Searches for ttH (Multilepton+Diphoton) Production in ATLAS

André Sopczak on behalf of the ATLAS Collaboration
Institute of Experimental and Applied Physics, Czech Technical University in Prague
E-mail: andre.sopczak@cern.ch

ABSTRACT

After the discovery of a Higgs boson, the measurements of its properties are at the forefront of research. The determination of the associated production of a Higgs boson and a pair of top quarks is of particular importance as the ttH Yukawa coupling is large and thus an excellent probe for physics beyond the Standard Model (SM). For the complete LHC Run-1 dataset, the combined ATLAS and CMS signal strength (defined as the ratio of measured cross-section to the SM prediction) was $\mu = 2.3^{+0.7}_{-0.6}$, which indicated a mild excess with respect to the SM expectation. The results of the ttH search with multilepton and diphoton signatures are presented for the first LHC Run-2 data (13 fb$^{-1}$ at 13 TeV) recorded by the ATLAS experiment: $\mu = 2.5^{+1.3}_{-1.1}$ for the multilepton analyses and $\mu = -0.3^{+1.2}_{-1.0}$ for the diphoton analyses. Both measurements are in agreement with the SM expectation.

Contribution to the Annual Workshop: Higgs Coupling 2016, Menlo Park, CA, USA, 9-12 November 2016
1. Introduction

An important motivation of the top Yukawa coupling (ttH) research is the fact that at present it is the only quantity which can help us to get an idea about the scale of New Physics [1]. A fundamental prediction of the Brout-Englert-Higgs model is that the Yukawa coupling is proportional to the fermion mass \( m_f \). As \( m_t/m_b \approx 35 \), the ttH coupling is much stronger than the other couplings to lighter fermions.

The ATLAS and CMS collaborations [2, 3] have extensive programmes to search for the ttH signal and measure the ttH coupling strength. The results of both collaborations using the data of LHC Run-1 were combined and gave as ratio of signal strengths \( \mu = \sigma_{ttH,\text{obs}}/\sigma_{ttH,\text{SM}} = 2.3^{+0.7}_{-0.6} \). The measurement is \( 2.3\sigma \) above the SM expectation. The combined ttH signal strength is shown in Fig. 1 (from [4]) compared to other Higgs boson signal strength measurements. Figure 1 also shows the impressive confirmation within the uncertainties of the linearity between the Yukawa coupling and the fermion mass. As the top Yukawa coupling appears not only in the direct measurements, but also for example in the gluon-gluon-fusion process, the best fit of the top Yukawa coupling lies slightly below the diagonal line, indicating an overall smaller value than expected in the SM. Thus, there is some tension between the indirect (loop) ttH measurements and the direct (tree-level) ttH measurements.

![Figure 1](image-url)

**Figure 1.** Left: best fit results for the production signal strengths for the combination of ATLAS and CMS data, as well as from each experiment. The error bars indicate the 1\( \sigma \) (thick lines) and 2\( \sigma \) (thin lines) intervals. The measurements of the global signal strength \( \mu \) are also shown. Right: best fit values as a function of particle mass for the combination of ATLAS and CMS data in the case of the parameterisation described in Ref. [4], with parameters defined as \( \kappa_F \cdot m_F/v \) for the fermions, and as \( \sqrt{\kappa_V} \cdot m_V/v \) for the weak vector bosons, where \( v = 246 \) GeV is the vacuum expectation value of the Higgs field. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The solid (red) line indicates the best fit result to the \([M, \epsilon]\) phenomenological model of Ref. [5] with the corresponding 68% and 95% CL bands.

Regarding the ATLAS ttH multilepton searches in LHC Run-1, there were five channels: one lepton with two hadronic taus, two same-charge leptons with no tau, two same-charge leptons plus one tau, three leptons and four leptons which are tau inclusive. In the initial Run-2 data analysis, there are only the later four channels.
Regarding the ATLAS Higgs diphoton searches, the Run-1 results show a clear peak in the invariant mass spectrum for all Higgs boson production modes, while the ttH production mode sensitivity is limited by the small statistics, as shown in Fig. 2 (from [6]).

