Search for H $\rightarrow WW^*$ Production at CDF Using 3.0 fb^{-1} of Data

Text for the public web page – CDF note 9500

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: August 25, 2008)

We report a search for standard model Higgs boson to WW^* production in the two charged lepton (e,μ) and two neutrino final state in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The data were collected with the CDF II detector at the Tevatron collider at Fermilab and correspond to an integrated luminosity of approximately 3.0 fb^{-1}. In order to maximize sensitivity, three neural networks are separately trained to distinguish signal from background processes in final states with either zero, one, or two or more jets. For 0-jet events, likelihood ratios based on Matrix Element calculations are used as additional inputs to the neural network to further strengthen its discriminating power. In order to take advantage of the full signal acceptance, we also consider associated production with a W or Z boson and Higgs boson production via vector boson fusion for final states with one or more jets. We find a total of 918 candidates with an expectation of 894.8 ± 113.2 background and 17.5 ± 2.2 signal for a Higgs mass of 160 GeV/c^2. The median expected 95% C.L. limit for a Higgs mass of 160 GeV/c^2 is $1.7^{+0.7}_{-0.5}$ times the standard model prediction, while the observed limit is 1.6 times the standard model prediction.

PACS numbers:

I. INTRODUCTION

The Higgs boson is introduced into the standard model to explain electroweak symmetry breaking and the origins of particle mass. Current precision electroweak measurements constrain the mass of a standard model Higgs boson to be less than 160 GeV/c^2 (one-sided 95% C.L. upper limit) or 190 GeV/c^2 when including the LEP-2 direct search lower bound limit of 114 GeV/c^2 at 95% C.L. [1]. For this analysis we focus on the Higgs decaying to WW^*, which is the dominant decay channel for $m_H > 135$ GeV/c^2 [2]. The small cross section of the dominant gluon fusion production mechanism ($\sigma_{NNLL}(gg \to H) = 0.431$ pb for m_H =160 GeV/c^2 [3]) makes observation of the signal difficult within the hadron collider environment.

In this note, we present a search of $gg \to H \to WW^* \to l^+l^-\nu\bar{\nu}$, where $l^\pm =$ e, μ or τ with final states e^+e^-, $e^\pm\mu^\mp$ and $\mu^+\mu^-$. For events with one or more jets, we extend our search to include Higgs bosons produced from vector boson fusion, as well as associated production with a Z boson ($ZH \to ZWW^*$) or a W boson ($WH \to WWW^*$). We use approximately 3.0 fb^{-1} of integrated luminosity collected by the CDF II detector at the Fermilab Tevatron with $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV.

The analysis strategy is to use a common event selection for 14 different Higgs mass hypotheses between 110 GeV/c^2 and 200 GeV/c^2. For each higgs mass point, a NeuroBayes® [4] neural network is trained to distinguish signal from background on a weighted combination of known signal and background events. To maximize sensitivity, a different choice of kinematic inputs to the neural network is made for events with zero jets in the final state, one jet in the final state, and two or more jets in the final state. In the 0-jet analysis, additional inputs to the neural networks are based on Matrix Element calculations for the event probability densities for the different background and signal processes using leading order matrix elements from the MCFM [5] package. In the two or more jets events, an additional selection criteria is made to veto events with a jet that has been tagged as coming from a b quark in order to suppress the dominant $t\bar{t}$ background. Templates from the neural network output in each of the three jet multiplicity categories corresponding to the expected background and signal contributions are then compared to the neural network data templates. A likelihood method is used to calculate the 95% C.L. limits on the Higgs production cross section based on the predicted and observed neural network output templates.

II. DETECTOR DESCRIPTION

The components of the CDF II detector relevant to this analysis are described briefly here; a more complete description can be found elsewhere [6]. The detector geometry is described by the azimuthal angle ϕ and the pseudo-rapidity $\eta \equiv -\ln(\tan\theta/2)$, where θ is the polar angle of a particle with respect to the proton beam axis (positive z-axis). The pseudo-rapidity of a particle originating from the center of the detector is referred to as $\eta_{\rm det}$.

The trajectories of charged particles are reconstructed using silicon micro-strip detectors [7, 8] and a 96-layer open-cell drift chamber (COT) [9] embedded in a 1.4 T solenoidal magnetic field. For $|\eta_{\rm det}| \leq 1$, a particle traverses all 96 layers of the COT; this decreases to zero at $|\eta_{\rm det}| \approx 2$. The silicon system provides coverage with 6 (7) layers with radii between 2.4 cm and 28 cm for $|\eta_{\rm det}| < 1.0$ ($1.0 < |\eta_{\rm det}| < 2.0$). Outside of the solenoid are electromagnetic (EM) and hadronic (HAD) sampling calorimeters segmented in a projective tower geometry. The first hadronic interaction length (λ) of the calorimeter, corresponding to 19-21 radiation lengths (X_0), uses lead absorber for measuring the electromagnetic component of showers, while the section extending to 4.5-7 λ uses iron to contain the hadronic component. The calorimeters are divided in a central ($|\eta_{\rm det}| < 1$) and forward ($1.1 < |\eta_{\rm det}| < 3.64$) region. Shower maximum detectors (SMX) embedded in the electromagnetic calorimeters at approximately $6X_0$ help in the position measurement and background suppression for electrons. Outside of the central calorimeters are scintillators and drift chambers for identifying muons as minimum ionizing particles. We use three complementary track pattern recognition algorithms which are distinguished by their starting point in COT, silicon, or projection from calorimeter energy cluster to interaction region.

III. LEPTON IDENTIFICATION

In order to maximize signal acceptance and suppress backgrounds from jets and photons misidentified as leptons, we use two(three) categories of electrons(muons). A sixth category, based on central tracks that are not fiducial to calorimeters or muon detectors, is used as either an electron or muon in forming $H \to WW^*$ candidates. The resulting six categories exploit essentially all the tracks and electromagnetic calorimeter clusters available.

All leptons are required to be isolated such that the sum of the E_T for the calorimeter towers in a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ around the lepton is less than 10% of the electron E_T or muon p_T. The transverse energy E_T

of a shower or calorimeter tower is $E\sin\theta$, where E is the associated energy. Similarly, p_T is the component of track momentum transverse to the beam line.

Electron candidates are required to have a ratio of HAD energy to EM energy consistent with originating from an electromagnetic shower and are further divided into central and forward categories. The central electron category requires a well-measured COT track satisfying $p_T > 10$ GeV/c that is fiducial to the central SMX and matched to a central EM energy cluster. The candidate is also required to have a matching cluster in the shower maximum detector, minimal energy sharing between towers, and a ratio for shower energy E to track momentum p of less than $2.5 + 0.0015E_T$. A forward electron is required to be fiducial to the forward SMX detector and have an energy deposition in both the calorimeter towers and SMX detector consistent with an electron shower shape. One of the calorimeter seeded tracks is required to be matched with a silicon track to reduce the photon background.

