DØ Note 6223-CONF

Search for the Standard-Model Higgs Boson in the $ZH \to \nu\bar{\nu}b\bar{b}$ Channel in 8.4 fb^{-1} of $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration
URL: http://www-d0.fnal.gov

(Dated: August 12, 2011)

A search is performed for the standard-model Higgs boson in 8.4 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV, collected with the D0 detector at the Fermilab Tevatron Collider. The final state considered contains a pair of b jets and is characterized by an imbalance in transverse energy, as expected from $p\bar{p} \to ZH \to \nu\bar{\nu}b\bar{b}$ production. The search is also sensitive to the $WH \to \ell\nu b\bar{b}$ channel when the charged lepton is not identified. For a Higgs-boson mass of 115 GeV, a limit is set at the 95% C.L. on the cross section $\sigma(p\bar{p} \to [Z/W]H)$, assuming standard-model branching fractions, that is a factor of 3.2 larger than the theoretical standard-model value, consistent with the expected factor of 4.0. The search was also reinterpreted as a search for WZ and ZZ production, which resulted in the measurement of a cross section scale factor of $1.5 \pm 0.3(stat) \pm 0.4(syst)$ with an observed significance of 2.8 σ, which is consistent with the expected significance of 1.9 σ. This translates into a measured cross section of $6.9 \pm 1.3(stat) \pm 1.8(syst)$ pb, which is consistent with the predicted SM cross section of 4.6 pb.

Preliminary Results for 2011 Summer Conferences

I. INTRODUCTION

The existence of the Higgs boson is the only fundamental element of the standard model (SM) that has yet to be confirmed. Its observation would be a key step in establishing the mechanism of electroweak symmetry breaking and mass generation. Associated ZH production in $p\bar{p}$ collisions, with $Z \to \nu\bar{\nu}$ and $H \to b\bar{b}$, is among the most sensitive processes for seeking a Higgs boson with a mass $m_H \lesssim 135$ GeV at the Fermilab Tevatron Collider [1]. The D0 Collaboration published a search for this process based on 5.2 fb^{-1} of integrated luminosity [2]. A lower limit of 114.4 GeV was set by the LEP experiments on the mass of the Higgs boson from searches for the reaction $e^+e^- \to ZH$ [5], while an indirect upper limit of 158 GeV can be inferred from precision electroweak data [6]. These limits and those given below are all defined at the 95% confidence level (C.L.).

The final-state topology consists of a pair of b jets from $H \to b\bar{b}$ decay and missing transverse energy ($\not\!\!E_T$) from $Z \to \nu\bar{\nu}$. The search is therefore also sensitive to the WH process when the charged lepton from $W \to \ell\nu$ decay is not identified. The main backgrounds arise from (W/Z)+heavy-flavor jets (jets initiated by b or c quarks), top-quark production, and multijet (MJ) events with $\not\!\!E_T$ arising from mismeasurement of jet energies.

Compared to the preliminary result released for the Moriond 2011 conference [3] there are three main changes. Firstly, the signal cross sections and branching ratios have been updated to more recent predictions [4], this resulted in an approximate 7% loss in the signal yield. Secondly, to improve the modeling of the backgrounds and the senstivity of the analysis, the analysis was split into two and three jet exclusive channels (previously a combined two and three jet channel was used). The analysis of the three-jet channel did not converge in time for the conference and so only the 2-jet channel is included here. This resulted in an additional approximate loss of senstivity of 6%. These two effects were countered by an increase in integrated luminosity from 6.2 fb^{-1}to 8.4 fb^{-1}, which improves the senstivity by approximately 16%. The overall senstivity therefore will remain similar to the Moriond 2011 preliminary result.

To validate the techniques used in this Higgs search and the combination of similar channels across the Tevatron, this analysis has also been reinterpreted as a measurement of diboson production. For the diboson search the analysis is kept as close as possible to the Higgs search, with only the training of the multivariant event selection changed to reflect the altered signal.

II. DATA AND SIMULATED SAMPLES

The D0 detector is described in Ref. [7]. The data used in this analysis were recorded using triggers designed to select events with jets and $\not\!\!E_T$ [8]. After imposing data quality requirements, the total integrated luminosity [9] is 8.4 fb^{-1}.

The Tevatron Run II data taking is split into two periods, one prior to March 2006 which is referred to as Run IIa, while the period after is referred to as Run IIb. This division corresponds to the installation of an additional layer of silicon vertex detector, trigger upgrades, and a significant increase in the rate of delivered luminosity. The Run IIb period is further split into two to reflect another significant increase in the rate of delivered luminosity, referred to as Run IIb1 and Run IIb2. The data is modeled in each of these periods using dedicated Monte Carlo samples, designed to replicate both the instantaneous luminosity profile and detector performance in the respective periods.

The analysis relies on (i) charged particle tracks, (ii) calorimeter jets reconstructed in a cone of radius 0.5, using the iterative midpoint cone algorithm [10], and (iii) electrons or muons identified through the association of tracks with electromagnetic calorimeter clusters or with hits in the muon detector, respectively. The $\not\!\!E_T$ is reconstructed as the opposite of the vectorial sum of transverse components of energy deposits in the calorimeter and is corrected for identified muons. Jet energies are calibrated using transverse energy balance in photon+jet events [11], and these corrections are propagated to the $\not\!\!E_T$.

Backgrounds from SM processes are determined through Monte Carlo simulation, while the instrumental MJ background is estimated from data. Events from (W/Z)+jets processes are generated with ALPGEN [12], interfaced with PYTHIA [13] for initial and final-state radiation and for hadronization. The p_T spectrum of the Z is reweighted to match the D0 measurement [14]. The p_T spectrum of the W is reweighted using the same experimental input, corrected for the differences between the Z and W p_T spectra predicted in next-to-next-to-leading order (NNLO) QCD [15]. For $t\bar{t}$ and electroweak single top quark production, the ALPGEN and COMPHEP [16] generators, respectively, are interfaced with PYTHIA, while vector boson pair production is generated with PYTHIA. The ZH and WH signal processes are generated with PYTHIA for Higgs-boson masses (m_H) from 100 to 150 GeV, in 5 GeV steps. All these simulations use CTEQ6L1 parton distribution functions (PDFs) [17].

The absolute normalizations for (W/Z)+jets production are obtained from NNLO calculations of total cross sections based on Ref. [18], using the MRST2004 NNLO PDFs [19]. The heavy-flavor fractions are obtained using MCFM [20] at next-to-leading order (NLO). Cross sections for other SM backgrounds are taken from Ref. [21], or calculated with MCFM, and the cross sections for signal are taken from Ref. [4].

Signal and background samples are passed through a full GEANT3-based simulation [22] of the detector response and processed with the same reconstruction program as used for data. Events from randomly selected beam crossings are overlaid on simulated events to account for detector noise and contributions from additional $p\bar{p}$ interactions. Parameterizations of the trigger efficiencies are determined using events collected with independent triggers based on information from the muon detectors. Weight factors compensating for residual differences between data and simulation are applied for electron, muon and jet identification. Jet energy calibration and resolution are adjusted in simulated events to match those measured in data.

III. EVENT SELECTION

A preselection that greatly reduces the overwhelming background from multijet events is performed as follows. The interaction vertex must be reconstructed within the acceptance of the silicon vertex detector, and at least three tracks must originate from that vertex. Jets with associated tracks (using only tracks that meet minimal quality criteria to ensure that the b-tagging algorithm operates efficiently) are denoted as "taggable" jets. The leading (highest p_T) jet must be taggable and there must be exactly two taggable jets; the Higgs candidate is formed from these jets. These jets must have transverse momentum $p_T > 20$ GeV and pseudorapidity $|\eta| < 2.5$ [23]. The two taggable jets must not be back-to-back in the plane transverse to the beam direction: $\Delta\phi(\mathrm{jet}_1, \mathrm{jet}_2) < 165^\circ$. Finally, $\not\!\!E_T > 30$ GeV is required.

The RunIIa data taking period had looser triggers which resulted in a larger multijet background component in the analysis. To reduce the multijet background in the RunIIa component of the analysis, two additional cuts are introduced which replicate the additional trigger terms introduced in RunIIb. These are $\Delta\phi(Jets, \not\!\!E_T) > 23^\circ$ and $\not\!\!E_{T\,Trig} > 30$ GeV, where $\not\!\!E_{T\,Trig}$ is the $\not\!\!E_T$ as calculated in the trigger without the energy in the outermost hadronic calorimeter taken into account.

Additional selection criteria define four distinct samples: (i) an "analysis" sample used to search for a Higgs-boson signal, (ii) an "electroweak (EW) control" sample, enriched in $W(\rightarrow \mu\nu)$+jets events where the jet system has a topology similar to that of the analysis sample, which is used to validate the SM background simulation, (iii) a "MJ-model" sample, dominated by multijet events, used to model the MJ background in the analysis sample, and (iv) a large "MJ-enriched" sample, used to validate this modeling procedure.

