

# Search for the Higgs boson in fermionic channels using the ATLAS detector

Stephan Hageböck<sup>1,a</sup>, on behalf of the ATLAS Collaboration

<sup>1</sup>University of Bonn, Physikalisches Institut, Nussallee 12, 53115 Bonn

**Abstract.** Since the discovery of the Higgs boson by the ATLAS and CMS experiments at the LHC, the emphasis has shifted towards measurements of its properties. Of particular importance is the direct observation of the coupling of the Higgs boson to fermions. A review of ATLAS results in the search for the Higgs boson in tau, muon and b-quark pairs is presented.

## 1 Introduction

In 2012, the Higgs boson was discovered by the ATLAS [1] and CMS experiments at LHC in decay channels involving bosons while decays to fermions were not observed [2]. The reason is that decays to fermions are hidden by large backgrounds and for light fermions the branching fractions are very low.

At LHC, three different fermionic decay channels are accessible with the following branching fractions<sup>1</sup>:

**$\mu\mu$ , BR=0.02%** Because the muon is light in comparison to the heavy quarks or the  $\tau$ , Higgs decays to muons have a very small branching fraction. However, reconstruction efficiency and invariant mass resolution is the best in comparison to the other fermionic channels.

**$\tau\tau$ , BR=6.3%** The second most abundant decay to fermions is the  $H \rightarrow \tau\tau$  decay. Because the  $\tau$  decays before it can be detected,  $\tau$  identification and reconstruction of the invariant mass are more difficult. Due to various backgrounds, many discriminating variables are used in a multivariate analysis to maximise the sensitivity.

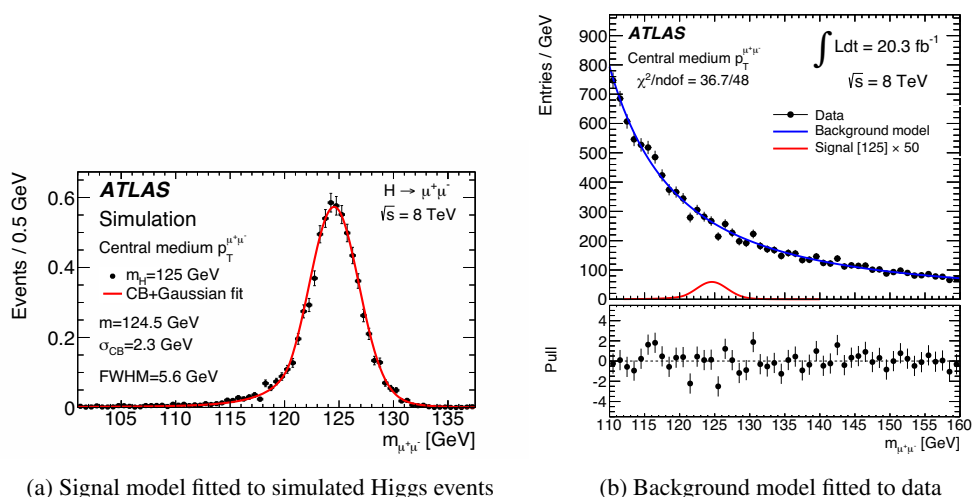
**$bb$ , BR=57%** Because the b-quark is the fermion with the highest mass<sup>2</sup>, the most abundant decay is  $H \rightarrow bb$ . However, b-tagging and the measurement of the  $bb$ -pair invariant mass are challenging. As very large backgrounds contribute to this final state, analyses focus on associated production with a vector boson or a pair of top quarks. This review will focus on a cut based analysis targeting associated production with vector bosons.

---

<sup>a</sup>e-mail: hageboeck@physik.uni-bonn.de

<sup>1</sup>Assuming  $m_H = 125.4$  GeV [3], values from [4])

<sup>2</sup>The top quark is kinematically not accessible.



(a) Signal model fitted to simulated Higgs events

(b) Background model fitted to data

Figure 1: Fit model for the  $H \rightarrow \mu\mu$  analysis. Signal and background model for the central  $\eta$  category, medium  $p_T^{\mu\mu}$ , 8 TeV are shown [5].

Though the  $H \rightarrow cc$  decay has a non-negligible branching fraction of 2.9%, it is not accessible at LHC. As backgrounds with charm jets in the final state are much more abundant than  $bb$  final states and as charm tagging is not pure, the  $H \rightarrow cc$  channel is completely hidden by background processes. Therefore, only the three final states mentioned above will be discussed in more detail.

