

Search for the Standard Model Higgs boson produced in association with top quarks with the ATLAS detector

Yang Qin^{1,a} On behalf of the ATLAS Collaboration

¹University of Manchester

Abstract. After the discovery of a new particle consistent with the Higgs boson in 2012, the direct observation and measurement of the coupling of this particle to top quarks has become increasingly important. In this talk, a review of the latest ATLAS results on the search for the Higgs boson produced in association with top quarks, $t\bar{t}H$, is presented. In particular, three analyses aiming at different $t\bar{t}H$ final states are reviewed. These include a search with the Higgs boson decaying to $b\bar{b}$, a search in multileptonic final states and a search with $H \rightarrow \gamma\gamma$. The results from these analyses are combined, giving a $t\bar{t}H$ signal strength of 1.8 ± 0.8 and excluding a $t\bar{t}H$ signal 3.2 larger than the SM expectation at 95% confidence level for a SM Higgs boson with a mass of 125.36 GeV.

1 Introduction

The discovery of a new particle in consistency with the Standard Model (SM) [1–3] Higgs boson [4–7] was reported by the ATLAS [8] and the CMS [9] collaborations in July 2012. Ever since the discovery, the properties of the scalar boson have been extensively studied. One of the most important aspects is to measure the coupling of the Higgs boson to fermions, i.e. the Yukawa coupling. As the top quark is the heaviest fermion, it gives the strongest Yukawa coupling with a coupling constant at the order of one. The measurement of the top-Higgs Yukawa coupling, y_t , is of particular importance for both understanding the SM electroweak symmetry breaking and testing Beyond Standard Model (BSM) theories. With a value close to unity, y_t is also expected to provide insight into the scale of new physics [10].

The $t\bar{t}H$ process, shown in Figure 1(a) and 1(b), provides a direct probe to y_t , while other processes are only sensitive through loop effects, for example the gluon-gluon fusion shown in Figure 1(c). However, the small production rate of the $t\bar{t}H$ process makes the relevant search very challenging.

Another process that is sensitive to the top Yukawa coupling is the Higgs production associated with a single top quark, tH ¹. As an example, the Feynman diagrams for the t -channel tH production are shown in Figure 2, where the Higgs boson can be emitted from either the top quark or the W boson. The SM predicts an almost perfect destructive interference between the two processes shown in Figure 2(a) and Figure 2(b), resulting in a suppressed total tH cross-section. The tH production therefore is sensitive not only to the magnitude of y_t , but also to the relative sign of y_t and g_{HWW} , which represents the coupling between the Higgs boson and the W boson. In BSM theories, y_t and g_{HWW} can have an anomalous correlation, leading to a boosted inclusive tH production cross-section.

^ae-mail: yang.qin@cern.ch

¹For simplicity, tH also refers to $t\bar{t}H$ throughout these proceedings.

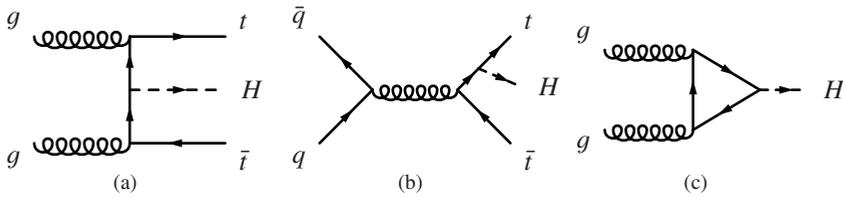


Figure 1. The Feynman diagrams of (a,b) tree level $t\bar{t}H$ processes demonstrating a direct coupling of the top quark and the Higgs boson and (c) the Higgs production in gluon-gluon fusion with a top-Higgs coupling at the loop level.

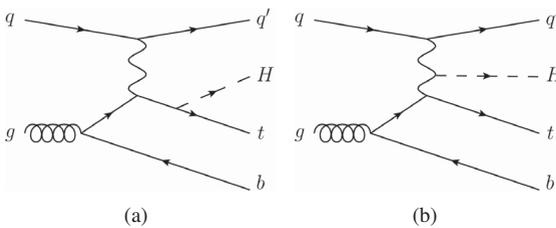


Figure 2. The Feynman diagrams of the t -channel tH production, where the Higgs boson can be emitted from either (a) the top quark or (b) the W boson [11].