The analysis sample is selected by requiring the scalar sum of the two leading taggable jets $p_T > 80$ GeV, $\not\!\!E_T > 40$ GeV and a measure of the $\not\!\!E_T$ significance $\mathcal{S} > 5$ [24]. Larger values of $\mathcal{S}$ correspond to $\not\!\!E_T$ values that are less likely to be caused by fluctuations in jet energies. The $\mathcal{S}$ distribution is shown for the analysis and EW-control samples in Fig. 1;

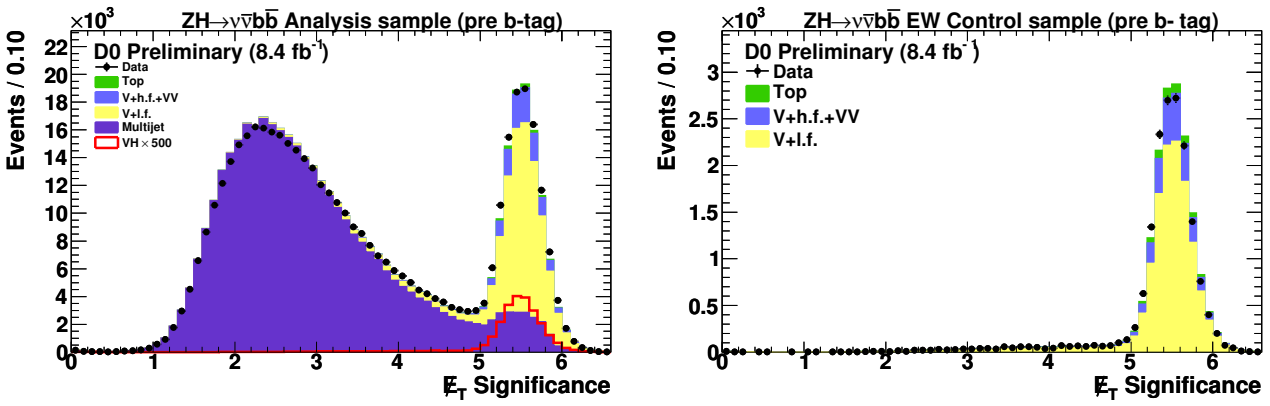
In signal events $\Delta\phi(\not\!p_T, \mathrm{Jet_L})$ and $\Delta\phi(\not\!p_T, \mathrm{Jet_{NL}})$ are quite correlated (where $\not\!p_T$ is the missing p_T calculated from the reconstructed charged particle tracks, L refers to the leading taggable jet and NL refers to the next-to-leading taggable jet when sorted in p_T order), with the result that the vast majority of signal events have $D > \pi/2$, where $\mathcal{D} = (\Delta\phi(\not\!p_T, \mathrm{Jet_L}) + \Delta\phi(\not\!p_T, \mathrm{Jet_N L}))/2$, whereas background events tend to be symmetrically distributed around $\pi/2$. Advantage is taken of this feature by requiring $\mathcal{D} > \pi/2$. The reverse of the $\mathcal{D}$ requirement is also used to define the MJ-model sample (described in more detail below). As $\mathcal{D}$ is calculated taking into account only information from the reconstructed charged particle tracks and the direction of the calorimeter jets, it reduces any bias between the MJ-model and analysis sample due to fluctuations in the jet energy measurement.

Events containing an isolated electron or muon [25] with $p_T > 15$ GeV are rejected to reduce backgrounds from W+jets, top quark, and diboson production.

The EW-control sample is selected in a similar manner to the analysis sample, except that an isolated muon with $p_T > 15$ GeV is required. The multijet content of this sample is rendered negligible by requiring that the transverse mass of the muon and $\not\!\!E_T$ system is larger than 30 GeV, and that the $\not\!\!E_T$, calculated taking account of the μ from the W decay, is greater than 20 GeV. To ensure similar jet topologies for the analysis and EW-control samples, $\not\!\!E_T$, not corrected for the selected muon, is required to exceed 40 GeV. The number of selected events is in good agreement with the SM expectation. All the kinematic distributions are also well described once a reweighting of the distribution of $\Delta\eta$ between the two taggable jets is performed, as suggested by a simulation of (W/Z)+jets using the SHERPA generator [26]. Four representative distributions in the EW-control sample are shown in Fig. 2.

The MJ-model sample, used to determine the MJ background, is selected in the same manner as the analysis sample, except that the requirement that $\mathcal{D} > \pi/2$ is inverted. The small contribution from non-MJ SM processes in the $\mathcal{D} < \pi/2$ region is subtracted, and the resulting sample is used to model the MJ background in the analysis sample. The MJ background in the region $\mathcal{D} > \pi/2$ is normalized by performing a fit of the MJ and SM backgrounds to the data in the analysis sample.

The MJ-enriched sample is used to test the validity of this approach and is defined in the same manner as the analysis sample, except that the $\not\!\!E_T$ threshold is reduced to 30 GeV and no requirement is imposed on $\mathcal{S}$. As a result,



(a) (b)

FIG. 1: Missing E_T significance in (a) the analysis and (b) the EW-control samples without the requirement that it be larger than 5. The distributions for signal (VH) are multiplied by a factor of 500 and include ZH and WH production for $m_H = 115$ GeV. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production.

the MJ background dominates the entire range of $\mathcal{D}$ values, and this sample is used to verify that the events with $\mathcal{D} < \pi/2$ correctly model those with $\mathcal{D} > \pi/2$. Representative distributions in the MJ-enriched sample are shown in Fig. 3.

A boosted decision tree algorithm [28] designed to discriminate b from light (u,d,s,g) jets is used to select events with one or more b quark candidates. The algorithm is an upgraded version of the neural network b-tagging algorithm described in [27]. The new algorithm includes more information relating to the lifetime of the jet and results in a better discrimination between b and light jets. The algorithm provides an output (L_b) between 0 and 1 for all jets, with a value closer to one indicating a higher probability that the jet originated from a b quark.

To improve the sensitivity of the analysis, the information provided by the algorithm is exploited in a more sophisticated manner than placing a simple cut on L_b. The analysis sample is divided into two channels, where exactly one (single tag) or two (double tag) of the leading taggable jets satisfy a very loose cut on L_b (referred to as an operating point) that accepts $\approx$ 80% of b jets and $\approx$ 10% of light jets for $p_T \approx 45$ GeV and $|\eta| \approx 0.8$. For the surviving jets, L_b is used as an additional input to the decision tree used to separate signal events from background events (described in Sec. IV). The output from the algorithm measured on simulated events is adjusted to match the output measured on dedicated data samples as described in more detail in Ref. [27].

IV. ANALYSIS USING DECISION TREES

A boosted decision tree (DT) technique is employed to take advantage of differences in signal and background processes to improve their separation. First, a "MJ DT" (multijet-rejection DT) is trained to discriminate between signal and MJ-model events before any b tagging is applied, for each m_H, using thirty kinematic variables. All variables which provide some discrimination have been chosen for the MJ DT, but those directly related to the Higgs mass are avoided. The full list of input variables to the MJ DT is given in Table I.

The MJ DT output, which ranges between -1 and $+1$, is shown for the analysis and EW control samples for $m_H = 115$ GeV in Fig. 4. Good agreement is found between data and the predicted background, with any residual difference covered by the systematic errors (Sec. V). A value of the multijet discriminant in excess of 0 is required (multijet veto), which removes 80% of the multijet background and 23% of the non-MJ SM backgrounds, while retaining 88% of the signal. The number of expected signal and background events, as well as the number of observed events, are given in Table II, after imposing the multijet veto. Distributions in the analysis sample after the multijet veto are shown in Fig. 5 before any b-tagging requirement and in Fig. 6 for b-tagged events. Good agreement is found between data and the predicted background for all variables with any residual difference covered by the systematic errors.

