Search for the Higgs Boson Produced in Association with $Z \to \ell^+\ell^-$ in $pp$ Collisions at $s^\sqrt{=1.96}$ TeV

CDF Collaboration

CLARK, Allan Geoffrey (Collab.), et al.

Abstract

We present a search for the Higgs boson in the process $qq\to ZH\to \ell^+\ell^-bb$. The analysis uses an integrated luminosity of 1 fb$^{-1}$ of $pp$ collisions produced at $s^\sqrt{=1.96}$ TeV and accumulated by the upgraded Collider Detector at Fermilab (CDF II). We employ artificial neural networks both to correct jets mismeasured in the calorimeter and to distinguish the signal kinematic distributions from those of the background. We see no evidence for Higgs boson production, and set 95% C.L. upper limits on $\sigma_{ZHB}(H\to bb)$, ranging from 1.5 to 1.2 pb for a Higgs boson mass ($m_H$) of 110 to 150 GeV/c².

Reference

CDF Collaboration, CLARK, Allan Geoffrey (Collab.), et al. Search for the Higgs Boson Produced in Association with $Z \to \ell^+\ell^-$ in $pp$ Collisions at $s^\sqrt{=1.96}$ TeV. Physical review letters, 2008, vol. 101, no. 25, p. 251803

DOI : 10.1103/PhysRevLett.101.251803
Search for the Higgs Boson Produced in Association with $Z \rightarrow \ell^+ \ell^-$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV  

We present a search for the Higgs boson in the process $qar{q} \rightarrow ZH \rightarrow \ell^+\ell^- b\bar{b}$. The analysis uses an integrated luminosity of 1 fb$^{-1}$ of $p\bar{p}$ collisions produced at $\sqrt{s} = 1.96$ TeV and accumulated by the upgraded Collider Detector at Fermilab (CDF II). We employ artificial neural networks both to correct jets mismeasured in the calorimeter and to distinguish the signal kinematic distributions from those of the background. We see no evidence for Higgs boson production, and set 95% C.L. upper limits on $\sigma_{ZH}B(H \rightarrow b\bar{b})$, ranging from 1.5 to 1.2 pb for a Higgs boson mass ($m_H$) of 110 to 150 GeV/$c^2$. 

The Higgs boson is the only particle predicted by the standard model (SM) of particle physics which has not yet been discovered. It is the physical manifestation of the mechanism which provides mass to fundamental particles [1,2]. Direct searches have excluded the SM Higgs boson for masses $m_H < 114.4$ GeV/$c^2$ at the 95% C.L. [3]. The Higgs boson mass is indirectly constrained from precise electroweak measurements to $m_H = 76^{+33}_{-22}$ GeV/$c^2$ [4]. A number of extensions to the SM predict a SM-like Higgs boson; in particular, Ref. [5] predicts a SM-like Higgs boson with 68% posterior probability to be between 115.4 and 120.4 GeV/$c^2$. Only the Tevatron collider experiments are currently capable of extending the limits on a Higgs boson for $m_H > 114.4$ GeV/$c^2$.

This Letter presents the first CDF II search for a Higgs boson in the process $p\bar{p} \rightarrow Z' \rightarrow ZH \rightarrow \ell^+\ell^- b\bar{b}$, where $\ell$ is $e$ or $\mu$, with a data set of 1 fb$^{-1}$, almost 3 times that of the previously reported analysis [6]. CDF and D0 have previously presented Higgs boson searches in other decay modes [7–13].

CDF II [14] is a general purpose detector. Its coordinate system and quantities used throughout this Letter are defined in Ref. [15]. At its center is a cylindrical silicon detector which tracks charged particles from a radius of 1.35 to 29 cm for $|\eta| \leq 2$. Around this is a cylindrical wire drift chamber which tracks charged particles from 43 to 132 cm for $|\eta| \leq 1.3$. A superconducting solenoid surrounds the tracking volume providing a 1.4 T magnetic field for momenta measurements. Segmented electromagnetic and hadronic sampling calorimeters surrounding the solenoid measure energies of interacting particles with $|\eta| < 3.6$. A system of drift chambers and scintillation counters outside the calorimeters detect muon candidates for $|\eta| < 1.5$.