Figure 2. Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data. Left: all Higgs boson production modes (untagged). Right: ttH production mode. In each plot the contribution from the different categories in each group is weighted according to the S/B ratio in each category. The error bars represent 68% confidence intervals of the weighted sums. The solid red line indicates the fitted signal plus background model when the Higgs boson mass is fixed at \( m_H = 125 \text{ GeV} \). The background component of each fit is shown with a dotted blue line. Both the signal plus background and background-only curves reported here are obtained from the sum of the individual curves in each category weighted by their signal-to-background ratio. The bottom plot in each figure represents the data relative to the background component of the fitted model.

2. The SM ttH Production and Higgs Boson Decays

The ttH production cross-section, Higgs boson decay branching fractions and the top-antitop decay branching fractions are given in Fig. 3 (from [7]). The ttH production cross-section increased by about a factor four from LHC Run-1 (8 TeV) to Run-2 (13 TeV). A 125 GeV Higgs boson can decay into various pairs of particles. This defines the signatures in the detector, together with the top-antitop decays into alljets, lepton+jets and dileptons final states.

Figure 3. Left: SM Higgs boson production cross-sections as a function of the LHC center-of-mass energy. Center: SM Higgs boson branching ratios as a function of the Higgs boson mass. Right: SM top-antitop branching ratios.

Figure 4 illustrates the Higgs boson production via a loop process involving the top Yukawa coupling, as well as the direct ttH production. Although the highest Higgs boson production rate is via a loop (gg \( \rightarrow \) H), the tree-level direct measurements of pp \( \rightarrow \) ttH are more sensitive to physics beyond the SM.
Some characteristics of the ttH analyses are given, ordered according to the Higgs boson decay modes:

- $H \rightarrow WW, ZZ, \tau\tau$ produce multilepton events,
- $H \rightarrow \gamma\gamma$ has a narrow bump in the diphoton invariant mass spectrum, and
- $H \rightarrow bb$ has a large background rate.

Figure 4. Feynman diagrams of Higgs boson production involving the ttH coupling. Left: loop production. Center: direct production. Right: ATLAS 2016 delivered and recorded integrated luminosity.

3. LHC Luminosity

The measurements of the Higgs boson properties require high statistics datasets. The LHC performance was excellent during the first phase of Run-2 in delivering proton-proton collisions, and ATLAS recorded them with high efficiency. In 2015, $3.2 \, fb^{-1}$, and in 2016, $36.0 \, fb^{-1}$ data were recorded. The results presented here are based on the analysed data up to July 2016 corresponding to $3.2 + 10.0 = 13.2 \, fb^{-1}$. Figure 4 (from [8]) shows the increase of integrated luminosity delivered to and recorded by the ATLAS experiment.

For a luminosity of $13.2 \, fb^{-1}$ and a ttH production cross-section of $507.1 \, fb$ [7], about 6700 ttH events were expected to be produced. The experimental challenge was to detect these events in the about $10^{15}$ background events with a production cross-section of $78.1 \pm 2.9 \, mb$ [9], thus, finding one ttH event in about 154 billion background events.

4. ttH (Multilepton) Agreement in Validation Regions, and Signal and Background Compositions

In order to search for a small number of signal events, it is essential to understand the background reactions with precision. The background reactions were studied in dedicated validation regions which are orthogonal to the signal regions. The validation regions are close to the signal regions and demonstrate the good agreement between the observed and simulated ttZ, WZ+1bjet, and ttW events, as shown in Fig. 5 (from [10]), and listed in Table 1 (from [10]).

