DØ Note 5828-CONF
February 5^{th}, 2009

Search for WH associated production using a combined Neural Network and Matrix Element Approach with 2.7 fb^{-1} of Run II data

A search for WH production in $p\bar{p}$ collisions at a center of mass energy of $\sqrt{s} = 1.96\,\text{TeV}$ is presented, combining the full Tevatron Run IIa data with a part of the Run IIb data. The data correspond to an integrated luminosity of 2.7 fb^{-1} accumulated by the DØ experiment. Events containing one lepton, missing transverse energy and one or two b-tagged jets are selected using a neural network to enhance the potential WH signal over standard model (SM) background, A discriminant based on matrix element calculations of the signal and the background is used as one of the inputs of a neural network. In the single and double b-tagged samples, good agreement between data and the background prediction is observed. We set a combined upper limit at 95% C.L. on $\sigma(p\bar{p} \to WH) \times B(H \to b\bar{b})$ with an observed (expected) ratio to the SM cross section of 6.7 (6.4) for $m_H = 115$ GeV.

Preliminary Results for Winter 2009 Conferences

I. INTRODUCTION

For Higgs searches, the most sensitive production channel at the Tevatron for a Higgs mass below $\sim 125\,\text{GeV}$ is the associated production of a Higgs boson with a W boson. Several searches for WH production have been published at a center-of-mass energy of $\sqrt{s} = 1.96\,\text{TeV}$. Three [1–3] used subsamples (0.17 fb^{-1}, 0.44 fb^{-1}, and 1.1 fb^{-1}) of the data analyzed in this conference note, while two others, from the CDF collaboration, are based on 0.32 fb^{-1} and 0.95 fb^{-1} of integrated luminosity [4, 5].

This note presents a new search using a combined neural network plus matrix element approach on 2.7 fb^{-1} of data. The search is based on events with one lepton (e or μ), missing transverse energy $\not\!\!E_T$ to account for the neutrino in the W boson decay, and exactly two or three jets, with one or two of these jets being b-tagged, since $H \to b\bar{b}$ is the dominant SM Higgs decay at low mass.

In double b-tagged events, the dominant backgrounds to WH are $Wb\bar{b}$ production, $t\bar{t}$ and single-top quark production. In single b-tagged events, multijet events and W production in association with c and/or light jets also provide important contributions to the background.

The channels are separated in events having exactly one "tight" b-tagged jet, and those having two "loose" b-tagged jet (with no overlap). We then apply a neural network (NN) to the selected events to separate the standard model background from the signal, and search for an excess at high values of the discriminant. The neural network takes kinematics variables and a matrix element discriminant as inputs. The matrix element input, which was not used in our previous WH results, allows for an additional gain in sensitivity ($\sim 5\%$).

The resulting four $W + 2$ jet channels (e and μ for 1 b-tag and 2 b-tag) are analyzed independently to optimize the sensitivity and then combined. In parallel, the four equivalent channels in $W + 3$ jet events are also analyzed, using the dijet invariant mass as discriminant variable, and also combined with the previous four channels.

The data set corresponds to an integrated luminosity of 2.7 fb^{-1} after requiring good quality criteria and has been recorded between April 2002 and February 2006 (full Run IIa $\sim 1.1\,\text{fb}^{-1}$) and data recorded between June 2006 and April 2008 (part of Run IIb $\sim 1.6\,\text{fb}^{-1}$), i.e. after the Spring 06 Tevatron shutdown, during which the DØ detector was upgraded. In particular the triggering system was improved, and a new layer of silicon sensors of the microvertex detector was installed closer to the beam pipe, giving more precise track reconstruction. The two data periods are analyzed independently and then combined.

II. DATA SAMPLE

The analysis relies on the following components of the DØ detector: a central-tracking system, which consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet [6]; a liquid-argon/uranium calorimeter consisting of a central section (CC) covering pseudorapity $|\eta|$ up to ≈ 1.1, and two endcap calorimeters (EC) extending coverage to $|\eta| < 4.0$, all housed in separate cryostats [7], with scintillators between the CC and EC cryostats providing sampling of developing showers at $1.1 < |\eta| < 1.4$; a muon system which resides beyond the calorimetry, and consists of layers of tracking detectors and scintillation trigger counters before and after the 1.8 T iron toroids.

The luminosity is measured using plastic scintillator arrays located in front of the EC cryostats, at $2.7 < |\eta| < 4.4$. The uncertainty on the measured luminosity is 6.1% [8]. We reject data periods in which the tracking (CFT and SMT), calorimeter or muon information may be compromised.

The trigger and data acquisition systems are designed to accommodate the large instantaneous luminosity of Run II. The events used in this analysis are triggered in the electron channel by a single trigger requiring an electromagnetic (EM) object and also, in Run IIa, using one EM objet and one jet (EM+jet). Triggers are taken into account in the simulation through event reweighting with an efficiency derived from data, and parametrized as a function of lepton azimuthal angle φ and η, and jet p_T. This efficiency is on average $\sim 95\%$ (90%) in Run IIa (IIb) for the events passing our requirements.

In the muon analysis we accept events from any trigger, since we expect close to 100% of our events to be triggered by our redundant triggering system (single muons, muon+jets, topological triggers). This is verified using a combination of single-muon triggers for which a full analysis is completed. The efficiency of this trigger combination is measured on di-muon Z events with an uncertainty of 3%, and is $\sim 70\%$ for the events passing our requirements.This efficiency is consistent with the overall increase in statistics (47% on average) that we observe when accepting events from any trigger, both in data and in simulation, within $\pm 3\%$ for the high statistics channels, and within $\pm 10\%$ for the (low statistics) b-tagged channels. The shape of all distributions remains unchanged within the acceptance systematic uncertainty originating mainly from the jet energy scale uncertainty. We thus use all triggers for the final analysis, after attributing an additional 5% uncertainty on the trigger systematic in this channel.

III. SIMULATED DATASETS

Background and signal processes have been generated using the CTEQ6L1 [11] leading-order parton distribution functions, using different event generators as listed below. The PYTHIA [10] event generator, has been used to generate the following processes:

- Diboson processes: Inclusive decay of WW, WZ and WZ.
- $WH \to \ell\nu b\bar{b}$ production ($\ell = e$, μ, or τ)
- $ZH \to \ell\ell b\bar{b}$ production ($\ell = e$, μ, or τ)

The following processes are simulated using other generators:

- $W+$ jets and $Z+$ jets events are generated with ALPGEN [12] (interfaced to PYTHIA for parton showering and hadronization) since ALPGEN yields a better description of processes with high jet multiplicity. ALPGEN samples have been produced using the MLM parton-jet matching prescription [12] and are generated in bins of light parton multiplicity. $W(Z)+$ jets samples contain $W(Z)jj$ and $W(Z)cj$ processes, while $W(Z)b\bar{b}$ and $W(Z)c\bar{c}$ are generated separately, also with ALPGEN.
- $t\bar{t}$ (lepton+jet and dilepton channels) production is generated with ALPGEN.
- Single-top events (s-channel (tb) and t-channel (tbq)) are generated using COMPHEP [13], and PYTHIA for the hadronization.

All the generated events were processed through the DØ detector simulation (based on GEANT [14]) and the same reconstruction software as for data.

Simulated events are reweighted for trigger efficiencies and efficiency ratios of data to simulation (scale factors), depending on the analyzed dataset. Depending on the efficiency considered these scale factors are constant or have a dependence on the event kinematics which is then taken into account.

The simulated background processes are absolutely normalized to the SM prediction of their cross section except for $W+$ jets which are normalized to data in the pre-tagged sample, where the signal contamination is expected to be negligible. The Data/MC normalization factor is $\sim 1.3\pm 0.15$ when comparing to the Next-to-Leading-Order (NLO) prediction for the $W+$ jet expectation, as determined using the program MCFM [15]. The normalization factor is derived after subtracting all other expected background processes from data, and its value greater than 1 is originating from higher order effects, mismodeling of the p_T spectrum of the W in ALPGEN, and potential instrumental effects (only the uncertainty on the latter is included in the quoted Data/MC uncertainty).

For the $Wb\bar{b}$ and $Wc\bar{c}$ samples we have determined an additional multiplicative K_{hf} factor of 1.4 ± 0.3 by adjusting the total number of background events to the data in the $W + 2$ jet sample with at least 1 b-tagged jets. This additional multiplicative factor is consistent with the ratio of the K-factors for $Wb\bar{b}$ and $W+$ light jets obtained from MCFM.

The same K-factors have been used for the corresponding $Z+$ jet processes.

IV. EVENT SELECTION

The analysis is based on the selection of events with exactly one electron with $p_T > 15\,\text{GeV}$ and detector $|\eta| < 1.1$ or $1.5 < |\eta| < 2.5$, or exactly one muon ($p_T > 15\,\text{GeV}$ and $|\eta| < 2.0$). Events are also required to have missing transverse energy $\not\!\!E_T >$ 20 GeV or 25 GeV in case of electron with $1.5 < |\eta| < 2.5$; two or three jets, with $p_T > 20$ GeV, (all after jet energy calibration corrections) and $|\eta| < 2.5$, with the further requirements that the leading jet has to have $p_T > 25$ GeV, and that the sum (H_T) of the p_T of the jets exceeds 60 GeV (90 GeV in the 3-jet sample).

