Standard Model Higgs Searches at the Tevatron

Ralf Bernhard
Physikalisches Institut, Albert-Ludwigs Universität Freiburg

Abstract
The latest searches for the Standard Model Higgs boson at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV with the DØ and the CDF detectors at the Fermilab Tevatron collider are presented. For the first time since the LEP experiments the sensitivity for a Standard Model Higgs boson has been reached at a Higgs boson mass of 170 GeV/c^2.

1 Introduction

In the Standard Model (SM) of particle physics the Higgs mechanism is responsible for breaking electroweak symmetry, thereby giving mass to the W and Z bosons. It predicts the existence of a heavy scalar boson, the Higgs boson, with a mass that can not be predicted by the SM. Direct searches for the Higgs Boson were performed at the LEP experiments in the process $e^+e^- \rightarrow ZH$ with a centre of mass energy of 206.6 GeV. A direct mass limit at $m_H >$ 114.4 GeV/c^2 [1] was set at the 95% confidence level (CL)[1]. This limit is slightly below the maximum available kinematic limit due to a small excess observed in the LEP data.

Indirect limits have been placed on the Higgs boson mass by the LEP, SLD and Tevatron experiments from electroweak precision measurements [2]. The main contribution to these indirect constraints from the Tevatron experiments, DØ and CDF, are the measurements of the W Boson and top quark masses [2]. The dependence of the Higgs mass on these measurements is shown in Figure 1 on the left and the Higgs mass dependence on the measured electroweak precision parameters in Figure 1 on the right. The SM fit yields a best value of $m_H = 84^{+34}_{-26}$ GeV/c^2 [3]. The upper limit on the Higgs mass at 95% CL is $m_H < 154$ GeV/c^2. If the direct mass limit is also taken into account this limit is increased to $m_H <$ 185 GeV/c^2.

2 Higgs Searches at the Tevatron

The Tevatron experiments CDF [4] and DØ [5] search for direct Higgs boson production in the mass range above the LEP limit using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The relevant processes at these energies are associated Higgs production ($qq' \rightarrow WH$, $q\bar{q} \rightarrow ZH$) and gluon fusion ($gg \rightarrow H$). Typical cross-sections are $\sigma \simeq 0.7 - 0.15$ pb for gluon fusion and $\sigma \simeq 0.2 - 0.02$ pb for associated production at Higgs masses in the range $115 - 200$ GeV/c^2.

The Higgs boson predominantly decays into $b\bar{b}$ quark pairs in the low mass range below 135 GeV/c^2. Hence the signal in the $gg \rightarrow H$ channel is overwhelmed by multi-jet background. This makes the process $gg \rightarrow H$ therefore not a viable search channel at low Higgs boson masses. The WH and ZH channels, where the vector boson decays into leptons, have much lower cross-sections but the lepton tag from the decay of the $W \rightarrow \ell\nu$ or $Z \rightarrow \ell\ell$ and selections

