Searches for the Standard Model Higgs Boson at the Tevatron

D. Bortoletto
Department of Physics, Purdue University
West Lafayette, Indiana, U.S.A.

The results from the search for a standard model Higgs boson using entire data delivered by the Fermilab Tevatron collider are presented. The data corresponding to 10 fb^{-1} of proton-antiproton collisions at a center-of-mass energy of 1.96 TeV were recorded by the CDF and D0 Detectors between March 2001 and September of 2011. A broad excess between $115 < m_H <$ 145 GeV/c^2 with a global significance of 2.2 standard deviations relative to the background-only hypothesis is observed.

1 Introduction

The Higgs boson is a critical missing element of the standard model (SM) of elementary particles and interactions. Within the SM, vector boson masses arise from the spontaneous breaking of electroweak symmetry due to the existance of the Higgs particle. Although the value of the Higgs mass is not predicted by the SM indirect constraints can be set through precision measurements of electroweak observabled such as m_{top} and m_W. These measurements indicates m_H to be less than $\approx$ 145 GeV/c^2 at 95% confidence level (C.L.). The latest results from the LHC and the Tevatron experiments have excluded wide regions of the possible m_H ranges. The most interesting region to search for the Higgs is the mass range between 115 and 127 GeV/c^2 where the both the ATLAS and the CMS experiments have found some excesses with respect to the background-only hypothesis. The Tevatron experiments can contribute to the understanding of this region by studying the production of the Higgs boson in association with an W or a Z boson followed by the decay $H \to b\bar{b}$.

Here, we present the latest results for the $H \to b\bar{b}$ searches at the Tevatron. Both CDF and D0 have developed new advanced analysis techniques to improve the sensititivity of their searches and have almost completed the analysis of the full data sample of 10 fb^{-1} of data delivered by the Tevatron. Multivariate techniques have been implemented to separate signal from QCD and electroweak backgrounds. To obtain the best expected sensitivities to SM Higgs production, the b-tagging, dijet invariant mass, and lepton identification algorithms have been re-optimized to improve the discriminantion of the Higgs signal from background processes. At CDF, a new b-tagging algorithm called HOBIT [8] has been used in most of the mainstream $H \to b\bar{b}$ search channels. This multivariate algorithm, trained on $H \to b\bar{b}$ events for a Higgs mass of $m_H = 120$ GeV/c^2, increases the Higgs sensitivity by roughly 10% for a given search channel. Many analysis improvements were also made by the D0 collaboration, including increasing signal acceptance by relaxing variable definitions and futher optimization of their b tagging algorithm.

2 Low Mass Higgs Searches at the Tevatron

Both CDF and D0 are searching for the Higgs in a variaty of final states [4, 5] . The complete list of channels entering the Higgs Tevatron combination is given in [6], along with a complete description of the limit-setting procedure and handling of systematic uncertainties. The most sensititive low-mass Higgs searches at the Tevatron rely on three optimized analysis according to the decay of the W and Z boson produced in association with the Higgs. The first considers decays of the $Z \rightarrow \ell^+\ell^-$ and therefore requires final states with two leptons. The second requires a lepton from the W decay and transverse missing energy ($\not\!\!E_{\mathrm{T}}$). The last one studies final states with large $\not\!\!E_{\mathrm{T}}$. This includes ZH production where $Z \rightarrow \nu\overline{\nu}$ and the neutrinos ν escape detection or $Z \rightarrow \ell\ell$ when both leptons ℓ are undetected or give rise to jets. For WH production it accepts events where $W \rightarrow e\bar{\nu}_e$ when the electron e is misidentified as a jet, $W \rightarrow e\nu/\mu\nu$ when the e or the muon μ is undetected and $W \rightarrow \tau\bar{\nu}$ when the τ lepton decays hadronically and is detected as a jet.