Events with an additional muon or electron, isolated from jets and having a transverse momentum above $p_T > 20$ (> 15 GeV in the μ channel) GeV are rejected to decrease the Z and $t\bar{t}$ dilepton background. Only events having a primary z-vertex within $\pm 60\,\text{cm}$ of the nominal interaction point and at least three attached tracks are retained for analysis.

A. Lepton reconstruction and identification

The leptons used in the analysis are identified in two steps. 1) The lepton candidates are first required to pass the "loose" identification criteria, which for the electron are: energy fraction deposited in the EM calorimeter > 0.9, ratio of the energy in the hollow cone having external and internal radii of $R = 0.4$ and 0.2 around the electron candidate direction, divided by the candidate energy < 0.15, shower shape requirements, and matching of an EM cluster to a track having $p_T > 5$ GeV. For the muon, we require hits in each layer of the muon system, scintillator hits timing cuts to veto cosmics, matching between the muon hits and a central track, and isolation from jets to reject muons from semi-leptonic hadron decays. 2) The loose leptons then undergo a final, "tight" selection: tight electrons have to satisfy a likelihood test developed on $Z \to ee$ samples, that takes as input seven quantities sensitive to the EM nature of the particles [9]; tight muons must satisfy stricter isolation criteria requiring low calorimeter and tracking activity around the muon candidate. The inefficiencies induced by the lepton identification and isolation criteria are measured using $Z \to \ell\ell$ data samples.

The final selection uses only tight leptons, whilst the sample of loose leptons is used for instrumental and semi-leptonic background determination.

B. Instrumental and semi-leptonic background

The instrumental and semi-leptonic backgrounds, referred to as multijet background below, are estimated from the data. The instrumental background is important in the electron channel, where a jet with high EM fraction can pass the electron identification criteria, or since a photon can be misidentified as an electron. In the muon channel, the multijet background is due to events in which a muon from a semi-leptonic heavy quark decay is mis-identified as being isolated.

To estimate the number of events containing a jet passing the final electron identification criteria we determine the probability p_{tight}^{loose} for a loose electron candidate originating from a jet to pass the likelihood test. This is done on data, using the sample of events used for the analysis after the preselection requirements, i.e. with one loose lepton and two jets, but with low $\not\!\!E_T (< 10$ GeV), and in which one of the jets has an EM fraction smaller than 0.7, is in the central calorimeter ($|\eta| < 1.1$) and far from the calorimeter module boundaries. The probability p_{tight}^{loose} is obtained by dividing the number of events containing at least one electron candidate passing the likelihood test by the total number of events of the sample. This probability is determined as a function of the p_T of the candidate electron. We proceed similarly in the muon channel to determine the semi-leptonic background. We use the same selection criteria, but require a loose muon to be back-to-back in φ with one of the jets.

The multijet background is then estimated for every differential distribution: this p_T-dependent probability is used in the " matrix method" that we apply to our final sample and to the loose sample. This method allows for the derivation of the multijet background directly from data, once p_{tight}^{loose} and the efficiency to identify as tight true loose leptons are known [9]. The p_T distribution of the lepton in the $W + 2$ jet sample is shown in Fig.1a and compared to the expectation. The shape and magnitude of the distribution is well reproduced by the `ALPGEN` simulation of the W + jets processes, after adding the multijet background and other standard model (SM) backgrounds detailed in the previous section.

C. Missing E_T and Jet properties

To select W decays we require large missing transverse momentum, $\not\!\!E_T > 20$ GeV. $\not\!\!E_T$ is calculated from the calorimetric cells except unclustered cells in the coarse hadronic layers and is corrected for the presence of any muons. All energy corrections to electrons or to jets are propagated into $\not\!\!E_T$.

The transverse mass $m_T = \sqrt{2p_T^\ell p_T^\nu (1 - \cos(\varphi_\ell - \varphi_\nu))}$ of the W boson can be reconstructed from the charged lepton and neutrino (ν) kinematics quantities, in which the neutrino transverse momentum is approximated by the missing transverse energy. Its distribution is shown in Fig.1b. The distributions of H_T and missing transverse energy are shown in Fig.1c,d and compared to the expectation.

Jets are reconstructed using a midpoint cone algorithm [16] with a radius of $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta\varphi)^2} = 0.5$, where y is the jet rapidity. Identification requirements ensure that the jet energy distribution in the various layers of the calorimeter is reasonable and that the jets are not due to spurious energy deposits. The difference in efficiency of the jet identification requirement and of the jet resolution between data and simulation is taken into account in the overall jet reconstruction efficiency scale factor. Comparison of `ALPGEN` with other generators and with data show discrepancies in distributions of the jet pseudorapidity and dijet angular separation distributions [17] The data are

	W + 2 jets	W + 2 jets (1 b tag)	W + 2 jets (2 b tag)	W + 3 jets	W + 3 jets (1 b tag)	W + 3 jets (2 b tag)
WH, ZH	15.8 ± 2.7	6.8 ± 1.3	3.9 ± 0.7	3.8 ± 0.6	1.6 ± 0.3	1.0 ± 0.2
WW, WZ, ZZ	1453 ± 244	87 ± 16	13.7 ± 2.6	302 ± 51	23.1 ± 4.2	3.9 ± 0.7
$W/Z + b\bar{b}$	1769 ± 353	592 ± 109	138 ± 27	471 ± 94	174 ± 32	49.2 ± 9.4
$t\bar{t}$	581 ± 98	242 ± 45	96.9 ± 19	926 ± 155	394 ± 73	211 ± 41
Single top	290 ± 49	123 ± 23	31.5 ± 6.1	91 ± 15	38.5 ± 7.0	16.9 ± 3.2
Multijet	3575 ± 629	189 ± 38	16.7 ± 4.0	1228 ± 216	92.8 ± 18	14.4 ± 3.4
W/Z+ jets	44464 ± 570	942 ± 226	44.5 ± 9.8	8357 ± 105	239 ± 57	25.7 ± 5.8
Total expectation	52148 (n.t.d.)	2182 ± 348	345 ± 51	11379 (n.t.d.)	963 ± 152	322 ± 49
Observed Events	52148	2174	336	11379	912	321

TABLE I: *Summary table for the W+ 2,3 jet final states. Observed events in data are compared to the expected number of W+ jet events before tagging, with exactly one tight b-tagged jet, and with exactly 2 loose b-tagged jets. First three columns are for the $W + 2$ jet channel, the last three columns for the $W + 3$ jet channel. Expectation originates from the simulation of WH and ZH (with $m_H = 115$ GeV), dibosons (WW, WZ, ZZ, labeled WZ in the table), $Wb\bar{b}$ production, top production ($t\bar{t}$ and single-top), multijet background and "W+ jet" production, which contains light and c quarks. All Z processes are fully simulated, and included in the corresponding W categories. The processes $W(Z)b\bar{b}$ and $W(Z)$ + light and/or c jets are counted separately. "n.t.d." stands for "normalized to data". The uncertainties given include statistics and systematics.*

therefore used to correct these quantities in the ALPGEN W+ jets and Z+ jets MC samples: the simulated η, $\Delta\eta$ and $\Delta\varphi$ for the two leading jets in the W/Z+ jets samples are reweighted with simple polynomial functions to provide agreement of the total background with the data of the high statistics pre-tagged sample.

We have studied standard kinematic distributions and for example the p_T distributions of the leading jet and next to leading jet in $W + 2$ jet events are shown in Fig. 2a and b. The distribution of the $\eta - \varphi$ distance ΔR between the two jets in the $W + 2$ jet events is shown in Fig. 2c. The dijet invariant mass is shown in Fig. 2d. The shape of the distributions are described over the complete kinematic range for all jets.

V. b-TAGGING RESULTS

Since we search for WH production with two b-jets in the final state, efficient identification of b-jets in our selected events is a central part of the analysis. The DØ neural network (NN) b-tagging algorithm has been used for tagging heavy flavored jets. It is based on the combination of seven variables sensitive to the presence of tracks or secondary vertices significantly displaced from the primary vertex. All tagging and mistagging efficiencies have been determined from data and dedicated simulated samples.

We start with a "loose" NN operating point, which corresponds to a fake rate, i.e. the fraction of "light" partons (u, d, s, g) mistakenly tagged as heavy-flavored jets by the tagger, of about 1.5% for a jet p_T of 50 GeV. If two jets are tagged the event is selected as double-tagged. Otherwise the operating point is tightened to a value corresponding to a fake rate of about 0.5%, and the event can then be selected as an "exclusive" single b-tag, simply called single b-tag in the following. We are thus left with two orthogonal samples, one "loose" double-tag (DT) and one "tight" single-tag (ST). This simplifies the combination of the results of the Higgs search in the two samples which is performed in order to improve the sensitivity to a potential signal. The operating points have been selected based on the optimal combined sensitivity to a WH signal.

