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


(CDF Collaboration)

1Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China
2Argonne National Laboratory, Argonne, Illinois 60439, USA
3University of Athens, 157 71 Athens, Greece
4Institut de Fisica d’Altes Energies, Universitat Autonoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain
5Baylor University, Waco, Texas 76798, USA
6Istituto Nazionale di Fisica Nucleare Bologna, University of Bologna, I-40129 Bologna, Italy
7Brandeis University, Waltham, Massachusetts 02254, USA
8University of California, Davis, Davis, California 95616, USA
9University of California, Los Angeles, Los Angeles, California 90024, USA
10University of California, San Diego, La Jolla, California 92093, USA
11University of California, Santa Barbara, Santa Barbara, California 93106, USA
12Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain
13Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
14Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA
15Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia
16Joint Institute for Nuclear Research, RU-141980 Dubna, Russia
17Duke University, Durham, North Carolina 27708, USA
18Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
19University of Florida, Gainesville, Florida 32611, USA
20Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy
21University of Geneva, CH-1211 Geneva 4, Switzerland
22Glasgow University, Glasgow G12 8QQ, United Kingdom
23Harvard University, Cambridge, Massachusetts 02138, USA
24Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
25University of Illinois, Urbana, Illinois 61801, USA
We present a search for a Higgs boson decaying to two W bosons in \( p\bar{p} \) collisions at \( \sqrt{s} = 1.96 \text{ TeV} \) center-of-mass energy. The data sample corresponds to an integrated luminosity of 3.0 fb\(^{-1} \) collected with the CDF II detector. We find no evidence for production of a Higgs boson with mass between 110 and 160 GeV, and determine upper limits on the production cross section. For the mass of 144 GeV 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 \( \sqrt{s} = 1.96 \text{ TeV} \) produced by the Fermilab
The events are reconstructed in the $ll\nu\nu$ final state, whose branching fraction is 6.0% of the $WW^* \rightarrow ll\nu\nu$ 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 $m_H = 170$ GeV [3]. Previous searches set limits ranging from 10 to 40 times 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 [5]. The geometry is described using the azimuthal angle $\phi$ and the pseudorapidity $\eta = -\ln \tan(\theta/2)$, 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 $\slashed{E}_T > 15$ GeV. The variable $\slashed{E}_T$, used to infer the presence of neutrinos, is defined as $\sum E_{T,i} n_{T,i}$ and $\slashed{E}_T = |\slashed{E}_T|$, where $n_{T,i}$ 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 $W +$ 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 [7]. 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 the 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 is off shell below the $W$ boson mass and produces a lower $p_T$ lepton. We also require a dilepton invariant mass $m_{\ell\ell} > 16$ 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 ($W +$ 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. The DY background is suppressed by requiring sufficiently large $\slashed{E}_T$ in the event to remove contributions from mismasured leptons or jets. This is achieved by requiring $\slashed{E}_T > 25$ GeV, where $\slashed{E}_T$ is $E_T \sin \Delta \phi_{E_T,i,\text{jet}}$ if $\Delta \phi_{E_T,i,\text{jet}} < \pi/2$ and $\slashed{E}_T$ otherwise, and $\Delta \phi_{E_T,i,\text{jet}}$ is the angle between the $\slashed{E}_T$ direction and the nearest lepton or jet. The observed $\slashed{E}_T$ is corrected for muons and track-only lepton candidates, because they do not deposit all of their energy in the calorimeter. Except for $W +$ jets, acceptances are determined using data simulated with the MC@NLO program for $WW$ [8], PYTHIA for $H \rightarrow WW^{(*)}$, 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 GEANT4-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 $Z \rightarrow l^+l^-$ events. An additional correction is applied to the $W\gamma$ background estimate based on a measurement in data of the photon-conversion veto efficiency. The $W +$ jets contribution is estimated from data by extrapolating from a sample of events that contain an identified lepton and a jet $j_i$ containing a track or EM energy cluster similar to those required in the lepton identification. The contribution of each event to the total yield is scaled by the probability that
the $j_i$ is identified as a lepton. This probability $p(p_T; j_i)$ is determined for each lepton type as a function of the $j_i$ $p_T$ using multijet events collected with jet-based triggers. A correction to $p(p_T; j_i)$ is applied for the small real lepton contribution using Monte Carlo simulation of single $W$ and $Z$ boson production.

