ATLAS-NOTE ATL-COM-PHYS-2004-014
November, 2004

Search for a light Standard Model Higgs in the channels $pp \rightarrow t\bar{t}H$,WH,ZH

B.King, S.Maxfield, J.Vossebeld.

Abstract

The feasability of observing a light Standard Model Higgs in the channels $pp \rightarrow t\bar{t}H$, WH, ZH is investigated.

1 Introduction

We study the feasibility of observing a light Standard Model Higgs boson in the processes $pp \rightarrow t\bar{t}H, \mathrm{WH}$, and ZH. The $t\bar{t}H$ channel is expected to contribute significantly to the ATLAS discovery potential for a Standard Model Higgs with a mass just above the LEP mass exclusion limit [1], (~114 GeV/c^2), whereas background processes to the WH and ZH channels are a major challenge [2].

The main goals of this study are to confirm the relative usefulness of each of these three channels and to test a number of ideas on how to approach these analyses. For this reason we have performed our study using the ATLAS fast simulation program ATLFAST [3] to simulate the detector response and we apply relatively simple selection and reconstruction procedures.

For the $t\bar{t}H$ analysis, we have studied the case where: the Higgs decays to two b quarks (the dominant branching fraction in the mass region under study); the top quarks decay to a b-quark and a W-boson; one W-boson decays semi-leptonically with an electron or a muon in the final state and the other W-boson decays hadronically. This means the final state will contain at least 4 b-quark jets, 2 light quark jets together with an electron or muon. The energetic lepton from the semi-leptonic W-decay provides a means to trigger the event and an effective discrimination against QCD backgrounds.

Since the final state contains four b-quark jets, a full reconstruction of the final state is needed, in order to attribute the correct b-jets to the Higgs decay, thereby reducing the combinatorial background when reconstructing the Higgs mass.

This will also strongly reduce the number of processes that form a significant background to the signal. We find that if we place sufficiently stringent constraints on the reconstructed top and W masses, the only significant backgrounds come from processes with a top quark pair in the final state. We simulated these backgrounds using two generator programs: The AcerMC [4] program was used to simulate events with a $t\bar{t}b\bar{b}$ final state. In addition a very large sample of inclusive $t\bar{t}$ event was simulated to account for background events where light quark or gluon jets are misidentified as b-jets. The latter sample was simulated using the Pythia [5] MC program.

For the ZH analysis, we have concentrated on the case where the Z has decayed to either an electron or muon pair and the Higgs to a $b\bar{b}$ pair. Similarly, in the WH analysis, we try to select events with two b-jets from the Higgs decay and one lepton from the W decay.

Large samples of all significant background processes were simulated, including some not actually containing a b-quark, but where misidentification of a light quark jet could result in contamination of the selected sample.

2 Simulation of signal and background samples

a) $t\bar{t}H$

Three samples of $t\bar{t}H$ signal events were generated with Higgs masses of 115, 125 and 135 GeV/c^2 respectively. The generator program used was Pythia 6.203.

Background processes with top quark pairs in the final state were simulated using two generator programs. The AcerMC program was used to simulate events with a $t\bar{t}b\bar{b}$ final state. These events include both QCD diagrams and diagrams with Z and W exchange. In addition, a very large sample of inclusive $t\bar{t}$ event was simulated to account for background events where light quark or gluon jets are misidentified as b-jets. This sample was generated using Pythia. A summary of the signal and background cross-sections and the number of generated events is shown in Table 1. These cross-sections were checked to be consistent with those used in reference [6].

b) ZH, WH.

The signal samples were generated using Pythia 6.203 with an input Higgs mass of 115 GeV/c^2. Large statistical samples of all potential backgrounds were also generated, in particular the Z+Jets and W+Jets processes which are troublesome due to their relatively large cross-sections. Table 2 summarises the processes simulated for the ZH and WH analyses. All Z+Jets and W+Jets backgrounds were generated using Pythia 6.203.

All events were passed through the ATLFAST detector simulation and reconstruction code (version 2.60). Throughout this analysis a cone algorithm has been used to define jets. The efficiency to tag jets originating from b-quarks is taken to be 60%, independent of the jet transverse momentum, for all jets with $|\eta| < 2.5$. The associated probabilities to misidentify jets from c quarks or light quarks/gluons were taken to be 10% and 1% respectively.

2.1 Selection and reconstruction

a) $t\bar{t}H$

Simulated events were first subjected to a set of pre-selection requirements:

- at least one electron or muon is observed with $p_T > 20$ GeV/c and $\eta < 2.5$
- 6 or more jets are found with $p_T > 20$ GeV/c of which:
  - 4 or more are identified as b-jets and have $p_T > 25$ GeV/c and $\eta < 2.5$,
  - 2 or more are not identified as b-jets.

The expected number of selected events for the signal and background processes is given in Table 3 for an integrated luminosity of 30 fb^{-1}. Using all events in the pre-selected samples we attempt to do a full reconstruction of the top quark decays taking the following steps:

- In order to fully reconstruct the semi-leptonic top quark decay we have to determine the neutrino momentum. The x and y components of the neutrino momentum are set equal to the missing energy in the x and y directions. We estimate the longitudinal component of the neutrino momentum, p_z^ν by solving the mass equation

  $$m_W^2 = (E^l + E^\nu)^2 - (p^l + p^\nu)^2.$$

  This equation has either 2 or no solutions for p_z^ν. If there are two solutions, both are subsequently considered in the analysis. If there is no solution it means that using only the x and y components of the momentum, $m_{l\nu}$ already overshoots the W mass. In such cases we choose p_z^ν such as to not add any further to the invariant mass. This is done by giving the neutrino the same polar angle as the lepton. In Figure 1 we compare our estimated neutrino momentum components with the true neutrino momentum.

