

**Search for pair production of Higgs bosons in the  
 $b\bar{b}\tau^+\tau^-$  final state using proton-proton collisions at  
 $\sqrt{s} = 13$  TeV with the ATLAS detector**

---

**Emily Graham**

**First Year Report**



UNIVERSITY OF  
LIVERPOOL

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>The ATLAS Experiment</b>	<b>2</b>
2.1	The LHC	2
2.2	The ATLAS detector	2
<b>3</b>	<b>Motivation</b>	<b>3</b>
3.1	Resonant and non-resonant di-Higgs production	3
3.2	The $b\bar{b}\tau^+\tau^-$ decay channel	5
<b>4</b>	<b>Previous Studies</b>	<b>6</b>
4.1	Run I ATLAS combined result	6
4.2	Run II CMS $b\bar{b}\tau^+\tau^-$ result	6
4.3	Run II ATLAS $b\bar{b}b\bar{b}$ result	6
<b>5</b>	<b>Background Processes</b>	<b>7</b>
<b>6</b>	<b>Event Selection</b>	<b>8</b>
6.1	Event pre-selection	8
6.2	Cut-based analysis	9
6.2.1	Method	9
6.2.2	Transverse mass, $m_T^{lv}$	9
6.2.3	$E_T^{\text{miss}}$ $\phi$ centrality	10
6.2.4	Fake top veto	12
6.2.5	Higgs mass windows	12
6.3	Multivariate analysis	16
6.3.1	Boosted decision trees	16
6.3.2	BDT training	18
<b>7</b>	<b>Limit Setting</b>	<b>19</b>
7.1	The $CL_s$ method	19
7.2	Limits for $b\bar{b}\tau^+\tau^-$	20
<b>8</b>	<b>Outlook and Conclusion</b>	<b>20</b>

---

# 1 Introduction

This report presents a contribution to the search at the ATLAS detector for resonant and non-resonant di-Higgs production with a  $b\bar{b}\tau^+\tau^-$  final state in 13 TeV proton-proton collisions.

Since the discovery of a Higgs-like boson (here denoted as  $h$ ) at the LHC in July 2012 [1, 2], the efforts of the ATLAS and CMS experiments have turned towards verifying that the particle's spin and couplings are consistent with those predicted by the Standard Model (SM).

The Higgs boson is a consequence of electroweak symmetry breaking (EWSB), which also predicts self-coupling between Higgs bosons. This self-coupling is one mechanism for Higgs pair-production. Higgs boson pairs can also be produced non-resonantly through Higgs-fermion Yukawa interactions in the SM. The Higgs pair-production cross section is enhanced in many extensions to the Standard Model, both by resonant and non-resonant processes. Resonant production processes include the decay of a heavy CP-even neutral Higgs boson,  $H$ , or a Randall-Sundrum graviton,  $G$ , to a pair of lighter Higgs bosons,  $h$ . A search was conducted for a heavy resonance in the mass range 260 – 1000 GeV.

## 2 The ATLAS Experiment

### 2.1 The LHC

The LHC [3] ran between 2010 and 2013 with a centre-of-mass collision energy of  $\sqrt{s} = 7 - 8$  TeV (Run I), with ATLAS collecting approximately  $5 \text{ fb}^{-1}$  of data in 2011 at  $\sqrt{s} = 7$  TeV and  $20 \text{ fb}^{-1}$  in 2012 at  $\sqrt{s} = 8$  TeV. After a period of shut-down, Run II collisions restarted in 2015 with an increased energy of  $\sqrt{s} = 13$  TeV and increased luminosity.

The following analysis was performed for Run II data with integrated luminosity of  $3.2 \text{ fb}^{-1}$  and  $\sqrt{s} = 13$  TeV.

### 2.2 The ATLAS detector

The ATLAS detector is a multi-purpose particle physics detector with forward-backward symmetric, cylindrical geometry [4]. Its various components are each optimised for the detection and/or measurement of different types of particles.

The inner detector (ID) tracks particles using its silicon pixel detector, silicon microstrip detector (SCT), and transition radiation tracker (TRT). Surrounding the ID is a thin superconducting solenoid producing a magnetic field of 2 T, which bends the particle tracks to aid identification, and a liquid argon electromagnetic (EM) calorimeter, which measures the energy of an EM shower. The next component in the track of a particle is the hadron calorimeter, which measures the energy of hadronic particles. By combining tracking and calorimetry, particles can be detected and identified and the event can be reconstructed. The muon spectrometer (MS), which identifies muons and measures their momentum, is the outermost component of the ATLAS detector - all other particles, apart from neutrinos, are absorbed by the inner detector and calorimeters before they reach the MS.

Muons are not usually absorbed by the calorimeter due to the fact that they do not interact via the strong interaction and they do not lose energy in the same way as electrons due to their greater mass.

Due to the high number of interactions occurring per bunch crossing in the ATLAS detector, triggers are employed in order to select only the most interesting physics events.

### 3 Motivation

#### 3.1 Resonant and non-resonant di-Higgs production

Electroweak theory requires a local gauge symmetry which requires all gauge bosons to be massless. However, this is contradicted by the fact that W and Z bosons are known to be massive. A solution to this is the existence of a scalar Higgs field, the interactions of which are responsible for the masses of intermediate vector bosons and fermions through spontaneous symmetry breaking [5]. The electroweak symmetry breaking mechanism implies the existence of a scalar boson - the Higgs boson.

In July 2012, the ATLAS and CMS detectors at the LHC announced the discovery of a Higgs-like particle with a mass of 126 GeV. Studies have so far concluded that the particle behaves as expected for the SM Higgs boson.

The potential of the Higgs field is given by:

$$V(|\phi|^2) = \mu^2|\phi|^2 + \lambda|\phi|^4, \quad (3.1)$$

where  $\lambda > 0$  and  $\mu^2 < 0$ . The minimum value of the Higgs potential, the vacuum expectation value (VEV) is:

$$v^2 = -\frac{\mu^2}{\lambda}. \quad (3.2)$$

After spontaneous symmetry breaking, i.e.

$$\phi = \frac{(v + H^0)}{\sqrt{2}}, \quad (3.3)$$

where  $H^0$  denotes the excitation from the VEV, the Higgs boson acquires mass, given by:

$$H = 2\lambda v^2. \quad (3.4)$$

The Higgs potential can now be written as:

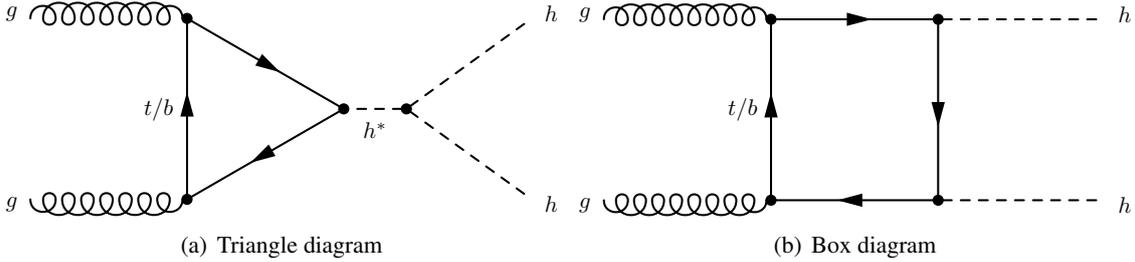
$$\begin{aligned} V(H^0) &= 2\lambda v^2 \frac{(H^0)^2}{2} + 6\lambda v \frac{(H^0)^3}{3!} + 6\lambda \frac{(H^0)^4}{4!} - \frac{v^4 \lambda}{4} \\ &\equiv m_H^2 \frac{(H^0)^2}{2} + \lambda_{HHH} \frac{(H^0)^3}{3!} + \lambda_{HHHH} \frac{(H^0)^4}{4!} - \frac{v^4 \lambda}{4}, \end{aligned} \quad (3.5)$$

where  $\lambda_{HHH}$  and  $\lambda_{HHHH}$  are the Higgs trilinear and quartic couplings, respectively. The trilinear and quartic couplings are uniquely fixed in terms of the Higgs mass:

$$\lambda_{HHH} = \frac{3m_H^2}{v}, \lambda_{HHHH} = \frac{3m_H^2}{v^2} [6]. \quad (3.6)$$

The Higgs trilinear coupling,  $\lambda_{HHH}$ , can only be studied at the LHC through the non-resonant production of Higgs boson pairs. At leading order, Higgs pair-production occurs in gluon-gluon fusion through a box diagram,  $gg \rightarrow HH$ , or a triangle diagram,  $gg \rightarrow H^* \rightarrow HH$ . These are shown in Figures 1(b) and 1(a), respectively. In the triangle diagram, the off-shell Higgs decays to two further Higgs bosons and, therefore, the process is sensitive to the trilinear Higgs self-coupling. Other non-resonant processes for di-Higgs production exist; however, like the box diagram, these do not depend on  $\lambda_{HHH}$ . **These additional processes dilute the dependence of the di-Higgs production cross-section on  $\lambda_{HHH}$**  [7].

