ATL-PHYS-2001-022
20/12/2001

The $gg, q\bar{q} \to t\bar{t}b\bar{b}$ process:
matrix element implementation to PYTHIA 6.1

Borut Paul Kersevan

CERN, EP Division, 1211 Geneva 23, Switzerland.

Elżbieta Richter-Wąs[1]

Institute of Computer Science, Jagellonian University;
Institute of Nuclear Physics
30-055 Krakow, ul. Kawiory 26a, Poland.

Abstract

The $t\bar{t}b\bar{b}$ production at LHC is a dominant irreducible background for observability of the Standard Model (SM) or Minimal Super-symmetric Standard Model (MSSM) Higgs boson in the associated production, $t\bar{t}H$, followed by the $H \to b\bar{b}$ decay. In this Note we describe implementation[2] of the complete event generation using massive matrix element for the process $gg, q\bar{q} \to t\bar{t}b\bar{b}$. This implementation is prepared as an *external process* in PYTHIA 6.1 Monte Carlo. The matrix element has been coded using MADGRAPH package, the phase space generation algorithm was adapted following importance sampling technique developed for NEXTCALIBUR package with multi-channel self-optimising approach. For the final optimisation also the VEGAS package has been used. The efficiency of the order of 10% was achieved for generating unweighted events. In the Note we discuss also the cross-section dependencies on the QCD energy scale defined for the process. In addition we show comparison between differential distributions expected for the $t\bar{t}H$ signal and the $t\bar{t}b\bar{b}$ background.

1 Introduction

The $t\bar{t}b\bar{b}$ production at LHC is a dominant irreducible background for the Standard Model (SM) and Minimal Super-symmetric Standard Model (MSSM) Higgs boson search in the associated production, $t\bar{t}H$, followed by the decay $H \to b\bar{b}$.

The potential for the observability of this channel has been carefully studied and documented in [1] and [2]. Proposed analysis requires identifying four b-jets, reconstruction of both top-quarks in the hadronic and leptonic mode and visibility of the peak in the invariant mass distribution of the remaining b-jets. The irreducible $t\bar{t}b\bar{b}$ background contributes about 60-70% of the total background from the $t\bar{t}$ events ($t\bar{t}b\bar{b}$, $t\bar{t}bj$, $t\bar{t}jj$). Expected significance is of 3.6 σ for the Higgs mass of 120 GeV and integrated luminosity $30fb^{-1}$. The expected signal-to-background ratio in the mass window is 24%. The contribution from the ttZ and $Wjjjjjj$ background was estimated to be below 10% of that from the $t\bar{t}$ events.

In the analysis presented in [1] the irreducible background was estimated with PYTHIA 5.7 Monte Carlo. The hard process $gg, q\bar{q} \to t\bar{t}$ was generated and two out of four b-quarks in the event have been generated in the process of the shower cascade from the initial or final state. It required very CPU intensive computing to collect sufficient statistics of the $t\bar{t}b\bar{b}$ background events passing selection procedure. The parton shower approach explored in the procedure gives estimates to which large uncertainties should be assigned, particularly if events are than studied with such exclusive analysis as proposed for this channel. There are several components of the parton shower model, difficult to control systematically, which could lead to biased topology and particles multiplicities, so acceptances of the selection criteria. Let as specify just few of them: effective branching ratio of the cascade $g^* \to (g)b\bar{b}(g)$ splitting, energy scale for the initial and final state shower evolution, angular correlations in the cascade splitting, distribution of the virtuality of the cascading gluon, etc. etc.

The dedicated study [3], comparing cross-sections and acceptances between matrix element calculations evaluated with COMPHEP package [5] and parton shower predictions with PYTHIA 5.7, roughly confirmed the reliability of the parton shower approach. The final estimates for the expected number of background events turned out to be similar in both cases with predictions from the parton shower generation stream being even slightly higher.

In this paper we present complete event generation of the $gg, q\bar{q} \to t\bar{t}b\bar{b}$ process implemented into PYTHIA 6.1 as an *external process*. The gluon-gluon fusion is the dominant production mechanism, the quark-antiquark annihilation contributes only small percentage of the total cross-section. This is similar as for the the $t\bar{t}$ production, where the $q\bar{q} \to t\bar{t}$ process contributes less than 15% of the total cross-section. The matrix element has been coded using MADGRAPH package [6], the phase space generation algorithm was adapted following importance sampling technique developed for NEXTCALIBUR package[12] with multi-channel self-optimising approach following method developed in [8]. For the final optimisation also VEGAS package [15] was used. The efficiency of 10% has been achieved for generating unweighted events [3].

---

[3]Calculation of the matrix elelement itself is rather slow, it takes 8 hours to generate 20000 unweighted (cca 200000 weighted) events on a Pentium III 800 MHz (with 256 k cache).

The paper is organised as follows: In Section 2 we discuss expected cross-sections and differential distributions for the $gg, q\bar{q} \to t\bar{t}b\bar{b}$ production. In Section 3 we compare several distributions for the $t\bar{t}b\bar{b}$ background and the $t\bar{t}H$ signal with an attempt to draw some conclusions concerning possible further optimisation of the selection algorithm proposed in [2]. Section 4 gives detailed description of the algorithms used for the phase space generation. The technical details concerning matrix element implementation into PYTHIA 6.1 and guidances for the users are collected in Section 5.

2   The $gg, q\bar{q} \to t\bar{t}b\bar{b}$ process

The processes $gg, q\bar{q} \to t\bar{t}b\bar{b}$ have been simulated for pp collision with 14 TeV centre-of-mass energy and CTEQ5L structure functions, using events generation with massive $2 \to 4$ matrix element implemented as an external process to PYTHIA 6.1 (see Section 4 and 5). The decays of the top-quarks have been left under control of PYTHIA 6.1 generator.

The total cross-section is very sensitive to the choice of the QCD energy scale used for calculation of that process, just indicating potential large contributions from higher order corrections. The same definition for the factorisation and renormalisation scale is used here. The example values of the total cross-section for implemented choices of the QCD energy scale are given in Table 1[4]. The $q\bar{q} \to t\bar{t}b\bar{b}$ subprocess contributes less than 10% of the total cross-section.

Table 1: *Cross-sections for the $gg, q\bar{q} \to t\bar{t}b\bar{b}$ production and different choices of the QCD energy scale. The 14 TeV centre-of-mass energy and CTEQ5L structure functions were used for the simulation. The $m_H = 120$ GeV and $m_t = 175$ GeV are used for calculating the Q^2_{QCD} in the last row of this table.*

QCD scale	Process	Cross-section This implementation	Cross-section COMPHEP
$Q^2_{QCD} = \hat{s}$	gg	3.94 pb	3.91 pb
	$q\bar{q}$	0.29 pb	0.29 pb
	Total	4.23 pb	4.20 pb
$Q^2_{QCD} = \sum (p_T^{i\,2} + m_i^2)/4 = <m_T^2>$	gg	10.18 pb	10.17 pb
	$q\bar{q}$	0.61 pb	0.61 pb
	Total	10.79 pb	10.78 pb
$Q^2_{QCD} = \sum (p_T^{i\,2})/4 = <p_T^2>$	gg	16.89 pb	16.92 pb
	$q\bar{q}$	0.90 pb	0.91 pb
	Total	17.79 pb	17.83 pb
$Q^2_{QCD} = (m_t + m_H/2)^2$	gg	8.08 pb	8.06 pb
	$q\bar{q}$	0.51 pb	0.51 pb
	Total	8.59 pb	8.57 pb

---

[4]Estimates on the cross-section, that can be found in the literature, have been calculated with COMPHEP code prepared by the authors of the COMPHEP. The [4] quotes cross-section of 3.1 pb for QCD scale $Q^2_{QCD} = (m_t + <p_T>)^2$, CTEQ4L structure functions and after requiring 4 b-quarks with $p_T > 15$ GeV and $|\eta| < 2.5$. For the same QCD scale, but CTEQ2L structure functions and asking only b-quarks from gluon cascade to pass selection $p_T > 15$ GeV and $|\eta| < 3.0$ the quoted cross-section [3] is 11.9 pb.

As a cross-check, the process $gg, q\bar{q} \rightarrow t\bar{t}b\bar{b}$ have been coded independently using COMPHEP package [5]. The same set of diagrams was selected and only the integrating part of the package was used to calculate total cross-section. The choices for the QCD energy scale were kept consistently. A very good agreement between the cross-sections obtained with two independent calculation streams prepared for this study has been achieved.

One can observe very strong scale dependence of the cross-section for the $gg, q\bar{q} \rightarrow t\bar{t}b\bar{b}$ process. Factor four (!!) can be expected on the predicted cross-section when changing from the scale $Q_{QCD}^2 = \hat{s}$ to the scale $Q_{QCD}^2 = <p_T^2>$. This very strong dependency on the energy scale is also observed in the case of the $t\bar{t}H$ production, for recent discussion see [11]. There, the recommended *central* factorisation and renormalisation energy scale is $\mu_0 = (m_t + m_H/2)$. Having in mind that the primary interest of evaluating this background is the Higgs search in the $t\bar{t}H$ production, so with the b-quark system being produced with the invariant mass of the expected Higgs boson, we have also introduced this *central* energy scale, with $m_H = 120$ GeV, as the one of the possible choices.
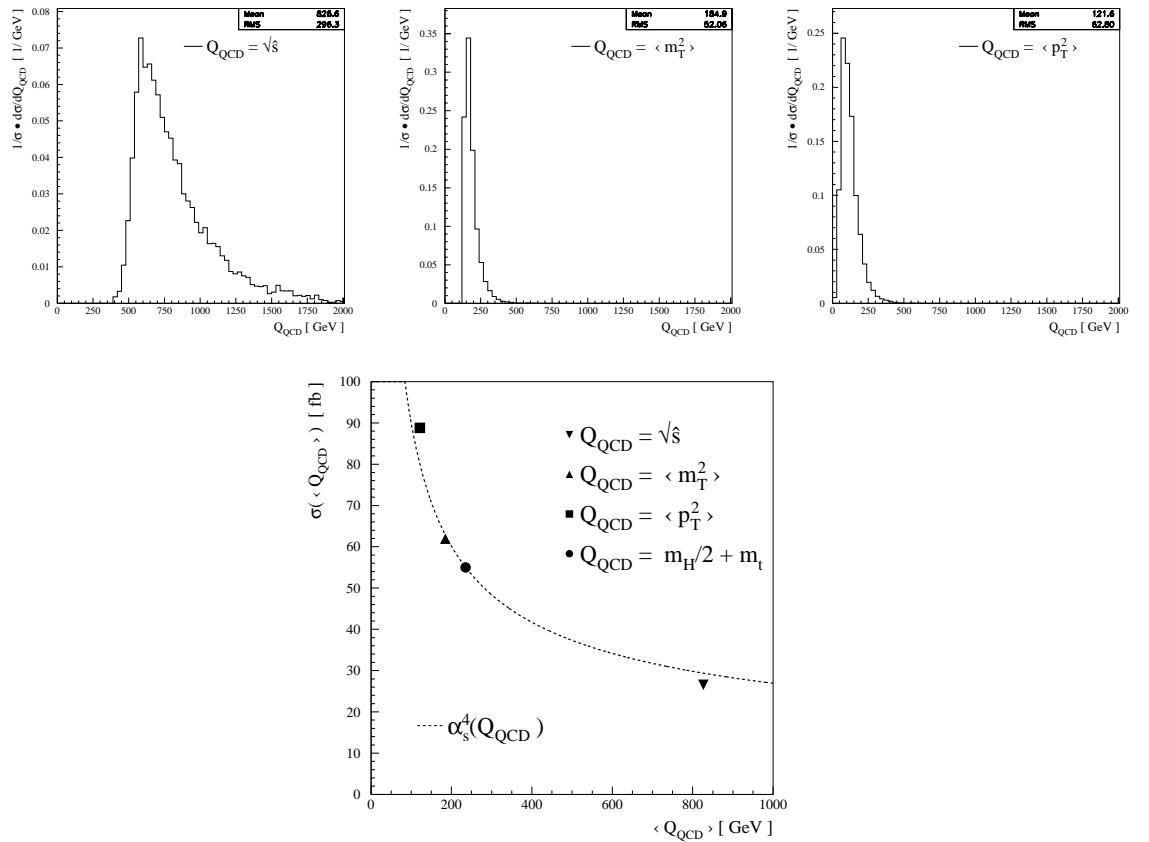