The $t\bar{t}H$ process produces complex final states, since both the $t\bar{t}$ pair and the Higgs boson can undergo multiple possible decays, which in combination provides a variety of final states. In these proceedings, three analyses recently performed by the ATLAS experiment probing different $t\bar{t}H$ final states are reviewed. All three analyses discussed in these proceedings were designed to be orthogonal, and contributed to the combined Higgs boson property measurement described in Ref. [12].

2 Search for $t\bar{t}H$ with $H \rightarrow b\bar{b}$ at $\sqrt{s} = 8$ TeV

The dominant decay mode for a Higgs boson with a mass around 125 GeV is $H \rightarrow b\bar{b}$, with a branching ratio of approximately 58%. However, it usually suffers from a large background associated with heavy flavour jets, especially the $t\bar{t}+b\bar{b}$ background which has the same final states as the $t\bar{t}H$ signal. In this analysis, two channels based on the decay of the $t\bar{t}$ pair are included, the single lepton channel and the dilepton channel. In the single lepton channel one of the top quarks decays leptonically, $t \rightarrow Wb \rightarrow l\nu b$, while the other decays hadronically, $t \rightarrow Wb \rightarrow q\bar{q}'b$ ². In the dilepton channel, both the top and anti-top quark decay leptonically. The lepton in both cases exclusively refers to an electron or muon given the complexity introduced by hadronically decaying taus.

Since this analysis is heavily dominated by background, the selected events are categorised into different regions defined by the number of jets (j) and number of b -tagged jets (b) in order to improve the search sensitivity. In the single lepton channel, nine independent regions are considered: (4j, 2b), (4j, 3b), (4j, 4b), (5j, 2b), (5j, 3b), (5j, $\geq 4b$), (6j, 2b), (6j, 3b) and (6j, $\geq 4b$). In the dilepton channel, six regions are considered: (2j, 2b), (3j, 2b), (3j, 3b), ($\geq 4j$, 2b), ($\geq 4j$, 3b) and ($\geq 4j$, $\geq 4b$). As is shown in Figure 3, the signal-to-background ratios S/B and the sensitivities, represented by S/\sqrt{B} , are very low in all regions, where S and B refer to the expected signal for a Higgs boson with $m_H = 125$ GeV and the total background respectively. The regions with $S/B > 1\%$ and $S/\sqrt{B} > 0.3$ are referred to as “signal-rich regions”, The remaining regions are referred to as “signal-depleted regions”, and are used to constrain systematic uncertainties on the background processes.

²Charge conjugate states are implied here.

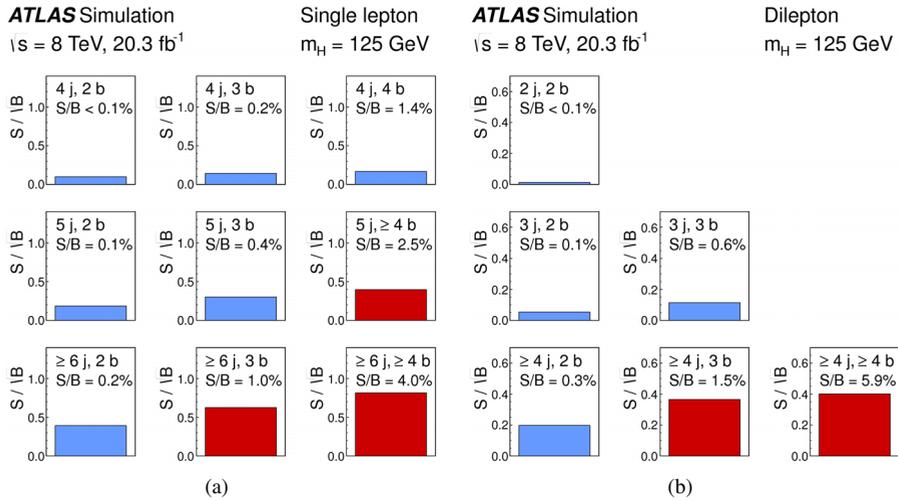


Figure 3. The S/B and S/\sqrt{B} in each region considered in (a) the single lepton channel and (b) the dilepton channel. The “signal-rich regions” are shown in red, and the “signal-depleted regions” in blue [13].