The signal compositions (Higgs boson decay modes) in the four multilepton channels are given in Table 2 (from [10]). For example, the simulations show that in the 2$\ell$1$\tau_{had}$ category, the $H \rightarrow \tau\tau$ decay is enriched with 51% contribution. The background compositions in six categories of the four cut-and-counting analyses are illustrated in Fig. 6 (from [10]).

‡ computed at NLO in QCD and electroweak couplings. It has uncertainties of $\pm 5.8\%$ from QCD renormalization/factorization scale choice and $\pm 3.6\%$ from parton distribution function uncertainties (including $\alpha_s$ uncertainties).
Figure 5. Left: invariant mass of leptons 0 and 1 for a loose ttZ validation region. The leptons are labeled in the same way as for the 3ℓ signal region. The events away from the Z peak are those satisfying the Z selection with leptons 0 and 2. Center: jet multiplicity in the WZ+1bjet validation region. Right: lepton flavor composition for events in the ttW validation region. Non-prompt lepton and charge mis-reconstruction backgrounds (indicated as “QMisReco”) are estimated using data as described in Section 6.2 of Ref. [10].

Table 1. Expected and observed event yields in validation regions (VR). The quoted uncertainties in the expectations include all systematic uncertainties. “Purity” indicates the fraction of events in the VR expected to arise from the targeted process (ttZ for the first VR, WZ for the second, and ttW for the third).

<table>
<thead>
<tr>
<th>VR</th>
<th>Purity (%)</th>
<th>Expected</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loose ttZ</td>
<td>58</td>
<td>91 ± 12</td>
<td>89</td>
</tr>
<tr>
<td>WZ+1bjet</td>
<td>33</td>
<td>137 ± 27</td>
<td>147</td>
</tr>
<tr>
<td>ttW</td>
<td>22</td>
<td>51 ± 10</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 2. Fraction of the expected ttH signal arising from different Higgs boson decay modes in each analysis category, and acceptance ($A$) times efficiency ($\epsilon$). The decays contributing to the “other” column are dominantly $H \rightarrow \mu\mu$ and $H \rightarrow b\bar{b}$. Rows may not add to 100% due to rounding. The acceptance times efficiency includes Higgs boson and top quark branching fractions, detector acceptance, and reconstruction and selection efficiency, being computed relative to inclusive ttH production.

<table>
<thead>
<tr>
<th>Higgs boson decay mode (%)</th>
<th>Category</th>
<th>WW*</th>
<th>ττ</th>
<th>ZZ*</th>
<th>Other</th>
<th>$A \times \epsilon$ (×10^{-4})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2ℓ0τ_{had}</td>
<td>77</td>
<td>17</td>
<td>3</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2ℓ1τ_{had}</td>
<td>46</td>
<td>51</td>
<td>2</td>
<td>1</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>3ℓ</td>
<td>74</td>
<td>20</td>
<td>4</td>
<td>2</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>4ℓ</td>
<td>72</td>
<td>18</td>
<td>9</td>
<td>2</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Figure 6. Expected contribution to the background in each channel from various sources, using values of the background estimates before the fit. Charge mis-reconstruction backgrounds are indicated as “QMisReco”.
5. Two Same-charge Light Leptons Plus One Hadronic Tau Final State

In the two same-charge light leptons plus one hadronic tau analysis channel, as illustrated in Fig. 7, one light lepton arises from the tau decay, and one from the top decay. The other tau and the other top decay hadronically, leading to the signature with multijets and b-jets. A selected candidate event display is also shown in Fig. 7 (from [10]).

![Feynman diagram of the two same-charge leptons plus one hadronic tau analysis channel (2ℓτ_{had} category).](image1)

Figure 7. Left: Feynman diagram of the two same-charge leptons plus one hadronic tau analysis channel (2ℓτ_{had} category). Right: event display for a candidate eμτ_{had} event in the 2ℓτ_{had} category. The blue track is the selected electron; the red track is the selected muon; and the white cone is the τ_{had} candidate. The azure cone is the selected b-tagged jet, and the three yellow cones are the non-b-tagged jets. Green and yellow bars indicate energy deposits in the electromagnetic (liquid argon) and hadronic (tile) calorimeters, respectively.

