Search for the Brout-Englert-Higgs Boson at D0

J. Haley for the D0 Collaboration
Department of Physics, Northeastern University,
Boston, MA, USA

We present a search by the D0 collaboration for the standard model Brout-Englert-Higgs boson in the mass range $100 \leq m_H \leq 200$ GeV/c^2. The production modes include gluon-gluon fusion ($gg \to H$), associated production ($q\bar{q} \to VH$), and vector boson fusion ($q\bar{q} \to q\bar{q}H$). An array of distinct final states are analyzed, making use of integrated luminosities up to 9.7 fb^{-1}. We set 95% C.L. upper limits on the production rate relative to the SM of 2.17 (0.94) for $m_H = 115$ (165) GeV/c^2. The expected limit is found to be a factor of 1.58 (0.76) times the standard model prediction for the same mass. We exclude at the 95% C.L. the region $159 < m_H < 166$ GeV/c^2 with an *a priori* expected exclusion of $157 < m_H < 172$ GeV/c^2.

1 Introduction

In the standard model (SM) of particle physics, the mechanism by which electroweak symmetry is broken involves the introduction of a complex doublet of scalar fields that generate the masses of elementary particles via their mutual interactions. This so-called Brout-Englert-Higgs mechanism[1] also gives rise to a single scalar boson (henceforth referred to as simply the Higgs boson, H); however, the mass of the Higgs boson is not predicted. Searches at the Large Electron Positron (LEP) collider provide a lower mass limit[2] of $m_H > 114.4$ GeV/c^2 at 95% confidence level (C.L.). Precision electroweak data yield the indirect constraint[3] $m_H < 161$ GeV/c^2 ($m_H < 185$ GeV/c^2 when also considering the limit from LEP) at 95% C.L., indicating that the range $100 \leq m_H \leq 200$ GeV/c^2 is the most important search region for a SM Higgs boson. Recent results from the CMS and ATLAS experiments now limited the existence of a SM Higgs boson to have a mass between 115.5 GeV and 127 GeV at 95% C.L.[4]

We present the results of combining the direct searches for SM Higgs bosons in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV recorded by the D0 experiment[5]. The analyses seek signals of Higgs bosons produced through gluon-gluon fusion (GGF) ($gg \to H$), in association with vector bosons ($q\bar{q} \to VH$, where $V = W, Z$), and through vector boson fusion (VBF) ($q\bar{q} \to q\bar{q}H$). The analyses utilize data corresponding to integrated luminosities ranging from 4.3 to 9.7 fb^{-1}, collected during the data taking period 2002-2011 (Run II).

1.1 Overview of Contributing Analyses

The Higgs boson channels studied are $WH \to \ell\nu b\bar{b}$ [6], $ZH \to \nu\bar{\nu}b\bar{b}$ [7], $ZH \to \ell\ell b\bar{b}$ [8], $H \to W^+W^- \to \ell^\pm\nu\ell^\mp\nu$ [9], $VH \to e^\pm\mu^\pm + X$ [10], $VH \to ee\mu/\mu\mu e + X$ [11], $VH \to \tau\tau\mu + X$ [12], $H \to W^+W^- \to \ell\nu q\bar{q}$ [13], $H + X \to \mu^\pm\tau^\mp_{had} + \leq 1j$ [14], $H + X \to \ell^\pm\tau^\mp_{had}jj$ [14], $H \to \gamma\gamma$ [15]. Each analysis is further organized into sub-channels based on different production, decay, and/or final state particle configurations, each designed to maximize the sensitivity. To facilitate proper

combination of signals, the analyses were constructed to be mutually exclusive after analysis selections.

To isolate $H \to b\bar{b}$ decays in the $WH \to \ell\nu b\bar{b}$, $ZH \to \nu\bar{\nu}b\bar{b}$, and $ZH \to \ell\ell b\bar{b}$ analyses, an algorithm for identifying jets consistent with containing the decay of a b-quark is applied to each jet (b-tagging). Several kinematic variables sensitive to displaced jet vertices and jet tracks with large transverse impact parameters relative to the hard-scatter vertices are combined in a new boosted decision tree based b-tagging discriminant. Each analysis defines two or three b-tagged sub-channels that depend on how many and how strongly the jets in the event are b-tagged. Typical per-jet efficiencies (fake rates) for the b-tag selections are 50–80% (1–10%).

