



## Search for the Standard-Model Higgs Boson in the $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ Channel in $6.2 \text{ fb}^{-1}$ of $p\bar{p}$ Collisions at $\sqrt{s} = 1.96 \text{ TeV}$

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

(Dated: March 9, 2011)

A search is performed for the standard-model Higgs boson in  $6.2 \text{ fb}^{-1}$  of  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$ , collected with the D0 detector at the Fermilab Tevatron Collider. The final state considered contains a pair of  $b$  jets and is characterised by an imbalance in transverse energy, as expected from  $p\bar{p} \rightarrow ZH \rightarrow \nu\bar{\nu}b\bar{b}$  production. The search is also sensitive to the  $WH \rightarrow \ell\nu b\bar{b}$  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 multiplied by branching fraction  $\sigma(p\bar{p} \rightarrow [Z/W]H) \times BR(H \rightarrow b\bar{b})$  that is a factor of 3.4 larger than the theoretical standard-model value, consistent with the expected factor of 4.0.

*Preliminary Results for Winter 2011 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  $Z \rightarrow \nu\bar{\nu}$  and  $H \rightarrow b\bar{b}$ , 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  $5.2 \text{ fb}^{-1}$  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  $e^+e^- \rightarrow ZH$  [3], while an indirect upper limit of 158 GeV can be inferred from precision electroweak data [4]. These limits and those given below are all defined at the 95% confidence level (C.L.).

The data taking period prior to March 2006 is referred to as Run IIa, while IIb denotes the period after. 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. For the result described in this note the Run IIb dataset was re-analyzed and the integrated luminosity was increased by  $2.0 \text{ fb}^{-1}$  beyond that used in Ref. [2] to give a total integrated luminosity in RunIIb of  $6.2 \text{ fb}^{-1}$ . The Run IIa data is currently in the process of being re-analyzed including all the improvements developed during the analysis of Run IIb (the two data taking epochs need to be analyzed separately because of significant changes in the D0 detector), and will be combined with the current result in the future.

The final-state topology consists of a pair of  $b$  jets from  $H \rightarrow b\bar{b}$  decay and missing transverse energy ( $\cancel{E}_T$ ) from  $Z \rightarrow \nu\bar{\nu}$ . The search is therefore also sensitive to the  $WH$  process when the charged lepton from  $W \rightarrow \ell\nu$  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  $\cancel{E}_T$  arising from mismeasurement of jet energies.

## II. DATA AND SIMULATED SAMPLES

The D0 detector is described in Ref. [5]. The data used in this analysis were recorded using triggers designed to select events with jets and  $\cancel{E}_T$  [6]. After imposing data quality requirements, the total integrated luminosity [7] is  $6.2 \text{ fb}^{-1}$ .

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 [8], and (iii) electrons or muons identified through the association of tracks with electromagnetic calorimeter clusters or with hits in the muon detector, respectively. The  $\cancel{E}_T$  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 [9], and these corrections are propagated to the  $\cancel{E}_T$ .

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 [10], interfaced with PYTHIA [11] for initial and final-state radiation and for hadronization. The  $p_T$  spectrum of the  $Z$  is reweighted to match the D0 measurement [12]. 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 [13]. For  $t\bar{t}$  and electroweak single top quark production, the ALPGEN and COMPHEP [14] 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) [15].

The absolute normalizations for  $(W/Z)$ +jets production are obtained from NNLO calculations of total cross sections based on Ref. [16], using the MRST2004 NNLO PDFs [17]. The heavy-flavor fractions are obtained using MCFM [18] at next-to-leading order (NLO). Cross sections for other SM backgrounds are taken from Ref. [19], or calculated with MCFM, and the cross sections for signal are taken from Ref. [20].

Signal and background samples are passed through a full GEANT3-based simulation [21] 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 at least two and no more than three taggable jets in total, with the Higgs candidate always formed from the two leading jets. These jets must have transverse momentum  $p_T > 20$  GeV and pseudorapidity  $|\eta| < 2.5$  [22]. The two leading taggable jets must not be back-to-back in the plane transverse to the beam direction:  $\Delta\phi(\text{jet}_1, \text{jet}_2) < 165^\circ$ . Finally,  $\cancel{E}_T > 30$  GeV is required.

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, that 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  $\cancel{E}_T > 40$  GeV and a measure of the  $\cancel{E}_T$  significance  $\mathcal{S} > 5$  [23]. Larger values of  $\mathcal{S}$  correspond to  $\cancel{E}_T$  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; there is non-perfect modeling in the analysis sample at low  $\mathcal{S}$ , which is dominated by the MJ background, although it should be noted that a 25% systematic error is assigned to the modeling of this background and the good agreement in the EW sample demonstrates that the SM (and thereby the signal) efficiencies are well modeled. In signal events, the missing track  $p_T$ ,  $\cancel{p}_T$ , defined as the opposite of the vectorial sum of the charged particle transverse momenta, is expected to point in a direction close to that of the  $\cancel{E}_T$ . Such a strong correlation is not expected in multijet events, where the  $\cancel{E}_T$  originates mainly from mismeasurement of jet energies. Advantage is taken of this feature by requiring  $\mathcal{D} < \pi/2$ , where  $\mathcal{D} = \Delta\phi(\cancel{E}_T, \cancel{p}_T)$ . Events containing an isolated electron or muon [24] with  $p_T > 15$  GeV are rejected to reduce backgrounds from  $W$ +jets, top quark, and diboson production.