[1] All limits given in this paper are at 95% CL

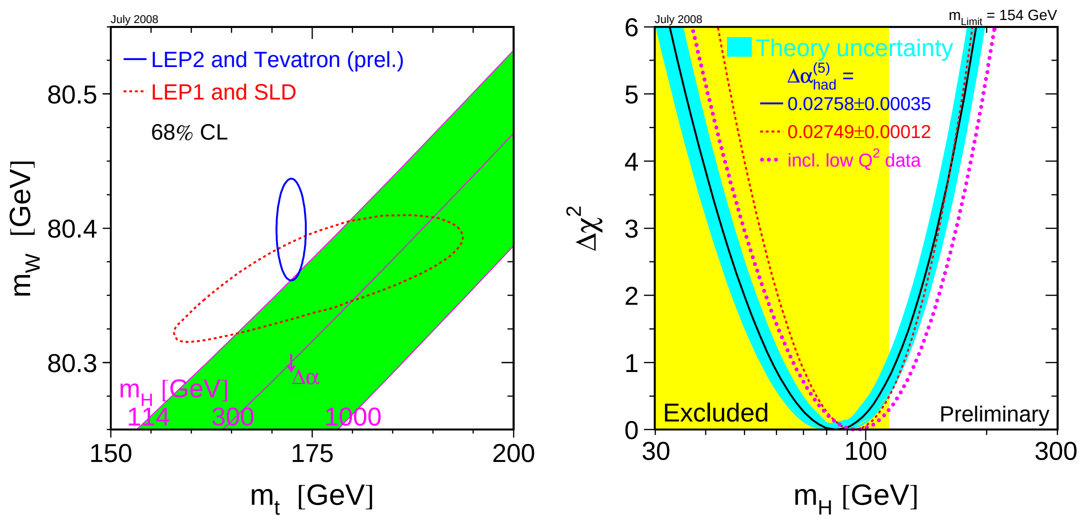

Fig. 1: Constraints on the Higgs mass from precision top and W mass measurements (left) and fit for the Higgs Mass from the W data showing the direct search LEP limit (right)

on missing transverse energy from the neutrino in the decays $W \to \ell\nu$ or $Z \to \nu\nu$ help to reduce the background significantly.

At higher masses, around $m_H = 165$ GeV/c^2, the Higgs boson will predominantly decay into WW pairs. Leptons from the decays of the W bosons and the missing transverse energy are used to reject background, making the channel $gg \to H \to WW$ the most promising search channel in this mass region. A 'hybrid' channel, the associated production with subsequent Higgs decay into (virtual) W pairs, $qq' \to WH \to WWW$, also contributes in the intermediate mass region.

2.1 The Tools

The main tools employed in Higgs searches at the Tevatron are lepton identification and - especially in the low Higgs mass region - jet reconstruction and b jet tagging. The experiments use b jet tagging algorithms that exploit the long lifetime of b hadrons. These algorithms are applied to each jet, searching for tracks with large transverse impact parameters relative to the primary vertex and for secondary vertices formed by tracks in the jet.

To further improve the b jet tagging these variables are used as input to a artificial Neural Network (NN) jet-flavor separator. The NN is trained to separate b quark jets from light flavour jets. By adjusting the minimum requirement on the NN output variable, a range of increasingly stringent b tagging operating points is obtained, each with a different signal efficiency and purity. Using this tool at DØ, b tagging efficiencies have been improved by 33% while keeping the rate of falsely identified light flavor jets (mistags) low. The efficiencies range between 40-70% for b

jets at a low mistag rates between 0.5-3% for light flavor jets.

Almost all Higgs searches at the Tevatron employ advanced analysis techniques like artificial Neural Networks (NN), boosted decision trees (BDT) or matrix element techniques (ME) to combine kinematic characteristics of signal and background events into a single discriminant. These techniques improve the separation of signal to background over the invariant Higgs boson mass distribution which is the most important single variable. Careful validation of all input variables is mandatory for robust results.

Events with neutrinos in the final state are identified using missing transverse energy. The reconstruction of all these variables require excellent performance of all detector components.

2.2 Signal and Background

The Higgs signal is simulated with PYTHIA [6]. The signal cross-sections are normalised to next-to-next-to-leading order (NNLO) calculations [7, 8] and branching ratios from HDECAY [9].

There are many types of background to the Higgs search. An important source of background are multi-jet events (often labeled "QCD background"). This background and the instrumental background due to mis-identified leptons or b jets is either simulated with PYTHIA (only for the CDF $ZH \to \nu\nu b\bar{b}$ analysis) or is taken directly from data, since it is not very well simulated by Monte Carlo. Determining this background from data is done using control samples with no signal content.

Electroweak background processes such as di-boson production, $p\bar{p} \to VV(V = W, Z)$, V+jets or $t\bar{t}$ pair production often dominate at the final stages of the selection; these are simulated using leading order Monte Carlo programs such as PYTHIA, ALPGEN, HERWIG or COMPHEP. The normalisation of these processes is obtained either from data or from NLO calculations.