## 2 $H \rightarrow \mu\mu$

Due to the excellent invariant mass resolution of approximately 2% for a muon pair (fig. 1a), this analysis [5] focuses on extracting the Higgs signal from the invariant mass spectrum of muon pairs. The main backgrounds are  $Z/\gamma^* \rightarrow \mu\mu$  decays. Figure 1b shows data and the background model from the high tail of the  $Z$  peak.

Events are sorted into 7 categories to maximise the mass resolution and to optimise the signal to background ratio. In each of these categories, a signal and a background fit model similar to those in fig. 1 are constructed and fitted to the invariant mass of the muon pairs. Categorisation and fit model will now be discussed in more detail.

### 2.1 Event Selection and Categorisation

To discriminate signal from backgrounds, the strongest selection requirement is the presence of two oppositely charged muons with transverse momenta  $p_T^{\mu 1} > 25$  GeV and  $p_T^{\mu 2} > 15$  GeV. The first of the seven categories targets Higgs production in vector boson fusion (VBF). For VBF events two forward quark jets with a large gap in pseudorapidity are expected. Therefore, two jets with  $\Delta\eta > 3$  and large invariant mass of  $m_{jj} > 500$  GeV are required.

All events failing this requirement are categorised according to the pseudorapidity of the muons into *central*, where both muons have  $|\eta| < 1$ , and *non-central*, where at least one muon has  $|\eta| \geq 1$ .

This uses the fact that the resolution is best in the central part of the detector. These two categories are further splitted according to the transverse momentum of the  $\mu\mu$  system: The backgrounds show a softer  $p_T^{\mu\mu}$  spectrum than the signal. 3  $p_T^{\mu\mu}$  bins are therefore used. This amounts to 7 categories:  $1 \times \text{VBF}$ ,  $3 \times p_T^{\mu\mu} : < 15; 15 - 50; > 50 \text{ GeV}$ , split in central and non-central  $\eta$ .

## 2.2 Fit Model

For each of the 7 categories, an empirical fit model is constructed. It is designed to reproduce the features of the observed invariant mass distribution of the muon pairs (see fig. 1):

**Signal:** the signal model is derived using simulated Higgs events. These events are fitted using the sum of a Crystal Ball and Gaussian function. Parameters like Higgs mass and width of the peak are determined in the fit to the simulation and kept fixed for the final fit on data. These fits are repeated for different simulated Higgs masses and interpolated linearly in order to be able to search in a mass range of  $110 \text{ GeV} < m_{\mu\mu} < 160 \text{ GeV}$ .

Figure 1a shows this model fitted in the central category, medium  $p_T^{\mu\mu}$ . The invariant mass spectrum is well described. A similar level of agreement is also found in the other categories.

**Background:** the background model is developed on simulated events. The background sample is almost completely dominated by  $Z/\gamma^* \rightarrow \mu\mu$  decays. To make the model flexible enough to fit all of the categories outlined above, multiple analytical functions were constructed and tested: Best results were obtained with a Breit-Wigner distribution (to model the Z peak) convolved with a Gaussian (to model the resolution) times an exponential function divided by  $m_{\mu\mu}^3$ :

$$B(m_{\mu\mu}) = f \times [BW * G](m_{\mu\mu}) + (1 - f) \times C \times \frac{e^{A \cdot m_{\mu\mu}}}{m_{\mu\mu}^3} \quad (1)$$

The parameters of the Breit-Wigner distribution are fixed to the Z mass and width, the resolution of the Gaussian is fixed to the mass resolution obtained on simulated  $Z/\gamma^* \rightarrow \mu\mu$  events (separately for each category), the parameters  $f$ ,  $A$  and the overall normalisation  $C$  are determined in fits to data. Figure 1b shows such a fit for the same category as the signal fit (fig. 1a). For the VBF category, a Breit-Wigner times exponential model was found to be more appropriate.

## 2.3 Results

The signal and background models are combined in the form  $\mu \times S + B$  and fitted simultaneously to 7 and 8 TeV data in all 7 categories. For  $m_H = 125.5 \text{ GeV}$ , the observed (expected) upper limit (95% CL) on the SM Higgs production is 7.0 (7.2) times the Standard Model prediction. This translates to an upper limit on the  $H \rightarrow \mu\mu$  branching ratio of  $1.5 \times 10^{-3}$ . The data are consistent with both the signal+background and the background only hypothesis. With the LHC Run 1 dataset of approximately  $25 \text{ fb}^{-1}$  the analysis is clearly limited by data statistics.