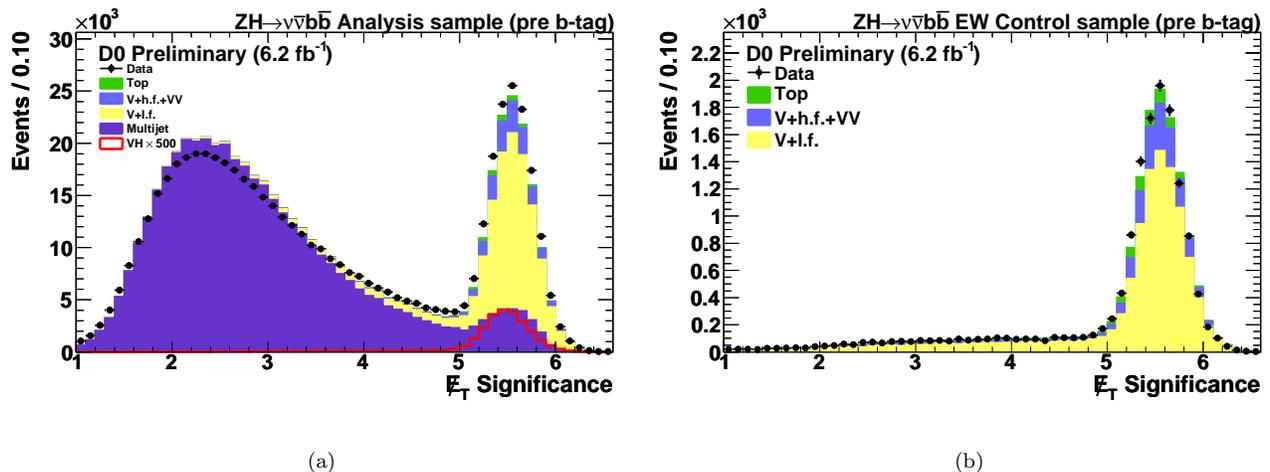


FIG. 1: Missing  $E_T$  significance in (a) the analysis and (b) the EW-control samples without the requirement that it be larger than 5. The distributions for signal (VH) are multiplied by a factor of 500 and include  $ZH$  and  $WH$  production for  $m_H = 115$  GeV. Data is shown as points and the 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 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 the transverse mass of the muon and  $\cancel{E}_T$  system to be larger than 30 GeV. To ensure similar jet topologies for the analysis and EW-control samples,  $\cancel{E}_T$ , not corrected for the selected muon, is required to exceed 40 GeV. The number of selected events is in excellent 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 leading taggable jets is performed, as suggested by a simulation of  $(W/Z)$ +jets using the SHERPA generator [25]. Four representative distributions in the EW-control

