1 Search for neutral supersymmetric Higgs bosons in multijet events at $\sqrt{s}$ =1.96 TeV

Andrew Haas[1], *Avto Kharchilava*[2] *and Marine Michaut*[3]
(For the DØ Collaboration)
[1]Columbia University, USA
[2]State University of New York at Buffalo, USA
[3]DAPNIA/SPP, SACLAY, France

2 Introduction

In two-Higgs-doublet models of electroweak symmetry breaking, such as the minimal supersymmetric extension of the standard model (MSSM) [1], there are five physical Higgs bosons: two neutral CP-even scalars, h^0 and H^0, with H^0 being the heavier state; a neutral CP-odd state, A^0; and two charged states, $H^{0\ \pm}$. The ratio of the vacuum expectation values of the two Higgs fields is defined as $\tan\beta = v_u/v_d$, where v_u and v_d refer to the fields that couple to the up-type and down-type fermions, respectively. At tree level, the coupling of the A boson to down-type quarks, such as the b quark, is enhanced by a factor of $\tan\beta$ relative to the standard model (SM), and the production cross section is therefore enhanced by $\tan^2\beta$ [2]. At large $\tan\beta$, this is also true either for the h or H boson depending on their mass.

For several representative scenarios of the MSSM, LEP experiments have excluded at the 95% C.L. a light Higgs boson with mass $m_h < 92.9$ GeV [3]. At hadron colliders, neutral Higgs bosons can be produced in association with b quarks, leading to final states containing three or four b jets. The CDF experiment at the Tevatron Collider performed a search for these events in data from Run I [4]. Although their analysis was excluding some large values of $\tan\beta$ for m_A up to 250 GeV, the PDFs used to simulate the signal have been superseded. The resulting limits with the most recent PDFs would be much less stringent, hence we do not compare our current results to those from Run I.

Higgs boson production in association with b quarks in $p\bar{p}$ collisions can be calculated in two ways: in the five-flavor scheme [5], only one b quark has to be present, while in the four-flavor scheme [6], two b quarks are explicitly required in the final state. Both calculations are now available at next-to-leading order (NLO), and agree within their respective theoretical uncertainties [7, 8]. Figure 1 illustrates these processes for h^0 production at leading order (LO), and analogous diagrams can be drawn for the H^0 and A^0 bosons.

This analysis assumes CP-conservation in the Higgs sector. The masses, widths, and branching fractions for the neutral Higgs bosons into $b\bar{b}$ pairs are calculated using the CPsuperH program [9, 10]. The current analysis is sensitive to $\tan\beta$ in

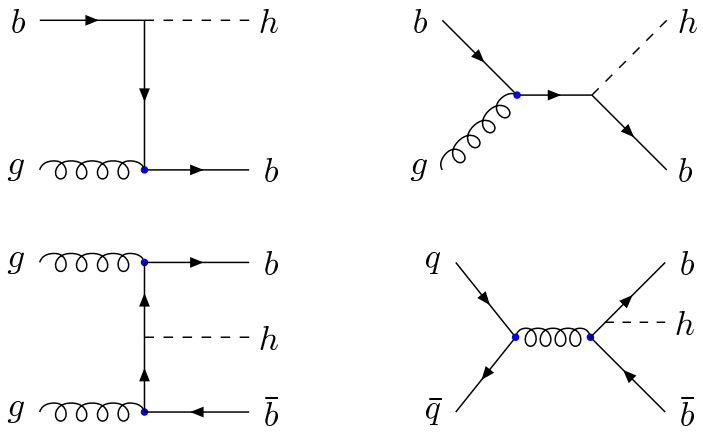

Figure 1: Leading-order Feynman diagrams for neutral Higgs boson production in the five-flavor scheme (top) and four-flavor scheme (bottom).

the range 50 – 100, and depends on the Higgs boson mass. In this region of $\tan\beta$, the A^0 boson is nearly degenerate in mass with either the h^0 or the H^0 boson, and their widths are small compared to the di-jet mass resolution. Consequently, we cannot distinguish between the h^0/H^0 and the A^0, and the total cross section for signal is assumed to be twice that of the A^0 boson. In the region of m_A from 100 to 130 GeV, all three neutral Higgs bosons can be degenerate in mass and produced simultaneously [11]. Nevertheless, the total cross section still remains twice that of the A^0 boson. Using data collected by the DØ detector from November 2002 to June 2004, corresponding to an integrated luminosity of about 260 pb^{-1}, we search for an excess in the invariant mass distribution of the two leading transverse momentum (p_T) jets in events containing three or more b quark candidates.

3 DØ detector

The DØ detector has a magnetic central tracking system surrounded by a uranium/liquid-argon calorimeter, contained within a muon spectrometer. The tracking system consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T solenoidal magnet [12]. The SMT and CFT have designs optimized for tracking and vertexing at pseudorapidities $|\eta| < 2.5$, where $\eta = -\ln(\tan(\theta/2))$ and θ is the polar angle with respect to the proton beam direction (z). The calorimeter has a central section (CC) covering up to $|\eta| \approx 1.1$, and two end calorimeters (EC) extending coverage to $|\eta| \approx 4.2$, all housed in separate cryostats [13]. The calorimeter is divided into an electromagnetic part followed by fine and coarse hadronic sections. Scintillators between the CC and EC cryostats provide additional sampling of devel-

oping showers for $1.1 < |\eta| < 1.4$. The muon system consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroidal magnets, followed by two similar layers behind the toroids, which provide muon tracking for $|\eta| < 2$. The luminosity is measured using scintillator arrays located in front of the EC cryostats, covering $2.7 < |\eta| < 4.4$. The trigger system comprises three levels (L1, L2, and L3), each performing an increasingly detailed event reconstruction in order to select the events of interest.

4 Data Sample

The large cross section for multijet production necessitates a specialized trigger to maximize signal acceptance while providing reasonable rates. This trigger at L1 requires signals in at least three calorimeter towers of size $\Delta\eta \times \Delta\varphi = 0.2 \times 0.2$ (where φ is the azimuthal angle), each with transverse energy $E_T > 5$ GeV; three clusters and $H_T^{L2} > 50$ GeV at L2 ($H_T^{L2} \equiv$ scalar sum of the L2 clusters E_T with $E_T >$5 GeV), and three jets with $p_T > 15$ GeV at L3. A total of 87 million events were selected off-line with one jet of $p_T > 20$ GeV and at least two more jets with $p_T > 15$ GeV. Jets are reconstructed using a Run II cone algorithm [14] with radius $\Delta\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} < 0.5$, and are then required to pass a set of quality criteria. To be accepted for further analysis, jets with $p_T > 15$ GeV must have $|\eta| < 2.5$. The jet energies are corrected to the particle level using η-dependent scale factors. Events with up to five jets are selected if they have a primary vertex position $|z| < 35$ cm and at least three jets with corrected $p_T > 35$, 20, and 15 GeV. The final selections are chosen to optimize the expected signal significance, defined as $S/\sqrt{B}$, where S (B) refers to the number of signal (background) events. The window size used for optimization is Higgs boson mass dependent, taken to be $\pm$ 1.5 sd of the signal peak, and varies from 24 to 36 GeV, for masses from 90 to 150 GeV.

Jets containing b quarks are identified using a secondary vertex (SV) tagging algorithm. A jet is tagged as a b-jet if it has at least one SV within $\Delta\mathcal{R} < 0.5$ of the jet axis and a transverse displacement from the primary vertex that exceeds five times the displacement uncertainty. Jets are b tagged up to $|\eta| < 2.5$, although the b tagging is about twice as efficient in the central region ($|\eta| < 1.1$) because of the CFT coverage. The b tagging efficiency is $\approx 55\%$ for central b-jets of $p_T > 35$ GeV, with a light quark (or gluon) tag rate of about 1%.

