DØ Note 5586-CONF

Search for the standard model Higgs boson in the $HZ \to b\bar{b}\nu\bar{\nu}$ channel in 2.1 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration
URL: http://www-d0.fnal.gov
(Dated: February 29, 2008)

A search for the standard model Higgs boson has been performed in 2.1 fb^{-1} of $p\bar{p}$ collisions at 1.96 TeV, collected with the DØ detector at the Fermilab Tevatron. The final state considered is a pair of b jets with large missing transverse energy, as expected from the reaction $p\bar{p} \to HZ \to b\bar{b}\nu\bar{\nu}$. The search is also sensitive to the $HW \to b\bar{b}\ell\nu$ channel, when the charged lepton is not identified. Boosted decision trees were used to discriminate the signal from the backgrounds, dominated by Wbb and Zbb final states. For a Higgs boson mass of 115 GeV, a limit has been set at 95% C.L. on the cross section times branching fraction of $(p\bar{p} \to H(Z/W)) \times (H \to b\bar{b})$, which is 7.5 times larger than the standard model value.

Preliminary Results for Winter 2008 Conferences

I. INTRODUCTION

The $p\bar{p} \to HZ$ reaction, with $H \to b\bar{b}$ and $Z \to \nu\bar{\nu}$, is among the most promising for the discovery of a low mass Higgs boson at the Fermilab Tevatron [1]. The DØ collaboration published a search for this process, based on 260 pb^{-1} [2], and a preliminary result based on 0.93 fb^{-1} was subsequently reported [3]. In this note, an extension of this search to a 2.1 fb^{-1} data set is presented. A lower mass limit of 114.4 GeV was set by the LEP experiments for the Higgs boson from analyses of the reaction $e^+e^- \to HZ$ [4], while an upper limit of 144 GeV can be inferred from precision electroweak data [5]. Here and in the following, all limits quoted are at the 95% confidence level.

The final state topology considered in this analysis is a pair of b jets from the decay of the Higgs boson, with missing transverse energy ($\not\!\!E_T$) due to the neutrinos from the Z decay. The search is therefore also sensitive to the HW channel, with $W \to \ell\nu$ when the charged lepton from the W decay is not detected. The main backgrounds arise from $(W/Z)b\bar{b}$, from (W/Z)+(non-b jets) due to flavor misidentification (mistagging), from top quark production, e.g., $t\bar{t} \to \ell\nu bq\bar{q}'\bar{b}$ or $t(q)\bar{b} \to \ell\nu b(q)\bar{b}$, from diboson production such as $WZ \to q\bar{q}'\nu\bar{\nu}$ or $ZZ \to b\bar{b}\nu\bar{\nu}$, and from multijet events produced by the strong interaction, with fake $\not\!\!E_T$ resulting from fluctuations in jet energy measurements and with real b or mistagged light parton jets.

A kinematic selection is first applied to reject most of the multijet events. The two jets expected from the Higgs boson decay are next required to be tagged as b jets, using a neural network b-tagging algorithm. Finally, discrimination between the signal and the remaining backgrounds is achieved by means of a boosted decision tree technique.

II. DATA AND SIMULATED SAMPLES

The DØ detector [6] consists of a silicon microstrip vertex detector and a fiber tracker, both located within a 2 T superconducting solenoidal magnet. A liquid-argon and uranium sampling calorimeter, housed in three cryostats, provides pseudorapidity coverage out to $|\eta| = 4.2$ with pseudo projective towers of size 0.1×0.1 in (η, ϕ), where η is calculated with respect to the proton beam direction, and ϕ is the azimuthal angle. It is segmented longitudinally into four electromagnetic and up to five hadronic layers. Additional shower sampling is provided by scintillating tiles located at the boundaries between cryostats. Beyond the calorimeter, a muon detector consists of tracking detectors and scintillation trigger counters before and after 1.8 T iron toroids.

Online event selection is accomplished with a three-level trigger system. For this analysis, the data were recorded using a set of triggers designed to select events with jets and missing transverse energy. At the highest trigger level, the main requirement through winter 2006 (Run IIa) was $\not\!\!H_T > 30$ GeV, where $\not\!\!H_T = |-\Sigma \overrightarrow{p}_{T\mathrm{jet}}|$ is a measure of $\not\!\!E_T$ based on jets only. For data collected since spring 2006 (Run IIb), the trigger system was improved [7], allowing in particular the $\not\!\!E_T$ to be used at the first level, and the $\not\!\!H_T$ threshold to be lowered to 25 GeV. Other changes occurred between Run IIa and Run IIb: an additional layer for the silicon tracker was installed close to the beam pipe, thus improving charged particle momentum resolution and heavy flavor identification; the quality and the stability of the calibration of the scintillating tiles of the inter-cryostat detector (ICD) were also greatly improved.

After data quality requirements, the total integrated luminosity available for this analysis was (2.08 ± 0.13) fb^{-1} [8], of which 57% was collected during Run IIb. The physics objects used were charged particle tracks, the primary interaction vertex, calorimeter jets reconstructed in a cone of radius 0.5 by the iterative midpoint cone algorithm [9] and with $p_T > 15$ GeV, electrons or muons identified by the association of charged particles with electromagnetic calorimeter clusters or with hits in the muon detector, respectively. The $\not\!\!E_T$ and total E_T were reconstructed as the opposite of the vectorial sum and as the scalar sum of the transverse energies of all energy deposits in the calorimeter, respectively. Unless otherwise specified, the $\not\!\!E_T$ was corrected for reconstructed muons. The jet energy calibration was performed by requiring transverse energy balance in photon+jet events, and this calibration was propagated to the $\not\!\!E_T$.