The $H \rightarrow boverlineb$ search sensitivity is consirably increased by requiring the two leading jets in the event to be b-tagged. Additional sensitivity can be gained by also considering events where one but not both leading jets are tagged. Several analyses also accept events with a third jet in the final state in addition to the two b jets from the Higgs boson decay. The third jet is produced either from radiation from initial or final state partons or when an e or τ from the W boson decay is reconstructed as a jet. Note that the $\not\!\!E_{\mathrm{T}}$ final state at CDF [7] is not yet using HOBIT and it will be updated for the Summer 2012 conferences.

In this contribution I will highlight a few analysis were new techniques or significant improvements have occurred and I will present the results of the latest TEVATRON combination.

2.1 The dilepton final state at CDF

One of the latest major improvement to this CDF analysis has been the utilization of new multivariate algorithms to distinguish between ZH signal and background processes. To isolate ZH signal from $t\bar{t}$ an *expert NN*, trained to distinguish ZH from top is employed. Similarly a second *expert NN* , denoted as Z+jet expert, separates ZH from Z + light flavor jets and $Z + c\bar{c}$ backgrounds. CDF has now introduced a third *expert NN* trained to distinguish ZZ and WZ from ZH signal. The three expert networks are utilized to assign events to distinct regions in the final event discriminant used in the extraction of upper limits.

In addition to the three *expert NN*, an additional network is trained to simultaneously separate ZH signal from all backgrounds. CDF emploies 26 versions of this NN, designated as *final discriminants*, optimized for different values of M_H and separately for 2 and 3 jet events. Once an event receives a region classification, it is evaluated by the final discriminant and assigned to a bin corresponding to the final discriminant score within the region.

CDF does not observe a significant excess over the number of events predicted by the background model and uses MCLIMIT quantify the maximum allowed ZH component. CDF evaluates 95% C.L. upper limits on $ZH \times BR(H \rightarrow b\bar{b})$ and computes observed limits for Higgs masses between 90 and 150 GeV/c^2 in 5 GeV intervals. For a Standard Model Higgs boson mass of 120 GeV, the expected 95% C.L. is 3.1 times the Standard Model prediction with an observed limit of 5.7.

2.2 The $\not\!\!E_{\mathrm{T}}$ final state at D0

The D0 collaboration has significantly refined the b-tagging and the multivariate techniques used in the $\not\!\!E_{\mathrm{T}}$ final state. A multivariate b-tagging discriminant [9], using several boosted decision trees as inputs, is used to select events with one or more b quark candidates. The new algorithm includes more information relating to the lifetime of the jet and results in a better discrimination between b and light jets. It provides an output between 0 and 1 for each jet, with a value closer to one indicating a higher probability that the jet originated from a b quark. From this continuous output, twelve operating points (L_b) are defined, with untagged jets having $L_b = 0$ and b purity increasing with L_b from 1 to 12. The typical per-jet efficiency and fake rate for the loosest (tightest) b-tag operating point are about 80% (50%) and 10% (1%), respectively. To improve the sensitivity of the analysis, two high signal purity samples are defined from the analysis sample using the variable $L_{bb} = L_{b,L} + L_{b,NL}$. A tight (medium) b-tag sample: $L_{bb} \geq 18 (17 \geq L_{bb} \geq 11)$ The medium b-tag sample contains events with two loosely b-tagged jets, as well as events with one tightly b-tagged jet and one untagged jet. The signal-to-background ratios for a Higgs-boson mass of 115 GeV in the pre, medium and tight b-tag samples, after applying a multijet veto, are respectively 0.05%, 0.3% and 1.5%.

Since 50% of the signal in this final states is from WH, D0 has improved the efficiency of this search by excluding isolated tracks from the definition of the missing p_T, a variable similar to $\not\!\!E_{\mathrm{T}}$, calculated from the reconstructed charged particle tracks. The combination of these changes increase the sensititvity by 25% while luminosity alone would have given increase of only 6%. For $m_H = 115$ GeV, the observed and expected limits on the combined cross section of ZH and WH production are factors of 2.5 and 3.0 larger than the theoretical standard-model value, for an expected factor of 3.0.

