $\gamma = E/m$, a property that can be used to discriminate electrons from heavier particles. Nevertheless, the tracking capability has been preferred over the electron discrimination in the TRT design – the latter would have in fact benefitted from a greater path length and fewer straw detectors.

As already noted above, the resolution of TRT modules is not as good as the one of the pixel and SCT components, but this is compensated by the large number of hits expected. There is also no measurement of the third coordinate, since only the $R - \phi$ dimension can be exploited through the measurement of the drift time, with an intrinsic accuracy of $130\ \mu\text{m}$ per tube.

3.2.2 Calorimeters

After the inner detector, where charged particles are tracked and their interaction vertex and momentum are measured, particles coming from the IP reach the calorimeter system. It comprises electromagnetic and hadronic calorimeters, which aim to measure the energy of charged and neutral particles with high precision, and to provide a first classification. The latter goal is achieved analyzing the lateral and longitudinal shapes of the energy depositions inside the calorimeters. As shown in Figs. 3.6 and 3.7 (b), the inner part of the calorimeter system is constituted by a presampler, followed by the barrel and end-cap electromagnetic calorimeters (EMCal and EMEC) covering the pseudorapidity region $|\eta| < 3.2$, the hadronic barrel calorimeter covering $|\eta| < 1.7$, the hadronic end-cap calorimeters covering $1.5 < |\eta| < 3.2$ and finally the forward calorimeters (FCal), serving both as electromagnetic and hadronic detectors, in the region $3.1 < |\eta| < 4.9$. They are all sampling calorimeters, made of alternated absorbers and active material layers. In order to measure the energy of the originary particle in a reliable way, the shower has to be fully contained in the detector: to this extent, the calorimeter system extends up to $|\eta| < 4.9$ and its size corresponds roughly to 22 radiation lengths⁽³⁾ X_0 in the barrel and more than 24 in the end-caps. The containment of hadronic showers, which are on average broader and less regular, is achieved in the hadronic calorimeter through a total thickness of about 10 interaction lengths⁽⁴⁾. Such a thickness is found to be adequate also to provide excellent resolution for high-energy jets, while the wide coverage ensures an equally good measurement of missing transverse energy.

Electromagnetic calorimeters

Both the barrel and end-cap electromagnetic calorimeters use lead as absorber and liquid argon (LAr) as active material. The layers are accordion-shaped, ensuring a complete azimuthal coverage, without cracks. The size of the LAr gaps is constant in the barrel (2.1 mm), while it varies in the end-caps in order to compensate for increasing radii.

The barrel calorimeter ($|\eta| < 1.475$) is divided into two identical halves by a crack (6 mm) at $z = 0$, while each end-cap calorimeter consists of two coaxial wheels, with the inner and outer wheels covering the regions $2.5 < |\eta| < 3.2$ and $1.375 < |\eta| < 2.5$, respectively, as shown in Fig. 3.7 (b). Since both the barrel and the end-cap calorimeters are contained inside a cryostat, in the transition region

³A radiation length is defined as the distance travelled in the medium by a high-energy electron before losing all but $1/e$ of its energy via bremsstrahlung.

⁴An interaction length is defined as the mean distance travelling which, hadrons passing through the medium are reduced by a factor $1/e$ via inelastic hadronic interactions.

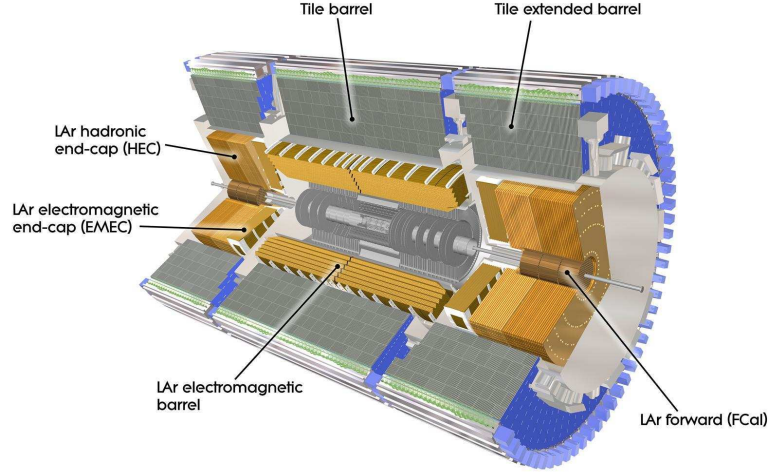


Figure 3.6: Cutaway view of the ATLAS calorimeter system. Shown are the LAr electromagnetic calorimeter, the tile hadronic one and the FCal [25].

between the two there is a sizeable amount of material; for this reason, the region between 1.37 and 1.52 in $|\eta|$ is not used for precision measurements but rather for storing cables and services (crack region).

Between the solenoid which surrounds the ID and the EMCal, there is an additional layer of LAr (1.1 cm thick in the barrel, 0.5 cm in the end-caps, covering up to $|\eta| < 1.8$) which serves as presampler, quantifying the energy lost by photons and electrons in traversing the ID material.

Over the region devoted to precision physics ($|\eta| < 2.5$), the EMCal components are segmented into three layers, so that the granularity is maximum closer to the IP (see Fig. 3.7 (a)). The first layer is the most precise one, being constituted by strips as small as 0.003×0.1 in units of $\Delta\Phi \times \Delta\eta$; it is responsible for providing a precise position measurement along the first $\sim 4 X_0$, so that electrons and photons can be better distinguished from e.g. pions. The second layer is made of square cells with a granularity of 0.025×0.025 ; most of the shower energy from electrons and photons is deposited inside this layer, thanks to its thickness of $\sim 16 X_0$. Finally, the third section has a coarser structure, with towers of 0.025×0.05 , and a thickness varying between 2 and $12 X_0$, and plays an important role in discriminating between electromagnetic and hadronic showers. In the end-cap region, the layers are only two and have overall a coarser granularity.

The intrinsic energy resolution provided by the EMCal is, after noise subtraction, $\sigma(E)/E = 0.1/\sqrt{E} \oplus 0.002$, dominated by the stochastic term $0.1/\sqrt{E}$.

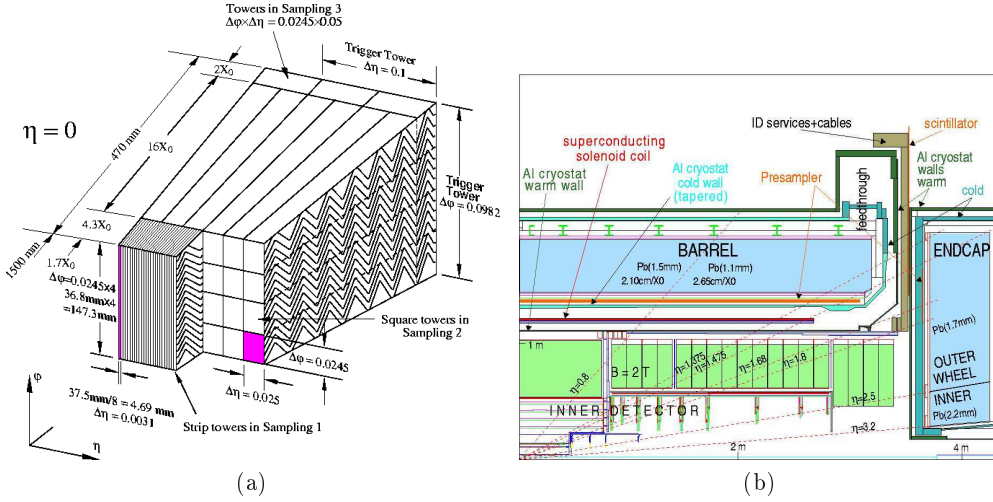


Figure 3.7: (a) Schematic view of a component of the electromagnetic calorimeter, displaying its accordion shape and the sizes of the three samplings. (b) Cutaway of the calorimeter system, showing the η coverage of the various subcomponents, and the location of the services. Both figures are taken from Ref. [25].

Hadronic calorimeters

The hadronic calorimeters are different in the barrel and in the end-caps, to cope with the strongly varying requirements and radiation environments found over the large η range. The barrel hosts a tile calorimeter, made of iron and scintillating plastic tiles. Together with the extended barrel components, it covers the region up to $|\eta| < 1.7$. The tiles are 3 mm thick, arranged in a structure which is periodic in z ; the signal they produce is collected by wavelength shifting fibers and read out into photomultipliers. In the end-caps, each hadronic calorimeter consists of two independent wheels: the first wheel is equipped with 25 mm copper plates, the second one, further from the IP, less precisely with 50 mm plates. Together, they extend the coverage up to $|\eta| < 3.5$.

The outermost region, at a distance of 4.7 m from the IP, is covered by the FCal, which is located on the beam axis, in the same cryostat of the EMCal and HCAL end-cap calorimeters. This design allows an uniform coverage, but to cope with the extreme radiation doses, the FCal front face is recessed of about 1.2 m with respect to the EMCal, minimizing neutron albedo. In order to avoid shower leaks, the FCal needs also to be highly dense. Such a compact and radiation-hard detector is in fact accomplished by a threefold structure: one section in the front made of copper, two sections in the back made of tungsten, all filled with LAr. The first is optimized for detecting electromagnetic showers, the other two for hadronic showers.

3.2.3 Muon spectrometer

After travelling through the ID and the calorimeter system, most visible particles have been tracked, measured and absorbed, the only exception being muons (apart from rare cases of punch-throughs from the calorimeters). A muon is a so-called *minimum ionizing particle*, since it produces negligible bremsstrahlung radiation due to its relatively large mass. In order to measure the momentum of muons with energies in the range $\mathcal{O}(1 - 100)$ GeV, the muon spectrometer (MS) is equipped with eight large superconducting air-core toroids in the barrel, surrounding radially the beam axis, and two smaller toroids in the end-caps. They are 25 m long in the barrel (5 m each in the end-caps), with a radial extension from 9 to 20 m (1.6 – 11 m), and generate a magnetic field varying from 0.15 to 2.5 Tesla (0.2 – 3.5 Tesla), depending on the space point. Such a configuration ensures that the magnetic field is mostly orthogonal to the muon trajectories, at the same time minimizing the effect of multiple scattering thanks to the air-core structure.

The individual components of the muon spectrometer can be seen in Fig. 3.8. Between and above the toroids, precision-tracking chambers are placed. In the barrel region, they are organized in three concentric layers (stations), each one containing 16 chambers (the 16-fold azimuthal symmetry reflects the 8-fold structure of the magnets), at distances of 5, 7.5 and 10 m from the beam axis, while in the end-caps they consist of four wheels (the small wheels located at $|z| = 7.4$ and 10.8 m, the big wheel at $|z| = 14$ m, the outer wheel at $|z| = 21.5$ m).

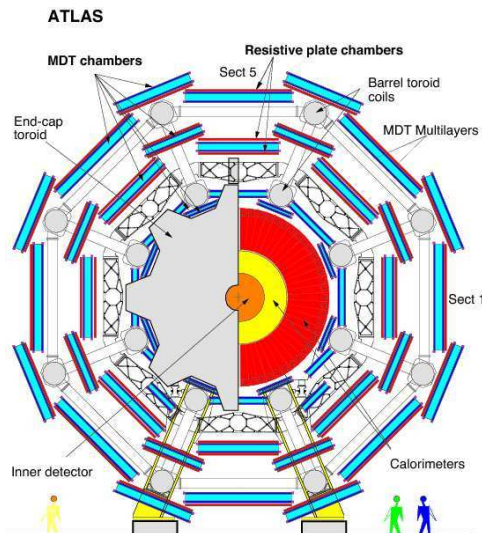


Figure 3.8: Frontal view of the ATLAS muon spectrometer, showing the arrangement of the various components: RPCs, MDTs and toroids [25].

Momenta are measured by Monitored Drift Tube chambers (MDT) up to $|\eta| < 2.7$,

except for the innermost layer in the forward region, where the counting rates and radiation doses are outside the MDTs operational constraints, so they are substituted by high-granularity Cathode Strip Chambers (CSC) ($2.0 < |\eta| < 2.7$). MDTs are pressurized drift tubes, filled with a mixture of 97% Ar and 3% CO₂, with 30 mm diameter and a single-wire resolution of $\sim 80 \mu\text{m}$. CSCs are multi-wire proportional chambers, with good tracking capabilities and fast response. The coordinate of the hit is measured through the charge induced on the cathode by the avalanche formed on the anode wire. A small neutron sensitivity is achieved by the small gas volume used and the absence of hydrogen.

The outermost region is covered by dedicated trigger chambers, consisting of Resistive Plate Chambers (RPCs) in the barrel and Thin Gap Chambers (TGC) in the end-caps, covering the region $|\eta| < 2.4$. Besides triggering, they provide a measurement of the track multiplicity and a first momentum estimation, together with bunch crossing identification. RPCs comprise two parallel resistive plates, with a relative distance of 2 mm, filled with gas, which is ionized by incoming muons, producing primary electrons that get multiplied and form an avalanche due to the high electric field at the wire. Like MDTs, TGCs are multi-wire proportional chambers, with the difference that the anode wire pitch is larger than the cathode-anode distance; they are used to trigger and to determine the azimuthal coordinate of the tracks, complementing the MDT radial information.

3.2.4 Trigger system

Given the extremely high interaction rate at the LHC, one major challenge for the ATLAS experiment is the online selection of interesting events. Starting from an interaction rate of $\sim 10^9$ Hz at a luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, corresponding to a bunch crossing of 40 MHz, the rate of selected events must be reduced to 200–300 Hz for permanent storage. This is achieved by the trigger system, organized in three levels [160], as shown in Fig. 3.9.

The first level, the level-1 (LVL1) trigger, which is hardware-based, performs an initial selection using only the information from a reduced subset of detectors. High- p_T muons are identified using only their dedicated trigger chambers (see Section 3.2.3), while a coarse calorimeter granularity is used to tag high- p_T electrons and photons, jets, τ -leptons, missing and total transverse energies. Isolation cuts may be applied, depending on the object, and generally trigger information is provided for a number of p_T thresholds. LVL1 defines the so-called regions of interest (ROIs), providing a measurement of η , ϕ and p_T for the selected objects, besides uniquely identifying the bunch crossing.

The second level, the level-2 (LVL2) trigger, which is software-based, focuses

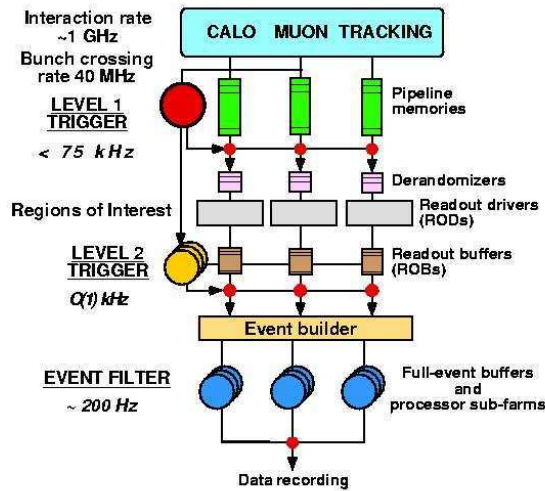


Figure 3.9: Scheme of the ATLAS trigger system [161]. Level-1, -2 and EF trigger stages are shown, with an indication of the characteristic rates.

on the ROIs identified by the LVL1. This reduces the trigger rate from ~ 75 kHz (upgradable to 100 kHz) entering LVL2 to the kHz range. LVL2 refines the LVL1 selection, adding p_T requirements and isolation cuts; the full detector granularity is exploited.

The final selection is performed by the Event Filter (EF), which applies up-to-date calibrations, improved alignment and magnetic field maps by means of offline algorithms, and selects the events which are finally written to mass storage for offline analysis. The EF output rate is around 200 Hz, corresponding to 200 Mb/s. The electron and muon triggers specific to the $H \rightarrow WW^{(*)}$ analysis are discussed in Sections 3.3.3 and 3.3.4.

3.3 Performance

3.3.1 Track reconstruction

The detection of tracks is the starting point for reconstruction and identification of charged particles [162]. It relies on the inner detector, whose coverage extends up to $|\eta| < 2.5$ (see Section 3.2.1). Two algorithms are used: the *inside-out* algorithm, which goes from the silicon detectors to the TRT, and the *back-tracking* algorithm, in the opposite direction.

The inside-out algorithm is the main algorithm for tracking primary particles, i.e. those directly produced in p-p interactions. It starts from 3-point seeds in the pixel and SCT detectors and adds hits moving away from the IP, using a combina-

torial Kalman filter. The reconstructed track is extended into the TRT, provided that its transverse momentum is $p_T > 400$ MeV. Finally, a fit using the whole ID information is performed. Recovery of secondary tracks from conversions, material interactions and long-lived particle decays is achieved, on the other hand, through the back-tracking algorithm, which begins with TRT track segments and extends them inwards to search for silicon hits.

Both tracking algorithms have been designed to minimize the impact of pile-up. In fact, the increase of per-event detector occupancy⁽⁵⁾ can result in resolution and efficiency degradation, together with an increase in fake tracks from random hit combinations. The impact of pile-up is kept under control by means of robust requirements on the track quality, namely a reconstructed track needs to have at least nine hits in the silicon detectors and exactly zero holes⁽⁶⁾ in the pixels (for the inside-out algorithm, similar requirements exist also for the back-tracking one).

3.3.2 Vertex reconstruction

The reconstruction of vertices happens in two steps [163]: first, the vertex finding algorithm associates reconstructed tracks to the vertex candidates, then the vertex *adaptive* fitting algorithm [164] computes the exact position of the vertex and refits the tracks so that they originate from it. More in detail, preselection criteria are applied to the reconstructed tracks originating from the interaction region, in order to discard secondary particles; from these, a vertex seed is found, corresponding to the global maximum of their z distribution, and tracks are associated with it by means of a χ^2 algorithm (tracks incompatible with the vertex by more than 7σ are used to seed a new vertex). The iterative process goes on until no unassociated tracks are left in the event. Out of the many reconstructed vertices, each event comprises one primary vertex, which is the supposed vertex of the hard interaction process and has the largest $\sum p_T^2$ of the associated tracks.

Pile-up increases not only the number of near-by vertices to be reconstructed (see Fig. 3.10 (a)), but also the number of fake tracks. The selection discussed above reduces the amount of non-primary tracks (secondary and fake tracks) in primary vertices by a factor of 2 – 5, while decreasing vertex reconstruction efficiency by no more than 5%. It also improves the stability of track and vertex efficiencies against pile-up up to $\langle\mu\rangle = 40$.

The estimated vertex resolution in the transverse plane can be seen in Fig. 3.10 (b); for vertices with 70 associated tracks, it amounts to $\sigma(x) \sim 23 \mu\text{m}$ in 2012 data

⁵Occupancy is defined as the fraction of detector channels with a hit in a local area. Occupancy increases in the ID going from the pixel detector (high granularity thus low occupancy), to the SCT strips, to the TRT (smaller granularity, thus higher occupancy).

⁶A hole is a non-existing but expected hit given a certain trajectory.

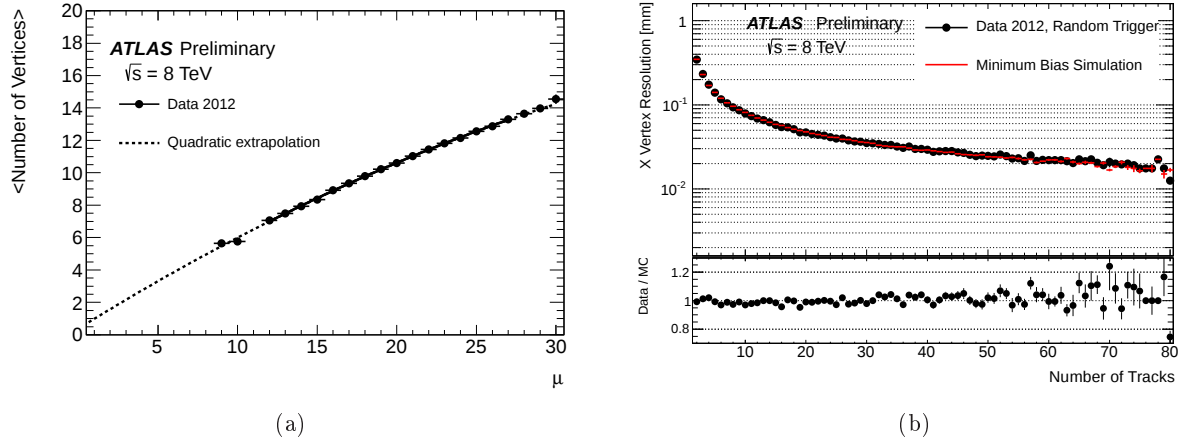


Figure 3.10: (a) Average number of reconstructed primary vertices as a function of the average number of p-p interactions per bunch crossing, measured in 2012 data. (b) Vertex position resolution in the transverse plane, for data (black) and MC (blue), as function of the number of tracks in the vertex fit. Both figures are taken from Ref. [165].

and $\sigma(x) \sim 30 \mu\text{m}$ in 2011 ($\sigma(z) \sim 40 \mu\text{m}$ and $\sigma(z) \sim 50 \mu\text{m}$ along the beam axis, respectively) [165].

3.3.3 Electrons

Electron reconstruction exploits information from both the ID and the EMCal. In order to distinguish real electrons, resulting from interesting physics processes, from fakes (hadronic jets misidentified as electrons) or photon conversions, cuts on track quality, isolation and transverse momentum are applied. In the following, the individual steps for reconstructing and identifying electrons are briefly described, with a focus on the requirements specific to the $H \rightarrow WW^{(*)}$ analysis.

Trigger

As described in Section 3.2.4 for the general case, the ATLAS trigger system comprises three levels [160], going from higher to smaller rates, and from coarser to finer resolutions. For electrons, in particular, the LVL1 trigger uses the so-called calorimeter trigger towers, shown in Fig. 3.7 (a), with sizes of $\Delta\Phi \times \Delta\eta = 0.1 \times 0.1$. A cut is applied on the transverse energy of the region (2×2 towers) centred around the electromagnetic cluster, with a threshold specific to the chosen trigger. Afterwards, in LVL2, tracks are matched to clusters, both reconstructed with dedicated algorithms, which are similar to the offline ones but satisfying looser criteria (for cluster formation, e.g., only the second EMCal layer is used).

Three standard selections are defined, with increasing background rejection power: *loose*, *medium* and *tight*. If tighter cuts on cluster and track variables are applied, they are named *loose1*, *medium1* and *tight1*, respectively. The label *vh*, on the other hand, signals that two additional features are employed: an η -dependent optimization on energy thresholds and a cut on hadronic leakage. Finally, the label *i* indicates that isolation criteria are used during online event selection (to be reinforced offline).

The $H \rightarrow WW^{(*)}$ analysis of $\sqrt{s} = 7$ TeV data, presented in Chapter 4, makes use of the following single-electron triggers: *EF_e20_medium*, *EF_e22_medium* and *EF_e22vh_medium1*. The number after *EF_e* refers to the p_T threshold, which is raised from 20 GeV to 22 GeV to cope with increasing instantaneous luminosity in 2011 data, together with the electron quality, going from *medium* to *vh_medium1*.

In 2012, requirements have been tightened further: the nominal analysis uses *EF_e24vhi_medium* and *EF_e60_medium*, taking an “or” of the two to recover for losses in efficiency of high- p_T electrons. For *EF_e24vhi_medium*, relative track isolation is required to be $\sum p_T/p_T(e) < 0.1$.

Reconstruction

In order for an object to be reconstructed as an electron, energy deposits (clusters) in the central region of the EMCal ($|\eta| < 2.47$) have to be associated with tracks from the ID.

Clusters are first searched with an iterative procedure, the *sliding-window* algorithm [166], which looks for energy deposits in towers of 3×5 cells from the middle layer of the calorimeter (where each cell has a size of 0.025×0.025 in units of $\Delta\Phi \times \Delta\eta$, as described in Section 3.2.2). If the total transverse energy inside the window exceeds 2.5 GeV, the cluster is considered an electron or photon candidate. MC simulations show that, for $E_T > 15$ GeV, such an algorithm has an efficiency close to 100%. In addition, electron candidates from the EMCal need to have an associated track in the ID. Once matched, the cluster is rebuilt using a larger window (3×7 cells in the barrel, 5×5 cells in the end-caps). The cluster energy is then determined by summing to the energy deposit of the cluster itself three more contributions: the estimated energy deposits in the material in front of and beyond the EMCal, and the estimated lateral leakage outside the cluster. The ϕ and η coordinates are taken instead from the corresponding track parameters at the vertex.

Outside the central region, $2.5 < |\eta| < 4.9$, only energy deposits in the FCal are available, due to the absence of tracking detectors. For this reason, the reconstruction of forward electrons is based only on clusters. Forward electrons are not used in the $H \rightarrow WW^{(*)}$ analysis and are therefore not discussed further.

A critical issue for electron reconstruction is represented by radiative energy

losses when traversing the substantial amount of material in the ID, which can give rise to deviations from the original path, and hence to significant inefficiencies. These issues have been addressed in 2012 through the implementation of an improved track refitting procedure, the Gaussian Sum Filter (GSF) algorithm [167], a non-linear generalization of the Kalman filter technique.

Identification

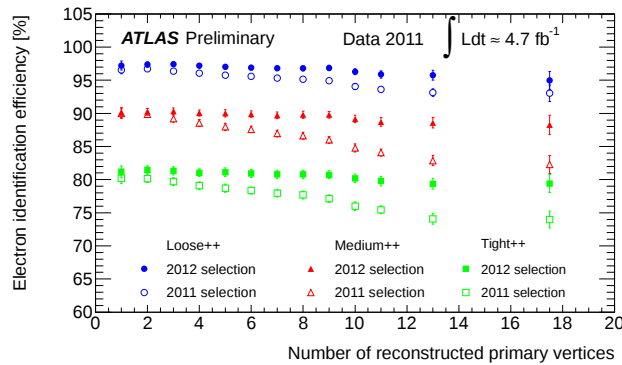


Figure 3.11: Electron identification efficiency, as a function of the number of primary vertices in the event, measured in 2011 data applying 2011 and re-optimised 2012 selections [168].

Electron identification relies on a cut-based method combining ID information with shower shape characteristics. Real isolated electrons have to be distinguished from:

- background electrons originating from photon conversions, Dalitz decays or light mesons decays,
- non isolated electrons which are part of a b- or c-quark jet,
- fake electrons which are hadronic jets misidentified as electrons.

Three sets of cuts are used in ATLAS to identify electrons: *loose*, *medium* and *tight* [169], [166]. The *loose* selection uses only shower shapes characteristics from the EM second layer and hadronic leakage variables, the *medium* selection adds to these the information coming from the EM strip layer, track quality requirements and track-cluster matching, while the *tight* selection exploits the full information available in the detector, such as E/p requirements, particle identification in the TRT and discrimination against photon conversions via a B-layer hit requirement. The corresponding cuts have been optimized in bins of η and cluster E_T , according to pile-up conditions.

The background rejection power increases going from loose, to medium, to tight, from 500, to 5000, to 50000, respectively, based on MC simulations. The signal efficiency decreases accordingly, between $\sim 97\%$ for loose and $\sim 80\%$ for tight.

To cope with increasing pile-up in 2012 data, pile-up sensitive criteria have been relaxed, such as the fraction of energy leaking in the hadronic calorimeter, while pile-up robust ones have been tightened. In this way, the efficiency dependence on pile-up is largely mitigated, but rejection power is not degraded. The improved categories are named *loose++*, *medium++* and *tight++*, the latter used by the $H \rightarrow WW^{(*)}$ analysis, both in 2011 and 2012. As can be seen in Fig. 3.11, applying 2012 selection to 2011 data, an identification efficiency above 95%, 88% and 79% is achieved for the three categories, for $\langle\mu\rangle$ up to 20 [168].

Isolation

Isolation is a fundamental requirement for distinguishing real electrons from fakes and strongly depends on the properties of the final state under consideration, therefore it is analysis-dependent and no default isolation criteria can be included in the standard set of identification cuts discussed above. Two definitions of isolation variables can be used, track-based and calorimeter-based. Track-based isolation pertains to charged particles only; one commonly used variable is the scalar sum of transverse momenta of the tracks in a cone of radius ΔR around the electron candidate (tracks $< \Delta R$):

$$p_T^{\text{cone}} = \sum_{\text{tracks} < \Delta R} p_T. \quad (3.4)$$

Only tracks are considered, that satisfy specific quality requirements, excluding the electron candidate itself. A derived variable is the relative isolation, p_T^{cone}/p_T .

Calorimeter isolation, E_T^{cone} , is sensitive to both charged and neutral particles, which deposit their energy in cells around the electron cluster. Also in this case, the sum is done inside a cone of radius ΔR , from which the electron candidate itself is removed. Corrections to this quantity are pile-up dependent, since the presence of candidates from subsequent events in neighbouring cells may bias the computed isolation.

The typical value for the cone radius, both for track and calorimeter isolation, is $\Delta R = 0.3$, which is a compromise between robustness against pile-up and leakage of the electron shower into the isolation cone – this is indeed the value used in the $H \rightarrow WW^{(*)}$ analysis.

3.3.4 Muons

Unlike electrons, muons do not lose most of their energy when traversing the electromagnetic calorimeter. They are detected in the muon spectrometer, which has also dedicated trigger chambers, as explained in Section 3.2.3.

Trigger

The LVL1 muon trigger uses dedicated trigger chambers, RPCs in the barrel ($|\eta| < 1.05$) and TGCs in the end-caps ($1.05 < |\eta| < 2.4$), to search for a sequence of hits compatible with a muon track, and to roughly compute its η and ϕ coordinates, defining a corresponding region of interest (ROI) [160]. High- p_T muons are expected to reach the third RPC layer, while low- p_T muons do not. At the following stage (LVL2), the full MS granularity within the ROIs is exploited to refine the p_T measurement. This new estimation can be combined with the one coming from the ID, using a fast track combination algorithm. Offline, at the EF stage, the muon candidates are accepted or discarded; two main methods are used for this purpose, an inside-out method going from the ID track to the MS and an outside-in algorithm in the opposite direction, similar to those described for track reconstruction in Section 3.3.1.

In 2011, the $H \rightarrow WW^{(*)}$ analysis uses two single-muon triggers: *EF_mu18_MG* (up to period I) and *EF_mu18_MG_medium*, both with a 18 GeV p_T threshold. The suffix *medium* indicates that the corresponding L1 trigger has been tightened from *L1_MU10* to *L1_MU11*, in order to cope with increasing instantaneous luminosity.

In 2012, the trigger menu moves to *EF_mu24i_tight* and *EF_mu36_tight*, where the suffix *i* indicates that isolation has been applied during online event selection, and *tight* that LVL1 threshold has been increased to *L1_MU15*.

Reconstruction and identification

ATLAS analyses use three main strategies to reconstruct muons: *stand-alone*, *combined* and *segmented tagged* muons [170].

Stand-alone muons rely only on information coming from the MS; the hits left in the three stations are combined to form segments, which are then linked together to form a track. Once the track is reconstructed, it is extrapolated back to the IP, correcting for energy losses in the calorimeters. This strategy offers the widest coverage (up to $|\eta| < 2.7$), with an efficiency close to 100%, removing the inefficient regions around $\eta = 0$ and $|\eta| \sim 1.2$.

Combined muons use both MS and ID tracks, matched by means of a χ^2 test. In this case, the coverage is limited by the ID one ($|\eta| < 2.5$) but momenta have higher

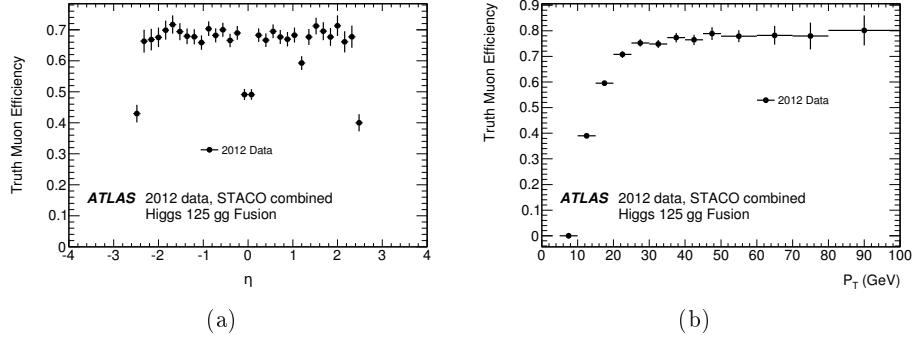


Figure 3.12: Muon reconstruction efficiencies for truth muons from a MC sample of $gg \rightarrow H(m_H = 125) \rightarrow WW^{(*)}$ events, as a function of η (a) and p_T (b). The efficiency is around 70% on the plateau. Both figures are taken from Ref. [173].

resolution due to the combination of the two systems. η and ϕ coordinates for the track are provided by the ID. The efficiency of this strategy has been measured in $Z \rightarrow \mu^+ \mu^-$ events in data to be $> 94\%$, again ignoring the inefficient regions.

Finally, segmented tagged muons are reconstructed by propagating ID tracks above a certain momentum threshold to the first MS station and searching for nearby segments; since it uses the first MS station only, this strategy is best suited to reconstruct low- p_T muons.

Two algorithm chains exist, that contain all these three strategies, STACO [171] and MUID [172]. The baseline method for physics analyses in ATLAS is currently the STACO combined one. Additional requirements are imposed in the $H \rightarrow WW^{(*)}$ analysis: in 2011 data, tracks should have at least two hits in the pixel detector (one of them in the first layer) and at least six hits in the SCT, no more than two holes and no excessive amount of outlier hits in the TRT. Cuts are placed also on the longitudinal and transverse impact parameters, $|z_0| < 1$ mm and $|d_0/\text{sig}(d_0)| < 3$, respectively, where all positions are computed with respect to the primary vertex. These requirements have been loosened in 2012 data, to recover some small efficiency losses.

Fig. 3.12 shows the reconstruction efficiency for STACO combined muons in a MC Higgs boson signal sample, in bins of (a) η and (b) p_T , after all selection cuts have been applied, including the isolation ones discussed hereafter. The average efficiency is around 70%.

Isolation

As discussed for electrons in Section 3.3.3, also for muons isolation criteria are based on the fact that muons from signal processes are generally not accompanied by

nearby particles, unlike those from heavy-quark or hadron decays.

Track-based isolation exploits the scalar sum of transverse momenta of tracks in a ΔR cone around the muon candidate, excluding the muon itself, $p_T^{\text{cone}} = \sum_{\text{tracks} < \Delta R} p_T$, and the corresponding relative isolation, p_T^{cone}/p_T . Calorimeter isolation sums on cells in a ΔR cone around the muon cluster, removing the muon itself, defining E_T^{cone} and the relative quantity E_T^{cone}/p_T .

The $H \rightarrow WW^{(*)}$ analysis cuts on both relative tracking and calorimeter isolations, namely $p_T^{\text{cone}}/p_T < 0.13$ and $E_T^{\text{cone}}/p_T < 0.14$. To account for data/MC differences arising from these cuts and for pile-up corrections in the calorimeter, isolation scale factors are applied, which are consistent with unity within 1%.

3.3.5 Jets

The most important ATLAS subdetector for jet reconstruction is the calorimeter. Jets are observed as topologically connected energy deposits, associated with tracks of charged particles in the ID. Real jets have to be distinguished from fakes, which do not come from collision events, i.e. beam-gas events (where one proton of the beam collides with the residual gas within the beam pipe), beam-halo events, cosmic ray muons overlapping in-time with collision events and finally calorimeter noise.

Jet reconstruction is based on *topological clusters* [174], which are 3D structures representing the shower evolution of a particle entering the calorimeter. The seeds are cells with a signal-to-background ratio $S/B > 4$, to which all the directly neighbouring cells are added if $S/B > 2$, and finally a ring of guard cells is merged into the cluster regardless of any threshold.

Clusters are then formed into jets using dedicated algorithms [175,176]. A good jet clustering algorithm needs to be both collinear- and infrared-safe, i.e. the presence of collinear or soft radiation close to the hard particle should not influence the jet reconstruction; moreover, it is desirable to have jets with regular shape, to facilitate their calibration, but not too soft-resilient, in order to be able to accommodate successive QCD branchings. Three algorithms satisfying these requirements are the k_t [177], anti- k_t [178] and Cambridge-Aachen [179] ones. The discriminating criterium is, for all of them, the distance d_{ij} between two objects i and j (particles or pseudoclusters), defined as:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \quad (3.5)$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the distance in the $y - \phi$ coordinates and k_t is the transverse momentum. The choice $p = 0$ corresponds to the Cambridge-Aachen algorithm, which does not depend on the particle momentum, $p = 1$ to the

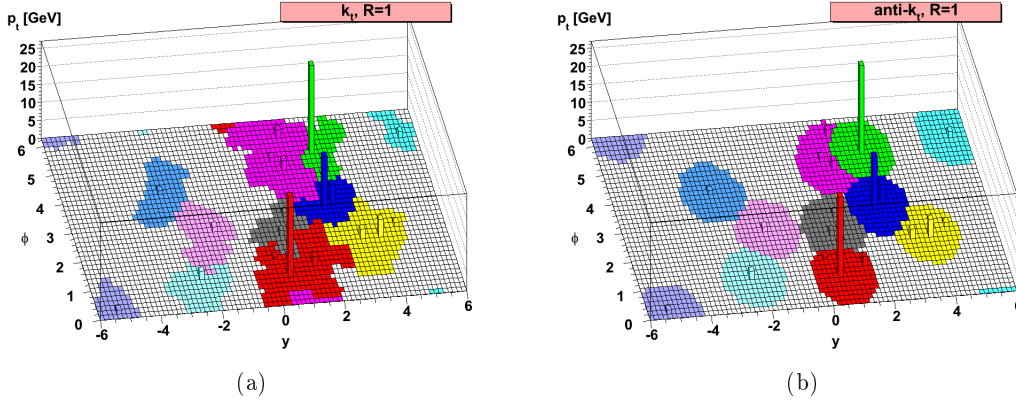


Figure 3.13: Parton level event generated with Herwig, together with random soft radiation, and clustered with the k_t (a) or anti- k_t (b) algorithms. Shown is the structure and the (ϕ, y) location of the resulting jets [178].

k_t algorithm (which shares its properties with all the $p > 0$ cases) while $p = -1$ corresponds to the anti- k_t algorithm, which is the one used by the $H \rightarrow WW^{(*)}$ analysis.

Its behaviour can be understood considering an event with only one hard particle, with momentum k_{t1} , and many soft particles. The d_{1i} distance between 1 and a soft particle i is determined by the $1/k_{t1}$ term, since it is $1/k_{t1} < 1/k_{ti}$. For the same reason, $d_{1i} < d_{iB}$, i.e. the soft particle is closer to the hard one than to the beam B , thus the anti- k_t algorithm will merge all the soft particles i , within a cone of radius R , to the hard particle 1. If there is a second hard particle in the event, with a relative distance $R < \Delta_{12} < 2R$, there will be two hard jets, but at least one will not be perfectly conical; in particular, if $k_{t1} \gg k_{t2}$, jet 1 will be conical, while jet 2 will be deformed, since it will miss the overlapping region. Such a case is illustrated in Fig. 3.13 by the two jets around $(\phi, y) = (5, 2)$. The figure also shows that the shapes produced by the anti- k_t algorithm (b) are more regular than those produced by, e.g., the k_t one (a).

The $H \rightarrow WW^{(*)}$ analysis uses anti- k_t jets, as noticed above; they are referred to as “anti- k_t EM+JES” jets, with $\Delta R = 0.4$. The label “EM+JES” refers to the energy scale. Jets are reconstructed at the electromagnetic (EM) scale, which is the signal scale of the calorimeters, established using test-beam measurements and validated using the invariant mass of $Z \rightarrow e^+e^-$ collision events. In order to calibrate jet energies, the EM scale has to be translated into the hadronic one. This is done via data-driven corrections and calibration constants derived from truth jets in MC. Corrections due to pile-up and to non-compensation of the calorimeter are

also applied. The uncertainties on the jet energy scale (JES) [180] and resolution (JER) [181] are among the most important systematic uncertainties for many physics analyses, as it is the case for $H \rightarrow WW^{(*)}$.

3.3.6 b -Jets

Jets originating from b -quarks are tagged by specific algorithms, that can be divided into two main categories: vertex-based and impact parameter-based. All rely on the observation that B-hadrons possess a relatively long lifetime ($\tau \sim 1.5$ ps), and hence are able to travel a significant path in the detector before decaying, on the order of millimeters in the transverse plane. Therefore, the identification of primary and secondary vertices, and the measurement of their distances with respect to the interaction point, are fundamental in order to identify b -jets.

Two parameters are used to assess the performance of a b -tagging algorithm: the tagging efficiency, in presence of real b -jets, and the mistag efficiency, for light jets mistaken as b -jets.

The vertex-based algorithms [182] aim to reconstruct the whole decay chain of b -quarks: starting from the tracks that constitute a selected jet (satisfying the criteria $\Delta R < 0.4$, $p_T > 20$ GeV, $|\eta| < 2.5$), an iterative search is performed, building all two-track pairs that form a good vertex, far enough from the IP. The selected two-track vertices are then combined into a single inclusive vertex, using various quality criteria. In particular, the *JetFitter* algorithm [183] employs a Kalman filter to find a common path on which the primary and secondary vertices lie, and compute their positions.

The impact parameter-based algorithms, on the other hand, combine the impact parameter significances of all the tracks in the jet, with various approaches: the IP3D algorithm, in particular, uses a multivariate method (the likelihood ratio technique) to compare a set of input variables (transverse and longitudinal impact parameters and their significances) to templates for b - and light jets, obtained from MC simulations.

While the vertex-based algorithms are affected by inefficiencies in secondary vertex finding, the impact parameter-based ones have a higher mistag rate, so that combining the two classes offers many advantages. The combination of IP3D and JetFitter is called *JetFitterCombNN* [182] and uses a neural network to fully exploit the correlation among input variables. The $H \rightarrow WW^{(*)}$ analysis uses this latter algorithm, tuned to achieve a b -jet identification efficiency of 80% in top-quark pair events, together with a light-jet mistag rate around 6% [182]. Fig. 3.14 (a) compares the performance of the JetFitterCombNN algorithm with a few others, in terms of light jet rejection versus b -jet efficiency, using simulated top-quark pair events at

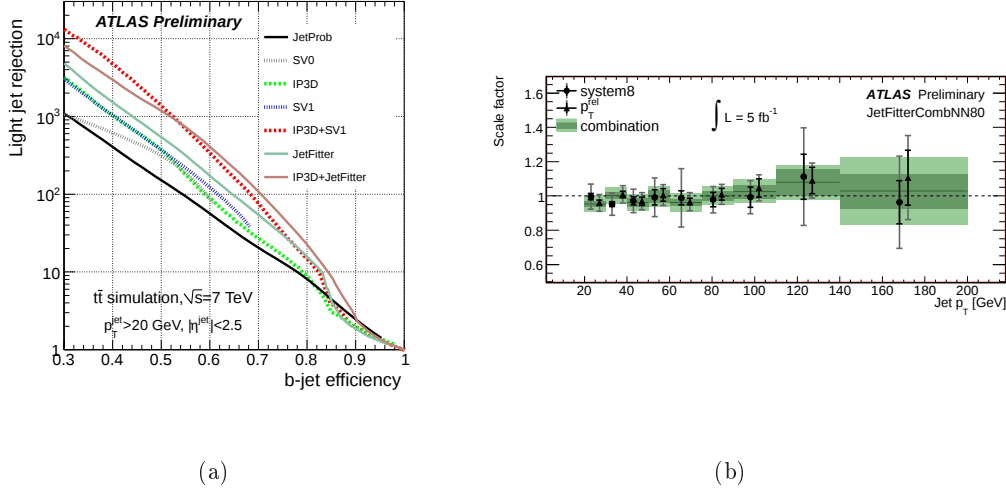


Figure 3.14: (a) Light-jet rejection as a function of the b-jet tagging efficiency for various tagging algorithms, based on top-quark pair MC events [182]. (b) Scale factors for the JetFitterCombNN algorithm at 80% efficiency. The dark green and light green bands represent, respectively, the statistical and total uncertainties of the combined scale factors [184].

$\sqrt{s} = 7$ TeV.

The b -tag scale factors, provided in bins of jet- p_T , are illustrated in Fig. 3.14 (b); they are consistent with unity within 10%.

3.3.7 Missing transverse energy

An excellent measurement of the missing transverse energy, E_T^{miss} , is essential for many physics analyses, such as supersymmetry searches, and in particular for the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channel, where the two charged leptons are accompanied by two neutrinos, that escape detection. Due to momentum conservation in the transverse plane, the E_T^{miss} vector must balance the sum of transverse momenta carried by visible particles, provided that fake sources of E_T^{miss} have been identified and corrected: limited detector coverage, presence of dead regions and noisy read-out channels. All of these effects, together with the finite detector resolution, limit the E_T^{miss} resolution. Several algorithms have been developed to suppress the effect of pile-up interactions [185].

E_T^{miss} reconstruction is based on energy depositions in calorimeter cells, after the corresponding calibration has been carried out (at the EM scale for electromagnetic objects, at the hadronic scale for jets), on low- p_T tracks and on MS-matched tracks for muons. In summary, the following terms are taken into account when including

cells in the $E_{x,y}^{\text{miss}}$ computation, in this order:

$$E_{x,y}^{\text{miss}} = E_{x,y}^e + E_{x,y}^\tau + E_{x,y}^{\text{b-jets}} + E_{x,y}^{\text{jets}} + E_{x,y}^\mu + E_{x,y}^{\text{Cell Out}} \quad (3.6)$$

where the $E_{x,y}^{\text{Cell Out}}$ term refers to cells with no associated object (here, also soft jets, i.e. with $7 \text{ GeV} < p_T < 20 \text{ GeV}$, are included). Coverage in pseudorapidity is $|\eta| < 4.9$ for all objects apart from muons, whose tracks are searched inside $|\eta| < 2.7$. From the x and y components, the missing E_T value is computed as:

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

The quality of the E_T^{miss} reconstruction, i.e. its linearity and resolution, is assessed through samples of leptonically decaying W and Z bosons, without accompanying jets. The former provides a clean signature for linearity studies, since no hadronic contribution is expected, while the latter should in principle contain no real E_T^{miss} , thus is used to signal biases in the reconstruction algorithm [185, 186].

Systematic uncertainties on E_T^{miss} depend on the topology of the final state under consideration, since they are evaluated propagating the systematic uncertainties of the visible objects present in the event, such as leptons, hard and soft jets.

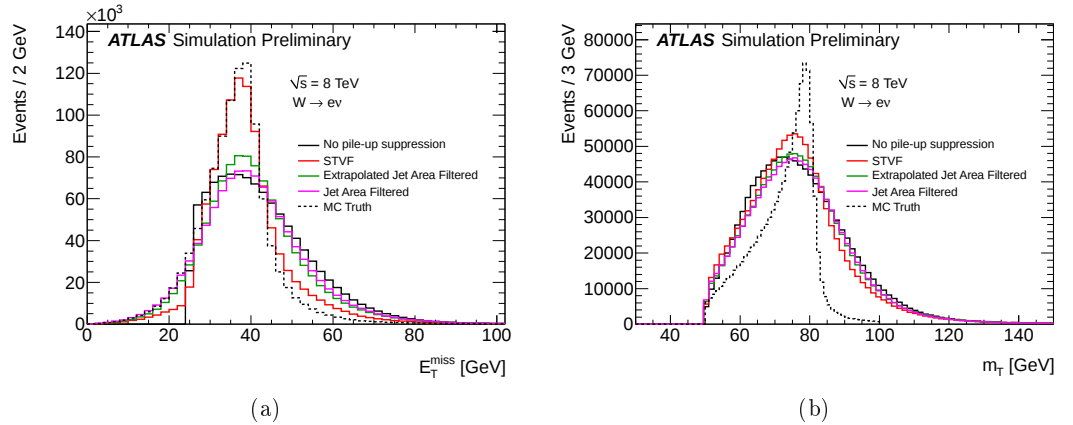


Figure 3.15: Distributions of E_T^{miss} (a) and m_T (b) (defined in Eq. 4.1) before and after pile-up suppression with different algorithms in MC $W \rightarrow e\nu$ events. The true distributions are also shown for comparison [186].

Fig. 3.15 shows the improvement in E_T^{miss} (a) and m_T (b) (which is the transverse mass of the W system, defined in Eq. 4.1), before and after pile-up suppression with different algorithms, in MC $W \rightarrow e\nu$ events. The improvement in E_T^{miss} is

substantial, especially using the STVF algorithm⁽⁷⁾, while it is more moderate in m_T , due to the fact that the transverse mass computation depends both on the modulus and direction of the missing transverse energy vector.

⁷The Soft Terms Vertex Fraction (STVF) algorithm is based on the fraction of soft tracks associated to the primary vertex with respect to the total. It is analogous to the JVF algorithm described in Section 4.4.6, but considers only the terms which do not satisfy jet requirements.

Chapter 4

Search for the $H \rightarrow WW^{(*)}$ Decay Using a Multivariate Analysis

In this Chapter a multivariate analysis of the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channel with the $\sqrt{s} = 7$ TeV dataset, collected in 2011 with the ATLAS detector, is presented. It uses boosted decision trees, which are meant to classify the *candidate* events⁽¹⁾ as signal-like or background-like, through a series of binary decisions. They are in this respect similar to rectangular cuts; however, whereas a standard cut-based analysis is able to select only one hypercube out of the phase space, decision trees are able to split the events into a larger number of hypercubes, allowing for a better discrimination between the signal and the background. They are expected to take advantage of the correlations among the variables, which may otherwise be broken when applying rectangular cuts.

Sections 4.1 and 4.2 discuss the topologies of the different signal and background components, respectively, while Section 4.3 summarizes the Monte Carlo samples used to simulate them. Sections 4.4 and 4.5 present the object definition and the event preselection, which are largely in common with those of the standard cut-based analysis. In Section 4.6 the method and the usage of the boosted decision tree technique are discussed, with particular emphasis on the choice of the settings and of the input variables. Section 4.7 shows the resulting BDT responses in the signal and control regions, in the different jet bins; the backgrounds are further discussed in Sections 4.8 and 4.9, the latter concentrating on the systematic uncertainties. Section 4.10 outlines the statistical method and the fit procedure used in this analysis, and finally Section 4.11 contains the results for the full 2011 dataset, in terms of CL_s exclusion limits, fitted signal strengths and local p_0 -values, as a function of the

¹A candidate event is defined as an event satisfying a specific set of cuts, most often the full selection criteria.

Higgs boson mass. The results are compared with those from the standard cut-based analysis.

Additional details and cross-checks are given in the Appendices. In Appendix A, the boosted decision tree is compared to other multivariate techniques, and in particular with the Neural Network one, which has been developed in the context of this thesis. Appendix B motivates the choice of the BDT input variables, whose distributions are shown for the individual lepton flavours in Appendix C. Appendices D and E show the performance of the BDT analysis when trained with a reduced set of input variables, and when trained with the WW background as signal, respectively. Appendices F and G focus on the correlation between the BDT and the cut-based analyses: the former presents the topology of the candidate events in the signal-like and in the background-like regions, defined through a cut on the BDT response, the latter shows the BDT score for the cut-based candidates. Finally, Appendices H and I give further details on the fit results.

The details specific to the cut-based analysis can be found in Ref. [28] and are not repeated here. This analysis has been published as an ATLAS conference note [27].

The work documented in this thesis represents a leading contribution to the development and the finalization of the $H \rightarrow WW^{(*)}$ BDT analysis, in all its phases: the optimization of the BDT parameters and of the input variables, the definition of the selection for the signal and the control regions, the scrutiny of the output variables, the assessment of the systematic uncertainties, the validation of the limit-setting framework, the statistical results and their comparison with the cut-based ones. Moreover, all the studies presented in the Appendices pertain to this work and have been fundamental to validate the usage of the BDT algorithm, which had not been used before in the context of the ATLAS Higgs working group.

4.1 Signal topology and analysis categories

The search for the decay of a SM Higgs boson into a pair of W bosons is performed by the ATLAS collaboration in two main final states: purely leptonic ($\ell^+\nu\ell^-\bar{\nu}$, shortened as $\ell\nu\ell\nu$ in the following) and hadronic-leptonic ($\ell\nu qq$) [187]. Despite the larger branching ratio for a W boson to decay into quarks ($\text{BR}(W \rightarrow \ell\nu) = 10.8\%$, $\text{BR}(W \rightarrow \text{hadrons}) = 67.6\%$, [47]), and the possibility to fully reconstruct the Higgs boson mass in this channel, the leptonic final state is experimentally cleaner due to the presence of two isolated, high- p_T leptons, and can be exploited over the mass range $110 \text{ GeV} < m_H \leq 600 \text{ GeV}$.

As discussed in Section 2.1.1, the main production mechanism of a Higgs boson in the Standard Model is gluon-fusion, depicted in Fig. 4.1; in this case, jets can

arise in the final state due to QCD bremsstrahlung. The Higgs boson production via vector boson fusion, on the other hand, is characterized by two hard jets, which are the remnants of the colliding partons, as depicted in Fig. 4.2.

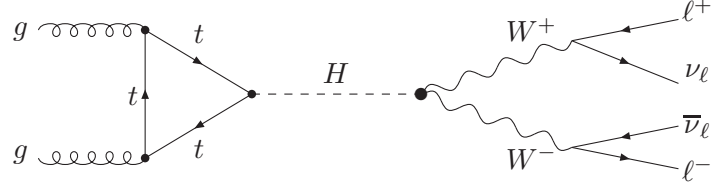


Figure 4.1: Feynman diagram at leading order for the production of a Higgs boson via gluon-fusion (top-quark loop) and its subsequent decay into a WW pair, where both W bosons decay leptonically.

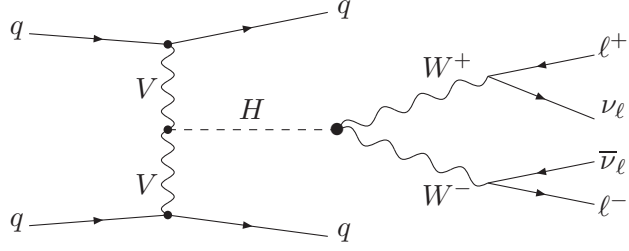


Figure 4.2: Feynman diagram at leading order for the production of a Higgs boson via vector boson fusion ($V = W, Z$) and its subsequent decay into a WW pair, where both W bosons decay leptonically.

The analysis is divided into nine categories, according to the number of jets in the final state ($H + 0, 1, \geq 2$ jets, denoted as 0-jet, 1-jet and 2-jet, respectively) and to the lepton flavours ($e^+\nu_e e^-\bar{\nu}_e$, $\mu^+\nu_\mu \mu^-\bar{\nu}_\mu$ and $\mu\nu_\mu e\nu_e$, denoted as ee , $\mu\mu$ and $e\mu$, respectively). The term “same flavour” refers to the ee and $\mu\mu$ channels, while “opposite flavour” to $e\mu$. The case in which e or μ arise from the decay of a τ -lepton is also taken into account.

The topology of this channel has been described in detail in Section 2.3; in short, due to the $J^P = 0^+$ nature of the SM Higgs boson, the two charged leptons are expected to be close to each other both in the azimuthal plane (small $\Delta\phi_{\ell\ell}$) and in the polar direction (small $\Delta\theta_{\ell\ell}$, hence small $m_{\ell\ell}$), which also induces a larger value for the transverse momentum of the dilepton pair (large $p_{T,\ell\ell}$). The two neutrinos tend to be emitted in the opposite direction with respect to the charged leptons, thus producing a sizeable missing transverse energy. The $\ell\nu\ell\nu$ final state cannot be fully reconstructed, since this would require 4×3 independent quantities, but only 2×3 are available from the charged leptons, plus two from the x, y components of

the missing transverse energy, thus four degrees of freedom are missing. For this reason, rather than the invariant mass of the system, the transverse mass is used; multiple definitions of this quantity exist, all characterized by a broad peak around the true Higgs boson mass, terminating in a narrow edge for $m_T > m_H$ and a longer tail for $m_T < m_H$, while backgrounds do not show such a peak. This analysis uses in particular the following definition [188]:

$$m_T = \sqrt{E_{T,\ell\ell+\nu\nu}^2 - |\vec{p}_{T,\ell\ell+\nu\nu}|^2} = \sqrt{(E_{T,\ell\ell} + E_T^{\text{miss}})^2 - |\vec{p}_{T,\ell\ell} + \vec{p}_T^{\text{miss}}|^2}. \quad (4.1)$$

The topology of the final state is contained into the four *topological* variables, $m_{\ell\ell}$, m_T , $p_{T,\ell\ell}$ and $\Delta\phi_{\ell\ell}$, which are used as inputs for the BDT training.

The properties discussed so far strictly apply to the case of a light Higgs boson only, since for masses above ≈ 200 GeV, the boost with which the W bosons are produced, partly breaks the angular correlations among the final leptons.

4.2 Background topologies

Under the category of *background*, all processes are contained, which possess the same topology as the signal, namely two isolated, high- p_T leptons and missing transverse energy, with or without jets. This signature can be either physical or induced by detector effects, such as object misidentification, limited coverage, bremsstrahlung or pile-up events. The most important background is the non-resonant W^+W^- production, which is said to be *irreducible* since it contains two W bosons alike the signal. Other backgrounds are $V + \text{jets}$ ($V = W, Z/\gamma^*$), top-quark production ($t\bar{t}$ and single top-quark channels), diboson production ($ZZ, WZ, W\gamma$) and multi-jet events. In the following, additional details are given on the background topologies, comparing the distributions of the topological variables with those of the signal.

4.2.1 The W^+W^- background

The continuous W^+W^- production, or WW in short, is mediated by quark-antiquark annihilation, depicted in Fig. 4.3 (a) for the s -channel and (b) for the t -channel, using the language of the Mandelstam invariants. The contribution from gluon-fusion production is subleading.

The ATLAS collaboration has measured the inclusive cross section for W^+W^- production to be 53.4 ± 2.1 (stat) ± 4.5 (sys) ± 2.1 (lumi) pb with an integrated luminosity of 4.7 fb^{-1} , at $\sqrt{s} = 7$ TeV, a value which is 10% higher than the corresponding NLO prediction [189]. Multiplying the inclusive cross section with the W leptonic branching ratio, a value of $q\bar{q}/g \rightarrow \sigma(W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}) = 4.68$ pb

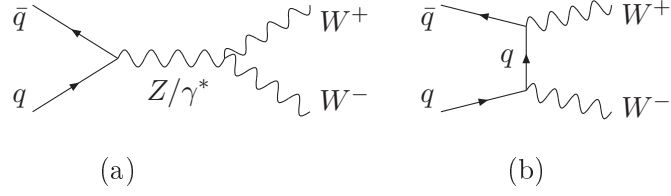


Figure 4.3: Feynman diagrams at leading order for the production of W^+W^- pairs initiated by a quark-antiquark pair: (a) production in the s -channel, (b) production in the t -channel.

is obtained, to be compared with 0.347 pb for the resonant Higgs boson production via gluon-fusion, assuming $m_H = 125$ GeV (see Table 4.1).

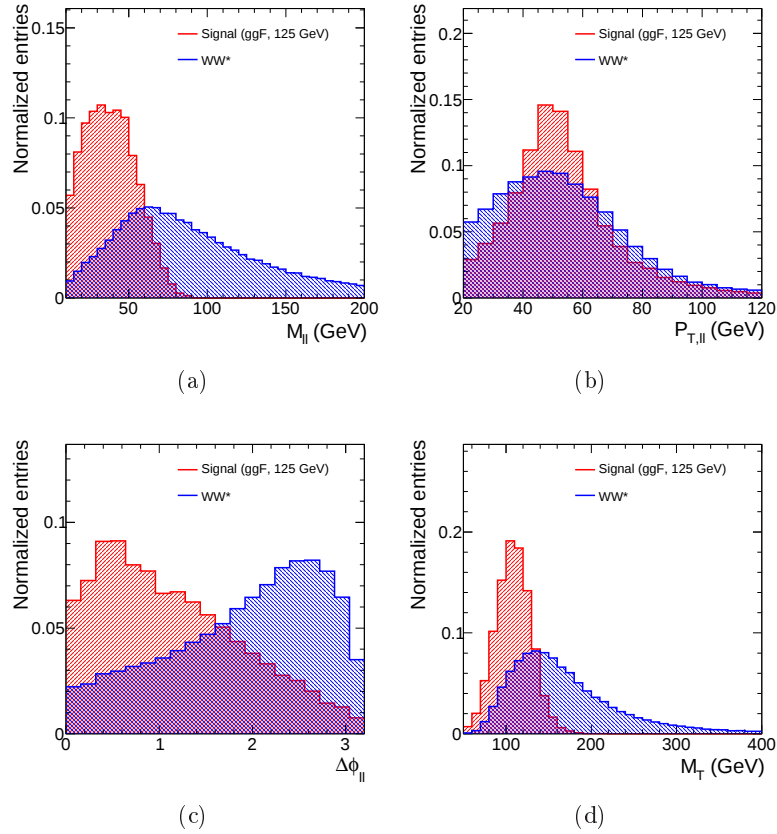


Figure 4.4: Normalized distributions after the preselection (see Section 4.5) comparing the signal shape (ggF only, in red) with that of the continuous WW background (in blue). The gluon-initiated processes are not considered, since subleading. Only the $e\mu$ final state is shown. (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T .

The WW system has access to three spin states (0,+1,-1) when produced by a quark-antiquark state, thus the (+1,-1) states modify the profile of the (0) case. This lack

of correlation causes, as illustrated in Fig. 4.4 for the $e\mu$ final state, both the $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$ distributions to be larger for the WW background than for the signal. For what concerns the transverse mass, its distribution extends up to 400 GeV and does not exhibit a Jacobian peak, unlike the signal.

4.2.2 The top-quark backgrounds

Subleading in the 0-jet category, the top-quark backgrounds are the most important processes when jets are present in the final state. As illustrated in Figs. 4.5 and 4.6, respectively, top-quarks can be produced either in pairs or singularly: in the first case, mostly through gluon-gluon interaction (annihilation or scattering), with the quark-antiquark initiated processes accounting for 10% of the total rate, in the second case with contributions from the t - and s -channels and in association with a W boson. The theoretical values for the cross sections are given in Table 4.1; this analysis uses 167 pb for $t\bar{t}$ and 85 pb for single top-quark production as reference values.

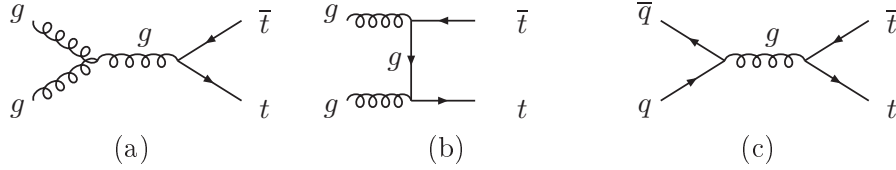


Figure 4.5: Feynman diagrams at leading order for the production of $t\bar{t}$ pairs through (a) gluon-gluon annihilation and (b) scattering, and through (c) quark-antiquark annihilation.

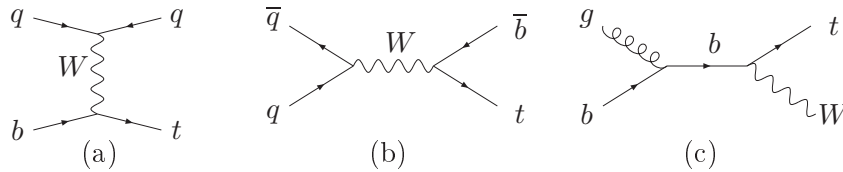


Figure 4.6: Feynman diagrams at leading order for the production of single top-quarks, in (a) the t -channel, (b) the s -channel and (c) in association with a W boson.

The production of $t\bar{t}$ pairs proceeds via the strong interaction, while the single top-quark production proceeds via the electroweak one; subsequently, the top-quark decays into a Wb pair (decays into lighter quarks are suppressed by the corresponding CKM elements), resulting in a final state with one or two W bosons, that can mimic the signal signature in case they decay leptonically. The jets originate from b -quark decays.

Fig. 4.7 compares the topological variables for the signal and the top-quark backgrounds, in the $e\mu$ final state. The main difference resides in the dilepton azimuthal angle, which tends to be larger and flatter for the top-quark backgrounds than for the signal; also the invariant mass is found to be larger, while the transverse mass does not show the characteristic Jacobian peak around the Higgs boson mass; this holds for both the top- and single top-quark production modes.

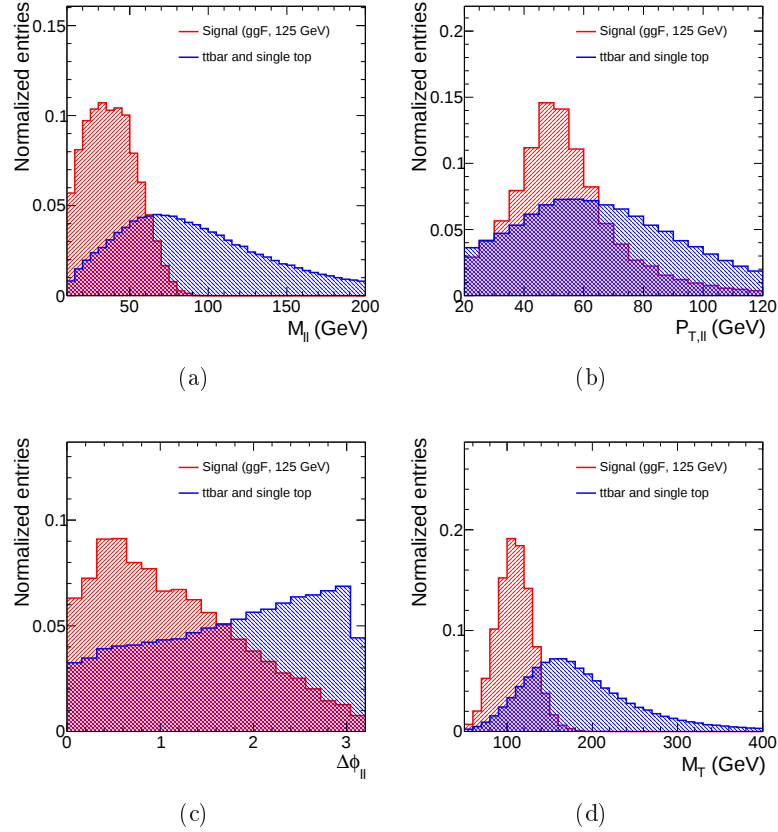


Figure 4.7: Normalized distributions after the preselection (see Section 4.5) comparing the signal shape (ggF only, in red) with that of the top-quark backgrounds (in blue). Both $t\bar{t}$ and single top-quark processes are considered. Only the $e\mu$ final state is shown. (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T .

4.2.3 The $Z + \text{jets}$ background

An important background, especially in the same-flavour channels, is the one where a Z/γ^* boson decays into two isolated, oppositely charged, high- p_T leptons. As shown in Fig. 4.8, the intermediate boson can be produced via $q\bar{q}$ annihilation as well as $qq, q\bar{q}$ scattering. The cross section, multiplied by the leptonic branching ratio, is about 10^4 times larger than the one predicted for the signal (cf. Table 4.1).

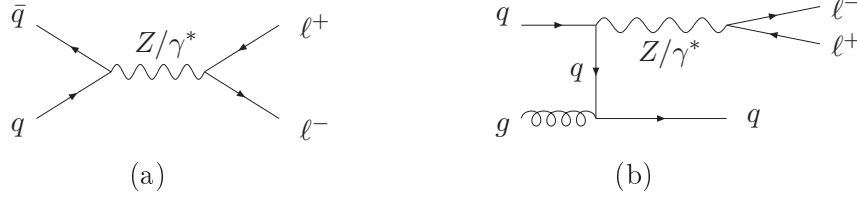


Figure 4.8: Feynman diagrams at leading order for the production and the leptonic decay of a Z/γ^* boson, via (a) $q\bar{q}$ or (b) gg . In the latter case, an associated jet is also produced.

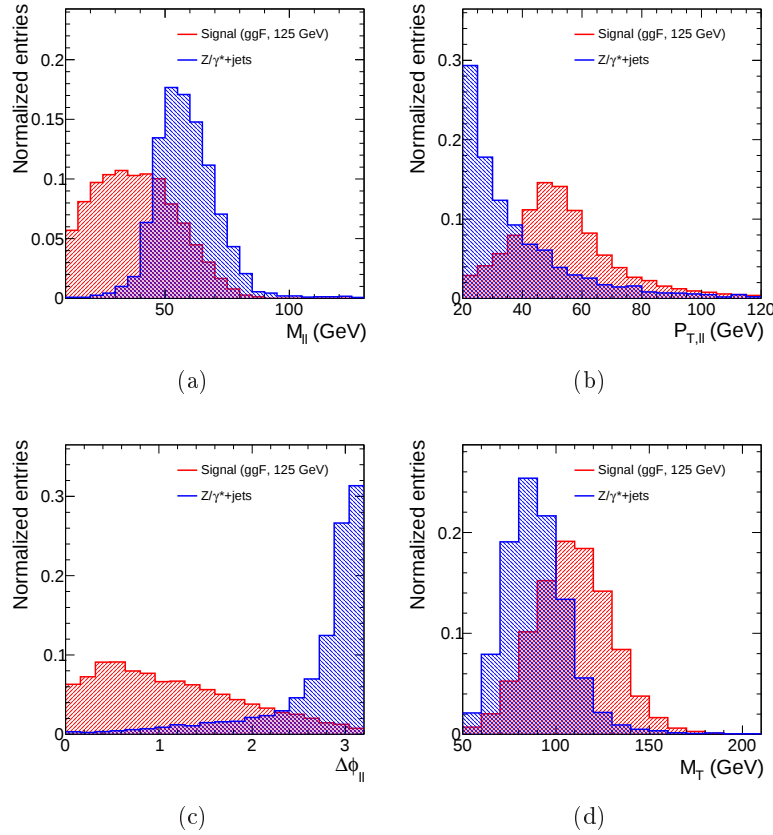


Figure 4.9: Normalized distributions after the preselection (see Section 4.5) comparing the signal shape (ggF only, in red) with that of the $Z/\gamma^*+\text{jets}$ background (in blue). Only the $e\mu$ final state is shown. (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T .

Fig. 4.9 compares the topological variables for the signal and the $Z+\text{jets}$ backgrounds, in the $e\mu$ final state. No true missing transverse energy is expected in these events, unlike in the signal case, but a certain fraction can still be present, due to mismeasured jets or leptons; moreover, the energy resolution gets worse with increasing pile-up. Setting a lower threshold on the missing transverse energy, there-

fore, helps reducing this background. Besides, in the same-flavour channels, the dilepton invariant mass peaks around the Z boson mass, and the leptons tend to be back-to-back in the azimuthal plane, resulting in a large $\Delta\phi_{\ell\ell}$ angle. Also the opposite-flavour category can contain some Z/γ^* background, if the intermediate boson decays into a $\tau^+\tau^-$ pair which subsequently decays into $e^\pm\mu^\mp$.

4.2.4 The W + jets background

The production and decay of a W boson in association with jets, shown in Fig. 4.10, can be mistaken for a $H \rightarrow WW^{(*)}$ event if one of the jets is reconstructed as a lepton (thus being called a *fake* lepton). The missing energy in this case is provided by the neutrino resulting from the W leptonic decay. The estimation of this background is particularly challenging because its size is comparable to that of the signal⁽²⁾ and also the shape of the characteristic variables after the selection is similar, as shown in Fig. 4.11.

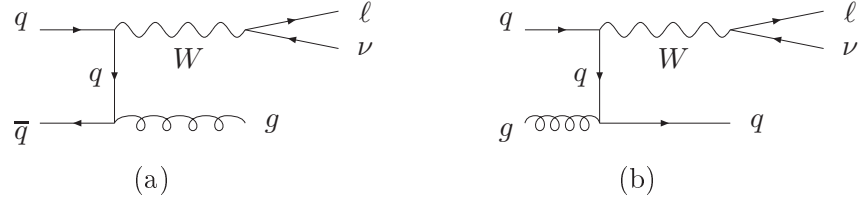


Figure 4.10: Feynman diagrams at leading order for the production and leptonic decay of a W boson, via (a) $q\bar{q}$ or (b) qg . In both cases, an associated jet is also produced.

4.2.5 The diboson ($WZ, ZZ, W\gamma$) backgrounds

Besides the WW background, other diboson backgrounds in this channel are WZ, ZZ and $W\gamma^{(*)}$, with the gauge bosons decaying leptonically. These final states are expected to contain three or more leptons, so a veto on additional leptons is quite effective in reducing them. As shown in Table 4.1, the cross sections range from 345 pb for $W\gamma$ to 5.6 pb for ZZ ; all of these backgrounds are found to be small after the kinematical cuts, and are estimated directly from Monte Carlo simulations, after validation in dedicated control regions.

²As a rough estimate, the inclusive cross section is a factor 10^4 larger than the signal, as for $Z/\gamma^* + \text{jets}$, but the rate for jets to fake leptons is 10^{-4} , yielding ~ 1 as a result.

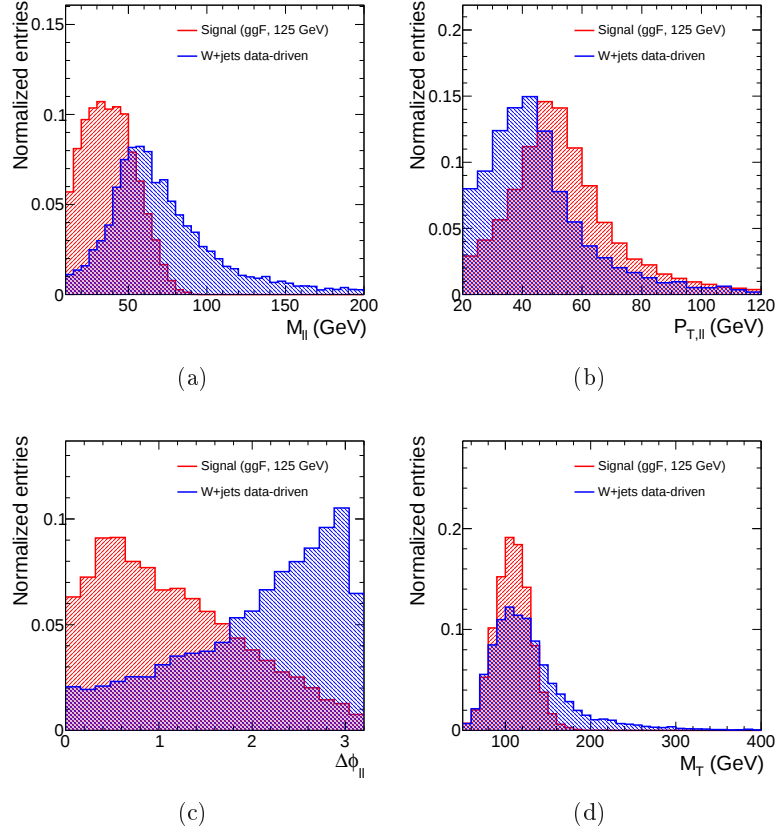


Figure 4.11: Normalized distributions after the preselection (see Section 4.5) comparing the signal shape (ggF only, in red) with that of the data-driven W +jets background (in blue). (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T .

4.3 Monte Carlo samples

Several Monte Carlo generators are used to simulate the signal and background processes; they are tuned according to the knowledge gained in 2011 on the ATLAS detector performance, e.g. its alignment, pile-up, minimum bias conditions and material distribution. Different MC generators take care of the hard scattering process and of the following stage of showering and hadronization (cf. Section 1.4.3): in particular, as summarized in Table 4.1, **HERWIG** [79] (in association with **JIMMY** [84] for the simulation of the underlying event) is interfaced to **MC@NLO** [190] and **GG2WW** [191], while **PYTHIA** [78] to the others.

Concerning the Higgs boson signal, besides the main gluon-fusion production mode (denoted ggF in the following), also the vector boson fusion (VBF) and the associated production (WH, ZH) modes are considered. The branching ratios into the $WW^{(*)} \rightarrow \ell^+ \nu \ell^- \bar{\nu}$ final state are provided by the **HDECAY** [133] program. Several

Process	MC generator	PDF sets	$\sigma \cdot \text{BR}$ (pb)
ggF ($m_H = 125$ GeV)	POWHEG+ PYTHIA	MRSTMCa1	0.347
VBF ($m_H = 125$ GeV)	POWHEG+ PYTHIA		$27 \cdot 10^{-3}$
WH, ZH ($m_H = 125$ GeV)	PYTHIA		$20 \cdot 10^{-3}$
$q\bar{q}/g \rightarrow WW$	MC@NLO+ HERWIG	CT10	4.68
$gg \rightarrow WW$	GG2WW+ HERWIG		0.14
$t\bar{t}$	MC@NLO+ HERWIG		167
single top-quark tW	AcerMC+PYTHIA	MRSTMCa1	85
single top-quark t -channel			
single top-quark s -channel			
W +jets	ALPGEN+ PYTHIA	CTEQ6L1	$32 \cdot 10^3$
Z/γ^* +jets	ALPGEN+ PYTHIA		$15 \cdot 10^3$
ZZ	Sherpa		5.6
WZ	MC@NLO+ HERWIG	CT10	18.0
$W\gamma$	ALPGEN	CTEQ6L1	345
$W\gamma^*$	MadGraph		6.5

Table 4.1: Monte Carlo generators and PDF sets used to simulate the different signal processes for a Higgs boson with $m_H = 125$ GeV, and the background processes. Also reported are the various cross sections as used in the analysis; the leptonic decay modes (charged only for Z/γ^* production) are summed over, apart from $t\bar{t}$, single top-quark, WZ and ZZ production, where the inclusive cross sections are quoted. For the signal processes, the production cross sections include the branching ratios into the $\ell\nu\ell\nu$ final state. See the text for further details and references.

hypotheses for the Higgs boson mass are considered, ranging from 110 to 600 GeV, in steps of 5 GeV for $m_H \leq 200$ GeV and of 20 GeV above.

The ggF hard process is simulated with the POWHEG generator [192, 193]; its cross section is computed to NNLO in QCD (NLO: [94, 194, 195], NNLO: [196–198]) resumming soft gluons up to NNLL [199]. The Higgs boson p_T spectrum is reweighted in such a way that the one obtained with the POWHEG generator matches the HqT [200] one. The VBF process is also simulated with POWHEG; its cross section is calculated up to NLO [99, 201, 202], and includes the leading NNLO QCD terms [100]. Finally, PYTHIA is used for the associated WH, ZH production; the cross sections quoted in Table 4.1 are calculated up to NNLO [104, 203]. In all cases, the electroweak corrections, which are assumed to be factorized from the QCD ones, are included up to NLO [109, 204, 205].

Concerning the backgrounds, WW and $t\bar{t}$ are simulated using MC@NLO (the $gg \rightarrow WW$ process with GG2WW), while ALPGEN [206] has been chosen for the inclusive W and Z/γ^* production³) and for $W\gamma$. In the latter case, kinematical filters have been applied at generator level: the photon must have $p_T > 10$ GeV and be separated from the charged lepton ($\Delta R > 0.1$). Kinematical criteria have been applied also

³For inclusive production of single gauge bosons in association with jets, the MLM matching scheme [207] is used to combine MC samples with different jet multiplicities, up to five for $W/Z +$ jets and up to three for $Wb\bar{b}$.

in the generation of Z/γ^* events ($m_{\ell\ell} > 10$ GeV) and of $W\gamma^*$ events (the leading and subleading transverse momenta of the leptons from the γ^* decay must exceed 15 and 5 GeV, respectively). The latter process is simulated with **MadGraph** [208, 209]. Moreover, ZZ employs **Sherpa** [210], WZ **MC@NLO**, single top-quark **AcerMC** [211]. Finally, **PYTHIA** is used to model the multi-jet background, which, despite its large cross section (4270 pb for $b\bar{b}$ production, requiring at least two leptons with $p_T > 10$ GeV), is heavily suppressed by the quality criteria for lepton identification and preselection cuts; for this reason, it is ignored in the following.

Concerning PDFs, as indicated in Table 4.1, the following sets are used: CT10 [71] for **MC@NLO** and all samples interfaced to **HERWIG**, CTEQ6L1 [212] for **ALPGEN**, **Sherpa** and **MadGraph**, **MRSTMC**al [69] for **AcerMC** and the **PYTHIA** samples.

4.4 Object selection

4.4.1 Data-taking conditions

All MC events are corrected for detector effects going through the full chain of the ATLAS simulation, provided by the **GEANT4** package [85]. In particular, a realistic treatment of the pile-up conditions in the 2011 run is included, overlaying simulated minimum bias events to the hard one. The 2011 data-taking has been split into two periods by a technical stop in September, in the following way:

	$\int L \, dt$	$\beta^*(\text{m})$	$\langle\mu\rangle$
March-August 2011	2.3	1.5	6.3
October-December 2011	2.4	1.0	11.6

where β^* characterizes the beam size (more specifically, it is the distance to the IP where the beam width has doubled). Decreasing it from 1.5 to 1.0 metres, causes the mean number of interactions per beam crossing, $\langle\mu\rangle$, to almost double. These conditions are propagated to the MC events through a reweighting procedure, in order to match them to the real data.

4.4.2 Event cleaning

Before applying the object selection, requirements are placed on the quality of the event itself, to ensure that all ATLAS subdetectors have been fully efficient at the time of data-taking (the data events must be flagged as “AllGood” [213], according to the definition given in Section 3.1.1) and that an hard interaction is actually present in the event. The latter condition is achieved through an “event cleaning” algorithm [214], which suppresses non-collision backgrounds such as cosmic muons,

noise in the calorimeters, beam-halo and beam-gas events. After the application of the so-called “GoodRunList”, the integrated luminosity available for physics analyses amounts to 4.7 fb^{-1} , with an uncertainty of $\pm 3.7\%$ [215].

Subsequently, quality criteria are applied to the primary vertex of the event (defined as in Section 3.3.2), requiring at least three associated tracks with $p_T > 400 \text{ MeV}$.

4.4.3 Trigger requirements

A preliminary selection of signal events is carried out requiring the presence of triggering leptons. As explained in Sections 3.3.3 for electrons and 3.3.4 for muons, only unprescaled, single-lepton triggers are used in this analysis, in order to maximize the acceptance. In particular, the triggers named EF_e20_medium , EF_e22_medium and $EF_e22vh_medium1$ are applied for electrons, while EF_mu18_MG and $EF_mu18_MG_medium$ for muons, tightening the corresponding quality and p_T requirements as the pile-up level increases. In the case of the $e\mu$ channel, an “or” of electron and muon triggers is used.

Scale factors are applied to the MC events to simulate trigger inefficiencies. First, per-lepton trigger efficiencies are computed ($\varepsilon_{\text{Data},\ell}$ and $\varepsilon_{\text{MC},\ell}$, for the data and MC events, respectively), using the tag-and-probe method on $Z \rightarrow \ell^+\ell^-$ events, and are found to be above 90% for all triggers. From these, per-lepton (SF_ℓ) and per-event scale factors are derived as a (η, ϕ) map :

$$\text{Trigger SF}_{\text{event}} = \frac{1 - \varepsilon_{\text{Data},\ell_1} \cdot \varepsilon_{\text{Data},\ell_2}}{1 - \varepsilon_{\text{MC},\ell_1} \cdot \varepsilon_{\text{MC},\ell_2}} = \frac{1 - (\varepsilon_{\text{MC},\ell_1} \cdot \text{SF}_{\ell_1}) \cdot (\varepsilon_{\text{MC},\ell_2} \cdot \text{SF}_{\ell_2})}{1 - \varepsilon_{\text{MC},\ell_1} \cdot \varepsilon_{\text{MC},\ell_2}}. \quad (4.2)$$

Out of the two selected leptons (according to the offline criteria outlined in the following), at least one must be matched to a triggering lepton, within a $\Delta R < 0.15$ cone (*trigger-matching*), otherwise the event is discarded. MC events passing the trigger-matching get an additional luminosity weight, to ensure the correct time exposure for the selected trigger.