Except for the background from multijet production, which was estimated from the data, all backgrounds from standard model (SM) processes were determined by Monte Carlo simulation. The (W/Z)+jets processes were generated with ALPGEN [10] interfaced with PYTHIA [11] for initial and final state radiation, and for jet hadronization. A matching algorithm [12] was applied to avoid double counting in phase space regions which can be populated both by ALPGEN or by PYTHIA. Light (u, d, s, g) and heavy (c, b) flavor production in association with W/Z were generated separately, and care was taken to avoid double counting between heavy flavor jets produced directly by ALPGEN or subsequently by PYTHIA. For $t\bar{t}$ and for electroweak single top production, the ALPGEN and COMPHEP [13] generators interfaced with PYTHIA were used, respectively, while the vector-boson pair production processes were generated with PYTHIA. The signal processes (HZ and HW) were generated with PYTHIA for Higgs boson masses ranging from 105 to 145 GeV, in 10 GeV steps. In all these simulations, the CTEQ6L1 parton distribution functions were used [14]. For the (W/Z)+jets processes, the absolute normalizations were determined from the data before b-tagging, as explained below. The heavy flavor fractions were obtained using MCFM [15]. The cross sections for the other background pro-

cesses were taken from Refs. [16], or calculated with MCFM, and the signal production cross sections were taken from Ref. [17].

Signal and background samples were passed through a full GEANT3-based simulation [18] of the detector geometry and response, and processed with the same reconstruction program as the data. Real events from randomly selected beam crossings were overlaid on simulated events to account for additional minimum bias interactions. The trigger conditions on jets and $\not\!\!E_T$ were not included in the simulation, but trigger efficiency parameterizations were applied, as determined from triggers based only on information from the muon detectors, therefore independent of those used in this analysis. Weight factors were further applied to compensate for residual differences between data and simulation for luminosity profile, primary vertex longitudinal distribution, electron, muon and jet identification. The jet energy calibration and resolution were adjusted in simulated events so as to match those measured in data.

III. EVENT SELECTION

The goal of the selection criteria applied in a first stage is to eliminate the bulk of the otherwise overwhelming background from multijet events, while retaining a high efficiency for the signal. In a second stage, further enhancement of the search sensitivity is achieved using heavy flavor tagging. The analysis was optimized for a Higgs boson mass of 115 GeV.

The selection follows closely the one described in Ref. [3]. The event primary vertex has to be reconstructed within the acceptance of the silicon vertex detector ($|z_{\rm PV}| < 60$ cm, where z is measured from the detector center along the beam direction), and at least three charged particle tracks have to originate from that vertex. There has to be exactly two or three jets with $p_T > 20$ GeV and within $|\eta_{\rm det}| < 2.5$, where $\eta_{\rm det}$ is the pseudorapidity measured from the detector center rather than from the primary interaction vertex. Furthermore, for Run IIa data, the two leading jets, i.e. those with the highest p_T, are required not to point toward the ICD region, defined as $1.1 < |\eta_{\rm det}| < 1.4$. The two leading jets must not be back-to-back in the plane transverse to the beam direction ($\Delta\phi({\rm jet}_1, {\rm jet}_2) < 165^\circ$), and there has to be charged particle tracks associated with those jets, meeting minimal quality criteria such that the b-tagging algorithm can operate efficiently. Finally, a large missing transverse energy is required ($\not\!\!E_T > 40$ GeV).

After this preselection, further criteria are used to define three distinct samples: an analysis sample in which the search for a Higgs boson signal is performed; a control sample, enriched in $(W \to \mu\nu)$+jets events where the jet system displays a topology similar to that of the analysis sample, and which is used to validate the SM background simulation; and a "QCD sample", highly enriched in multijet events, and which is used to model the multijet background in the analysis sample.

To select the analysis sample, the threshold for the missing transverse energy is increased ($\not\!\!E_T > 50$ GeV). This criterion is tightened in case the direction of the missing E_T is close to the direction of one of the jets in the plane transverse to the beam direction, a configuration expected from the multijet background: $\not\!\!E_T({\rm GeV}) > 80 - 40 \times \min\Delta\phi(\not\!\!E_T, \text{any jet})$, where the angle is measured in radians. As was shown in Ref. [2], the asymmetry $\mathcal{A} = (\not\!\!E_T - \not\!\!H_T)/(\not\!\!E_T + \not\!\!H_T)$ provides good discrimination between signal and multijet events. It is therefore required that $-0.1 < \mathcal{A} < 0.2$. In signal events, the missing track-p_T, $\not\!\!p_T$, defined as the opposite of the vectorial sum of the charged particle transverse momenta, is expected to point in a direction similar to that of the $\not\!\!E_T$, while this is not expected in multijet events where a jet energy is mismeasured. Advantage is taken of this feature by requiring $\Delta\phi(\not\!\!E_T, \not\!\!p_T) < \pi/2$.

