Search for a Higgs Boson Decaying to Two W Bosons at CDF

Aaltonen, T.
American Institute of Physics
2008


http://hdl.handle.net/10138/23973
http://dx.doi.org/10.1103/PhysRevLett.102.021802

Downloaded from Helda, University of Helsinki institutional repository.
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.
Search for a Higgs Boson Decaying to Two W Bosons at CDF

We present a search for a Higgs boson decaying to two W bosons in pp collisions at $\sqrt{s} = 1.96$ TeV center-of-mass energy. The data sample corresponds to an integrated luminosity of $3.0 \text{ fb}^{-1}$ collected with the CDF II detector. We find no evidence for production of a Higgs boson with mass between 110 and 200 GeV/c$^2$, and determine upper limits on the production cross section. For the mass of 160 GeV/c$^2$, where the analysis is most sensitive, the observed (expected) limit is 0.7 pb (0.9 pb) at 95% Bayesian credibility level which is 1.7 (2.2) times the standard model cross section.

PACS numbers: 13.85.Rm 14.80.Bn
The Higgs boson in the standard model (SM) breaks the electroweak $SU(2)_L \otimes U(1)_Y$ symmetry. While this symmetry is now well established, the mechanism of the symmetry breaking has not yet been identified. Direct searches at the LEP experiments have set a lower limit on the Higgs boson mass $m_H$ of 114.4 GeV/c$^2$ at 95% C.L. in the context of the SM [4]. Precision measurements provide the indirect upper limit $m_H < 144$ GeV/c$^2$ at 95% C.L. through radiative corrections to the SM predictions of the particle masses and couplings [2]. However, these indirect limits assume no significant contributions to the radiative corrections due to as-yet unobserved processes.

In this letter, we report a search for the process $gg \rightarrow H \rightarrow WW(*)$ in a 3.0 fb$^{-1}$ integrated luminosity sample of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV produced by the Fermilab Tevatron and collected by the CDF II detector. For an SM Higgs boson with a mass that is not directly excluded by the LEP experiments, the dominant production mechanism at the Tevatron is gluon-gluon fusion which proceeds via a virtual top quark loop [3]. For $m_H > 135$ GeV/c$^2$, the SM Higgs boson decays primarily to the $WW^*$ [2], where one of the final state $W$ bosons is virtual for $m_H$ below two times the W mass.

The events are reconstructed in the $ll\nu\nu$ final state, whose branching fraction is 6.0% of the $WW^*$ decays, where $l$ is either an electron $e$ or a muon $\mu$, including those from $\tau$ leptons produced in the $W$ decays. The SM Higgs boson branching fraction to $WW^*$ varies from 7.5% at 115 GeV/c$^2$ to 73.5% at 200 GeV/c$^2$ with a maximum of 96.5% at $\approx 170$ GeV/c$^2$ [2]. Previous searches set limits ranging from 10 to 40 times than the predicted SM rate, depending on the value of $m_H$ [4].

The CDF II detector is a multipurpose solenoidal spectrometer surrounded by calorimeters and muon detectors [2]. The geometry is described using the azimuthal angle $\phi$ and the pseudorapidity $\eta \equiv -\ln[\tan(\theta/2)]$, where $\theta$ is the polar angle with respect to the proton beam axis (positive $z$-axis). The transverse energy $E_T$ is $E \sin \theta$, where $E$ is the energy associated with a calorimeter element or energy cluster. Similarly, $p_T$ is the track momentum component transverse to the beam line.

The events we consider must pass one of four online selections, triggers, before being recorded. One electron trigger requires an electromagnetic (EM) energy cluster in the central ($|\eta| < 1.1$) calorimeter with $E_T > 18$ GeV matched to a track found in the drift-chamber with $p_T > 8$ GeV/c. A second electron trigger requires an EM energy cluster with $E_T > 20$ GeV in the forward (1.2 < $|\eta| < 2.0$) calorimeter and the missing transverse energy $E_T > 15$ GeV. The variable $E_T$, used to infer the presence of neutrinos, is defined as $\sum_i E_{T,i}$ and $E_T \equiv |E_T|$, where $E_{T,i}$ is the transverse component of the unit vector pointing from the interaction point to calorimeter element $i$. Muon triggers are based on track segments in the muon chambers matched to a drift-chamber track with $p_T > 18$ GeV/c. Trigger efficiencies are measured using leptonic $W$ and $Z$ data samples [6].

To improve the signal acceptance while maintaining acceptable background rejection for the $W+\text{jets}$ and $W\gamma$ processes where a jet or $\gamma$ is misidentified as a lepton, we use a modified version of the lepton identification strategy developed for the $WZ$ observation analysis [2]. Candidate leptons are separated into six mutually exclusive categories: two for electrons; three for muons; and one for tracks that extrapolate outward to detector regions with insufficient calorimeter coverage for energy measurement. The electron categories are central ($|\eta| < 1.1$) using a drift-chamber-based tracking algorithm and forward ($1.2 < |\eta| < 2.0$) using a silicon-detector-based tracking algorithm. One of the muon categories uses the muon chambers and other two use tracks matched with energy deposits consistent with minimum ionization in the central or forward calorimeters.