Muons are identified by either a charged track matched to a reconstructed track segment ("stub") in muon chambers or as a stubless minimum ionizing particle fiducial to calorimeters. In addition, stubless muons are required to have at least 0.1 GeV in total calorimeter energy. For $|\eta_{\rm det}| < 1.2$, strict requirements on the number of COT hits and the χ^2 of the track fit are placed on the muon tracks in order to suppress kaon decay-in-flight backgrounds. The category of stubless muons with $|\eta_{\rm det}| > 1.2$ requires that at least 60% of the COT layers crossed by the track have hits. In order to suppress background from cosmic rays, the track's point of closest approach to the beamline must be consistent with originating from the beam.

The final category of leptons are constructed from tracks which are not fiducial to the SMX detectors nor identified as stubbed muons. The requirements for the tracks are the same as stubless muons with $|\eta_{\rm det}| < 1.2$, but without any of the calorimeter requirements. Due to the lack of calorimeter information, electron and muons cannot be reliably differentiated in this region, and are therefore treated as having either flavor in the Higgs candidate selection. If an electron or non-fiducial track candidate is consistent with being due to a photon conversion as indicated by the presence of an additional nearby track, the candidate is vetoed.

To identify the presence of a Higgs boson decaying to two neutrinos, we use the missing transverse energy $\not\!\!E_T = |\sum_i E_{T,i}\ \hat{n}_{T,i}|$, where the $\hat{n}_{T,i}$ is the transverse component of the unit vector pointing from the interaction point to calorimeter tower i. The $\not\!\!E_T$ is corrected for muons which do not deposit all of their energy in the calorimeter.

The Higgs candidate events are required to pass one of four online trigger selections implemented in three successively more stringent levels. The final central electron requirement is an EM energy cluster with $E_T > 18$ GeV matched to a track with $p_T > 8$ GeV/c. Muon triggers are based on information from muon chambers matched to a track with $p_T > 18$ GeV/c. The trigger for forward electrons requires an $E_T > 20$ GeV EM energy cluster and an online measurement of the $\not\!\!E_T > 15$ GeV [10].

IV. BASE SELECTION

We define a common Base selection for the different 14 Higgs mass hypotheses. The $ll\nu\overline{\nu}$ candidates are selected from two opposite-sign leptons. At least one lepton is required to satisfy the trigger 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 leptons to increase Higgs kinematic acceptance. The z-positions of all the leptons in a candidate at the point of closest approach to the beam-line are required to be within 4 cm of each other.

There are several sources of background: Drell-Yan where the measured large $\not\!\!E_T$ is due to resolution tails, $WZ \rightarrow lll\nu$ where one lepton is lost, $WW \rightarrow ll\nu\overline{\nu}$, $t\overline{t} \rightarrow b\overline{b}ll\nu\overline{\nu}$, and $W\gamma$ and W+jets where a photon or jet is misidentified as a lepton. Because of the significant backgrounds from $W\gamma$ and W+jets, we apply an additional isolation requirement on the leptons that the sum of the track p_T in a cone for $\Delta R < 0.4$ around each lepton is less than 10% of the electron E_T or muon p_T.

To suppress the Drell-Yan background, we require $\not\!\!E_{T\,\rm spec} > 25$ GeV for dielectron and dimuon events and $\not\!\!E_{T\,\rm spec} > 15$ GeV for electron-muon events where $\not\!\!E_{T\,\rm spec}$ is defined to be:

$$\not\!\!E_{T\,\rm spec} \equiv \begin{cases} \not\!\!E_T & \text{if } \Delta\phi(\not\!\!E_T, \text{lepton}, \text{jet}) > \frac{\pi}{2} \\ \not\!\!E_T \sin(\Delta\phi(\not\!\!E_T, \text{lepton}, \text{jet})) & \text{if } \Delta\phi(\not\!\!E_T, \text{lepton}, \text{jet}) < \frac{\pi}{2} \end{cases}$$

This definition is a requirement that the $\vec{\not\!\!E}_T$ transverse to each lepton or jet in the event is greater than the minimum threshold if $\vec{\not\!\!E}_T$ points along that object, so that losing energy from just one such object in an event would not allow it to enter the sample. We further require the candidates to have $M_{\ell^+\ell^-} > 16$ GeV in order to suppress heavy flavor contributions, and exactly 2 leptons to suppress WZ contributions with a third lepton.

We consider individually final states with no jets in the event, one jet, or two or more jets where a jet is required to have $p_T > 15$ GeV and $|\eta| < 2.5$. This allows us to tune our multivariate discriminants on the different mixes of signal

TABLE I: Expected and observed yields for background and signal for events with 0 jets in the final state.

$M_H = 160$ GeV/c^2			
$t\bar{t}$	0.96	$\pm$	0.19
DY	66.88	$\pm$	15.20
WW	280.42	$\pm$	38.99
WZ	12.17	$\pm$	1.93
ZZ	17.29	$\pm$	2.74
W+jets	83.61	$\pm$	20.09
$W\gamma$	79.15	$\pm$	21.12
Total Background	**540.48**	$\pm$	**64.81**
$gg \to H$	8.38	$\pm$	1.29
Total Signal	**8.38**	$\pm$	**1.29**
Data	**552**		

HWW 0 Jet

TABLE II: Expected and observed yields for background and signal for events with 1 jet in the final state. In addition to gluon fusion, associated Higgs production with either a W or Z boson and production from vector boson fusion are considered as signal.

$M_H = 160$ GeV/c^2			
$t\bar{t}$	24.57	$\pm$	4.77
DY	71.21	$\pm$	16.19
WW	75.10	$\pm$	10.11
WZ	12.71	$\pm$	2.02
ZZ	4.53	$\pm$	0.72
W+jets	26.23	$\pm$	6.78
$W\gamma$	11.35	$\pm$	3.00
Total Background	**225.69**	$\pm$	**28.21**
$gg \to H$	4.08	$\pm$	0.63
WH	0.57	$\pm$	0.08
ZH	0.21	$\pm$	0.03
VBF	0.33	$\pm$	0.05
Total Signal	**5.18**	$\pm$	**0.69**
Data	**227**		

HWW 1 Jet

and background contributions within each jet multiplicity bin. In order to suppress the dominant $t\bar{t}$ background in the two or more jets bin, all events containing jets with a tight, secondary vertex b-tag are rejected.

The expected and observed yields after base selection cuts have been applied are shown for the three jet multiplicity categories in Tables I-III.

V. DATA MODELING

The geometric and kinematic acceptance for the WW, WZ, ZZ, $W\gamma$, Drell-Yan (DY), $t\bar{t}$, and all of the signal processes ($gg \to H$, WH, ZH, VBF) processes are determined using a Monte Carlo calculation of the collision followed by a GEANT3-based simulation of the CDF II detector [11] response. The Monte Carlo generator used for WW is MC@NLO [12], for WZ, ZZ, DY, $t\bar{t}$, and the signal processes it is PYTHIA [13], for $W\gamma$ it is the generator described in [14]. We use the CTEQ5L parton distribution functions (PDFs) to model the momentum distribution of the initial-state partons [15].

