DØnote 5503-CONF

Search for Neutral Higgs Bosons at High $\tan\beta$ in multi-jet Events

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: July 25, 2006)

The full RunIIa data sample recorded at DØ has been analyzed to search for Neutral Higgs bosons produced in association with b-quarks at high $\tan\beta$ within the MSSM framework. The search has been performed in the three b-quarks channel using multi-jet triggered events corresponding to an integrated luminosity of ~ 0.9 fb^{-1}. No excess of events with respect to the predicted background is observed in the final selected three b-tag sample, so limits are set in the MSSM parameter space.

Preliminary Results for Summer 2006 Conferences

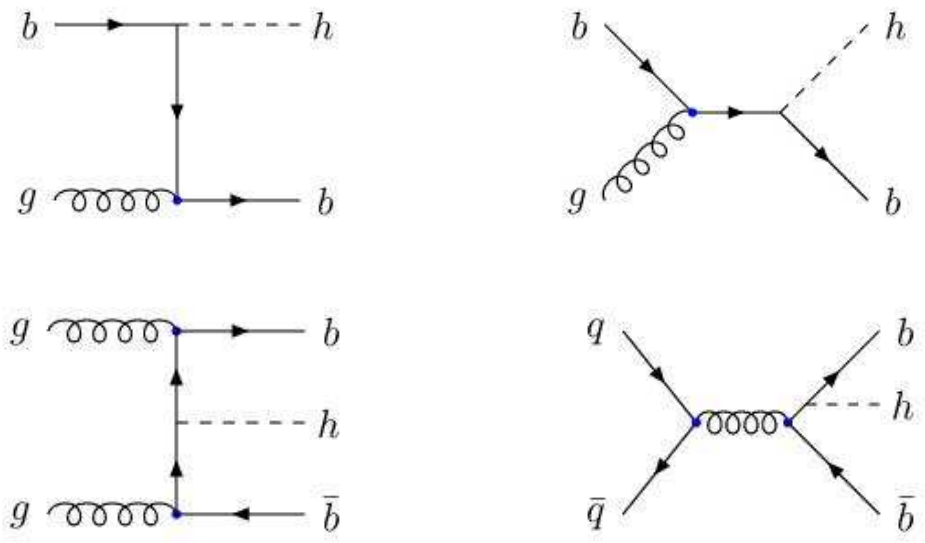

FIG. 1: Leading order Feynman diagrams for neutral Higgs boson production in the five-flavor scheme (top) and four-flavor scheme (bottom).

I. INTRODUCTION

In two-Higgs-doublet models of electroweak symmetry breaking, such as the minimal supersymmetric extension of the standard model (MSSM), there are five physical Higgs bosons : two neutral *CP*-even scalars, h and H, H being the heavier one; a neutral *CP*-odd state A; and two charged states H^+ and H^-. The coupling of the A boson to the down quarks, such as the b quark, is enhanced by a factor of $\tan\beta$ compared to the Standard Model (SM) one, where $\tan\beta$ is defined as the ratio of the vacuum expectation values of the two Higgs doublets. At high $\tan\beta$, this is also true for either h or H. Depending on their mass, h and A bosons or H and A bosons or even h, H and A have a degenerated mass in the high $\tan\beta$ limit. Thus, in large $\tan\beta$ scenarii, the Higgs bosons production in association with b quarks is enhanced by a factor $2 \times \tan^2\beta$ compared to the SM one. The Higgs decays are also dominated by the $b\overline{b}$ production.

For several representative scenarii of the MSSM, LEP experiments have excluded a light Higgs boson with $m_h < 93$ GeV/c^2 at the 95 % confidence level [1]. The CDF experiment at the Tevatron Collider performed a search for Higgs bosons produced in association with b quarks in data from Run I [2]. MSSM Higgs bosons reconstructed in the channel $\tau^+\tau^-$ has also been searched for by both the CDF experiment [3] and the DØ experiment [4].