Based on the procedure above, we expect $768 \pm 91$ background events and observe 779 in the selected region. The expected composition of the background is $356 \pm 49$ $WW$, $24.9 \pm 3.9$ $WZ$, $21.8 \pm 3.5$ $ZZ$, $25.5 \pm 5.0$ $t\bar{t}$, $138 \pm 31$ $DY$, $90.5 \pm 24.1$ $W\gamma$, and $111 \pm 27$ $W + \text{jets}$, where the indicated uncertainties include the systematic uncertainties described below. As a cross-check of the background model, we measure the fraction of $q\bar{q} \to WW$ events in the sample with a similar method to the signal extraction described below and find it to be consistent with the expectation. Table I shows the dependence of expected yield on $m_H$. Since the level of non-$WW^{(\ell\ell)}$ background depends on the lepton identification categories, this information is used to divide the sample into high and low signal-to-background ($S/B$) classes. The low $S/B$ category largely consists of candidates with a forward lepton.

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 parameterization of the detector response is used, and for some small ($WZ$ and $t\bar{t}$) or 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 $W + \text{jets}$ model.

The event probability density for the ME method is

$$P(\tilde{x}_{\text{obs}}) = \frac{1}{\langle \sigma \rangle} \int \frac{d\sigma_{LO}(\tilde{y})}{d\tilde{y}} e(\tilde{y}) G(\tilde{x}_{\text{obs}}, \tilde{y}) d\tilde{y},$$

where the elements of $\tilde{y}$ ($\tilde{x}_{\text{obs}}$) are the true (observed) values of the lepton momenta and $E_T$, $d\sigma_{LO}/d\tilde{y}$ is the parton level differential cross section [12], $e(\tilde{y})$ is a parameterization of detector acceptance and efficiency function, and $G(\tilde{x}_{\text{obs}}, \tilde{y})$ is the transfer function representing the detector resolution and a PYTHIA-based estimate of transverse momentum of the $\ell\ell E_T$ 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 $LR_{H \to WW^{(\ell\ell)}}(\tilde{x}_{\text{obs}}) = \frac{P_{H}(\tilde{x}_{\text{obs}})}{\sum_{k} P_{k}(\tilde{x}_{\text{obs}})}$, where $k_i$ are the expected background fractions of $WW$, $ZZ$, $W\gamma$, and $W + \text{jets}$. The $LR$ distributions are shown in Fig. 1(a). Additional ME likelihood ratios $LR_{WW}$, $LR_{ZZ}$, $LR_{W\gamma}$, and $LR_{W + \text{jets}}$ are defined analogously to $LR_{H \to WW^{(\ell\ell)}}$.

For the final results, a 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, a NN is trained on signal events and an appropriately weighted composition of background events. The NN classifies events as signal or background based on the inputs $LR_{H \to WW^{(\ell\ell)}}$, $LR_{WW}$, $LR_{ZZ}$, $LR_{W\gamma}$, $LR_{W + \text{jets}}$, $\Delta \phi_\ell\ell$, $\Delta R_\ell\ell$, $m_{\ell\ell}$, $E_T$, $\Delta \phi_E_{\text{jet},\ell\ell}$, and $\Delta E_{\text{T,rel}}$, where $\Delta \phi_\ell\ell$ and $\Delta R_\ell\ell$ are the separation between the two leptons in $\phi$ and $\Delta R$, respectively. We find that the most discriminating input variables are $LR_{H \to WW^{(\ell\ell)}}$, $\Delta R_{\ell\ell}$, and $E_T$. An example of the NN output is shown in Fig. 1(b). The presented results use the NEUROBAYES [13] program. Comparable results are also obtained using the TMV A-MLP [14] program, demonstrating the robustness of the technique.

