



A Search for the Standard Model Higgs Boson in the Channel $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration
URL <http://www-d0.fnal.gov>
(Dated: November 1, 2007)

This note describes a search at DØ for the Standard Model Higgs boson produced in association with a Z boson, where the Higgs decays to a $b\bar{b}$ pair and the Z boson decays to neutrinos. An integrated luminosity of $0.93 \pm 0.06 \text{ fb}^{-1}$ was used. The $p\bar{p} \rightarrow ZH \rightarrow \nu\bar{\nu}b\bar{b}$ channel is one of the most sensitive ways to search for a light mass Higgs, but is problematic at hadron colliders due to the absence of visible leptons and the presence of only two jets in the final state. We thus require a large missing transverse energy and that the jets be b -tagged. The WH processes in which the charged lepton from the W decays is either undetected or decays hadronically (τ) also exhibit the jets+ \cancel{E}_T signature and are therefore considered part of the signal. Artificial neural networks are used to separate signal from background. In the absence of a significant excess in the data above the background expectation, we set limits on $\sigma(p\bar{p} \rightarrow VH) \times (H \rightarrow b\bar{b})$ at the 95% confidence level of 2.6 pb – 2.3 pb for Higgs boson masses in the range 105 – 135 GeV, where $V = W, Z$. The corresponding expected limits range from 2.8 pb – 2.0 pb.

Preliminary Result

I. INTRODUCTION

The Higgs mechanism is the best candidate to explain electroweak symmetry breaking. It predicts the existence of the Higgs boson, which is as yet undiscovered. The LEP experiments have placed a lower limit on its mass at 114.4 GeV at 95% confidence level (CL) [1]. Electroweak global fits prefer a light mass Higgs, currently $m_H < 144$ GeV at 95% CL [2]. At these masses, the Tevatron has significant discovery potential, details of which can be found elsewhere [3, 4].

The $p\bar{p} \rightarrow ZH \rightarrow \nu\bar{\nu}b\bar{b}$ channel is one of the most sensitive ways to search for a light mass Higgs because of the large $H \rightarrow b\bar{b}$ and $Z \rightarrow \nu\bar{\nu}$ branching ratios. The two b -jets from the Higgs are boosted along the direction of the Higgs momentum and so tend to be more acoplanar than the dijet background. There are two major sources of background: (i) physics backgrounds such as Z +jets, W +jets, electroweak diboson production or top quark production with missed leptons and jets and (ii) the instrumental background resulting from calorimeter mismeasurements which can lead to large \cancel{E}_T signals with the presence of jets from QCD processes. Selecting events with a relatively large \cancel{E}_T and two b -tagged jets eliminates much of the physics background. This analysis updates DØ's previously published result in this channel [5].

II. DATA SAMPLE AND EVENT SELECTION

The Run II DØ detector has a central-tracking system, consisting of a silicon microstrip tracker and a central fiber tracker, both located within a 2 T superconducting solenoidal magnet, with designs optimised for tracking and vertexing at pseudorapidities $|\eta| < 3$ and $|\eta| < 2.5$, respectively. The liquid-argon and uranium calorimeter has a central section (CC) covering pseudorapidities $|\eta|$ up to ≈ 1.1 , and two end calorimeters (EC) that extend coverage to $|\eta| \approx 4.2$, with all three housed in separate cryostats. An outer muon system, at $|\eta| < 2$, consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroids, followed by two similar layers after the toroids. Luminosity is measured using plastic scintillator arrays placed in front of the EC cryostats. Full details of the Run II DØ detector are given elsewhere [6].

Dedicated triggers designed to select events with acoplanar jets and large missing E_T were used. After data quality cuts the total data sample is 0.93 fb^{-1} [7]. As the data were taken with different trigger versions, the exact criteria vary, but typical requirements at the highest level trigger were a $\cancel{H}_T > 30$ GeV (where $\cancel{H}_T = |-\sum \vec{p}_{T\text{jet}}|$) and an azimuthal angle, ϕ , between the two leading jets of $\Delta\phi(\text{jet}_1, \text{jet}_2) < 170^\circ$. A parameterised trigger simulation was used to model the effects of the trigger requirements on the simulated events. Each trigger version was simulated separately and the different versions combined with the appropriate luminosity weighting.

The basic event selection is as follows:

- Two jets with $p_T > 20$ GeV and $|\eta| < 1.1$ or $1.4 < |\eta| < 2.5$ ¹.
- $\Delta\phi(\text{jet}_1, \text{jet}_2) < 165^\circ$.
- $\cancel{E}_T > 50$ GeV.
- The presence of a primary vertex with a Z position, $|Z_{PV}| < 35$ cm, with at least three attached tracks.
- $H_T < 240$ GeV (where H_T is the scalar sum of the jet p_T).
- No isolated leptons (electron or muon)².

The first three requirements select the basic signal topology and requiring a good vertex maximises the b -tagging discrimination potential. The last two requirements are designed to reduce the contribution of $t\bar{t}$ events; if the W from the top quark decays to jets then the total scalar sum of the jets in the event will be large and conversely if the W decays leptonically then the isolated lepton cut will reject these. The isolated lepton cut also eliminates a significant proportion of W/Z +jets processes with W/Z decays into charged leptons. The jets are required to pass

¹ Jets in the intercryostat region ($1.1 < |\eta| < 1.4$) have worse energy resolution and so are excluded from this analysis.