The dominant background in this analysis is $t\bar{t}$ +jets, which has been decomposed into $t\bar{t}$ +light, $t\bar{t}+c\bar{c}$ and $t\bar{t}+b\bar{b}$. In the signal-rich regions the largest component is the $t\bar{t}+b\bar{b}$ background. Therefore the understanding of this background is vital to this analysis. The modelling of the $t\bar{t}+b\bar{b}$ background is done using an inclusive $t\bar{t}$ sample generated with POWHEG+PYTHIA6 [14–17]. However, since POWHEG+PYTHIA only models $t\bar{t}+b\bar{b}$ with parton shower, it is then reweighted to a fully matched next-to-leading order (NLO) prediction with massive b quarks [18], which generates $t\bar{t}+b\bar{b}$ processes at NLO following the four-flavour scheme using the SHERPA+OPENLOOPS framework [19, 20].

A simultaneous profiled likelihood fit is performed across all regions. This allows the impact of the systematic uncertainties on the search sensitivity to be reduced by exploiting the constraining power provided by the high statistics in the background-dominated regions. Given the extremely small signal expectation, the analysis uses Neural Networks (NN) to exploit the separation between signal and background in each signal-rich region. Various categories of variables are considered in the construction of the NN, including object kinematics, global event variables, event shape variables and object pair properties. In addition to the kinematic variables, two variables calculated by the Matrix Element Method (MEM) [21] are used to construct the NN in the single lepton channel, which carry significant separation power into the NN discriminants. These two variables are the Neyman-Pearson likelihood ratio and the logarithm of the summed signal likelihoods. A maximum of twelve (ten) variables are used to build the NN in the single lepton (dilepton) channel. An additional NN is built in the (3j, 3b) region in the dilepton channel despite of its low S/\sqrt{B} , because its relatively high S/B ratio can further improve the sensitivity. Furthermore, the (5j, 3b) region in the single lepton channel has a significant amount of both $t\bar{t}$ +light and $t\bar{t}$ +HF (HF for “heavy flavour”) background, where $t\bar{t}$ +HF refers inclusively to $t\bar{t}+c\bar{c}$ and $t\bar{t}+b\bar{b}$. A dedicated NN is built in this region to separate the $t\bar{t}$ +light and $t\bar{t}$ +HF background. This provides strong constraints on the normalisation of the $t\bar{t}$ +HF background. In other background-dominated regions where the sensitivity is low, H_T^{had} and H_T are used as the discriminant variables respectively in the single lepton channel and the dilepton channel,

where H_T^{had} is defined as the scalar sum of the jet p_T ³, and H_T is the scalar sum of the jet and lepton p_T .

The post-fit NN distribution in the most sensitive region in the two channels are shown in Figure 4. A reasonable agreement between data and the prediction as well as the separation power provided by the NN between the total background and the $t\bar{t}H$ signal are presented in both channels. The results of the statistical analysis are summarised in Figure 5. The best-fit signal strength is μ ($m_H = 125$ GeV) = 1.5 ± 1.1 , shown in Figure 5(a). The observed (expected) significance of the signal is 1.4 (1.1) standard deviations, corresponding to an observed (expected) p -value of 8 (15) %. Figure 5(b) shows the observed upper limits on the $t\bar{t}H$ signal strength and those expected with and without assuming a SM Higgs boson with $m_H = 125$ GeV. Using the CL_s method [22–24], an observed (expected) 95% confidence level (CL) upper limit on the signal strength of 3.4 (2.2) is obtained. The expected upper limit on the signal strength with the presence of a SM Higgs boson is found to be 3.1.