- For all possible pairs of light-quark jets we calculate m_{jj}. All combinations are subsequently considered in the analysis.

- In reconstructing the hadronically decaying top quark, each light quark pair is combined with each identified b-jet and the invariant mass of the three jet system, m_{bjj}, is calculated. Similarly, we construct the invariant mass of each b-jet, lepton and neutrino combination, $m_{bl\nu}$, in order to reconstruct the semi-leptonically decaying top quark. In this procedure, we loop over all b-jets and all p_z^ν solutions.

  Figure 3 and Figure 4 show the reconstructed top quark masses obtained by this procedure for the fully hadronic decay and the semi-leptonic decay respectively.

  For each combination we then calculate a χ^2:

$$\chi^2_{\text{masses}} = \left(\frac{m_W - m_{jj}}{\sigma\left(m_{jj}\right)}\right)^2 + \left(\frac{m_t - m_{bjj}}{\sigma\left(m_{bjj}\right)}\right)^2 + \left(\frac{m_t - m_{bl\nu}}{\sigma\left(m_{bl\nu}\right)}\right)^2$$

  where $\sigma(m_{jj})$, $\sigma(m_{bjj})$, $\sigma(m_{bl\nu})$ are the mass resolutions for the different cases (as shown in Fig 2-4). We obtain:

$$\sigma\left(m_{jj}\right) = 8.5 \text{ GeV/c}^2$$

$$\sigma\left(m_{bjj}\right) = 10.9 \text{ GeV/c}^2$$

$$\sigma\left(m_{bl\nu}\right) = 20.3 \text{ GeV/c}^2$$

- Finally we select the combination with the lowest χ^2_{masses} value and require:

$$\chi^2_{\text{masses}} < 6.$$

- Having assigned two of our identified b-jets to the top quark decays, we calculate $m_{b\bar{b}}$ for the remaining b-jet pair. If more than two identified b-jets remain, we use the jets with the highest p_T to obtain $m_{b\bar{b}}$. This invariant mass distribution is shown in Figure 5 for the $t\bar{t}H$signal and backgrounds for the three input Higgs masses of 115, 125 and 135 GeV/c^2.

b) ZH

To select potential candidates for this process we first demand that at least two identified leptons (electrons or muons) are present in the event. Two leptons must be of the same species but of opposite charge. One of the leptons is required to have a transverse momentum of at least 20 GeV/c and to be isolated from all other tracks by an opening angle of at least 10 degrees. Finally, we require the invariant mass of the lepton pair to be compatible with the Z mass: 60 $< m_{ll} <$ 120 GeV/c^2.

In the present analysis we demand the presence of at least two reconstructed jets in the event, at least one of which must be tagged as a b-jet. One jet must have a transverse momentum of at least 50 GeV/c whilst the second must have transverse momentum greater than 40 GeV/c. Only 1 b-jet is required in order to retain as much signal as possible. In a real analysis, one may choose to demand one tightly tagged b-jet and one loosely tagged b-jet, though the exact final requirements would clearly depend on the b-tagging performance on real data.

We require the reconstructed visible transverse energy to be greater than 150 GeV and the missing transverse momentum to be less than 50 GeV/c.

Since the direction of the b quarks arising from a Higgs decay are expected to be isotropic in the Higgs rest frame, whereas some backgrounds are forward peaked, we cut on:

$\cos\theta_b < 0.8$, where θ_b is the angle between the Higgs direction and the b quark direction, calculated in the Higgs rest frame.

We have also investigated the angular distribution of the Z (determined from the lepton pair) in the centre of mass frame of the collision: i.e. the frame in which the Higgs and the Z are back to back. This distribution is quite different for signal and backgrounds, even after detector effects are taken into account. We cut at:

$\cos\theta_z < 0.7$, where θ_z is the angle between the Z direction and the collision frame direction, boosted into the collision rest frame.

A summary of the number of signal and background events expected in an integrated luminosity of 30 fb^{-1} is shown in Table 4.

c) WH

For the WH selection we demand at least one isolated electron or muon in the final state with a transverse momentum of at least 20 GeV/c. Any second lepton, if present, must have transverse momentum less than 6 GeV/c. We further require at least two reconstructed jets with transverse momentum greater than 25 GeV/c both of which must have been tagged as a b-jet. A summary of the number of signal and background events expected in an integrated luminosity of 30 fb^{-1} is shown in Table 5.

2.2 Results

We have used a mass window of 30 GeV/c^2 around the input Higgs mass to assess the discovery potential of each channel.

a) $t\bar{t}H$

The sensitivity of the $t\bar{t}H$ channel for a Higgs discovery is shown in Table 6 for integrated luminosities of 30 fb^{-1} and 100 fb^{-1}. The total number of events retained after all cuts differs from the analysis in reference [6], in which a comprehensive likelihood analysis is performed. However, we found that cuts can be varied to substantially increase or decrease the statistics whilst keeping $S/\sqrt{B}$ fairly constant. The $S/\sqrt{B}$ obtained with our relatively simple selection is similar to that obtained with the full likelihood based analysis. Figure 6 compares the Higgs sensitivity we have obtained in our analysis with that of Cammin and Schuhmacher [6] as a function of Higgs mass. It is apparent that both analyses achieve similar sensitivities.

b) ZH

The expected number of events in the mass window around the generated Higgs mass is shown in Table 7. It is clear that the backgrounds to this channel are huge. Furthermore, the overwhelming background (Z+Jets) is incompletely modelled and the predicted background from this process is currently very uncertain. We find a sensitivity $S/\sqrt{B} \sim 0.7$ indicating that this is not a promising Higgs discovery channel.

c) WH

Table 8 shows the final number of selected signal and background events for 30 fb^{-1} for this channel. The applied mass window is wider than that used in earlier analyses. Applying the same mass window as that in reference [2] yields similar numbers of signal and background events. The principal background arises from the W+Jets channel, which is currently poorly modelled. We find a sensitivity of ~ 2.0 for this channel. The $b\bar{b}$ invariant mass distribution for signal and background is shown in Figure 7.