In the analysis presented here, we search for the production of Higgs bosons, decaying into a $b\overline{b}$ pair, in association with one or two b quarks in $p\overline{p}$ collisions with a center of mass energy of 1.96 TeV. Such a production can be described by two different theoretical approaches : the five-flavor scheme [5] and the four-flavor scheme [6]. Both calculations agree within their uncertainties [7, 8]. The corresponding Feynman diagrams are shown on Figure 1 in the h production case. Similar diagrams hold for A and H productions.

Using 880 pb^{-1} of data collected by the DØ experiment at the Tevatron Collider, we search for an excess in the invariant mass m_{01} of the two leading transverse momentum (p_T) jets in events containing three of more b quarks candidates. The analysis presented here exploits a substantially larger data sample than the original DØ result [17]. It also employs a more sophisticated b tagging algorithm which eventually increases the statistical significance of our result.

II. DATA AND MONTE CARLO SAMPLES

A. Trigger

Due to the high cross section of multijet events, a specialized trigger for the three trigger levels (L1, L2, L3) was designed to maximize signal acceptance while remaining within data acquisition constraints. The trigger demanded at least three calorimeter towers of size $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ at L1, where ϕ is the azimuthal angle. It required three jets with $p_T > 8$ or 6 GeV, and $H_T^{L2} > 50$ or 70 GeV at L2 ($H_T^{L2} \equiv$ scalar sum of the p_T of the L2 jets). At L3, the triggering condition is three jets with $p_T > 15$ GeV, two with $p_T > 25$ GeV and the probability that the event contains only light or gluon jets less than 0.05. This b-tagging event condition at L3 is formed by combinig the impact parameter significances of the L3 tracks belonging to the six leading L3 jets.

B. Data Selection

A total of 75 million of events corresponding to an integrated luminosity of 880 pb^{-1} were preselected with one reconstructed jet of $p_T > 20$ GeV and two more jets with $p_T > 15$ GeV, all with $|\eta| < 2.6$. Jets are reconstructed using the cone algorithm [11] with radius of 0.5 and are then required to pass a set of identification cuts. The jet energies are corrected to the particle level. Only jets with corrected $p_T > 15$ GeV in $|\eta| < 2.5$ are considered when selecting on the number of jets. Events are preselected with at least three jets with corrected $p_T > 40$, 25 and 15 GeV, and no more than five jets.

Jets initiated by b quarks are identified using a neural network b-tagger. This is based upon seven input variables: the decay length significance of the secondary vertex (resulting from the b decay), a weighted combination of the track impact parameter significances, the probability that the jet originates from the primary vertex, the χ^2 per degree of freedom of the fit used to reconstruct the secondary vertex, as well as the number of tracks used in this reconstruction, the mass of the secondary vertex and the number of secondary vertices found in the jet. The average b-tagging efficency is 48.6 % while the corresponding mistag rate is 0.33 %.

C. Monte Carlo

Events of the expected signal and background are first generated by Pythia [12] or ALPGEN [13]. Then the Pythia showering is applied followed by DØ full detector GEANT [14] simulation. To simulate the impact of additionnal spectator interactions occuring in the same beam crossing, real "zero bias" data events are overlaid to the simulated ones.

Additional corrections are applied to the simulated quantities so that the jet reconstruction efficiency, the jet energy scale and resolution as well as the b-tagging efficiency match those measured in real data events.

1. hb Signal simulation

hb events in which the Higgs bosons decay into $b\bar{b}$ are generated by Pythia at different Higgs boson masses ranging from 100 to 170 GeV. For each simulated Higgs boson mass, to correct for Next to Leading Order effects, the Higgs boson momentum and rapidity spectra obtained from the NLO MCFM generator [15] are used to reweight the events. The MCFM cross-sections are taken as the signal cross-sections as well.

2. Simulation of background events

The main backgrounds for high multiplicity final states with 3 b-tagged jets arise from the QCD multi-jet production: $p\bar{p} \to jjj(j)$, $p\bar{p} \to bjj(j)$, $p\bar{p} \to b\bar{b}j(j)$, $p\bar{p} \to b\bar{b}b(b)(j)$, where j stands for light parton either $u,\ d,\ s,\ c$ or g and (j) means "plus $n \geq 0$ light partons ". The background is determined from the data as explained in Section III, but the observed multi-jet production is also compared to the simulation as a cross-check.