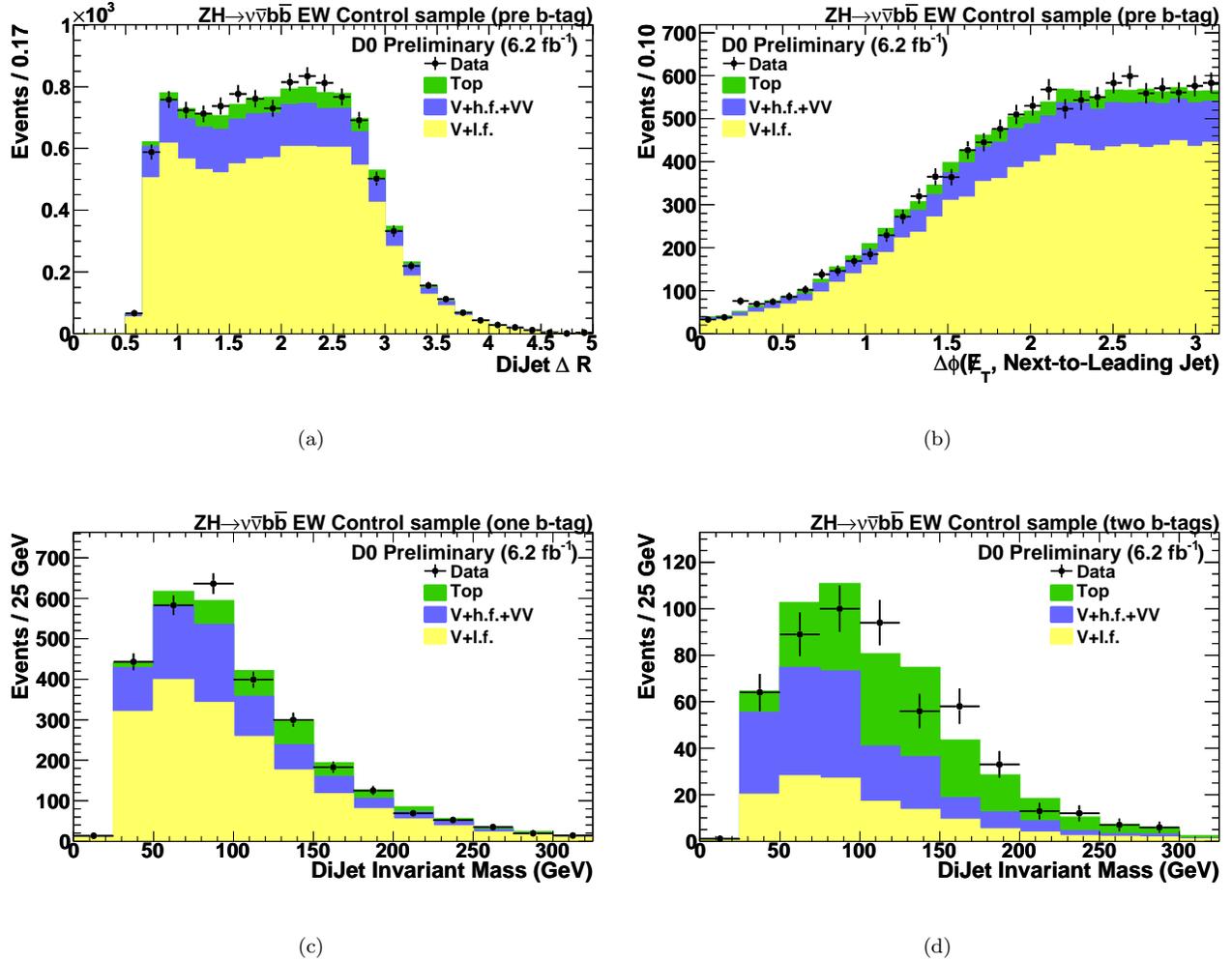


FIG. 2: Representative variable distributions in the EW-control sample: (a) dijet  $\Delta R$  in the pre-tag sample, (b)  $\Delta\phi(\cancel{E}_T, \text{Next-to-Leading Jet})$  in the pre-tag sample, (c) dijet invariant mass in the single-tag sample, (d) dijet invariant mass in the double-tag sample. Data is 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.

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. After adding contributions from SM backgrounds, the MJ background is normalized so that the expected number of events is identical to the number observed 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  $\cancel{E}_T$  threshold is reduced to 30 GeV and no requirement is imposed on  $\mathcal{S}$ . As a result, 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 [27] 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 [26]. 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.

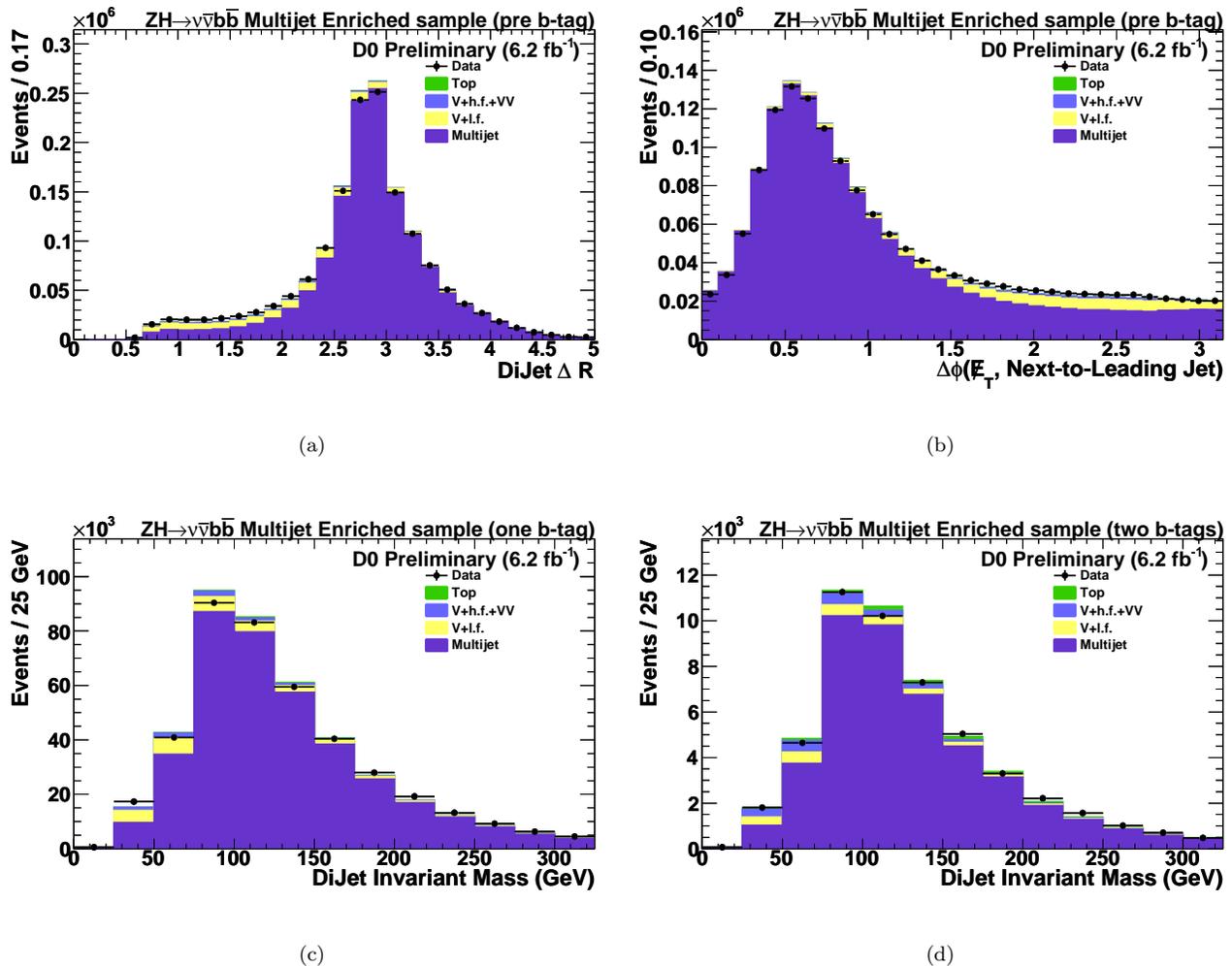


FIG. 3: Representative variable distributions in the MJ-enriched sample: (a) dijet  $\Delta R$  in the pre-tag sample, (b)  $\Delta\phi(\cancel{E}_T, \text{Next-to-Leading Jet})$  in the pre-tag sample, (c) dijet invariant mass in the single-tag sample, (d) dijet invariant mass in the double-tag sample. Data is 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.