5 Signal

Signal events were simulated using the PYTHIA [15] event generator followed by the full DØ detector simulation and reconstruction chain. PYTHIA minimum-bias events

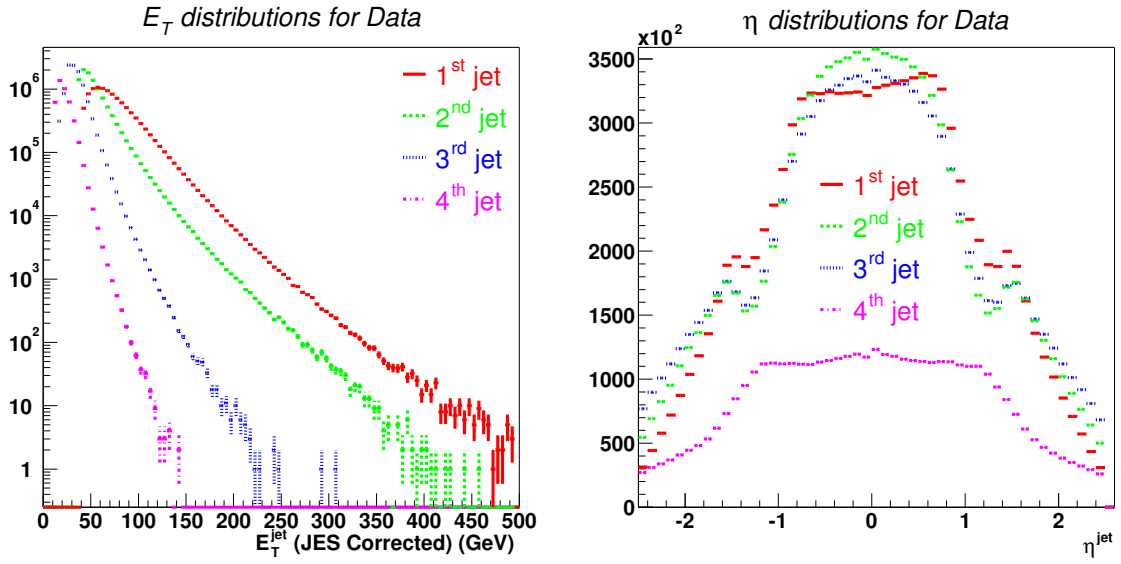

Figure 2: Left: The E_T distributions of the first, second , third , and fourth highest E_T jets in each event. Right: The η distributions of the first, second, third, and fourth highest E_T jets in each event. For both plots, jets must pass all quality cuts and taggability requirements, each event must pass one of the multi–jet triggers, jet E_T's are corrected for jet–energy scale, and the events must pass E_T cuts of 40, 35, and 15 GeV for the first, second, and third leading jets.

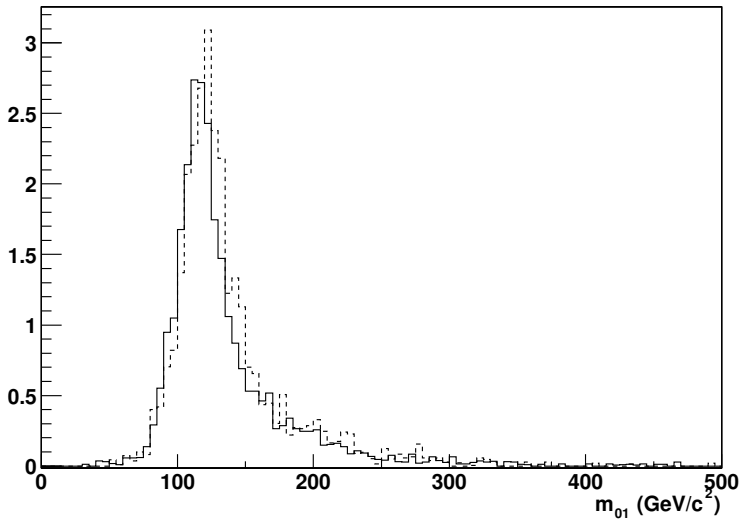

Figure 3: Comparison of bh (solid) and bbh (dashed) simulations of the triple b–tagged signal in 3–jet events.

were added to all generated events, using a Poisson probability with a mean of 0.4 events to match the instantaneous luminosities at which the data were taken ($1-6\times 10^{31}\mathrm{cm}^{-2}\mathrm{s}^{-1}$).

There are two related leading-order (LO) processes of MSSM neutral Higgs production with three or more high–p_T b–jets in the final state.

- $q\bar{q}$,gg$\rightarrow b\bar{b}$ h$\rightarrow b\bar{b}$ $b\bar{b}$, the "4-flavor" prescription.
- $q\bar{q}$,gg$\rightarrow$bg$\rightarrow$bh$\rightarrow$b$b\bar{b}$, the "5-flavor" prescription.

Radiative effects at NLO, such as the emission of hard gluons, alters the kinematics of each simulation such that they are nearly indistinguishable in both the 3-jet and 4-jet channels at higher orders [16] [17] [18]. This behavior is to be expected since both prescriptions are being used to calculate the same physical process, and their agreement at NLO is encouraging. Figures 3 compare the NLO predictions for Higgs signals using each prescription, for a Higgs mass of 120 GeV.

At LO, with initial-state radiation (ISR) and final-state radiation (FSR) turned off in Pythia, the bh and bbh spectator b quarks (the b from the initial state which does not radiate the Higgs) have very different properties, as seen in top plots in Figure 4. The bh process always has a spectator b in this case at very high absolute η, whereas all b-quarks in the bbh process are more central. Thus, it would seem that the signals could be combined without double-counting. However, when ISR and FSR are included, the signals are no longer significantly kinematically different, as seen in

bottom plots in Figure 4. Because the bh and bbh significantly overlap kinematically after radiation effects, a procedure must be chosen to avoid double counting. We chose to simply use one signal process. (Another solution would be to match the two processes by the p_T of the outgoing spectator bottom quarks.) The bh signal is a simpler calculation at NLO theoretically, and suffers from smaller scale uncertainties, so we chose bh for simulating signal events.

The bh events, with $h^0 \rightarrow b\bar{b}$, were generated for Higgs boson masses from 90 to 150 GeV. Reconstructed jets in simulated events were corrected to match the jet reconstruction and identification efficiencies in data. The energy of simulated jets was smeared to match the measured jet energy resolution. Plots showing the same basic quantities checked for data quality are shown in Figures 5 for the bh Monte Carlo samples (with $m_h = 120$ GeV).

Figure 6 shows NLO ...

The following settings have been used for the SM bh production cross section calculations. Besides the new kinematic cuts appropriate for current analysis, it is basically corresponds to settings for published results.

* Mh=90-150 GeV in 10 GeV steps

* MSbar scheme for the b Yukawa coupling (with running b mass, evolved at 2 loops), mb_{pole}=4.62 GeV

* SM b and t Yukawa couplings: y_b(SM)=$-mb(mu_r) * sqrt(G_F sqrt(2))$ $y_t(SM) = -mtop * sqrt(G_F sqrt(2))$ with G_F=1.16639 10^-5 1/GeV2.

* $E_C M$=1.96 TeV

* CTEQ6M

* —eta—¡2.5 and ptcut=15 GeV

* $mu_r = mu_f = 0.5 * mu_0$ (=central scale choice with $mu0 = mb_pole + mh/2$)

* $alphas(mu_r)$ with the top quark decoupled (alphas(mz=91.1875 GeV)=0.1179816073295792)

* cut on the separation of the gluon and b-quark in the pseudo-rapidity-azimuthal-angle plane, dR, ie when dR¡0.4 the hard gluon 4-momentum is combined with the b-quark 4-momentum.

* mtop=178 GeV

The p_T and rapidity spectra of the Higgs bosons from PYTHIA were compared to those from the NLO calculation [5]. The shapes agree well, as shown in Figures 7 and 8. The agreement indicates that the PYTHIA kinematics are approximately correct, especially after ISR/FSR effects in Pythia, which mimic the behavior of higher–order processes (such as gluon radiation). The simulated events were weighted to match the p_T spectrum of the Higgs boson given by NLO, resulting in a 10% reduction of the overall signal efficiency. The weighting function used is shown in Figure 9. The effect on the total number of events in the signal peak is shown in Figure 10. There is a 13% reduction in the total number of events passing analysis cuts caused by the re–weighting of the events to match the NLO Higgs p_T spectrum.