## 3 $H \rightarrow \tau\tau$

This section summarises a multivariate analysis to search for  $H \rightarrow \tau\tau$  decays [6]. If the Higgs boson decays into a pair of  $\tau$  leptons, these decay further into hadrons or lighter leptons and neutrinos before they are recorded by the ATLAS detector. The estimation of the  $\tau$  four vectors becomes the more uncertain, the more neutrinos are produced in the decay. A maximum likelihood technique, called

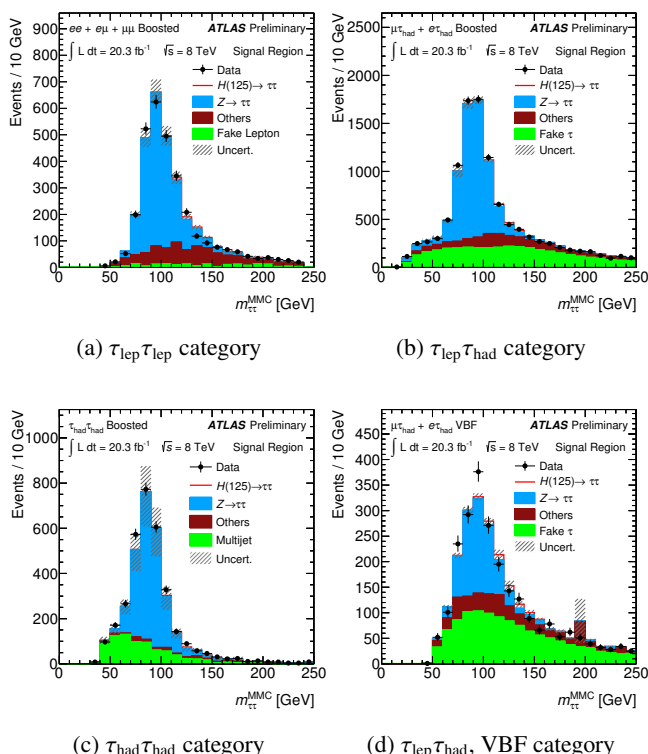


Figure 2: MMC estimate for the  $\tau\tau$  invariant mass using the visible decay products [6]. Background and signal predictions from the global fit (see section 3.2). Most important backgrounds are  $Z \rightarrow \tau\tau$  (blue) and multijet events with jets faking  $\tau_{\text{had}}$  candidates (green).

the missing mass calculator (MMC) [7], is used to obtain the best estimate for the invariant mass. An invariant mass resolution of 14% to 21% (17.6 to 26.3 GeV for  $m_H = 125.4$  GeV) is achieved, depending on the number of neutrinos.

Compared to the  $H \rightarrow \mu\mu$  channel (sect. 2), this is considerably worse but on the other hand, the branching fraction of 6.3% (see sect. 1) is much higher. The sensitivity is maximised by using several discriminating variables in a multivariate approach: Boosted Decision Trees (BDTs) [8] are trained in six different categories. BDTs are a powerful multivariate classification algorithm designed to combine multiple variables into a single discriminant with maximal statistical sensitivity.

### 3.1 Event Selection and Categorisation

Because  $\tau$  leptons can decay either leptonically or hadronically, the analysis is splitted into three categories according to the decay mode:

- $\tau_{\text{lep}}\tau_{\text{lep}}$ : Both  $\tau$  leptons decay leptonically. These leptons must be oppositely charged. Hadronic  $\tau$  candidates “ $\tau_{\text{had}}$ ” are vetoed. Cuts on the invariant mass and transverse momenta of the  $\tau$  leptons as well as on  $E_T^{\text{miss}}$  are applied to suppress backgrounds like bottomonium or other resonances as well

as processes not involving  $\tau$  decays. The most important backgrounds are  $Z \rightarrow \tau\tau$ ,  $Z \rightarrow ee$  or  $\mu\mu$  decays,  $t\bar{t}$  and single top production (see fig. 2a).