A correction of up to 10% per lepton is applied to the simulation based on measurements of the lepton reconstruction and identification efficiencies in data using Z decays. Additional 5% and 10% corrections based on $Z \to \ell\ell$ cross section measurements are applied to stubless muons reconstructed from only central tracks and muons reconstructed from minimum ionizing energy deposits in the forward calorimeter, respectively, to account for the known poor modeling of these lepton types. A further correction is applied to the $W\gamma$ background estimate using a measurement of the photon conversion veto efficiency in data. Also, 8% and 4% corrections are applied to the WW background in the 0 and 1 jet multiplicity bins, respectively, to account for the limited run range of the simulated sample used to estimate the yield. This correction was found to be negligible in the 2-jet bin. However, in the 2-jet analysis, a scale factor of

TABLE III: Expected and observed yields for background and signal for events with 2 or more jets in the final state. In addition to gluon fusion, associated Higgs production with either a W or Z boson and production from vector boson fusion are considered as signal.

$M_H = 160$ GeV/c^2			
$t\bar{t}$	70.34	$\pm$	14.46
DY	27.74	$\pm$	6.75
WW	15.68	$\pm$	2.47
WZ	3.33	$\pm$	0.53
ZZ	1.35	$\pm$	0.21
W+jets	8.38	$\pm$	2.27
$W\gamma$	1.80	$\pm$	0.47
Total Background	**128.62**	$\pm$	**20.19**
$gg \to H$	1.52	$\pm$	0.26
WH	1.18	$\pm$	0.16
ZH	0.59	$\pm$	0.08
VBF	0.61	$\pm$	0.10
Total Signal	**3.90**	$\pm$	**0.45**
Data	**139**		

HWW $\geq$ 2 Jets

M_H (GeV)	$\sigma^{NNLL}_{gg\to H} \times \delta_{EW}$	σ_{WH}	σ_{ZH}	σ_{VBF}	$\mathrm{Br_{H\to WW}}$
110	1.2808 $\times$ 1.0400	0.2075	0.1236	0.0841	0.0441
120	1.0062 $\times$ 1.0510	0.1529	0.0927	0.0717	0.1320
130	0.8013 $\times$ 1.0580	0.1141	0.0705	0.0611	0.2869
140	0.6455 $\times$ 1.0650	0.0860	0.0542	0.0521	0.4833
145	0.5818 $\times$ 1.0685	0.0749	0.0477	0.0481	0.5731
150	0.5251 $\times$ 1.0730	0.0654	0.0421	0.0445	0.6817
155	0.4750 $\times$ 1.0765	0.0572	0.0373	0.0412	0.8007
160	0.4310 $\times$ 1.0690	0.0510	0.0331	0.0382	0.9011
165	0.3920 $\times$ 1.0450	0.0441	0.0294	0.0354	0.9566
170	0.3566 $\times$ 1.0310	0.0389	0.0261	0.0329	0.9653
175	0.3254 $\times$ 1.0250	0.0244	0.0233	0.0305	0.9505
180	0.2972 $\times$ 1.0200	0.0306	0.0208	0.0283	0.9345
190	0.2493 $\times$ 0.9940	0.0243	0.0166	0.0244	0.7761
200	0.2105 $\times$ 0.9850	0.0193	0.0135	0.0210	0.7347

TABLE IV: Cross sections for the signal processes considered in this analysis (given in *pb*) along with the branching ratio for $H \to WW$ for each Higgs mass investigated. δ_{EW} refers to the electroweak corrections discussed in [16]

.

1.07 is applied to the $t\bar{t}$ sample to account for the difference in efficiency between the data and simulation of the b-tag veto requirement. Trigger efficiencies are determined from $W \to e\nu$ data for electrons and from $Z \to \mu^+\mu^-$ data for muons.

The background from W+jets is estimated from a sample of events with an identified lepton and a jet that is required to pass loose isolation requirements and contain a track or 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 jet is identified as a lepton. This probability is determined from multijet events collected with a set of jet-based triggers. A correction is applied for the small real lepton contribution using single W and Z boson Monte Carlo simulation.

In this analysis we consider contributions to the Higgs signal production from four sources. The gluon fusion cross section ($\sigma_{gg\to H}$) has been calculated to NNLL [3]. Additionally, two-loop electroweak corrections (δ_{EW}) [16] of up to 7.7% are applied to these NNLL cross sections. Associated production (WH and ZH) cross sections were obtained from [17] and vector boson fusion (VBF) cross sections can be obtained from MCFM [5] or [18]. Table IV gives the signal cross sections (in *pb*) and branching ratio for $H \to WW^*$ used in this analysis as well as the electroweak corrections δ_{EW} for each mass investigated.

VI. MULTIVARIATE DISCRIMINANTS

For each of the three jet categories, a NeuroBayes® neural network is trained on a weighted combination of known signal and background events from Monte Carlo independently for each of the 14 Higgs mass points. Each neural network has three layers consisting of input nodes, hidden nodes, and one output node. Once the neural network has been trained, templates are created for signal and background. For the zero and one jet categories, high S/B and low S/B templates based on the signal to background ratio of the different dilepton combinations are considered separately. Although this method acheives greater sensitivity than keeping all dilepton pairs in one cagetory, it was not done in the 2-jet bin because of low statistics. The templates are used as the final discriminant in calculating the 95% C.L. limits.

A. 0-jet Analysis

The 0-jet analysis uses 5 input variables, 6 hidden nodes, and 1 output node. The inputs are: the ΔR between the leptons, the $\Delta\phi$ between the leptons, the scalar sum of leptons E_T and $\not{E}_T$ (H_T), the likelihood ratio for Higgs to WW production (LRHWW) and the likelihood ratio for WW production (LRWW). These likelihoods were calculated using an event-by-event calculation of the probability density $P_m(x_{obs})$. Five modes (m) are modeled including WW, ZZ, $W\gamma$, W+jet, and $H \to WW$ (for all 14 masses). The probability density for any given mode m is given by:

$$P_m(x_{obs}) = \frac{1}{< \sigma_m >} \int \frac{d\sigma_m^{th}(y)}{dy} \epsilon(y) G(x_{obs}, y) dy \tag{1}$$

where

x_{obs}	are the observed "leptons" and $\vec{\not{E}}_T$,
y	are the true lepton four-vectors (include neutrinos),
σ_m^{th}	is leading-order theoretical calculation of the cross-section for mode m,
$\epsilon(y)$	is total event efficiency $\times$ acceptance
$G(x_{obs}, y)$	is an analytic model of resolution effects, and
$\frac{1}{<\sigma_m>}$	is the normalization.