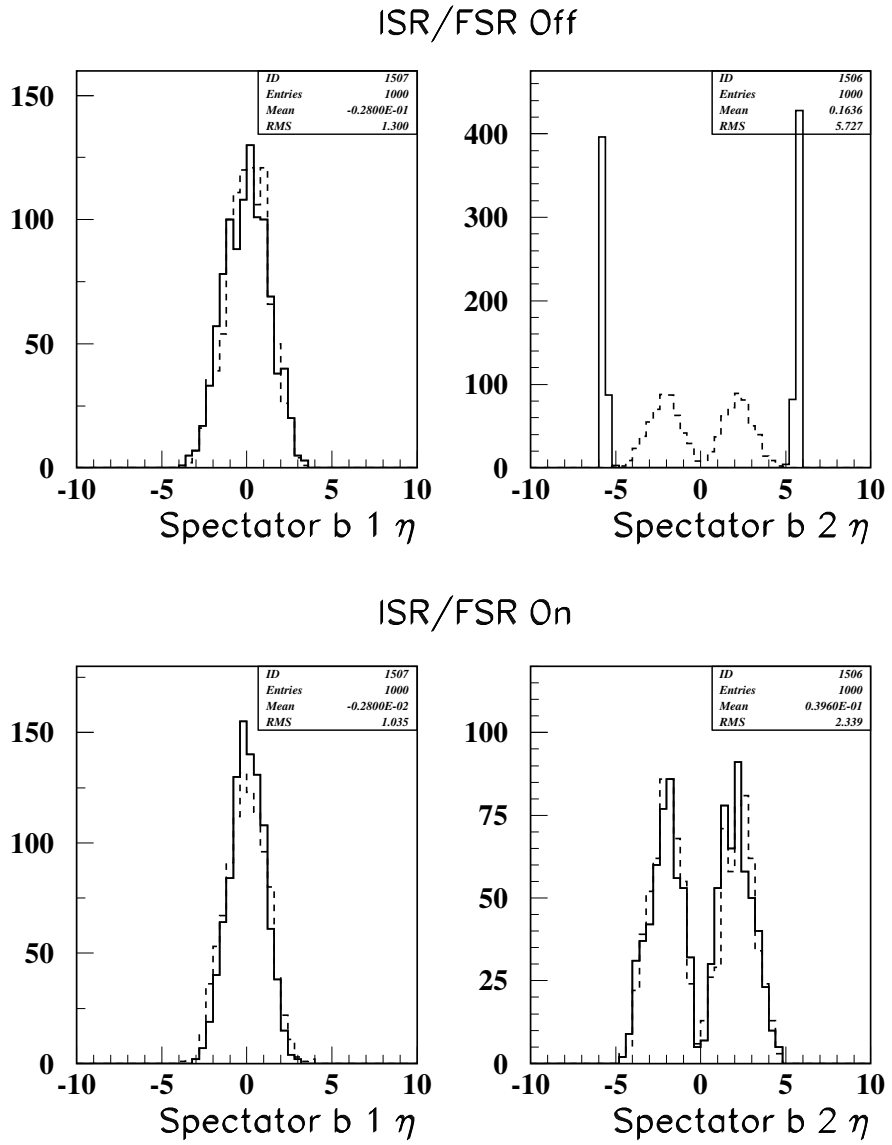

Figure 4: A comparison of the η spectra of the bh (solid) and bbh (dashed) MC signals from Pythia with ISR and FSR turned OFF(top) and ON(bottom). The spectator b 1 is the b parton which radiated the Higgs, and is thus at higher p_T and more central, typically than spectator b 2 which is simply from the initial "gluon splitting" and does not interact further, aside from radiation. With ISR and FSR, the spectator 2 b parton has very similar kinematics in bh and bbh.

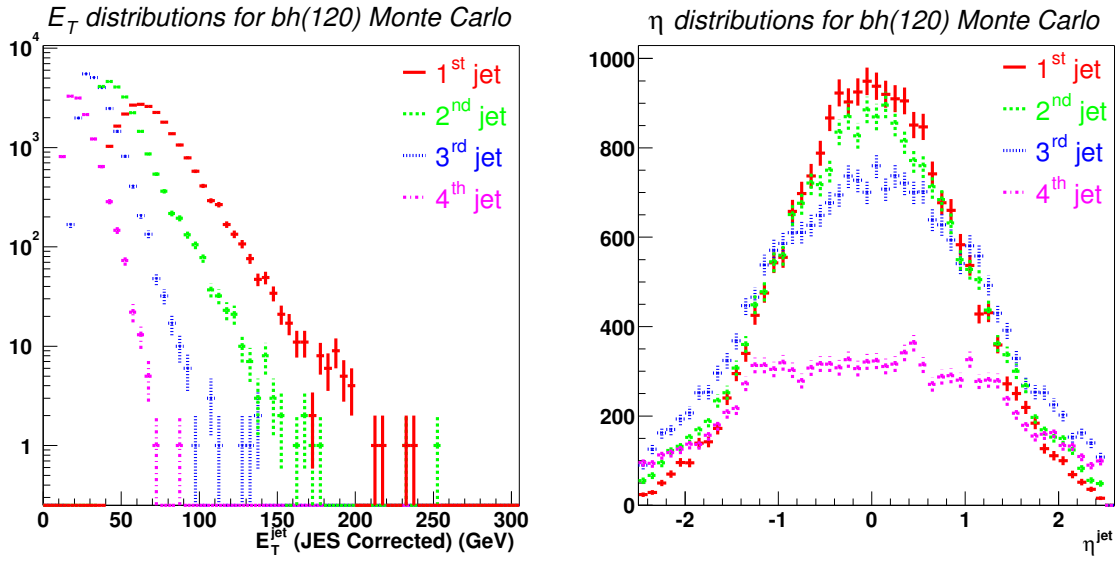

Figure 5: Left: The E_T distributions of the first, second, third, and fourth highest E_T jets in each event of the generated bh Monte Carlo sample with $m_h = 120$ GeV. Right: The η distributions of the same jets. For both plots, jets must pass all quality cuts and taggability requirements, each event must pass one of the (simulated) multi–jet triggers, jet E_T's are corrected for jet–energy scale, and the events must pass E_T cuts of 40, 35, and 15 GeV for the first, second, and third leading jets.

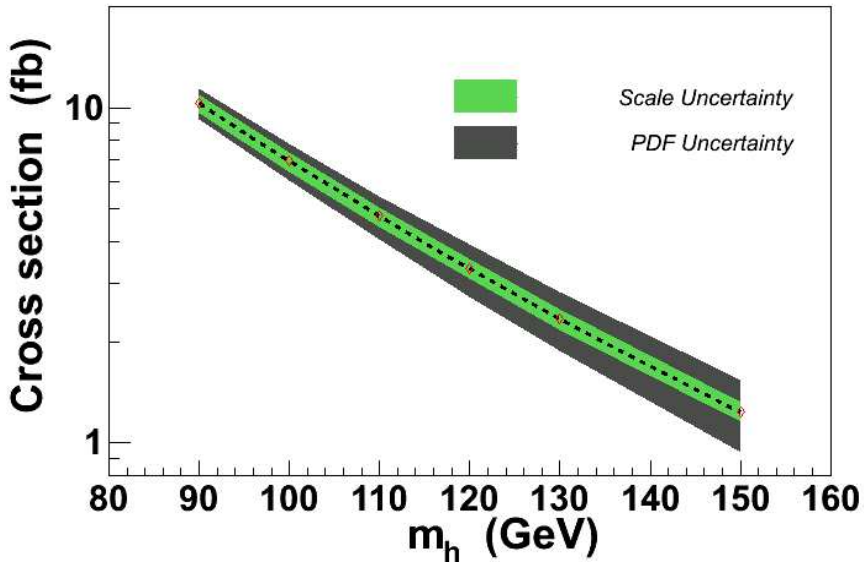

Figure 6: NLO cross section ... .

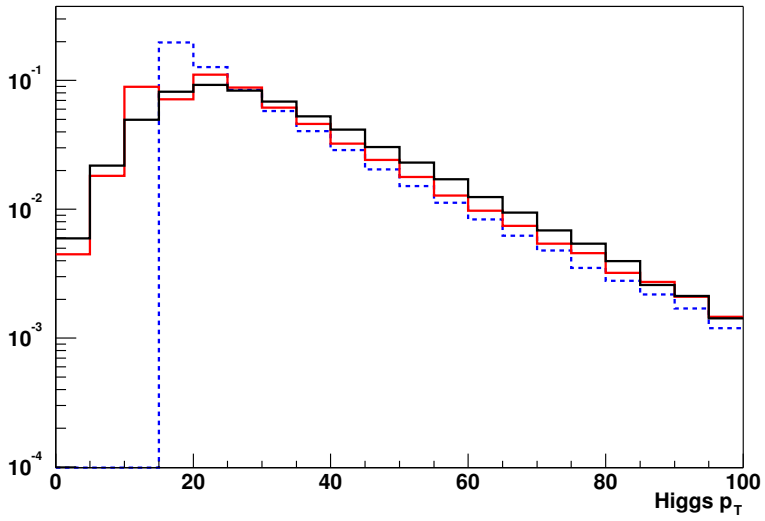

Figure 7: Comparison between the p_T spectra of the simulated Higgs in MCFM at LO (dashed–blue), MCFM at NLO (solid–red), and our Pythia simulation (solid–black).

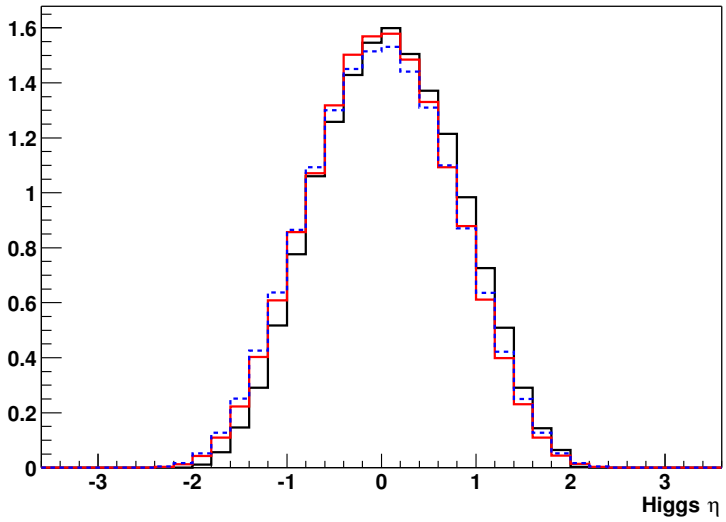

Figure 8: Comparison between the η spectra of the simulated Higgs in MCFM at LO (dotted–blue), MCFM at NLO (solid–red), and our Pythia simulation (solid–black).