The processes with b-quark production have been simulated with ALPGEN, based on LO matrix elements. A summary of cross-sections obtained with ALPGEN, as well as the kinematic cuts is given in table I. These cross-sections are obtained using the renormalization and factorization scales $\mu^2 = \sum p_T^2$ where the sum is over the outgoing partons. When the scale is varied to $\mu^2 = \frac{1}{N_{part}} \sum p_T^2$ the cross-sections are nearly doubled, indicating that the level of

uncertainty is close to $50\% - 100\%$. Thus the magnitude for these processes are normalized to the data in Section III

TABLE I: Cross-sections for the generated background events.

Process	Cross-section (pb)	Generator cuts (p_T in GeV)
$b\bar{b}j$	3810	$p_T(j) > 15$, $p_T(b) > 25$ $\|\eta\| < 3$, $\Delta R < 0.4$
$b\bar{b}jj(j)$ [a]	2540	$p_T(j) > 15$, $p_T(b) > 25$, $\|\eta\| < 3$, $\Delta R < 0.4$
$b\bar{b}b\bar{b}(j)$	120	2 b's with $p_T(b) > 25$ 3b's with $p_T(b) > 15$

[a] MLM matching [16] is used to avoid double counting of events between the $b\bar{b}j$ and the $b\bar{b}jj(j)$ samples.

Simulation of other sources of backgrounds such as $Z + b$ or $t\bar{t}$ productions are judged as unnecessary, as their cross-sections are much smaller than the uncertainty arising from the choice of the scale in the QCD processes.

III. ANALYSIS

A. Background estimation

Because of the large uncertainty on the theoretical computation of the multi-jet process cross-sections, the expected background is calculated from the data. The final data sample is made of events containing at least three b-tagged jets, and is called triple b-tagged data sample. In a similar manner, events containing at least two b-tagged jets define the double b-tagged sample. The following paragraph describes how we estimate the background in the triple b-tagged sample from the double b-tagged sample.

First the probability to b-tag a jet when two other jets are already b-tagged is measured as a function of the transverse momentum of the jet. This function is called "tag rate function". To avoid a possible contamination by a $hb\bar{b}$ signal, the tag rate function is measured for jets outside a "signal region". It is defined as the window $|m_{01} - m_H| < \sigma_H$ where m_H and σ_H are respectively the mean and the resolution obtained from a Gaussian fit to the two leading jets invariant mass spectrum in the signal Monte Carlo sample. This tag rate function is calculated for each Higgs boson mass hypothesis.

Then, for each event in the double b-tagged sample the tag rate function is used to determine the probability that the event survives the three b-tag requirement. Thus we obtain the expected triple b-tagged background m_{01} distribution.

This triple b-tagged background distribution is eventually normalized outside the signal region to the triple b-tagged data sample, taking possible signal tails into account. As a consequence, this has to be done for each tested $\tan\beta$ when setting the limit.

B. Acceptance Systematics

We split the selection uncertainties into three different categories: the trigger level, the kinematic cuts ($\mathrm{p_T}$, η), and the b-tagging. Table II shows the acceptance for each set of cuts at each studied Higgs boson mass.

The signal acceptance uncertainties are estimated for these masses and are listed in Table III. Systematic uncertainties due to jet reconstruction and identification efficiencies, jet energy resolution, jet energy scale and b-tagging efficiency are computed by repeating the analysis, varying their values by $\pm 1\sigma$. In addition a theoretical uncertainty that arises from the NLO ($\mathrm{p_T}$,η) reweighting procedure, the cross-section uncertainty obtained when varying the renormalization and factorization scale and the parton distribution function uncertainty is quoted. The uncertainties on the integrated luminosity and on the trigger efficiency are also taken into account.