² For muons we require no objects with segments in the A and BC layer of the muon system [6] matched or not with a central track with $p_T > 8$ GeV and separated from jets by $\Delta R(\text{muon}; \text{jet}) > 0.5$. Electrons are identified as having a calorimeter EM fraction of 0.9, the EM cluster must be matched to a track having $p_T > 5$ GeV and an 'electron likelihood' > 0.2 . If the isolation parameter, namely the ratio of the difference of the total energy in a cone with radius $R=0.4$ and the EM energy inside a cone with $R=0.2$ over the EM energy inside the $R=0.2$ cone, is greater than 0.2 then the event is rejected.

basic quality cuts to remove fake jets. Corrections are applied as a function of η and p_T to correct the jet energies to the particle level. As in Ref [5] additional requirements are made to reduce the instrumental background and are detailed in Section IV.

III. SIMULATED EVENT SAMPLES

The samples listed below were used to determine the number of expected signal and background events:

- Signal samples, $ZH \rightarrow \nu\nu b\bar{b}$, $WH \rightarrow e\nu_e b\bar{b}$, $WH \rightarrow \mu\nu_\mu b\bar{b}$ and $WH \rightarrow \tau\nu_\tau b\bar{b}$ were generated for Higgs masses, m_H , from 105 to 135 GeV using PYTHIA 6.323 [8]. The WH processes in which the charged lepton from the W decays is either undetected or decays hadronically (τ) also exhibit the jets+ \cancel{E}_T signature and are therefore considered part of the signal.
- $t\bar{t}$ production with up to 4 jets, generated with ALPGEN v2.05 [9].
- W +jets (including jj , $b\bar{b}$ and $c\bar{c}$ jets separately) and Z +jets (including $Z \rightarrow \nu\bar{\nu}$ and $Z \rightarrow \tau\tau$ processes for jj , $b\bar{b}$ and $c\bar{c}$ jets) samples were generated using ALPGEN.
- Diboson processes, namely WW , WZ and ZZ , were generated with PYTHIA.

Those samples generated with ALPGEN were processed through PYTHIA for showering and hadronisation. NLO cross sections were used for normalization for all processes. All samples were processed through the DØ detector simulation, the readout simulation and the reconstruction software. Prior to b -tagging the jets are required to be ‘taggable’ i.e. to satisfy certain tracking and vertexing criteria³. The fraction of taggable jets was investigated as a function of p_T , η and the z -position of the primary vertex (Z_{PV}) using a W +jets data sample and Monte Carlo (MC). The scale factor, the ratio of the taggability in data and MC, was found to only depend on η . Scale factors of 0.97 ± 0.01 and 0.95 ± 0.03 (statistical error only) are used for the central and endcap regions respectively. Residual differences between data and simulation are taken into account by smearing the simulation.

IV. INSTRUMENTAL BACKGROUND ESTIMATION

The following additional requirements are introduced to reduce the instrumental background:

- The minimum of the difference in azimuthal angle, ϕ , between the direction of the \cancel{E}_T and any of the jets, $\min\{\Delta\phi_i(\cancel{E}_T, \text{jet}_i)\} > 0.15$, and the related cut, \cancel{E}_T (GeV) $> -40 \times \min\{\Delta\phi_i(\cancel{E}_T, \text{jet}_i)\} + 80$.
- The difference in azimuthal angle between the direction of \cancel{E}_T and p_T^- : $\Delta\phi(\cancel{E}_T, p_T^-) < \pi/2$. p_T^- is the magnitude of the negative of the vector sum of all track p_T s.
- The asymmetry between \cancel{E}_T and \cancel{H}_T : $-0.1 < A(\cancel{E}_T, \cancel{H}_T) < 0.2$, where $A(\cancel{E}_T, \cancel{H}_T) \equiv (\cancel{E}_T - \cancel{H}_T)/(\cancel{E}_T + \cancel{H}_T)$.

For events originating from hard scattering with genuine missing transverse energy the \cancel{H}_T , \cancel{E}_T and p_T^- all point in the same direction and are correlated. By contrast, dijet events in which one of the jets has been mismeasured to give a \cancel{E}_T signal have $\Delta\phi(\cancel{E}_T, p_T^-)$ in the same direction as one of the jets. The variables mentioned above were studied in W +jets data, an orthogonal sample to that used in the final analysis, to ensure that they are well modeled. Distributions of the p_T for the leading and next-to-leading jet, along with \cancel{E}_T and $\Delta\phi(\cancel{E}_T, p_T^-)$, are shown in Fig. 1.

Whilst the requirements above reduce the instrumental background, its absolute scale still needs to be determined from data to calculate any residual contribution. The $A(\cancel{E}_T, \cancel{H}_T)$ and $\Delta\phi(\cancel{E}_T, p_T^-)$ distributions are used to make this estimate. The simulated signal and physics backgrounds peak around zero for both variables. The signal region is defined as having $\Delta\phi(\cancel{E}_T, p_T^-) < \pi/2$ and the sideband region as having $\Delta\phi(\cancel{E}_T, p_T^-) > \pi/2$. To estimate the scale of the instrumental background we consider the $A(\cancel{E}_T, \cancel{H}_T)$ distribution for the signal and sideband regions. The shape of the physics backgrounds is, for either region, taken directly from simulation. We fit a sixth-order polynomial to the $A(\cancel{E}_T, \cancel{H}_T)$ distribution in the sideband region to determine the shape for the instrumental background

³ For a jet to be taggable, it is required to have at least 2 tracks, with $p_T > 1$ and 0.5 GeV, each with ≥ 2 SMT hits and $\Delta R(\text{track}, \text{jet}) < 0.5$.