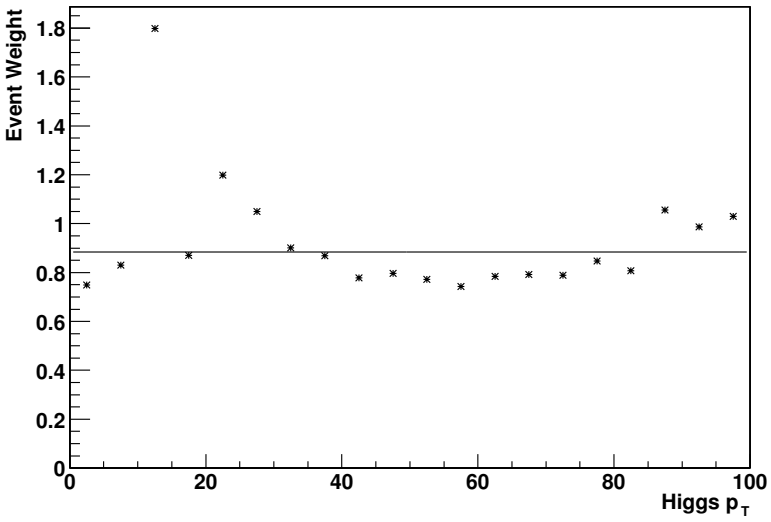

Figure 9: The weight given to each event in the bh Monte Carlo such that the correct NLO Higgs p_T spectrum is reproduced. Also shown is a fit to a constant (the average weight).

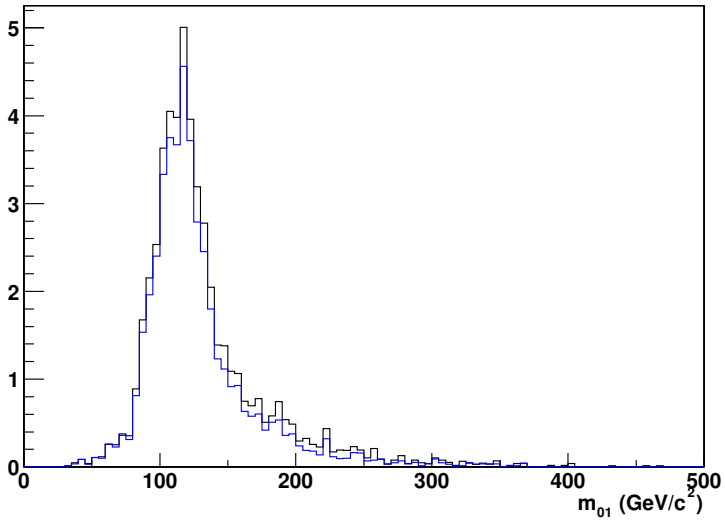

Figure 10: Comparison between the number of events in the non–b–tagged invariant mass peak of the bh Monte Carlo sample (with m_h=120 GeV) passing cuts before (black, higher) and after (blue, lower) re–weighting the leading–order events to match the MCFM NLO Higgs p_T spectrum.

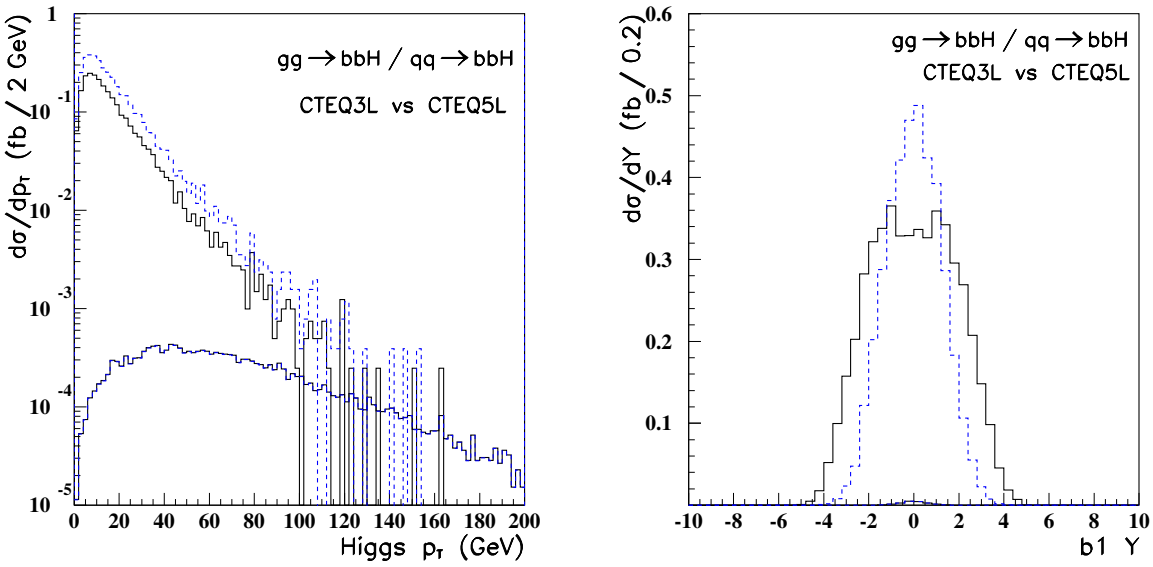

Figure 11: Higgs boson p_T distributions (left) and the b-quarks rapidity distributions (right) as obtained by Pythia with CTEQ3L and CTEQ5L PDFS and for two processes: gg, qq $\rightarrow$ bbH. For the rapidity distribution, one b-quark is picked up in the event that has smaller —Y— value.

6 Signal simulations with CTEQ3L and CTEQ5L PDFs

This section describes signal simulation results obtained with two PDFs: CTEQ3L [ref...] that have been used in CDF Run I analysis that is compared with a more recent set, CTEQ5L [ref.]

Simulations were done using Pythia 6.2 for the process number xxx.

Figure **??** Higgs boson p_T ...

Clearly, ...

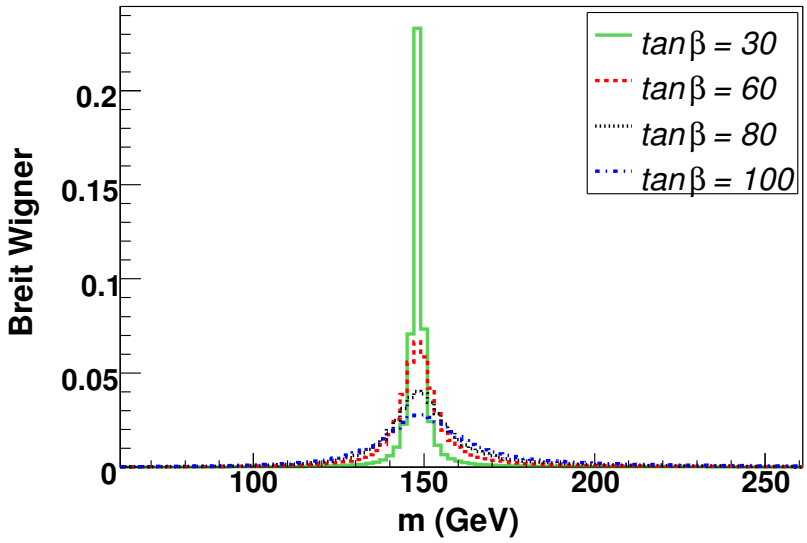

Figure 12: Breit Wigner distribution for $m_H = 150$ GeV.

7 Modifications Due to the Higgs Bosons Widths

7.0.1 Event spectrum and production cross section

Given the integrated luminosity available, this analysis is only sensitive to the production of Higgs bosons at high tan β. The Higgs bosons' widths, which are proportional to $\tan^2\beta$, are large at high $\tan\beta$, thus their possible effects must be studied. For the heaviest mass studied of $m_H = 150$ GeV, the $\tan\beta$ values excluded at 95% C.L. are greater than 93. The 150 GeV Higgs boson' width is 22 GeV for such a high value of $\tan\beta$ as given by HDECAY.