The efficiencies for identifying a jet containing a b hadron of the loose and tight operating points are about $59 \pm 1\%$ and $48 \pm 1\%$, respectively, for a jet p_T of 50 GeV. The efficiency has been determined relative to taggable jets, i.e jets having at least 2 good quality tracks, of which one has $p_T > 1$ GeV and another $p_T > 0.5$ GeV, and a sufficient number of SMT hits. The jet taggability is typically 80% in a two jet sample with an uncertainty of 3% per jet. Simulated events are corrected to have the same fraction of jets satisfying the taggability and b-tagging requirements as in data. The systematic uncertainty on this scale factor is 4-7% for heavy quarks (b, c) and 25% when mis-tagging "light" partons.

The distributions of the p_T of the first and second jet, of ΔR and of the dijet mass for the $W + 2$ jet sample in which exactly one jet is b-tagged are shown in Figs. 3a–d. The observed agreement in both cases indicates that simulation, which includes the different standard model processes, describes the data well. The same distributions are shown for the $W + 2$ jet sample in which the two jets are b-tagged, in Figs. 4a–d.

The number of observed events are well described by the Monte Carlo simulation as detailed in Table I.

VI. NEURAL NETWORK SELECTION

To optimize the sensitivity of the search in $W + 2$ jet events, differences between the kinematic properties of the objects in the event (lepton, jets, $\not\!\!E_T$) are exploited using a neural network trained to separate Higgs signal events from background events. The neural network uses the following seven kinematic variables as inputs: $p_{\rm T}$(leading jet), $p_{\rm T}(2^{nd}$ jet), ΔR(jets), $\Delta\phi$(jets), $p_{\rm T}$(dijet system), dijet invariant mass, $p_{\rm T}(\ell$-$\not\!\!E_T$ system), together with a matrix element discriminant.

The most discriminant kinematic variable is the dijet invariant mass (shown in Figs. 5 and 6 for the different 1 b-tag and 2 b-tag samples), so it is directly used as the final discriminant variable in the $W+$ 3 jet sample.

The matrix-element discriminant uses the 4-vectors of the lepton and of the two jets, integrating over the unmeasured momentum of the neutrino and convoluting with the resolution function of the detector to calculate the relative probability for each event to come from a WH decay or from the backgrounds, using the leading order matrix elements for the processes considered. The distributions of the ME discriminant for m_H=115 GeV are shown in Figs. 7a–d. After having verified that the distribution of the ME discriminant is well reproduced by the expectation, we then use it as an input to the neural network.

The training of the neural network is performed using simulated samples of WH and Wbb events, for each of eight channels: (e, μ), (ST, DT), (Run IIa and b). Separate networks are trained and used for the different Higgs masses considered in the analysis.

Distributions of the neural network output trained on the $m_H = 115$ GeV sample, after all selection cuts, are shown in Figs. 8a–d in the pre-tag, exclusive single-tag and double-tag samples. They are well described by the simulation. In Fig. 9a–d, the same distributions are shown with a logarithmic vertical scale and the contribution of WH is indeed peaking at high values of the neural network output, contrarily to the background. The neural network output is then fitted to separate enhance the signal over the background as explained in section VIII. The gain in sensitivity compared to the analysis using the dijet mass as discriminant (measured in terms of the final cross section expected limit) using this matrix-element + neural network approach is of the order of 20% depending on the Higgs mass, equivalent to a gain of about 40% in integrated luminosity.

VII. SYSTEMATIC UNCERTAINTIES

The experimental systematic uncertainties due to efficiencies (i.e. the uncertainty on the ratio data/simulation of the efficiencies) or to the propagation of other systematic uncertainties (trigger, energy calibration, resolution), which affect the signal and standard model backgrounds (QCD background excepted) are summarized as follows:

- 3-5% uncertainty for the trigger efficiency derived from the data sample used in this analysis;
  5-6% uncertainty for the lepton identification and reconstruction efficiency.
- 2-6% uncertainty on the acceptance due to the jet identification and jet energy calibration and resolution uncertainty. 5% for the acceptance uncertainty due to jet modelling (fragmentation).
- 3% for the jet taggability, 2-5% for the b-tagging efficiency, per heavy quark jet. For the light quark jets the uncertainty is 25%; this translates into an uncertainty on the total background of the exclusive single-tag sample of 7% (negligible for double-tag).
- An uncertainty ($\sim$ 10%) on the neural network output, originating from the modeling uncertainty of the $W+$ jets background. This uncertainty is obtained by comparing the original distribution to the one obtained after "reweighting" the simulated events (as described above) in the pre-tag sample, in which the signal is expected to be negligible. An additional uncertainty (5-10%) on the shape of the $Wb\bar{b}$ dijet invariant mass distribution, based on the Data/MC difference in the pre-tag W+jet sample (since there is not enough statistics to determine it from the tagged samples) is also added.

Overall, the experimental systematic uncertainty on the acceptance varies between 16 and 28% depending on the process and the channel ($\sim$ 18% for WH in the DT channel).

The uncertainty on the cross sections of the background processes is 11% for $t\bar{t}$ production and single-top production, 6% for diboson production. The uncertainty on $W+$ heavy flavor production is 20%.

VIII. WH CROSS SECTION LIMIT

The expected contribution from the $b\bar{b}$ decay of a standard model Higgs boson of 115 GeV, produced in association with a W boson is shown in Figs. 5a, b for the ST W+2, 3 jet channel, and in Figs. 5c, d for the DT W+2, 3 jet channels. The same distributions are also shown in log-scale, in Figs. 6a–d.

As no excess is observed compared to the expectation, we proceed to set limits using the neural network output of the $W + 2$ jet events and the dijet mass of the $W + 3$ jet events. Each channel is analyzed independently. Limits are derived from the sixteen individual analyses (e, μ, ST, DT) done with the $W+$ 2/3 jet events in the two different data taking periods (Run IIa and b), and then combined.

Limits are calculated at 95% confidence level using the modified Frequentist CL_s approach with a Poisson log-likelihood ratio test statistic [18, 19]. The impact of systematic uncertainties is taken into account, and all correlations in systematic uncertainties are maintained amongst channels and between signal and background. The expected distributions for background are evaluated by minimizing a profile likelihood function, which takes into account the shape and rate of the observed distributions in the sideband regions.

The log-likelihood ratio (LLR) distributions for the $WH \to \ell\nu b\bar{b}$, (i.e. after combining the sixteen individual channels) is shown in Fig. 10a. Included in the figure are the LLR values for the signal-plus-background hypothesis (LLR_{s+b}), background-only hypothesis (LLR_b), and the observed data (LLR_{obs}). The shaded bands represent the 1 and 2 standard deviation departures for LLR_b. The cross section limit obtained for $\sigma(p\bar{p} \to WH) \times B(H \to b\bar{b})$ is 6.7 times above the SM expectation, at 95% C.L. for a Higgs boson mass of 115 GeV, The corresponding expected upper limit normalized to the SM expectation is 6.4. The same study is performed for ten other Higgs mass points between 100 and 150 GeV. The corresponding observed and expected 95% CL Limits/SM are given in Table II. These 95% CL Limits/SM are shown in Fig. 10b.

m_H (GeV)	expected 95% CL Limit/SM	observed 95% CL Limit/SM
100	5.3	5.2
105	4.9	4.2
110	5.8	5.1
115	6.4	6.7
120	7.5	8.2
125	9.5	9.8
130	13.7	16.7
135	16.1	17.3
140	23.0	23.3
145	36.1	43.7
150	56.0	52.4

TABLE II: *95% C.L. expected and observed limits on $\sigma(p\bar{p} \to WH) \times B(H \to b\bar{b})$ divided by the corresponding SM expectation, as a function of the Higgs mass.*

The improvement in sensitivity obtained with the current analysis which combines extended acceptance/luminosity and combined multivariate techniques is significant, in particular in the region where we have best sensitivity for low Higgs mass discovery, i.e. 115–135 GeV, with a ratio of observed (expected) limit to the SM cross section of 6.4 (6.7) for $m_H = 115$ GeV.

IX. SUMMARY

The $\ell+ \not\!\!E_T$ + 2 or 3 jets final states have been analyzed in the search for WH production in 2.7 fb^{-1} of data collected between April 2002 and April 2008, split in two periods separated by the Spring '06 Tevatron shutdown. During the Tevatron shutdown DØ's triggering and b-tagging capabilities were upgraded.

The production rate of the single and double b-tagged events is in agreement with the expected standard model cross sections, within statistical and systematic errors, both in $W + 2$ and $W + 3$ jet events.

To search for a low mass Higgs boson, we have combined all channels (e, μ, ST,DT) and derived limits from the neural network discriminant distribution in $W + 2$ jet events and from the dijet mass in the $W + 3$ jet events, using the CL_S method. We set upper cross section limits on $\sigma(p\bar{p} \to WH) \times B(H \to b\bar{b})$ divided by the SM expectation

between 4 and 10 at 95% C.L. (5 to 10 for the corresponding expected limits) for Higgs masses between 100 and 125 GeV, with the ratio increasing up to ~ 50 with m_H increasing up to 150 GeV. For m_H =115 GeV, the observed (expected) 95% CL Limit/SM is 6.7 (6.4).

