

Search for the Standard Model Higgs boson in the $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$ decay modes with the ATLAS detector

A. Farilla on behalf of the ATLAS Collaboration

INFN Sezione di Roma Tre, Via Della Vasca Navale 84, 00146, Roma, IT

E-mail: ada.farilla@roma3.infn.it

Abstract. We present an updated search for the Standard Model Higgs boson in the $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$ decay modes performed by the ATLAS experiment using 4.7 fb⁻¹ of LHC data at $\sqrt{s} = 7$ TeV and 13.0 fb⁻¹ at $\sqrt{s} = 8$ TeV delivered in 2011 and 2012 respectively. No significant excess is observed. For $m_H = 125$ GeV the observed (expected) upper limit at 95% confidence level on the cross section times the branching ratio is found to be 1.8 (1.9) times the Standard Model prediction for $H \rightarrow b\bar{b}$ and 1.9 (1.2) for $H \rightarrow \tau^+\tau^-$.

1. Introduction

The search for the Standard Model (SM) Higgs boson is one of the most important endeavours of the Large Hadron Collider (LHC) [1]. The discovery of a Higgs-like boson decaying to a pair of photons or heavy vector bosons has recently been reported by the ATLAS [2] and CMS [3] collaborations. The precise determination of the nature of this boson is crucial for establishing if it is indeed the SM Higgs boson. The fermionic decays $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$ will play an important role in such determination. This note presents an update of the results reported in Ref. [4] and Ref. [5], which were based on 4.7 fb⁻¹ of LHC data at $\sqrt{s} = 7$ TeV. A detailed description of the ATLAS detector can be found in [6].

2. $H \rightarrow b\bar{b}$

The decay to b -quarks, for a SM Higgs boson with a mass of $m_H = 125$ GeV, is expected to be the dominant decay mode ($\text{BR}(H \rightarrow b\bar{b}) \simeq 58\%$). Therefore, an observation in this channel is crucial in order to provide a direct constraint on the dominant decay mode. The three processes with the largest SM Higgs boson production cross-section at the LHC are gluon-gluon fusion (ggF), W or Z boson fusion (VBF) and associated production with a W or a Z boson (VH). The cross-sections for each process at 7 TeV and 8 TeV, for $m_H = 125$ GeV, are summarized in Table 1.

To reduce the huge background from multijet production and to have a clean signature for triggering, the search for $H \rightarrow b\bar{b}$ has been performed with the Higgs boson produced in association with a W or Z boson in the decay channels $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ (0-lepton channel), $WH \rightarrow l\nu b\bar{b}$ (1-lepton channel) and $ZH \rightarrow l^+l^-b\bar{b}$ (2-lepton channel), where l refers to either an electron or a muon. More details on the analysis can be found in Ref. [7]. Monte Carlo (MC) event samples produced with the full ATLAS detector simulation [8] based on the GEANT4 [9]



Table 1. Cross sections for Higgs production processes for pp collisions at $\sqrt{s}=7$ TeV and $\sqrt{s}=8$ TeV at LHC, for $m_H = 125$ GeV [10], [11].

\sqrt{s}	ggF	VBF	VH
7 TeV	15.3 pb	1.22 pb	0.89 pb
8 TeV	19.5 pb	1.57 pb	1.09 pb

program, corrected for all known detector effects, are used to model the Higgs boson signal and most backgrounds. The total production cross sections and associated uncertainties, computed at next-to-leading order (NLO), are taken from Ref. [10] with NLO corrections as a function of the transverse momentum of the vector boson, p_T^V [12], [13]. The decay branching ratios are calculated with HDECAY [14].