The increased width of the Higgs resonance was studied using a relativistic Breit Wigner function:

$$\mathrm{BW}(\mathrm{m_H}, \mathrm{m}, \tan\beta) \propto \frac{\mathrm{m}\,\Gamma(\mathrm{m}, \tan\beta, \mathrm{m_H})}{(\mathrm{m}^2 - \mathrm{m_H^2})^2 + \Gamma(\tan\beta, \mathrm{m_H})^2 \mathrm{m_H^2}} \tag{1}$$

where $\Gamma(\mathrm{m}, \tan\beta, \mathrm{m_H}) = \Gamma(\mathrm{m_H}, \tan\beta)(1 - 4\frac{\mathrm{m_b^2}}{\mathrm{m^2}})^{\frac{3}{2}}\frac{\mathrm{m}}{\mathrm{m_H}}$ is the width of the Higgs boson of mass $\mathrm{m_H}$ at the scale m. The Breit Wigner shapes for four different values of $\tan\beta$ are shown in Figure 12.

To obtain the event spectrum, the Breit Wigner is then weighted by the bh production cross section:

$$\mathrm{BW}(\mathrm{m_H}, \mathrm{m}, \tan\beta) \cdot \sigma(\mathrm{m}, \tan\beta, \mathrm{p_{T\,b}}) \tag{2}$$

where $\sigma(\mathrm{m}, \tan\beta, \mathrm{p_{T\,b}})$ is the bh production cross section for a zero-width Higgs boson, and $\mathrm{p_{T\,b}}$ is the transverse momentum of the b-quark produced in association

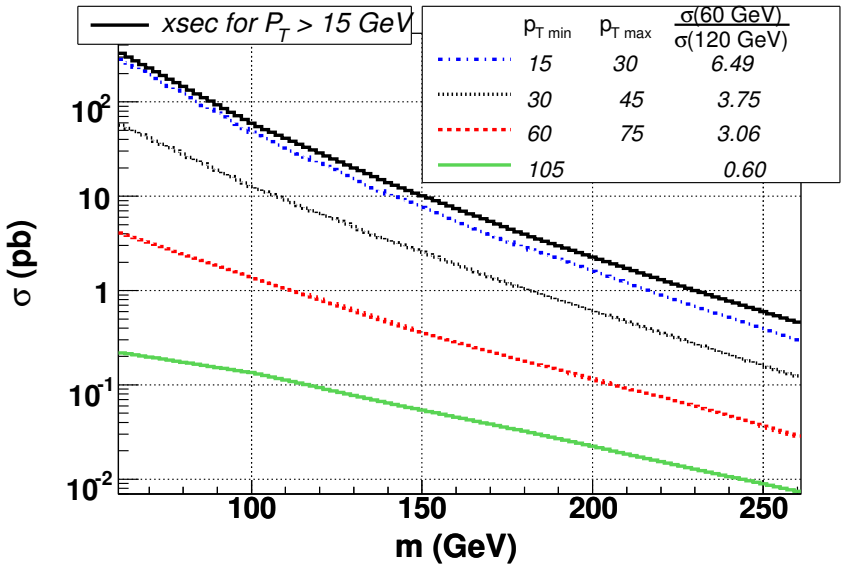

Figure 13: bh process production cross section at NLO. The ratio $\frac{\sigma(60\ \mathrm{GeV})}{\sigma(120\mathrm{GeV})}$ for each $p_{T\,b}$ bin is shown in the legend.

with the Higgs. This cross section is computed at NLO using the program MCFM [**?**] for six Higgs masses in 15 GeV $p_{T\,b}$ bins, as shown in Figure 13 for $\tan\beta$ of 100. Note that the cross-section varies less with the Higgs mass for high p_{Tb} events. It has been checked that the values obtained for the integrated cross section over $p_{T\,b}$ are in agreement with those values used in the analysis. The event spectrum using the integrated cross section is shown in Figure 14 for a Higgs mass of 150 GeV and four different values of $\tan\beta$. Since the bh production cross-section varies by two orders of magnitude between 60 and 120 GeV, the event spectrum for a heavy Higgs and a $\tan\beta$>60 show a tail at low masses that can yield a loss in signal efficiency. Because of the variation of the cross section with $p_{T\,b}$, this tail is smaller at high $p_{T\,b}$, as seen in Figure 15 which represents the different event spectra for each $p_{T\,b}$ bin, a Higgs mass of 150 GeV and $\tan\beta = 100$. Due to the larger Higgs width, one could think that the width effect should be more sizable for heavy Higgs. But this is toned down by the harder p_{Tb} spectrum.

When the width is taken into account, the cross section for the Higgs production is: $\int \mathrm{BW}(\mathrm{m}, \mathrm{m_H})\sigma(\Gamma = 0, \mathrm{m})\,\mathrm{dm}$. For practical reasons, the integration was numerically performed over the range [60-280] GeV. The results are shown in Figure 16 for Higgs masses of 110, 130 and 150 GeV. The width effect always increases the cross section. The loss in efficiency is therefore compensated by a higher production cross section.

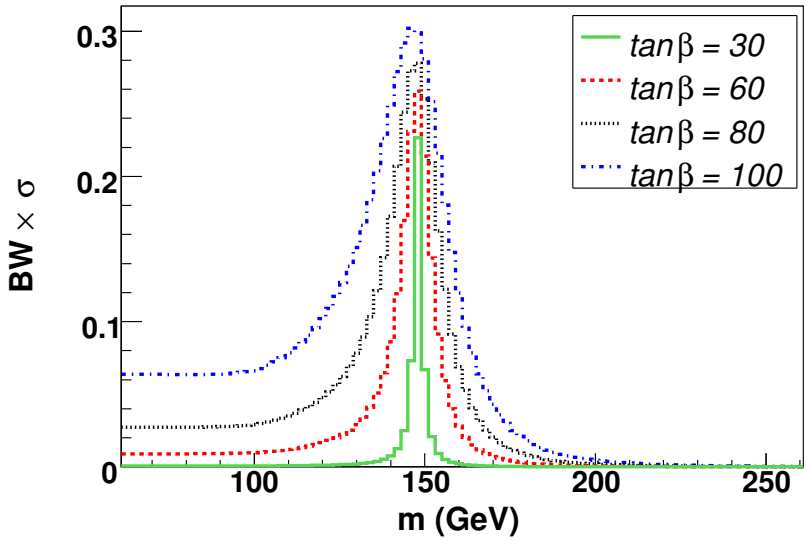

Figure 14: Event spectrum for $m_H = 150$ GeV and four $\tan\beta$ values, using the integrated cross section over $p_{T\,b}$.

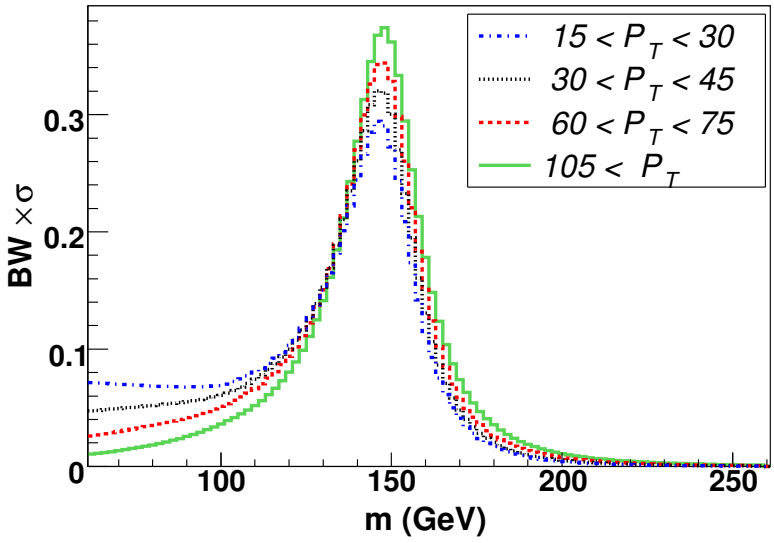

Figure 15: Event spectrum for $m_H = 150$ GeV and $\tan\beta = 100$, using the $p_{T\,b}$ dependent cross section.

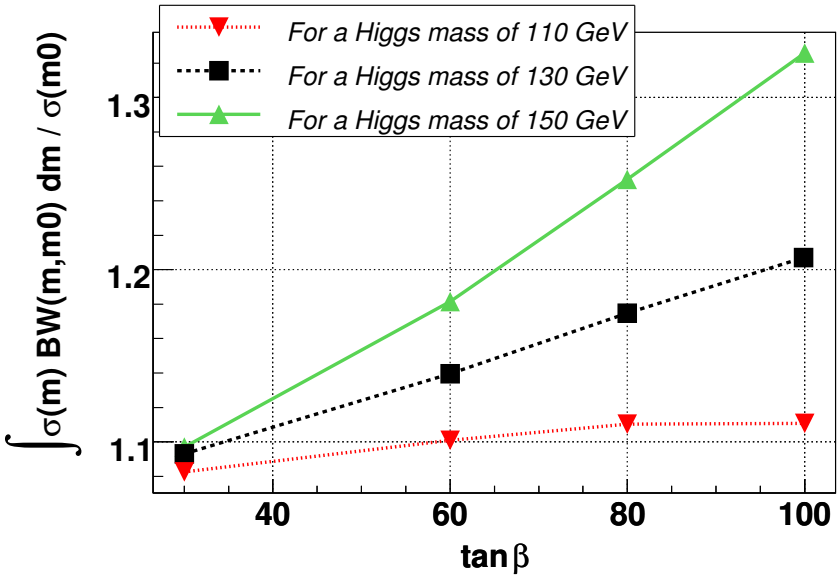

Figure 16: Ratio between the cross sections obtained with and without width effect versus tan β for m_H = 110, 130 and 150 GeV.