2.3 Search for $WH \to \ell\nu b\bar{b}$

One of the most sensitive channels for a low Higgs boson mass is the decay $WH \to \ell\nu b\bar{b}$. This final state consists of two b jets from the Higgs boson and a charged lepton ℓ and a neutrino from the W boson. All three leptonic decays of the W boson are analysed at DØ, with the most sensitive being the decays to electrons and muons. Events are selected with one or two b tagged jets an isolated electron or muon and missing transverse energy. The main backgrounds after selection are W+jets and $t\bar{t}$ production. The di-jet invariant mass distribution for events with two b-tags is shown in Figure 2 on the left side. To improve the separation between the signal and the irreducible background a NN is trained which takes a number of kinematic and topological variables as input. The output of this NN is used to extract limits on Higgs production and is shown in Figure 2 on the right side. The analysis uses 1.7 fb $^{-1}$ of recorded data and sets an observed (expected) limit on σ_{95}/σ_{SM}= 9.1(8.5) for a Higgs boson mass $m_H = 115$ GeV/c^2 (where σ_{SM} is the cross section predicted for this process by the Standard Model). A dedicated search for $W^{\pm}H \to \tau^{\pm}\nu b\bar{b}$ with hadronic τ decays has been added at DØ. Using the di-jet mass distribution to separate signal from background an observed (expected) limit on σ_{95}/σ_{SM}= 35.4 (42.1) for a Higgs boson mass $m_H = 115$ GeV/c^2 has been obtained in that channel. At

CDF a similar analysis using 2.7 fb^{-1} of data with a NN discriminant and a combined ME+BDT technique is performed. The analysis sets an observed (expected) limit on σ_{95}/σ_{SM}= 5.0 (5.8) for the NN analysis and σ_{95}/σ_{SM}= 5.8 (5.6) for the ME+BDT analysis for a Higgs boson mass $m_H = 115$ GeV/c^2.

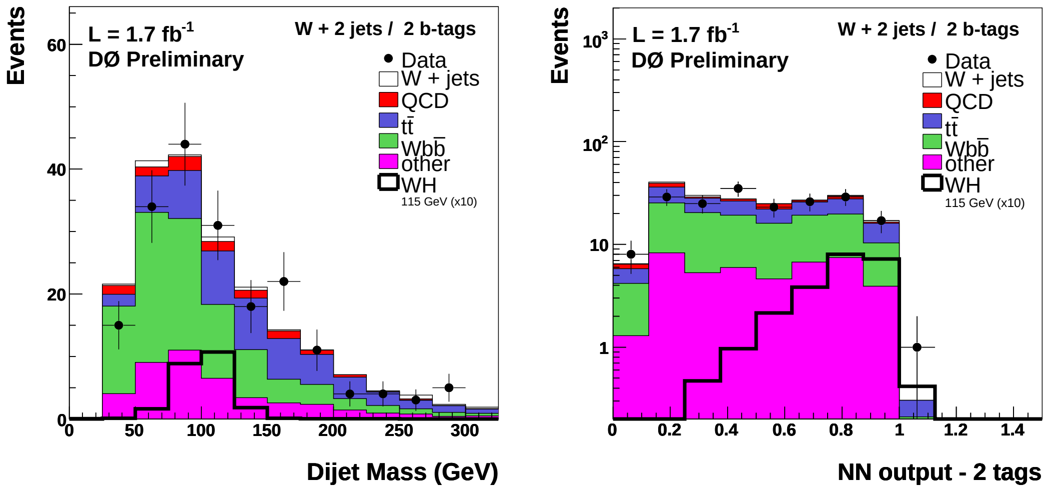

Fig. 2: DØ $WH \to \ell\nu b\bar{b}$ channel: Di-jet invariant mass distribution for events with two b-tags and the NN distribution at the final stage of the selection.