2.1. Event selection

Three categories of leptons, as defined in lepton reconstruction, are used and are denoted in increasing order of purity as loose, medium and tight leptons. Events containing no loose leptons are assigned to the 0-lepton channel, those with one tight and no additional loose leptons are assigned to the 1-lepton channel. Events with one medium and one additional loose lepton of the same flavour but opposite charge are assigned to the 2-lepton channel. Jets are reconstructed using the anti- k_t algorithm [15] with a radius parameter $R = 0.4$ and are required to have $p_T > 20$ GeV and $|\eta| < 4.5$. The two jets used to reconstruct the Higgs boson candidate are required to satisfy $|\eta| < 2.5$ and the leading jet is required to have $p_T > 45$ GeV.

b -quark jets are identified using the MV1 b -tagging algorithm [16], [17], tuned to provide an approximate efficiency of 70% for b -jets and rejection factors of 150 and 5 for light and c -quark jets respectively. Selected events are required to have exactly two b -tagged jets, which are used to reconstruct the mass of the Higgs boson candidate, $m_{b\bar{b}}$. In 3-jets events, the leading two jets are required to be b -tagged.

The missing transverse energy, E_T^{miss} , is measured as the negative vector sum of the transverse momenta associated with cluster energies in the calorimeters with $|\eta| < 4.9$. A track-based missing transverse momentum, p_T^{miss} , is calculated as the negative vector sum of the transverse momenta of tracks associated to the primary vertex. The transverse momentum of the vector boson, p_T^V , where V denotes the W or the Z boson, is reconstructed as the E_T^{miss} in the 0-lepton channel, as the vector sum of the lepton and the E_T^{miss} in the 1-lepton channel and the vector sum of the two leptons in the 2-lepton channel. In order to increase the analysis sensitivity, the events are split in different categories, for a total of 16 categories. Events in the 0-lepton channel are split into six categories depending on whether they contain two or three jets and on the amount of missing transverse energy. Events with one or two leptons are divided into five categories of p_T^V . Topological cuts, optimized for each category, are used to enhance the signal and reject background.

The multijet background in the 0-lepton channel is suppressed by cuts on E_T^{miss} , p_T^{miss} and their directions with respect to each other and the reconstructed jets. In the 1-lepton channel, requirements are made on E_T^{miss} and the transverse mass on the W boson, m_W , to select events consistent with the presence of a W boson. The background from top production is reduced in the 2-lepton channel by cuts on the di-lepton invariant mass, $83 < m_{ll} < 99$ GeV, and $E_T^{miss} < 60$ GeV.

2.2. Background Estimation

The dominant backgrounds depend on the event category: the 2-lepton channel is dominated by Z +jets events, the 1-lepton channel is dominated by W +jets and top (both $t\bar{t}$ and single top) while all three types of backgrounds (Z +jets, W +jets and top) contribute to the 0-lepton channel. Background from multijet and diboson production (WW , ZZ , WZ) is also present. Most background shapes are taken from simulation and normalised using control regions in data except for multijet, which is estimated entirely from data, and the diboson production for which both normalisation and shape are taken from simulation.

Control regions are used to normalise the top background from data in the 1- and 2-lepton channels. In the 1-lepton channel the control region is defined by requiring an additional jet and for the 2-lepton channel by inverting the requirements on m_{ll} and E_T^{miss} . Control distributions for the V +jets background are categorised by the number of b -tagged jets. The 0-tag and 1-tag control regions contain exactly 0 or 1 b -tagged jets.

The normalisations of the $V + c$ -jets and $V + \text{light-jet}$ backgrounds are determined by a maximum likelihood fit to control and signal regions in the 1- and 2-lepton channels. Events containing 0, 1 and 2 b -tagged jet and the top control regions of each channel are used. This flavour fit exploits the fact that the b -tagging algorithm has very different efficiencies for b , c and light jets, such that the flavour composition varies significantly with the number of b -tagged jets. The normalisations of V +light, V + c , V + b and top are allowed to float in the fit. The distributions of the signal regions for each of the three channels, after having been normalised by the result of the flavour fit, are shown in Figure 1 for the $\sqrt{s} = 8$ TeV data.