In the SM, the di-Higgs production cross section is very small ( $\sim 40$  fb at a 13 TeV centre-of-mass energy [8]). **Although  $\lambda_{HHH}$  depends on  $m_H$  in the SM (Equation 3.6), in new physics (NP) models, this value may differ.** Therefore, it is important to measure the non-resonant Higgs pair production cross section in order to test the SM.

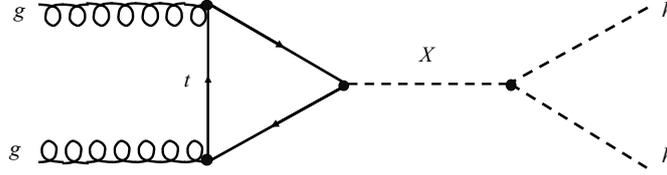


**Figure 1.** Feynman diagrams showing the production of SM Higgs boson pairs at leading order in QCD perturbation theory. In diagram (a), an off-shell Higgs state is produced in gluon-gluon fusion and decays to a pair of Higgs bosons. In diagram (b), the Higgs boson pair is produced through a Higgs-fermion Yukawa interaction; this process is not sensitive to the Higgs trilinear self-coupling. The loops in each of these diagrams are mainly top-quark loops as they couple strongly to the Higgs. Figures taken from [9].

Furthermore, physics beyond the Standard Model (BSM) could greatly enhance the di-Higgs production cross-section (and alter the event kinematics) through the introduction of heavy particles, here denoted by  $X$ , which decay to a Higgs boson pair, as shown in Figure 2. This process is referred to as resonant di-Higgs production. This study focuses on two examples:

- In two-Higgs-doublet models (2HDM) [10], a second Higgs doublet is added - giving a total of five Higgs states: two CP-even, neutral scalars,  $h$  and  $H$  (where  $H$  is more massive than  $h$ ); a CP-odd pseudoscalar,  $A$ ; and two charged scalars,  $H^+$  and  $H^-$ . The heavier neutral scalar,  $H$ , can decay to a pair of lighter (SM-like) neutral scalars,  $h$ , i.e.  $H \rightarrow hh$ , if  $m_H > 2m_h$  [11].
- The Randall-Sundrum (RS) model [12] was postulated to solve the hierarchy problem by

suggesting that spacetime is made up of five dimensions with ‘warped’ geometry. The RS model predicts a series of gravitons with spin-2 and masses on the TeV scale, which may be detectable at the LHC. Therefore, as for the 2HDM heavy Higgs boson,  $H$ , the Randall-Sundrum graviton,  $G$ , may decay to a pair of Higgs bosons, i.e.  $G \rightarrow hh$ .



**Figure 2.** Feynman diagram showing the resonant production of a pair of Higgs bosons. A heavy BSM particle,  $X$ , is produced through gluon-gluon fusion and subsequently decays to a Higgs pair.

### 3.2 The $b\bar{b}\tau^+\tau^-$ decay channel

The decay channels of a Higgs boson pair are given in Table 1, alongside the SM predictions for their branching fractions.

Decay channel	Branching fraction
$b\bar{b}b\bar{b}$	33%
$b\bar{b}W^+W^-$	25%
$b\bar{b}\tau^+\tau^-$	7.4%
$W^+W^-\tau^+\tau^-$	5.4%
$ZZb\bar{b}$	3.1%
$ZZW^+W^-$	1.2%
$\gamma\bar{b}$	0.3%
$\gamma\gamma WW^*$	0.1%
$\gamma\gamma\gamma\gamma$	0.001%

**Table 1.** Table showing the di-Higgs decay channels and their branching fractions, as predicted by the SM.

This analysis is performed for the final state where one Higgs boson decays to a  $b\bar{b}$  pair and the other to a  $\tau^+\tau^-$  pair. The  $bb\tau\tau$  decays are categorized into the three most sensitive  $\tau\tau$  final states:

- $X \rightarrow hh \rightarrow b\bar{b}e\tau_{had}$
- $X \rightarrow hh \rightarrow b\bar{b}\mu\tau_{had}$
- $X \rightarrow hh \rightarrow b\bar{b}\tau_{had}\tau_{had}$

where  $t_{had}$  is a  $\tau$  lepton which decays into hadrons. The analysis is performed in two parts: one for the  $bb\tau_{had}\tau_{had}$  final state and the other for the  $bb\tau_{had}\tau_{lep}$  final state, where  $\tau_{lep}$  refers to a  $\tau$  lepton which decays to either an electron or muon. The work described in this report is for the  $bb\tau_{had}\tau_{lep}$  channel.

## 4 Previous Studies

### 4.1 Run I ATLAS combined result

Searches for both resonant and non-resonant Higgs pair production were performed in the  $hh \rightarrow b\bar{b}\tau^+\tau^-$ ,  $\gamma^*$ ,  $b\bar{b}b\bar{b}$  and  $\gamma\gamma b\bar{b}$  channels, using  $20.3 \text{ fb}^{-1}$  of data from Run I [9]. No evidence for either di-Higgs production process was observed in any final state. The results of these analyses were combined to set 95% confidence level (see Section 7.1) upper limits on the production cross sections. For non-resonant production, the expected (observed) upper limit is 0.69 (0.47) pb which corresponds to 70 (48) times the  $gg \rightarrow hh$  cross section predicted by the SM.

For the resonant production, exclusion limits on  $\sigma(gg \rightarrow H) \times \mathcal{B}(H \rightarrow hh)$  were set for a heavy 2HDM Higgs,  $H$ , as a function of mass,  $m_H$ . The observed (expected) limits range from 2.1 (1.1) pb for  $m_H = 260 \text{ GeV}$  to 0.011 (0.018) pb for  $m_H = 1 \text{ TeV}$ ; this is shown in Figure 3.

From this limit plot it can be seen that limits were set in the range 260 GeV to 1 TeV for the  $bb\tau\tau$  channel. In the range  $260 < m_X < 500 \text{ GeV}$ , there are also limits from the  $WW\gamma\gamma$  and  $bb\gamma\gamma$  final states; the  $bb\tau\tau$  limits are of a similar order of magnitude to those for  $bb\gamma\gamma$ . For  $m_X > 500 \text{ GeV}$ , there are limits on  $\sigma(gg \rightarrow H) \times \mathcal{B}(H \rightarrow hh)$  from the  $bbbb$  analysis which are of a much smaller order of magnitude than the other channels; the  $bbbb$  channel has the highest sensitivity in this region. This results in a large improvement in the combined limit for  $m_X > 500 \text{ GeV}$ .

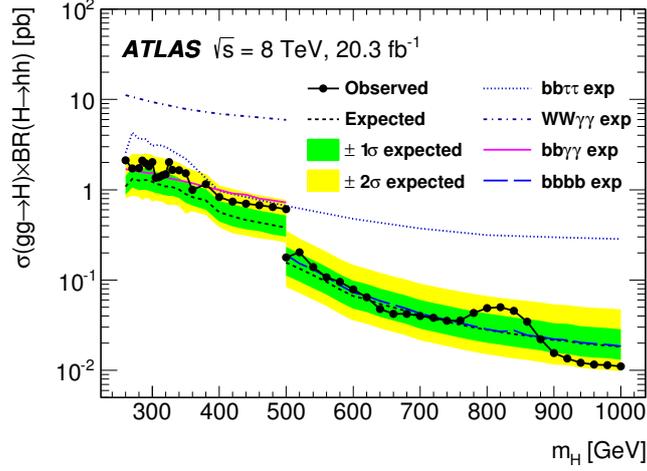
### 4.2 Run II CMS $b\bar{b}\tau^+\tau^-$ result

Results from the CMS Experiment  $pp \rightarrow H \rightarrow hh \rightarrow bb\tau\tau$  analysis, where  $H$  is a heavy scalar Higgs boson, were released in March 2016 [13]. This analysis was performed for three  $\tau\tau$  final states:  $\tau_{had}\tau_{had}$ ,  $e\tau_{had}$  and  $\mu\tau_{had}$ , where  $\tau_{had}$  is a tau lepton which decays hadronically. The analysis uses data collected in 2015 (i.e. Run II of the LHC) with  $\sqrt{s} = 13 \text{ TeV}$  and integrated luminosity of  $2.7 \text{ fb}^{-1}$ .

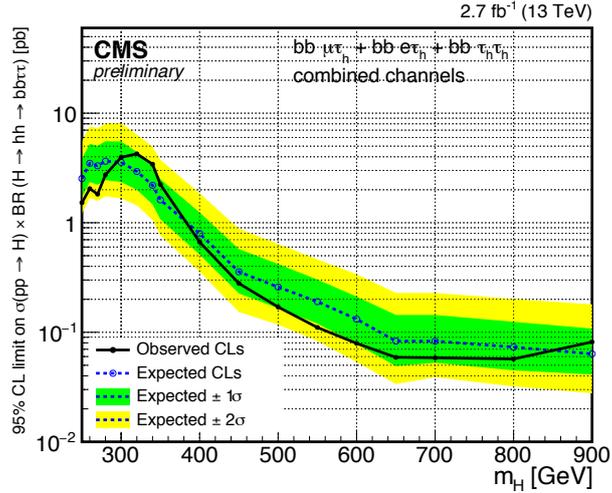
Figure 4 shows the model-independent expected and observed 95% CL limits on  $\sigma(pp \rightarrow H) \times \mathcal{B}(H \rightarrow hh \rightarrow bb\tau\tau)$ . No excess over the SM prediction was observed for the resonant  $hh$  production.