2.4 $ZH \to \nu\nu b\bar{b}$

The channel $ZH \to \nu\nu b\bar{b}$ has very good sensitivity since the branching ratios for $Z \to \nu\nu$ and $H \to b\bar{b}$ decays are large. With the two b-jets being boosted in the transverse direction, the signature for the final state are acoplanar di-jets and large missing transverse energy. Thus is in contrast to most background di-jet events which are expected to be back-to-back in the transverse plane. The main background sources in this search channel are W boson or Z boson production in association with heavy flavour jets, multi-jet events and $t\bar{t}$ pairs.

The basic selection requires at least one (CDF) or two jets (DØ) with a b tag, large missing transverse energy ($E_T^{\rm miss} > 50$ GeV), and a veto on any isolated muon or electron in the event.

In the CDF analysis, the final sample is divided into three samples, one sample with exactly one tight secondary vertex b tag, the second sample with one tight secondary vertex b tag and one tag with the JetProb algorithm and a third sample with two tight secondary vertex b tags. Two NNs are trained one against the dominant QCD background (see Figure 3 on the left side for the second b-tag sample) and one against di-boson and $t\bar{t}$ background (see Figure 3 on the right side for the second sample), which is also used to extract limits on the production cross section.

In the case of DØ, events with two NN b tags are used to construct a BDT for identifying signal events. Asymmetric operating points, one loose and one tight, are chosen for the two b

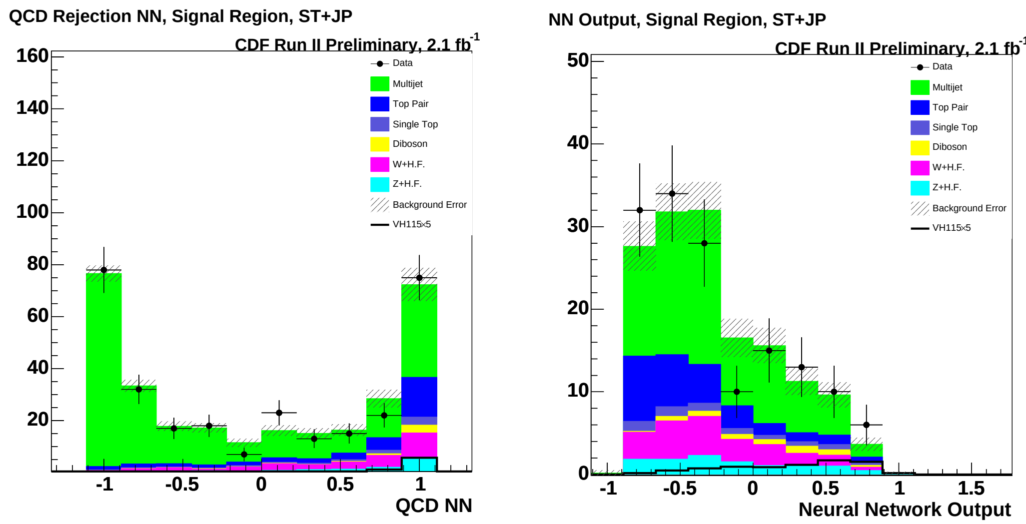

Fig. 3: CDF $ZH \rightarrow \nu\nu b\bar{b}$ channel: NN output distribution to separate against the dominate QCD background (left) and the NN distribution for the remaining backgrounds (right).

tags. The output distributions of the BDT, retrained for every Higgs mass, is shown in Figure 4 on the right side.