Next, to discriminate signal from the other SM backgrounds, two "SM DTs" (SM-rejection DTs) are trained for

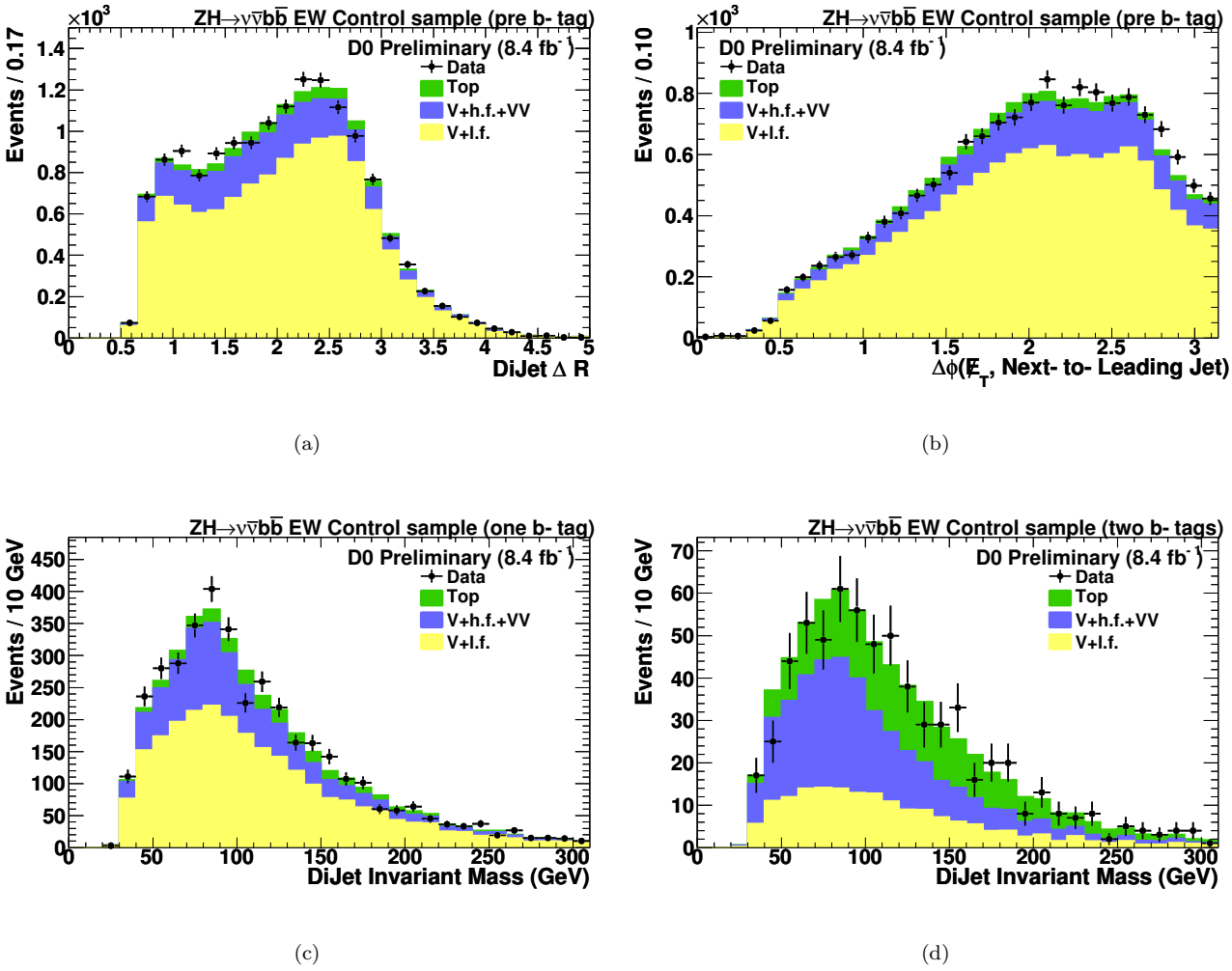


FIG. 2: Representative variable distributions in the EW-control sample: (a) dijet ΔR in the pre-tag sample, (b) $\Delta\phi(\not\!\!E_T, \text{Next-to-Leading Jet})$ in the pre-tag sample, (c) dijet invariant mass in the single-tag sample, (d) dijet invariant mass in the double-tag sample. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production.

each m_H, one in the single tag channel and one in the double tag channel. The same variables are used as for the MJ DT, with additional kinematic variables related to the Higgs mass and the b-tagging output of the b-tagged jets, the full list of variables are again shown in Table III. The SM DT outputs, which range between -1 and $+1$, are used as final discriminants. Their distributions are shown in Fig. 7 for $m_H = 115$ GeV.

V. SYSTEMATIC UNCERTAINTIES

Experimental uncertainties arise from trigger simulation (2%), jet energy calibration and resolution (1-3%), jet reconstruction and taggability (2%), lepton identification (1%), the modeling of the MJ background (25%, which translates into a 1% uncertainty on the total background) and b tagging (from 1% for background in the single-tag sample to 6% for signal in the double-tag sample). Their impact is assessed on overall normalizations, as shown in Table III, and on the shapes of distributions in the final discriminants. Correlations among systematic uncertainties in signal and background are taken into account in extracting the final results, including a 6.1% uncertainty on the integrated luminosity.

Theoretical uncertainties on cross sections for SM processes are estimated as follows. For (W/Z)+jets production,

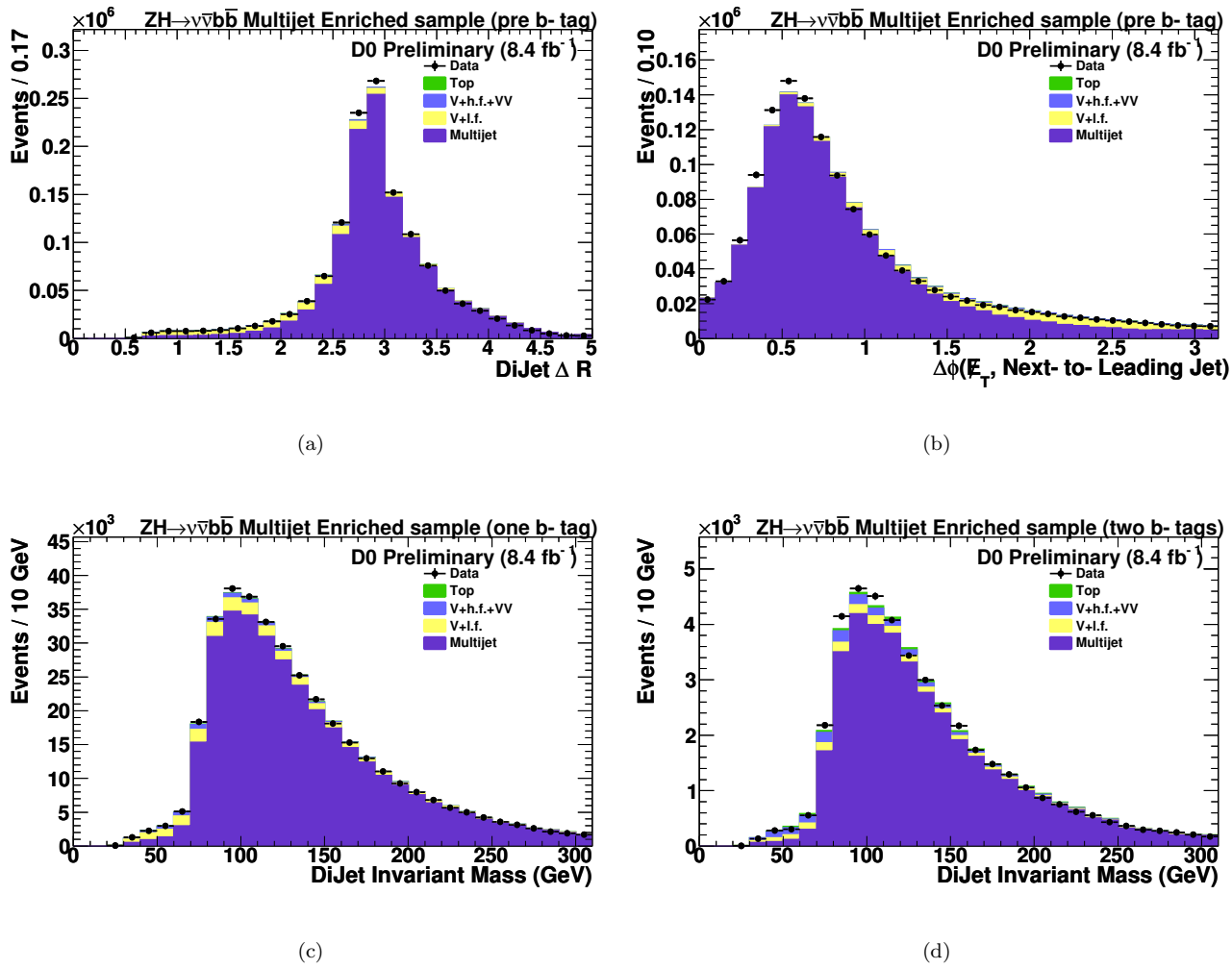
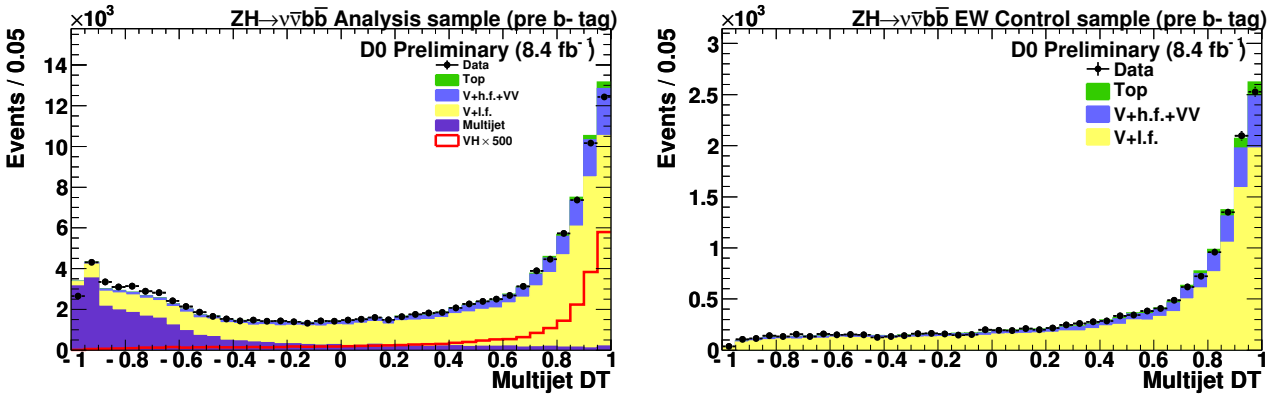