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 one (single tag) or two (double tag) of the leading taggable jets satisfy a very loose cut on  $L_b$  that accepts  $\approx 80\%$  of  $b$  jets and  $\approx 10\%$  of light jets with  $p_T \approx 45$  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). This approach results in a  $\approx 14\%$  increase in the expected limit compared to just placing optimized cuts on  $L_b$  and making no further use of the information. 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. [26].

#### 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 twenty kinematic variables. These variables include the number of jets, jet  $p_T$ , dijet  $p_T$ ,  $\cancel{E}_T$ , angles between jets, between the jets and  $\cancel{E}_T$  and the dijet

mass. The following, more elaborate variables are also included: (i) a color-flow variable (described in Ref. [28]) to estimate if the jets are from decays of color singlets, or initiated by gluon splitting as in the main backgrounds, (ii) the cosine of the angle between the direction of the Higgs-candidate and the leading jet, boosted to the rest frame of the Higgs-candidate; this distribution is expected to be uniform for the decay of a scalar, in contrast to gluon splitting, and (iii) the  $p_T$  weighted distance in  $(\eta - \phi)$  space between the second leading jet and other jets. The full list of input variables to the decision trees is given in Table I.

The MJ DT output, whose distributions range 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 over 90% of the multijet background and 72% of the non-MJ SM backgrounds, while retaining 84% 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.

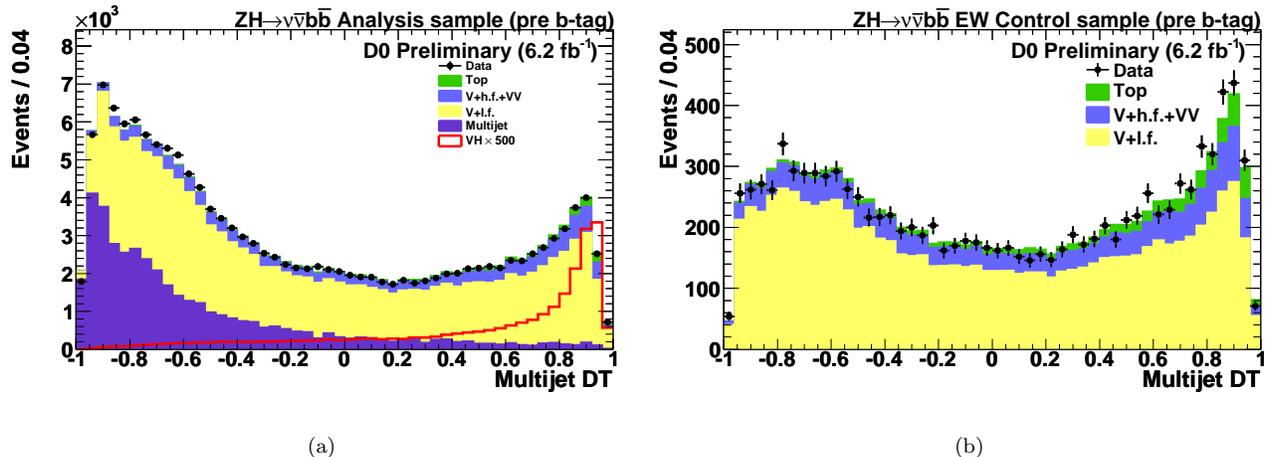


FIG. 4: MJ DT output for  $m_H = 115$  GeV before any  $b$ -tagging requirement in the (a) signal sample and (b) EW-control sample. The distribution for signal (VH) is multiplied by a factor of 500 and includes  $ZH$  and  $WH$  production. Data is shown as points and the background contributions as 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.

Next, to discriminate signal from the other SM backgrounds, two “SM DTs” (SM-rejection DTs) are trained for each  $m_H$ , one in the single tag channel and one in the double tag channel. The same kinematic variables are used as for the MJ DT, with the addition of the  $b$ -tagging output of the  $b$ -tagged jets comprising the dijet system. The SM DT outputs, ranging 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 (2-3%), jet reconstruction and taggability (3%), lepton identification (0.5%), the modeling of the MJ background (25%, which translates into a 2% uncertainty on the total background) and  $b$  tagging (from 4% for background in the single-tag sample to 10% 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, an uncertainty of 6% is assigned to the total cross sections, and an uncertainty of 20% on the heavy-flavor fractions (estimated from MCFM at NLO [18]). For other SM backgrounds, uncertainties are taken from Ref. [19] or from MCFM, and range from 6% to 10%. The uncertainties on cross sections for signal (6% for  $m_H = 115$  GeV) are taken

TABLE I: Variables used as input to the decision trees.

