Search for Higgs boson production in association with a top quark pair in pp collisions [1]

Andrew Brinkerhoff for the CMS Collaboration

University of Notre Dame
Notre Dame, IN

E-mail: andrewbrink@gmail.com

Abstract. We present a search for the standard model Higgs boson produced in association with a top quark pair in 5 fb^{-1} of 7 TeV pp collision data. We look for events where the Higgs boson decays to $b\bar{b}$ and the top quark pair decays to either lepton plus jets ($t\bar{t} \rightarrow \ell\nu jjb\bar{b}$) or dileptons ($t\bar{t} \rightarrow \ell\nu\ell\nu b\bar{b}$), where ℓ is an electron or muon. The major background to our signal is top pair production. We use artificial neural networks to discriminate between background and signal events. We perform a simultaneous fit for signal and background fractions using the neural network output distributions. Our expected limit on Higgs boson production cross section times $H \rightarrow b\bar{b}$ branching ratio for a Higgs boson mass of 125 GeV/c^2 is 4.6 times the standard model expectation, while the observed limit is 3.8 times the standard model expectation.

1. Introduction

In this note, we present a search for Higgs boson production in the $t\bar{t}H$ channel using data taken in 2011 by the CMS detector [2], where $H \rightarrow b\bar{b}$ and the top pair decays semi-leptonically ($t\bar{t} \rightarrow \ell\nu jjb\bar{b}$) or fully leptonically ($t\bar{t} \rightarrow \ell\nu\ell\nu b\bar{b}$). The $t\bar{t}H$ process is interesting for a number of reasons. The rate of this process depends on the couplings of the Higgs boson to both the top quark and the bottom quark, the two largest couplings in the SM. If a Higgs-like excess is observed in the data, measuring the values of these couplings will help to determine whether the object responsible for the excess has the properties expected from a SM Higgs boson. Because of the large mass of the top quark, $t\bar{t}H$ production represents the only opportunity to directly probe the $t\bar{t}H$ vertex without making assumptions on contributions beyond the SM. In addition, $t\bar{t}H$ production provides a probe of the Higgs boson coupling to the b quark that is complementary to the VH channel, with $t\bar{t}$+jets instead of W+jets providing the dominant background contribution. Should the Higgs boson be discovered, $t\bar{t}H$ production will play an important role in determining its properties.

2. Event reconstruction and selection

This analysis selects events consistent with the production of a top quark pair in association with a Higgs boson, where the top quark pair decays by either the lepton + jets ($t\bar{t} \rightarrow \ell\nu jjb\bar{b}$) or dilepton ($t\bar{t} \rightarrow \ell\nu\ell\nu b\bar{b}$) modes and the Higgs boson decays to a pair of bottom quarks. For the lepton + jets case, we select events containing an energetic, isolated, charged electron or muon, and at least four energetic jets, two or more of which should be identified as containing a bottom quark hadron (b-tagged). For the dilepton case, we require a pair of oppositely charged energetic leptons (two electrons, two muons or one electron and one muon) and two or more jets, with at least two of the jets being b-tagged. In both channels, a significant amount of missing transverse energy (MET) should be present as the result of one or more energetic neutrinos. For both channels, we use a common set of criteria for selecting individual objects (electrons, muons, and jets).

Both lepton + jets and dilepton events are required to have one tight muon or electron. The second charged lepton in dilepton events is allowed to be either tight or loose. For lepton + jets events, tight muons are required to have $p_T > 30$ GeV/c and $|\eta| < 2.1$. For dilepton events, tight muons are allowed with $p_T > 20$ GeV/c and $|\eta| < 2.4$. Loose muons must have $p_T > 10$ GeV/c and $|\eta| < 2.4$. The degree of isolation for the muon is assessed by comparing the sum of the transverse energy of particles in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ around the muon direction, excluding the muon itself, to the p_T of the muon. For tight muons, the ratio of the neighboring energy deposits to the muon p_T is required to be less than 0.125. For loose muons, this ratio must be less than 0.2. Tight electrons in lepton + jets events are required to have $E_T > 30$ GeV/c^2, while in dilepton events they must have $E_T > 20$ GeV/c^2. Loose electrons must have $E_T > 15$ GeV/c^2. All electrons are required to have $|\eta| < 2.5$. Electrons that fall into the transition region between the barrel and endcap of the electromagnetic calorimeter ($1.442 < |\eta| < 1.566$) are rejected. Electron isolation is calculated in a similar manner to muon isolation. Tight electrons must have an isolation less than 0.1, while loose electrons may have an isolation less than 0.2.

Jets are required to have corrected $p_T > 30$ GeV/c and $|\eta| < 2.4$. However, in lepton + jets events, because of the jet requirement in the electron trigger, the leading three jets are required to have $p_T > 40$ GeV/c. For consistency, this requirement is also applied to lepton + jets events containing a muon, even though there is no jet component in the trigger for those events. Jets are identified as containing a b quark hadron using the Combined Secondary Vertex (CSV) algorithm. This algorithm combines information about the impact parameter of tracks and reconstructed secondary vertices within the jets in a multivariate algorithm to separate jets containing b quark hadrons from jets containing charm quarks (c), light quarks (uds), or gluons (g). Both categories of events use the medium operating point of the CSV algorithm. For this operating point, the b-tag efficiency is 70% (20%) for jets originating from a b (c) quark, and the probability of mis-tagging for jets originating from light quarks or gluons is 2%.