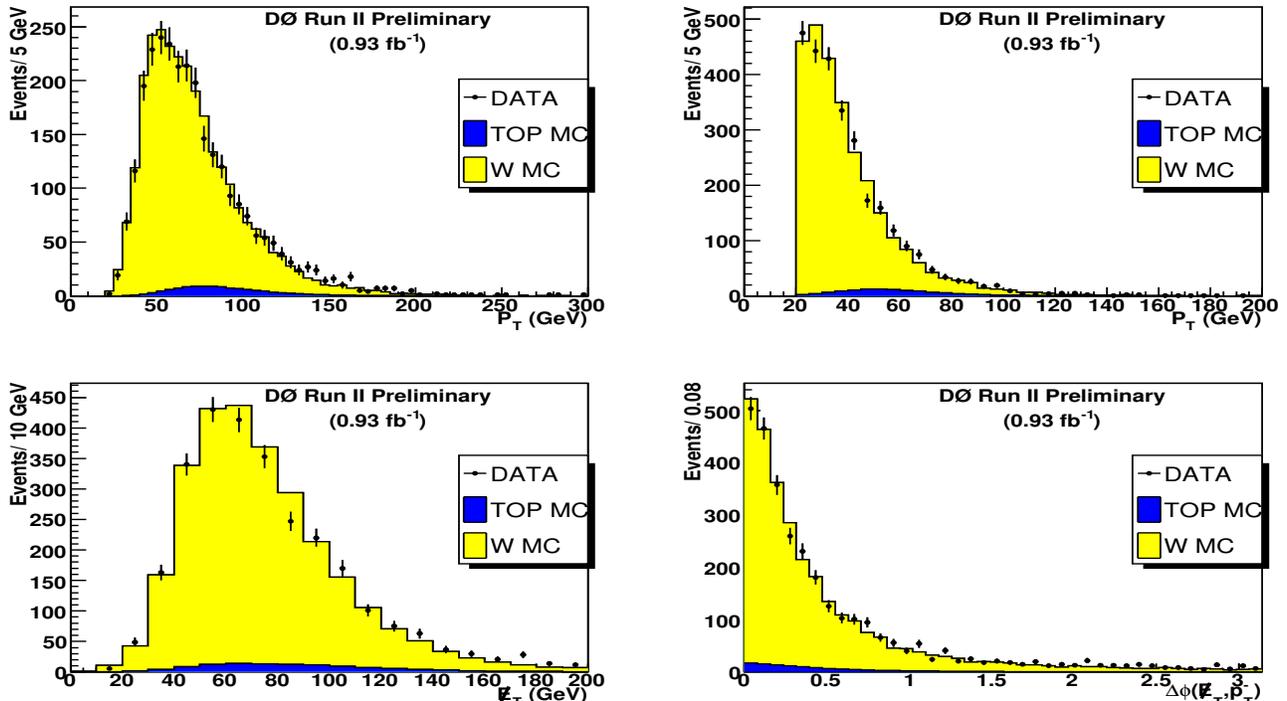


FIG. 1: Distributions of the p_T for the leading (upper left) and next-to-leading jet (upper right), \cancel{E}_T (lower left) and $\Delta\phi(\cancel{E}_T, p_T^-)$ (lower right) for W +jets events before b -tagging.

(having subtracted the physics background contamination.) Thus, having determined the shape of the instrumental $A(\cancel{E}_T, \cancel{H}_T)$ distribution we then make a combined (Monte Carlo and instrumental) fit to data in the signal region, as shown in Fig. 2. The sixth-order polynomial fits the data well. For this combined fit the simulation and instrumental shapes are fixed to those from the previous fits and only the absolute scale of the two types of background is allowed to float. The normalisation of the physics background is found to be 1.06 ± 0.02 (statistical error only) in good agreement with the expected cross sections. For all the other distributions the instrumental background also has to have its shape adjusted for the physics background contamination in the sideband region. Distributions of p_T for the leading and next-to-leading jet, \cancel{E}_T and dijet mass are shown after selection cuts, taggability corrections and the normalisation of the instrumental and physics backgrounds in Fig. 3.

V. B-TAGGED RESULTS

The standard $D\bar{O}$ neural net b -tagging algorithm was used [10]; this uses as input the output from several lifetime based b -tagging algorithms. The final selection uses the following optimum b -tagging combination: one tight b -tag (b -tag efficiency $\sim 50\%$ for a mistag rate of $\sim 0.4\%$) and one loose b -tag (b -tag efficiency $\sim 70\%$ for a mistag rate of $\sim 4.5\%$). Table I shows the number of expected events from simulation and instrumental background, along with the number seen in data, before and after b -tagging requirements. After b -tagging 134 ± 18 events are expected, and 140 events are observed. Fig. 4 shows variable distributions after applying the optimised b -tagging.

VI. NEURAL NETWORK

To improve the separation of signal from backgrounds, an artificial neural network (NN) is used. The NN is trained and tested using half of the background and signal MC events, the rest being used to derive signal cross section limits. The QCD background contribution is not taken into account during training. The signal sample used for training is a combination of the ZH and WH contributions. Events were weighted such that the total contribution of each sample made up that expected after b -tagging. A separate NN is trained for each value of m_H . The NN input variables