| Variable  |
|---|
| Number of taggable jets   |
| Leading jet $p_T$   |
| Second jet $p_T$  |
| Leading jet $b$ -tagging output (only used in the SM DT)                                |
| Second jet $b$ -tagging output (only used in the SM DT)                                 |
| $\Delta\eta$ (Leading Jet, Next-to-Leading Jet)   |
| $\Delta\phi$ (Leading Jet, Next-to-Leading Jet)   |
| $\Delta R$ (Leading Jet, Next-to-Leading Jet)   |
| $p_T$ weighted $\Delta R$ (Next-to-Leading Jet, Other Jets)                             |
| $\cancel{E}_T$  |
| $\Delta\phi(\cancel{E}_T, \text{Next-to-Leading Jet})$                                  |
| $\max \Delta\phi(\cancel{E}_T, \text{Jet}) + \min \Delta\phi(\cancel{E}_T, \text{Jet})$ |
| $\max \Delta\phi(\cancel{E}_T, \text{Jet}) - \min \Delta\phi(\cancel{E}_T, \text{Jet})$ |
| $H_T$ (scalar sum of jet $p_T$ )  |
| $\vec{H}_T$ (vectorial sum of jet $p_T$ )   |
| $\vec{H}_T / H_T$   |
| Dijet $p_T$   |
| Dijet mass  |
| Dijet transverse mass   |
| $\cos$ (Higgs decay angle)  |
| Color flow leading jet  |
| Color flow second jet   |

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    | $20.69 \pm 0.08$ | $19.8 \pm 0.13$ | 34 306    | 13 155    | 2 139 | 1 753 | 4 558    | $55\,767 \pm 207$ | 56 017   |
| Single Tag | $8.66 \pm 0.05$  | $8.36 \pm 0.09$ | 9 272     | 3 479     | 1 127 | 552   | 1 588    | $16\,018 \pm 116$ | 15 863   |
| Double Tag | $9.95 \pm 0.06$  | $9.41 \pm 0.09$ | 1 183     | 514       | 636   | 83    | 187      | $2\,603 \pm 42$   | 2 501    |

from Ref. [20]. 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 [31] to ALPGEN interfaced with PYTHIA, and (ii) the choice of PDFs, estimated using the prescription of Ref. [15].

## 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 [29] is used to set limits on the cross section for SM Higgs-boson production multiplied by the branching fraction for  $H \rightarrow b\bar{b}$ , 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 systematic uncertainties on the sensitivity of the analysis is reduced by maximizing a “profile” likelihood function [30] in which these uncertainties are given Gaussian constraints associated with their priors. Figure 8 shows a comparison of the DT distribution after profiling between background-subtracted data and the expected signal scaled by the observed limit in the Run IIb analysis 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.