FIG. 3: Representative variable distributions in the MJ-enriched sample: (a) dijet ΔR in the pre-tag sample, (b) $\Delta\phi(\not\!\!E_T, \text{Next-to-Leading Jet})$ in the pre-tag sample, (c) dijet invariant mass in the single-tag sample, (d) dijet invariant mass in the double-tag sample. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes (W/Z)+(u, d, s, g) jets, "V+h.f." includes (W/Z)+(b, c) jets and "Top" includes pair and single top quark production.

an uncertainty of 10% is assigned to the total cross sections, and an uncertainty of 20% on the heavy-flavor fractions (estimated from MCFM at NLO [20]). For other SM backgrounds, uncertainties are taken from Ref. [21] or from MCFM, and range from 6% to 10%. The uncertainties on cross sections for signal (6% for $m_H = 115$ GeV) are taken from Ref. [4]. Uncertainties on the shapes of the final discriminants arise from (i) the modeling of (W/Z)+jets, assessed by varying the renormalization-and-factorization scale and by comparing results from ALPGEN interfaced with HERWIG [33] to ALPGEN interfaced with PYTHIA, and (ii) the choice of PDFs, estimated using the prescription of Ref. [17].

VI. LIMIT SETTING PROCEDURE

Agreement is found between data and the predicted background, both in the number of selected events (Table II) and in the distribution of final discriminants (Fig. 7), once systematic uncertainties are taken into account (Table III). A modified frequentist approach [31] is used to set limits on the cross section for SM Higgs-boson production, where the test statistic is a log-likelihood ratio (LLR) for the background-only and signal+background hypotheses. The result is obtained by summing LLR values over the bins in the final discriminants shown in Fig. 7. The impact of



(a) (b)

FIG. 4: MJ DT output for $m_H = 115$ GeV before any b-tagging requirement in the (a) signal sample and (b) EW-control sample. The distribution for signal (VH) is multiplied by a factor of 500 and includes ZH and WH production. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production.

systematic uncertainties on the sensitivity of the analysis is reduced by maximizing a "profile" likelihood function [32] in which these uncertainties are given Gaussian constraints associated with their priors. Figure 8 shows a comparison of the SM DT distributions after profiling along with the background-subtracted data and expected signal for the $m_H = 115$ GeV hypothesis. In this plot the background prediction and its uncertainties have been determined from the fit to data under the background-only hypothesis.

VII. HIGGS SEARCH RESULTS

The results of the updated analysis using 8.4 fb^{-1} of data are given in terms of LLR values in Fig. 9(a) and as limits in Table IV and Fig. 9(b). For $m_H = 115$ GeV, the observed and expected limits on the combined cross section of ZH and WH production are factors of 3.2 and 4.0 larger than the SM value, respectively.

VIII. DIBOSON SEARCH RESULTS

This analysis is also sensitive to various decay channels of ZZ and WZ production. Due to the similarity of the decay modes of the Higgs and diboson searches, the latter can be used to validate the techniques used in the former. The only modification to the Higgs search is the re-training of the three decision trees using the ZZ and WZ samples as the signal with the remaining diboson process, WW, kept as background, assuming that there is no VH signal present. The single and double tag SM DTs are shown in Fig. 10.

The LLR distribution for the full Run II single and double tag combined result can be found in Fig. 11. A cross section scale factor of $1.5 \pm 0.3(stat) \pm 0.4(syst)$ is measured with respect to the predicted standard-model value of 4.6 pb, with an observed significance of 2.8 σ (1.9 σ expected). This translates into a measured cross section of $6.9 \pm 1.3(stat) \pm 1.8(syst)$ pb, which is consistent with the predicted SM cross section of 4.6 pb. Figure 12 shows a comparison of the SM DT distributions, along with the background-subtracted data, after the background prediction and its uncertainties have been determined from the fit to data under the signal+background hypothesis.

The measurement of the diboson cross section has also been carried out using the dijet invariant mass as the final discriminant (as opposed to the SM DT). The LLR distribution for the full Run II single and double tag combined result can be found in Fig. 13. A cross section scale factor of $1.5 \pm 0.4(stat) \pm 0.6(syst)$ is measured with respect to the predicted standard-model value of 4.6 pb, with an observed significance of 2.2 σ (1.4 σ expected). This translates into a measured cross section of $6.9 \pm 1.6(stat) \pm 2.6(syst)$ pb, which is consistent with the predicted SM cross section of 4.6 pb and the cross section measured using the SM DTs as the final discriminant. Figure 14 shows a comparison of the dijet invariant mass distributions, along with the background-subtracted data, after the background prediction and its uncertainties have been determined from the fit to data under the signal+background hypothesis.

TABLE I: Variables used as input to the decision trees, where the angles θ and ϕ are the polar and azimuthal angles defined with respect to the proton beam direction. j_1 refers to the leading taggable jet, j_2 refers to the next-to-leading taggable jet, $j_{\rm all}$ refers to all jets in the event with $p_T > 15$ GeV and pseudorapidity $|\eta| < 3.2$. The thurst axis is the direction obtained from the difference of the transverse momenta of the leading and next-to-leading jets. The recoil is defined in the plane transverse to the beam using i) either the amount of missing transverse energy that remains after removal of the two leading jets, ii) or the sum of all good jets in the half plane on the opposite side of the thurst axis to the dijet system. Among these two possible recoil definitions, the one chosen is that has the larger component along the normal axis. The color flow variables are described in detail in [29] and further detail on the Higgs decay angle variables can be found in [30].

Variables used in the MJ DT and in the SM DT
$\Delta\eta(j_1, j_2)$
$\Delta\phi(j_1, j_2)$
$\Delta R((j_1, j_2))$
η of j_1
η of j_2
p_T weighted $\Delta R(j_1, j_{\rm all})$
p_T weighted $\Delta R(j_2, j_{\rm all})$
$\not\!\!E_T$
$\not\!\!E_T$ significance
$\Delta\phi(\not\!\!E_T, j_1)$
$\Delta\phi(\not\!\!E_T, j_2)$
$\Delta\phi(\not\!\!E_T, \text{dijet})$
$\min\Delta\phi(\not\!\!E_T, j_{\rm all})$
$\max\Delta\phi(\not\!\!E_T, j_{\rm all}) + \min\Delta\phi(\not\!\!E_T, j_{\rm all})$
$\max\Delta\phi(\not\!\!E_T, j_{\rm all}) - \min\Delta\phi(\not\!\!E_T, j_{\rm all})$
$\not\!\!H_T$(vectorial sum of $j_{\rm all}$ p_T)
$\not\!\!H_T$/ H_T (with H_T the scalar sum of $j_{\rm all}$ p_T)
Asymmetry between $\not\!\!E_T$ and $\not\!\!H_T$
$\not\!\!E_T$ component along the thrust axis
$\not\!\!E_T$ component perpendicular to the thrust axis
Sum of the signed components of the dijet and recoil momenta along the thrust axis
Sum of the signed components of the dijet and recoil momenta perpendicular to the thrust axis
Dijet p_T
Scalar sum of j_1 and j_2 p_T
Centrality (ratio of the scalar sum of j_1 and j_2 p_T to the sum of j_1 and j_2 energy)
Effective mass (sum of $\not\!\!E_T$ and of the scalar sum of j_1 and j_2 p_T)
θ angle of j_1 boosted to the dijet rest frame
θ angle of the dijet system
Polar angle of j_1 boosted to the dijet rest frame with respect to the dijet direction in the laboratory
Azimuthal angle of j_1 boosted to the dijet rest frame with respect to the dijet direction in the laboratory
Color flow j_1
Color flow j_2
Variables used only in the SM DT
Dijet mass
Dijet transverse mass
j_1 p_T
j_2 p_T
H_T
j_1 b-tagging output
j_2 b-tagging output

IX. SUMMARY

A search is performed for the standard-model Higgs boson in 8.4 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, collected with the D0 detector at the Fermilab Tevatron Collider. The final state considered contains a pair of b jets and is characterized by an imbalance in transverse energy, as expected from $p\bar{p} \to ZH \to \nu\bar{\nu}b\bar{b}$ production. The search is also sensitive to the $WH \to \ell\nu b\bar{b}$ channel when the charged lepton is not identified. For a Higgs-boson mass of 115 GeV, a limit is set at the 95% C.L. on the cross section $\sigma(p\bar{p} \to [Z/W]H)$, assuming standard-model branching fractions, that is a factor of 3.2 larger than the theoretical standard-model value, consistent with the expected factor of 4.0.