We also consider Higgs decays to two W bosons for the three dominant production mechanisms: gluon-gluon fusion, associated production, and vector-boson fusion. In the case of production via gluon-gluon fusion and vector-boson fusion, we search for leptonic W boson decays with five final states of opposite-signed leptons: $WW \to e^+\nu e^-\nu$, $e^\pm\nu\mu^\mp\nu$, $\mu^+\nu\mu^-\nu$, $e^\pm\nu\tau^\mp_{had}\nu$ and $\mu^\pm\nu\tau^\mp_{had}\nu$, where τ_{had} denotes a hadronic tau decay. In addition we consider final states originating from Higgs boson production in association with a vector boson (WH or ZH), where leptons may originate from the vector boson or Higgs boson decay. We classify events according to their jet multiplicity in order to isolate particular signal production mechanisms and optimize the discrimination between signal and background. The $H \to W^+W^- \to \ell^\pm\nu\ell^\mp\nu$ ($l = e, \mu$) analyses separate events in three final states with 0 jets, 1 jet, and 2 or more jets. Analyses identifying hadronic tau candidates select events with ≤ 1 jets, mainly sensitive to the gluon-gluon fusion signal, or with ≥ 2 jets, also sensitive to associated production and vector-boson fusion. At high mass, the dominant signal contribution to both tau analyses originates from $H \to W^+W^- \to \mu^\pm\nu\tau^\mp\nu$. At lower masses, the tau analyses requiring at least two jets have significant signal contributions from $ZH \to \tau\tau b\bar{b}$ and $VH \to q\bar{q}\tau\tau$. Another analysis considers the semileptonic decay $H \to W^+W^- \to \ell\nu q\bar{q}$. In all $H \to W^+W^-$ decays with $m_H < 2m_W$, at least one of the W bosons will be off mass shell. For $VH \to VWW$ production, we consider final states with three charged leptons ($ee\mu$, $\mu\mu e$, and $\tau\tau\mu$), as well as the dilepton final state containing an electron and muon with the same charge ($e^\pm\mu^\pm + X$), which benefits greatly from the suppression of Drell-Yan background. Finally, we include an analysis that searches for Higgs bosons decaying to two photons and produced via gluon-gluon fusion, vector boson fusion, and associated production mechanisms.

1.2 Background Estimations

The backgrounds from multijet production are measured in data. The other backgrounds were generated by PYTHIA [16], ALPGEN [17], and COMPHEP [18], with PYTHIA providing parton-showering and hadronization. Drell-Yan, W, and diboson background cross sections are normalized either to next-to-leading order (NLO) calculations from MCFM [19] or, when possible, to data control samples. Top pair and single top production are normalized to approximate next-to-NLO [20] and next-to-next-to-NLO [21] calculations, respectively.

1.3 Signal Predictions

A common approach for the signal predictions and associated uncertainties is followed by both the CDF and D0 Collaborations. The Monte Carlo signal simulation is provided by the PYTHIA generator. We reweight the Higgs boson p_T spectra in our PYTHIA Monte Carlo samples to that predicted by HQT [22] when making predictions of differential distributions of GGF signal events. We normalize our Higgs boson signal predictions to the most recent high-order calculations available. Further details on the signal predictions and uncertainties can be found in the references [5].

2 Limit Calculations

We combine results using the CL_s method with a negative log-likelihood ratio (LLR) test statistic [23]. Separate channels and bins are combined by summing LLR values over all bins and channels. This method provides a robust means of combining individual channels while maintaining individual channel sensitivities and incorporating systematic uncertainties. Systematic uncertainties are treated as Gaussian uncertainties on the expected number of signal and background events. The CL_s approach used in this combination utilizes binned final-variable distributions rather than a single-bin (fully integrated) value for each contributing analysis.

2.1 Systematic Uncertainties

The systematic uncertainties differ between analyses for both the signals and backgrounds. Here we summarize only the largest contributions. Most analyses carry an uncertainty on the integrated luminosity of 6.1% [24], while the overall normalization of other analyses is determined from the NNLO Z/γ^* cross section in data events near the peak of $Z \to \ell\ell$ decays. The $H \to b\bar{b}$ analyses have an uncertainty on the b-tagging rate of 1-10%. The uncertainty on the jet measurement and acceptance is $\sim$ 7%. All analyses include uncertainties associated with lepton measurement and acceptances, which range from 1-9% depending on the final state. The largest contribution for all analyses is the uncertainty on the background cross sections at 4-30% depending on the analysis channel and specific background. These values include both the uncertainty on the theoretical cross section calculations and the uncertainties on the higher order correction factors. The uncertainty on the expected multijet background is dominated by the statistics of the data sample from which it is estimated, and is considered separately from the other cross section uncertainties. The $H \to \gamma\gamma$ and $H \to W^+W^- \to \ell\nu q\bar{q}$ analyses also assign two uncertainties to the NNLO GGF Higgs production cross section associated with the accuracy of the inclusive cross section calculation due to PDF model and scale choice. The $H \to W^+W^- \to \ell^+\nu\ell^-\nu$ ($\ell = e, \mu$) analyses divide the data by jet multiplicity and apply uncertainties for each jet multiplicity final state. In addition, several analyses incorporate shape-dependent uncertainties on the kinematics of the dominant backgrounds in the analyses. These shapes are derived from the potential variations of the final variables due to generator and background modeling uncertainties.