At the Tevatron the cross section for $q\bar{q} \rightarrow Z' \rightarrow ZH$ production for a Higgs boson with mass $m_H = 115$ GeV/$c^2$ is 1.04 pb [16], and the branching ratio $B(H \rightarrow b\bar{b})$ is 73% [17]. To identify candidate $ZH$ events, we first search for $Z$ candidates decaying to electron or muon pairs. The full selection criteria are described in Ref. [18]; the most salient features are described here. Events are collected using a trigger which identifies a primary electron (muon) with $E_T > 18$ GeV ($p_T > 18$ GeV/$c$) within the central region $|\eta| < 1.0$. The requirements for the second electron are relaxed to $E_T > 10$ GeV in the central region, and maintained at 18 GeV for electrons with $1.0 < |\eta| < 2.4$. The second muon must have $p_T > 10$ GeV/$c$. Energy deposits from leptons must be isolated from other energy deposits within $\Delta R < 0.4$.

From the measured lepton energies and momenta, we reconstruct the invariant mass of the $Z$ candidate and require it to be between 76 and 106 GeV/$c^2$. This requirement is 92% efficient for real $ZH$ candidates, but helps remove non-$Z$ backgrounds. We require oppositely charged leptons for muon pairs and for electron pairs when the second electron has $|\eta| < 1.0$ due to improved tracking efficiency in the central region.

Higgs candidates are then selected by requiring a jet with $E_T > 25$ GeV and an additional jet with $E_T > 15$ GeV, both with $|\eta| < 2.0$. Jets are corrected for calorimeter response, multiple interactions, and energy loss in the uninstrumented detector regions [19]. To enhance signal significance, we implement an algorithm to identify the decay of a long-lived hadron containing a $b$ quark by reconstructing a significantly displaced secondary vertex [20,21]. The efficiency is 40%–50% for $b$-quark jets and 1%–3% for $u, d, s, or g$ (light parton (LP)) jets. We consider events in which one or both jets are “tagged” by this algorithm.

Backgrounds originating from $W, Z$, and $t\bar{t}$ production are determined using leading order Monte Carlo (MC) calculations, normalized to next to leading order, followed by a detailed simulation of the CDF II detector. We model the $Z + b\bar{b}, Z + c\bar{c}$, and $Z + LP$ processes by first producing the exact leading order multiparton final states with the ALPGEN [22] MC program, and then using HERWIG [23] to model the hadronization and parton showering. In addition we use an inclusive PYTHIA [24] Z MC sample to compare with the observed data and evaluate systematic uncertainties. We model $ZZ$, $ZW$, and $t\bar{t}$ background contributions using PYTHIA. A “fake lepton” background arises from jets being misidentified as leptons, and we estimate this contribution from observed data [25]. The contribution from false tags of LP jets is evaluated from data by applying a parametrization of the false tagging rate to jets passing $E_T$ and $\eta$ requirements [20,21].

The acceptance, with statistical uncertainties, of $ZH \rightarrow \ell^+\ell^- b\bar{b}$ events is (10.8 ± 0.1)% for $m_H = 120$ GeV/$c^2$ and is evaluated using the PYTHIA MC program followed by a detailed CDF II simulation for Higgs boson masses from 110 to 150 GeV/$c^2$.

Table I shows the expected background and signal contributions with systematic uncertainties in 1 fb$^{-1}$ of data compared to the number observed after dividing events

<table>
<thead>
<tr>
<th>Sample</th>
<th>Single tagged</th>
<th>Double tagged</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z + b\bar{b}$</td>
<td>35.1 ± 14.6</td>
<td>6.3 ± 2.5</td>
</tr>
<tr>
<td>$Z + c\bar{c}$</td>
<td>21.8 ± 8.5</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>$Z + LP$</td>
<td>32.3 ± 5.5</td>
<td>1.0 ± 0.2</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>5.2 ± 1.0</td>
<td>2.8 ± 0.6</td>
</tr>
<tr>
<td>$ZZ$</td>
<td>4.0 ± 0.8</td>
<td>1.3 ± 0.3</td>
</tr>
<tr>
<td>$ZW$</td>
<td>1.2 ± 0.2</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>Non-$Z$</td>
<td>1.9 ± 1.4</td>
<td>0.2 ± 0.2</td>
</tr>
</tbody>
</table>

**Table I.** Expected and observed numbers of events in 1 fb$^{-1}$ for electron and muon decay modes combined, compared to the expected $ZH$ signal for $m_H = 120$ GeV/$c^2$.  