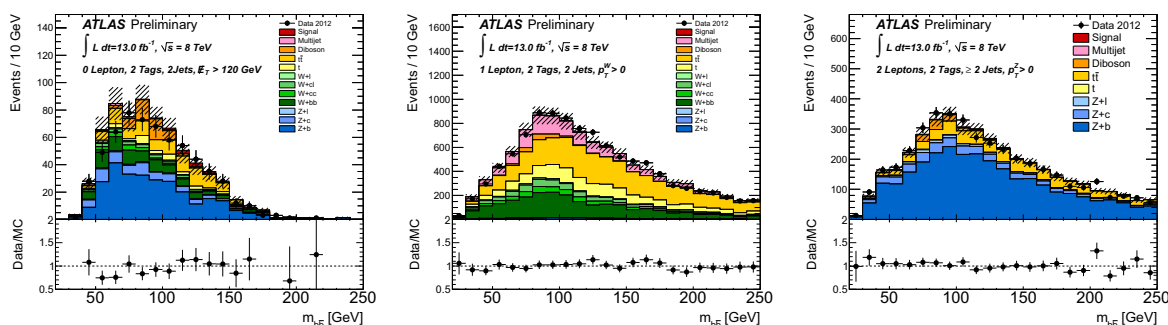


Figure 1. The $m_{b\bar{b}}$ signal distribution integrated over bins of p_T^V : (left) 0-lepton channel, (middle) 1-lepton channel, (right) 2-lepton channel. The error bands indicate the size of the combined statistical and systematic uncertainties before the profile likelihood fit [7].

Systematic uncertainties are determined on all experimental and theoretical sources. The most significant are on the b -tagging efficiency ($\simeq 7\%$), the calibration of the jet energy (7–20%) and the statistical uncertainty on the MC (4–8%).

2.3. Results

The statistical analysis of the data employs a binned profile likelihood function constructed as a product of the likelihood terms for each category. The likelihood in each category is a product over bins in the distribution of $m_{b\bar{b}}$. A detailed description of the statistical analysis of LHC data can be found in Ref. [18], [19].

The event categories that enter the profile likelihood fit are the 16 individual 2 b -tag signal regions and the top control regions for the 1- and 2-lepton channels. The signal strength parameter, μ , multiplying the expected signal yield in each bin, is the parameter of interest

in the fit procedure. The μ parameter is defined as the ratio of the measured cross section normalised to the Standard Model Higgs boson production cross section times the branching ratio for $H \rightarrow b\bar{b}$. The dependence of the signal and background predictions on the systematic uncertainties is described by nuisance parameters. The normalisation of the top, $Z + b$ -jet and $W + b$ -jet backgrounds floats without constraint in the fit, while the other backgrounds have their uncertainties applied as a constraint.

Diboson production (WZ or ZZ) with a Z boson decaying to a pair of b -quarks, has a signature very similar to the Higgs associated production but with lower p_T^Z and $m_{b\bar{b}}$ and a cross section 5 times larger. Therefore a separate fit was made to validate the analysis procedure. In the fit the normalisation of the diboson contribution is allowed to vary with a multiplicative scale μ_D with respect to the SM expectation.

Figure 2 (top) shows the distribution in data after subtracting all backgrounds except the diboson and SM Higgs contribution. An excess in the data compared to the background is observed at the expected mass for the diboson signal. We measure $\mu_D = 1.05 \pm 0.32$ corresponding to a significance of 4.0σ and agrees with the SM expectation of $\mu_D = 1$.