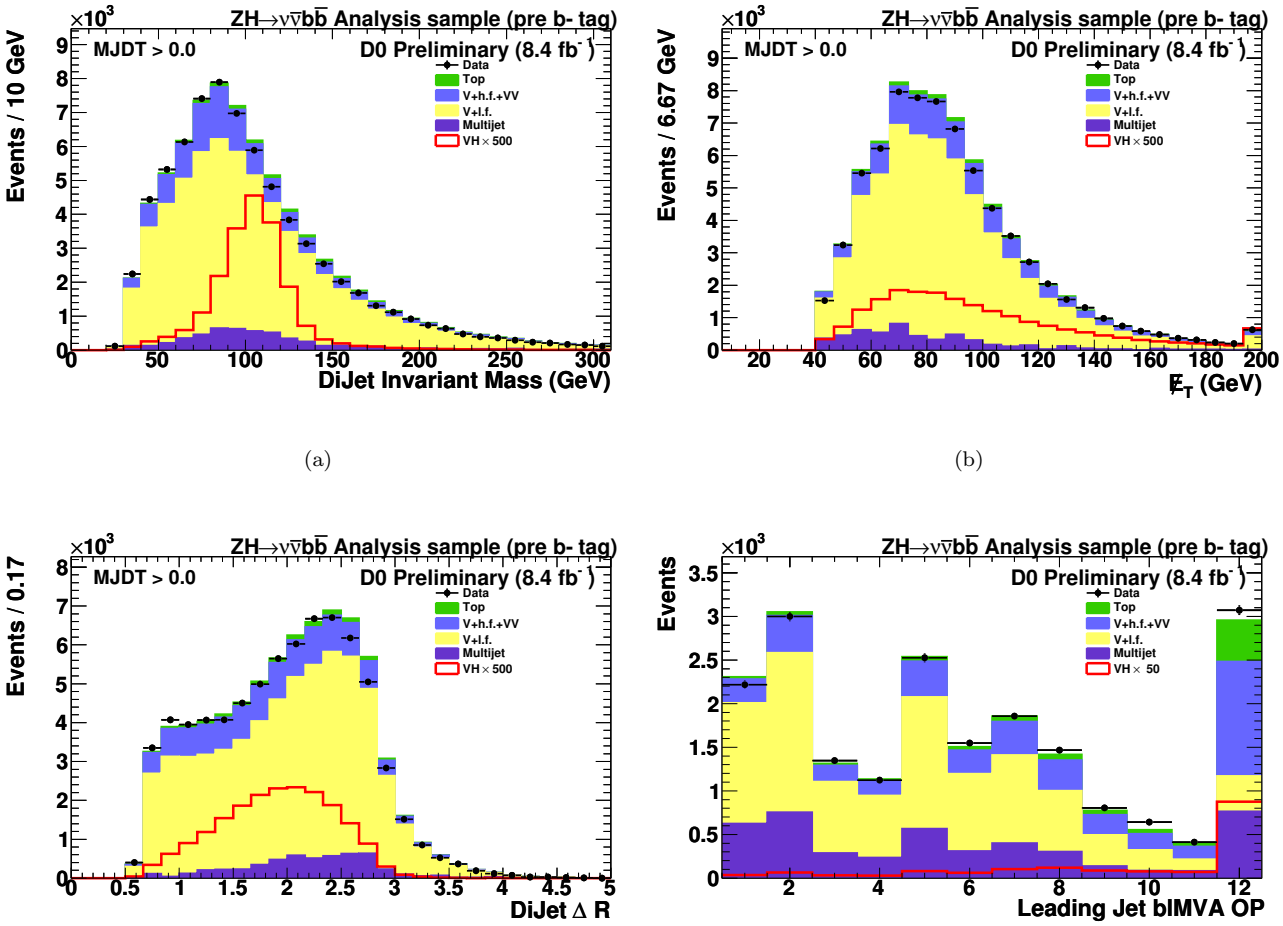
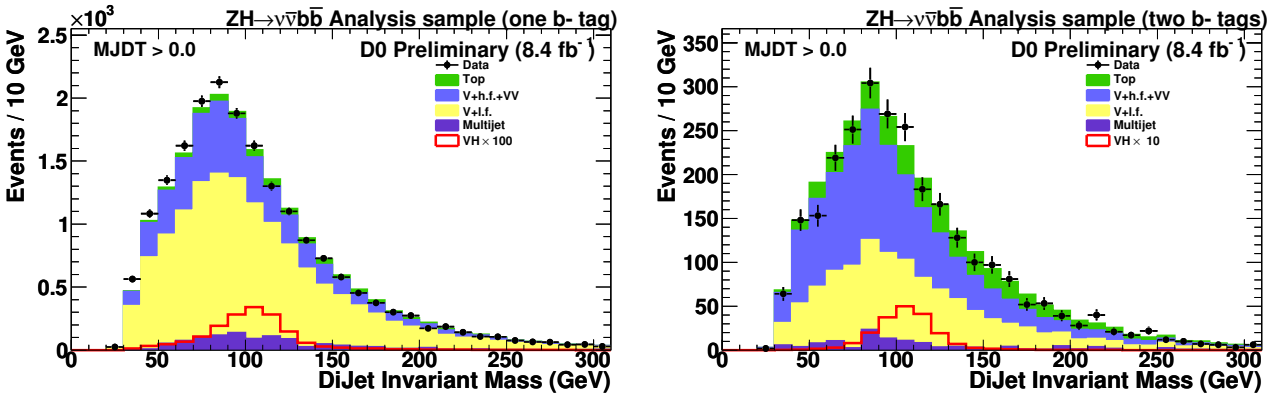


FIG. 5: Representative variable distributions in the analysis sample after the multijet veto and before any b tagging requirement: (a) dijet invariant mass, (b) missing E_T, (c) dijet ΔR, (d) b-tagging algorithm output (L_b) plotted in bins of the twelve operating points defined for the algorithm (12 is the tightest cut on L_b, 0 is the loosest, which is surpressed in this plot due to the large number of entries, mostly from light jets). The distributions for signal (VH), which are multiplied by a factor of 500 for (a)-(c) and 50 for (d), include ZH and WH production for $m_H = 115$ GeV. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production.

The search was also reinterpreted as a search for WZ and ZZ production, to validate the techniques used in the Higgs search, which results in the measurement of a cross section scale factor of $1.5 \pm 0.3(stat) \pm 0.4(syst)$ with an observed significance of 2.8 σ, which is consistent with the expected significance of 1.9 σ. This translates into a measured cross section of $6.9 \pm 1.3(stat) \pm 1.8(syst)$ pb, which is consistent with the predicted SM cross section of 4.6 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); 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 and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada);

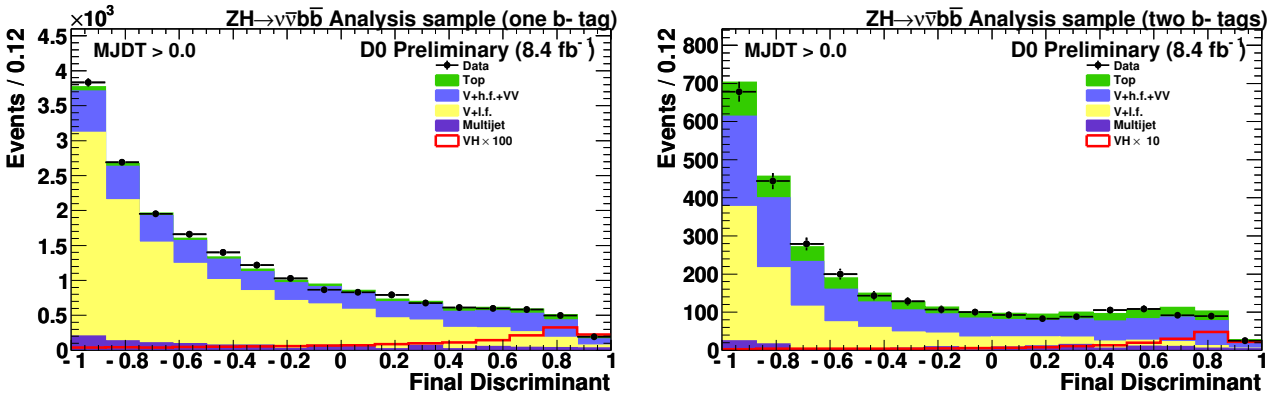


(a) (b)

FIG. 6: Dijet invariant mass in the analysis sample after the multijet veto for (a) single tag and (b) double tag. The distributions for signal (VH), which are multiplied by a factor of 100 for single tag and 10 for double tag respectively, include ZH and WH production for $m_H = 115$ GeV. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes (W/Z)+(u, d, s, g) jets, "V+h.f." includes (W/Z)+(b, c) jets and "Top" includes pair and single top quark production.