6. Two Same-charge Light Leptons Without Hadronic Tau Final State, Three and Four Light Leptons Final States

For the ttH search with two same-charge light leptons without hadronic tau final state, and three leptons, Fig. 8 (from [10]) shows the number of data, SM signal and background composition. For the four light leptons final state, no data event passes the selection.

![Lepton flavor composition in the 2ℓ0τ_{had} signal region.](image2)

Figure 8. Left: lepton flavor composition in the 2ℓ0τ_{had} signal region. The signal is set to the SM expectation (μ_{ttH} = 1) while the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin. Charge mis-reconstruction backgrounds are indicated as “QMisReco”. Right: lepton flavor composition in the 3ℓ signal region. The signal is set to the SM expectation (μ_{ttH} = 1) and the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin.
7. ttH (Multilepton) Results

The results of the ttH search can be expressed as numbers of selected data, simulated signal and determined background events, as shown in Fig. 9 (from [10]) for pre-fit and post-fit, and pre-fit details are presented in Table 3 (from [10]). The post-fit gives \( \mu_{ttH} = 2.5^{+1.3}_{-1.1} \) indicating a mild data excess compared to the SM prediction. The comparison of the obtained signal strengths \( \mu_{ttH} \) and their upper limits are given in Fig. 10 (from [10]).

**Table 3.** Expected and observed yields in the six signal region categories in 13.2 fb\(^{-1}\) of data at \( \sqrt{s} = 13 \text{ TeV} \). The uncertainties in the background expectations due to systematic effects and MC statistics are also shown. “Other” backgrounds include tZ, tWZ, tHqb, tHW, tttt, ttWW, and triboson production. Values are obtained pre-fit, i.e., using the initial values of background systematic uncertainty nuisance parameters.

<table>
<thead>
<tr>
<th>Category</th>
<th>2(\ell)had(ee)</th>
<th>2(\ell)had(\mu)</th>
<th>2(\ell)had(\mu)μ</th>
<th>3(\ell)had</th>
<th>4(\ell)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ttW</td>
<td>2.9 ± 0.7</td>
<td>9.1 ± 2.5</td>
<td>6.6 ± 1.6</td>
<td>0.8 ± 0.4</td>
<td>6.1 ± 1.3</td>
</tr>
<tr>
<td>tt(Z/γ*)</td>
<td>1.55 ± 0.29</td>
<td>4.3 ± 0.9</td>
<td>2.6 ± 0.6</td>
<td>1.6 ± 0.4</td>
<td>11.5 ± 2.0</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.38 ± 0.25</td>
<td>2.5 ± 1.4</td>
<td>0.8 ± 0.5</td>
<td>0.20 ± 0.15</td>
<td>1.8 ± 1.0</td>
</tr>
<tr>
<td>Non-prompt lepton</td>
<td>12 ± 6</td>
<td>12 ± 5</td>
<td>8.7 ± 3.4</td>
<td>1.3 ± 1.2</td>
<td>20 ± 6</td>
</tr>
<tr>
<td>Ch.misID</td>
<td>6.9 ± 1.3</td>
<td>7.1 ± 1.7</td>
<td>—</td>
<td>0.24 ± 0.03</td>
<td>—</td>
</tr>
<tr>
<td>Other</td>
<td>0.81 ± 0.22</td>
<td>2.2 ± 0.6</td>
<td>1.4 ± 0.4</td>
<td>0.63 ± 0.15</td>
<td>3.3 ± 0.8</td>
</tr>
<tr>
<td>Total background</td>
<td>25 ± 6</td>
<td>38 ± 6</td>
<td>20 ± 4</td>
<td>4.8 ± 1.4</td>
<td>43 ± 7</td>
</tr>
<tr>
<td>ttH (SM)</td>
<td>2.0 ± 0.5</td>
<td>4.8 ± 1.0</td>
<td>2.9 ± 0.6</td>
<td>1.43 ± 0.31</td>
<td>6.2 ± 1.1</td>
</tr>
<tr>
<td>Data</td>
<td>26</td>
<td>59</td>
<td>31</td>
<td>14</td>
<td>46</td>
</tr>
</tbody>
</table>