7.0.2 Effect on the analysis

The bh Monte Carlo was smeared using the spectra of the Figure 15, assuming that the two leading jets come from the Higgs, and the third leading jet comes from the b produced in association with the Higgs. Depending on the third jet transverse momentum, the associated event spectrum was used to re-scale the transverse momenta of the two leading jets in each event. The difference in signal acceptance before and after the smearing of the Monte Carlo to include the Higgs width is shown in Figure 17. The effect is small after the kinematic analysis cuts, and yields only a slight shift of the invariant mass peak. These signal spectra were used to recompute the expected limits for Higgs masses of 90, 110, 130, 150 GeV. The limits are only slightly weaker than if the Higgs' widths are neglected, and differ by typically one or two tan β unit.

A cross check was performed to verify that smearing at the reconstructed level, as opposed to the parton level, and assuming that the two leading jets come from the Higgs, yields a similar invariant mass shape. Each event in Pythia was re-scaled at the parton level, before the Higgs decay and b quark fragmentation, and then passed into the full DØ simulation. No significant difference was seen.

To conclude, the Higgs width can be neglected up to a Higgs mass of 150 GeV and tan β<100.

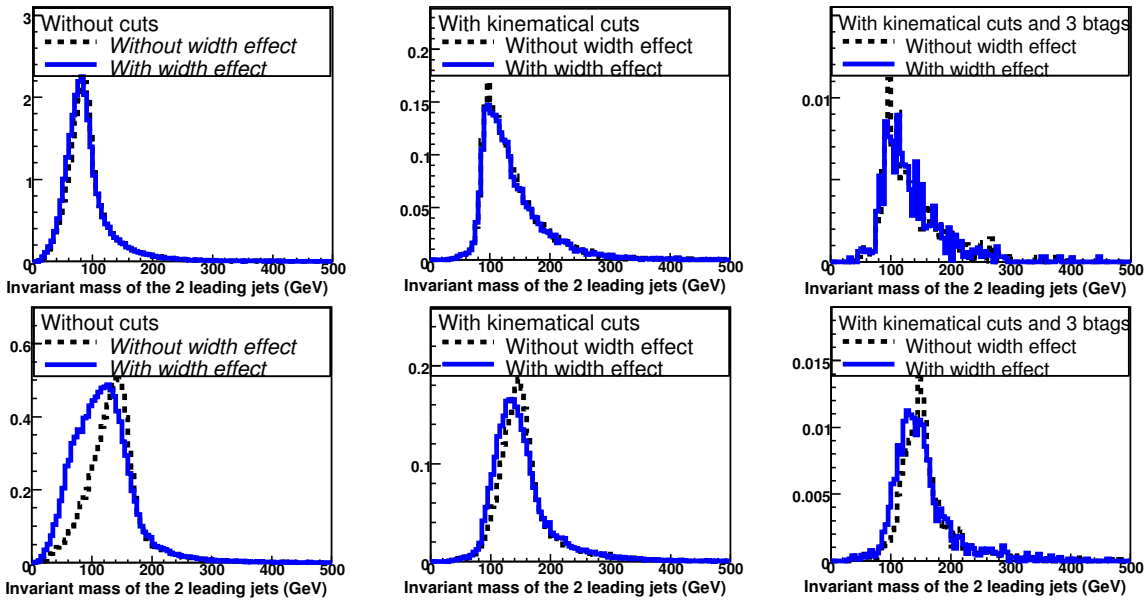

Figure 17: Comparison between the invariant mass of the leading two jets without cuts (left) after kinematical cuts of the analysis (middle) requiring three or more b-tagged jets (right), before and after the addition of the Higgs width, for m_H = 90 GeV, $\tan\beta = 60$ (top row) and m_H = 150 GeV, $\tan\beta = 100$ (bottom row).

8 Background

Of all SM processes, multijet production is the major source of background. This background is determined from data by normalizing distributions outside of the signal region. As a cross-check, we also compare data with simulations. ALPGEN [19] is used to generate three samples of events for $b\bar{b}j$ and $b\bar{b}jj$ with j corresponding to up, down, strange or charm quarks, or gluons, and $b\bar{b}$ $b\bar{b}$ final states with generator-level requirements: $p_T^b > 25$ GeV, $p_T^j > 15$ GeV, $|\eta| < 3.0$, and $\Delta\mathcal{R} > 0.4$ between any two final-state partons. These selections do not introduce significant bias because the final sample contains much harder jets, after the application of trigger and b-tagging requirements. The cross–sections given by ALPGEN for these processes are listed in Table 1.Samples of $b\bar{b}j$ and $b\bar{b}jj$ are added together, but the $b\bar{b}jj$ sample is weighted by 0.85 to match the jet multiplicity observed in doubly b-tagged data. The cross sections obtained from ALPGEN are 8.9 nb, 3.9 nb, and 60 pb, for the respective three states. All other backgrounds are expected to be small and are simulated with PYTHIA: $p\bar{p} \rightarrow Z(\rightarrow b\bar{b})$+jets, $p\bar{p} \rightarrow Zb$, and $p\bar{p} \rightarrow t\bar{t}$. Cross sections of 1.2 nb, 40 pb [20], and 7 pb are assumed, respectively.

There are two main categories of multijet background. One contains genuine heavy-flavor (HF) jets, while the other has only light-quark or gluon jets that are

Process	Cross–section (pb)	Generator Cuts (p_T in GeV)
$b\bar{b}$ jj (ALPGEN)	3810	p_T(b)>25, p_T(j)>15, $\|\eta\|$<3.0, ΔR>0.4
$b\bar{b}$ jj (MADGRAPH)	6100	“”
$b\bar{b}$ j (ALPGEN)	8940	“”
$b\bar{b}$ $b\bar{b}$ (ALPGEN)	58	p_T(b)>15, $\|\eta\|$<3.0, ΔR>0.4
t$\bar{\text{t}}$ (Pythia)	7	none
Z($b\bar{b}$)+jets (Pythia)	1180	none
Zb (MADGRAPH)	10	p_T(b)>15, $\|\eta\|$<3.0, ΔR>0.4
Z$b\bar{b}$ (MADGRAPH)	3	“”

Table 1: Background Monte Carlo cross–sections and generator-level cuts.

mistakenly tagged as b-quark jets, or correspond to gluons that branch into nearly collinear $b\bar{b}$ pairs. Using the selected data sample, before the application of b-tagging requirements, the probability to b-tag a jet is measured as a function of its p_T in three $|\eta|$ regions. These functions are called “mis-tag” functions. They are corrected for the contamination from true HF events by subtracting the estimated fraction of $b\bar{b}j(j)$ events in the multijet data sample (1.2%), obtained from an initial fit to the doubly b-tagged data. These corrected mis-tag functions are then used to estimate the mis-tagged background, by applying them to every jet reconstructed in the full data sample.

In order to test the modeling of the mis-tag background, the high statistics doubly b-tagged data is compared to simulations first, before extrapolating to the triply b-tagged background. The expected signal contribution to the doubly b-tagged data is negligible. The comparison in invariant mass spectrum of the two jets of highest p_T (not necessarily the two b-tagged jets) in the doubly b-tagged data with the expected background is shown in Fig. 18. The b-tagging in this analysis does not distinguish between contributions from bottom and charm events. However, the efficiency for tagging a c-jet is known from simulations to be about 1/4 of that for tagging a b-jet. Therefore, when two b-tags are required, the fraction of $c\bar{c}j(j)$ events relative to $b\bar{b}j(j)$ events will be a factor of ≈ 16 lower after tagging. We have estimated the fractions of $c\bar{c}jj$ to $b\bar{b}jj$ prior to b-tagging using the MADGRAPH Monte Carlo generator [21]. The $c\bar{c}jj$ cross section is 22% higher than $b\bar{b}jj$ for the same generator-level selections. Therefore, the contribution of $c\bar{c}j(j)$ in the doubly b-tagged data sample is expected to represent about 8% of the events. Thus, when we refer to the $b\bar{b}j(j)$ normalization, it should be understood that approximately 8% of the events are from the $c\bar{c}j(j)$ process. After these corrections for $c\bar{c}j(j)$ events, the HF multijet processes are only a factor of 1.08 higher in data than predicted by ALPGEN. The shape of the estimated background agrees well with the data over the entire invariant mass region.