4.4.4 Electron definition

As explained in Section 3.3.3, electron candidates are identified starting from the *tight++* requirement and adding further cuts:

- cuts on the longitudinal and transverse impact parameters and their significances, defined with respect to the primary vertex:

$$\begin{aligned} & - |z_0| < 1 \text{ mm}, \\ & - |d_0| < 10 |\sigma(d_0)|, \end{aligned}$$

- cuts on isolation (in a $\Delta R < 0.3$ cone, see Eq. 3.4 for the definition of p_T^{cone}):

$$\begin{aligned} & - p_T^{\text{cone}}/p_T < 0.13, \\ & - E_T^{\text{cone}}/p_T < 0.14. \end{aligned}$$

Isolation scale factors, consistent with unity within 1%, are applied to correct the small differences arising between the data and the MC expectation after the isolation requirements. Electrons are reconstructed within the EMCal geometrical acceptance ($|\eta| < 2.47$, excluding the crack regions $1.37 < |\eta| < 1.52$), provided that they have $p_T > 15$ GeV.

4.4.5 Muon definition

Muons are reconstructed in the active region of the muon spectrometer ($|\eta| < 2.4$), with $p_T > 15$ GeV. They are defined according to the STACO combined algorithm, with some additional quality cuts:

- cuts on the longitudinal and transverse impact parameters and their significances, defined with respect to the primary vertex:

$$\begin{aligned} & - |z_0| < 1 \text{ mm}, \\ & - |d_0| < 3 |\sigma(d_0)|, \end{aligned}$$

- cuts on isolation (in a $\Delta R < 0.3$ cone):

$$\begin{aligned} & - p_T^{\text{cone}}/p_T < 0.13, \\ & - E_T^{\text{cone}}/p_T < 0.14. \end{aligned}$$

- cuts on the track quality (see Section 3.3.4 for details).

4.4.6 Jet definition

Jets are defined according to the “anti- k_t EM+JES” algorithm, with $\Delta R = 0.4$, as anticipated in Section 3.3.5. The geometrical acceptance requires $|\eta| < 4.5$ and $p_T > 25$ GeV ($p_T > 30$ GeV in the transition regions $2.75 < |\eta| < 3.25$). To reduce the contribution of pile-up jets, a cut is placed on the *jet vertex fraction* (JVF) – which is the fraction of tracks in the jet pointing to the primary vertex of the event – requiring it to be above 75%. The performance of the JVF algorithm does not depend on the number of reconstructed primary vertices, therefore no correction is applied.

Moreover, b -jets are tagged with the help of the JetFitterCombNN algorithm (see Section 3.3.6), for which different working points can be chosen; this analysis,

in particular, requires 80% efficiency for b -jets identification. Vetoing b -jets gives a substantial handle in rejecting the top-quark backgrounds.

4.5 Event preselection

In the following, the preselection criteria for the $H \rightarrow WW^{(*)}$ multivariate analysis are discussed. They are applied after the GoodRunList, primary vertex and event cleaning criteria, in order to reduce the backgrounds which are easier to reject, since localized in phase space; the BDT technique is then used to train and classify the surviving topologies. These criteria are mostly in common with the preselection of the standard cut-based analysis — differences are stressed in the following. The number of signal and background events at various stages of the preselection is shown in Table 4.2, together with the number of observed events in data. Scale factors are applied for all backgrounds which are normalized in dedicated control regions, as discussed in Section 4.7.

1. **Opposite-sign leptons, p_T threshold, third lepton veto:** exactly two opposite-sign leptons must be present in the event, and at least one has to be trigger-matched. The threshold for the leading lepton p_T is increased to 25 GeV with respect to the identification requirement of 15 GeV, while the subleading one stays at 15 GeV. An overlap removal routine is applied, in the following way: a muon (jet) candidate is discarded if found inside a $\Delta R < 0.1$ (0.3) cone centered around the direction of the electron candidate; no overlap removal between muons and jets is applied, since its effect would be negligible. At this point, the background is dominated by the Z/γ^* +jets background (accounting for 90% of the total), while the S/B ratio is $6 \cdot 10^{-5}$.
2. **Cut on the dilepton invariant mass:** $m_{\ell\ell} > 10, 12$ GeV. The invariant mass of the dilepton system is required to be above 10 or 12 GeV, in the opposite- and same-flavour channels, respectively. This suppresses mainly the leptons coming from the γ^* and Υ resonances, and has a negligible impact on the overall number of events.
3. **Z veto:** to reduce the Z +jets background, events with an invariant mass closer than 15 GeV to the Z boson mass are rejected. This cut applies only to the same-flavour channels, and is quite efficient, since it rejects 90% of the Z/γ^* events and only 1% of the signal, lowering the S/B ratio to $6 \cdot 10^{-4}$.
4. **Cut on the missing transverse energy:** $E_{T,\text{rel}}^{\text{miss}} > 25, 45$ GeV. In order to further suppress the Z +jets background, a cut on $E_{T,\text{rel}}^{\text{miss}}$, a derived quantity

of the missing transverse energy, is employed. $E_{T,\text{rel}}^{\text{miss}}$ is identical to E_T^{miss} if there is no object (candidate lepton or jet) within $\Delta\phi \geq \pi/2$ from the $\vec{p}_T^{\text{miss}}$ vector, otherwise $E_{T,\text{rel}}^{\text{miss}}$ gets projected on the direction of the object itself (the nearest in case there are several):

$$E_{T,\text{rel}}^{\text{miss}} = E_T^{\text{miss}} \cdot \sin \Delta\phi \quad \text{if} \quad \Delta\phi < \pi/2. \quad (4.3)$$

Thanks to this procedure, the events with mismeasured jets or leptons, and therefore fake missing E_T close to the object, get a lower value of $E_{T,\text{rel}}^{\text{miss}}$, thus can be efficiently rejected by a cut on this quantity. The threshold value is 45 GeV for same-flavour channels, while 25 GeV is enough for $e\mu$, where the pollution from the Z +jets events, due only to $Z \rightarrow \tau\tau$ pairs, is less severe. After this cut, the S/B ratio goes down to $7 \cdot 10^{-3}$, with only 0.5% of the Z +jets events surviving; all background components are reduced by this cut (around 50% of events surviving for WW , $t\bar{t}$ and single top-quark events, 41% for W +jets, 32% for diboson processes), and also the signal is reduced to half of its previous size, with around 91 events expected for $m_H = 125$ GeV in 4.7 fb^{-1} of integrated luminosity.

Fig. 4.12 shows the $E_{T,\text{rel}}^{\text{miss}}$ (left) and $\mathbf{m}_{\ell\ell}$ (right) logarithmic distributions, for the data and the MC expectation, before the Z veto cut, broken down into lepton flavours. All backgrounds are normalized to their MC expectations, apart from W +jets which is data-driven (cf. Section 4.8.3). The lower part of each plot shows the ratio between the data and the background expectation, with the yellow band indicating the total normalization uncertainty (i.e. the shape uncertainties⁽⁴⁾ are not included). The data appear to be well modelled by the MC simulation, with an agreement well below the 10% level for both the $\mathbf{m}_{\ell\ell}$ and the $E_{T,\text{rel}}^{\text{miss}}$ distributions.

Table 4.2 gives, at each stage of the preselection, the expected numbers of events in 4.7 fb^{-1} of integrated luminosity. The two rightmost columns show the observed numbers of events in data and the data/MC ratio; the latter is compatible with unity for all cuts. In the bottom part, the predictions after the preselection are split into the lepton channels; it can be seen that, for all channels, the S/B ratios are around 0.7%.

⁽⁴⁾In the following, the name *shape uncertainty* refers to an uncertainty which does not affect the normalization of a variable, but only its shape. The shape uncertainties are never shown in the plots.

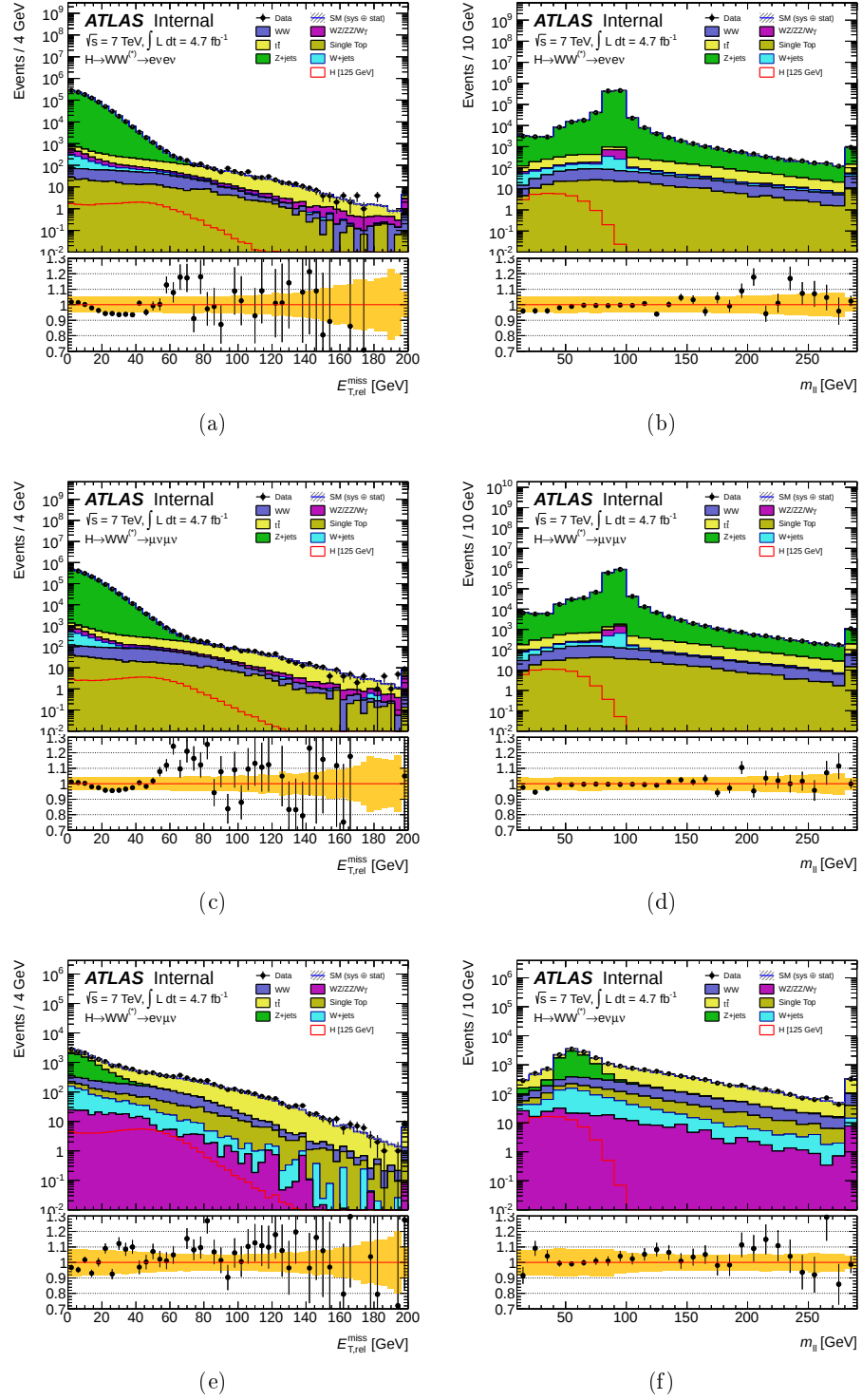


Figure 4.12: Distributions of the relative missing transverse energy $E_{T,rel}^{miss}$ (left) and of the dilepton invariant mass $m_{\ell\ell}$ (right), in logarithmic scale, for the individual lepton channels, before the Z veto: (a,b) ee , (c,d) $\mu\mu$, (e,f) $e\mu$. The lower part of each plot shows the data/MC ratio, with the yellow band indicating all uncertainties apart from the shape ones. All backgrounds are normalized as in Table 4.2. The final bin includes the overflow.

	Signal [125 GeV]	WW	WZ/ZZ/Wgamma	ttbar	single top	Z+jets	W+jets (d-d)	Total Bkg. (d-d)	Observed	Data/MC (d-d)
lepton p_T	180.87 ± 0.40	3871.06 ± 9.55	3213.65 ± 20.73	16332.68 ± 27.65	1753.31 ± 11.35	2800788.09 ± 1309.25	3159.63 ± 37.12	2829118.43 ± 1310.32	2823123	1.00 ± 0.00
OS leptons	178.43 ± 0.39	3856.00 ± 9.53	2733.59 ± 15.79	16240.28 ± 27.57	1734.85 ± 11.30	2793617.99 ± 1307.64	2373.89 ± 36.19	2820556.60 ± 1308.61	2816240	1.00 ± 0.00
$m_{\ell\ell} > 12, 10$ GeV	175.14 ± 0.39	3848.08 ± 9.52	2716.34 ± 15.68	16211.72 ± 27.55	1731.84 ± 11.29	2790757.50 ± 1307.32	2194.45 ± 35.91	2817459.92 ± 1308.28	2806551	1.00 ± 0.00
Z veto (for $ee, \mu\mu$)	173.11 ± 0.39	3421.39 ± 8.97	559.69 ± 8.00	14456.35 ± 26.02	1542.36 ± 10.66	278756.95 ± 404.20	1403.39 ± 20.67	300140.13 ± 405.88	298691	1.00 ± 0.00
$E_{T,rel}^{\text{miss}} > 45, 25$ GeV	90.82 ± 0.28	1828.61 ± 113.91	177.88 ± 4.82	7322.28 ± 497.99	839.24 ± 57.60	1523.00 ± 33.78	341.94 ± 5.92	12032.95 ± 688.35	12231	1.02 ± 0.06
$ee : E_{T,rel}^{\text{miss}} > 45$ GeV	10.24 ± 0.10	208.11 ± 13.14	26.70 ± 1.88	929.80 ± 63.55	107.98 ± 7.89	289.44 ± 14.91	56.34 ± 1.10	1618.38 ± 91.86	1652	1.02 ± 0.06
$\mu\mu : E_{T,rel}^{\text{miss}} > 45$ GeV	22.31 ± 0.14	389.62 ± 24.42	26.72 ± 0.84	1536.65 ± 104.78	173.20 ± 12.31	698.71 ± 23.59	24.67 ± 2.87	2849.57 ± 154.06	2854	1.00 ± 0.06
$e\mu : E_{T,rel}^{\text{miss}} > 25$ GeV	58.28 ± 0.23	1230.87 ± 76.74	124.45 ± 4.35	4855.83 ± 330.36	558.06 ± 38.49	534.85 ± 13.88	260.93 ± 5.05	7565.00 ± 445.09	7725	1.02 ± 0.06

Table 4.2: The expected numbers of signal (in the hypothesis $m_H = 125$ GeV) and background events after the preselection requirements listed in the first column, as well as the observed numbers of events in data. The rightmost column shows the ratio between the data and the MC estimation. The W +jets background is estimated entirely from data, whereas MC predictions normalized to data in the control regions (before the final fit) are used for the WW , Z/γ^* +jets, $t\bar{t}$, and $tW/tb/tqb$ processes. Contributions from other background sources are taken directly from the MC predictions. The total uncertainties (statistical, normalization systematic uncertainties, shape systematic uncertainties) are also given. The bottom part of the table lists the number of expected and observed events for each lepton channel after the preselection.

Fig. 4.13 shows the jet multiplicity distributions after the preselection, for all lepton channels combined as well as for the individual ones. It appears evident that the background composition depends strongly on the number of jets in the final state. In particular the background is dominated by WW and Z/γ^* events when no jets are present, and by the top- and single top-quark backgrounds when at least one jet is identified. As for the signal, in the 0-jet channel the main production mode is gluon-fusion, while in the 1- and 2-jet channels a larger fraction of events is produced via VBF.

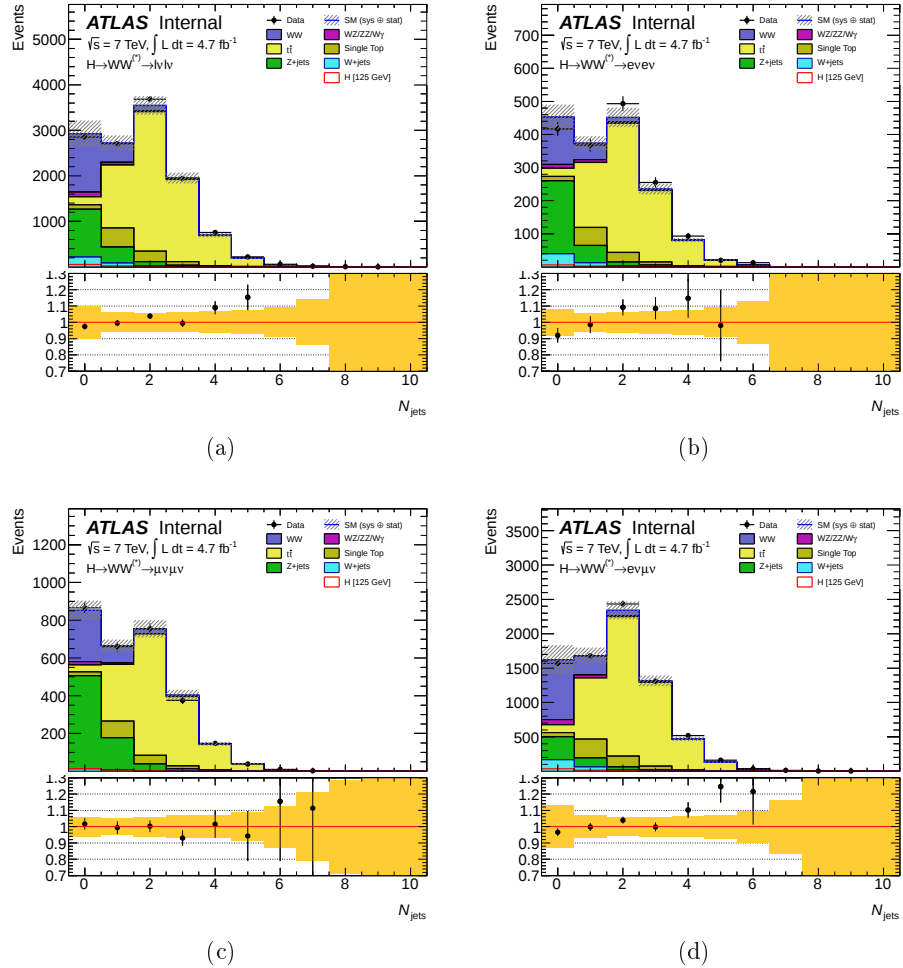


Figure 4.13: Distributions of the jet multiplicities after the preselection, for the combined case and the individual lepton channels: (a) $\ell\ell$, (b) ee , (c) $\mu\mu$, (d) $e\mu$. The lower part of each plot shows the data/MC ratio, with the yellow band indicating all uncertainties apart from the shape ones. All backgrounds are normalized as in Table 4.2.

As a consequence, after the preselection, additional selection criteria are applied according to the jet multiplicity. As noticed above, leptons arising from a heavy Higgs boson would not be characterized by a small opening angle; for this reason, the analysis is also split according to the signal mass hypothesis: the multivariate analysis uses a low-mass and a high-mass selection (below and above 200 GeV, respectively), while the cut-based analysis uses three (low-mass below 200 GeV, intermediate-mass between 200 GeV and 300 GeV, high-mass above 300 GeV).

Due to the lower number of signal events expected in 4.7 fb^{-1} of integrated luminosity, no multivariate analysis has been developed for the 2-jet channel. On the other hand, in 2012, with increased integrated luminosity and energy in the centre-of-mass, a BDT analysis has been developed for the 2-jet channel, starting from the analysis presented in this thesis.

4.6 The Boosted Decision Tree technique

After the preselection, the overwhelming $Z/\gamma^* + \text{jets}$ background has been reduced, so that its contribution is comparable with those of the WW and top-quark components; the total S/B ratio is 2% and 0.8% in the 0- and 1-jet channels, respectively. At this point, a boosted decision tree technique is employed, to classify the events as signal-like or background-like, both in the MC simulation and in the data, and then the *BDT response* (or *output*) is fitted to extract the final exclusion limits.

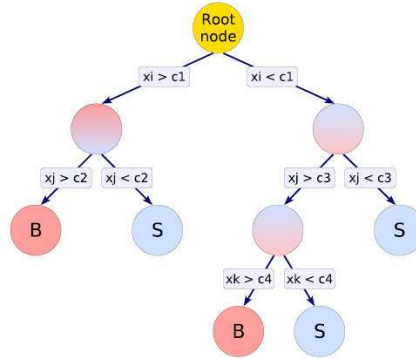


Figure 4.14: Schematic view of a decision tree [216]. In this example, three levels of splittings are performed: the first uses the variable x_i to divide the training events in two nodes, the second uses x_j on each node, the third uses x_k . These variables are chosen because they give the best separation between signal and background when being cut on. At the end of the procedure, leaves are defined as signal- or background-like, depending on the majority of the events they contain.

A decision tree can be represented as in Fig. 4.14. It is based on a set of input variables, whose shapes discriminate between signal and background events. A *training*

procedure is performed on (a subset of) MC samples, consisting in a series of subsequent splittings: at each step, the algorithm finds the most discriminating variable and the optimal cut on it, in order to achieve the best *separation* between the signal and the background. The separation is defined as [216]:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{[f_S(y) - f_B(y)]^2}{f_S(y) + f_B(y)} dy \quad (4.4)$$

where f_S and f_B are the signal and background PDFs, respectively. The separation is $\langle S^2 \rangle = 0$ for identical distributions and $\langle S^2 \rangle = 1$ if there is no overlap, hence the algorithm maximizes $\langle S^2 \rangle$.

At each step, the events are split between the *signal-leaf* and the *background-leaf*, until the maximum number of allowed splittings (the *depth* of the tree) is reached; at this point, an event is correctly identified if it is a signal (background) event and is located into a signal-leaf (background-leaf). Events which have been wrongly identified, on the other hand, receive an increased weight (or *boost*) with respect to the others, such that a second tree can focus on them. This procedure is repeated multiple times, until a *forest* is grown. The idea of building a whole forest, rather than a single tree, is meant as a protection against statistical fluctuations in the input samples. At the end, a weighted average (with the weight given by the boost) is taken across all trees, to form the BDT response. Moreover, boosting the trees also improves the algorithm performance [216].

The training information is contained inside an *xml* file, which consists, for each event, in a mapping between the N input variables and the 1-d BDT response. In the subsequent *classification* step, this mapping is applied to all events.

The BDT response is distributed between 0 and 1, where, by definition, signal-like (background-like) events tend to have values close to 1 (0), as can be seen e.g. in Fig. 4.19. The more they are separated from each other, the better the BDT performance. Once the final discriminant is built, one can either cut on it, selecting only the events above the cut, which are expected to have the best S/B ratio, or fit the whole shape, using the leftmost part, which is signal-depleted, to constrain the backgrounds. This analysis employs the second technique.

An optimization has been carried out both on input variables (see Appendix B) and on tree parameters. Concerning the latter, some changes with respect to the default TMVA settings [216] have been applied, in order to minimize the effect of the statistical fluctuations. More details are given in Section 4.6.1. Moreover, Appendix A contains a comparison of the BDT and the Neural Network performances, in several configurations. Here, only the final results for the chosen configuration are discussed.

4.6.1 BDT settings

The BDT training is performed in steps of 5 GeV at low mass ($110 \text{ GeV} < m_H \leq 200 \text{ GeV}$) and of 20 GeV at high-mass ($200 \text{ GeV} < m_H \leq 600 \text{ GeV}$), for a total of 39 mass points. From the technical point of view, training and classification are done through the TMVA v4.1.2 software [216], which is part of the ROOT v5.30.02 release. Several implementations of the boosting technique can be found; this analysis uses the *bagged gradient boosting* one.

When estimating a function $F(x)$ through boosting, one can expand it as a weighted sum of parametrised base functions, in the following way [216]:

$$F(x; P) = \sum_{m=0}^M \beta_m f(x; a_m) \quad (4.5)$$

where the base functions $f(x; a_m)$ correspond to the decision trees, and there are M parameters $P \in (\beta_m; a_m)$. The distance between the “true” function y , obtained from the training, and the model $F(x)$, can be estimated by means of a *loss function*, which, in case of gradient boosting, is a binomial log-likelihood:

$$L(F, y) = \ln \left(1 + e^{-2F(x)y} \right). \quad (4.6)$$

Decision trees are grown by the algorithm in a way that the loss function is minimised. Gradient boosting is typically resilient against overtraining, a characteristic that can be enhanced by reducing the learning rate through the `shrinkage` parameter, which controls the weight of the individual trees: the smaller the shrinkage, the more trees will be grown, slowing down the convergence but improving the accuracy of the prediction. The value chosen by this analysis, 0.1, is indeed small.

The usage of the option `UseBaggedGrad = True` means that *stochastic gradient* boosting is enabled: random subsets of the training samples are chosen to grow the trees; the fraction of events at each iteration is controlled by the parameter `GradBaggingFraction`, 0.5 in this case. The stochastic re-sampling allows for additional stability of the algorithm.

The above mentioned BDT settings, together with the others which have been changed with respect to the default, are listed in Table 4.3. The choice of `MaxDepth = 2` indicates that only two levels of splitting are allowed for each tree. The parameter `nEventsMin` corresponds to the *stopping criterion* of the tree growing: the splitting ends when an additional step would bring the number of events in each leaf below 1% of raw MC signal events. The parameter `IgnoreNegWeightsInTraining = True` means that, if a MC event has an associated negative weight (as it is the case for

MC@NLO samples), the negative sign is ignored, and the weight is taken to be positive. This choice is due to the fact that no proper way of treating negative-weighted events has been implemented in TMVA yet; ignoring their weights does not bias the result, since they have the same topology of positive-weighted events, and hence can be safely added to the latter.

BDT setting	Value
BoostType	Gradient
MaxDepth	2
NTrees	1000
nEventsMin	1% of signal events
IgnoreNegWeightsInTraining	True
UseBaggedGrad	True
GradBaggingFraction	0.5
Shrinkage	0.1

Table 4.3: BDT settings used in the 2011 analysis, which have been changed with respect to the default.

4.6.2 BDT input variables

The input variables are the same in all cases (0- and 1-jet analyses, low- and high-mass), namely the four topological variables defined in Section 4.2:

- the dilepton azimuthal opening angle $\Delta\phi_{\ell\ell}$
- the dilepton invariant mass $m_{\ell\ell}$
- the dilepton transverse momentum $p_{T,\ell\ell}$
- the transverse mass of the total system m_T .

Their choice resides in two aspects. On one hand, they are used also in the cut-based analysis to cut on, implying that they have been extensively studied and validated for that purpose, also for what concerns the impact of systematic uncertainties. On the other hand, given the limited number of MC events available after the preselection, especially for the signal samples, the BDT performance is found to be saturated when using four variables: adding more variables does not add discrimination. As shown in Appendix B, since no other set of four variables significantly outperforms the topological ones, they result to be the most convenient choice. As a side note, it should be mentioned that many multivariate techniques have been tested before choosing the boosted decision tree one, using the four cut-based variables as common inputs, as documented in Appendix A.

Training in the 0-jet channel

In the 0-jet channel, the training is performed after the preselection, with an additional cut on $p_{T,\ell\ell}$ for all lepton flavours: $p_{T,\ell\ell} > 30$ GeV. This cut is used also in the cut-based analysis, but only after an additional cut on $m_{\ell\ell}$ ($m_{\ell\ell} < 50$ GeV). The cut on $p_{T,\ell\ell}$ is meant to strongly reduce the $Z/\gamma^* + \text{jets}$ background; in general, whereas a rectangular cut is efficient in rejecting a particular background, it is useful to apply it before the training, to avoid loading the BDT algorithm with topologies which are anyway very easy to distinguish. It should be noticed that the cut on $p_{T,\ell\ell}$ is similar to the one applied in the preselection on $E_{T,\text{rel}}^{\text{miss}}$, due to the fact that, in the 0-jet channel, the E_T^{miss} and $p_{T,\ell\ell}$ vectors should balance each other.

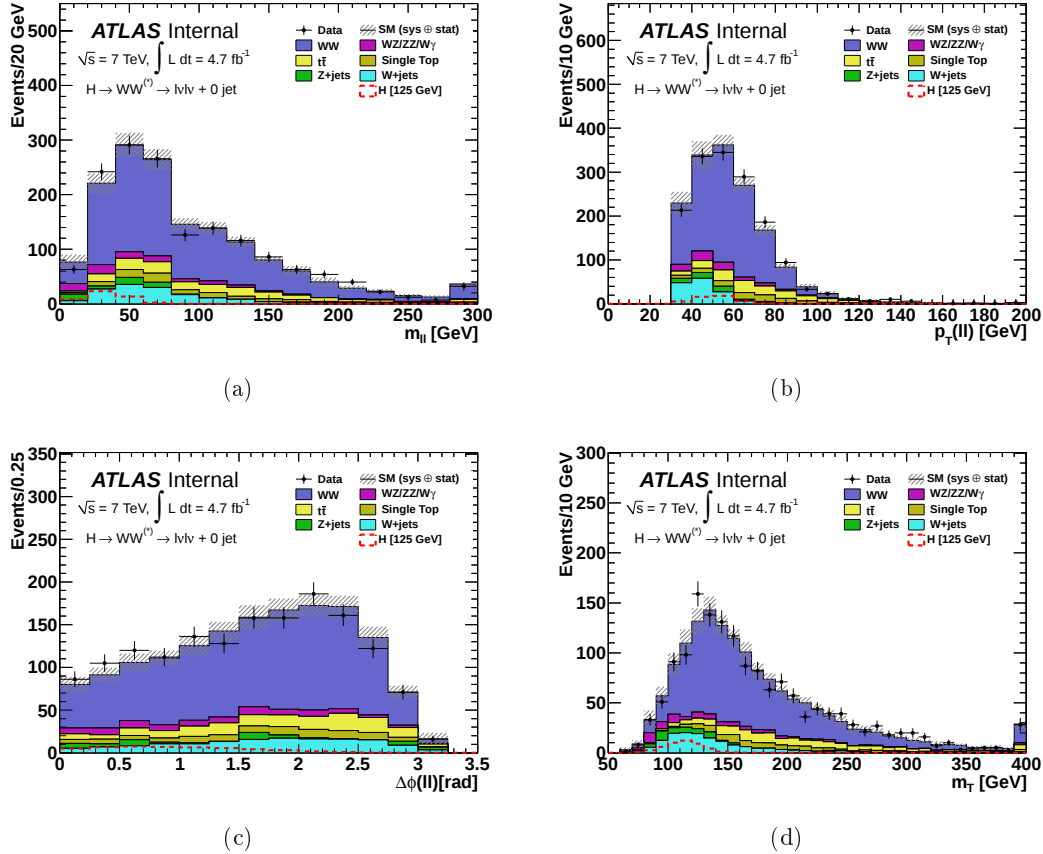


Figure 4.15: Distributions of the four input variables, after the $p_{T,\ell\ell}$ requirement, as used for the BDT training, in the 0-jet channel: (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . Only combined distributions are shown, the individual lepton channels can be found in Appendix C. The shaded bands indicate the total normalization systematic uncertainty. The final bin includes the overflow. All backgrounds are normalized as in Table 4.2.

The distributions of the four input variables, after the $p_{T,\ell\ell}$ requirement, are shown in Fig. 4.15 for the combined lepton flavours, and in Appendix C for the individual ones. No significant deviation between the data and the MC expectation is observed, and the data points lay mostly inside the shaded band provided by the total MC normalization uncertainty.

Given the similarities of the ee and $\mu\mu$ channels, they are trained together in order to increase the number of MC events, while the $e\mu$ channel is trained separately. Nevertheless, the final fit is done individually for each channel. Not all events are used for training: each sample is split into two halves, according to the event number, and one half is used for training, the other half for classification. This is done in order to avoid biasing the BDT learning procedure, otherwise the MC events would be used twice.

The MC samples with a very low number of events are excluded from the training: this is the case for the $\gamma^* + \text{jets}$ (Z/γ^* events with $m_{\ell\ell} < 40$ GeV) and the $W\gamma$ samples. Also $W + \text{jets}$ events are not used, since they are obtained through a data-driven technique, as discussed in Section 4.8.3, and no information from real data should enter the learning step of the BDT algorithm. All other MC samples are used for the training, each one weighted by the corresponding cross section.

At the end of the training, each input variable is ranked according to its discrimination power, counting how many times that variable has been chosen for splitting a node, weighted by the number of events in the leaf and the gain in S/B separation. The ranking order is:

- m_T , $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$ and $p_{T,\ell\ell}$ for the same-flavour channels,
- $\Delta\phi_{\ell\ell}$, m_T , $m_{\ell\ell}$ and $p_{T,\ell\ell}$ for the opposite-flavour channels.

Fig. 4.16 shows the linear correlation coefficients among the four input variables, for the signal (a,c) and the background (b,d) samples, for the same-flavour (a,b) and the opposite-flavour (c,d) channels. The largest correlations, either positive or negative, are found to be between $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$ and $p_{T,\ell\ell}$ for the signal, as expected, and between $m_{\ell\ell}$ and m_T for the background.

Training in the 1-jet channel

In the 1-jet channel, the training is performed immediately after the preselection, without additional cuts. At this stage, as noticed above and as can be seen in Table 4.5, the background is dominated by the top-quark contributions, which account for 70% of the total. The same four variables are used as inputs, despite the presence of a jet in the final state, since optimization studies have shown that adding jet-related

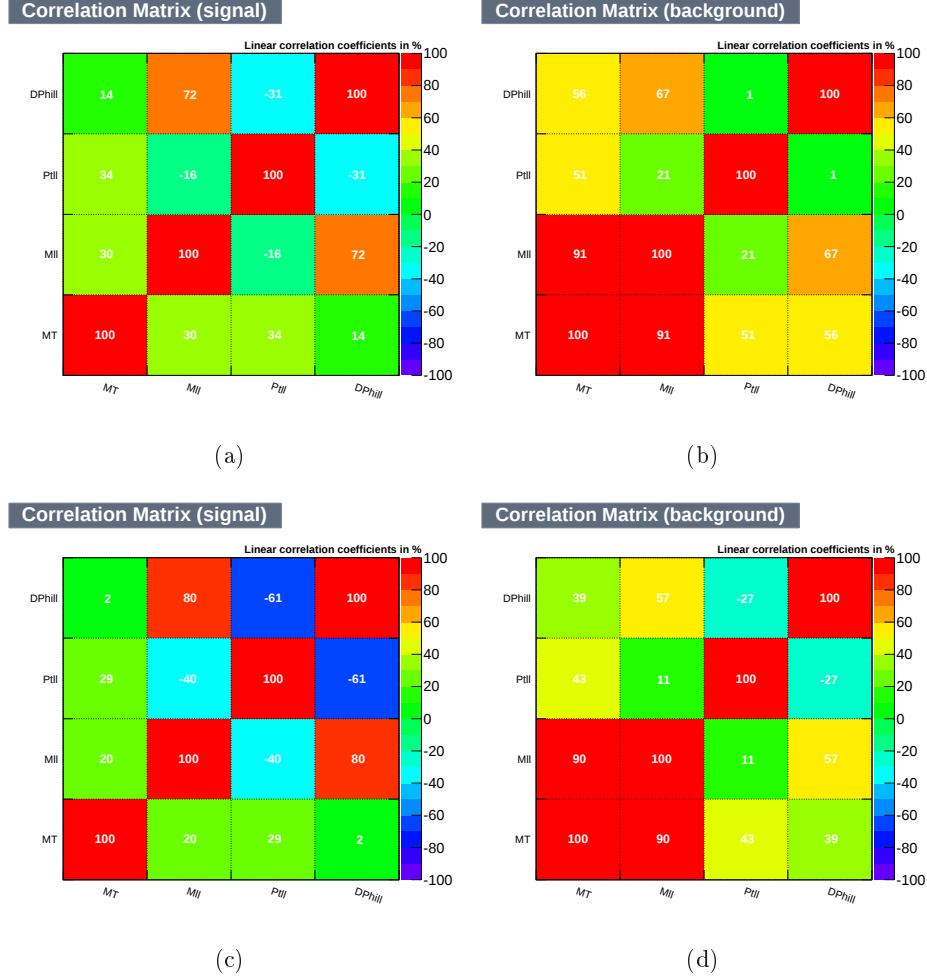


Figure 4.16: Linear correlation coefficients among the four input variables, in the 0-jet channel, for the signal (a,c) and the total background (b,d), in the same-flavour channels (a,b) and in the opposite-flavour channels (c,d).

variables (e.g. jet p_T) does not improve significantly the separation between the signal and the background.

The distributions of the discriminating variables are shown, at the end of the selection (but before the last $m_{\ell\ell}$ cut), in Fig. 4.17 for the combined lepton channels, and in Appendix C for the individual lepton flavours. Also in this case, the agreement between the data and the MC simulation is satisfying; the shaded band is larger where there is a residual Z/γ^* contribution, given the uncertainty on the mismeasurement of E_T^{miss} .

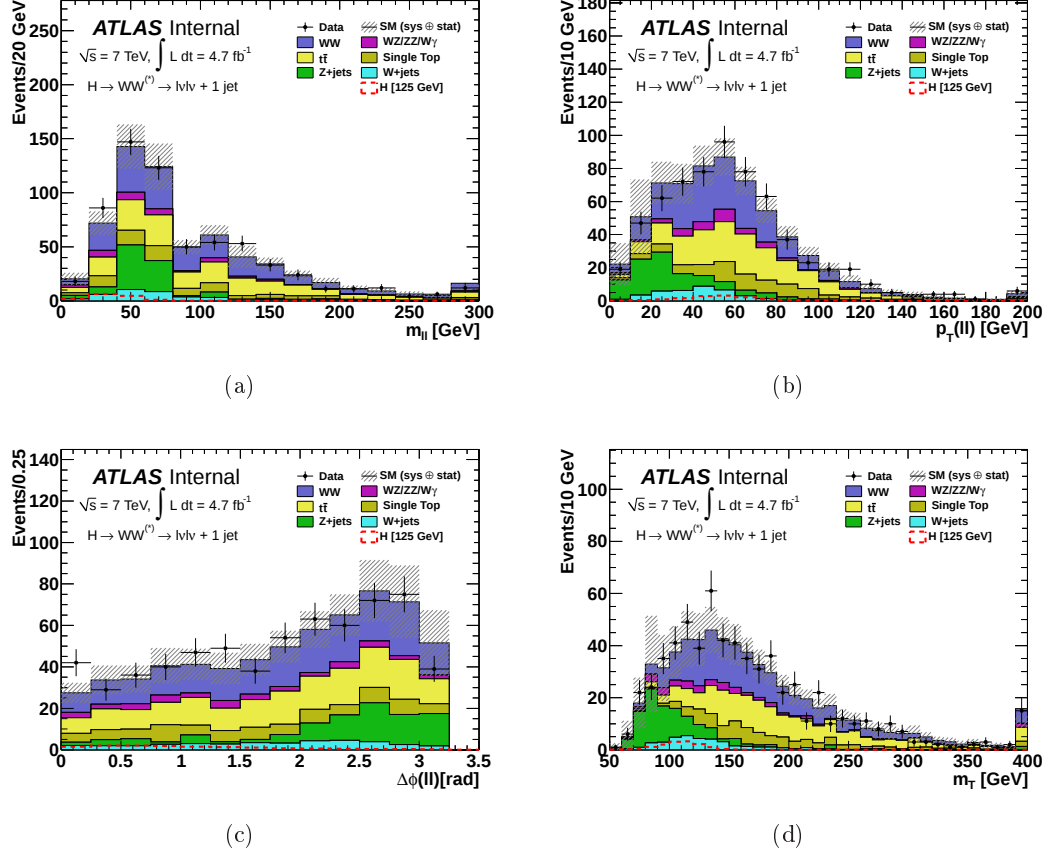


Figure 4.17: Distributions of the four input variables in the 1-jet signal region (before the $m_{\ell\ell}$ cut, cf. Section 4.7.4): (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . Only combined distributions are shown, those for the individual lepton channels can be found in Appendix C. The shaded bands indicate the total normalization systematic uncertainty. The final bin includes the overflow. All backgrounds are normalized as in Table 4.2.

The same training details given for the 0-jet channel hold here. Concerning the variable ranking, for the same-flavour channels it stays the same as for 0-jets (m_T , $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$ and $p_{T,\ell\ell}$), while for $e\mu$ it is $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, m_T and $p_{T,\ell\ell}$. The linear correlation coefficients among the four input variables are shown in Fig. 4.18, for the signal (a,c) and the background (b,d) samples, for the same-flavour (a,b) and the opposite-flavour (c,d) channels. Again, the largest correlations, either positive or negative, are found between $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$ and $p_{T,\ell\ell}$ for the signal and between $m_{\ell\ell}$ and m_T for the background.

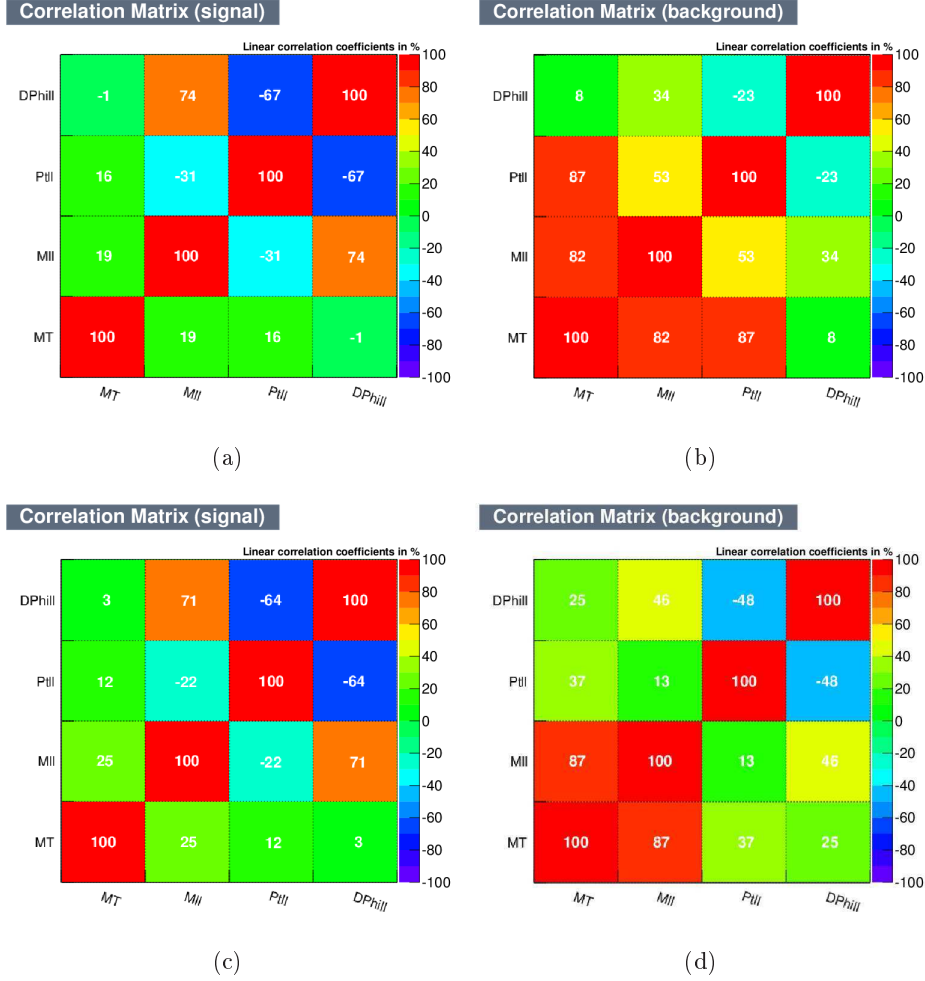


Figure 4.18: Linear correlation coefficients among the four input variables, in the 1-jet channel, for the signal (a,c) and the background (b,d), in the same-flavour channels (a,b) and in the opposite-flavour channels (c,d).

4.7 BDT response

4.7.1 Signal region for the cut-based analysis

To ease the comparison with the cut-based analysis, its selection criteria are summarized in Table 4.4. They are mostly similar to those of the BDT analysis, commented in the following.

For completeness, also the details of the 2-jet analysis are given in the Table; they target the VBF topology for the signal, aiming to distinguish it, in particular, from the $t\bar{t}$ background. The two jets with highest p_T are labelled as *tagging* jets, and the following cuts refer to them. They must be reconstructed in opposite hemispheres in η , and have a separation from each other $\Delta\eta_{jj} > 3.8$. The presence of additional

central jets is vetoed (in order to distinguish the VBF topology from the QCD one, cf. Section 2.1.2), while the two tagging jets cannot be identified as originating from the decay of a b or c -quark. Afterwards, cuts on the dijet invariant mass and on the total transverse momentum of the system are imposed, and the $Z \rightarrow \tau\tau$ veto is applied, as explained in Section 4.7.4 for the 1-jet channel. Finally the topological selection is applied, with cuts on $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$.

At the end of the selection, a fit on the whole shape of the m_T variable is performed. In the so-called cut-and-count analysis, on the other hand, a cut on m_T is performed, depending on the mass of the Higgs boson signal, $0.75 m_H < m_T < m_H$, and then the number of MC events in the mass window is computed and compared with the corresponding number in the data.

0-jet channel	• no identified jets • $p_T^{\ell\ell} > 45$ GeV for the same-flavour channels • $p_T^{\ell\ell} > 30$ GeV for the opposite-flavour channels
1-jet channel	• one identified jet • b-jet veto for the identified jet • $\vec{p}_T^{\text{TOT}} = \vec{p}_{T,\ell_1} + \vec{p}_{T,\ell_2} + \vec{p}_{T,\text{jet}} + \vec{p}_{T,\text{miss}} < 30$ GeV • $Z \rightarrow \tau\tau$ veto (cf. Section 4.7.4)
2-jet channel	• at least two identified jets • $\Delta\eta_{\text{jet},1} \times \Delta\eta_{\text{jet},2} < 0$ • no additional jets with $p_T > 25$ GeV ($p_T > 30$ GeV) inside $\eta < 2.5$ ($2.5 < \eta < 3.2$) • $\Delta\eta_{jj} > 3.8$ • $m_{jj} > 500$ GeV • $\vec{p}_T^{\text{TOT}} = \vec{p}_{T,\ell_1} + \vec{p}_{T,\ell_2} + \vec{p}_{T,\text{jet},1} + \vec{p}_{T,\text{jet},2} + \vec{p}_{T,\text{miss}} < 30$ GeV • $Z \rightarrow \tau\tau$ veto (cf. Section 4.7.4)
Topological cuts	• $m_{\ell\ell} < 50$ GeV if $m_H < 200$ GeV in the 0- and 1-jet channels • $m_{\ell\ell} < 80$ GeV if $m_H < 200$ GeV in the 2-jet channel • $m_{\ell\ell} < 150$ GeV if $200 \text{ GeV} \leq m_H \leq 300$ GeV • $\Delta\phi_{\ell\ell} < 1.8$ if $m_H < 200$ GeV

Table 4.4: Selection criteria of the cut-based analysis [28]. See the text for further details.

4.7.2 Overtraining checks

The TMVA software provides a tool to check for *overtraining*, which automatically compares the BDT response shape of the trained and classified samples. When the BDT performance is affected by overtraining, the cause resides in a low number of MC events, for which reason the tree growth has been based on statistical fluctuations: applying the learned function to the second half of events (the odd-numbered ones), results in a different shape of the BDT response. This can be checked performing a Kolmogorov-Smirnov (KS) compatibility test between the response shape of the even- and odd-numbered events. A KS value of x_{KS} indicates that the probability of two distributions to come from the same parent distribution is $P \geq x_{KS}$;

the compatibility is considered acceptable if $x_{KS} \geq 5\%$.

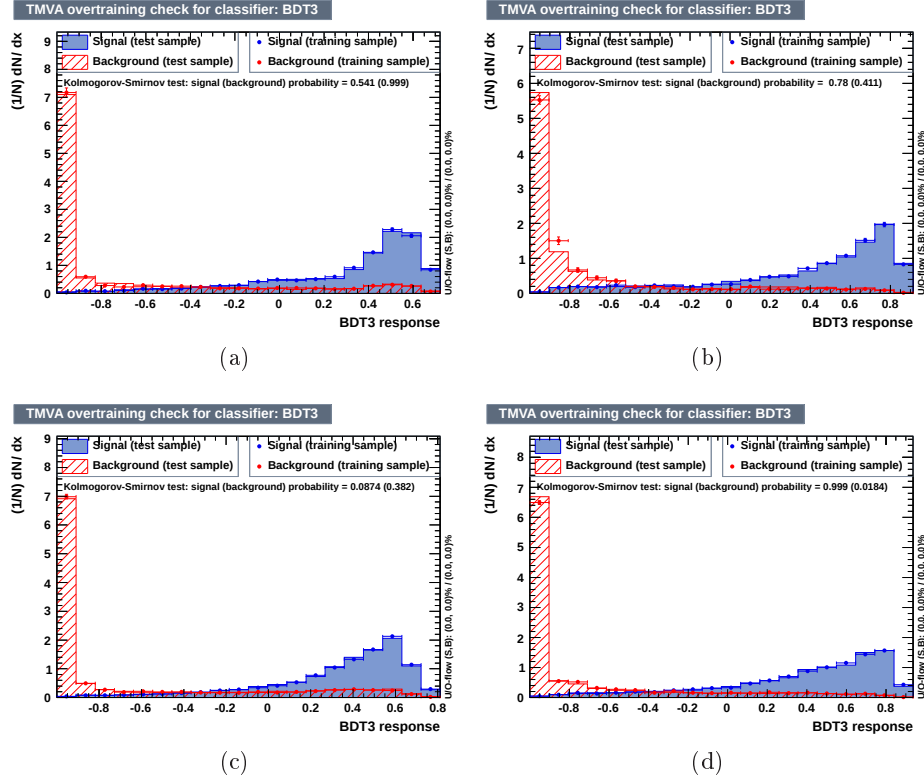


Figure 4.19: Distributions of the BDT response from the training (dots) and testing (full areas) samples, for the signal (blue) and the background (red) components. A Kolmogorov-Smirnov (KS) test has been performed, in order to assess their compatibility, and the results are given in the plots. Shown are: (a) $ee, \mu\mu$ 0-jet, (b) $ee, \mu\mu$ 1-jet, (c) $e\mu$ 0-jet, (d) $e\mu$ 1-jet.

Examples of these distributions can be seen in Fig. 4.19, for (a) $ee, \mu\mu$ 0-jet, (b) $ee, \mu\mu$ 1-jet, (c) $e\mu$ 0-jet and (d) $e\mu$ 1-jet. The training and testing distributions are compared: in most cases, the KS values are above 5%, and no large fluctuations are observed, so the chosen training configuration can be considered satisfying.

4.7.3 Signal region in 0-jet

As explained above, loose cuts are applied to MC events before the training. Afterwards, in the classification step, the so-obtained BDT response is applied to the odd-numbered events of all signal and background samples, and to the real data. In this phase, more stringent cuts are performed, in order to define a *signal region*, where the S/B ratio is maximal, and a set of *control regions*, in order to correctly assess the normalization of the most important background components. The expected

number of signal and background events, at each stage of the selection, is given in Table 4.5, where, for completeness, also the 2-jet cut-based analysis is shown. Also shown is the number of observed data events.

$H + 0\text{-jet}$	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
Jet Veto	56.7 ± 0.2	1273 ± 79	97 ± 4	174 ± 12	95 ± 7	1039 ± 28	217 ± 4	2893 ± 115	2849
$p_T^{\ell\ell}$ cut	47.7 ± 0.2	1041 ± 66	77 ± 3	151 ± 10	84 ± 6	45 ± 8	146 ± 3	1543 ± 76	1559
$m_{\ell\ell} < 80$ GeV	47.4 ± 0.2	557 ± 36	51 ± 3	60 ± 4	40 ± 3	37 ± 7	98 ± 2	844 ± 39	862
$H + 1\text{-jet}$	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
1 jet	22.7 ± 0.1	343 ± 54	56 ± 3	1438 ± 60	436 ± 19	357 ± 17	85 ± 3	2715 ± 142	2706
$b\text{-jet veto}$	20.9 ± 0.1	319 ± 50	52 ± 3	412 ± 18	139 ± 7	332 ± 16	76 ± 3	1330 ± 84	1369
$ \mathbf{p}_T^{\text{tot}} < 30$ GeV	14.0 ± 0.1	226 ± 35	34 ± 2	181 ± 8	80 ± 4	108 ± 8	37 ± 2	666 ± 51	684
$Z \rightarrow \tau\tau$ veto	14.0 ± 0.1	220 ± 34	34 ± 2	173 ± 8	77 ± 4	85 ± 7	37 ± 2	627 ± 50	644
$m_{\ell\ell} < 80$ GeV	13.9 ± 0.1	110 ± 17	22 ± 2	77 ± 3	40 ± 2	79 ± 6	27 ± 1	355 ± 28	374
$H + 2\text{-jet}$	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
opp. hemispheres	4.2 ± 0.1	46 ± 1	7 ± 1	149 ± 3	21 ± 1	32 ± 3	9 ± 1	264 ± 5	269
$ \Delta\eta_{jj} > 3.8$	1.8 ± 0.1	8.4 ± 0.4	0.9 ± 0.2	23.2 ± 1.0	2.2 ± 0.4	5.8 ± 1.7	1.7 ± 0.4	42.2 ± 2.1	40
$m_{jj} > 500$ GeV	1.3 ± 0.1	3.9 ± 0.3	0.4 ± 0.1	10.4 ± 0.6	1.0 ± 0.3	0.7 ± 0.4	0.9 ± 0.3	17.3 ± 0.9	13
$m_{\ell\ell} < 80$ GeV	0.9 ± 0.1	1.1 ± 0.2	0.1 ± 0.1	1.4 ± 0.2	0.4 ± 0.1	0.2 ± 0.2	0.2 ± 0.2	3.2 ± 0.4	2
$\Delta\phi_{\ell\ell} < 1.8$	0.8 ± 0.1	0.8 ± 0.1	0.1 ± 0.1	0.9 ± 0.2	0.1 ± 0.1	negl.	negl.	1.8 ± 0.3	1
Control Regions	Signal [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
WW 0-jet	0.28 ± 0.03	471 ± 3	26 ± 1	87 ± 2	42 ± 2	7 ± 2	49 ± 2	682 ± 5	697
WW 1-jet	0.12 ± 0.02	128 ± 2	12 ± 1	89 ± 2	34 ± 2	9 ± 2	11 ± 1	282 ± 4	270
Top 1-jet	1.16 ± 0.03	20 ± 1	1.9 ± 0.5	434 ± 4	169 ± 4	7 ± 2	4 ± 1	635 ± 6	676

Table 4.5: The expected numbers of signal and background events after the low-mass requirements listed in the first column, as well as the observed numbers of events in the data. The $W + \text{jets}$ background is estimated entirely from the data, whereas scale factors from control regions are used for the WW , $Z/\gamma^* + \text{jets}$ and top-quark processes. The contributions from the other backgrounds are taken directly from the MC predictions. Only statistical uncertainties are shown. The same is also shown in the control regions, but here no scale factors are applied.

In the 0-jet channel, the signal region (SR) is defined in the following way, after the jet veto:

- $p_{T,\ell\ell} > 45$ GeV for $ee, \mu\mu$, $p_{T,\ell\ell} > 30$ for $e\mu$,
- $m_{\ell\ell} < 80$ GeV if $m_H < 200$ GeV.

Cutting on $p_{T,\ell\ell}$ strongly reduces the Z/γ^* background, bringing the S/B ratio to 3%, where 2/3 of the total background are now constituted by WW . In the same-flavour case, the cut is reinforced with respect to the one used for training. The last cut is different for the BDT and the cut-based low-mass analyses: whereas the former cuts on $m_{\ell\ell} < 80$ GeV, the latter applies a tighter cut, $m_{\ell\ell} < 50$ GeV, and afterwards an additional cut on $\Delta\phi_{\ell\ell}$ (cf. Table 4.4). At the end of the selection, the S/B ratio is 5.6%, with the WW background still accounting for 66% of the total; with respect to the beginning of the selection, 26% of the signal events survived, while 99.97% of the background was rejected. 47 signal events are expected for a

$m_H = 125$ GeV Higgs boson, 3/4 of which come from the $e\mu$ channel, which is also driving the S/B ratio, as can be seen in Table 4.6. A total of $B = 844 \pm 39$ background events are expected, compatible with the observed number of events in data (862), so that exclusion limits on the presence of a Higgs boson signal can be set, with a fit to the BDT response shape. Fig. 4.20 shows the distributions of the BDT response in the signal region for the 0-jet channel, for the individual lepton flavours and their combination. Shown are the various background components, the signal expectation for $m_H = 125$ GeV and the events found in the data.

Lepton Channels	0-jet ee	0-jet $\mu\mu$	0-jet $e\mu$	1-jet ee	1-jet $\mu\mu$	1-jet $e\mu$
Total bkg.	106 ± 8	190 ± 16	549 ± 22	50 ± 7	85 ± 9	218 ± 17
Signal	4.4 ± 0.1	10.2 ± 0.1	32.9 ± 0.2	1.60 ± 0.04	3.29 ± 0.05	9.0 ± 0.1
Observed	95	220	547	51	88	235

Table 4.6: Expected and observed numbers of events, for each lepton channel, in the signal region, before the final fit. Uncertainties on MC predictions are given by the total normalization error.

The data appear to follow the shape of the predicted background-only shape, with the first bin $(-1.0, -0.8)$, i.e. the most background-like, being the most populated. In this bin, the data show a slight overfluctuation over the MC prediction, partly covered by the shaded uncertainty band: the fit can accomodate this excess floating the background normalization. An overfluctuation can also be seen in the signal-like region for the $\mu\mu$ channel, compensated by a downward fluctuation in the ee channel. As shown in the figure, the BDT response is divided into ten bins, of which the rightmost one $(0.8, 1.0)$ is empty; in fact, it could be populated only by events that got classified as signal-like in almost 100% of the tree splittings. The WW background is evenly distributed across the ten bins, while the top-quark backgrounds are classified mostly into the first bin, which means that their topologies are easier to recognize for the BDT. The small fraction of $Z/\gamma^* + \text{jets}$ events which survives the selection is classified as signal-like; these events are, however, almost absent in the $e\mu$ channel, the most sensitive. This, on the other hand, is polluted by a significant contribution of $W + \text{jets}$ events, which are particularly tricky to handle since their BDT response has almost the same shape as the signal one, and a yield which is double in size. Only a few signal events have got a low BDT response ($\text{BDT} \leq -0.2$).

What has been described up to now, holds for low-mass signal hypotheses; since for each mass point a dedicated training is performed, the BDT response changes as a function of the mass, as illustrated in Fig. 4.21 for a few examples ($m_H =$

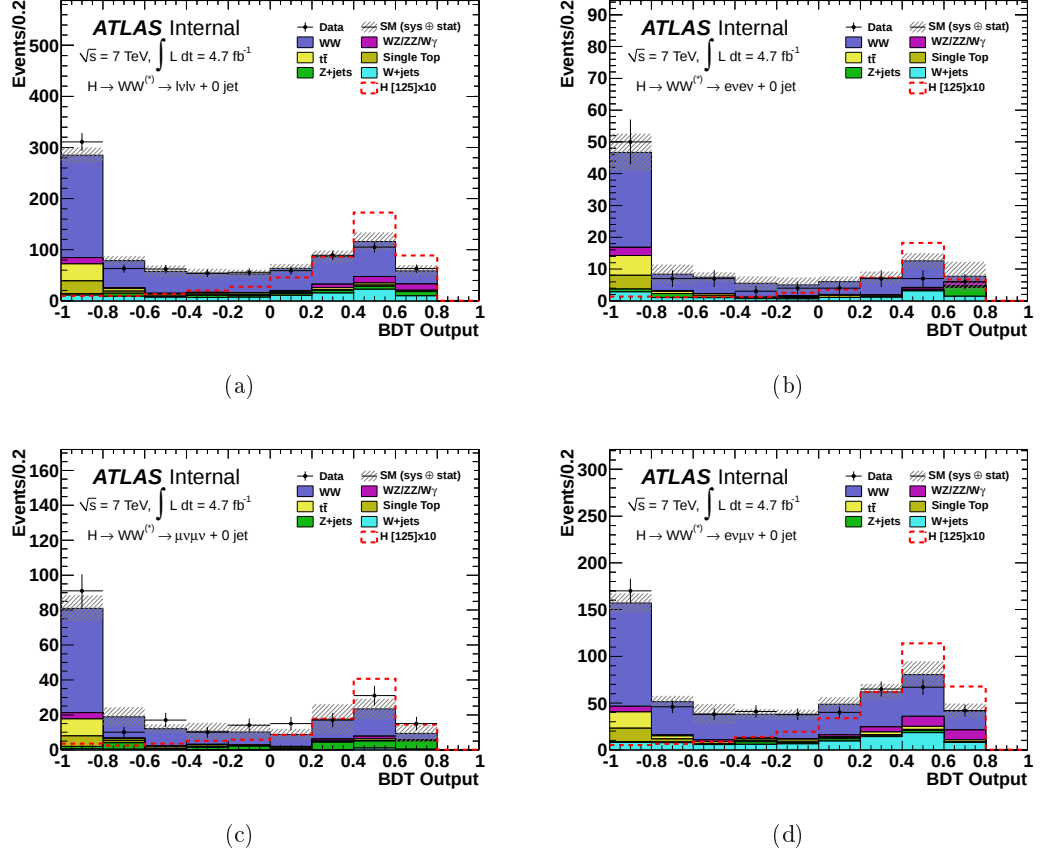


Figure 4.20: Distributions of the BDT response in the data and in the MC simulation, after all selection cuts, in the 0-jet channel. The plots show (a) $\ell\ell$, (b) ee , (c) $\mu\mu$, (d) $e\mu$. The signal contribution for $m_H = 125 \text{ GeV}$ has been multiplied by a factor of ten for better visibility. All backgrounds are normalized as in Table 4.5.

125, 400, 600 GeV). In these plots, the distributions are shown in five bins and the leftmost bin has been scaled down by 0.1 for better visibility, while for masses other than 125 GeV, the signal has been multiplied by ten, to cope with its decreasing rate. It can be recognized that, as the mass increases, it gets more and more difficult for the BDT to discriminate between the signal and the background hypotheses, since the boost of the WW system breaks the spin correlations between the leptons; for this reason, more and more signal events get classified as background-like.

4.7.4 Signal region in 1-jet

In the 1-jet channel, after the requirement for one identified jet, additional cuts are applied:

- b -jet veto: the identified jet cannot be b -tagged, i.e. its b -tag weight cannot

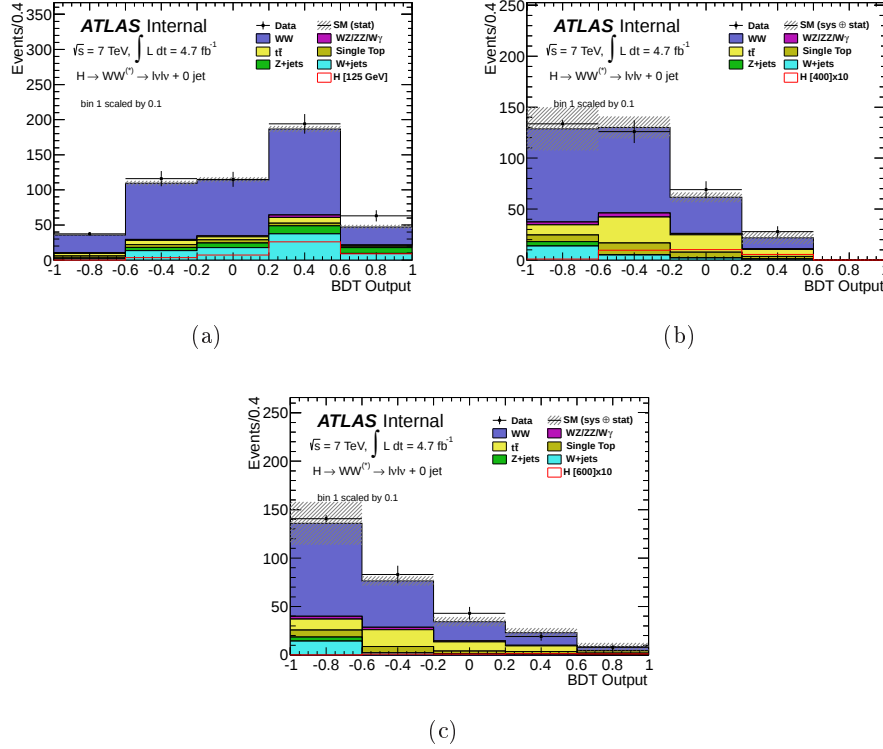


Figure 4.21: Distributions of the BDT response in real data and in MC simulation, in the signal region after all selection cuts, in the 0-jet channel, for three mass hypotheses: (a) $m_H = 125$ GeV, (b) $m_H = 400$ GeV, (c) $m_H = 600$ GeV. In the second and third cases, the signal contribution has been multiplied by a factor of ten for better visibility, while in all cases the first bin has been scaled down by 0.1. The difference in the background shape is caused by the fact that each mass point is trained independently.

exceed -1.25 (corresponding to a working point of 80% efficiency for b -jets identification in MC, see Section 4.4.6),

- $p_{T,TOT} < 30$ GeV,
- $Z \rightarrow \tau\tau$ veto: the event is discarded if $|m_{\tau\tau} - m_Z| < 25$ GeV,
- $m_{\ell\ell} < 80$ GeV if $m_H < 200$ GeV.

All cuts are in common with the cut-based selection, apart from the $m_{\ell\ell}$ threshold (cf. Table 4.4). The corresponding number of signal and background events can be seen in Table 4.5, at each stage of the selection, together with the number of the observed events in data. As expected, the b -tag veto rejects most of the top-quark backgrounds (namely 70%), thus halving the total background. Then, a cut on the

total transverse momentum of the event is applied, defined as:

$$\vec{p}_{T, \text{TOT}} = \vec{p}_{T, \ell_1} + \vec{p}_{T, \ell_2} + \vec{p}_{T, \text{jet}} + \vec{p}_{T, \text{miss}}$$

in order to veto the presence of additional soft jets, which may be present in the event but fail the jet identification criteria. Basically all backgrounds are reduced by this requirement, halving again the total yield; also the signal is affected. Afterwards, in order to explicitly remove the remaining $Z/\gamma^* + \text{jets}$ events, which at this point are mostly given by $Z \rightarrow \tau\tau$ decays, a $Z \rightarrow \tau\tau$ veto is applied, requiring the mass of the hypothetical $\tau\tau$ system not to be within 25 GeV from the m_Z mass. Under the hypotheses that the two leptons are τ decay products, that the two neutrinos are *collinear* with the leptons (collinear approximation [217]) and that there is no other source of missing transverse energy, the $\tau\tau$ invariant mass is computed as $m_{\tau\tau} = m_{\ell\ell}/\sqrt{x_1 x_2}$, where $x_{1,2}$ are the energy fractions carried by the visible decay products of the two τ leptons, and are given by $x_{1,2} = p_{\text{vis},1,2}/(p_{\text{vis},1} + p_{\text{vis},2})$ [218]. These energy fractions have to be in the range (0,1) for the veto to apply. This cut rejects around 20% of the Z/γ^* background. Finally, the same cut on the dilepton invariant mass as in the 0-jet analysis is applied.

At the end of the selection, the S/B ratio is 3.9%, while it used to be 0.8% after the 1-jet requirement; 14 signal events are expected for $m_H = 125$ GeV, together with 355 events for the total background; the relative uncertainty on the total background is $\pm 8\%$, larger than in the 0-jet case (where it is $\pm 4.6\%$). 374 events are observed in the data. As noticed already for the 0-jet result, most of the candidates belong to the $e\mu$ category, as can be seen in Table 4.6. Fig. 4.22 shows the distributions of the BDT response in the signal region for the 1-jet channel, for the individual lepton flavours and their combination. Shown are the various background components, the signal expectation for $m_H = 125$ GeV and the events found in the data. Five bins are used here, as done in the fit. The data and the MC predictions appear to be in good agreement. Contrarily to the 0-jet case, here the leftmost bin is found to have a ratio between the data and the MC prediction close to unity; on the other hand, a similar situation arises for the signal-like region, showing slight fluctuations, downwards for ee and upwards for $\mu\mu$.

Fig. 4.23 shows how the BDT response changes in shape when going to higher Higgs boson masses. As in the 0-jet case, increasing m_H causes the classified events to shift to the left side, with a worsening of the S/B discrimination.

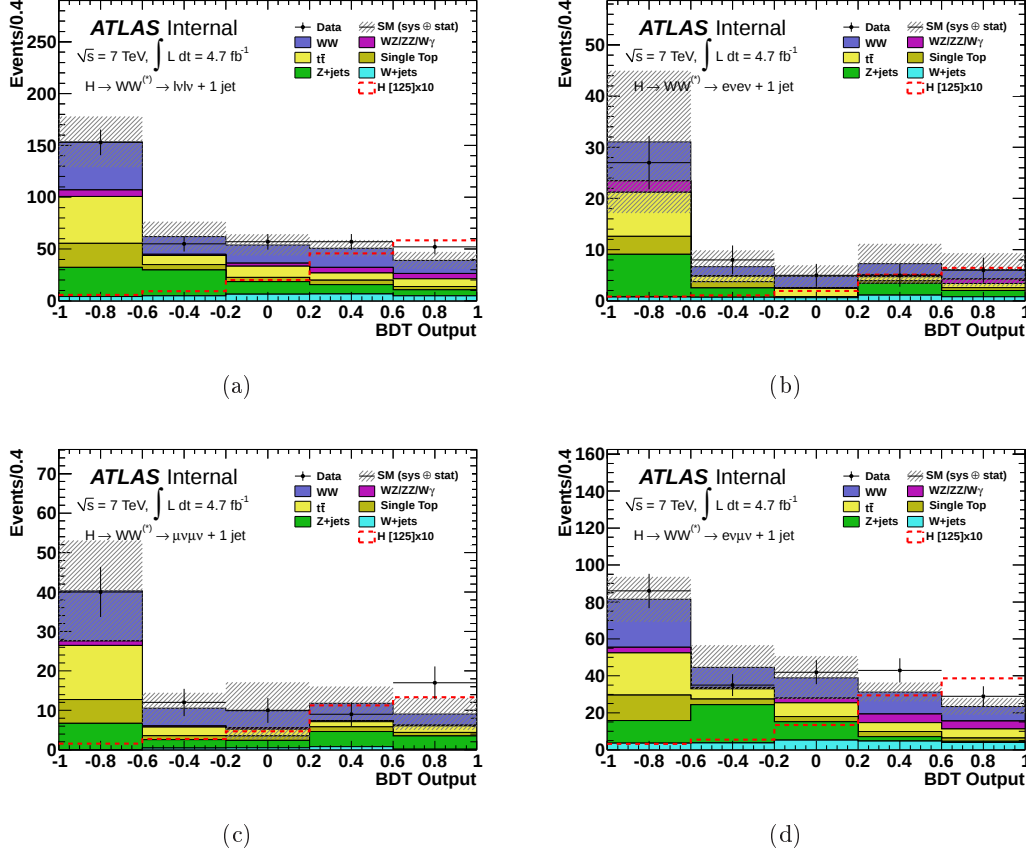


Figure 4.22: Distributions of the BDT response in the data and in the MC simulation, after all selection cuts, in the 1-jet channel. The plots show (a) $\ell\ell$, (b) ee , (c) $\mu\mu$, (d) $e\mu$. The signal contribution for $m_H = 125$ GeV has been multiplied by a factor of ten for better visibility. All backgrounds are normalized as in Table 4.5.

4.7.5 WW control region in 0-jet

In the 0-jet channel, the normalization of the WW background is fixed to the data by means of a dedicated control region. This is defined applying all cuts up to the last $m_{\ell\ell}$ requirement and inverting it⁽⁵⁾ ($m_{\ell\ell} > 80$ GeV).

As can be seen again in Table 4.5, the WW background dominates in its CR as expected, with some contributions from the top-quark backgrounds. The normalization factor, α_{WW} , is estimated as the ratio between WW_{data} and WW_{MC} , where the latter is simply the MC prediction, while the former is obtained sub-

⁽⁵⁾In the cut-based analysis, where it holds $m_{\ell\ell} < 50$ GeV for the signal region and $m_{\ell\ell} > 80$ GeV for the WW control region, the two regions are not adjacent, requiring a careful investigation of the theoretical uncertainties on the extrapolation factors. In this analysis, on the other hand, since cuts are relaxed (no $\Delta\phi_{\ell\ell}$ and no m_T cuts are used in the signal region), the two regions are defined as adjacent and no extrapolation uncertainty is considered.

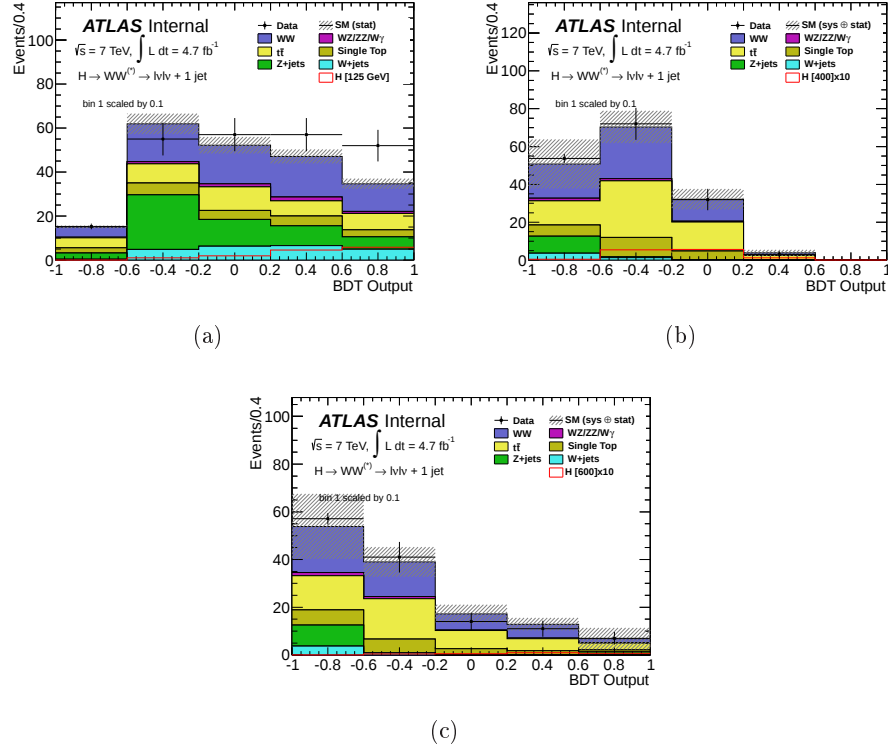


Figure 4.23: Distributions of the BDT response in real data and in MC simulation, in the 1-jet signal region after all selection cuts, for three mass hypotheses: (a) $m_H = 125$ GeV, (b) $m_H = 400$ GeV, (c) $m_H = 600$ GeV. In the second and third cases, the signal contribution has been multiplied by a factor of ten for better visibility, while in all cases the first bin has been scaled down by 0.1. The difference in the background shape is caused by the fact that each mass point is trained independently.

tracting from the data the sum of all backgrounds except WW itself: the result is $\alpha_{WW} = 1.03 \pm 0.06$, in agreement with the cut-based value of $\alpha_{WW}^{\text{c.b.}} = 1.05 \pm 0.06$. The normalization obtained in the fit to the BDT response, which can be found in Table I.1, is 1.01 ± 0.04 .

Even if the BDT response is actually fed into the limit setting machinery as a single bin, it is instructive to look at the shape of the BDT response also in the control regions, as shown in Fig. 4.24 for the individual lepton flavours and their combination. All events are squeezed to the leftmost bins, with no event beyond -0.8. This can be interpreted as a good performance of the BDT algorithm, which is able to correctly identify all events as very background-like, given their orthogonal topology with respect to the signal. For what concerns the high-mass analysis, as already explained, no signal-free control region can be identified, thus the WW contribution is taken directly from the MC prediction.

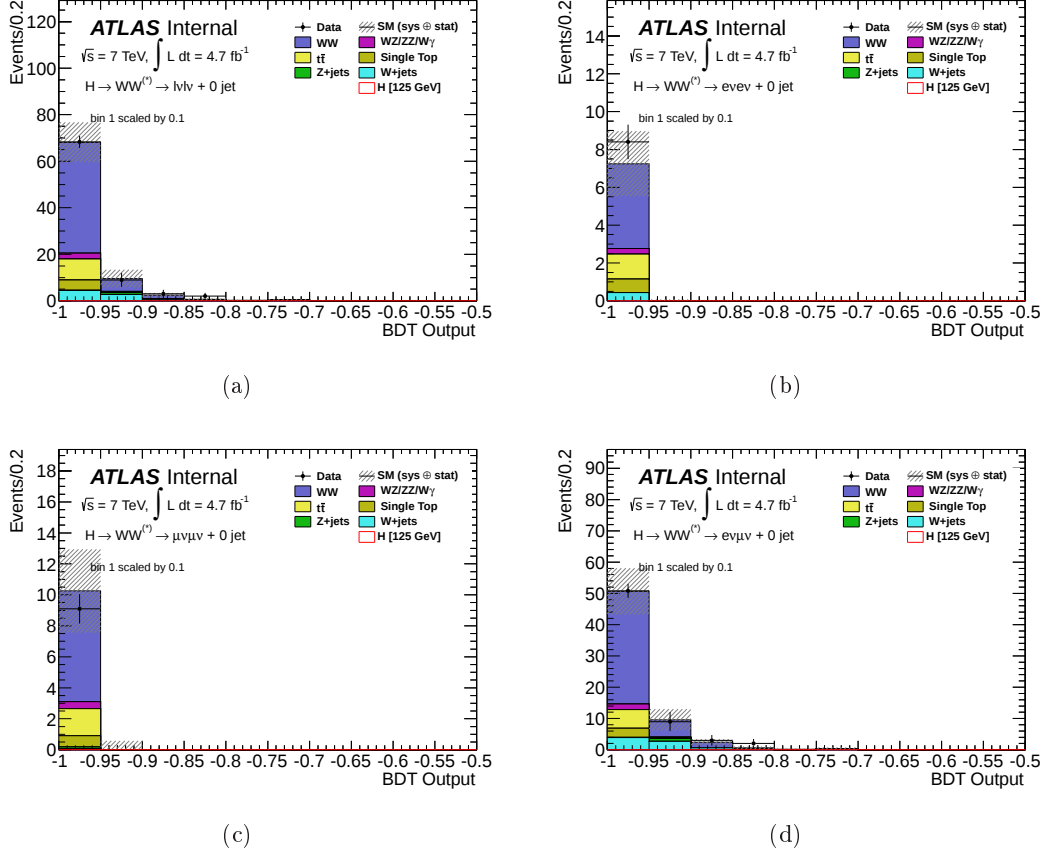


Figure 4.24: Distributions of the BDT response in the data and in the MC simulation, in the 0-jet WW control region. The plots show (a) $\ell\ell$, (b) ee , (c) $\mu\mu$, (d) $e\mu$. The signal contribution for $m_H = 125$ GeV has been multiplied by a factor of ten for better visibility. All backgrounds are normalized as in Table 4.5, except the WW one, which corresponds to the pure MC prediction.

4.7.6 WW control region in 1-jet

Similarly to the 0-jet channel, also in the 1-jet channel the WW background is normalized to the data by means of a dedicated control region, which is obtained by inverting the last $m_{\ell\ell}$ requirement. The pollution from the top-quark backgrounds is more severe in this case, such that the WW events account for less than half of the total (see Table 4.5). The WW contributions observed in the data and predicted by the MC simulation appear to be in very good agreement with each other (127 and 128 events, respectively), so that the α factor is compatible with 1 ($\alpha_{WW} = 0.99 \pm 0.15$), while the cut-based analysis finds $\alpha_{WW}^{\text{c.b.}} = 0.95 \pm 0.14$. The normalization obtained in the fit to the BDT response, which can be found in Table I.1, is 1.19 ± 0.22 . As for the 0-jet case, the BDT response, shown in Fig. 4.25, is squeezed to the left;

large shaded bands reflect the non negligible size of the total uncertainties. Finally, also in 1-jet no WW CR is used for the high-mass analysis.

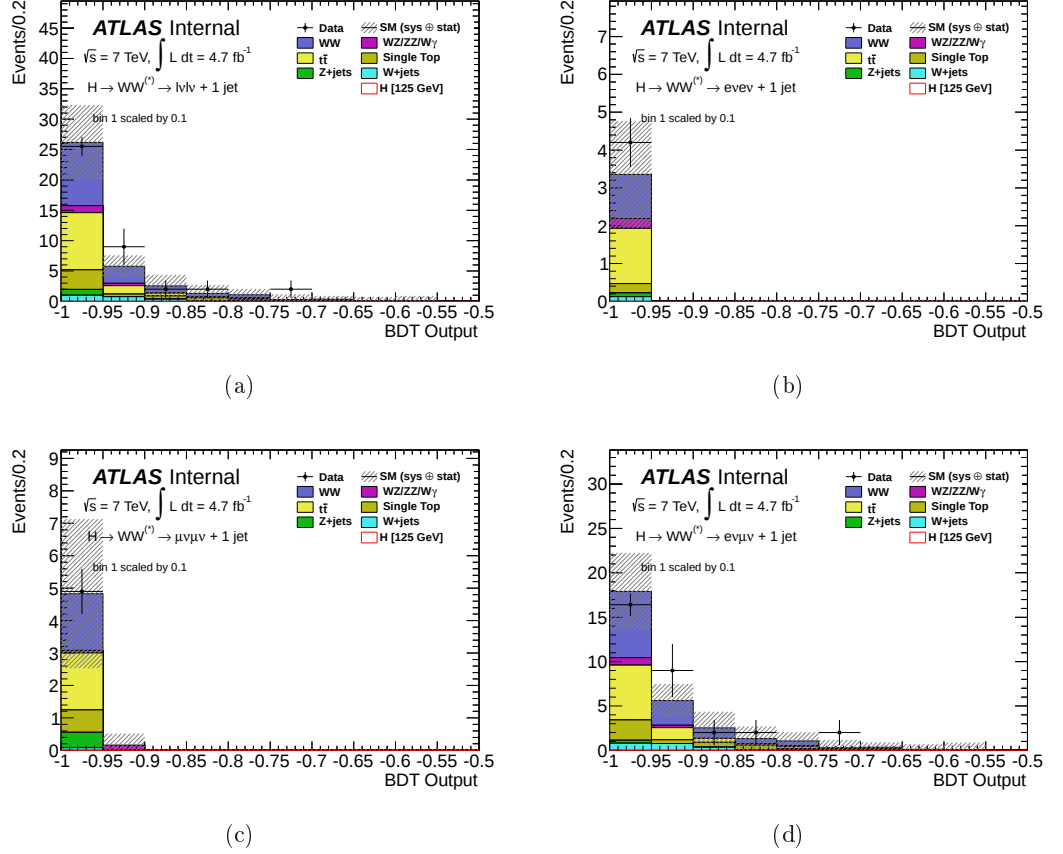


Figure 4.25: Distributions of the BDT response in the data and in the MC simulation, in the 1-jet WW control region. The plots show (a) $\ell\ell$, (b) ee , (c) $\mu\mu$, (d) $e\mu$. The signal contribution for $m_H = 125$ GeV has been multiplied by a factor of ten for better visibility. All backgrounds are normalized as in Table 4.5, except the WW one, which corresponds to the pure MC prediction.