To reject backgrounds from W+jets, top, and diboson production, events containing an isolated electron or muon are rejected. In order to keep this analysis orthogonal to the search in the $HW \to b\bar{b}\ell\nu$ channel, these electrons and muons are selected as in Ref. [19]. In particular, they have to fulfill $p_T > 15$ GeV, $|\eta| < 2$ for muons, and $|\eta_{\rm det}| < 1.1$ or $1.5 < |\eta_{\rm det}| < 2.5$ for electrons.

The control sample is selected in a similar way as the analysis sample, except that the muon veto is removed, and that the presence of a muon with $p_T > 20$ GeV within $|\eta| < 2$ is required instead. In order to ensure that the jet system has similar topologies in the analysis and control samples, the 50 GeV threshold is applied in the control sample to a missing E_T calculated without the correction for reconstructed muons. Since there is little multijet background in the control sample, the value of the cut on $\Delta\phi(\not\!\!E_T, \not\!\!p_T)$ is increased to $\pi - 1$ radian. The global normalization factor for the W+jets background is determined such that the numbers of predicted and observed events are equal in the control sample. The normalization for the Z+jets background is taken from Ref. [20]. These normalization factors of SM backgrounds are kept throughout the analysis. Examples of distributions in the control sample are given in Fig. 1, showing good agreement between data and prediction.

The variable $\Delta\phi(\not\!\!E_T, \not\!\!p_T)$ is used to define a sample dominated by the multijet background. It is selected in the same way as the analysis sample, except that the cut $\Delta\phi(\not\!\!E_T, \not\!\!p_T) < \pi/2$ is now inverted. The small contribution from SM processes in that $\Delta\phi(\not\!\!E_T, \not\!\!p_T) > \pi/2$ region is subtracted. The "QCD sample" thus obtained is used as a model of the multijet background in the analysis sample, i.e., in the $\Delta\phi(\not\!\!E_T, \not\!\!p_T) < \pi/2$ region, where it is normalized in such a way that, after addition of the SM backgrounds, the numbers of events expected and observed in the analysis sample are