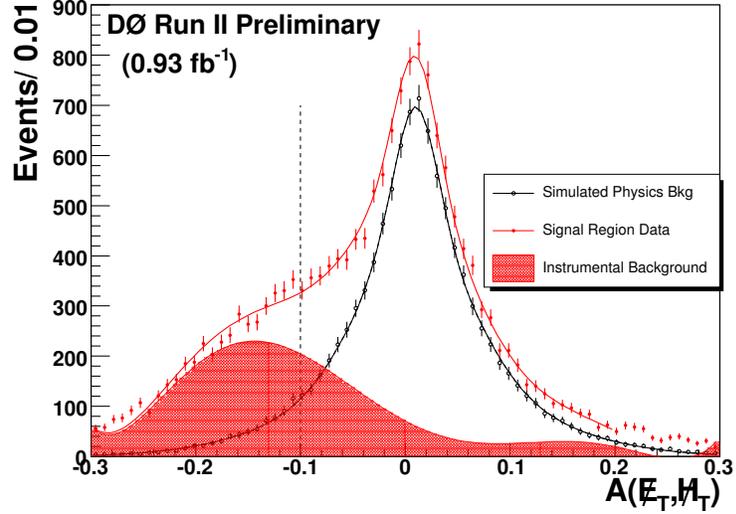


FIG. 2: $A(E_T, \#H)$ for data, physics background and instrumental background in the signal region before b -tagging. The selection cut is $-0.1 < A(E_T, \#H) < 0.2$, the line at -0.1 indicates the lower edge of this range.

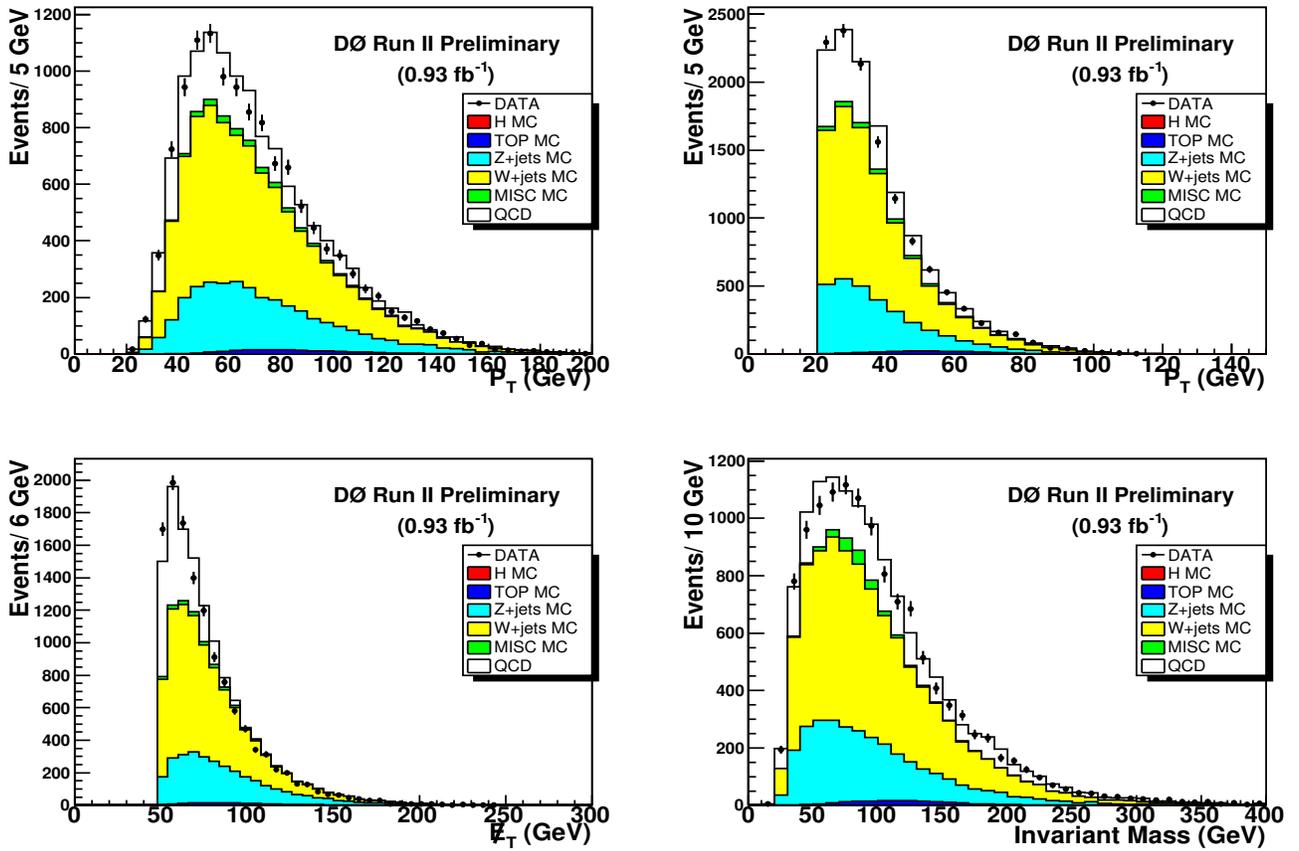
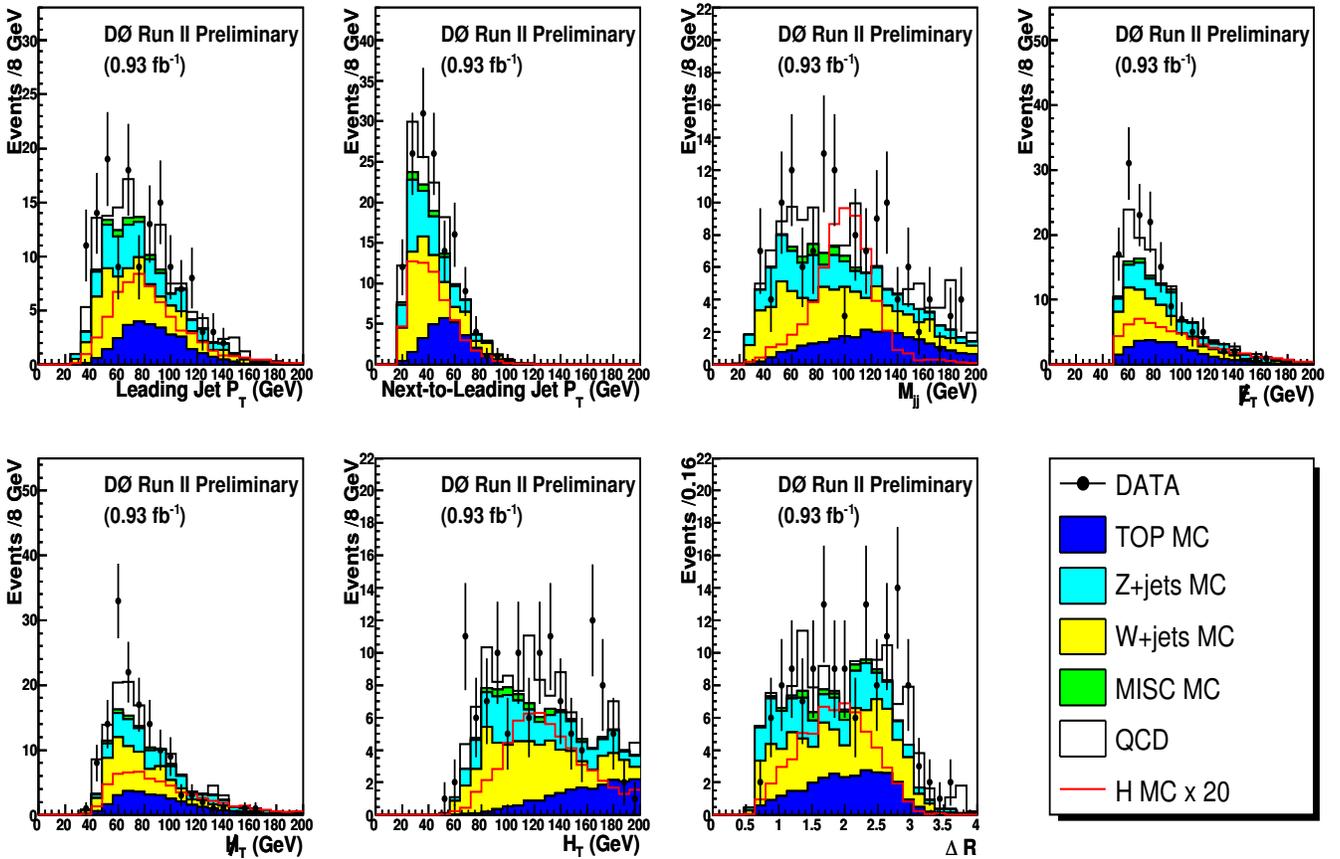