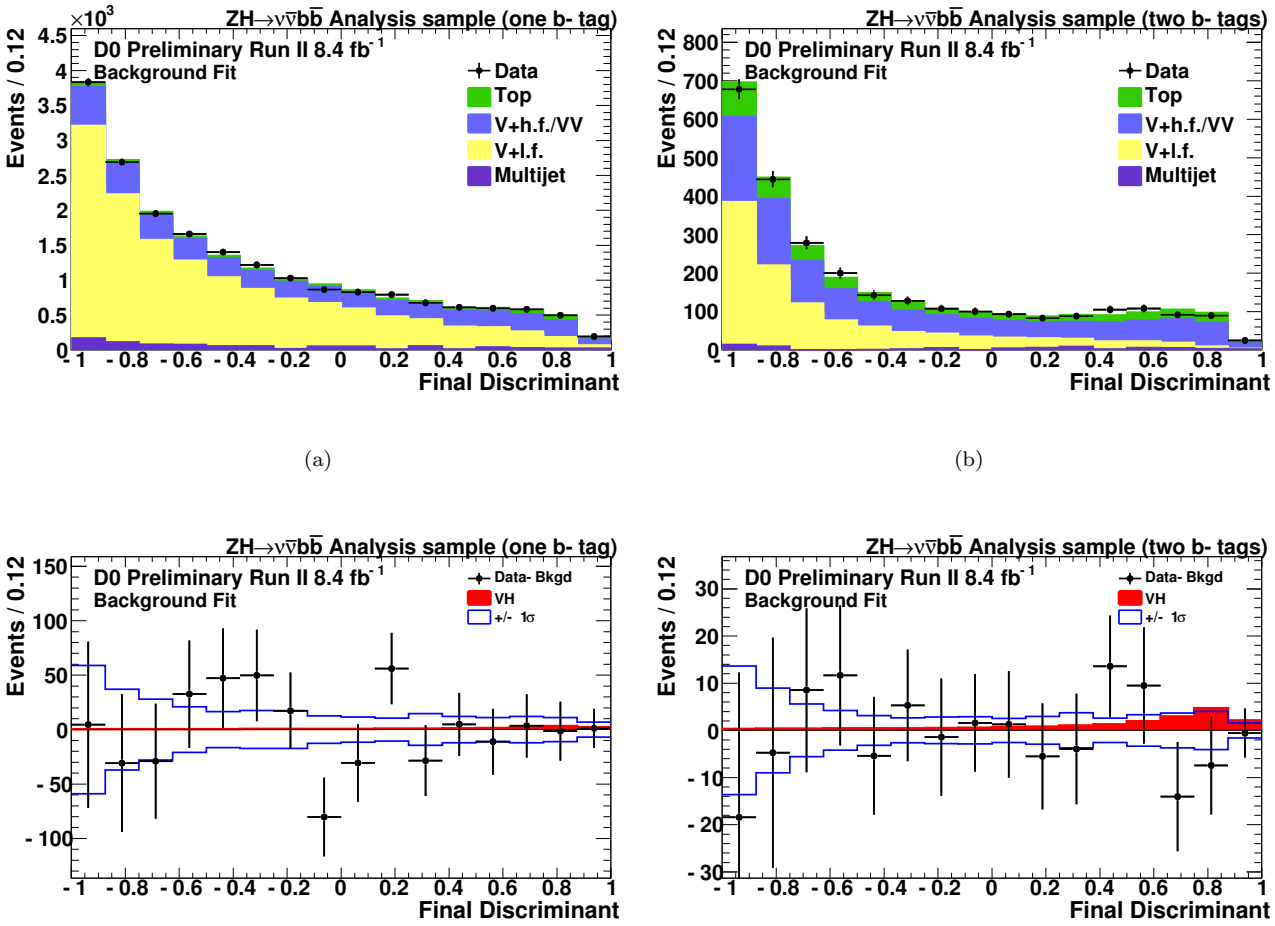
TABLE II: The number of expected signal, expected background and observed data events after the multijet veto, for pre-tag, single and double b tagging requirements. The signal corresponds to $m_H = 115$ GeV, "Top" includes pair and single top quark production, and VV is the sum of all diboson processes. The quoted uncertainties correspond to the statistics of the simulation only.

Sample	ZH	WH	W+jets	Z+jets	Top	VV	Multijet	Total Background	Observed
Pre-tag	21.78 ± 0.09	18.49 ± 0.13	46070	19119	1342	2329	5291	74151 ± 230	72190
Single tag	8.99 ± 0.06	7.81 ± 0.09	12027	4695	681	711	1095	19209 ± 99	19426
Double tag	10.60 ± 0.06	8.80 ± 0.09	1479	673	440	121	125	2838 ± 36	2763



(a) (b)

FIG. 7: The SM DT output for the VH search where $m_H = 115$ GeV following the multijet veto for (a) single and (b) double tag prior to the fit to data. The distributions for signal are multiplied by a factor of 100 for single tag and 10 for double tag respectively, and includes ZH and WH production for $m_H = 115$ GeV. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV," "V+l.f." includes (W/Z)+(u, d, s, g) jets, , ignoring a possible Higgs signal "V+h.f." includes (W/Z)+(b, c) jets and "Top" includes pair and single top quark production.



(a) (b)

(c) (d)

FIG. 8: The SM DT output for SM VH signal, for $m_H = 115$ GeV, following the multijet veto and after the fit to the data under the background-only hypothesis in the (a) single and (b) double tag channels. The data are shown as points and the background contributions as histograms: dibosons are labeled as "VV", "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production. The SM VH signal expectation (red histogram) and the data after subtracting the fitted background (points) in the (c) the single and (d) double tag channels. Also shown is the ± 1 standard deviation band on the total background after fitting. The signal is scaled to the SM cross section.

BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).

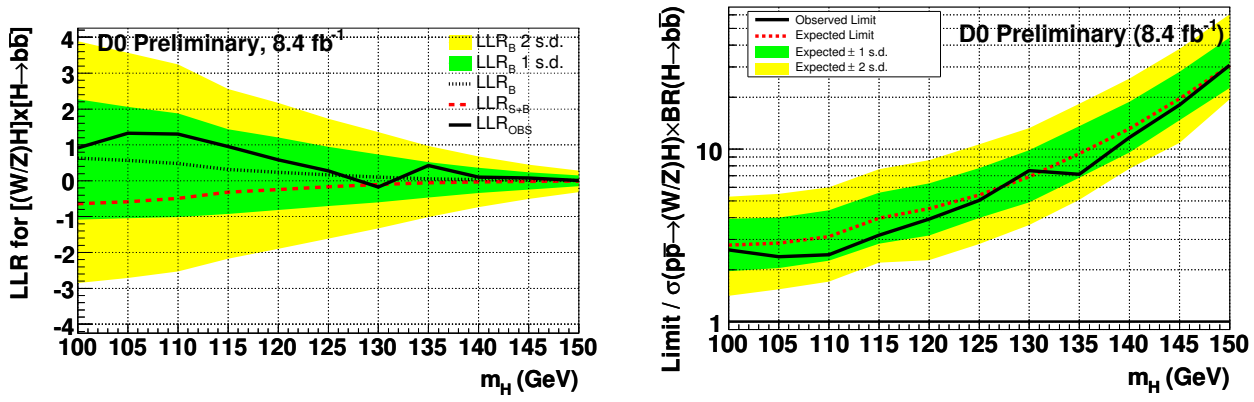
[1] M. Carena *et al.*, arXiv:hep-ph/0010338; CDF and D0 Collaborations, Report No. FERMILAB-PUB-03/320-E, 2003.
[2] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **104**, 071801 (2010).
[3] (D0 Collaboration), http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/HIGGS/H98/H98.pdf.
[4] J. Baglio and A. Djouadi, arXiv:hep-ph/1003.4266v2.
[5] R. Barate *et al.* (LEP Working Group for Higgs boson searches), Phys. Lett. B **565**, 61 (2003).
[6] LEP, Tevatron and SLD Electroweak Working Groups, arXiv:1012.2367.
[7] V.M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006); M. Abolins *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **584**, 75 (2008); R. Angstadt *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **622**, 298 (2010).
[8] C. Ochando, Report No. FERMILAB-THESIS-2008-78.
[9] T. Andeen *et al.*, Report No. FERMILAB-TM-2365, 2007.
[10] G.C. Blazey *et al.*, arXiv:hep-ex/0005012.