All lepton candidates are required to be isolated such that the sum of the $E_T$ for the calorimeter elements in a cone of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.4$ around the lepton is less than 10% of the $E_T$ (for electrons) or $p_T$ (for muons and track lepton candidates). For lepton types in the central region where the track finding efficiency is sufficient, we also apply a track-based isolation criterion which requires there is no more than 10% of the electron $E_T$ or muon $p_T$ in other tracks within a cone of $\Delta R < 0.4$ around the lepton track.

The Higgs boson candidates are selected from events with exactly two lepton candidates. At least one lepton is required to match a trigger lepton candidate and have $E_T > 20$ GeV ($p_T > 20$ GeV/c) for electrons (muons). We loosen this requirement to 10 GeV (GeV/c) for the other lepton to increase the kinematic acceptance, particularly for lower $m_H$ where one $W$ is off-shell below the $W$ boson mass and produces a lower $p_T$ lepton. We also require a dilepton invariant mass $m_{ll} > 16$ GeV/c$^2$ to suppress misidentified multijet events.

Aside from $H \rightarrow WW(*)$ production, other SM processes that can lead to two high-$p_T$ leptons include Drell-Yan (DY), $t\bar{t}$, $WW$, $WZ$, and ZZ production, and $W$ production in association with a photon ($W\gamma$) or a jet ($W+\text{jets}$) misidentified as a lepton. The $t\bar{t}$ contribution is suppressed by requiring fewer than two reconstructed jets with $E_T > 15$ GeV and $|\eta| < 2.5$ in the event. This
allows for up to one jet due to initial state radiation. The DY background is suppressed by requiring sufficiently large $\vec{E}_T$ in the event to remove contributions from mis-measured leptons or jets. This is achieved by requiring $E_{T,rel} > 25$ GeV, where

$$E_{T,rel} \equiv \begin{cases} E_T & \text{if } \Delta \phi_{\vec{E}_T,\ell} > \pi/2 \\ E_T \sin \Delta \phi_{\vec{E}_T,\ell} & \text{if } \Delta \phi_{\vec{E}_T,\ell} < \pi/2 \end{cases},$$

and $\Delta \phi_{\vec{E}_T,\ell}$ is the angle between the $\vec{E}_T$ direction and the nearest lepton or jet. The observed $E_{T,rel}$ is corrected for muons and track-only lepton candidates, because they do not deposit all of their energy in the calorimeter.

The acceptances for the $H \to WW^{(*)}$, $WW$, $WZ$, $ZZ$, $W\gamma$, DY, and $t\bar{t}$ processes are determined using simulated data. Events are simulated with the MC@NLO program for WW [8], PYTHIA for $H \to WW^{(*)}$, DY, $WZ$, $ZZ$, and $t\bar{t}$ [9], and the generator described in Ref. [10] for $W\gamma$. The response of the CDF II detector is then estimated with a GEANT-4-based simulation [11] to which an efficiency correction of up to 10% per lepton is applied based on measurements of the lepton reconstruction and identification efficiencies using observed $Z \to \ell^+\ell^-$ events. An additional correction is applied to the $W\gamma$ background estimate based on a measurement in data of the efficiency of requirement that supresses the fraction of photon contamination described below and find it to be consistent with the expected. The largest predicted SM Higgs boson yield is 11.6 events at $m_H=160$ GeV/$c^2$ which results from an acceptance of 16%. Table I shows the dependence of acceptance on the lepton identification categories, this information is used to divide the sample into high and low signal-to-background ($S/B$) classes.

After selection, the dominant background is the $q\bar{q} \to WW$ process which differs from the signal process only by spin, production mechanism, and resonant structure. Because of the two neutrinos in the final state, a simple mass peak cannot be used to isolate the resonant $qg \to H \to WW^{(*)}$ process from the backgrounds. Instead, we combine two different multivariate techniques to differentiate signal and background. One is a matrix element (ME) technique, which uses an event-by-event calculation of the probability density for each contributing process to produce the observed event. If all details of the collision properties and the detector response are modeled in the ME calculation, this method provides the optimal sensitivity to the signal. However, there are several approximations used in the calculations: theoretical differential cross sections are only implemented to leading order, a simple parameterization of the detector response is used, and for some small ($WZ$ and $t\bar{t}$) difficult to model (DY) backgrounds, we do not calculate a probability density. In order to improve these approximations, we extend the ME calculation with a neural network (NN) which exploits the more complete model implemented in the simulated data and $W$+jets model.