3 Conclusion

We have investigated the light Higgs discovery potential in the $t\bar{t}H$,WH and ZH production channels. We find that $t\bar{t}H$ is a promising discovery channel for a light Higgs and find similar sensitivities to other studies.

The WH and ZH investigations indicate that these are marginal discovery channels. A more sophisticated analysis exploiting likelihood functions may improve the situation somewhat. If these channels are to be useful, good mass resolution is vital and efficient b-tagging is needed. Furthermore, a solid understanding of the W+Jets and Z+Jets backgrounds is essential.

References

[1] E. Richter-Was,ATL-PHYS-2000-024.
E.Richter-Was, M. Sapisnki, ATL-PHYS-98-132.

[2] E. Richter-Was, ATL-PHYS-2000-023.

[3] http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/HIGGS/Atlfast.html.

[4] http://borut.home.cern.ch/borut.

[5] http://www.thep.lu.se/ torbjorn/Pythia.html.

[6] J. Cammin and M. Schumacher, ATL-PHYS-2003-024.

sample	$\sigma_{\text{incl.}}$ (pb)	$BR_{H\to b\bar{b}}$	σ_{reduced} (pb)	events generated	L (fb^{-1})
$t\bar{t}H$ m_H=115 GeV/c^2	0.618	0.731	0.452	959000	2122
$t\bar{t}H$ m_H=125 GeV/c^2	0.493	0.608	0.300	913500	3045
$t\bar{t}H$ m_H=135 GeV/c^2	0.396	0.434	0.172	544000	3163
$t\bar{t}b\bar{b}$ (QCD)			8.61	6950000	807
$t\bar{t}b\bar{b}$ (EW)			0.90	160000	178
$t\bar{t}X$			489	21000000	42.9

Table 1: Summary of the simulated event samples used in the $t\bar{t}H$ study.

sample	σ (pb)	events generated
WZ W$\to l\nu Z \to$ udsc	2.87	50 million
WZ W$\to l\nu Z \to b\bar{b}$	0.80	50 million
WZ W$\to JetsZ \to$ ll	1.01	50 million
W+Jets W$\to l\nu$	63000.	500 million
WW W$\to JetsW \to$ udsc	0.86	50 million
ZZ Z$\to llZ \to b\bar{b}$	0.23	50 million
Z+Jets Z$\to ll$	16550.	500 million
$t\bar{t}$ All decay modes.	489.0	100 million
t+X	204.7	100 million
ZH Z$\to llH \to b\bar{b}$	0.04	50 million
WH W$\to l\nu H \to b\bar{b}$	0.23	50 million

Table 2: Summary of the simulated event samples used in the WH and ZH studies.

sample	$t\bar{t}H$ m_H=115 GeV/c^2	$t\bar{t}H$ m_H=125 GeV/c^2	$t\bar{t}H$ m_H=135 GeV/c^2	$t\bar{t}b\bar{b}$ (QCD)	$t\bar{t}b\bar{b}$ (EW)	$t\bar{t}X$
pre-selection	204	142	87	1093	157	1187
$\chi^2 < 6$	133	94	57	659	107	612

Table 3: $t\bar{t}H$: Expected numbers of events (for a luminosity of 30 fb^{-1}).

Cut	Z + Jets	$t\bar{t}$	W + Jets	ZH
Triggered :	1.7 10^8	9.6 10^6	9.7 10^8	923
Lepton Cuts :	2.1 10^7	4.4 10^6	8.7 10^7	869
Jet Cuts :	1.9 10^7	4.3 10^6	8.1 10^7	826
Etvis,Ptmiss:	1.6 10^7	4.2 10^6	6.3 10^7	792
m_{ll} Cut :	8.7 10^5	5.3 10^4	3.0 10^3	371
Angle Cuts :	2.8 10^5	3.4 10^4	1.0 10^3	295

Table 4: ZH: Expected numbers of events (for a luminosity of 30 fb^{-1})

.

Cut	W + Jets	Z + Jets	WH
Triggered :	9.7 10^8	1.7 10^8	6978
2 high p_T Jets :	1.1 10^8	2.4 10^7	4164
Lepton Cuts :	1.0 10^8	1.1 10^7	3753
No 3rd Jet :	7.0 10^7	6.8 10^6	1855
B-Tag :	6.9 10^4	7.6 10^3	356

Table 5: WH: Expected numbers of events (for a luminosity of 30 fb^{-1})

.

test mass mass window	m_H=115 GeV/c^2 85–145 GeV/c^2	m_H=125 GeV/c^2 95–155 GeV/c^2	m_H=135 GeV/c^2 105–165 GeV/c^2
signal (30 fb^{-1})	72	48	27
background (30 fb^{-1})	480	451	409
$S/\sqrt{B}$ (30 fb^{-1})	3.3	2.3	1.3
$S/\sqrt{B}$ (100 fb^{-1})	6.0	4.2	2.4

Table 6: For the $t\bar{t}H$ analysis: expected numbers of events and significances in $m_{b\bar{b}}$ mass windows.