The systematic uncertainties for background rates are generally several times larger than the signal expectation itself and are an important factor in the calculation of limits. Each systematic uncertainty is folded into the signal and background expectations in the limit calculation via Gaussian distributions. These Gaussian values are sampled for each MC trial (pseudo-experiment) using Poisson distributions for the number of signal and background events. Several of the systematic uncertainties, for example the jet energy scale uncertainty, typically impact the shape of the final variable. These shape dependencies were preserved in the description of systematic fluctuations for each Poisson trial. Correlations between systematic sources are carried through in the calculation. For example, the uncertainty on the integrated luminosity is held to be correlated between all signals and backgrounds and, thus, the same fluctuation in the luminosity is common to all channels for a single pseudo-experiment.

3 Results

After setting limits in the individual channels [6–15], we perform a combined measurement using all of the channels to set a limit on the production of a SM Higgs bosons. Because the production and decay channels vary from channel to channel, the combined 95% C.L. upper limit is defined as a ratio to the SM prediction. In other words, the limits are set on the scale factor by which all SM Higgs boson cross sections would have to be scaled to be excluded at 95% C.L. The SM

prediction for Higgs boson production is therefore excluded at 95% C.L. if the limit falls below unity.

The combined limits are evaluated over the range $100 \leq M_H \leq 200$ GeV/c^2 with the contributions from individual analyses varying with the Higgs boson mass. The $H+X \to \ell^{\pm}\tau^{\mp}_{had}jj$ analysis contributes to the region $M_H \geq 105$ GeV/c^2, the $ZH \to \ell\ell b\bar{b}$ $ZH \to \nu\bar{\nu}b\bar{b}$ $WH \to \ell\nu b\bar{b}$ and $H \to \gamma\gamma$ analyses contribute for $M_H \leq 150$ GeV/c^2, the $VH \to \ell^{\pm}\ell^{\pm}+X$analyses contribute for $M_H \geq 115$ GeV/c^2, the $H \to W^+W^- \to (ee, \mu\mu, e\mu)\nu\nu$ analyses contribute for $M_H \geq 115$ GeV/c^2, and the $H \to W^+W^- \to \ell\nu q\bar{q}$ analyses contribute for $M_H \geq 155$ GeV/c^2.

Figure 1(a) shows the expected and observed 95% C.L. cross section limits as a ratio to the SM cross section and for the probed mass range $100 \leq m_H \leq 200$ GeV/c^2, with all analyses combined. The LLR distributions for the full combination are shown in Fig. 1(b). Included in these figures are the median LLR values for the signal-plus-background hypothesis (LLR_{s+b}), background-only hypothesis (LLR_b), and the observed data (LLR_{obs}). The shaded bands represent the one and two standard deviation (σ) departures for LLR_b. These distributions can be interpreted as follows:

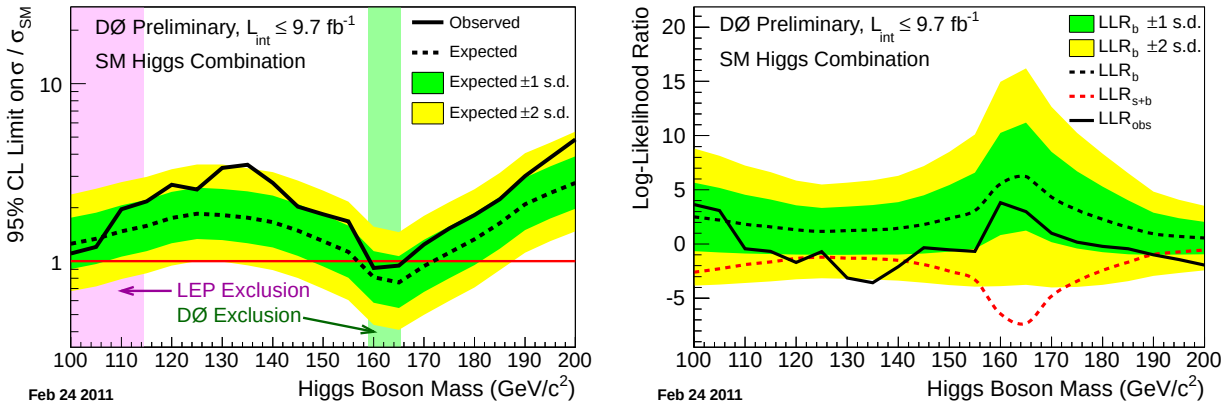
- The separation between LLR_b and LLR_{s+b} provides a measure of the discriminating power of the search. This is the ability of the analysis to separate the $s+b$ and $b-$only hypotheses.
- The width of the LLR_b distribution (shown here as one and two standard deviation (σ) bands) provides an estimate of how sensitive the analysis is to a signal-like background fluctuation in the data, taking account of the presence of systematic uncertainties. For example, the analysis sensitivity is limited when a 1σ background fluctuation is large compared to the signal expectation.
- The value of LLR_{obs} relative to LLR_{s+b} and LLR_b indicates whether the data distribution appears to be more like signal-plus-background or background-only. As noted above, the significance of any departures of LLR_{obs} from LLR_b can be evaluated by the width of the LLR_b distribution.