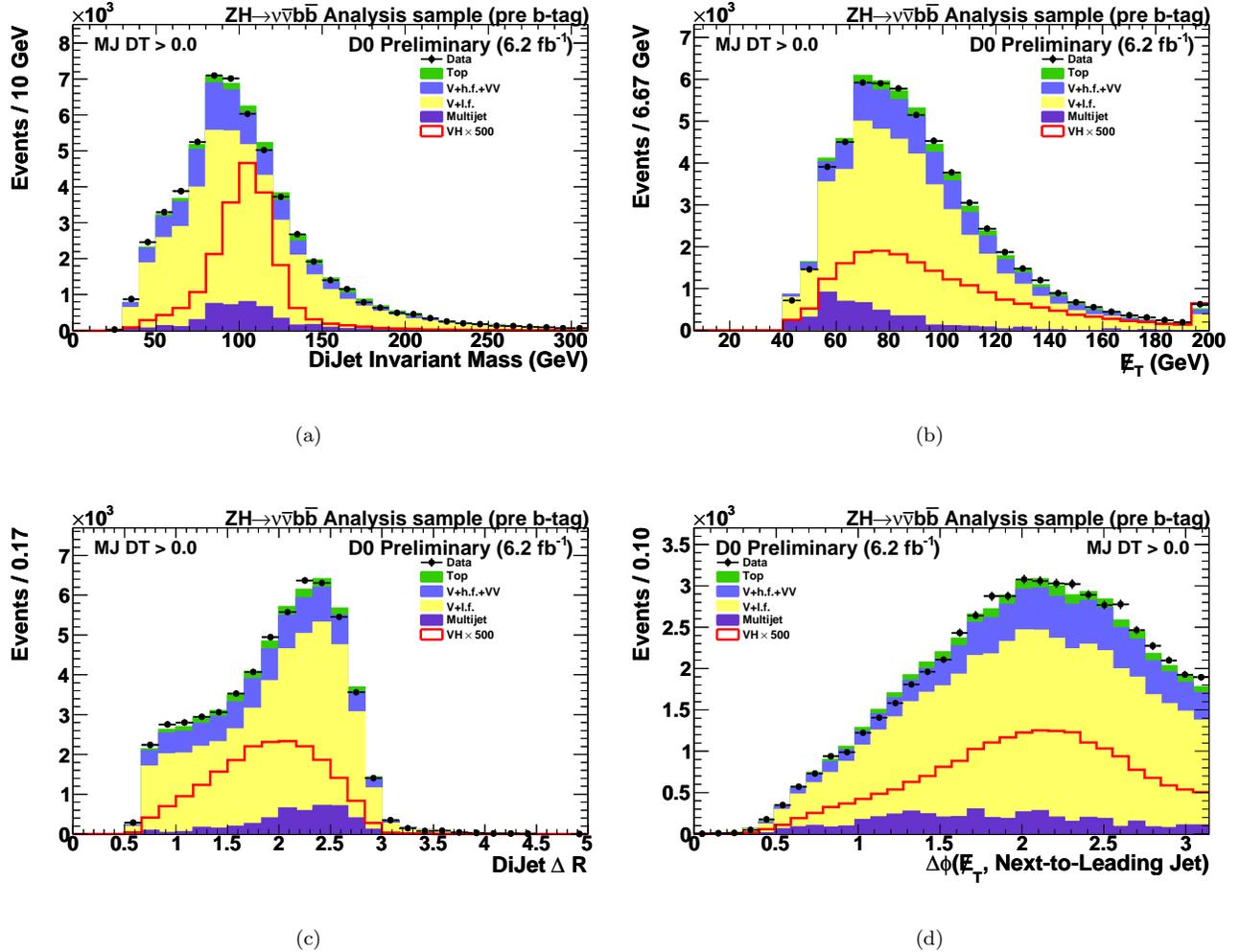


FIG. 5: Representative variable distributions in the analysis sample after the multijet veto and before any  $b$  tagging requirement: (a) dijet invariant mass, (b) missing  $E_T$ , (c) dijet  $\Delta R$ , (d)  $\Delta\phi(\cancel{E}_T, \text{Next-to-Leading Jet})$ . The distributions for signal (VH) are multiplied by a factor of 500 and include  $ZH$  and  $WH$  production for  $m_H = 115$  GeV. Data is 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.

## VII. RESULTS

The results of the updated analysis using  $6.2 \text{ fb}^{-1}$  of RunIIb 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, multiplied by the branching fraction for  $H \rightarrow b\bar{b}$ , are factors of 3.4 and 4.0 larger than the SM value, respectively.

## 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);

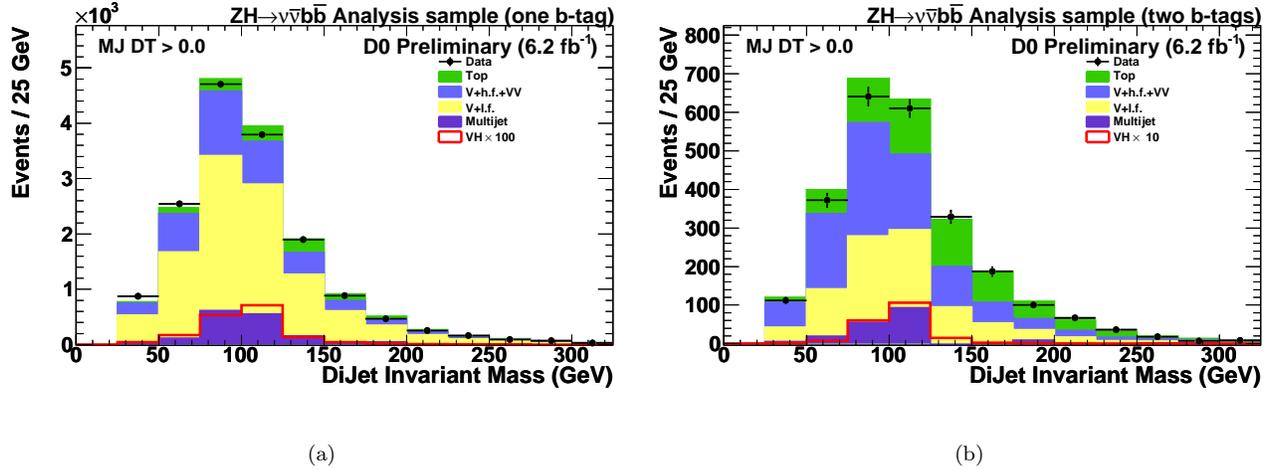


FIG. 6: Dijet invariant mass in the analysis sample after the multijet veto for (a) single tag and (b) double tag. The distributions for signal (VH) are multiplied by a factor of 100 for single tag and 10 for double tag respectively, and include  $ZH$  and  $WH$  production for  $m_H = 115$  GeV. Data is 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.