Acknowledgments

We thank the staff at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à L'Energie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry for Science and Technology and Ministry for Atomic Energy (Russia), CAPES, CNPq and FAPERJ (Brazil), Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Natural Sciences and Engineering Research Council and WestGrid Project (Canada), BMBF (Germany), A.P. Sloan Foundation, Civilian Research and Development Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

[1] D0 Collaboration, V.M. Abazov *et al.*, Phys. Rev. Lett. 94, 091802 (2005).
[2] D0 Collaboration, V.M. Abazov *et al.*, Phys. Lett. B 663, 26 (2008).
[3] D0 Collaboration, V.M. Abazov *et al.*, Accepted in Phys. Rev. Lett., arXiv:0809.5265 (2008).
[4] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. 94, 091802 (2005).
[5] CDF Collaboration, T. Aaltonen *et al.*, Phys. Rev. Lett. 100, 041801 (2008).
[6] DØ Collaboration, V. Abazov *et al.*, Nucl. Instrum. Meth., Res. A 565, 463 (2006).
[7] S. Abachi, *et al.*, Nucl. Instrum. Methods Phys. Res. A 338, 185 (1994).
[8] T. Andeen, *et al.*, FERMILAB-TM-2365, April 2007
[9] DØ Collaboration, V. Abazov *et al.*, Phys. Rev. D 76, 092007 (2007)
[10] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. 135, (2001) 238, versions 6.319, 6.323 and 6.409.
[11] H. L. Lai et al., Phys.Rev. D55 (1997) 1280, `http://hep.pa.msu.edu/people/wkt/cteq6/cteq6pdf.html`
J. Pumplin et al., JHEP07(2002)012, `arXiv:hep-ph/0201195v3`.
[12] M. Mangano et al.: `ALPGEN` hep-ph/0206293, JHEP07 (2003) 001, version 2.05, `http://mlm.web.cern.ch/mlm/alpgen`
[13] A. Pukhov *et al.*, `COMPHEP`, hep-ph/9908288 (1999)
[14] R. Brun and F. Carminati, CERN Program library Long Writeup, Report W5013 (1993).
[15] J. Campbell and K. Ellis, `MCFM`, http://mcfm.fnal.gov/
[16] G. Blazey *et al.*, in *Proceedings of the workshop "QCD and Weak Boson Physics in Run II"* edited by U. Baur, R.K. Ellis, and D. Zeppenfeld, Batavia (2000), p. 47. arXiv:hep-ex/0005012 (2000).
[17] J. Alwall et al., Eur. Phys. C 53, 473 (2008).
[18] T. Junk, Nucl. Instrum. Meth. A434, p. 435-443, 1999.
[19] W. Fisher, DØ-Notes 4975 and 5309.

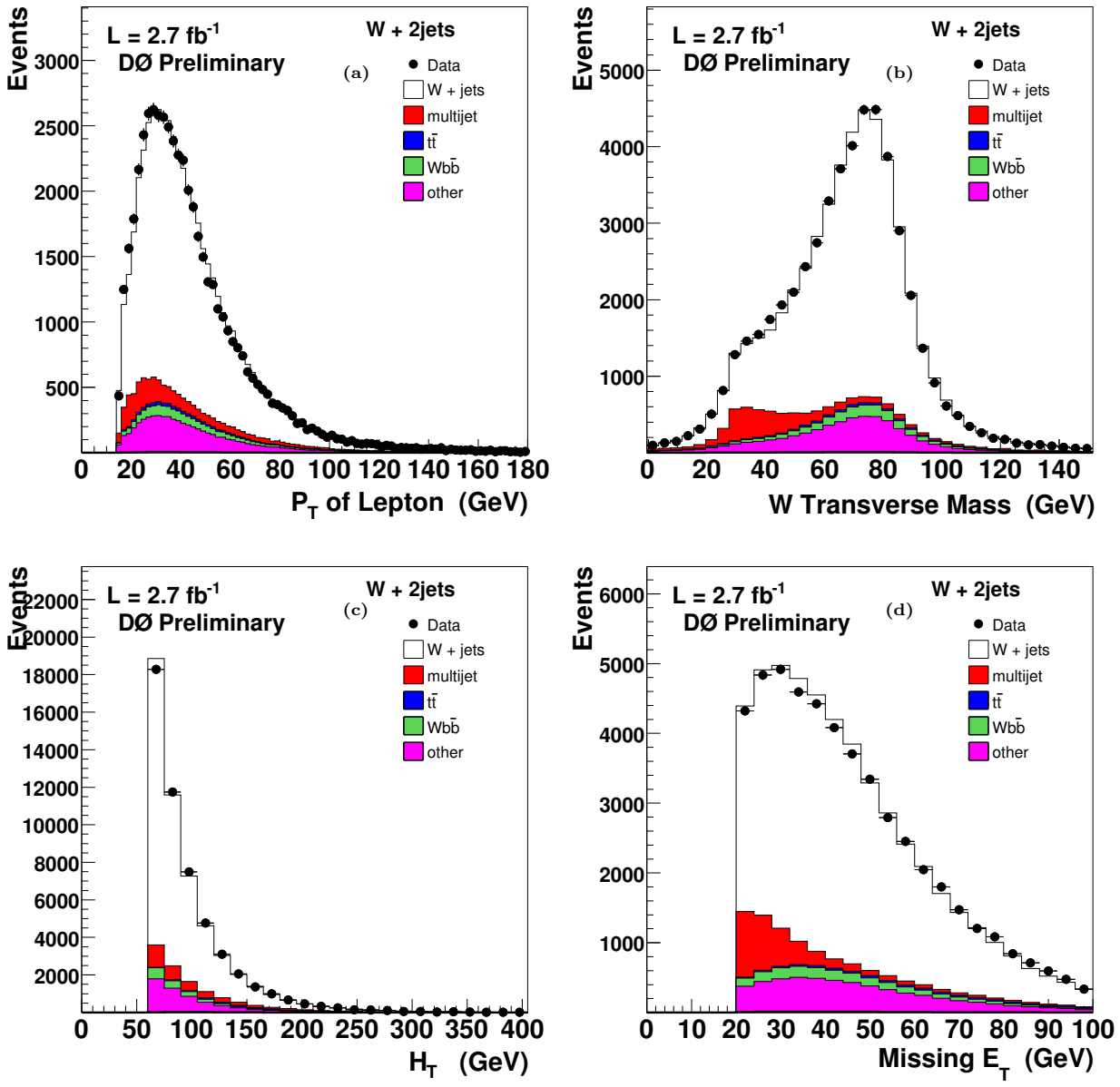

FIG. 1: *Distribution in the W + 2 jets sample of the (a) lepton momentum, (b) the transverse W mass, (c) the H_T variable and (d) missing transverse energy compared to the simulated expectation in the $W + 2$ jet event sample. The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+$ jets sample which is normalized on the "untagged sample" to the data, taking into account all the other backgrounds.*