### 4.3 Run II ATLAS $b\bar{b}b\bar{b}$ result

Results from the ATLAS Experiment  $b\bar{b}b\bar{b}$  search were released in March 2016 [14], using  $3.2 \text{ fb}^{-1}$  Run II proton-proton collision data. The data was found to be consistent with the estimated background; therefore, upper limits were set on  $\sigma(pp \rightarrow G) \times \mathcal{B}(G \rightarrow hh \rightarrow bb\tau\tau)$ , where  $G$  is a graviton with spin-0 in the Randall-Sundrum model.



**Figure 3.** Plot of the combined observed and expected 95% CL upper limits on  $\sigma(gg \rightarrow H) \times \mathcal{B}(H \rightarrow hh)$  as functions of  $m_H$ , where  $H$  is a heavy BSM Higgs boson. The expected limits from the individual decay channels are shown. The green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties around the combined expected limit. Figure taken from [9].

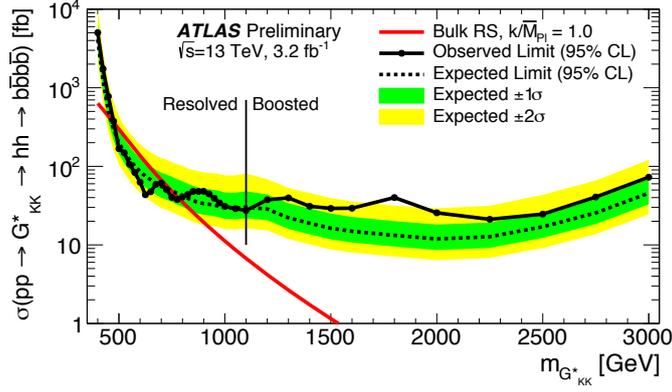


**Figure 4.** The expected and observed 95% CL limits on  $\sigma(pp \rightarrow H) \times \mathcal{B}(H \rightarrow hh \rightarrow bb\tau\tau)$  from the CMS Experiment. The dotted line is the expected limit and the green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties respectively. The solid line represents the observed limit. Figure taken from [13].

## 5 Background Processes

The following background processes are considered in the analysis:

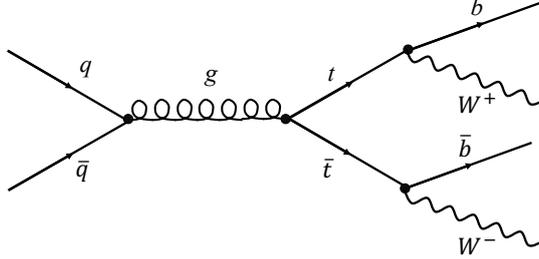
- $t\bar{t}$  production is the main background: top quarks can decay to  $W$  bosons and  $b$  quarks, i.e.  $t\bar{t} \rightarrow W^+ b W^- \bar{b}$ , and the  $W$  bosons can decay leptonically to give  $E_T^{\text{miss}}$  and electrons,



**Figure 5.** The expected and observed 95% CL limits on  $\sigma(pp \rightarrow G_{KK}^* \rightarrow hh \rightarrow bbbb)$ . The dotted line is the expected limit and the green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties respectively. The solid line represents the observed limit. The red line represents  $\sigma(pp \rightarrow G_{KK}^* \rightarrow hh \rightarrow bbbb)$  for a spin-0 graviton in the Randall-Sundrum model,  $G$ , with  $k/\bar{M}_{Pl} = 1$ . Figure taken from [14].

muons or taus. Furthermore, if a  $W$  boson decays hadronically, this can be mis-identified as a hadronically-decaying  $\tau$ . This is shown in Figure 6.

•



**Figure 6.** Feynman diagram showing an example of the  $t\bar{t}$  background. The final state of this process can resemble that of the signal process as the  $W$  boson can decay to either a lepton and  $E_T^{\text{miss}}$  or to hadrons, which can be mis-identified as a hadronically-decaying  $\tau$ .

## 6 Event Selection

Signal Monte Carlo samples at  $m_X = 260, 300, 400, 500, 600, 700, 800, 900, 1000$  GeV, both for the RSG,  $G$ , and 2HDM Higgs,  $H$ , were used.

### 6.1 Event pre-selection

In order to pass the pre-selection, each event requires:

- $\geq 2$  b-tagged jets with  $p_T > 45, 20$  GeV;
- $\tau_{had}$  with  $p_T > 20$  GeV and  $|\eta| < 2.3$ ;
- Exactly one  $\tau_{lep}$ , either:
  - $\mu$  with  $p_T > 21$  GeV and  $|\eta| < 2.5$ ;
  - $e$  with  $p_T > 25$  GeV and  $|\eta| < 2.4$ ;
- Opposite sign  $\tau_{lep}$  and  $\tau_{had}$ .

Two methods of discriminating between signal and background events were investigated in order to achieve higher sensitivity: a cut-based analysis and a multivariate analysis (MVA).

## 6.2 Cut-based analysis

### 6.2.1 Method

The first method of signal and background separation which was investigated was the cut-based analysis. Events which passed the pre-selection were subject to an optimised set of kinematic cuts, which were chosen by comparing the Monte Carlo signal and  $t\bar{t}$  background sets. The figure of merit used here is the significance,  $Z$ , which is given by

$$Z = \sqrt{2 \left( (s+b) \log \frac{1+s}{b} - s \right)}, \quad (6.1)$$

where  $s$  and  $b$  are the number of signal and background events, **respectively**.

In a cut-based analysis, variables are selected for which the signal and background events are well-separated. A cut is then applied to the data in order to keep only the signal events. A series of cuts can be applied, each increasing the significance,  $Z$ .

In this analysis, each cut was optimised by scanning over a range of values of each variable and calculating  $Z$  for a cut made at each value. The variable and corresponding cut value which gave the optimum  $Z$  was chosen, and the cut applied. This process was then repeated for the remaining events and the unused variables.

The optimised background rejection criteria are described in the following sections.

### 6.2.2 Transverse mass, $m_T^{l\nu}$

The transverse mass of the lepton and  $E_T^{\text{miss}}$  system,  $m_T^{l\nu}$ , is defined by

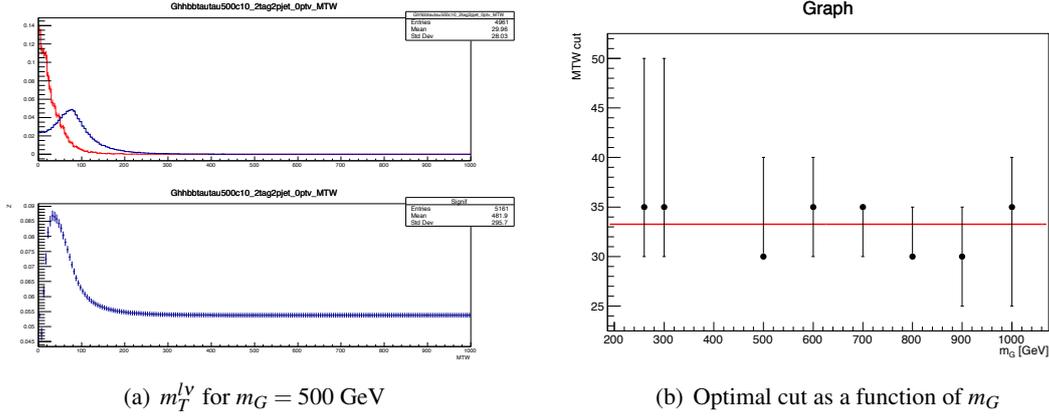
$$m_T^{l\nu} = \sqrt{2p_T^l E_T^{\text{miss}} (1 - \cos \Delta\phi)}, \quad (6.2)$$

where  $p_T^l$  is the transverse momentum of the lepton and  $\Delta\phi$  is the difference in  $\phi$  of the lepton and  $E_T^{\text{miss}}$ . **Signal events tend to have a lower  $m_T^{l\nu}$  than the  $t\bar{t}$  process.**

As explained in Section 6.2.1, the optimal cut for each value of  $m_X$  was chosen by varying the value of the cut and calculating the significance achieved, as shown in Figure 7(a). The top

plot shows the background (blue) and graviton signal with  $m_G = 500$  GeV (red) and the bottom plot shows how the significance varies with the cut value. This shows that, for  $m_G = 500$  GeV, the significance is highest when a requirement of  $m_T^{lV} > 30$  GeV is made.