- $\tau_{\text{lep}}\tau_{\text{had}}$ : One  $\tau$  lepton decays hadronically. The  $\tau$  candidates are required to be oppositely charged.  $W$ +jets backgrounds are reduced by requiring an  $m_{\text{T}}^W < 70$  GeV. The most important backgrounds are  $Z \rightarrow \tau\tau$  decays and multijet events with jets faking hadronically decaying  $\tau$  leptons (fig. 2b).
- $\tau_{\text{had}}\tau_{\text{had}}$ : Both  $\tau$  leptons decay hadronically. They are required to be oppositely charged. Events with additional leptons are rejected and multijet events are suppressed by applying  $\Delta R$  and  $\Delta\eta$  cuts between the  $\tau_{\text{had}}$  candidates. The most important backgrounds are again  $Z \rightarrow \tau\tau$  decays and multijet events faking hadronically decaying  $\tau$  (fig. 2c).

Depending on whether the Higgs boson is produced in a gluon-gluon fusion ( $\sigma = 19$  pb) or vector boson fusion, VBF ( $\sigma = 1.6$  pb) [4], different event topologies can be exploited. Therefore, the above categories are also splitted into two categories targeted at the production mechanism:

**VBF:** at least two jets, well separated in  $\eta$ , are required. Depending on the decay mode of the  $\tau$  leptons, the minimum  $\Delta\eta$  between the jets with highest transverse momenta ranges from 2.0 to 3.0. Though the cross section is much lower compared to the gluon-gluon fusion, this category is very sensitive to the Higgs signal because of the distinctive topology.

**Boosted:** events failing the requirements of the VBF category are required to have  $p_{\text{T}}^{\tau\tau} = p_{\text{T}}^H > 100$  GeV. Similar to the  $H \rightarrow \mu\mu$  analysis, the backgrounds show on average a softer spectrum of transverse momenta than the signal. This preselection cut therefore increases the signal to background ratio. Higgs events in this category are mostly from gluon-gluon fusion (ggF) but it also includes contributions from Higgs production in association with a vector boson (see sec. 4). However, no dedicated selection targeting the remnants of the vector boson is done.

A comparison of figure 2b with figure 2d illustrates the difference between the boosted and the vector boson fusion topology: The background composition is roughly the same but the event kinematics are significantly different. The preselection cuts are therefore tuned separately for each of these categories (for details see [6]). The BDTs are also trained separately.

## 3.2 Multivariate Analysis and Fit Model

In each of the six categories introduced above BDTs are trained using simulated signal events and partly data-driven, partly simulated background events.

Each decision tree splits the phase space in hypercubes by applying cuts on the discriminating variables. A weight proportional to the signal purity is assigned to each of these hypercubes. For a given event, the input variables are evaluated on all BDTs and the weighted sum of all BDT responses is transformed into the interval  $[-1; 1]$ . 1 is assigned to the most signal like events,  $-1$  to the most background like.

The number of discriminants per category ranges from 6 to 9. These variables describe kinematic properties of the  $\tau\tau$  system (e.g. the invariant mass, fig. 2), properties of the jet system (especially in the VBF category) and event quantities like  $E_{\text{T}}^{\text{miss}}$ . Figure 3 shows two of these six BDT distributions. As mentioned before, the VBF category is powerful because of the distinctive topology (see fig. 3a): The predicted signal  $S$  to background  $B$  ratio in the highest BDT bins of the most sensitive  $\tau_{\text{lep}}\tau_{\text{had}}$  category is 1.0, compared to 0.25 for the boosted category.

The BDT distributions from all six categories are used in a profile likelihood fit to extract the signal strength parameter  $\mu$ . For the categories with at least one leptonically decaying  $\tau$ , event yields in a top and a  $Z \rightarrow ll$  control region are also used to constrain the freely floating top and  $Z \rightarrow ll$