The function $\epsilon(y)$ describes the probabilities of a parton level object (e, μ, γ, or parton) to be reconstructed as one of the six lepton categories. The efficiency function is determined solely from Monte Carlo for e and μ but comes from a combination of Monte Carlo and data-driven measurements described in section V for γ and partons. The event probability densities are used to construct a discriminant:

$$LR_S(x_{obs}) \equiv \frac{P_S(x_{obs})}{P_S(x_{obs}) + \Sigma_i k_i P_i(x_{obs})},$$

where S is ehter WW or $H \to WW$, k_i is the expected fraction for each background and $\Sigma_i k_i = 1$.

Because of the missing neutrinos in the final state, the integral in Equation 1 integrates out the unobserved degrees of freedom (DOF) reducing the 12 DOF in y to the eight measured DOF. Distributions of the 5 neural network input variables with the expected and observed yields for signal and background are shown in Fig. 1. The NN output templates are shown in Fig. 2 for $M_H = 160$ GeV/c^2.

B. 1-jet Analysis

The 1-jet analysis uses 8 input variables, 9 hidden nodes, and 1 output node. The inputs are: the dilepton invariant mass, the transverse mass of the sum of leptons' 4-momenta and ($\not{E}_T$, MetX, MetY, 0) where MetX/Y are the x/y components of the $\vec{\not{E}}_T$ ($M_T(ll\not{E}_T)$), ΔR between the leptons, H_T, $\not{E}_{T\mathrm{spec}}$, the transverse momentum of the leading lepton ($p_T(l_1)$), transverse momentum of the second lepton ($p_T(l_2)$), and the energy of the leading lepton ($E(l_1)$). Distributions of these variables with expected and observed yields for signal and background are shown in Fig. 3. The NN output templates are shown in Fig. 4 for $M_H = 160$ GeV/c^2.

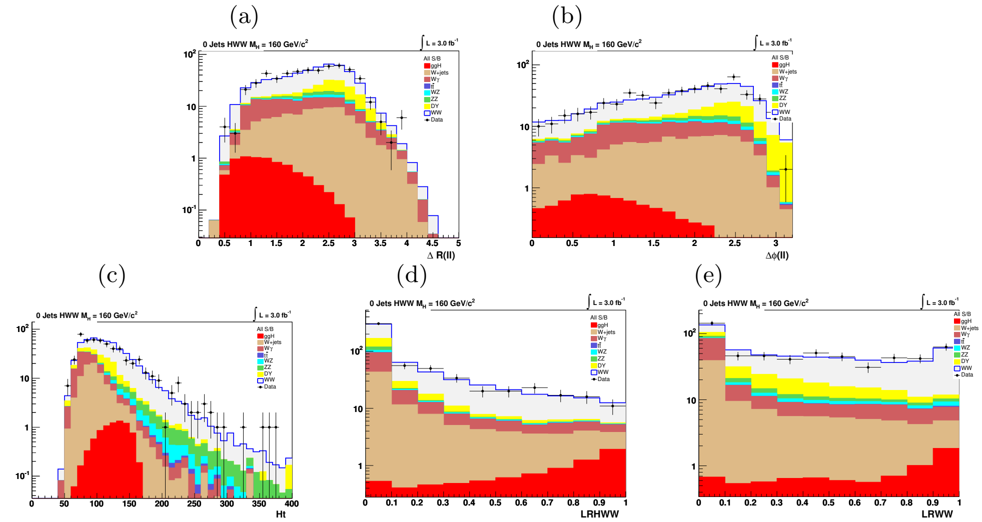

FIG. 1: Distributions of input variables to the neural network for events with zero jets in the final state: (a) ΔR of the leptons (b) $\Delta\phi$ of the leptons (c) $H_{\rm T}$ (d) LRHWW (e) LRWW

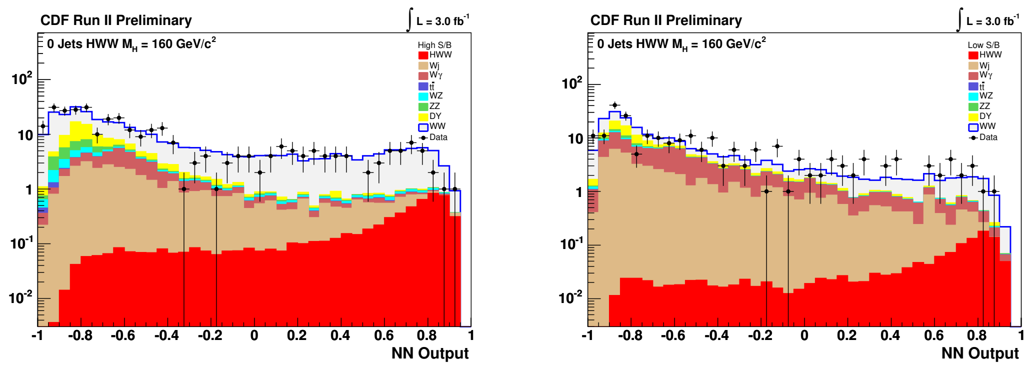

FIG. 2: Neural network templates for high S/B (*left*) low S/B (*right*) for $m_H = 160$ GeV/c^2 and zero jets in the final state.

C. 2-jet Analysis

The 2-jet analysis uses 8 input variables, 9 hidden nodes, and 1 output node. The inputs are: the dilepton invariant mass, the transverse energy of the leading lepton, the transverse energy of the second lepton, $H_{\rm T}$, ΔR between the leptons, $\Delta\phi$ between the leptons, $\Delta\phi$ between the vector sum of the two lepton momenta and the $\not\!\!E_T$ (dPhiLepSumMet), and the vector sum of first and second jet $p_{\rm T}$ (SumJetPt). Distributions of these variables with expected and observed yields for signal and background are shown in Fig. 5. The NN output template is shown in Fig. 6 for $M_H = 160$ GeV/c^2.

VII. SYSTEMATICS

Systematic uncertainties associated with the Monte Carlo simulation affect the Higgs, WW, WZ, ZZ, $W\gamma$, DY and $t\bar{t}$ acceptances taken from the simulated event samples. Uncertainties originating from lepton selection and trigger efficiency measurements are propagated through the acceptance calculation, giving uncertainties typically around 2%, but as high as 7% for the different signal and background processes. In the case of the specific muon categories where additional correction factors are applied (see Section V), half of the applied correction is incorporated as an additional uncertainty.