The optimal cut values were plotted as a function of  $m_X$  - see Figure 7(b). The errors on Figure 7(b) show the range for which the cut would achieve a significance within the errors of the optimal significance (Figure 7(a)). The optimal cut for the full mass range was chosen as the loosest cut which was within the error bars for all points. Therefore the requirement applied was  $m_T^{lV} < 35$  GeV.



**Figure 7.** Figure 7(a) shows  $m_T^{lV}$  for the  $t\bar{t}$  background (blue) and RSG signal with  $m_G = 500$  GeV. The bottom plot shows how the significance varies with the cut value. Figure 7(b) shows the optimal cut on  $m_T^{lV}$  as a function of  $m_G$ . The red line is a line of best fit; this line was not used to choose the optimal cut.

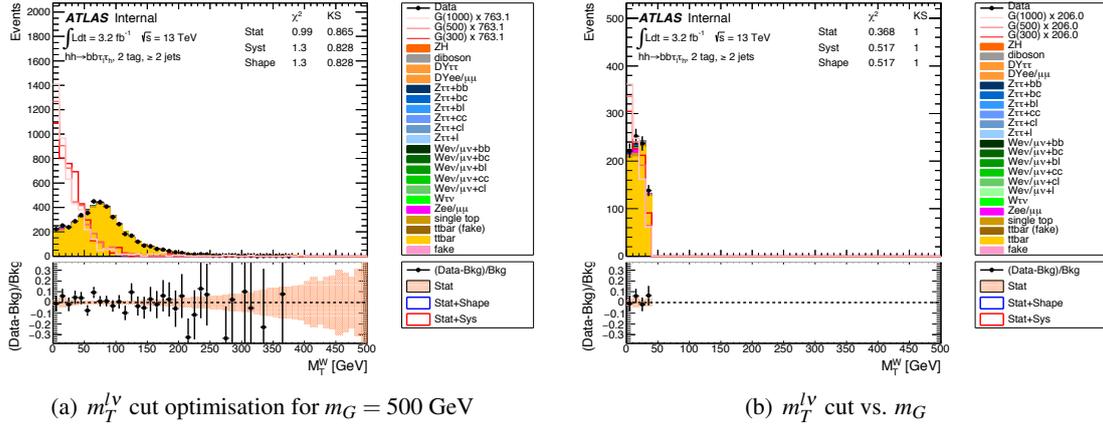
Figure 8 shows  $m_T^{lV}$  for signal, background and data. The signal shown is the graviton,  $G$ , with  $m_G = 300, 500, 1000$  GeV. The plots show how the  $m_T^{lV}$  spectrum varies only a little with  $m_G$ . Figure 8(a) is before the cut was applied and 8(b) is after.

### 6.2.3 $E_T^{\text{miss}}$ $\phi$ centrality

The  $E_T^{\text{miss}}$   $\phi$  centrality quantifies the position in  $\phi$  of the  $E_T^{\text{miss}}$  with respect to the visible decays of the two taus. It is equal to:

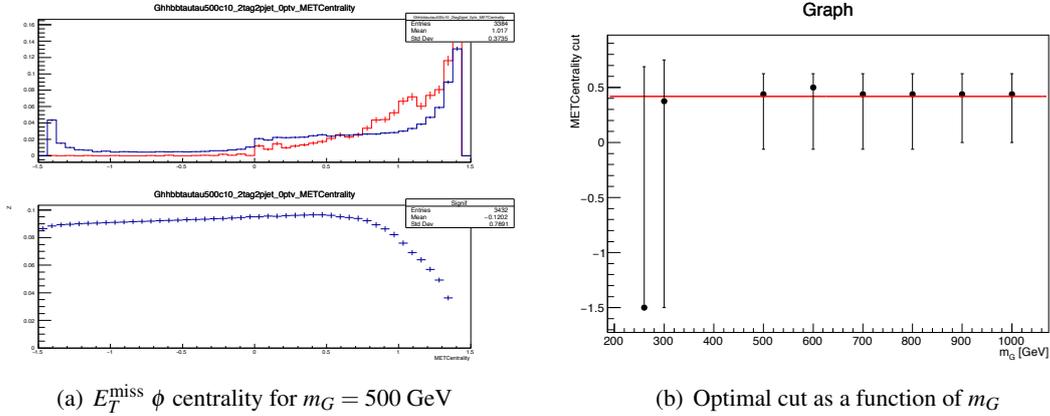
- $\sqrt{2}$  when the  $E_T^{\text{miss}}$  lies directly between the two taus;
- 1 if the  $E_T^{\text{miss}}$  is perfectly aligned with either of the taus; or
- $< 1$  if the  $E_T^{\text{miss}}$  lies outside of the two taus.

This variable is greater for signal events than for the  $t\bar{t}$  background, as can be seen in Figure 10(a). This is due to the fact that for the signal, the  $\tau$  leptons are produced as a pair from one  $H$ , alongside the  $E_T^{\text{miss}}$ , so the  $E_T^{\text{miss}}$  is most likely to be between the two taus. On the other hand, the neutrinos in the background process are produced alongside each  $W$  boson and in the decay of the  $\tau$  to leptons. Therefore, the  $E_T$  is slightly more evenly distributed.



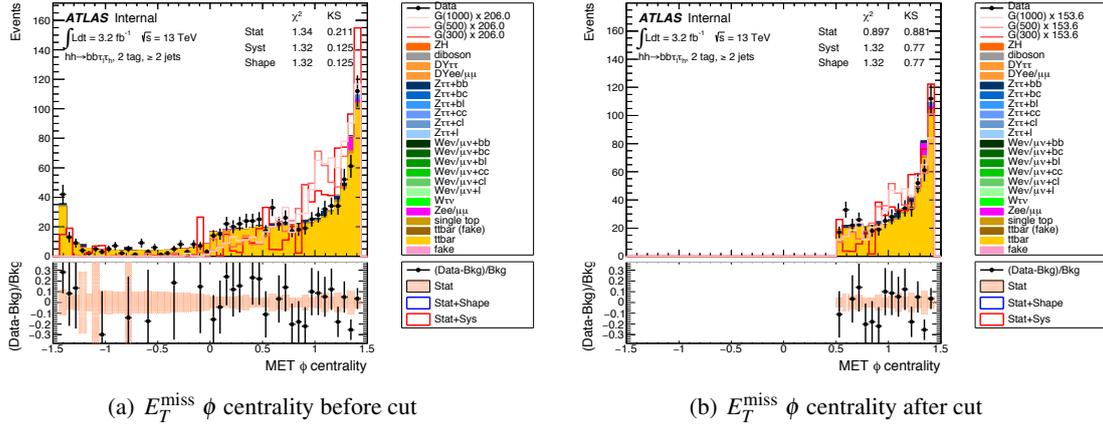
**Figure 8.** The  $m_T^{lV}$  spectra before and after the cut is made. The plots show the MC for the different background processes; the main background is  $t\bar{t}$ , which is shown in yellow. This is overlaid with the signal samples for  $m_X = 300, 500, 1000$  GeV. The signal has been scaled so that it can be more easily compared with the background. The data is also shown on these plots.

The same method of cut optimisation was applied as for the  $m_T^{lV}$  cut. In Figure 9(a), the significance achieved as the cut value is varied is shown for the graviton sample with  $m_G = 500$  GeV and the  $t\bar{t}$  background. Figure 9(b) shows how the optimal cut varies with  $m_X$ . Again, the optimal cut is relatively constant with mass. The cut chosen was  $E_T^{\text{miss}} \phi$  centrality  $> 0$ .



**Figure 9.** Figure 9(a) shows  $E_T^{\text{miss}} \phi$  centrality for the  $t\bar{t}$  background (blue) and RSG signal with  $m_G = 500$  GeV. The bottom plot shows how the significance varies with the cut value. Figure ?? shows the optimal cut on  $E_T^{\text{miss}} \phi$  centrality as a function of  $m_G$ . The red line is a line of best fit; this line was not used to choose the optimal cut.

Figure 10(a) shows the  $E_T^{\text{miss}} \phi$  centrality after the cut on  $m_T^{lV}$  and before the cut on  $E_T^{\text{miss}} \phi$  centrality. Figure 10(b) is the same plot after the cut.



**Figure 10.** The  $E_T^{\text{miss}}$   $\phi$  centrality before and after the cut is made. The plots show the MC for the different background processes; the main background is  $t\bar{t}$ , which is shown in yellow. This is overlaid with the signal samples for  $m_X = 300, 500, 1000$  GeV. The signal has been scaled so that it can be more easily compared with the background. The data is also shown on these plots.