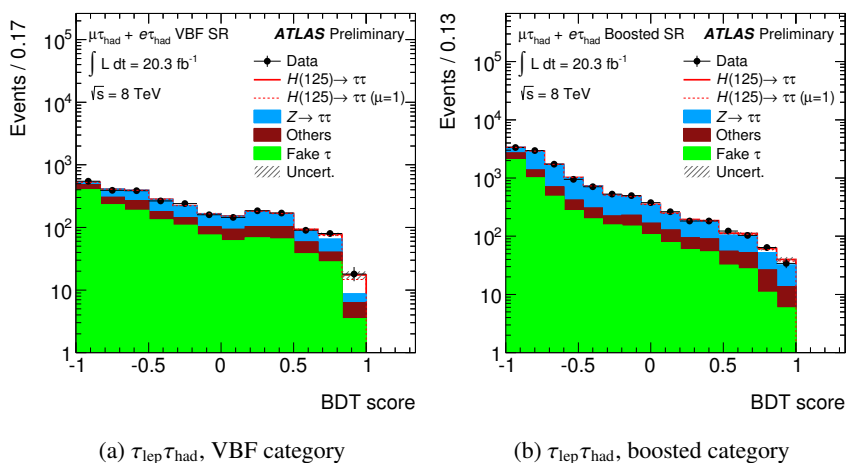


Figure 3: BDT response for the  $H \rightarrow \tau\tau$  analysis. The most sensitive  $\tau_{\text{lep}}\tau_{\text{had}}$  category after the global fit is shown. The Higgs signal (red line) accumulates at BDT scores close to 1 while the backgrounds are peaking at -1. The VBF category (left) has very pure bins in the most signal-like region close to 1 [6].

backgrounds. For the fully hadronic  $\tau$  decays, events failing the requirements for the VBF or Boosted categories (see sect. 3.1) are added in bins of  $\Delta\eta(\tau, \tau)$  to constrain the normalisation of the freely floating multijet background. The freely floating  $Z \rightarrow \tau\tau$  normalisation is constrained by the BDT distributions. The normalisation of other backgrounds is constrained in respective control regions and is only allowed to float within uncertainties determined in separate fits.

### 3.3 Results

Figure 4a summarises all bins of the analysis in a single histogram. The bins are sorted according to the signal  $S$  to background  $B$  ratio so that the most sensitive bins are shown on the right. An excess over the SM background prediction is observed.

The fitted signal strength is  $\mu = 1.43^{+0.31}_{-0.29}(\text{stat.})^{+0.41}_{-0.30}(\text{syst.})$  for  $m_H = 125$  GeV. The data are consistent with the presence of a Standard Model Higgs boson decaying into  $\tau$  leptons. In a background only scenario the probability of obtaining a result at least as signal like as this one is  $p_0 = 2.0 \times 10^{-5}$  or 4.1 standard deviations (3.2 expected). This is evidence for  $H \rightarrow \tau\tau$  decays.

As this analysis targets two different production modes, ggF and VBF, the results can also be expressed in two signal strength parameters: ggF and VBF. Figure 4b shows that the results are consistent with the standard model predictions within 1 standard deviation.

## 4 $H \rightarrow b\bar{b}$

With a branching fraction of 57% the  $H \rightarrow b\bar{b}$  decay is the most probable decay channel. However,  $H \rightarrow b\bar{b}$  decays induced by ggF or VBF will only produce jets in the final state. As the cross-section for production of b-quark pairs at LHC is roughly  $10^8$  pb, it is with the current dataset impossible

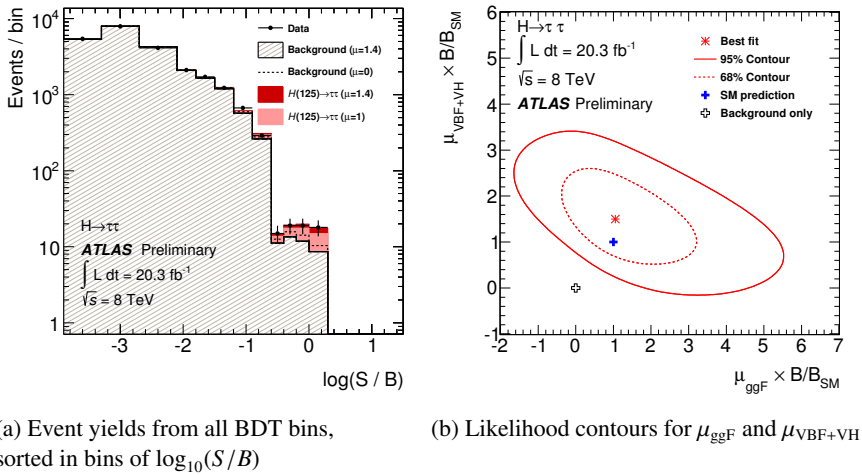


Figure 4: Event yields from all BDT bins and signal strength  $\mu$  for the  $H \rightarrow \tau\tau$  analysis [6].

to use ggF or VBF for the Higgs to fermion search. Instead, this analysis [9] exploits the associated production with vector bosons:  $\sigma_{WH} = 0.68$  pb,  $\sigma_{ZH} = 0.41$  pb. If the vector boson decays into muons, electrons or neutrinos, it creates clear signatures to be used for triggering and background suppression.