FIG. 3: Distributions of the p_T for the leading (upper left) and next-to-leading jet (upper right), $\#E_T$ (lower left) and dijet mass (lower right) before b -tagging.

considered are the invariant mass of the two leading jets in the event (M_{jj}), the ΔR between the two jets, the p_T of

Sample	No b -tag	Double b -tag
$ZH(m_H = 115 \text{ GeV})$	2.46	0.88 ± 0.12
$WH(m_H = 115 \text{ GeV})$	1.75	0.61 ± 0.08
$W + jets$	6750	52.3
$\rightarrow Wbb$	397	35.4
$\rightarrow Wcc$	1170	9.33
Zjj		
$Z \rightarrow \tau\tau$	107	0.25
$Z \rightarrow \nu\bar{\nu}$	2130	0.63
Zbb		
$Z \rightarrow \tau\tau$	6.39	0.63
$Z \rightarrow \nu\bar{\nu}$	229	24.9
Zcc		
$Z \rightarrow \tau\tau$	12.8	0.18
$Z \rightarrow \nu\bar{\nu}$	467	4.93
$t\bar{t}$	172	29.1
$Di - boson$	228	3.84
Total Physics Background	10100	117 ± 17
Instrumental Background	2560	17.2 ± 3.4
Total Background	12700	134 ± 18
Observed Events	12500	140

TABLE I: Number of events after final selection.

FIG. 4: Distributions of the NN input variables, the p_T for the leading and next-to-leading jet, dijet mass, E_T , H_T , H_T and ΔR , from left to right and top to bottom, after neural net b -tagging.

the leading jet, the p_T of the next-to-leading jet, the \cancel{E}_T , the \cancel{H}_T and H_T . Distributions of these variables are shown in Fig. 4. The input variables were selected for their ability to separate signal and background, as given by their impact on the training sample total error reduction. Each NN was trained for 200 epochs, using 14 hidden neurons (in a single layer), and one output neuron. The NN outputs for signal, background and data are shown in Fig. 5.