We also assign an acceptance uncertainty due to potential contributions from higher-order effects. In the case

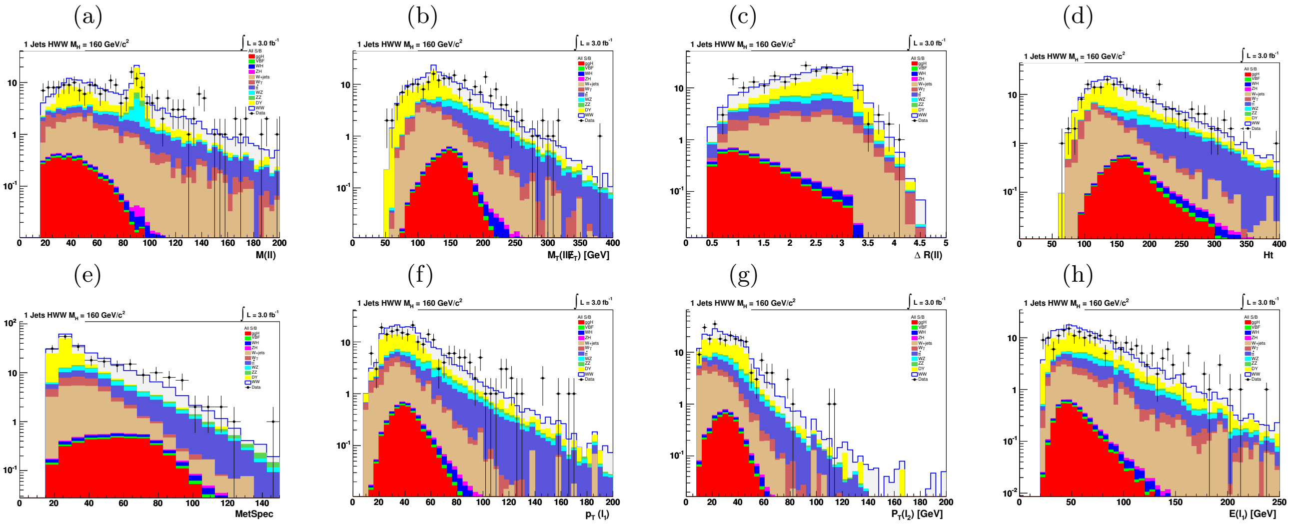

FIG. 3: Distributions of input variables to the neural network with one jet in the final state: (a) Dilepton invariant mass (b) transverse mass: $M_T(ll\not\!\!E_T)$ (c) ΔR of the leptons (d) $H_{\rm T}$ (e) $\not\!\!E_{T\rm spec}$ (f) $p_T(l_1)$ (g) $p_T(l_2)$ (h) $E(l_1)$

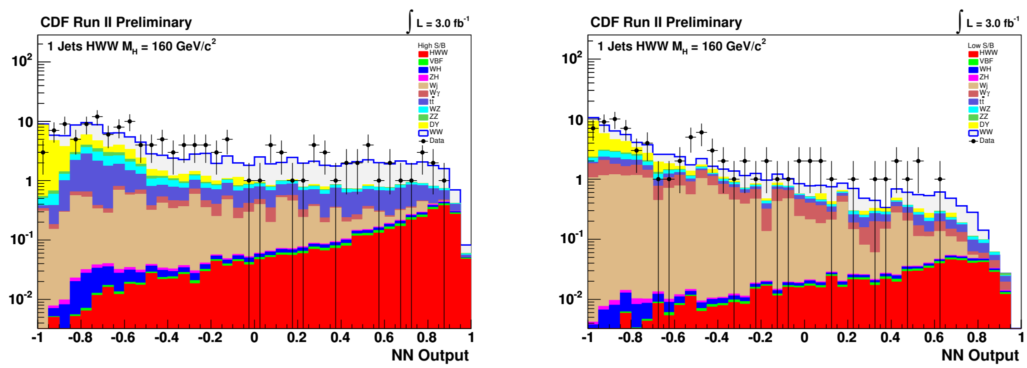

FIG. 4: Neural network templates for high S/B (*left*) low S/B (*right*) for $m_H = 160$ GeV/c^2 and one jet in the final state.

of WW we take half of the difference between the leading-order (PYTHIA-based [13]) and next-to-leading order (MC@NLO [12]) acceptances. The other processes are only simulated at leading-order and for these modes we assign the full difference observed in WW leading to a 10% uncertainty. For the Drell-Yan process, which has been studied extensively at CDF and for which the simulation has been tuned to reproduce the kinematic distributions in data, we reduce this assignment to 5%. For Monte Carlo samples not currently generated over the entire data run range, we take additional uncertainties on the acceptance for the corresponding signal/background process. This uncertainty is taken from the observed difference in the leading-order WW acceptance for the corresponding partial run range versus that for the full run range. The acceptance variations due to PDF model uncertainties is assessed to be on the order of 2% using the 20 pairs of PDF sets described in [19].

The $\not\!\!E_T$ resolution modeling uncertainty on the acceptance is determined from comparisons of the data and the Monte Carlo simulation in a sample of dilepton events. For WW, WZ, $ZZ \to ll\nu\overline{\nu}$, and $t\overline{t}$ production, where one or more neutrinos are responsible the observed $\not\!\!E_T$, we determined the uncertainty to be 1%. The uncertainty due to the $\not\!\!E_T$ modeling for the Drell-Yan background is much larger (20%) than for other final states because it depends on the non-Gaussian tails of the resolution function.

For the $W\gamma$ background contribution, there is an additional uncertainty of 20% from the detector material description and conversion veto efficiency. The systematic uncertainty on the W+jets background contribution is determined from differences in the measured 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.

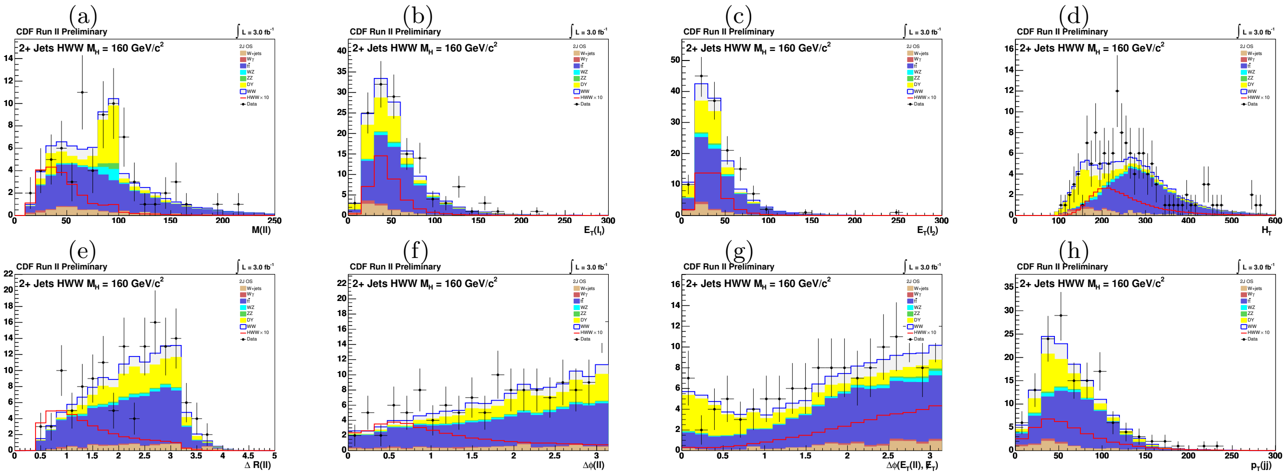

FIG. 5: Distributions of input variables to the neural network for events with two or more jets: (a) Dilepton invariant mass (b) Leading lepton E_{T} (c) Second leading lepton E_{T} (d) H_{T} (e) ΔR of the leptons (f) $\Delta\phi$ of the leptons (g) dPhiLepSumMet (h) sumJetPt

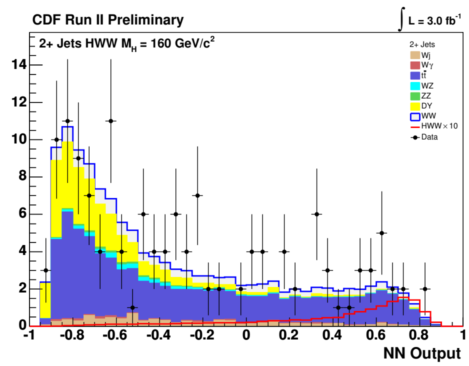

FIG. 6: Neural network template for events with two or more jets in the final state.