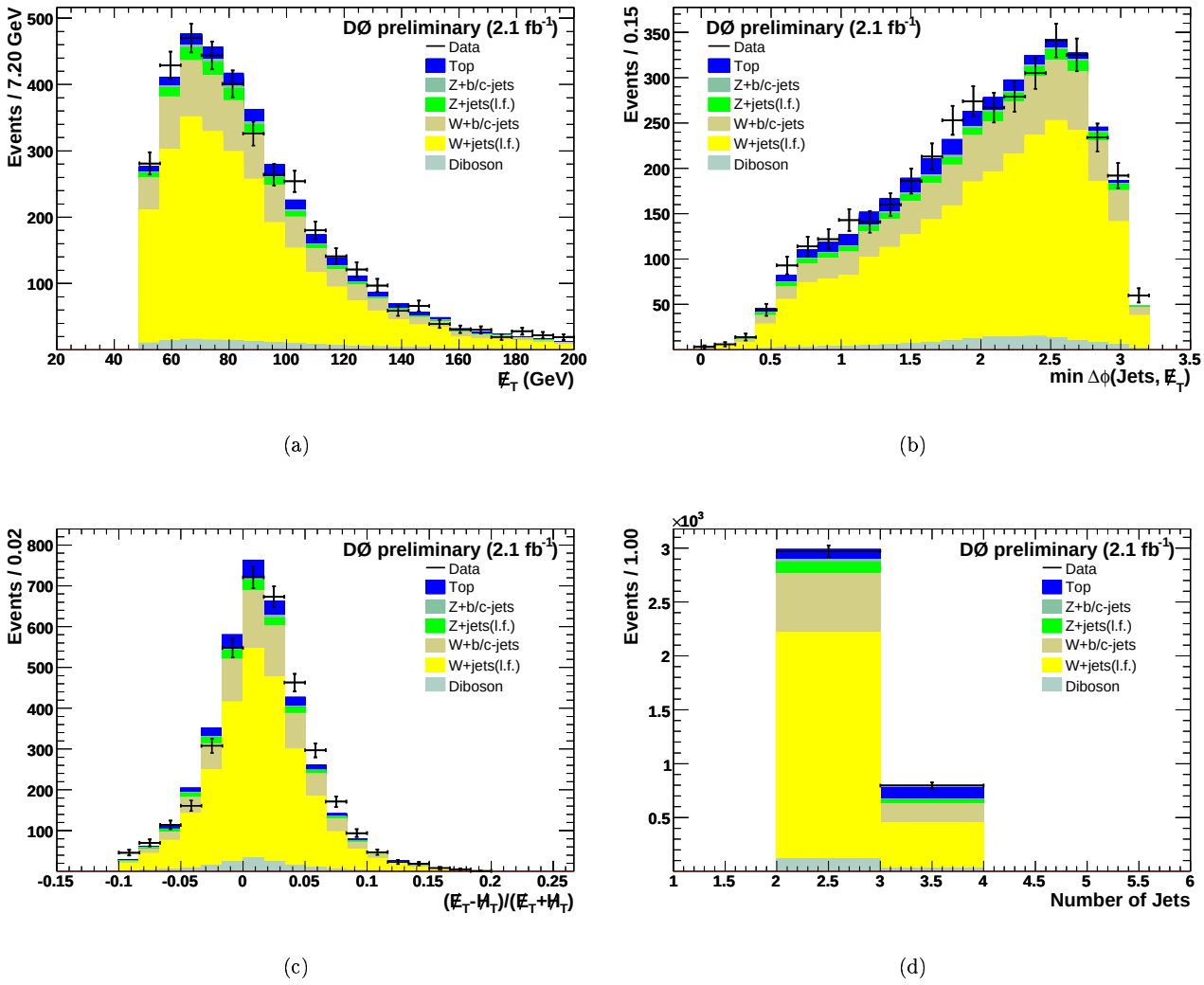


FIG. 1: Distributions in the $W \to \mu\nu$ control sample, before b tagging: a) $\not\!\!E_T$; b) min$\Delta\phi(\not\!\!E_T$, any jet); c) $\mathcal{A} = (\not\!\!E_T - \not\!\!H_T)/(\not\!\!E_T + \not\!\!H_T)$; d) jet multiplicity. The data are shown as points with error bars. The various SM background contributions are shown as histograms, with color codes as indicated in the frames.

equal. This normalization factor is found to be close to unity, as expected from Monte Carlo simulations of multijet events, which show that the distribution of $\Delta\phi(\not\!\!E_T, \not\!\!p_T)$ is almost symmetric around $\pi/2$. Examples of distributions in the analysis sample are shown in Fig. 2. It is seen that the combination of multijet and SM backgrounds provides a good description of the data.

Advantage is next taken of the large branching fraction for $H \to b\bar{b}$ by requiring that the two leading jets be b tagged. Asymmetric cuts on the outputs of the b tagging algorithm neural network [21] were applied to the two leading jets. They were chosen such that one of those jets is tagged with an efficiency of 73%, and the other with an efficiency of 48%. These values are indicative, and apply to jets with $p_T > 30$ GeV and $|\eta_{\rm det}| < 0.8$. The corresponding mistag rates, i.e., the probablilities to wrongly tag u, d, s, g jets, are 5% and 0.5%, respectively. This asymmetric tagging procedure was found to provide the best sensitivity to a Higgs boson signal. In order to cope with differences in track reconstruction efficiencies in data and in simulation, the b-tagging algorithm was applied directly only to the data, including the "QCD sample", while flavor-dependent tagging probabilities measured in dedicated data samples were applied to the simulated jets.

In the simulation, the (W/Z)+heavy-flavor jets backgrounds are normalized as follows. The leading order to next-to-leading order (NLO/LO) K-factors were determined with MCFM for (W/Z)+2-jets for both heavy and light flavors, and the ratio of those K-factors was applied to the (W/Z)+heavy-flavor jets events generated with ALPGEN, on top of the overall normalization factor determined from data before b tagging. As can be seen in Fig. 3 for the invariant