3. Artificial neural networks

The first step in separating signal from background is to divide the lepton + jets (dilepton) sample into 7 (2) categories based on the number of jets and tags in the event. Artificial neural networks (ANNs) are used to further improve signal-to-background sensitivity in all categories of the analysis. Separate ANNs are trained for different categories, utilizing information related to object kinematics, event shape, and the discriminant output from the b-tagging algorithm. The choice of variables is optimized for each category. In the lepton + jets (dilepton) categories, ten (five) variables are chosen from a pool of candidates based on initial separation between signal and background. In addition, the number of jets is employed as a variable in the dilepton channel for the category $\geq$3 jets + $\geq$3 b-tags.

In both lepton + jets and dilepton channels, the variables that discriminate best between signal and background directly involve b-tagging information, such as the average CSV output value for b-tagged jets. The variables that distinguish best between signal and $t\bar{t} + b\bar{b}$, a background very similar to the signal, are variables that include b-tagging information, kinematic information, and angular correlation information, such as the minimum ΔR between all pairs of b-tagged jets. The modeling of these variables were checked before using them as inputs to the ANN.

4. Results

Our background model matches data well for the event yields and for the shapes of the neural network input variables. We checked the modeling of our input variables in all categories in our search region. In general, we found that all of our input variables were well modeled within the statistics of the events in each category.

To search for the presence of signal in our data, we use a simultaneous maximum-likelihood fit to nine lepton+jets and dilepton neural network output distributions, some of which are shown in Figs. 1 and 2. These plots are input to the CMS standard "Higgs Combination" limit-setting package. The package uses a modified-frequentist approach. Our uncertainties are handled as nuisance parameters in the fit.

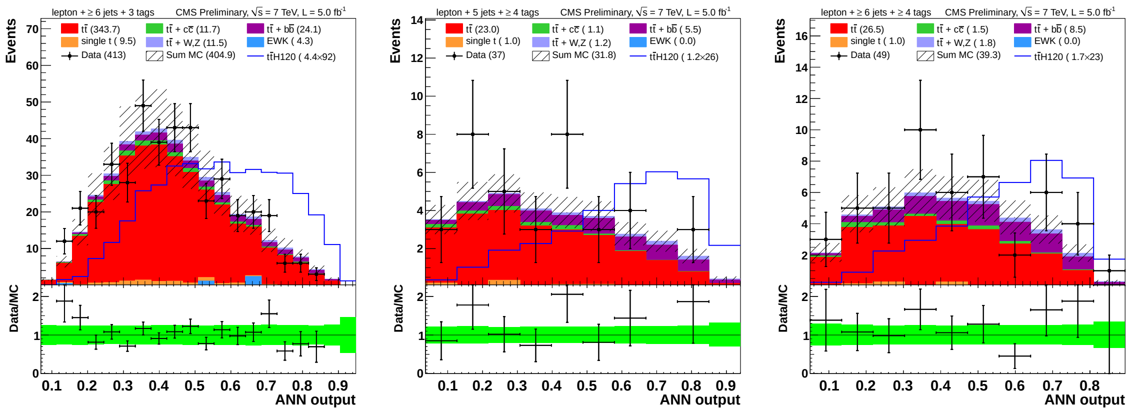

Figure 1. ANN output for lepton + jet events. Background-like events have a low neural network output value. Signal-like events have a high value. The uncertainty band includes statistical and systematic uncertainties that affect both the rate and shape of the background distributions. The plots read from left to right: ≥6 jets + 3 tags, 5 jets + ≥4 b-tags, and ≥6 jets + ≥4 b-tags. The $t\bar{t}H$ signal (m_H = 120 GeV/c^2) is normalized to the total background yield.