Figure 1: *Top: the Q_{QCD}^2 distributions for $t\bar{t}b\bar{b}$ events with the invariant mass of the b-jets system $m_{bb-jets} = 120 \pm 30$ GeV. Bottom: the total cross-section of accepted events as a function of the averaged Q_{QCD}^2 (for these events).*

3

Fig. 1 shows the distributions of the $Q_{QCD} = \sqrt{Q^2_{QCD}}$ (distributions have been normalised to one), for the $t\bar{t}b\bar{b}$ events with the invariant mass of the b-jets system[5] $m_{bb-jets} = 120 \pm 30$ GeV. The distribution is well collimated around average value when Q^2_{QCD} is defined as $< m^2_T >$ or $< p^2_T >$ while it is much more broad when Q^2_{QCD} is defined as $\hat{s}$. There are very similar in shape for separate $gg \to t\bar{t}b\bar{b}$ and $q\bar{q} \to t\bar{t}b\bar{b}$ contributions. The total cross-section for accepted events as a function of the averaged Q^2_{QCD} (for these events) is shown in the bottom plot. It can be noted that the cross-section decreases rather fast with increasing value of the average $< Q^2_{QCD} >$. Also shown is the $\alpha^4_s(Q_{QCD})$ dependence scaled to match the cross-section at $Q_{QCD} = (m_H/2 + m_t)$ with $m_H = 120$ GeV, it being the only calculated cross-section point with fixed scale. The other cross-sections are shown to follow the expected $\alpha^4_s(Q_{QCD})$ dependence rather well, while the deviations are induced by the structure function dependence on the Q^2_{QCD} scale, most noticeable at $Q^2_{QCD} = \hat{s}$ value. The deviations induced by the structure functions dependence on the Q^2_{QCD} scale is different for the gg and $q\bar{q}$ contributions, as can be concluded from results given in Table 1.

In the series of plots below the most relevant differential distributions for the top-quarks and b-quarks, obtained without imposing any kinematic acceptances, are shown. The presented kinematic features are interesting to keep in mind for the optimisation of the signal-and-background selection. On should however not forget that the study will approach the realistic case only when the QCD showering and detector effects are included and the same distributions are plotted for the reconstructed top-quarks and b-jets with the mass of the b-jet system in the range of interest for the $t\bar{t}H$ channel. The contribution from the $gg \to t\bar{t}b\bar{b}$ and $q\bar{q} \to t\bar{t}b\bar{b}$ subprocesses are shown separately. Plots are normalised to the unity cross-section and energy scale $Q^2_{QCD} = (m_t + mH/2)$ was used for the generation.

- Fig. 2: the transverse momenta distribution of the b-quarks from the hard process (left) and from the top-decays (right). As expected, the mean of the distribution is almost two times higher for b-quarks from top-quarks decays ($< p_T >= 75$ GeV) than for b-quarks from the hard process ($< p_T >= 35$ GeV).

- Fig. 3: the pseudo-rapidity distribution of the top-quarks (left), b-quarks from the hard process (middle) and b-quarks from the top-quark decays (right). The b-quarks from top-quarks decays are produced more centrally ($RMS = 1.4$) than b-quarks from the hard process ($RMS = 1.9$). As the acceptance for the b-jet identification is rather large, up to the pseudo-rapidity $|\eta| < 2.5$, this broader distribution might reduce overall acceptance at most by some percentage only.

- Fig. 4: the cone separation for the b-quarks from the hard process (left) and for the b-quarks from top-quarks decays (right). The b-quarks from the hard process are produced dominantly close to each other, while these from the top-quarks decays have tendency to be more back-to-back.

---

[5]This would makes distributions directly relevant for the $t\bar{t}H$ analysis. For details on the jet reconstruction see [17]. See also next Section.

- Fig. 5: the cone separation between the hard-process b-quarks and top-quarks (left); cone separation between $t\bar{t}$ and $b\bar{b}$ systems from the hard process (right). The tendency of the same-flavour quark systems to be produced back-to-back is observed also here.

- Fig. 6: the transverse momenta distribution of the top-quarks system (left), b-quarks system from the hard process (middle) and b-quarks system from the top-quarks decays (right). The b-quarks system from the top-quarks decays is produced with large transverse momenta, ($< p_T^{bb} >= 90$ GeV and RMS=45 GeV), the spectrum is harder than for the top-quarks system itself. Contrary, the b-quarks system from the hard-process is produced with relatively small transverse momenta, almost half of events with $p_T^{bb} < 60$ GeV. The spectra are noticeable harder for the $gg \rightarrow t\bar{t}b\bar{b}$ events than for the $q\bar{q} \rightarrow t\bar{t}b\bar{b}$ events, in particular the spectrum for the top-quark system.

- Fig. 8: the invariant mass distribution for all b-quarks pairs, ordered with increasing invariant mass.
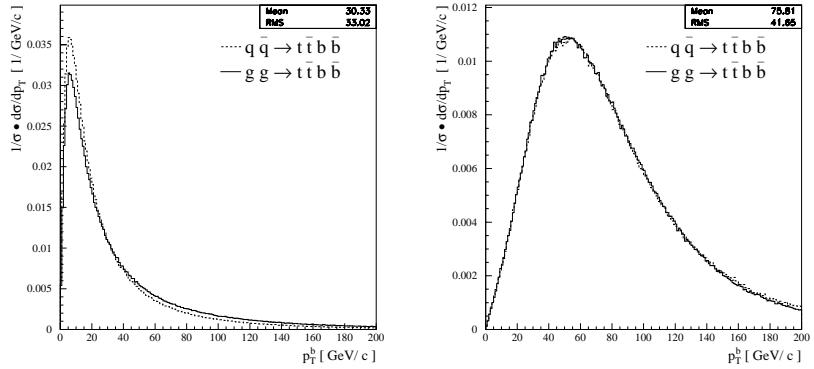
Figure 2: *The transverse momenta distribution of the b-quarks from the hard process (left) and of the b-quarks from the top-quarks decays (right).*
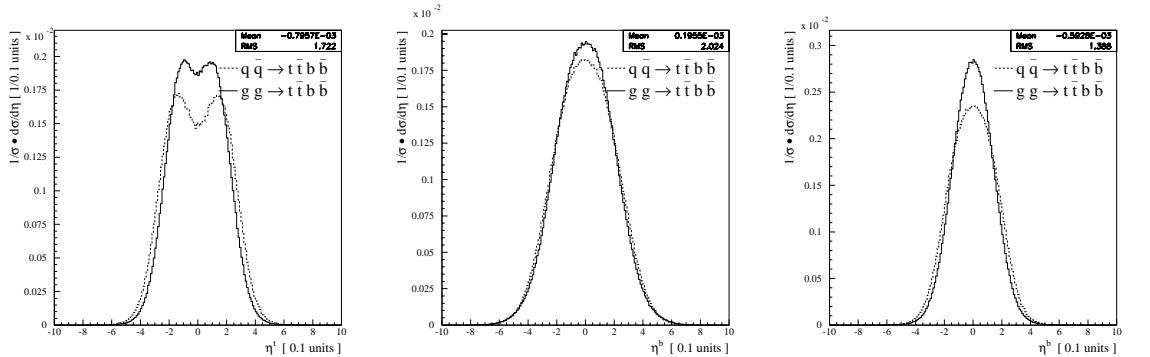
Figure 3: *The pseudo-rapidity distribution of the top-quarks (left), b-quarks from the hard process (middle) and b-quarks from the top-quarks decays (right).*
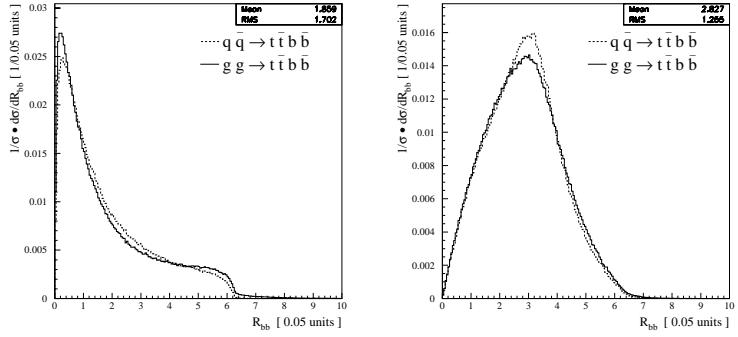
Figure 4: *The cone separation for b-quarks from the hard process (left) and for b-quarks from the top-quarks decays (right).*
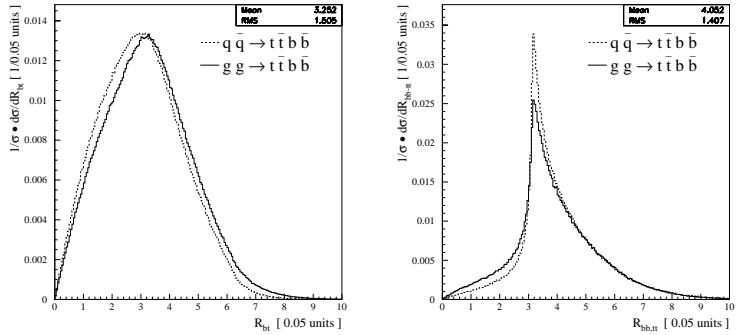
Figure 5: *The cone separation between b-quarks from the hard process and top-quarks (left) and between top-quarks and b-quarks systems from the hard process (right).*
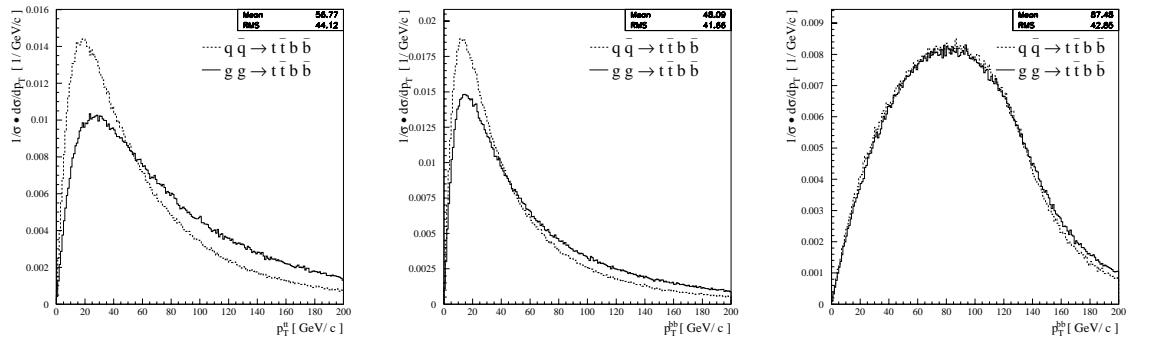
Figure 6: *The transverse momenta distribution of the top-quarks system (left), b-quarks system from the hard process (middle) and b-quarks system from the top-quarks decays (right).*
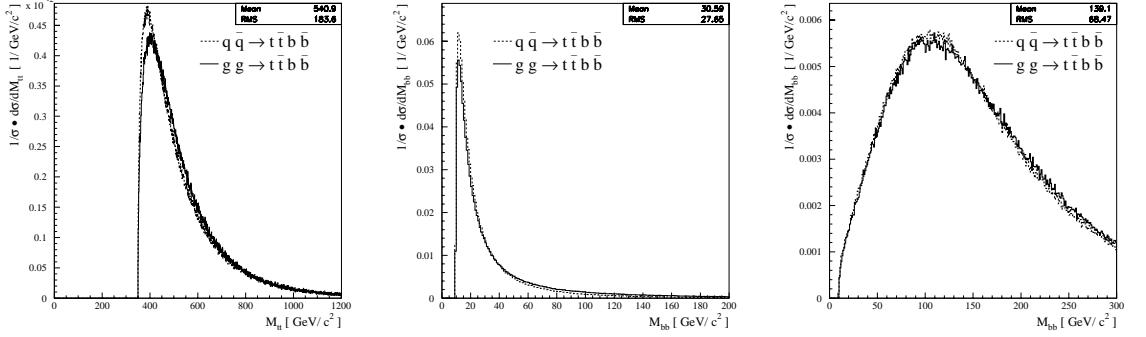
Figure 7: *The invariant mass distribution of the top-quarks system (left), b-quarks system from the hard process (middle) and b-quarks system from the top-quarks decays (right).*
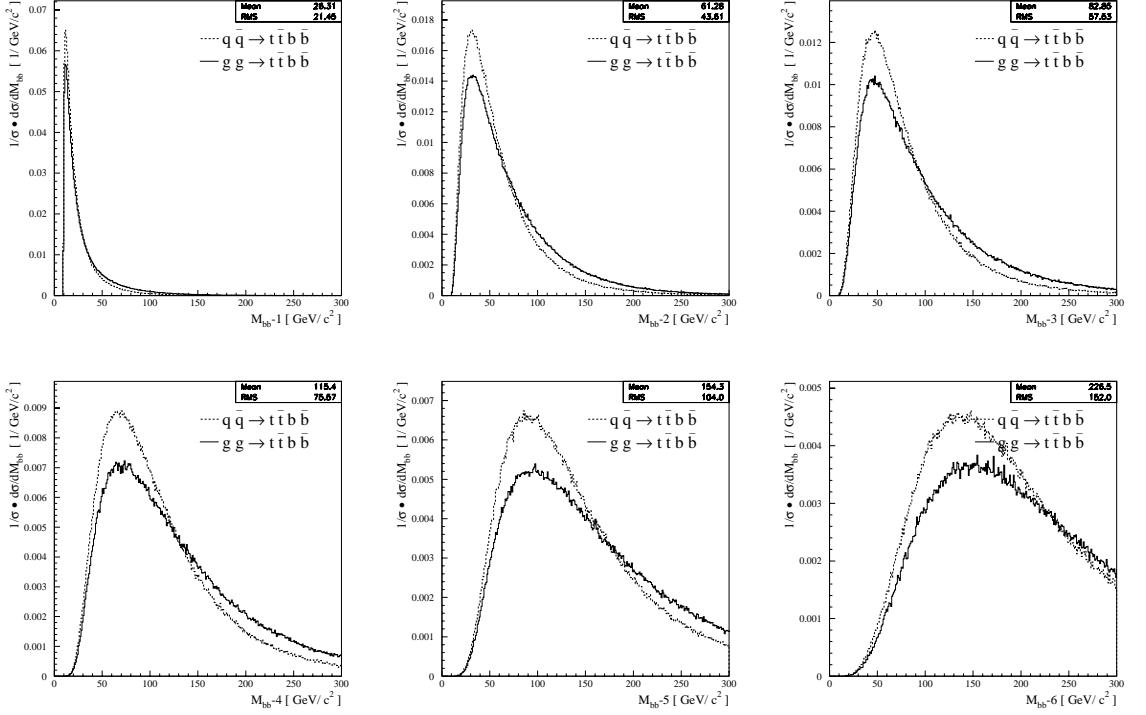
Figure 8: *The invariant mass distribution of all b-quarks pairs, ordered with increasing invariant mass of the system.*