We also compare the invariant mass spectrum of the two leading E_T *b-tagged*

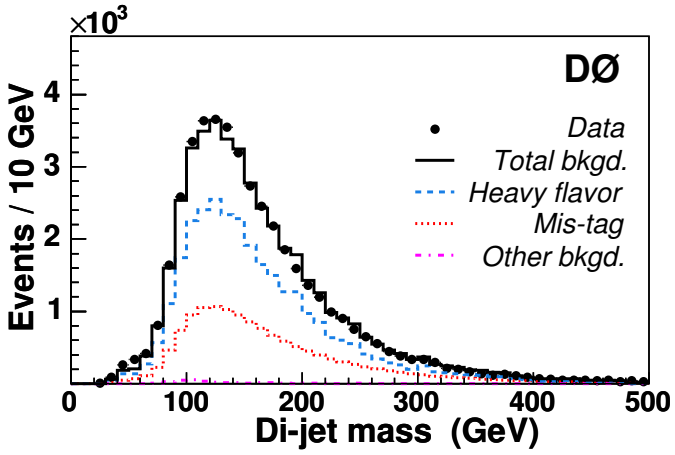

Figure 18: Fit of the invariant mass spectrum of the two leading p_T jets in the doubly b-tagged data to a sum of backgrounds: mis-tags derived from data (dotted), $b\bar{b}j(j)$ (dashed), and other backgrounds ($Z(\rightarrow b\bar{b})$+jets, Zb, $t\bar{t}$ and $b\bar{b}$ $b\bar{b}$) (dashed-dotted).

jets in the double b-tagged data to background estimates based on Monte Carlo and data, as shown in Figure 19. The $b\bar{b}$ j(j) events dominate the low invariant mass part of the spectrum, most likely due to gluon-slitting events. The fake b-tag probability, on the other hand, is not correlated with the angle between other b-tagged jets, thus the jets are more back-to-back on average, and the invariant mass tends to be larger. The fact that the same normalization for the $b\bar{b}$ j(j) Monte Carlo, derived in the leading E_T invariant mass spectrum, describes also this leading E_T b-tagged invariant mass distribution, gives us confidence that the background composition is properly modeled.

To estimate the background for triply b-tagged events, the mis-tag function is applied to the non-b-tagged jets in the doubly b-tagged events. This provides the shape of the multijet background distribution with at least three b-tagged jets. This neglects any contributions from processes with more than two true b-jets, such as from $b\bar{b}$ $b\bar{b}$ and $Z(\rightarrow b\bar{b})b\bar{b}$ production. However, the shapes of these backgrounds from simulations are similar to those of the doubly b-tagged spectra, and their rates are small. The overall background normalization is therefore determined by fitting the leading two jets invariant mass spectrum in triply b-tagged events outside of the hypothesized signal region to the estimated shape for triply b-tagged background. The systematic effect on the normalization of the background from any signal contributing outside the search window was studied and found to be small relative to other uncertainties, as described below.

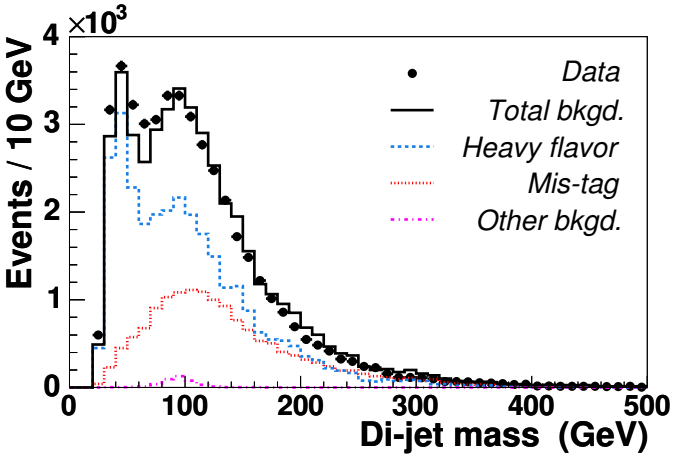

Figure 19: Fit of the double b–tagged data's two leading E_T *b-tagged* jets' invariant mass spectrum, after fake–tag parameterization correction, to a sum of backgrounds: multi–jet fakes from data (solid), ALPGEN $b\bar{b}$ jj MC (dashed), and other small backgrounds (Z($\rightarrow b\bar{b}$)+jets, $t\bar{t}$, and ALPGEN $b\bar{b}$ $b\bar{b}$) (dotted).

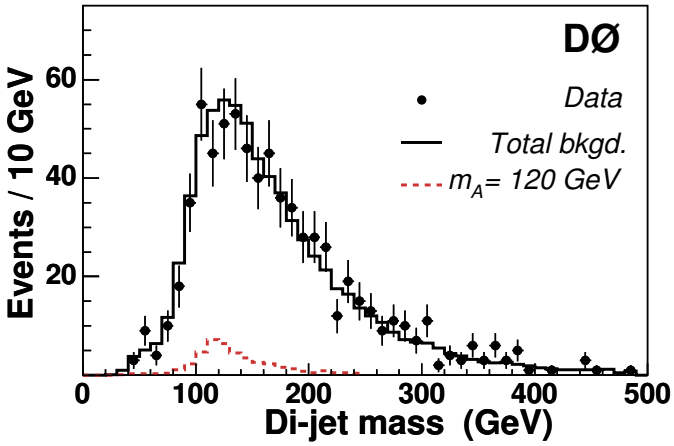

Figure 20: Invariant mass spectrum of two leading jets in events with at least three b-tagged jets, estimated background, and the signal for a 120 GeV Higgs boson that can be excluded at the 95% C.L.

m_A (GeV)	Trigger	Kinematic	b-tag	Total
90	44	18	3.5	0.3
100	45	24	3.5	0.4
110	56	24	3.9	0.5
120	60	27	4.2	0.7
130	65	29	4.3	0.8
150	76	31	4.4	1.0

Table 2: Signal acceptances for each set of criteria (in %).

9 Systematic errors

The selections in this analysis can be grouped into trigger level, kinematic (p_T, η, n_j), where n_j is the number of untagged jets, and b-tagging. Table 2 shows the acceptances for each set of criteria made in the analysis, for six values of Higgs boson mass. The systematic uncertainty on signal acceptance is nearly independent of assumed m_A, and is dominated by the uncertainty on b-tagging efficiency ($\pm$15%), followed by uncertainties on jet energy scale, resolution and identification efficiency ($\pm$9% in sum). These uncertainties are calculated by repeating the analysis with each value changed by $\pm$ one standard deviation (sd). The systematic uncertainties corresponding to uncertainties in p_T distributions for simulated signal at NLO, the integrated luminosity, and the trigger efficiency are found to be $\pm$5%, $\pm$6.5%, and $\pm$9%, respectively. These uncertainties, added in quadrature, result in a total systematic uncertainty of $\pm$21%. Signal acceptance uncertainties are listed in Table 3 for various masses.

The accuracy in modeling the shape of the background distribution can be estimated from the χ^2/dof between the estimated background and the data. The statistical error associated with the uncertainty in the normalization of the background (from the fit outside the signal region) is multiplied by $\sqrt{\chi^2/dof}$. The background uncertainty is estimated to be $\lesssim$ 3%. The systematic uncertainty arising from the width chosen for the search window is evaluated by varying it from $\pm$1 to $\pm$2 sd, centered on the peak value. The resulting change in background normalization is much smaller than from other sources of background uncertainties.

10 Interpretation within the framework of the MSSM

[Need to type in formula for MSSM enhancement factor]

[Describe ingredients and refer to Carena et al.]

Depending on the choice of model parameter ... Fig. 22 shows b-Higgs boson coupling enhancement factors as function of $\tan\beta$.

m_A (GeV)	b-tag	Jet E. Scale	Jet E. Res.	Jet-ID	Total
90	14	7.5	1.4	3.8	20.4
100	15	7.1	0.7	3.8	21.0
110	14	8.1	0.7	3.7	20.6
120	14	8.5	0.8	3.6	21.0
130	14	7.8	0.4	3.3	20.4
150	15	7.7	0.8	3.4	21.1

Table 3: The uncertainties from each source (in %). Other small uncertainties which do not depend on mass are also added in quadrature (see text).