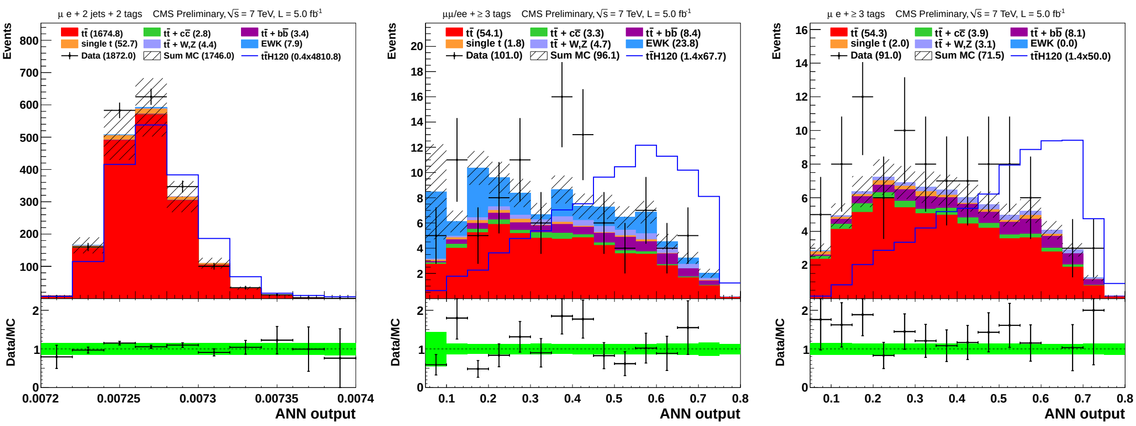

Figure 2. ANN output for dilepton events. The left plot is for $e\mu$ events in the 2 jets + 2 b-tags category. The right plots are for the ≥3 b-tag category, $\mu\mu$ or ee and $e\mu$ events, respectively.

m_H (GeV/c^2)	Obs limit	Median Exp limit	Expected range for 68% CL	Expected range for 95% CL
110	2.3	2.9	[2.1,4.0]	[1.6,5.4]
115	2.8	3.5	[2.4,4.7]	[1.8,6.2]
120	3.1	3.8	[2.8,5.3]	[2.1,7.1]
125	3.8	4.6	[3.3,6.4]	[2.5,8.5]
130	4.4	5.7	[4.1,7.9]	[3.1,10.5]
135	5.7	7.0	[5.1,9.8]	[3.8,13.0]
140	6.6	9.5	[6.8,13.2]	[5.1,17.5]

Table 1. Expected and observed upper limits for SM Higgs boson for the lepton + jets and dilepton channels combined.

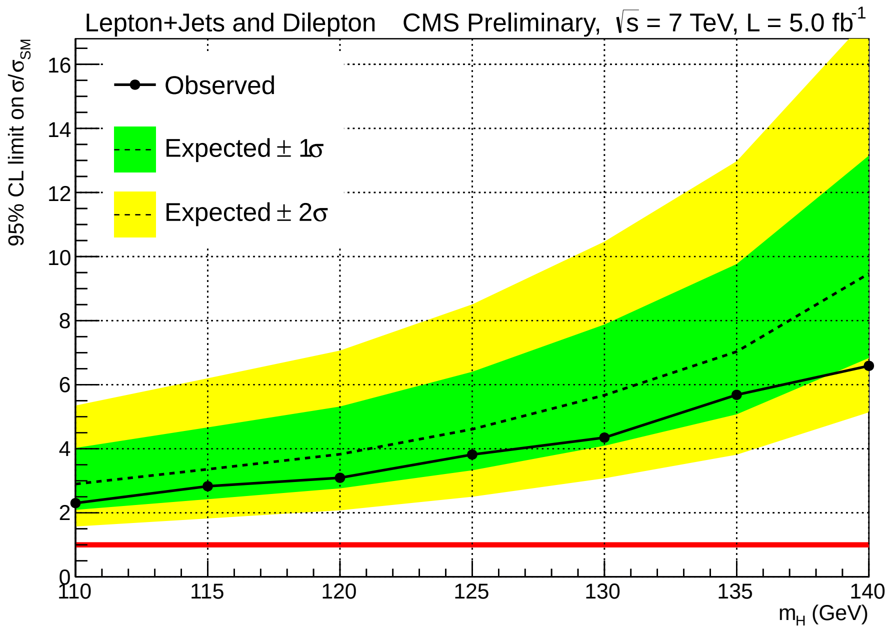

Figure 3. The expected and observed 95% CL upper limits on the signal strength parameter $\mu = \sigma/\sigma_{SM}$ for lepton + jets and dilepton channels combined.

5. Conclusions

The results presented in these proceedings represent the first CMS search for SM Higgs boson production in association with a top quark pair decaying semi-leptonically ($t\bar{t} \to \ell\nu jjb\bar{b}$) or fully leptonically ($t\bar{t} \to \ell\nu\ell\nu b\bar{b}$). Using the full 2011 dataset, we set limits on Higgs boson production comparable in sensitivity to the limits set in the CMS analysis searching for SM Higgs boson production in association with a vector boson [3]. For example, at a Higgs boson mass of 125 GeV/c^2, the expected limit is 4.6 times the SM expectation. The observed limit for $m_H = 125$ GeV/c^2 is 3.8 times the SM expectation.

References

[1] CMS Collaboration (CMS) 2012 Search for higgs boson production in association with a top quark pair in pp collisions CMS Physics Analysis Summary CMS-PAS-HIG-12-025 these TOP2012 conference proceedings are excerpted from the CMS PAS, which can be found at https://cds.cern.ch/record/1460423 URL `https://cds.cern.ch/record/1460423`
[2] CMS Collaboration (CMS) 2008 *JINST* **3** S08004 URL `http://dx.doi.org/10.1088/1748-0221/3/08/S08004`
[3] Chatrchyan S *et al.* (CMS) 2012 *Phys.Lett.* **B710** 284–306 (*Preprint* `1202.4195`)