These distributions are quite similar for the $gg \to t\bar{t}b\bar{b}$ and $q\bar{q} \to t\bar{t}b\bar{b}$ events. The noticeable difference is observed for the transverse momentum distribution of the top-quark system, which is much harder in the $gg \to t\bar{t}b\bar{b}$ events. Also the invariant mass distribution for all b-quarks pairs is harder in these events and the difference increases for the combinations with the higher invariant mass.

8

All above shown distributions are very similar for the very different choices of the definition of the Q^2_{QCD} energy scale. Roughly one can conclude that only the overall normalisation changes for different Q^2_{QCD} energy scale.

3   The $t\bar{t}H$ signal and $t\bar{t}b\bar{b}$ irreducible background.

The main interest in simulating the $t\bar{t}b\bar{b}$ background is search for the Higgs boson of the mass around 120 GeV in SM and MSSM model. Of the immediate interest is therefore understanding how the presented above distributions compare with these expected for the signal events. To make this comparison more realistic we process events first with the fast detector simulation and reconstruction [17] and compare distributions for events accepted by the simplified, with respect to that foreseen for the Higgs search [2], selection procedure. The simplified selection procedure requires:

- One isolated lepton and four identified b-jets. The kinematic thresholds are: $p_T^{jet}, p_T^{\ell} >$ 15 GeV and $|\eta| < 2.5$.

- Reconstruction of both top-quarks by taking true generated W-boson momenta and pairing with the b-jets by requiring chosen combinations to minimise the $\chi^2 = (m_1 - m_t)^2 + (m_2 - m_t)^2$.

- The invariant mass of the remaining pair of b-jets to be within window $m_{bb-jet} =$ 120 ± 30GeV.

The above procedure, although simplified with respect to the full one, is nevertheless adequate to study the final differences between topological features of signal and background events.

The cross-sections and acceptances of the above selection for the $t\bar{t}H$ signal and $t\bar{t}b\bar{b}$ background are summarised in Table 2. The signal events have been generated with PYTHIA 6.1 Monte Carlo and value of the cross-section for signal events should be considered as the reference one only [6]. Recent calculations including NLO corrections [11] indicates that one should expect the total cross-section to be around 640 fb at *central* energy scale, $Q^2_{QCD} = (m_H/2 + m_t)^2$, while the LO cross-section is 580 fb at *central* scale and 380 fb at scale $Q^2 = \hat{s}$.

In the series of plots below we show the most relevant differential distributions for the reconstructed top-quarks and b-jets, obtained after fast simulation and reconstruction and described above selection procedure, including requiring the mass window $m_{bb-jet} =$ 120 ± 30 GeV constraint on the remaining b-jets system. Plots are normalised to the unity cross-section in pb and energy scale $Q^2_{QCD} = (m_H/2 + m_t)^2$ with $m_H = 120$ GeV was used for the generation.

- Fig. 9: the transverse momenta distribution of the b-jets assigned to the Higgs decay (left) and these assigned to the top-decays (right). The mean of the distributions is almost comparable for both classes of b-jets and is only slightly harder for the signal than for the background events.

---

[6]In PYTHIA the Q^2 is defined as *default* with the MSTP(39)=2 - switch specific to this process: chooses the scale used for structure functions and ISR, set to maximal transverse mass of the two outgoing top quarks.

- Fig. 10: the pseudo-rapidity distribution of the reconstructed top-quarks (left), b-jets assigned to the Higgs decay (middle) and b-jets assigned to the top-quarks decays (right). The distributions are very similar for the signal and the background events, very similar are also distributions for both classes of b-jets.

- Fig. 11: the transverse momenta distribution of the reconstructed top-quarks system (left), b-jets system assigned to the Higgs decay (middle) and b-jets system assigned to the top-quarks (right). The transverse momenta spectrum of the b-jets system assigned to the Higgs decay is slightly softer for the signal than for the background events, reverse is observed for the transverse momenta of b-jets system assigned to the top-quarks.

- Fig. 12: the cone separation for the b-jets system assigned to the Higgs decay (right) and for the b-jets system assigned to the top-quarks (left). The system assigned to the Higgs decay preserves tendency observed at the parton level to have smaller separation than the system assigned to the top-quarks decay, but the difference is less apparent than shown in Fig. 4. This is easy to explain as the mass required of the system assigned to the Higgs decay is relatively high which forces larger cone separation. As a consequence of the kinematic selection the distributions from the signal and background events are very similar.

- Fig. 13: the cone separation for the b-jets assigned to the Higgs decay and reconstructed top-quarks (left) and between reconstructed top-quarks system and b-jets system assigned to the Higgs decay. In both cases cone separation is peaked at the back-to-back configurations; distribution is much more narrower for the $R_{bb,tt}$ than for the $R_{b,t}$ separations. Again, the distributions from the signal and background events are very similar.

- Fig. 14: the invariant mass distribution of the reconstructed top-quarks pair (left), b-jets pair assigned to the Higgs decays (middle) and b-jets pair assigned to the top-quarks decays (right). Explicitly middle plots shows the mass window for b-jet pair system assigned to Higgs decay used for events selection. The invariant masses of the b-jet pair assigned to the top-quarks decays and of the reconstructed top-quarks system are only slightly harder for signal than for the background.

- Fig. 15: the invariant mass distribution of the b-jet systems (any pair) ordered with increasing invariant mass. These distributions were indicated in [18] as having crucial discrimination potential for the $t\bar{t}b\bar{b}$ background rejection in searches scenarios at Tevatron energies. Plotted, using presented here implementation of the $t\bar{t}b\bar{b}$ matrix element show no noticeable differences for signal and background distributions in LHC running conditions.

- Fig. 16: the rapidity difference between the b-jets system assigned to the Higgs decay (left) and the cone separation between b-jet system and the top-quark nearest in the angle, calculated using rapidity (right). These two variables were proposed very recently in [19] to increase efficiency for background rejection. Again, the discrimination power observed on the plots seems not to be very strong.

The conclusions which can be drawn from the above comparisons are that the topology of accepted signal and background events is very similar, at least on the presented set of the distributions. The differences of the distributions for the $t\bar{t}b\bar{b}$ background presented in above plots, Figs. 9 - 16 and respective ones shown in Figs. 2 - 15 comes dominantly from not accepting events with initial low invariant mass of the b-jets system. Some the differences in the shape of the distributions for b-quark system from the hard process and from the top-quarks decays, are less pronounced now due to the detector effects and unavoidable smearing of the kinematic variables (angles, energies) during reconstruction procedure. It seems that there is not much room left for the further optimisation of the discrimination criteria with the classical cuts methods.

There is also obvious that the strong sensitivity of the background cross-section to the choice of the QCD energy scale is a serious problem for firmly specifying expected number of background events from the $t\bar{t}b\bar{b}$ production. Only calculating higher order corrections, like for the $t\bar{t}H$ signal [11], might help to resolve uncertainties which have to be assigned to the estimates for that background. It is also obvious that it is important to foresee control channels which will help to understand/calibrate this background with the experimental data.

Table 2: *Cross-sections and efficiencies for the $gg, q\bar{q} \to t\bar{t}b\bar{b}$ and $gg, q\bar{q} \to t\bar{t}H, H \to b\bar{b}$ production with one $W \to \ell\nu$ decay (both flavours). The 14 TeV centre-of-mass energy and CTEQ5L structure functions were used for the simulation. Fiducial cuts for lepton and b-jets are: transverse momenta $p_T > 15$ GeV and pseudo-rapidity $|\eta| < 2.5$. The variable y denotes rapidity, b-tagging and lepton identification efficiencies are not included, jets reconstruction and b-jet labelling efficiencies taken into account, see [17]. Numbers have been rounded to two digits only.*

Selection $gg \to t\bar{t}b\bar{b}$	$Q^2 = \hat{s}$	$Q^2 = <m_T^2>$	$Q^2 = <p_T^2>$	$Q^2 = (m_t + m_H/2)^2$
σ	3 940 fb	10 180 fb	16 890 fb	8 080 fb
$\sigma \times BR$	1 150 fb	2 970 fb	4 920 fb	2350 fb
$p_T^b > 15$ GeV $|\eta^b| < 2.5$	160 fb	380 fb	550 fb	330 fb
acceptance	14 %	13 %	11 %	14 %
$p_T^\ell > 15$ GeV $|\eta^\ell| < 2.5$	130 fb	310 fb	450 fb	270 fb
acceptance	81 %	81 %	81 %	81 %
$m_{bb-jets} = 120 \pm 30$ GeV	30 fb	60 fb	80 fb	50 fb
acceptance	19 %	19 %	19 %	19 %
$|y_{bb} - y_t| < 1$	20 fb	40 fb	50 fb	30 fb
acceptance	67 %	64 %	63 %	62 %