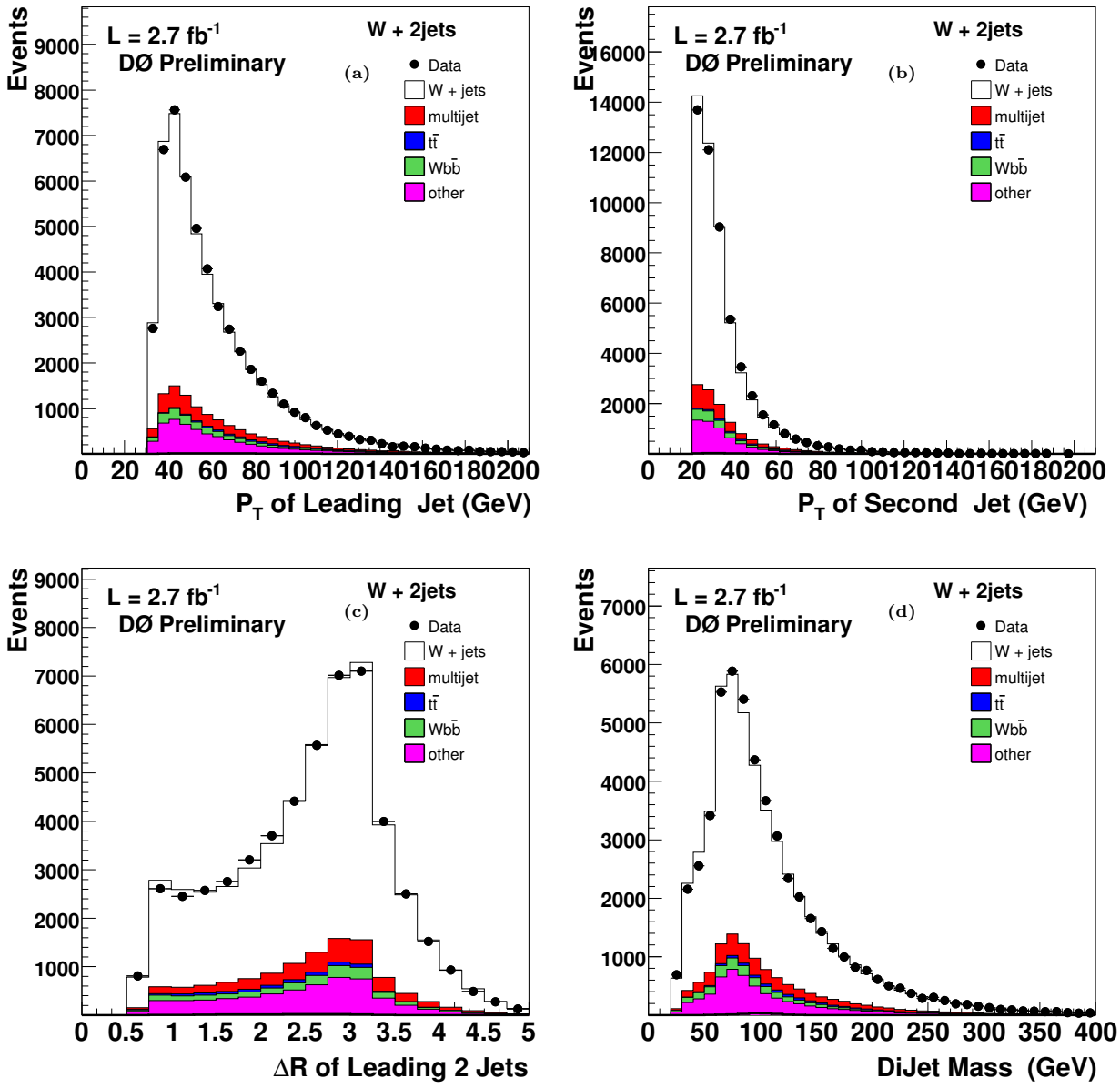

FIG. 2: *Distribution in the W + 2 jets sample of the (a) p_T of the leading and (b) next to leading jet, (c) of the distance in the $\eta - \varphi$ plane between the two jets and (d) of the dijet mass (d) between the two jets in the $W + 2$ jet sample compared with the simulated expectation. The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the W+ jets sample which is normalized on the "untagged sample" to the data, taking into account all the other backgrounds.*

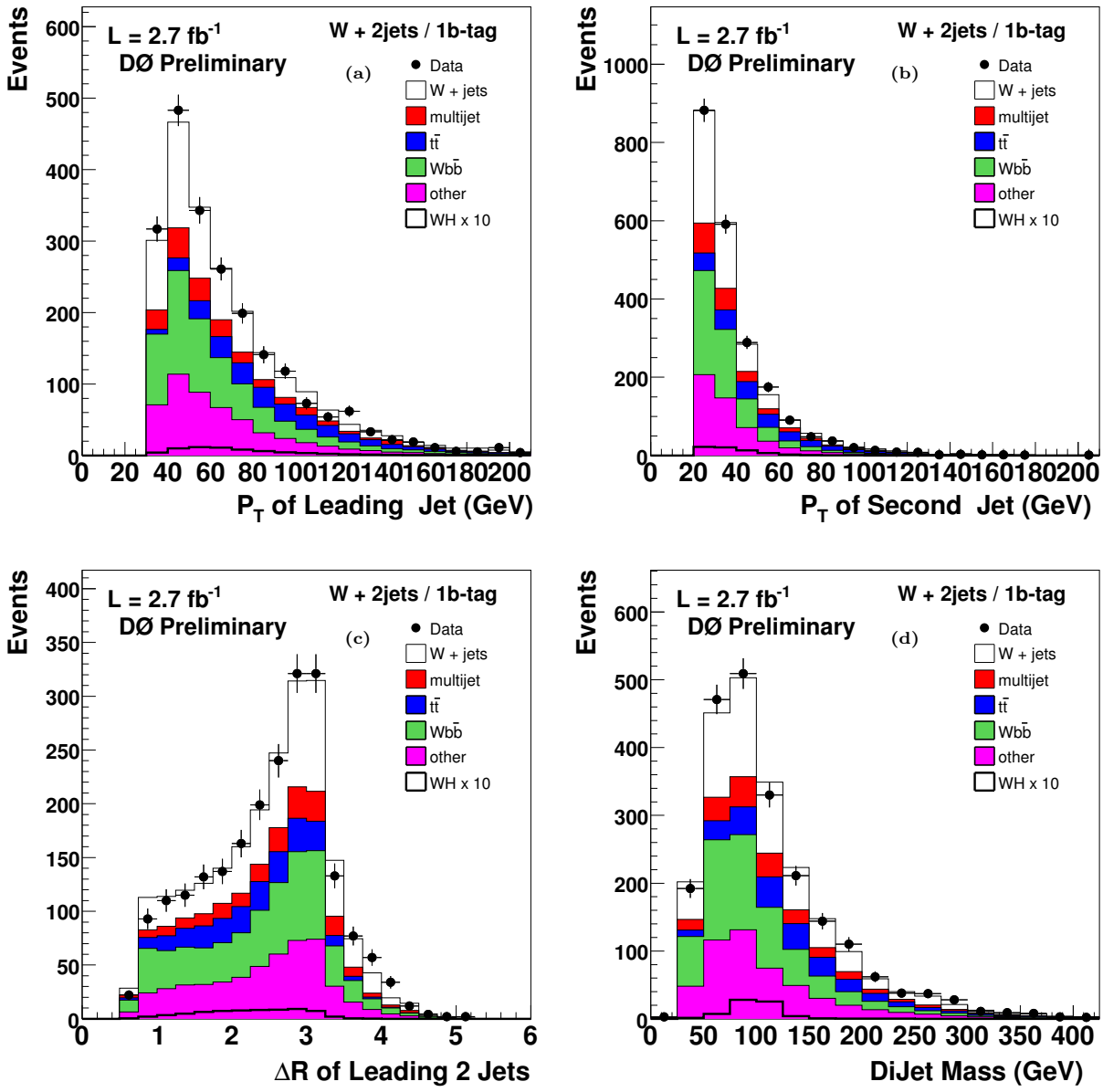

FIG. 3: *Distribution in the W + 2 jets sample with one b-tagged jet of (a) the p_T of the leading and (b) next to leading jet, (c) of the distance in the $\eta - \varphi$ plane between the two jets and (d) of the dijet mass between the two jets in the $W + 2$ jet sample compared with the simulated expectation. The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+$ jets sample which is normalized on the "untagged sample" to the data, taking into account all the other backgrounds. Also shown is the contribution expected for standard model WH production with $m_H = 115$ GeV, multiplied by a factor 10.*