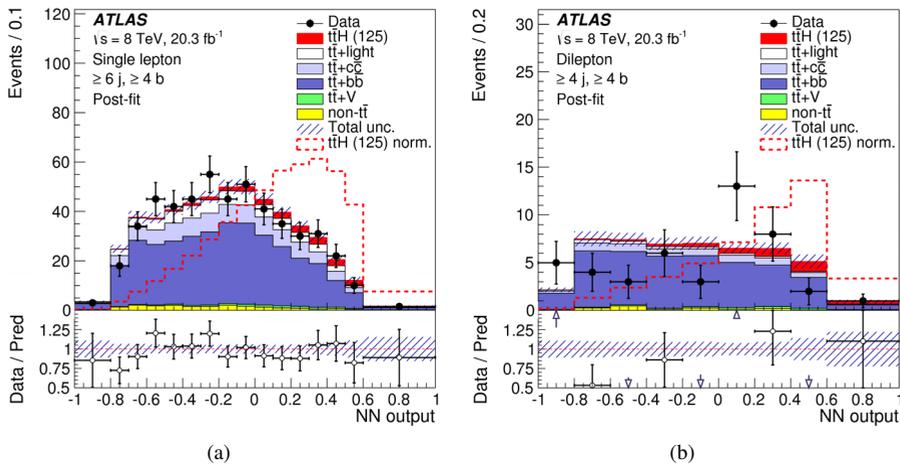


Figure 4. The post-fit NN distribution in (a) the ($6j, \geq 4b$) region in the single lepton channel and (b) the ($\geq 4j, \geq 4b$) region in the dilepton channel [13].

3 Search for $t\bar{t}H$ with multilepton final states at $\sqrt{s} = 8$ TeV

This analysis is primarily sensitive to the $H \rightarrow WW^*$, $H \rightarrow ZZ^*$ and $H \rightarrow \tau\tau$ decays, with one or two leptons from the decay of the $t\bar{t}$ pair. A total of five multileptonic final states are considered, with different sensitivities on the three Higgs decay modes. These include the final states with two same-charge-sign (SS) light leptons (e or μ) with no additional hadronically decaying τ lepton (denoted as τ_{had}), three light leptons, two SS light leptons with one τ_{had} , four light leptons and one lepton with two τ_{had} candidates. The five final states with their corresponding fraction of expected $t\bar{t}H$ signal yields from different Higgs decay modes are listed in Table 1.

³ ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z -axis coinciding with the axis of the beam pipe. The x -axis points from the IP to the centre of the LHC ring, and the y -axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln[\tan(\theta/2)]$. Transverse momentum and energy are defined as $p_T = p \sin \theta$ and $E_T = E \sin \theta$, respectively.

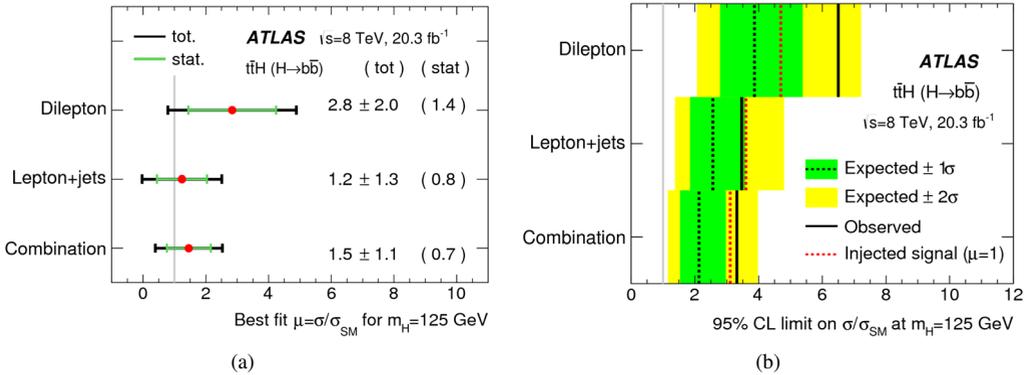


Figure 5. The (a) fitted values of the $t\bar{t}H$ signal strength and their uncertainties and (b) 95% CL upper limits on $\sigma(t\bar{t}H)$ relative to the SM prediction, σ/σ_{SM} , for the individual channels and their combination [13].

Table 1. Fraction of expected $t\bar{t}H$ signal from different Higgs decay modes in each of the five final states included in the analysis [25].