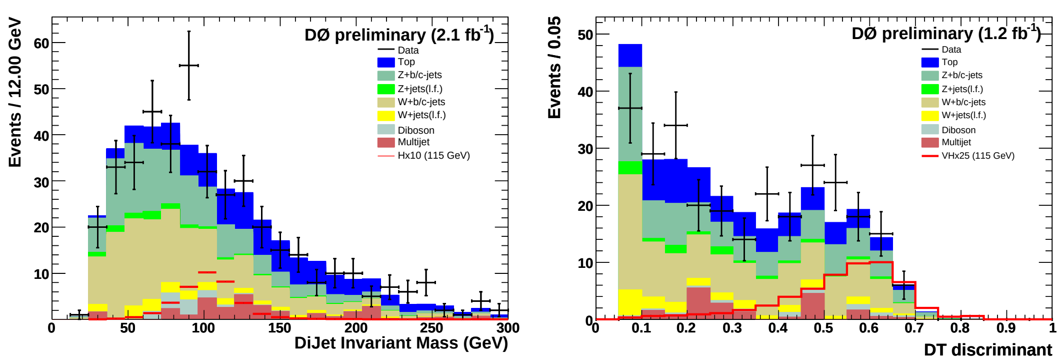

Fig. 4: DØ $ZH \rightarrow \nu\nu b\bar{b}$ channel: Invariant dijet Mass Distribution (left) and output distribution of the BDT variable (right).

To increase the sensitivity of this analysis, WH signal events where the charged lepton has not been identified are also included in the signal definition. This search yields a median observed (expected) upper limit on the $VH(V = W, Z)$ production cross-section of $\sigma_{95}/\sigma_{SM} = 7.9(6.3)$ for CDF and $7.5(8.4)$ for DØ at a Higgs mass of $m_H = 115$ GeV/c^2. The data set for both experiments corresponds to 2.1 fb^{-1} of analyzed data.

2.5 $ZH \to \ell\ell b\bar{b}$

In the $ZH \to \ell\ell b\bar{b}$ channel the Z boson is reconstructed through the decay into two high-p_T isolated muons or electrons. The reconstructed Z and two b-tagged jets are used to select the Higgs signal. The invariant mass of the two leptons is required to be in the Z mass range $70 < m_Z < 110$ GeV/c^2 (DØ) or $76 < m_Z < 106$ GeV/c^2 (CDF). Both experiments require two jets with either one tight b tag or two loose b tags.

The main background sources are Z production in association with heavy jets and $t\bar{t}$ production. ZZ production is an irreducible background, apart from the mass discriminant. CDF trains two separate NNs to reject these two background components. Slices of the output of these NNs, projected on the two axes, is shown in Figure 5. The di-jet mass resolution is improved by training a different NN using E_T^{miss} and the kinematics of both jets. The data set corresponds to an integrated luminosity of 2.4 fb^{-1}. The DØ analysis is performed with 2.3 fb^{-1} of data using a kinematic NN and two NN b tag samples with one tight b tag and two loose b tags.

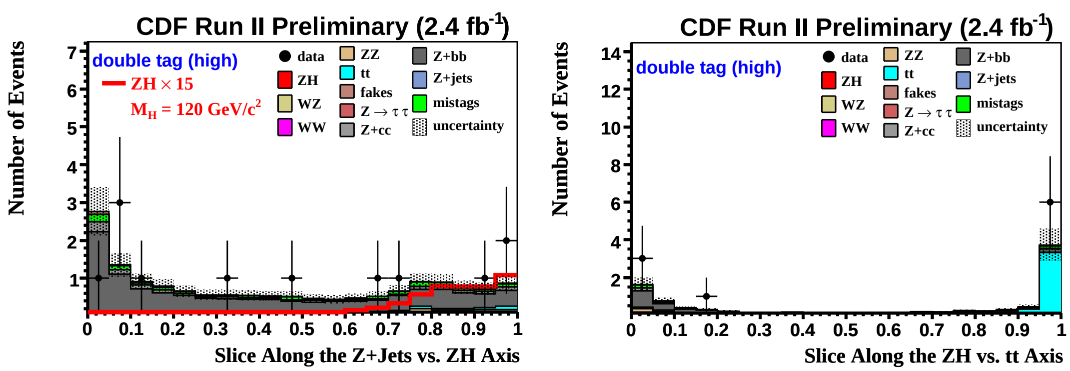

Fig. 5: $ZH \to \ell\ell b\bar{b}$ channel: NN output projection with $y \leq 0.1$ in the Z+Jets vs. ZH projections and $x \geq 0.9$ in the ZH vs. $t\bar{t}$ projection.