## 6.2.4 Fake top veto

The  $t\bar{t}$  background includes a ‘fake’ tau component, which comes from both  $W \rightarrow l\nu$  and  $W \rightarrow jj$ , where one jet can fake a tau. Therefore, the invariant mass of the  $W$  is formed from the ‘tau’ and its nearest jet. The top mass is then reconstructed by choosing the  $b$ -jet such that  $m_{lb} + m_{\tau b}$  is minimised. An elliptical selection of a 2D plane with the reconstructed  $W$  mass,  $m_{\tau \text{ jet}}$ , and the reconstructed top mass,  $m_{\tau \text{ jet } b\text{-jet}}$ , is made to reject the fake taus.<sup>1</sup>

Figure 11 shows  $m_{\tau \text{ jet}} (m_W)$  vs.  $m_{\tau \text{ jet } b\text{-jet}} (m_t)$  for the  $t\bar{t}$  background (Figure 11(a)) and the graviton signal with  $m_G = 500$  GeV (Figure 11(b)). The events which fall inside the ellipse are removed; it can be seen that there is a far higher density of background than signal in this region.

## 6.2.5 Higgs mass windows

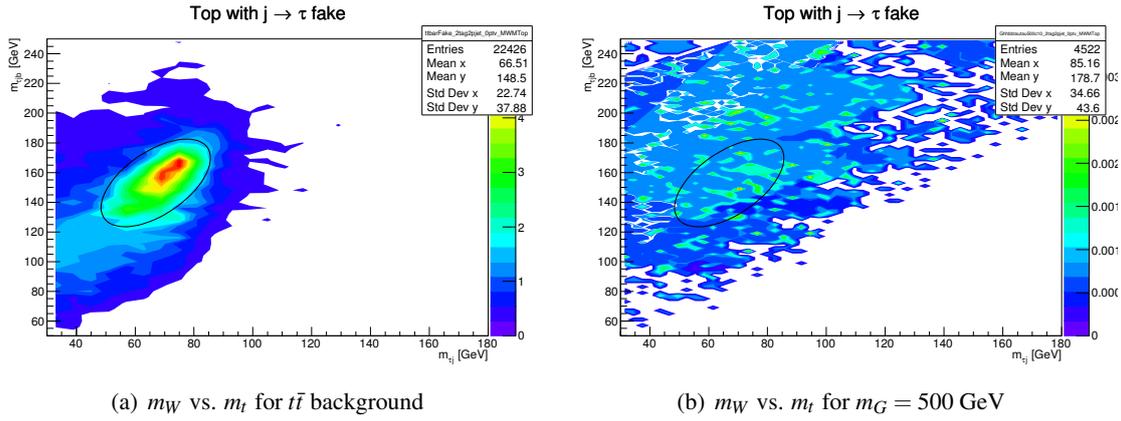
The di-tau and di-jet invariant masses,  $m_{\tau\tau}$  and  $m_{jj}$  respectively, are required to fall in a symmetric window around the Higgs mass defined by:

- $90 < m_{\tau\tau} < 140$  GeV
- $110 < m_{jj} < 140$  GeV.

These windows were chosen in the same way as the cuts above but by varying the size of a symmetrical window around the peak value of the reconstructed mass.

The invariant  $\tau\tau$  mass,  $m_{\tau\tau}$ , was reconstructed using the Missing Mass Calculator (MMC) [15]. For the  $\tau_{lep}\tau_{had}$  channel, the reconstructed  $m_{\tau\tau}$  includes 8 unknowns:  $x$ -,  $y$ - and  $z$ -components of the momentum of the neutrinos for each of the  $\tau$  leptons plus the invariant masses of the two neutrinos from the leptonic  $\tau$  decay.

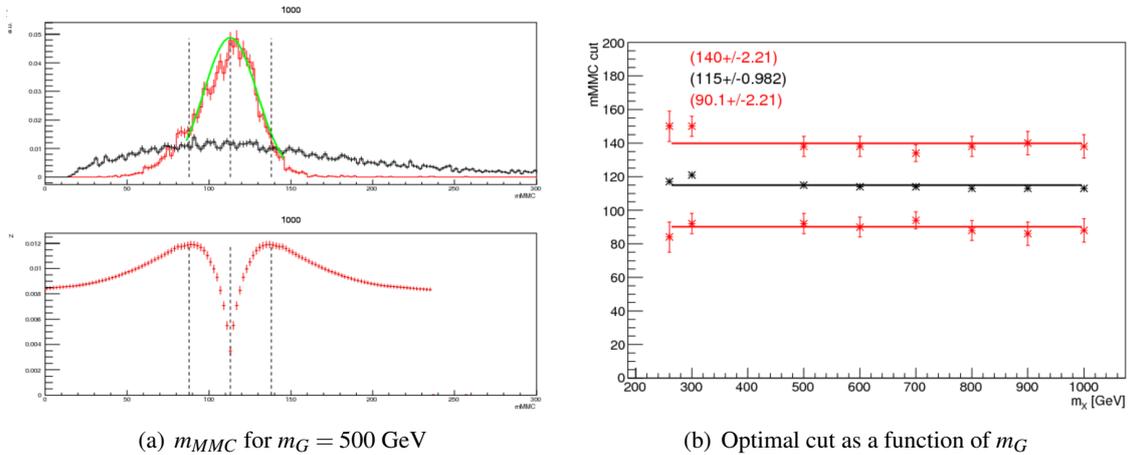
<sup>1</sup>This part of the analysis was performed by Carl Gwilliam. I am including it here for completeness.



**Figure 11.**  $m_{\tau_{\text{jet}}}$  ( $m_W$ ) vs.  $m_{\tau_{\text{jet } b\text{-jet}}}$  ( $m_t$ ) for: 11(a) the  $t\bar{t}$  background and 11(b) the graviton signal with  $m_G = 500$  GeV. The events which fall inside the ellipse are removed; it can be seen that there is a far higher density of background than signal in this region.

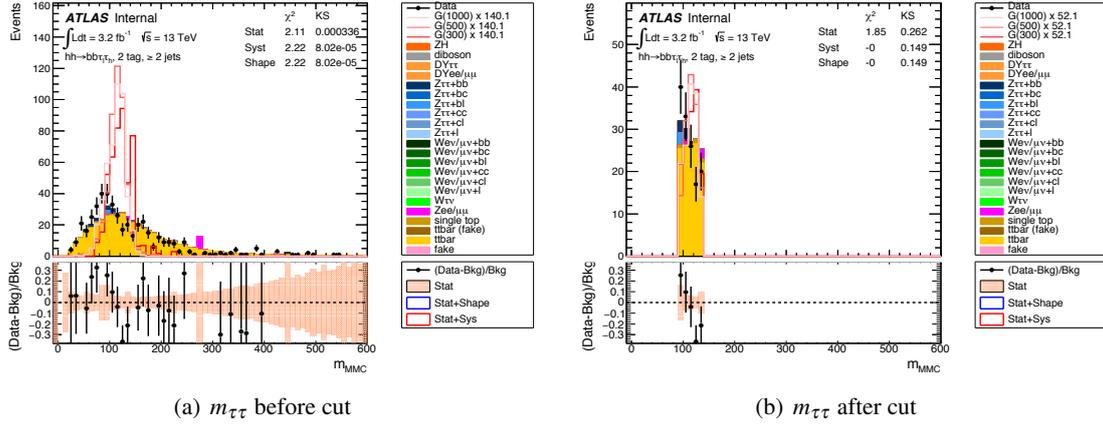
Figure 12(a) shows the reconstructed mass,  $m_{MMC}$ , for the  $t\bar{t}$  background (black) and the graviton signal with  $m_G = 500$  GeV (red). The green line is the fit applied to the signal - this was used to give the position of the peak. The bottom plot shows the significance,  $Z$ , as the size of the window varies. The vertical dotted lines show the size of the window for which  $Z$  is highest.

Figure 12(b) shows the optimal cut on  $m_{MMC}$  as a function of  $m_G$ . The black points are the peak  $m_{MMC}$  values and the red points above and below are the optimised upper and lower cuts. These are relatively constant with mass and so a line of best fit was used. The requirement chosen is  $90 < m_{MMC} < 140$  GeV.



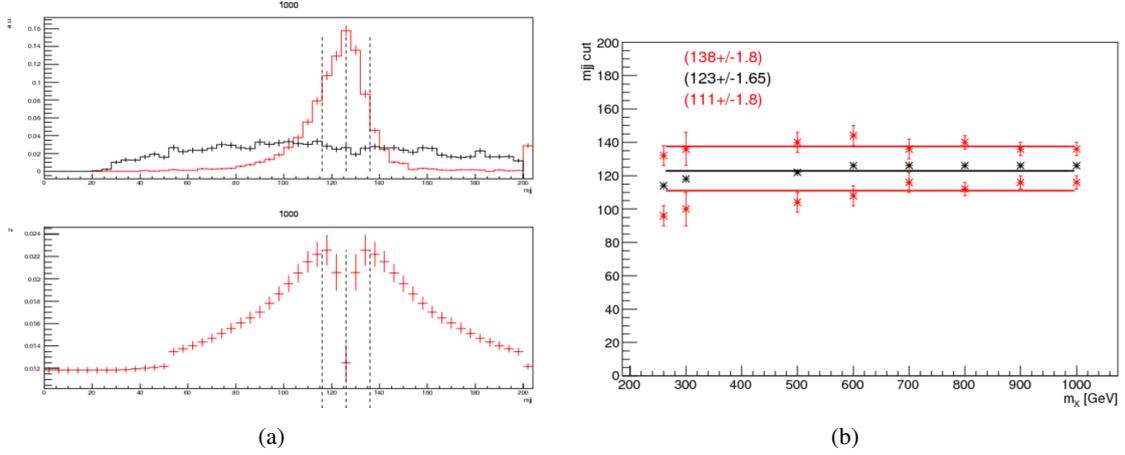
**Figure 12.** Figure 12(a) shows the reconstructed mass,  $m_{MMC}$ , for the  $t\bar{t}$  background in black and the graviton signal with  $m_G = 500$  GeV in red (top) and the significance,  $Z$ , as the size of the window varies (bottom). The green line is the fit applied to the signal. Figure 12(b) shows the optimal upper and lower cuts on  $m_{MMC}$  as a function of  $m_G$ .