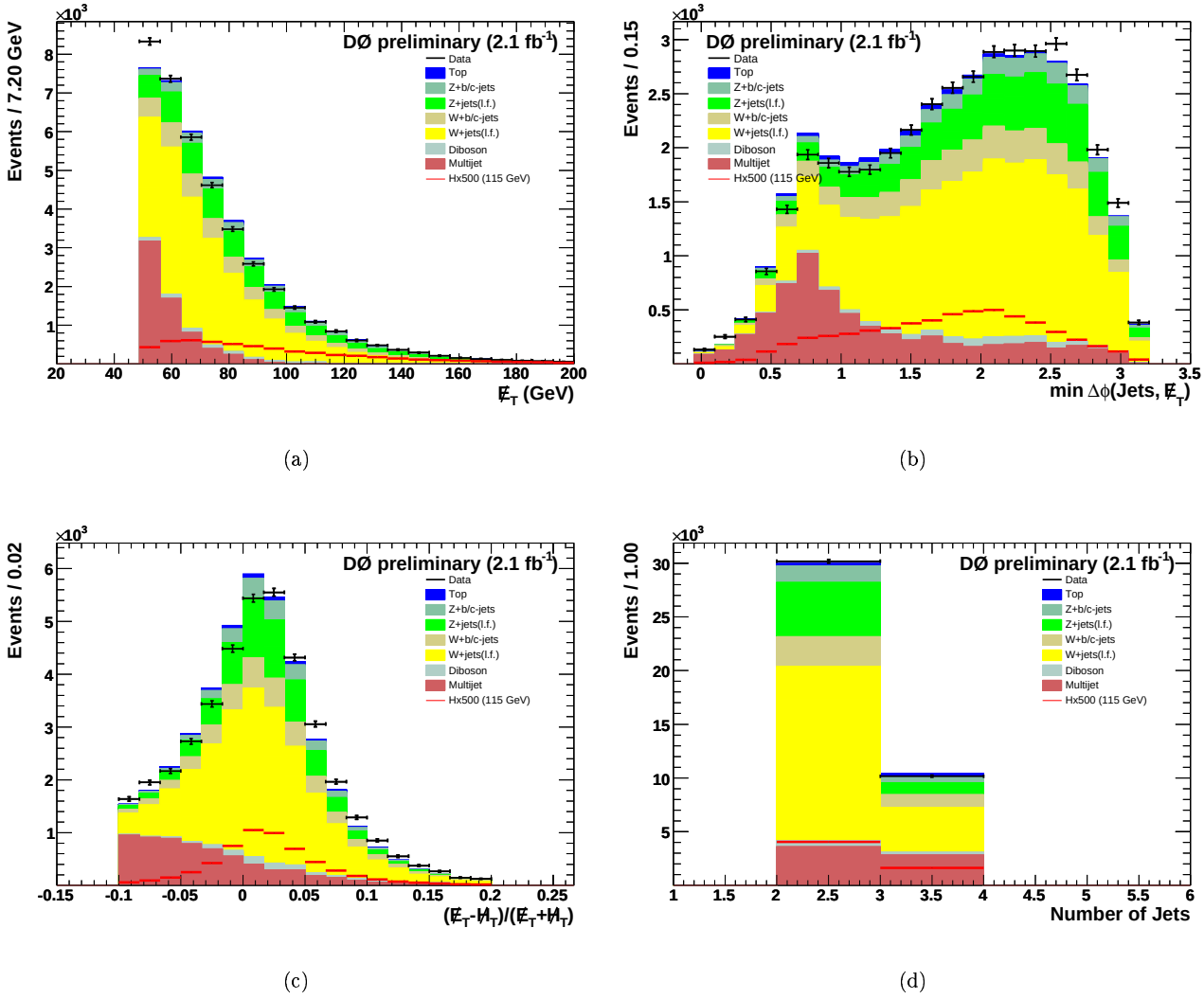
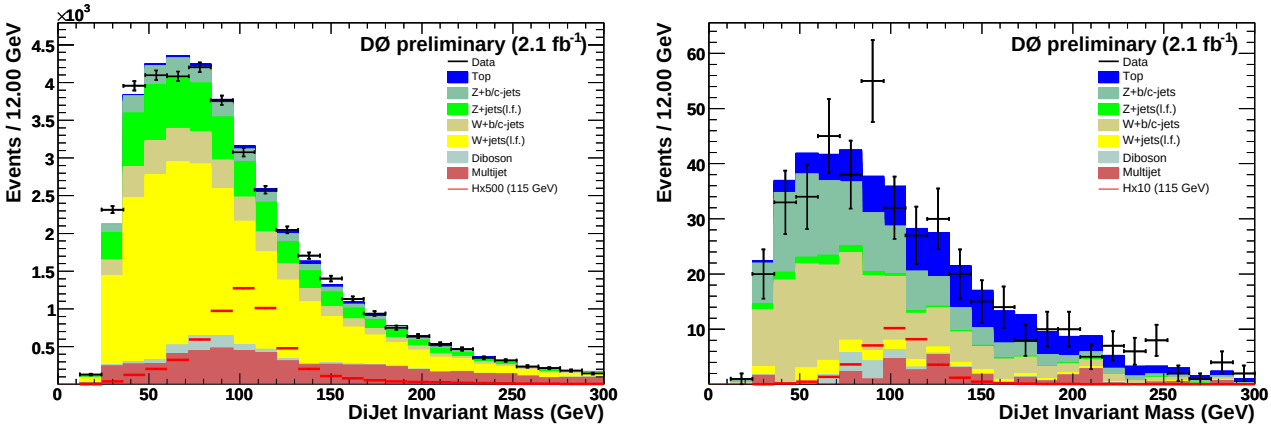


FIG. 2: Distributions in the analysis sample, before b tagging: a) $\not\!\!E_T$; b) min$\Delta\phi(\not\!\!E_T$, any jet); c) $\mathcal{A} = (\not\!\!E_T - \not\!\!H_T)/(\not\!\!E_T + \not\!\!H_T)$; d) jet multiplicity. The data are shown as points with error bars. The various background contributions (SM and multijet) are shown as histograms, with color codes as indicated in the frames. Distributions for a signal with a Higgs boson mass of 115 GeV are also shown, multiplied by 500.

TABLE I: Numbers of signal and background events expected, and numbers of events observed, before and after b tagging. The numbers for the signals are given for a Higgs boson mass of 115 GeV; "top" includes pair and single top production; VV stands for the sum of all diboson processes. The errors quoted are statistical only.

Sample	HZ	HW	W+jets	Z+jets	top	VV	multijet	Total	Observed
before	6.15 ± 0.03	5.20 ± 0.04	24 357	8 199	532	814	6 438	40 340	40 340
after	2.12 ± 0.01	1.58 ± 0.01	174.0	127.3	95.2	12.5	33.8	442.8 ± 1.1	439

mass of the dijet system built from the two leading jets, the simulation provides a good description of the data both before and after b tagging. Details of the signal and background expected numbers of events are given in Table I, as well as the numbers of selected events. After b tagging, the W+jets and Z+jets backgrounds are dominated by heavy flavor jet production, which contributes $\mathcal{O}(90\%)$.



(a) (b)

FIG. 3: Distributions in the analysis sample of the invariant mass of the two leading jets a) before and b) after b tagging. The data are shown as points with error bars. The various background contributions (SM and multijet) are shown as histograms, with color codes as indicated in the frames. Distributions for a signal with a Higgs boson mass of 115 GeV are also shown, multiplied by 500 before and by 10 after b tagging.