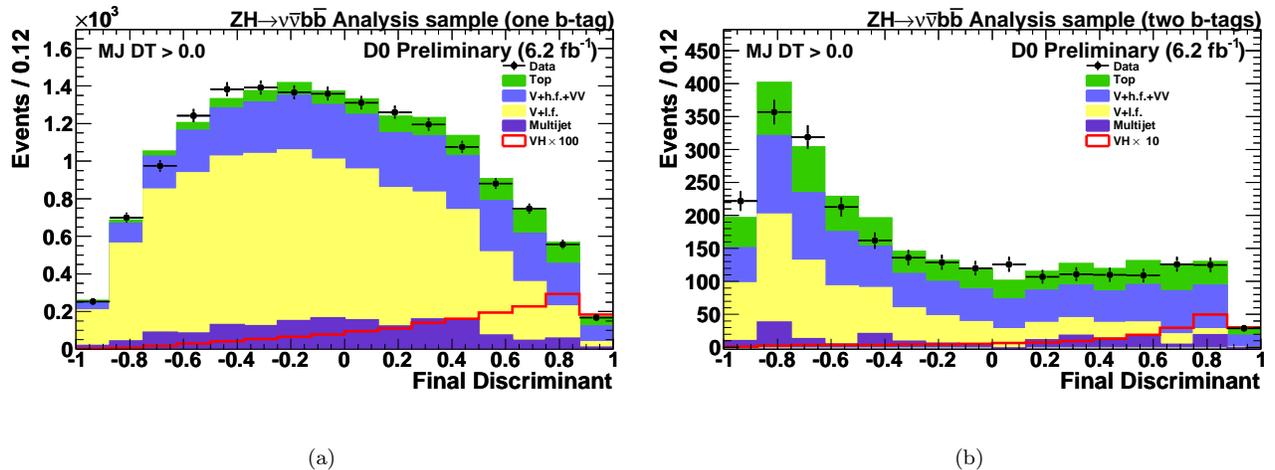


FIG. 7: SM DT output for  $m_H = 115$  GeV following the multijet veto for (a) single and (b) double tag. The distributions for signal (VH) are multiplied by a factor of 100 for single tag and 10 for double tag respectively, and includes  $ZH$  and  $WH$  production for  $m_H = 115$  GeV. Data is shown 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.

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](https://arxiv.org/abs/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] R. Barate *et al.* (LEP Working Group for Higgs boson searches), *Phys. Lett. B* **565**, 61 (2003).  
 [4] LEP, Tevatron and SLD Electroweak Working Groups, [arXiv:1012.2367](https://arxiv.org/abs/1012.2367).  
 [5] V.M. Abazov *et al.* (D0 Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **565**, 463 (2006); M. Abolins *et al.*,

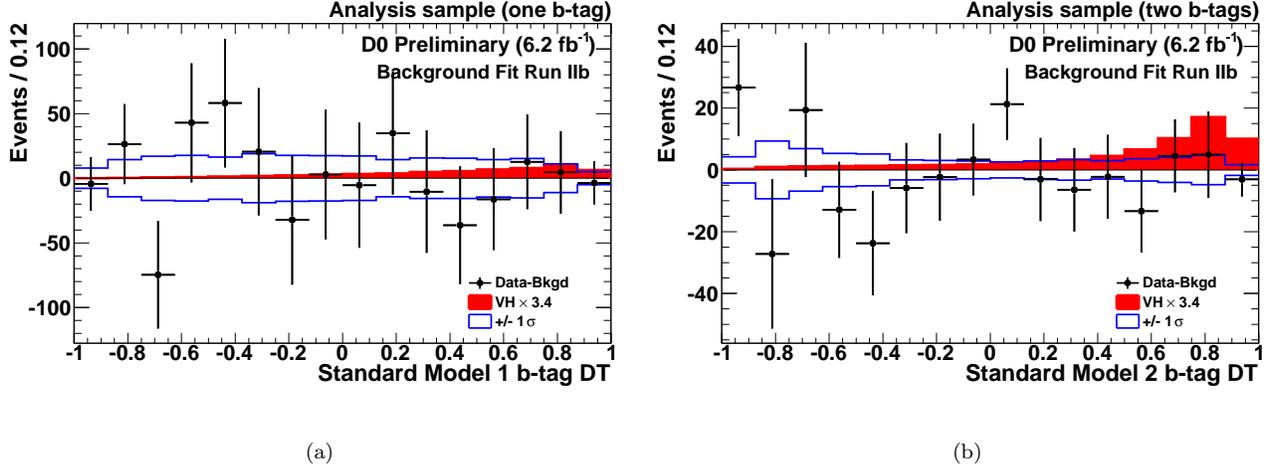


FIG. 8: Data after subtracting the fitted background (points) and scaled SM signal expectation (red histogram) as a function of the SM DT output for  $m_H = 115$  GeV for the (a) single and (b) double tag channels after the fit to the data under the background-only hypothesis. The histogram for signal is scaled by the observed limit in the Run IIb analysis. Also shown is the  $\pm 1$  standard deviation band on the total background after fitting.

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        | 2.1        | 2.9            |
| Jet R&T                | 3.4        | 2.9            |
| $b$ Tagging            | 7.0        | 4.3            |
| Trigger                | 2.0        | 1.8            |
| Lepton identification  | 0.4        | 0.3            |
| Heavy Flavor Fractions | -          | 4.2            |
| Cross Sections         | 6          | 6.4            |
| Luminosity             | 6.1        | 5.5            |
| Multijet Normalization | -          | 2.1            |
| Total                  | 11.9       | 10.8           |
| Double Tag             |            |                |
| Jet EC - Jet ER        | 0.4        | 1.1            |
| Jet R&T                | 3.3        | 3.2            |
| $b$ Tagging            | 9.5        | 7.4            |
| Trigger                | 2.0        | 1.9            |
| Lepton identification  | 0.4        | 0.5            |
| Heavy Flavor Fractions | -          | 7.3            |
| Cross Sections         | 6.0        | 7.2            |
| Luminosity             | 6.1        | 5.7            |
| Multijet Normalization | -          | 1.5            |
| Total                  | 13.4       | 13.4           |

TABLE IV: The observed and expected upper limits measured using  $6.2 \text{ fb}^{-1}$  of RunIIb data on the  $(W/Z)H$  production cross section multiplied by branching fraction for  $H \rightarrow b\bar{b}$  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 | 3.2 | 3.4 | 3.7 | 4.0 | 5.1 | 5.6 | 7.5 | 10.2 | 13.5 | 20.2 | 30.6 |
| Observed | 2.5 | 2.8 | 3.9 | 3.4 | 3.3 | 3.8 | 4.7 | 7.1  | 10.4 | 14.2 | 22.0 |