Category	Higgs boson decay mode			
	WW^*	$\tau\tau$	ZZ^*	Other
$2\ell 0\tau_{had}$	80%	15%	3%	2%
3ℓ	74%	15%	7%	4%
$2\ell 1\tau_{had}$	35%	62%	2%	1%
4ℓ	69%	14%	14%	4%
$1\ell 2\tau_{had}$	4%	93%	0%	3%

The observed and expected number of jets spectrum in each signal region of the five categories are shown Figure 6. In general, a reasonable agreement between data and prediction is presented, whereas in $2\ell 1\tau_{had}$ (6(c)) and 4ℓ (6(d)) categories only one event is observed. As can be seen from the histograms, the leading background is from vector boson production in association with top quark pairs, $t\bar{t}V$, which includes $t\bar{t}Z$ and $t\bar{t}W$. Another dominant background is $t\bar{t}$ with non-prompt leptons, including light leptons from the decay of heavy flavour hadrons and, in the categories with τ_{had} , misidentified τ_{had} .

Figure 7 shows the best-fit values of the $t\bar{t}H$ signal strength for a Higgs with $m_H = 125$ GeV extracted from the individual categories and their combination. The best-fit signal strength given by the combination is $\mu = 2.1_{-1.2}^{+1.4}$. The observed (expected) significance of signal under the background only hypothesis ($\mu = 0$) is 1.8 (0.9) standard deviations, corresponding to a p -value of 3.7 (18) %. This compares to an observed significance of 0.9 standard deviations assuming the presence of the SM $t\bar{t}H$ signal, corresponding to a p -value of 18%. Using the CL $_s$ method, the observed (expected) 95% CL upper limit on the signal strength is found to be 4.7 (2.4).

4 Search for $t\bar{t}H$ with $H \rightarrow \gamma\gamma$ and constraints on y_t at 7 TeV and 8 TeV

The search in the $H \rightarrow \gamma\gamma$ decay channel is a particular challenge, as this channel has an extremely small branching ratio. For a Higgs with a mass around 125 GeV, the branching ratio is $BR(H \rightarrow \gamma\gamma) = 2.28 \times 10^{-3}$. But the search in this channel benefits from the good diphoton invariant mass ($m_{\gamma\gamma}$)

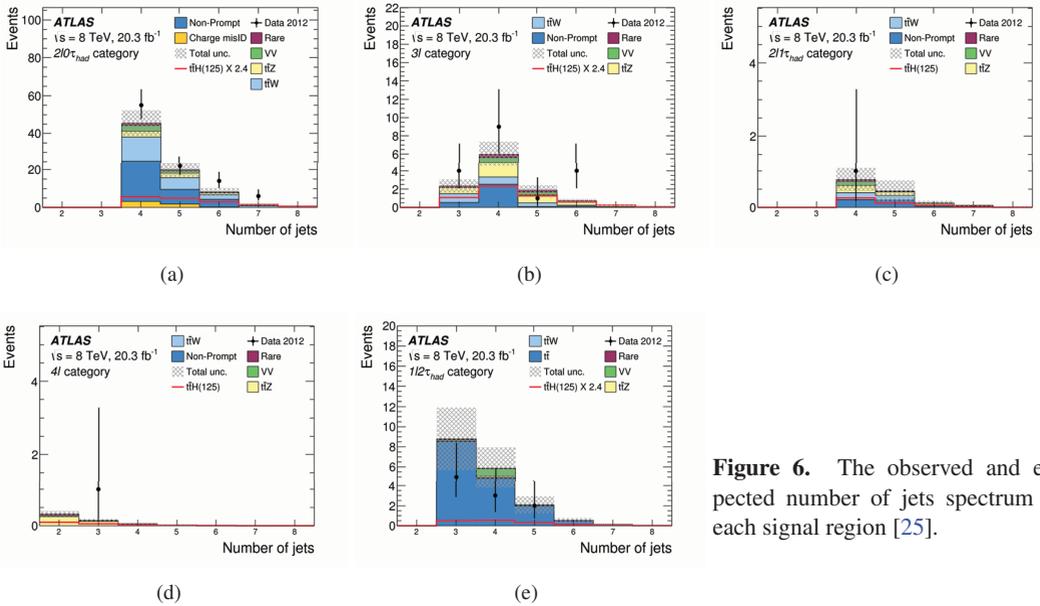


Figure 6. The observed and expected number of jets spectrum in each signal region [25].