IV. DISCRIMINANT ANALYSIS

In order to take full advantage of the different kinematic characteristics of the signal and background processes, a boosted decision tree technique [22] was used. Boosting is a technique whereby a decision tree (DT) is retrained, with a larger weight given to events which were misclassified in the previous iteration. For each of the Higgs boson masses probed, the DT training was performed on one third of the signal and SM background Monte Carlo samples, with twenty boosts, while the other two thirds were used to derive the results. The multijet background was not included in the DT training. Because of the differences in experimental conditions between Run IIa and Run IIb, a dedicated DT was trained for each of the two data taking periods.

The following variables were fed into the DTs: $\not\!\!E_T$, total E_T, $\not\!\!H_T$, $H_T = \Sigma p_T(\text{jets})$, $\not\!\!H_T/H_T$, $(\not\!\!E_T - \not\!\!H_T)/(\not\!\!E_T + \not\!\!H_T)$, $\Delta\phi(\text{jet}_1,\text{jet}_2)$, $\Delta R(\text{jet}_1,\text{jet}_2)$ (with $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$), $\Delta\phi(\not\!\!E_T,\text{jet}_1)$, $\Delta\phi(\not\!\!E_T,\text{jet}_2)$, $\min\Delta\phi(\not\!\!E_T,\text{any jet})$, $\max\Delta\phi(\not\!\!E_T,\text{any jet}) - \min\Delta\phi(\not\!\!E_T,\text{any jet})$, $\max\Delta\phi(\not\!\!E_T,\text{any jet}) + \min\Delta\phi(\not\!\!E_T,\text{any jet})$, $\Delta\phi(\not\!\!E_T,\text{dijet})$, $\Sigma p_T(\text{all jets with } p_T > 20 \text{ GeV})$, $p_T^{\text{jet}_1}$, $p_T^{\text{jet}_2}$, η^{jet_1}, η^{jet_2}, dijet invariant mass, dijet transverse mass, $(\not\!\!E_T - \not\!p_T)/(\not\!\!E_T + \not\!p_T)$, $\Sigma p_T(\text{tracks})$, $\Sigma p_T(\text{tracks from dijet})$, $(\Sigma p_T(\text{tracks}) - \Sigma p_T(\text{tracks from dijet}))/\Sigma p_T(\text{tracks})$.

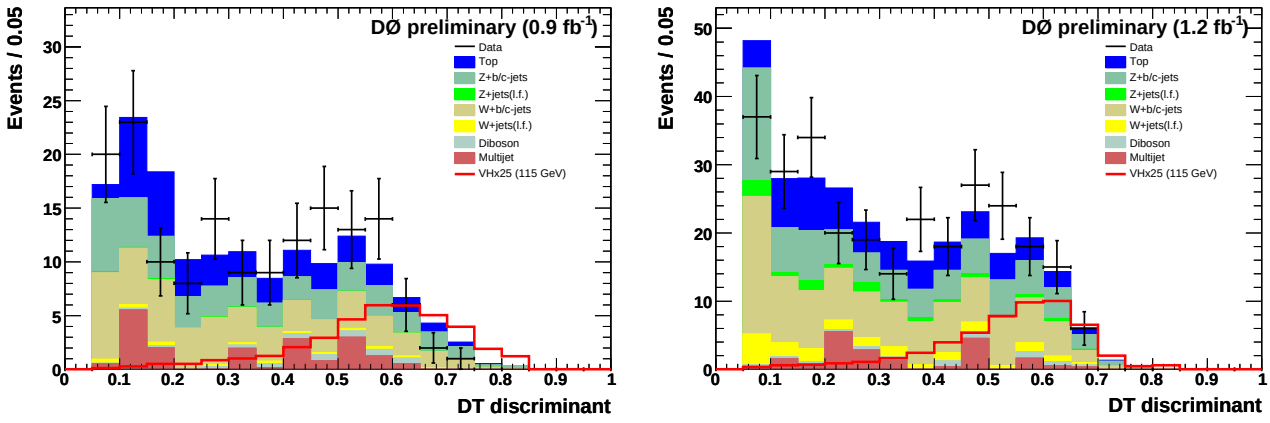
The DT discriminants are shown in Fig. 4 for a Higgs boson mass of 115 GeV.

V. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties originate from various sources. Experimental uncertainties arise from the trigger simulation, from the jet energy calibration, resolution, and reconstruction efficiency, and from b tagging. The corresponding parameterizations applied to the simulation were varied within their uncertainties, and the impact on both the normalizations and the shapes of the DT discriminants were assessed. The correlations between these systematic uncertainties in the signal and background processes were taken into account to derive the final results, as well as for the uncertainty on the integrated luminosity of the data sample.

The cross sections of the various SM processes suffer from theoretical uncertainties. For the W+jets process, the overall normalization was constrained by the data in the control sample before b tagging, and similarly for Z+jets, based on Ref. [20]. Uncertainties on the heavy flavor fractions can be partially obtained with MCFM. They were substantially inflated to cope with unaccounted effects, such as heavy quark masses at NLO. For the other background processes, theoretical uncertainties were taken from Refs. [16], or calculated with MCFM. The normalization of the multijet background is fully anticorrelated with the normalization of the SM backgrounds, as the sum is constrained by the data before b tagging. The uncertainties on the signal cross sections were taken from Ref. [17].