Selection $gg \to t\bar{t}H, H \to b\bar{b}$	$Q^2 = default$	$Q^2 = \hat{s}$
σ	270 fb	240 fb
$\sigma \times BR$	80 fb	70 fb
$p_T^b > 15$ GeV $\|\eta^b\| < 2.5$	40 fb	40 fb
acceptance	56 %	56 %
$p_T^\ell > 15$ GeV $\|\eta^\ell\| < 2.5$	40 fb	30 fb
acceptance	80 %	80 %
$m_{bb-jets} = 120 \pm 30$ GeV	30 fb	20 fb
acceptance	71 %	71 %
$\|y_{bb} - y_t\| < 1$	20 fb	20 fb
acceptance	84 %	85 %

Selection $q\bar{q} \to t\bar{t}b\bar{b}$	$Q^2 = \hat{s}$	$Q^2 = <m_T^2>$	$Q^2 = <p_T^2>$	$Q^2 = (m_t + m_H/2)^2$
σ	288 fb	605 fb	904 fb	513 fb
$\sigma \times BR$	84 fb	176 fb	263 fb	149 fb
$p_T^b > 15$ GeV $\|\eta^b\| < 2.5$	10 fb	21 fb	28 fb	19 fb
acceptance	12 %	12 %	11 %	13 %
$p_T^\ell > 15$ GeV $\|\eta^\ell\| < 2.5$	8 fb	17 fb	22 fb	15 fb
acceptance	80 %	80 %	80 %	80 %
$m_{bb-jets} = 120 \pm 30$ GeV	2 fb	3 fb	4 fb	3 fb
acceptance	20 %	20 %	20 %	20 %
$\|y_{bb} - y_t\| < 1$	1 fb	2 fb	3 fb	2 fb
acceptance	70 %	68 %	67 %	69 %

Selection $q\bar{q} \to t\bar{t}H, H \to b\bar{b}$	$Q^2 = default$	$Q^2 = \hat{s}$
σ	131 fb	126 fb
$\sigma \times BR$	38 fb	37 fb
$p_T^b > 15$ GeV $\|\eta^b\| < 2.5$	19 fb	19 fb
acceptance	50 %	50 %
$p_T^\ell > 15$ GeV $\|\eta^\ell\| < 2.5$	15 fb	15 fb
acceptance	80 %	80 %
$m_{bb-jets} = 120 \pm 30$ GeV	11 fb	11 fb
acceptance	73 %	72 %
$\|y_{bb} - y_t\| < 1$	9 fb	9 fb
acceptance	82 %	82 %
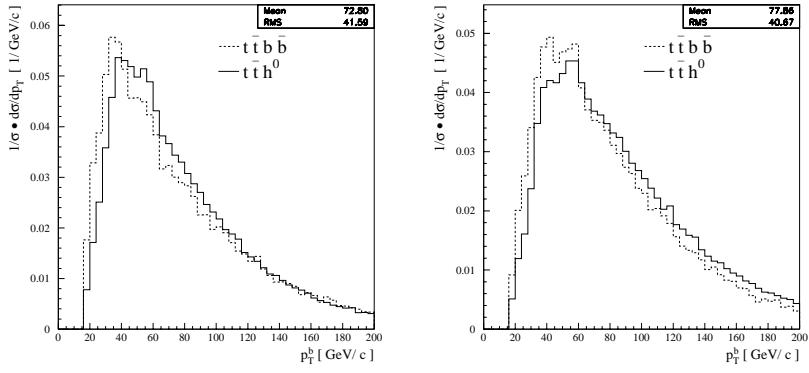
Figure 9: *The transverse momenta distribution of the b-jets assigned to the Higgs decay (left) and these assigned to the top-quarks decays (right).*
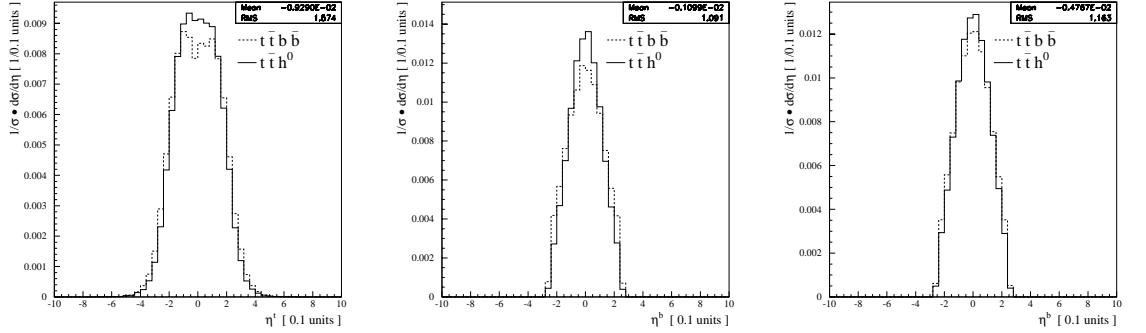
Figure 10: *The pseudo-rapidity distribution of reconstructed top-quarks (left), b-jets assigned to the Higgs decay (middle) and b-jets assigned to the top-quarks decays (right).*
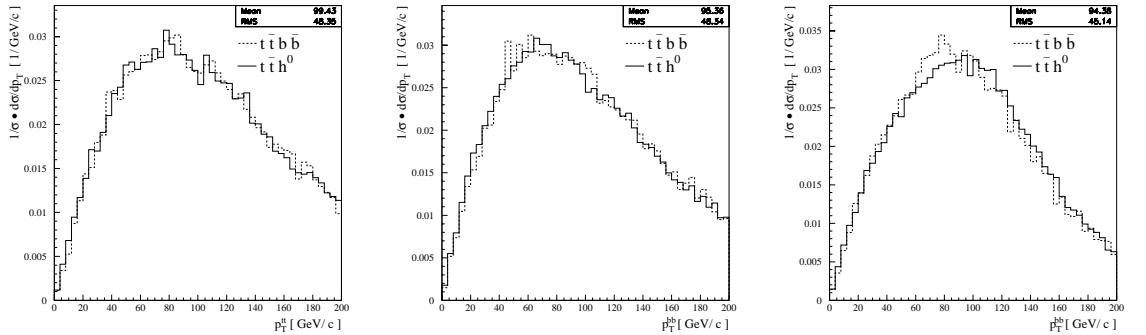
Figure 11: *The transverse momenta distribution of the reconstructed top-quarks system (left), b-jets system assigned to the Higgs decay (middle) and b-jet system assigned to the top-quarks decays (right).*

13
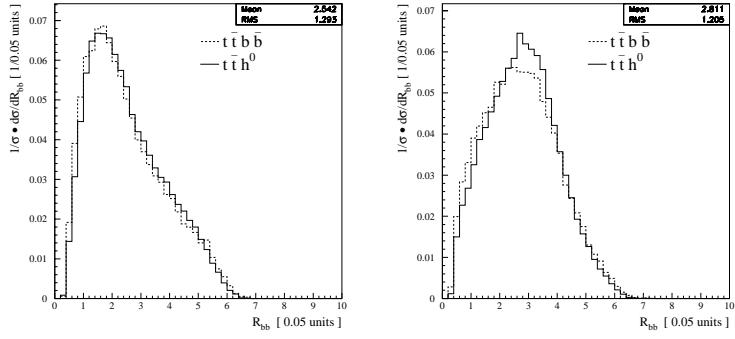
Figure 12: *The cone separation for the b-jets system assigned to the Higgs decay (right) and for the b-jets system assigned to the top-quarks pair (left).*
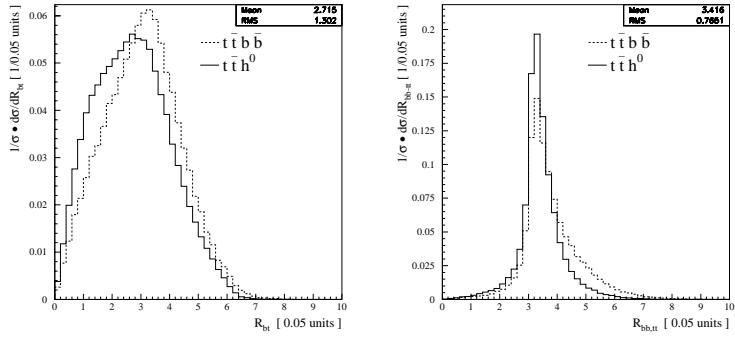
Figure 13: *The cone separation for the b-jets assigned to the Higgs decay and reconstructed top-quarks (left) and between reconstructed top-quarks system and b-jets system assigned to the Higgs decay.*
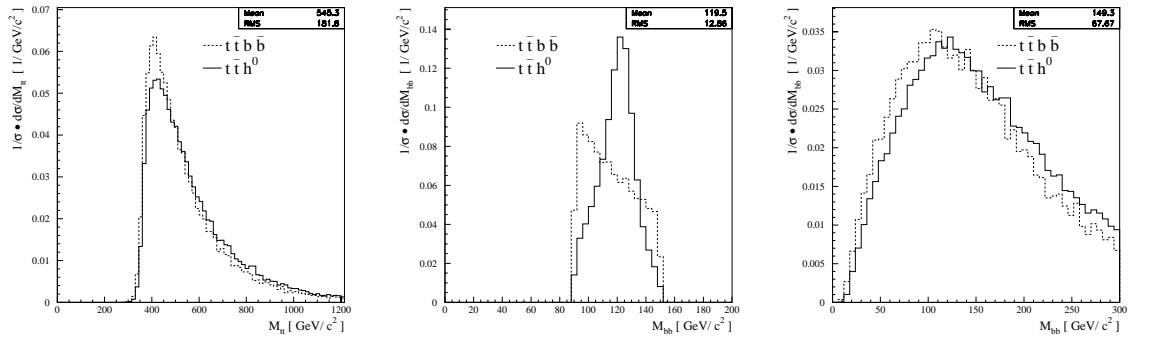
Figure 14: *The invariant mass distribution of the reconstructed top-quarks pairs (left), b-jets pair assigned to the Higgs decays (middle) and b-jets pair assigned to the top-quarks decays (right).*

14
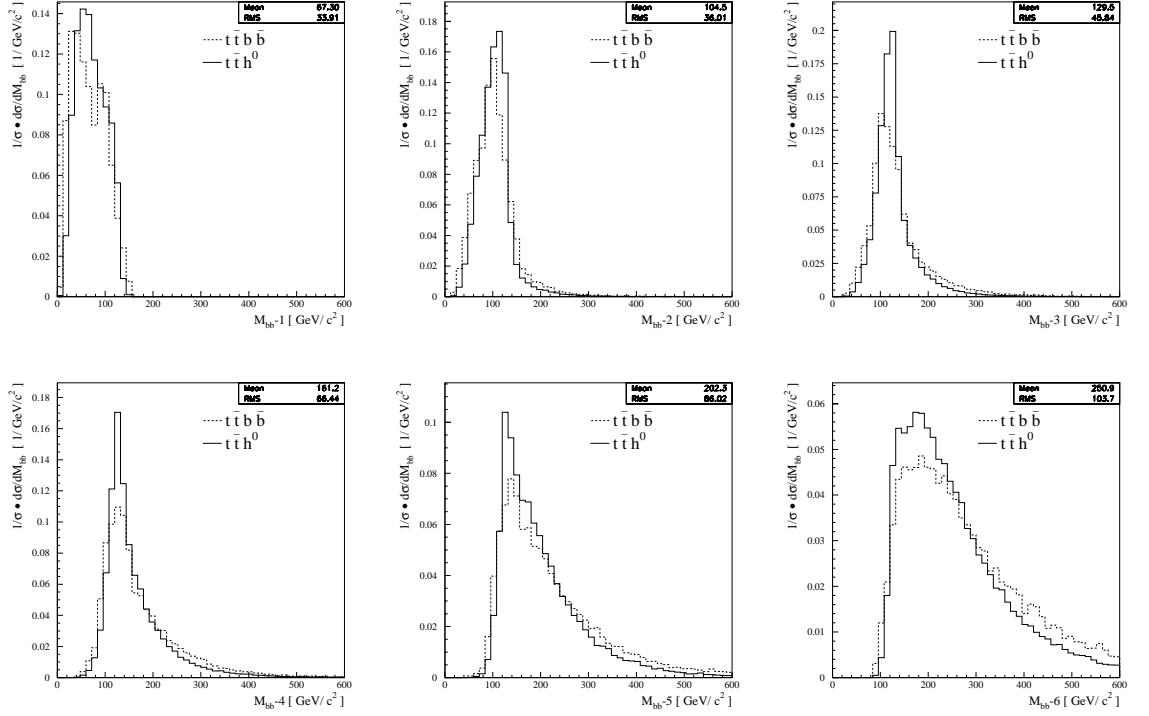
Figure 15: *The invariant mass distribution of the b-jets systems (any pair) ordered with the increasing invariant mass.*
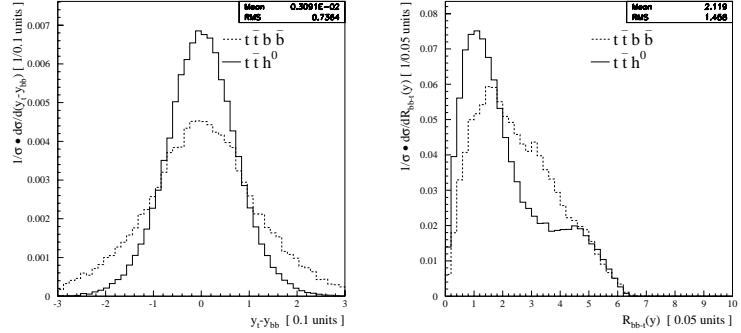
Figure 16: *The rapidity difference between the b-jets system assigned to the Higgs decay (left) and the cone separation between b-jet system and the top-quark nearest in the angle, calculated using rapidity (right).*