Channel	Number of Events.
ZH	162
ZZ	325
WZ	39
Z + Jets	47465
W + Jets	151
top	4794
WW	0.5
Total Background	52935

Table 7: For the ZH analysis: expected numbers of events in an $m_{b\bar{b}}$ mass window of 30 GeV/c^2.

Channel	Number of Events.
WH	324
WZ	585
W + Jets	21435
Z + Jets	2409
ZZ	41
Top	6924
WW	24
Total Background	30393

Table 8: For the WH analysis: expected numbers of events in an $m_{b\bar{b}}$ mass window of 30 GeV/c^2.

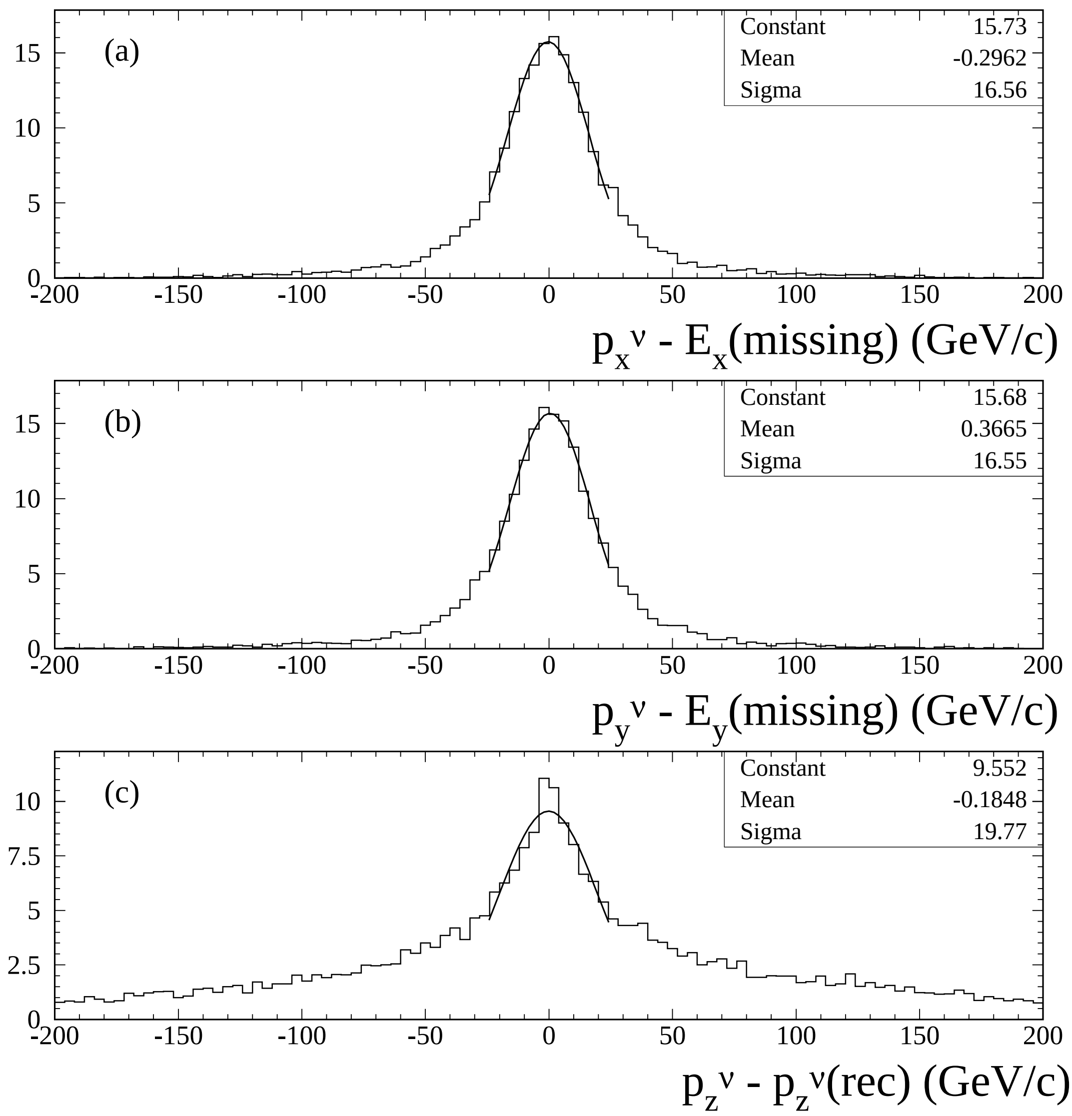

Figure 1: Reconstructed neutrino momentum. The x and y components of the neutrino momentum (figures a and b) are estimated from the total missing momentum in the detector. The z component (figure c) is estimated from solving the mass equation for $m_W = m_{l\nu}$.