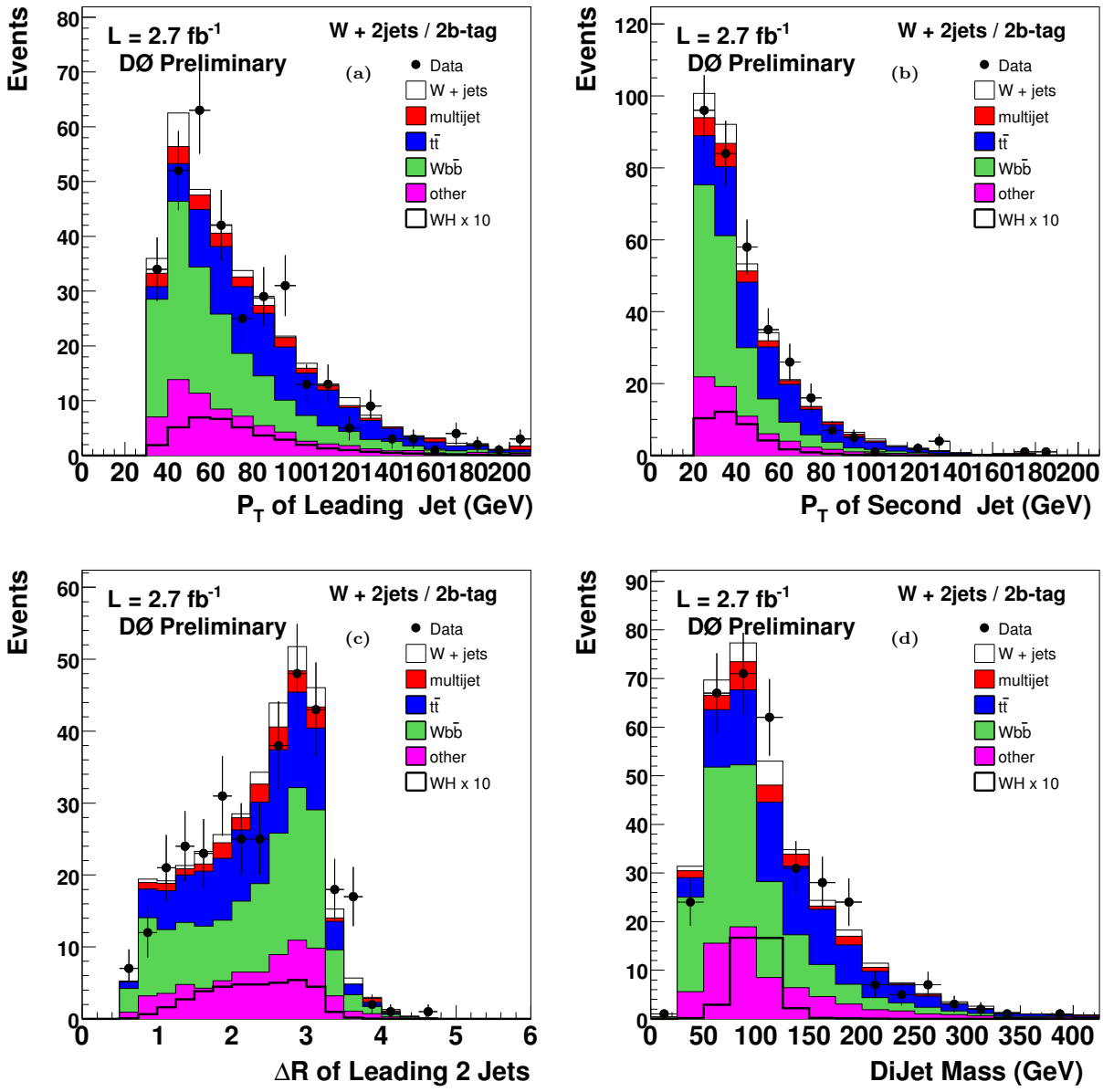

FIG. 4: *Distribution in the W + 2 jets sample with two b-tagged jets of the p_T of the (a) leading and (b) next to leading jet, (c) of the distance in the $\eta - \varphi$ plane between the two jets and (d) of the dijet mass between the two jets in the $W + 2$ jet sample compared with the simulated expectation. The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+$ jets sample which is normalized on the "untagged sample" to the data, taking into account all the other backgrounds. Also shown is the contribution expected for standard model WH production with $m_H = 115$ GeV, multiplied by a factor 10.*

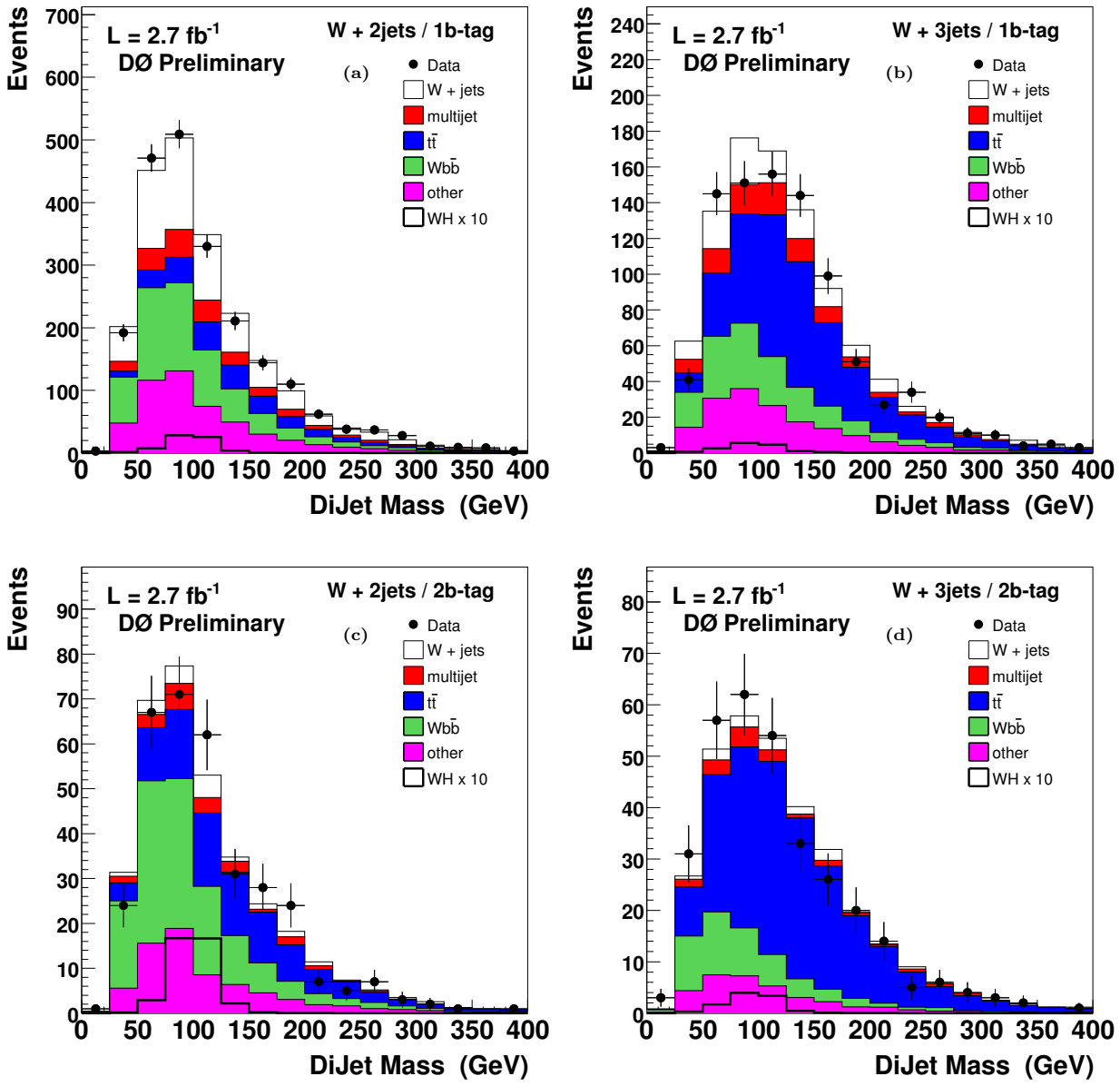

FIG. 5: *a) (b) dijet invariant mass in $W + 2(3)$ jet events when exactly one jet is b-tagged. c) (d) same distributions when at least 2 jets are b-tagged. Linear scale. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+$ jets sample which is normalized on the "untagged sample" to the data, taking into account all the other backgrounds. The backgrounds labelled as "other" in the figure are dominated by single-top production. Also shown is the contribution expected for standard model WH production with $m_H = 115$ GeV, multiplied by a factor 10.*