C. Background Systematics

Background systematics are summarized in table IV. There is a statistical error associated with the uncertainty in the normalization of the background : the error is given by $1/\sqrt{\mathrm{N_{event}}}$ where $\mathrm{N_{event}}$ is the number of events in the predicted background outside the signal window. There is an additional systematic from the tag rate function, since this function is used to propagate the shape of the double b-tagged data to the triple b-tagged data. To estimate it, the invariant mass shape of the data sample requiring at least two b-tagged jets is directly used to predict the

Higgs boson mass (GeV)	**100**	**110**	**120**	**150**	**170**
Trigger	61	64	64	64	67
Kinematic	31	35	39	40	43
BID	8.8	8.8	9.7	9.9	9.2

TABLE II: The relative acceptance (in %) of selection cuts for signal for each Higgs mass.

Higgs boson mass (GeV)	**100**	**110**	**120**	**150**	**170**
Lumi	6.5	6.5	6.5	6.5	6.5
Theoretical	12.3	12.0	12.1	13.0	13.5
Trigger	4.0	4.9	3.6	4.2	2.5
ID	0.3	0.5	0.4	0.4	0.4
JES	4.8	4.6	3.9	2.8	2.7
Reso	0.6	0.2	0.1	0.3	0.5
JET	4.9	4.7	3.9	2.8	2.8
b-ID	8.1	8.2	8.3	8.8	9.3
Total	17.2	17.3	17.0	17.7	18.0

TABLE III: The errors from each source (in %) which are added in quadrature to give the total errors on acceptance.

background. The difference between the number of events from these two background predictions give the systematic uncertainty.

D. Monte Carlo cross checks

To test our understanding of the m_{01} distributions and of b-identification algorithm performances, we use a Monte Carlo simulation of the data sample. As we can not rely on the cross-sections computed at the LO with ALPGEN, we only assume that their ratios are correct.

In a first step, we estimate the $jjj(j)$ and $bjj(j)$ contribution in the preselected sample by using the known b-tagging efficiencies (on light and b quarks) and assuming the j are only light jets initiated by: u, d, s or g. This estimation is corrected for the $bbj(j)$ production. The measured contributions demonstrate that these backgrounds are already very small in the two b-tagged sample and might be neglected. Subtracting the $jjj(j)$ and $bjj(j)$ contribution from the two b-tagged sample, and assuming the rest of the events arises from the sum of all simulated multi-jet background, we obtain a scale factor between the observation and the MC cross-sections: we find that the total cross-sections for simulated events is 2.7 higher in the data. This scale factor is not very sensitive to the way one deals with $jjj(j)$ and $bjj(j)$ productions as they are negligible. In a last step, we follow the same procedure in the three b-tagged sample and obtain a new scale factor between simulated and observed cross-sections of 2.8.

A comparison between data observation and Monte Carlo expectation [19] is shown in Figure 2 for the two b-tagged sample on the left and for the three b-tagged sample on the right. One should notice that in the two b-tagged sample the $bbb(b)$ background is negligible compared to $bbj(j)$. This is no more the case in the three b-tagged sample where the background contains a sizeable fraction of $bbb(b)$ events. Thus, the fact that the two measured scale factors are in agreement leads to the conclusion that not only we have a good knowledge of the b-tagger performances but also that the use of simulated cross-section ratios was a fair enough assumption.

A particular attention has been paid to the background involving jets initiated by c-quarks as they are identified by the b-tagging algorithm with an efficiency five times lower than for a b-quark jet. We find that processes such as $cjj(j)$, $ccj(j)$ and $ccc(c)$ are negligible in the two and three b-tagged samples and that their influences on the scale

Higgs boson mass (GeV)	**100**	**110**	**120**	**150**	**170**
Alternate method	1.8	2.3	2.4	1.9	1.9
Due to normalization	1.8	1.7	1.7	1.7	1.7
Total	2.4	2.8	2.8	2.6	2.6

TABLE IV: The errors from each source (in %) which are added in quadrature to give the total background systematics.