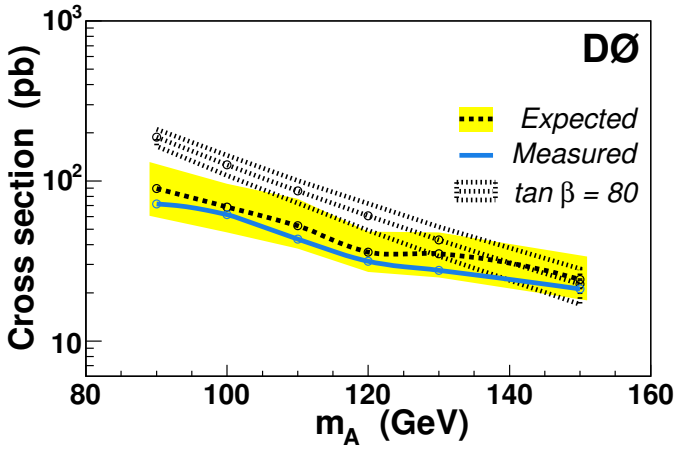

Figure 21: The expected and measured 95% C.L. upper limits on the signal cross section as a function of m_A. The band indicates the ± 1 sd range on the expected limit. Also shown is the cross section for the signal at $\tan\beta = 80$ in the "no mixing" scenario of the MSSM, with the theoretical uncertainty indicated by the overlaid band.

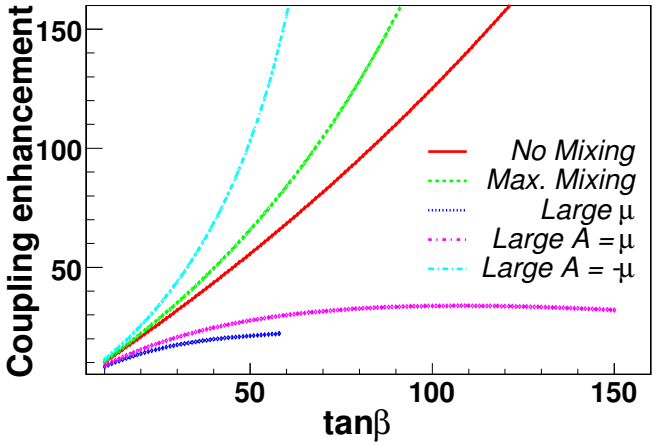

Figure 22: b-Higgs boson coupling enhancement factors as function of $\tan\beta$ for various MSSM scenarios.

As one can see in some parameter space enhancement is factor is mild thus no sensitivity to Higgs boson is expected with the current dataset. In the following we interpret our measurements in the following two scenarios: ... ...

[Would be good to have another exclusion plot for all 5 scenarios ?]

11 Results

A modified frequentist method is used to set limits on the production of signal [22]. The di-jet invariant mass distributions in triply b-tagged events of data, simulated signal, and the normalized background were used as inputs. The value of $\tan\beta$ was varied until the confidence level for signal (CL_S) was $< 5\%$. Figure 20 shows the data, background, and simulated signal at the exclusion limit, for $m_A = 120$ GeV. This is converted to a cross section limit for signal production in Fig. 21, which also shows the expected MSSM Higgs boson production cross section as a function of m_A for $\tan\beta = 80$, and the median expected limit with the background-only hypothesis along with its $\pm$ 1 sd range. The NLO cross sections and their uncertainties from parton distribution functions (PDF) and scale dependence are taken from Refs. [5, 8]. The MSSM cross section shown in Fig. 21 corresponds to no mixing in the scalar top quark sector [23], or $X_t = 0$, where $X_t = A_t - \mu\cot\beta$, A_t is the tri-linear coupling, and the Higgsino mass parameter $\mu = -0.2$ TeV. We also interpret our results in the "maximal mixing" scenario with $X_t = \sqrt{6} \times M_{SUSY}$, where M_{SUSY} is the mass scale of supersymmetric particles, taken to be 1 TeV.

Results for both scenarios of the MSSM are shown in Fig. 23 as limits in the $\tan\beta$ versus m_A plane. The present DØ analysis, based on 260 pb^{-1} of data, excludes a

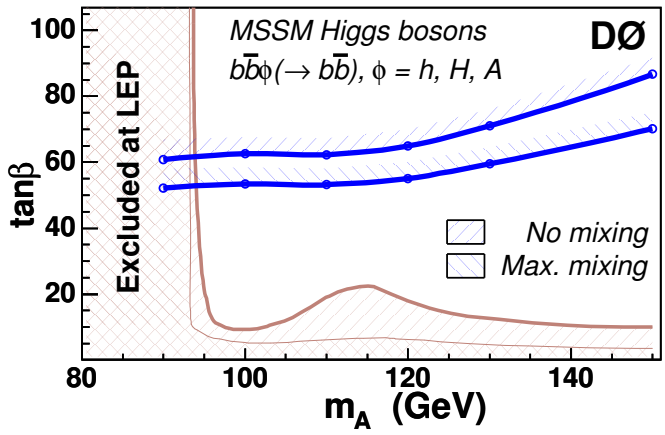

Figure 23: The 95% C.L. upper limit on $\tan\beta$ as a function of m_A for two scenarios of the MSSM, "no mixing" and "maximal mixing." Also shown are the limits obtained by the LEP experiments for the same two scenarios of the MSSM [3].

	Tevatron	LHC
	xxx	yyy
	xxx	yyy
	xxx	yyy
	xxx	yyy

Table 4: Alpgen cross section values for bh, bbj ... processes at the Tevatron and LHC .

significant portion of the parameter space, down to $\tan\beta$= 50, depending on m_A and the MSSM scenario assumed.

12 Prospects for LHC

Table 4 lists the cross section for the SM bh production and the dominant backround (at the Tevatron) processes.

Bla-bla ...

However, the real issue at LHC would be to design an efficient multi-b-jet trigger viable at relatively low Higgs boson masses.

...

13 Acknowledgments

We thank the authors of Refs. [5, 8, 23] for valuable discussions. 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 (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), A.P. Sloan Foundation, Research Corporation, Texas Advanced Research Program, Alexander von Humboldt Foundation, and the Marie Curie Fellowships.

References

[1] H. P. Nilles, Phys. Rept. **110**, 1 (1984); H. E. Haber and G. L. Kane, Phys. Rept. **117**, 75 (1985).

[2] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, "The Higgs Hunter's Guide," Addison-Wesley, 1990.

[3] The LEP Working Group for Higgs Boson Searches, LHWG-Note 2004-01.

[4] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **86**, 4472-4478 (2001).

[5] J. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, Phys. Rev. D **67**, 095002 (2003).

[6] S. Dawson, C. B. Jackson, L. Reina, and D. Wackeroth, Phys. Rev. D **69**, 074027 (2004); S. Dittmaier, M. Krämer, and M. Spira, Phys. Rev. D **70**, 074010 (2004).

[7] J. Campbell *et al.*, hep-ph/0405302.

[8] S. Dawson, C. B. Jackson, L. Reina, and D. Wackeroth, Phys. Rev. Lett. **94**, 031802 (2005).

[9] J. S. Lee *et al.*, Comp. Phys. Comm. **156**, 283 (2004).

[10] M. Carena and H. E. Haber, Prog. Part. Nucl. Phys. **50**, 63 (2003).

[11] E. Boos, A. Djouadi, M. Mühlleitner, and A. Vologdin, Phys. Rev. D **66**, 055004 (2002).

[12] DØ Collaboration, V. Abazov *et al.*, in preparation for submission to Nucl. Instrum. Methods Phys. Res. A.

[13] DØ Collaboration, S. Abachi *et al.*, Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994).

[14] G. C. 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, (Fermilab, Batavia, IL, 2000) p. 47; see Sec. 3.5 for details.

[15] T. Sjöstrand *et al.*, Comp. Phys. Comm. **135**, 238 (2001).

[16] A. Belyaev, E. Boos, "Single top quark tW+X production at the LHC: a closer look", hep-ph/0003260, Phys.Rev. D63 (2001) 034012.

[17] Private meeting with Steve Mrenna, summarized here: http://www-d0.fnal.gov/Run2Physics/higgs/d0_private/groups/hbb/talks/meet_1.html.

[18] J. Campbell, S. Dawson, S. Dittmaier, C. Jackson, M. Kramer, F. Maltoni, L. Reina, M. Spira, D. Wackeroth, S. Willenbrock, "Higgs Boson Production in Association with Bottom Quarks", hep-ph/0405302 v1, 31 May, 2004.

[19] M. L. Mangano *et al.*, JHEP **0307**, 001 (2003).

[20] J. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, Phys. Rev. D **69**, 074021 (2004).

[21] F. Maltoni and T. Stelzer, JHEP **0302**, 027 (2003).

[22] T. Junk, Nucl. Instrum. Methods Phys. Res. A **434**, 435 (1999).

[23] M. Carena, S. Mrenna, and C. E. M. Wagner, Phys. Rev. D **60**, 075010 (1999); ibid, D **62**, 055008 (2000).