Figure 13 shows  $m_{\tau\tau}$  before and after the cut is applied. Both plots are made after the  $m_T^{\text{ly}}$ ,  $E_T^{\text{miss}}$   $\phi$  centrality and top veto cuts described in the above sections.



**Figure 13.**  $m_{\tau\tau}$  before and after the cut is applied. The plots show the MC for the different background processes; the main background is  $t\bar{t}$ , which is shown in yellow. This is overlaid with the signal samples for  $m_\chi = 300, 500, 1000$  GeV. The signal has been scaled so that it can be more easily compared with the background. The data is also shown on these plots.

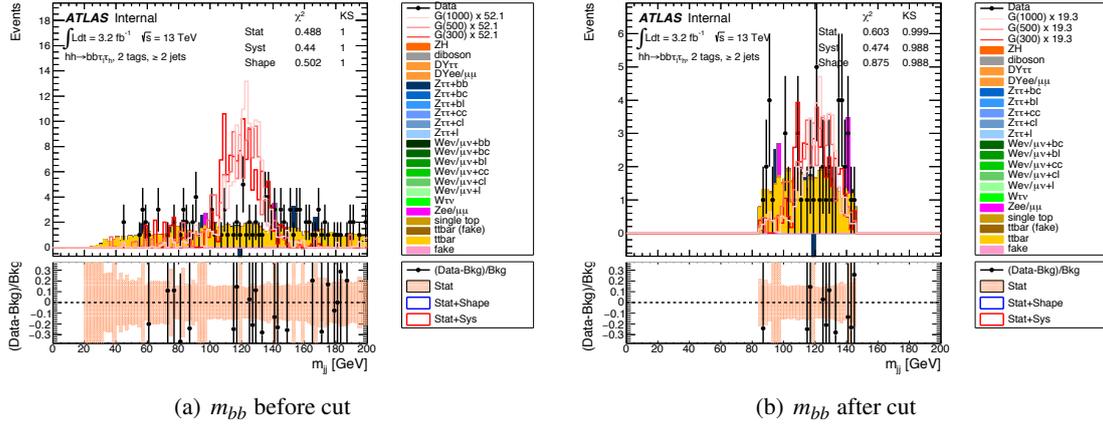
Figure 14 illustrates how the same method is applied to select the optimal window around the di-jet mass,  $m_{jj}$ . Again, the optimal window is relatively constant with  $m_\chi$  so it is required that  $110 < m_{jj} < 140$  GeV.



**Figure 14.** Figure 14(a) shows the reconstructed mass,  $m_{jj}$ , for the  $t\bar{t}$  background in black and the graviton signal with  $m_G = 500$  GeV in red (top) and the significance,  $Z$ , as the size of the window varies (bottom). The green line is the fit applied to the signal. Figure 14(b) shows the optimal upper and lower cuts on  $m_{jj}$  as a function of  $m_G$ .

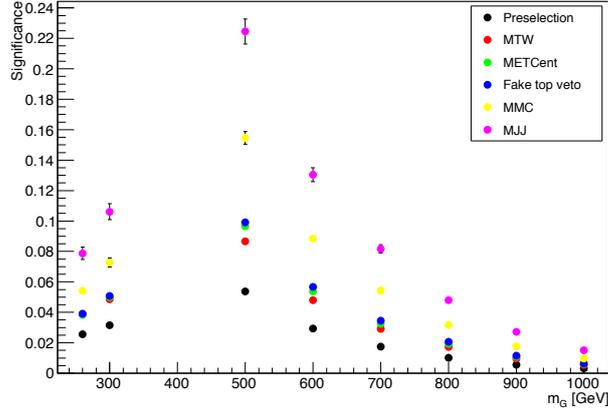
Figure 15 compares  $m_{jj}$  for the backgrounds and the RSG signal with  $m_G = 300, 500$  and  $1000$  GeV, before and after the cut is applied. In these plots, all of the previous selection criteria have also been applied.

Figure 16 shows the significance,  $Z$ , (see Equation 6.1) as a function of  $m_\chi$  after each selection criterion is applied. The plot shows an increase in  $Z$  for each successive cut across the full range of



**Figure 15.**  $m_{jj}$  before and after the cut is applied. The plots show the MC for the different background processes; the main background is  $t\bar{t}$ , which is shown in yellow. This is overlaid with the signal samples for  $m_X = 300, 500, 1000$  GeV. The signal has been scaled so that it can be more easily compared with the background. The data is also shown on these plots.

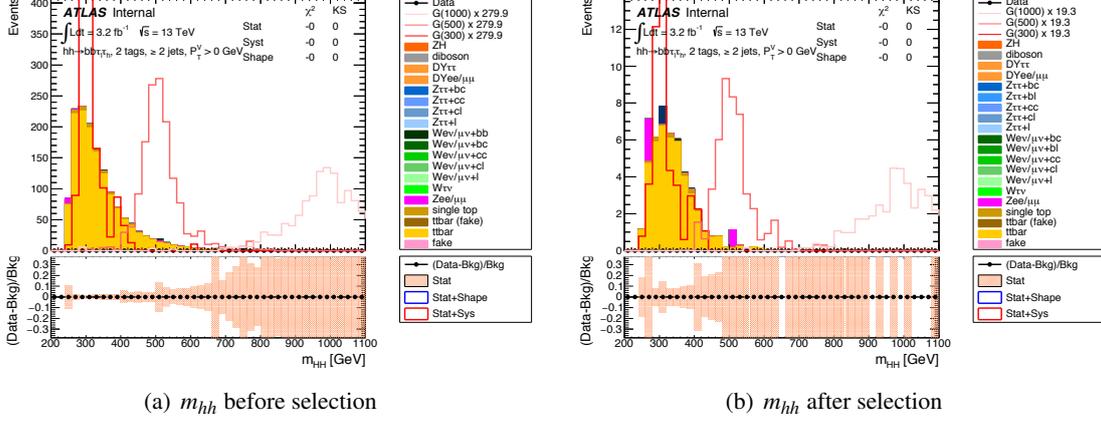
$m_X$ .



**Figure 16.** The significance,  $Z$ , as a function of  $m_X$  after each selection criterion is applied.

Figure 17 compares the reconstructed di-Higgs mass,  $m_{hh}$ , for the background and RSG signal with  $m_G = 300, 500$  and  $1000$  GeV, before and after the full list of selection criteria are applied. The data is not shown here. The main background is  $t\bar{t}$ , which is shown in yellow; this can be seen to decrease relative to the other backgrounds between Figures 17(a) and 17(b). The signal samples for a RSG with  $m_G = 300, 500$  and  $1000$  GeV are also shown. For  $m_X \sim 300$  GeV, the background events have similar values of  $m_{hh}$ , giving a lower sensitivity. For this reason, it may be preferable to perform a multivariate analysis (see Section 6.3) as this will achieve higher sensitivity.

Between Figures 17(a) and 17(b), the factor which the signal peaks are multiplied by for plotting decreases from 279.9 to 19.3. This illustrates how the selection criteria have increased the ratio between the signal and background.



**Figure 17.**  $m_{hh}$  before and after the full list of criteria have been applied. The plots show the MC for the different background processes; the main background is  $t\bar{t}$ , which is shown in yellow. This is overlaid with the signal samples for  $m_X = 300, 500, 1000$  GeV. The signal has been scaled so that it can be more easily compared with the background. The data is also shown on these plots.

## 6.3 Multivariate analysis

### 6.3.1 Boosted decision trees

The Toolkit for Multivariate Analysis (TMVA) package [16] was used to determine whether the analysis would benefit from the use of a Boosted Decision Tree (BDT) algorithm in order to discriminate between signal and background.

A BDT is a machine learning technique used in high energy physics in both event classification and particle identification. It is used to classify signal and background events as an extension of the cut-based method using a number of variables.

A BDT or other MVA method can be more successful than cut-based selection because the MVA can keep events which may be incorrectly rejected by a specific cut based on whether they pass other selection criteria. The decision tree splits the phase space into a large number of hypercubes (multi-dimensional ‘cubes’) where the cut-based analysis selects a signal hypercube region of phase space.