Similar to the  $H \rightarrow \mu\mu$  analysis, this analysis tries to extract a Higgs signal from the invariant mass spectrum of b-jet pairs. Depending on the region of phase space, the mass resolution for b-jet pairs ranges from approximately 12% to 18% (15 to 22.6 GeV for  $m_H = 125.4$  GeV). The mass distribution is obtained from simulation for signal and all backgrounds except for QCD multijet production where a data-driven technique is used.

#### 4.1 Event Selection and Categorisation

All signal events are expected to have two b-tagged jets. The first b-jet is required to have a transverse momentum of at least 45 GeV, other jets at least 20 GeV. One additional jet is allowed but must not be b-tagged. Due to the requirement of a leptonically decaying vector boson, the analysis naturally can be splitted using the lepton multiplicity<sup>3</sup>. These categories are illustrated in the columns of fig. 5:

**0 Leptons (fig. 5a, 5d):**  $\sigma \times \text{BR} = 48$  fb. The Z boson decays to neutrinos.  $p_T^Z \approx E_T^{\text{miss}} > 120$  GeV needs to be required to have a clear  $E_T^{\text{miss}}$  trigger signature. Main backgrounds are Z+jets and top production with contributions from W+jets. Backgrounds are suppressed using various  $\Delta\phi$  cuts between  $E_T^{\text{miss}}$  and the visible objects.

**1 Lepton (fig. 5b, 5e):**  $\sigma \times \text{BR} = 132$  fb. The W boson decays into electrons or muons. Main backgrounds are W+jets and top with contributions from QCD multijet production. Though this channel is the most probable, it also suffers from the presence of large backgrounds. Backgrounds without W production can be reduced using  $E_T^{\text{miss}}$  and  $m_T^W$  cuts.

<sup>3</sup>Here, “lepton” means either a muon or electron/positron.

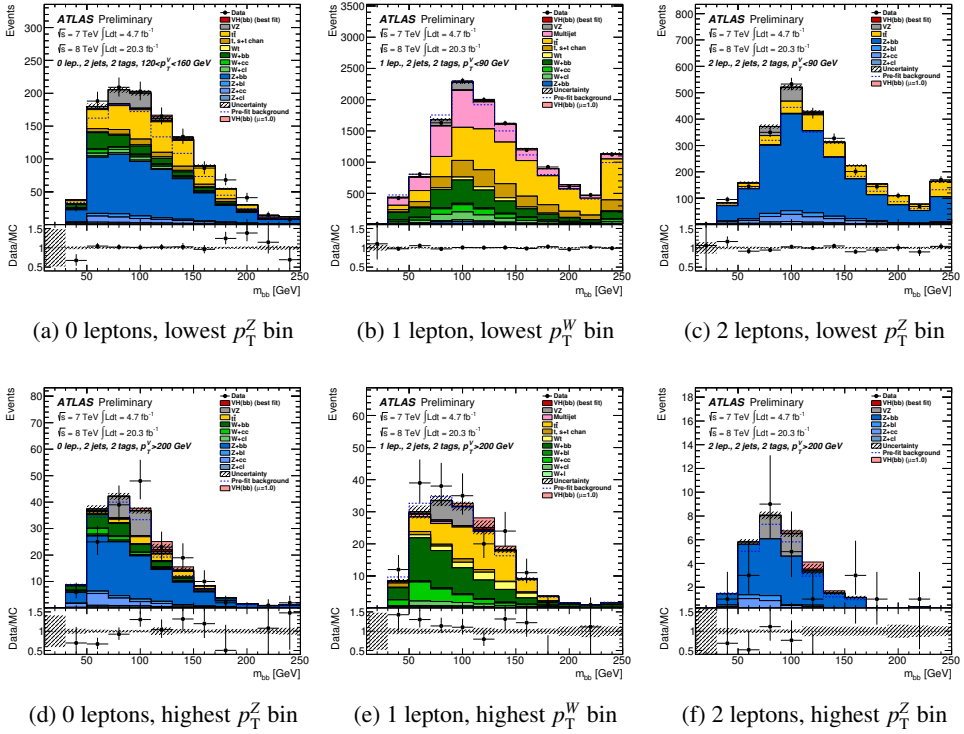


Figure 5: Invariant mass of b-jet pairs [9]. Left column: 0 lepton category. Middle column: 1 lepton category. Right column: 2 lepton category. Top row: Lowest  $p_T^V$ . Bottom row: Highest  $p_T^V$  (highest sensitivity). The Higgs signal is visible only in the high  $p_T^V$  categories. The low  $p_T^V$  categories mainly constrain the backgrounds.