3 The TEVATRON combination

The results from CDF and D0 on all direct searches for the standard model (SM) Higgs boson have been analyized by the TEVATRON combination group and the current status is presented in [6]. All analyses provide binned histograms of the final discriminant variables for the signal and background predictions. In order to preserve most sensitivity data and predictions are aggregated in bins of signal-to-background ratio, s/b. These distributions can be integrated from the high-s/b side downwards, showing the sums of signal, background, and data for the most pure portions of the selection of all channels added together. The integrated plots of the 100 highest s/b events show an excess consistent with signal for the analyses seeking a Higgs boson mass of 125 GeV/c^2, and a deficit of events in the highest-s/b bins for the analyses seeking a Higgs boson of mass 165 GeV/c^2 as shown in Fig. 1.

To gain confidence that the final result does not depend on the details of the statistical formulation, two types of combinations, using Bayesian and Modified Frequentist approaches are used to find the limits on the Higgs boson production rate. The two techniquese agree within 10% at each value of m_H, and within 1% on average. Systematic uncertainties enter on the predicted number of signal and background events as well as on the distribution of the discriminants in each analysis. Limits on the SM Higgs boson production $\sigma \times B(H \to X)$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV for $100 < m_H < 200$ GeV/c^2 are extracted. The comparisons with the SM is facilitated by dividing them by the SM Higgs boson production cross section, as a function of Higgs boson mass. A value of the combined limit ratio which is less than or equal to one indicates that that particular Higgs boson mass is excluded at the 95% C.L. The combinations of results [4, 5] of each single experiment, as used in the Tevatron combination [6], yield the following ratios of 95% C.L. observed (expected) limits to the SM cross section: 2.37 (1.16) for CDF and 2.17 (1.58) for D0 at m_H= 115 GeV/c^2, 2.90 (1.41) for CDF and 2.53 (1.85) for D0 at $m_H = 125$ GeV/c^2, and 0.42 (0.69) for CDF and 0.94 (0.76) for D0 at $m_H = 165$ GeV/c^2.

With up to 10 fb^1 of luminosity analyzed, the 95% C.L. median expected upper limits on Higgs boson production are factors of 0.94, 1.10, and 0.49 times the values of the SM cross section for Higgs bosons of mass m_H =115 GeV/c^2, 125 GeV/c^2,and 165 GeV/c^2, respectively. The TEVATRON experiments exclude, at the 95% C.L., a new and larger region at high mass

Figure 1: Integrated distributions of s/b, starting at the high s/b side, for Higgs boson masses of 125, and 165 GeV/c^2 . The total signal+background and background-only integrals are shown separately, along with the data sums. Data are only shown for bins that have data events in them.

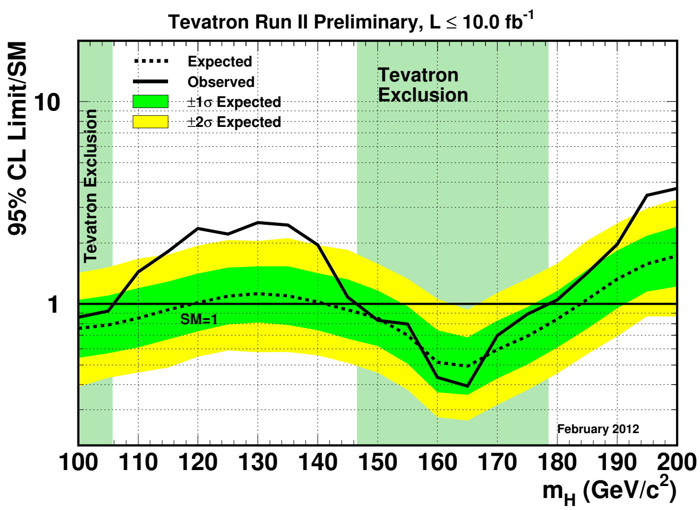

Figure 2: Observed and expected (median, for the background-only hypothesis) 95% C.L. upper limits on the ratios to the SM cross section, as functions of the Higgs boson mass for the combined CDF and D0 analyses. The bands indicate the 68% and 95% probability regions where the limits can fluctuate, in the absence of signal. The limits displayed in this figure are obtained with the Bayesian calculation.