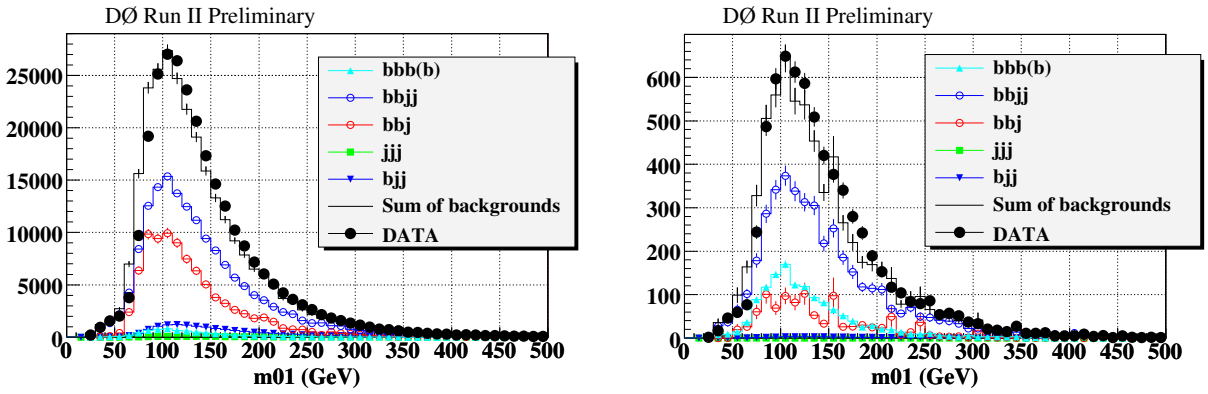

FIG. 2: Data-Monte Carlo comparison of the m_{01} distribution in the two b-tagged sample on the left and in the three b-tagged sample on the right. m01 stands for the invariant mass of the two leading jets. See text for details

Selection before b-tagging	19×10^6				
2 b-tags	269870				
3 b-tags	6749				
Higgs boson mass (GeV)	**100**	**110**	**120**	**150**	**170**
Normalization factor	0.997	0.993	0.992	0.992	0.990
Expected background	6947	6928	6926	6845	6687

TABLE V: The final number of events in the double and triple b-tagged data, and the expected background to the triple b-tagged data, as well as the normalization of the expected background. NN tight b-tagged is used.

factor measurements is negligible.

IV. RESULTS

Table V summarizes the final number of events in the data sample after requiring at least two and at least three b-tagged jets, as well as the expected number of tripple b-tagged background events for each signal mass.

The influence of the Higgs width was studied for the publication of [17]. We found that at tree level, the Higgs width yields to a slight decrease in the values of $\tan\beta$ excluded (around 2-3 unities in $\tan\beta$). Conservatively, we neglect the Higgs width in this conference note.

As no excess is seen in the data, the CL_s method, with $CL_s = CL_{s+b}/CL_b$ is used to set limits on signal production [18], with the full leading di-jet invariant mass shapes of the signal, expected background and data sample used as input. The signal histogram, derived assuming $\tan\beta = 1$ is scaled by $\tan\beta^2$. The value of $\tan\beta$ was varied until the confidence level for signal is less than 5 %. Table VI and Figure 3 show the observed and expected 95 % CL exclusion limits in the $\tan\beta - m_A$ plane that we are able to exclude with the current data.

Figure 4 shows the data, the normalized background, the Higgs signal (grey line) and the sum of the background and the Higgs signal for a higgs mass of 120 GeV at the observed 95 % CL exclusion limit.

This is converted to a cross section limit for signal production in figure 5.

Higgs boson mass (GeV)	**100**	**110**	**120**	**150**	**170**
Observed $\tan\beta$ limit	46	57	60	85	121
Expected $\tan\beta$ limit	50	58	62	84	104

TABLE VI: The observed and expected 95 % CL exclusion limits in the $\tan\beta - m_A$ plane at tree level in the MSSM.