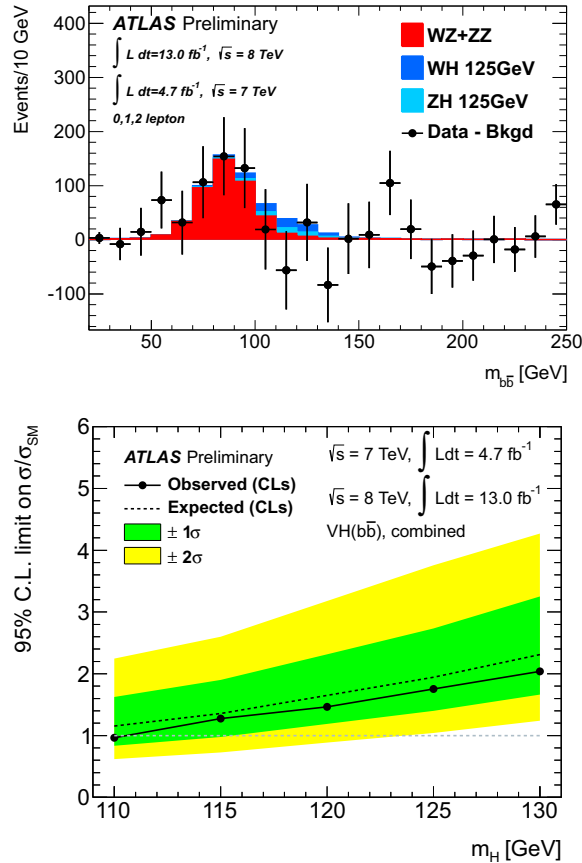


Figure 2. Top plot: the $m_{b\bar{b}}$ distribution after subtraction of all backgrounds except diboson processes (WZ , ZZ) and WH , ZH production. The MC backgrounds are normalised according to the results of the global fit. Only statistical uncertainties are shown. Bottom plot: expected (dashed) and observed (solid) 95% CL on the normalised signal strength as a function of m_H for the combination of the $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$ dataset [7].

The profile likelihood fit with the Higgs boson signal strength floating is performed with

the diboson contribution fixed to its SM expectation. Figure 2 (bottom) shows the results for the 95% confidence level (CL) exclusion limits, evaluated using the CL_S method [18], on the Higgs boson production cross section in the mass range 110-130 GeV. The observed limit for the combination of the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV datasets is 1.8 times the SM expectation with an expected limit of 1.9 times the SM. The probability, p_0 , that a fluctuation in the background can mimic the presence of a SM Higgs boson is studied by calculating the local p_0 value with respect to the background-only hypothesis. For $m_H = 125$ GeV the combination of both datasets yields an observed (expected) p_0 value of 0.64 (0.15). We measure $\mu = -2.7 \pm 1.1(stat) \pm 1.1(syst)$, $\mu = 1.0 \pm 0.9(stat) \pm 1.1(syst)$ and $\mu = -0.4 \pm 0.7(stat) \pm 0.8(syst)$ in the $\sqrt{s} = 7$ TeV, $\sqrt{s} = 8$ TeV and combined datasets, respectively.

3. $H \rightarrow \tau^+\tau^-$

The coupling of the new Higgs-like boson to leptons has not yet been established. In the SM, the Higgs coupling to tau leptons is significant ($BR(H \rightarrow \tau^+\tau^-) \simeq 6.3\%$), due to a relatively high mass of the tau. The search for $H \rightarrow \tau^+\tau^-$ is performed across the three main Higgs production processes, shown in Table 1, and this will ultimately lead to the measurement of a number of important Higgs couplings. Both leptonic τ_{lep} and hadronic τ_{had} decays of the tau leptons have been exploited, therefore the search has been performed in the following three Higgs decay channels: $H \rightarrow \tau_{lep}\tau_{lep}$, $H \rightarrow \tau_{lep}\tau_{had}$ and $H \rightarrow \tau_{had}\tau_{had}$. More details on the analysis can be found in Ref. [20].