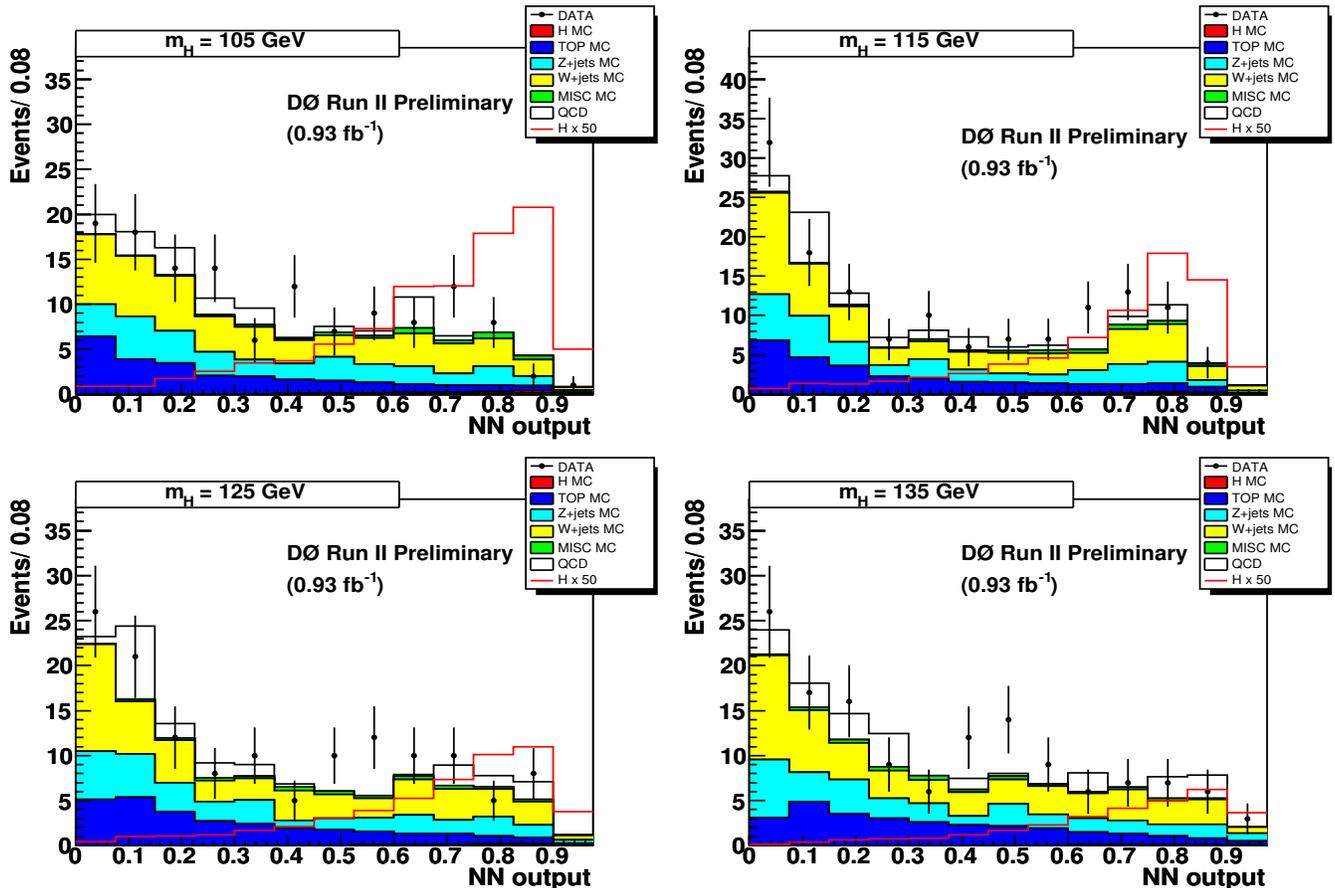


FIG. 5: NN output distributions for mass points in the range $105 \leq m_H \leq 135$ GeV.

VII. SYSTEMATIC ERRORS

Systematic errors associated with the luminosity, trigger efficiencies, jet ID, b -tagging, background MC cross section and QCD background are estimated and included in the limit setting procedure. All systematic errors are common and correlated between signal and background except for the uncertainty on the cross section and the QCD contribution. Table II lists these systematic errors.

Luminosity	Trigger	Jet ID	b -tagging	Background σ	QCD multijets
6.1	5	5	7	6-18	20

TABLE II: Table of systematic errors in %.

Errors are also estimated for the difference in the shape of the NN output when varying the Jet Energy Scale by $\pm 1\sigma$ separately for all signal and background MC samples at each mass point. The difference in the shape of the NN output because of the uncertainty in the shape of the MC dijet mass spectrum at each mass point is also taken into account. The former source of error was estimated to be $\leq 10\%$ and the latter $\leq 8\%$; both are included in the limit setting.

VIII. CROSS SECTION LIMITS

Higgs Mass (GeV)	$m_H = 105$	$m_H = 115$	$m_H = 125$	$m_H = 135$
$ZH(H \rightarrow bb)$ Expected Limit (pb)	1.6	1.5	1.4	1.2
$ZH(H \rightarrow bb)$ Observed Limit (pb)	1.5	1.5	1.4	1.3
$WH(H \rightarrow bb)$ Expected Limit (pb)	4.8	4.3	3.8	3.6
$WH(H \rightarrow bb)$ Observed Limit (pb)	4.4	5.0	4.4	4.2
$VH(H \rightarrow bb)$ Expected Limit (pb)	2.8	2.5	2.3	2.0
$VH(H \rightarrow bb)$ Observed Limit (pb)	2.6	2.7	2.5	2.3

TABLE III: Limits on the SM Higgs cross section.

As no significant excess is observed we proceed to set a limit on the Higgs production cross section. A modified frequentist approach with a Poisson log-likelihood ratio (LLR) statistic is used [11, 12]. The impact of systematic uncertainties is incorporated through marginalisation of the Poisson probability distributions for signal and background via Gaussian distributions. All correlations in the systematic errors are maintained between signal and background. The expected distributions for background are evaluated by minimising a profile likelihood function, referencing the shape and rate of the distributions in the sideband regions of the final discriminant (NN output). The observed and expected limits (pb) for the processes $\sigma(p\bar{p} \rightarrow ZH) \times (H \rightarrow b\bar{b})$, $\sigma(p\bar{p} \rightarrow WH) \times (H \rightarrow b\bar{b})$ and the combined $\sigma(p\bar{p} \rightarrow VH) \times (H \rightarrow b\bar{b})$ are given in Table III at 95% CL. The expected and observed limits for WH , ZH and combined VH production, divided by the SM cross section, are displayed in Fig. 6. The solid line at 1 corresponds to the expected SM Higgs cross section.