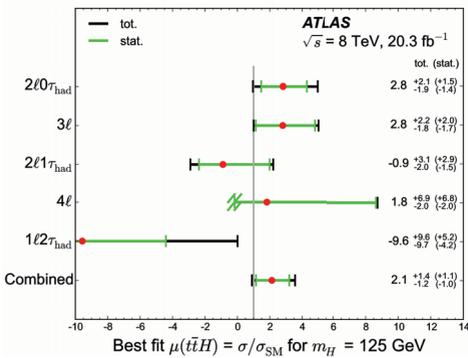


Figure 7. Best-fit values of the signal strength in each of the fitted categories. the truncated lower uncertainty for the 4l category is due to the negative expected total yield given by a non-physical signal strength of $\mu < -0.17$ [25].

resolution, which is estimated to be approximately 1.5 GeV. Additionally, the diphoton background has a continuously decreasing $m_{\gamma\gamma}$ spectrum, which can be constrained by sideband data. Hence, there is no need to rely on the theoretical assumptions on the modelling of the background.

Apart from the search for the $t\bar{t}H$ production, this analysis also provides constraints on the value of y_t . A scale factor, κ_t , is introduced to describe the relation between y_t and its expected value in the SM: $y_t = \kappa_t y_t^{SM}$. The variation of κ_t affects not only the $t\bar{t}H$ and tH production rate, but also the $H \rightarrow \gamma\gamma$ branching ratio and other Higgs boson production processes [26]. Figure 8 illustrates the dependence of the $t\bar{t}H$ and the tH production cross section and the $BR(H \rightarrow \gamma\gamma)$ on the value of κ_t . The value of $\kappa_t = 1$ represents the SM. The tH production has a minimum around the SM expectation given the destructive interference between y_t and g_{HWW} . For $\kappa_t = 0$, the $t\bar{t}H$ production is turned off, and the contribution from y_t to the tH production and $BR(H \rightarrow \gamma\gamma)$ vanishes, resulting in an increase in their values due to the removal of the destructive interference. For $\kappa_t < 0$, a further increase can be seen, because the interference between the contribution of top quarks and W bosons becomes constructive.

The minimum of $\text{BR}(H \rightarrow \gamma\gamma)$ appears at $\kappa_t = +4.7$ due to the cancellation between the contribution from top quarks and W bosons to the loop-induced $H \rightarrow \gamma\gamma$ decay.

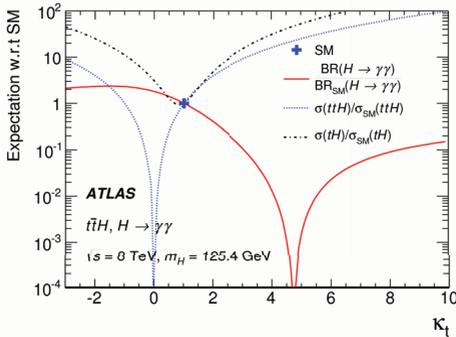


Figure 8. The $t\bar{t}H$ and tH production cross section as well as the $H \rightarrow \gamma\gamma$ branching ratio divided by their SM expectations as a function of κ_t [11].

This analysis was performed on both 7 and 8 TeV data. The selected events in this analysis are categorised based on the decay of the $t\bar{t}$ pair: the leptonic category has at least one lepton; the hadronic category has no lepton in the final state. The analysis searches for a local event excess in the $m_{\gamma\gamma}$ window of (120, 130) GeV. An unbinned signal-plus-background likelihood fit is performed in the range of (105, 160) GeV. Figure 9 shows the fitted distribution compared to the data at 8 TeV for both the leptonic and the hadronic category. In the $m_{\gamma\gamma}$ range of (120, 130) GeV, five candidate events are observed in total. The expected signal yield is 1.3 events, while the total background is estimated to be $4.6^{+1.3}_{-0.9}$ events.