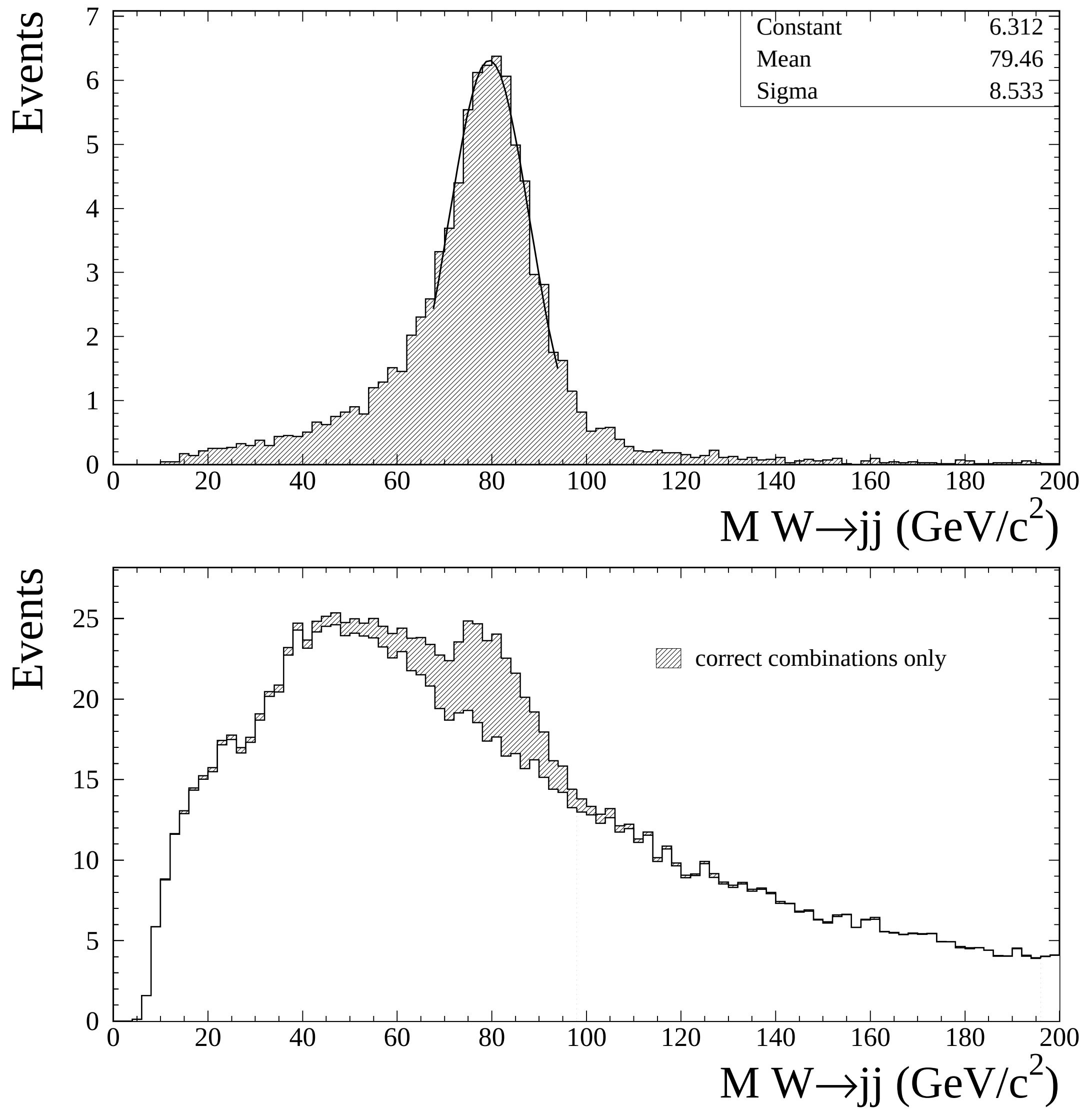

Figure 2: Mass reconstructed from pairs of light-quark jets. In (a) only the correct $W \to jj$ pairings are shown, while in (b) all possible light-quark jet pairings are shown, with the correct pairings indicated by the shaded histogram.