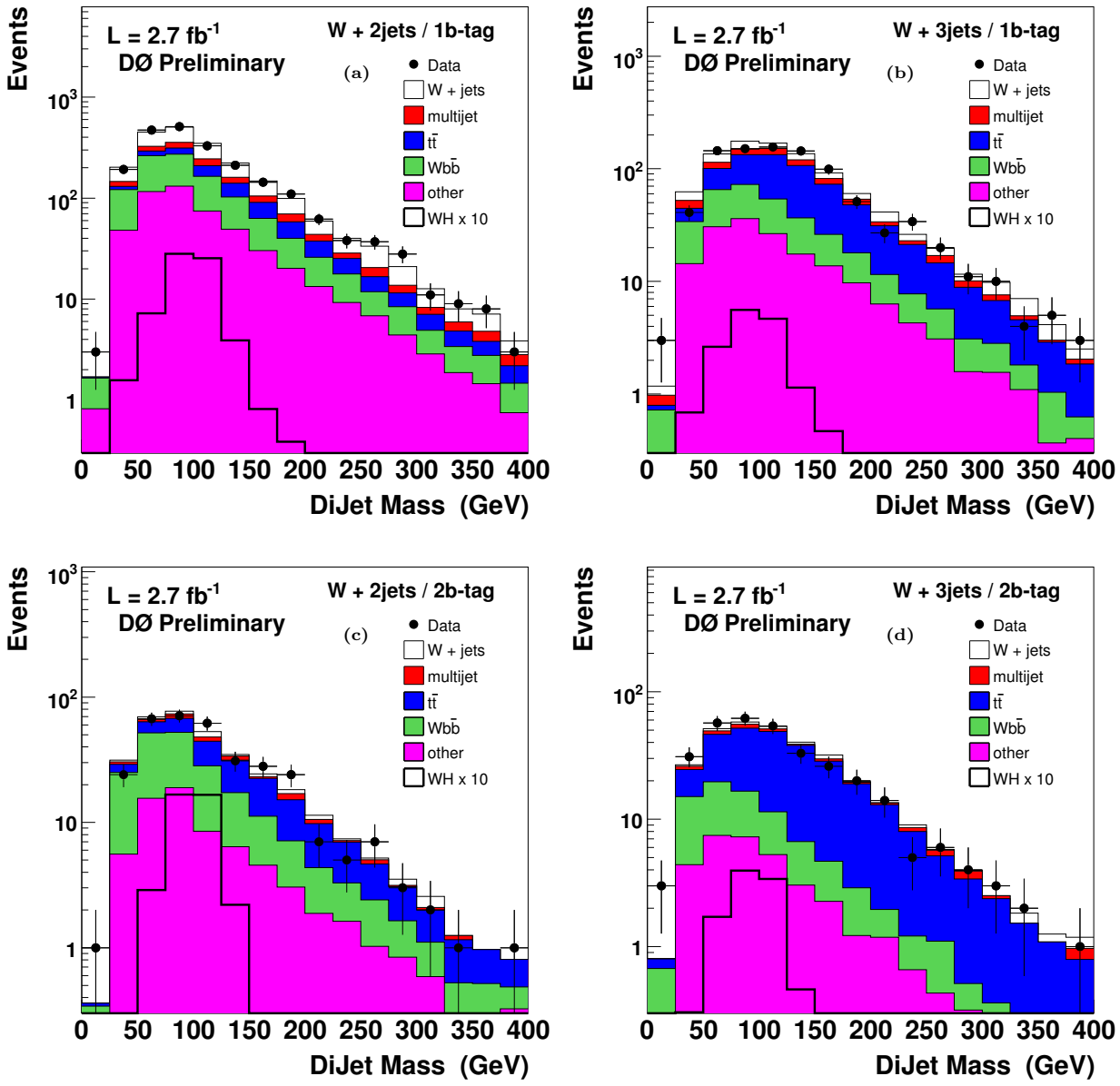

FIG. 6: *a) (b) dijet invariant mass in $W + 2(3)$ jet events when exactly one jet is b-tagged. c) (d) same distributions when at least 2 jets are b-tagged. Logarithmic scale. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+$ jets sample which is normalized on the "untagged sample" to the data, taking into account all the other backgrounds. The backgrounds labelled as "other" in the figure are dominated by single-top production. Also shown is the contribution expected for standard model WH production with $m_H = 115$ GeV, multiplied by a factor 10.*

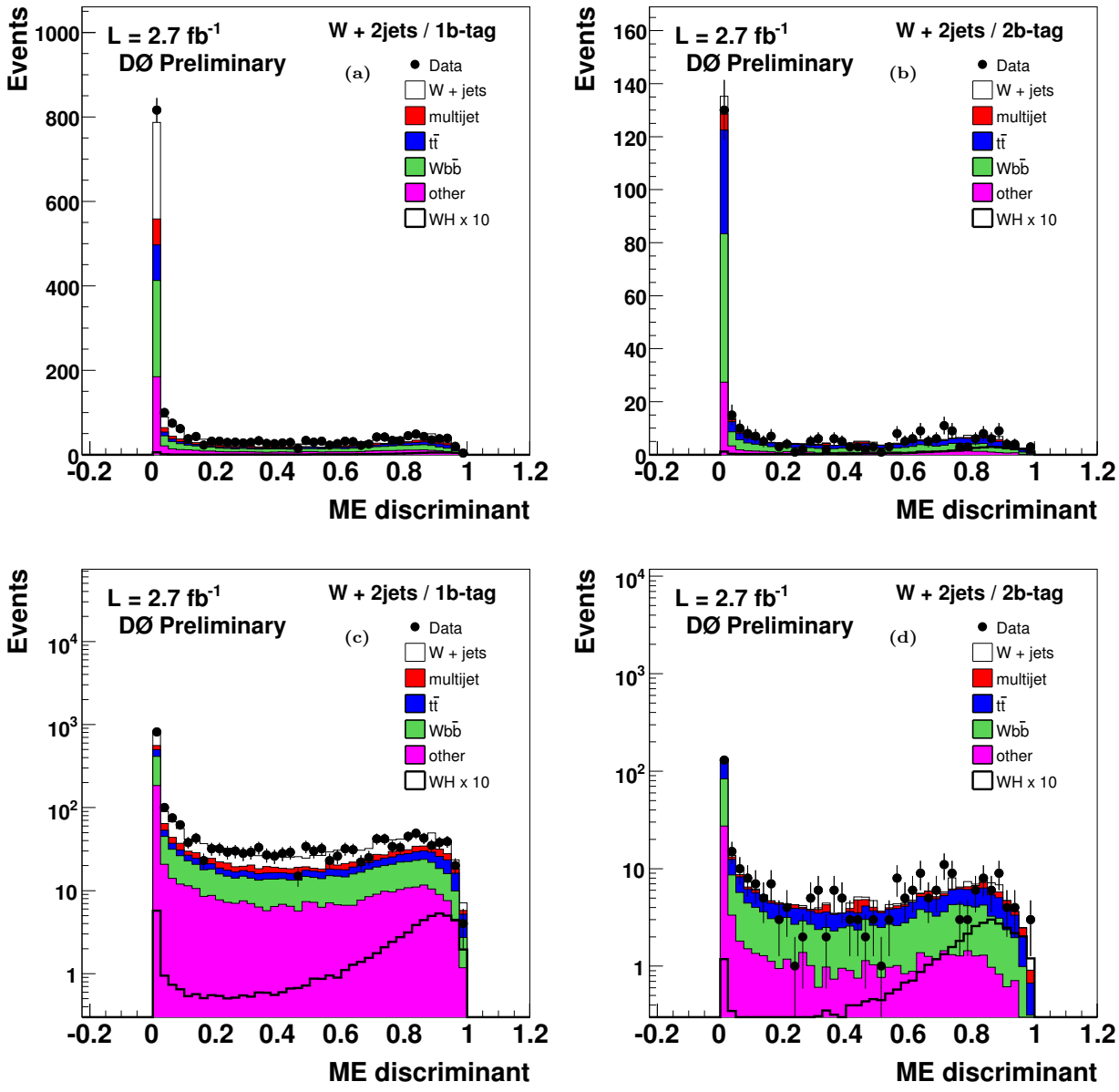

FIG. 7: *Distributions of the Matrix Element discriminant compared with the simulated expectation: a) in the single b−tag sample for the single-tag ME discriminant ; b) in the double b−tag sample for the double-tag ME discriminant ; c) log scale of figure (a); d) log scale of figure (b); The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+$ jets sample which is normalized on the "pre-tag sample" to the data, taking into account all the other backgrounds. The WH expected contribution which is scaled by a factor 10 is peaking at high values of the ME discriminant as shown in c) and d).*

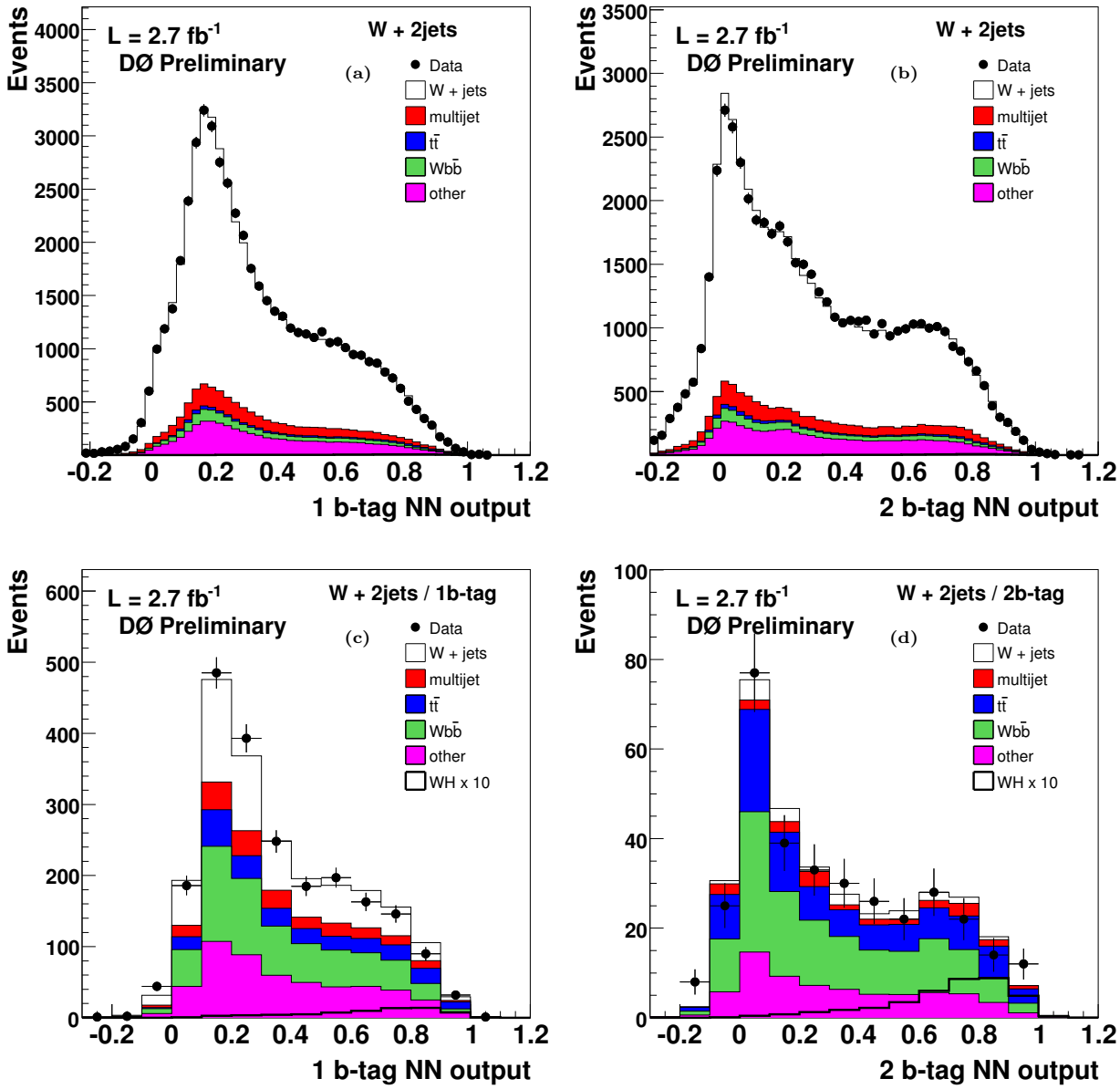

FIG. 8: *Distributions (with linear vertical scale) of the neural network output compared with the simulated expectation: a) before b-tagging for the single-tag NN ; b) before b-tagging for the double-tag NN ; c)in the single b−tag sample for the single-tag NN ; d) in the double b−tag sample for the double-tag NN ; The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the W+ jets sample which is normalized on the "pre-tag sample" to the data, taking into account all the other backgrounds. The WH expected contribution for* $m_H = 115$ *GeV, multiplied by a factor 10, is peaking at high values of the NN output as shown in c) and d).*