15

4 Phase-space generation

4.1 The $gg, q\bar{q} \rightarrow t\bar{t}b\bar{b}$ processes

There are in total 38+7 Feynman diagrams contributing to the $gg, q\bar{q} \rightarrow t\bar{t}b\bar{b}$ production, shown in Fig. 17. The squared matrix element of the processes were obtained by using the MADGRAPH [6] package and take properly into account the masses and helicity contributions of participating particles. The matrix elements were consequently modified to include the running coupling constant $\alpha_S(Q^2)$ calculated using subroutines coded in PYTHIA 6.1 [9].

The four-fermion phase space corresponding to the processes at hand was modelled using the importance sampling technique based on the procedures implemented in the e^+e^- event generators FERMISV [13], EXCALIBUR [14] and NEXTCALIBUR [12]. For the present implementation, the phase space was generated using eleven different kinematic diagrams (*channels*) for the $gg \rightarrow t\bar{t}b\bar{b}$ and four channels for the $q\bar{q} \rightarrow t\bar{t}b\bar{b}$ process. The relative weight between contributions of each sampling channel was subsequently obtained by using multi-channel self-optimising approach [8]. Eventually, additional smoothing of the phase space was obtained by using a modified VEGAS routine to improve the generation efficiency. The procedure of multi-channel importance sampling used in the event generation can briefly be outlined as follows. An analytically integrable function $g(\vec{\Phi})$, which aims to approximate the peaking behaviour of the differential cross-section dependence on various kinematic quantities is introduced as:

$$d\sigma = s(\vec{\Phi})\, d\vec{\Phi} = \frac{s(\vec{\Phi})}{g(\vec{\Phi})} \cdot g(\vec{\Phi})\, d\vec{\Phi} = w(\vec{\Phi})\, g(\vec{\Phi})\, d\vec{\Phi}, \tag{1}$$

The function $g(\vec{\Phi})$ has to be unitary, i.e. a normalised probability density:

$$\int g(\vec{\Phi}) d\vec{\Phi} = 1. \tag{2}$$

Since the peaking behaviour of $s(\vec{\Phi})$ can be very complex due to several possible topologies introduced by the large number of contributing Feynman diagrams, the function $g(\vec{\Phi})$ is composed as a weighted sum of several channels $g_i(\vec{\Phi})$, each adapted to a certain event topology:

$$\int g(\vec{\Phi}) = \sum_i \alpha_i \cdot g_i(\vec{\Phi}). \tag{3}$$

The values of relative weights α_i are determined from multi-channel self-optimisation procedure in order to minimise the variance of the weights $w(\vec{\Phi})$ [8]. The phase space points are than sampled from the function $g(\vec{\Phi})$, first by randomly choosing a channel i according to the relative frequencies α_i and then deriving the required four momenta from the chosen $g_i(\vec{\Phi})$ using unitary algorithms.

The differential of the total production cross-section for the process $gg, q\bar{q} \rightarrow t\bar{t}b\bar{b}$ can be expressed with the formula:

$$d\sigma = \sum_{a,b} f_a(x_1, Q^2) f_b(x_2, Q^2) \frac{|\mathcal{M}|^2}{(2\pi)^8 (2\hat{s})} dx_1 dx_2 d\vec{\Phi}, \tag{4}$$

where $f_{a,b}(x, Q^2)$ represent the gluon or (anti)quark structure functions, $|\mathcal{M}|^2$ the squared matrix element divided by the flux factor $2\hat{s}$ and $d\vec{\Phi}$ denotes the phase space differential
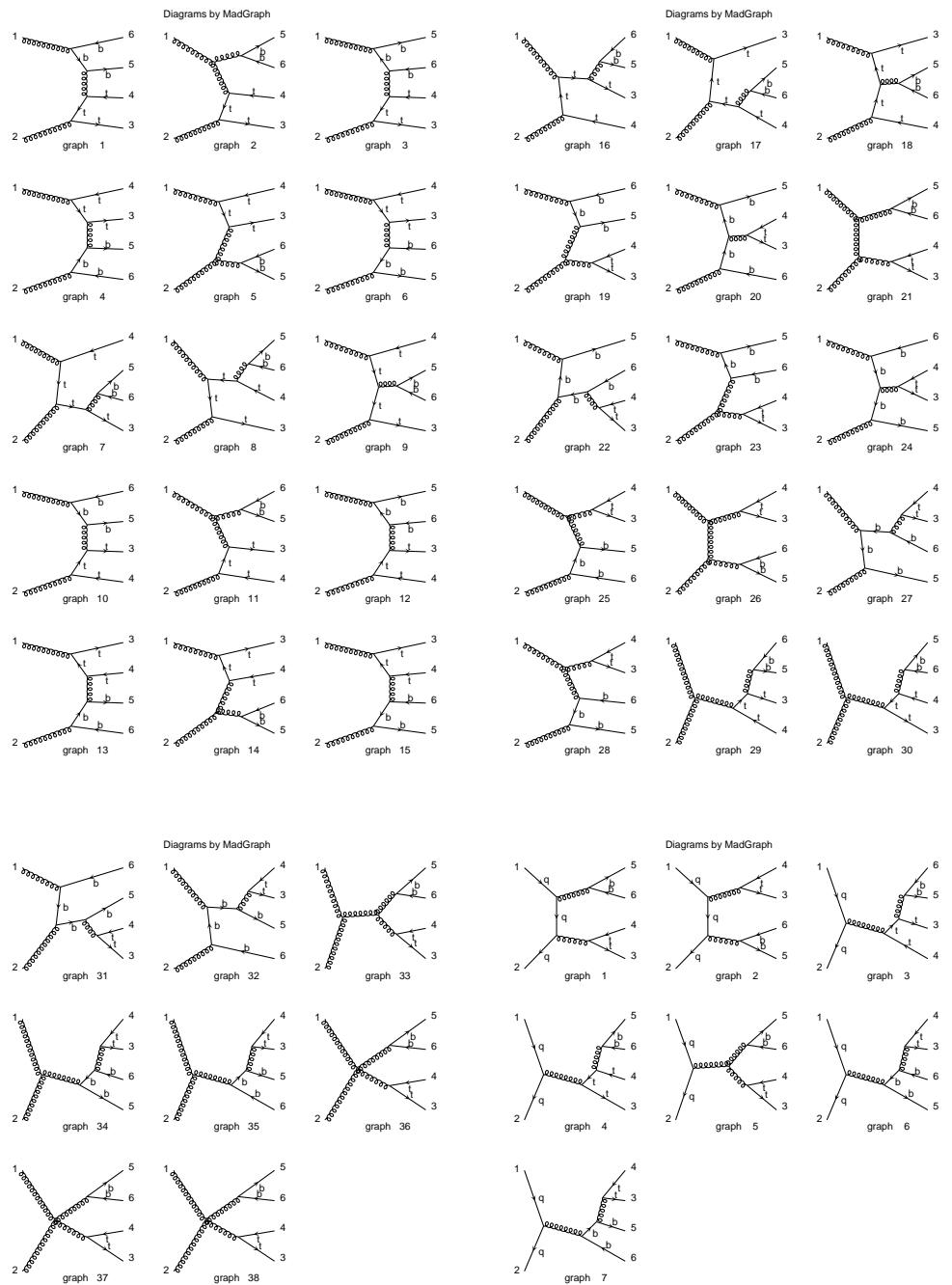
Figure 17: *The Feynman diagrams for the processes* $gg, q\bar{q} \to t\bar{t}b\bar{b}$.

of the four final-state particles. The quantity $\hat{s} = x_1 \, x_2 \, s$ is the effective centre-of-mass energy, and the sum $\sum_{a,b}$ runs in over all possible quark-antiquark combinations $(a, b = u, d, s, c, \bar{u}, \bar{d}, \bar{s}, \bar{c})$. In case of gg initial state the sum has only one term with $a = b = g$.

Alternatively, in order to use the `PYTHIA 6.1` built-in structure functions $xf(x, Q^2)$, one has to rewrite the differential cross-section to the form:

$$d\sigma = \sum_{a,b} x_1 f_a(x_1, Q^2) \, x_2 f_b(x_2, Q^2) \frac{|\mathcal{M}|^2}{(2\pi)^8 (2s)} dy \frac{d\tau}{\tau^2} d\vec{\Phi}, \tag{5}$$

with the two new variables given by $\tau = x_1 \cdot x_2$ and $y = 0.5 \log(x_1/x_2)$. The matrix element used in the calculation depends explicitly on the four-momenta of the incoming gluons and outgoing fermions:

$$\mathcal{M} = \mathcal{M}(q_1, q_2, p_1, p_2, p_3, p_4), \tag{6}$$

where the p_i represent the four-momenta of the final state particles and q_1, q_2 the four-momenta of the incident gluons. Consequently, all the four-momenta have to be explicitly generated.

While generating the $q\bar{q} \rightarrow t\bar{t}b\bar{b}$ events, an additional step is required to pick the flavour of the incoming quark-antiquark pair. The selection is again done using importance sampling, using as sampling weights:

$$\lambda_i = x_1 f_{q_i}(x_1, Q^2) \cdot x_2 f_{\bar{q}_i}(x_2, Q^2), \tag{7}$$

where the index i runs over all possible quark-antiquark combinations. The matrix element of the hard scattering process is for the given (pure QCD) processes flavour independent, since the the incoming quarks are treated as massless. Only four flavours (u, d, s, c) of the incoming quarks are considered at the moment, the contributions of the incoming b quark pair are excluded from calculation due to the very high suppression induced by the structure functions.

In the event generation procedure, a generation channel is chosen by weighted sampling using the set α_i. Next, the values of τ are sampled from a distribution:

$$\frac{1}{(\tau)^\mu} \qquad \mu \sim 1, \tag{8}$$

and y from the distribution

$$\frac{1}{\cosh(\nu \, y)} \qquad \nu \sim 1, \tag{9}$$

as used in `PYTHIA` standard phase-space algorithm [7]. From the two values x_1 and x_2, the momenta q_1, q_2 of incoming particles and the effective centre-of-mass energy $\sqrt{\hat{s}}$ are derived. In the following step, the four momenta of the final state particles p_i are sampled by re-parametrising the general four-body phase space:

$$d\vec{\Phi} = \left(\prod_{i=1}^{4} d^4 p_i \delta(p_i^2 - m_i^2) \theta(p_i^0) \right) \delta^4 (q_1 + q_2 - \sum_{i=1}^{4} p_i). \tag{10}$$

18

in terms of kinematic variables that are expected to exhibit the strongest peaking behaviour for a specific channel topology and consequently introducing appropriate sampling functions for these variables.