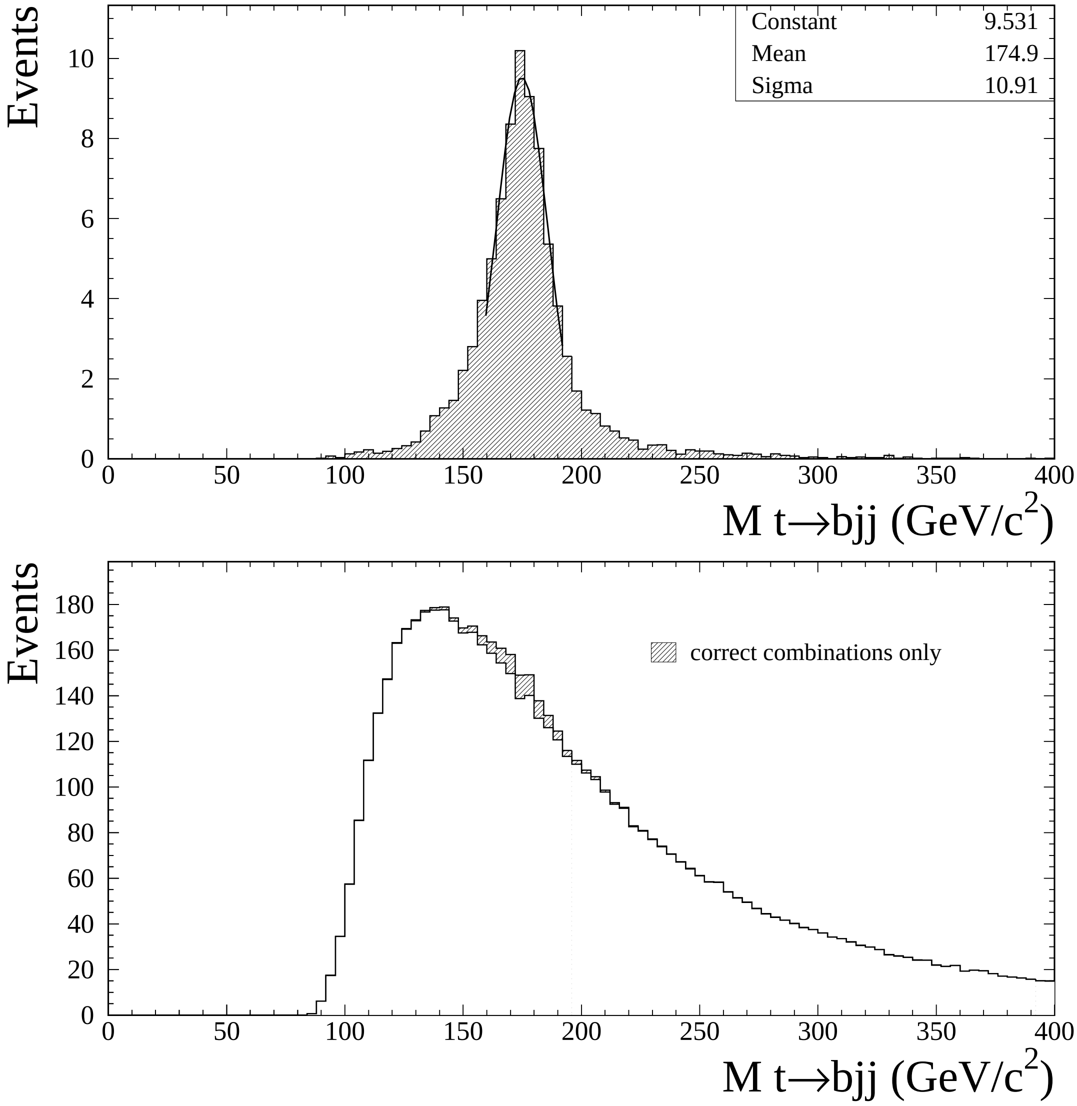

Figure 3: Mass reconstructed from possible combinations of 1 b-quark jet and 2 light quark jets. In (a) only the correct $t \to bjj$ combinations are shown, while in (b) all possible combinations are shown, with the correct ones indicated by the shaded histogram. These plots are for ttH signal events.

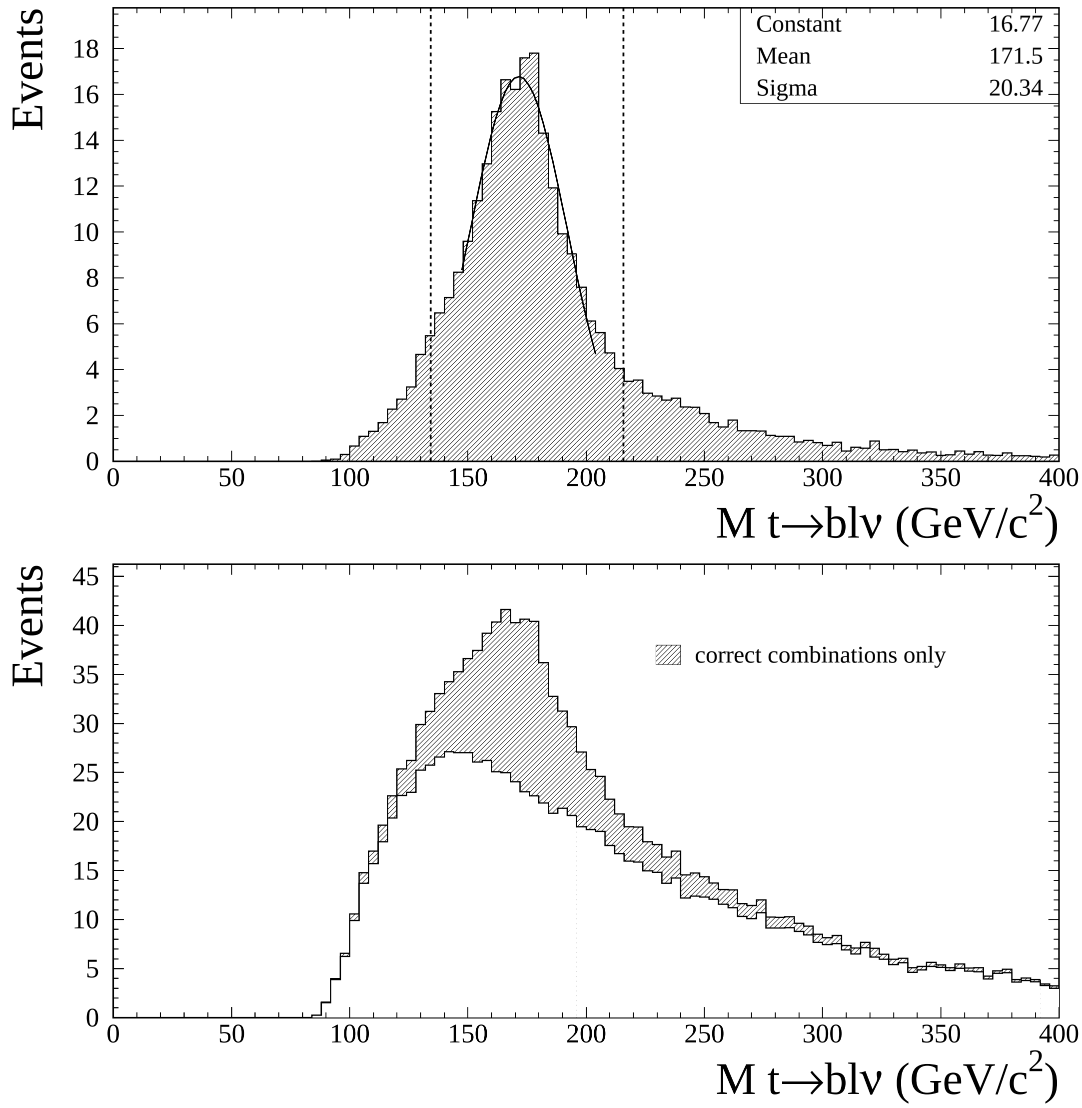

Figure 4: Mass reconstructed from possible combinations of a b-quark jet and the lepton and neutrino. In (a) only the correct $t \rightarrow bl\nu$ combinations are shown, while in (b) all possible combinations are shown, with the correct one indicated by the shaded histogram. These plots are for ttH signal events.