These searches yield a median observed (expected) upper limit on the ZH production cross-section of $\sigma_{95}/\sigma_{SM} = 11.6(11.8)$ for CDF and $11.0(12.3)$ for DØ at a Higgs mass of $m_H = 115$ GeV/c^2. Even though the limits are less stringent than for the $ZH \to \nu\nu b\bar{b}$ channel, they still provide an important input to increase the overall sensitivity of the analysis.

2.6 $W \to WW \to \ell\nu\ell\nu$

The dominant decay mode for higher Higgs masses is $H \to WW^{(*)}$. Leptonic decays of the W bosons are therefore used to suppress the QCD background. The signature of the $gg \to H \to WW^{(*)}$ channel is two high-p_T opposite signed isolated leptons with a small azimuthal separation, $\Delta\phi_{\ell\ell}$, due to the spin-correlation between the final-state leptons in the decay of the spin-0 Higgs boson. In contrast, the lepton pairs from background events, mainly WW events, are predominantly back-to-back in $\Delta\phi_{\ell\ell}$. This is shown in Figure 6 (left) for a preselected CDF data sample with zero reconstructed jets.

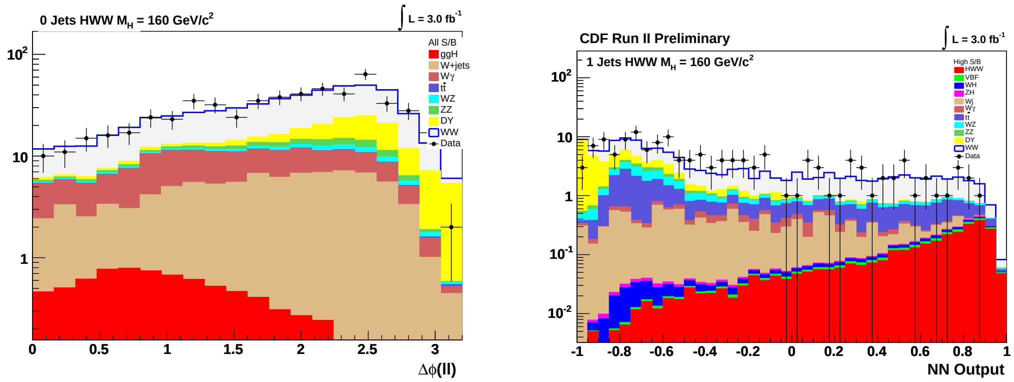

Fig. 6: CDF WW channel: azimuthal angle between the two leptons in the $H \to WW$ search. Due to spin correlations, the signal is at low $\Delta\phi_{\ell\ell}$, whereas the background is at high $\Delta\phi_{\ell\ell}$.

An additional selection requires $E_T^{\text{miss}} > 25$ GeV for CDF and $E_T^{\text{miss}} > 20$ GeV for DØ to account for the neutrinos in the final state. DØ defines three final states (e^+e^-, $e^\pm\mu^\mp$, and $\mu^+\mu^-$). CDF separates the $H \to W^+W^-$ events into five non-overlapping samples, first by separating the events by jet multiplicity (0, 1 or 2), then subdviding the 0 and 1 jet samples in two, one having a low signal/bacgkround (S/B) ratio, the other having a higher one. In these analyses, the final discriminants are neural-network outputs based on several kinematic variables. These include likelihoods constructed from matrix-element probabilities as input to the neural network for CDF and is shown on the right side of Figure 6. The background subtracted NN distribution for DØ is shown in Figure7 on the left side. This distribution has been used to extract median observed (expected) limits on the production cross-section of $\sigma_{95}/\sigma_{SM} = 1.9\,(2.0)$ for $m_H = 165$ GeV/c^2. The obtained limits on the production cross-section as a function of the Higgs boson mass are shown in Figure 7 on the right side. With the NN distributions CDF obtains $\sigma_{95}/\sigma_{SM} = 1.7(1.6)$ for $m_H = 165$ GeV/c^2. The data sets analyzed correspond to an integrated luminosity of 3 fb^{-1} for each experiment.