**2 Leptons (fig. 5c, 5f):**  $\sigma \times \text{BR} = 16 \text{ fb}$ . The Z boson decays into a pair of electrons or muons. The main background is Z+jets with contributions from top production. Other backgrounds are largely suppressed by requiring two leptons with an invariant mass near the Z mass. For this reason, this channel is very clean, yet the statistics are low.

The analysis is further splitted into a two and a three jet category. Most of the signal falls in the two jet category. The three jet category mainly serves as a top control region. Like in the  $H \rightarrow \mu\mu$  and the  $H \rightarrow \tau\tau$  analysis, the signal to background ratio increases with the transverse boost of the Higgs. Because of momentum conservation, a high transverse boost of the Higgs also implies a high transverse boost of the vector boson,  $p_T^V$ . For this reason, the phase space is further splitted into five  $p_T^V$  categories at  $p_T^V = 90, 120, 160, 200 \text{ GeV}^4$ . A comparison of the top and bottom rows in fig. 5 illustrates the difference between the lowest and highest  $p_T^V$  categories: At high  $p_T^V$  multijet and top backgrounds are reduced, the signal to background ratio is increased but the statistics are lower.

<sup>4</sup>The 0 lepton category does not include the bins below  $p_T^Z = 120 \text{ GeV}$  because of the trigger requirements.



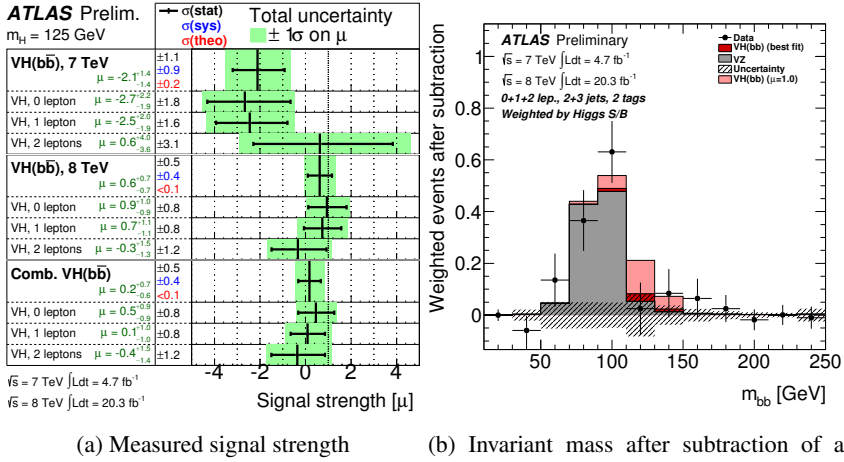


Figure 6: Signal strength  $\mu$  and invariant mass distribution after subtraction of all backgrounds for the  $H \rightarrow bb$  analysis [9].

With higher transverse boost, the jets are more collimated. Therefore,  $p_T^V$ -dependent  $\Delta R$  cuts are applied in all channels. In the 1 lepton channel, also  $p_T^W$ -dependent  $m_T^W$  and  $E_T^{\text{miss}}$  cuts are used to enhance the sensitivity.

## 4.2 Fit Model

The signal strength  $\mu$  is extracted from the invariant mass distribution  $m_{bb}$  in the 26 regions described above. Top, (W/Z)+b and (W/Z)+cl backgrounds are floating in the fit. Other backgrounds are allowed to float within uncertainties determined prior to the final fit like in the  $H \rightarrow \tau\tau$  analysis (see sect. 3.2). Background normalisations are constrained by adding event yields from 26 one b-tag control regions. These are defined as above, but exactly one b-tagged jet is required<sup>5</sup>. This particularly improves the constraints on the normalisation of light flavour V+jets backgrounds. An additional control region is constructed in the 2 lepton channel. This region is defined by requiring an  $e\mu$ -pair instead of  $ee$ - or  $\mu\mu$ -pairs for the leptons. Event yields from this region are added to constrain the top background.