**Figure 9.** Left: pre-fit background and signal predictions along with observed data yields for each signal region. The ttH prediction corresponds to the SM expectation (\( \mu_{ttH} = 1 \)). The charge mis-reconstruction backgrounds are indicated as “QMisReco”. Right: post-fit background and signal predictions as well as observed data yields for each signal region. The background expectations have been updated to reflect the values of systematic uncertainty nuisance parameters after the fit to data. The ttH prediction corresponds to the best-fit value (\( \mu_{ttH} = 2.5^{+1.3}_{-1.1} \)) and the displayed total uncertainties reflect the uncertainty in ttH as well as the backgrounds.

**Figure 10.** Left: best fit values of the ttH signal strength \( \mu_{ttH} \) by final state category and combined. The SM prediction is \( \mu_{ttH} = 1 \). For the 4\(\ell\) category, as zero events are observed, a 68% CLs upper limit is shown instead. Right: upper limits on the ttH signal strength ttH at 95% CL by final state category and combined. The SM prediction is \( \mu_{ttH} = 1 \). The median upper limit that would be set in the presence of a SM ttH signal \( (\mu = 1) \) is also presented.
8. ttH (H → γγ) Analysis

Owing to the excellent photon energy resolution of the ATLAS calorimeter, the Higgs boson signal manifests itself as a narrow peak in the diphoton invariant mass spectrum on top of a smoothly falling background, and the Higgs boson signal yield can be measured using an appropriate fit. The search for ttH (H → γγ) in the initial LHC Run-2 data [11] follows the Run-1 strategy, and is performed in two categories, one with associated leptons and one with associated hadrons. In both categories the ttH contents is above 90% among all Higgs boson production modes, as shown in Fig. 11 (from [11]).

The Higgs boson mass peak in the diphoton invariant mass distribution is clearly visible combining all production modes (Fig. 11 from [11]). Figure 12 (from [11]) illustrates that in both the ttH associated leptonic and hadronic production modes, the current statistics is too small to observe a peak in the diphoton mass spectrum. Table 4 (from [12]) lists signal, background and data events. The signal strength is shown in Fig. 13 (from [12]) for ttH (H → γγ) together with the results from the other Higgs boson decay modes.

![Figure 11](image1.png)

**Figure 11.** Left: expected signal composition of the event categories for the full kinematic phase space. Right: invariant mass spectrum combining all production mode categories. The black data points indicate the measured distribution, where each event is weighted by the signal-to-background ratio of the event category it belongs to. The blue dashed curve represents the result of a background-only fit to the data, the red curve shows the signal+background distribution based on the fitted signal yields, while the black curve shows the signal component. The bottom inset displays the residuals of the data with respect to the fitted background component (bkg).

![Figure 12](image2.png)

**Figure 12.** Invariant mass spectrum in the ttH (H → γγ) search. Left: hadronic production mode. Right: leptonic production mode. The black data points indicate the measured distribution, the blue dashed curve represents the result of a background-only fit to the data, the green curve marks the signal+background distribution based on the predicted SM signal for a Higgs boson mass of \( m_H = 125.09 \text{ GeV} \), and the red curve shows the signal+background distribution based on the fitted signal yields from the combined fit to all event categories.
Table 4. Expected signal yields (S) and background yields (B) in the ttH (H → γγ) search. The tHqb+WtH yields are also presented, and are included in the background yields. The effect of background is illustrated by providing the number of background events after the fit to data in the smallest diphoton mass interval expected to contain 90% of the SM signal events. This corresponds to a diphoton mass interval of [121.9-127.8] GeV and [121.9-127.9] GeV in the leptonic and hadronic channels, respectively.