The uncertainties on the $WW/WZ/ZZ$, $W\gamma$ and $t\bar{t}$ cross sections are assigned to be 10% [20], 10% [14], and 15% [21, 22], respectively. In addition, all signal and background estimates obtained from simulation have an additional 6% uncertainty originating from the luminosity measurement [23].

Most systematic uncertainties on the signal processes are assessed using the same techniques described for the background processes. Uncertainties on the theoretical cross sections vary for the different Higgs production mechanisms. Associated production cross sections are known to NNLO, so the theoretical uncertainty on these cross sections is small, less than 5% [18]. Vector Boson Fusion (VBF) production is known only to NLO, so the residual theoretical uncertainty is higher (on the order of 10% [18]). Gluon fusion is a QCD process, so although it is known to NNLO, the corresponding theoretical uncertainty is still significant. We use the HNNLO program [24–26], which computes the theoretical cross section for $gg \to H$ production at the Tevatron to NNLO, to assign theoretical uncertainties due to changing the renormalization and factorization scales (10.9%) and the gluon PDF model (5.1%). These are independent effects and give a total $gg \to H$ cross-section uncertainty of 12.0%.

PDF model uncertainties on the signal acceptances used for the associated and VBF Higgs production modes are taken from studies completed for the $H \to \tau\tau$ search performed at CDF [27]. We again use the HNNLO program to evaluate acceptance uncertainties for $gg \to H$ production for both changes in scale and the PDF model. To evaluate the change in acceptance, we re-weight the Higgs p_T and η distributions obtained from Pythia to match the HNNLO calculations. Re-weighting events based on Higgs p_T leads to changes in the relative acceptances for the different jet bins. The resulting anti-correlated uncertainties between different jet bins are indicated in our final uncertainty tables by negative signs on the corresponding uncertainties for the 1-jet and two or more jet bins. Re-weighting events based on Higgs η primarily affects the overall acceptance for leptons leading to correlated uncertainties between the different jet bins.

The complete set of systematic uncertainties are summarized in Table XI.

VIII. CROSS-CHECKS

We use several control regions to validate different aspects of our data modeling. The sample sizes of the control regions are designed to be large enough to give statistically meaningful tests. The most fundamental control sample is the Drell-Yan sample selected using the default candidate selection minus the $\not\!\!E_{T\text{spec}}$ and $\not\!\!E_{T,\text{sig}}$ requirements. This control sample provides a high statistics test of acceptance, lepton id efficiency, and trigger efficiency calculations. The same-sign control sample is selected based on the default selection but reversing the opposite-sign requirement. This sample is used to test the modeling of jets and photons misidentified as leptons which provide the dominate contributions to this sample. We also test our $\not\!\!E_T$ model by constructing two additional control samples whose dominant contributions come from Drell-Yan events on the tail of the $\not\!\!E_T$-resolution distribution. The two samples are the low-$\not\!\!E_{T\text{spec}}$ sample which is constructed using default selection with the exception of requiring $15 < \not\!\!E_{T\text{spec}} < 25$ GeV instead of $\not\!\!E_{T\text{spec}} > 25$ GeV, and the low-$\not\!\!E_{T,\text{sig}}$ sample for which the $\not\!\!E_{T,\text{sig}}$ requirement is reversed ($\not\!\!E_{T,\text{sig}} \equiv \not\!\!E_T/\sqrt{\sum E_T} < 2.5\ \text{GeV}^{\frac{1}{2}}$). The predicted and observed yields for these four control regions (for events with either 0 or 1 jet in the final state) are shown in Tables V-VIII.

TABLE V: Expected and observed yields in the Drell-Yan control region for events with zero or one jet in the final state.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	6.6	38.3	33.8	1.3	104622.3	1.9	517.1	105221.1	102213.0
$e\ \mu$	6.3	0.1	0.1	1.1	82.9	0.4	31.2	122.2	99.0
$\mu\ \mu$	6.1	31.3	29.4	1.4	82223.1	0.0	369.7	82661.1	80382.0
e trk	5.9	12.4	11.3	1.2	32387.7	0.7	533.4	32952.6	30240.0
μ trk	3.8	10.8	10.0	0.8	27673.3	0.0	312.0	28010.7	26750.0
Total	28.7	92.9	84.6	5.9	246989.3	3.0	1758.8	248963.0	239684.0

TABLE VI: Expected and observed yields in the Same-Sign control region for events with zero or one jet in the final state.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	4.2	2.2	0.6	0.2	3.9	41.3	32.9	85.4	83.0
$e\ \mu$	5.0	4.6	0.4	0.3	9.4	39.0	36.6	95.1	99.0
$\mu\ \mu$	0.0	1.5	0.2	0.0	0.1	0.0	3.2	5.1	5.0
e trk	1.7	1.6	0.2	0.1	1.5	9.8	9.7	24.5	27.0
μ trk	0.0	0.9	0.1	0.0	0.3	1.7	4.6	7.6	7.0
Total	10.9	10.8	1.4	0.6	15.2	91.8	84.9	215.5	221.0

TABLE VII: Expected and observed yields in the Low-$\not\!\!E_{T\text{spec}}$ control region for events with zero or one jet in the final state.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	5.4	1.4	1.1	0.6	60.4	3.7	7.1	79.7	60.0
$e\ \mu$	8.5	0.4	0.1	0.8	46.6	2.6	6.1	65.1	41.0
$\mu\ \mu$	4.0	1.0	0.9	0.6	48.8	0.0	2.1	57.3	73.0
e trk	4.3	0.7	0.5	0.5	59.4	0.8	4.8	71.0	58.0
μ trk	2.6	0.4	0.3	0.3	13.3	0.1	2.0	19.0	22.0
Total	24.9	3.8	2.9	2.8	228.5	7.2	21.7	291.8	254.0

For events with two or more jets, the dominant background is $t\bar{t}$. For this reason, we examine a $t\bar{t}$ control sample selected using the default selection with the b-tag veto requirement reversed such that only events with one or more jets identified as coming from a b quark are accepted. The predicted and observed yields for this control region are shown in Table IX.

TABLE VIII: Expected and observed yields in the Low-$\not\!\!E_{T,\mathrm{sig}}$ control region for events with zero or one jet in the final state.