The systematic uncertainties affecting the signal and SM background normalizations are reported in Table II.



(a) (b)

FIG. 4: Distributions of the decision tree discriminants for a Higgs boson mass of 115 GeV (HZ and HW signals combined) for a) Run IIa and b) Run IIb. The data are shown as points with error bars. The various background contributions (SM and multijet) are shown as histograms, with color codes as indicated in the frames. The distributions for the signal are multiplied by a factor of 25.

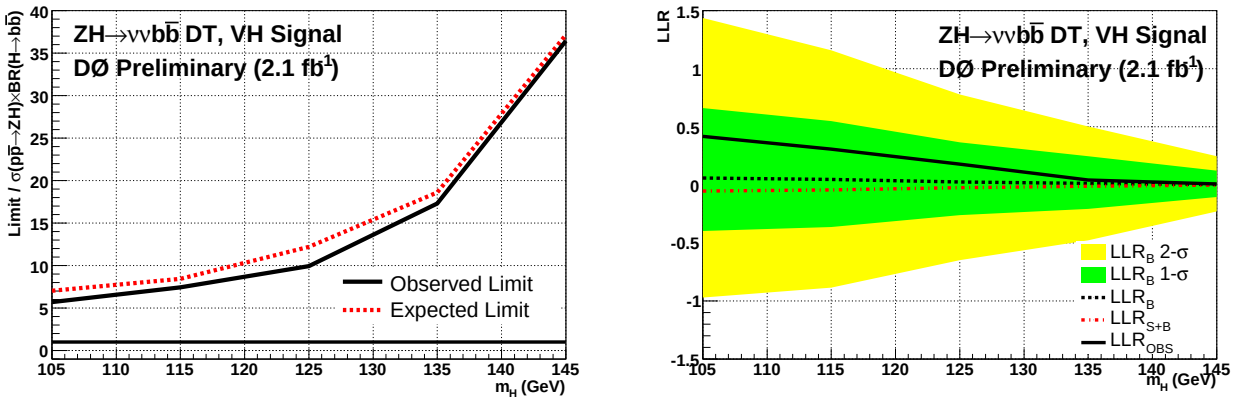
TABLE II: Systematic uncertainties on the numbers of predicted signal and SM background events, in percents: integrated luminosity (Lumi), trigger, jet energy calibration (J-C), resolution (J-R), and reconstruction efficiency (J-E), b tagging, cross sections (σ), heavy flavor fractions (HF). For the signal, a Higgs boson mass of 115 GeV is chosen.

	Lumi	Trigger	J-C	J-R	J-E	b tagging	σ	HF
Signal	6.1	5.5	2.2	1.1	2.2	5.8	6	
SM backgrounds	6.1	5.5	2.9	0.9	2.1	5.6	6 – 16	50

VI. RESULTS

Agreement between data and expectation from SM and multijet backgrounds is observed both in terms of numbers of events selected (Table I) and of DT discriminant shapes (Fig. 4). To set limits on the SM Higgs boson production cross section, a modified frequentist approach [23] was used, where the signal confidence level CL_s, defined as the ratio of the confidence level for the signal-plus-background hypothesis to the background-only hypothesis ($CL_s = CL_{s+b}/CL_b$), is calculated by integration of the distributions of a test statistic over the outcomes of pseudo-experiments, generated according to Poisson statistics, for the signal+background and background-only hypotheses. The test statistic is calculated as a joint log-likelihood ratio (LLR) obtained by summing LLR values over the bins of the DT discriminants. Systematic uncertainties were incorporated via Gaussian smearing of the Poisson probability distributions for signal and backgrounds within the pseudo-experiments. All correlations between signal and backgrounds were maintained. To reduce the impact of systematic uncertainties on the sensitivity of the analysis, the individual signal and background contributions were fitted to the data (and pseudo-data) for both the signal-plus-background and the background-only hypotheses independently by maximizing a profile likelihood function for each hypothesis [24]. The profile likelihood is constructed via a joint Poisson probability over the number of bins in the calculation and is a function of the nuisance parameters in the system and their uncertainties, which are given an additional Gaussian constraint associated with their prior predictions. The maximization of the likelihood function is performed over the nuisance parameters.

The results obtained are shown as a function of the Higgs boson mass in Fig. 5 and in Table III, in terms of ratios of excluded to SM production cross sections times branching fraction for $H \to b\bar{b}$. The LLRs are also shown in Fig. 5. The observed limit is in agreement with the expected limit, defined as the median of the limits obtained in background-only pseudo experiments. For a 115 GeV Higgs boson mass, the observed and expected limits on the cross section of combined HZ and HW production times branching fraction for $H \to b\bar{b}$ are 7.5 and 8.4 times larger than the SM value, respectively.



(a) (b)

FIG. 5: As a function of the Higgs boson mass, a) limit on the cross section of combined HZ and HW production times branching fraction for $H \to b\bar{b}$, relative to the standard model value, and b) log likelihood ratio. In a) the observed and expected limits are shown as solid and dashed curves, respectively. In b) the observed LLR is shown as a solid curve, the expected LLRs are shown as black and red dashed curves for the background-only and signal+background hypotheses, respectively, and the green and yellow areas correspond to the one and two σ deviations around the expected background-only LLR.