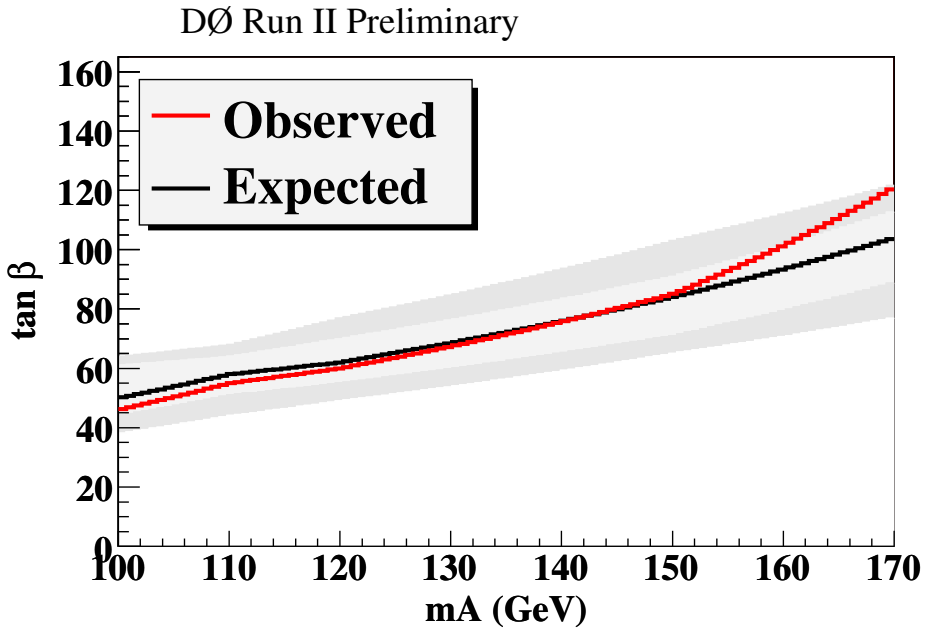

FIG. 3: The observed and expected 95 % CL limits on $\tan\beta$ as a function of m_A, assuming $\tan\beta^2$ cross section enhancement. The error bands indicate the $\pm1\sigma$ and $\pm2\sigma$ range of the expected limit.

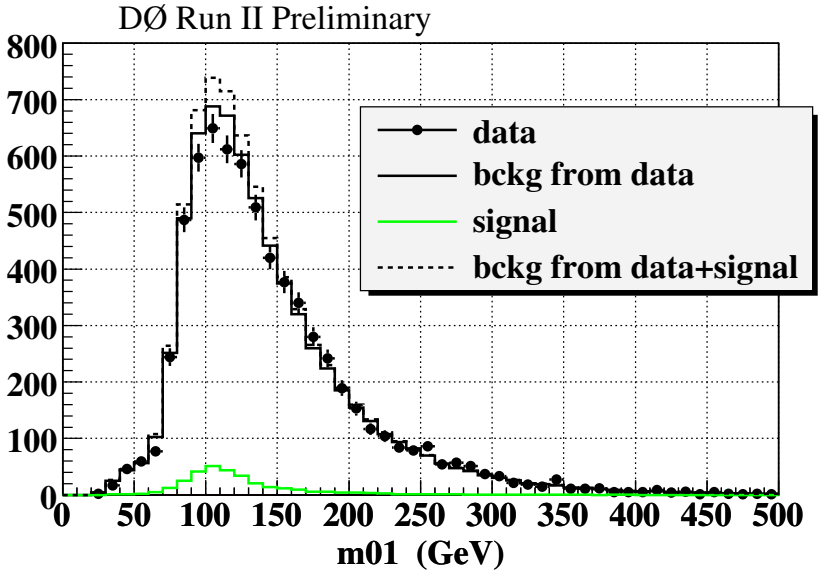

FIG. 4: The data (circles), the normalized background (solid black line), the Higgs signal (solid grey line)and the sum of the background and the Higgs signal (dashed black line) for $m_A = 120$ GeV at the observed 95 % CL exclusion limit ($\tan\beta = 60$).

The current DØ analysis, based on 880 pb^{-1}, excludes a significant portion of $\tan\beta - m_A$ plane.