TABLE III: Systematic uncertainties in percent of the overall signal and background yields. "Jet EC" and "Jet ER" stand for jet energy calibration and resolution respectively. "Jet R&T" stands for jet reconstruction and taggability. "Signal" includes ZH and WH production and is shown for $m_H = 115$ GeV.

Systematic Uncertainty	Signal (%)	Background (%)
Single Tag		
Jet EC - Jet ER	1.0	2.5
Jet R&T	2.6	2.6
b Tagging	3.2	1.3
Trigger	2	1.9
Lepton Identification	1.1	0.8
Heavy Flavor Fractions	–	4.1
Cross Sections	6	9.8
Luminosity	6.1	5.8
Multijet Normalilzation	–	1.3
Total	9.8	12.3
Double Tag		
Jet EC - Jet ER	0.7	2.3
Jet R&T	3.5	2.6
b Tagging	5.8	3.6
Trigger	2	1.9
Lepton Identification	1.1	1.0
Heavy Flavor Fractions	0	8.0
Cross Sections	6	9.8
Luminosity	6.1	5.8
Multijet Normalilzation	–	1.1
Total	10.9	13.9

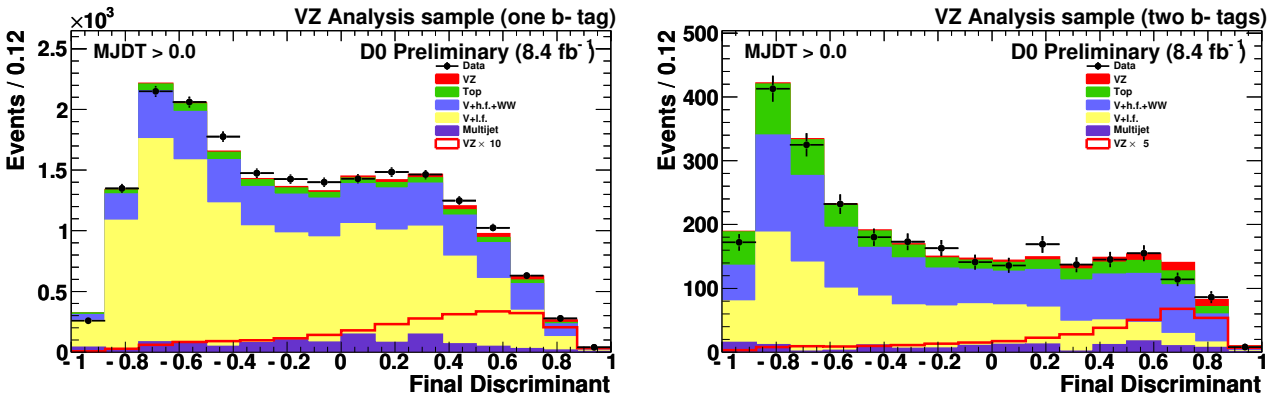
TABLE IV: The observed and expected upper limits measured using 8.4 fb^{-1} of data on the $(W/Z)H$ production cross section relative to the SM expectation as a function of m_H.

m_H	100	105	110	115	120	125	130	135	140	145	150
Expected	2.8	2.9	3.1	4.0	4.5	5.4	6.9	9.4	13.1	19.6	30.5
Observed	2.6	2.4	2.4	3.2	3.9	5.0	7.5	7.1	11.7	18.0	30.6



(a) (b)

FIG. 9: (a) The observed (solid black) and expected LLRs for the background-only (black dots) and signal+background hypotheses (red dashes). (b) Ratio of the observed (solid black) and expected (dotted red) exclusion limits to the SM production cross section for the VH search. Both are shown as a function of m_H with the heavy green and light yellow shaded areas corresponding to the 1 and 2 standard deviations (s.d.) around the background-only hypothesis.



(a) (b)

FIG. 10: The SM DT output for the WZ and ZZ diboson search following the multijet veto for (a) single and (b) double tag prior to the fit to data. The data are shown as points and the background contributions as histograms: "WW" denotes the diboson production considered as background, "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production. The distributions for signal are scaled by the SM cross section (filled red histogram) or multiplied by a factor of 10 for single tag and 5 for double tag (solid red line) respectively,

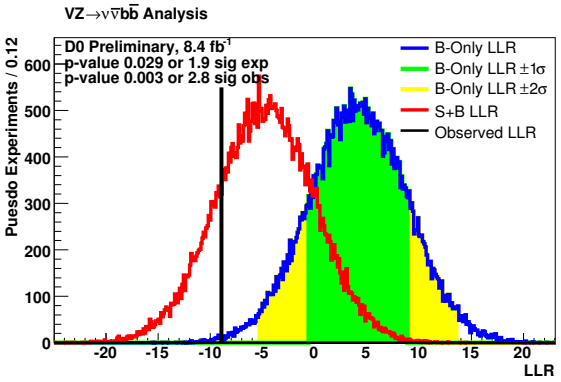


FIG. 11: The observed (solid black) and expected LLRs for the background-only (blue line) and signal+background hypotheses (red line), where WZ and ZZ are taken as the signal. The heavy green and light yellow shaded areas corresponding to the 1 and 2 standard deviations (s.d.) around the background-only hypothesis.

[11] M. Voutilainen, Report No. FERMILAB-THESIS-2008-19.
[12] M.L. Mangano *et al.*, J. High Energy Phys. **07**, 001 (2003); version 2.11 was used.
[13] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. **05**, 026 (2006); version 6.409, D0 Tune A, was used.
[14] V.M. Abazov *et al.*, (D0 Collaboration), Phys. Rev. Lett. **100**, 102002 (2008).
[15] K. Melnikov and F. Petriello, Phys. Rev. D **74**, 114017 (2006).
[16] E. Boos *et al.* (CompHEP Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **534**, 250 (2004).
[17] J. Pumplin *et al.*, J. High Energy Phys. **07**, 012 (2002); D. Stump *et al.*, J. High Energy Phys. **10**, 046, (2003).
[18] R. Hamberg, W.L. van Neerven, and W.B. Kilgore, Nucl. Phys. **B359**, 343 (1991); **B644**, 403 (2002).
[19] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Phys. Lett. B **604**, 61 (2004).
[20] J.M. Campbell and R.K. Ellis, Phys. Rev. D **60**, 113006 (1999).
[21] N. Kidonakis, Phys. Rev. D **74**, 114012 (2006). S. Moch and P. Uwer, Phys. Rev. D **78** 34003 (2008).
[22] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).
[23] The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle with respect to the proton beam direction.
[24] A. Schwartzman, Report No. FERMILAB-THESIS-2004-21.
[25] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **76**, 092007 (2007).
[26] T. Gleisberg *et al.*, J. High Energy Phys. 02 (2004) 056; J. Alwall *et al.*, Eur. Phys. J. C **53**, 473 (2008).