3.1. Event selection

Signal events are selected by requiring exactly two opposite-charge tau decays products which are two light leptons l (l =electron or muon) for the $H \rightarrow \tau_{lep}\tau_{lep}$ candidate, one light lepton and a $\tau_{had-vis}$ for $H \rightarrow \tau_{lep}\tau_{had}$ candidate and two $\tau_{had-vis}$'s for $H \rightarrow \tau_{had}\tau_{had}$ candidate, where $\tau_{had-vis}$ denotes the visible component of hadronically decaying tau lepton. To suppress backgrounds produced via purely QCD processes, signal events are selected by requiring large E_T^{miss} to account for the neutrinos produced in tau decays. The invariant mass of the two taus is the final discriminating observable in all the three channels and is reconstructed by means of the Missing Mass Calculator (MMC) [21]. This technique provides a reconstruction of event kinematics in the $\tau\tau$ final state with efficiency $> 99\%$ and 13–20% resolution in $m_{\tau\tau}$, depending on the event topology and final state (better resolution is obtained for events with high- p_T jets and boosted topology).

All three channels analyzed exploit both the VBF and Boosted categories. Events in the VBF category are selected requiring two jets of high invariant mass and high angular separation in η . Events in the Boosted category are selected requiring large transverse momenta of the $l+l+E_T^{miss}$ system (for the $H \rightarrow \tau_{lep}\tau_{lep}$ case), of $l+\tau_{had-vis}+E_T^{miss}$ system (for the $H \rightarrow \tau_{lep}\tau_{had}$ case) or a leading jet in the event (for $H \rightarrow \tau_{had}\tau_{had}$ case). The following selections and categorizations are applied for each decay channel to reduce backgrounds and to focus on a particular production process.

$H \rightarrow \tau_{lep}\tau_{lep}$: To search for a Higgs within the range $100 \text{ GeV} < m_H < 150 \text{ GeV}$ the dilepton invariant mass is required to be in the range $30 \text{ GeV} < m_{ll} < 100 \text{ GeV}$ for the $e\mu$ channel, whereas for the ee and $\mu\mu$ channel $30 \text{ GeV} < m_{ll} < 75 \text{ GeV}$ is required to reduce the contamination from $Z/\gamma^* \rightarrow ll$. Events are rejected if there is any jet, with $p_T > 25 \text{ GeV}$ and $|\eta| < 2.5$, identified as a b -quark jet. Events are also required to have $H_T^{miss} > 40 \text{ GeV}$, which is the missing transverse momentum defined by high- p_T objects, the di-lepton azimuthal opening angle $0.5 < \Delta\phi_{ll} < 2.5$ and $0.1 < x_{1,2} < 1.0$, where x_1 and x_2 denote the visible momentum fraction of the tau leptons carried by the leading and sub-leading light leptons, respectively. For the event selection the momentum of the tau lepton is estimated using the

collinear approximation [22]. Five mutually exclusive categories are then defined, depending on the jet multiplicity and kinematics: 2-jet VBF, Boosted, 2-jet VH, 1-jet and 0-jet.

$H \rightarrow \tau_{\text{lep}} \tau_{\text{had}}$: To suppress W +jets and non-resonant backgrounds, the transverse mass, m_T , of lepton and E_T^{miss} system, the $\Sigma\Delta\phi$ variable, defined as $\Delta\phi_{l, E_T^{\text{miss}}} + \Delta\phi_{\tau, E_T^{\text{miss}}}$ and the $\Delta(\Delta R)$ variable, defined as the difference between the angular separation $\Delta R_{l\tau_{\text{had-vis}}}$ and its expectation value for the Higgs decay, are used. Six exclusive categories are defined for this channel: VBF, Boosted, 1-jet ($e\tau_{\text{had-vis}}$, $\mu\tau_{\text{had-vis}}$) and 0-jet ($e\tau_{\text{had-vis}}$, $\mu\tau_{\text{had-vis}}$).

$H \rightarrow \tau_{\text{had}} \tau_{\text{had}}$: Two $\tau_{\text{had-vis}}$ candidates are required with $0.8 < \Delta R < 2.8$ and $\Delta\eta < 1.5$. Two exclusive categories are defined for this channel: VBF and Boosted.