Figure 2(a) contains the values for the observed 1-CL_b, which is the p-value for the background-only hypothesis. This p-values each provide information on the compatibility of the observed data with the background-only hypothesis. Small values indicate rejection of the hypothesis and values near unity indicate general agreement between the hypothesis in question and the data. The largest deviation from the background-only hypothesis is at $m_H = 135$ GeV/c^2, which has a p-value of 0.0164 corresponding to 2.14 Gaussian standard deviations. This p-value does not include a trials factor or look-elsewhere effect.

As a further investigation of this deviation from the background-only hypothesis, we present in Figure 2(b) the distribution of the best-fit Higgs signal cross section ratio to the SM prediction ($\sigma^{\mathrm{Fit}}/\sigma^{\mathrm{SM}}$). This value is obtained by performing a maximum likelihood fit over all search channels simultaneously, in which the fit is allowed to vary all nuisance parameters within their priors and with the Higgs signal rate as a free parameter. The result indicate a best-fit signal rate of roughly twice the SM Higgs predicted cross section for $m_H = 135$ GeV/c^2. There is also an excursion from zero cross section near $m_H = 200$ GeV/c^2; however, the p-value for the background-only hypothesis is less significant, corresponding to 1.68 Gaussian standard deviations for $m_H = 200$ GeV/c^2.

4 Conclusions

These proceedings presented a summary of the analyses searching for the SM Higgs boson at the D0 experiment. The data analyzed in these searches correspond to integrated luminosities of



(a) (b)

Figure 1: (a) Median expected and observed 95% C.L. cross section upper limit ratios for combined $WH/ZH/H, H \to b\bar{b}/W^+W^-/\gamma\gamma/\tau^+\tau^-$ analyses as a function of Higgs boson mass. (b) The corresponding log-likelihood ratio distribution for the combined analyses. In both plots, the green and yellow bands enclose the one and two standard deviation fluctuations of the background.

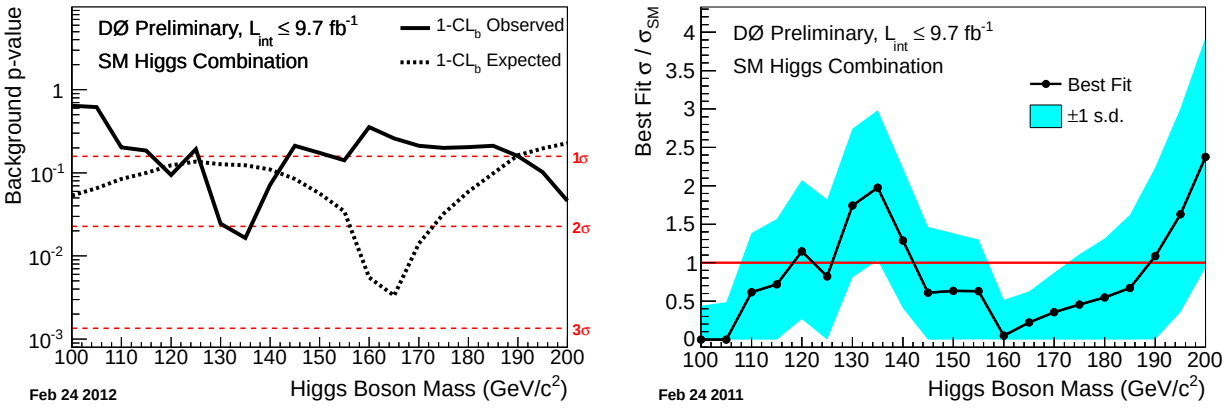
4.3 - 9.7 fb^{-1} from Run II of the Tevatron collider. By combining all of the analyses, we set a 95% C.L. upper limit on the SM Higgs boson cross section of 2.17 (0.94) times the predicted SM cross section for $m_H = 115$ (165) GeV/c^2. The median expected limit is found to be a factor of 1.58 (0.76) times the standard model prediction for the same mass. We exclude at the 95% C.L. the region $159 < m_H < 166$ GeV/c^2 with an *a priori* expected exclusion of $157 < m_H < 172$ GeV/c^2. We are also becoming sensitive to the low mass region where we exclude with 95% C.L. the region $100 < m_H < 105$ GeV/c^2, in agreement with the LEP exclusion. Lastly, we see a signal-like excesses for Higgs masses near both $m_H = 135$ GeV/c^2 and $m_H = 200$ GeV/c^2, with significances of 2.14 and 1.68 Gaussian standard deviations, respectively.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program and NSERC (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).