## 4.3 Results

The  $VH$  analysis is cross-checked using  $VZ$  production where the  $Z$  decays to b-quarks: It has the same signature but a higher cross-section and peaks at the  $Z$  mass (grey in fig. 5). The fitted signal strength for this diboson process is  $\mu_{VZ} = 0.9 \pm 0.2$ , well in agreement with standard model predictions. This process is confirmed with an observed significance of  $4.8\sigma$  ( $5.1\sigma$  expected).

The same measurement for the Higgs signal yields no significant excess:  $\mu_H = 0.2 \pm 0.5(\text{stat.}) \pm 0.4(\text{sys.})$ . The probability that a background-only experiment would yield the same or a greater value

<sup>5</sup>This can be viewed as fitting a  $p_T^V$  distribution with five bins in each of the 0,1,2 lepton and 2,3 jets categories.

Table 1: Summary of  $H \rightarrow ff$  results. SM expectations in brackets.

Channel	Exclusion limit $CL_S$	Signal strength $\mu_H$	Significance $\sigma$
$H \rightarrow \mu\mu$	7.0 (7.2)		
$H \rightarrow \tau\tau$		$1.4 \pm 0.5$ (1.0)	4.1 (3.2)
$H \rightarrow bb$	1.4 (1.3)	$0.2 \pm 0.7$ (1.0)	
$H \rightarrow bb$ (MVA)	1.2 (0.8)	$0.5 \pm 0.4$ (1.0)	1.4 (2.6)

is  $p_0 = 0.36$  (0.05 expected). Decomposing this result into 7 and 8 TeV data yields  $\mu_{H,2011} = -2.1 \pm 1.4$  and  $\mu_{H,2012} = 0.6 \pm 0.7$  (see fig. 6a). The low level of agreement between these values is caused by a deficit in the 7 TeV data around the Higgs mass.

The combined result for the full dataset is illustrated in fig. 6b: It shows the invariant mass of the bb-pairs. All post-fit backgrounds except for diboson and Higgs signal are subtracted. The data points confirm the presence of a diboson mass peak but no significant Higgs peak.

The exclusion limit for associated Higgs production in this channel is 1.4 times the Standard Model cross section (1.3 expected).

## 5 Summary

At LHC, three channels are accessible to prove that the Higgs couples to fermions. Table 1 summarises the results. There is direct evidence for  $H \rightarrow \tau\tau$  decays. The other lepton channel,  $H \rightarrow \mu\mu$ , is not sensitive, yet because the data statistics are too low. This proves that the boson does not couple universally to leptons, as predicted by the Standard Model. The data statistics need to be increased significantly to make this channel competitive.

The Higgs decay to b-quarks, produced in association with vector bosons, can neither be proven nor excluded with the dijet mass analysis. Recently, it has been superseded by a more sensitive multivariate analysis [10]: With a signal strength of  $\mu_H = 0.51 \pm 0.31(\text{stat.}) \pm 0.24(\text{syst.})$  and an excess with a significance of  $1.4 \sigma$  (2.6 expected) the presence or absence of this decay can still not be proven.

## References

- [1] ATLAS Collaboration, JINST **3**, S08003 (2008)
- [2] ATLAS Collaboration, Phys.Lett. **B716**, 1 (2012), 1207.7214
- [3] ATLAS Collaboration, Phys.Rev. **D90**, 052004 (2014), 1406.3827
- [4] S. Heinemeyer et al. (LHC Higgs Cross Section Working Group) (2013), 1307.1347
- [5] ATLAS Collaboration, Phys.Lett. **B738**, 68 (2014), 1406.7663
- [6] ATLAS Collaboration (2013), ATLAS-CONF-2013-108, <http://cds.cern.ch/record/1632191>
- [7] A. Elagin, P. Murat, A. Pranko, A. Safonov, Nucl.Instrum.Meth. **A654**, 481 (2011), 1012.4686
- [8] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, PoS **ACAT**, 040 (2007), [physics/0703039](https://arxiv.org/abs/physics/0703039)
- [9] ATLAS Collaboration (2013), ATLAS-CONF-2013-079, <http://cds.cern.ch/record/1563235>
- [10] ATLAS Collaboration, submitted to JHEP (2014), 1409.6212