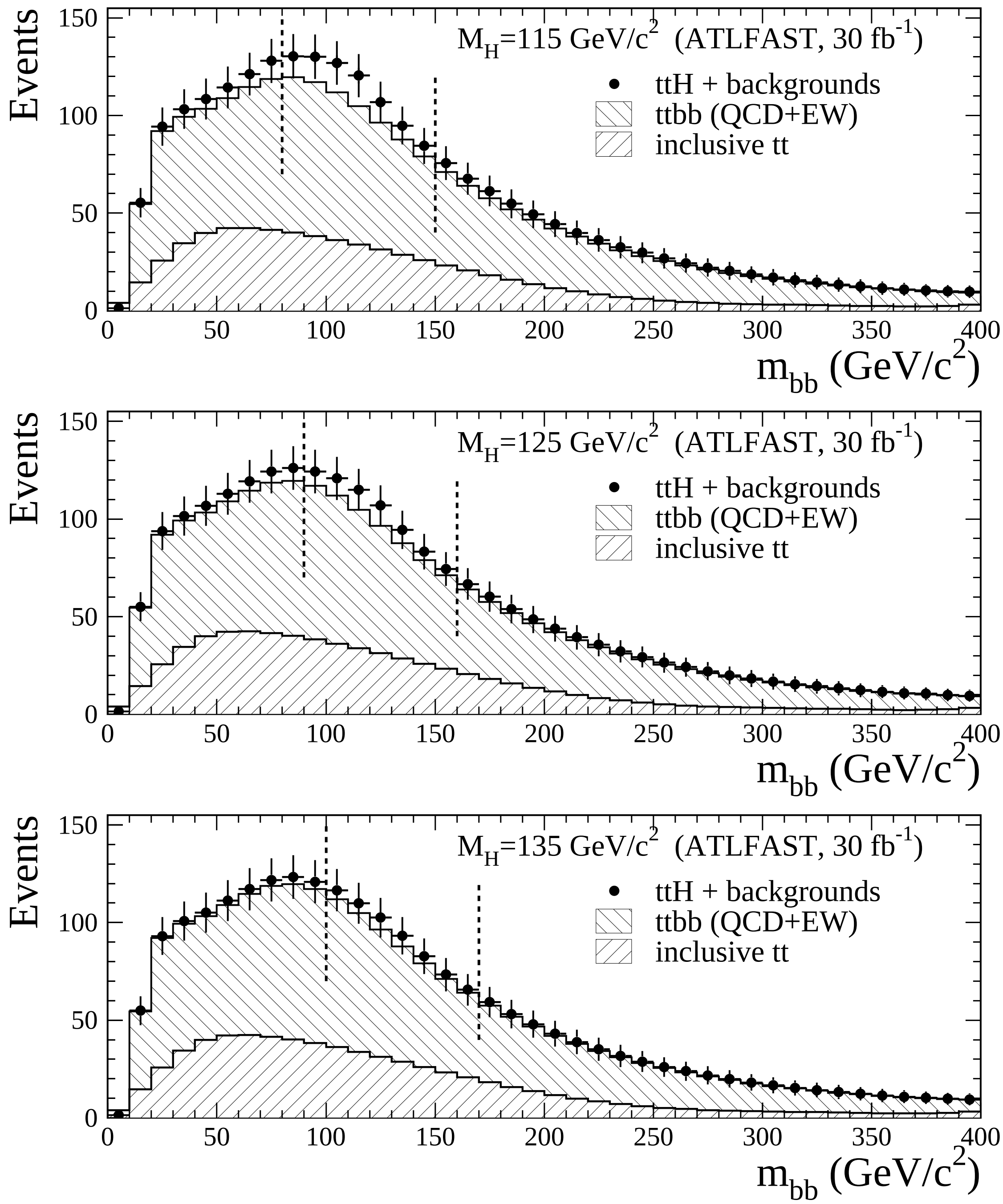

Figure 5: Mass reconstructed from the two highest p_T remaining b-quark jets. The shaded histograms show the distribution from the background processes, as detailed on the plot, while the open points show the contribution from the signal Monte Carlo sample. The Higgs mass used in the signal samples is 115 GeV/c^2 (a), 125 GeV/c^2 (b), 135 GeV/c^2 (c).