$Z + b\bar{b}$  $35.1 \pm 14.6$  $6.3 \pm 2.5$  
$Z + c\bar{c}$  $21.8 \pm 8.5$  $1.0 \pm 0.4$  
$Z + LP$        $32.3 \pm 5.5$  $1.0 \pm 0.2$  
$t\bar{t}$      $5.2 \pm 1.0$   $2.8 \pm 0.6$  
$ZZ$            $4.0 \pm 0.8$   $1.3 \pm 0.3$  
$ZW$            $1.2 \pm 0.2$   $0.04 \pm 0.01$  
Non-$Z$         $1.9 \pm 1.4$   $0.2 \pm 0.2$  

Expected  $101.5 \pm 32.0$  $12.7 \pm 4.1$  
Observed  $100$  $11$  
$ZH$       $0.44$   $0.23$  

251803-4
into those with only one $b$ tag (single tagged), and those with exactly two $b$ tags (double tagged). Electron decay modes account for 60% of the total expected and observed events.

The $Z +$ jets system is only expected to have $E_T$ due to jet mismeasurement, either from the limited calorimeter resolution, uninstrumented regions, or from the semileptonic decay of the jets. Therefore, we correct the jet energies using a multilayer perceptron artificial neural network (ANN) [26] which uses the missing transverse energy $E_T$ vector projected onto those of the jets in order to determine individual scale factors for each jet. The result is a dijet mass resolution improvement from 18% to 11% [18] which improves $ZH$ signal discrimination from the backgrounds.

To achieve a greater separation of signal and background we employ an additional ANN implemented with JETNET [27] to distinguish the kinematics of the signal from those of the backgrounds. Our ANN configuration is 8 input variables, 17 hidden nodes, and 2 output nodes, such that the output distribution is 2D, with one axis separating $ZH$ and $Z + b\bar{b}$, and the other axis separating $t\bar{t}$ and $ZH$. The variables chosen, in order of importance for minimizing the classification error, are the scalar sum of the transverse energies of the jets and leptons composing the Higgs and $Z$ candidates ($H_T$), $E_T$, dijet mass, $\Delta R$ between first jet and $Z$ candidate, $\Delta R$ between subleading jet and $Z$ candidate, $\Delta R$ between leading and subleading jets, sphericity, which is a measure of how isotropic the leptons and jets are, and $\eta$ of the subleading jet. The most important distributions are shown in Fig. 1.

The uncertainty on the amount of $Z + b\bar{b}$ and $Z + c\bar{c}$ background is taken to be 40% [25], and for the $t\bar{t}$, $WZ$, and $ZZ$ [28,29] it is 20%, including the uncertainties on the cross section and on the selection efficiencies of these processes and the top quark mass uncertainty. The uncertainty on the non-$Z$ background is 50%. The uncertainty on the shape of the background is evaluated by comparing $Z + b\bar{b}$ events between PYTHIA and ALPGEN. The signal

![Figure 1](image1.png)

**FIG. 1.** Expected and observed distributions for the three most important inputs to the signal discriminating ANN shown after ANN jet corrections have been applied, for events with 1 $b$ tag.

![Figure 2](image2.png)

**FIG. 2.** Expected and observed distributions for the ANN discriminant projected onto the axis which discriminates $Z +$ jets from $ZH$, after enriching signal by selecting the most signal-like 25% of events as determined by the $t\bar{t}$ vs $ZH$ ANN output. The top plot is the ANN discriminant before $b$-tagging requirements, demonstrating two models for the $Z +$ jets background normalized to data for shape comparison. The middle and lower plots are for single-tagged and double-tagged events, and are shown with a $ZH$ signal at the level of the observed 95% C.L. upper limit for $m_H = 120$ GeV/c$^2$. 

251803-5
TABLE II. Upper limits at the 95% C.L. on the cross section of $\sigma(ZH) \times B(H \rightarrow bb)$. Also shown is the expected and observed ratio of the limit compared to the SM cross section.