Category	WW	WZ	ZZ	tt	DY	$W\gamma$	W+jets	Total	Data
e e	4.1	0.9	0.7	0.3	33.2	3.7	6.0	49.0	44.0
e μ	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
μ μ	3.3	0.7	0.6	0.3	21.4	0.0	1.2	27.5	34.0
e trk	3.5	0.5	0.3	0.3	13.0	0.9	2.1	20.7	23.0
μ trk	2.0	0.3	0.2	0.2	6.9	0.1	1.0	10.7	6.0
Total	13.0	2.3	1.8	1.2	74.5	4.8	10.3	107.8	107.0

TABLE IX: Expected and observed yields in the $t\bar{t}$ channel for events with two or more jets in the final state.

DY	0.56 ± 0.12
WW	0.42 ± 0.05
W+jets	1.13 ± 0.29
$W\gamma$	0.07 ± 0.02
WZ	0.08 ± 0.01
ZZ	0.08 ± 0.01
Total background	**2.34 ± 0.32**
tt	91.44 ± 16.92
Data	**98**

IX. RESULTS

In order to determine the sensitivity, we construct a binned likelihood using the five neural network templates: two templates each for the 0-jet and 1-jet categories (high S/B and low S/B) and one template for the 2 or more jets bin. All the signals and backgrounds are allowed to float, but ratios of signal and background contributions are constrained to their expectations within a set of Gaussian constraints that are determined from the assumed correlations between the systematics uncertainties. The total signal yield is also allowed to float.

The 95% C.L. limit is determined with a set of 10,000 Monte Carlo background-only experiments based on expected yields varied within the assigned systematics. Again, correlations between the systematics for different backgrounds are included. For each experiment a test statistic is formed from the difference in the likelihood value for the background-only model versus that for the signal plus background model.

We find a total of 918 candidates with an expectation of 894.8 ± 113.2 background and 17.5 ± 2.2 signal for Higgs mass of 160 GeV/c^2. The median expected 95% C.L. limit at a Higgs mass of 160 GeV/c^2 is $1.7^{+0.7}_{-0.5}$ times the standard model prediction (including EWK corrections), while the observed limit is 1.6 times the standard model prediction.

TABLE X: Expected and observed 95% C.L. limits for 14 different Higgs masses for all jet multiplicity categories.

	110	120	130	140	145	150	155	160	165	170	175	180	190	200
$-2\sigma/\sigma_{SM}$	20.05	6.42	3.03	1.95	1.69	1.43	1.15	0.88	0.84	0.96	1.21	1.42	2.26	3.04
$-1\sigma/\sigma_{SM}$	27.38	8.80	4.20	2.70	2.34	1.93	1.58	1.20	1.14	1.34	1.67	1.96	3.17	4.23
Median/$\sigma_{\mathbf{SM}}$	**38.90**	**12.61**	**6.05**	**3.86**	**3.35**	**2.79**	**2.29**	**1.71**	**1.62**	**1.92**	**2.39**	**2.82**	**4.63**	**6.16**
$+1\sigma/\sigma_{SM}$	56.09	18.02	8.70	5.55	4.84	4.02	3.36	2.45	2.34	2.74	3.46	4.08	6.77	8.99
$+2\sigma/\sigma_{SM}$	78.70	25.44	12.21	7.74	6.75	5.64	4.71	3.40	3.36	3.89	4.86	5.68	9.56	12.79
Observed/$\sigma_{\mathbf{SM}}$	**57.89**	**13.98**	**6.13**	**4.03**	**3.35**	**3.25**	**2.33**	**1.56**	**1.72**	**1.91**	**2.01**	**2.82**	**5.26**	**10.35**

X. SUMMARY

We have searched for standard model Higgs boson decay to WW^* in the $ll\nu\nu$ final state using a combined Matrix Element and neural network technique. The observed 95% C.L. limits on the production cross section compare well with the expected limits over a wide range of potential Higgs mass values as shown in Table X and Fig. 7. We observe no sign of a significant signal excess or deficit over the studied range of Higgs masses.

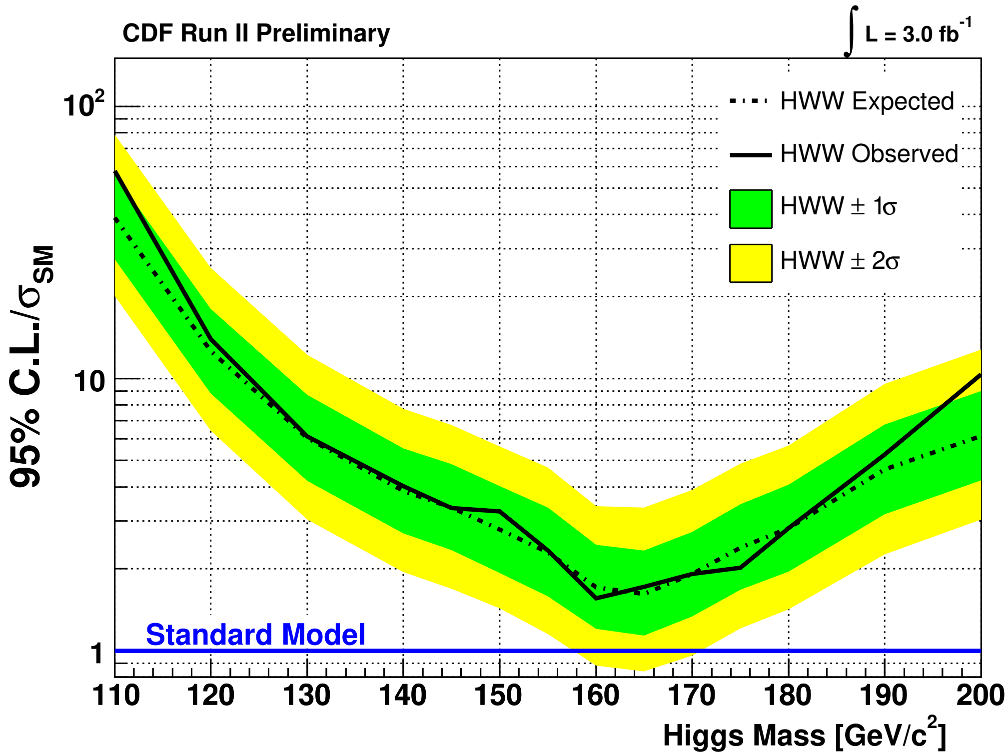

FIG. 7: The ratio of 95% C.L. production cross section limit to Standard Model prediction for 14 different Higgs masses for all jet bins.