Fig. 7 shows the LLR distributions for the ZH , WH and VH production in the $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ channel. Included in these figures are the LLR values for the signal-plus-background hypothesis (LLR_{s+b}), background-only hypothesis (LLR_b), and the observed data (LLR_{obs}). The shaded bands represent the 1 and 2 standard deviation (σ) departures for LLR_b . These distributions can be interpreted as follows:

- The separation between LLR_b and LLR_{s+b} provides a measure of the overall power of the search. This is the ability of the analysis to discriminate between the signal-plus-background and background-only hypotheses.
- The width of the LLR_b distribution (shown here as 1 and 2 standard deviation (σ) bands) provides an estimate of how sensitive the analysis is to a signal-plus-background-like fluctuation in data, taking account of the presence of systematic uncertainties. For example, when a 1- σ background fluctuation is large compared to the signal expectation, the analysis sensitivity is thereby limited.
- The value of LLR_{obs} relative to LLR_{s+b} and LLR_b indicates whether the data distribution appears to be more signal-plus-background-like or background-like. As noted above, the significance of any departures of LLR_{obs} from LLR_b can be evaluated by the width of the LLR_b distribution.

IX. SUMMARY

We have performed a search for the SM Higgs, produced in association with either a Z or W boson, in the final state topology requiring missing energy and 2 b -tagged jets using 0.93 fb^{-1} of data. In the absence of a significant excess in the data above the background expectation, we set limits on $\sigma(p\bar{p} \rightarrow VH) \times (H \rightarrow b\bar{b})$ at the 95% confidence level of 2.6 pb – 2.3 pb for Higgs boson masses in the range 105 – 135 GeV; the corresponding expected limits range from 2.8 pb – 2.0 pb.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland);

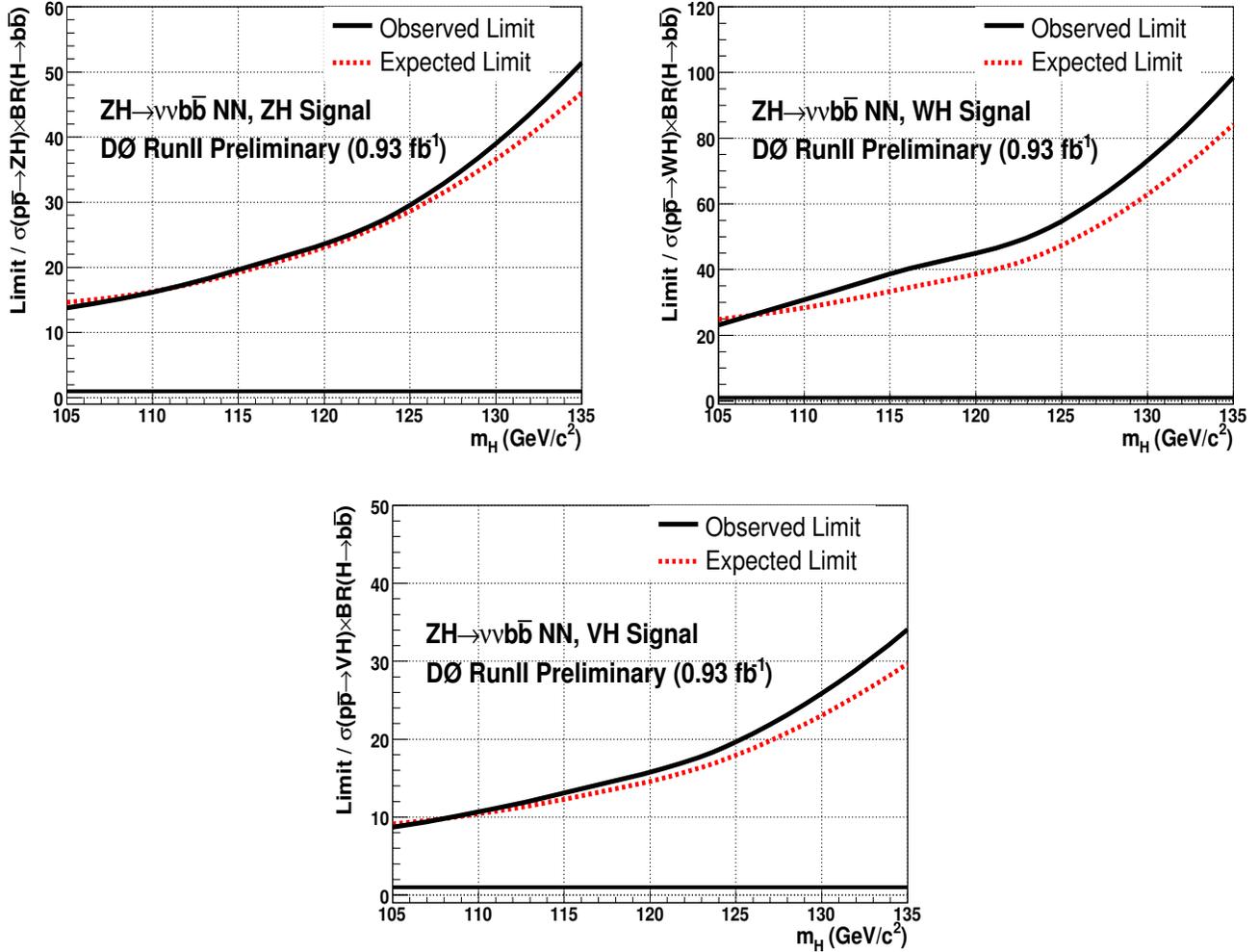


FIG. 6: Observed (solid) and expected (dashed) relative cross section limits at 95% CL on ZH production (left), WH production (right) and combined VH production (bottom).