<table>
<thead>
<tr>
<th>$m_H$ GeV/c$^2$</th>
<th>Observed $\sigma$/[pb]</th>
<th>Observed $\sigma/\sigma_{SM}$</th>
<th>Expected $\sigma/\sigma_{SM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>1.5</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>115</td>
<td>1.4</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>120</td>
<td>1.2</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>130</td>
<td>1.2</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>140</td>
<td>1.2</td>
<td>65</td>
<td>55</td>
</tr>
<tr>
<td>150</td>
<td>1.2</td>
<td>160</td>
<td>140</td>
</tr>
</tbody>
</table>

shape uncertainty is evaluated by varying the amount of initial and final state QCD radiation [30], and by changing the parton distribution functions using the 40 eigenvectors from CTEQ6 [31]. We evaluate both rate and shape uncertainties for the signal and backgrounds by varying the jet energy scale within its uncertainties [19]. In addition, both signal and background estimates are affected by the trigger efficiency uncertainty (1%), and the luminosity measurement uncertainty (6%) [32]. The b-tagging efficiency has an uncertainty of 8% for b-quark jets, 16% for c-quark jets, and 13% for LP jets. In double-tagged events, the uncertainties are 16%, 32%, and 24%, respectively. These values are updates obtained from the procedure found in Ref. [20].

The projections of the 2D ANN signal discriminant are shown in Fig. 2. We analyze the binned 2D ANN discriminant distribution test for a ZH signal in the presence of SM backgrounds using a Bayesian technique [33] and marginalize over variations in the systematic uncertainties. The expected and observed upper limits at the 95% C.L. are shown in Table II. Expected limits are obtained by generating pseudoexperiments from the expected SM ANN shapes to calculate the median ZH contribution which could be excluded at the 95% level with no ZH signal present. Since backgrounds are smaller for double-tagged events, we analyze them separately from single-tagged events, resulting in a 20% improvement in the expected limits.

The dominant systematic uncertainty is the b-quark identification efficiency which accounts for 12% of the total 14% increase in the expected limit due to systematic uncertainties.

In summary, we have extended the limits for a Higgs boson decaying to $b\bar{b}$ produced in association with a Z boson. This is the first Tevatron run II search for a Higgs boson to use a multivariate approach to separate signal and background kinematics, and results in a significant improvement in Higgs sensitivity over previous analyses in this decay mode [6]. The observed event kinematics and ANN signal discriminants show no significant excess above SM predictions. The improvement in limits using our approach of two ANNs is a factor of 1.8 compared to a fit of the uncorrected dijet mass distribution alone. This result finds the best limit on standard model Higgs production in the most favored Higgs boson mass range of less than 125 GeV/c$^2$ that was achieved with 1 fb$^{-1}$ of data as in Ref. [18]. Full sensitivity for this mass range will be achieved by combining this analysis with other CDF and D0 Higgs search channels and the combined Tevatron data set.

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, U.K.; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Educación y Ciencia and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

---

[a]Deceased.
[b]Visitor from Universiteit Antwerpen, B-2610 Antwerp, Belgium.
[c]Visitor from Chinese Academy of Sciences, Beijing 100864, China.
[d]Visitor from University of Bristol, Bristol BS8 1TL, United Kingdom.
[e]Visitor from University of California Irvine, Irvine, CA 92697, USA.
[f]Visitor from Cornell University, Ithaca, NY 14853, USA.
[g]Visitor from University of Cyprus, Nicosia CY-1678, Cyprus.
[h]Visitor from University College Dublin, Dublin 4, Ireland.
[i]Visitor from University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom.
[j]Visitor from Universidad Iberoamericana, Mexico D.F., Mexico.
[k]Visitor from University of Manchester, Manchester M13 9PL, England.
[m]Visitor from University de Oviedo, E-33007 Oviedo, Spain.
In the CDF coordinate system, the $z$ axis is the proton direction, $\theta$ is the polar angle, and $\phi$ is the azimuthal angle. Pseudorapidity $\eta$ is defined as $-\ln[\tan(\theta/2)]$. The transverse momentum of a particle is $p_T = p \sin \theta$, where $p$ is its momentum. The transverse energy $E_T$ is an analogous quantity defined from calorimeter energies and coordinates as $E_T = E \sin \theta$. The missing transverse energy $E_{\text{miss}}$ is defined as $-\sum |E_i^T \hat{n}_i|$ where $\hat{n}_i$ is the radial unit vector pointing from the interaction point to the energy deposition in a given calorimeter cell $i$. $\Delta R$ is defined as $\sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}$ and used to specify particle separation.