One of the principal sampling procedures used is an extension of the three-body phase-space sampling, developed for the phase space generation in `NEXTCALIBUR` program, as described in [12] (kinematic channel `MULTI1`, Appendix A). In order to describe the features implemented in the extension of the procedure one has first to recapitulate its basic principles (a detailed description is given in [12]): The procedure used in `MULTI1` starts with splitting the four-body phase space into a three-body times a two-body decay:

$$\int d\Phi_4 = \int ds_{34} \, I_3(p_1, p_2, p_{34}) \, I_2(p_3, p_4), \tag{11}$$

where the integral I_3 represents the three-body and I_2 the two-body phase space, the labels p_i, $i = 1, 4$ represent the four-momenta of the produced particles with $p_{34} = p_3 + p_4$ and $s_{34} = p_{34}^2$. The principal issue thus becomes how to parametrise the (more difficult) three body phase-space integral $I_3(P_1, P_2, P_3)$:

$$I_3(P_1, P_2, P_3) = \int \prod_{i=1}^{3} d^4 P_i \delta(P_i^2 - M_i^2)\theta(E_i). \tag{12}$$

It has to be stressed that the masses M_i in the above on-shell constraint do not necessarily correspond to physical masses of the elementary particles but can also be values derived by a pre-sampling procedure (e.g. $M_3^2 = s_{34}$ in the above example, i.e. the M_3 can represent the off-shell mass of a gluon decaying into p_3 and p_4). Defining the reduced energies x and y as:

$$E_1 = \frac{\sqrt{s}}{2} \, x, \qquad E_2 = \frac{\sqrt{s}}{2} \, y, \tag{13}$$

where the s is the square of centre-of-mass energy of the system, and setting $\mu_i = M_i^2/s$ one can re-parametrise the integral $I_3(P_1, P_2, P_3)$ as:

$$I_3(P_1, P_2, P_3) = \frac{s}{32} \int_{2\sqrt{\mu_1}}^{x_-} dx \int dy \, d\Omega_1 \, d\Omega_2 \, \delta[F(x, y)]. \tag{14}$$

In words, the integral is now parametrised with the energies and solid angles of two particles with a constraint in form of a $\delta[F(x, y)]$ due to the presence of the third particle. The integration limits on x can be easily derived from the Dalitz constraints, with x_- given by:

$$x_- = (1 + \mu_1) - (\sqrt{\mu_2} + \sqrt{\mu_3})^2. \tag{15}$$

The delta-function $\delta[F(x, y)]$ is actually a constraint on the value of the cosine of the angle between the particles 1 and 2:

$$F(x, y) = C(x, y) - c_{12}, \tag{16}$$

the full expression for $C(x, y)$ is given in [12]. A short calculation shows that, given values $x \in [2\sqrt{\mu_1}, x_-]$ and $c_{12} \in [-1, 1]$, the constraint of Equation 16 gives either two, one or no solutions for y, i.e. the event generation procedure is not necessarily unitary.
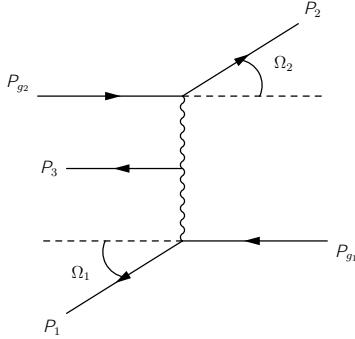
Figure 18: *A diagrammatic representation of a three-body decay as modelled by the procedure described in the text.*

In the original `MULTI1` channel implementation, the event generation nevertheless used the sampling sequence:

- Sample the value of x.

- Calculate the value of $\tilde{y}$ assuming $c_{12} = -1$ (which is favoured by multipheral singularity).

- Use the $(x, \tilde{y})$ to sample the four values $\cos(\theta_{1,2})$, $\phi_{1,2}$.

- From the angles calculate $c_{12} = \cos(\theta_{12})$ and hence determine the possible values of y using Equation 16. If there is no solution for y, the event is rejected.

As it turns out the rejection rate, very low in cases[1] of processed where it was applied originally [12], increases dramatically for the present QCD process where heavy quarks are involved. Therefore, in the present implementation the approach was extended to an unitary one (albeit at the cost of computation time).

The developed approach is based on a short calculation, which gives the existence of a turning point (maximum) at $y_T = y_T(x)$ in the $c_{12}(y) = C(x, y)$ dependence, i.e. at a certain value of x there is a maximal value $c_{12}(y_T)$. At a given value of x, the number of solutions for y, given a value of c_{12}, thus depends on the location and value of the extremum $c_T = c_{12}(y_T)$. Basing on these deductions, a new approach was developed:

- Sample the value of x from a distribution:

$$\frac{1}{(1-x)^\nu} \qquad \nu \sim 1, \tag{17}$$

- Sample the angles $\phi_1 \in [0, 2\pi]$ from a flat distribution and θ_1 from:

$$\frac{1}{a_1 - \cos(\theta_1)} \qquad a_1 = \frac{2E_1 E_{g_1} - s_1 - s_{g_1} + s_3}{2|\vec{P}_1||\vec{P}_{g_1}|}, \tag{18}$$

where the index g_1 represents the incoming particle that splits into the t-channel exchange (remember that $E_1 = \sqrt{s}/2 \cdot x$)

20

- Calculate the values of kinematic limits on $[y_{\min}(x), y_{\max}(x)]$ and the values $y_T(x)$. If the y_T is within kinematic limits $[y_{\min}(x), y_{\max}(x)]$, calculate the maximum limit c_T and hence the limits on ϕ_2 and θ_2.

- Calculate the value of $\tilde{y}$ assuming $c_{12} = -1$ (which is always allowed).

- Sample the value of θ_2 within the calculated limits from a distribution analogous to Equation 18 using $\tilde{y}$.

- Get the value of ϕ_2 from a flat distribution within the allowed limits.

- From the angles calculate $c_{12} = \cos(\theta_{12})$ and hence determine the possible values of y using Equation 16. If the calculations were done correctly, there should always be at least one solution.

The modified procedure is thus an unitary one, giving the four-vectors P_1, P_2, P_3.

Going back to the initial issue of the four body phase space (Eq. 11), the whole sampling procedure for such kinematic channels is as follows:

- Sample the squared invariant mass s_{34} from:

$$\frac{1}{(s_{34})^\mu}. \tag{19}$$

  The value of μ depends on the channel implementation (virtual gluon, a t-channel object, masses of the particles 3 and 4).

- The described above three-body decay procedure is implemented.

- The virtual P_{34} object is decayed according to the actual channel (a t-channel angular dependence, isotropic decay of a virtual gluon).

The detailed description of the other implementations of four-momenta sampling in all eleven kinematic channels is omitted for the sake of brevity; the above example should serve as a representative illustration.

Subsequently, the four-momenta constructed from the obtained set of kinematic variables are used in the matrix element calculation. Each event is further weighted by the appropriate phase space weight corresponding to the importance sampling procedure and calculated using the unitary algorithms; for further details on the applied method and unitary algorithms the reader is referred to the literature (e.g. [13, 12]).

Using the described multi-channel approach, the total generation efficiency amounts to about $3 - 4\%$. In order to further improve the efficiency, a set of modified VEGAS [15] routines was used as a random generator for the random numbers used in the sampling of the peaking quantities for each kinematic channel separately. After training all the sampling grids (of dimensions 5-7, depending on the kinematic channel), the generation efficiency increased to the order of $6 - 7\%$. The motivation for this approach was that in an unitary algorithms only a very finite set of simple sampling functions is available, since the functions have to have simple analytic integrals for which an inverse function also exists.

Consequently, the non-trivial kinematic distributions can not be adequately described by simple functions at hand in the whole sampling domain (e.g. the τ distribution, c.f. Figure 19) and some additional smoothing might be welcome. In addition, the random numbers distributions should, due to the applied importance sampling, have a reasonably flat behaviour to be approached by an adaptive algorithm such as VEGAS[7].
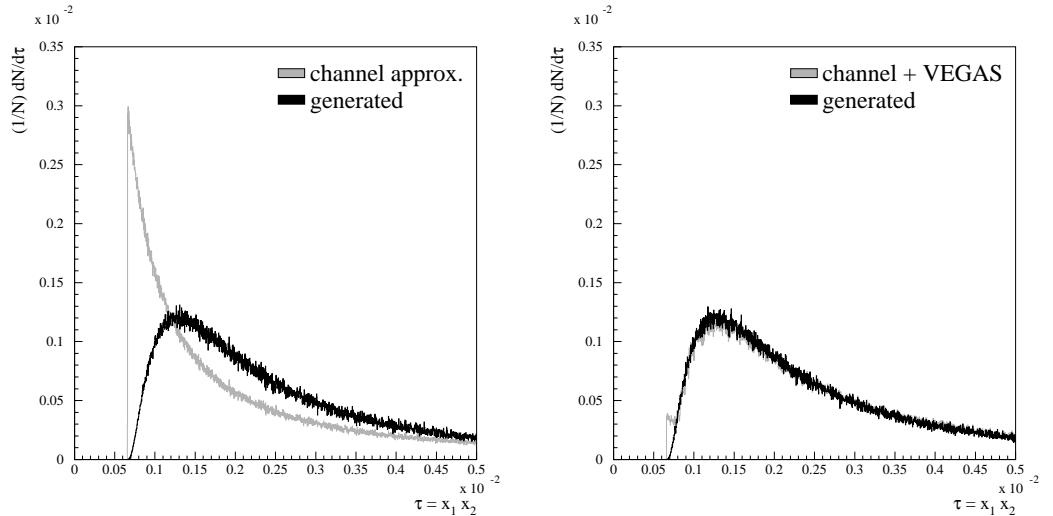
Figure 19: *Comparison between the sampling distribution for the* $\tau = \hat{s}/s \in [\tau_{\min}, 1]$ *variable before and after the application of modified* VEGAS *[15] smoothing procedure (see text).*

One of the sampling distributions is shown in Fig. 19 as a gray histogram (marked 'channel approximation') and the actual ('generated') differential cross-section dependence is drawn in black. In the first figure, the random variable used for sampling values from $1/\tau^{\mu}$ distribution was drawn from a flat probability in the interval $[0, 1]$; in the second plot the VEGAS algorithm was used to give an optimal grid for sampling the random variables needed for parameter generation (the grid is trained for each kinematic channel separately, the sum of all channels is shown in the plot). The improvement is evident; one has to stress that the use of VEGAS algorithm to generate the values of τ directly would be much less efficient since VEGAS gives a grid of 50 bins/dimension, which would give a very crude description of the τ distribution compared to the one at hand.

Observing the distributions of the event weights before and after the inclusion of the modified VEGAS algorithm (Figure 20) it is evident that VEGAS quite efficiently clusters the weights at lower values. Note that the principal effect of VEGAS is to cluster event weights about a central value (variance reduction); a few of the event weight however still retain their large values, thus reducing the generation efficiency. Given the difference

---

[7]At this point also a disadvantage of using the adaptive algorithms of the VEGAS type should be stressed, namely that these are burdened with the need of training them on usually very large samples of events before committing them to event generation.

in distributions, the observed increase of the generation efficiency for about 3% seems relatively modest. To better understand this result one should consider that the formula for the MC generation efficiency is given by:

$$\epsilon = \frac{< \text{wt} >}{wt_{\max}},\qquad(20)$$

where $< \text{wt} >$ is the average weight of the sample and equals the total event cross-section, while $wt_{\max}$ represents the maximum event weight in the applied generation procedure and is determined through a pre-sampling run with a high statistic. Since the average weight $< \text{wt} >$ equals the total cross-section of the process it remains (necessarily) unchanged after the application of the VEGAS refining; consequently the change of efficiency results in the reduction of the maximum weight $wt_{\max}$ by approximately a factor two, which is from technical point of view quite an achievement.
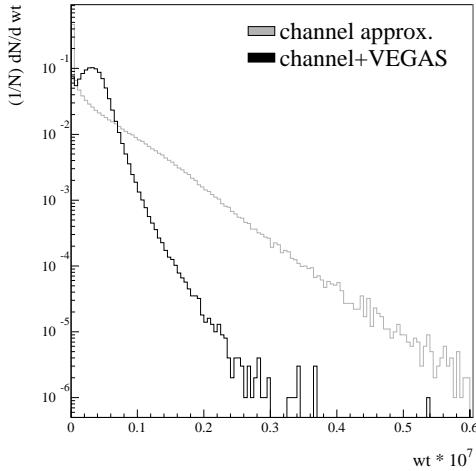
Figure 20: *The distribution of event weights before (gray histogram) and after (black histogram) the application of* VEGAS *algorithm at* $Q^2_{\text{QCD}} = \hat{s}$.