For the signal and backgrounds modeled by Monte Carlo simulation, the same procedures as Ref. [7] are used to assess the systematic uncertainties on the lepton selection efficiency, trigger efficiency, parton-distribution function, and luminosity to be $1.4\%$ to $2.0\%$, $2.1\%$ to $7.1\%$, $1.9\%$ to $4.1\%$, and $6\%$, respectively, depending on mode. The cross section uncertainties are $10\%$ for WW [12], $WZ$ [12], $ZZ$ [12], and $W\gamma$ [15], and $15\%$ for $t\bar{t}$ [16]. 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.

\begin{table}[h]
\centering
\begin{tabular}{l|cccccccccc}
\hline
\hline
$m_H$ (GeV/c^2) & 110 & 120 & 130 & 140 & 150 & 160 & 170 & 180 & 190 & 200 \\
\hline
Expected yield & 0.5 & 1.9 & 4.3 & 7.0 & 9.3 & 11.6 & 11.0 & 9.0 & 6.4 & 5.1 \\
\hline
\end{tabular}
\caption{Expected Higgs boson events as a function of $m_H$.}
\end{table}
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 $E_T^{\text{rel}}$ 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 $W\gamma$ 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 $\sigma(gg \to H) \times B(H \to WW^{(*)})$ with a uniform prior density. A 95% C.L. limit is then determined such that 95% of the posterior density for $\sigma \times B$ falls below the limit. Limits as a fraction of the SM cross section $\sigma(gg \to H)$ are calculated by including $\sigma_{\text{SM}}(gg \to H)$ as a parameter whose prior density is determined by the 10% theoretical uncertainty of its next-to-next-to-leading-log prediction [17]. The resulting cross section limits for both discriminants are shown in Table II and Fig. 2. Because the NN uses the ME calculation as input, it is the full result; the ME results are given only for comparative purposes.

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 two techniques independently achieve consistent results with a sensitivity approximately 0.1 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 1 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

<table>
<thead>
<tr>
<th>$m_H$ (GeV/$c^2$)</th>
<th>110</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using Matrix Element Only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected (pb)</td>
<td>3.6</td>
<td>2.6</td>
<td>2.2</td>
<td>1.9</td>
<td>1.5</td>
<td>0.9</td>
<td>0.9</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>Observed (pb)</td>
<td>2.8</td>
<td>1.5</td>
<td>1.1</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.7</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Expected/SM</td>
<td>63.7</td>
<td>19.6</td>
<td>9.4</td>
<td>6.0</td>
<td>4.3</td>
<td>2.4</td>
<td>2.6</td>
<td>3.8</td>
<td>6.0</td>
<td>8.2</td>
</tr>
<tr>
<td>Observed/SM</td>
<td>50.3</td>
<td>16.9</td>
<td>3.0</td>
<td>2.3</td>
<td>1.7</td>
<td>1.8</td>
<td>2.6</td>
<td>5.0</td>
<td>10.3</td>
<td>13.0</td>
</tr>
<tr>
<td>Using Neural Net Discriminator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected (pb)</td>
<td>3.0</td>
<td>2.3</td>
<td>1.9</td>
<td>1.7</td>
<td>1.4</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Observed (pb)</td>
<td>2.5</td>
<td>1.7</td>
<td>1.2</td>
<td>1.1</td>
<td>0.9</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>1.0</td>
<td>1.6</td>
</tr>
<tr>
<td>Expected/SM</td>
<td>54.0</td>
<td>17.1</td>
<td>8.4</td>
<td>5.4</td>
<td>3.9</td>
<td>2.2</td>
<td>2.4</td>
<td>3.5</td>
<td>5.6</td>
<td>7.7</td>
</tr>
<tr>
<td>Observed/SM</td>
<td>44.6</td>
<td>13.2</td>
<td>5.3</td>
<td>3.5</td>
<td>2.6</td>
<td>1.7</td>
<td>2.2</td>
<td>2.7</td>
<td>5.5</td>
<td>10.6</td>
</tr>
</tbody>
</table>
the $gg \rightarrow H$ coupling is enhanced by additional particles in the virtual loops of the production amplitude [18].

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 Nucleaire 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.

\[ \text{FIG. 2 (color online). Upper limits on } \frac{\sigma \times B}{\sigma \times B}_{\text{SM}} \text{ versus } m_H. \]

\[2\text{Higgs Mass [ GeV/c}^2]\]

\[\begin{array}{lllllll}
110 & 120 & 130 & 140 & 150 & 160 & 170 & 180 & 190 & 200 \\
\text{Standard Model} & \text{ME Observed} & \text{ME Median Expect} & \text{NN Observed} & \text{NN Median Expect} & \text{NN Median Expect } \pm 1\sigma & \text{NN Median Expect } \pm 2\sigma \\
\end{array}\]