To train a decision tree, the Monte Carlo is split into two sets of known events: one half for training and the other for testing. A decision tree makes a series of binary splits of the data. It begins with all events, both signal and background, on the first node (known as the root node). The algorithm iterates over each variable to find the variable and corresponding cut value which achieve the best separation between signal and background. From this, two new nodes are produced.

There are several ‘separation criteria’ which are used in choosing the best variable and corresponding value to split the node. This study uses the ‘gini index’:

$$\text{gini} = p(1 - p) \sum_i W_i, \quad (6.3)$$

where event  $i$  has weight  $W_i$  and  $p$  is the purity of a node, which is given by

$$P = \frac{s}{s + b}, \quad (6.4)$$

where  $s$  and  $b$  represent the number of signal and background events on the node, respectively.

For purity  $p = 0$  or  $p = 1$  (i.e. a data sample which is 100% signal or background),  $\text{gini} = 0$ ;  $\text{gini}$  is symmetric with respect to the event classes. Therefore, when choosing a splitting variable and value,  $\text{gini}_{\text{left node}} + \text{gini}_{\text{right node}}$  is minimised.

The algorithm is repeated recursively on each node; the next node chosen to split is the one for which splitting will give the maximum change in  $\text{gini}$ . A variable may be used more than once. The decision tree is built when the stopping criteria is reached. The final nodes are known as leaves. Depending on the user-defined configuration, the algorithm may be terminated when:

- The minimum leaf size is reached
- Insufficient improvement is made from further splitting
- The maximal tree depth is reached
- The maximum number of final nodes are produced.

Decision trees are known to be unstable with respect to statistical fluctuations in the training sample. For example, if two variables give similar separation, a relatively small fluctuation in the training sample could lead to a different variable being chosen and, therefore, a substantially different tree being grown from that node.

This is avoided by the use of boosting. Events in the training sample which were misclassified in the original tree have their weights increased (i.e. ‘boosted’) and a new tree is grown. This procedure is repeated recursively for the new tree such that a ‘forest’ of trees are built up, each using a different boosted version of the training data. A weighted average of the trees in the forest is taken to produce the final classification.

This study uses the adaptive boost, or AdaBoost, method. The weights of the misclassified events are multiplied by a boost weight,  $\alpha$ , given by

$$\alpha = \frac{1 - \text{err}}{\text{err}}, \quad (6.5)$$

where misclassification rate of the tree,  $\text{err}$ , is given by:

$$\text{err} = \frac{\text{weight of misclassified events}}{\text{total weight of tree}}. \quad (6.6)$$

The weights are then renormalised so that the sum of the weights of the training events remains constant.

An event is given a score,  $h$ , according to its classification by an individual tree;  $h = +1$  or  $h = -1$  if it is classified as signal or background, respectively. The boosted classification,  $y_{\text{boost}}$  is then given by

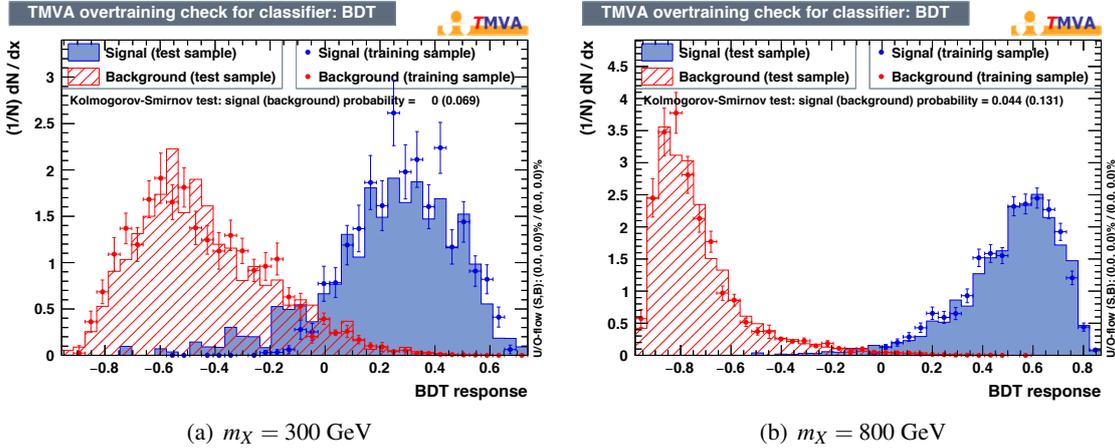
$$y_{\text{boost}} = \frac{1}{N_{\text{trees}}} \sum_{i=1}^{N_{\text{trees}}} h_i \ln \alpha_i, \quad (6.7)$$

where  $N_{\text{trees}}$  is the number of trees in the forest.

Due to the simplicity of this method, where each step of the algorithm involves a simple binary split based on a one-dimensional cut optimisation, the BDT is thought to be the best ‘out of the box’ method for event classification.

### 6.3.2 BDT training

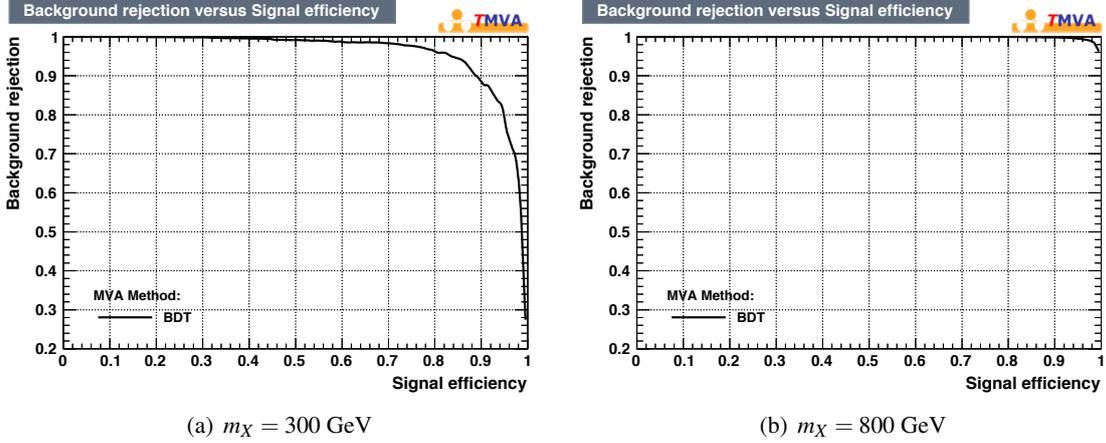
The BDT was trained separately for each mass point in the range from 260 – 1000 GeV. Figure 18 shows the separation achieved by the BDT for  $m_X = 300$  and  $m_X = 800$  GeV; the signal is in blue and the background is in red. As expected, the separation is much greater for the higher  $m_X$ . The solid lines on the plot are the test sample and the individual points are the training sample. A large disparity between the two shows that the BDT is overtrained. In this case, the BDT is overtrained for  $m_X = 300$  GeV. This can be rectified by increasing the number of events in the Monte Carlo sample.



**Figure 18.** These plots show how the BDT separates signal and background for  $m_X = 300$  GeV and  $m_X = 800$  GeV. The solid lines show the signal (blue) and background (red) for the test sample, whereas the individual points are for the training sample. The level of disagreement between the test and training sample for  $m_X = 300$  GeV shows that the BDT is overtrained.

Figure 19 shows how the background rejection varies with the signal efficiency for  $m_X = 300$  GeV and  $m_X = 800$  GeV. For  $m_X = 800$  GeV, greater background rejection can be achieved for

a given signal efficiency than for  $m_X = 300$  GeV - as expected from the signal and background separation shown in Figure 18. The BDT classifier for higher  $m_X$  can achieve a higher significance than for lower  $m_X$ .



**Figure 19.** These graphs show the background rejection that can be achieved as the signal efficiency is increased when using the BDT. The graph on the left is for  $m_X = 300$  GeV and that on the right is for  $m_X = 800$  GeV. If the BDT achieves greater separation then more background rejection can be achieved whilst keeping a higher signal efficiency.

## 7 Limit Setting

95% confidence level expected limits have been produced for the resonant production on  $\sigma(X) \times \mathcal{B}(X \rightarrow hh)$  [17, 18]. For  $\sigma(X) \times \mathcal{B}(X \rightarrow hh)$  below this value, no signal is observable; this is the value at which an upper limit can be set if no excess is seen.

### 7.1 The $CL_s$ method

In the search for new physics, there are conventions for the statistical significance required for an observation to be claimed as evidence, a discovery or exclusion. The statistical significance is classified in terms of sigma,  $\sigma$ ; a  $3\sigma$  deviation from the SM prediction is required for an observation to be classed as evidence, whereas  $5\sigma$  is required to claim a discovery. The number of sigma,  $Z$ , is

$$Z = \sqrt{2} \Phi^{-1}(1 - p), \quad (7.1)$$

where  $\Phi^{-1}$  is the inverse of the cumulative distribution. The  $p$ -value,  $p$ , is the probability that the measurement is compatible with the SM (the null hypothesis) and therefore the observed deviation is within the SM uncertainty. A discovery, or  $Z = 5$ , corresponds to a probability of  $p = 2.87 \times 10^{-7}$ .