A further step to profit from the clustering of weights induced by VEGAS is to adopt a re-definition of the MC generation efficiency as proposed by [20, 21]. In this approach, the alternative definition of $wt_{\max}$ is: For a given precision level $\alpha << 1$, the $wt_{\max}$ is determined from the total weight distribution in such a way that the contribution of the events exceeding this value to the total weight sum (i.e. cross-section integral) equals α. Such a quantity is referred to as $wt^{\alpha}_{\max}$ and the efficiency expression becomes:

$$\epsilon = \frac{< \text{wt} >}{wt^{\alpha}_{\max}}.\qquad(21)$$

The argument presented in [20, 21] seems to be quite reasonable since the 'true' event weight is in any case only estimated from a finite sample of events and the new definition simply takes into account a certain level of accuracy in the maximum weight determination. In addition, certain very weak singularities that might exist in the simulated process and might occasionally result in a very high event weight are automatically taken into account.

23

Using this new definition of efficiency one can thus profit by the weight clustering induced by VEGAS. In the example shown in the Figure 20, setting the maximum weight to $wt^\alpha_{\max} = 0.3 \cdot 10^{-7}$ gives after a short calculation $\alpha \simeq 10^{-4}$ and is thus a very reasonable choice (another quantification might also be the probability that an event has a weight higher than $wt^\alpha_{\max}$, which is estimated to be approximately $p \simeq 10^{-5}$). The new $wt^\alpha_{\max}$ consequently results in a generation efficiency of about $\epsilon \simeq 10\%$, which is a significant improvement in terms of time needed for MC generation.

4.2   Colour Flow in $gg, q\bar{q} \to t\bar{t}b\bar{b}$

For the process $gg \to t\bar{t}b\bar{b}$ six colour flow configurations are possible, as shown in Figure 21. With 36+2 Feynman diagrams contributing to the process and at least half of them participating in two or more colour flow configurations, calculations by hand would prove to be very tedious. Consequently, a slightly modified colour matrix summation procedure from MADGRAPH [6] was used to determine the colour flow combinations of the diagrams and the corresponding colour factors. The thus derived squared matrix elements for separate colour flow combinations $|\mathcal{M}_{\mathrm{flow}}|^2$ were used as sampling weights on an event-by-event basis to decide on a colour flow configuration of the event before passing it on to PYTHIA 6.1 for showering and fragmentation. The procedure was verified to give identical results regarding the colour flow combinations and corresponding colour factors when applied to the processes published in [16].
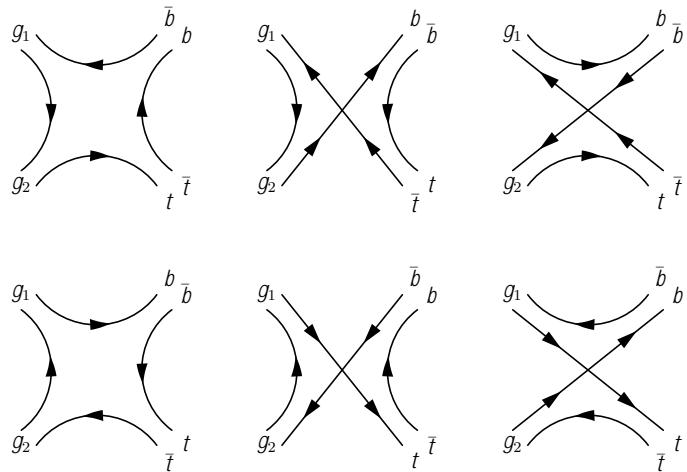
Figure 21: *A diagrammatic representation of the six colour flow configurations in the process $gg \to t\bar{t}b\bar{b}$.*

The colour flow configuration in the $q\bar{q} \to t\bar{t}b\bar{b}$ channel is much simpler since only two colour flow topologies exist (Fig. 22); the choice between the two has been solved in a manner identical to the one for the $gg \to t\bar{t}b\bar{b}$ process, as described above.
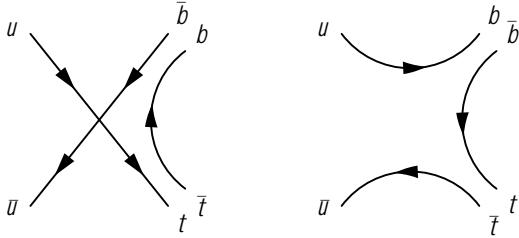
Figure 22: *A diagrammatic representation of the two colour flow configurations in the process $q\bar{q} \to t\bar{t}b\bar{b}$.*

5   Technical details on implementation in `PYTHIA 6.1`

The described implementation of the $gg, q\bar{q} \to t\bar{t}b\bar{b}$ events generation is interfaced into `PYTHIA 6.1` using an *external process* interface [7]. The implementation is done similarly as for the $q\bar{q} \to Wb\bar{b}$ process [22] and the $gg, q\bar{q} \to Z/\gamma^* bb$ processes [23]. There is a common `demo.f` main Fortran file and common libraries for the $Wb\bar{b}$, $Zb\bar{b}$ and $t\bar{t}b\bar{b}$ implementations. The interface in `demo.f` is somewhat temporary, we plan that it will evolve during next months to a simplified final version.

- The implementation offers full massive treatment of quarks, no momenta rescaling is needed for `PYTHIA` interface.

- Only the total cross-section of the process is output by `PYTHIA`, in case specific decay modes of the t- or b-quarks are required the resulting $\sigma \times BR$ has to be calculated manually, i.e the cross-section has to be calculated consistently by multiplying the total cross-section with the branching rations of the allowed decay modes as provided in `PYTHIA`.

The use of the process implementation in one's own program is basically trivial, there are only a couple of extra switches. The general procedure to put the process in one's own main Fortran file is as follows:

- One has to include the commons:
  ```
  DOUBLE PRECISION CMS,XWC
  INTEGER VLINIT,NOPTIM,NMAXSTEP,KFL
  COMMON/VLTRANS/CMS,ICH,VLINIT,NOPTIM,NMAXSTEP,KFL,XWC
  SAVE /VLTRANS/

  DOUBLE PRECISION VLCMS,VLPHMASS
  INTEGER VLHBOOK,VLCOD,VLPHOT
  COMMON/VLSETS/VLHBOOK,VLCOD,VLCMS,VLPHOT,VLPHMASS
  SAVE /VLSETS/
  ```

- Only one of these parameters is of relevance to the user and have to be set:

  - VLCMS - c.m.s. energy in GeV (DOUBLE PRECISION)

- There is an additional switch which determines the calculation of the Q^2 scale. For this purpose the so far unused PYTHIA switch MSTP(191) has been adopted:
  MSTP(191): (D=1)
  1 - $Q^2 = \hat{s}$
  2 - $Q^2 = \sum (p_T^{i\,2} + m_i^2)/4\ = < m_T^2 >$
  3 - $Q^2 = \sum (p_T^{i\,2})/4\ = < p_T^2 >$
  4 - $Q^2 = (m_t + m_H/2)^2, \quad m_H = 120\ \mathsf{GeV}/c^2$

- Note that the t and b quark masses are taken from PYTHIA tables; if the user wants to modify these it should do it through PYTHIA switches

- The external routine needs an initialisation & finalisation call, which is passed on to the PYUPEV routine by the variable VLINIT (=0 - initialisation, =-1 - normal mode =-2 finalisation).

```
C initialisation call for one process only
VLINIT=0
MSUB(435)=1
MINT(1)=435
CALL PYUPEV(ISUB,SIGMAX)
C ISUB code 435 for external process g + g to t tbar b bbar (MG)
CALL PYUPIN(435,'g + g → t tbar b bbar (MG)',SIGMAX)
MINT(1)=0
VLINIT=-1


C similar for more processes
VLINIT=0
MSUB(435)=1
MSUB(436)=1
MINT(1)=435
CALL PYUPEV(ISUB,SIGMAX)
C ISUB code 435 for external process g + g to t tbar b bbar (MG)
CALL PYUPIN(435,'g + g → t tbar b bbar (MG)',SIGMAX)
MINT(1)=436
CALL PYUPEV(ISUB,SIGMAX)
C ISUB code 436 for external process q + qbar to t tbar b bbar (MG)
CALL PYUPIN(435,'q + qbar → t tbar b bbar (MG)',SIGMAX)
MINT(1)=0
VLINIT=-1
```

```
C... your code, generate events by calling CALL PYEVNT, as usual ...


C finalization call
VLINIT=-2
CALL PYUPEV(ISUB,SIGMAX)
```

The codes for the included processes are:

431 - g + g → Z b bbar (Kleiss ME)

432 - g + g → Z b bbar (Madgraph ME)

433 - q + qbar → Z b bbar (Madgraph ME)

434 - q + qbar' → W b bbar (Madgraph ME)

435 - g + g → t tbar b bbar (Madgraph ME)

436 - q + qbar → t tbar b bbar (Madgraph ME)

Two additional codes are included in demo.f, namely KEYPRO=440, which is a dummy code that activates both processes 432 and 433; and KEYPRO=441, which initialises the processes 435 and 436. This provides a guideline in case users want to add additional (PYTHIA internal) channels; in case one wants to add an external channel some additional hacking is needed; one of the options of how to deal with it is to consult the author (*borut.kersevan@cern.ch*).

- Compiling & linking:

  The external process itself with its routines is compiled in a library `libxxbb.a` with the additional object file `vlpyupev.o` that has to be linked explicitly to the main program. The other required libraries are the `libpythia.a` with object files `pydata.o` and `pyr.o`, helas (`libdhelas.a`) and *cernlib* libraries. The provided PYTHIA library is labelled `libpythia6162.a` (the 6.157 PYTHIA version with CTEQ5 bug fixed) and symlinked to `libpythia.a`. An example implementation can be found at ATLAS CVS repository,
  `/afs/cern.ch/atlas/software/cvs/offline/atlfast/demo/demo-xxbb`.

  For generating the $gg, q\bar{q} \rightarrow t\bar{t}b\bar{b}$ processes, another directory named `grids` has to be symlinked in the same directory where executable is executed; this directory contains the trained VEGAS grids which improve the generation efficiency. In the `atlfast` cvs repository the files are already provided in the directory `vegas_grids`.

- The process is already implemented in the `./demo.f`, which reads the input card `./prod/run.dat`. Format of this input card is common for steering generation of $W b\bar{b}$ and $Z b\bar{b}$ processes (last three entries on the list below are dummy for $t\bar{t}b\bar{b}$ process):

```
 61111   — RANDOM NUMBER
   435   — KEYPRO
     3   — KEYHAD
     7   — KEYSTR
100000   — NEVENT
    11   — LEPTON CODE
     0   — FULL Z/GAMMA PROP
    10   — MIN. PHOTON MASS (GeV)
```

For the description of other parameters see the provided ./README.

- Present features & limitations:

  - The `PYTHIA CKIN` cuts are *not* implemented.
  - It is not possible to specify decay modes for W^+ and W^- separately.
  - The external process is verified to work correctly with the `PYTHIA` version 6.162 (the 6.157 `PYTHIA` version with CTEQ5 bug fixed, please note that this version is not distributed from the `PYTHIA` web page).