The event probability density for the ME method is

$$P(\vec{x}_{obs}) = \frac{1}{\langle \sigma \rangle} \int \frac{d\sigma_{LO}(\vec{y})}{d\vec{y}} \epsilon(\vec{y}) G(x_{obs}, \vec{y}) d\vec{y},$$

where the elements of $\vec{y}$ ($\vec{x}_{obs}$) are the true (observed) values of the lepton momenta and $E_{\vec{T}}$, $d\sigma_{LO}/d\vec{y}$ is the parton level differential cross section [12], $\epsilon(\vec{y})$ is a parameterization of detector acceptance and efficiency function, and $G(x_{obs}, \vec{y})$ is the transfer function representing the detector resolution and a PYTHIA-based estimate of transverse momentum of the $\ell\vec{E}_T$ system due to the initial state radiation. The constant $\langle \sigma \rangle$ normalizes the total event probability to unity. This calculation integrates the theoretical differential cross section over the missing information due to two unobserved neutrinos in the final state. We form a likelihood ratio discriminant which is the signal probability divided by the sum of signal and background probabilities

$$LR_H - WW^{(*)}(\vec{x}_{obs}) \equiv \frac{P_H(\vec{x}_{obs})}{\sum_k P_k(\vec{x}_{obs})}$$

where $k_i$ are the expected background fractions of $WW$, $ZZ$, $W\gamma$, and $W$+jets. The $LR$ distributions are shown in Figure II. Additional ME likelihood ratios $LR_{WW}$, $LR_{ZZ}$, $LR_{W\gamma}$, and $LR_{W+}$ are defined analogously to $LR_H - WW^{(*)}$.

For the final results, an NN discriminant is used to extend the ME calculation using as input the ME likelihood ratios in addition to various kinematic variables. For each of the Higgs boson masses investigated, an NN is trained on signal events and an appropriately weighted compo-
sition of background events. The NN classifies events as high SM expectation. The distributions are shown only for the high S/B class, which provides the majority of the sensitivity.

The systematic uncertainties associated with the Monte Carlo simulation affect all of the signal and background expectations similarly, except for W+jets. The uncertainties from the lepton selection and trigger efficiency measurements are propagated through the analysis, giving uncertainties from 1.4% to 2.0% and 2.1% to 7.1% for the respective efficiencies of the different signal mass points and background processes. The detector acceptance variation due to parton-distribution function (PDF) uncertainties is assessed to be 1.9% - 4.1% using the 20 pairs of PDF sets described in Ref. [13].

The cross section uncertainties are 10% for WW [12], ZZ [12], and Wγ [17], and 15% for t¯t [18]. Based on a comparison of simulated WW events generated with Mc@nlo and PYTHIA, we assign a systematic uncertainty on the acceptance due to higher order QCD effects of 5.5% for WW events and 10% for the other modes which are only simulated at leading order.

The systematic uncertainty on the W+jets background is estimated to be 24% from differences in the observed probability that a jet is identified as a lepton for jets collected using different jet E_T trigger thresholds. These variations correspond to changing the parton composition of the jets and the relative amount of contamination from real leptons. Because only the E_T requirement suppresses the DY background, there is an uncertainty due to the E_T resolution modeling, which is estimated to be 20% based on comparisons of the data and Monte Carlo simulation in a sample of dilepton events. For the Wγ background contribution, there is an additional uncertainty of 20% from the detector material description and photon-conversion veto efficiency.

A Bayesian credibility level (C.L.) is calculated for each m_H hypothesis based on the combined binned likelihood of the discriminant distributions for the high and low S/B samples. A posterior density is obtained by multiplying this likelihood by Gaussian prior densities for the background normalizations and systematic uncertainties leaving σ(gg → H) × B(H → WW(∗)) with a uniform prior density. A 95% C.L. limit is then determined such that 95% of the posterior density for σ(gg → H) × B(H → WW(∗)) falls below the limit. Limits as a fraction of the SM cross section σ_SM(gg → H) are calculated by including σ_SM(gg → H) as a parameter whose prior density is determined by the 10% theoretical uncertainty of its next-to-next-to-leading-log prediction [19]. The resulting cross section limits for both discriminants are shown in Table II and Figure 2.

In conclusion, we have presented limits on the production of a Higgs boson through gluon fusion followed by its decay to a pair of W bosons. A combination of matrix element and neural network techniques is used to discriminate signal from background. Studies using the

![FIG. 1: The likelihood ratio distribution from a) the ME discriminate alone and b) the full NN score for m_H = 160 GeV/c^2. The Higgs boson distribution is normalized to the SM expectation. The distributions are shown only for the high S/B class, which provides the majority of the sensitivity.](image-url)
two techniques independently achieve consistent results with a sensitivity approximately 0.1σ_{SM} worse than the combination at \(m_H = 160 \text{ GeV}/c^2\). The consistency of results obtained with different algorithms provides evidence of the robustness of the multivariate techniques. At the most sensitive value of \(m_H = 160 \text{ GeV}/c^2\), the observed limit is 1.7 times the SM prediction where the median expected limit is 2.2, corresponding to a downward fluctuation slightly larger than one standard deviation. Compared to an optimized selection and a likelihood based on the \(\Delta \phi_{\ell \ell}\) variable, the multivariate discriminators gain a factor of 1.7 to 2.5 in effective integrated luminosity depending on \(m_H\). This measurement also constrains alternative models in which the \(gg \to H\) coupling is enhanced by additional particles in the virtual loops of the production amplitude [20].

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucléaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, Spain; the Slovak R&D Agency; and the Academy of Finland.

---