4.7.7 Top-quark control region in 1-jet

For the 1-jet channel only, where it is most abundant, the top-quark background is normalized by means of a dedicated control region. All cuts are applied as for the signal region, apart from the b -tag veto which is reversed, i.e. the identified jet must be a b -jet; moreover, the final $m_{\ell\ell}$ cut, specific to the signal topology, is not applied. As shown in Table 4.5, a high purity is reached for this region; the contributions estimated in the data and in the MC agree closely with each other (603 and 602 events, respectively); the normalization obtained in the fit to the BDT response,

which can be found in Table I.1, is 1.03 ± 0.06 .

Fig. 4.26 shows the distributions of the BDT response in this region, in five bins and with the first bin scaled down by 0.1. Even if, in this case, the events are distributed across the whole range and not squeezed to the left as for the WW control region, the good agreement between the data and the MC expectation indicates that the background-like data events are being correctly classified by the algorithm.

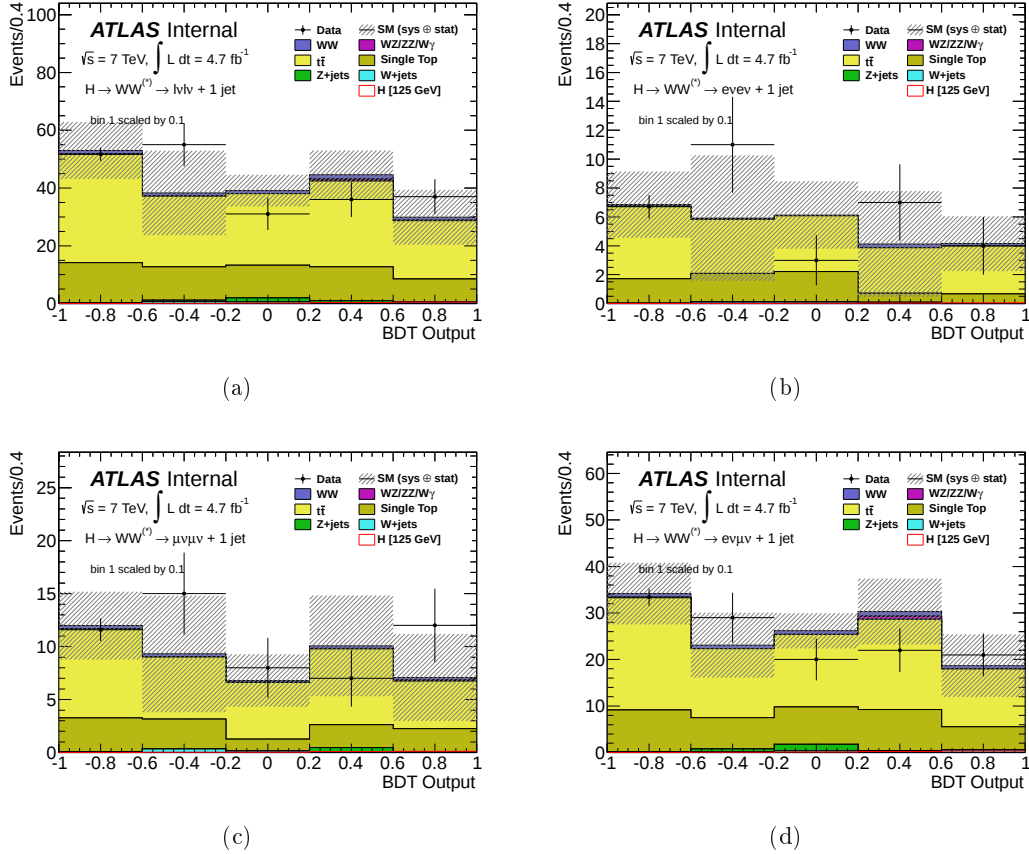


Figure 4.26: Distributions of the BDT response in the data and in the MC simulation, in the 1-jet top-quark control region. The plots show (a) $\ell\ell$, (b) ee , (c) $\mu\mu$, (d) $e\mu$. The signal contribution for $m_H = 125$ GeV has been multiplied by a factor of ten for better visibility. All backgrounds are normalized as in Table 4.5, except the top-quark backgrounds, which correspond to the pure MC prediction.

4.8 Background estimation

Whereas a dedicated control region is not present, the background estimation is a mixture of MC predictions and data-driven samples, following mostly the cut-based approach. In the following, the backgrounds for which a data-driven or pseudo data-driven technique is employed are described; the others rely on pure MC predictions.

4.8.1 Top-quark backgrounds in 0-jet

While in 1-jet the top-quark background is estimated in its control region, as discussed in Section 4.7.7, in 0-jet it is not possible to define a background-enriched region, since there is no identified jet to tag as b -jet. In fact, the residual top-quark background consists mostly of events whose b -jets have either failed to be reconstructed as such, or have not passed the p_T threshold. To correct the pure MC prediction, then, the estimation of the jet veto efficiency is needed. This is done scaling the number of $t\bar{t}$ events after the $E_{T,\text{rel}}^{\text{miss}}$ requirement (background-subtracted data) to their number in the 0-jet signal region. This scale factor is taken to be the square of the efficiency for one top-quark decay to satisfy the jet veto, efficiency which is estimated from another top-quark control sample enriched in b -jets, and which is corrected for the subleading contribution of single top-quark decays.

Top-quark events have to be rescaled through this jet veto efficiency by about 4%, the data-driven scale factor being $\alpha_{\text{top}}^{0\text{-jet}} = 1.04 \pm 0.07 \text{ (stat)} \pm 0.17 \text{ (sys)}$, combining the lepton channels together. This estimation is taken directly from the cut-based analysis [28] (the approximation is good enough given its uncertainty).

4.8.2 $Z/\gamma^* + \text{jets}$ background

The $Z/\gamma^* + \text{jets}$ background is estimated, both in the 0- and 1-jet channels, correcting the MC prediction by means of an E_T^{miss} *mismodelling factor*, accounting for the fact that this background can pollute the signal region only because of fake E_T^{miss} . The mismodelling factors are obtained comparing the data and the MC prediction in a two-dimensional plane $(m_{\ell\ell}, E_{T,\text{rel}}^{\text{miss}})$, where $m_{\ell\ell}$ is inside or outside a 10 GeV window from m_Z , for each lepton flavour separately. This method is similar but less complex than the $ABCD$ method employed by the cut-based analysis [28].

4.8.3 $W + \text{jets}$ background

The $W + \text{jets}$ contribution in the signal region derives from misidentification of jets as leptons, with the other lepton being the real one from the W decay. The misidentification rate needs to be estimated in data; to this extent, a data-driven technique

has been developed in the context of the cut-based analysis, as described in the following. The number of W +jets events in the SR is given by:

$$(W+jets)^{SR} = (W+jets)^{CR} \times \text{fake factor} = (W+jets)^{CR} \times \left(\frac{\text{identified leptons}}{\text{anti-identified leptons}} \right)^{\text{dijet}}$$

where $(W+jets)^{CR}$ is computed in a data control sample, while the fake factor uses an inclusive dijet sample. For the former, the control sample is defined as having one identified lepton, as in the main analysis, and one *anti-identified* lepton which satisfies only a loosened set of criteria, but fails the standard ones (as defined in Sections 4.4.4 and 4.4.5), as normally happens for a jet faking a lepton. The loosened criteria are:

- for electrons:
 - standard p_T and η requirements,
 - standard cut on the longitudinal impact parameter, $|z_0| < 1$ mm,
 - no cut on the transverse impact parameter significance, $|d_0/\sigma(d_0)|$,
 - standard track isolation $p_T^{\text{cone}}/p_T < 0.13$,
 - loosened calorimeter isolation $E_T^{\text{cone}}/p_T < 0.3$,
- for muons:
 - standard p_T , η and track quality requirements,
 - standard cut on the longitudinal impact parameter, $|z_0| < 1$ mm,
 - no cut on the transverse impact parameter significance, $|d_0/\sigma(d_0)|$,
 - no cut on track isolation,
 - loosened calorimeter isolation $E_T^{\text{cone}}/p_T < 0.3$.

Afterwards, the same selection criteria as in the cut-based signal region are applied (the BDT analysis relies on the cut-based estimation, whose uncertainty is large enough to cope with possible differences [28]).

The fake factor, on the other hand, is estimated using an inclusive dijet sample; it is given by the ratio between the number of fully identified and anti-identified lepton candidates, in bins of p_T of the anti-identified lepton. From the dijet sample, the contamination from W and Z decays is removed imposing specific cuts, while that from $W\gamma$ and $W\gamma^*$ is subtracted using the MC prediction.

Table 4.7 shows the number of W +jets in the signal region and in the control sample, for the individual lepton channels and jet numbers. It can be seen that the electron and muon fake factors, on average 0.01 and 0.1, respectively, differ

significantly; this is due, on one hand, to the different rate of jets to fake electrons and muons, on the other hand to the different probability for the jet to pass the anti-identification criteria for electrons and for muons.

Region	ee	$e\mu$	$\mu\mu$
H + 0 jet			
W+jets control	$86.0 \pm 9.8(\text{stat})$	$219 \pm 15.6(\text{stat})$	$5.74 \pm 2.65(\text{stat})$
Signal	$0.87 \pm 0.10(\text{stat}) \pm 0.10(\text{sys})$	$2.81 \pm 0.43(\text{stat}) \pm 1.22(\text{sys})$	$1.00 \pm 0.45(\text{stat}) \pm 0.45(\text{sys})$
H + 1 jet			
W+jets control	$15.10 \pm 4.24(\text{stat})$	$54.80 \pm 8.00(\text{stat})$	$0.70 \pm 1.00(\text{stat})$
Signal	$0.12 \pm 0.04(\text{stat}) \pm 0.05(\text{sys})$	$0.51 \pm 0.21(\text{stat}) \pm 0.17(\text{sys})$	$0.15 \pm 0.20(\text{stat}) \pm 0.02(\text{sys})$

Table 4.7: Number of W +jets events predicted in the W +jets control sample and in the cut-based signal region, for the individual lepton channels and jet bins [28]. The control sample is defined reverting the quality criteria for one of the leptons (anti-identification) as explained in the text.

This data-driven estimation also includes implicitly contributions from the multi-jet background, for which both leptons are faked by jets. For this reason, and since the isolation and quality criteria on the lepton identification make this background negligible, as noted above, this contribution is ignored in the signal region.

4.9 Systematic uncertainties

In general, the systematic uncertainties, both from experimental and theoretical sources, are taken into account in the analysis computing the relative change in the number of selected events that they induce; the change in normalization can be expressed as:

$$\delta_{\text{up}} = \frac{|N_{\text{nom}} - N_{\text{sys up}}|}{N_{\text{nom}}} \quad , \quad \delta_{\text{down}} = \frac{|N_{\text{nom}} - N_{\text{sys down}}|}{N_{\text{nom}}} \quad (4.7)$$

where the label “nom” refers to the nominal yield, while “sys up” and “down” are the variations which are typically obtained shifting the systematic upwards and downwards of 1σ , respectively. For each variation, the BDT classification is performed again, using the nominal training. The normalization uncertainties are used as nuisance parameters in the limit extraction procedure.

A few systematic uncertainties induce also a substantial change in shape in the BDT response; for these, the shape variation is taken into account besides the normalization one.

4.9.1 Experimental uncertainties

Systematic uncertainties are associated with every step of the lepton reconstruction and identification, with the missing E_T measurement, with the jet calibration and tagging procedures, with the trigger and the luminosity estimation.

The efficiency-related systematic uncertainties, such as b -tagging efficiency, induce only a change in the MC event weight; in this case, the nominal BDT response is reweighted up and down, without reclassifying the events, and the corresponding systematic is treated as a shape uncertainty.

Table 4.10 summarizes the treatment of experimental systematic uncertainties, indicating also if a shape component is present. Additional details on jet-, lepton- and missing E_T -related systematic uncertainties are given in the following. The names in parentheses correspond to the names of the nuisance parameters as presented in Table 4.10.

Moreover, an uncertainty of $\pm 3.9\%$ is assigned to the integrated luminosity ($\pm 3.9\%$ is the standard value for ATLAS analyses before the publication of the final recommendation for the $\sqrt{s} = 7$ TeV dataset, which is $\pm 3.7\%$ [215]).

Jet systematic uncertainties

The main experimental systematic uncertainty is the *jet energy scale* (JES) one. As briefly discussed in Section 3.3.5, jets are reconstructed at the EM scale, as measured by the calorimeters. The EM scale energy does not account for calorimeter non-compensation, leakage, dead material, inefficiencies. The goal of the JES calibration [180] is to correct the jet energies and momenta, using as a reference the kinematics of the corresponding Monte Carlo truth jets, in bins of p_T and η . Uncertainties on, e.g., the dead material budget, the noise description and the beam spot position form the corresponding JES uncertainty, which varies from $\pm 14\%$ to $\pm 2\%$ as a function of jet transverse momenta and pseudorapidities, for jets with $p_T > 25$ GeV and $|\eta| < 4.5$. In particular, it is at most $\pm 4\%$ for central jets. Pile-up effects have to be taken into account separately; their contribution is less than $\pm 5\%$ in all bins.

Also the uncertainty on the *jet energy resolution* (JER), $\sigma(p_T)/p_T$, has to be taken into account; it is obtained via in-situ techniques and is found to vary between $\pm 5\%$ and $\pm 2\%$ [181].

Finally, a shape systematic is added on the b -tagging procedure [219], with variations with respect to the nominal case ranging between $\pm 5\%$ and $\pm 14\%$. Besides the b -tagging efficiency (B_EFF), the same algorithm provides also uncertainties for c -tagging (C_EFF) and for misidentification of light jets (L_EFF); both are treated as shape systematic uncertainties.

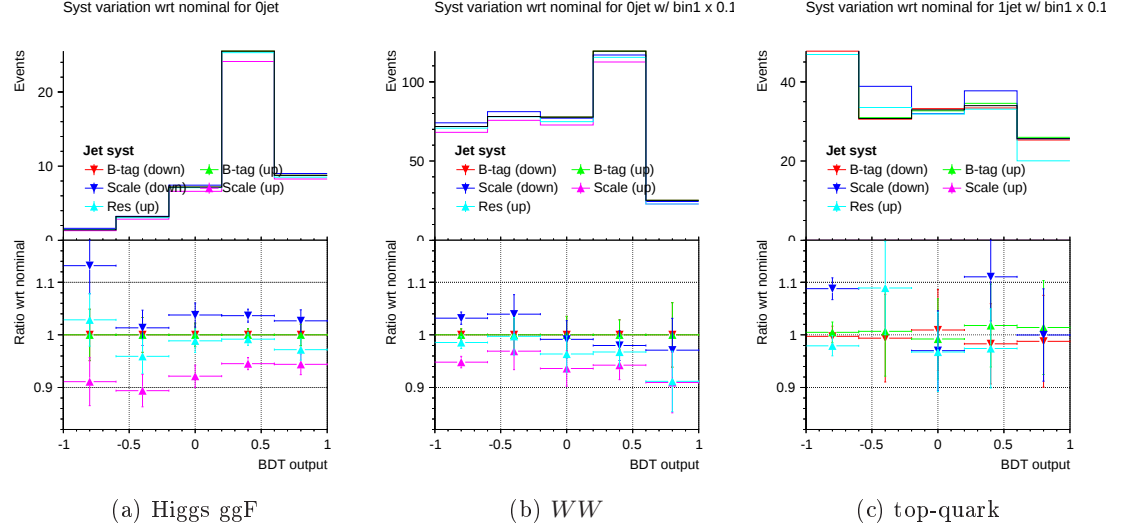


Figure 4.27: Variations in the BDT response shape due to the jet-related systematic uncertainties (JES, JER, b -tagging efficiency). The samples shown are: (a) the Higgs boson signal (gluon-fusion production mode) in 0-jet, (b) the WW background in 0-jet, (c) the top-quark backgrounds in 1-jet. In the latter two cases, the first bin is scaled down by a factor 0.1 for better visibility. In each plot, the top frame shows the overlay of the nominal shape with the systematic variations, while the bottom frame shows the ratio with respect to the nominal.

Fig. 4.27 shows the impact on the BDT response of the various jet systematic uncertainties, for the ggF Higgs boson signal and the WW and top-quark backgrounds. For the other backgrounds, given the limited number of MC events, variations are not significant. For the signal and the WW background, plots are shown in the 0-jet channel, while for the top-quark backgrounds in the 1-jet channel; in all cases, the lepton flavours have been combined. It can be seen that the largest variations with respect to the nominal shape are due to JES, especially for the top-quark samples.

Lepton systematic uncertainties

The lepton systematic uncertainties are related to the uncertainties on their selection efficiency, energy scale and resolution.

For electrons, the selection efficiency includes multiple effects, as discussed in Section 3.3.3. Contributions from identification, reconstruction and isolation are added in quadrature [220,221]; the identification has an uncertainty up to $\pm 8\%$ in the low- p_T range, but within $\pm 1\%$ for $p_T > 30$ GeV, while the reconstruction and isolation uncertainties are less than $\pm 1\%$ for $p_T > 15$ GeV. All together, including also a trigger uncertainty of about $\pm 1\%$, the electron efficiency uncertainty (E_EFF)

ranges between $\pm 2\%$ and $\pm 5\%$, as a function of p_T and η . The electron energy scale (E_SCALE) and resolution (E_RES) systematic uncertainties are negligible, being smaller than $\pm 1\%$ each.

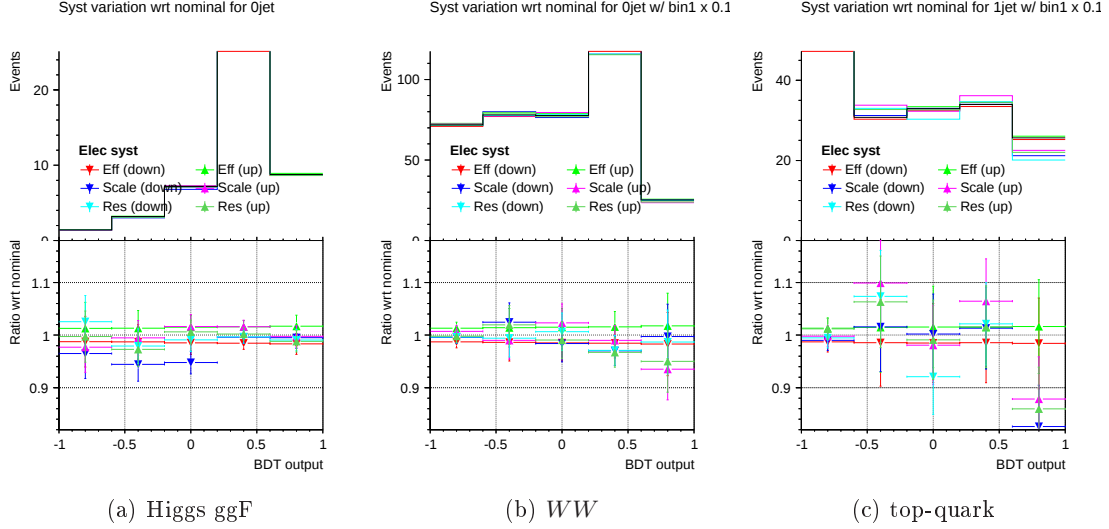


Figure 4.28: Variations in the BDT response shape due to the electron-related systematic uncertainties (efficiency, energy scale, energy resolution). The samples shown are: (a) the Higgs boson signal (gluon-fusion production mode) in 0-jet, (b) the WW background in 0-jet, (c) the top-quark backgrounds in 1-jet. In the latter two cases, the first bin is scaled down by a factor 0.1 for better visibility. In each plot, the top frame shows the overlay of the nominal shape with the systematic variations, while the bottom frame shows the ratio with respect to the nominal.

For muons [222,223], the same considerations bring to a selection efficiency systematic (M_EFF) within $\pm 1\%$ on the whole p_T and η ranges, and to a momentum scale (M_SCALE) and resolution systematic uncertainties also smaller than $\pm 1\%$. Concerning resolution, two independent effects have been implemented, one accounting for track resolution in the Inner Detector (ID_RES), the other in the Muon Spectrometer (MS_RES).

Figs. 4.28 and 4.29 show the impact on the BDT response of the various electron and muon systematic uncertainties, respectively, for the Higgs boson signal (in 0-jet) and the WW (in 0-jet) and top-quark (in 1-jet) backgrounds, with the lepton flavours combined. It can be seen that the induced distortion of the BDT shape is very small for the signal sample and the WW background, and inside 10% for the top-quark backgrounds.

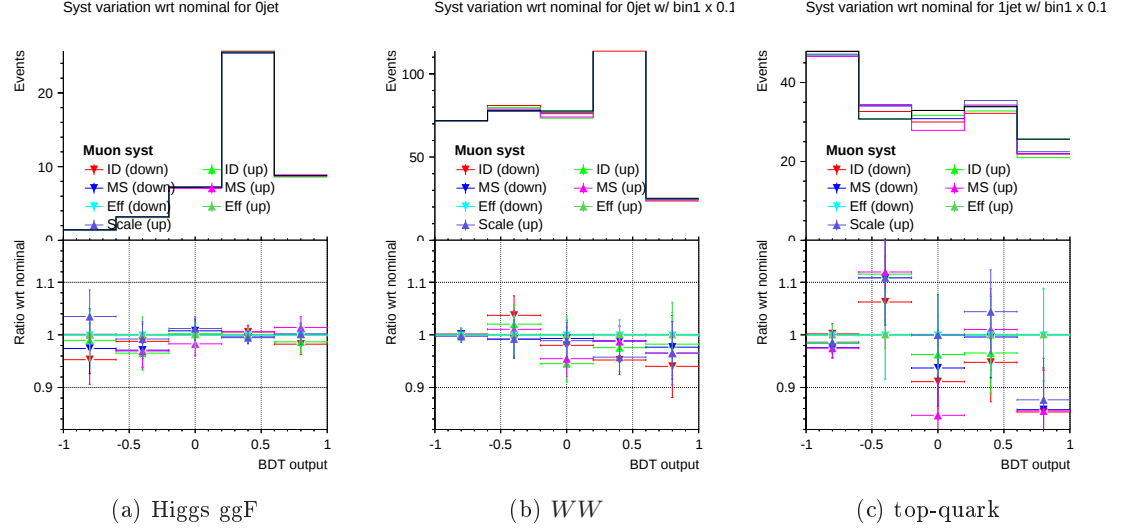


Figure 4.29: Variations in the BDT response shape due to muon-related systematic uncertainties (efficiency, energy scale, ID resolution, MS resolution). The samples shown are: (a) the Higgs boson signal (gluon-fusion production mode) in 0-jet, (b) the WW background in 0-jet, (c) the top-quark backgrounds in 1-jet. In the latter two cases, the first bin is scaled down by a factor 0.1 for better visibility. In each plot, the top frame shows the overlay of the nominal shape with the systematic variations, while the bottom frame shows the ratio with respect to the nominal.

Missing E_T systematic uncertainties

The main systematic uncertainties on the missing transverse energy (the CLUSTER term) include the propagation of the lepton and jet energy scale uncertainties, and the effects of sub-threshold jets and of soft calorimeter depositions not associated with reconstructed physics objects [224]. The variation induced on the total background yield goes from $\pm 1\%$ to $\pm 8\%$. Pile-up effects on missing E_T are taken into account separately (PILEUP), varying the pile-up modelling; the impact on the background yield ranges in this case between $\pm 1\%$ and $\pm 5\%$.

Fig. 4.30 shows the effect on the BDT response of the missing E_T systematic uncertainties, for the Higgs boson signal (in 0-jet) and the WW (in 0-jet) and top-quark (in 1-jet) backgrounds, with the lepton flavours combined. As before, the induced distortions of the BDT shape are very small for the signal sample and the WW background, larger for the top-quark backgrounds.

4.9.2 Theoretical uncertainties

The main sources of theoretical systematic uncertainties include underlying event (UE) modelling, initial and final state radiation (ISR/FSR), renormalization and

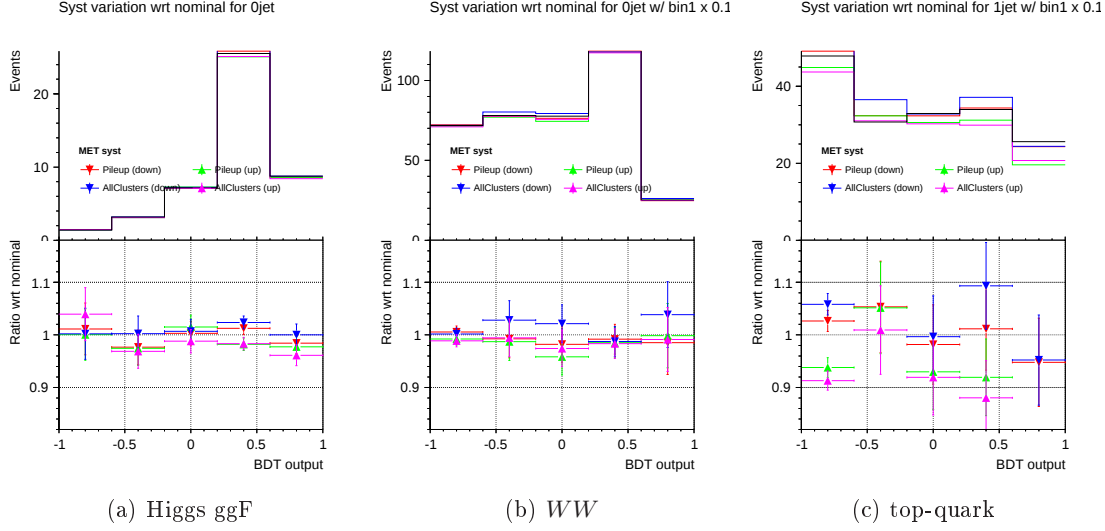


Figure 4.30: Variations in the BDT response shape due to the missing E_T -related systematic uncertainties (the cluster and pile-up terms). The samples shown are: (a) the Higgs boson signal (gluon-fusion production mode) in 0-jet, (b) the WW background in 0-jet, (c) the top-quark backgrounds in 1-jet. In the latter two cases, the first bin is scaled down by a factor 0.1 for better visibility. In each plot, the top frame shows the overlay of the nominal shape with the systematic variations, while the bottom frame shows the ratio with respect to the nominal.

factorization scales, MC modelling and PDFs. They are listed in Table 4.10 after the experimental ones.

For what concerns the uncertainties on the signal production cross section, they are determined following the prescriptions in Refs. [87,225]. The main contribution is obtained varying the QCD renormalization and factorization scales up and down by a factor of two ($\mu_R, \mu_F = m_H/2, 2m_H$). Independent results are obtained for the uncertainty on the ggF cross section with 0, ≥ 1 or ≥ 2 jets in the final state: as summarized in Table 4.8, for the 0-jet channel uncertainties range from $\pm 25\%$ at low and intermediate masses, to $\pm 47\%$ at high masses. For the 2-jet cut-based analysis, the impact of the QCD scale variations is computed also on the VBF cross section, resulting in an additional $\pm 5\%$ term over the whole mass range; this also includes the impact of the uncertainty on the third jet veto requirement.

For masses above 300 GeV, where the width of the Higgs boson starts to be larger than 10 GeV, the uncertainty on the description of the Higgs boson lineshape, as implemented in the POWHEG generator, is also taken into account. It is added in quadrature for both the ggF and VBF processes, and amounts to $150\% \times (m_H/1 \text{ TeV})^3$.

The profile of the Higgs boson p_T spectrum is reweighted from the POWHEG prediction to the HqT one, which includes a NLO fixed order calculation and NNLL

	$m_H = 125 \text{ GeV}, m_H = 240 \text{ GeV}$	$m_H = 600 \text{ GeV}$
QCD scales μ_R, μ_F		
0-jet analysis	$\pm 25\%$	$\pm 47\%$
1-jet analysis	$\pm 37\%$	$\pm 43\%$
2-jet analysis (ggF)	$\pm 25\%$	$\pm 25\%$
2-jet analysis (VBF)	$\pm 5\%$	$\pm 6\%$
Higgs boson lineshape	$150\% \times (m_H/1 \text{ TeV})^3$	
PDF uncertainties		
ggF	$\pm 8\%$	$\pm 8\%$
VBF	$\pm 2\%$	$\pm 4\%$

Table 4.8: Theoretical relative uncertainties on the MC yields for the ggF and VBF signal processes, at low, intermediate and high Higgs boson masses [28].

resummation. The reweighting needs to be taken into account because it is expected to change the lepton kinematics and the number of jets.

The PDF uncertainties are estimated by the variation associated with their error sets, added to the variation across different sets, independently for the gluon-gluon, gluon-quark and quark-quark initiated processes; the relative uncertainty for the ggF term amounts to $\pm 8\%$, while for VBF no extra term is required.

For what concerns the MC modelling, the accuracy of the signal prediction is checked comparing the baseline generator, **POWHEG**, to an alternative one, namely **MC@NLO**; the corresponding systematic uncertainty is labelled as “UEPS” (see Table 4.10) since it includes the effects of different models for the underlying event (UE) and the parton shower (PS). Also the uncertainty on the modelling of the background processes is estimated by alternative generators: **ALPGEN** for the WW production, **POWHEG** for $t\bar{t}$, **PYTHIA** for the $Z/\gamma^* + 2$ jets process. All uncertainties for $t\bar{t}$ are assumed to be the same also for the subleading single top-quark component.

4.9.3 Uncertainties from data-driven methods and control regions

Some additional systematic uncertainties arise from the usage of data-driven methods, pseudo data-driven methods and control regions. They affect, respectively, the W +jets, Z/γ^* +jets, WW and $t\bar{t}$ backgrounds.

The uncertainty on the data-driven estimation of the W +jets background is rather consistent, accounting for $\pm 50\%$ in the signal region. Its components include the difference in jet composition between the dijet sample used to estimate the fake factor and the actual W +jets background, the impact of the limited number of events

in the data sample used to define the corresponding control region ($(W+\text{jets})^{\text{CR}}$, see Section 4.8.3) and the contamination in real leptons from W and Z decays. A pair of systematic ntuples is obtained shifting the fake rate upwards and downwards (of a quantity which has the loose meaning of “ 1σ ”); after running the BDT classification, their difference to the nominal sample is used as a shape uncertainty on the BDT response (Fake_Rate). The size of this systematic uncertainty limits the accuracy of the whole analysis, especially given the fact that, at the end of the selection, the $W+\text{jets}$ yield is double the signal yield (see Table 4.5), with a very similar shape.

Concerning the $Z/\gamma^*+\text{jets}$ background, only a statistical uncertainty is assigned on the extraction of the mismodelling scale factors (Z_SCALEF), for each channel independently (see Appendix I).

The WW and $t\bar{t}$ backgrounds are obtained from control regions (for WW , only if $m_H < 200$ GeV), as explained in Section 4.7; the theoretical error on the $\alpha_{WW,t\bar{t}}$ factors has to be evaluated, while the experimental one is negligible. For WW , some of the diagrams and off-shell spin correlations are missing from the **MCONLO** implementation, while they are included in the **MCFM** one. Small differences between the two computations result in different lepton kinematics ($m_{\ell\ell}$ in particular), thus in differences in α_{WW} , yielding $\alpha(\text{MCONLO})/\alpha(\text{MCFM}) = 0.980 \pm 0.015$, with the error being purely statistical; a conservative estimate of $\pm 3.5\%$ is used in the analysis. Further uncertainties on α are obtained using different PDF sets and varying the QCD scales, both for the main qq, qg initiated processes and the gg component; the results are summarized in Table 4.9.

The total pre-fit⁽⁶⁾ uncertainties on the WW background in the signal region amount to $\pm 9\%$ in 0-jets, dominated by the statistical term, and to $\pm 22\%$ in 1-jet, dominated by the contamination of $t\bar{t}$ events in the WW CR.

Finally, the total uncertainties on the $t\bar{t}$ background in the signal region are larger, being $\pm 23\%$ and $\pm 40\%$ in the 1- and 2-jets channels, respectively. They are driven by the uncertainty on the b -tagging efficiency, used to define both the signal and the control region. For the 0-jet bin, a $\pm 22\%$ error is assigned to the $t\bar{t}$ estimation.

4.10 Statistical model and fit procedure

At the end of the selection, as detailed in Section 4.7, the number of events in real data has to be compared with the sum of the signal and background events, in order to establish limits on $\sigma_{\text{meas}}/\sigma_{\text{SM}}(m_H)$, namely the measured cross section for the

⁽⁶⁾Here and in the following, “pre-fit” refers to values computed before performing the fit, while “post-fit” to the final results, taking into account all systematic uncertainties.

	QCD scales	PDFs	Modelling
qq, qg, α_{WW}^{0j}	2.5%	3.7%	3.5%
qq, qg, α_{WW}^{1j}	4%	2.9%	3.5%
gg, α_{WW}^{0j}	6%	4.4%	—
gg, α_{WW}^{1j}	9%	4.6%	—

Table 4.9: Uncertainties on α_{WW} from various sources: QCD scales ($\mu_R, \mu_F = m_H/2, 2m_H$), different PDF prescriptions and MC modelling. The relative uncertainties are reported for both the main qq, qg initiated processes and the gg component.

production of a Higgs boson of mass m_H , in units of the corresponding theoretical SM cross section.

To exploit the full power of the BDT technique, the limit extraction takes advantage of the classification into signal-like and background-like events, through a shape fit to the BDT response. As anticipated in the previous Sections, in the low-mass case shapes are binned into ten and five bins, for the 0- and 1-jet channels, respectively, while for high-mass 20 bins are used and a remapping is performed, as discussed in Section 4.10.1. For comparison, the cut-based analysis uses five, three and one bin, for 0-, 1- and 2-jet channels, respectively. Normalizations from control regions are always estimated via a single-bin fit.

All systematic uncertainties — due to the limited number of MC events, to the experimental effects and to the theoretical approximations — are taken into account in the form of nuisance parameters. The fit is performed into 18 categories: three lepton flavours ($ee, \mu\mu, e\mu$), three jet channels (0, 1, 2 or more jets), two data-taking periods (BK, LM). The 2-jet likelihood function is taken from the cut-based analysis and multiplied to the 0-jet and 1-jet BDT ones to get the final combined result. The likelihood is built using the HistFactory framework [226], implemented in ROOT, and the final results are extracted by means of the common RooStats and RooFit tools [227].

4.10.1 Remapping

Since, as shown later, the likelihood function is built as a product on independent bins, the latter can be resorted without any bias to the final result. It is useful to *remap* the BDT response, in a way that empty bins are removed, in order to improve

the stability of the fit convergence. In particular, the low number of MC events can generate a troublesome bin looking like:

$$S \neq 0, B = 0, \text{data} \neq 0$$

which would induce a spuriously large result for the likelihood. Too pronounced bin-to-bin fluctuations, moreover, can cause the fit to be highly unstable. A solution to instabilities resides in dynamically changing the bin size in such a way that the total background is uniformly distributed across the bins. This reduces the impact of the MC statistical fluctuations, without losing the signal shape information. However, this is done in the cut-based analysis, but not in the BDT analysis. As shown already, at low-mass the BDT response is distributed between -1 and ≈ 0.8 , without discontinuities (on the contrary, m_T , the variable fitted in the cut-based analysis, peaks close to m_H and then tapers off in the tails), so there are no severe fluctuation problems to cure. Moreover, given the fact that the BDT algorithm classifies events as signal-like or background-like, and one is interested in the shapes of both the signal and the background, remapping the background to be flat means losing this essential information. For these reasons, the final version of the BDT analysis for $\sqrt{s} = 7$ TeV data does not use remapping at low-mass. Fig. 4.31 shows a few typical examples of the BDT response, in various channels and jet categories, with (a,b) and without (c,d) the remapping applied. The corresponding improvement in the CL_s limit on $\sigma_{\text{meas}}/\sigma_{\text{SM}}$ (see Section 4.10.4) is on average between 10% and 20% for increasing m_H , when dropping the remapping procedure.

In the high-mass range, on the other hand, the BDT response shrinks more and more to the left side as m_H increases; this can create a tail on the right side, giving rise to a similar situation to the one described for m_T . In particular, a discontinuity in the expected exclusion limit is found for $m_H = 380$ GeV, as a result of MC statistical fluctuations. To avoid this, a remapping is performed for $m_H > 200$ GeV, such that the first, leftmost bin contains exactly half of the total background events, while the remaining half is evenly distributed across 19 bins, up to $\text{BDT} = 1$. The signal events and the real data, on the other hand, are left free to adjust themselves on the so obtained bin pattern. The effect of this procedure can be seen in Fig. 4.32, where the BDT response for the $e\mu$ 1-jet channel (the source of the above mentioned discontinuity) is shown before and after remapping. Having 20 remapped bins in the high-mass signal region, gives to the analysis stability and improved sensitivity at the same time.

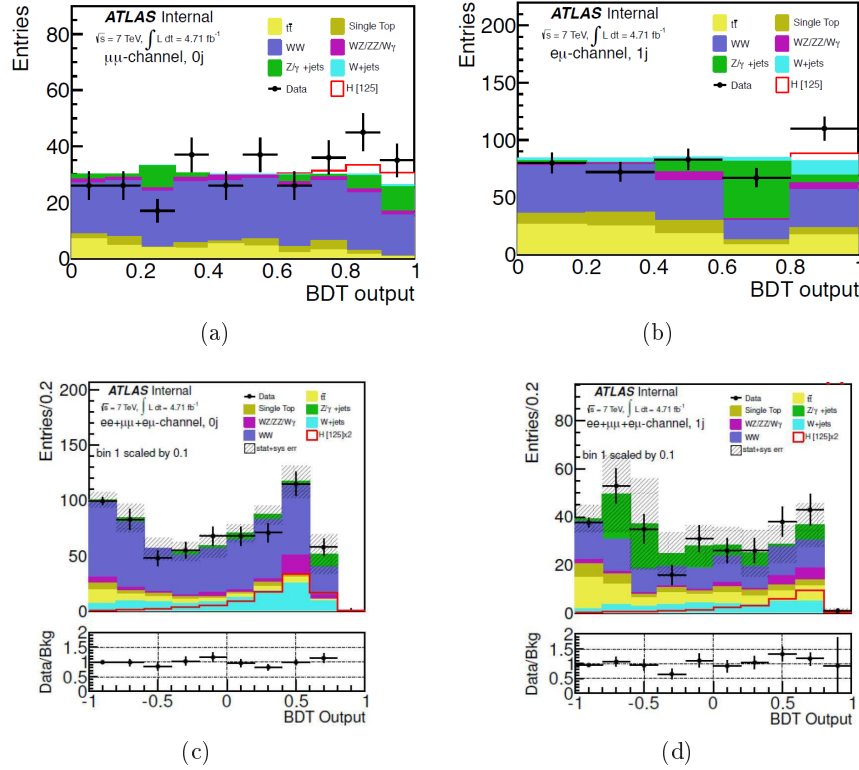


Figure 4.31: Distributions of the BDT response, in various lepton channels and jet numbers, before (a,b) and after (c,d) dropping the remapping procedure. The latter is designed in a way that the total background is evenly distributed across the bins, while the signal and the data are left free to adjust themselves on the so-obtained bin pattern. While, with the remapping applied, the signal is constrained to cover only a few bins, after dropping it, the signal spans a wider range and its shape can be fully exploited.

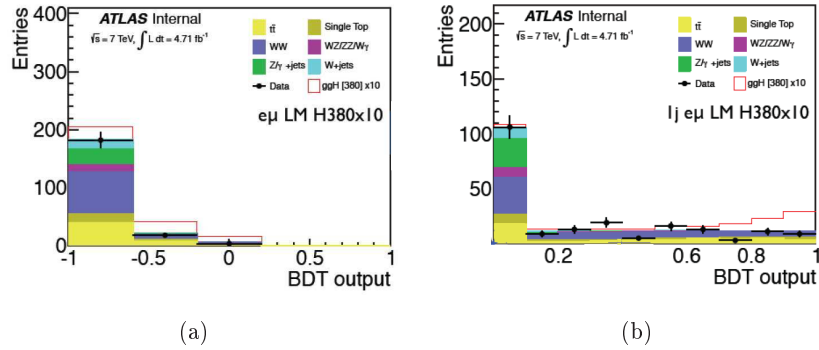


Figure 4.32: Distributions of the BDT response in the $e\mu$ 1-jet channel (period LM only), before (a) and after (b) the high-mass remapping strategy described in Section 4.10.1.

4.10.2 Construction of the likelihood function

The likelihood function is the basic building block of any statistical analysis, since it contains all analysis details in a compact form. Its basic form is the following:

$$\mathcal{L}(\mu) = P(N | \mu S + B) \quad (4.8)$$

where N is the observed number of events, while S and B are the expected numbers of signal and background events, respectively. The signal strength is given by μ , which is the *parameter of interest* (POI), to be fitted in a way that $\hat{\mu} = (N - B)/S$ (the hat on a variable indicates its best fit value). The case $\mu = 0$ corresponds to the *null hypothesis*, while $\mu = 1$ to the nominal signal hypothesis. In this simple model, the likelihood is just given by the Poisson probability as a function of μ .

Often, as in the case of the $H \rightarrow WW^{(*)}$ analysis, additional measurements are provided on the various background components, namely as control regions which help constraining their normalizations. The likelihood incorporates this information multiplying the Poissonian term for the SR with a Poissonian term for the CR [228, 229]:

$$\mathcal{L}(\mu, \theta) = P(N | \mu S + \theta B) P(N_{\text{CR}} | \theta B_{\text{CR}}). \quad (4.9)$$

Here the *nuisance parameter* (NP) θ has been introduced, which serves to fit the *target* background in its dedicated control region ($\hat{\theta} = N_{\text{CR}}/B_{\text{CR}}$), and to relate the corresponding measurements between SR and CR. On the contrary, all the information concerning μ is contained in the first Poissonian term. The α factors mentioned previously for the CR normalizations, e.g. α_{WW} , are defined as $\alpha = B/B_{\text{CR}}$. The auxiliary constraint can be written compactly as $F(\tilde{\theta}, \theta)$, with $\tilde{\theta}$ being an auxiliary measurement; in this analysis, it is taken to be either a unit Gaussian or a Poissonian.

Since the true values of the POI and the NPs are unknown, they are approximated as the values they take in correspondence with the maximum of the likelihood (*maximum likelihood estimator*, MLE).

In this analysis, many single-likelihood terms as in Eq. 4.9 are multiplied together: one for each bin i of the BDT distribution, one for each jet category j , one for each lepton flavour k and one for each data-taking period l , resulting in [228, 229]:

$$\mathcal{L}(\mu, \theta) = \left\{ \prod_{l=BK, LM} \prod_{k=ee, e\mu, \mu\mu} \prod_{j=0}^{N_{\text{jets}}} \prod_{i=1}^{N_{\text{bins}}} P(N_{ijkl} | \mu S_{ijkl} + \sum_b B_{ijklb}) \right\} \times \prod_{m=1}^{N_{\theta}} F(\tilde{\theta}_m, \theta_m). \quad (4.10)$$

The sum of the B_{ijklb} terms includes all background components; to better disentangle the contribution of control regions, the likelihood in braces can be factorized

as $\mathcal{L}^{\text{SR}}(\mu, \theta) \times \prod P(N_i^{\text{CR}}|\mu, \theta)$, where the product runs over all control regions, each having its own Poissonian term.

4.10.3 Normalization, shape and statistical uncertainties

The predictions for the numbers of signal and background events are functions of the nuisance parameters in a factorized way, i.e. for the signal (and similarly for the background) $S = S_0 \times \nu(\theta)$, where S_0 is the nominal expected rate, and the form of $\nu(\theta)$ depends on the systematic uncertainty. As mentioned already throughout the text, three kinds of systematic uncertainties are used in the analysis: normalization, shape and statistical uncertainties [228, 229].

- Normalization uncertainties are flat, since they change only the overall yield of a distribution. The constraint on θ is a unit Gaussian.
- Shape systematic uncertainties on a distribution affect the content of each bin individually. If they have a normalization component, by construction this is treated separately, so that the shape component does not have any effect on the overall rate; the two components, however, share the same θ , i.e. $\nu_{\text{norm}}(\theta)$ and $\nu_{\text{shape}}(\theta)$. The associated nuisance parameter is distributed as $\nu(\theta) = 1 + \epsilon\theta$, with a constraint such that $\nu(\theta) > 0$, θ being distributed as a truncated Gaussian.
- The statistical uncertainty, due to the limited number of MC events, is described by the Poissonian term $P(\tilde{\theta}|\theta\lambda) = (\theta\lambda)^{\tilde{\theta}}e^{-\theta\lambda}/(\tilde{\theta}!)$, where $\theta\lambda$ is the expected number of events. This uncertainty concerns also data-driven samples.

As explained in Section 4.9.1, all efficiency-related systematic uncertainties are associated with a shape uncertainty (however, they are neglected if the variation they induce is $< 1\%$ in all bins). The other systematic uncertainties have been checked for the presence of shape deviations on the samples with a substantial number of MC events (the ggF signal and the WW and top-quark backgrounds, the latter in the 1-jet channel only) but, as shown in Figs. 4.27, 4.28, 4.29 and 4.30, no significant shape deviation has been found. In fact, most systematic uncertainties *do* induce a shape variation, but on the total background rather than on its components. Table 4.10 summarizes names and definitions of the nuisance parameters used in the analysis, indicating also if a shape component is present.

For what concerns the normalization systematic uncertainties, they are neglected if smaller than $\pm 0.5\%$, to improve the speed and the convergence stability of the fit. The MC statistical uncertainties are applied to the total background in each bin.

An additional case, on top of the three discussed above, applies to the backgrounds normalized in dedicated control regions; the constraint is also a Poissonian, with λ substituted by $\lambda(\theta)$. The latter is the expected number of events in the CR, and is given by $\lambda(\theta) = \mu S + \theta B_{\text{target}} + \sum B_i$, where the sum runs on the backgrounds not targetted by the CR. The strength parameter θ is used in any region where B_{target} appears.

The BDT analysis does not use any nuisance parameters specific to the data-taking period (BK or LM), unlike the standard analysis, where the $Z/\gamma^* + \text{jets}$ estimation gives rise to statistical uncertainties which are period-related. For this reason, all BDT nuisance parameters are fully correlated among the periods.

4.10.4 Test statistic and CL_s procedure

In order to distinguish between the null hypothesis (B) and the alternative one ($S+B$), a test statistic Q has to be defined. Following the standard LHC prescription [230,231], this analysis uses the *alternate profile likelihood ratio* (PLR) $\tilde{\lambda}(\mu)$:

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} & , \hat{\mu} > 0 \\ \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(0, \hat{\theta}(0))} & , \hat{\mu} \leq 0 \end{cases} \quad (4.11)$$

where the signal strength μ is bounded to be physical. The Neyman-Pearson lemma [232] motivates the choice of the likelihood ratio, stating that, among all data functions, this is the one with the highest statistical power, defined as the probability to reject the null hypothesis if false. The *profile likelihood test statistic* (PLTS) is constructed from the PLR, under the requirement for μ to be physical:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu) & , \hat{\mu} < \mu \\ 0 & , \hat{\mu} \geq \mu \end{cases} . \quad (4.12)$$

In case $\hat{\mu}$ is larger than the tested μ , i.e. in case of an overfluctuation of the signal-like events in data, this expression for the test statistic ensures that no evidence against the signal can be extracted.

The distribution of $Q = \tilde{q}_\mu$ is labelled as $f(Q|B)$ in the null hypothesis, $f(Q|S+B)$ in the alternate; they are schematically represented in Fig. 4.33. The solid vertical line corresponds to the observed value. The p -value of the $S+B$ hypothesis, corresponding to the blue filled area, is defined as the probability of finding a Q value equally or less compatible with the $S+B$ model than the one measured in

Parameter Name	Description	Type	Shape Variation
SigXsecOverSM	$\sigma_{\text{meas}}/\sigma_{\text{SM}}$	Parameter of Interest	—
sampleNorm_ggf125	Normalization uncertainty for the ggF signal sample	Statistical	—
sampleNorm_vbf125	Normalization uncertainty for the VBF signal sample	Statistical	—
sampleNorm_wh125	Normalization uncertainty for the WH signal sample	Statistical	—
sampleNorm_zh125	Normalization uncertainty for the ZH signal sample	Statistical	—
sampleNorm_ttbar	Normalization uncertainty for the $t\bar{t}$ sample	Statistical	—
sampleNorm_st	Normalization uncertainty for the single top-quark sample	Statistical	—
sampleNorm_ww	Normalization uncertainty for the WW sample	Statistical	—
sampleNorm_wzzz	Normalization uncertainty for the diboson sample	Statistical	—
sampleNorm_zjets	Normalization uncertainty for the $Z/\gamma^* + \text{jets}$ sample	Statistical	—
sampleNorm_wjets	Normalization uncertainty for the $W + \text{jets}$ sample	Statistical	—
norm_WW0j	WW normalization in 0-jet	Control Region	—
norm_WW1j	WW normalization in 1-jet	Control Region	—
norm_top1j	Top-quark backgrounds normalization in 1-jet	Control Region	—
B_EFF	B -tagging efficiency	Experimental	✓
C_EFF	C -tagging efficiency	Experimental	✓
L_EFF	Mis-tagging efficiency	Experimental	✓
JES	Jet energy scale	Experimental	—
JER	Jet energy resolution	Experimental	—
E_EFF	Electron selection efficiency	Experimental	✓
M_EFF	Muon selection efficiency	Experimental	✓
E_SCALE	Electron energy scale	Experimental	—
M_SCALE	Muon energy scale	Experimental	—
E_RES	Electron energy resolution	Experimental	—
ID_RES	Muon ID track resolution	Experimental	—
MS_RES	Muon Spectrometer track resolution	Experimental	—
TRIGGER_SF	Trigger efficiency uncertainty	Experimental	✓
CLUSTER	MET uncertainty	Experimental	—
PILEUP	MET uncertainty due to pile-up	Experimental	—
LUMI	Luminosity uncertainty	Experimental	—
Fake_Rate	Uncertainty on the $W + \text{jets}$ fake rate	Data-driven	✓
TOP_SCALEF_0j	Theoretical uncertainty on 0-jet top-quark estimate	Data-driven (th)	—
TOP_SCALEF_0j_hpt	Statistical uncertainty on 0-jet top-quark estimate	Data-driven (stat)	—
Z_SCALEF	Statistical uncertainty on the data-driven $Z + \text{jets}$ estimate	Data-driven (stat)	—
QCDscale_V	$q\bar{q} \rightarrow V$ scale uncertainty	Theory	—
QCDscale_VV	$q\bar{q} \rightarrow VV$ scale uncertainty	Theory	—
QCDscale_VV_ACCEPT	$q\bar{q} \rightarrow VV$ scale uncertainty on acceptance	Theory	—
QCDscale_ggH	$gg \rightarrow H$ scale uncertainty	Theory	—
QCDscale_ggHlin	$gg \rightarrow H$ scale uncertainty (for ≥ 1 jet)	Theory	—
QCDscale_ggH_ACCEPT	$gg \rightarrow H$ scale uncertainty on acceptance	Theory	—
QCDscale_qqH	$qq \rightarrow qqH$ scale uncertainty	Theory	—
QCDscale_qqH_ACCEPT	$qq \rightarrow qqH$ scale uncertainty on acceptance	Theory	—
QCDscale_ttbar	Uncertainty on inclusive $t\bar{t}$ cross section	Theory	—
UEPS	Uncertainty on ggF signal shape from POWHEG/MC@NLO comparison	Theory	✓
pdf_gg	gg PDF uncertainty	Theory	—
pdf_qqbar	$q\bar{q}$ PDF uncertainty	Theory	—
pdf_qqbar_ACCEPT	$q\bar{q}$ PDF acceptance uncertainty	Theory	—
pdf_qqbar	$q\bar{q}$ PDF uncertainty	Theory	—

Table 4.10: Naming scheme for the nuisance parameters used in this analysis, from experimental and theoretical sources, limited number of MC events, control regions and data-driven methods. The rightmost column indicates if a shape component is present in addition to the normalization one.

data, assuming $S + B$ to be true:

$$p_{S+B} = P(Q \geq Q_{\text{obs}} | S + B) = \int_{Q_{\text{obs}}}^{\infty} f(Q | S + B) dQ. \quad (4.13)$$

Following the CL_{s+b} approach, the signal hypothesis is considered as excluded with a confidence level of $1 - \alpha$ if $p_{S+B} < \alpha$. A 2σ exclusion corresponds to the choice $\alpha = 0.05$.

In the same way, the p -value of the null hypothesis, corresponding to the red filled area in Fig. 4.33, is defined as the integral to the left-side on the background-only distribution:

$$p_B = P(Q \leq Q_{\text{obs}}|B) = \int_{-\infty}^{Q_{\text{obs}}} f(Q|B) dQ. \quad (4.14)$$

In the CL_{s+b} approach outlined above, if $S \ll B$ (thus $S + B \approx B$ so that the distributions of the two hypotheses largely overlap with each other), a pathological situation can arise. In this case, a fluctuation in the data can lead to exclude S even without exclusion power. This bias can be cured defining the CL_s criterium:

$$\text{CL}_s = \frac{p_{S+B}}{1 - p_B} < \alpha \quad (4.15)$$

in such a way that the p -value in Eq. 4.13 gets a penalty of $1 - p_B$, which is negligible in case the two distributions are well separated, but becomes more consistent the more they overlap, reducing substantially in this case the rejection power of the analysis. The requirement $\text{CL}_s < \alpha$ is more conservative than the previous CL_{s+b} criterium, since $\text{CL}_s > p_{S+B}$. The CL_s technique, also known as the *modified frequentist* method, is employed in this analysis to extract limits on the observed Higgs boson cross section.

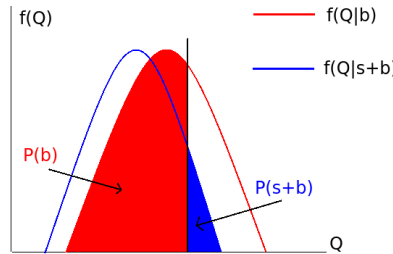


Figure 4.33: Schematic representation of the $f(Q|S+B)$ and $f(Q|B)$ distributions and definition of p_B and p_{S+B} , labelled as $P(b)$ and $P(s+b)$ in the figure, respectively).

4.11 Results

This Section presents an overview of the final results for the BDT analysis with the full 2011 luminosity. CL_s exclusion limits are shown and compared to the cut-based results, both for a shape fit of m_T and a cut-and-count analysis. Moreover, results for the fitted signal strength $\hat{\mu}(m_H)$ and the p_0 significance for the background-only expectation are presented. Only the main plots are given here; a larger collection, broken down into lepton flavours, jet bins and data-taking periods, can be found in

Appendix H.

4.11.1 CL_s exclusion limits

After the application of all selection criteria, the number of events in the data, signal and background samples is shown in Table 4.11, where a sum over lepton flavours has been carried out. These numbers are the same shown in Table 4.5, but with the addition of the total pre-fit uncertainties, including the constraints from the control regions. The 2-jet bin is shown for completeness. All numbers larger than 1 are rounded to the closest integer. Comparing the total uncertainties with the statistical ones in Table 4.5, the effect of the systematic uncertainties can be appreciated: it is substantial for all processes (apart from $Z/\gamma^* + \text{jets}$, due to its limited number of MC events), and in particular for $W + \text{jets}$ it is close to $\pm 50\%$ as already noticed.

	Signal	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	$tW/tb/tqb$	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$	Total Bkg.	Obs.
0-jet	47 ± 12	557 ± 53	51 ± 10	60 ± 13	40 ± 13	37 ± 7	98 ± 51	844 ± 78	862
1-jet	14 ± 5	110 ± 33	22 ± 4	77 ± 7	40 ± 15	79 ± 28	27 ± 13	355 ± 60	374
2-jet	0.8 ± 0.2	0.8 ± 0.6	0.1 ± 0.1	0.9 ± 0.4	negl.	negl.	negl.	1.8 ± 1.1	1

Table 4.11: Expected numbers of signal ($m_H = 125$ GeV) and background events at the end of the selection, as well as the observed numbers of events in the data, summed over the lepton flavours. The total pre-fit uncertainties are shown, taking into account the constraints from the control regions.

The number of candidates in the data and in the MC expectation, as well as the fit to the BDT response, show no evidence of a SM Higgs boson, so limits on $\sigma_{\text{meas}}/\sigma_{\text{SM}}$ can be set, as a function of m_H , following the modified frequentist method outlined in Section 4.10.4. Fig. 4.34 shows the individual results for the 0- and 1-jet bins, in the full range ($110 < m_H < 600$ GeV) (a,b) and zoomed below 300 GeV (c,d) and 150 GeV (e,f). Fig. 4.35 shows the combined result for all jet bins, with the 2-jet result taken from the cut-based analysis. The impact of the individual lepton flavours can be seen in Fig. 4.36, in the intermediate mass range.

Comparing Figs. 4.34, 4.35 and 4.36, it appears evident that the sensitivity of the combined result is driven by that of the $e\mu$ channel in the 0-jet bin, and the performance in 1-jet is in general weaker than in 0-jet. While in the $e\mu$ 0-jet channel the observation agrees with the background-only expectation, in the $\mu\mu$ 0-jet and $e\mu$ 1-jet channels a moderate overfluctuation is observed in the data, at low-mass, with a $\sim 2\sigma$ significance (see also Table 4.6), while the ee 0-jet channel presents an underfluctuation in the same range. Overall, as shown in Fig. 4.34, in the 1-jet channel the observation is 2σ apart from the background-only hypothesis.

Combining all channels and jet bins, a SM Higgs boson is expected to be excluded

in the range:

$$127 \text{ GeV} \leq m_H \leq 254 \text{ GeV} \quad \text{at 95\% CL} \quad (\text{BDT expected}).$$

The observed excluded range, with 4.7 fb^{-1} of integrated luminosity, is:

$$130 \text{ GeV} \leq m_H \leq 281 \text{ GeV} \quad \text{at 95\% CL} \quad (\text{BDT observed}).$$

The corresponding cut-based results are:

$$127 \text{ GeV} \leq m_H \leq 233 \text{ GeV} \quad \text{at 95\% CL} \quad (\text{cut-based expected})$$

$$133 \text{ GeV} \leq m_H \leq 258 \text{ GeV} \quad \text{at 95\% CL} \quad (\text{cut-based observed}).$$

The comparison between the BDT and the cut-based results can be visualized in Fig. 4.37, for 0-jet, 1-jet and combined limits (the actual values are given in Table 4.12 for the combined case, and in Appendix H for 0- and 1-jet). A ratio smaller than 1 indicates that the BDT analysis is improving with respect to the cut-based one. The improvement is substantial in the high-mass range, namely $> 20\%$ for $m_H \gtrsim 250 \text{ GeV}$; for smaller masses, however, it is less significant, and negligible for $m_H \lesssim 145 \text{ GeV}$, where the BDT analysis confirms the cut-based results. This behaviour can be related mainly to two facts:

- the small number of signal candidates at low-mass (for $m_H = 125 \text{ GeV}$, there are < 20 expected Higgs boson events in the $e\mu$ 0-jet channel, which is the one leading the sensitivity,
- the improvement gained by the cut-based analysis when moving from a cut-and-count experiment to a shape fit of m_T .

Comparing the BDT limits with the standard cut-and-count ones (with a cut on the transverse mass, $0.75 m_H < m_T < m_H$), shows a larger improvement at low-mass for the multivariate result, as can be seen in Fig. 4.38 for the combined channels (the actual values are given in Table 4.13). In this case, the improvement increases monotonically for $m_H \lesssim 300 \text{ GeV}$ (up to 52% at $m_H = 340 \text{ GeV}$), and then starts to decrease, as the shape of the BDT response shrinks going to larger masses. For observed limits, on the other hand, the improvement is more stable: around 30% for intermediate masses and above 40% for high masses, while the low-mass range is dominated by statistical fluctuations in the data.

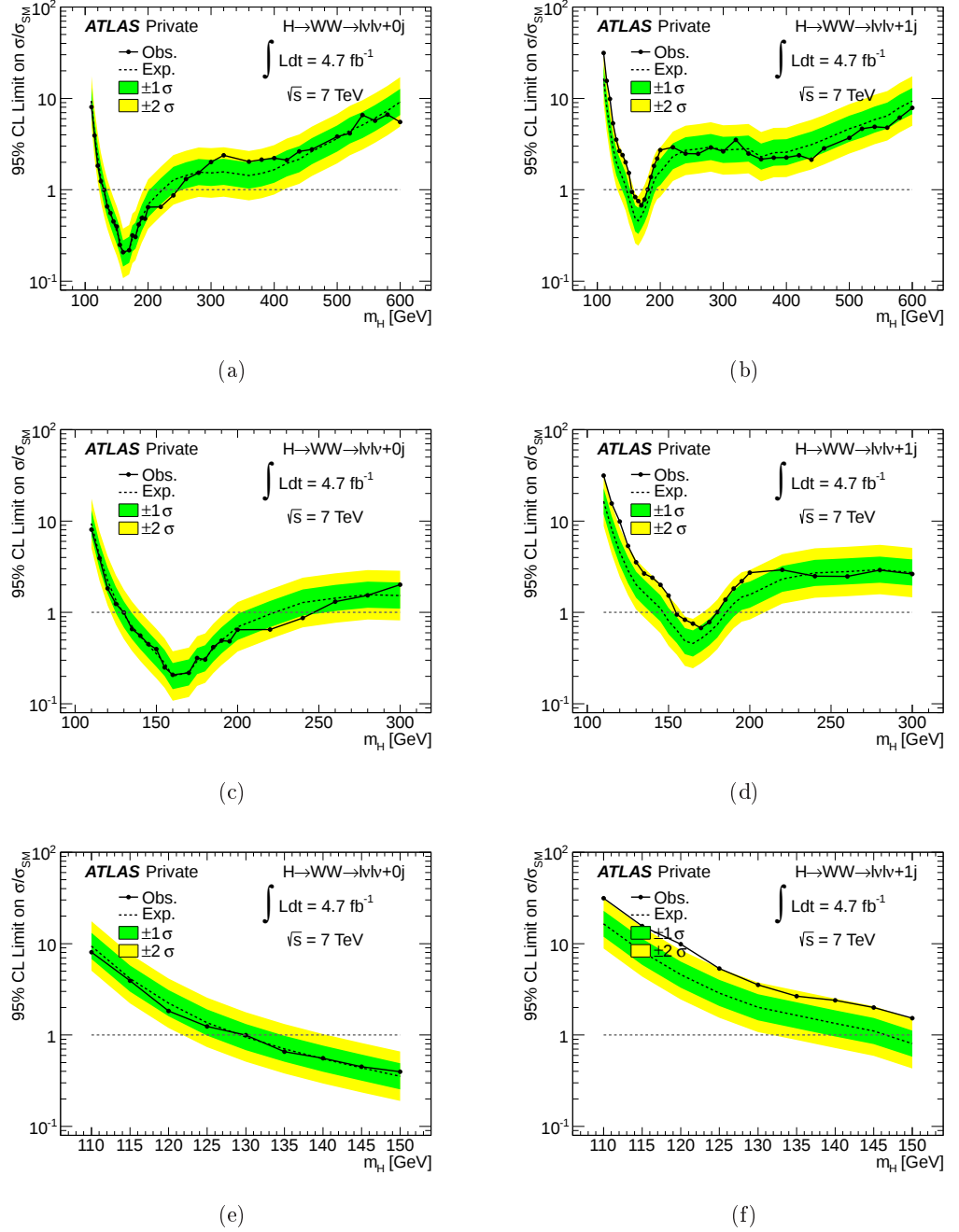


Figure 4.34: Expected (dashed) and observed (solid) 95% CL upper limits on the Higgs boson production cross section, normalized to the SM cross section, as a function of m_H , obtained using a shape fit to the BDT response on 4.7 fb^{-1} of integrated luminosity. The green and yellow bands show the 1σ and 2σ expected uncertainty ranges. The left and right columns contain the limits for the 0- and 1-jet channels, respectively; in (a,b) the full mass range ($110 < m_H < 600$ GeV) is shown, in (c,d) the range $110 \text{ GeV} < m_H < 300$ GeV is selected, in (e,f) the range $110 \text{ GeV} < m_H < 150$ GeV.

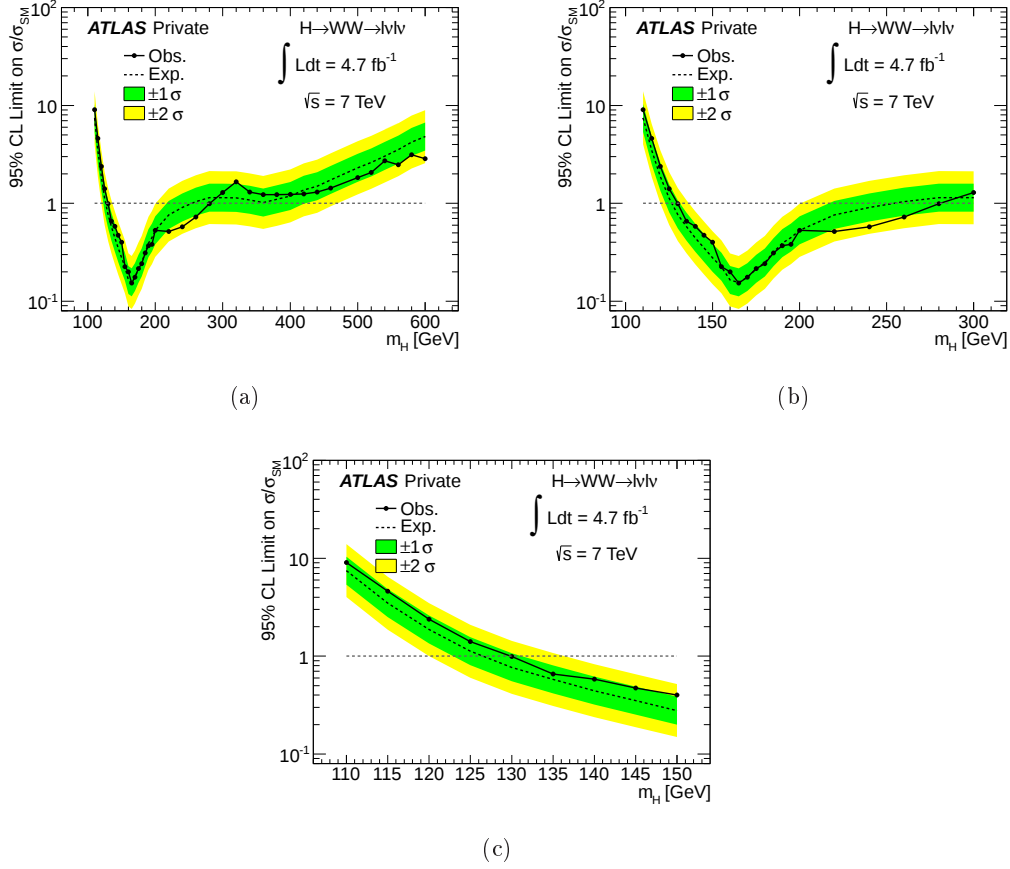


Figure 4.35: Expected (dashed) and observed (solid) 95% CL upper limits on the Higgs boson production cross section, normalized to the SM cross section, as a function of m_H , (a) in the full mass range, (b) zoomed up to 300 GeV, (c) zoomed up to 150 GeV, obtained using a shape fit to the BDT response on 4.7 fb⁻¹ of integrated luminosity. The green and yellow bands show the 1σ and 2σ expected uncertainty ranges.

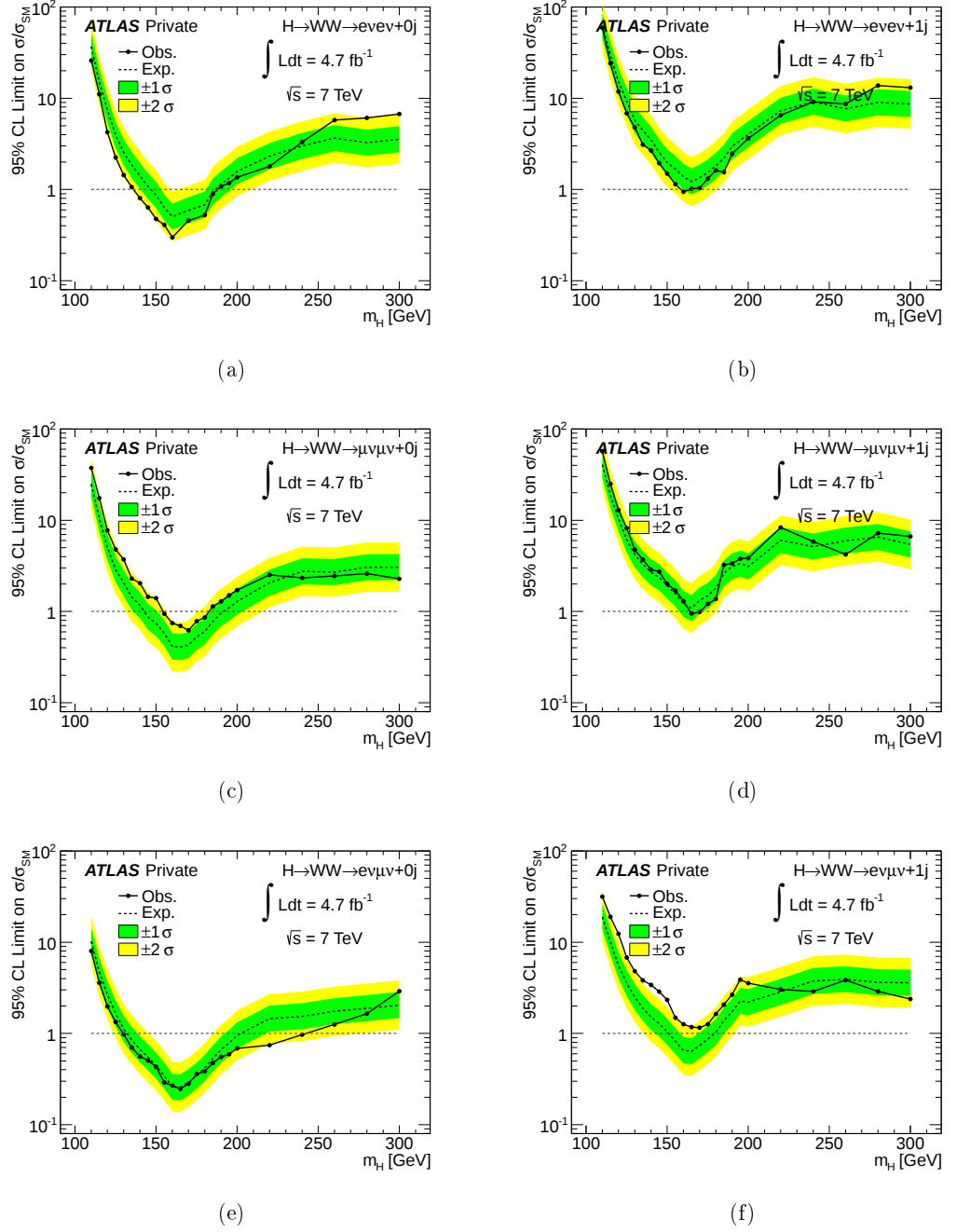


Figure 4.36: Expected (dashed) and observed (solid) 95% CL upper limits on the Higgs boson production cross-section, normalized to the SM cross-section, in the intermediate-mass region ($110 \text{ GeV} < m_H < 300 \text{ GeV}$). The left and right columns contain the limits for the 0- and 1-jet channels, respectively; the lepton flavours shown are (a,b) ee , (c,d) $\mu\mu$, (e,f) $e\mu$.

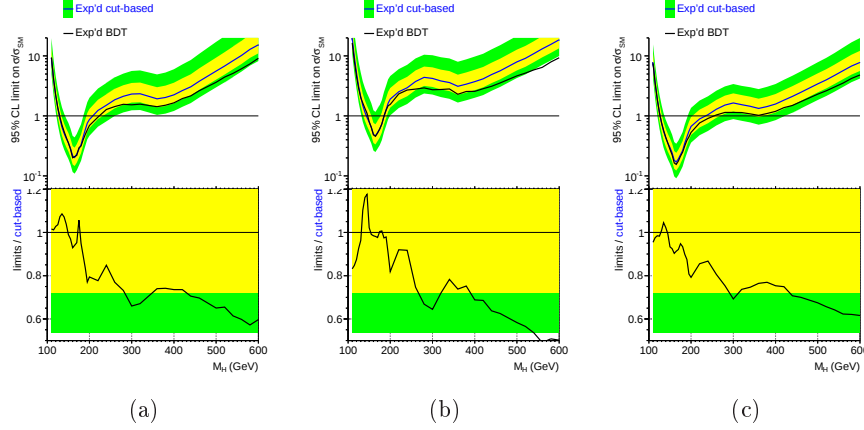


Figure 4.37: Comparison of expected limits between the BDT and the cut-based analyses, broken down into jet bins: (a) 0-jet, (b) 1-jet, (c) combined. The bottom pads contain the ratio of the corresponding two curves. The yellow and green bands refer to the 1σ and 2σ envelopes of the cut-based limit.

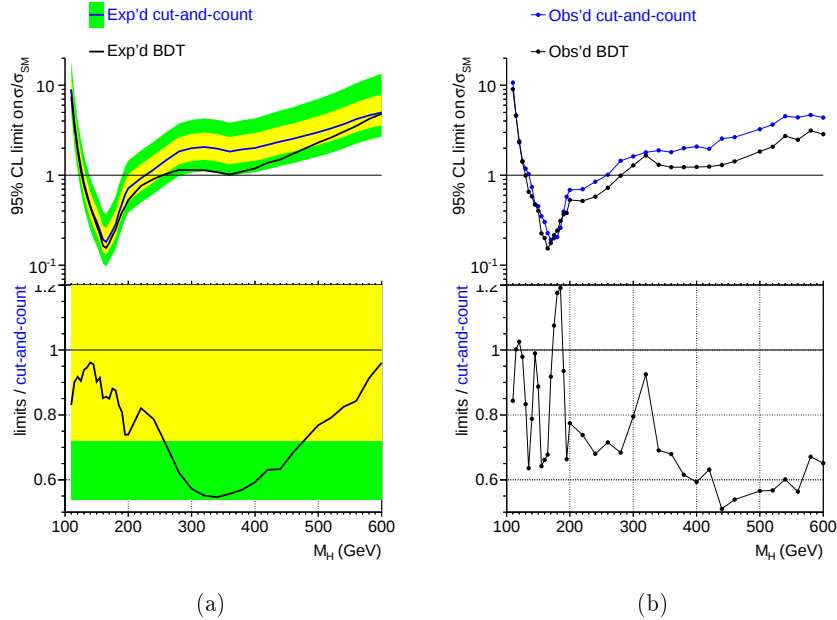


Figure 4.38: Comparison of (a) expected and (b) observed limits between the BDT and the cut-based analyses (the latter using a cut on the transverse mass, $0.75 m_H < m_T < m_H$), with all three jet multiplicities combined. The bottom pads contain the ratio of the corresponding two curves. The yellow and green bands refer to the 1σ and 2σ envelopes of the cut-based expected limit.

m_H (GeV)	BDT		Cut-based m_T shape fit		Ratio of expected limits (BDT/cut-based)
	Observed	Expected	Observed	Expected	
110	9.03	7.48	10.15	8.08	0.93
115	4.61	3.48	4.68	3.67	0.95
120	2.38	1.87	2.63	1.96	0.95
125	1.41	1.12	1.69	1.18	0.95
130	0.99	0.77	1.14	0.78	0.98
135	0.66	0.58	0.87	0.57	1.01
140	0.58	0.44	0.66	0.45	0.99
145	0.47	0.35	0.52	0.36	0.96
150	0.40	0.28	0.41	0.31	0.91
155	0.23	0.22	0.31	0.25	0.90
160	0.20	0.16	0.21	0.19	0.87
165	0.15	0.16	0.17	0.18	0.88
170	0.18	0.18	0.19	0.2	0.89
175	0.22	0.21	0.21	0.23	0.92
180	0.24	0.25	0.24	0.27	0.90
185	0.31	0.32	0.31	0.36	0.87
190	0.37	0.39	0.39	0.46	0.85
195	0.38	0.45	0.45	0.57	0.78
200	0.53	0.53	0.65	0.68	0.78
220	0.52	0.76	0.77	0.9	0.84
240	0.58	0.91	0.86	1.07	0.85
260	0.72	1.04	1.03	1.31	0.79
280	0.99	1.14	1.21	1.54	0.74
300	1.29	1.14	1.28	1.67	0.68
320	1.66	1.14	1.34	1.55	0.73
340	1.30	1.08	1.3	1.46	0.74
360	1.23	1.02	1.21	1.34	0.76
380	1.23	1.10	1.34	1.44	0.76
400	1.23	1.19	1.44	1.59	0.75
420	1.25	1.37	1.69	1.85	0.74
440	1.30	1.50	1.91	2.14	0.70
460	1.43	1.74	2.3	2.51	0.69
480	-	-	-	-	-
500	1.83	2.31	3.02	3.48	0.66
520	2.07	2.62	3.64	4.06	0.65
540	2.73	3.04	4.32	4.81	0.63
560	2.48	3.53	5.07	5.77	0.61
580	3.13	4.22	6.27	6.93	0.61
600	2.85	4.81	7.1	7.97	0.60

Table 4.12: Comparison of observed and expected limits from the BDT analysis with those from the cut-based m_T fit analysis, with all jet multiplicities combined. The last column contains the ratio of the expected limits.

m_H (GeV)	BDT		Cut-based cut-and-count		Ratio of expected limits (BDT/cut-based)
	Observed	Expected	Observed	Expected	
110	9.03	7.48	10.67	8.93	0.84
115	4.61	3.48	4.52	3.87	0.90
120	2.38	1.87	2.32	2.06	0.91
125	1.41	1.12	1.46	1.24	0.91
130	0.99	0.77	1.19	0.84	0.92
135	0.66	0.58	1.03	0.63	0.92
140	0.58	0.44	0.74	0.51	0.87
145	0.47	0.35	0.49	0.37	0.95
150	0.40	0.28	0.45	0.31	0.89
155	0.23	0.22	0.35	0.25	0.88
160	0.20	0.16	0.3	0.2	0.83
165	0.15	0.16	0.23	0.2	0.80
170	0.18	0.18	0.19	0.21	0.84
175	0.22	0.21	0.2	0.24	0.88
180	0.24	0.25	0.2	0.29	0.87
185	0.31	0.32	0.26	0.38	0.83
190	0.37	0.39	0.39	0.49	0.81
195	0.38	0.45	0.57	0.61	0.74
200	0.53	0.53	0.67	0.73	0.73
220	0.52	0.76	0.7	0.96	0.79
240	0.58	0.91	0.84	1.17	0.78
260	0.72	1.04	1.02	1.49	0.70
280	0.99	1.14	1.5	1.89	0.60
300	1.29	1.14	1.65	2.03	0.56
320	1.66	1.14	1.84	2.13	0.53
340	1.30	1.08	1.95	2.08	0.52
360	1.23	1.02	1.86	1.9	0.54
380	1.23	1.10	2.04	2	0.55
400	1.23	1.19	2.17	2.15	0.55
420	1.25	1.37	2.04	2.24	0.61
440	1.30	1.50	2.6	2.52	0.59
460	1.43	1.74	2.7	2.56	0.68
480	-	-	-	-	-
500	1.83	2.31	3.35	3.31	0.70
520	2.07	2.62	3.9	3.99	0.66
540	2.73	3.04	4.68	4.21	0.72
560	2.48	3.53	4.61	4.78	0.74
580	3.13	4.22	4.89	5.43	0.78
600	2.85	4.81	4.6	5.66	0.85

Table 4.13: Comparison of observed and expected limits from the BDT analysis with those from the cut-based analysis (the latter using a cut on the transverse mass, $0.75 m_H < m_T < m_H$), with all jet multiplicities combined. The last column contains the ratio of the expected limits. Some of the numbers are missing due to fit failures.

4.11.2 Fitted signal strength

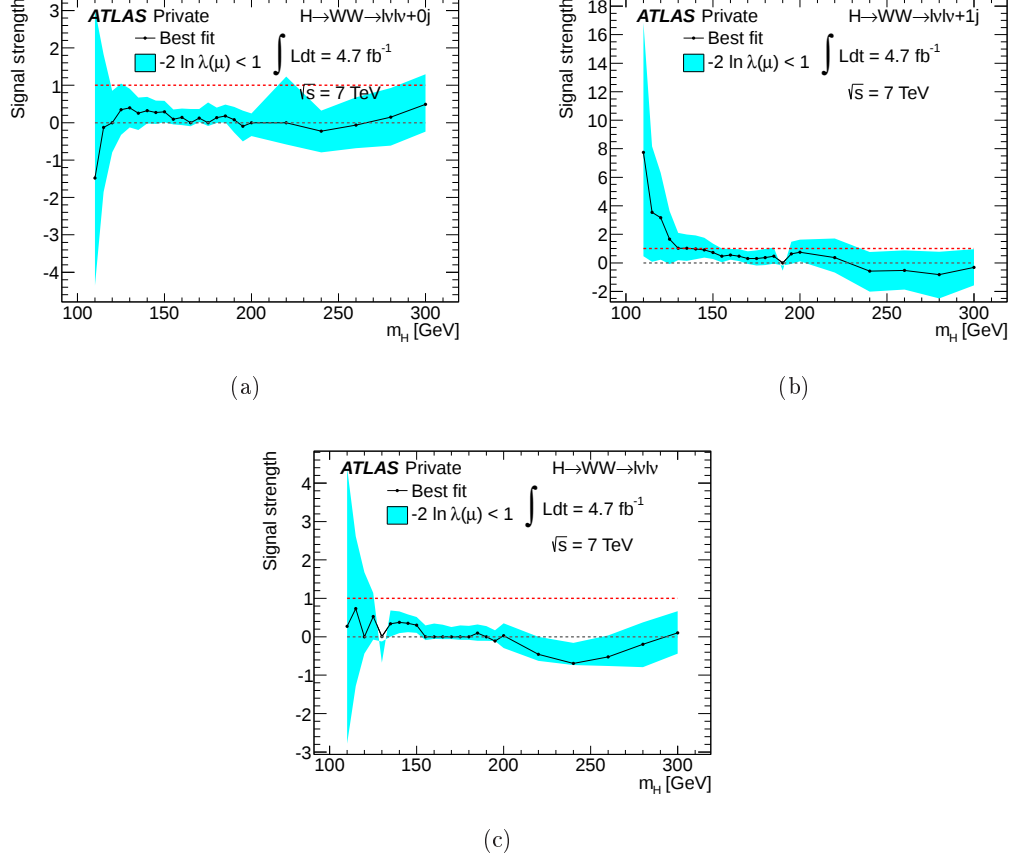


Figure 4.39: The fitted signal strength as a function of m_H , for $m_H < 300$ GeV: (a) 0-jet, (b) 1-jet, (c) combined. The solid line shows the observed result, the shaded cyan band encloses the systematic area $-2\ln(\mu) < 1$. For two points ($m_H = 180$ GeV for 1-jet, $m_H = 130$ GeV for combined) the result should not be considered due to fit failures.

Fig. 4.39 shows the fitted values for the signal strength μ as a function of m_H , for $m_H < 300$ GeV. The shaded cyan band encloses the systematic area $-2\ln(\mu) < 1$. The results for $\hat{\mu}$, in units of the SM cross section, at $m_H = 125$ GeV, are summarized in Table 4.14, where they are compared with corresponding values for the cut-based analysis.