[1] The LEP Electroweak Working Group, http://lepewwg.web.cern.ch/LEPEWWG/.
[2] M. Spira (1998), hep-ph/9810289.
[3] S. Catani, D. de Florian, M. Grazzini, and P. Nason, JHEP **07**, 028 (2003), hep-ph/0306211.
[4] M. Feindt and U. Kerzel, Nucl. Instrum. Methods A **559**, 190 (2006).
[5] J. Campbell and K. Ellis, MCFM - Monte Carlo for FeMtobarn processes, http://mcfm.fnal.gov/.
[6] R. Blair, *et al.* (CDF Collaboration) (1996), FERMILAB-PUB-96/390-E.
[7] A. Sill *et al.*, Nucl. Instrum. Methods A **447**, 1 (2000).
[8] A. Affolder *et al.*, Nucl. Instrum. Methods A **453**, 84 (2000).
[9] T. Affolder *et al.*, Nucl. Instrum. Methods A **526**, 249 (2004).
[10] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, Phys. Rev. D **48**, 2182 (1993).
[11] R. Brun, R. Hagelberg, M. Hansroul, and J. Lassalle, CERN-DD-78-2-REV and CERN-DD-78-2.
[12] S. Frixione and B. R. Webber, JHEP **06**, 029 (2002), hep-ph/0204244.
[13] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05**, 026 (2006).
[14] U. Baur and E. L. Berger, Phys. Rev. **D47**, 4889 (1993).
[15] H. L. Lai *et al.* (CTEQ), Eur. Phys. J. **C12**, 375 (2000).
[16] U. Aglietta, B. Bonciani, G. Degrassi, and A. Vivini (2006), hep-ph/0610033.
[17] K. A. *et. al.*. Assamagan (2004), hep-ph/0406152.
[18] TEV4LHC Working Group, http://maltoni.home.cern.ch/maltoni/TeV4LHC/SM.html.
[19] S. Kretzer, H. L. Lai, F. I. Olness, and W. K. Tung, Phys. Rev. **D69**, 114005 (2004).
[20] J. M. Campbell and R. K. Ellis, Phys. Rev. **D60**, 113006 (1999).
[21] N. Kidonakis and R. Vogt, Phys. Rev. **D68**, 114014 (2003).
[22] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, JHEP **04**, 068 (2004).
[23] D. Acosta *et al.*, Nucl. Instrum. Meth. **A494**, 57 (2002).
[24] S. Catani and M. Grazzini, Phys. Rev. Lett. **98**, 222002 (2007), hep-ph/0703012.
[25] M. Grazzini, JHEP **0802**, 043 (2008), 0801.3232 [hep-ph].
[26] HNNLO program, http://theory.fi.infn.it/grazzini/codes.html.
[27] CDF collaboration (2008), http://www-cdf.fnal.gov/physics/new/hdg/results/smhtt_080222/.

TABLE XI: Systematics

0 Jet Uncertainties	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet	$gg \to H$	WH	ZH	VBF
Cross Section											
Scale								10.9%			
PDF Model								5.1%			
Total	*10.0%*	*10.0%*	*10.0%*	15.0%	5.0%	10.0%		12.0%			
Acceptance											
Scale (leptons)								2.5%			
Scale (jets)								4.6%			
PDF Model (leptons)	1.9%	2.7%	2.7%	2.1%	4.1%	2.2%		1.5%			
PDF Model (jets)								0.9%			
Higher-order Diagrams	**5.5%**	**10.0%**	**10.0%**	10.0%	**5.0%**	**10.0%**					
Missing Et Modeling	1.0%	1.0%	1.0%	1.0%	20.0%	1.0%		1.0%			
Conversion Modeling						20.0%					
Jet Fake Rates											
(Low S/B)							21.5%				
(High S/B)							27.7%				
MC Run Dependence	3.9%			4.5%		4.5%		3.7%			
Lepton ID Efficiencies	2.0%	1.7%	2.0%	2.0%	1.9%	1.4%		1.9%			
Trigger Efficiencies	2.1%	2.1%	2.1%	2.0%	3.4%	7.0%		3.3%			
Luminosity	5.9%	5.9%	5.9%	5.9%	5.9%	5.9%		5.9%			
1 Jet Uncertainties	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet	$gg \to H$	WH	ZH	VBF
Cross Section											
Scale								10.9%			
PDF Model								5.1%			
Total	*10.0%*	*10.0%*	*10.0%*	15.0%	5.0%	10.0%		12.0%	5.0%	5.0%	10.0%
Acceptance											
Scale (leptons)								2.8%			
Scale (jets)								-5.1%			
PDF Model (leptons)	1.9%	2.7%	2.7%	2.1%	4.1%	2.2%		1.7%	1.2%	0.9%	2.2%
PDF Model (jets)								-1.9%			
Higher-order Diagrams	**5.5%**	**10.0%**	**10.0%**	10.0%	**5.0%**	**10.0%**			*10.0%*	*10.0%*	*10.0%*
Missing Et Modeling	1.0%	1.0%	1.0%	1.0%	20.0%	1.0%		1.0%	1.0%	1.0%	1.0%
Conversion Modeling						20.0%					
Jet Fake Rates											
(Low S/B)							22.2%				
(High S/B)							31.5%				
MC Run Dependence	1.8%			2.2%		2.2%		2.6%	2.6%	1.9%	2.8%
Lepton ID Efficiencies	2.0%	2.0%	2.2%	1.8%	2.0%	2.0%		1.9%	1.9%	1.9%	1.9%
Trigger Efficiencies	2.1%	2.1%	2.1%	2.0%	3.4%	7.0%		3.3%	2.1%	2.1%	3.3%
Luminosity	5.9%	5.9%	5.9%	5.9%	5.9%	5.9%		5.9%	5.9%	5.9%	5.9%
≥ 2 Jets Uncertainties	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet	$gg \to H$	WH	ZH	VBF
Cross Section											
Scale								10.9%			
PDF Model								5.1%			
Total	*10.0%*	*10.0%*	*10.0%*	15.0%	5.0%	10.0%		12.0%	5.0%	5.0%	10.0%
Acceptance											
Scale (leptons)								3.1%			
Scale (jets)								-8.7%			
PDF Model (leptons)	1.9%	2.7%	2.7%	2.1%	4.1%	2.2%		2.0%	1.2%	0.9%	2.2%
PDF Model (jets)								-2.8%			
Higher-order Diagrams	**10.0%**	**10.0%**	**10.0%**	10.0%	**10.0%**	**10.0%**			*10.0%*	*10.0%*	*10.0%*
Missing Et Modeling	1.0%	1.0%	1.0%	1.0%	20.0%	1.0%		1.0%	1.0%	1.0%	1.0%
Conversion Modeling						20.0%					
b-tag Veto				7.0%							
Jet Fake Rates							27.1%				
MC Run Dependence	1.0%			1.0%		1.0%		1.7%	2.0%	1.9%	2.6%
Lepton ID Efficiencies	1.9%	2.9%	1.9%	1.9%	1.9%	1.9%		1.9%	1.9%	1.9%	1.9%
Trigger Efficiencies	2.1%	2.1%	2.1%	2.0%	3.4%	7.0%		3.3%	2.1%	2.1%	3.3%
Luminosity	5.9%	5.9%	5.9%	5.9%	5.9%	5.9%		5.9%	5.9%	5.9%	5.9%