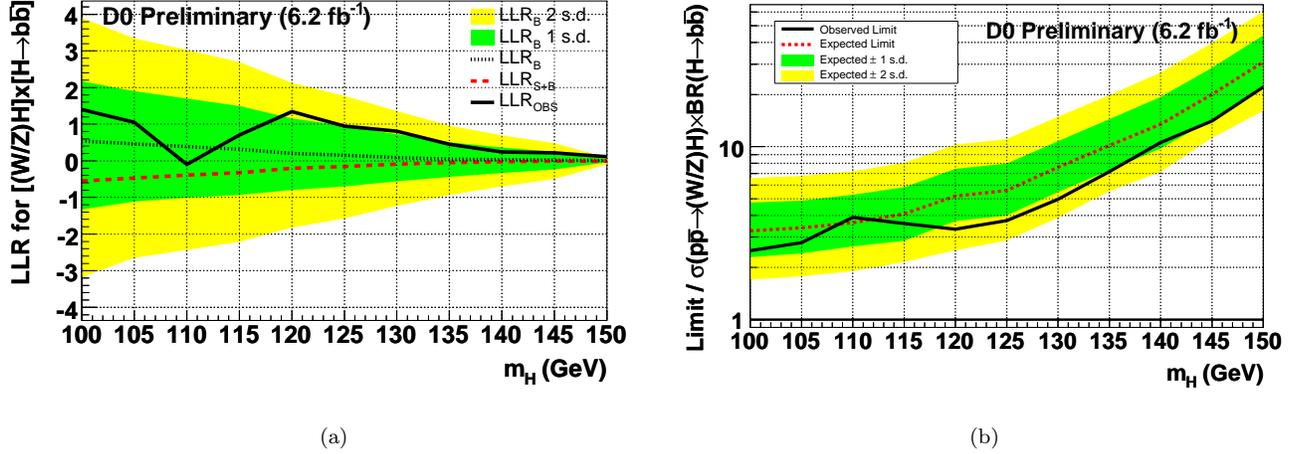


FIG. 9: (a) The observed (solid black) and expected LLRs for the background-only (black dots) and signal+background hypotheses (red dashes). (b) Ratio of the observed (solid black) and expected (dotted red) exclusion limits to the SM production cross section multiplied by branching fraction for  $H \rightarrow b\bar{b}$ . 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 (a) background-only hypotheses and (b) expected limit.

Nucl. Instrum. Methods Phys. Res., Sect. A **584**, 75 (2008); R. Angstadt *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **622**, 298 (2010).

- [6] C. Ochando, Report No. FERMILAB-THESIS-2008-78.  
 [7] T. Andeen *et al.*, Report No. FERMILAB-TM-2365, 2007.  
 [8] G.C. Blazey *et al.*, arXiv:hep-ex/0005012.  
 [9] M. Voutilainen, Report No. FERMILAB-THESIS-2008-19.  
 [10] M.L. Mangano *et al.*, J. High Energy Phys. **07**, 001 (2003); version 2.11 was used.  
 [11] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. **05**, 026 (2006); version 6.409, D0 Tune A, was used.  
 [12] V.M. Abazov *et al.*, (D0 Collaboration), Phys. Rev. Lett. **100**, 102002 (2008).  
 [13] K. Melnikov and F. Petriello, Phys. Rev. D **74**, 114017 (2006).  
 [14] E. Boos *et al.* (CompHEP Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **534**, 250 (2004).  
 [15] J. Pumplin *et al.*, J. High Energy Phys. **07**, 012 (2002); D. Stump *et al.*, J. High Energy Phys. **10**, 046, (2003).  
 [16] R. Hamberg, W.L. van Neerven, and W.B. Kilgore, Nucl. Phys. **B359**, 343 (1991); **B644**, 403 (2002).  
 [17] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Phys. Lett. B **604**, 61 (2004).  
 [18] J.M. Campbell and R.K. Ellis, Phys. Rev. D **60**, 113006 (1999).  
 [19] N. Kidonakis, Phys. Rev. D **74**, 114012 (2006). S. Moch and P. Uwer, Phys. Rev. D **78** 34003 (2008).  
 [20] J. Baglio and A. Djouadi, arXiv:1003.4266 [hep-ph].  
 [21] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).  
 [22] The pseudorapidity is defined as  $\eta = -\ln[\tan(\theta/2)]$ , where  $\theta$  is the polar angle with respect to the proton beam direction.  
 [23] A. Schwartzman, Report No. FERMILAB-THESIS-2004-21.  
 [24] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **76**, 092007 (2007).  
 [25] T. Gleisberg *et al.*, J. High Energy Phys. **02** (2004) 056; J. Alwall *et al.*, Eur. Phys. J. C **53**, 473 (2008).  
 [26] V. M. Abazov *et al.* (The D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. Sect. A **620**, 400 (2010).  
 [27] 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.  
 [28] J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett. **105**, 022001 (2010); V. M. Abazov *et al.* (D0 Collaboration), arXiv:1101.0648 [hep-ex].  
 [29] T. Junk, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999); A. Read, J. Phys. G **28**, 2693 (2002).  
 [30] W. Fisher, Report No. FERMILAB-TM-2386-E, 2006.  
 [31] G. Corcella *et al.*, J. High Energy Phys. **01**, 010 (2001).