The values are compatible inside the uncertainties for 0- and 1-jet bins, and in very good agreement for the combination. The corresponding combined result for the cut-and-count standard analysis is $\hat{\mu}(125 \text{ GeV}) = 0.17^{+0.70}_{-0.74}$, in units of the SM cross section, which is also compatible. It should be noticed that, since both the BDT and the cut-based analyses find no evidence of a SM Higgs boson at low-mass,

	0-jet	1-jet	Combined
BDT	$0.35^{+0.69}_{-0.67}$	$1.67^{+1.97}_{-1.75}$	$0.53^{+0.61}_{-0.60}$
Cut-based	$-0.04^{+0.62}_{-0.65}$	$2.90^{+2.59}_{-1.49}$	$0.49^{+0.72}_{-0.66}$

Table 4.14: Results for Higgs boson the signal strenght, in units of the SM cross section, as fitted by the BDT analysis and by the standard cut-based analysis. The results are shown for the individual 0- and 1-jet bins, and for the combination, which includes, both for the BDT and the cut-based cases, the cut-based 2-jet channel.

given their respective sensitivities, the fitted signal strength is compatible with 0 for both of them; in 1-jet, the central values are larger, but in this channel the expected CL_s limit is above the SM expectation, so no conclusion can be reached on the presence of a signal, as also shown by the large errors.

4.11.3 Local p_0 -values

As defined in Eq. 4.14, the p_0 (alternatively, p_B) value for the background-only hypothesis quantifies the probability of finding a value of the test statistic at least equal to the one observed in data, assuming the null hypothesis to be true. The results are shown in Fig. 4.40 for the individual jet bins and their combination, while Fig. 4.41 shows the individual lepton flavours.

As already noticed in Section 4.11.1, in the 1-jet channel an excess of events is observed with respect to the background-only hypothesis, and this can be confirmed looking at the p_0 plots presented here: in the 1-jet bin, the local⁽⁷⁾ significance of the excess is around 2σ , and compatible with the presence of a signal in a broad range $130 \text{ GeV} \lesssim m_H \lesssim 145 \text{ GeV}$. Breaking the result down into lepton flavours, the excess appears localized in the $e\mu$ 1-jet channel, where 235 data events are found, to be compared with 218 background and 9 signal events expected from the simulation (see Table 4.6). In the rest of the mass range, up to $m_H = 600 \text{ GeV}$, the local p_0 -values are below 1σ , as can be seen in Appendix H, apart from some statistical fluctuations in the $e\mu$ 0-jet channel.

⁷To estimate the correct significance, the *look-elsewhere effect* should be taken into account, normalizing the significance found for a specific m_H hypothesis to the full range where the Higgs boson is searched for.

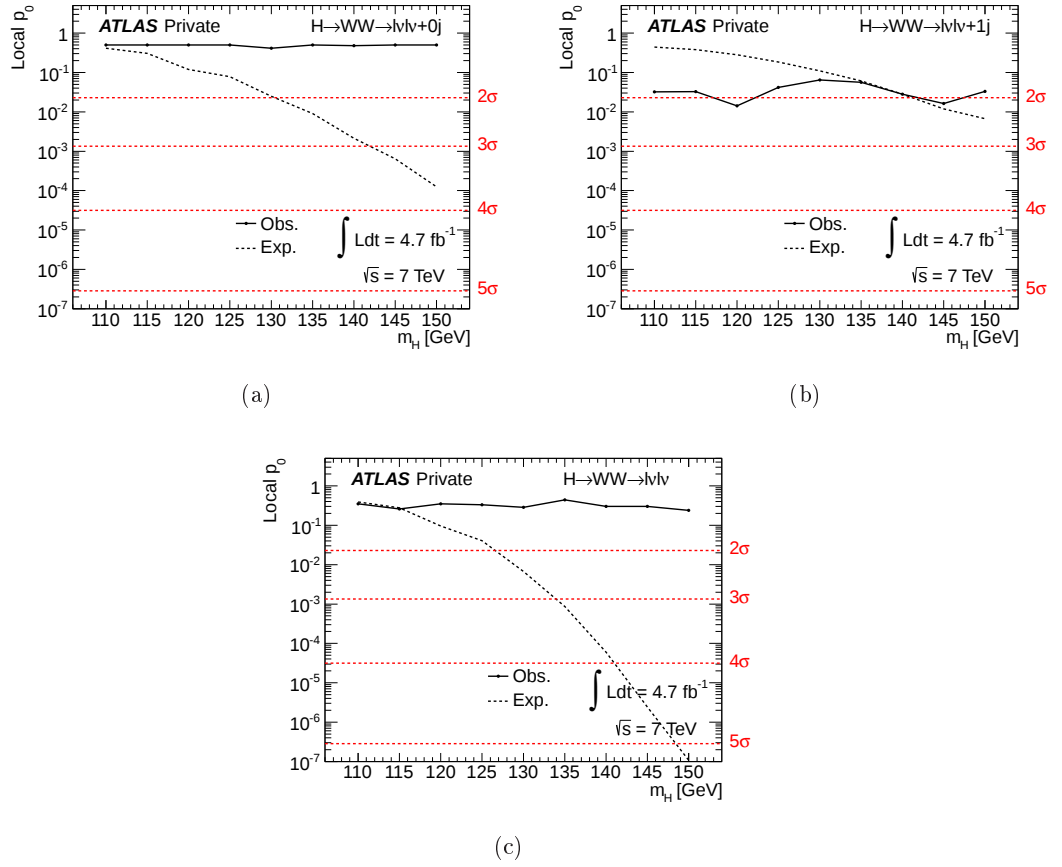


Figure 4.40: Local p_0 -values as a function of m_H , in the low-mass range: (a) 0-jet, (b) 1-jet, (c) combined. The observed probability for the background-only scenario is shown as a solid line, while the dashed line corresponds to the expectation for the signal+background hypothesis at the given value of m_H .

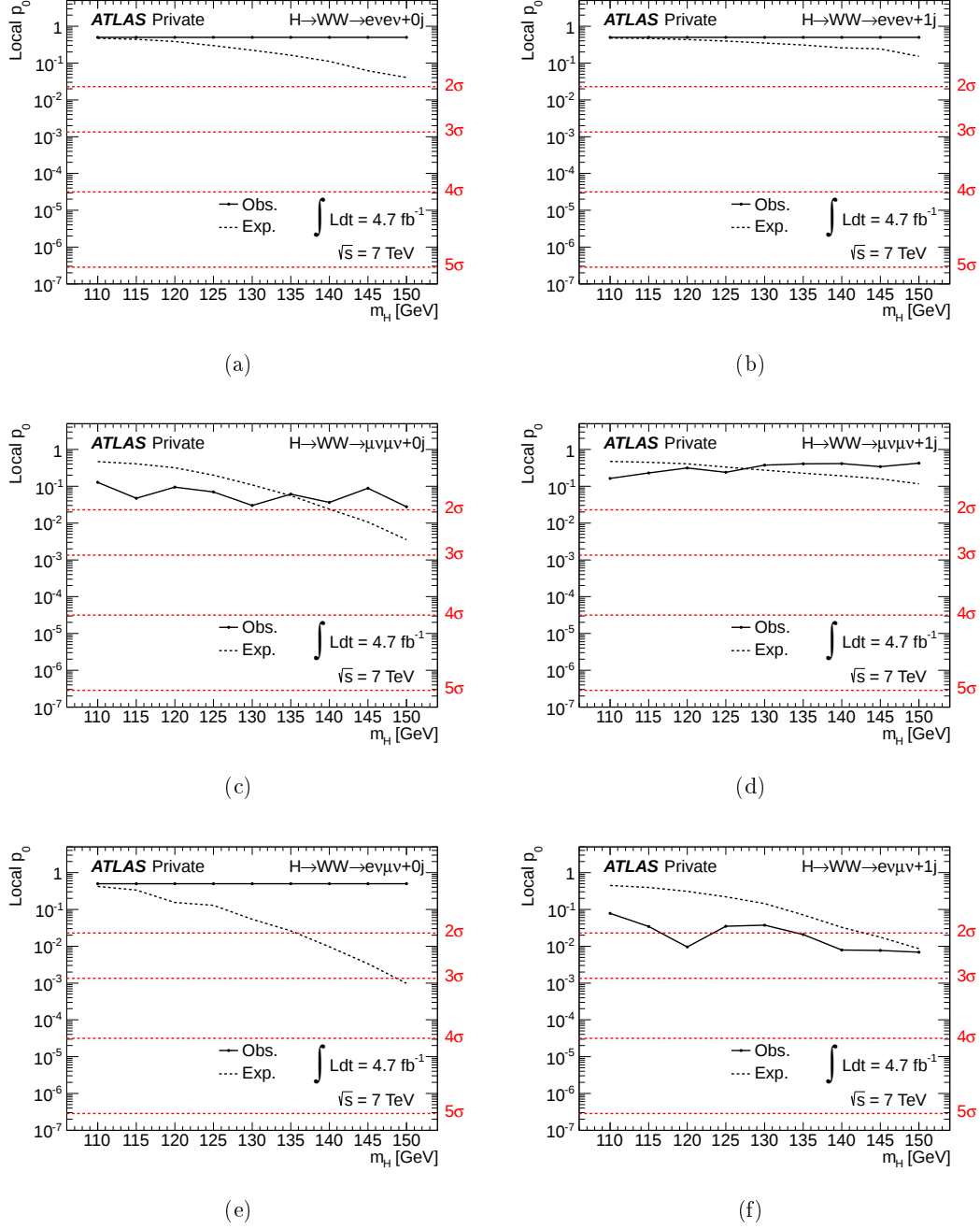


Figure 4.41: Local p_0 -values as a function of m_H , in the low-mass range. The observed probability for the background-only scenario is shown as a solid line, while the dashed line corresponds to the expectation for the signal+background hypothesis at the given value of m_H . The left and right columns contain the results for the 0- and 1-jet channels, respectively; the lepton flavours shown are (a,b) ee , (c,d) $\mu\mu$, (e,f) $e\mu$.

Chapter 5

Study of the Spin of the Higgs Boson in the $H \rightarrow WW^*$ Channel

In this Chapter a multivariate study of the spin of the Higgs boson, with the full dataset collected in 2012 with the ATLAS detector, at $\sqrt{s} = 8$ TeV, is presented. Starting from the knowledge acquired with the $\sqrt{s} = 7$ TeV analysis performed in 2011 and documented in Chapter 4, a BDT analysis has been developed, to explore the spin and parity⁽¹⁾ quantum numbers of the boson discovered by the ATLAS and the CMS collaborations at a mass close to 125 GeV [119, 233]. In particular, the Standard Model hypothesis, $J^P = 0^+$, is tested against the alternative $1^+, 1^-, 2^+$ hypotheses, which have been described in Chapter 2. The $J^P = 2^+$ model is the minimal one for Kaluza-Klein-like gravitons, as discussed in Section 2.3.3, while $J^P = 1^\pm$ follows Section 2.3.2, for decays into two massive vector bosons. The $J^P = 1^\pm$ models are considered even if the Landau-Yang theorem forbids the decay of such a boson into two on-shell photons [37, 38], and thus the evidence of the decay in the $H \rightarrow \gamma\gamma$ channel [39] seems to contradict this hypothesis. The latter is, however, still worthwhile to be tested, given the fact that there could be two distinct resonances, close in mass, decaying into photons and vector bosons, respectively. The $J^P = 0^-$ model has not been investigated so far, given the low discrimination expected in the $H \rightarrow WW^*$ channel for this hypothesis, in particular compared to the $H \rightarrow ZZ^* \rightarrow 4\ell$ one. The $J^P = 2^-$ model is also not covered here, since such a model is not uniquely defined from the theoretical point of view.

The spin analysis follows the analysis of the Higgs boson rate (in the following referred to simply as “rate analysis”) described in Refs. [32, 234] for what concerns object definition, event selection, background estimation and systematic uncertain-

¹For brevity, spin and parity are referred to only as spin in the following, unless explicitly needed.

ties, with the following differences: only the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ channel is used, with no accompanying jets, and the selection has been loosened, in order to enlarge the acceptance for different spin hypotheses. Moreover, whereas the rate analysis is a standard cut-based one, the spin analysis exploits the BDT technique.

The analysis has been first developed for the $J^P = 2^+$ hypothesis, and then extended to the $J^P = 1^\pm$ case: for this reason, most details and cross-checks are given for $J^P = 2^+$, but also hold for $J^P = 1^\pm$.

For what concerns the $J^P = 2^+$ hypothesis, since the production mechanism is expected to affect the angular and mass distributions of the final state particles, as anticipated in Section 2.3.3, the analysis is performed for various couplings of the Higgs-like boson to gg and $q\bar{q}$, in five fractions $f_{q\bar{q}} = 0.00, 0.25, 0.50, 0.75, 1.00$; the five analyses follow the same strategy, so in the main text results are mostly shown for one of them (namely $f_{q\bar{q}} = 0.25$).

Sections 5.1, 5.2 and 5.3 describe, respectively, the Monte Carlo generators used to simulate the signal and background samples, the object definition and the event preselection. Section 5.4 is dedicated to the input variables and the technical settings used in the BDT training, while Section 5.5 shows the resulting BDT responses in the signal and control regions. Section 5.6 concentrates on systematic uncertainties, with emphasis on the shape uncertainty assigned to the WW background, which is peculiar to the spin analysis. Section 5.7 discusses the statistical model and the fit procedure. Results are given in Sections 5.8 and 5.9, for $J^P = 2^+$ and $J^P = 1^\pm$ analyses, respectively. Finally, Section 5.10 presents an overview of the spin combination, performed by the ATLAS collaboration, based on this analysis and the corresponding $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ ones: the different channels are briefly discussed and the final results are presented.

Additional details and cross-checks are given in the Appendices. In Appendix J, the choice of the training configuration is justified. Appendices K and L are both dedicated to the theoretical uncertainty on the WW shape (the UEPS systematic), given its impact on the final result: the former details its baseline treatment, the latter shows additional studies on the subject, describing an alternative treatment and its drawbacks. Appendix M, similarly to Appendix E for the 2011 analysis, shows the BDT performance when trained with one of the leading backgrounds as signal (WW , $t\bar{t}$ and $Z/\gamma^* + \text{jets}$). Appendix N compares the chosen test statistic with the profile likelihood ratio, discussing their differences. Finally, Appendix O gives the results for all $J^P = 2^+$ fractions.

This analysis has been first published as an ATLAS conference note [40], and later as part of the combination among the bosonic channels performed by the ATLAS collaboration [41, 235]. Most of the fundamental aspects of the $H \rightarrow WW^*$

spin analysis have been developed in the context of this thesis, in particular: the choice of the input variables, the optimization of the BDT settings, the definition of the signal and the control regions, the scrutiny of the output variables, the choice of the fit model and of the statistical prescription, the assessment of the systematic uncertainties (with a particular dedication to the most important uncertainty, the one related to the shape of the WW background) and the statistical results. Moreover, all the studies presented in the Appendices pertain to this work. This also applies to the additional cross-checks which have been performed to harmonize the $H \rightarrow WW^*$ results with those of the other bosonic channels, in order to allow for their combination.

5.1 Monte Carlo samples

The choice of Monte Carlo generators in 2012 follows closely the description given in Section 4.3 for the 2011 analysis, with a few differences:

- PYTHIA8 [236] substitutes PYTHIA6 [78] in a few cases,
- $q\bar{q}/g \rightarrow WW$ uses POWHEG rather than MC@NLO,
- for single top-quark production, AcerMC [211] is used only for the t -channel, while the s -channel and tW use ALPGEN [206],
- for ZZ production, Sherpa [210] is substituted by POWHEG [192, 193],
- $W\gamma^*$ and $WZ^{(*)}$ are generated together, in order to account for interferences, with the boundary given by the invariant mass: $m_{\ell\ell} < 7$ GeV (MadGraph+PYTHIA6 [208]) and $m_{\ell\ell} \geq 7$ GeV (POWHEG+PYTHIA8).

In addition, a few filters are applied in the generation: for $W\gamma$ events, the photon must have $p_T > 8$ GeV and be separated from the charged lepton by $\Delta R > 0.25$; for $W\gamma^*$ events, at least two leptons with $p_T > 5$ GeV are required in the event, with $|\eta| < 3$ ($|\eta| < 5$) for $ee, \mu\mu$ (for $\tau\tau$), while in the $WZ^{(*)}$ case the cut on lepton pseudorapidities is $|\eta| < 2.7$ for all flavours. Moreover, $Z^{(*)}Z^{(*)} \rightarrow 4\ell$ events are generated with $m_{\ell\ell} > 4$ GeV for the leading lepton pair. Details are summarized in Table 5.1, which also includes the cross sections at $\sqrt{s} = 8$ TeV. For these, the leptonic decays of the W and Z bosons are assumed for all processes, and lepton flavours are summed over, with the exception of top-quark production, for which the inclusive cross section is quoted.

Concerning PDFs, the following sets are used: CT10 [71] for MC@NLO and POWHEG samples, CTEQ6L1 [212] for ALPGEN, MadGraph and PYTHIA6/8 samples.

Process	MC generator	$\sigma \cdot \text{BR}$ (pb)
ggF ($m_H = 125$ GeV)	POWHEG+PYTHIA8	0.441
$gg, q\bar{q} \rightarrow H(J^P \neq 0^+)$ ($m_H = 125$ GeV)	JHU 2.0.2+PYTHIA8	0.441
$q\bar{q}/g \rightarrow WW$	POWHEG+PYTHIA6	5.68
$gg \rightarrow WW$	GG2WW+HERWIG	0.16
$t\bar{t}$	MC@NLO+HERWIG	238
single top-quark tW, tb	MC@NLO+HERWIG	28
single top-quark tqb	AcerMC+PYTHIA6	88
W +jets	ALPGEN+HERWIG	$37 \cdot 10^3$
Z/γ^* +jets	ALPGEN+HERWIG	$16 \cdot 10^3$
ZZ	POWHEG+PYTHIA8	0.73
$W(Z/\gamma^*)(m_{Z/\gamma^*} > 7 \text{ GeV})$	POWHEG+PYTHIA8	0.83
$W(Z/\gamma^*)(m_{Z/\gamma^*} \leq 7 \text{ GeV})$	MadGraph+PYTHIA6	11
$W\gamma$	ALPGEN+HERWIG	369

Table 5.1: Monte Carlo generators used to simulate the different components of the signal ($m_H = 125$ GeV) and the background processes and their corresponding cross sections as used in the analysis. The quoted values assume the leptonic decays of the W and Z bosons for all processes, and lepton flavours are summed over, with the exception of top-quark production, for which the inclusive cross section is reported. For the signal processes, the production cross sections include the branching ratios into the $\ell\nu\ell\nu$ final state. See the text for further details and references.

POWHEG, the baseline generator for the Higgs boson production, is not able to simulate alternative spin configurations, thus it is replaced by the JHU generator [35,36] for this purpose. Only the ggF production mechanism is considered, for all spin hypotheses, given the fact that VH and $t\bar{t}H$ are negligible, and, since considering only events in the 0-jet category, also VBF is.

Moreover, since no predictions for the $J^P = 1^\pm$ or $J^P = 2^+$ cross sections exist, all spin hypotheses are normalized to the SM $J^P = 0^+$ cross section, in both tables and plots.

The nominal JHU sample for the process $gg \rightarrow H(J^P = 2^+) \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ corresponds to the production fraction $f_{q\bar{q}} = 0.25$, which is taken as arbitrary default value. After distinguishing, at generation level, the $q\bar{q}$ events from the gg ones, all MC weights are reweighted in order to generate the other fractions ($f_{q\bar{q}} = 0.00, 0.50, 0.75, 1.00$) and the events are then processed with the fast simulation of the ATLAS detector, ATLFAST [86].

A comparison has been performed between POWHEG and JHU on a sample of SM Higgs boson events, at generation level and before the preselection, in order to validate the angular and kinematical distributions produced by the JHU generator interfaced with ATLFAST. The variables compared are the five Cabibbo-Maksymowicz variables defined in Section 2.3.1; in all cases, no significant deviation between the two generators has been found.

Furthermore, thanks to an update in the `POWHEG` generator, the latter can now reproduce the Higgs boson p_T spectrum provided by the `HqT` generator [200] within a 5% accuracy; for this reason, no reweighting is needed for the `POWHEG` events in the 2012 analysis, unlike in the 2011 one (cf. Section 4.3).

5.2 Object selection

5.2.1 Data-taking conditions

Alike 2011 data, 2012 data can be divided into periods, according to the data-taking conditions, for a total of 20.7 fb^{-1} of integrated luminosity; in this case, however, no period-dependent splitting is performed in the likelihood. The number of interactions per bunch-crossing in Monte Carlo simulation is reweighted to the one found in data, in order to get a better agreement on the minimum bias vertex multiplicity between the `PYTHIA8` tuning in MC and data. This rescaling is associated with a systematic uncertainty on the minimum bias modelling, obtained by varying the scale factor in a range from 0.80 to 1.00.

Before the trigger requirements, events not satisfying the “AllGood” criterium [213] are discarded, in order to ensure that all ATLAS subdetectors have been fully efficient at the time of data-taking (cf. Sections 3.1.1 and 4.4.2).

5.2.2 Trigger requirements

As anticipated in Section 3.3.3 for electrons and in Section 3.3.4 for muons, the following triggers are used, with a logical “or” in each case:

e trigger	EF_e24vhi_medium or EF_e60_medium
μ trigger	EF_mu24i_tight or EF_mu36_tight

The corresponding efficiencies are determined with the tag-and-probe method on $Z \rightarrow ee, \mu\mu$ events, in bins of p_T , η and data-taking period, and are found to be around 90% for electrons and 90% (70%) for muons in the end-caps (barrel). As in 2011 (see Section 4.4.3), at least one of the two selected leptons must be trigger-matched, otherwise the event is discarded.

5.2.3 Electron definition

Electron identification is improved with respect to the 2011 analysis thanks to the use of the GSF algorithm (cf. Section 3.3.3), which recovers for bremsstrahlung losses, thus allowing for tighter isolation criteria, both on track (now p_T -dependent) and calorimeter quantities. Similarly, tighter requirements are placed on the transverse

and longitudinal impact parameters. Starting from the *tight++* menu, the following requirements are added:

- cuts on longitudinal and transverse impact parameters and their significances, defined with respect to the primary vertex:
 - $|z_0 \sin \theta| < 0.4$ mm,
 - $|d_0| < 3 |\sigma(d_0)|$,
- cuts on isolation (in a $\Delta R < 0.3$ cone):
 - $p_T^{\text{cone}}/p_T < 0.12$ (0.16) for $15 \text{ GeV} < p_T < 25 \text{ GeV}$ ($p_T > 25 \text{ GeV}$),
 - $E_T^{\text{cone}}/p_T < 0.16$.

The corresponding criteria for the 2011 analysis can be found in Section 4.4.4. The requirement on the geometrical acceptance is unchanged ($|\eta| < 2.47$, excluding the crack regions $1.37 < |\eta| < 1.52$, and $p_T > 15 \text{ GeV}$). Small differences between the data and the MC prediction, after the application of all identification criteria, are recovered by scale factors which are consistent with unity within $\pm 2\%$.

5.2.4 Muon definition

Muons are reconstructed according to the STACO combined algorithm, as for the 2011 analysis, within a geometrical acceptance which is extended in $|\eta|$ up to 2.5. Quality criteria are tightened with respect to those discussed in Section 4.4.5 for the 2011 analysis, similarly to what just described for electrons:

- cuts on the longitudinal and transverse impact parameters and their significances, defined with respect to the primary vertex:
 - $|z_0 \sin \theta| < 1$ mm,
 - $|d_0| < 3 |\sigma(d_0)|$,
- cuts on isolation (in a $\Delta R < 0.3$ cone):
 - $p_T^{\text{cone}}/p_T < 0.01 p_T - 0.105$ and $p_T^{\text{cone}}/p_T < 0.15$,
 - $E_T^{\text{cone}}/p_T < 0.014 p_T - 0.15$ and $E_T^{\text{cone}}/p_T < 0.20$.

The projection of the longitudinal impact parameter is introduced in order to account for larger uncertainties at high values of η . Isolation requirements are further enhanced rejecting muon candidates in case they are closer than $\Delta R < 0.3$ from a reconstructed jet. Scale factors, computed using a $Z \rightarrow \mu^+ \mu^-$ tag-and-probe sample, are found to be consistent with unity within $\pm 1\%$.

5.2.5 Jet definition

The same jet definition and acceptance criteria defined for the 2011 analysis (cf. Section 3.3.5), are used in the 2012 analysis (“anti- k_t EM+JES”, in a cone of $\Delta R < 0.4$, requiring $|\eta| < 4.5$). Two main differences arise: an η -dependent p_T threshold and a tighter JVF criterium, both aiming to reduce the impact of pile-up jets. The JVF algorithm is fully efficient only inside the tracking volume of the ATLAS detector ($|\eta| < 2.4$): in this region, jets must satisfy the condition $\text{JVF} > 50\%$, if their transverse momentum is $30 \text{ GeV} < m_H < 50 \text{ GeV}$ (no cut on JVF for $p_T > 50 \text{ GeV}$), while outside this region a lower p_T threshold of 25 GeV is applied. This selection is found to be robust against pile-up, based on studies of the jet multiplicity as a function of the number of reconstructed vertices, in $Z/\gamma^* + \text{jets}$ events: while no dependence is found after the application of the JVF requirement, a significant slope can be observed without it [32]. The uncertainty on the JVF efficiency is around $\pm 1\%$, considerably smaller than the other uncertainties on bin migration, so it is not considered in the statistical analysis.

b -jets are tagged with the help of the MV1 algorithm [184], which is a derivation of the JetFitterCombNN algorithm (see Section 3.3.6), combined with primary vertex and impact parameter information, via a Neural Network technique. It requires 85% efficiency for the identification of b -jets. No additional details are given here since the spin analysis, contrarily to the rate analysis, uses only events with exactly zero jets, and therefore does not need to veto for the presence of b -jets.

5.3 Event preselection

After applying the GoodRunList, primary vertex and event cleaning criteria described above, the preselection and selection cuts are applied, as summarized in Table 5.2. The cuts are in common with both the rate and the 2011 BDT analyses, apart from some changes in the actual values (cf. Section 4.5). Specifically, the following cuts are applied:

1. **Opposite-sign, opposite-flavour leptons, p_T threshold, third lepton veto:** exactly two opposite-sign, opposite-flavour leptons must be present in the event, and at least one has to be trigger-matched. The p_T thresholds for the leading and subleading leptons are set to 25 GeV and 15 GeV, respectively.
2. **$m_{\ell\ell} > 10 \text{ GeV}$:** the invariant mass of the dilepton system is required to be above 10 GeV.
3. **$E_{T,\text{rel}}^{\text{miss}} > 10 \text{ GeV}$:** the cut on $E_{T,\text{rel}}^{\text{miss}}$ aims mostly to suppress the $Z/\gamma^* + \text{jets}$ back-

ground, which is reduced by about 80%. Before this cut, the $t\bar{t}$ and $Z/\gamma^* + \text{jets}$ contributions dominate the background composition, while afterwards $Z/\gamma^* + \text{jets}$ becomes comparable to WW . The $E_{T,\text{rel}}^{\text{miss}}$ requirement halves the total background and has an impact also on the signal, leaving the S/B ratio almost unchanged (from 0.4% to 0.6%).

The cut on $E_{T,\text{rel}}^{\text{miss}}$ is looser than the corresponding one used in the rate analysis, namely 20 GeV rather than 25: this is motivated by the comparison in Fig. 5.1 (a), which shows the $E_{T,\text{rel}}^{\text{miss}}$ distribution for the $J^P = 0^+$ and $J^P = 2^+$ signals, after the dilepton requirement. The flat behaviour of $J^P = 2^+$ at low missing transverse energy values, $E_{T,\text{rel}}^{\text{miss}} \lesssim 40$ GeV, suggests to loosen the cut on this quantity as much as possible, in order to preserve the $J^P = 2^+$ signal; the choice of 20 GeV is a compromise between this attempt, the recommendations of the ATLAS working groups on the “combined performance” and the need to suppress the $Z \rightarrow \tau\tau$ background. The same motivation leads to loosen the cuts on $p_{T,\ell\ell}$, $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$ later in the selection (cf. Section 5.5.1).

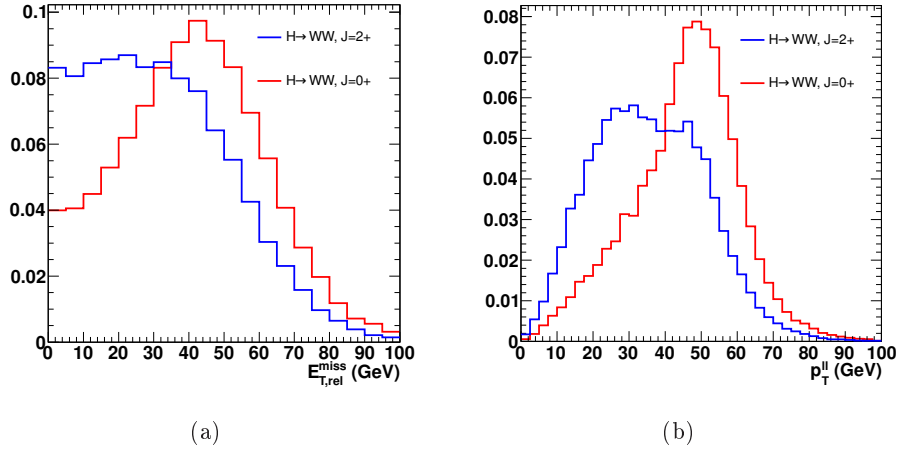


Figure 5.1: Distributions, normalized to unity, of (a) $E_{T,\text{rel}}^{\text{miss}}$ and (b) $p_{T,\ell\ell}$, for the SM Higgs boson (red) and the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ hypothesis (blue). The distributions are drawn after requiring two opposite-sign leptons.

After the $E_{T,\text{rel}}^{\text{miss}}$ cut, the requirement for exactly zero jets is applied, due to two reasons: on one hand, most of the sensitivity comes from the 0-jet channel, as discussed in Chapter 4 for the 2011 analysis, on the other hand the VBF production for alternative spin hypotheses is not implemented in the JHU generator (see Section 5.1). For what concerns the choice of restricting the analysis to the opposite-flavour channels, it can be understood looking at the jet multiplicity distribution for the same-flavour channels in Fig. 5.2 (b), which is drawn after the preselection defined

	0^+ [125 GeV]	2^+ [125 GeV]	WW	$WZ/ZZ/W\gamma$	$t\bar{t}$	Single Top	$Z/\gamma^*+\text{jets}$	$W+\text{jets}$	Total Bkg.	Observed	Data/MC
lepton p_T	422.44 ± 2.32	391.87 ± 1.59	10153.22 ± 25.16	3208.09 ± 41.84	51021.44 ± 79.41	5239.91 ± 30.30	29338.73 ± 112.84	2504.04 ± 19.69	101465.43 ± 150.76	107737	1.06 ± 0.00
OS leptons	421.98 ± 2.32	391.39 ± 1.59	10128.22 ± 25.13	1580.12 ± 28.66	50899.97 ± 79.32	5197.84 ± 30.17	28821.87 ± 104.61	2504.04 ± 19.69	99132.06 ± 141.37	103713	1.05 ± 0.00
$m_{\ell\ell} > 10$ GeV	418.34 ± 2.31	390.30 ± 1.59	10117.76 ± 25.11	1549.43 ± 28.43	50859.54 ± 79.28	5193.20 ± 30.15	28810.90 ± 104.61	2496.51 ± 19.67	99027.34 ± 141.29	103579	1.05 ± 0.00
Scale factors			NF = 1.08		NF = 1.05	NF = 1.05	NF = 0.91				
$E_{T,\text{rel}}^{\text{miss}} > 20$ GeV	307.82 ± 1.99	235.78 ± 1.23	7710.95 ± 22.79	912.39 ± 23.24	34206.27 ± 66.55	3725.61 ± 26.13	5795.38 ± 45.93	1039.31 ± 11.49	53389.91 ± 91.72	53865	1.01 ± 0.00
Scale factors			NF = 1.08		NF = 1.07	NF = 1.07	NF = 0.92				
jet veto	175.63 ± 1.50	130.85 ± 0.91	5078.77 ± 18.57	421.01 ± 13.53	539.63 ± 8.28	329.57 ± 7.63	4181.31 ± 41.36	580.96 ± 8.39	11131.25 ± 49.35	11493	1.03 ± 0.01
$p_{T,\ell\ell} > 20$ GeV	170.14 ± 1.48	118.86 ± 0.87	4622.63 ± 17.71	378.38 ± 12.63	514.89 ± 8.09	314.47 ± 7.46	667.02 ± 24.02	455.53 ± 6.58	6952.93 ± 34.85	7261	1.04 ± 0.01
$m_{\ell\ell} < 80$ GeV	169.60 ± 1.48	116.92 ± 0.86	2212.68 ± 12.23	230.67 ± 8.78	178.35 ± 5.00	123.66 ± 4.63	610.19 ± 22.83	303.64 ± 5.62	3659.18 ± 28.73	3964	1.08 ± 0.02
$\Delta\phi_{\ell\ell} < 2.8$	168.16 ± 1.47	113.38 ± 0.85	2190.39 ± 12.17	226.91 ± 8.76	176.12 ± 4.97	122.15 ± 4.61	291.05 ± 14.99	276.67 ± 4.62	3283.28 ± 22.73	3615	1.10 ± 0.02
$Z \rightarrow \tau\tau$ control region	3.60 ± 0.21	9.13 ± 0.24	95.31 ± 2.42	11.73 ± 1.22	5.09 ± 0.80	2.72 ± 0.69	3532.67 ± 37.48	94.57 ± 5.13	3742.09 ± 37.94	3474	0.93 ± 0.02
WW control region	0.54 ± 0.09	1.95 ± 0.11	2225.92 ± 11.83	147.71 ± 9.08	336.54 ± 6.36	190.82 ± 5.84	56.83 ± 7.49	151.89 ± 3.43	3109.71 ± 19.10	3297	1.06 ± 0.02

Table 5.2: The expected numbers of $J^P = 0^+$, $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ and background events after the requirements listed in the first column, as well as the observed numbers of events in data. The rightmost column shows the ratio between the data and the MC (background-only) estimation. Both signal hypotheses are normalized to the SM Higgs boson cross section, for $m_H = 125$ GeV. The backgrounds are estimated from their MC predictions, apart from the $W+\text{jets}$ background, which is data-driven, and the WW and $Z/\gamma^*+\text{jets}$ backgrounds, whose MC predictions are normalized to the data in their respective control regions (which can be seen, before such normalization, in the lower part of the Table). The normalizations factors (NFs) are obtained before the fit. Only statistical uncertainties are shown.

for the rate analysis (the same as discussed before for the spin analysis, with an additional cut $E_{T,\text{rel}}^{\text{miss}} > 45$ GeV). Even after such a tight cut on the missing transverse energy (which would actually be *too* tight for the spin analysis, heavily reducing the $J^P = 2^+$ and $J^P = 1^\pm$ signal yields), the background in the 0-jet bin is dominated by Z/γ^* +jets events, due to the increased pile-up in the 2012 data. The rate analysis employs also the same-flavour channels thanks to the development of a dedicated technique (the PacMan method [234]), which is tailored on the $J^P = 0^+$ signal region, and therefore cannot be applied to a generic spin hypothesis. For all these reasons, the spin analysis presented here uses the $e\mu$ 0-jet channel only.

After the jet veto requirement, the total background, whose main components are WW and Z/γ^* +jets at this stage, is reduced by about 80%, while the $J^P = 0^+$ signal is only reduced by 50%. The resulting S/B ratio is 1.6%.

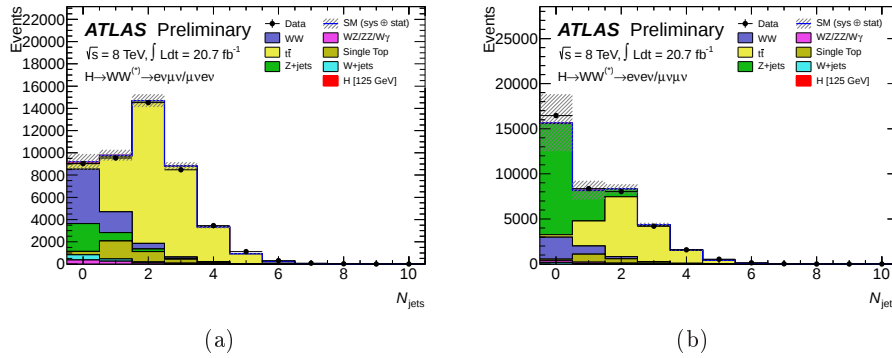


Figure 5.2: Distributions of the jet multiplicities at the end of the preselection defined for the rate analysis, in (a) the opposite-flavour channels and (b) the same-flavour channels. At this stage, the SM Higgs boson signal is too small to be seen. The shaded areas represent the total normalization uncertainties [234].

5.4 BDT training

Details on the BDT technique can be found in Section 4.6 and in Ref. [216]; the spin analysis follows mostly what developed for the 2011 BDT analysis presented in Chapter 4, so only the differences are discussed in the following.

The choice of the training configuration has undergone a detailed optimization. At first, a standard $J^P = 0^+$ versus background BDT was trained, adding a few kinematical variables to the four used in 2011 (as shown in Appendix B, with the available number of MC events a saturation in BDT performance is reached already with four variables, so it is not useful to add many others). Using spin-dependent variables, however, could introduce an a-priori asymmetry towards one of the two

hypotheses, if only $J^P = 0^+$ is used for training. To avoid this, several alternative training configurations have been tried, removing one or more of the spin-dependent dilepton variables ($\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, $p_{T,\ell\ell}$) from the inputs, and then fitting one of these latter variables to compute the final significances, after a cut on the BDT response. As documented in Appendix J, all alternative training configurations explored could not reach the expected (i.e. MC only) sensitivities for discriminating $J^P = 0^+$ from $J^P = 2^+$ as the standard four variables. The cut on the BDT output, needed to reject the backgrounds, in general reduces the discrimination power of the variable to be fitted. Given these observations, the standard four variables have been chosen again, but now two independent trainings are performed:

1. the SM Higgs boson signal sample is trained against the total background, obtaining the “BDT0” response,
2. in the same way, the alternative signal sample is trained against the total background, obtaining the “BDT2” response⁽²⁾.

With this procedure, both spin hypotheses are treated at the same level, without any bias towards one of the two.

The BDT responses are then fitted in their 2-dimensional plane, as discussed later. The various background components are expected to get classified as background-like in both trainings, thus obtaining low values for both BDT responses, while the two signal hypotheses are expected to get larger values, possibly with some shape difference in the 2D plane, which allows for their discrimination, as shown schematically in Fig. 5.6.

5.4.1 Input variables

Out of the four variables, $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$ and $p_{T,\ell\ell}$ are expected to discriminate at the same time between the different signal hypotheses and between the signals and the background, while m_T has a stronger background rejection power but weaker spin discrimination. Their shapes are shown in Fig. 5.3, which compares the $J^P = 0^+$ hypothesis with the alternative ones after the jet veto cut has been applied, and in Fig. 5.4, which compares the $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ signals with the main background components, after the jet veto and $p_{T,\ell\ell}$ requirements.

After the preselection and jet veto requirements, two additional cuts are applied before the training:

- $p_{T,\ell\ell} > 20$ GeV,

²In this case, the label “2” does not refer to the $J^P = 2^+$ model only, but rather is a generic indication of the alternative spin hypothesis. Occasionally, the name BDT1 is used, referring specifically to the $J^P = 1^\pm$ analyses.

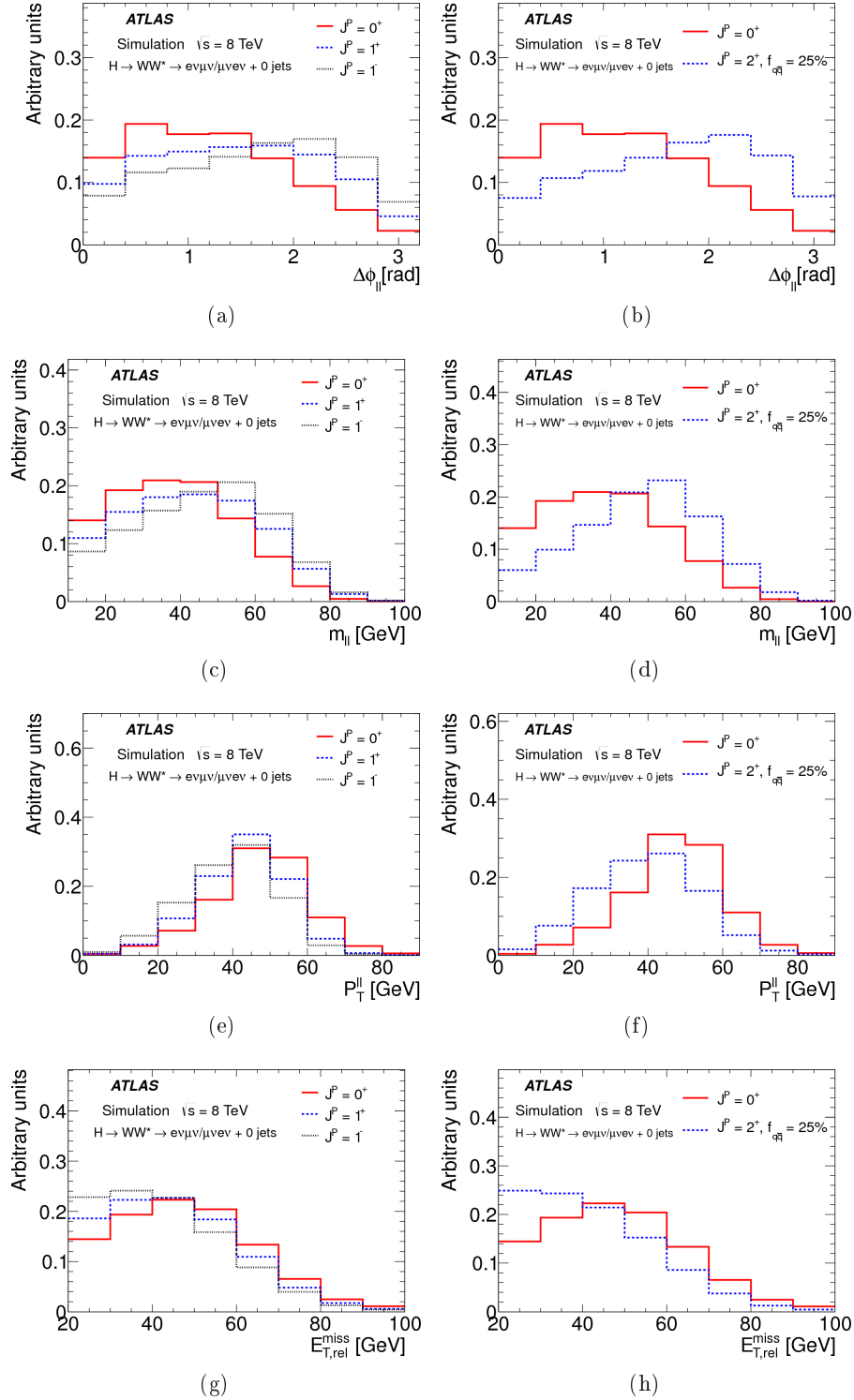


Figure 5.3: Distributions, normalized to unity, of (a,b) $\Delta\phi_{\ell\ell}$, (c,d) $m_{\ell\ell}$, (e,f) $p_{T,\ell\ell}$ and (g,h) $E_{T,\text{rel}}^{\text{miss}}$, comparing the SM Higgs boson (solid red line) to the $J^P = 1^{\pm}$ hypotheses (left column, 1^+ dashed blue line, 1^- dashed black line) and to the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ hypothesis (right column, dashed blue line). The distributions are shown after requiring two opposite-sign leptons, $E_{T,\text{rel}}^{\text{miss}} > 20$ GeV and zero jets.

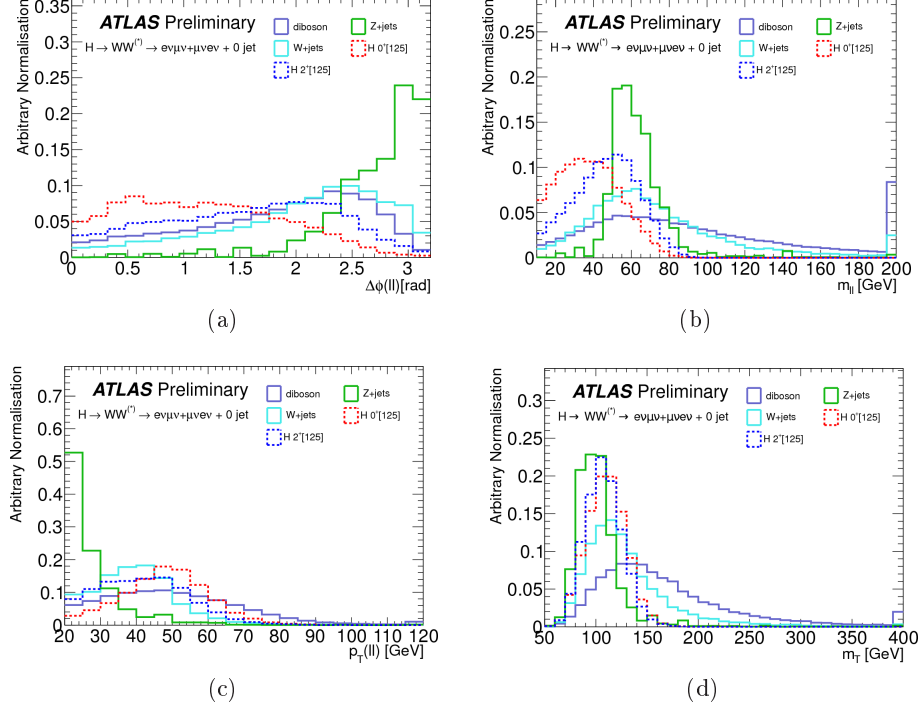


Figure 5.4: Distributions, normalized to unity, of (a) $\Delta\phi_{\ell\ell}$, (b) $m_{\ell\ell}$, (c) $p_{T,\ell\ell}$ and (d) m_T , comparing the SM Higgs boson (dashed red line) to the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ hypothesis (dashed blue line) and to the main background components (dibosons including WW , W +jets, Z/γ^* +jets). The distributions are shown after the jet veto and $p_{T,\ell\ell}$ requirements.

- $m_{\ell\ell} < 100$ GeV.

The cut on the dilepton transverse momentum is looser than the one applied in both the 2011 BDT and 2012 rate analyses (30 GeV), for the reasons discussed in Section 5.3. The 100 GeV threshold on $m_{\ell\ell}$, moreover, is set in order to restrict the BDT training to the phase space where the signals are present (cf. Fig. 5.4 (b)).

The linear correlation coefficients among the input variables are shown in Fig. 5.5, for the $J^P = 0^+$, $J^P = 2^+$ and background cases. The pattern is very similar for $J^P = 0^+$, $J^P = 2^+$ and also $J^P = 1^\pm$ (the latter is not shown here), with large positive correlations between $\Delta\phi_{\ell\ell}$ and $m_{\ell\ell}$.

The ranking order of input variables is:

- $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, m_T and $p_{T,\ell\ell}$ for the $J^P = 0^+$ training,
- $\Delta\phi_{\ell\ell}$, m_T , $m_{\ell\ell}$ and $p_{T,\ell\ell}$ for the $J^P = 2^+$ training.

As a side note, the ranking for the $J^P = 2^+$ training is the same as found in 2011 for the $e\mu$ channel (see Section 4.6.2).

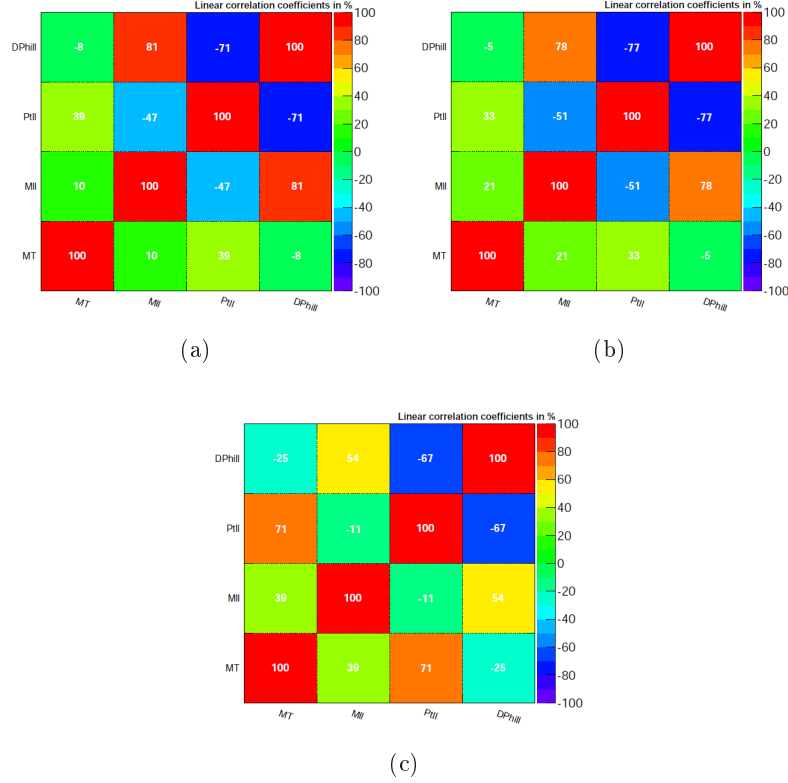


Figure 5.5: Linear correlation coefficients among the BDT input variables for (a) $J^P = 0^+$, (b) $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ and (c) background events.

5.4.2 BDT settings

Table 5.3 summarizes the BDT settings which have been changed with respect to the TMVA defaults; most are in common with the 2011 analysis (cf. Table 4.3). The various choices are motivated by the dual need of optimizing and stabilizing the performance of the BDT algorithm, reducing the impact of statistical fluctuations. From the technical point of view, training and classification use the latest TMVA version (v4.1.3), which is part of the ROOT v5.34 release.

In particular, given the limited number of signal events for the JHU samples, the number of trees is decreased to 400 (from 1000 as in 2011), while the minimum number of events in each leaf is increased to 2% of signal events (rather than 1%), corresponding to about 2000 events.

All background samples are used in the training, including the W +jets data-driven events, contrarily to what done for the 2011 analysis (cf. Section 4.6.2); no overtraining is observed. In all cases, only the even-numbered events are used for the training, while the odd-numbered ones are used for the classification and to extract

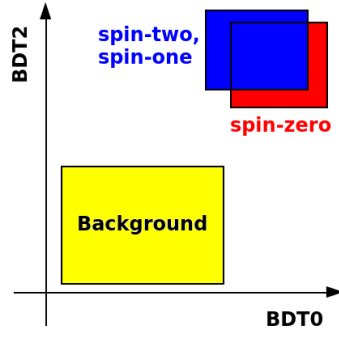


Figure 5.6: Sketch illustrating the 2D fit concept, with the expected localizations of the total background and of the various signal hypotheses.

BDT setting	Value
BoostType	Gradient
MaxDepth	3
NTrees	400
nEventsMin	2% of signal events
IgnoreNegWeightsInTraining	True
UseBaggedGrad	True
GradBaggingFraction	0.5
Shrinkage	0.1

Table 5.3: BDT settings used in the spin analysis.

the final results.

5.5 BDT Output

5.5.1 BDT output in the signal region

After the preselection cuts described in Section 5.3 and the jet veto, the signal region is defined by adding the following requirements:

- $p_{T,\ell\ell} > 20$ GeV,
- $m_{\ell\ell} < 80$ GeV,
- $\Delta\phi_{\ell\ell} < 2.8$.

Table 5.2 shows the expected numbers of signal and background events after the preselection, as well as the observed numbers of events in data. All these cuts have been loosened with respect to the rate analysis, as discussed in Section 5.3; the actual comparison is shown in Table 5.4.

The cut on the dilepton transverse momentum targets the $Z/\gamma^* + \text{jets}$ background, removing more than 80% of it, thus halving the total background. Subsequently, the cut on $m_{\ell\ell}$ is tightened with respect to the one used in training (100 GeV), in order to remove 50% of WW and sizeable amounts of all other backgrounds. Also in this case, the total background is reduced by a factor two. Finally, a loose cut on $\Delta\phi_{\ell\ell}$ allows to remove more than half of the remaining $Z/\gamma^* + \text{jets}$ events, due to the fact that, for this background, leptons are emitted preferentially back-to-back (see e.g. Section 4.2.3).

At the end of the selection, 168 events are expected for $J^P = 0^+$ and 113 for $J^P = 2^+$ $f_{q\bar{q}} = 0.25$: despite the facts that the two samples have been normalized initially to the same cross section, and the selection has been loosened, the acceptance for $J^P = 2^+$ events is smaller than for $J^P = 0^+$ events, namely 67%. This fact, however, has no consequences on the fit, since both signal normalizations are floated.

Most of the background is constituted by WW events, with minor contributions from $Z/\gamma^* + \text{jets}$, $W + \text{jets}$ and other components. In total, 3283 background events are expected (3% of the initial yield, with a S/B ratio of 5%) while 3615 candidates are observed in data, yielding a data/MC (background-only) ratio of 1.10 ± 0.02 (stat).

Fig. 5.7 shows the distributions of the input variables in the signal region, with the SM $J^P = 0^+$ signal stacked on top of the total background. Looking in particular at $p_{T,\ell\ell}$ and m_T , it can be seen how the predicted SM rate is not enough to account for the excess in data, thus anticipating an observed signal strength $\hat{\mu}$ larger than one. In Appendix O, the distributions of the input variables are compared between the $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ hypotheses (Figs. O.2 and O.3), while Fig. O.1 shows the shapes of the input variables for all spin-2 fractions. Already visually, a SM Higgs boson seems easier to reconcile with the excess found in data than the alternative $J^P = 2^+$ hypotheses.

Variable	Spin Analysis	Rate Measurement
$E_{T,\text{rel}}^{\text{miss}}$	> 20 GeV	> 25 GeV
N_{jets}	0 jets	0, 1, ≥ 2 jet selections
$p_{T,\ell\ell}$	> 20 GeV	> 30 GeV
$m_{\ell\ell}$	< 80 GeV	< 50 GeV
$\Delta\phi_{\ell\ell}$	< 2.8	< 1.8

Table 5.4: List of selection cuts applied for the spin (left) and rate (right) analyses, after the common preselection.

Spin-2 analysis

Fig. 5.8 shows the BDT responses in the signal region, with the $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ hypotheses alternatively superimposed to the total background. Both signals are normalized to the expected SM cross section.

BDT0 exhibits the best discrimination between the two signal hypotheses, while BDT2 helps to disentangle both signals from the background. This can be seen explicitly in Fig. 5.9, where the shapes of the two signals and of the total background are compared: BDT0 is much flatter for $J^P = 2^+$ than for $J^P = 0^+$, while BDT2

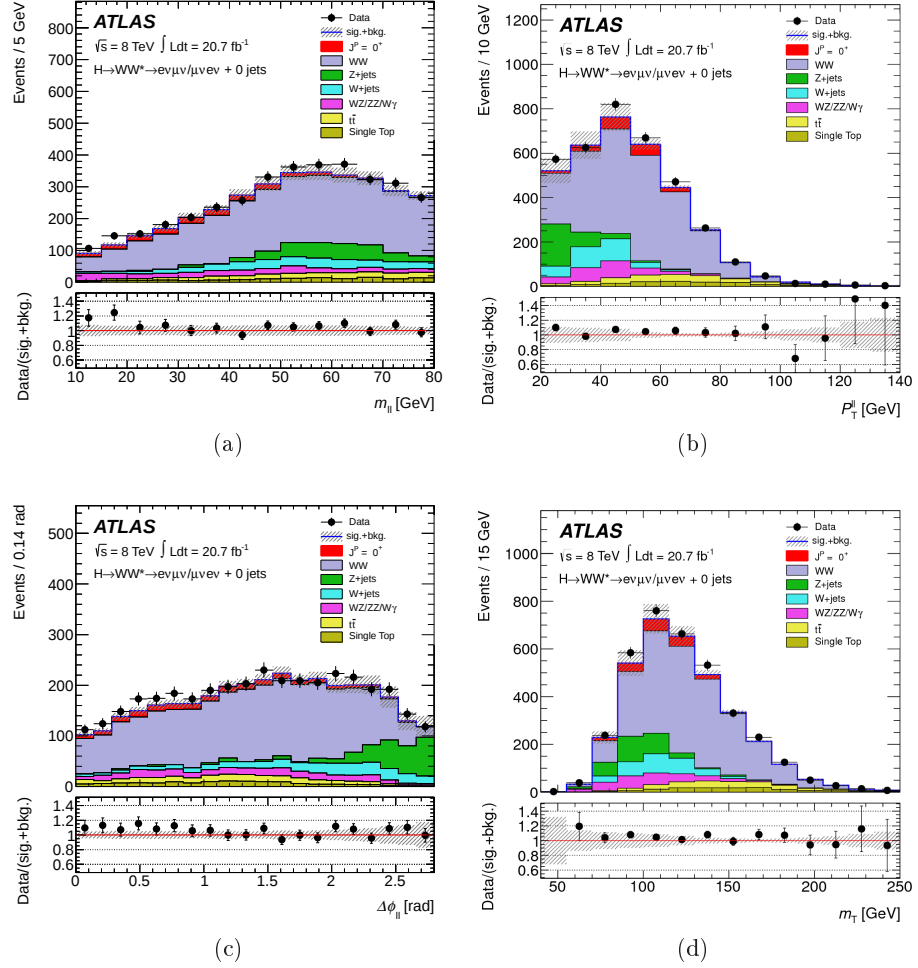


Figure 5.7: Distributions of the four input variables in the signal region: (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . The SM Higgs boson ($m_H = 125$ GeV) is shown on top of the total background. All backgrounds are normalized as in Table 5.2. The shaded bands indicate the total normalization systematic uncertainty. In the lower part of the figures, the ratio between the data and the sum of signal (with $\mu = 1$) and background events is shown.

shows similar shapes. Hence, the 2-dimensional fit exploits more information than the simple 1-dimensional fit to BDT0.

The shapes of the BDT responses for all fractions are compared in Fig. 5.10; the distribution of BDT0 gets flatter when increasing $f_{q\bar{q}}$, thus more different with respect to the $J^P = 0^+$ case, and also the BDT2 shape changes substantially according to the production mode.

Fig. 5.11 shows the 2-dimensional correlation between BDT0 and BDT2 in the signal region, for the two signals, the total background and the data. As expected (cf. Fig. 5.6), a substantial part of the background events gets a low response in

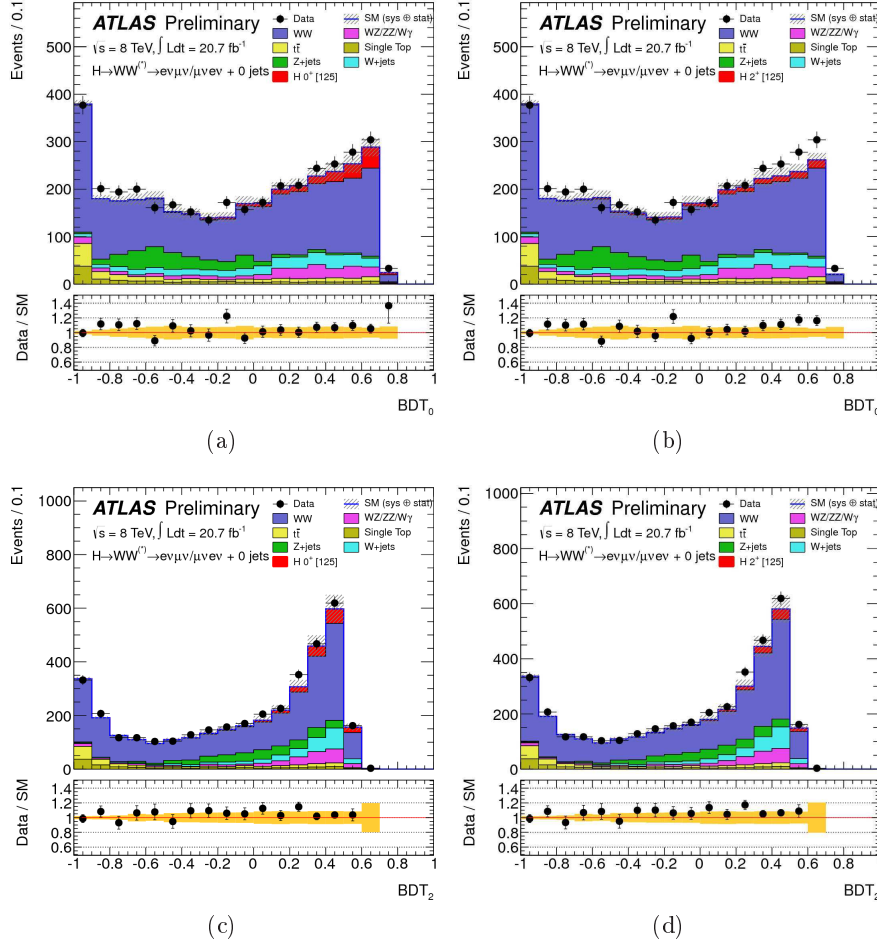


Figure 5.8: Distributions of the BDT responses for the $J^P = 2^+$ analysis, with $f_{q\bar{q}} = 0.25$, in the signal region. The upper row shows the BDT0 distribution with (a) the $J^P = 0^+$ signal or (b) the $J^P = 2^+$ signal superimposed to the total background, while the lower row shows the BDT2 distribution with (c) the $J^P = 0^+$ signal or (d) the $J^P = 2^+$ signal superimposed. Both signals are scaled to the expected SM cross section. All backgrounds are normalized as in Table 5.2. The shaded areas represent the total normalization uncertainties on the sum of the signal and background yields; the “SM” label in the ratio boxes refers to the sum as well.

both BDT0 and BDT2, and the same holds for the data, while the two signals get larger responses, with some differences that allow for discriminating one from the other.

Spin-1 analysis

The same distributions shown for the $J^P = 2^+$ case, are given here for the $J^P = 1^\pm$ analyses: the BDT responses in the signal region for the $J^P = 1^+$ analysis in Fig.

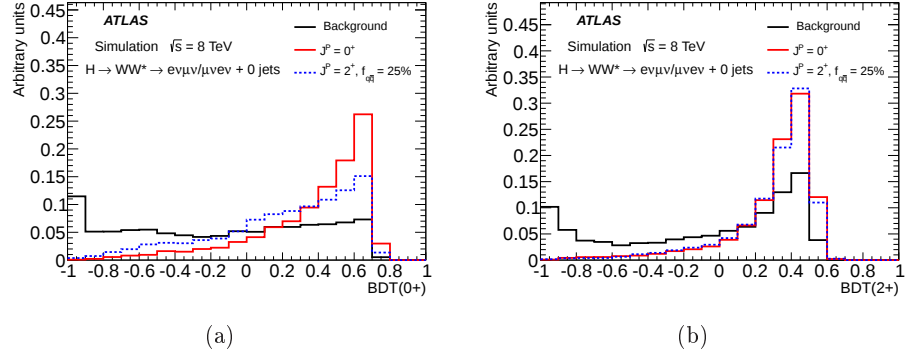


Figure 5.9: Comparison of shapes (normalized to unity) for (a) the BDT0 and (b) the BDT2 outputs for the $J^P = 0^+$ signal (solid red line), the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ signal (dotted black line) and the total background (solid blue line).

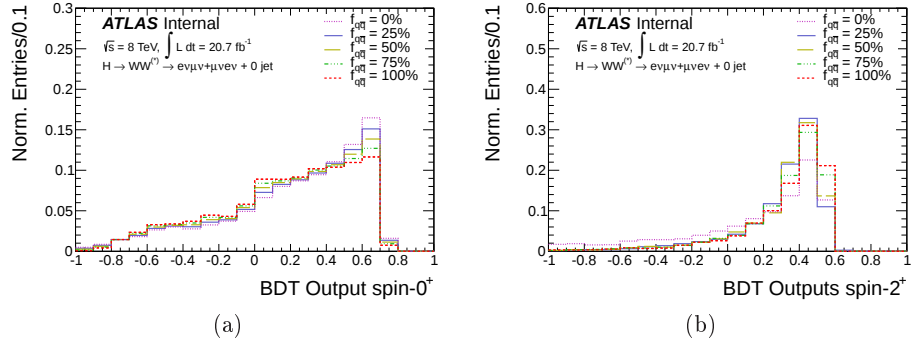


Figure 5.10: Shape distributions, normalized to unity, for the output variables, (a) BDT0 and (b) BDT2, comparing the five $J^P = 2^+$ fractions.

5.12 and for the $J^P = 1^-$ analysis in Fig. 5.13, while Fig. 5.14 contains the 2-dimensional correlations for $J^P = 1^+$. The same comments as made in the $J^P = 2^+$ case mostly hold for $J^P = 1^\pm$; as expected from the input distributions shown in Fig. 5.3, the shapes of the BDT responses look more similar between the $J^P = 0^+$ and $J^P = 1^\pm$ signals, than between the $J^P = 0^+$ and $J^P = 2^+$ signals.

5.5.2 WW control region

The normalization of the WW background is obtained through the fit in a dedicated control region, defined inverting the cut on $m_{\ell\ell}$ and removing that on $\Delta\phi_{\ell\ell}$:

- $p_{T,\ell\ell} > 20$ GeV,
- $m_{\ell\ell} > 80$ GeV.

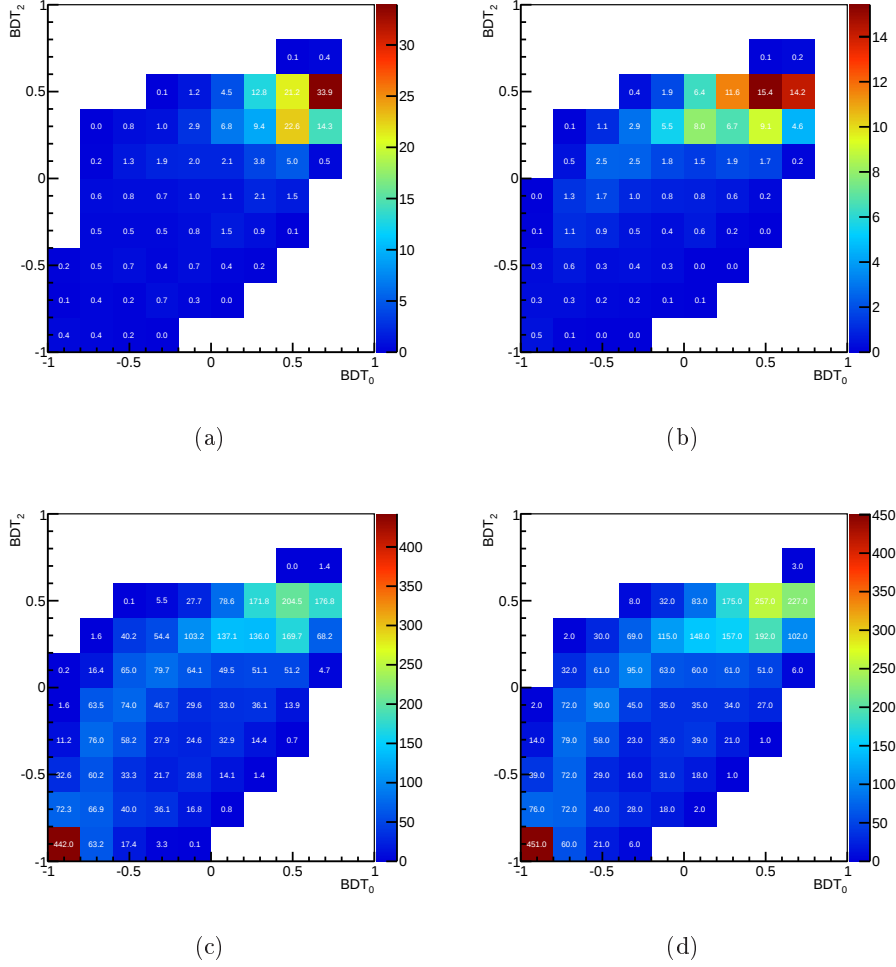


Figure 5.11: 2-dimensional distributions in the (BDT0, BDT2) plane for the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ analysis, in the signal region. The upper row shows the signals ((a) $J^P = 0^+$, (b) $J^P = 2^+$), the bottom row shows the total background (c) and the data candidates (d).

This is different from what done in the rate analysis, where, given the signal region requirement of $m_{\ell\ell} < 50$ GeV, the cut $50 \text{ GeV} < m_{\ell\ell} < 100$ GeV is applied in the WW control region.

A good data/MC agreement is found, as can be seen in Figs. 5.15 and 5.16 for the $J^P = 2^+$ and $J^P = 1^\pm$ analyses, respectively; the BDT responses are found to be squeezed to the left side in all cases, signaling a good capability of the algorithm to distinguish between signal- and background-like topologies. The corresponding number of events in MC and in data is shown in Table 5.2: the WW background dominates as expected, with some contributions from the top-quark backgrounds, while the signal yield is negligible. As explained in Section 4.7.5, a normalization

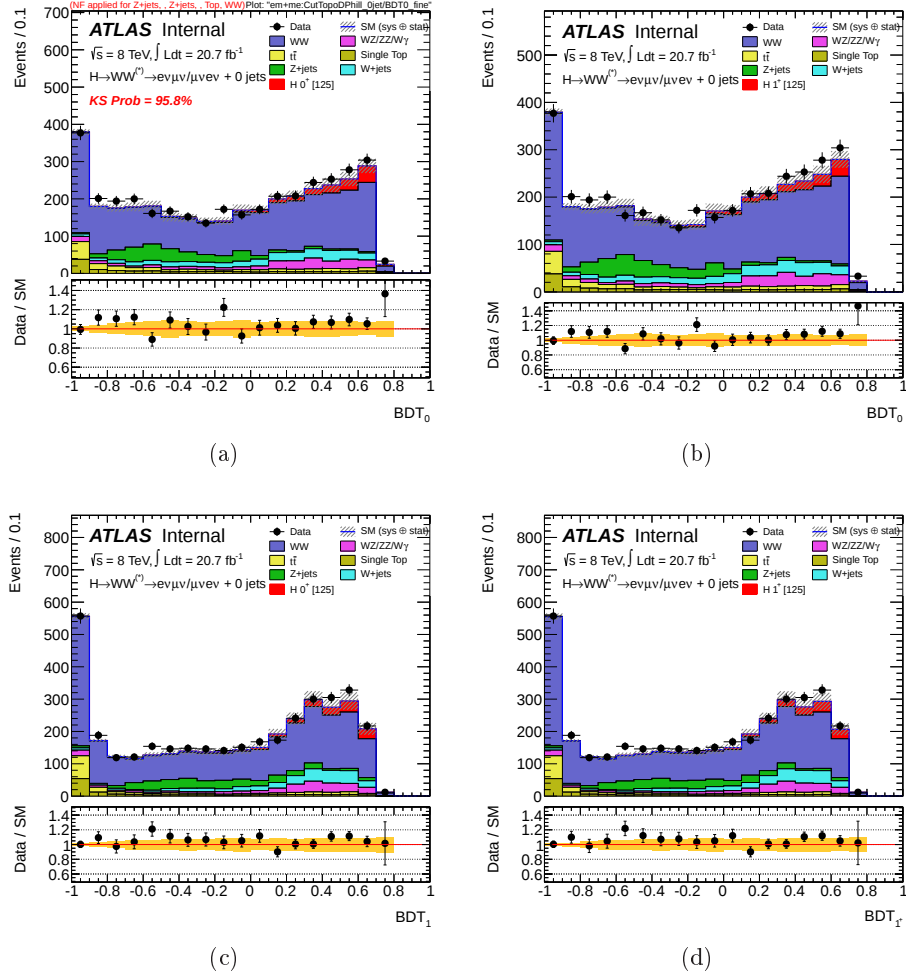


Figure 5.12: Distributions of the BDT response for the $J^P = 1^+$ analysis, in the signal region. The upper row shows the BDT0 distribution with the (a) $J^P = 0^+$ signal or (b) the $J^P = 1^+$ signal superimposed to the total background, while the lower row shows the BDT2 distribution with (c) the $J^P = 0^+$ signal or (d) the $J^P = 1^+$ signal superimposed. Both signals are scaled to the expected SM cross section. All backgrounds are normalized as in Table 5.2. The shaded areas represent the total normalization uncertainties on the sum of signal and background yields; the “SM” label in the ratio boxes refers to the sum as well.

factor, α_{WW} , is extracted from the ratio between WW_{data} and WW_{MC} , where the former is obtained by subtracting from the data the sum of all backgrounds which are not WW itself: the result is $\alpha_{WW} = 1.09 \pm 0.04$, compatible with the cut-based value of $\alpha_{WW}^{\text{c.b.}} = 1.16 \pm 0.04$, and with the fitted values which can be found in Tables 5.13 and 5.18, for the $J^P = 2^+$ and $J^P = 1^\pm$ fits, respectively.

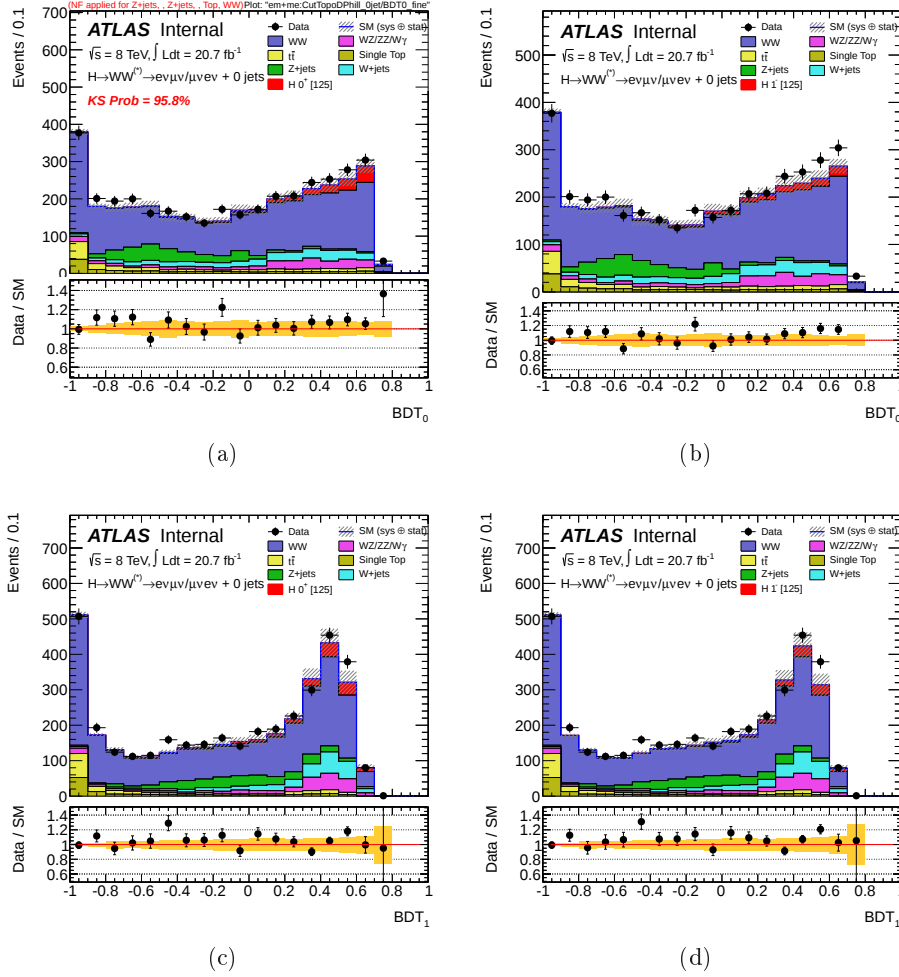


Figure 5.13: Distributions of the BDT response for the $J^P = 1^-$ analysis, in the signal region. The upper row shows BDT0 with (a) the $J^P = 0^+$ signal or (b) the $J^P = 1^-$ signal superimposed to the total background, while the lower row shows BDT1 with (c) the $J^P = 0^+$ signal or (d) the $J^P = 1^-$ signal superimposed. Both signals are scaled to the expected SM cross section. All backgrounds are normalized as in Table 5.2. The shaded areas represent the total normalization uncertainties on the sum of the signal and background yields; the “SM” label in the ratio boxes refers to the sum as well.

5.5.3 $Z \rightarrow \tau\tau$ control region

Unlike the 2011 analysis, the spin analysis uses a $Z/\gamma^* + \text{jets}$ control region, which is actually a $Z \rightarrow \tau\tau$ control region (with both the τ leptons decaying leptonically). This is defined inverting the $\Delta\phi_{\ell\ell}$ requirement since, as already noticed, the leptons from Z boson decays tend to go back-to-back:

- $m_{\ell\ell} < 80 \text{ GeV}$,

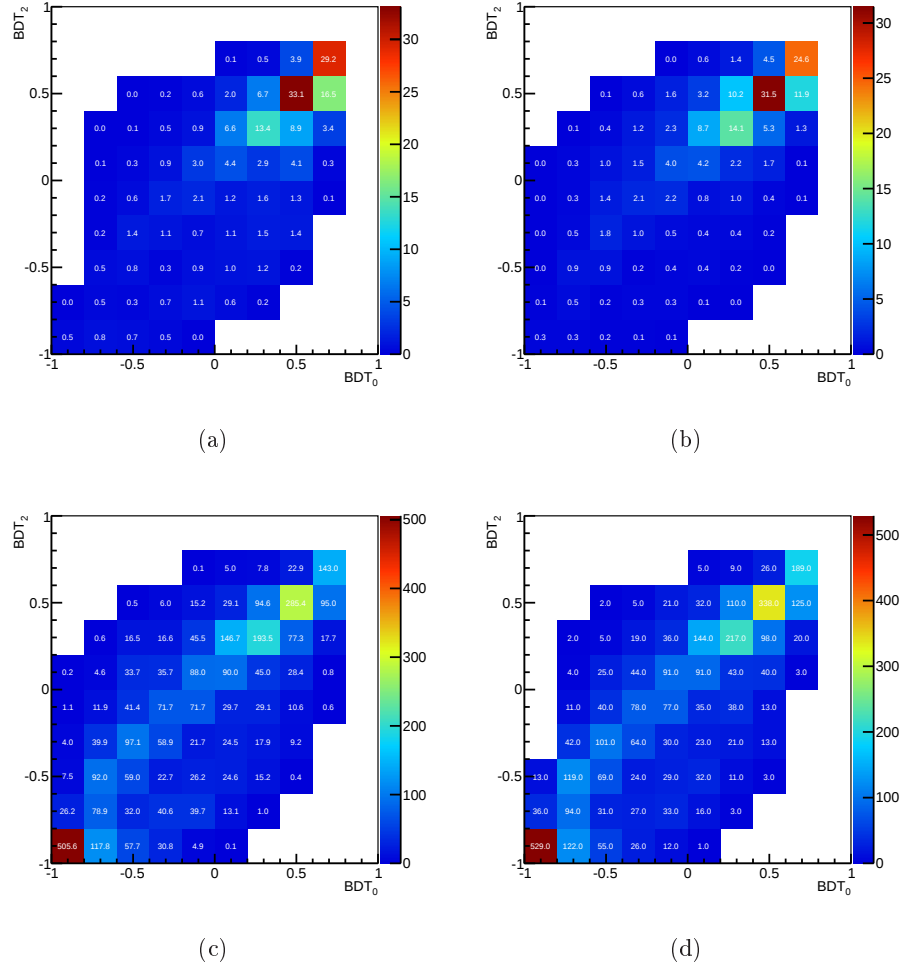


Figure 5.14: 2-dimensional distributions in the $(\text{BDT}_0, \text{BDT}_2)$ plane responses for the $J^P = 1^+$ analysis, in the signal region: (a) $J^P = 0^+$, (b) $J^P = 1^+$, (c) total background, (d) data.

- $\Delta\phi_{\ell\ell} > 2.8$.

The $p_{T,\ell\ell}$ cut is not applied since it would have the same effect of the $\Delta\phi_{\ell\ell}$ one. There is no dedicated $Z \rightarrow \tau\tau$ control region in the rate analysis, since the tighter $E_{T,\text{rel}}^{\text{miss}}$ threshold removes this background more efficiently. As shown in Figs. 5.17 and 5.18 for the $J^P = 2^+$ and $J^P = 1^\pm$ analyses, respectively, the data and the MC predictions agree well for both the input and the output variables in this region. For plots and tables, normalization factors of 0.91 ± 0.11 and 0.92 ± 0.11 are used before and after the $E_{T,\text{rel}}^{\text{miss}}$ cut, respectively (cf. values in Table 5.2). The same values are retained by the fit (cf. Tables 5.13 and 5.18).

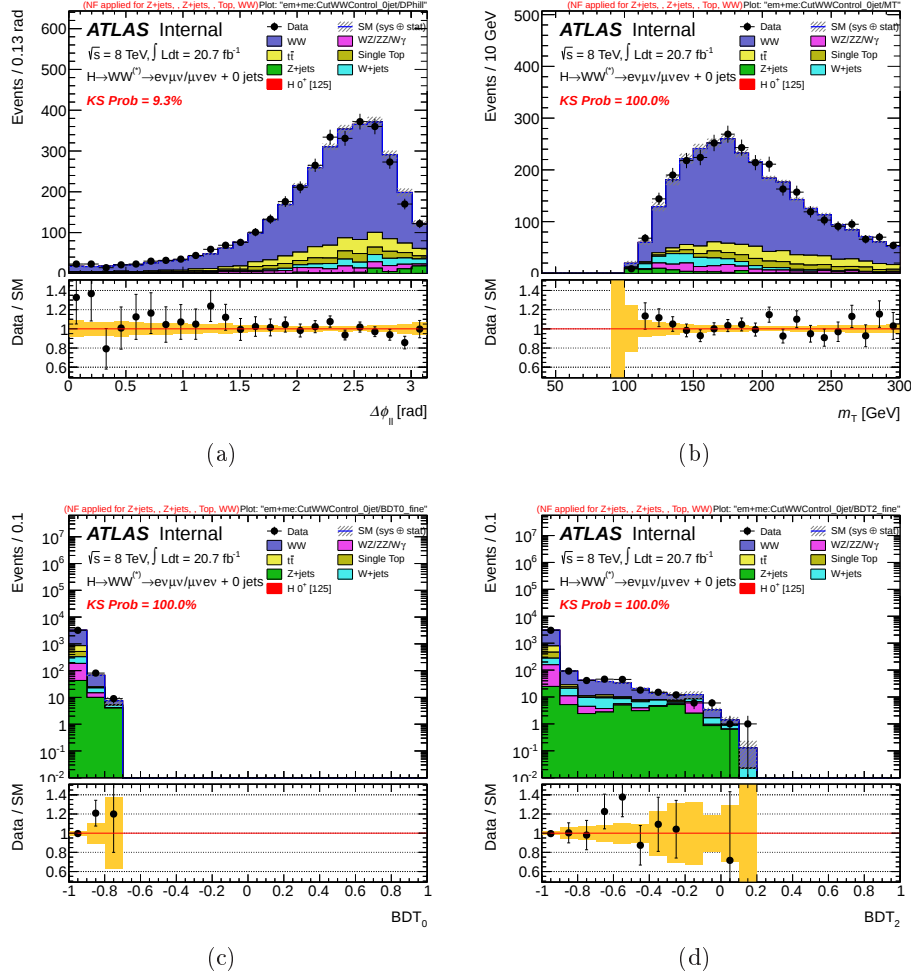


Figure 5.15: Distributions in the WW control region for the $J^P = 2^+ f_{q\bar{q}} = 0.25$ analysis: (a) $\Delta\phi_{\ell\ell}$, (b) m_T , (c) BDT_0 , (d) BDT_2 . All backgrounds are normalized as in Table 5.2, except the WW one, which corresponds to the pure MC prediction. The shaded areas represent the total normalization uncertainties. In the lower part of the figures the ratio between the data and the total background is shown; the signal is negligible in all cases.