3.2. Background Estimation

The background composition and normalisation have been determined using both data-driven and simulation based methods. The main background for this search is due to the largely irreducible $Z/\gamma^* \rightarrow \tau\tau$. Other background processes are multijet, W +jets, Z +jets, top (both $t\bar{t}$ and single top) and diboson production (WW , ZZ , WZ). Since it is not possible to select a signal-free $Z/\gamma^* \rightarrow \tau\tau$ sample directly from data, this background has been modeled in a partly data-driven way by selecting a control sample of $Z/\gamma^* \rightarrow \mu\mu$ data, where the expected signal contamination is negligible. In this sample the muon tracks and associated calorimeter cells are replaced by simulated tau leptons (embedding method). In this way only the tau decays and the corresponding detector response are taken from the simulation, while jets, underlying event and pile-up effects are obtained from data, thus resulting in a very good agreement between the MMC mass distribution obtained with the embedded method and the one obtained with the simulation [20].

The contributions of multijet and W +jets backgrounds are estimated using data-driven methods, either using template fitting of kinematics distributions or derived from a control region defined by inverting the opposite charge requirement. For the top background the normalisation is obtained from control regions, defined by inverting the m_T cut and requiring at least one b -tagged jet, whereas kinematic shapes are obtained by simulation. A similar approach, based on selected control regions, is used for the Z +jets and diboson backgrounds.

Systematic uncertainties are determined on all experimental and theoretical sources. The dominant systematic uncertainties are those on the theoretical predictions for the signal production cross-sections and decay branching ratios (3-28 %), on tau energy scale (2-15%), on the tau identification (1-10%), as well as those on the jet energy scale (1-9%) [20] and on the measurement of integrated luminosity (3.6-3.9%).

3.3. Results

The most sensitive categories are the Boosted and VBF categories for all three channels. Figure 3 and Figure 4 show distributions for the final discriminant variables MMC $m_{\tau\tau}$ of the selected events. No significant excess is observed in the data compared to the SM background-only expectation in any of the channels. The statistical analysis of the data employs a binned likelihood function constructed as a product of the likelihood terms for each category. The likelihood in each category is a product over bins in the distributions of $m_{\tau\tau}$.

Figure 5 (left) shows the expected and observed cross section limits for a combination of all three channels as a function of the Higgs boson mass at the 95% CL. The combined expected limit varies between 1.2 and 3.4 times the SM cross section times branching ratio for the mass range 100-150 GeV. The corresponding observed limits are in the range between 1.9 and 3.3 times the SM prediction. For $m_H = 125$ GeV specifically, the expected limit is 1.2 and the observed 1.9. The best fit value for the signal strength is $\mu = 0.7 \pm 0.7$. The combination of both datasets yields an observed (expected) p_0 value of 0.135 (0.047) for $m_H = 125$ GeV. The contributions from the three different production modes in the $H \rightarrow \tau^+ \tau^-$ channel have

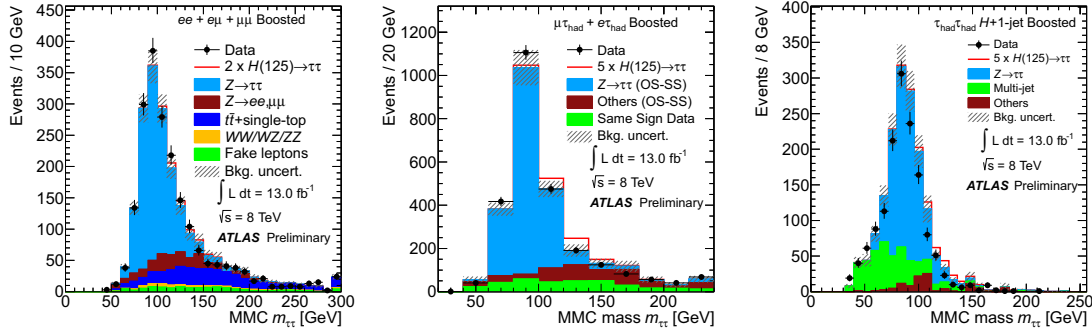


Figure 3. Reconstructed $m_{\tau\tau}$ of the selected events in the 8 TeV sample for the Boosted category: (left) $\tau_{lep}\tau_{lep}$, (middle) $\tau_{lep}\tau_{had}$, (right) $\tau_{had}\tau_{had}$ [20].