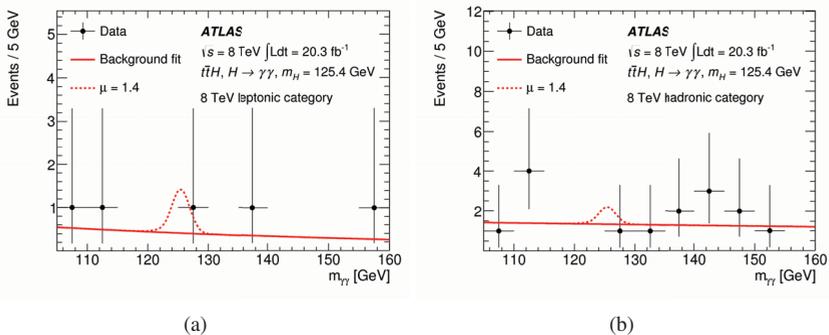


Figure 9. The $m_{\gamma\gamma}$ distribution for (a) the leptonic channel and (b) the hadronic channel for data at 8 TeV. An unbinned signal-plus-background likelihood fit to the full spectrum is used to estimate the continuum background (dashed line) and the SM Higgs boson production with $m_H = 125.4$ GeV [11].

The best-fit signal strength for $t\bar{t}H$ is $\mu_{t\bar{t}H}(m_H = 125.4 \text{ GeV}) = 1.3^{+2.5}_{-1.7}(\text{stat.})^{+0.8}_{-0.4}(\text{syst.})$, assuming SM values for all other $H \rightarrow \gamma\gamma$ contribution. The results interpreted as 95% CL limits for a Higgs boson with $m_H = 125.4$ GeV are summarised in Figure 10. The observed and expected upper limits on the $t\bar{t}H$ signal strength times $\text{BR}(H \rightarrow \gamma\gamma)$ are shown in Figure 10(a). In combination the observed (expected) limit is 6.7 (4.9). The limits as a function of κ_t are shown in Figure 10(b), where the continuum background plus the SM Higgs boson production ($\kappa_t = +1$) is used as the alternative hypothesis. The shape of the distribution is a result of the combined dependence of the $t\bar{t}H$ and tH

production cross section and $\text{BR}(H \rightarrow \gamma\gamma)$, as is shown in Figure 8. The observed (expected) lower and upper limits on κ_t are -1.3 and +8.0 (-1.2 and +7.8).

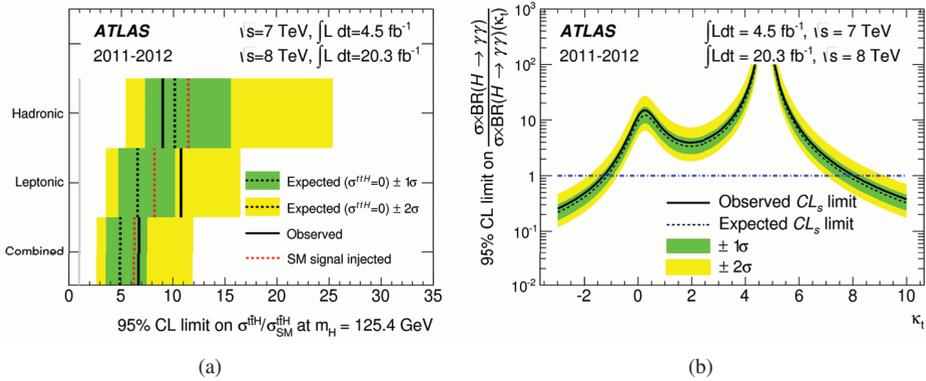


Figure 10. The observed and expected 95% CL limits on (a) the $t\bar{t}H$ signal strength times $\text{BR}(H \rightarrow \gamma\gamma)$ and (b) κ_t (using $\kappa_t = +1$ as the alternative hypothesis) for a Higgs boson with $m_H = 125.4$ GeV [11].