5.5.4 Other backgrounds

Unlike the WW and $Z \rightarrow \tau\tau$ cases, the other backgrounds do not use dedicated control regions. The top-quark background is estimated using the same method outlined in Section 4.8.1. The correction to be applied to MC events, inherited from the rate analysis, is 1.07 in the signal region, as reported in Table 5.2. The W +jets background is estimated with the data-driven method illustrated in Section 4.8.3; the fake factors are also in this case of the order of 0.01 and 0.1, for electrons and muons, respectively, with a relative uncertainty of $\pm 50\%$ on the yield in the signal

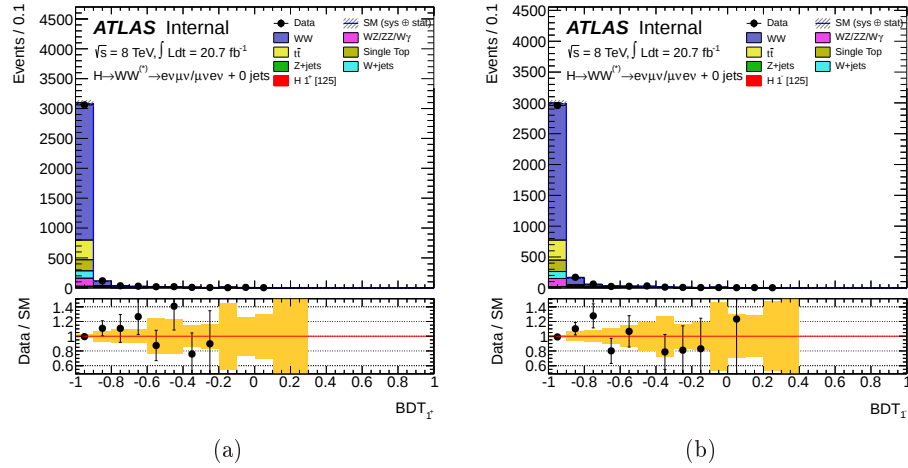


Figure 5.16: Distributions of the BDT responses in the WW control region for the $J^P = 1^\pm$ analyses: BDT1 (a) for $J^P = 1^+$ and (b) for $J^P = 1^-$. All backgrounds are normalized as in Table 5.2, except the WW one, which corresponds to the pure MC prediction. The shaded areas represent the total normalization uncertainties. In the lower part of the figures the ratio between the data and the total background is shown; the signal is negligible in all cases.

region.

5.6 Systematic uncertainties

Table 5.5 summarizes the definition and the treatment of the experimental, theoretical and statistical uncertainties. The “2012” tag labels the systematic components which are decorrelated between the 2011 and 2012 analyses (this is important for the rate analysis, which uses both datasets), while the “HWW” tag identifies systematic uncertainties which are analysis-specific and thus decorrelated among decay channels in the final combination. Most of the uncertainties are in common with the 2011 analysis, discussed in Section 4.9; differences are stressed.

5.6.1 Experimental uncertainties

As discussed in Section 4.9.1 for the 2011 analysis, experimental uncertainties are associated with leptons, jets, missing E_T , triggers and luminosity ($\pm 3.6\%$). Not all efficiency-related systematic uncertainties are treated as shape uncertainties, rather only those whose impact is not negligible. Another difference with respect to the 2011 analysis, is the introduction of an uncertainty on the $\langle\mu\rangle$ rescaling, which is applied to the MC events to reproduce different data-taking periods.

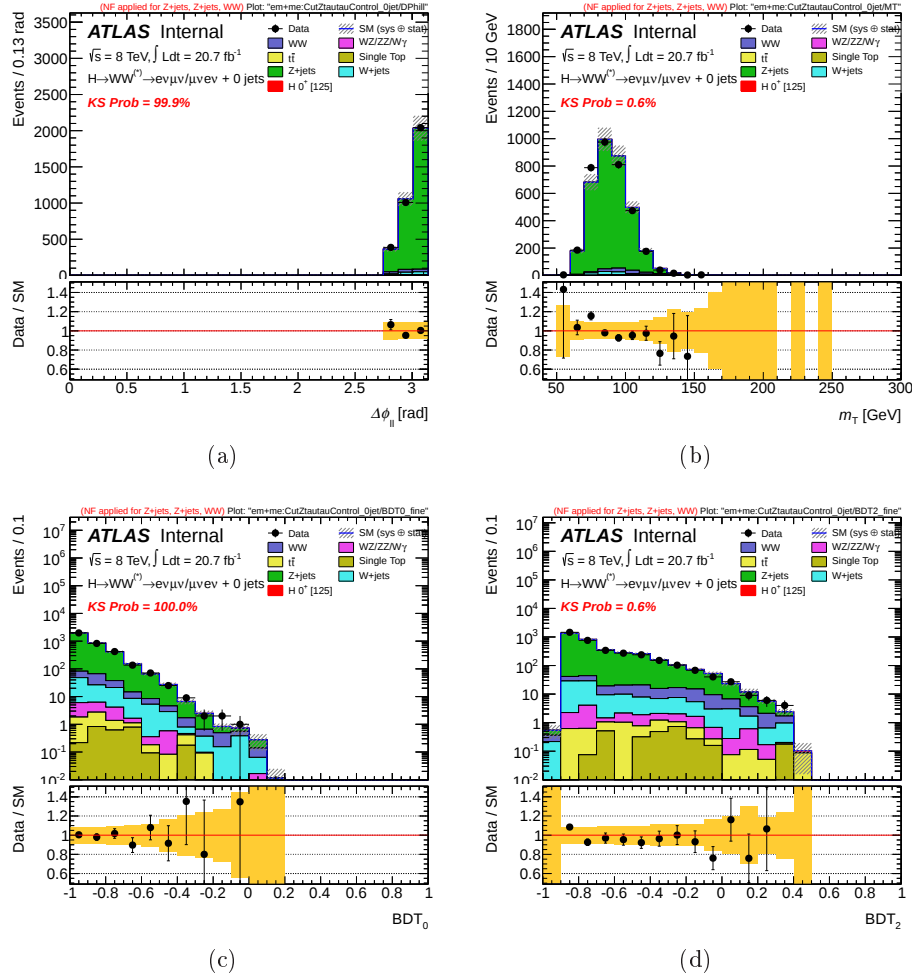


Figure 5.17: Distributions in the $Z \rightarrow \tau\tau$ control region for the $J^P = 2^+ f_{q\bar{q}} = 0.25$ analysis: (a) $\Delta\phi_{\ell\ell}$, (b) m_T , (c) BDT_0 , (d) BDT_2 . All backgrounds are normalized as in Table 5.2, except the $Z \rightarrow \tau\tau$ one, which corresponds to the pure MC prediction. The shaded areas represent the total normalization uncertainties. In the lower part of the figures the ratio between the data and the total background is shown; the signal is negligible in all cases.

Additional details on jet-, lepton- and missing E_T -related systematic uncertainties are given in the following.

Jet systematic uncertainties

The JES uncertainty is split into 15 independent components [237], which can be divided into the following categories: in-situ jet energy correction, η intercalibration, high- p_T jets, non-closure, close-by jets, flavour composition and response for non- b jets, b -jets, event pile-up. All together, the JES uncertainty ranges from $\pm 2\%$ to

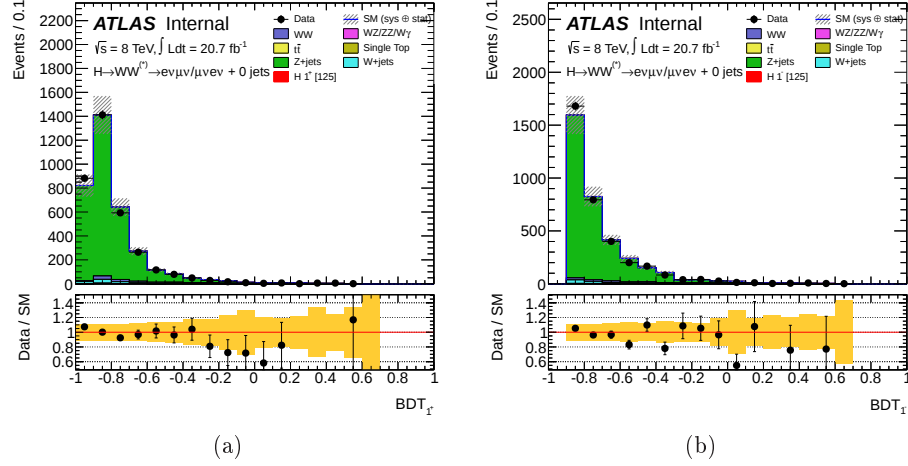


Figure 5.18: Distributions of the BDT responses in the $Z \rightarrow \tau\tau$ control region for the $J^P = 1^\pm$ analysis: (a) BDT1 for $J^P = 1^+$, (b) BDT1 for $J^P = 1^-$. All backgrounds are normalized as in Table 5.2. The shaded areas represent the total normalization uncertainties. In the lower part of the figures the ratio between the data and the total background is shown; the signal is negligible in all cases.

$\pm 9\%$, depending on η and p_T of the jet, and on the sample composition.

The uncertainty on the jet energy resolution is found to vary between $\pm 25\%$ and $\pm 5\%$, as a function of jet η and p_T , as for the 2011 analysis (cf. Section 4.9.1).

Finally, the uncertainty on the JVF criterium used for jet identification (see Section 5.2.5) is found to have a negligible impact on the analysis.

Lepton systematic uncertainties

Lepton systematic uncertainties follow what is described in Section 4.9.1 for the 2011 analysis, and are divided into uncertainties on the selection efficiency (EL_EFF and MU_EFF, for electrons and muons, respectively), on the energy scale (EL_ESCALE, MU_ESCALE) and on the resolution (EL_RES, MU_ID_RES, MU_MS_RES).

The first term includes the uncertainties on identification (for electrons $\pm 3\%$ for $p_T \sim 15$ GeV and decreasing above, for muons smaller than $\pm 1\%$ on the whole range), reconstruction (smaller than $\pm 1\%$) and trigger (around $\pm 2\%$), for a total impact between $\pm 2\%$ and $\pm 5\%$. For what concerns the energy scale and resolution, their uncertainties are both smaller than $\pm 1\%$, for both electron and muons.

Missing E_T systematic uncertainties

Also the treatment of missing E_T systematic uncertainties is similar to what is done in 2011 (Section 4.9.1), but less conservative, thanks to the knowledge acquired in

the meanwhile by the ATLAS collaboration [186] on the impact of soft objects (reconstructed jets with $p_T < 20$ GeV and soft depositions in the calorimeter) on the missing E_T computation. The uncertainty resides now on two independent terms: the first term is the resolution one (MET_RESOSOFT), which covers data/MC differences in the description of soft objects, and is a function of p_T and the number of primary vertices, the second term (MET_SCALESOFT) accounts for uncertainties on the soft energy scale, with a ± 1 GeV smearing of the scale itself.

5.6.2 Theoretical uncertainties

The theoretical uncertainties mostly follow what discussed in Section 4.9.2 for the 2011 analysis, with the caveat that only the uncertainties specific to the 0-jet bin are relevant for this analysis and are hence summarized here.

Concerning the uncertainties on the $J^P = 0^+$ production cross section, they are determined following the prescriptions in Refs. [87, 225, 238]. QCD scale uncertainties are obtained varying μ_R and μ_F up and down by a factor of two ($\mu_R, \mu_F = m_H/2, 2m_H$), while for PDFs two sources are taken into account: the envelope of error sets and the difference between the central values of different prescriptions. Variations of QCD scales and PDFs change the $J^P = 0^+$ cross section by $\pm 17\%$ and $\pm 8\%$, respectively.

Uncertainties on the $W\gamma$ cross section amount to $\pm 11\%$ [238], while for $W(Z/\gamma^*)$, with $m_{Z/\gamma^*} < 7$ GeV, the MadGraph prediction is normalized to the NLO POWHEG one by a K -factor of 1.4 ± 0.2 .

As in the 2011 analysis, the modelling accuracy of the signal and background processes is checked using different MC generators, namely MC@NLO for $J^P = 0^+$ (their difference is the “UEPS” signal systematic) and POWHEG for $t\bar{t}$.

5.6.3 Uncertainties from data-driven methods and control regions

Some additional systematic uncertainties arise from the usage of data-driven methods, pseudo data-driven methods and control regions. They affect, respectively, the W +jets, $t\bar{t}$ and WW backgrounds.

For what concerns the W +jets background, a conservative uncertainty of $\pm 50\%$ on its estimation in the signal region is applied, as in 2011; however, given the difference in electron and muon fake factors, their contributions are now decorrelated.

The uncertainty on the $t\bar{t}$ and single top-quark yields in the signal region is in common with the rate analysis. It is described by four independent nuisance parameters (TOP_SCALEF_0j): the experimental and theoretical uncertainties amount together to $\pm 11.4\%$, an additional $\pm 4\%$ takes care of the fact that the scale

Parameter Name	Description	Type	Shape Variation
SigXsecOverSM	$\sigma_{\text{meas}}/\sigma_{\text{SM}}$	Parameter of Interest	–
sampleNorm_ggf125	Normalization uncertainty for the ggF signal sample	Statistical	–
sampleNorm_spin2	Normalization uncertainty for the VBF signal sample	Statistical	–
sampleNorm_ttbar	Normalization uncertainty for the $t\bar{t}$ sample	Statistical	–
sampleNorm_st	Normalization uncertainty for the single top-quark sample	Statistical	–
sampleNorm_ww	Normalization uncertainty for the WW sample	Statistical	–
sampleNorm_wzzz	Normalization uncertainty for the diboson sample	Statistical	–
sampleNorm_zjets	Normalization uncertainty for the $Z/\gamma^* + \text{jets}$ sample	Statistical	–
sampleNorm_wjets	Normalization uncertainty for the $W + \text{jets}$ sample	Statistical	–
norm_WW0j	WW normalization	Control Region	–
norm_Z0j	$Z \rightarrow \tau\tau$ normalization	Control Region	–
BTAG_BEFF	B -tagging efficiency	Experimental	–
BTAG_CEFF	C -tagging efficiency	Experimental	–
BTAG_LEFF	Mis-tagging efficiency	Experimental	–
JES	Jet energy scale, split into 15 independent components	Experimental	–
JER	Jet energy resolution	Experimental	–
EL_EFF	Electron selection efficiency	Experimental	✓
MU_EFF	Muon selection efficiency	Experimental	✓
EL_ESCALE	Electron energy scale	Experimental	–
MU_ESCALE	Muon energy scale	Experimental	–
EL_RES	Electron energy resolution	Experimental	–
MU_ID_RES	Muon ID track resolution	Experimental	–
MU_MS_RES	Muon Spectrometer track resolution	Experimental	–
ISO	Lepton isolation	Experimental	✓
MU_RESCALE_lv1v_2012	Pile-up rescaling	Experimental	✓
TRIGGER_SF	Trigger efficiency uncertainty	Experimental	✓
MET_RESOFT	MET uncertainty	Experimental	–
MET_SCALESOFT	MET uncertainty due to pile-up	Experimental	–
LUMI	Luminosity uncertainty	Experimental	–
Fake_Rate_EL	Uncertainty on the $W + \text{jets}$ fake rate for electrons	Data-driven	✓
Fake_Rate_MU	Uncertainty on the $W + \text{jets}$ fake rate for muons	Data-driven	✓
TOP_SCALEF_0j	Theoretical uncertainty on 0-jet top-quark estimate (four NPs)	Data-driven (th)	–
UEPS_SHAPE_PS	Uncertainty on the WW shape given by the parton shower modelling	Theory	✓
UEPS_SHAPE_PDF	Uncertainty on the WW shape given by the PDFs	Theory	✓
UEPS_SHAPE_SCALE	Uncertainty on the WW shape given by μ_R, μ_F	Theory	✓
QCDscale_V	$q\bar{q} \rightarrow V$ scale uncertainty	Theory	–
QCDscale_VV	$q\bar{q} \rightarrow VV$ scale uncertainty	Theory	–
QCDscale_VV_ACCEPT	$q\bar{q} \rightarrow VV$ scale uncertainty on acceptance	Theory	–
QCDscale_Wg_ACCEPT0j	$W\gamma$ scale uncertainty on acceptance	Theory	–
QCDscale_Wgs_ACCEPT0j	$W\gamma^*$ scale uncertainty on acceptance	Theory	–
QCDscale_ggH	$gg \rightarrow H$ scale uncertainty	Theory	–
QCDscale_ggHlin	$gg \rightarrow H$ scale uncertainty (for ≥ 1 jet)	Theory	–
QCDscale_ggH_ACCEPT	$gg \rightarrow H$ scale uncertainty on acceptance	Theory	–
Higgs_UEPS	Uncertainty on ggF signal shape from POWHEG/MC@NLO comparison	Theory	✓
pdf_Higgs_gg	gg PDF uncertainty	Theory	–
pdf_Higgs_qq	$q\bar{q}$ PDF uncertainty	Theory	–
pdf_qqbar_ACCEPT	$q\bar{q}$ PDF acceptance uncertainty	Theory	–
pdf_qqbar	$q\bar{q}$ PDF uncertainty	Theory	–
pdf_Wg_ACCEPT	$W\gamma$ PDF acceptance uncertainty	Theory	–
pdf_Wgs_ACCEPT	$W\gamma^*$ PDF acceptance uncertainty	Theory	–

Table 5.5: Naming scheme for the nuisance parameters used in the spin analysis, from experimental and theoretical sources, control regions, data-driven methods and the limited number of MC events. The rightmost column indicates if a shape component is used in the fit in addition to the normalization one.

factor is computed at a different cut stage (after the $E_{T,\text{rel}}^{\text{miss}}$ cut) and not in the SR itself, and finally $\pm 3\%$ due to the limited number of MC events.

There is no specific theoretical uncertainty on the $Z/\gamma^* + \text{jets}$ control region prediction, since, given its limited contribution in the signal region, its theoretical modelling is not critical for the extraction of the final results; moreover, no shape information is extracted from the CR, only the normalization. For WW , on the

other hand, the error on α_{WW} , i.e. the extrapolation factor from the control to the signal region, is evaluated as discussed in Section 4.9.3. Various sources are considered, from the MC generator side (comparing the nominal POWHEG+PYTHIA6 to POWHEG+PYTHIA8 and POWHEG+HERWIG, in order to investigate the effect of the parton shower), and to MCFM for the matrix element), to the QCD scales, to the PDFs (CT10 error sets and comparison between CTEQ and MRST). The sum in quadrature of the single terms yields a total uncertainty on α_{WW} of $\pm 4.3\%$ ($\pm 3.0\%$ from modelling, $\pm 2.5\%$ from PDFs). For the shape component of the WW background uncertainty, a dedicated technique has been developed, as discussed in the following.

5.6.4 Theoretical uncertainty on the WW background shape

The uncertainty on the shape of the WW background has to be taken into account with particular care, since this is the main background in the signal region (constituting 2/3 of the total expected background, cf. Table 5.2) and sensitivities are obtained via a shape fit. For this reason, in addition to the theoretical uncertainties on the yield discussed above, more nuisance parameters are introduced, to let the shape of the WW background free to float in the fit. This is done in particular with three independent nuisance parameters, which correspond to the uncertainties on the PDFs, on the QCD scales and on the underlying event/parton shower model (UEPS_SHAPE_PDF, _SCALE and _PS, respectively). In the following, only the basic idea of the procedure is described; additional details can be found in Appendix K.

For each variation, new samples of WW events have been produced at generator-level, under the assumption that the properties of generator-level quantities follow those of the reconstructed ones and that, in particular, relative changes are coherent. For QCD scales, samples with $(\mu_R, \mu_F = m_H/2, 2m_H)$ have been produced, for UE/PS POWHEG+HERWIG has been compared to the nominal POWHEG+PYTHIA6, finally, for PDFs, reweighted sets have been created for the CTEQ error sets and the MSTW and NNPDF central values. In all cases, a standard training is performed, using the topological variables computed with generator-level lepton quantities and smeared neutrino momenta reproducing $E_{T,rel}^{\text{miss}}$.

The difference between the nominal and the up/down BDT responses is used to create an envelope, inside which the fitted shape can float. In order to have a continuous 1-dimensional envelope, BDT0 and BDT2 (or BDT1 for the $J^P = 1^\pm$ analysis) are first decorrelated, computing the direction of their correlation (θ_{corr})

and performing a rotation accordingly, so that two new variables are defined:

$$\begin{aligned} \text{BDT}_x &= \text{BDT}_0 \cdot \sin \theta_{\text{corr}} + \text{BDT}_{1,2} \cdot \cos \theta_{\text{corr}} \\ \text{BDT}_y &= \text{BDT}_0 \cdot \cos \theta_{\text{corr}} - \text{BDT}_{1,2} \cdot \sin \theta_{\text{corr}} \end{aligned} \quad (5.1)$$

Hence, BDT_x and BDT_y are mostly uncorrelated, and two independent envelopes can be built on each of them. Examples of these envelopes are shown in Fig. 5.19, for the UE/PS and QCD scale variations, for the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ analysis. The fit uses in each case the largest envelope.

This treatment minimizes overconstraints and pulls on the UEPS nuisance parameters, and for this reason has superseded the previous one (also adopted in 2011) which consisted in the comparison between the `MCONLO` and `POWHEG` shapes. One important point is that the envelope is applied to an event only if it has both BDT_0 and BDT_2 (or BDT_1) > -0.2 , i.e. only in the signal-like region. Applying it on the whole BDT response, on the contrary, the WW shape would be constrained more strongly in the background-like region and then forced to follow the same trend also in the signal-like region, possibly biasing the result (as discussed further in Appendix L).

5.6.5 Higgs boson transverse momentum

As anticipated in Section 5.1, the latest version of the `POWHEG` generator does not need to be reweighted to the `HqT` output, since the two Higgs boson p_T spectra are compatible within $\pm 5\%$. For generating the alternative spin hypotheses, however, `JHU` is used, which requires to evaluate the effect of the missing orders. A test is performed for the extreme configuration, reweighting $J^P = 2^+$ $f_{q\bar{q}} = 1.00$ events with the following function: $f(p_T) = f^{\text{HqT}}(p_T)/f^{\text{JHU}}(p_T)$. While, before the topological selection, the reweighted and non-reweighted distributions are different (as shown in Fig. 5.20 (a) for $\Delta\phi_{\ell\ell}$, which is drawn after the $E_{T,\text{rel}}^{\text{miss}}$ cut), after applying the whole selection they become almost identical (Fig. 5.20 (b)), so no systematic uncertainty is assigned on this effect.

5.6.6 Impact on the analysis

Most of the systematic uncertainties discussed above have a negligible impact on the analysis; the others are listed in Table 5.6, which shows the relative impact on the yield in the signal region (the largest between the up and down variations). The most important systematic effects are represented by the jet energy scale, the W +jets fake factor, the lepton selection efficiency and the missing E_T terms. The Z/γ^* +jets background is more affected by the systematic uncertainties because of

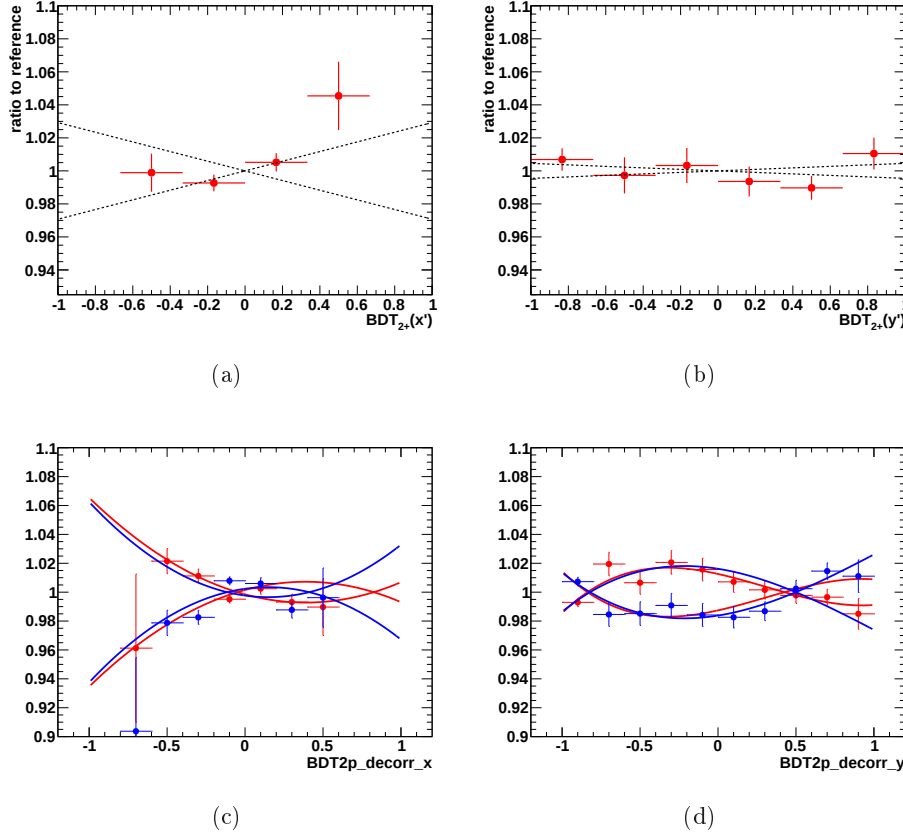


Figure 5.19: Envelopes on the rotated BDT responses, as defined in Eq. 5.1, for the $J^P = 2^+ f_{q\bar{q}} = 0.25$ analysis. In (a,b) the UE/PS variations are shown, as a ratio to the nominal configuration, while in (c,d) the largest variations given by the modified QCD scales are shown (μ_R and μ_F up and down by a factor of two ($\mu_R, \mu_F = m_H/2, 2m_H$)).

its low number of MC events after the selection. The effect on the total background, however, is around or smaller than $\pm 1\%$ for all systematic uncertainties, apart from $\pm 3.4\%$ for the data-driven fake factor. The latter has an impact also on the WW background because of the W +jets contamination in the WW control region.

Fig. 5.21 shows the Kolmogorov-Smirnov probabilities for compatibility between the nominal shape and the up/down variations, for each source of experimental uncertainties. The shapes considered here are the BDT0 and BDT2 ones, for the total background and the $J^P = 0^+$ signal (similar results are obtained for the $J^P = 2^+$ signal and in the $J^P = 1^\pm$ analysis). The resulting values are inside $\pm 60\%$, apart from a few isolated cases affected by statistical fluctuations, so that a shape component of the uncertainty is implemented only for the efficiency-related ones, the W +jets fake factor and the WW UEPS terms.

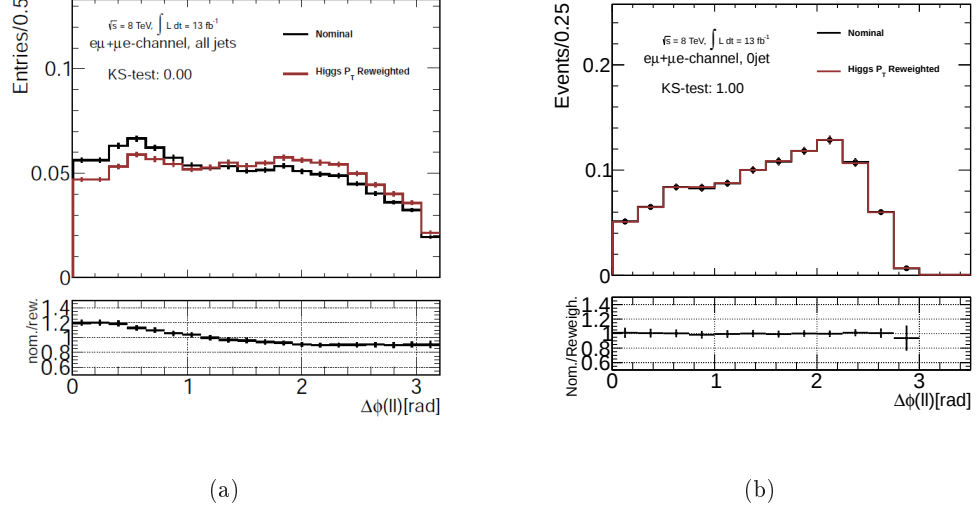


Figure 5.20: Distributions of the dilepton azimuthal angle $\Delta\phi_{\ell\ell}$, before and after the reweighting of the Higgs boson p_T spectrum provided by the JHU generator to the HqT one, for $J^P = 2^+$ $f_{q\bar{q}} = 1.00$ events. The plot in (a) is shown after the $E_{T,\text{rel}}^{\text{miss}}$ cut, the one in (b) after the full selection.

Source of Uncertainty	$J^P = 0^+$	$J^P = 2^+$	$J^P = 1^+$	$J^P = 1^-$	WW	$W+\text{jets}$	$Z/\gamma^*+\text{jets}$	Top-quark	Total Bkg
Jet Energy Scale	4.7%	4.2%	2.8%	2.9%	0.8%	–	9.6%	2.9%	1.0%
$W+\text{jets}$ Fake Factor	–	–	–	–	2.9%	50%	–	–	3.4%
Lepton ID Efficiency	3.2%	3.1%	3.2%	3.3%	1.0%	–	3.7%	1.0%	1.1%
$E_{T,\text{rel}}^{\text{miss}}$	0.3%	0.3%	0.3%	0.3%	0.1%	–	6.3%	1.4%	0.5%
Number of MC events	0.9%	0.8%	0.6%	0.7%	0.6%	–	5.2%	2.3%	0.7%

Table 5.6: Largest systematic (normalization) uncertainties affecting the yields of the various signal and background components in the signal region.

5.7 Statistical model and fit procedure

5.7.1 Likelihood and test statistic

The definition of the likelihood function follows Section 4.10.2 and is, in particular, analogous to Eq. 4.10, with the product running now only on one data-taking period, one jet category and one lepton flavour:

$$\mathcal{L}(\mu, \varepsilon, \theta) = \left\{ \prod_{i=1}^{N_{\text{bins}}} P \left[N_i \mid \mu(\varepsilon S_i^{(0)} + (1 - \varepsilon) S_i^{(\text{ALT})}) + \sum_b^{N_{\text{bkg}}} B_{ib} \right] \right\} \times \prod_{m=1}^{N_\theta} F(\tilde{\theta}_m, \theta_m). \quad (5.2)$$

The other main difference is that the parameter of interest is not μ (which is now treated as a nuisance parameter), but ε , the fraction of $J^P = 0^+$ events, such that

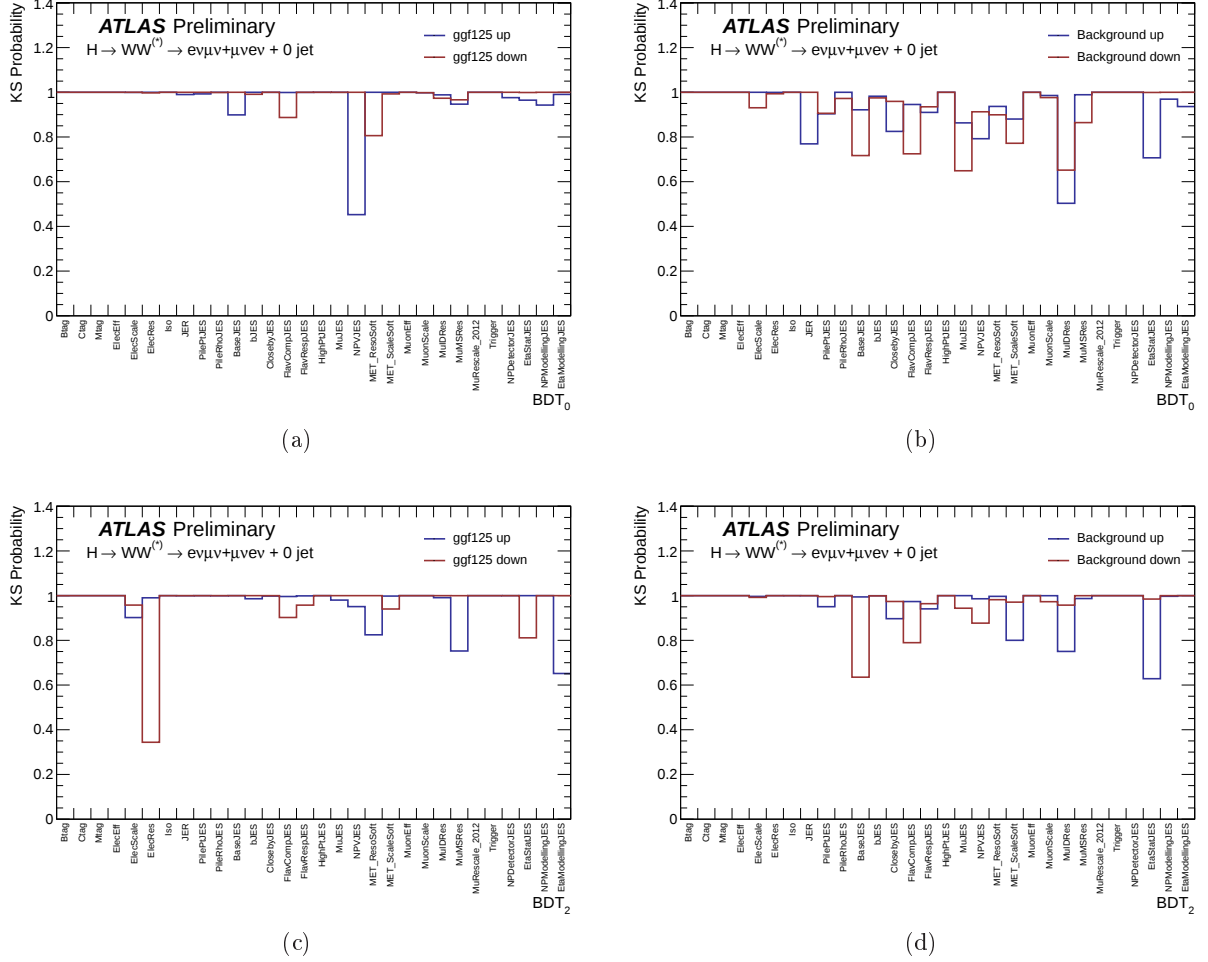


Figure 5.21: Kolmogorov-Smirnov test probabilities for compatibility between the nominal shape and the up/down variations of each experimental uncertainties (cf. Table 5.5). The top row shows BDT0, the bottom row BDT2; the left column shows the $J^P = 0^+$ signal, the right column the total background.

$\varepsilon = 0$ corresponds to a pure $J^P = 2^+$ or $J^P = 1^\pm$ signal, while $\varepsilon = 1$ to a pure SM signal. As before, N_θ are the other nuisance parameters, with $F(\tilde{\theta}_m, \theta_m)$ being the auxiliary measurements that constrain the systematic uncertainties.

The test statistic is defined as the logarithmic ratio of the two likelihoods, for the null (the SM) and alternative hypotheses under consideration:

$$q_\mu = \ln \frac{\mathcal{L}(\varepsilon = 1, \hat{\mu}_{\varepsilon=1}, \hat{\theta}_{\varepsilon=1})}{\mathcal{L}(\varepsilon = 0, \hat{\mu}_{\varepsilon=0}, \hat{\theta}_{\varepsilon=0})}. \quad (5.3)$$

The symbols $\hat{\mu}$ and $\hat{\theta}$ refer to the best fit values (MLEs) of the signal strength and the nuisance parameters, obtained maximizing the likelihood in correspondence of

each ε hypothesis.

The distribution of q , and consequently the p_0 -values that quantify the rejection power for alternative spin hypotheses, are defined as described in Section 4.10.4, with $f(Q|B)$ and $f(Q|S+B)$ substituted by $f(Q|\varepsilon=1)$ and $f(Q|\varepsilon=0)$, respectively. Their distributions are obtained by means of pseudo-experiments (also known as *toy* experiments), since the asymptotic approximation is not applicable, given the constraint on the POI; they are shown schematically in Fig. 5.22, in (a) for the MC expectation, in (b) for the observation in real data. The p_0 -value for compatibility with the alternative hypothesis, corresponding to the blue filled area (dashed for expectation), is defined as the probability of finding a Q value equally or less compatible with the “ $2^+ + B$ ” model than the one measured in data (than the “ $0^+ + B$ ” median), assuming “ $0^+ + B$ ” to be true:

$$p_0(\text{ALT}) = P(Q \geq Q_{\text{obs}} | \varepsilon = 0) = \int_{Q_{\text{obs}}}^{\infty} f(Q | \varepsilon = 0) dQ. \quad (5.4)$$

At the same time, $1 - p_0(\text{ALT})$ can be regarded as the probability of rejecting the “ $2^+ + B$ ” hypothesis, assuming that “ $0^+ + B$ ” is true: for instance, having $p_0(\text{ALT}) = 0.05$ corresponds to a 95% rejection probability for the $J^P = 2^+$ hypothesis. The confidence level for excluding the alternative hypothesis in favour of the SM is evaluated by means of a CL estimator [230]:

$$\text{CL}_s = \frac{p_0(\text{ALT})}{1 - p_0(0^+)} \quad (5.5)$$

which normalizes the rejection power of the alternative hypothesis to the compatibility with the SM case, analogously to Eq. 4.15 for the 2011 analysis.

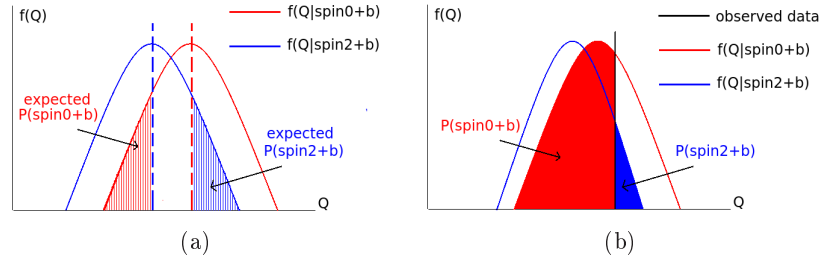


Figure 5.22: Schematic representation of the $f(Q|0^+ + B)$ and $f(Q|2^+ + B)$ distributions and definition of $p_0(0^+)$, $p_0(\text{ALT})$ (labelled as $P(\text{spin}0+b)$ and $P(\text{spin}2+b)$ in the figures, respectively), for (a) MC expectation and (b) observation in real data.

A different definition of the test statistic considers both distributions ($\varepsilon = 0, 1$)

normalized to the one with ε free, in the following way:

$$q_{\varepsilon=0} = \ln \frac{\mathcal{L}(\varepsilon = 0, \hat{\mu}_{\varepsilon=0}, \hat{\theta}_{\varepsilon=0})}{\mathcal{L}(\hat{\varepsilon}, \hat{\mu}_{\hat{\varepsilon}}, \hat{\theta}_{\hat{\varepsilon}})} \quad , \quad q_{\varepsilon=1} = \ln \frac{\mathcal{L}(\varepsilon = 1, \hat{\mu}_{\varepsilon=1}, \hat{\theta}_{\varepsilon=1})}{\mathcal{L}(\hat{\varepsilon}, \hat{\mu}_{\hat{\varepsilon}}, \hat{\theta}_{\hat{\varepsilon}})} \quad (5.6)$$

where in the denominators $\hat{\mu}_{\hat{\varepsilon}}$ and $\hat{\theta}_{\hat{\varepsilon}}$ are computed in correspondance with the best fit value of ε . This is the standard LHC prescription for computing exclusion limits [230,231], as discussed in Section 4.10.4, i.e. the profile likelihood ratio (PLR).

The drawback of this method is that, in principle, $\hat{\varepsilon}$ can take a continuous spectrum of values, while the model this analysis wants to explore considers $\varepsilon = 0, 1$ as the only two physical scenarios. Nevertheless, the rejection powers achieved by the two methods are expected to be comparable, as long as ε is close either to 0 or 1, which has been verified in different configurations, as shown in Appendix N, where the residual differences are investigated.

In the following, p_0 -values are computed with the first method; they can be converted into significances, Z , defined such that a Gaussian-distributed variable, found Z standard deviations above its mean, has an upper-tail probability equal to p . That is:

$$Z = \Phi^{-1}(1 - p) \quad , \quad (5.7)$$

where Φ^{-1} is the Gaussian *quantile*, i.e. the inverse of the cumulative distribution [239]. In particular, $Z = 1.64$ corresponds to a threshold p_0 -value of 0.05 (95% CL).

5.7.2 Choice of the binning

As already discussed, sensitivities are computed with a 2-dimensional fit in the (BDT0, BDT2) plane (or BDT1 for the $J^P = 1^\pm$ analysis). The signal region uses a (10, 5) binning, while for the control regions (WW and $Z/\gamma^* + \text{jets}$) only a single bin is used. The (10, 5) binning has been chosen starting from two considerations: firstly, the 2011 BDT analysis uses ten bins (and its response basically corresponds to BDT0 in 2012), secondly, BDT0 accounts for most of the spin discrimination, and as a consequence it is more finely binned.

A scan using five, seven and ten bins in each dimension has been performed on a reduced dataset (13fb^{-1}), and the results are summarized in Tables 5.7 – 5.11, where expected sensitivities for $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ are shown, for $\mu = 1$ and $\mu = \hat{\mu}$, together with the fitted $\hat{\mu}$ values. The results do not take systematic uncertainties into account. The best sensitivities are obtained in all cases for ten bins in BDT0; having more than five bins in BDT2 does not improve the results in a significant and coherent way, so that there is no clear gain by using a finer binning. The fitted signal strengths are constant within uncertainties.

BDT2 bins BDT0 bins	5	7	10
5	0.65	0.57	0.62
7	0.66	0.71	0.67
10	0.69	0.69	0.72

Table 5.7: Expected sensitivities for $J^P = 0^+$ rejection, with $\mu = 1$.

BDT2 bins BDT0 bins	5	7	10
5	1.41	1.31	1.36
7	1.42	1.49	1.48
10	1.49	1.54	1.58

Table 5.8: Expected sensitivities for $J^P = 2^+$ rejection, with $\mu = 1$.

BDT2 bins BDT0 bins	5	7	10
5	1.51	1.35	1.52
7	1.51	1.56	1.41
10	1.57	1.59	1.57

Table 5.9: Expected sensitivities for $J^P = 0^+$ rejection, with $\mu = \hat{\mu}$.

BDT2 bins BDT0 bins	5	7	10
5	1.78	1.62	1.82
7	1.69	1.87	1.82
10	1.98	1.95	2.22

Table 5.10: Expected sensitivities for $J^P = 2^+$ rejection, with $\mu = \hat{\mu}$.

BDT2 bins BDT0 bins	5	7	10
5	$\hat{\mu}_0 = 1.46 \pm 0.30$	$\hat{\mu}_0 = 1.39 \pm 0.29$	$\hat{\mu}_0 = 1.46 \pm 0.34$
	$\hat{\mu}_2 = 2.19 \pm 0.67$	$\hat{\mu}_2 = 2.13 \pm 0.67$	$\hat{\mu}_2 = 2.19 \pm 0.67$
7	$\hat{\mu}_0 = 1.41 \pm 0.40$	$\hat{\mu}_0 = 1.40 \pm 0.33$	$\hat{\mu}_0 = 1.39 \pm 0.28$
	$\hat{\mu}_2 = 2.11 \pm 0.68$	$\hat{\mu}_2 = 2.02 \pm 0.67$	$\hat{\mu}_2 = 1.96 \pm 0.67$
10	$\hat{\mu}_0 = 1.51 \pm 0.37$	$\hat{\mu}_0 = 1.46 \pm 0.37$	$\hat{\mu}_0 = 1.51 \pm 0.36$
	$\hat{\mu}_2 = 2.18 \pm 0.68$	$\hat{\mu}_2 = 2.04 \pm 0.68$	$\hat{\mu}_2 = 2.00 \pm 0.66$

Table 5.11: Fitted $\hat{\mu}$ values for the $J^P = 0^+$ and $J^P = 2^+$ signals, with different binnings. All the results in Tables 5.7 – 5.11 correspond to an integrated luminosity of 13 fb^{-1} and use no systematic uncertainties.

5.8 Results for the spin-2 analysis

Most of the results in this Section are given for $f_{q\bar{q}} = 0.25$, more details for the other fractions can be found in Appendix O.

5.8.1 Post-fit yields

Table 5.12 compares pre-fit and post-fit event yields for the backgrounds and the different signal hypotheses, for $J^P = 0^+$ and $J^P = 2^+$ fits, in signal and control regions. The background breakdown is the same as used in the fit. The pre-fit numbers are in principle equal to those in Table 5.2, keeping into account two differences: firstly, the latter use all MC events, while those in Table 5.12 correspond to half of them

(only the odd-numbered events) scaled to the full luminosity, secondly, normalization factors are not applied for the WW and $Z/\gamma^*+\text{jets}$ backgrounds, since they are recovered in the fit from their control regions (the normalization factors are, on the other hand, applied for the top-quark backgrounds).

After the fit, the two signals are scaled up by their $\hat{\mu}$ values. The WW background is increased of around 10%, in agreement with the pre-fit NF (1.08), more in the $J^P = 2^+$ case: the reduced increase is compensated, in the $J^P = 0^+$ case, by an increase in the number of $Z/\gamma^*+\text{jets}$ and $W+\text{jets}$ events. The post-fit $Z/\gamma^*+\text{jets}$ NF is 0.91, in agreement with the pre-fit one (0.92). The $W+\text{jets}$ background is decreased by the fit for both signal hypotheses, more for $J^P = 2^+$.

	0^+	2^+	WW	WZ/ZZ	$W\gamma$	$W\gamma^*$	$t\bar{t}$	Single Top	$Z/\gamma^*+\text{jets}$	$W+\text{jets}$
pre-fit	170	112	2028	63	83	83	171	130	309	278
post-fit $J^P = 0^+$	265	—	2223	63	76	83	169	129	334	266
post-fit $J^P = 2^+$	—	265	2302	63	82	78	166	126	285	240
pre-fit $Z \rightarrow \tau\tau$ CR	4	9	93	6	5	2	4	3	3582	95
post-fit $J^P = 0^+$ $Z \rightarrow \tau\tau$ CR	6	—	101	6	4	2	4	3	3262	90
post-fit $J^P = 2^+$ $Z \rightarrow \tau\tau$ CR	—	22	81	6	5	2	4	3	3252	81
pre-fit WW CR	1	2	2212	56	34	42	333	192	56	153
post-fit $J^P = 0^+$ WW CR	1	—	2427	55	35	48	330	191	67	148
post-fit $J^P = 2^+$ WW CR	—	5	2476	55	33	39	323	187	54	134

Table 5.12: Pre-fit and post-fit event rounded yields for the various background components (as used in the fit) and the two alternative signal hypotheses, for $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ fits, in signal and control regions. The statistical and systematic uncertainties are all taken into account.

	$J^P = 0^+$ fit			$J^P = 2^+$ fit		
	$\hat{\mu}$	WW norm	Z norm	$\hat{\mu}$	WW norm	Z norm
$f_{q\bar{q}} = 0.00$	$1.59^{+0.61}_{-0.48}$	$1.10^{+0.05}_{-0.06}$	$0.92^{+0.13}_{-0.08}$	$3.05^{+1.23}_{-1.02}$	$1.11^{+0.05}_{-0.06}$	$0.92^{+0.11}_{-0.11}$
$f_{q\bar{q}} = 0.25$	$1.59^{+0.64}_{-0.50}$	$1.09^{+0.05}_{-0.06}$	$0.93^{+0.02}_{-0.02}$	$2.36^{+1.12}_{-0.96}$	$1.12^{+0.03}_{-0.03}$	$0.92^{+0.11}_{-0.08}$
$f_{q\bar{q}} = 0.50$	$1.53^{+0.62}_{-0.47}$	$1.09^{+0.05}_{-0.06}$	$0.92^{+0.01}_{-0.08}$	$1.99^{+1.06}_{-0.88}$	$1.12^{+0.05}_{-0.05}$	$0.92^{+0.10}_{-0.08}$
$f_{q\bar{q}} = 0.75$	$1.55^{+0.60}_{-0.49}$	$1.10^{+0.05}_{-0.06}$	$0.93^{+0.12}_{-0.08}$	$1.39^{+0.94}_{-0.79}$	$1.13^{+0.05}_{-0.05}$	$0.92^{+0.10}_{-0.08}$
$f_{q\bar{q}} = 1.00$	$1.62^{+0.62}_{-0.48}$	$1.10^{+0.05}_{-0.06}$	$0.93^{+0.12}_{-0.08}$	$0.88^{+0.82}_{-0.73}$	$1.13^{+0.05}_{-0.05}$	$0.92^{+0.10}_{-0.08}$

Table 5.13: Best-fit values for the signal strenghts, in units of the expected SM cross section, and the background normalizations obtained from the control regions (WW and $Z \rightarrow \tau\tau$), for the $J^P = 0^+$ and $J^P = 2^+$ fits, with all systematic uncertainties included.

5.8.2 Post-fit BDT distributions

Fig. 5.23 shows the 2-dimensional BDT response (BDT0, BDT2) after its unrolling in 1D. All normalizations (for the signal hypotheses and for the various background components) are the post-fit ones, summarized in Table 5.12. In both cases, the agreement between the data and the MC expectation is good on the whole range; the relevant signal contributions are concentrated in four bins, where, already visually, $J^P = 0^+$ seems to match better the data/MC excess than $J^P = 2^+$, with a significance that can be computed by means of toy experiments.

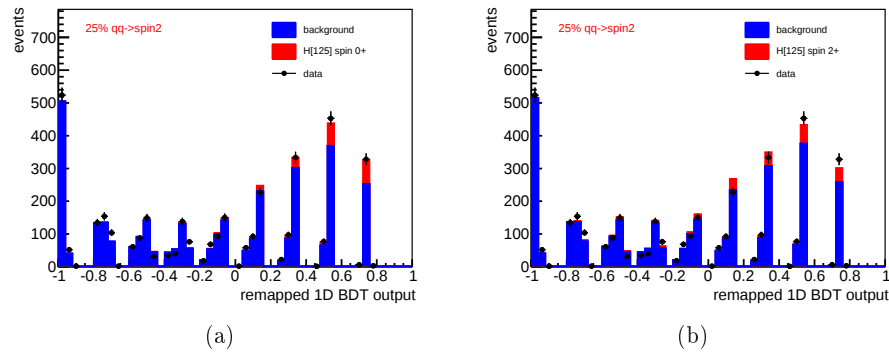


Figure 5.23: Unrolled 1-dimensional distributions for the (a) $J^P = 0^+$ and (b) $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ signal hypotheses, as well as the total background and the data. The statistical and systematic uncertainties are all taken into account. The normalizations are the post-fit ones.

The difference between the results of the two fits can be better visualized by means of the *background-subtracted data* plots, where the excess (computed as the number of data events after subtracting the post-fit background) is compared with the expected distribution for the signal, as shown in Fig. 5.24. In this case the unrolled 1-dimensional distribution is used, but the bins are ordered according to the (increasing) number of expected signal events, and all bins with expected content below 0.1 are removed. After this reordering, the signal excess is peaking on the righthand side, and it looks easier to be reconciled with the expected $J^P = 0^+$ shape than with the $J^P = 2^+$ one.

5.8.3 Pulls in nuisance parameters

The fit result has to be evaluated also in terms of the fitted values of the nuisance parameters. As explained also in Appendix I, all systematic NPs have pre-fit values of 0 ± 1 (since they represent the variation with respect to the nominal configuration, which is expected to float inside the range defined by the initial prescription), while

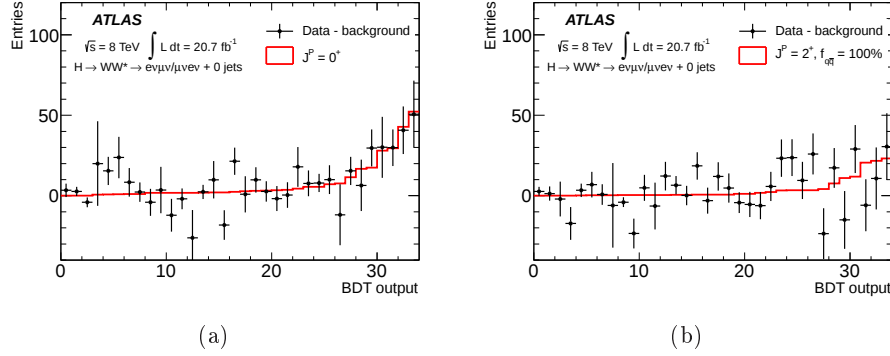


Figure 5.24: Unrolled 1-dimensional BDT distributions for the (a) $J^P = 0^+$ and (b) $J^P = 2^+$ $f_{q\bar{q}} = 1.00$ signal hypotheses, after the background subtraction from data, using the post-fit normalizations for all backgrounds and signals. The bins have been reordered according to the number of expected signal events, increasing from left to right. The statistical and systematic uncertainties are all taken into account.

the normalizations from the control regions (the α factors) have 1 ± 0 . In the former case, a post-fit error much smaller than 1 indicates that the fit is overconstraining the systematic uncertainty more than the initial prescription, possibly biasing the result. Fig. 5.25 shows the pulls for all NPs listed in Table 5.5, after discarding those that have a negligible impact on the final result. Pulls are defined as the difference between the post- and pre-fit values of the parameter, divided by its post-fit error. Parameters are marked in red if their best fit value is more than 40% away from 0 (the parameter is pulled) or if their error is less than 75% of the pre-fit one (the parameter is overconstrained).

For what concerns the $J^P = 0^+$ fit, shown in Fig. 5.25 (b), the only parameter that needs additional checks is the JES_2012_Detector1 one, which has a post-fit value of $0.60^{+0.79}_{-1.70}$, i.e. it is both pulled and asymmetrically constrained. This is most likely due to statistical fluctuations in the background-enriched part of the BDT response, as suggested by a dedicated study: with reference to the unrolled $J^P = 0^+$ distribution in Fig. 5.24 (a), the first 25 bins are combined into one, three or five bins, while the remaining bins, i.e. those responsible for most of the spin discrimination, are left untouched. The sensitivities slightly decrease (as expected, since some of the differential knowledge on the backgrounds is lost) and at the same time the pull on the JES NP disappears. The likelihood profiled as a function of the parameter shows, in the standard configuration, an asymmetric structure with two minima, the absolute one being moved away from 0, in the standard configuration (Fig. 5.26 (a)), while when the first 25 bins are combined into three (b) a symmetric minimum centered at zero is restored. In conclusion, removing background bin-to-

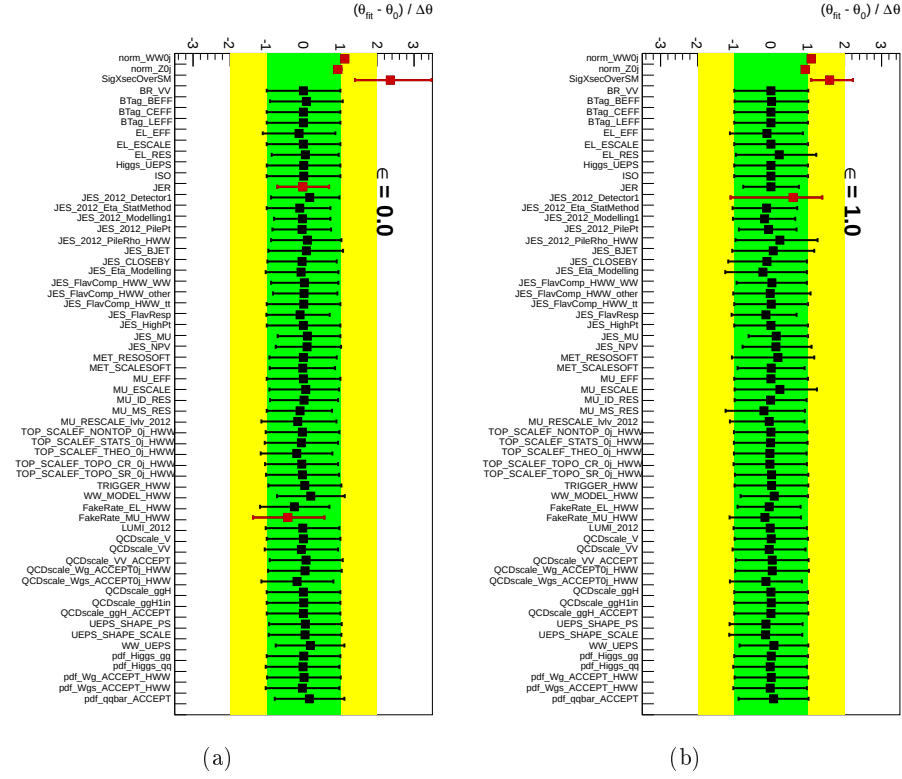


Figure 5.25: Post-fit pulls in the nuisance parameters, for the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ analyses: (a) $\varepsilon = 0$ fit, (b) $\varepsilon = 1$ fit. The statistical and systematic uncertainties are all taken into account, but only those which are non negligible appear in the plots. The CR normalizations are labelled “norm_WW0” and “norm_Z0”, “SigXsecOverSM” is the signal strength, all other parameters follow the naming convention in Table 5.5.

bin fluctuations, i.e. combining bins together, causes the pull on the JES parameter to disappear, at the price of a loss in sensitivity. On the other hand, just performing the fit fixing the JES parameter to 0 ± 1 , does not alter the final results significantly, proving that the JES pull is not the source of any bias.

For what concerns the $J^P = 2^+$ fit (Fig. 5.25 (a)), on the other hand, a slight overconstraint on the JER parameter appears (70%), and more interestingly both Fake Rate parameters are somewhat pulled ($\text{Fake_Rate_EL} = -0.25^{+0.95}_{-0.94}$, $\text{Fake_Rate_MU} = -0.43^{+0.99}_{-0.94}$). In fact, as already noticed commenting Table 5.12, the W +jets background is reduced by the $J^P = 2^+$ fit more than by the $J^P = 0^+$ fit, since the former tends to accomodate more WW events.

It should be noticed that, for both signal hypotheses, the UEPS nuisance parameters are well behaved, and the PDF one is discarded by the fit since its impact is

negligible.

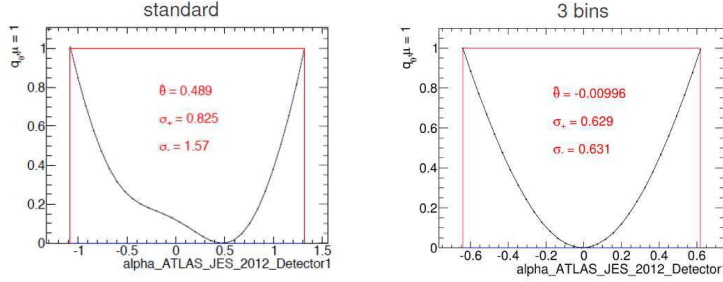


Figure 5.26: Profiled likelihood as a function of the JES_2012_Detector1 nuisance parameter, for the $J^P = 0^+$ fit, in the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ analysis. In the standard configuration (a) an asymmetric structure appears, corresponding to the pulled and asymmetric uncertainty on the post-fit NP. Combining the first 25 bins into three (b), a symmetric minimum, centered at zero, is restored.

5.8.4 Signal strength and compatibility with the rate measurement

Even if the spin analysis is not designed to provide a precise measurement of the signal strength, a value of $\hat{\mu}$ is obtained as a result of the fit, and it is used to compute the sensitivities, since the test statistic q_μ is a function of μ (cf. Eq. 5.3). Table 5.14 summarizes the post-fit normalizations for the signal, in the $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ hypotheses, and the corresponding WW and $Z \rightarrow \tau\tau$ CR normalizations, for different configurations of the systematic uncertainties. The last line, in particular, refers to the configuration where only the main systematic uncertainties are considered: the UEPS, Fake Rate and theoretical ones and the uncertainties due to the limited number of MC events. All values are compatible within uncertainties. Moreover, the error on $\hat{\mu}$ increases consistently when adding the systematic uncertainties, as expected.

The signal strength in the $J^P = 0^+$ hypothesis, $\hat{\mu} = 1.59^{+0.64}_{-0.50}$, appears to be larger than the one obtained by the rate analysis, 1.26 ± 0.35 [234] (which is computed with all lepton flavours and jet bins included). A study on the compatibility of the two results has been performed, generating 100 toy experiments under the assumption that the $J^P = 0^+$ hypothesis is true (with $\mu = 1$), fitting them in two ways: once according to the rate measurement prescription (fit m_T after the full cutbased selection, but for the $e\mu$ 0-jet channel only), once according to the spin analysis prescription. The results are presented in Fig. 5.27 (a), where the distribution of $|\hat{\mu}_{\text{spin}} - \hat{\mu}_{\text{rate}}|$ is shown, together with the observed value. Integrating to the right of the data line, the compatibility between the two measurements is found to be 1.65σ . The expected values are linearly dependent, as shown in Fig. 5.27 (b).

	$J^P = 0^+$ fit			$J^P = 2^+$ fit, $f_{q\bar{q}} = 0.25$		
	$\hat{\mu}$	WW norm	Z norm	$\hat{\mu}$	WW norm	Z norm
All systematic uncertainties	$1.59^{+0.64}_{-0.50}$	$1.09^{+0.05}_{-0.06}$	$0.93^{+0.02}_{-0.02}$	$2.36^{+1.12}_{-0.96}$	$1.12^{+0.03}_{-0.03}$	$0.92^{+0.11}_{-0.08}$
No systematic uncertainties	$1.52^{+0.28}_{-0.29}$	$1.11^{+0.02}_{-0.02}$	$0.91^{+0.02}_{-0.02}$	$2.41^{+0.51}_{-0.51}$	$1.11^{+0.02}_{-0.02}$	$0.90^{+0.02}_{-0.02}$
Limited number of MC events only	$1.50^{+0.31}_{-0.31}$	$1.11^{+0.02}_{-0.02}$	$0.91^{+0.02}_{-0.02}$	$2.41^{+0.57}_{-0.57}$	$1.11^{+0.03}_{-0.02}$	$0.90^{+0.02}_{-0.02}$
Theory systematic uncertainties only	$1.45^{+0.47}_{-0.39}$	$1.10^{+0.02}_{-0.02}$	$0.91^{+0.04}_{-0.04}$	$2.27^{+0.83}_{-0.69}$	$1.10^{+0.02}_{-0.02}$	$0.90^{+0.02}_{-0.02}$
Detector systematic uncertainties only	$1.57^{+0.45}_{-0.39}$	$1.12^{+0.05}_{-0.05}$	$0.88^{+0.07}_{-0.06}$	$2.73^{+0.83}_{-0.80}$	$1.15^{+0.05}_{-0.05}$	$0.87^{+0.06}_{-0.05}$
Main systematic uncertainties only	$1.50^{+0.53}_{-0.44}$	$1.11^{+0.04}_{-0.04}$	$0.91^{+0.04}_{-0.04}$	$2.36^{+1.10}_{-0.92}$	$1.12^{+0.04}_{-0.04}$	$0.90^{+0.02}_{-0.02}$

Table 5.14: Best-fit values for the signal strenghts, in units of the expected SM cross section, and normalizations for the backgrounds obtained from their respective control regions (WW and $Z \rightarrow \tau\tau$), for the $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ fits. Different configurations of the systematic uncertainties are considered, in particular the “main systematic uncertainties” one uses only the UEPS, Fake Rate and theoretical contributions and the uncertainty due to the limited number of MC events.

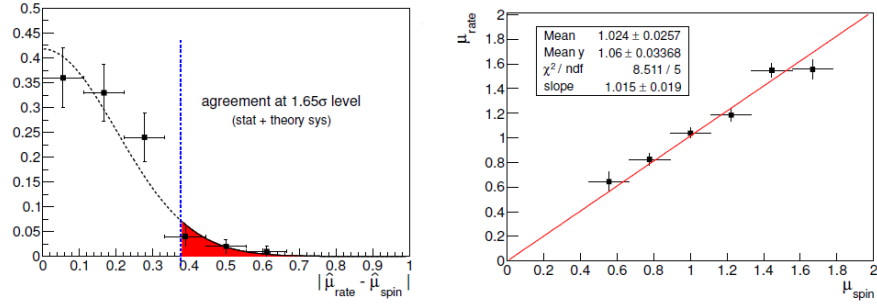


Figure 5.27: Compatibility between the signal strengths measured by the spin and rate analyses, in both cases in the $e\mu$ 0-jet channel only. Left: Distribution of $|\hat{\mu}_{\text{spin}} - \hat{\mu}_{\text{rate}}|$ obtained by means of toy experiments, generated assuming the $J^P = 0^+$ hypothesis. The observed value is also shown, see the text for further details. Right: Linear correlation between the two measurements.

5.8.5 Sensitivities

As explained in Section 5.7, the sensitivities are computed by means of toy experiments, generated under the assumption that either the $J^P = 0^+$ or the $J^P = 2^+$ hypotheses are true. Each toy is then fitted for each of the two hypotheses, computing the corresponding value of the test statistic q_μ , defined in Eq. 5.3 as the logarithmic ratio between the two likelihoods: $\mathcal{L}(\varepsilon = 1, \hat{\mu}_{\varepsilon=1}, \hat{\theta}_{\varepsilon=1}) = \mathcal{L}(J^P = 0^+)$ and $\mathcal{L}(\varepsilon = 0, \hat{\mu}_{\varepsilon=0}, \hat{\theta}_{\varepsilon=0}) = \mathcal{L}(J^P = 2^+)$. Then, as illustrated in Fig. 5.22, the expected sensitivities are computed integrating on each toy distribution starting from the median of the other one (for expected) or from the q_μ value found in data (for observed). In this way, the rejection probabilities are obtained in terms of p_0 -

values, Z significances and finally CL_s confidence levels (cf. Eqs. 5.4, 5.7 and 5.5, respectively); they are shown in Tables 5.15 and 5.16 for the five $f_{q\bar{q}}$ fractions. All values use the best-fit signal strengths and the best-fit nuisance parameters. As expected (see Section 5.5.1 and in particular Fig. 5.10), the sensitivities increase with increasing $f_{q\bar{q}}$. The expected and observed sensitivities for $J^P = 2^+$ rejection are compatible, while the observed p_0 -values for $J^P = 0^+$ rejection are larger than 0.5, and negative in terms of Z significances, because the data cannot reject the $J^P = 0^+$ hypothesis. The CL_s values range from 95.2% for $f_{q\bar{q}} = 0.00$ (2.2σ for $J^P = 2^+$ rejection) to 99.94% for $f_{q\bar{q}} = 1.00$ (3.4σ).

Table 5.16 also shows the numbers of fitted $J^P = 0^+$ and $J^P = 2^+$ events: almost constant in the former case, anticorrelated with the increase in sensitivity for $J^P = 2^+$, which confirms that the latter is due to a better shape discrimination for increasing $f_{q\bar{q}}$ and not to the freedom in the choice of $\hat{\mu}$.

Fig. 5.28 contains the five toy distributions of the test statistic, normalized to unity. The $J^P = 0^+$ distribution (in blue) gets broader as $f_{q\bar{q}}$ increases, which can be understood in the following way: since $q_\mu = \ln(\mathcal{L}(J^P = 0^+)/\mathcal{L}(J^P = 2^+))$, as $f_{q\bar{q}}$ increases it becomes harder and harder to fit a $J^P = 2^+$ signal on toys generated according to the $J^P = 0^+$ hypothesis (and, conversely, easier to fit $J^P = 0^+$ on them), so that it becomes more likely to obtain high q_μ values. For $f_{q\bar{q}} = 1.00$ and 0.75, in particular, only a few toy experiments contribute to the integral: in order to populate the tails, more toy experiments are generated for these fractions (see Table 5.16), resulting in uncertainties on the observed $J^P = 2^+$ sensitivities that range from 8% for $f_{q\bar{q}} = 0.00$ to 45% for $f_{q\bar{q}} = 1.00$ (see Appendix O for all the relative uncertainties). The observed value is found in all cases close to the $J^P = 0^+$ expected median, which indicates again their compatibility.

5.9 Results for the spin-1 analysis

In the following, results for the $J^P = 1^\pm$ analyses are given. Details are omitted if equivalent to what discussed in Section 5.8 for the $J^P = 2^+$ analysis.

5.9.1 Post-fit yields

The pre- and post-fit normalizations of the signal and background processes are shown in Table 5.17 for the $J^P = 1^+$ analysis (the $J^P = 1^-$ case is very similar and hence omitted here). This is done for the signal and control regions, with the same background breakdown as used in the fit. In the pre-fit case, the normalization factors are applied for all backgrounds apart from WW and $Z/\gamma^* + \text{jets}$, which are

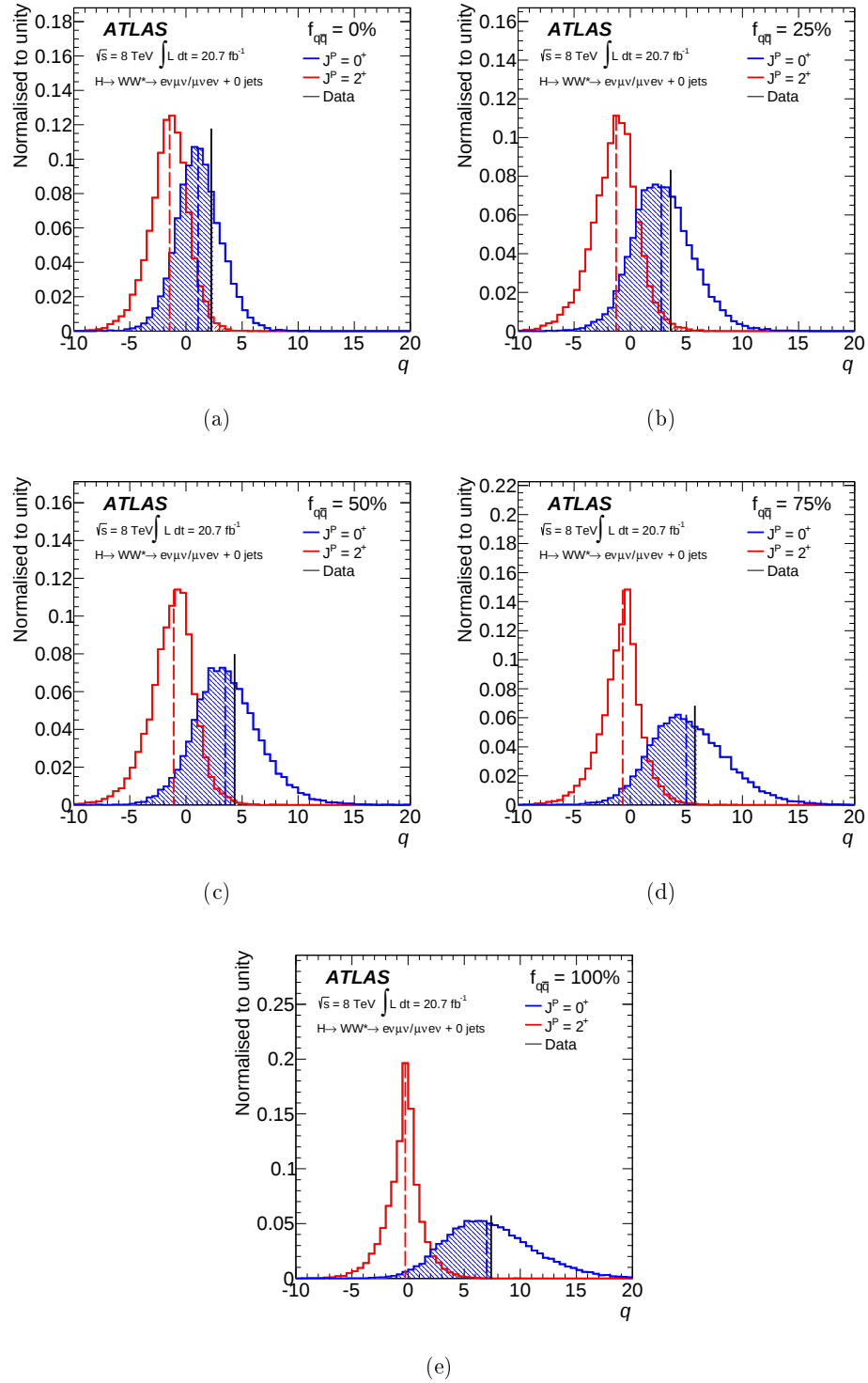


Figure 5.28: Expected distributions of the test statistic $q = \ln(\mathcal{L}(J^P = 0^+)/\mathcal{L}(J^P = 2^+))$ for the five $f_{q\bar{q}}$ fractions, given by toy experiments generated under the $J^P = 0^+$ (blue dotted) or $J^P = 2^+$ (red dashed) hypotheses. The dashed lines indicate the expected medians, while the observed value is marked by a solid black line. The coloured areas correspond to the integrals used to compute the observed p_0 -values for rejecting each hypothesis.

$f_{q\bar{q}}$	exp. $p_0(0^+)$	exp. $p_0(2^+)$	obs. $p_0(0^+)$	obs. $p_0(2^+)$	1-CL _s (2^+)
1.00	0.013	$3.6 \cdot 10^{-4}$	0.541	$1.7 \cdot 10^{-4}$	0.9994
0.75	0.028	0.003	0.586	0.001	0.997
0.50	0.042	0.009	0.616	0.003	0.992
0.25	0.048	0.019	0.622	0.008	0.980
0.00	0.086	0.054	0.731	0.013	0.952

Table 5.15: Summary of p_0 -values for the five $f_{q\bar{q}}$ fractions, with all systematic uncertainties taken into account, in terms of $\hat{\mu}$. The second and third columns show the expected sensitivities for $J^P = 0^+$ and $J^P = 2^+$ rejection, respectively, assuming the other hypothesis to be true, while the fourth and fifth columns show the corresponding observed sensitivities. The last column contains the 1-CL_s results for $J^P = 2^+$ rejection.

normalized in their respective control regions; the signals are scaled up by their $\hat{\mu}$ values.

As for the $J^P = 2^+$ analysis, the WW background is increased by the fit more in the $J^P = 1^\pm$ than in the $J^P = 0^+$ case, with Z/γ^* +jets and W +jets events compensating for the difference. Concerning the W +jets background, it is decreased by the fit for both signal hypotheses. In general, the numbers are very similar for the two $J^P = 1^\pm$ analyses.

5.9.2 Post-fit BDT distributions

The 1-dimensional unrolled BDT response is shown in Fig. 5.29, with the signal and the background distributions stacked on top of each other, using the post-fit normalizations. The data distribution agrees well with the MC one and, as for the $J^P = 2^+$ case, the signal contribution appears concentrated in a few bins only. Moreover, Fig. 5.30 shows the BDT0 variable in the data after background subtraction, compared with the expected signal shape (in a finer binning than the one used in the fit). The shape looks flatter for $J^P = 1^-$ than for $J^P = 0^+$ and $J^P = 1^+$, thus anticipating a higher sensitivity for the $J^P = 1^-$ analysis.

5.9.3 Pulls in nuisance parameters

Fig. 5.31 displays the pulls for (a,b) the $J^P = 1^+$ and (c,d) the $J^P = 1^-$ fits. The parameters marked in red (i.e. with a pull $> 40\%$ or an overconstraint $> 75\%$) are the same as those already mentioned for the $J^P = 2^+$ analysis in Section 5.8.3: the

$f_{q\bar{q}}$	$N_{\text{fit}}(0^+)$	$N_{\text{fit}}(2^+)$	exp. $Z(0^+)$	exp. $Z(2^+)$	obs. $Z(0^+)$	obs. $Z(2^+)$	$N_{\text{toys}} (10^3)$
1.00	270^{+100}_{-80}	110^{+110}_{-90}	2.24	3.34	-0.09	3.43	16.8
0.75	250^{+100}_{-80}	170^{+110}_{-100}	1.90	2.76	-0.22	3.05	18.2
0.50	250^{+100}_{-80}	230^{+140}_{-100}	1.72	2.37	-0.29	2.72	13.9
0.25	260^{+110}_{-80}	260^{+130}_{-110}	1.66	2.07	-0.31	2.42	13.5
0.00	260^{+100}_{-80}	320^{+130}_{-110}	1.36	1.60	-0.62	2.23	12.5

Table 5.16: Summary of Z significances (in units of standard deviations) for the five $f_{q\bar{q}}$ fractions, corresponding to the p_0 -values in Table 5.15. Additionally, the second and third columns show the fitted numbers of $J^P = 0^+$ and $J^P = 2^+$ signal events, respectively, while the last column gives the numbers of toy experiments generated for each fraction.

	0^+	1^+	WW	WZ/ZZ	$W\gamma$	$W\gamma^*$	$t\bar{t}$	Single Top	$Z/\gamma^* + \text{jets}$	$W + \text{jets}$
pre-fit	170	157	2029	63	83	83	171	130	309	278
post-fit $J^P = 0^+$	260	—	2271	63	79	83	169	130	306	247
post-fit $J^P = 1^+$	—	241	2340	63	81	80	167	127	286	217
pre-fit $Z \rightarrow \tau\tau$ CR	4	7	93	6	5	2	4	3	3582	95
post-fit $J^P = 0^+$ $Z \rightarrow \tau\tau$ CR	6	—	102	6	4	2	4	3	3263	86
post-fit $J^P = 1^+$ $Z \rightarrow \tau\tau$ CR	—	11	103	6	5	2	4	3	3264	77
pre-fit WW CR	1	2	2212	56	34	42	333	192	56	153
post-fit $J^P = 0^+$ WW CR	1	—	2454	55	34	43	330	190	57	140
post-fit $J^P = 1^+$ WW CR	—	3	2490	55	33	39	325	187	53	126

Table 5.17: Pre-fit and post-fit rounded event yields for the various background components (as used in the fit) and the two alternative signal hypotheses, for $J^P = 0^+$ and $J^P = 1^+$ fits, in the signal and control regions. The statistical and systematic uncertainties are all taken into account.

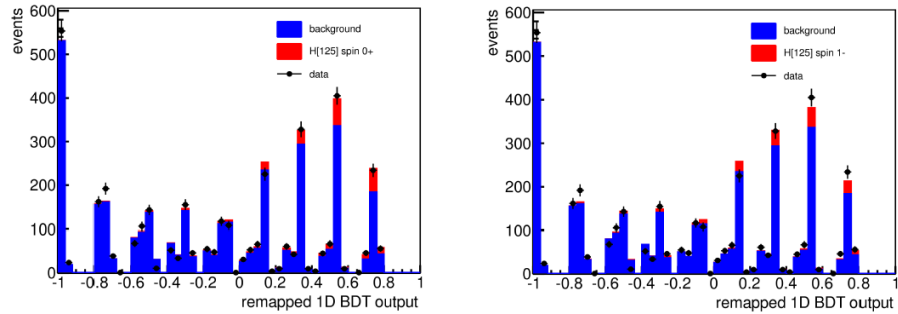


Figure 5.29: Unrolled 1-dimensional distributions for the (a) $J^P = 0^+$ and the (b) $J^P = 1^-$ signal hypotheses, as well as the total background and the data. The statistical and systematic uncertainties are all taken into account. The normalizations are the post-fit ones.

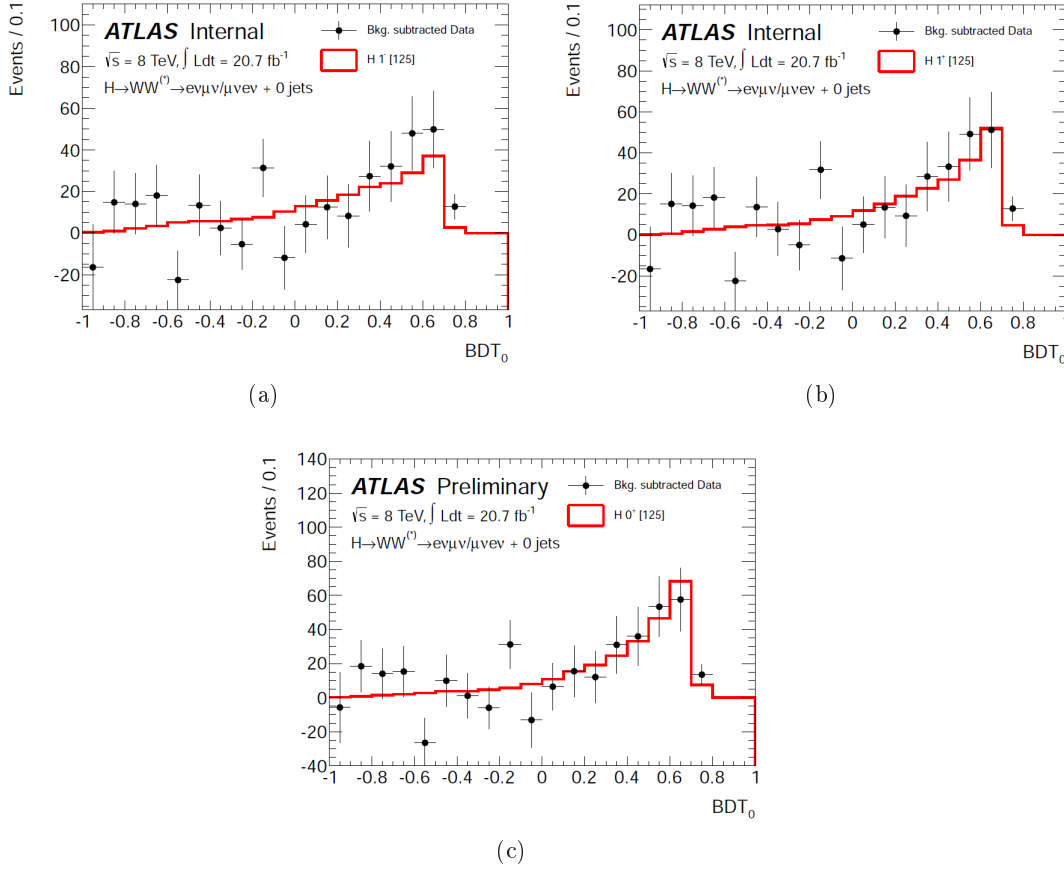


Figure 5.30: BDT0 distributions for the (a) $J^P = 1^-$, (b) $J^P = 1^+$ and (c) $J^P = 0^+$ signal hypotheses, after the background subtraction from data, using the post-fit normalizations for all backgrounds and signals. The binning is finer than the one actually used in the fit. The statistical and systematic uncertainties are all taken into account.

fake factor uncertainties and various JES and JER terms. The reason resides in the same statistical fluctuations as discussed for the $J^P = 2^+$ case, given the fact that the background-like part of BDT0 is the same for all spin analyses (cf. Fig. 5.30). Also in this case, the UEPS PDF term is not significant and is therefore discarded by the fit.