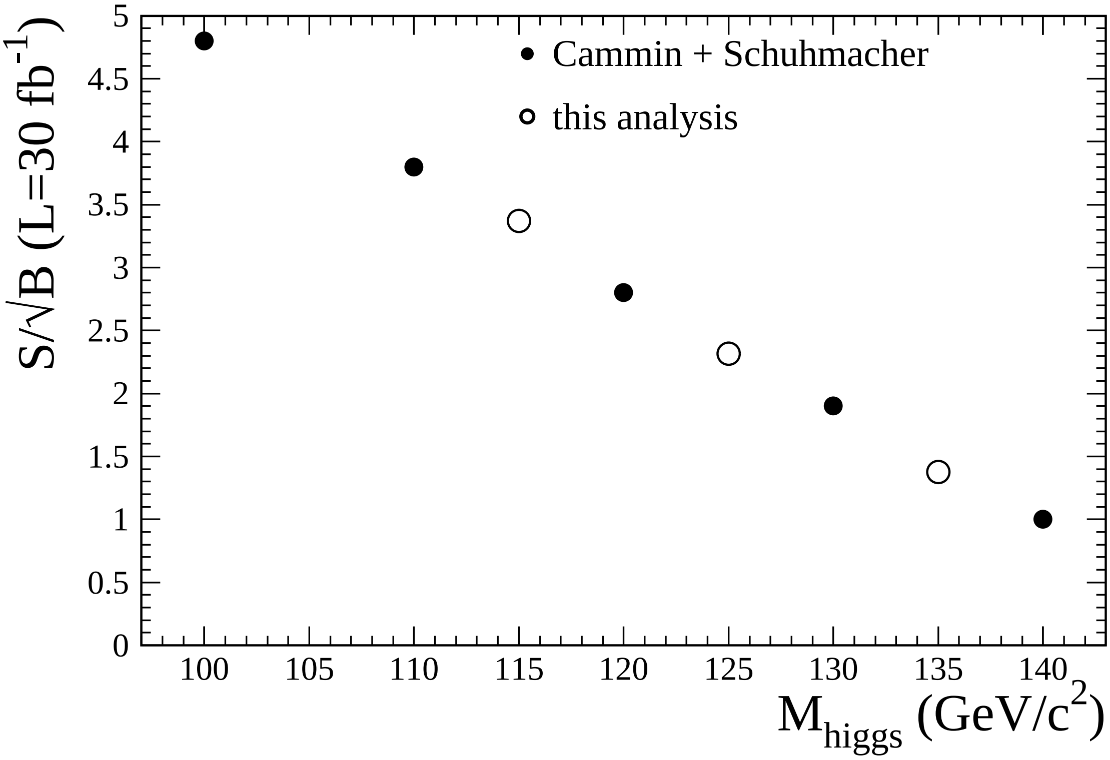

Figure 6: The statistical significance $S/\sqrt{B}$ for different Higgs masses obtained in this analysis and in a complementary analysis, based on an integrated luminosity of 30 fb^{-1}.

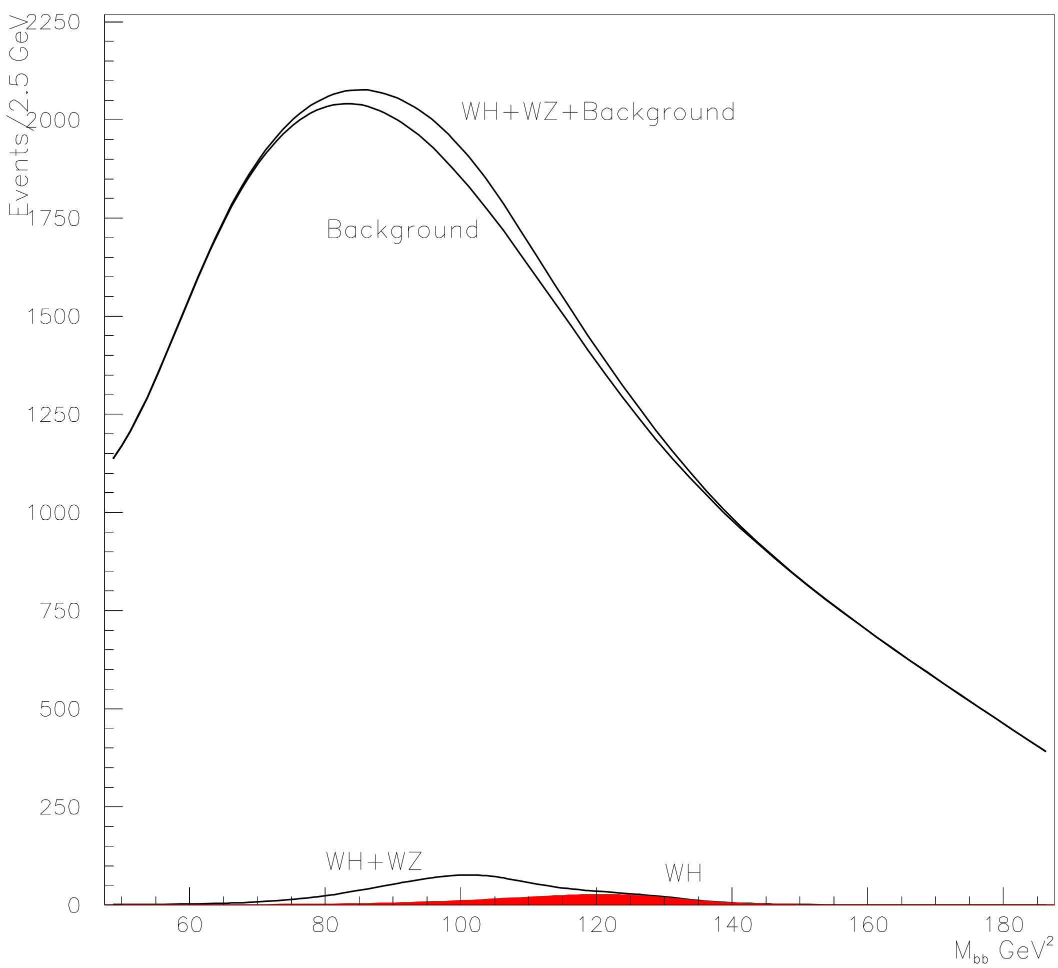

Figure 7: Mass reconstructed from pairs of b-quark jets in the WH analysis. The histogram shows (a) the background contribution, (b) the WH signal contribution, (c) the sum of the WZ and WH contributions, and (d) the total sum of background, WZ and WH contributions.