5 Summary

The ATLAS Collaboration has performed searches for the SM Higgs boson produced in association with top quarks in various final states and have exploited constraints on the top Yukawa coupling, y_t . All results so far are consistent with the SM prediction. A combined measurement of the Higgs boson production, decay and coupling has been performed, taking into account the results from the three analyses reviewed in these proceedings, with minor modifications to the analyses to have a consistent SM Higgs boson mass of 125.36 GeV and treatment against the $t\bar{t}H$ processes. The combined $t\bar{t}H$ signal strength is measured to be 1.8 ± 0.8 . A $t\bar{t}H$ signal 3.2 larger than the expectation of the SM prediction has been excluded. In LHC Run 2 starting from June 2015, a much higher $t\bar{t}H$ production rate is expected given the increased luminosity and the collision energy of 13 TeV. Moreover, the relative increase of the $t\bar{t}H$ production cross section with respect to that of the background $t\bar{t}$ process is expected to be 3 times larger, which significantly improves the search sensitivity. Therefore, the observation and measurement of top Yukawa coupling in Run 2 is highly promising.

References

- [1] S. L. Glashow, *Nucl. Phys.* **22** (1961) 579.
- [2] S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [3] A. Salam, Proc. of the 8th Nobel Symposium (1969) 367.
- [4] F. Englert and R. Brout, *Phys. Rev. Lett.* **13** (1964) 321.
- [5] P. W. Higgs, *Phys. Rev. Lett.* **13** (1964) 508.
- [6] P. W. Higgs, *Phys. Lett.* **12** (1964) 132.
- [7] G. Guralnik, C. Hagen, and T. Kibble, *Phys. Rev. Lett.* **13** (1964) 585.
- [8] ATLAS Collaboration, *Phys. Lett.* **B716** (2012) 1, [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex].

- [9] CMS Collaboration, *Phys. Lett.* **B716** (2012) 30, [arXiv:1207.7235 \[hep-ex\]](#).
- [10] F. Bezrukov and M. Shaposhnikov, [arXiv:1411.1923 \[hep-ph\]](#).
- [11] ATLAS Collaboration, *Phys. Lett.* **B740** (2015) 222–242, [arXiv:1409.3122 \[hep-ex\]](#).
- [12] ATLAS Collaboration, submitted to *Eur. Phys. J. C* (2015), [arXiv:1507.04548 \[hep-ex\]](#).
- [13] ATLAS Collaboration, *Eur. Phys. J.* **C75** no. 7, (2015) 349, [arXiv:1503.05066 \[hep-ex\]](#).
- [14] P. Nason, *JHEP* **11** (2004) 040, [arXiv:hep-ph/0409146](#).
- [15] S. Frixione, P. Nason, and C. Oleari, [arXiv:0709.2092 \[hep-ph\]](#).
- [16] S. Alioli, P. Nason, C. Oleari, and E. Re, *JHEP* **06** (2010) 040, [arXiv:1002.2581 \[hep-ph\]](#).
- [17] T. Sjöstrand, S. Mrenna, and P. Skands, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [18] F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini, and F. Siegert, *Phys. Lett.* **B734** (2014) 210–214, [arXiv:1309.5912 \[hep-ph\]](#).
- [19] T. Gleisberg et al., *JHEP* **0902** (2009) 007, [arXiv:0811.4622 \[hep-ph\]](#).
- [20] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Phys. Rev. Lett.* **108** (2012) 111601, [arXiv:1111.5206 \[hep-ph\]](#).
- [21] K. Kondo, *J.Phys.Soc.Jap.* **57** (1988) 4126–4140.
- [22] T. Junk, *Nucl. Instr. Meth. A* **434** (1999) 435, [arXiv:hep-ex/9902006](#).
- [23] A. L. Read, *J. Phys. G* **28** (2002) 2693.
- [24] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71** (2011) 1554, [arXiv:1007.1727 \[physics.data-an\]](#).
- [25] ATLAS Collaboration, *Phys. Lett.* **B749** (2015) 519–541, [arXiv:1506.05988 \[hep-ex\]](#).
- [26] LHC Higgs Cross Section Working Group Collaboration, J. R. Andersen et al., [arXiv:1307.1347 \[hep-ph\]](#).