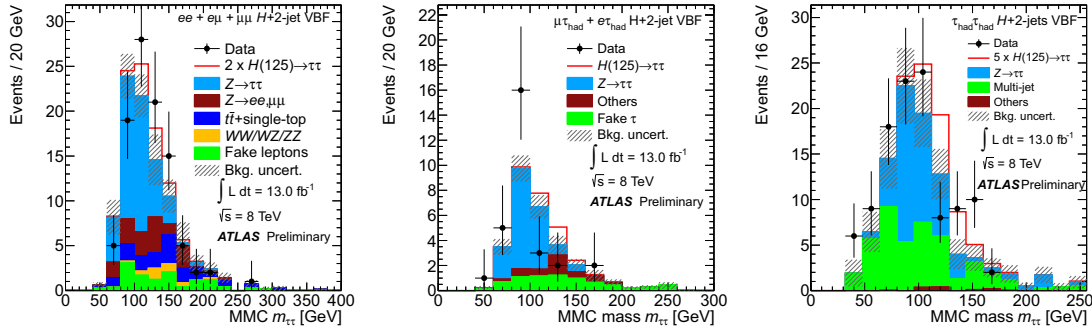


Figure 4. Reconstructed $m_{\tau\tau}$ of the selected events in the 8 TeV sample for the VBF category: (left) $\tau_{lep}\tau_{lep}$, (middle) $\tau_{lep}\tau_{had}$, (right) $\tau_{had}\tau_{had}$ [20].

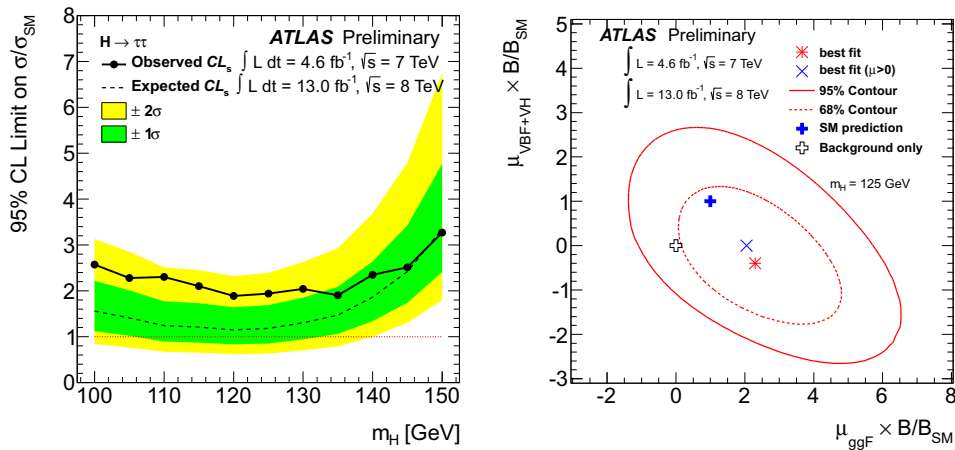


Figure 5. Left plot: expected (dashed) and observed (solid) 95% CL on the normalised signal strength as a function of m_H for the combination of the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV dataset. Right plot: likelihood contours for $H \rightarrow \tau^+\tau^-$ in the $\mu_{ggF} \times B/B_{SM}$, $\mu_{VBF+VH} \times B/B_{SM}$ plane are shown for the 68% and 95% CL by dashed and solid lines, respectively. The best fit to the data is shown for the case when both parameters have been constrained to be non-negative and for the unconstrained case, as explained in the legend of the plot [20].