between $147 < m_H < 179$ GeV/c^2, with an expected exclusion region of $100 < m_H < 119$ GeV/c^2 and $141 < m_H < 184$ GeV/c^2. There is an excess of data events with respect to the background estimation in the mass range $115 < m_H < 135$ GeV/c^2 which causes the limits to not be as stringent as expected. At $m_H = 120$ GeV/c^2, the p-value for a background fluctuation to produce this excess is $\approx 3.5 \times 10^3$, corresponding to a local significance of 2.7 σ. The global significance for such an excess anywhere in the full mass range is approximately 2.2 σ. The searches for $H \to b\bar{b}$ and $H \to W^+W^-$ are also combined separately and show that the excess is concentrated in the $H \to b\bar{b}$ channel, although the results in the $H \to W^+W^-$ channel are also consistent with the possible presence of a low-mass Higgs boson.

4 Conclusions

The CDF and D0 Collaborations have combined their results to give a Tevatron-wide combination of the upper limits of the SM Higgs production at 95% C.L. After combining all channels across the range $100 < m_H < 200$ GeV/c^2, a broad excess is observed in data relative to the background-only hypothesis, corresponding to a 2.2 standard-deviatios is found in the region of

M_H between 120-130 GeV/c^2 If one considers only the $H \to b\bar{b}$ final state the excess correspnds to a 2.6 standard-deviation departure from the background-only prediction. The two collaborations are still improving their tools. The final Tevatron combination will be presented in summer 2012.

Acknowledgments

The author thanks the Moriond QCD workshop organizers and conveners for an enjoyable and rewarding experience. She acknowledges support from the DOE and thanks her colleaguses from CDF and D0.

References

1. F. Englert and R. Brout, *Phys. Rev.* D **13**, 321 (1964); P. W. Higgs, *Phys. Rev. Lett.* **13**, 508 (1964); G. S. Guralnik et al., *Phys. Rev. Lett.* **13**, 585 (1964).
2. CMS Collaboration, S. Chatrchyan et al., "Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV", arXiv:1202.1488v1 [hep-ex], submitted to PRL.
3. ATLAS Collaboration, "An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$TeV", 2012. ATLAS-CONF-2012-019
4.] CDF Collaboration, "Combination of CDF's searches for the standard model Higgs boson with up to 10.0 fb^{-1} of data", 2012. CDF Note 10804
5. D0 Collaboration, "Combined Search for the Standard Model Higgs Boson from the D0 Experiment in up to 9.7 fb^{-1} of Data", 2012. D0 Note 6304-CONF
6. The TEVNPH Working Group, for the CDF and D0 Collaborations, "Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data", arXiv:1203.3774v1 [hep-ex].
7. The CDF collboration, "Seach for the SM Higgs boson in the $\not\!\!E_{\rm T}$+b-jets signature with relaxed kinematic cuts in 9.45 fb^1 of data at CDF". CDF Note 10798.
8. J. Freeman et al., "Introduction to HOBIT, a b-Jet Identification Tagger at the CDF Experiment Optimized for Light Higgs Boson Searches", arXiv:1205.1812v1 [hep-ex].
9. L. Breiman et al., Classification and Regression Trees (Wadsworth, Belmont, CA, 1984). Y. Freund and R.E. Schapire, Experiments with a new boosting algorithm, in Machine Learning: Proceedings of the Thirteenth International Conference, Bari, Italy, 1996, edited by L. Saitta (Morgan Kaufmann, San Francisco, 1996), p148.