The Swedish Research Council (Sweden); Research Corporation; Alexander von Humboldt Foundation; and the Marie Curie Program.

-
- [1] LEP Collaborations, Phys. Lett. B **565**, 61 (2003).
 - [2] <http://lepewwg.web.cern.ch/LEPEWWG/>
 - [3] M. Carena *et al.*, “Report of the Tevatron Higgs Working Group”, arXiv:hep-ph/0010338.
 - [4] CDF and DØ Collaborations, “Results of the Tevatron Higgs Sensitivity Study”, FERMILAB-PUB-03/320-E.
 - [5] V. Abazov *et al.* (DØ Collaboration), Phys. Rev. Lett. **97**, 161803 (2006).
 - [6] V. Abazov *et al.* (DØ Collaboration), “The Upgraded DØ Detector”, Nucl. Instrum. Methods Phys. Res. A **565**, 463 (2006).
 - [7] T. Andeen *et al.*, FERMILAB-TM-2365-E (2007).
 - [8] T. Sjöstrand, L. Lönnblad, S. Mrenna and P. Skands, “PYTHIA 6.206”, arXiv:hep-ph/0108264.
 - [9] M. L. Mangano *et al.*, “ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions”, arXiv:hep-ph/0206293.
 - [10] T. Scanlon, “ b -Tagging and the Search for Neutral Supersymmetric Higgs Bosons at DØ”, FERMILAB-THESIS-2006-43.
 - [11] T. Junk, Nucl. Instrum. Methods A **434**, 435 (1999).
 - [12] W. Fisher, “Systematics and Limit Calculations”, FERMILAB-TM-2386-E (2007).

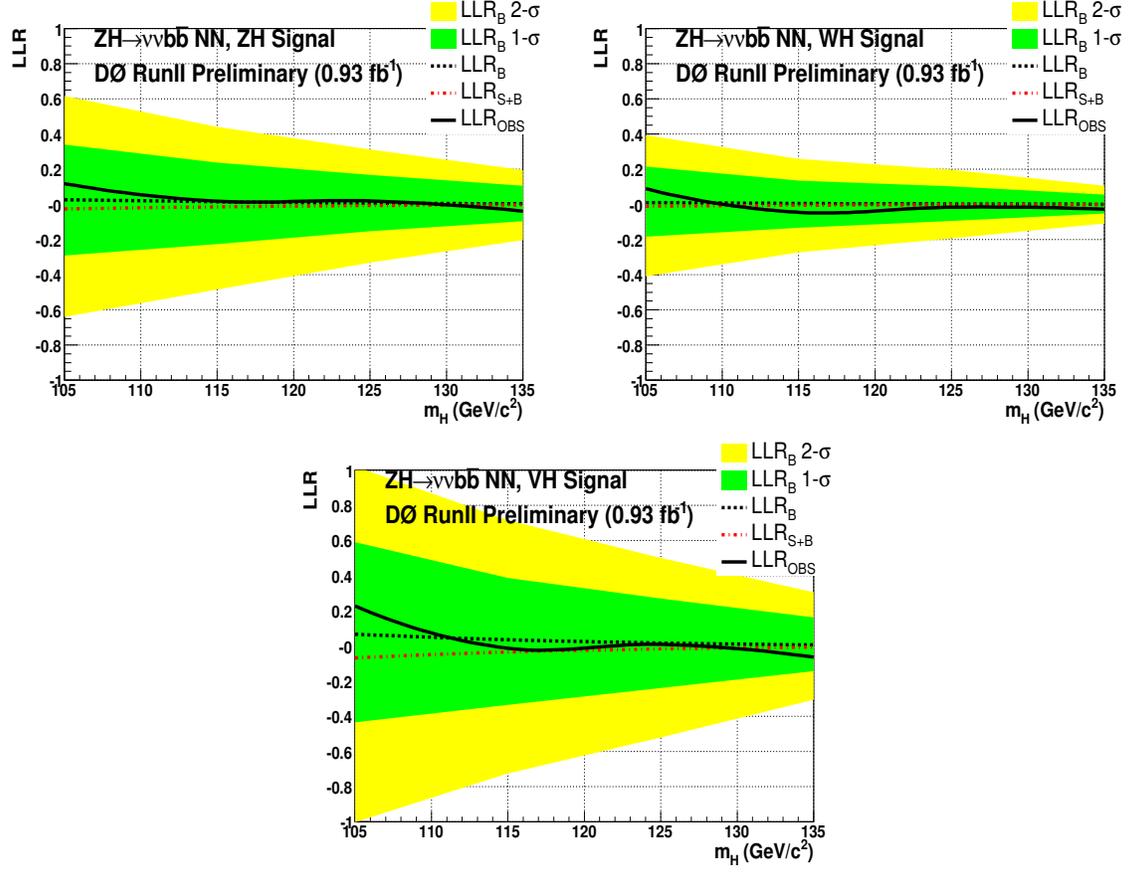


FIG. 7: LLR values for the signal-plus-background hypothesis (LLR_{s+b}), background-only hypothesis (LLR_b), and the observed data (LLR_{obs}). The shaded bands represent the 1 and 2 standard deviation (σ) departures for LLR_b .