TABLE III: For various Higgs boson masses, observed and expected ratios of excluded to SM production cross sections times branching fraction for $H \to b\bar{b}$.

mass	105	115	125	135	145
observed	5.7	7.5	9.9	17.3	36.4
expected	7.0	8.4	12.2	18.6	37.1

VII. SUMMARY

A search for the standard model Higgs boson has been performed in 2.1 fb^{-1} of $p\bar{p}$ collisions at 1.96 TeV. The topology analysed consists of a pair of b jets with large $\not\!\!E_T$, as expected from $p\bar{p} \to HZ \to b\bar{b}\nu\bar{\nu}$. The search is also sensitive to HW production, where the W decays leptonically and the charged lepton is undetected. No deviation from the expectation from standard model backgrounds was observed. A boosted decision tree technique was used to derive an upper limit on the cross section of the $p\bar{p} \to HZ$ and $p\bar{p} \to HW$ processes combined, as a function of the Higgs boson mass. For a mass of 115 GeV, this limit is a factor 7.5 larger than the standard model cross section.

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); STFC (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); CAS and CNSF (China); and the Alexander von Humboldt Foundation.

[1] M. Carena *et al.*, *"Report of the Tevatron Higgs Working Group"*, arXiv:hep-ph/0010338; CDF and DØ Collaborations, *"Results of the Tevatron Higgs Sensitivity Study"*, FERMILAB-PUB-03/320-E.

[2] V. Abazov *et al.* (DØ Collaboration), Phys. Rev. Lett. **97**, 161803 (2006).

[3] The DØ Collaboration, *"A Search for the Standard Model Higgs Boson in the Channel $ZH \to \nu\bar{\nu}b\bar{b}$ at $\sqrt{s} = 1.96$ TeV"*, DØ Note 5506-CONF.

[4] R. Barate *et al.* [LEP Working Group for Higgs boson searches], Phys. Lett. B **565**, 61 (2003).

[5] The LEP Collaborations: ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, the LEP Electroweak Working Group, *"Precision Electroweak Measurements and Constraints on the Standard Model,"* CERN-PH-EP/2007-039, arXiv:0712.0929v2.

[6] V.M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. A **565**, 463 (2006).

[7] M. Abolins *et al.*, Nucl. Instrum. Methods in Phys. Res. A **584/1**, 75 (2007).

[8] T. Andeen *et al.*, FERMILAB-TM-2365 (2007).

[9] G.C. Blazey *et al.*, in *Proceedings of the Workhop: "QCD and Weak Boson Physics in Run II,"* edited by U. Baur, R.K. Ellis, and D. Zeppenfeld (Fermilab, Batavia, IL, 2000), p. 47; see Sec. 3.5 for details.

[10] M.L. Mangano *et al.*, JHEP **0307**, 001 (2003); version 2.05 (2.11) was used for Run IIa (Run IIb) simulation.

[11] T. Sjöstrand, S. Mrenna and P. Skands, JHEP **0605**, 026 (2006); version 6.323 (6.409) was used for Run IIa (Run IIb) simulation.

[12] S. Höche *et al.*, *"Matching Parton Showers and Matrix Elements"*, in Proceedings of the Workshop on the Implications of HERA for LHC Physics, ed. A. De Roeck and H. Jung (CERN, Geneva, 2005), p288.

[13] E. Boos *et al.* (CompHEP Collaboration), Nucl. Instrum. Methods in Phys. Res. A **534**, 250 (2004).

[14] J. Pumplin *et al.*, JHEP **0207**, 012 (2002); D. Stump *et al.*, JHEP **0310**, 046 (2003).

[15] J.M. Campbell and R.K. Ellis, Phys. Rev. D **60**, 113006 (1999).

[16] M. Cacciari *et al.*, JHEP **404**, 068 (2004); N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003); N. Kidonakis, Phys. Rev. D **74**, 114012 (2006).

[17] S. Catani *et al.*, JHEP **0307**, 028 (2003).

[18] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).

[19] The DØ Collaboration, *"Search for WH Production using a Neural Network Approach in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV with 1.7 fb^{-1} of Data"*, DØ Note 5472-CONF.

[20] The DØ Collaboration, *"Search for $ZH \to \ell^+\ell^- b\bar{b}$ in $p\bar{p}$ collisions at $\sqrt{s} = 1960$ GeV"*, DØ Note 5482-CONF.

[21] T. Scanlon, *"b-Tagging and the Search for Neutral Supersymmetric Higgs Bosons at DØ"*, FERMILAB-THESIS-2006-43.

[22] L. Breiman *et al.*, *"Classification and Regression Trees"*, Wadsworth (1984); Y. Freund and R.E. Schapire, *"Experiments with a new boosting algorithm"*, in Machine Learning: Proceedings of the Thirteenth International Conference, pp. 148-156 (1996); V.M. Abazov *et al.*, Phys. Rev. Lett. **98**, 181802 (2007).

[23] T. Junk, Nucl. Instrum. Methods in Phys. Res. A **434**, 435 (1999); A. Read, in *"1st Workshop on Confidence Limits,"* CERN Report No. CERN-2000-005, 2000.

[24] W. Fisher, FERMILAB-TM-2386-E.