For exclusion, the requirement is reduced to  $p < 0.05$  (around  $2\sigma$ ) which is described as the 95% confidence level.

The likelihood,  $L$ , is

$$L(\mu, \theta) = \prod_{i=1}^N \left( \frac{(\mu s_i + b_i)^{n_i}}{n_i} e^{-(\mu s_i + b_i)} \prod_j^M \mathcal{N}_{i,j}(\theta) \right), \quad (7.2)$$

where  $\mu$ , the signal strength ( $\mu = \sigma_{fitted}/\sigma_{predicted}$ ), is the ratio of the fitted cross-section to the predicted cross-section.  $n_i$ ,  $s_i$  and  $b_i$  are the number of data, signal and background events in bin  $i$ , such that the number of events in bin  $i$  is given by  $\mu s_i + b_i$ .  $\theta$  is the set of nuisance parameters representing the systematic uncertainties.

The signal confidence level,  $CL_s$ , is given by

$$CL_s(\mu) = \frac{CL_{s+b}(\mu)}{CL_b} \quad (7.3)$$

where  $CL_b$  refers to the background-only confidence level and  $CL_{s+b}$  is the signal-plus-background confidence level. The 95% confidence level on  $\mu$  is achieved when  $CL_s = 0.05$ .  $CL_s$  can be found from the likelihood function,  $L$ ;  $CL_{s+b}$  corresponds to the  $p$ -value  $\tilde{p}_\mu$  (the probability that the data agrees with a given value of  $\mu$ ), whereas  $CL_b = 1 - \tilde{p}_b$ , where  $\tilde{p}_b$  is the  $p$ -value for the background only hypothesis (or  $\mu = 0$ ).

## 7.2 Limits for $b\bar{b}\tau^+\tau^-$

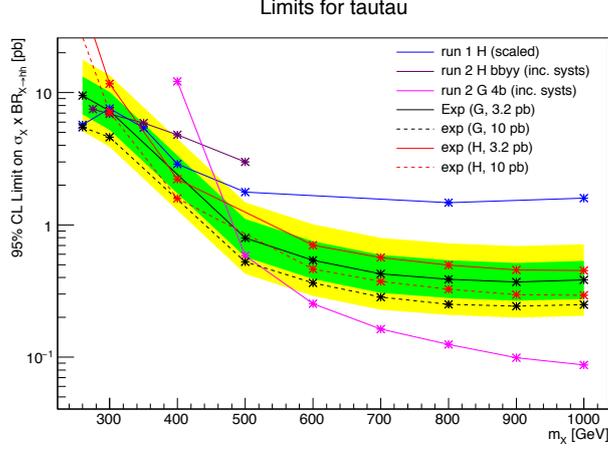
Figure 20 shows the 95% confidence level expected limits on  $\sigma(pp \rightarrow X) \times \mathcal{B}(X \rightarrow hh)$  for the  $bb\tau\tau$  channel as a function of  $m_X$ . The black solid line is the expected limit for the graviton,  $G$ , with  $3.21 \text{ fb}^{-1}$  data and the red solid line is that for the 2HDM Higgs,  $H$ . This plot also shows the result for the  $hh \rightarrow bb\tau\tau$  analysis from Run I; at this stage, the limit has significantly improved for the Run II analysis for  $m_X > 300 \text{ GeV}$ .<sup>2</sup>

Furthermore, the plot shows the limits achieved by two Run II analyses for different di-Higgs final states:  $hh \rightarrow bb\gamma\gamma$  and  $hh \rightarrow bbbb$ .

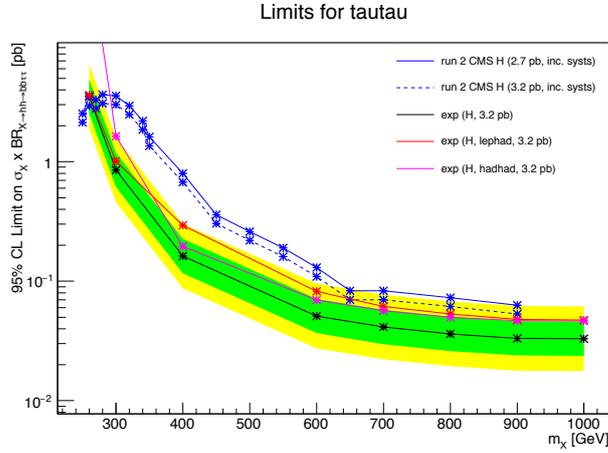
Figure 21 shows the expected limits on  $\sigma(pp \rightarrow X) \times \mathcal{B}(X \rightarrow hh \rightarrow bb\tau\tau)$ . In this case the limits have been multiplied by the  $hh \rightarrow bb\tau\tau$  branching fraction in order to allow for a direct comparison with the results from the Run II  $hh \rightarrow bb\tau\tau$  analysis from the CMS experiment. The black line shows the expected limit for the 2HDM Higgs,  $H$ , for the ATLAS  $bb\tau\tau$  analysis and the dashed blue line is that for the CMS analysis with  $3.2 \text{ fb}^{-1}$  data, allowing for a direct comparison between the two. The CMS limit does include systematics; when the ATLAS result also includes these, the limit can be expected to decrease.

## 8 Outlook and Conclusion

<sup>2</sup>The limit plots were produced by Carl Gwilliam.



**Figure 20.** Plot showing the expected limits on  $\sigma(pp \rightarrow X) \times \mathcal{B}(X \rightarrow hh)$  for the  $bb\tau\tau$  channel for the RSG,  $G$ , in black and the 2HDM Higgs,  $H$ , in black. This is shown for 3.2 pb data (solid line) and 10 pb (dashed line). For comparison, the Run I  $bb\tau\tau$  limit is also shown (blue) as well as the Run II results for the  $bb\gamma\gamma$  (purple) and  $bbbb$  (pink) channels. The  $bb\gamma\gamma$  limit is for a 2HDM Higgs,  $H$ , and the  $bbbb$  limit is for a graviton,  $G$ .



**Figure 21.** Plot showing the expected limits on  $\sigma(pp \rightarrow X) \times \mathcal{B}(X \rightarrow hh \rightarrow bb\tau\tau)$ . Multiplying the limit by the  $hh \rightarrow bb\tau\tau$  branching fraction allows for direct comparison with the limits achieved by the CMS experiment. The combined limit from the  $bb\tau_{lep}\tau_{had}$  and  $bb\tau_{had}\tau_{had}$  channels for this analysis is in black, while the same for the CMS analysis is in blue (the 3.2 pb limit is shown by a dashed line).

## References

- [1] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys.Lett.* **B716** (2012) 1–29, [[1207.7214](#)].
- [2] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys.Lett.* **B716** (2012) 30–61, [[1207.7235](#)].
- [3] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [4] ATLAS collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [5] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509.
- [6] M. Slawinska, W. van den Wollenberg, B. van Eijk and S. Bentvelsen, *Phenomenology of the trilinear Higgs coupling at proton-proton colliders*, [1408.5010](#).
- [7] J. Baglio, A. Djouadi, R. Grber, M. M. Mhlleitner, J. Quevillon and M. Spira, *The measurement of the Higgs self-coupling at the LHC: theoretical status*, *JHEP* **04** (2013) 151, [[1212.5581](#)].
- [8] D. de Florian and J. Mazzitelli, *Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*, *Phys. Rev. Lett.* **111** (2013) 201801, [[1309.6594](#)].
- [9] ATLAS collaboration, G. Aad et al., *Searches for Higgs boson pair production in the  $hh \rightarrow bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$  channels with the ATLAS detector*, *Phys. Rev.* **D92** (2015) 092004, [[1509.04670](#)].
- [10] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher and J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, *Phys. Rept.* **516** (2012) 1–102, [[1106.0034](#)].
- [11] L.-C. L., C. Du, Y. Fang, H.-J. He and H. Zhang, *Searching for Heavier Higgs Boson via Di-Higgs Production at LHC Run-2*, *Phys. Lett.* **B755** (2016) 509–522, [[1507.02644](#)].
- [12] L. Randall and R. Sundrum, *A large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- [13] CMS collaboration, *Search for resonant Higgs boson pair production in the  $b\bar{b}\tau^+\tau^-$  final state*, Tech. Rep. CMS-PAS-HIG-16-013, CERN, Geneva, 2016.
- [14] *Search for pair production of Higgs bosons in the  $b\bar{b}b\bar{b}$  final state using proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-017, CERN, Geneva, Mar, 2016.
- [15] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, *Nucl. Instrum. Meth.* **A654** (2011) 481–489, [[1012.4686](#)].
- [16] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, *PoS ACAT* (2007) 040, [[physics/0703039](#)].
- [17] A. L. Read, *Presentation of search results: the  $CL_s$  technique*, *J. Phys.* **G28** (2002) 2693–2704.
- [18] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554, [[1007.1727](#)].