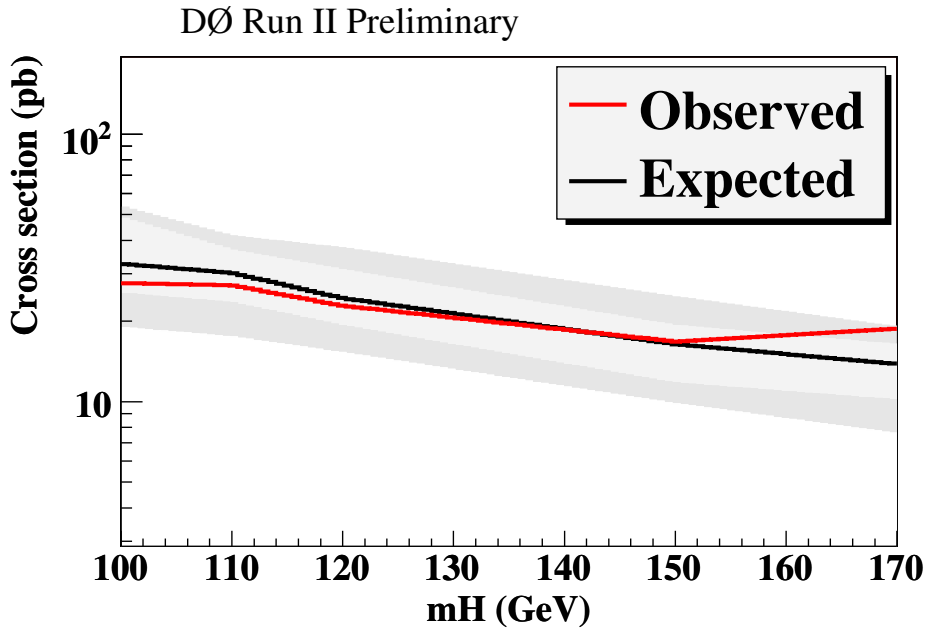

FIG. 5: The observed and expected 95 % CL limits on standard model $bH^0 \to b\bar{b}$ cross section as a function of m_H. The error bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ range of the expected limit.

Acknowledgments

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

[1] LEP Higgs Working Group (July 2005), Note2005-01 http://lephiggs.web.cern.ch/LEPHIGGS/papers/index.html
[2] The CDF collaboration, Phys. Rev. Lett. **86** (2001) 4472.
[3] The CDF collaboration, Phys. Rev. Lett. **96** (2006) 011802.
[4] The DØ collaboration, hep-ex/0605009.
[5] J.M. Campbell, R.K. Ellis, F. Maltoni and S. Willenbrock, *Higgs boson production in association with a single bottom quark*, Phys. Rev. D **67** (2003) 095002.
[6] S. Dittmaier, M. Kramer and M. Spira, Phys. Rev. D **70** (2004) 074010. S. Dawson, C. Jackson, L. Reina and D. Wackeroth, Phys. Rev. D **69** (2004) 074027.
[7] J.M. Campbell *et al.*, *Higgs boson production in association with bottom quarks*, hep-ph/0405302.
[8] S. Dawson, C. Jackson, L. Reina and D. Wackeroth, Phys. Rev. Lett. **94** (2005) 031902.
[17] The DØ collaboration, Phys. Rev. Lett. **95** (2005) 15801.
[10] DØ Collaboration, F. Last, Nucl. Instrum. Methods A **123**, 456 (2004).
[11] G.C. Blazey *et al.*, *Run II Jet Physic*, hep-ex/0005012, 10 May 2000
[12] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands, *Physics and manual*, hep-ph/0308153. We use Pythia 6.323.
[13] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," JHEP **0307**, 001 (2003) [arXiv:hep-ph/0206293].
[14] R. Brun *et al.*, CERN Program Library Long Writeup W5013 (1993)
[15] J.M. Campbell, R.K. Ellis, http://mcfm/.
[16] S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, A. Schalicke and S. Schumann, "Matching parton showers and matrix elements," arXiv:hep-ph/0602031.

[17] The DØ collaboration, Phys. Rev. Lett. **95** (2005) 15801.
[18] T. Junk, Nucl. Instrum. Methods in Phys. Res. A 434, 435 (1999)
[19] Monte Carlo expectations are corrected by the global cross-sections scale factor already mentioned