<table>
<thead>
<tr>
<th>Category</th>
<th>ttH (S)</th>
<th>Background (B)</th>
<th>tHqb+WtH</th>
<th>S/B</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>all-hadronic</td>
<td>1.58</td>
<td>8.27</td>
<td>0.10</td>
<td>0.19</td>
<td>9</td>
</tr>
<tr>
<td>leptonic</td>
<td>1.16</td>
<td>2.42</td>
<td>0.10</td>
<td>0.48</td>
<td>2</td>
</tr>
</tbody>
</table>

9. ttH Initial Run-2 Results and Comparison with Run-1 Results

The results of the ttH (multilepton) and ttH (H → γγ) searches are shown in Fig. 13 (from [12]) together with the results from the (H → bb) search [13]. The uncertainties in the ttH (multilepton) search are dominated by the systematic uncertainties arising from non-prompt background (heavy-flavor), mis-reconstruction of electric charge, and irreducible background (diboson, ttW, ttZ simulation). With the increasing statistics from the growing dataset, these uncertainties will be reduced, together with the statistical uncertainties. In the ttH (H → γγ) search, an analytic function is used to fit the data sidebands in order to estimate the continuum background. The resulting systematic uncertainty is much smaller than the statistical uncertainty, and the increase of analysed data will directly lead to a significant reduction of the total uncertainty.

The ATLAS measurement of \( \mu_{ttH} \) with the initial LHC Run-2 data agrees very well with the result obtained with the complete LHC Run-1 dataset, both for the mean value and its upper limit, as summarized in Fig. 13 (from [12]).
10. Conclusions

The direct \( \texttt{t} \texttt{t} \texttt{H} \) measurement is the key to determine the top Yukawa coupling independent of the physics in the gluon-Higgs sector. Several analysis channels contribute sensitivity. The excellent initial (2015 and 2016) LHC Run-2 operation and the efficient ATLAS detector operation allowed to record about 39 fb\(^{-1}\) data of which about 13 fb\(^{-1}\) are analysed up to now. The \( \texttt{t} \texttt{t} \texttt{H} \) multilepton and diphoton analyses do already have a higher sensitivity with initial Run-2 data if compared to the complete Run-1 results. There is no significant deviation from the SM expectation. The initial Run-2 results are very well in agreement with the Run-1 results, both having a mild data excess. The prospects of increasing the \( \mu \texttt{t} \texttt{t} \texttt{H} \) measurement precision with the full Run-2 dataset are excellent. This could lead to strong exclusions of models predicting non-SM \( \texttt{t} \texttt{t} \texttt{H} \) rates or an indication of New Physics.

Acknowledgments

I would like to thank the colleagues from the ATLAS and CMS Higgs working groups, and the theorists and phenomenologists present at the conference for the fruitful discussions, as well as the organizers of HC2016 for their invitation and hospitality. The project is supported by the Ministry of Education, Youth and Sports of the Czech Republic under projects number LG 15052 and LM 2015058.

References

[8] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2
[10] ATLAS Collaboration, Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector, ATLAS-CONF-2016-058.
[11] ATLAS Collaboration, Measurement of fiducial, differential and production cross-sections in the \( H \to \gamma\gamma \) decay channel with 13.3 fb\(^{-1}\) of 13 TeV proton-proton collision data with the ATLAS detector, ATLAS-CONF-2016-067.
[12] ATLAS Collaboration, Combination of the searches for Higgs boson production in association with top quarks in the \( \gamma\gamma \), multilepton, and \( bb \) decay channels at \( \sqrt{s} = 13 \) TeV with the ATLAS detector, ATLAS-CONF-2016-068.