5.9.4 Signal strengths

The signal strengths and the background normalizations from the control regions, shown in Table 5.18, are similar to those found for the $J^P = 2^+$ analysis in Section 5.8.4. For $J^P = 0^+$, in particular, the values of $\hat{\mu} = 1.54^{+0.62}_{-0.48}$ ($J^P = 1^+$ analysis) and $\hat{\mu} = 1.57^{+0.65}_{-0.50}$ ($J^P = 1^-$ analysis) are in excellent agreement with the best-fit

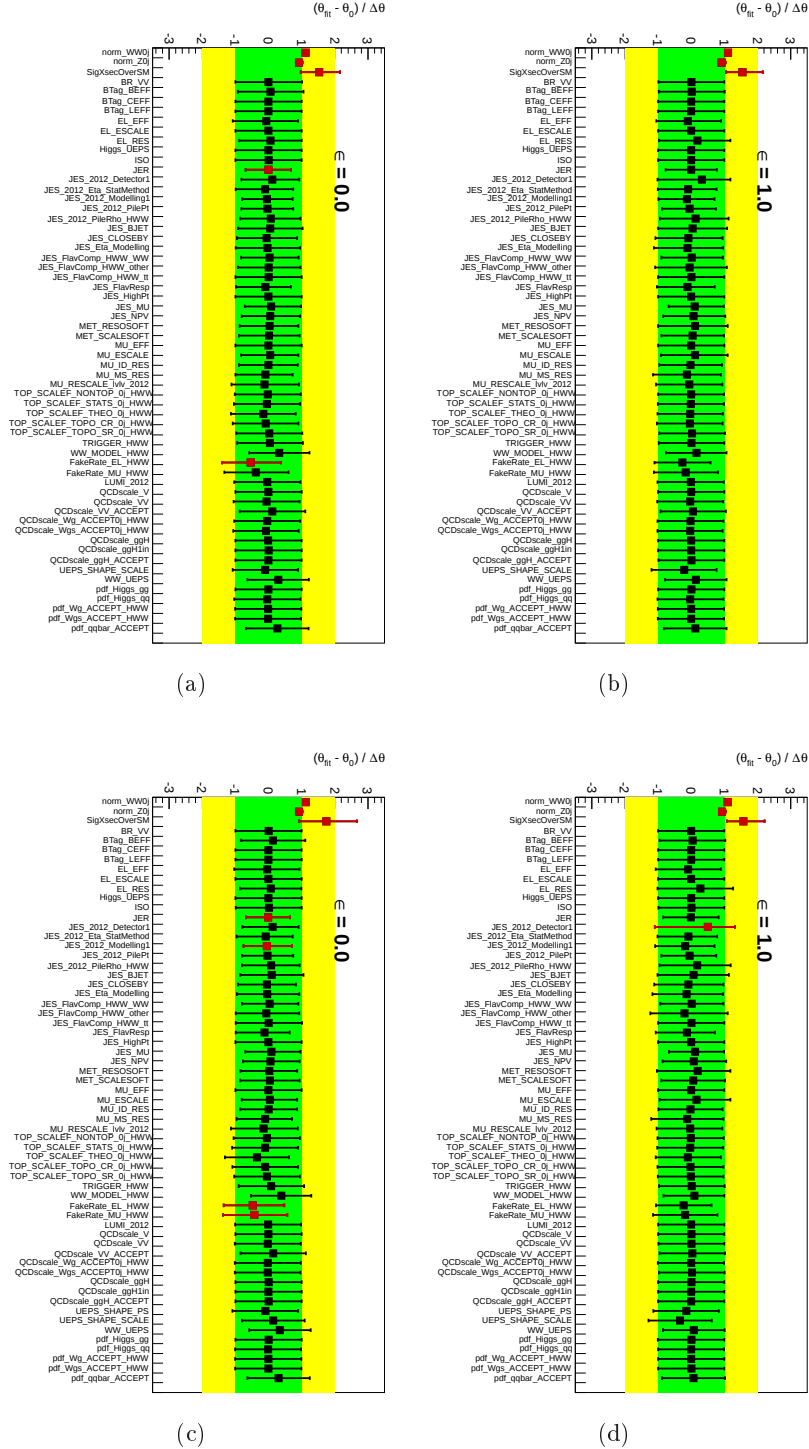


Figure 5.31: Post-fit pulls in nuisance parameters: $J^P = 1^+$ (a,b) and $J^P = 1^-$ (c,d), for the $\varepsilon = 0$ (a,c) and $\varepsilon = 1$ (b,d) hypotheses. The statistical and systematic uncertainties are all taken into account, but only those which are non negligible appear in the plots. The CR normalizations are labelled “norm_WW0j” and “norm_Z0j”, “SigXsecOverSM” is the signal strength, all other parameters follow the naming convention in Table 5.5.

value obtained in the $J^P = 2^+ f_{q\bar{q}} = 0.25$ analysis, $\hat{\mu} = 1.59^{+0.64}_{-0.50}$. Also the relative changes when using only subsets of the systematic uncertainties are similar and hence not repeated here.

	$J^P = 0^+$ fit			$J^P = 1^\pm$ fit		
	$\hat{\mu}$	WW norm	Z norm	$\hat{\mu}$	WW norm	Z norm
$J^P = 1^+$	$1.54^{+0.62}_{-0.48}$	$1.11^{+0.05}_{-0.06}$	$0.92^{+0.11}_{-0.08}$	$1.53^{+0.63}_{-0.56}$	$1.12^{+0.05}_{-0.03}$	$0.92^{+0.11}_{-0.08}$
$J^P = 1^-$	$1.57^{+0.65}_{-0.50}$	$1.10^{+0.05}_{-0.06}$	$0.92^{+0.12}_{-0.08}$	$1.74^{+0.92}_{-0.82}$	$1.13^{+0.04}_{-0.04}$	$0.92^{+0.11}_{-0.08}$

Table 5.18: Best-fit values for the signal strenghts, in units of the expected SM cross section, and normalizations for the backgrounds obtained from their respective control regions (WW and $Z \rightarrow \tau\tau$), for the $J^P = 0^+$ and $J^P = 1^\pm$ fits, with all systematic uncertainties included.

5.9.5 Sensitivities

The sensitivities for the $J^P = 1^\pm$ analysis are shown in Tables 5.19 and 5.20, in terms of p_0 -values, Z significances and CL_s confidence levels.

The sensitivities are found to be larger for $J^P = 1^-$ than for $J^P = 1^+$. The expected and observed results for $J^P = 1^\pm$ are compatible, yielding rejection probabilities of 98.3% and 92.1%, in the $J^P = 1^-$ and $J^P = 1^+$ cases, respectively. The relative uncertainties can be found in Appendix O. The fitted number of $J^P = 0^+$ signal events is the same for all spin analyses, confirming the fact that it depends mostly on BDT0.

Fig. 5.32 contains the q_μ toy distributions, normalized to unity. The observed value for the test statistic is found in both cases close to the $J^P = 0^+$ expected median.

Dependence of the sensitivities on the binning of the background distributions

As just discussed, large sensitivities are achieved by the BDT analysis, in particular an observed rejection of 2.5σ for the $J^P = 1^-$ signal hypothesis. This result may not be immediately apparent looking e.g. at the distributions in Fig. 5.30. In order to prove that sensitivities are not enhanced by pulls and overconstraints in the nuisance parameters, or more in general by informations extracted by the background fluctuations, an analogous study to the one discussed in Section 5.8.3 has been repeated for the spin 1^- analysis: the first 25 bins of the unrolled 1-dimensional BDT response are combined into one, three or five bins, and the whole fit procedure is performed

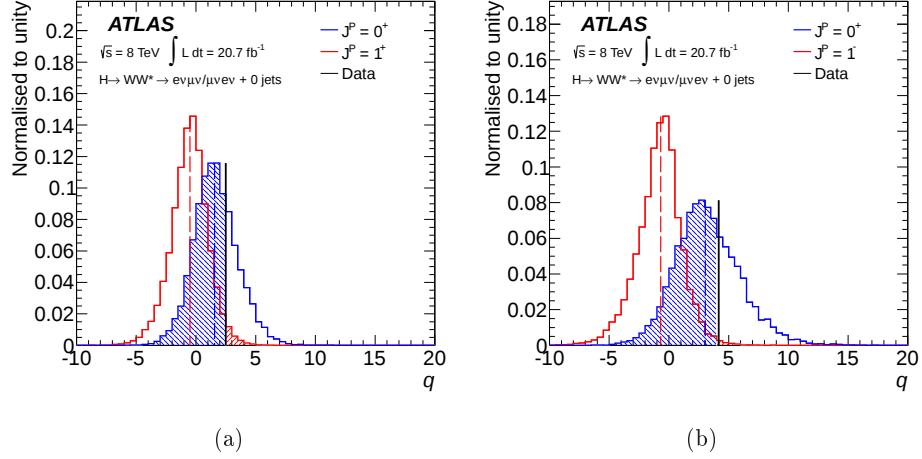


Figure 5.32: Expected distributions of the test statistic $q = \ln(\mathcal{L}(J^P = 0^+)/\mathcal{L}(J^P = 1^\pm))$ for the two $J^P = 1^\pm$ analyses, given by toy experiments generated under the $J^P = 0^+$ (blue) or $J^P = 1^\pm$ (red) hypotheses. The dashed lines indicate the expected medians, while the observed value is marked by a solid black line. The coloured areas correspond to the integrals used to compute the observed p_0 -values for rejecting each hypothesis.

again. The resulting sensitivities are shown in Table 5.21, where they are compared with the standard configuration including all systematic uncertainties. The expected sensitivities are not influenced by the binning of the background-enriched part; a finer binning provides, on the other hand, a more accurate knowledge of the various background components, and hence an increase of the observed rejections.

5.10 Spin Combination

After the finalization of the individual spin analyses in the three bosonic channels – $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$, $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow 4\ell$ – a combination has been performed, in order to fully exploit their correlations. The same approach as described for the $H \rightarrow WW^*$ analysis, is used for what concerns the definition of the $J^P = 2^+$ fractions, the theoretical models to test, most of the systematic uncertainties (apart from those analysis-dependent), the treatment of the Higgs boson p_T spectrum and the statistical approach.

Descriptions and results for the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ analyses are taken, respectively, from Refs. [138,240], while the combination is published in Refs. [41,235].

The combined best-fit value obtained by the ATLAS collaboration is used for the Higgs boson mass ($125.5 \pm 0.6 \text{ GeV}$ [234]); the spin results are insensitive to

	exp. $p_0(0^+)$	exp. $p_0(1)$	obs. $p_0(0^+)$	obs. $p_0(1)$	1-CL _s (1)
$J^P = 1^+$	0.106	0.076	0.698	0.024	0.921
$J^P = 1^-$	0.056	0.018	0.662	0.006	0.983

Table 5.19: Summary of p_0 -values for the two $J^P = 1^\pm$ analyses, with all systematic uncertainties taken into account, in terms of $\hat{\mu}$. The second and third columns show the expected sensitivities for rejecting the $J^P = 0^+$ and $J^P = 1^\pm$ hypotheses, respectively, assuming the other hypothesis to be true, while the fourth and fifth columns show the corresponding observed sensitivities. The last column contains the 1-CL_s results for $J^P = 1^\pm$ rejection.

	N _{fit} (0 ⁺)	N _{fit} (1)	exp. $Z(0^+)$	exp. $Z(1)$	obs. $Z(0^+)$	obs. $Z(1)$	N _{toys} (10 ³)
$J^P = 1^+$	260 ⁺¹⁰⁰ ₋₈₀	240 ⁺¹¹⁰ ₋₁₀₀	1.25	1.43	-0.52	1.98	17.7
$J^P = 1^-$	260 ⁺¹⁰⁰ ₋₈₀	210 ⁺¹⁰⁰ ₋₉₀	1.56	2.10	-0.42	2.52	12.5

Table 5.20: Summary of Z significances (in units of standard deviations) for the $J^P = 1^\pm$ analysis, corresponding to the p_0 -values in Table 5.19. Additionally, the second and third columns show the fitted numbers of $J^P = 0^+$ and $J^P = 1^\pm$ signal events, respectively, while the last column gives the numbers of toy experiments generated for each fraction.

variations of this parameter within the current experimental accuracy (0.6 GeV).

5.10.1 The $H \rightarrow \gamma\gamma$ analysis

The decay into two on-shell photons is not attainable by a $J^P = 1^\pm$ Higgs boson, due to the Landau-Yang theorem, as outlined in Section 2.3.2. For this reason, the $H \rightarrow \gamma\gamma$ analysis focuses on the $J^P = 2^+$ hypothesis only.

Analysis strategy

A standard cut-based approach is used in this case, with a fit to an angular variable in bins of the diphoton invariant mass. More in detail, the analysis is carried out in the Collins-Soper frame [241], where the incoming colliding partons travel along the z axis, the Higgs boson decays at rest and θ^* is the photon polar angle. This choice is expected to minimize the impact of the initial state radiation and to improve the spin discrimination with respect to other definitions. Before the cuts, the distribution of the $J^P = 0^+$ decay products is by definition isotropic in this frame, while for $J^P = 2^+$ it depends on θ^* , in a way that changes with $f_{q\bar{q}}$. In particular, it is $dN/d\cos\theta^* =$

	exp. $p_0(0^+)$	exp. $p_0(1^-)$	obs. $p_0(0^+)$	obs. $p_0(1^-)$	1-CL _s (1^-)
standard	0.06	0.02	0.67	0.01	0.983
5 bins	0.06	0.02	0.61	0.01	0.969
3 bins	0.06	0.03	0.56	0.02	0.958
1 bin	0.05	0.02	0.51	0.02	0.958

Table 5.21: Significances for the $J^P = 1^-$ analysis, with all systematic uncertainties taken into account, for the standard configuration and three alternative binnings: the first 25 bins of the unrolled 1-dimensional distribution are combined into one, three or five bins, respectively.

$1 + \cos^4 \theta^* + 6 \cos^2 \theta^*$ for pure gg production and $dN/d \cos \theta^* = 1 - \cos^4 \theta^*$ for pure $q\bar{q}$ production [35].

The object definition and event selection follow those of the corresponding rate analysis [39], requiring events with two photons, with an invariant mass in the range $105 \text{ GeV} < m_{\gamma\gamma} < 160 \text{ GeV}$, which is divided further into a central SR and two sidebands serving as CRs. The cosine of the polar angle is given by:

$$\cos \theta^* = \frac{\sinh(\eta_{\gamma_1} - \eta_{\gamma_2})}{\sqrt{1 + (p_T^{\gamma\gamma}/m_{\gamma\gamma})^2}} \cdot \frac{2p_T^{\gamma_1} p_T^{\gamma_2}}{m_{\gamma\gamma}^2}. \quad (5.8)$$

Its distribution is shown in Fig. 5.33 (a): an hypothetical $J^P = 2^+$ signal produced by gluon-fusion would imitate the background shape, peaking in the forward-backward direction, while $J^P = 0^+$ (not isotropic after the selection) tends to lower values. It can be noticed that the $J^P = 2^+$ shape becomes more similar to the $J^P = 0^+$ one when increasing $f_{q\bar{q}}$, thus a maximum sensitivity for pure gluon-fusion production is expected.

After decorrelating $m_{\gamma\gamma}$ and $\cos \theta^*$ by means of specific cuts, it is possible to define the likelihood as:

$$-\ln \mathcal{L} = (n_S + n_B) - \sum_{\text{events}} \ln [n_S \cdot f_S(|\cos \theta^*|) \cdot f_S(m_{\gamma\gamma}) + n_B \cdot f_B(|\cos \theta^*|) \cdot f_B(m_{\gamma\gamma})] \quad (5.9)$$

where $f_{S,B}$ are the PDFs for the signal and the background, respectively. For the signal, both $f_S(|\cos \theta^*|)$ and $f_S(m_{\gamma\gamma})$ are taken from MC simulation; for the invariant mass, in particular, the same distribution is assumed for both spin hypotheses, so that the polar angle is the discriminating variable. For the background, $f_B(|\cos \theta^*|)$ is obtained from a fit in the sidebands, while $f_B(m_{\gamma\gamma})$ is modelled with a fifth-degree

polynomial.

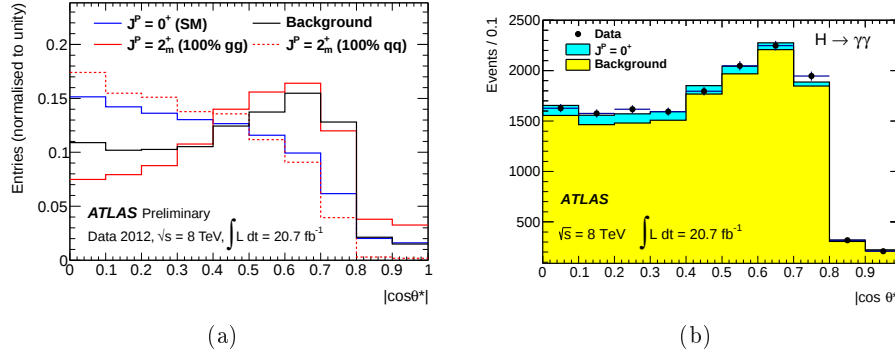


Figure 5.33: (a) Expected distributions of $|\cos \theta^*|$ for the signal ($J^P = 0^+$ in blue, $J^P = 2^+$ $f_{q\bar{q}} = 0.00$ in solid red, $J^P = 2^+$ $f_{q\bar{q}} = 1.00$ in dashed red), compared with the observed background shape, obtained from a fit in the sidebands [240]. (b) Observed distribution of $|\cos \theta^*|$ in the signal region: the data are overlaid with the post-fit sum of $J^P = 0^+$ and of the total background [235].

For backgrounds, the most important systematic effect comes from the limited number of events in the sidebands, and is of the order of a few percent.

For the signal, on the other hand, the most important systematic uncertainties are the interference with the continuous $gg \rightarrow \gamma\gamma$ background and the invariant mass resolution ($\pm 20\%$ uncertainty). For $J^P = 2^+$, the Higgs boson p_T is described up to NLO when reweighting its spectrum from JHU to POWHEG, without the need for additional uncertainties. For the quark-initiated process, however, the impact of such a reweighting is not negligible and it decreases the $\gamma\gamma$ sensitivity by about 30% for pure $q\bar{q}$ production.

Results

The post-fit (under the $J^P = 0^+$ hypothesis) $|\cos \theta^*|$ distribution is shown in Fig. 5.33 (b): in the signal region, 14977 events are observed, to be compared with 14300 and 370 expected for the total background and the $J^P = 0^+$ signal with $\mu = 1$, respectively. The fitted signal strength is almost double the SM expectation, with 690 ± 150 fitted events (620 ± 160 for $J^P = 2^+$ $f_{q\bar{q}} = 0.00$). Fig. 5.34 displays the background-subtracted data, for (a) the $J^P = 0^+$ and (b) the $J^P = 2^+$ $f_{q\bar{q}} = 0.00$ signal hypotheses.

Table 5.22 summarizes the sensitivities obtained in the $H \rightarrow \gamma\gamma$ channel, in terms of p_0 -values and CL_s confidence levels. A $J^P = 2^+$ Higgs boson can be excluded with probabilities ranging between 99.3% and 66.3%; as expected, the sensitivities decrease with increasing $f_{q\bar{q}}$, as the $|\cos \theta^*|$ distributions of $J^P = 0^+$ and $J^P = 2^+$

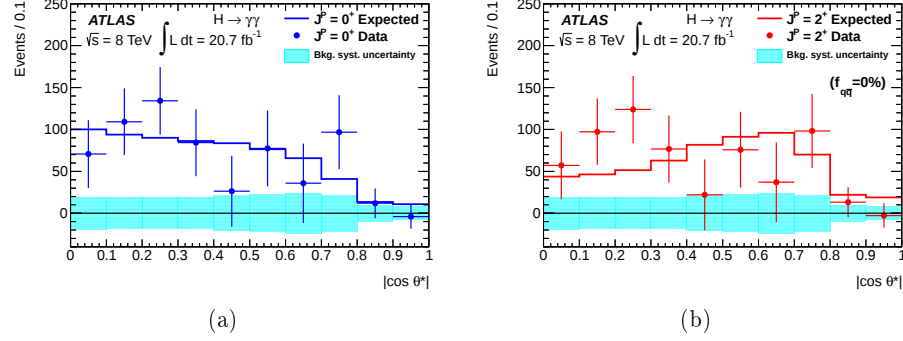


Figure 5.34: Background-subtracted data as a function of $|\cos \theta^*|$ in the signal region, for (a) the $J^P = 0^+$ fit and (b) the $J^P = 2^+$ $f_{q\bar{q}} = 0.00$ one. The cyan bands correspond to the envelope of the systematic uncertainties on the background modelling before the fits, including also the statistical uncertainty on the data sidebands. The figures are taken from Ref. [235].

become more and more similar. The results for all fractions are in good agreement with the $J^P = 0^+$ hypothesis.

$f_{q\bar{q}}$	exp. $p_0(0^+)$	exp. $p_0(2^+)$	obs. $p_0(0^+)$	obs. $p_0(2^+)$	1-CL _s (2^+)
1.00	0.148	0.135	0.798	0.025	0.876
0.75	0.319	0.305	0.902	0.033	0.663
0.50	0.198	0.187	0.708	0.076	0.740
0.25	0.052	0.039	0.609	0.021	0.946
0.00	0.012	0.005	0.588	0.003	0.993

Table 5.22: Summary of p_0 -values for the $H \rightarrow \gamma\gamma$ analysis, with all systematic uncertainties taken into account, in terms of $\hat{\mu}$ [235]. The second and third columns show the expected sensitivities for the $J^P = 0^+$ and $J^P = 2^+$ rejection, respectively, assuming the other hypothesis to be true, while the fourth and fifth columns show the corresponding observed sensitivities. The last column contains the 1-CL_s results for $J^P = 2^+$ rejection.

5.10.2 The $H \rightarrow ZZ^* \rightarrow 4\ell$ analysis

As discussed in Section 2.3, the decay of a Higgs boson into vector bosons, with their subsequent decay into four fermions, $H \rightarrow VV^* \rightarrow (f_1\bar{f}_2)(f_3\bar{f}_4)$, can be described by five quantities, the Cabibbo-Maksymowicz variables [139]. These are in particular the polar and azimuthal angles of f_1 and f_3 in the rest frame of the vector bosons,

$(\theta_1, 0)$ and (θ_3, ϕ_3) , respectively, and the masses of the two vector bosons. This property, that cannot be used in the $H \rightarrow WW^*$ case due to the presence of the two neutrinos, is exploited in the $H \rightarrow ZZ^* \rightarrow 4\ell$ (shortened as $H \rightarrow ZZ^*$ in the following) spin analysis, which tests the SM 0^+ hypothesis against the $J^P = 0^-, 1^\pm, 2^+$ hypotheses, the latter in five $f_{q\bar{q}}$ fractions.

Analysis strategy

The object definition, event selection and background estimation follow those of the rate analysis, also documented in Ref. [138]. Both 2011 and 2012 data are combined (corresponding to an integrated luminosity of 4.6 fb^{-1} at 7 TeV and of 20.7 fb^{-1} at 8 TeV, respectively). For each lepton quadriplet, all same-flavour, opposite-sign combinations are built, and the two are chosen, whose invariant mass is closer to the Z boson mass; afterwards, cuts on the invariant masses of the two pairs are applied. Four subchannels are considered, $4e$, 4μ , $2e2\mu$ and $2\mu2e$. The spin analysis uses only the mass region $115 \text{ GeV} < m_{4\ell} < 130 \text{ GeV}$, splitting it subsequently into a region with a high signal-to-background ratio ($121 \text{ GeV} < m_{4\ell} < 127 \text{ GeV}$) and two sidebands ($115 \text{ GeV} < m_{4\ell} < 121 \text{ GeV}$ and $127 \text{ GeV} < m_{4\ell} < 130 \text{ GeV}$).

A BDT analysis is performed on the selected events, training the SM signal hypothesis against each one of the others. The following variables are used for training: θ_1 , θ_2 , Φ , Φ_1 , θ^* , m_{12} and m_{34} , as defined in Fig. 2.10. For $J^P = 0^+$ versus 0^- , however, Φ_1 and θ^* are removed, since for a $J^P = 0^+$ Higgs boson such angles cannot be defined.

Systematic uncertainties affect mainly the shapes of the BDT response and the normalizations of the high and low S/B mass regions; they are due to uncertainties on the lepton identification efficiencies, energy scale and resolution (all lepton systematic uncertainties are in fact in common with the $H \rightarrow WW^*$ analysis). The experimental uncertainty on m_H translates into an additional $\pm 10\%$ error on the normalizations of the mass regions. The effect of the Higgs boson p_T reweighting is negligible alike in the $H \rightarrow WW^*$ analysis and hence discarded.

Results

At the end of the selection, 43 data events are observed, to be compared with 16 events expected from the backgrounds, mainly ZZ^* , and 18 signal events predicted for a SM Higgs boson. The corresponding BDT responses are shown in Fig. 5.35, for the 0^+ versus 0^- , 0^+ versus 1^+ and 0^+ versus 2^+ ($f_{q\bar{q}} = 0.00$) analyses. All signals are normalized using their post-fit strengths. The shapes of the BDT responses look most different for the $J^P = 0^+$ vs 0^- analysis; concerning the $J^P = 0^+$ vs 2^+

analysis, little dependence is found on the production mode, so that the results are very similar for the five fractions.

Table 5.23 contains a summary of the rejection probabilities for all J^P hypotheses in the $H \rightarrow ZZ^*$ channel. A pure pseudoscalar Higgs boson can be excluded with 97.8% confidence level, a $J^P = 1^\pm$ Higgs boson with 99.8% and 94%, for the $J^P = 1^+$ and $J^P = 1^-$ states, respectively, while a $J^P = 2^+$ Higgs boson can be excluded with confidence levels ranging between 83% and 97.4%. As in the $H \rightarrow \gamma\gamma$ and $H \rightarrow WW^*$ cases, the observed rejections for the alternative hypotheses tend to be larger than the expected ones, primarily due to the fitted signal strengths, which are larger than the SM one in all cases.

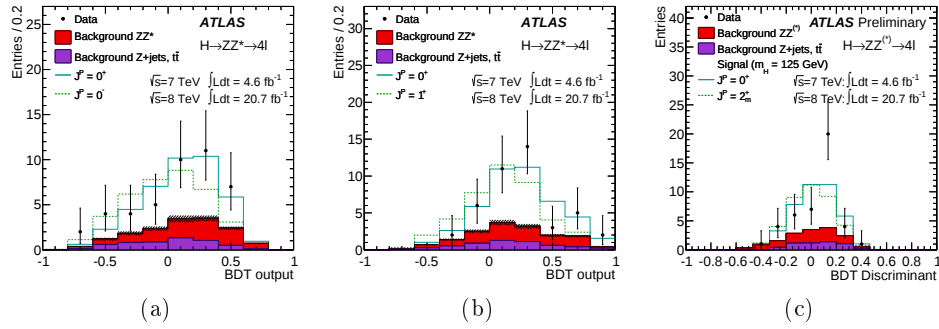


Figure 5.35: Distributions of the BDT responses in the signal region for (a) 0^+ versus 0^- , (b) 0^+ versus 1^+ and (c) 0^+ versus 2^+ $f_{q\bar{q}} = 0.00$, using the combined 2011 and 2012 datasets. The SM signal corresponds to the green solid line, the alternative hypothesis to the dotted one. All signals are normalized using their post-fit strengths. The hatched areas represent the total background uncertainty. The figures are taken from Refs. [138, 235].

5.10.3 Combined results

In order to perform the statistical combination, one crucial aspect to take into account resides in the common systematic uncertainties and their correlations. Systematic uncertainties on leptons (electron and muon selection efficiencies, as well as their energy and momentum resolution) are common to both the $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$ analyses; moreover, the uncertainties on the energy scale are highly correlated for electrons and photons.

The overall impact of the systematic uncertainties is evaluated by comparing the baseline results, with all nuisance parameters profiled on data, to those obtained keeping them fixed at their nominal values: the difference in the combined rejection significances is at the level of 0.3σ .

	exp. $p_0(0^+)$	exp. $p_0(\text{ALT})$	obs. $p_0(0^+)$	obs. $p_0(\text{ALT})$	$\text{CL}_s (\text{ALT})$
Spin- 0^-	$1.5 \cdot 10^{-3}$	$3.7 \cdot 10^{-3}$	0.31	0.015	0.022
Spin- 1^+	$4.6 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	0.55	$1.0 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$
Spin- 1^-	$0.9 \cdot 10^{-3}$	$3.8 \cdot 10^{-3}$	0.15	0.051	0.060
Spin- 2^+ , $f_{q\bar{q}} = 1.00$	0.102	0.082	0.962	0.001	0.026
Spin- 2^+ , $f_{q\bar{q}} = 0.75$	0.117	0.099	0.923	0.003	0.039
Spin- 2^+ , $f_{q\bar{q}} = 0.50$	0.129	0.113	0.943	0.002	0.035
Spin- 2^+ , $f_{q\bar{q}} = 0.25$	0.125	0.107	0.944	0.002	0.036
Spin- 2^+ , $f_{q\bar{q}} = 0.00$	0.099	0.092	0.532	0.079	0.169

Table 5.23: Summary of p_0 -values for the $H \rightarrow ZZ^* \rightarrow 4\ell$ analysis, with all systematic uncertainties taken into account, in terms of $\hat{\mu}$ [235]. The second and third columns show the expected sensitivities for rejecting one hypothesis, assuming the other to be true, while the fourth and fifth columns show the corresponding observed sensitivities. The last column contains the CL_s confidence levels for rejecting the alternative hypothesis.

Concerning the systematic uncertainty on the Higgs boson p_T spectrum, as already explained, only the $H \rightarrow \gamma\gamma$ channel is sensitive to it, especially at high values of $|\cos\theta^*|$; however, given the fact that the $J^P = 2^+$ sensitivities are dominated by the $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$ final states, the combined results are impacted at the level of 0.1σ .

The combined results are shown in Figs. 5.37, 5.38 and in Table 5.24, for the $J^P = 1^\pm$ and $J^P = 2^+$ analyses (since only the $H \rightarrow ZZ^*$ channel tests the $J^P = 0^-$ hypothesis, no combination is needed for it).

Concerning $J^P = 2^+$, the dependence of the sensitivities on the production fraction is illustrated in Fig. 5.37 for the individual channels and their combination: as already discussed, while for $H \rightarrow WW^*$ the sensitivities increase with increasing $f_{q\bar{q}}$, they decrease for $H \rightarrow \gamma\gamma$ and are stable for $H \rightarrow ZZ^*$. As a result, the combined values also slightly increase, with the observed values nicely following the expected ones. The yellow and green bands in these plots correspond to one and two standard deviations from the central expected values (assuming the SM hypothesis) and are obtained computing the quantiles of the corresponding toy q_μ distributions (approximated as Gaussians for this purpose). Two examples of the latter are shown in Fig. 5.36, in logarithmic scale: millions of toys are generated for each spin hypothesis, in order to simulate the tails properly.

Moreover, Fig. 5.38 translates the confidence levels into standard deviations and shows a summary for all the J^P hypotheses considered. In particular, the result for

$J^P = 2_m^+$ refers to the minimal hypothesis of leading-order couplings, $f_{q\bar{q}} = 0.04$, and is obtained by interpolating the results in Table 5.24.

The combined sensitivities reach 99.97% CL for $J^P = 1^+$ and 99.7% CL for $J^P = 1^-$, and are above 99.9% CL for all $J^P = 2^+$ fractions. In all cases, the observed values agree with the SM assignment.

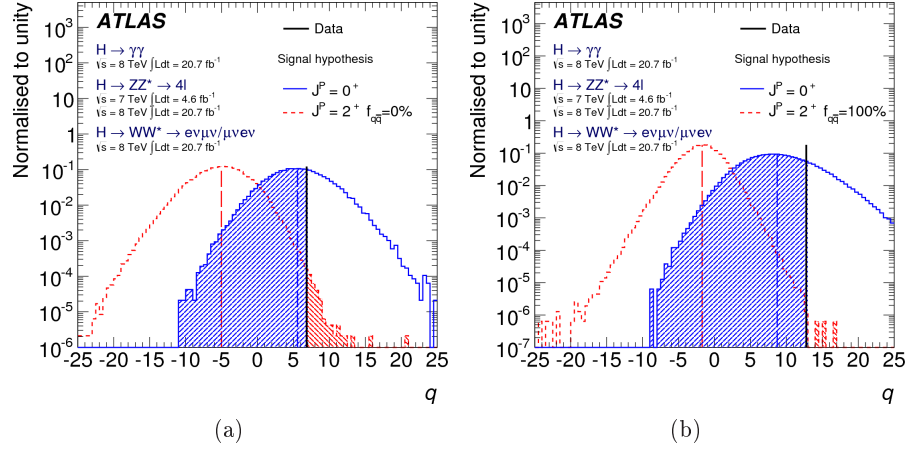


Figure 5.36: Expected distributions of the test statistic $q = \ln(\mathcal{L}(J^P = 0^+)/\mathcal{L}(J^P = 2^+))$, for the combination of the three bosonic channels, for the two extreme fractions ((a) $f_{q\bar{q}} = 0.00$, (b) $f_{q\bar{q}} = 1.00$). The distributions are obtained by means of toy experiments under the $J^P = 0^+$ (solid blue) or $J^P = 2^+$ (red dotted) hypotheses. The dashed lines indicate the expected medians, while the observed value is marked by a solid black line. The coloured areas correspond to the integrals used to compute the observed p_0 -values for rejecting each hypothesis. The figures are taken from Ref. [235].

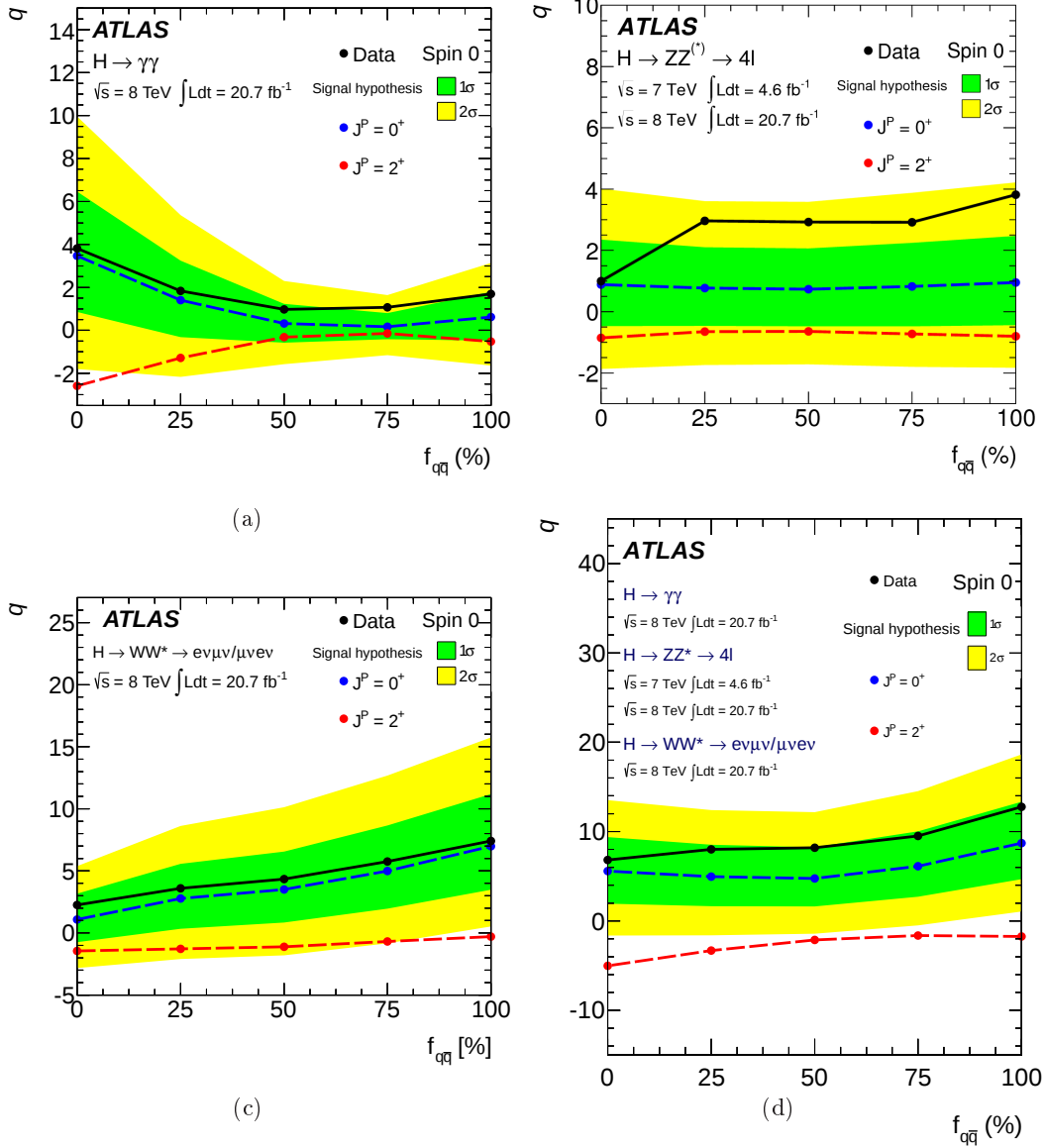


Figure 5.37: Summary of the observed values (black solid line) of the test statistic $q = \ln(\mathcal{L}(J^P = 0^+)/\mathcal{L}(J^P = 2^+))$ for the five $f_{q\bar{q}}$ fractions of $J^P = 2^+$ production. The blue and red dashed lines indicate the expected values for the medians of the sampling distributions obtained from pseudo-experiments, for the $J^P = 0^+$ and $J^P = 2^+$ signals, respectively. Moreover, the green and yellow bands correspond to one and two standard deviations around the $J^P = 0^+$ median. The channels shown are: (a) $H \rightarrow \gamma\gamma$, (b) $H \rightarrow ZZ^*$, (c) $H \rightarrow WW^*$, (d) their combination. The figures are taken from Ref. [235].

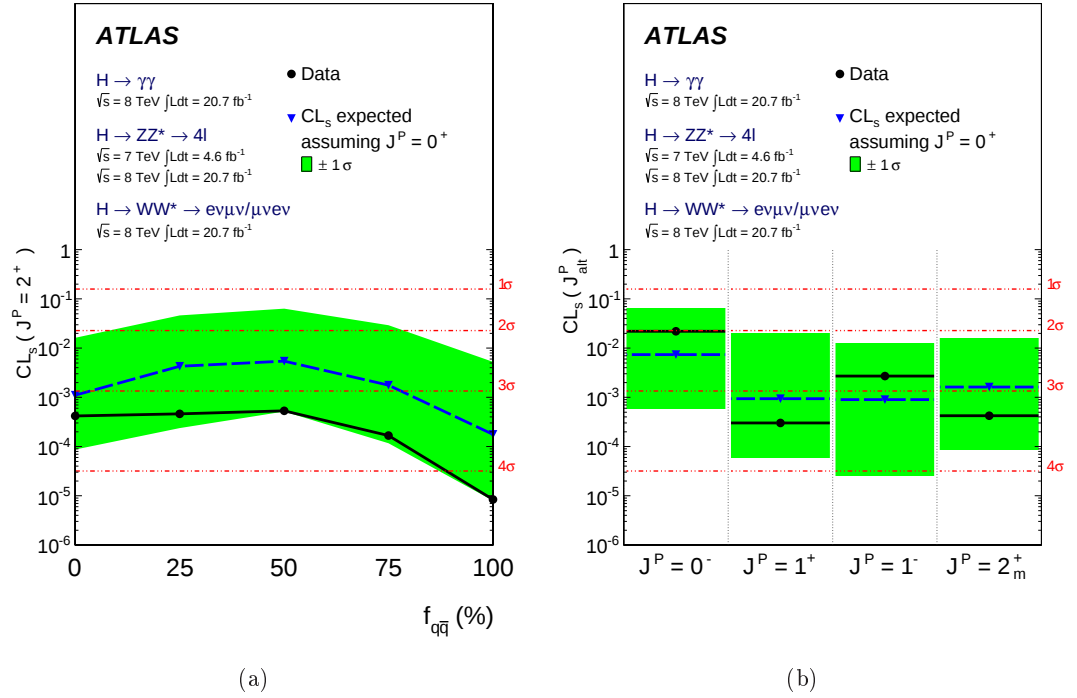


Figure 5.38: (a) Expected (blue triangles/dashed line) and observed (black circles/solid line) confidence levels for rejecting the $J^P = 2^+$ hypothesis as a function of the $f_{q\bar{q}}$ fraction. The green bands represent the 68% expected exclusion range assuming a $J^P = 0^+$ signal. On the right vertical axis, the corresponding standard deviations are given. (b) The same for all spin hypotheses. The $J^P = 2^+$ value refers to the model with LO couplings, $f_{q\bar{q}} = 0.04$, obtained by interpolating the results in Table 5.24. The figures are taken from Ref. [235].

	exp. $p_0(0^+)$	exp. $p_0(\text{ALT})$	obs. $p_0(0^+)$	obs. $p_0(\text{ALT})$	$\text{CL}_s (\text{ALT})$
Spin-1 ⁺ : $H \rightarrow ZZ^*$	$4.6 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	0.55	$1.0 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$
Spin-1 ⁺ : $H \rightarrow WW^*$	0.11	0.08	0.70	0.02	0.08
Spin-1 ⁺ : Combination	$2.7 \cdot 10^{-3}$	$4.7 \cdot 10^{-4}$	0.62	$1.2 \cdot 10^{-4}$	$3.0 \cdot 10^{-4}$
Spin-1 ⁻ : $H \rightarrow ZZ^*$	$0.9 \cdot 10^{-3}$	$3.8 \cdot 10^{-3}$	0.15	0.051	0.060
Spin-1 ⁻ : $H \rightarrow WW^*$	0.06	0.02	0.66	0.006	0.017
Spin-1 ⁻ : Combination	$1.4 \cdot 10^{-3}$	$3.6 \cdot 10^{-4}$	0.33	$1.8 \cdot 10^{-3}$	$2.7 \cdot 10^{-3}$
Spin-2 ⁺ $f_{q\bar{q}} = 1.00$, $H \rightarrow \gamma\gamma$	0.148	0.135	0.798	0.025	0.124
Spin-2 ⁺ $f_{q\bar{q}} = 1.00$, $H \rightarrow ZZ^*$	0.102	0.082	0.962	0.001	0.026
Spin-2 ⁺ $f_{q\bar{q}} = 1.00$, $H \rightarrow WW^*$	0.013	$3.6 \cdot 10^{-4}$	0.541	$1.7 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$
Spin-2 ⁺ $f_{q\bar{q}} = 1.00$, Combination	$3.0 \cdot 10^{-3}$	$8.8 \cdot 10^{-5}$	0.81	$1.6 \cdot 10^{-6}$	$0.8 \cdot 10^{-5}$
Spin-2 ⁺ $f_{q\bar{q}} = 0.75$, $H \rightarrow \gamma\gamma$	0.319	0.305	0.902	0.033	0.337
Spin-2 ⁺ $f_{q\bar{q}} = 0.75$, $H \rightarrow ZZ^*$	0.117	0.099	0.923	0.003	0.039
Spin-2 ⁺ $f_{q\bar{q}} = 0.75$, $H \rightarrow WW^*$	0.028	0.003	0.586	0.001	0.003
Spin-2 ⁺ $f_{q\bar{q}} = 0.75$, Combination	$9.5 \cdot 10^{-3}$	$8.8 \cdot 10^{-4}$	0.81	$3.2 \cdot 10^{-5}$	$1.7 \cdot 10^{-4}$
Spin-2 ⁺ $f_{q\bar{q}} = 0.50$, $H \rightarrow \gamma\gamma$	0.198	0.187	0.708	0.076	0.260
Spin-2 ⁺ $f_{q\bar{q}} = 0.50$, $H \rightarrow ZZ^*$	0.129	0.113	0.943	0.002	0.035
Spin-2 ⁺ $f_{q\bar{q}} = 0.50$, $H \rightarrow WW^*$	0.042	0.009	0.616	0.003	0.008
Spin-2 ⁺ $f_{q\bar{q}} = 0.50$, Combination	$1.3 \cdot 10^{-2}$	$2.7 \cdot 10^{-3}$	0.84	$8.6 \cdot 10^{-5}$	$5.3 \cdot 10^{-4}$
Spin-2 ⁺ $f_{q\bar{q}} = 0.25$, $H \rightarrow \gamma\gamma$	0.052	0.039	0.609	0.021	0.054
Spin-2 ⁺ $f_{q\bar{q}} = 0.25$, $H \rightarrow ZZ^*$	0.125	0.107	0.944	0.002	0.036
Spin-2 ⁺ $f_{q\bar{q}} = 0.25$, $H \rightarrow WW^*$	0.048	0.019	0.622	0.008	0.020
Spin-2 ⁺ $f_{q\bar{q}} = 0.25$, Combination	$6.4 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	0.80	$0.9 \cdot 10^{-4}$	$4.6 \cdot 10^{-4}$
Spin-2 ⁺ $f_{q\bar{q}} = 0.00$, $H \rightarrow \gamma\gamma$	0.012	0.005	0.588	0.003	0.007
Spin-2 ⁺ $f_{q\bar{q}} = 0.00$, $H \rightarrow ZZ^*$	0.099	0.092	0.532	0.079	0.169
Spin-2 ⁺ $f_{q\bar{q}} = 0.00$, $H \rightarrow WW^*$	0.086	0.054	0.731	0.013	0.048
Spin-2 ⁺ $f_{q\bar{q}} = 0.00$, Combination	$2.1 \cdot 10^{-3}$	$5.5 \cdot 10^{-4}$	0.63	$1.5 \cdot 10^{-4}$	$4.2 \cdot 10^{-4}$

Table 5.24: Summary of p_0 -values for $J^P = 1^\pm$ and $J^P = 2^+$ analyses, for the individual channels ($H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^*$ and $H \rightarrow WW^*$) as well as their combinations, according to the $f_{q\bar{q}}$ fraction [235]. Results include all systematic uncertainties and use best-fit values for all nuisance parameters, including $\hat{\mu}$. The second and third columns show the expected sensitivities for rejecting one hypothesis, assuming the other to be true, while the fourth and fifth columns show the corresponding observed sensitivities. The last column contains the CL_s probabilities for rejecting the alternative hypothesis.

Appendix A

Comparison with other multivariate techniques

The choice of BDT as the baseline multivariate method results from a comparison with other techniques, which BDT is able to outperform. The comparison is based on ROC curves, and secondarily on the expected exclusion limits (the latter not conclusive, since carried out with a preliminary and incomplete statistical model, and hence not reported here). ROC curves measure the background rejection as a function of the signal efficiency: they are typically used to compare different methods for a specified signal efficiency, e.g. 80%, looking for the best background rejection. Examples of these comparisons are shown in Fig. A.1, for two mass hypotheses, $m_H = 130$ GeV and $m_H = 240$ GeV, and different lepton flavours in the 0-jet channel. The techniques compared here are chosen because they are the fastest, most stable and best performing ones (all implemented in the TMVA package [216]): the boosted decision tree BDT, the *likelihood* LH and two implementations of the *neural network* technique, the *multilayer perceptron* MLP and the *k-nearest neighbor* method kNN. For this exercise, the following input variables have been used:

- BDT: $m_T, m_{\ell\ell}, R_{\ell\ell}, \Delta\phi(\ell_2, E_T^{\text{miss}}), E_{T,\text{rel}}^{\text{miss}}, p_T^{\ell_1}, p_T^{\ell_2},$
- LH: $m_T, R_{\ell\ell},$
- MLP: $m_T, m_{\ell\ell}, p_{T,\ell\ell}, \Delta\phi_{\ell\ell}, R_{\ell\ell}, p_T^{\ell_1}, p_T^{\ell_2}, E_{T,\text{rel}}^{\text{miss}}, \Delta\phi(\ell_2, E_T^{\text{miss}}),$
- NN: $p_T^{\ell_1}, p_z^{\ell_1}, p_x^{\ell_2}, p_y^{\ell_2}, p_z^{\ell_2}, E_x^{\text{miss}}, E_y^{\text{miss}}.$

As can be seen in Fig. A.1, the BDT technique outperforms all others, in all configurations under study. The performance of the MLP neural network is the closest to the BDT one, but it turns out not to be so competitive when looking at expected limits, while the simple LH method has a substantial worse performance for all

the values of the signal efficiency. These considerations lead to choose the boosted decision tree as the baseline multivariate method.

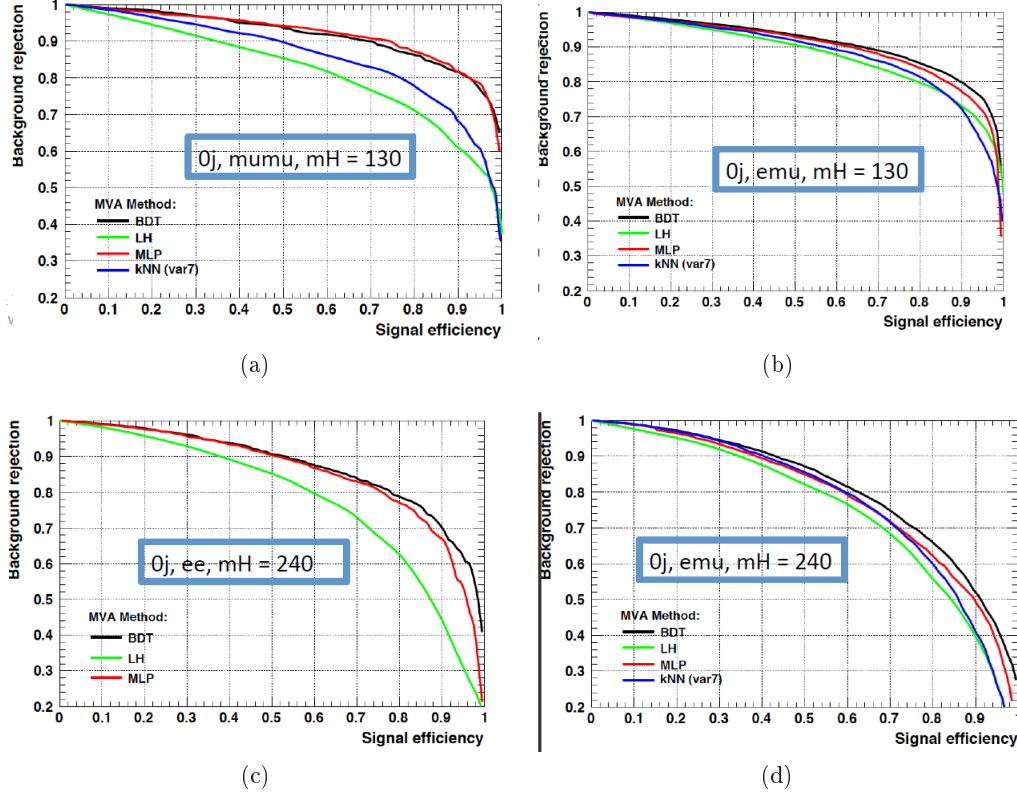


Figure A.1: ROC curves comparing different multivariate techniques: the boosted decision tree BDT, the likelihood LH, the multilayer perceptron MLP and the k-nearest neighbor kNN. Input variables are described in the text. Various configurations have been tested, all in the 0-jet channel: (a) $m_H = 130$ GeV in the $\mu\mu$ channel, (b) $m_H = 130$ GeV in the $e\mu$ channel, (c) $m_H = 240$ GeV in the $\mu\mu$ channel, (d) $m_H = 240$ GeV in the $e\mu$ channel.

Appendix B

Choice of the input variables

The choice of the input variables for the boosted decision tree algorithm is based on the separation between the signal and the background, defined in Eq. 4.4. In total, 11 variables have been used as inputs for the training, in all possible combinations, $\binom{n}{1}, \binom{n}{2} \dots \binom{n}{n}$; the best separations achieved for each subset of n variables are summarized in Table B.1 and in Fig. B.1. The signal mass hypothesis used here is $m_H = 130$ GeV, and the training is performed in the 0-jet channel, after the jet veto. The variables are labelled in the following way:

- a) m_T b) $m_{\ell\ell}$ c) $\Delta R_{\ell\ell}$ d) $\Delta\phi_{\ell\ell}$ e) $\Delta\phi(E_T^{\text{miss}}, \ell_2)$ f) $E_{T,rel}^{\text{miss}}$
g) p_{T,ℓ_1} h) p_{T,ℓ_2} i) $p_{T,\ell_2} \times \sin(\Delta\phi_{\ell\ell})$ j) $p_{T,\ell\ell}$ j) $p_{T,tot}^{\text{scalar}}$.

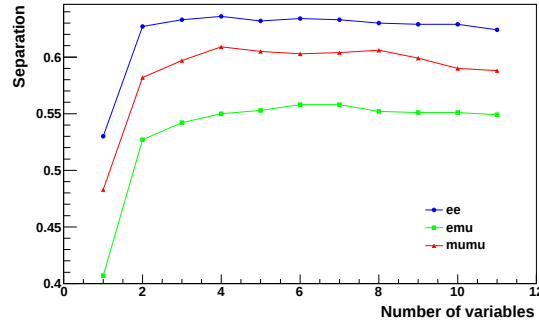


Figure B.1: Best separation for subsets of n variables, for each lepton flavour independently, in the 0-jet channel with $m_H = 130$ GeV.

As can be seen, a saturation in the separation is reached already with four variables, given the limited number of MC events and the correlations among the variables

themselves. The best performing variables are found to be:

$$\begin{aligned} ee : m_T, m_{\ell\ell}, \Delta\phi(E_T^{\text{miss}}, \ell_2), p_{T,\ell\ell} \quad , \quad e\mu : m_T, m_{\ell\ell}, \Delta\phi_{\ell\ell}, p_{T,\ell\ell} \\ \mu\mu : m_T, m_{\ell\ell}, p_{T,\ell_1}, p_{T,\ell\ell} \end{aligned}$$

with three of them ($m_T, m_{\ell\ell}, p_{T,\ell\ell}$) being in common for all lepton flavours. The fourth one is chosen to be $\Delta\phi_{\ell\ell}$ because of two considerations: it is shared with the cut-based analysis, and it pertains to the $e\mu$ 0-jet channel, which is the most sensitive. In this way, 98% of the maximal separation can be achieved for the ee final state, and 99% for the $\mu\mu$ final state.

	ee		$e\mu$		$\mu\mu$	
N variables	Variable set	Separation	Variable set	Separation	Variable set	Separation
1	c	0.530	b	0.407	c	0.483
2	ab	0.627	ac	0.527	ab	0.582
3	abj	0.633	abd	0.542	bjk	0.597
4	abej	0.636	abdj	0.550	abgj	0.609
5	bcfgj	0.632	abdej	0.553	abgij	0.605
6	abegij	0.634	abdehi	0.558	abfhij	0.603
7	bcdefij	0.633	abdeijk	0.558	abfghij	0.604
8	bcdefgij	0.630	abdeghjk	0.552	abceghik	0.606
9	bcdefghij	0.629	abdfghijk	0.551	acefghijk	0.599
10	abcdefgijk	0.629	abdefghijk	0.551	abcdefghijk	0.590
11	abcdefghijk	0.624	abcdefghijk	0.549	abcdefghijk	0.588

Table B.1: Best separation for each subset of n variables, with the indication of which variables belong to every subset, for each lepton flavour independently, in the 0-jet channel. The signal mass hypothesis is $m_H = 130$ GeV.

Repeating the exercise for $m_H = 200$ GeV, the same saturation in performance is observed, with the best variables found to be:

$$\begin{aligned} ee : m_{\ell\ell}, p_{T,\ell_2}, \Delta\phi_{\ell\ell}, p_{T,\ell\ell} \quad , \quad e\mu : m_{\ell\ell}, p_{T,\ell_1}, p_{T,\ell_2}, p_{T,\ell\ell} \\ \mu\mu : m_T, \Delta R_{\ell\ell}, p_{T,\ell_1}, p_{T,\ell_2} \end{aligned}$$

which are partly in common with those found for $m_H = 130$ GeV. Training the $m_H = 200$ GeV case with the optimal set found for $m_H = 130$ GeV, however, degrades the separation only of a few percent (2% for ee and $\mu\mu$, 6% for $e\mu$), so that the following conclusion can be reached: the four topological variables ($m_T, m_{\ell\ell}, \Delta\phi_{\ell\ell}, p_{T,\ell\ell}$) can be used as BDT inputs for all lepton flavours and for all mass points, without any substantial degradation in performance.

Appendix C

Input variables for the individual lepton flavours

The distributions of the four input variables in each lepton channel are shown here, corresponding to Figs. 4.15 and 4.17 for the combined case, in the 0- and 1-jet bins, respectively. A good agreement is found in all cases, inside the statistical and systematic uncertainties.

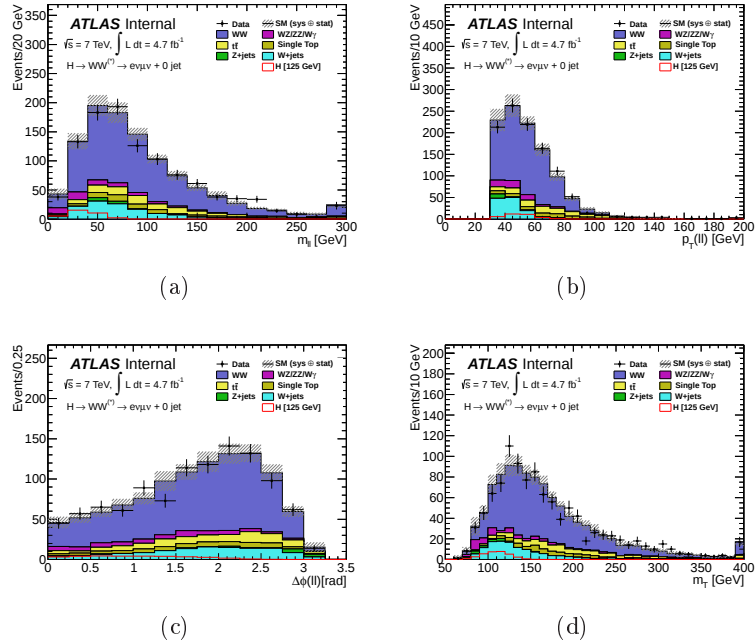


Figure C.1: Distributions of the four input variables, after the $p_{T,\ell\ell}$ requirement, as used for the BDT training, in the 0-jet $e\mu$ channel: (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . The shaded bands indicate the total normalization systematic uncertainty. The final bin includes the overflow. All backgrounds are normalized as in Table 4.2.

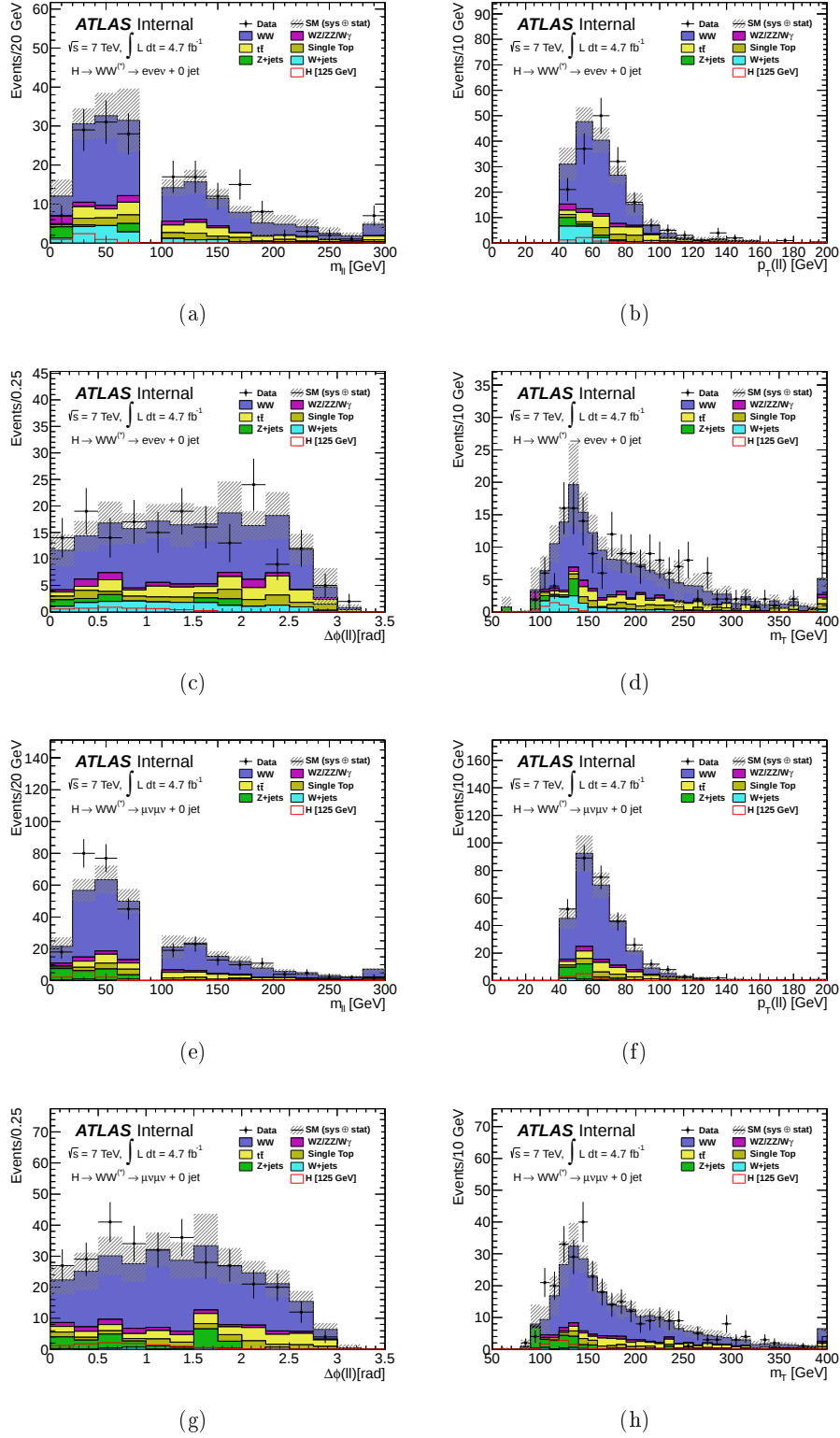


Figure C.2: Distributions of the four input variables, after the $p_{T,\ell\ell}$ requirement, as used for the BDT training, in the 0-jet same-flavour channels: (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . The shaded bands indicate the total normalization systematic uncertainty. The final bin includes the overflow. All backgrounds are normalized as in Table 4.2.

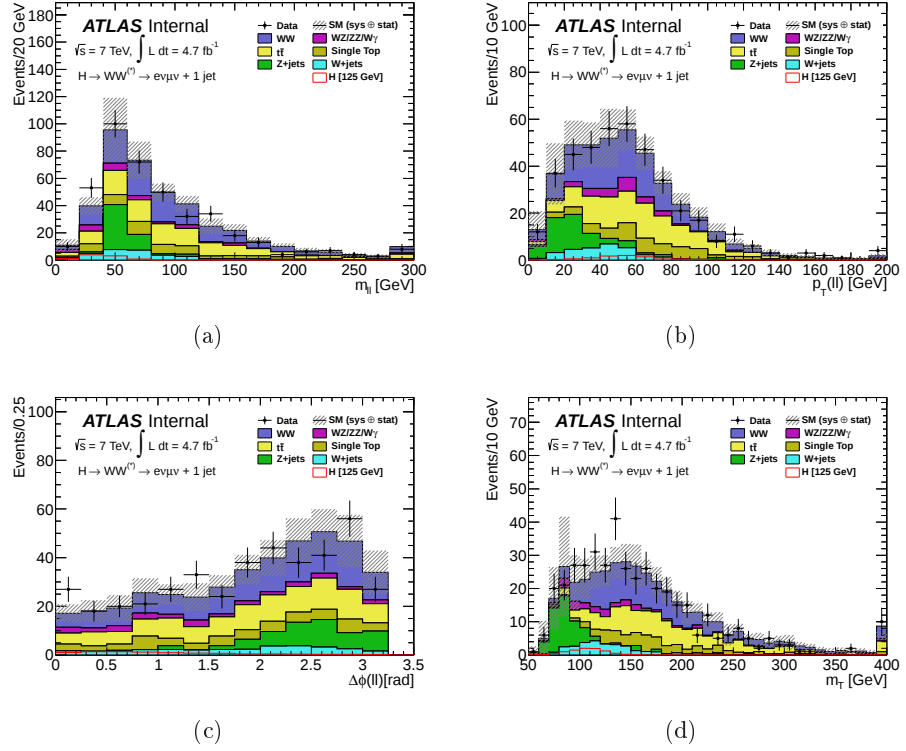


Figure C.3: Distributions of the four input variables, after the b-tag veto, the $p_{T,TOT}$ requirement and the $Z\tau\tau$ veto, in the 1-jet $e\mu$ channel: (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . The shaded bands indicate the total normalization systematic uncertainty. The final bin includes the overflow. All backgrounds are normalized as in Table 4.2.

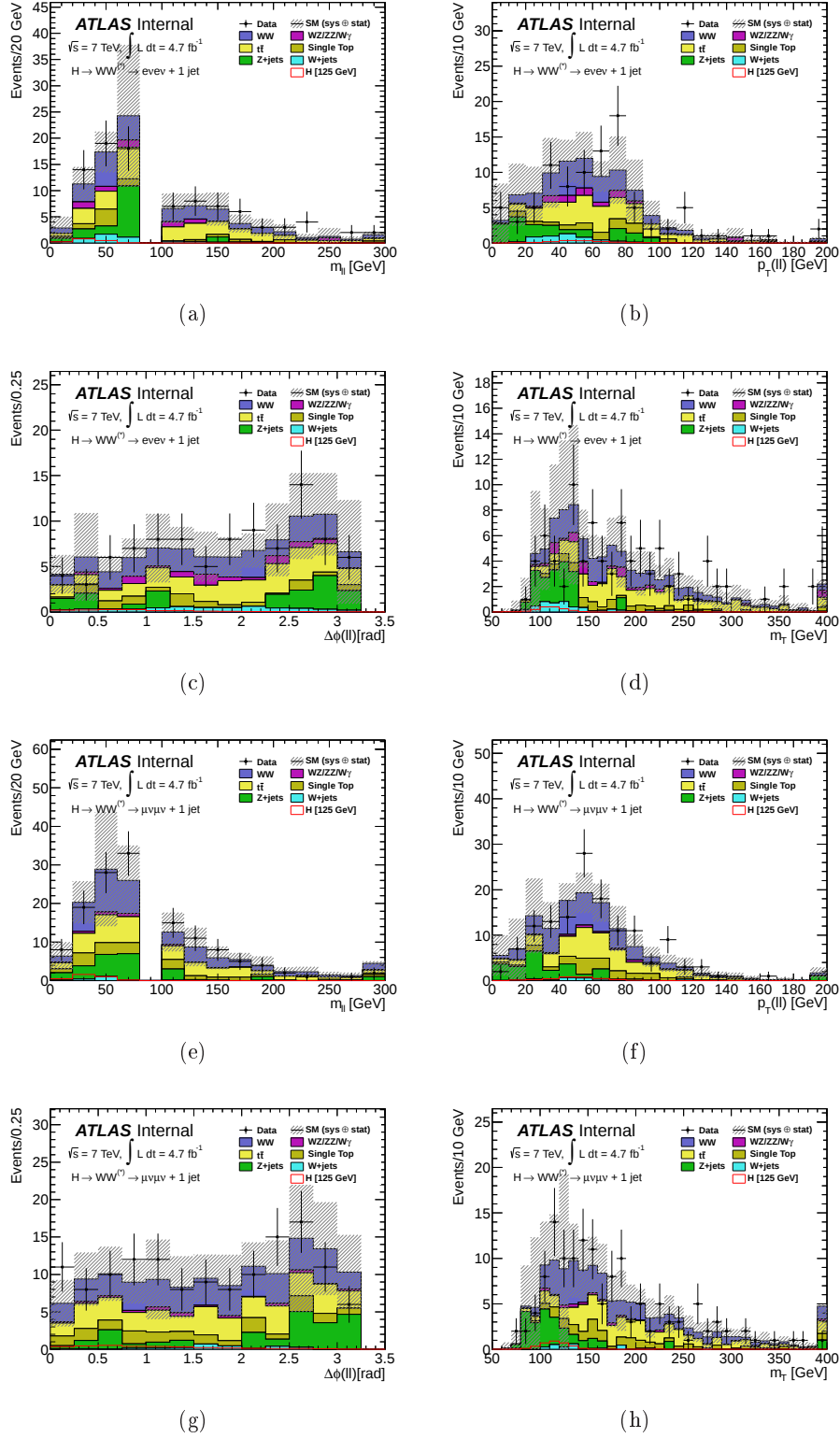


Figure C.4: Distributions of the four input variables, after the b-tag veto, the $p_{T,TOT}$ requirement and the $Z\tau\tau$ veto, in the 1-jet same-flavour channels: (a) $m_{\ell\ell}$, (b) $p_{T,\ell\ell}$, (c) $\Delta\phi_{\ell\ell}$, (d) m_T . The shaded bands indicate the total normalization systematic uncertainty. The final bin includes the overflow. All backgrounds are normalized as in Table 4.2.

Appendix D

BDT with two variables

A cross-check has been performed using only two variables to train the boosted decision tree, namely the transverse mass m_T and the azimuthal opening angle $\Delta\phi_{\ell\ell}$, keeping the rest of the analysis unchanged. This test is meant to demonstrate that the sensitivities computed with two variables are in between those obtained by the cut-based analysis (which uses only m_T) and the standard BDT analysis with four variables. For this study, five mass hypotheses have been considered: 125 GeV, 145 GeV, 165 GeV, 200 GeV and 240 GeV.

	BDT 2vars	BDT 4vars	Cut-based	BDT4/BDT2	Cut-based/BDT2
$m_H = 125$ GeV	1.10	1.05	1.15	0.95	1.05
$m_H = 145$ GeV	0.35	0.33	0.37	0.94	1.06
$m_H = 165$ GeV	0.18	0.16	0.18	0.89	1.00
$m_H = 200$ GeV	0.62	0.53	0.72	0.85	1.16
$m_H = 240$ GeV	1.22	1.06	1.10	0.87	1.11

Table D.1: Expected exclusion limits for five mass points in three configurations: BDT trained with two variables, standard BDT trained with four variables, cut-based analysis. The ratios of the latter two to the two-variable BDT are also shown.

Figs. D.1 and D.2 show the distributions of the BDT response in the 0- and 1-jet bins, respectively; their characteristics are the same as those already commented for the baseline analysis in Section 4.7. Table D.1 presents a comparison of the combined exclusion limits, $\sigma_{\text{meas}}/\sigma_{\text{SM}}$ at 95% CL, in three different configurations: the BDT

trained with two variables, the standard BDT trained with four variables⁽¹⁾ and the cut-based analysis. The four-variable BDT is always performing better than the two-variable BDT, while the cut-based analysis is always performing worse (or equally good at $m_H = 165$ GeV), demonstrating, on one hand, that the usage of $\Delta\phi_{\ell\ell}$ brings additional information with respect to m_T , on the other hand that the discrimination improves further adding $p_{T,\ell\ell}$ and $m_{\ell\ell}$.

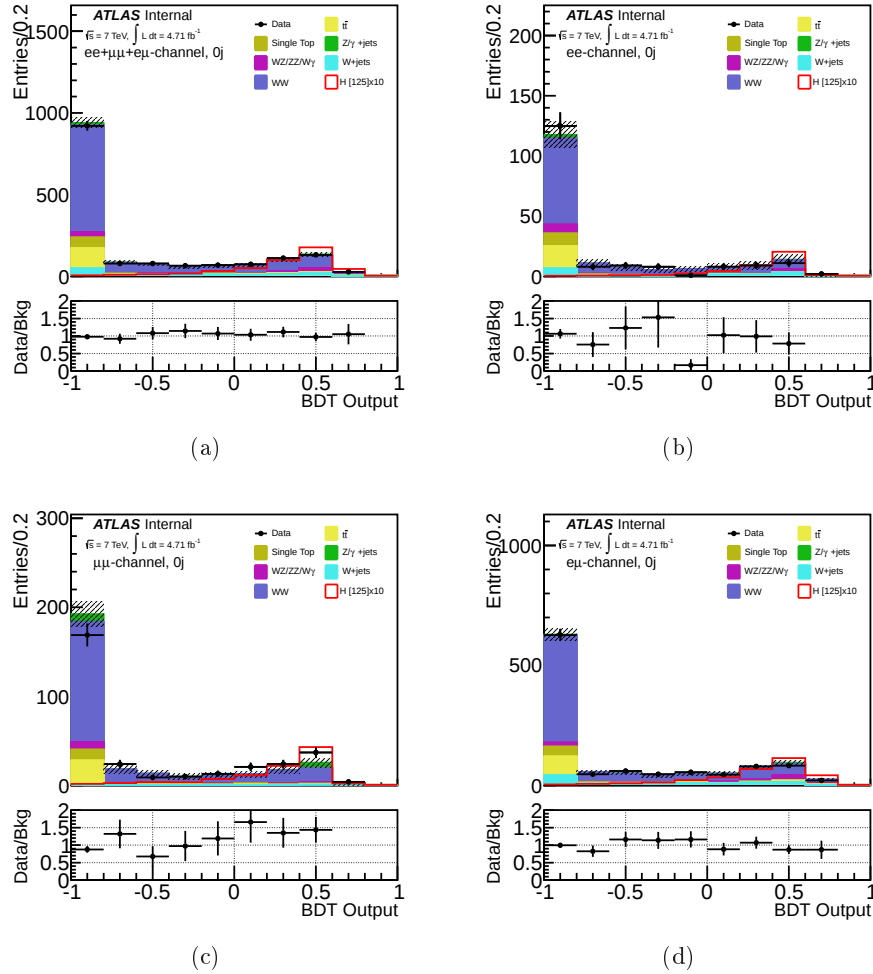


Figure D.1: Distributions of the BDT response in the data and in the MC simulation, after all selection cuts, in the 0-jet channel, obtained from a two-variable BDT training. Shown are the (a) $\ell\ell$, (b) ee , (c) $\mu\mu$ and (d) $e\mu$ final states, respectively. The signal contribution, shown for $m_H = 125$ GeV, has been multiplied by a factor ten for better visibility.

¹It should be noticed that the numbers for the standard four-variable BDT cannot be directly compared with those given in Table 4.12 because this study has been performed with a previous analysis version (small differences in recommendations for object definitions, and in MC samples).

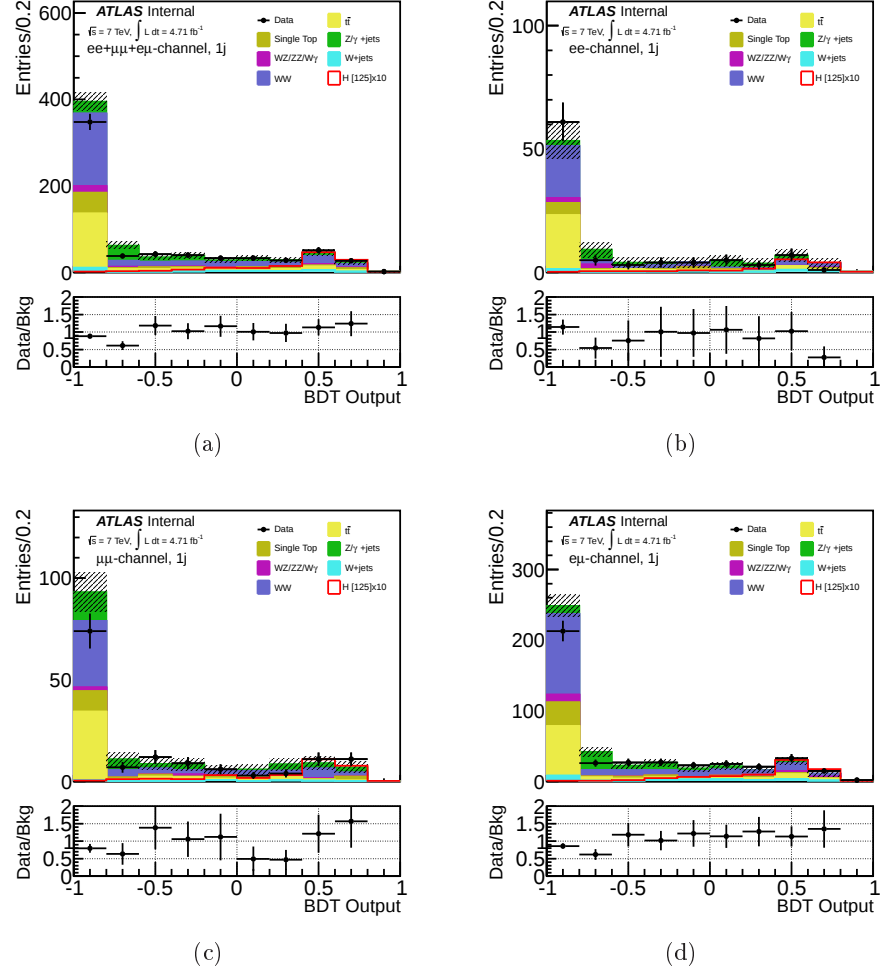


Figure D.2: Distributions of the BDT responses in the data and in the MC simulation, after all selection cuts, in the 1-jet channel, obtained from a two-variable BDT training. Shown are the (a) $\ell\ell$, (b) ee , (c) $\mu\mu$ and (d) $e\mu$ final states, respectively. The signal contribution, shown for $m_H = 125$ GeV, has been multiplied by a factor ten for better visibility.

Appendix E

BDT with the WW background as signal

A cross-check has been performed substituting the Higgs boson signal with the non-resonant WW background for training the BDT in the 0-jet channel. This study is meant to investigate the quality of the background modelling by the BDT algorithm, and especially of the WW background. All selection criteria are the standard ones, apart from the cut on $p_{T,\ell\ell}$, which has been removed to increase the number of the Z/γ^* +jets events in the training. The Higgs boson signal is not present at all, while backgrounds used for training are Z/γ^* +jets, top- and single top-quark processes, WZ and ZZ . Fig. E.1 shows the BDT response in the individual lepton flavours and their combination. In all cases, the WW events appear to be correctly classified as signal-like, with a large separation from the rest of the background events. In the $e\mu$ channel, where less Z/γ^* +jets events are present, hence more fluctuations are possible, some background contaminates the rightmost BDT bins, but overall the data/MC agreement is good, suggesting that the BDT analysis is able to understand and properly reproduce the background modelling.

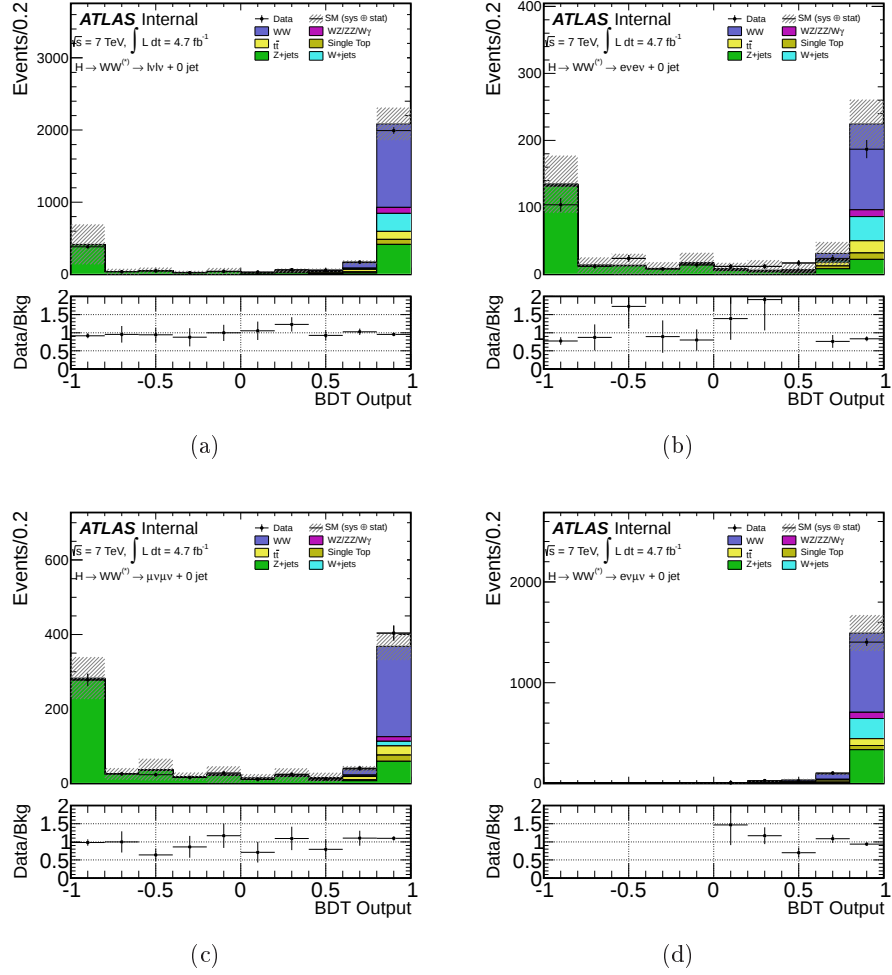


Figure E.1: Distributions of the BDT responses in the data and in the MC simulation, after all selection cuts, in the 0-jet channel, obtained training all backgrounds apart from WW versus WW (no Higgs boson signal is present). Shown are the (a) $\ell\ell$, (b) ee , (c) $\mu\mu$ and (d) $e\mu$ final states, respectively. The shaded areas indicate the total uncertainty on the background prediction.

Appendix F

Topology of signal-like and background-like events

The topology of the signal-like and the background-like events has been checked imposing a cut on the BDT response: $\text{BDT} > 0$ defines the signal-like region, while $\text{BDT} \leq 0$ defines the background-like region. The resulting shapes of the four topological variables, $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, $p_{T,\ell\ell}$, and m_T , are shown in Figs. F.1, F.2, F.3 and F.4, respectively, in the 0- and 1-jet bins, combining all lepton flavours together. As expected, the $\Delta\phi_{\ell\ell}$ distribution is shifted towards low (high) values in the signal-like (background-like) region, while the other variables have narrower distributions in the signal-like case than in the background-like case, as expected for the signal. It should be noticed that the Higgs boson signal (in the hypothesis $m_H = 125$ GeV) has to be magnified ten times in the background-like case in order to be visible, while there is no scale factor required in the signal-like region.

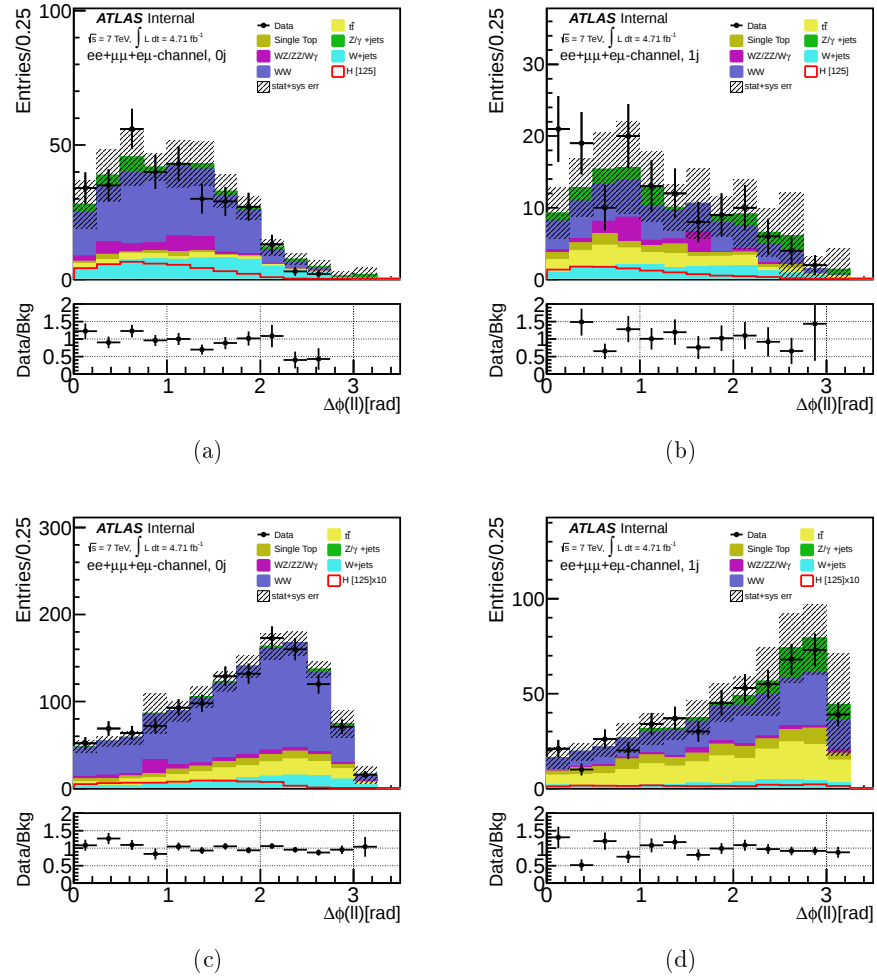


Figure F.1: Distributions of the azimuthal opening angle, $\Delta\phi_{\ell\ell}$, after all selection criteria, with an additional cut on the BDT response: (a,b) $\text{BDT} > 0$, (c,d) $\text{BDT} \leq 0$. The 0-jet channel is shown on the left, the 1-jet channel on the right; all lepton flavours have been combined. The shaded areas indicate the total uncertainty on the background prediction.