2.7 $WH \to WWW^* \to \ell\nu\ell'\nu q\bar{q}$

In the process $WH \to WWW^* \to \ell\nu\ell'\nu q\bar{q}$ the Higgs boson is produced in association with a W boson and subsequently decays into a WW pair. This process is important in the intermediate mass range. The signature is at least two isolated leptons from the W decays with $p_T > 15$ GeV and identical charge. The associated W and one of the two W bosons from the Higgs decay should have the same charge. For the final signal selection DØ used a two-dimensional likelihood based on the invariant mass of the two leptons, the missing transverse energy and their azimuthal angular correlations.

This same-sign charge requirement is very powerful in rejecting background from Z production. The remaining background is either due to di-boson production or due to charge mis-measurements. The rate of charge mis-measurements for muons is determined by comparing the independent charge measurements within the solenoidal and in the toroidal fields of the DØ

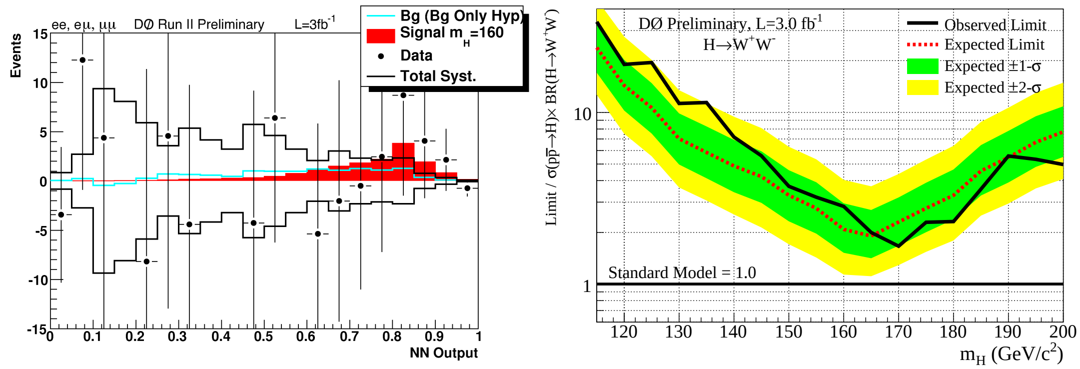

Fig. 7: DØ WW channel: The background subtracted distribution of the NN (left) and the obtained median observed and expected limits on the production cross-section (right).

detector. For electrons the charge mis-measurement rate is determined by comparing the charge measurement from the solenoid with the azimuthal offset between the track and the calorimeter cluster associated to the electron.

The expected cross-section ratio in the mass range 140 GeV/c^2 to 180 GeV/c^2 is $\sigma_{95}/\sigma_{SM} \simeq 20$, i.e. this channel makes a significant contribution at the limit in this mass range.

3 Combined Tevatron Limit

The data of both experiments have been combined using the full set of analyses with luminosities up to 3.0 fb^{-1}. To gain confidence that the final result does not depend on the details of the statistical method applied, several types of combination were performed, using both Modified Frequentist (sometimes called the LEP CL_s method) and Bayesian approaches. The results agree within about 10%. Both methods use Poisson likelihoods and rely on distributions of the final discriminants, e.g. NN output or di-jet mass distributions, not only on event counting.

Systematic uncertainties enter as uncertainties on the expected number of signal and background events, as well as on the shape of the discriminant distributions. The correlations of systematic uncertainties between channels, different background sources, background and signal and between experiments are taken into account. The main sources of systematic uncertainties are, depending on channel, the luminosity and normalisation, the estimates of the multi-jet backgrounds, the input cross-sections used for the MC generated background sources, the higher order corrections (K factors) needed to describe heavy flavour jet production, the jet energy scale, b tagging and lepton identification.