been studied in order to assess any tension between data and the signal production cross section as predicted by the SM. The VBF and VH production modes have been grouped together as they both scale with the WWH or ZZH coupling in the SM, denoted as $\mu_{\text{VBF+VH}}$ and the strength parameter for the remaining process is denoted as μ_{ggF} . Two-dimensional contours in the $\mu_{\text{VBF+VH}} \times B/B_{\text{SM}}$ and $\mu_{\text{ggF}} \times B/B_{\text{SM}}$ plane are shown in Figure 5 (right), where B and B_{SM} are the measured and SM branching ratio for $H \rightarrow \tau^+\tau^-$. The best fit values are $\mu_{\text{VBF+VH}} \times B/B_{\text{SM}} = -0.4$ and $\mu_{\text{ggF}} \times B/B_{\text{SM}} = 2.4$ and the result is consistent with both the SM Higgs and SM background-only hypothesis within the 95% CL.

3.4. Conclusions

A search for the Standard Model Higgs boson decaying in the $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$ channels has been performed by the ATLAS experiment at the LHC using 4.7fb^{-1} of pp collision data recorded at $\sqrt{s} = 7$ TeV and 13.0fb^{-1} recorded at $\sqrt{s} = 8$ TeV. No significant excess is observed. For $m_H = 125$ GeV the observed (expected) upper limit at 95% CL on the cross section times the branching ratio is found to be 1.8 (1.9) times the Standard Model prediction for $H \rightarrow b\bar{b}$ and 1.9 (1.2) for $H \rightarrow \tau^+\tau^-$ for the combination of the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV datasets.

References

- [1] L. Evans and P. Bryant (Editors), JINST **3** (2008) S08001
- [2] ATLAS Collaboration, Phys. Lett. B **716** (2012) 1
- [3] CMS Collaboration, Phys. Lett. B **716** (2012) 30
- [4] ATLAS Collaboration, Phys. Lett. B **718** (2012) 369
- [5] ATLAS Collaboration, JHEP **09** (2012) 070
- [6] ATLAS Collaboration, JINST **3** (2008) S08003
- [7] ATLAS Collaboration, ATLAS-CONF-2012-161 (2012), <https://cds.cern.ch/record/1493625>
- [8] ATLAS Collaboration, Eur. Phys. J. **C70**, 823 (2010)
- [9] S. Agostinelli et al. (GEANT4), Nucl. Instr. and Meth. **A506** (2003) 250
- [10] S. Dittmaier et al., LHC Higgs Cross Section Working Group, CERN-2011-002 (CERN, Geneva, 2011)
- [11] LHC Higgs Cross Section Working Group, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>
- [12] S. Dittmaier et al., LHC Higgs Cross Section Working Group, CERN-2012-002 (CERN, Geneva, 2012)
- [13] G. Ferrera, M. Grazzini and F. Tramontano, Phys. Rev. Lett. **107** (2011) 152003
- [14] A. Djouadi, J. Kalinowski and M. Spira, Comp. Phys. Commun. **108** (1998) 56
- [15] M. Cacciari, G.P. Salam and G. Soyez, JHEP **04** (2008) 063
- [16] ATLAS Collaboration, ATLAS-CONF-2012-040 (2012), <https://cds.cern.ch/record/1435194>
- [17] ATLAS Collaboration, ATLAS-CONF-2012-043 (2012), <https://cds.cern.ch/record/1435197>
- [18] A. L. Read, J. Phys. G **28** (2002) 2693
- [19] G. Cowan, K. Cranmer, E. Gross and O. Vitells, Eur. Phys. J. C **71** (2011) 1554
- [20] ATLAS Collaboration, ATLAS-CONF-2012-160 (2012), <https://cds.cern.ch/record/1493624>
- [21] A. Elagin, P. Murat, A. Pranko and A. Safonov, Nucl. Instr. and Meth. **A654** (2011) 481
- [22] R. K. Ellis, I. Hinchliffe, M. Soldate and J. J. Van der Bij, Nucl. Phys. B **297** (1988) 221