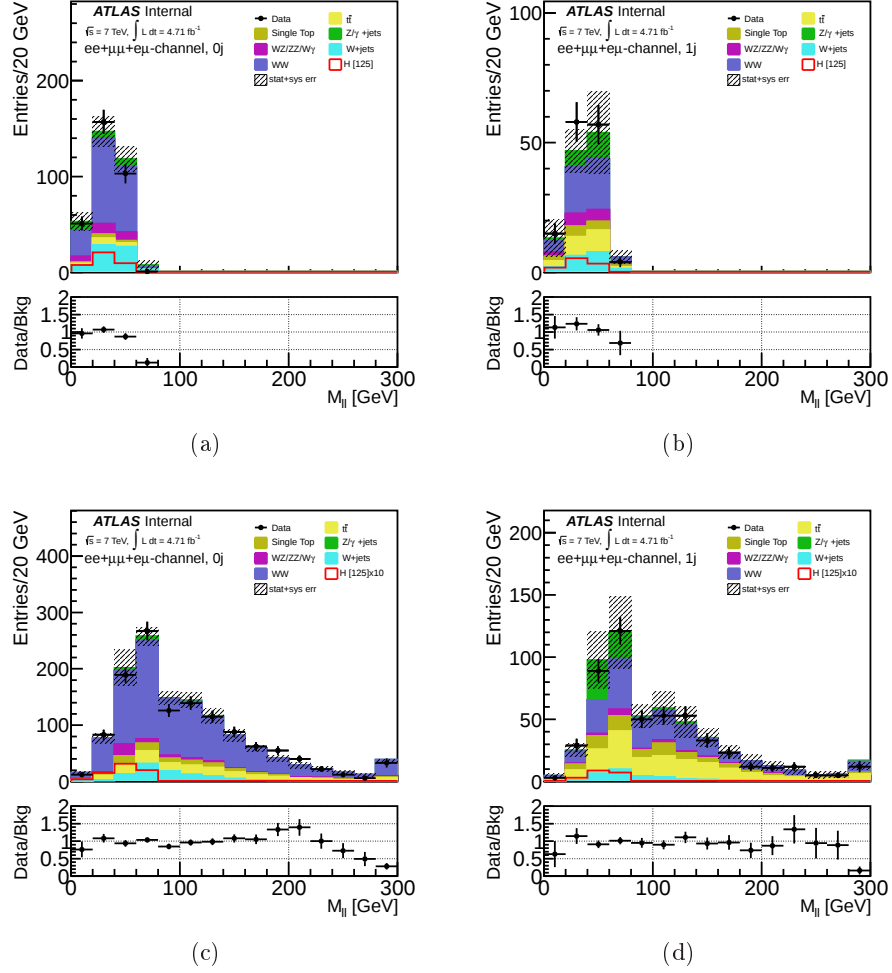


Figure F.2: Distributions of the dilepton invariant mass, $m_{\ell\ell}$, after all selection criteria, with an additional cut on the BDT response: (a,b) $\text{BDT} > 0$, (c,d) $\text{BDT} \leq 0$. The 0-jet channel is shown on the left, the 1-jet channel on the right; all lepton flavours have been combined. The shaded areas indicate the total uncertainty on the background prediction.

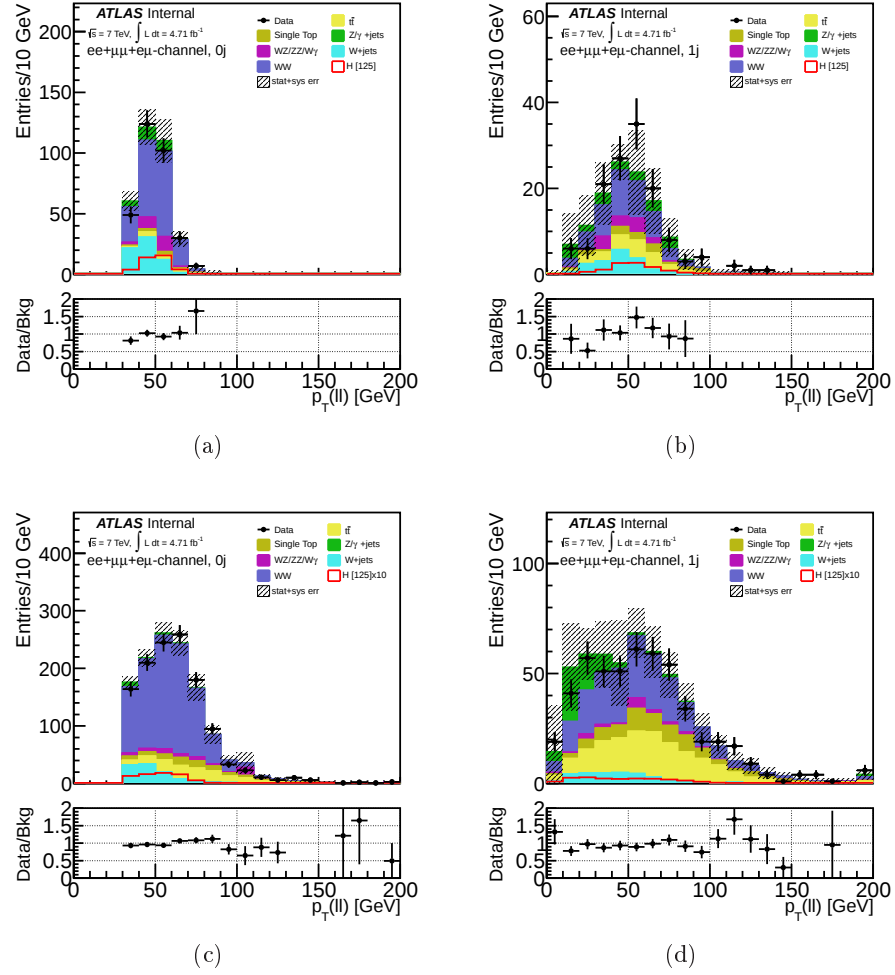


Figure F.3: Distributions of the dilepton transverse momentum, $p_{T,\ell\ell}$, after all selection criteria, with an additional cut on the BDT response: (a,b) BDT > 0, (c,d) BDT $\leq$ 0. The 0-jet channel is shown on the left, the 1-jet channel on the right; all lepton flavours have been combined. The shaded areas indicate the total uncertainty on the background prediction.

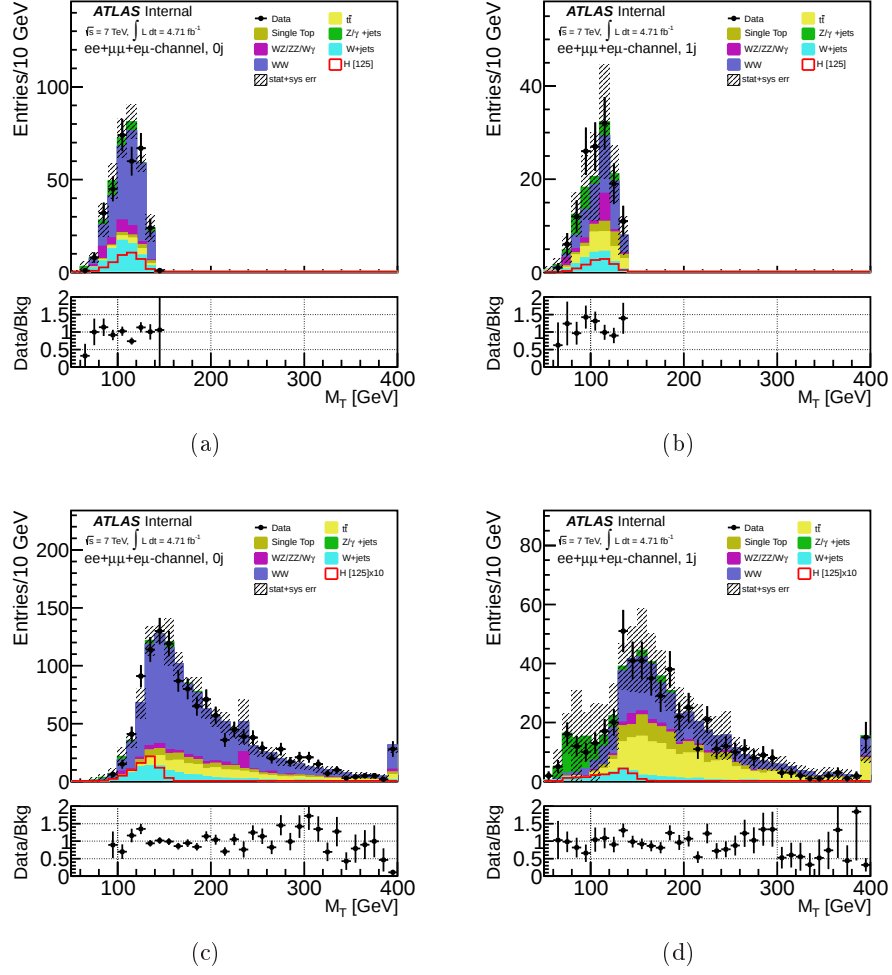


Figure F.4: Distributions of the transverse mass, m_T , after all selection criteria, with an additional cut on the BDT response: (a,b) $\text{BDT} > 0$, (c,d) $\text{BDT} \leq 0$. The 0-jet channel is shown on the left, the 1-jet channel on the right; all lepton flavours have been combined. The shaded areas indicate the total uncertainty on the background prediction.

Appendix G

BDT response for the cut-based candidates

Reversing the argument of Appendix F, it is interesting to look at the BDT response for the cut-based candidates, i.e. the events that fulfill all the standard selection criteria in the 0- and 1-jet channels, i.e. $m_{\ell\ell} < 50$ GeV and $\Delta\phi_{\ell\ell} < 1.8$ (cf. Table 4.4), with an additional cut on the transverse mass like in the cut-and-count analysis, $0.75 m_H < m_T < m_H$ with $m_H = 125$ GeV. After these requirements, all but three events have $\text{BDT} > 0$. Their distribution is shown in Fig. G.1, in the 0- and 1-jet bins. As expected, the cut-based candidates are classified by the BDT as signal-like, clustering at high BDT values.

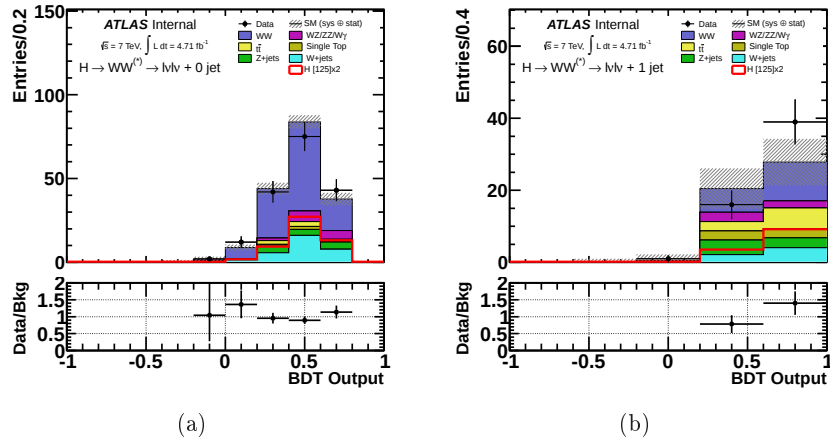


Figure G.1: Distributions of the BDT response for the cut-based candidates, selected applying the full selection of the standard cut-based analysis. The distributions are shown for (a) the 0-jet bin and (b) the 1-jet bin; all lepton flavours have been combined. The shaded areas indicate the total uncertainty on the background prediction.

Appendix H

Additional details on limits

In the following, additional plots and tables are shown, detailing the results presented in Section 4.11:

- Fig. H.1: limits broken down into lepton flavours and jet categories, in the full mass range ($110 < m_H < 600$ GeV),
- Fig. H.2: limits broken down into lepton flavours and jet categories, in the low mass range ($110 < m_H < 150$ GeV),
- Fig. H.3: combined limits for the BK and the LM data-taking periods, in the full, intermediate and low mass ranges,
- Fig. H.4: comparison between BDT and cut-based limits, broken down into lepton flavours and jet categories,
- Fig. H.5: local p_0 -values as a function of m_H in the full mass range, broken down into lepton flavours and jet categories,
- Table H.1: actual values for the comparison between BDT and cut-based limits, in the 0-jet channel,
- Table H.2: actual values for the comparison between BDT and cut-based limits, in the 1-jet channel.

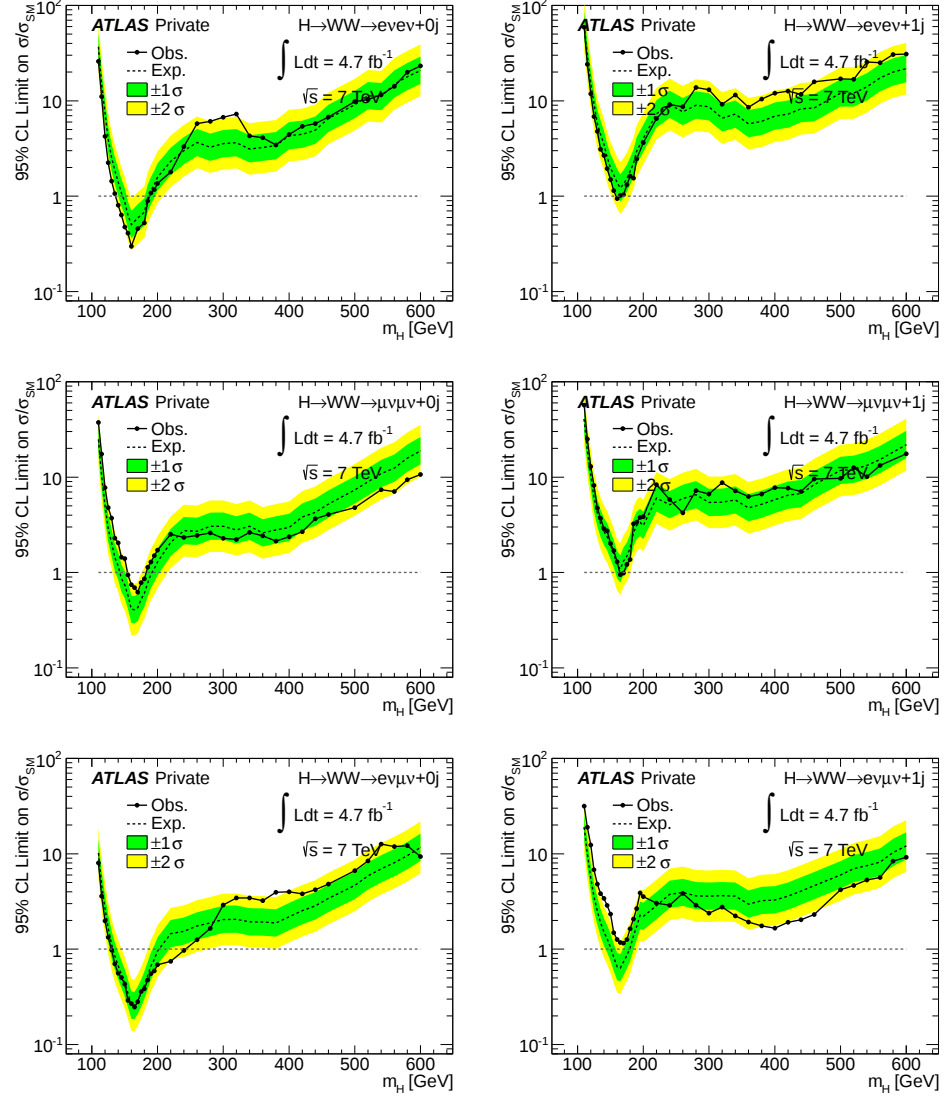


Figure H.1: Expected (dashed) and observed (solid) 95% CL upper limits on the Higgs boson production cross section, normalized to the SM cross section, in the full mass range ($110 < m_H < 600 \text{ GeV}$). The left and right columns contain the limits for the 0- and 1-jet channels, respectively; starting from above, the first line shows the ee channel, the second line the $\mu\mu$ channel, the third line the $e\mu$ channel.

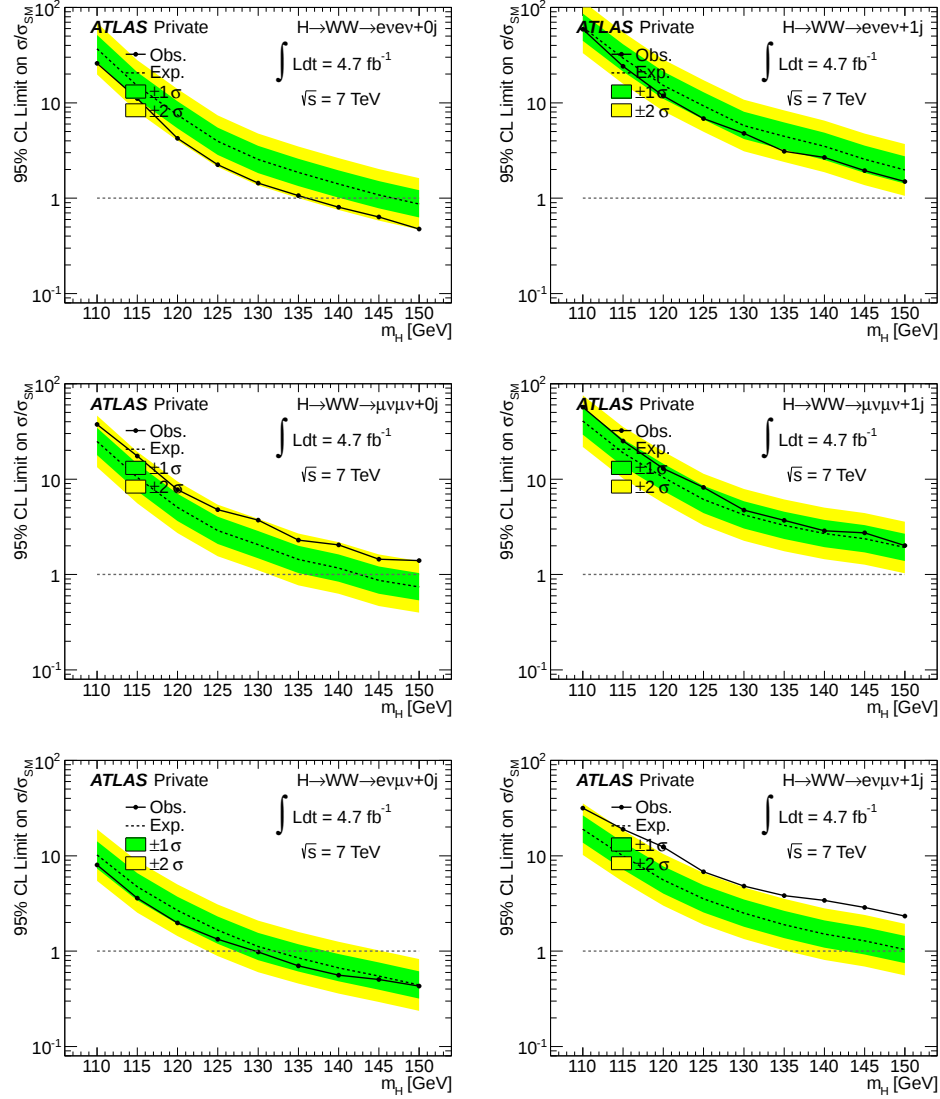


Figure H.2: Expected (dashed) and observed (solid) 95% CL upper limits on the Higgs boson production cross section, normalized to the SM cross section, in the low-mass region ($110 < m_H < 150$ GeV). The left and right columns contain the limits for the 0- and 1-jet channels, respectively; starting from above, the first line shows the ee channel, the second line the $\mu\mu$ channel, the third line the $e\mu$ channel.

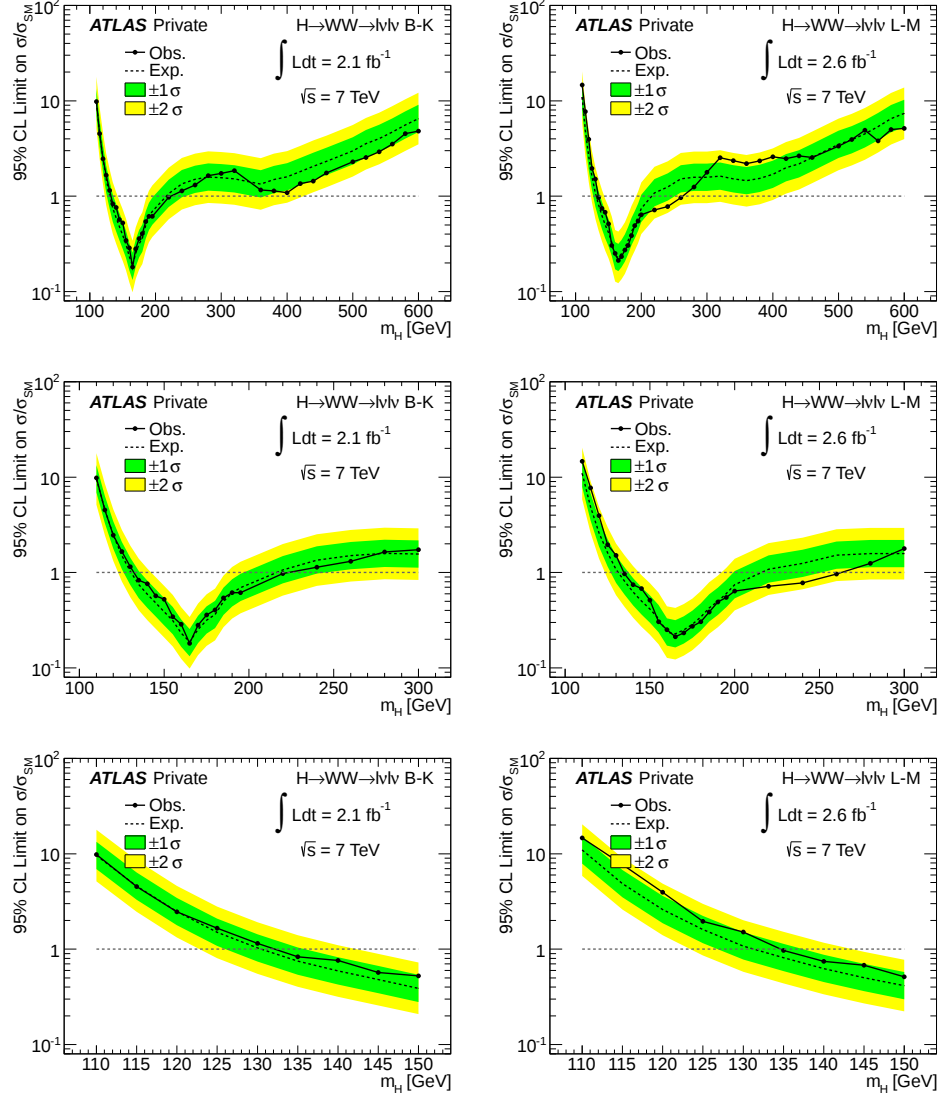


Figure H.3: Expected (dashed) and observed (solid) 95% CL upper limits on the Higgs boson production cross section, normalized to the SM cross section, in the low-mass region ($110 < m_H < 150$ GeV). All results are combined over lepton channels and jet bins. The left and right columns contain the limits for the BK and LM data-taking periods, respectively; starting from above, the first line shows the full mass range ($110 < m_H < 600$ GeV), the second line is zoomed in the range $110 < m_H < 300$ GeV, the third line is further zoomed in the range $110 < m_H < 150$ GeV.

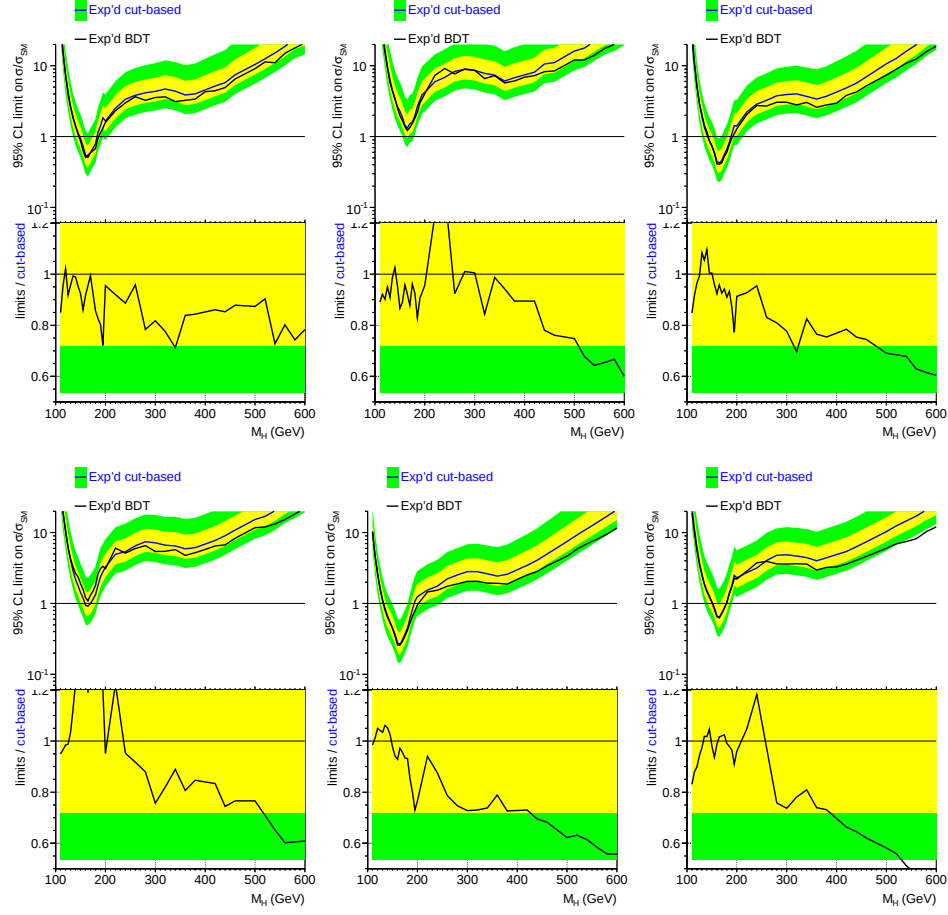


Figure H.4: Comparison of expected limits between the BDT and the cut-based analyses, broken down into lepton flavours and jet bins. First line, from left to right: ee 0-jet, ee 1-jet, $\mu\mu$ 0-jet. Second line, from left to right: $\mu\mu$ 1-jet, $e\mu$ 0-jet, $e\mu$ 1-jet. The bottom pads contain the ratio of the two corresponding curves. The yellow and green bands refer to the 1σ and 2σ envelopes of the cut-based limit.

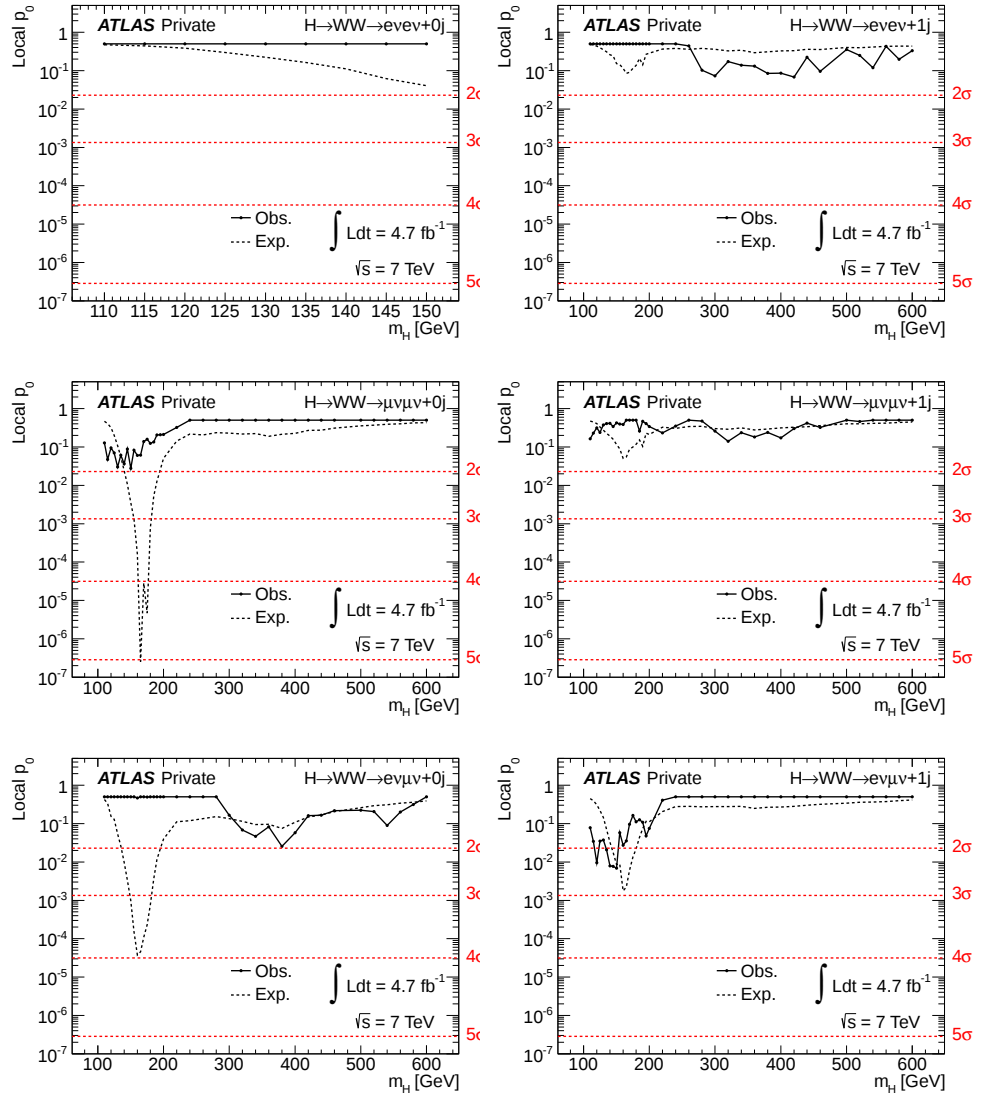


Figure H.5: Local p_0 -values as a function of m_H , in the full mass range. The observed probability for the background-only scenario is shown as a solid line, while the dashed line corresponds to the expectation for the signal+background hypothesis at the given value of m_H . The left and right columns contain the results for the 0- and 1-jet channels, respectively; starting from above, the first line shows the ee channel, the second line the $\mu\mu$ channel, the third line the $e\mu$ channel.

m_H (GeV)	BDT		Cut-based m_T shape fit		Ratio of expected limits (BDT/cut-based)
	Observed	Expected	Observed	Expected	
110	8.04	9.42	8.93	9.36	1.01
115	3.93	4.14	3.81	4.14	1.00
120	1.83	2.23	2.05	2.19	1.02
125	1.24	1.36	1.29	1.33	1.02
130	1.00	0.95	0.9	0.89	1.06
135	0.66	0.71	0.71	0.66	1.08
140	0.56	0.55	0.58	0.52	1.06
145	0.45	0.44	0.48	0.43	1.03
150	0.4	0.35	0.41	0.36	0.98
155	0.25	0.28	0.32	0.29	0.96
160	0.21	0.20	0.23	0.22	0.92
165	-	-	0.2	0.21	-
170	0.22	0.22	0.23	0.23	0.94
175	0.32	0.29	0.27	0.28	1.05
180	0.30	0.31	0.32	0.34	0.94
185	0.42	0.40	0.43	0.45	0.88
190	0.49	0.49	0.57	0.6	0.83
195	0.48	0.59	0.72	0.77	0.76
200	0.64	0.69	0.81	0.89	0.78
220	0.65	0.96	1.07	1.26	0.76
240	0.87	1.28	1.31	1.54	0.83
260	1.31	1.43	1.64	1.89	0.76
280	1.54	1.56	2.08	2.15	0.73
300	2.01	1.53	2.32	2.33	0.66
320	2.38	1.57	2.7	2.37	0.66
340	-	-	2.56	2.16	-
360	2.04	1.43	2.3	1.97	0.73
380	2.14	1.51	2.55	2.08	0.72
400	2.22	1.66	2.79	2.32	0.71
420	2.11	1.95	3.27	2.73	0.71
440	2.63	2.16	3.83	3.17	0.68
460	2.77	2.64	4.81	3.94	0.67
480	-	-	5.74	4.7	-
500	3.82	3.63	7.15	5.84	0.62
520	4.16	4.45	8.75	7.14	0.62
540	6.59	5.11	10.98	8.79	0.58
560	5.71	6.10	13.65	10.81	0.56
580	6.66	7.36	17.7	13.73	0.54
600	5.53	9.14	21.38	16.42	0.56

Table H.1: Comparison of observed and expected limits from the BDT analysis with those from the cut-based m_T fit analysis in the 0-jet bin. The last column contains the ratio. Some of the numbers are missing due to fit failures.

m_H (GeV)	BDT		Cut-based m_T shape fit		Ratio of expected limits (BDT/cut-based)
	Observed	Expected	Observed	Expected	
110	31.43	16.48	37.68	20.27	0.81
115	15.56	8.15	19.26	9.89	0.82
120	9.88	4.58	11.14	5.41	0.85
125	5.33	2.88	6.69	3.23	0.89
130	3.54	2.01	4.51	2.17	0.92
135	2.67	1.64	3.24	1.56	1.05
140	2.40	1.34	2.35	1.21	1.11
145	2.01	1.11	1.82	0.99	1.12
150	1.53	0.80	1.4	0.83	0.97
155	0.95	0.64	1.07	0.68	0.94
160	0.83	0.49	0.73	0.52	0.93
165	0.75	0.46	0.62	0.49	0.93
170	0.67	0.52	0.65	0.56	0.92
175	0.79	0.61	0.73	0.64	0.95
180	1.01	0.74	0.82	0.78	0.95
185	1.38	0.98	1.12	1.06	0.92
190	1.83	1.25	1.39	1.35	0.92
195	2.19	1.47	1.57	1.73	0.85
200	2.72	1.56	2.66	2.04	0.76
220	2.92	2.33	3.24	2.72	0.86
240	2.49	2.70	3.47	3.21	0.84
260	2.48	2.80	4.11	4.12	0.68
280	2.91	2.94	4.02	4.82	0.61
300	2.64	2.72	3.6	4.65	0.59
320	3.53	2.76	3.14	4.02	0.69
340	2.49	2.82	2.96	3.78	0.74
360	2.16	2.30	2.67	3.29	0.70
380	2.25	2.55	2.83	3.58	0.71
400	2.26	2.57	3.07	3.96	0.65
420	2.38	2.85	3.52	4.4	0.65
440	2.13	3.11	4.03	5.19	0.60
460	2.85	3.54	4.96	6.03	0.59
480	-	-	5.72	7.14	0
500	3.71	4.65	6.64	8.51	0.55
520	4.67	5.18	8.09	9.78	0.53
540	4.89	5.87	9.49	11.68	0.50
560	4.78	6.45	11.25	14.05	0.46
580	6.15	7.92	13.84	16.64	0.48
600	7.90	9.34	16.28	19.78	0.47

Table H.2: Comparison of observed and expected limits from the BDT analysis with those from the cut-based m_T fit analysis in the 1-jet bin. The last column contains the ratio. Some of the numbers are missing due to fit failures.

Appendix I

Post-fit nuisance parameters

The best fit values for the nuisance parameters in the combined $\sqrt{s} = 7$ TeV fit are shown in Table I.1, as well as their relative uncertainties.

The first four parameters are the α factors for control regions normalizations, with a pre-fit value of 1 ± 0 , while the others represent the systematic variation with respect to the nominal and their pre-fit value is 0 ± 1 . In the latter case, a post-fit error much smaller than 1 indicates an *overconstraint*, i.e. the fit is constraining that systematic more than the initial prescription. This situation is undesirable and could lead to fakely large sensitivities; it happens only for a few parameters in this analysis, such as E_SCALE and PILEUP.

The parameters appearing in Table 4.10 but not in Table I.1 have been discarded automatically by the fit because their variation with respect to nominal is under threshold, as detailed in Section 4.10.3.

Parameter	Best Fit	Error Up	Error Down
0 norm Top1j	1.03	0.0644	-0.0644
1 norm Top2j	1.02	0.0272	-0.0272
2 norm WW0j	1.01	0.0397	-0.0397
3 norm WW1j	1.19	0.218	-0.218
4 B EFF	0.697	0.636	-0.636
5 CLUSTER	0.31	0.48	-0.48
6 C EFF	-0.0115	0.966	-0.966
7 E EFF	0.663	0.736	-0.736
8 E RES	-0.253	0.388	-0.388
9 E SCALE	0.000113	0.00804	-0.00804
10 ID RES	-0.112	0.487	-0.487
11 JER	0.00101	0.0115	-0.0115
12 JES	-0.104	0.208	-0.208
13 L EFF	0.0216	0.896	-0.896
14 MS RES	-0.541	0.437	-0.437
15 M EFF	-0.0133	0.987	-0.987
16 M SCALE	-0.0012	0.0407	-0.0407
17 PILEUP	-5.96e-05	0.00642	-0.00642
18 TOP SCALEF 0j	0.576	0.73	-0.73
19 TOP SCALEF 0j hpt	0.165	0.944	-0.944
20 TRIGGER SF	-0.196	0.956	-0.956
21 Z SCALEF EE0j	-0.00511	0.978	-0.978
22 Z SCALEF EE1j	-0.226	0.93	-0.93
23 Z SCALEF MUMU0j	0.0798	0.943	-0.943
24 Z SCALEF MUMU1j	0.172	0.966	-0.966
25 top TH1j	0.215	0.891	-0.891
26 top TH2j	-0.0172	0.992	-0.992
27 ww model	0.256	0.796	-0.796
28 Fake Rate	-0.574	0.442	-0.442
29 LUMI	0.209	0.837	-0.837
30 QCDscale V	-0.0105	0.988	-0.988
31 QCDscale VV	0.0937	0.982	-0.982
32 QCDscale VV2in	-0.0823	0.975	-0.975
33 QCDscale VV ACCEPT	0.226	0.857	-0.857
34 QCDscale ggH	-0.0459	0.255	-0.255
35 QCDscale ggH1in	0.127	1.03	-1.03
36 QCDscale ggH2in	-0.0825	1.02	-1.02
37 QCDscale ggH ACCEPT	0.00063	0.989	-0.989
38 QCDscale qqH	-0.000225	0.993	-0.993
39 QCDscale qqH ACCEPT	-0.00294	0.993	-0.993
40 QCDscale ttbar	0.0665	0.178	-0.178
41 UEPS	0.0439	0.875	-0.875
42 ZJETS 2j	-0.159	0.938	-0.938
43 pdf gg	0.29	0.859	-0.859
44 pdf qqbar	0.0311	0.928	-0.928
45 pdf qqbar ACCEPT	0.187	0.865	-0.865

Table I.1: Best fit values and relative uncertainties for the nuisance parameters in the combined fit, for the 2011 $\sqrt{s} = 7$ TeV analysis.

Appendix J

Choice of the training configuration for the spin analysis

As introduced in Section 5.4, the optimization of the BDT configuration starts from the observation that using only one of the signals for training, namely $J^P = 0^+$, with spin-dependent variables as inputs, could introduce biases in the algorithm response. Hence, several alternative training configurations have been explored, which:

- use spin-independent variables for training $J^P = 0^+$ against the total background,
- cut on the BDT response,
- fit one of the spin-dependent variables (e.g. $\Delta\phi_{\ell\ell}$).

An example is shown in Fig. J.1, where the performance of the standard four topological variables is compared with those of alternative sets. The comparison uses ROC curves, which measure background rejection as a function of signal efficiency (see also Appendix A). The variables considered are $p_{T,\ell_1}, p_{T,\ell_2}, \Delta\phi(E_T^{\text{miss}}, \ell_1), \Delta y_{\ell\ell}$ and m_z ⁽¹⁾, where the latter two are invariant for boosts in the z direction (thus they are supposed to preserve the spin discrimination as in the Higgs boson rest frame). Out of the four configurations considered here, the best signal efficiency, for a fixed background rejection, is achieved with the four topological variables; the configurations which use only spin-independent variables, labelled as “BDT0” and “BDT2” in the figure, have a worse performance but this can be possibly improved after a cut on the BDT response.

Fig. J.2 (b) shows the BDT response after preselection, in 5.8 fb^{-1} of integrated luminosity. The training uses in this case the following variables: $\Delta y_{\ell\ell}, m_z, m_{\ell\ell}$,

¹The longitudinal dilepton mass, m_Z , is defined as:

$$m_z = \sqrt{(\vec{p}_{\ell_1} + \vec{p}_{\ell_2})^2 - (p_{z,\ell_1} + p_{z,\ell_2})^2} = \sqrt{m_{\ell\ell}^2 - (p_{x,\ell_1} + p_{x,\ell_2})^2 - (p_{y,\ell_1} + p_{y,\ell_2})^2}.$$

m_T and $p_{T,\ell_1} - p_{T,\ell_2}$. The aim is to fit $\Delta\phi_{\ell\ell}$, depicted in Fig. J.2 (a), after a cut on the BDT response. As shown in the figure, there is a sizeable amount of $Z/\gamma^* + \text{jets}$ events peaking at high $\Delta\phi_{\ell\ell}$ values, and hence at low BDT values, which can be removed requiring $\text{BDT} > -0.6$. However, such a cut greatly reduces the shape difference in the azimuthal angle to be fitted, as can be seen comparing the plots in Fig. J.2 (c) and (d), which show the $\Delta\phi_{\ell\ell}$ shapes for the $J^P = 0^+$ and $J^P = 2^+$ signals before and after the cut on BDT, respectively. Consequently, the expected sensitivities are significantly worse for this configuration (and for all those featuring spin-independent variables) than for the standard four-variables one.

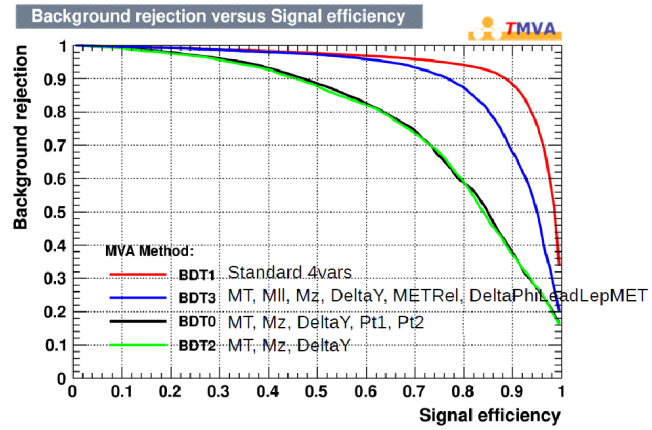


Figure J.1: Background rejection versus signal efficiency when training BDT ($J^P = 0^+$ against the total background) with different sets of input variables.

One additional approach explored proceeds on a different path:

1. firstly, reduce the background applying tight cuts or training a dedicated BDT and cutting on it,
2. train a second BDT to separate $J^P = 0^+$ from $J^P = 2^+$, using spin-dependent variables,
3. fit the latter BDT response.

This approach suffers from various limitations: after reducing the background with a cut on the BDT response, only a few hundreds of raw MC events are left for the signal samples, which makes a new training not straightforward, and besides, the shape of spin-sensitive variables has been distorted, so that again separation is reduced.

At an advanced stage of the analysis, it has been found out that an alternative configuration does exist, which seems promising in terms of expected sensitivities.

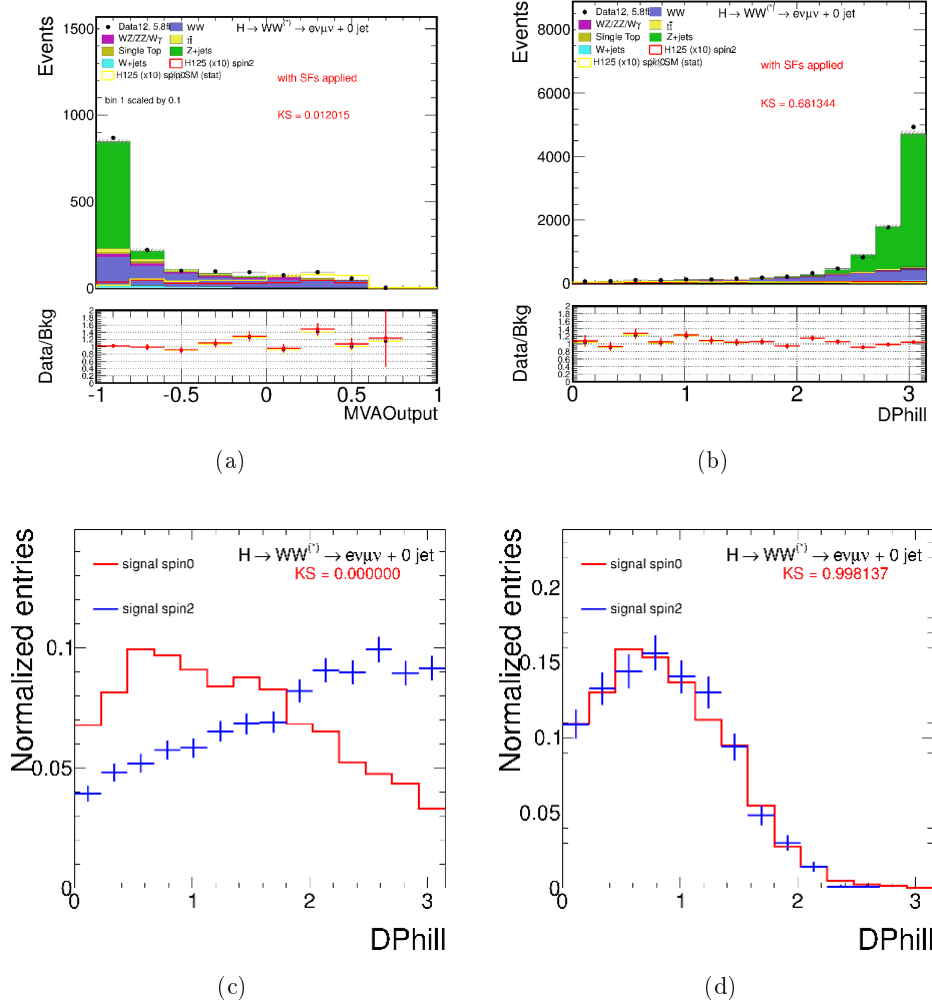


Figure J.2: (a) Distribution of the BDT response after the preselection, obtained training the $J^P = 0^+$ signal against the total background, using the following variables: $\Delta y_{\ell\ell}$, m_z , $m_{\ell\ell}$, m_T and $p_{T,\ell_1} - p_{T,\ell_2}$. The $\Delta\phi_{\ell\ell}$ variable, shown in (b), has been removed from the inputs so that it can be fitted after a cut the BDT response itself. The dataset used for this study corresponds to an integrated luminosity of 5.8 fb^{-1} . (c,d) Comparison of the $\Delta\phi_{\ell\ell}$ shape between the $J^P = 0^+$ and $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ signal samples, before (c) and after (d) a cut on the BDT response, $\text{BDT} > -0.6$.

It requires to train $J^P = 0^+$ against the sum of the $J^P = 2^+$ signal and the total background, and $J^P = 2^+$ against the sum of the $J^P = 0^+$ signal and the total background, resulting in signal shapes which are more different from each other, as can be seen in Fig. J.3. Preliminary results (based on an integrated luminosity of 13 fb^{-1} , and a treatment of systematic uncertainties which is not the final one) show that improvements in expected sensitivities up to 25% are in fact achievable. However, due to lack of time, this alternative has not been pursued further.

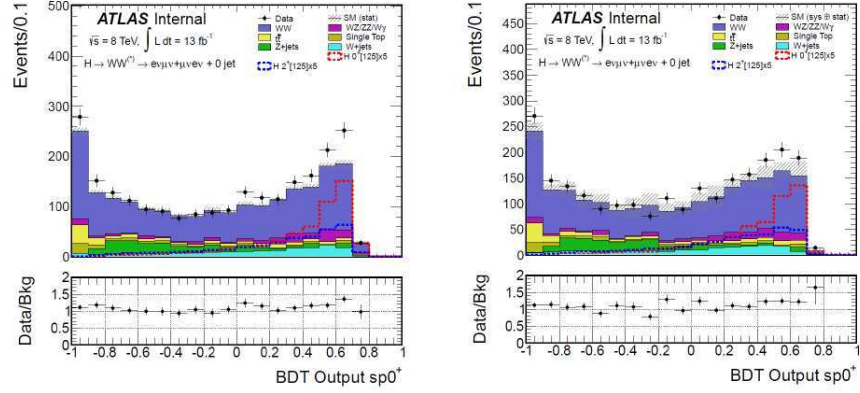


Figure J.3: Comparison between BDT0 responses: standard ($J^P = 0^+$ against background) on the left, improved ($J^P = 0^+$ against background plus $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ on the right). In both cases, the standard four variables are used. Only statistical uncertainties are shown. The dataset used for this study corresponds to an integrated luminosity of 13 fb^{-1} .

Appendix K

The theoretical systematic uncertainty on the shape of the WW background

This Appendix provides additional details on the shape uncertainty of the WW background, following what discussed in Section 5.6.4. Three effects are considered: PDFs, QCD scales and underlying event/parton shower model, corresponding to three independent nuisance parameters.

For each variation, new samples of WW events have been produced at generator-level:

- for uncertainties on PDFs, reweighted events for CTEQ (CT10) error sets, MSTW (2008) and NNPDF (2.0) central sets,
- for QCD scales, events with $(\mu_R, \mu_F = m_H/2, 2m_H)$,
- for UE/PS, **POWHEG+HERWIG** events to be compared with the nominal using **PYTHIA6**.

After smearing the neutrino momenta, in order to simulate the E_T^{miss} resolution, the four topological variables have been computed and used for the standard training, obtaining a BDT response for each one of the up and down variations, which envelopes the nominal one. An example of the varied BDT responses is shown in Fig. K.1 (a,b) for the QCD scale variations, which also compares the distributions of fully simulated MC samples with those of generator-level ones. From the figure, a few observations can be made: first, events generated with **POWHEG+PYTHIA6** or with **MC@NLO+HERWIG** give rise to very similar BDT responses (dashed green and red lines), thus the relative difference found between **MC@NLO** samples with varying scales

can be applied to the nominal POWHEG events. Moreover, the difference between fully- and non fully-simulated samples corresponds to the difference between the dashed green line and the solid black line ($\mu_R, \mu_F = (m_H, m_H)$), respectively; it is visible only in the leftmost bin, not in the signal-like region. The plots on the bottom line of Fig. K.1 show the ratio of the scale-varying configurations with respect to the nominal: they float inside a few percent in all cases.

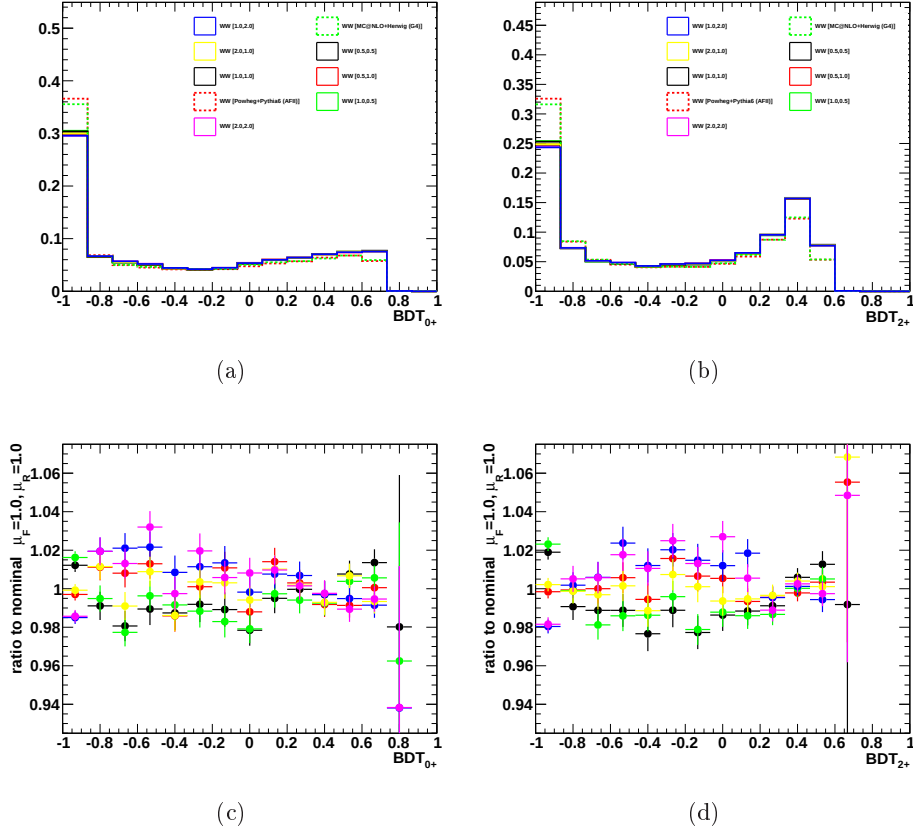


Figure K.1: (a,b) Distributions of the BDT response for different choices of μ_F and μ_R at generator-level, for the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ analysis. The dashed lines indicate fully simulated MC samples (POWHEG in green, MC@NLO in red), while the solid lines indicate the MC@NLO samples at generation-level. (c,d) Ratios of the configurations with varying scales with respect to the nominal one; the colour coding is the same as in the upper plots.

For what concerns the UE/PS variations, POWHEG+HERWIG has been compared to the nominal POWHEG+PYTHIA6, and the result is shown in Fig. K.2 for BDT0 and BDT2⁽¹⁾: the ratio lies inside a few percent, which is the case also for the $J^P = 1^\pm$

¹In the case of the $J^P = 2^+$ analysis, the computations are performed for $f_{q\bar{q}} = 0.25$ and applied to all the other fractions.

analysis. Finally, the PDF uncertainties are computed with the recommended master formula (Eq. 1.53), which sums the largest positive and negative deviations from the nominal. As shown in Fig. K.3 for BDT0, differences are of the order of a few percent for all the PDF prescriptions considered: in particular the HERA and NNPDF variations are very similar to each other, and the MSTW one is negligible, so the total PDF uncertainty is computed as the sum in quadrature of the CTEQ and NNPDF ones, the former being a flat $\pm 1\%$.

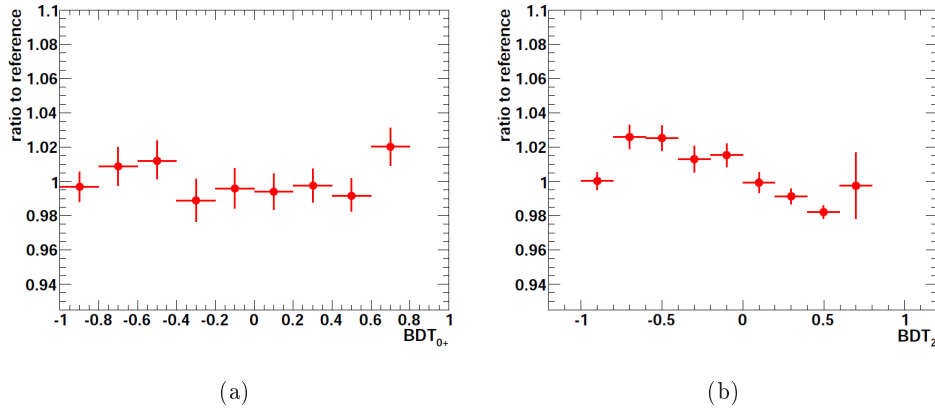


Figure K.2: Ratios of POWHEG+HERWIG WW events with respect to the nominal POWHEG+PYTHIA6 configuration, for the (a) BDT0 and (b) BDT2 ($f_{q\bar{q}} = 0.25$) distributions.

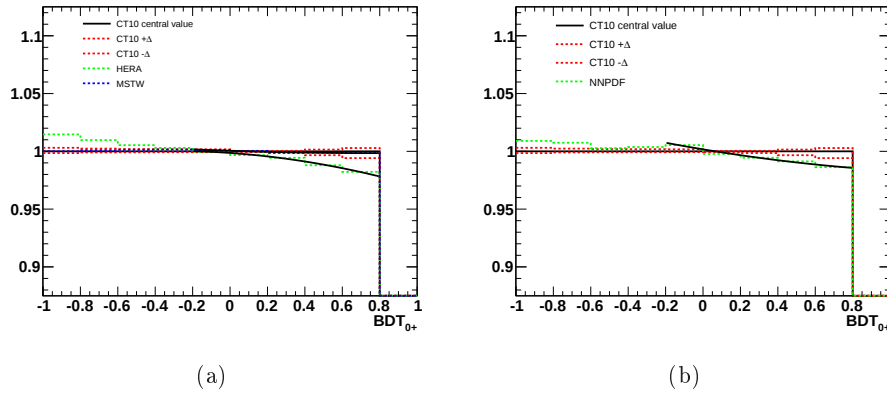


Figure K.3: Variations induced by different PDF prescriptions on the BDT0 distribution. The PDFs considered here are the CT10 error set and the central values of MSTW2008, HERA6.5 and NNPDF2.0 PDFs. The curved black lines represent the fit to the largest deviations ((a) HERA, (b) NNPDF).

For fitting purposes, the envelope induced by each one of the above mentioned

variations is computed in 1D, after a rotation of BDT0 and BDT2 (or BDT1) along the direction of maximal correlation. Thus, two uncorrelated directions are defined, BDT_x and BDT_y (cf. Eq. 5.1), using the following values for the correlation angles:

$$\theta_{\text{corr}} = 0.740 \ (J^P = 2^+) , 0.735 \ (J^P = 1^+) , 0.708 \ (J^P = 1^-) .$$

An example of the 2D correlations before and after the rotation is shown for the $J^P = 1^+$ analysis in Fig. K.4 (a) and (b), respectively. The correlation between the two directions decreases from 0.9 to 0.1, with the latter being concentrated in the x direction.

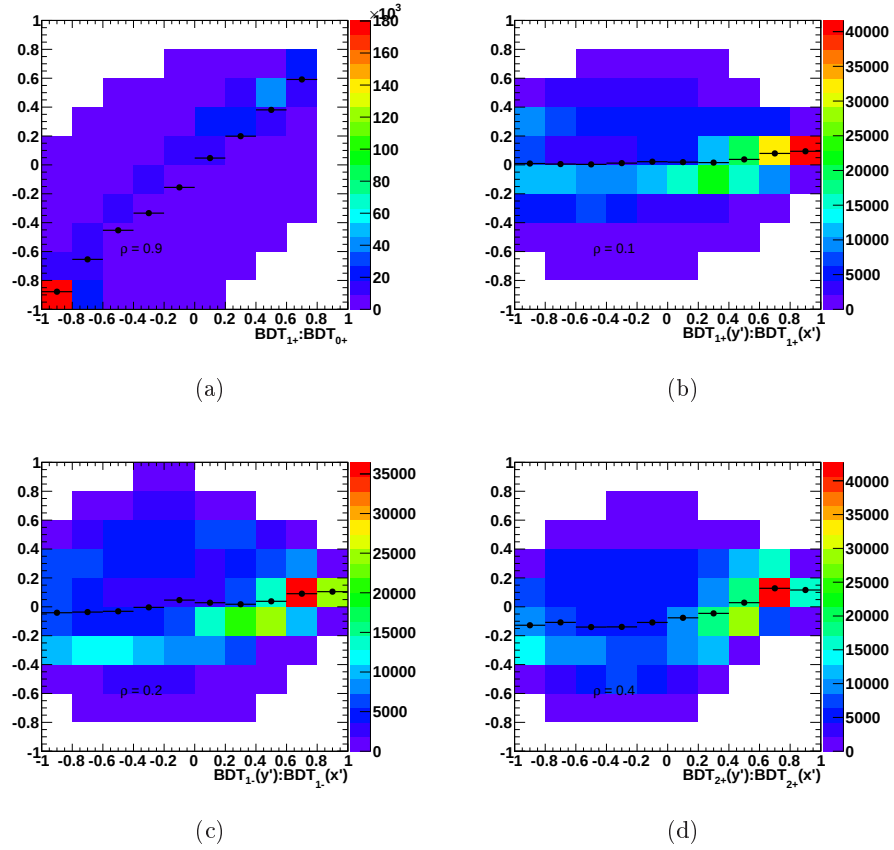


Figure K.4: 2-dimensional BDT distributions, with a label indicating the corresponding correlation coefficient. (a) BDT1 as a function of BDT0 for the $J^P = 1^+$ analysis before the rotation, (b) the same but after the rotation in Eq. (5.1). Bottom line: distributions after the rotation, for the (c) $J^P = 1^-$ and (d) $J^P = 2^+$ analyses.

The same rotation is applied to all systematic samples, and envelopes are constructed by means of polynomial functions, using the largest variation in each case. More in details, the following envelopes are used:

- PDFs: $\delta = \sqrt{\delta_{\text{CTEQ}}^2 + \delta_{\text{NNPDF}}^2}$:

$$\begin{aligned}
- \delta_{\text{CTEQ}}^{(1^+, 1^-, 2)} &= 0.01 * \text{BDT}_y, \\
- \delta_{\text{NNPDF}}^{(2)} &= (-0.0227) * \text{BDT}_y + (0.0149) * \text{BDT}_y^2, \\
- \delta_{\text{NNPDF}}^{(1^+)} &= (-0.0157) * \text{BDT}_y + (-0.0117) * \text{BDT}_y^2, \\
- \delta_{\text{NNPDF}}^{(1^-)} &= (-0.0106) * \text{BDT}_y + (-0.0238) * \text{BDT}_y^2,
\end{aligned}$$

- UE/PS:

$$\begin{aligned}
- \delta_{\text{UE/PS}}^{(2)} &= |(0.019) * \text{BDT}_x + (-0.03) * \text{BDT}_y|, \\
- \delta_{\text{UE/PS}}^{(1^+)} &= (0.0145) * \text{BDT}_x + (0.009) * \text{BDT}_y, \\
- \delta_{\text{UE/PS}}^{(1^-)} &= |(0.05) * \text{BDT}_x + (-0.01) * \text{BDT}_y|,
\end{aligned}$$

- QCD scales:

$$\begin{aligned}
- \delta_{\text{scale}}^{(2)} &= |(-0.0025) + (-0.015) * \text{BDT}_x + (0.05) * \text{BDT}_x^2| + |(0.016) + \\
&\quad (-0.017) * \text{BDT}_y + (-0.035) * \text{BDT}_y^2 + (0.01) * \text{BDT}_y^3|, \\
- \delta_{\text{scale}}^{(1^+)} &= |(-0.00165) + (-0.00485) * \text{BDT}_x + (0.05) * \text{BDT}_x^2| + |(0.018) + \\
&\quad (-0.0135) * \text{BDT}_y + (-0.04) * \text{BDT}_y^2 + (0.004) * \text{BDT}_y^3|, \\
- \delta_{\text{scale}}^{(1^-)} &= |(-0.0018) + (0.05) * \text{BDT}_x + (-0.05) * \text{BDT}_x^2| + |(0.019) + (-0.025) * \\
&\quad \text{BDT}_y + (0.043) * \text{BDT}_y^2 + (0.018) * \text{BDT}_y^3|.
\end{aligned}$$

The variation δ is added to the event weight, as

$$\text{weight}_{\text{up, down}} = \text{weight}_{\text{nominal}} * (1 \pm \delta)$$

only if both BDT0 and BDT2 (or BDT1) are > -0.2 , i.e. only in the signal-like region. This is because, otherwise, the fit would constrain the WW shape more strongly in the background-like region and then force it to follow the same trend also in the signal-like region, possibly affecting the post-fit NPs and the final sensitivities. These envelopes are shown in Figs. K.5 and K.6 for the QCD scales and UE/PS uncertainties, respectively; the fit uses in each case the largest envelope.

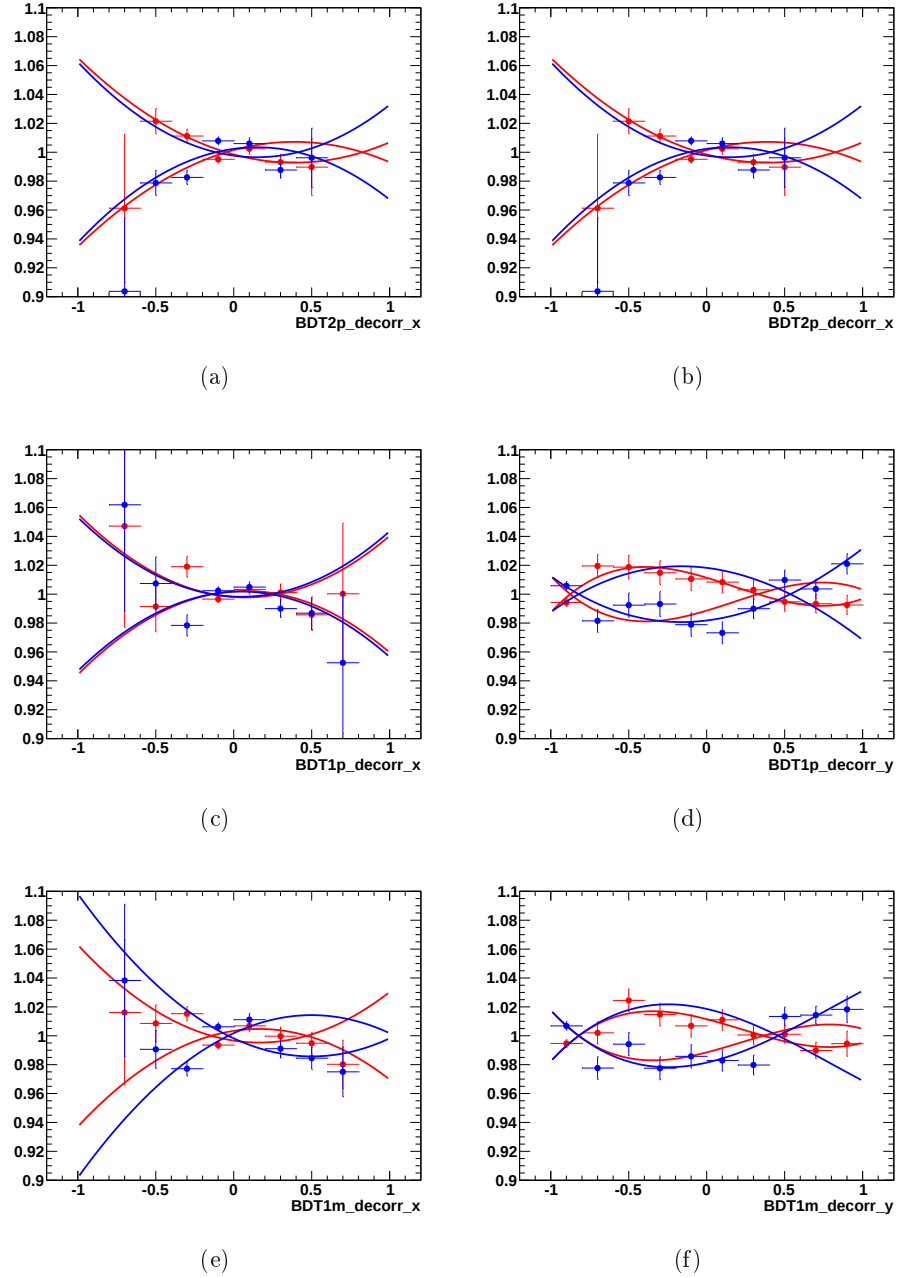


Figure K.5: Total up and down envelopes for the QCD scale uncertainties, for the (a,b) $J^P = 2^+$, (c,d) $J^P = 1^+$ and (e,f) $J^P = 1^-$ analyses. The fit uses in each case the largest envelope.

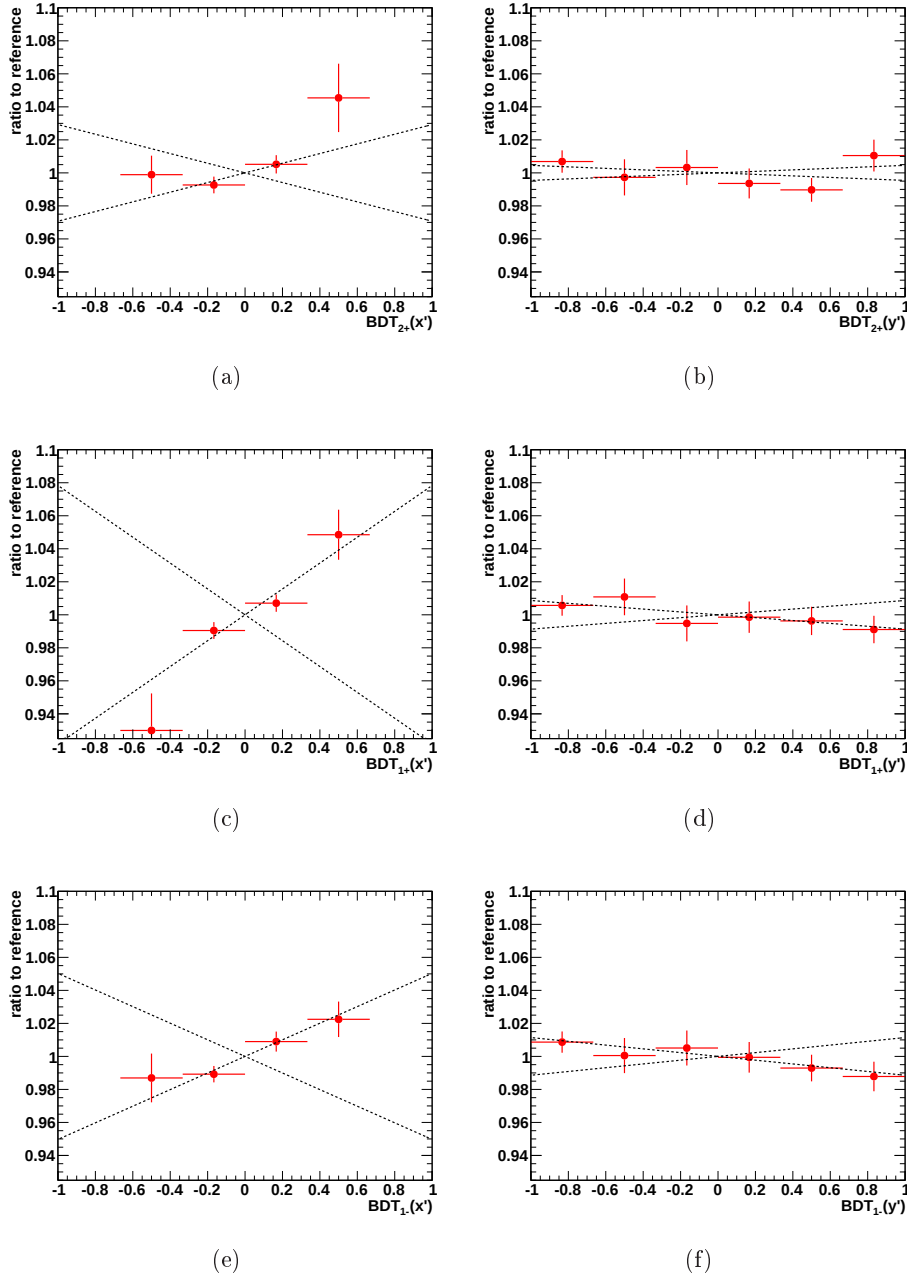


Figure K.6: Total up and down envelopes for the UE/PS uncertainties, for the (a,b) $J^P = 2^+$, (c,d) $J^P = 1^+$ and (e,f) $J^P = 1^-$ analyses. The fit uses in each case the largest envelope.

Appendix L

Additional studies on the shape of the WW background

Before the implementation of the baseline treatment for the UEPS uncertainty described in Section 5.6.4 and Appendix K, the spin analysis used a definition directly inherited from the 2011 analysis. This recipe consisted in letting the shape of the WW background free to float from its nominal configuration (the `POWHEG` one) to the `MC@NLO` one, using the net difference between the two generators as an envelope.

This simple approach, known to be conservative, causes overconstraints and pulls in the corresponding nuisance parameters (see Fig. L.1 for an example), a situation which is not desirable, especially given the fact that this systematic is, together with the Fake Rate uncertainty, the most important of the whole analysis. More specifically, Table L.1 shows the effect of turning off one systematic source at a time, on the fitted values of the signal strengths and ε , with the old UEPS definition. Referring in particular to the $J^P = 0^+$ signal strength, its best-fit value goes down from $1.45\sigma_{\text{SM}}$ to $1.35\sigma_{\text{SM}}$ when removing the UEPS uncertainty, and even lower if only the shape component is removed. This means that the fit substitutes the signal component with an additional WW one, when it is not given the freedom of adjusting the WW shape in the signal-like region. This also shows how crucial a correct treatment of this systematic is, in order to measure the presence of the signal in data and to distinguish among different spin hypotheses.

L.1 Compatibility of the `MC@NLO` and `POWHEG` predictions with the data

Apart from being conservative, the old recipe is also possibly wrong. In fact, in a previous study based on an integrated luminosity of 10.6fb^{-1} , it has been shown

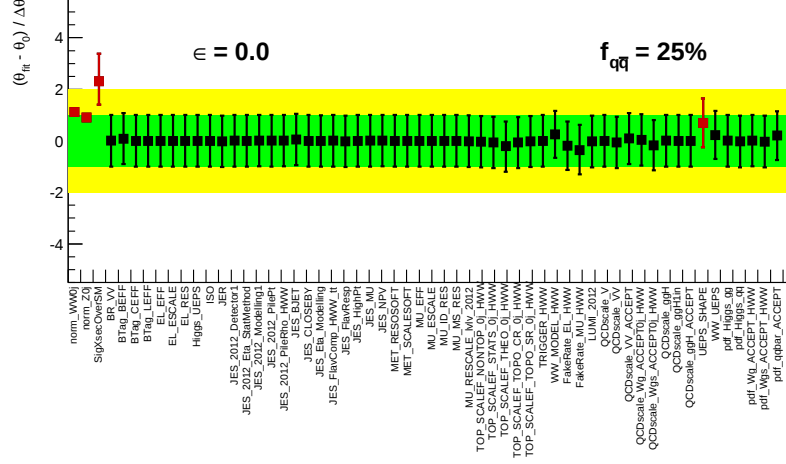


Figure L.1: Pull distribution for a preliminary version of the $J^P = 2^+$ $f_{q\bar{q}} = 0.25$ analysis (the fit assumes $\varepsilon = 0$), exemplifying the pathological behaviour of the UEPS shape nuisance parameter (marked in red) in case the correct treatment is not performed.

	$\hat{\mu}_0$	$\hat{\mu}_2$	$\hat{\varepsilon}$
All systematic uncertainties	$1.45^{+0.69}_{-0.58}$	$1.99^{+1.28}_{-1.17}$	$1-0.502$
All systematic uncertainties - Fake Rate	$1.41^{+0.68}_{-0.56}$	$1.59^{+1.14}_{-1.00}$	$1-0.438$
All systematic uncertainties - UEPS (shape + norm)	$1.35^{+0.66}_{-0.54}$	$1.97^{+1.20}_{-1.11}$	$1-0.541$
All systematic uncertainties - UEPS (shape)	$1.31^{+0.67}_{-0.55}$	$1.86^{+1.22}_{-1.14}$	$1-0.538$
All systematic uncertainties - UEPS (norm)	$1.47^{+0.71}_{-0.56}$	$2.11^{+1.25}_{-1.13}$	$1-0.506$
All systematic uncertainties - MC stats	$1.50^{+0.66}_{-0.54}$	$2.20^{+1.22}_{-1.08}$	$1-0.463$
All systematic uncertainties - JES	$1.44^{+0.68}_{-0.57}$	$2.01^{+1.25}_{-1.16}$	$1-0.515$

Table L.1: Best fit values for a preliminary version of the $J^P = 2^+$ analysis, showing the effect on the fitted signal strengths and ε of turning off one systematic source at a time. Only the most important effects are shown. “MC stats” refers to the uncertainty due to the limited number of MC events, “shape” and “norm” to the shape and normalization components of the UEPS uncertainty.

that the data rule out the MC@NLO configuration with a significance close to 3σ . This study combined information from more kinematical variables (m_T , $m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$) and used the tighter rate analysis selection. It is interesting to repeat this study with increased luminosity (13fb^{-1}) and with the looser spin cuts.

The model to be tested is indeed similar to the one used in this analysis, but replacing the spin hypotheses with the two different MC simulations for the WW background:

$$\varepsilon \times \text{POWHEG} + (1 - \varepsilon) \times \text{MC@NLO}. \quad (\text{L.1})$$

The analysis is performed in the WW control region (defined in Section 5.5.2); the fit is done on each of the input and output variables (m_T , $m_{\ell\ell}$, $p_{T,\ell\ell}$, $\Delta\phi_{\ell\ell}$, BDT0 and BDT2), without systematic uncertainties. Pseudo-experiments are generated like in the standard spin analysis, without any signal component; no control region is used in the fit, and the WW normalization is left free to float. Also the POI is left free to float, like in the standard LHC prescription (which is different from the one used for the spin analysis, cf. Section 5.7.1), in the range $-10 < \varepsilon < 10$, so that the boundaries are not hit.

Fig. L.2 shows the toy distributions obtained from the various fits, for the two hypotheses under consideration (POWHEG in red, MC@NLO in blue) and for real data (black line). The expected and observed sensitivities are reported on the plots themselves, and summarized in Table L.2 for observed, which also contains the fitted ε values. The data seem to prefer the POWHEG configuration in the m_T and $m_{\ell\ell}$ cases (with significances of 1.76 and 2.99 standard deviations, respectively, for MC@NLO rejection), but the MC@NLO configuration when fitting $p_{T,\ell\ell}$, $\Delta\phi_{\ell\ell}$ or BDT2 (with significances of 1.60, 1.55 and 1.81 standard deviations, respectively, for POWHEG rejection). The fit to BDT0 lies in between the two hypotheses, and so does the 2-dimensional fit to (BDT0, BDT2). The same conclusions are confirmed when applying the tighter rate analysis cuts for defining the WW control region (cf. Table 5.4), the salient difference being that in this case also BDT0 favours POWHEG.

In conclusion, since BDT0 and BDT2 tend to different directions, the 2-dimensional fit slightly prefers POWHEG, but with less than 1σ significance, so no one of the two generators can be discarded based on these studies.

	Rejection of POWHEG		Rejection of MC@NLO		$\hat{\varepsilon}$
	$1 - p_0$	Z	$1 - p_0$	Z	
m_T	0.25	-0.66	0.96	1.76	$1_{-0.02}^{-0.02}$
$m_{\ell\ell}$	0.15	-1.04	0.999	2.99	$1_{-0.75}^{+0.75}$
$\Delta\phi_{\ell\ell}$	0.94	1.55	0.36	-0.37	$-0.33_{-0.97}^{+0.85}$
$p_{T,\ell\ell}$	0.95	1.60	0.26	-0.66	$0_{-0.02}^{+0.02}$
BDT0	0.83	0.94	0.62	0.31	$0.31_{-0.80}^{+0.71}$
BDT2	0.96	1.81	0.36	-0.35	$-0.19_{-0.69}^{+0.67}$
(BDT0, BDT2)	0.66	0.41	0.71	0.55	$0.52_{-0.48}^{+0.48}$

Table L.2: Sensitivities for rejecting the POWHEG and MC@NLO hypotheses, with the spin analysis selection, in the WW control region. Only statistical uncertainties are used, and the POI is fitted in the range $-10 < \varepsilon < 10$ (apart from the last line, where it is $0 \leq \varepsilon \leq 1$).

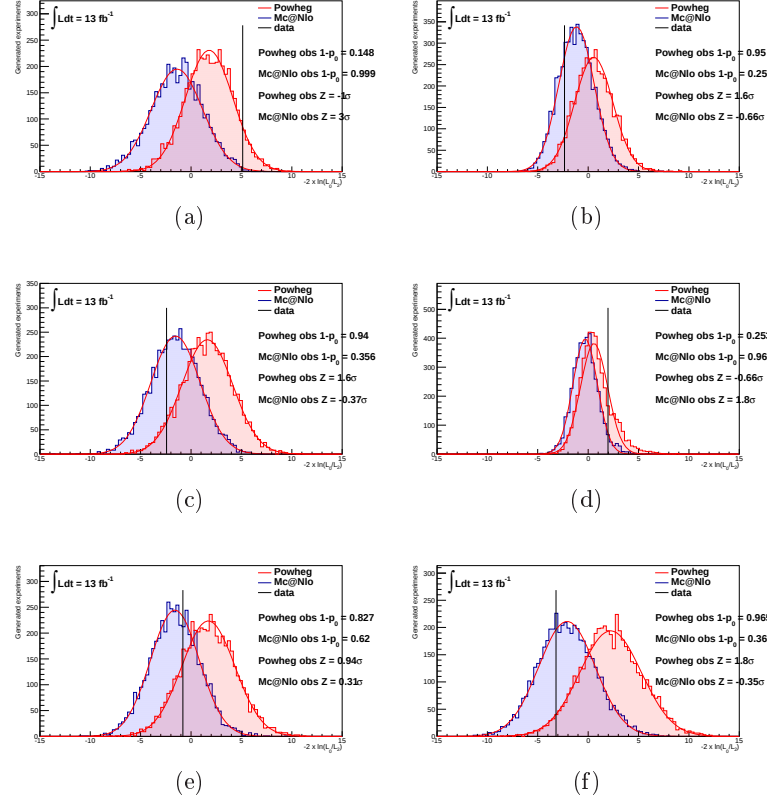


Figure L.2: Toy distributions for testing the POWHEG hypothesis (red) versus the MC@NLO one (blue), with 13fb^{-1} of integrated luminosity. The upper row shows toys obtained fitting (a) $m_{\ell\ell}$ and (b) $p_{T,\ell\ell}$, the middle row shows (c) $\Delta\phi_{\ell\ell}$ and (d) m_T , the lower row shows (e) BDT0 and (f) BDT2 ($f_{q\bar{q}} = 0.25$). See the text for further details.