The combinations of results of each single experiment, yield the following ratios of 95% C.L. observed (expected) limits to the SM cross section: 4.2 (3.6) for CDF and 5.3 (4.6) for DØ at $m_H = 115$ GeV/c^2, and 1.8 (1.9) for CDF and 1.7 (2.3) for DØ at $m_H = 170$ GeV/c^2.

The ratios of the 95% C.L. expected and observed limit to the SM cross section are shown in Figure 8 for the combined CDF and DØ analyses on the left side. The observed and median expected values are 1.2 (1.2) at $m_H = 165$ GeV/c^2, 1.0 (1.4) at $m_H = 170$ GeV/c^2 and 1.3

(1.7) at $m_H = 175$ GeV/c^2. On the right side in Figure 8 the 1-CL_S distribution as a function of the Higgs boson mass, which is directly interpreted as the level of exclusion of the search. For instance, both the observed and expected results exclude a Higgs boson with $m_H = 165$ GeV/c^2 at ≈ 92% C.L. The green and yellow bands show the one and two sigma bands for background fluctuations. We exclude at the 95% C.L. the production of a standard model Higgs boson with mass of 170 GeV/c^2.

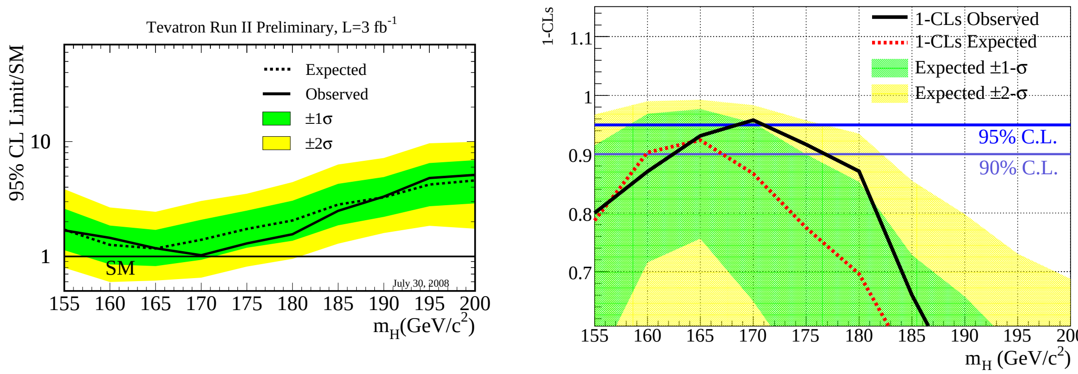

Fig. 8: Expected and observed 95% CL cross-section ratios for the combined CDF and DØ analyses. (status July 2008).

References

[1] LEP Working Group for Higgs boson searches Collaboration, R. Barate *et al.*, Phys. Lett. **B565**, 61 (2003). hep-ex/0306033.

[2] Tevatron Electroweak Working Group Collaboration (2008). http://tevewwg.fnal.gov/.

[3] LEP Electroweak Working Group Collaboration (2008). http://lepewwg.web.cern.ch/LEPEWWG/.

[4] CDF Collaboration, D. E. Acosta *et al.*, Phys. Rev. **D71**, 032001 (2005). hep-ex/0412071.

[5] T. D. M. G. V. M. Abazov, NUCL.INSTRUM.METH.A **552**, 372 (2005).

[6] T. Sjostrand, L. Lonnblad, and S. Mrenna (2001). hep-ph/0108264.

[7] S. Catani, D. de Florian, M. Grazzini, and P. Nason, JHEP **07**, 028 (2003). hep-ph/0306211.

[8] Higgs Working Group Collaboration, K. A. Assamagan *et al.* (2004). hep-ph/0406152.

[9] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56 (1998). hep-ph/9704448.