References

1. F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).
   P. Higgs, Phys. Rev. Lett. **13**, 508 (1964).
   G. Guralnik, C. R. Hagen, T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).
2. R. Barate *et al.* [LEP Working Group for Higgs boson searches], Phys. Lett. B **565**, 61 (2003). [arXiv:hep-ex/0306033].
3. The LEP Electroweak Working Group, http://lepewwg.web.cern.ch/LEPEWWG/.
4. CMS Collaboration, arXiv:1202.1488 [hep-ex] (2012);
   CMS Collaboration, arXiv:1202.4195 [hep-ex] (2012).
   ALTAS Collaboration, Phys. Rev. Lett. **107** 231801 (2011) [arXiv:1109.3615 [hep-ex]];
   ALTAS Collaboration, arXiv:1202.1408 (2012).



(a) (b)

Figure 2: (a) The background p-value ($1 - CL_b$) distribution for the combined $WH/ZH/H, H \rightarrow b\bar{b}/W^+W^-/\gamma\gamma/\tau^+\tau^-$ analyses versus Higgs boson mass. The three horizontal dashed lines indicate the p-values corresponding to significances of 1, 2 and 3 standard deviations. (b) The best-fit signal cross section relative to the standard model Higgs prediction (σ^{Fit}/σ^{SM}) for the combined analyses. This indicates the Higgs cross section that would best match the observed data in a global fit over all nuisance parameters. The Higgs cross section is treated as a free parameter, bounded at zero. The light-blue band indicates the ± 1 standard deviation region from the fit.

5. TEVNPHWG (for the CDF and D0 Collaborations), arXiv:1203.3774 [hep-ex] (2012), arXiv:1107.5518 [hep-ex] (2011), arXiv:1103.3233 [hep-ex] (2011).
6. D0 Collaboration, D0 Note 6309-CONF (2012).
7. D0 Collaboration, D0 Note 6299-CONF (2012).
8. D0 Collaboration, D0 Note 6296-CONF (2012).
9. D0 Collaboration, D0 Note 6302-CONF (2012).
10. D0 Collaboration, D0 Note 6301-CONF (2012).
11. D0 Collaboration, D0 Note 6276-CONF (2012).
12. D0 Collaboration, D0 Note 6286-CONF (2012).
13. D0 Collaboration, Phys. Rev. Lett. **106** 171802 (2011) [arXiv:1101.6079v2 [hep-ph]].
14. D0 Collaboration, D0 Note 6305-CONF (2012).
15. D0 Collaboration, D0 Note 6295-CONF (2012).
16. T. Sjöstrand, L. Lonnblad and S. Mrenna, arXiv:hep-ph/0108264 (2001).
17. M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, JHEP **0307**, 001 (2003) [arXiv:hep-ph/0206293].
18. A. Pukhov *et al.*, arXiv:hep-ph/9908288 (1999).
19. J. M. Campbell, R. K. Ellis, Nucl. Phys. Proc. Suppl. **205-206**, 10-15 (2010) [arXiv:1007.3492 [hep-ph]].
20. N. Kidonakis and R. Vogt, Phys. Rev. D **78**, 074005 (2008).
21. N. Kidonakis, Phys. Rev. D **74**, 114012 (2006).
22. G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, Phys. Lett. B **564**, 65 (2003); G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, Nucl. Phys. B **737**, 73 (2006).
23. T. Junk, Nucl. Intrum. Meth. A **434**, 435 (1999); A.Read, CERN 2000-005 (May 2000).
24. T. Andeen *et al.*, Report No. FERMILAB-TM-2365 (2007).