	Rejection of POWHEG		Rejection of MC@NLO		$\hat{\epsilon}$
	$1 - p_0$	Z	$1 - p_0$	Z	
m_T	0.21	-0.82	0.98	2.05	$1_{-0.64}$
m_{ll}	0.16	-1.00	1.00	2.70	$1_{-0.22}$
$\Delta\phi_{ll}$	0.65	0.39	0.86	1.10	$0.74^{+0.26}_{-0.68}$
$p_{T,ll}$	0.96	1.80	0.13	-1.14	$0^{+0.56}$
BDT0	0.02	-1.95	1.00	2.89	$1_{-0.27}$
BDT2	0.95	1.68	0.45	-0.14	$0.09^{+0.54}$
(BDT0, BDT2)	0.32	-0.46	0.76	0.71	1_{-1}

Table L.3: Sensitivities for rejecting the POWHEG and MC@NLO hypotheses, with the rate analysis selection, in the WW control region. Only statistical uncertainties are used, and the POI is fitted in the range $0 \leq \epsilon \leq 1$.

L.2 Constraint to UEPS from the signal-like region only

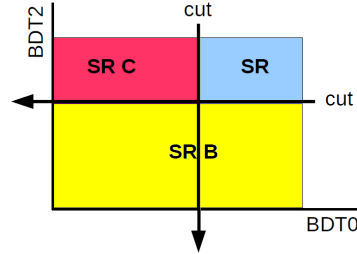


Figure L.3: Schematic illustration of the signal region splitting for the UEPS studies documented in the text.

In order to understand how much of the pull and overconstraint to the UEPS nuisance parameters (in the old approach) is coming from the signal-like region, i.e. from the rightmost part of the BDT response, the SR is split into different regions, as illustrated in Fig. L.3: the same cut is applied on BDT0 and BDT2, defining “SRB” and “SRC”, which are used as two single-bin measurements in the fit, while the remaining phase space, “SR”, is binned in $(10, 5)$ bins like in the standard case, thus retaining the whole shape information. The 2-dimensional phase space is scanned as a function of the cut, as shown in Table L.4, for the fitted values of the UEPS nuisance parameters (in the $\varepsilon = 1$ hypothesis) and the $J^P = 0^+$ signal strength. Starting from the standard configuration (no splitting applied) and increasing the value of the cut up to 0.2, the overconstraint on UEPS_SHAPE decreases (from 24% to 51%), while the pull on UEPS_norm improves (from 42% to 15% in absolute values); at the same time, the fitted signal strength increases, resulting in 19% more signal with respect to the standard case when defining the SR with $\text{BDT0} > 0.2$ and $\text{BDT2} > 0.2$.

These numbers show that the background-like part of the signal region is in fact constraining the UEPS parameters in the old approach, a situation which is not optimal since it limits the freedom of the fit to adjust the shape and normalization of the WW background in the signal-like region, and has a direct consequence on the signal yield computed by the fit. The baseline treatment of the UEPS systematic uncertainty, on the other hand, overcomes these issues thanks to its additional degrees of freedom.

	UEPS_SHAPE	UEPS_norm	$\hat{\mu}(J^P = 0^+)$
standard	0 ± 0.24	0.42 ± 0.93	$1.43^{+0.53}_{-0.44}$
cut -0.9	0 ± 0.24	0.40 ± 0.94	$1.39^{+0.52}_{-0.45}$
cut -0.7	0 ± 0.25	0.28 ± 0.95	$1.42^{+0.53}_{-0.46}$
cut -0.5	0 ± 0.35	0.28 ± 0.96	$1.43^{+0.55}_{-0.46}$
cut -0.2	0 ± 0.40	-0.08 ± 0.97	$1.58^{+0.58}_{-0.48}$
cut 0.0	0 ± 0.39	-0.17 ± 0.98	$1.59^{+0.58}_{-0.49}$
cut 0.2	0 ± 0.51	-0.15 ± 0.99	$1.70^{+0.60}_{-0.52}$

Table L.4: Fitted values for the UEPS nuisance parameters (shape and normalization, in the old prescription) and for the $J^P = 0^+$ signal strength, as a function of the cut which defines the splitting between SR, SRB and SRC (see Fig. L.3). The “standard” configuration has no splitting applied. Only the main systematic uncertainties are used for this study (the theoretical, UEPS and Fake Rate uncertainties, plus the error associated with the limited number of MC events). All values correspond to the $\varepsilon = 1$ fit.

Appendix M

Training on backgrounds

Similarly to what shown in Appendix E for the WW background, trained as signal versus the remaining backgrounds, in the context of the 2012 analysis, also for the spin analysis each one of the main backgrounds is trained versus all the others, one at a time. This allows to validate the BDT modelling of shapes and correlations in the input variables, and to gain confidence in the algorithm itself. Backgrounds under consideration are WW , $t\bar{t}$ and $Z/\gamma^* + \text{jets}$. The phase space used in each case is similar to that of the standard selection, but targetting the selected background more specifically. This study uses a reduced dataset corresponding to an integrated luminosity of 5.8fb^{-1} and no systematic uncertainties.

M.1 WW background

When training the WW background against the others, no $m_{\ell\ell}$ cut is applied, since the latter is meant to focus on the signal (cf. Section 5.4.1). Moreover, the minimum number of events per leaf is not specified manually; all the rest follows the standard training. The resulting BDT response is shown in Fig. M.1 (a): an excellent separation between the signal and the background is achieved, and no sign of overtraining is visible, thanks to the high number of raw MC events for the signal (around ten times that of the Higgs boson samples). Fig. M.1 (b) shows the BDT response in the data, with all backgrounds normalized by means of their pre-fit scale factors, after an additional cut on $m_{\ell\ell}$, $m_{\ell\ell} > 50$ GeV. A $\sim 10\%$ disagreement between the data and the MC expectation is observed in a few central bins, recovered by the fit; the shapes look similar, with a 66% compatibility assessed by the KS test.

M.2 $t\bar{t}$ background

Similarly to the WW case, no $m_{\ell\ell}$ cut is applied when training the $t\bar{t}$ background against the others; moreover, the b -tag veto is removed and the minimum number

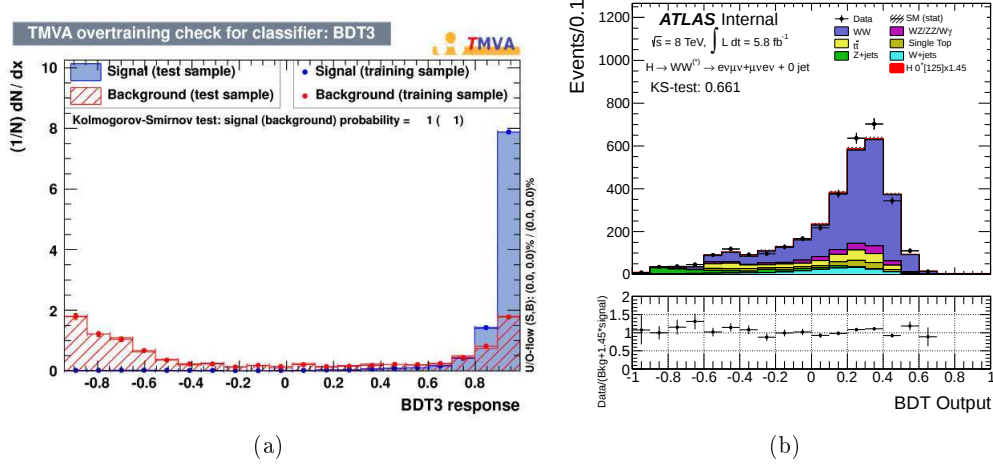
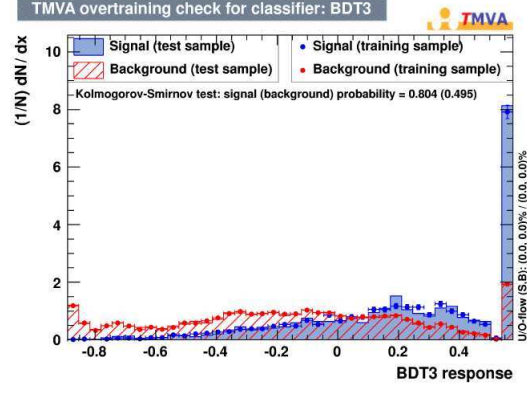


Figure M.1: Distributions of the BDT response obtained training the WW background against the others. This study uses a reduced dataset corresponding to an integrated luminosity of 5.8fb^{-1} and no systematic uncertainties. (a) Overtraining check, showing the compatibility of the BDT responses for the training (dots) and testing (full areas) samples. (b) BDT response in data and MC, with the $J^P = 0^+$ signal (normalized by an approximate post-fit strength, $\hat{\mu} = 1.45\sigma_{\text{SM}}$) stacked on top of the total background.

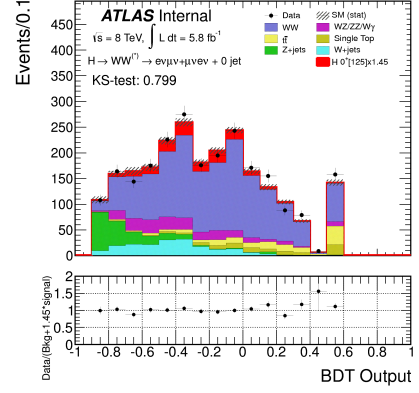
of events per leaf is chosen so that no overtraining appears (the number of MC $t\bar{t}$ events used for training, which is around 6000, is divided by 400). Fig. M.2 shows the resulting BDT response distribution, for (a) the overtraining check and (b) the comparison between the data and the MC expectation. Due to the low number of MC events for the $t\bar{t}$ background in the 0-jet channel, and the usage of the four standard topological variables, the separation achieved is not ideal, but still most of the signal is peaking on the righthand side. A good agreement is reached between the data and the MC prediction, with a compatibility close to 80%.

M.3 $Z/\gamma^* + \text{jets}$ background

In this case, both the $m_{\ell\ell}$ and $p_{T,\ell\ell}$ cuts are dropped in training, and the minimum number of events per leaf is chosen to be equal to the number of MC $Z/\gamma^* + \text{jets}$ events used for training, i.e. around 21600, divided by 50. Fig. M.3 shows the resulting BDT response: the overtraining check (a) exhibits an excellent separation between the signal and the background, as also visible in the data/MC comparison (b).

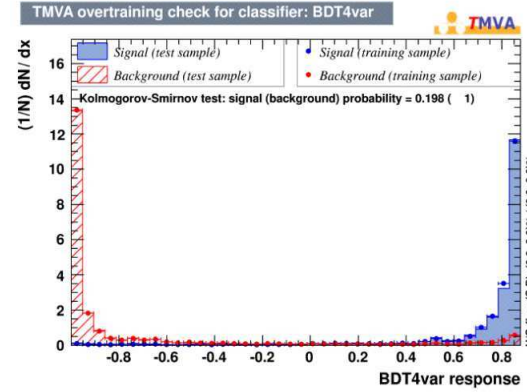


(a)

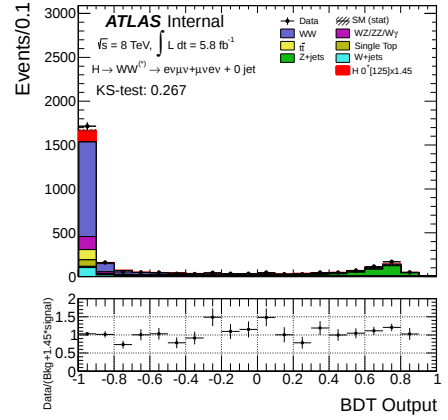


(b)

Figure M.2: Distributions of the BDT response obtained training the $t\bar{t}$ background against the others, with 5.8fb^{-1} and no systematic uncertainties. (a) Overtraining check, showing the compatibility of the BDT responses for the training (dots) and testing (full areas) samples. (b) BDT response in data and MC, with the $J^P = 0^+$ signal ($\mu = 1.45\sigma_{\text{SM}}$) stacked on top of the total background.



(a)



(b)

Figure M.3: Distributions of the BDT response obtained training the $Z/\gamma^* + \text{jets}$ background against the others, with 5.8fb^{-1} and no systematic uncertainties. (a) Overtraining check, showing the compatibility of the BDT responses for the training (dots) and testing (full areas) samples. (b) BDT response in data and MC, with the $J^P = 0^+$ signal ($\mu = 1.45\sigma_{\text{SM}}$) stacked on top of the total background.

Appendix N

Definition of the test statistic: comparison to the PLR method

Two approaches can be used to define the test statistic for spin rejection, as anticipated in Section 5.7.1: the one adopted in this analysis compares the likelihoods for the $\varepsilon = 0$ and $\varepsilon = 1$ hypotheses:

$$q = \ln \frac{\mathcal{L}(\varepsilon = 1, \hat{\mu}_{\varepsilon=1}, \hat{\theta}_{\varepsilon=1})}{\mathcal{L}(\varepsilon = 0, \hat{\mu}_{\varepsilon=0}, \hat{\theta}_{\varepsilon=0})},$$

while the alternative one compares both of them with the ε -free case (PLR) [230,231]:

$$q_{\varepsilon=0} = \ln \frac{\mathcal{L}(\varepsilon = 0, \hat{\mu}_{\varepsilon=0}, \hat{\theta}_{\varepsilon=0})}{\mathcal{L}(\hat{\varepsilon}, \hat{\mu}_{\hat{\varepsilon}}, \hat{\theta}_{\hat{\varepsilon}})} \quad , \quad q_{\varepsilon=1} = \ln \frac{\mathcal{L}(\varepsilon = 1, \hat{\mu}_{\varepsilon=1}, \hat{\theta}_{\varepsilon=1})}{\mathcal{L}(\hat{\varepsilon}, \hat{\mu}_{\hat{\varepsilon}}, \hat{\theta}_{\hat{\varepsilon}})}.$$

The two approaches are expected to yield compatible results, but a few caveats should be considered:

1. firstly, the $H \rightarrow WW$ analysis includes more than 100 nuisance parameters, so that pathological situations may easily trigger instabilities in the fit,
2. secondly, the physical meaning of letting ε free is not well defined: the model under investigation does *not* take into account a mixing between different J^P hypotheses, so that only the two extreme cases, $\varepsilon = 0$ *or* 1, should be considered,
3. finally, the interplay between the fit to ε and the fit to μ can lead to problematic situations, especially when one of them hits the boundaries.

A first look at the results computed with the two approaches, in fact, does not confirm a full compatibility, as can be seen in Table N.1, where the expected sensi-

tivities, using only statistical uncertainties, are shown for three different fractions.

	PLR	standard	PLR-standard /PLR (%)
$f_{q\bar{q}} = 0.00$ $Z(0^+)$	1.50	1.52	1.7
$f_{q\bar{q}} = 0.00$ $Z(2^+)$	1.73	1.68	2.7
$f_{q\bar{q}} = 0.25$ $Z(0^+)$	1.63	1.74	6.7
$f_{q\bar{q}} = 0.25$ $Z(2^+)$	2.02	1.99	1.3
$f_{q\bar{q}} = 1.00$ $Z(0^+)$	1.81	2.18	20.4
$f_{q\bar{q}} = 1.00$ $Z(2^+)$	3.16	2.99	5.3

Table N.1: Expected sensitivities for $J^P = 0^+$ and $J^P = 2^+$ rejection, with the PLR and the standard test statistics. The numbers are shown for different $f_{q\bar{q}}$ fractions, without systematic uncertainties. $Z(0^+)$ and $Z(2^+)$ are the sensitivities (in terms of standard deviations) for $J^P = 0^+$ and $J^P = 2^+$ rejection.

While numbers agree well for pure gg production, and in general for $J^P = 2^+$ rejection sensitivities, the same is not true for $J^P = 0^+$ rejection as the $f_{q\bar{q}}$ fraction increases, with a discrepancy between the two methods up to 20%. Adding the systematic uncertainties, the disagreement is further enhanced. In order to track down this effect, different configurations have been explored, as summarized in Table N.2. The first line corresponds to the standard case already shown in Table N.1. In the second line, ε and μ are not forced to be positive in the fit, so that the minimization procedure does not hit the boundaries; the discrepancy in $Z(0^+)$ for pure $q\bar{q}$ production diminishes from 20% to 11% (even if the discrepancy for pure gg production increases). The third and fourth lines exhibit the largest deviations, and are both computed forcing $\mu = 1$ when generating the toy distributions (but still allowing it to float in the final fit). Finally, the fifth line shows almost no disagreement, for both fractions, and corresponds to the case of $\mu = 1$ both in toy generation and fitting.

	standard $f_{q\bar{q}} = 0.00$	standard $f_{q\bar{q}} = 1.00$	PLR $f_{q\bar{q}} = 0.00$	PLR $f_{q\bar{q}} = 1.00$	$f_{q\bar{q}} = 0.00$ diff to PLR (%)	$f_{q\bar{q}} = 1.00$ diff to PLR (%)
$\varepsilon[0, 1], \mu[0, 10]$	Z0 : 1.52 Z2 : 1.68	Z0 : 2.18 Z2 : 2.99	Z0 : 1.50 Z2 : 1.73	Z0 : 1.81 Z2 : 3.16	Z0 : 1.7 Z2 : 2.7	Z0 : 20 Z2 : 5.3
$\varepsilon[-5, 5], \mu[-10, 10]$	Z0 : 1.48 Z2 : 1.74	Z0 : 2.13 Z2 : 2.85	Z0 : 1.60 Z2 : 1.77	Z0 : 1.92 Z2 : 3.09	Z0 : 7.5 Z2 : 1.6	Z0 : 11 Z2 : 7.8
$\varepsilon[0, 1], \mu_{\text{GEN}} = 1, \mu_{\text{TEST}} \text{ free}$	Z0 : 0.70 Z2 : 1.26	Z0 : 1.41 Z2 : 1.94	Z0 : 0.51 Z2 : 1.29	Z0 : 1.19 Z2 : 1.95	Z0 : 38 Z2 : 2.7	Z0 : 18 Z2 : 0.5
$\varepsilon[-5, 5], \mu_{\text{GEN}} = 1, \mu_{\text{TEST}} \text{ free}$	Z0 : 0.70 Z2 : 1.26	Z0 : 1.30 Z2 : 1.66	Z0 : 0.63 Z2 : 1.26	Z0 : 1.31 Z2 : 1.95	Z0 : 11 Z2 : 0.3	Z0 : 33 Z2 : 26
$\varepsilon[-5, 5], \mu_{\text{GEN}} = \mu_{\text{TEST}} = 1$	Z0 : 1.56 Z2 : 1.58	Z0 : 1.97 Z2 : 1.98	Z0 : 1.53 Z2 : 1.56	Z0 : 1.99 Z2 : 1.93	Z0 : 1.5 Z2 : 1.3	Z0 : 0.6 Z2 : 2

Table N.2: Expected sensitivities in different configurations, including only statistical uncertainties. For brevity's sake, the significances $Z(0^+)$ and $Z(2^+)$ are labelled as Z0 and Z2.

These results seem to indicate the presence of a correlation between the discrepancy and the fit to the signal strength; in particular, since the basic difference

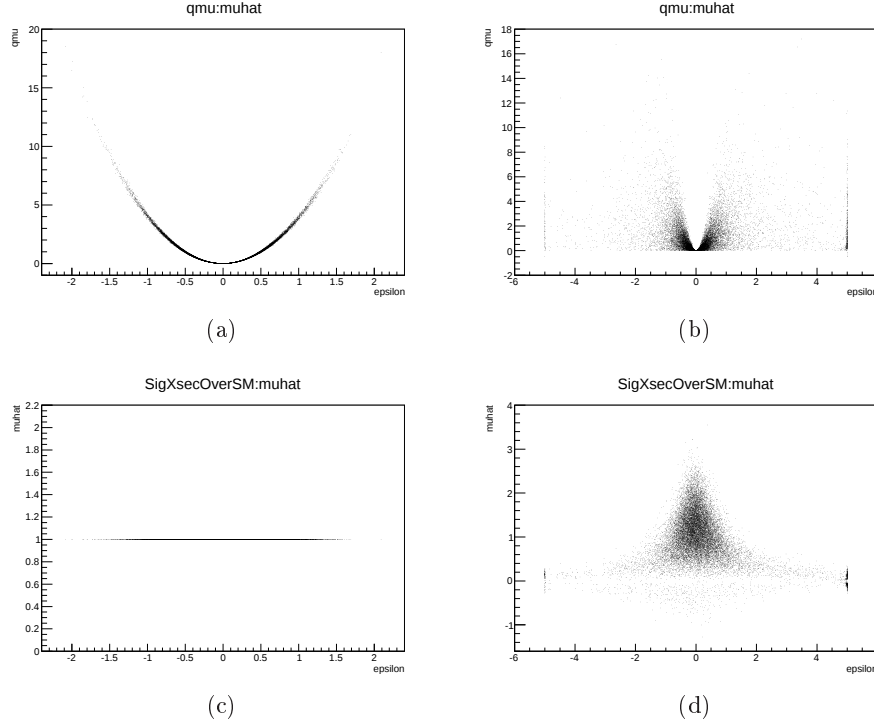


Figure N.1: Toy distributions of $q_\mu(\epsilon)$ for $\mu_{\text{GEN}} = \mu_{\text{TEST}} = 1$ with $\epsilon_{\text{GEN}} = 0$ and then fitted under the assumption (a) $\epsilon = 0$ (corresponding to the fifth line in Table N.2) and (b) with μ_{TEST} free (fourth line). (c,d) Distributions of $\hat{\mu}(\epsilon)$ in the same configurations.

between the two methods resides in the fit to ϵ , the issue may emerge when fitting ϵ and μ at the same time. More specifically, Fig. N.1 (a) shows the distribution of the test statistic q_μ as a function of ϵ in case $\mu_{\text{GEN}} = \mu_{\text{TEST}} = 1$ (fifth line in Table N.2), with $\epsilon_{\text{GEN}} = 0$ and then fitted under the assumption that $\epsilon = 0$ is true: there is a biunivocal correspondance between q_μ and ϵ . On the other hand, Fig. N.1 (b) shows the same but with μ_{TEST} free (fourth line in Table N.2): the biunivocal correspondance (q_μ, ϵ) is broken, and the difference between the two methods is maximal. The associated distributions of $\hat{\mu}$ as a function of ϵ are shown in Fig. N.1 (c,d): it appears evident that, in the μ_{TEST} -free case, the distribution gets broader as $\hat{\mu}$ approaches to zero, causing those toys to lose discrimination on ϵ . The smaller the best fit value for $\hat{\mu}$, the more toys are close to zero, i.e. in the degenerate region, which also explains why the effect under study gets more pronounced for smaller $q\bar{q}$ fractions.

In conclusion, the two methods yield the same results when fixing the signal strength in the fit. When profiling it on data, on the other hand, a discrepancy appears, which is a function of the $J^P = 2^+$ fraction, being maximal for $f_{q\bar{q}} = 1.00$:

this is due to the fact that, through the profiling, $q_\mu(\varepsilon)$ becomes degenerate, thus the global minimum of the likelihood is not well-defined anymore. This also means that, having broken the asymptotic approximation, the two methods are not supposed to converge asymptotically to the same result. Finally, when the mean value of the fitted μ is far enough from zero, less toys end up in the degenerate region, so that the biunivocal correspondance is restored, and the PLR method converges better to the standard one.

Appendix O

Results for all the $J^P = 2^+$ fractions

This Appendix presents additional results for all the $J^P = 2^+$ fractions.

Table O.1 summarizes the relative uncertainties on expected and observed p_0 -values, given by the limited toy statistics. As $f_{q\bar{q}}$ increases, less toys populate the tails, so the errors become larger. Fig. O.1 compares the shapes of the input variables for the five fractions: the biggest difference is visible in $\Delta\phi_{\ell\ell}$, where $f_{q\bar{q}} = 1.00$ peaks to larger values, thus being maximally different from $J^P = 0^+$. Figs. O.2 and O.3 show the input variables, in the signal region, for the $J^P = 0^+$ and the $J^P = 2^+$ hypotheses. Already visually, a SM Higgs boson seems easier to reconcile with the excess found in data than the alternative $J^P = 2^+$ hypotheses under consideration. Finally, Figs. O.4 and O.5 show the post-fit pulls of the NPs. The parameters exhibiting overconstraints and pulls (and thus marked in red) are the same discussed for $f_{q\bar{q}} = 0.25$ (primarily Fake Rate, JES and JER). An exception is EL_RES, that is overconstrained in the $f_{q\bar{q}} = 0.75$ fit with $\varepsilon = 1$; it has been checked that this overconstraint is due to statistical fluctuations in data. For what concerns the $J^P = 2^+$ fit, it becomes more and more difficult to reconcile the data with the alternative hypothesis as $f_{q\bar{q}}$ increases: for $f_{q\bar{q}} = 1.00$, in particular, various NPs show a pathological behaviour.

production fraction	exp. $p_0(0^+)$	exp. $p_0(2^+)$	obs. $p_0(0^+)$	obs. $p_0(2^+)$
$f_{q\bar{q}} = 1.00$	0.013 (6.9%)	$3.6 \cdot 10^{-4}$ (37.8%)	0.541 (1.0%)	$1.7 \cdot 10^{-4}$ (44.7%)
$f_{q\bar{q}} = 0.75$	0.028 (4.4%)	0.003 (13.9%)	0.586 (1.0%)	0.001 (21.8%)
$f_{q\bar{q}} = 0.50$	0.042 (4.1%)	0.009 (9.0%)	0.616 (1.1%)	0.003 (14.9%)
$f_{q\bar{q}} = 0.25$	0.048 (3.9%)	0.019 (6.2%)	0.622 (1.1%)	0.008 (9.8%)
$f_{q\bar{q}} = 0.00$	0.086 (3.0%)	0.054 (3.8%)	0.731 (1.0%)	0.013 (7.8%)

Table O.1: Summary of p_0 -values for the five $f_{q\bar{q}}$ fractions (as given in Table 5.15), with their relative uncertainties, due to limited toy statistics.

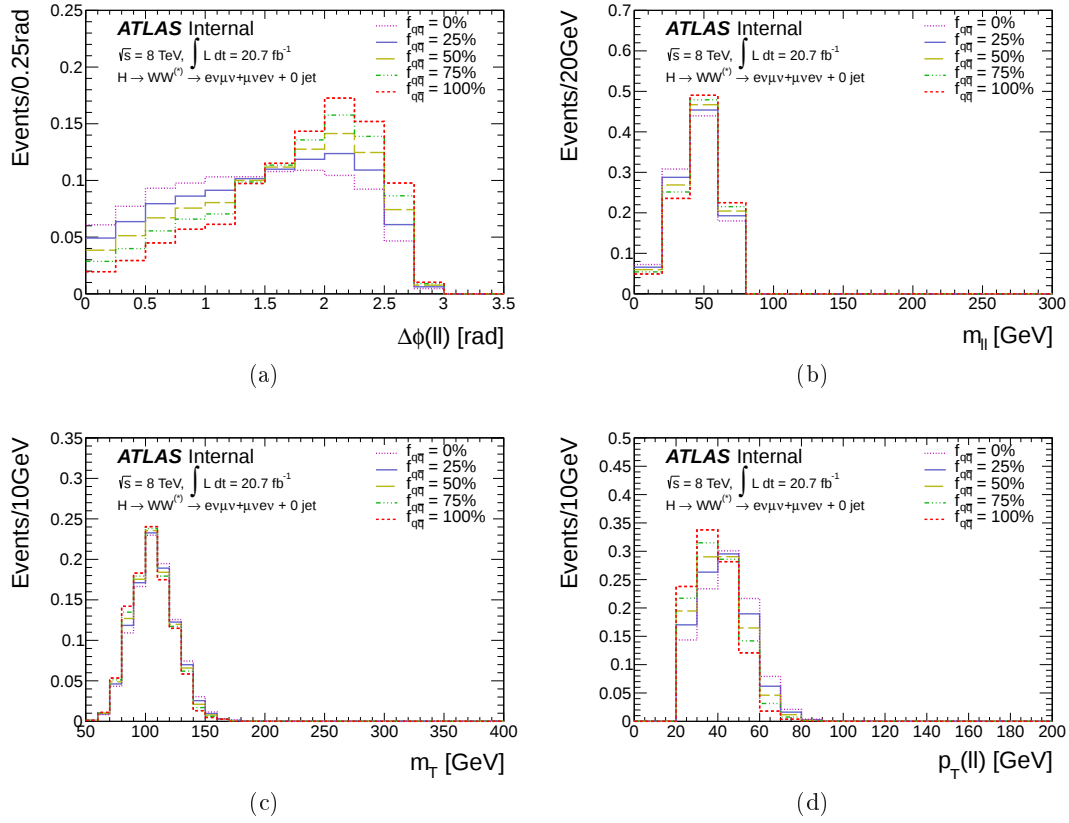


Figure O.1: Shape distributions, normalized to unity, comparing the five $J^P = 2^+$ fractions, for the four input variables: (a) $\Delta\phi_{\ell\ell}$, (b) $m_{\ell\ell}$, (c) m_T , (d) $p_{T,\ell\ell}$.

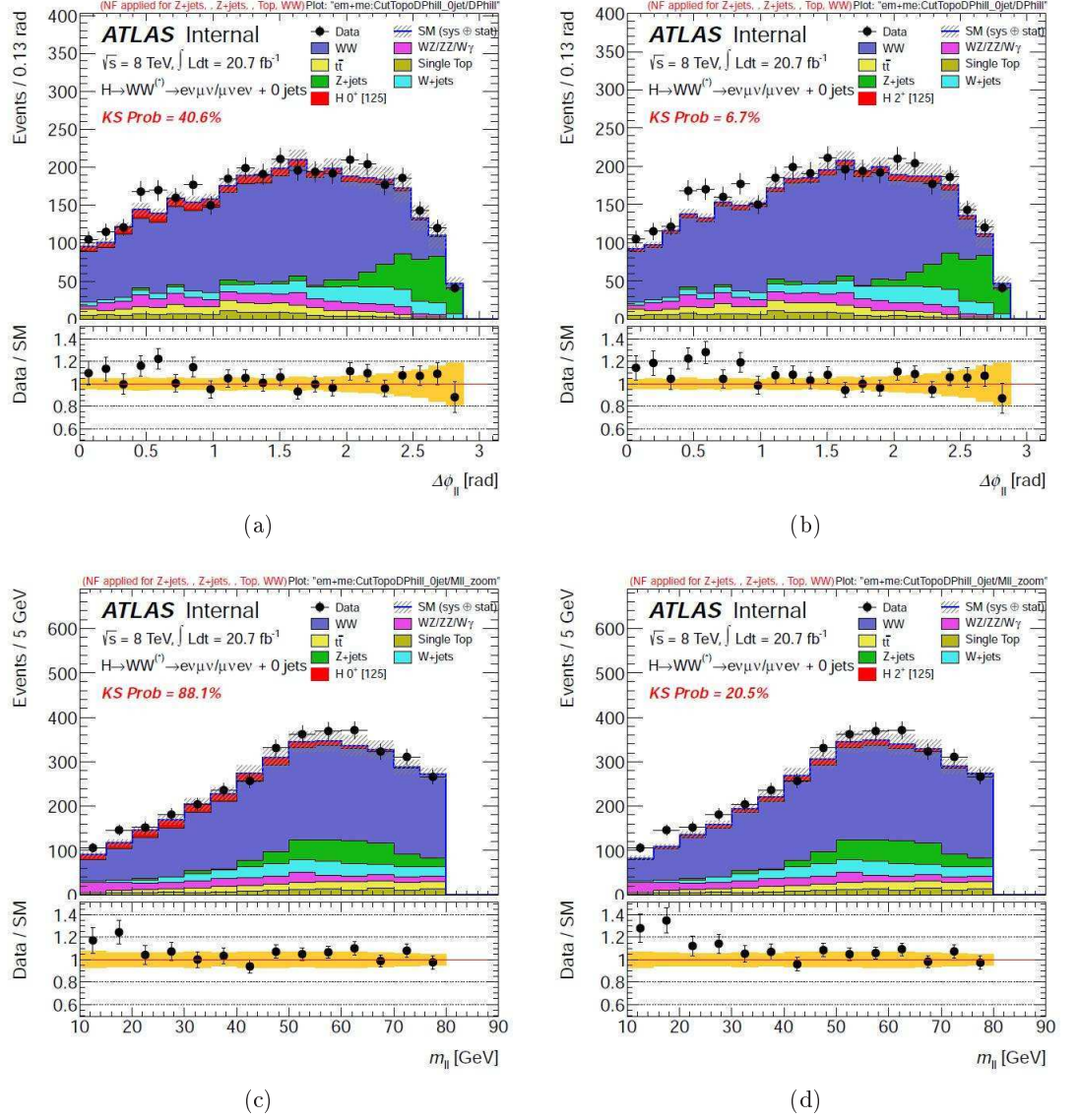


Figure O.2: Distributions of the four input variables in the 0-jet signal region, comparing the $J^P = 0^+$ and the $J^P = 2^+$ signal expectations (shown on the left and on the right, respectively): (a,b) $\Delta\phi_{\ell\ell}$, (c,d) $m_{\ell\ell}$. All backgrounds are normalized as in Table 5.2. The shaded bands indicate the total normalization systematic uncertainty.

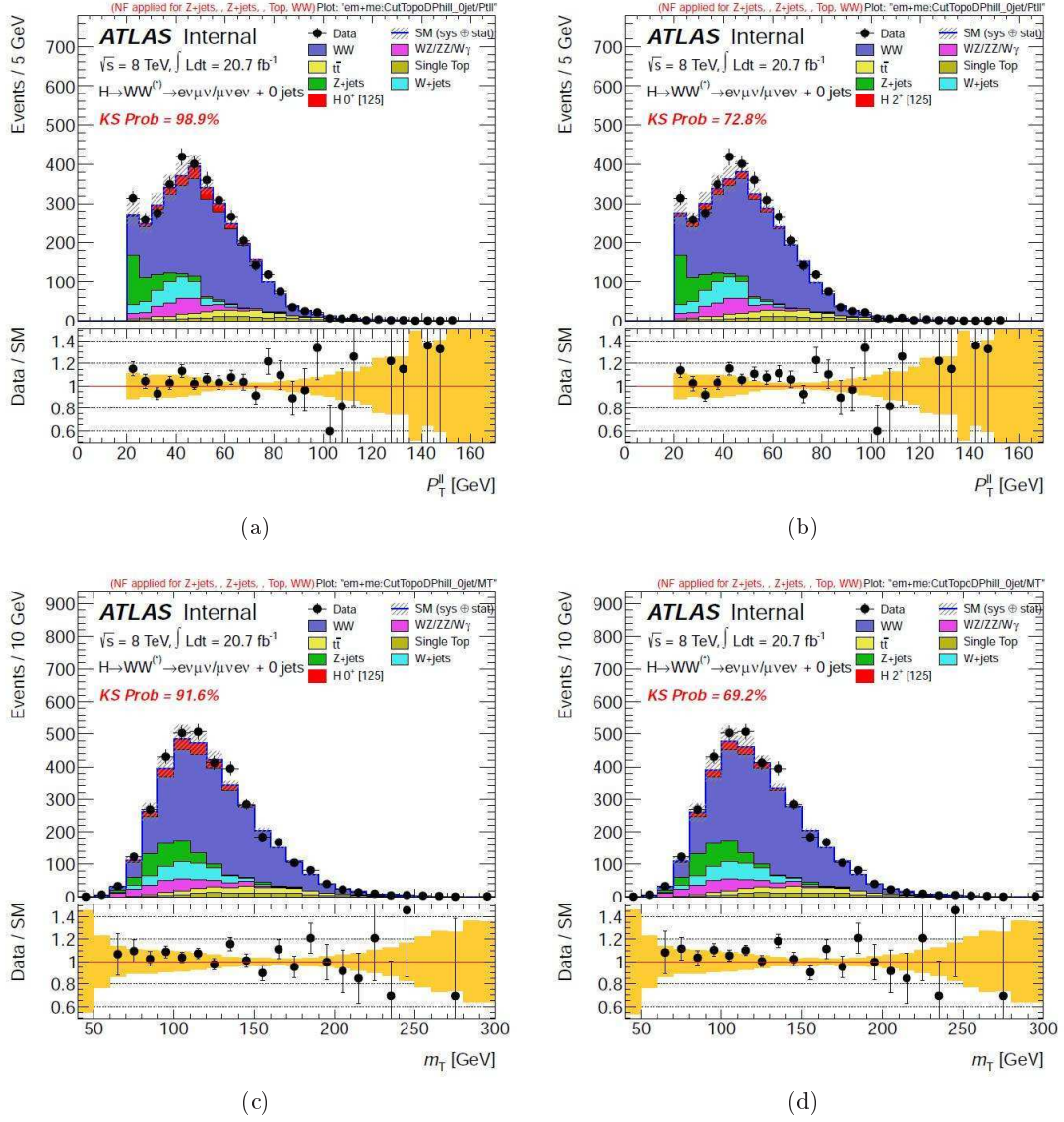
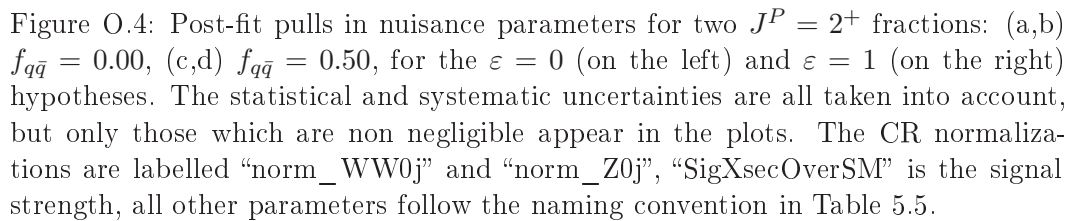


Figure O.3: Distributions of the four input variables in the 0-jet signal region, comparing the $J^P = 0^+$ and the $J^P = 2^+$ signal expectations (shown on the left and on the right, respectively): (a,b) $p_{T,\ell\ell}$, (c,d) m_T . All backgrounds are normalized as in Table 5.2. The shaded bands indicate the total normalization systematic uncertainty.



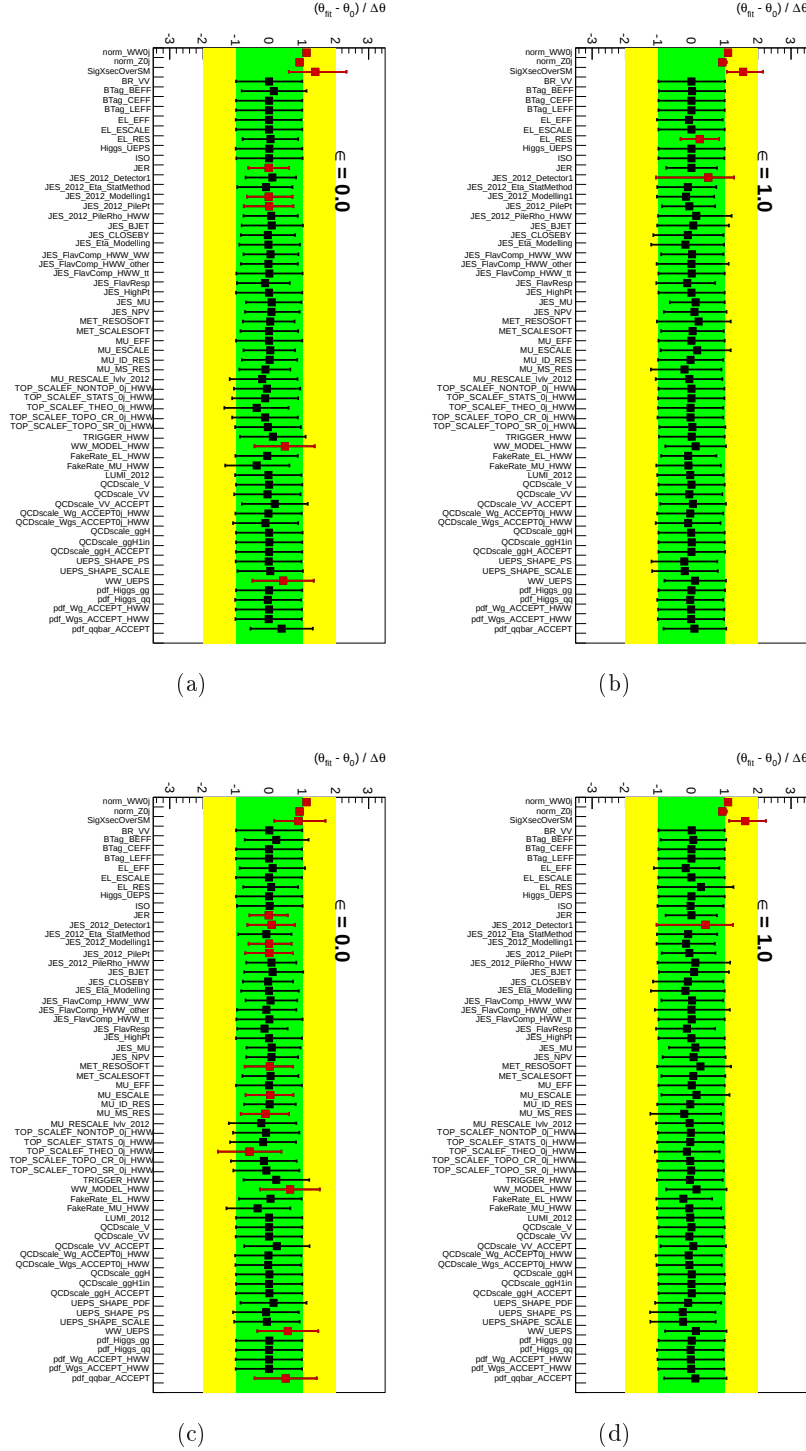


Figure 0.5: Post-fit pulls in nuisance parameters for two $J^P = 2^+$ fractions: (a,b) $f_{q\bar{q}} = 0.75$, (c,d) $f_{q\bar{q}} = 1.00$, for the $\varepsilon = 0$ (on the left) and $\varepsilon = 1$ (on the right) hypotheses. The statistical and systematic uncertainties are all taken into account, but only those which are non negligible appear in the plots. The CR normalizations are labelled “norm_WW0j” and “norm_Z0j”, “SigXsecOverSM” is the signal strength, all other parameters follow the naming convention in Table 5.5.

Conclusions

The work described in this thesis has been devoted to the study of the $WW^{(*)} \rightarrow \ell^+ \nu \ell^- \bar{\nu}$ decay channel for a Standard Model Higgs boson, with proton-proton collision data collected in 2011 and 2012 with the ATLAS experiment at the LHC. This study can be divided into two parts. First, a multivariate search has been developed, over the mass range $110 \text{ GeV} < m_H \leq 600 \text{ GeV}$, and applied to data corresponding to an integrated luminosity of 4.7 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$. Later, after the discovery by the ATLAS and the CMS collaborations of a new resonance at a mass close to 125 GeV [119, 233], its properties have been explored, again using a multivariate approach, based on data corresponding to an integrated luminosity of 20.7 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$.

As discussed in Chapter 4, the search carried out in 2011 uses boosted decision trees to improve the exclusion power of the corresponding cut-based analysis, exploiting the correlations among the variables which characterize the final-state particles. The first attempt of designing a multivariate search in this channel dates back to the beginning of 2011, and, as also documented in this thesis, it exploited a Neural Network discriminant. Since then, several multivariate techniques have been considered, all of which the BDT was able to outperform; for this reason, a decision has been made, to convey all the efforts of the $H \rightarrow WW$ ATLAS working group into developing it further. The analysis presented in this thesis constitutes a major contribution in this direction. All the salient parts of the analysis have been developed, from the choice of the input variables, to the optimization of the BDT settings, to the definition of the signal and control regions, to the treatment and impact of the various theoretical and experimental uncertainties, to the statistical model used for the fit, up to the final results and their comparison with the cut-based ones. Before the publication of the results of this analysis in Ref. [27], the most stringent constraints on the mass of a Standard Model Higgs boson came from direct searches, which excluded the high-mass range and set a lower threshold of $m_H > 114.4 \text{ GeV}$ [18, 19]. The BDT analysis presented in this thesis substantially improves these results, excluding the ranges $130 \text{ GeV} \leq m_H \leq 281 \text{ GeV}$ at 95% CL, to be compared with

an expected exclusion in the range $127 \text{ GeV} \leq m_H \leq 254 \text{ GeV}$. The results of the cut-based analysis in the same final state are comparable or worse, being limited to the range $133 \text{ GeV} \leq m_H \leq 258 \text{ GeV}$ at 95% CL. Besides confirming the cut-based results at low mass, the BDT analysis is able to significantly improve the sensitivity in the high-mass range, lowering the expected CL_s limits of more than 20% for $m_H \gtrsim 250 \text{ GeV}$. This analysis has been the first multivariate search approved by the ATLAS Higgs working group, requiring many months of deep scrutiny, validation and additional cross-checks.

In summer 2012, the same analysis, run on a dataset corresponding to 13.0 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$ (the same used in the $\gamma\gamma$ and $ZZ^* \rightarrow 4\ell$ final states for the first observation in July 2012 [29,30]), has been able to detect the presence of an excess of events at a mass close to 125 GeV, compatible with the long-sought Standard Model Higgs boson, again confirming what measured by the $H \rightarrow WW^*$ cut-based analysis, namely a signal strength of $\mu = 1.5 \pm 0.6$ for $m_H = 125 \text{ GeV}$, in units of the Standard Model cross section [32]. Rather than focusing on the measurement of the rate, for which the cut-based analysis seems to be good enough, the BDT analysis has been demonstrated, also thanks to the efforts reported in this thesis, to be the best suited for exploring the quantum numbers of the new resonance, and in particular its spin and parity assignments. This, together with the measurement of its couplings to bosons and fermions [242], is crucial to allow its identification with the Higgs boson as predicted in the Standard Model.

As described in Chapter 5, a new BDT analysis has been developed, starting from the experience gained in 2011, to test the SM value of $J^P = 0^+$ against the best motivated alternative hypotheses, namely $1^+, 1^-, 2^+$. A possible $J^P = 0^-$ term is not investigated, due to the reduced sensitivity of the $H \rightarrow WW^*$ final state for this hypothesis, in particular compared to the $H \rightarrow ZZ^* \rightarrow 4\ell$ one. The $J^P = 2^+$ model follows the minimal model for Kaluza-Klein-like gravitons, 2_m^+ [33–36], and has been investigated in five configurations for the couplings to quarks and gluons (from pure gg production to pure $q\bar{q}$). From the experimental point of view, it is worthwhile to explore in this final state ..,also the $J^P = 1^\pm$ hypothesis, despite the fact that a spin-1 Higgs boson cannot decay into two on-shell photons, due to the Landau-Yang theorem, thus contradicting the observation in the $\gamma\gamma$ final state [39].

This thesis covers all aspects of the $H \rightarrow WW^*$ spin analysis, from its first conception to the final results, discussing also several of the additional studies which have been performed in order to validate, consolidate and optimize the analysis. The results have been published individually [40], and later as a basic part of the spin combination performed by the ATLAS collaboration [41]. The $H \rightarrow WW^*$ analysis alone rejects the $J^P = 2^+$ hypothesis with probabilities ranging from 95.2% for pure

gg production to 99.7% for pure $q\bar{q}$ production, and rejects the $J^P = 1^\pm$ hypotheses with probabilities equal to 92.1% for $J^P = 1^+$ and 98.3% for $J^P = 1^-$.

Combining them with the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ ones, the $J^P = 2^+$ hypothesis is excluded with a confidence level larger than 99.9% for all production modes, while the $J^P = 1^\pm$ hypothesis is excluded with a CL above 99.7%. Hence, these results can be regarded as the first evidence for the spin-0 nature of the new particle, and thus its compatibility with the Standard Model hypothesis, a statement which motivates the award of the 2013 Nobel Prize in Physics to Peter Higgs and François Englert [42], which first predicted its existence forty years ago.

Bibliography

- [1] M. Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys.Lett.*, 8:214–215, 1964.
- [2] G. Zweig. An $SU(3)$ model for strong interaction symmetry and its breaking. Version 1. CERN-TH-401, 1964.
- [3] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys.Lett.*, B47:365–368, 1973.
- [4] D.J. Gross and F. Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys.Rev.Lett.*, 30:1343–1346, 1973.
- [5] H.D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys.Rev.Lett.*, 30:1346–1349, 1973.
- [6] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl.Phys.*, 22:579–588, 1961.
- [7] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.
- [8] A. Salam. in *Elementary Particle Theory*, p. 367. Almqvist and Wiksell, Stockholm, 1968.
- [9] S. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys.Rev.*, D2:1285, 1970.
- [10] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321, 1964.
- [11] P.W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys.Rev.Lett.*, 13:508–509, 1964.
- [12] P.W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys.Rev.*, 145:1156–1163, 1966.

- [13] G.S. Guralnik, C.R. Hagen, and T.W. Kibble. Global Conservation Laws and Massless Particles. *Phys.Rev.Lett.*, 13:585, 1964.
- [14] T.W. Kibble. Symmetry breaking in nonAbelian gauge theories. *Phys.Rev.*, 155:1554, 1967.
- [15] G. 't Hooft. Renormalization of Massless Yang-Mills Fields. *Nucl.Phys.*, B33:173–199, 1971.
- [16] C.H. Llewellyn Smith. High-Energy Behavior and Gauge Symmetry. *Phys.Lett.*, B46:233–236, 1973.
- [17] J.S. Bell. High-energy behavior of tree diagrams in gauge theories. *Nucl.Phys.*, B60:427–436, 1973.
- [18] G. Abbiendi et al. Search for the Standard Model Higgs boson at LEP. *Phys. Lett.*, B565:61, 2003.
- [19] Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data. *arXiv:hep-ex/1207.0449*, 2012.
- [20] The LEP Electroweak Working Group. <http://lepewwg.web.cern.ch/LEPEWWG/>.
- [21] <http://home.web.cern.ch/about/accelerators/large-hadron-collider>.
- [22] C. Lefevre. Lhc: the guide. Feb 2009.
- [23] M. Benedikt, P. Collier, V Mertens, J. Poole, and K. Schindl. *LHC Design Report*. Number CERN-2004-003-V-1,2,3. CERN, Geneva, 2004.
- [24] The ATLAS Collaboration. *Letter of Intent for a General-Purpose pp Experiment at the Large Hadron Collider at CERN*. LHC Tech. Proposal. CERN, Geneva, 1992.
- [25] The ATLAS Collaboration. *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*. LHC Tech. Proposal. CERN, Geneva, 1994.
- [26] The ATLAS Collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. *arXiv:hep-ex/0901.0512*, 2008.
- [27] The ATLAS Collaboration. Search for the Standard Model Higgs boson in the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ decay mode using Multivariate Techniques with 4.7 fb^{-1} of ATLAS data at $\sqrt{s} = 7 \text{ TeV}$. *ATLAS-CONF-2012-060*.

- [28] The ATLAS Collaboration. Search for the Standard Model Higgs boson in the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ decay mode with 4.7 fb^{-1} of ATLAS data at $\sqrt{s} = 7$ TeV. *Phys.Lett.*, B716:62–81, 2012.
- [29] The ATLAS Collaboration. Observation of an excess of events in the search for the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channel with the ATLAS detector. *ATLAS-CONF-2012-092*.
- [30] The ATLAS Collaboration. Observation of an excess of events in the search for the Standard Model Higgs boson in the gamma-gamma channel with the ATLAS detector. *ATLAS-CONF-2012-091*.
- [31] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.
- [32] The ATLAS Collaboration. Update of the $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ Analysis with 13 fb^{-1} of $\sqrt{s} = 8$ TeV Data Collected with the ATLAS Detector. *ATLAS-CONF-2012-158*, 2012.
- [33] L. Randall and R. Sundrum. A Large mass hierarchy from a small extra dimension. *Phys.Rev.Lett.*, 83:3370–3373, 1999.
- [34] L. Randall and R. Sundrum. An Alternative to compactification. *Phys.Rev.Lett.*, 83:4690–4693, 1999.
- [35] Y. Gao, A.V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, et al. Spin determination of single-produced resonances at hadron colliders. *Phys.Rev.*, D81:075022, 2010.
- [36] S. Bolognesi, Y. Gao, A.V. Gritsan, K. Melnikov, M. Schulze, et al. On the spin and parity of a single-produced resonance at the LHC. *arXiv:hep-ph/1208.4018*, 2012.
- [37] L.D. Landau. On the angular momentum of a two-photon system. *Dokl.Akad.Nauk Ser.Fiz.*, 60:207–209, 1948.
- [38] Y. Chen-Ning. Selection Rules for the Dematerialization of a Particle Into Two Photons. *Phys.Rev.*, 77:242–245, 1950.
- [39] The ATLAS Collaboration. Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data. *ATLAS-CONF-2013-012*, 2013.

- [40] The ATLAS Collaboration. Study of the spin properties of the Higgs-like particle in the $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ channel with 21 fb^{-1} of $\sqrt{s} = 8 \text{ TeV}$ data collected with the ATLAS detector. *ATLAS-CONF-2013-031*, 2013.
- [41] The ATLAS Collaboration. Evidence for the spin-0 nature of the Higgs boson using ATLAS data. *Phys.Lett.*, B726:120–144, 2013.
- [42] The Nobel Prize in Physics 2013 - Advanced Information. <http://www.nobelprize.org/nobel-prizes/physics/laureates/2013/advanced.html>.
- [43] A. Pich. The Standard Model of Electroweak Interactions. *arXiv:hep-ph/1201.0537*, 2012.
- [44] David Griffiths. *Introduction to elementary particles*. John Wiley and Sons, 2008.
- [45] A. Pich. Aspects of quantum chromodynamics. *arXiv:hep-ph/0001118*, pages 53–102, 1999.
- [46] A. Djouadi. The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model. *Phys.Rept.*, 457:1–216, 2008.
- [47] The Particle Data Group Collaboration. Review of Particle Physics. *Phys. Rev. D*, 86:010001, Jul 2012.
- [48] O.W. Greenberg. Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons. *Phys.Rev.Lett.*, 13:598–602, 1964.
- [49] M.Y. Han and Y. Nambu. Three Triplet Model with Double SU(3) Symmetry. *Phys.Rev.*, 139:B1006–B1010, 1965.
- [50] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys.Rev.Lett.*, 10:531–533, 1963.
- [51] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog.Theor.Phys.*, 49:652–657, 1973.
- [52] Y. Nambu. Quasiparticles and Gauge Invariance in the Theory of Superconductivity. *Phys.Rev.*, 117:648–663, 1960.
- [53] J. Goldstone. Field Theories with Superconductor Solutions. *Nuovo Cim.*, 19:154–164, 1961.

- [54] J. Goldstone, A. Salam, and S. Weinberg. Broken Symmetries. *Phys.Rev.*, 127:965–970, 1962.
- [55] J. Elias-Miro, J.R. Espinosa, G.F. Giudice, G. Isidori, A. Riotto, et al. Higgs mass implications on the stability of the electroweak vacuum. *Phys.Lett.*, B709:222–228, 2012.
- [56] K. Peters. Search for the Higgs boson in the $\gamma\gamma$ final state at the Tevatron. *PoS*, ICHEP2010:071, 2010.
- [57] The D0 Collaboration. Search for Higgs boson production in oppositely charged dilepton and missing energy final states in 9.7 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$. *Phys.Rev.*, D88:052006, 2013.
- [58] The CDF Collaboration. *CDF Note 10760*.
- [59] J.M. Campbell, J.W. Huston, and W.J. Stirling. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept.Prog.Phys.*, 70:89, 2007.
- [60] J. Stirling. 7 TeV LHC (x,Q2) parton kinematics. <http://www.hep.ph.ic.ac.uk/~wstirling/plots/lhcgrid7.eps>.
- [61] S.D. Drell and Tung-Mow Yan. Partons and their Applications at High-Energies. *Annals Phys.*, 66:578, 1971.
- [62] J.C. Collins and D.E. Soper. The Theorems of Perturbative QCD. *Ann.Rev.Nucl.Part.Sci.*, 37:383–409, 1987.
- [63] L.N. Lipatov. The parton model and perturbation theory. *Sov.J.Nucl.Phys.*, 20:94–102, 1975.
- [64] V.N. Gribov and L.N. Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:438–450, 1972.
- [65] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl.Phys.*, B126:298, 1977.
- [66] Y.L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.
- [67] R.K. Ellis, W.J. Stirling, and B.R. Webber. *QCD and collider physics*, volume 8. 1996.

- [68] S. Moch, J. A. M. Vermaseren, and A. Vogt. The three-loop splitting functions in QCD: The non-singlet case. *Nucl.Phys.*, B688:101–134, 2004.
- [69] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.
- [70] The ATLAS Collaboration. Measurement of the inclusive $W^\pm$ and Z/γ cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys.Rev.*, D85:072004, 2012.
- [71] H.L. Lai, M. Guzzi, J. Huston, Z. Li, P.M. Nadolsky, et al. New parton distributions for collider physics. *Phys.Rev.*, D82:074024, 2010.
- [72] A. Sherstnev and R. S. Thorne. Parton distributions for the LHC. *EPJ*, C55:553, 2009.
- [73] The NNPDF Collaboration. Parton distributions with LHC data. *Nucl.Phys.*, B867:244–289, 2013.
- [74] K. Nowak et al. QCD NLO analysis of inclusive, charm and jet data (HERA-PDF 1.7). pages 403–406, 2012.
- [75] P.M. Nadolsky and Z. Sullivan. PDF uncertainties in WH production at Tevatron. *eConf*, C010630:P510, 2001.
- [76] Z. Sullivan. Fully differential W' production and decay at next-to-leading order in QCD. *Phys.Rev.*, D66:075011, 2002.
- [77] M.A. Dobbs, S. Frixione, E. Laenen, K. Tollefson, H. Baer, et al. Les Houches guidebook to Monte Carlo generators for hadron collider physics. *arXiv:hep-ph/0403045*, pages 411–459, 2004.
- [78] T. Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 0605:026, 2006.
- [79] G. Corcella et al. HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including super-symmetric processes). *JHEP*, 0101:010, 2001.
- [80] The ATLAS Collaboration. Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC. *New J.Phys.*, 13:053033, 2011.
- [81] The ATLAS Collaboration. ATLAS Monte Carlo tunes for MC09. *ATL-PHYS-PUB-2010-002*, Mar 2010.

- [82] The ATLAS Collaboration. New ATLAS event generator tunes to 2010 data. *ATL-PHYS-PUB-2011-008*, Apr 2011.
- [83] The ATLAS Collaboration. Charged particle multiplicities in p p interactions at $\sqrt{s} = 0.9$ and 7 TeV in a diffractive limited phase-space measured with the ATLAS detector at the LHC and new PYTHIA6 tune. *ATLAS-CONF-2010-031*, 2010.
- [84] J.M. Butterworth, Jeffrey R. Forshaw, and M.H. Seymour. Multiparton interactions in photoproduction at HERA. *Z.Phys.*, C72:637–646, 1996.
- [85] S. Agostinelli et al. GEANT 4, A Simulation Toolkit. *Nucl. Instrum. Meth.*, A506:250, 2003.
- [86] M. Duehrssen and K. Jakobs. *Study of Higgs bosons in the WW final state and development of a fast calorimeter simulation for the ATLAS experiment*. PhD thesis, Freiburg U., Freiburg, 2009. Presented on 21 Dec 2009.
- [87] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.
- [88] The LHC Higgs Cross Section Working Group. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSectionsFigures>.
- [89] H. Georgi, S. Glashow, M. Machacek, and D. Nanopoulos. Higgs Bosons from Two Gluon Annihilation in Proton Proton Collisions. *Phys.Rev.Lett.*, 40:692, 1978.
- [90] G. Altarelli, R.K. Ellis, and G. Martinelli. Leptonproduction and Drell-Yan Processes Beyond the Leading Approximation in Chromodynamics. *Nucl. Phys.*, B157:461, 1979.
- [91] A. Kubar and F. Paige. Gluon Corrections to the Drell-Yan Model. *Phys.Rev.*, D19:221, 1979.
- [92] R. Furmanski, W. Petronzio. Lepton - Hadron Processes Beyond Leading Order in Quantum Chromodynamics. *Z.Phys*, C11:293, 1982.
- [93] D. Graudenz, M. Spira, and P. M. Zerwas. QCD corrections to Higgs boson production at proton proton colliders. *Phys.Rev.Lett.*, 70:1372, 1993.
- [94] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas. Higgs boson production at the LHC. *Nucl.Phys.*, B453:17–82, 1995.

- [95] R. Harlander and W. B. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys.Rev.Lett.*, 88:201801, 2002.
- [96] C. Anastasiou and K. Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nucl.Phys.*, B646:220, 2002.
- [97] V. Ravindran, J. Smith, and W. van Neerven. NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions. *Nucl.Phys.*, B665:325, 2003.
- [98] R.D. Ball, M. Bonvini, S. Forte, S. Marzani, and G. Ridolfi. Higgs production in gluon fusion beyond NNLO. *Nucl.Phys.*, B874:746–772, 2013.
- [99] M. Ciccolini, A. Denner, and S. Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. *Phys.Rev.*, D77:013002, 2008.
- [100] P. Bolzoni, F. Maltoni, S. Moch, and M. Zaro. Higgs production via vector-boson fusion at NNLO in QCD. *Phys.Rev.Lett.*, 105:011801, 2010.
- [101] 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.
- [102] M. Spira. QCD effects in Higgs physics. *Fortsch.Phys.*, 46:203–284, 1998.
- [103] T. Figy, C. Oleari, and D. Zeppenfeld. Next-to-leading order jet distributions for Higgs boson production via weak boson fusion. *Phys.Rev.*, D68:073005, 2003.
- [104] T. Han and S. Willenbrock. QCD correction to the $pp \rightarrow WH$ and ZH total cross- sections. *Phys. Lett.*, B273:167–172, 1991.
- [105] H. Baer, B. Bailey, and J. Owens. $O(\alpha_s)$ Monte Carlo approach to $W + \text{Higgs}$ associated production at hadron supercolliders. *Phys.Rev.*, D47:2730, 1993.
- [106] A. Djouadi and M. Spira. SUSY - QCD corrections to Higgs boson production at hadron colliders. *Phys.Rev.*, D62:014004, 2000.
- [107] O. Brein et al. Precision calculations for associated $W H$ and $Z H$ production at hadron colliders. *arXiv:hep-ph/0402003*, 2004.
- [108] S. L. Glashow, Dimitri V. Nanopoulos, and A. Yildiz. Associated Production of Higgs Bosons and Z Particles. *Phys. Rev.*, D18:1724–1727, 1978.

- [109] M. Ciccolini, S. Dittmaier, and M. Kramer. Electroweak radiative corrections to associated WH and ZH production at hadron colliders. *Phys. Rev.*, D68:073003, 2003.
- [110] J.R. Ellis, M.K. Gaillard, and D.V. Nanopoulos. A Phenomenological Profile of the Higgs Boson. *Nucl.Phys.*, B106:292, 1976.
- [111] L. Resnick, M.K. Sundaresan, and P.J.S. Watson. Is there a light scalar boson? *Phys.Rev.*, D8:172–178, 1973.
- [112] E. Braaten and J.P. Leveille. *Phys. Rev.*, D22:715, 1980.
- [113] N. Sakai. *Phys. Rev.*, D22:2220, 1980.
- [114] T. Inami and T. Kubota. *Nucl.Phys.*, B179:171, 1981.
- [115] S.G. Gorishny, A.L. Kataev, and S.A. Larin. *Sov. J. Nucl. Phys.*, 40:329, 1984.
- [116] M. Gavela, G. Girardi, C. Malleville, and P. Sorba. *Nucl. Phys.*, B193:257, 1981.
- [117] L. Okun. *Leptons and Quarks*. Ed. North Holland, Amsterdam, 1982.
- [118] Vainshtein A.I., M.B. Voloshin, V.I. Zakharov, and M.A. Shifman. *J. Nucl. Phys.*, 30:711, 1979.
- [119] 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.
- [120] H. Zheng and D. Wu. *Phys.Rev.*, D42:3760, 1990.
- [121] A. Djouadi, M. Spira, J. van der Bij, and P.M. Zerwas. *Phys.Lett.*, B257:187, 1991.
- [122] A. Djouadi, M. Spira, and P.M. Zerwas. *Phys.Lett.*, B311:255, 1993.
- [123] S. Dawson and R.P. Kauffman. *Phys.Rev.*, D47:1264, 1993.
- [124] K. Melnikov and O. Yakovlev. *Phys.Lett.*, B312:179, 1993.
- [125] J. Fleischer and O.V. Tarasov. *Phys.Lett.*, B584:294, 2004.
- [126] Pocsik, G. and Torma, T. *Z.Phys.*, C6:1, 1980.
- [127] Rizzo, T.G. *Phys.Rev.*, D22:722, 1980.

- [128] Keung, W.Y. and Marciano, W.J. *Phys.Rev.*, D30:248, 1984.
- [129] Fleischer, J. and Jegerlehner, F. *Phys.Rev.*, D23:2001, 1981.
- [130] Kniehl, B.A. *Nucl.Phys.*, B352:1, 1991.
- [131] Bardin, D.Y. and Khristova, P.K. and Vilensky, B.M. *Sov. J. Nucl. Phys.*, 54:833, 1991.
- [132] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber. Precise predictions for the Higgs-boson decay $H \rightarrow WW/ZZ \rightarrow 4$ leptons. *Phys. Rev.*, D74:013004, 2006.
- [133] 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, 1998.
- [134] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, et al. Theory and phenomenology of two-Higgs-doublet models. *Phys.Rept.*, 516:1–102, 2012.
- [135] H.E. Haber and G.L. Kane. The search for supersymmetry: Probing physics beyond the standard model. *Physics Reports*, 117(2–4):75 – 263, 1985.
- [136] R.D. Peccei and H.R. Quinn. CP Conservation in the Presence of Instantons. *Phys.Rev.Lett.*, 38:1440–1443, 1977.
- [137] J.E. Kim. *Phys.Rep.*, 150:1, 1987.
- [138] The ATLAS Collaboration. Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data. Technical Report ATLAS-CONF-2013-013, CERN, Geneva, 2013.
- [139] N. Cabibbo and A. Maksymowicz. Angular Correlations in K_{e4} Decays and Determination of Low Energy $\pi - \pi$ Phase Shifts. *Phys. Rev.*, 137:B438–B443, Jan 1965.
- [140] V. Barger, K. Cheung, A. Djouadi, B. Kniehl, and P. Zerwas. *Phys.Rev.*, D49:79, 1994.
- [141] A. Soni and R.M. Xu. *Phys.Rev.*, D48:5259, 1993.
- [142] J.F. Gunion, A. Stange, and S. Willenbrock. Weakly coupled Higgs bosons. *arXiv:hep-ph/9602238*, 1995.

- [143] N. Arkani-Hamed, S. Dimopoulos, and G.R. Dvali. The Hierarchy problem and new dimensions at a millimeter. *Phys.Lett.*, B429:263–272, 1998.
- [144] P. de Aquino, K. Hagiwara, Q. Li, and F. Maltoni. Simulating graviton production at hadron colliders. *JHEP*, 1106:132, 2011.
- [145] K. Cheung and T.C. Yuan. Could the excess seen at 124-126 GeV be due to the Randall-Sundrum Radion? *Phys.Rev.Lett.*, 108:141602, 2012.
- [146] M. Fierz and W. Pauli. On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc.Roy.Soc.Lond.*, A173:211–232, 1939.
- [147] R. Fok, C. Guimaraes, R. Lewis, and V. Sanz. It is a Graviton! Or maybe not. *JHEP*, 1212:062, 2012.
- [148] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 1106:128, 2011.
- [149] J. Ellis and D.S. Hwang. Does the ‘Higgs’ have Spin Zero? *JHEP*, 1209:071, 2012.
- [150] Plot kindly provided by Ioannis Tsirikos and Fabio Maltoni, Université catholique de Louvain. Private communication, June 2013.
- [151] The CMS Collaboration. CMS, the Compact Muon Solenoid: Technical proposal. *CERN/LHCC 94-38*.
- [152] The ALICE Collaboration. ALICE: Technical proposal for a large ion collider experiment at the CERN LHC. *CERN/LHCC 95-71*.
- [153] The LHCb Collaboration. LHCb technical proposal. *CERN/LHCC 98-004*.
- [154] <http://public.web.cern.ch/public/en/research/AccelComplex-en.html>.
- [155] L. Evans and P. Bryant. LHC Machine. *JINST*, 3:S08001, 2008.
- [156] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [157] <http://www-conf.slac.stanford.edu/ssi/2012/Presentations/Zimmermann.pdf>.
- [158] E. Prebys. Bunch spacing and pileup in high energy hadron colliders. SNOWMASS 2013.
- [159] http://project-atlas-lucid.web.cern.ch/project-atlas-lucid/taskforce/inner_detector_8.pdf.

- [160] The ATLAS Collaboration. Performance of the ATLAS Trigger System in 2010. *Eur.Phys.J.*, C72:1849, 2012.
- [161] The ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1999. with updated numbers.
- [162] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). *ATL-SOFT-PUB-2007-007*, Mar 2007.
- [163] The ATLAS Collaboration. Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}=7$ TeV in the ATLAS experiment. *ATLAS-CONF-2010-069*, Jul 2010.
- [164] R Frühwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. *CMS-NOTE-2007-008*, Mar 2007.
- [165] <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/IDTRACKING/PublicPlots/ATL-COM-PHYS-2012-474/>.
- [166] The ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *Eur.Phys.J.*, C72:1909, 2012.
- [167] The ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. *ATLAS-CONF-2012-047*, May 2012.
- [168] The ATLAS Collaboration. Electron identification efficiency dependence on pile-up. *ATL-COM-PHYS-2012-260*, Mar 2012.
- [169] The ATLAS Collaboration. Expected electron performance in the ATLAS experiment. *ATL-PHYS-PUB-2011-006*, 2011.
- [170] The ATLAS Collaboration. Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector. *ATLAS-CONF-2011-063*, Apr 2011.
- [171] The ATLAS Collaboration. A muon identification and combined reconstruction procedure for the atlas detector at the lhc at cern. *IEEE Trans. Nucl. Sci.*, 51:3030–3033, Oct 2003.

- [172] S. Hassani, L. Chevalier, E. Lancon, J.F. Laporte, R. Nicolaidou, et al. A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages. *Nucl.Instrum.Meth.*, A572:77–79, 2007.
- [173] The ATLAS Collaboration. Object selection for the $H \rightarrow WW$ search with the ATLAS detector at $\sqrt{s} = 8$ TeV. *ATL-COM-PHYS-2012-861*, 2012.
- [174] W. Lampl et al. Calorimeter clustering algorithms: Description and performance. *ATL-LARG-PUB-2008-002*, Apr 2008.
- [175] The ATLAS Collaboration. Selection of jets produced in proton-proton collisions with the atlas detector using 2011 data. *ATLAS-CONF-2012-020*, Mar 2012.
- [176] The ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C73:2304, 2013.
- [177] 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.
- [178] M. Cacciari, G. P. Salam, and G. Soyez. The Anti-KT jet clustering algorithm. *JHEP*, 0804:063, 2008.
- [179] Y.L. Dokshitzer, G.D. Leder, S. Moretti, and B.R. Webber. Better jet clustering algorithms. *JHEP*, 9708:001, 1997.
- [180] The ATLAS Collaboration. Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s}=7$ TeV and measured with the ATLAS detector. Technical Report ATLAS-CONF-2010-056, 2010.
- [181] 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.*, C73:2306, 2013.
- [182] The ATLAS Collaboration. Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. *ATLAS-CONF-2011-102*, Jul 2011.
- [183] Piacquadio, G. and Weiser, C. A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS. *Journal of Physics: Conference Series*, 119(23), 2008.

- [184] The ATLAS Collaboration. Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector. *ATLAS-CONF-2012-043*, Mar 2012.
- [185] The ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS. *Eur.Phys.J.*, C72:1844, 2012.
- [186] The ATLAS Collaboration. Performance of missing transverse momentum reconstruction in atlas studied in proton-proton collisions recorded in 2012 at 8 tev. Technical Report ATLAS-CONF-2013-082, CERN, Geneva, Aug 2013.
- [187] The ATLAS Collaboration. Search for the Higgs boson in the $H \rightarrow WW^{(*)} \rightarrow \ell\nu jj$ decay channel at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS detector. *Phys.Lett.*, B718:391–410, 2012.
- [188] A.J. Barr, B. G., and C.G. Lester. Measuring the Higgs boson mass in dileptonic W-boson decays at hadron colliders. *JHEP*, 0907:072, 2009.
- [189] The ATLAS Collaboration. Measurement of the W^+W^- Production Cross Section in Proton-Proton Collisions at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS Detector. *ATLAS-CONF-2012-025*, 2012.
- [190] S. Frixione and B.R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002.
- [191] N. Kauer T. Binoth, M. Ciccolini and M. Kramer. Gluon-induced W-boson pair production at the LHC. *JHEP*, 0612:046, 2006.
- [192] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 0904:002, 2009.
- [193] P. Nason and C. Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 1002:037, 2010.
- [194] S. Dawson. Radiative corrections to Higgs boson production. *Nucl.Phys.*, B359:283–300, 1991.
- [195] A. Djouadi, M. Spira, and P. M. Zerwas. Production of Higgs bosons in proton colliders: QCD corrections. *Phys. Lett.*, B264:440–446, 1991.
- [196] R. Harlander and W. B. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys. Rev. Lett.*, 88:201801, 2002.

- [197] C. Anastasiou and K. Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nucl.Phys.*, B646:220–256, 2002.
- [198] V. Ravindran, J. Smith, and W. van Neerven. NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions. *Nucl.Phys.*, B665:325, 2003.
- [199] S. Catani, D. de Florian, M. Grazzini, and P. Nason. Soft-gluon re-summation for Higgs boson production at hadron colliders. *JHEP*, 0307:028, 2003.
- [200] D. de Florian, G. Ferrera, Ma. Grazzini, and D. Tommasini. Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC. *JHEP*, 1111:064, 2011.
- [201] M. Ciccolini, A. 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.
- [202] K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, et al. VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons. *Comput.Phys.Commun.*, 180:1661–1670, 2009.
- [203] O. Brein, A. Djouadi, and R. Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys.Lett.*, B579:149, 2004.
- [204] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini. Two-loop light fermion contribution to Higgs production and decays. *Phys. Lett.*, B595:432, 2004.
- [205] S. Actis, G. Passarino, C. Sturm, and S. Uccirati. NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders. *Phys.Lett.*, B670:12, 2008.
- [206] M. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A.D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003.
- [207] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, et al. Matching parton showers and matrix elements. *arXiv:hep-ph/0602031*, 2006.
- [208] J. Alwall et al. MadGraph/MadEvent v4: The New Web Generation. *JHEP*, 0709:028, 2007.
- [209] R.C. Gray, C. Kilic, M. Park, S. Somalwar, and S. Thomas. Backgrounds To Higgs Boson Searches from $W\gamma^* \rightarrow \ell\nu\ell(\ell)$ Asymmetric Internal Conversion. *arXiv:hep-ph/1110.1368*, 2011.

- [210] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009.
- [211] B. P. Kersevan and E. Richter-Was. The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5. *arXiv:hep-ph/0405247*, 2004.
- [212] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 0207:012, 2002.
- [213] J. Adelman et al. ATLAS offline data quality monitoring. *Journal of Physics: Conference Series*, 219(4):042018, 2010.
- [214] The ATLAS Collaboration. Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV. *ATLAS-CONF-2010-038*, 2010.
- [215] The ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011. *ATLAS-CONF-2011-116*, 2011.
- [216] A. Hoecker and others. TMVA 4: Toolkit for Multivariate Data Analysis with ROOT. *arXiv:physics/0703039v5*, 2009.
- [217] R.K. Ellis, I. Hinchliffe, M. Soldate, and J.J. van der Bij. Higgs Decay to $\tau^+\tau^-$: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC. *Nucl.Phys.*, B297:221, 1988.
- [218] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A New Mass Reconstruction Technique for Resonances Decaying to di-tau. *Nucl.Instrum.Meth.*, A654:481–489, 2011.
- [219] The ATLAS Collaboration. Calibrating the b-tag efficiency and mistag rate in 35 pb^{-1} of data with the atlas detector. *ATLAS-CONF-2011-089*, Jun 2011.
- [220] The ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *arXiv:1110.3174*, 2011.
- [221] M. Hance, D. Olivito, and H. Williams. Performance studies for e/gamma calorimeter isolation. Technical Report ATL-COM-PHYS-2011-1186, CERN, Geneva, Sep 2011.

- [222] The ATLAS Collaboration. A measurement of the muon reconstruction efficiency in 2010 ATLAS data using J/Ψ decays. ATLAS-CONF-2012-125.
- [223] The ATLAS Collaboration. Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector. ATLAS-CONF-2011-063.
- [224] The ATLAS Collaboration. Performance of the Missing Transverse Energy Reconstruction and Calibration in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2010-057.
- [225] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs cross sections: 2. Differential Distributions. *arXiv:1201.3084*, 2012.
- [226] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical Report CERN-OPEN-2012-016, New York U., Jan 2012.
- [227] Verkerke, W. and others. The RooFit Toolkit for Data Modelling. <http://roofit.sourceforge.net>.
- [228] A. Armbruster et al. Statistical analysis aspects of the search for the Standard Model Higgs boson in the $H \rightarrow WW(*) \rightarrow \ell\nu\ell\nu$ decay mode with 4.7 fb-1 of ATLAS data at ECM = 7 TeV. Technical Report ATL-COM-PHYS-2012-350, CERN, Geneva, Mar 2012.
- [229] A. Armbruster and J. Qian. *Discovery of a Higgs Boson with the ATLAS Detector*. PhD thesis, Michigan U., May 2013.
- [230] A. L. Read. Modified frequentist analysis of search results (the CL_s method). Technical Report CERN-OPEN-2000-205, <http://cdsweb.cern.ch/record/451614/files/p81.pdf>, 2000.
- [231] A. L. Read. Presentation of search results: the CL_s technique. *J. Phys. G*, 28:2693, 2002.
- [232] J. Neyman and E.S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, 231:289–337, 1933.
- [233] The CMS Collaboration. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 1306:081, 2013.

- [234] The ATLAS Collaboration. Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \rightarrow \ell\nu\ell\nu$ decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data. *Phys. Lett.*, B726(ATLAS-CONF-2013-030):88–119, Mar 2013.
- [235] The ATLAS Collaboration. Study of the spin of the new boson with up to 25 fb^{-1} of ATLAS data. *ATLAS-CONF-2013-040*, 2013.
- [236] T. Sjostrand, S. Mrenna, and P.Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [237] ATLAS JES recommendation. <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties2012>.
- [238] I. Stewart and F. Tackmann. Theory uncertainties for Higgs mass and other searches using jet bins. *Phys. Rev.*, D85:034011, 2012.
- [239] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *EPJ*, C71:1554, 2010.
- [240] The ATLAS Collaboration. Study of the spin of the Higgs-like boson in the two photon decay channel using 20.7 fb^{-1} of pp collisions collected at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector. Technical Report ATLAS-CONF-2013-029, CERN, Geneva, Mar 2013.
- [241] John C. Collins and Davison E. Soper. Angular distribution of dileptons in high-energy hadron collisions. *Phys. Rev. D*, 16:2219–2225, Oct 1977.
- [242] Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data. *ATLAS-CONF-2013-034*, Mar 2013.

Acknowledgements

At the end of this remarkable journey, I wish to thank all the people that helped me along the way.

First of all, this work would not have been possible without the support of Prof. Karl Jakobs: his guidance, help, deep knowledge, patience and sense of humour have been really precious to me. I did not ever ask a better question than the one (about BDT!) at that school in Frascati. My gratitude goes to Prof. Jakobs also for giving me the possibility to spend a long time at CERN, which was one of the best periods of my life.

In the second place, a sincere thanks to Christina Skorek, that helped me solving innumerable administrative and practical issues, always with a smile and a kind word.

In these years, I shared the office with different people, and all of them contributed to enrich me in a way, both from the scientific and personal point of view: Mirjam Fehling and Stefan Winkelmann at the beginning, Kristin Lohwasser and Inga Ludwig later, Tuan Vu Ahn and Romain Madar at CERN, Evelyn Schmidt, Susanne Kühn and Hannah Arnold in the last year in Freiburg. Somebody gave me flowers, somebody else tried to kill them, and every one of them is dear to me.

To collaborate with the Freiburg $H \rightarrow WW$ group was a great and valuable opportunity; I would like to especially thank Tuan Vu Ahn and Ralf Bernhard for their supervision and for guiding me through the good times and the hard times.

In addition to those already mentioned, a special thanks to my other friends and colleagues in Freiburg. It is impossible to quote all of them since they have been so many through the years, so I am forced to select only a few, in random order: Luisa, Riccardo, Christian, Iacopo, Karsten, Francesca, Claudia, Markus, Edoardo, Manfredi, Simone, Valerio, Vakhtang, Georges, Liv, Chris, Emily, Andi, Carsten, Urban, Stan, Xavi, Daniel, Mario, Sara, Giulia, Lorenzo, Maria, Noel and Mostafa. Thank you guys for having been there when I needed you, I hope our paths will cross again one day.

Freiburg would not have been the same without Silvia Negri, who laughed and

cried with me, taught me the rules of life and some philosophy, and was at my side no matter what, my dearest friend, my home in this city.

The ATLAS Collaboration includes many exceptional people, with whom I had the pleasure to work and from whom I could receive support and inspiration. In particular, I feel grateful to Aleandro Nisati and Pamela Ferrari, for being always there for me, whenever I needed an advice or just some comfort.

The analyses presented in this thesis would not have been possible without the other members of the BDT $H \rightarrow WW$ working group: the first group of pioneers (in particular Tatjana Lenz, Justin Keung, Tae Min Hong, Lashkar Kashif), with whom I shared the funniest working hours (and inappropriate Skype chats), and later the spin team (besides Tatjana, Pamela and Tuan, I had the pleasure to work with Doug Schouten, Christiam Schmitt and Nikos Karastathis), which stood still among the waves. I would like to thank personally also the Higgs (sub)convenors, namely Biagio di Micco, Pierre Savard, Marumi Kado and Eilam Gross, for their support, encouragement and useful critiques of this research work.

Many other colleagues deserve an acknowledgement for the useful discussions we shared, for the fruitful collaboration and for all the beautiful memories: again in random order, Fabio Cerutti, Daniel Froidevaux, Aaron Armbruster, Stefan Gadatsch, Roberto di Nardo, Andrea Gabrielli, Kirill Prokofiev, Matthias Schott, Olivier Arnaez, Corrinne Mills, Kostas Nikolopoulos, Magda Chelstowska, Joana Machado Miguéns, Johanna Bronner, Rikard Sandström, Sara Diglio, Ada Farilla, Valerio Bortolotto, Sandro de Cecco, Altan Cakir, Nicola Venturi, Richard Polifka, Jonathan Long, Gemma Wooden, Petra Van Mulders and Jane Cummings and all those that I forgot. Thank you guys, without you working and walking around in B40 and going to workshops and conferences would not have been that fun!

As everybody working at CERN knows, one central part of the day, apart from the never-ending meetings and approvals and running jobs, is constituted by the lunches at Restaurant 1, and when it comes to this I have undoubtedly to thank my MBFs: Sara Borroni, Francesca Bellini, Ludovica Aperio Bella, Andrea Agostinelli, Davide Caffarri and Francesco Guescini. And, not belonging to the original core but also very precious to me, Daniele De Gruttola, Paolo Francavilla, Ennio Salvioni, Andrea Coccaro, Anna Rolando, Andrea Beraudo and Alessio Marrani. Moreover, my life in Geneva has been enriched by the presence of the best flatmates, Inanna Boesch and Herbert Wirtz, that had to listen to many complaints and physics-related issues, evening after evening, and always offered me their advice and a good glass of wine rather than throwing me out.

Finally, even if separated by many hundreds of kilometers, my family has always been present during my Ph.D. years, as a constant and solid reference point for

me. So, thank you to my brothers, Lorenzo and Alessandro (for pretending they understood the meaning of particles, collisions and “x” bosons), and to my parents, my father Mario (for never forgetting to ask about the missing chapters, even if sometimes losing the count) and my mother Antonella (for all the rest).