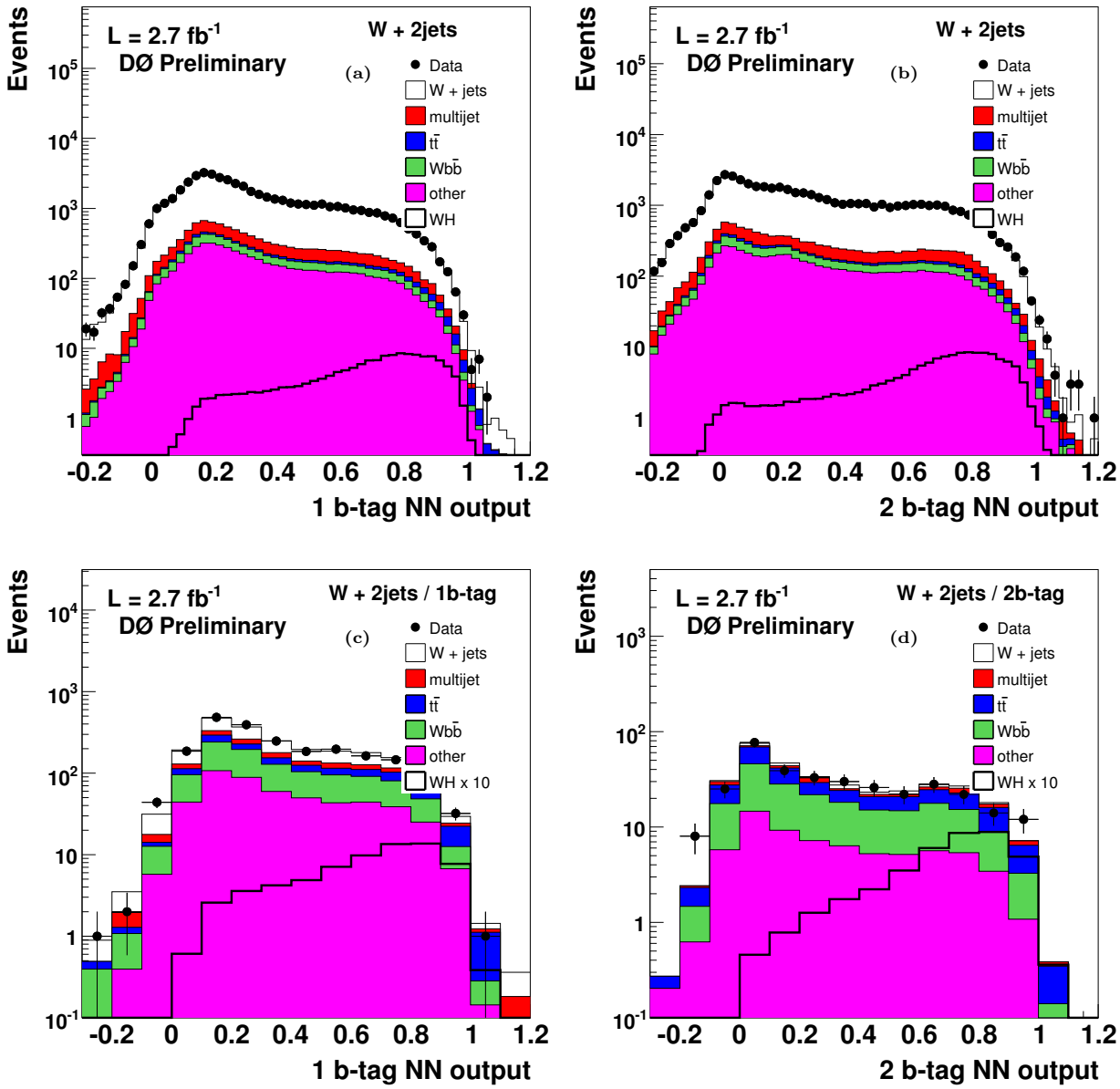

FIG. 9: *Distributions (with logarithmic vertical scale) of the neural network output compared with the simulated expectation: a) before b-tagging for the single-tag NN ; b) before b-tagging for the double-tag NN ; c)in the single b−tag sample for the single-tag NN ; d) in the double b−tag sample for the double-tag NN ; The simulation is normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the W+ jets sample which is normalized on the "pre-tag sample" to the data, taking into account all the other backgrounds. The WH expected contribution for* $m_H = 115$ *GeV, multiplied by a factor 10, is peaking at high values of the NN output as shown in c and d).*

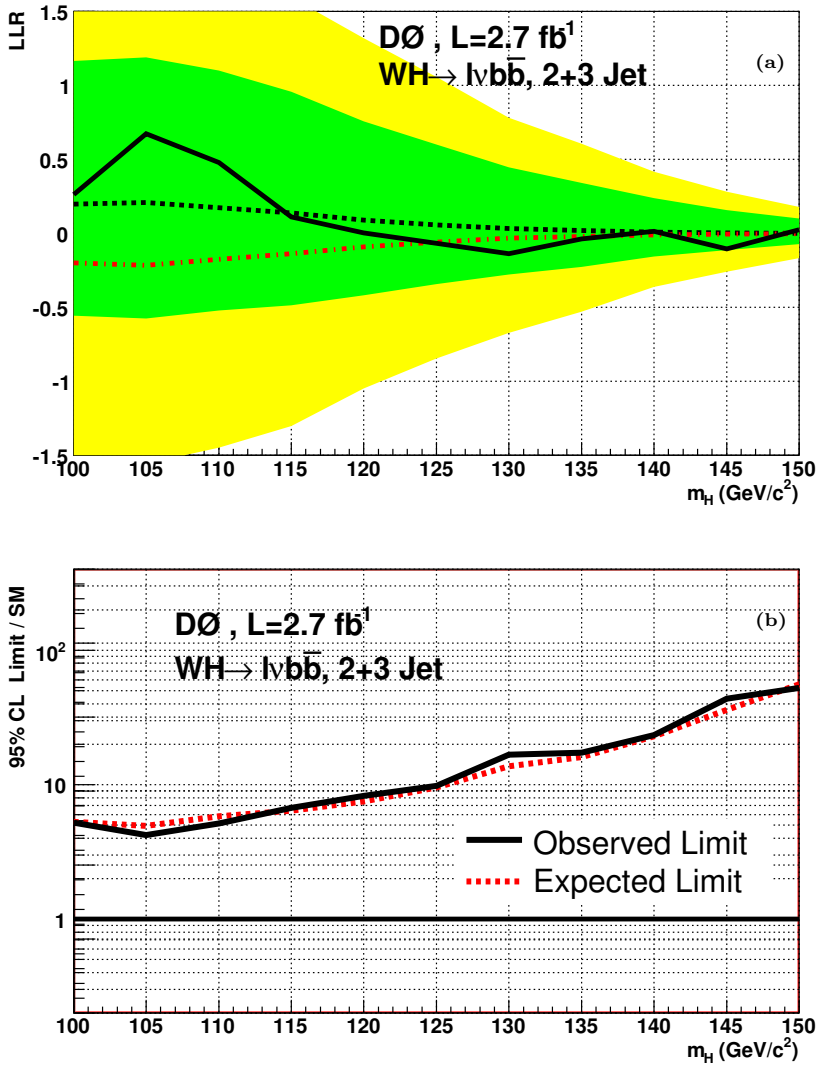

FIG. 10: *a) Log-likelihood ratio distribution for the $WH \to \ell\nu b\bar{b}$ 2+3 jet channel. The dashed lines represent LLR_b (upper line) LLR_{s+b} (lower line), the full line represents LLR_{obs}; for LLR_b, 1 and 2 standard deviation bands are drawn, cf section 8. b) Expected (median) and observed 95% C.L. upper limits on the cross section ratios for the combined $WH \to e, \mu\nu b\bar{b}$ $W+2$ jet analyses (single- and double-tag, CC electrons + muons) in the $m_H = 105 - 145$ GeV mass range. The solid line at $y = 1$ indicates the 95% C.L. exclusion of the standard model expectation.*