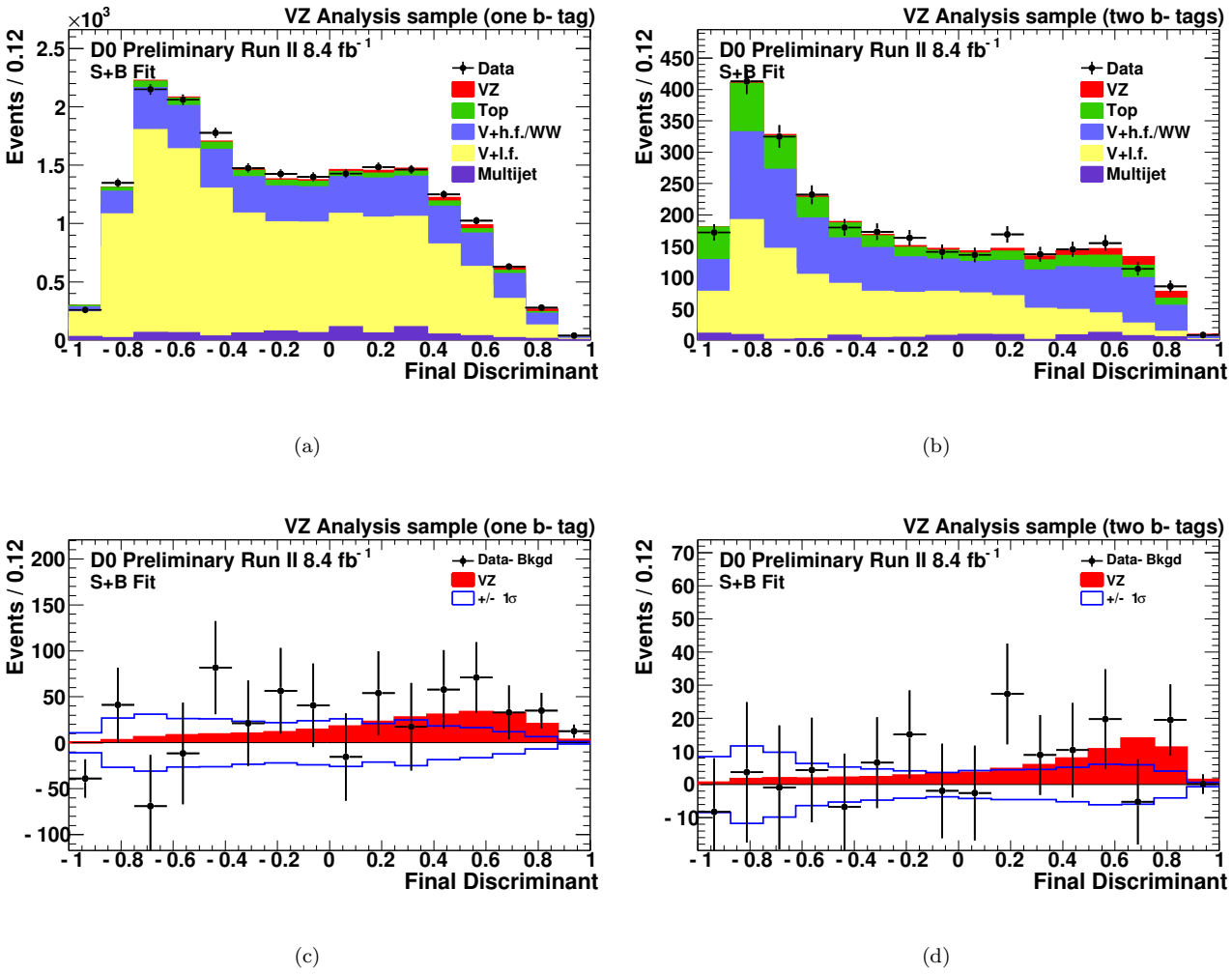


FIG. 12: The SM DT output for the WZ and ZZ diboson search, following the multijet veto, and after the fit to the data under the signal+background hypothesis in the (a) single and (b) double tag channels. The data are shown as points and the background contributions as histograms: "WW" denotes the diboson production considered as background, "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production. The SM diboson signal expectation (red histogram) and the data after subtracting the fitted background (points) in the (c) the single and (d) double tag channels. Also shown is the ± 1 standard deviation band on the total background after fitting. The signal is scaled to the SM cross section.

[27] V. M. Abazov *et al.* (The D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. Sect. **A 620**, 400 (2010).

[28] L. Breiman *et al.*, *"Classification and Regression Trees"* (Wadsworth, Belmont, CA, 1984).
Y. Freund and R.E. Schapire, *"Experiments with a new boosting algorithm"*, in *"Machine Learning: Proceedings of the Thirteenth International Conference", Bari, Italy, 1996,* edited by L. Saitta (Morgan Kaufmann, San Francisco, 1996), p148.

[29] J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett. **105**, 022001 (2010); V. M. Abazov *et al.* (D0 Collaboration), `arXiv:hep-ex/1101.0648`.

[30] K. Black *et al.*, `arXiv:hep-ph/1010.3698v2`.

[31] T. Junk, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999); A. Read, J. Phys. G **28**, 2693 (2002).

[32] W. Fisher, Report No. FERMILAB-TM-2386-E, 2006.

[33] G. Corcella *et al.*, J. High Energy Phys. **01**, 010 (2001).

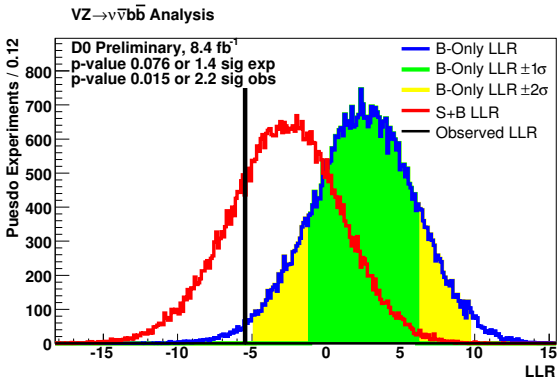


FIG. 13: The observed (solid black) and expected LLRs for the background-only (blue line) and signal+background hypotheses (red line), where WZ and ZZ are taken as the signal for the diboson search using the dijet invariant mass as the final discriminant. The heavy green and light yellow shaded areas corresponding to the 1 and 2 standard deviations (s.d.) around the background-only hypothesis.

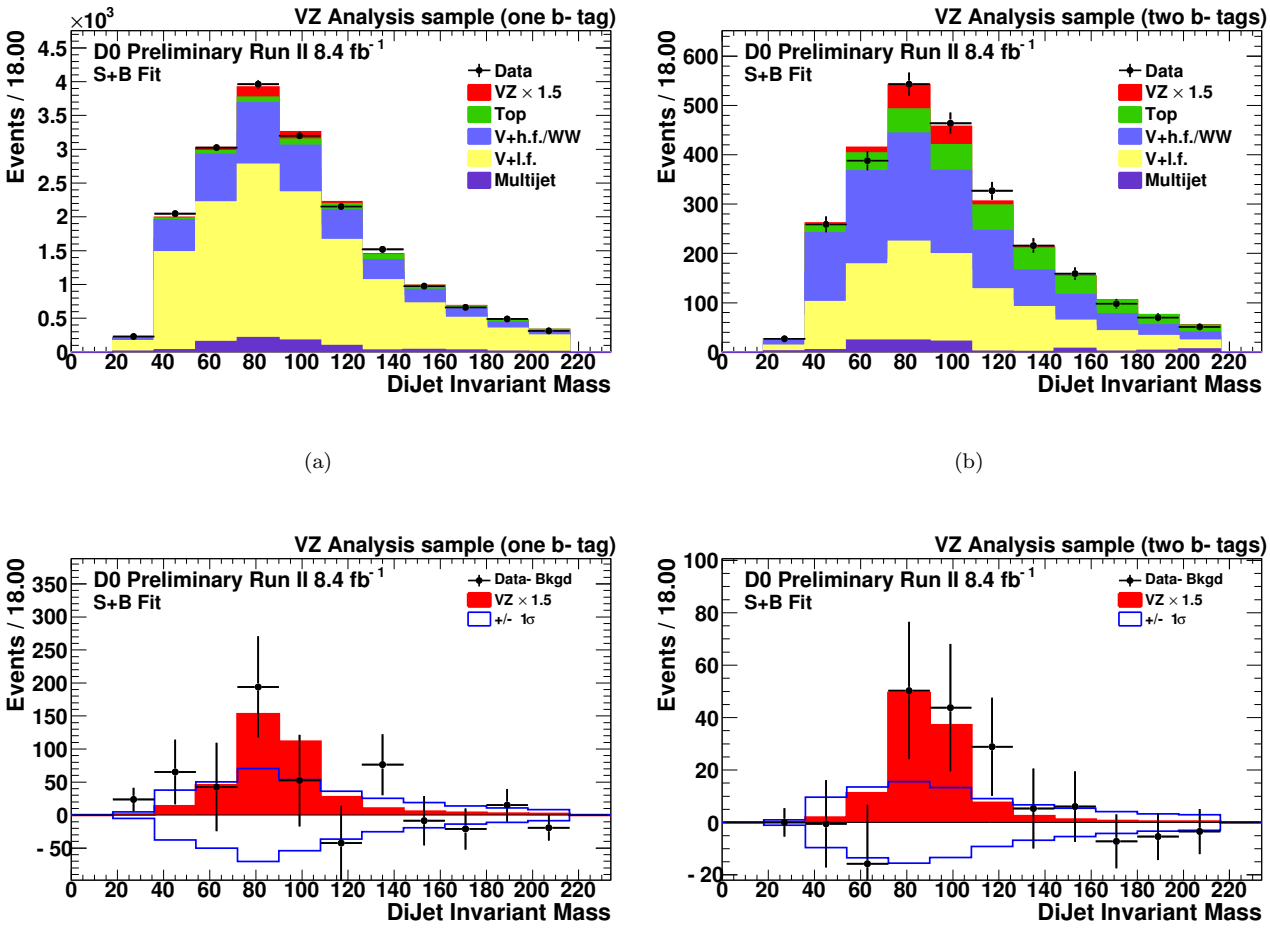


FIG. 14: The dijet invariant mass for the WZ and ZZ diboson search, following the multijet veto, and after the fit to the data under the signal+background hypothesis in the (a) single and (b) double tag channels. The data are shown as points and the background contributions as histograms: "WW" denotes the diboson production considered as background, "V+l.f." includes $(W/Z)+(u,d,s,g)$ jets, "V+h.f." includes $(W/Z)+(b,c)$ jets and "Top" includes pair and single top quark production. The SM diboson signal expectation (red histogram) scaled to the measured cross section and the data after subtracting the fitted background (points) in the (c) the single and (d) double tag channels. Also shown is the ± 1 standard deviation band on the total background after fitting.