Example output from demo.f

```
****************************************************************************
*                *********************************                        *
*                *         *****************         *                    *
*                *         *** START PYTHIA ***         *                 *
*                *         *****************         *                    *
*                *********************************                        *
*                general switches     ....                               *
*                435                  process           KEYPRO     X1 *
*                  3                  hadroniz. option   KEYHAD     X1 *
*                 13                  Z^0 leptons type   NPAR(13)   X1 *
*                  0                  Z^0/gamma propagator NPAR(16) X1 *
*                 10                  photon cutoff (GeV) NPAR(17) X1 *
*                masses ....                                              *
*          175.00000                 top  mass          XPAR(2)    X1 *
*          100.00000                 Higgs mass         XPAR(3)    X1 *
****************************************************************************


****************************************************************************
****************************************************************************
**                                                                        **
**                                                                        **
**            *......*                    Welcome to the Lund Monte Carlo!  **
**        *:::!!:::::::::::*                                               **
**     *::::::!!:::::::::::::*           PPP   Y   Y TTTTT H   H III   A    **
**   *::::::::!!:::::::::::::::*         P  P  YY    T   H   H  I  A A   **
**  *:::::::::!!:::::::::::::::::*       PPP   Y     T   HHHHH  I  AAAAA **
**  *:::::::::!!:::::::::::::::::*       P     Y     T   H   H  I  A   A **
**   *::::::::!!:::::::::::::::::*!      P     Y     T   H   H III A   A **
**      *:::::!!:::::::::::::::* !!                                        **
**      !! *:::!!:::::::::::* !!         This is PYTHIA version 6.157      **
**      !!   !* ->< - *    !!           Last date of change: 5 Feb 2001 **
**      !!    !!           !!                                              **
**      !!    !!           !!           Now is ** Dec 1932 at  0:0*:0*  **
**      !!                 !!                                              **
**      !!      ep         !!           Disclaimer: this program comes  **
**      !!                 !!           without any guarantees. Beware  **
**      !!            pp   !!           of errors and use common sense  **
**      !! e+e-           !!           when interpreting results.      **
**      !!                 !!                                              **
**      !!                              Copyright T. Sjostrand (2000)   **
**                                                                        **
** An archive of program versions and documentation is found on the web: **
** http://www.thep.lu.se/~torbjorn/Pythia.html                           **
**                                                                        **
****************************************************************************
****************************************************************************
1****************** PYINIT: initialization of PYTHIA routines ***************
****************************************************************************
****************************************************************************


 ============================================================================
 I                                                                          I
 I            PYTHIA will be initialized for a p on p collider              I
 I                at 14000.000 GeV center-of-mass energy                    I
 I                                                                          I
 ============================================================================

 ******** PYMAXI: summary of differential cross-section maximum search ********
```

```
===========================================================
I                                         I            I
I  ISUB  Subprocess name                  I  Maximum value  I
I                                         I            I
===========================================================
I                                         I            I
I   96   Semihard QCD 2 -> 2              I   1.8700E+03   I
I  435   g + g -> t tbar b bbar (MG)      I   4.0000E-08   I
I  436   q + qbar -> t tbar b bbar (M     I   2.7000E-09   I
I                                         I            I
===========================================================
```

********************** PYINIT: initialization completed **********************

                          Event listing (summary)

   I particle/jet KS   KF   orig   p_x   p_y   p_z   E   m

   1 !p+!   21 2212   0   0.000   0.000 7000.000 7000.000   0.938
   2 !p+!   21 2212   0   0.000   0.000-7000.000 7000.000   0.938
==
   3 !g!   21   21   1   0.120   -1.026 1293.232 1293.232   0.000
   4 !u!   21   2   2   0.186   -0.859-3025.105 3025.105   0.000
   5 !g!   21   21   3 -15.498   21.730 544.526 545.180   0.000
   6 !g!   21   21   4   -4.019 -13.675 -336.318 336.620   0.000
   7 !t!   21   6   0 144.369 -39.174   41.129 233.867 175.000
   8 !tbar!   21   -6   0   68.176   93.062 307.173 371.872 175.000
   9 !b!   21   5   0 -35.188   -9.252 -30.623   47.797   4.800
  10 !bbar!   21   -5   0 -196.874 -36.580 -109.471 228.264   4.800
==
  11 (W+)   11   24   7 162.446   -7.553   6.074 182.627   82.886
  12 (tau+)   11   -15   11   0.932 -18.929   -6.773   20.204   1.777
  13 nu_tau   1   16   11 161.514   11.376   12.847 162.423   0.000
  14 (W-)   11   -24   8 -25.529   8.997 121.417 147.302   78.887
  15 (u)   A 12   2   4   1.596   -0.318 -985.903 985.904   0.330
  16 (g)   I 12   21   4   0.021   -0.631 -14.298   14.312   0.000
  17 (g)   I 12   21   4   -0.086   1.155 -633.722 633.723   0.000
  18 (g)   I 12   21   4   1.119   0.177 -554.759 554.760   0.000
  19 (g)   I 12   21   4   0.168   0.231   -5.998   6.005   0.000
  20 (g)   I 12   21   4   0.473   -1.588 -147.411 147.421   0.000
  21 (g)   I 12   21   4   -1.070   -0.371 -67.021   67.031   0.000
  22 (g)   I 12   21   10   -5.363   -0.336   -2.819   6.068   0.000
  23 (bbar)   V 11   -5   10 -192.096 -36.434 -107.311 223.085   4.800
  24 (b)   A 12   5   9 -34.603   -9.061 -29.964   46.908   4.800
  25 (g)   I 12   21   0   4.012   -2.941   -1.822   5.298   0.000
  26 (g)   I 12   21   0   -0.890   -3.134   9.826   10.352   0.000
  27 (g)   I 12   21   3   0.259   -0.081   0.046   0.275   0.000
  28 (g)   I 12   21   3   -2.147   0.713   -0.280   2.280   0.000
  29 (g)   I 12   21   3   0.374   0.248   -0.341   0.564   0.000
  30 (g)   I 12   21   3   0.286   0.178   0.241   0.414   0.000
  31 (g)   I 12   21   3   2.902   -6.302   67.111   67.469   0.000
  32 (g)   I 12   21   3   7.690 -15.422 117.547 118.803   0.000
  33 (g)   I 12   21   3   2.439   -4.755   30.674   31.136   0.000
  34 (g)   I 12   21   3   0.841   -0.097   4.254   4.337   0.000
  35 (g)   I 12   21   3   0.040   0.188   1.165   1.181   0.000
  36 (g)   I 12   21   3   -1.166   0.349   0.164   1.228   0.000
  37 (g)   I 12   21   3   1.306   -0.397   17.134   17.188   0.000
  38 (g)   I 12   21   3   0.742   0.943   11.818   11.879   0.000
  39 (g)   I 12   21   3   2.041   1.469 494.016 494.023   0.000
  40 (ud_0)   V 11 2101   1   -0.013   0.422 5048.248 5048.248   0.579
  41 (d)   A 12   1   0   2.937   -0.625   6.661   7.307   0.000
  42 (dbar)   V 11   -1   0   -2.937   0.625   44.052   44.154   0.000
  43 (d)   A 12   1   0   -2.840   -0.603 235.257 235.275   0.000
  44 (dbar)   V 11   -1   0   2.840   0.603   68.705   68.766   0.000
```

```
45 (u)       A  12      2    1   -0.107    0.604    8.876     8.903    0.330
46 (g)       I  12     21    8    5.235    3.080    7.386     9.562    0.000
47 (g)       I  12     21    8    3.190    2.921    5.872     7.293    0.000
48 (bbar)    V  11     -5    8   85.280   78.063  172.499   207.715    4.800
49 (ud_0)    A  12   2101    2   -0.186    0.859-3503.421  3503.421    0.579
50 (g)       I  12     21    4   -0.069    0.192   -0.614     0.647    0.000
51 (g)       I  12     21    4    3.663    1.679  -94.543    94.629    0.000
52 (g)       I  12     21    4    0.847   -0.380  -21.964    21.984    0.000
53 (g)       I  12     21    4    0.181   -1.418  -35.349    35.378    0.000
54 (g)       I  12     21    4    0.060   -1.631   -1.752     2.394    0.000
55 (g)       I  12     21    4    1.126    1.076  -20.303    20.363    0.000
56 (g)       I  12     21    4   -4.105   12.489  -79.337    80.418    0.000
57 (g)       I  12     21    4    0.294    2.363  -20.656    20.793    0.000
58 (g)       I  12     21    0    0.890    3.134 -452.692   452.704    0.000
59 (g)       I  12     21    0   -4.012    2.941  -14.028    14.884    0.000
60 (g)       I  12     21    7    0.421   -2.530    2.347     3.477    0.000
61 (b)       V  11      5    7  -18.498  -29.091   32.707    47.763    4.800
62 (d)       A  12      1   14   -2.163    3.794    0.369     4.396    0.330
63 (g)       I  12     21   14   -1.170    3.769   -0.304     3.959    0.000
64 (g)       I  12     21   14   -2.183   10.247    1.256    10.551    0.000
65 (g)       I  12     21   14    1.805    5.696    3.645     6.999    0.000
66 (g)       I  12     21   14    1.525    1.667    1.374     2.644    0.000
67 (g)       I  12     21   14   -0.493    0.291    0.686     0.893    0.000
68 (g)       I  12     21   14   -2.892   -2.827   12.438    13.079    0.000
69 (ubar)    V  11     -2   14  -19.958  -13.640  101.954   104.781    0.330
70 (g)       A  12     21    0   -3.505   -2.622  160.045   160.104    0.000
71 (g)       V  11     21    0    3.505    2.622  122.167   122.245    0.000
================================================================================
             sum:  2.00          0.00     0.00     0.00 14000.00 14000.00
```

```
1********* PYSTAT:  Statistics on Number of Events and Cross-sections *********

================================================================================
I                                    I                          I          I    I
I             Subprocess             I      Number of points     I  Sigma   I    I
I                                    I                          I          I    I
I-----------------------------------I--------------------------I  (mb)    I    I
I                                    I                          I          I    I
I N:o Type                          I      Generated    Tried  I          I    I
I                                    I                          I          I    I
================================================================================
I                                    I                          I          I    I
I   0 All included subprocesses      I      10000       99063  I  4.227E-09 I    I
I 435 g + g -> t tbar b bbar (MG)    I       9379       92947  I  3.939E-09 I    I
I 436 q + qbar -> t tbar b bbar (M   I        621        6116  I  2.872E-10 I    I
I                                    I                          I          I    I
================================================================================

********* Fraction of events that fail fragmentation cuts =  0.00010 *********
```

Acknowledgments

We would like to thank Daniel Froidevaux and Torbjorn Sjostrand for several very valuable discussions.

References

[1] Atlas Collaboration, *ATLAS Detector and Physics Performance TDR*, CERN-LHCC/99-15 (1999).

[2] E. Richter-Was and M. Sapinski, ATLAS Internal Note, ATL-PHYS-98-132 (1998).

[3] M. Sapinski, ATLAS Internal Note, ATL-PHYS-2000-020 (2000).

[4] M. Beneke et al., *Top quark physics*, hep-ph/0003033.

[5] A. Pukhov et al., *CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space*, hep-ph/9908288 (1999).

[6] T. Stelzer and W. F. Long, Comput. Phys. Commun. **81** (1994) 357.

[7] T. Sjostrand, Comp. Phys. Comm. **82** (1994) 74.

[8] R. Kleiss and R. Pittau, Comput. Phys. Commun. **83** (1994) 141.

[9] T. Sjostrand et al., *High energy physics generation with PYTHIA 6.1*, eprint hep-ph/0010017, LU-TP 00-30, October 2000.

[10] Conference Info A. Pukhov et al. *COMPHEP: A package for evaluation of Feynman diagrams and integration over multiparticle phase*, INP-MSU-98-41-542, Aug 1999. 126pp, hep-ph/9908288

[11] W. Beenakker et al., *Higgs radiation off top quarks at the Tevatron and at the LHC* preprint DESY-01-077, hep-ph/0107081.

[12] F. A. Berends, C. G. Papadopoulos and R. Pittau, *NEXTCALIBUR: a four fermion generator for electron-positron collision*, hep-ph/0011031.

[13] J. Hilgart, R. Kleiss, F. Le Dibider, Comp. Phys. Comm. **75** (1993) 191.

[14] F.A. Berends, R. Kleiss, Comput. Phys. Commun. **85** (1996) 11.

[15] G.P. Lepage, J. Comput. Phys. **27** (1978) 192.

[16] H.-U. Bengtsson, Comput. Phys. Commun. **31** (1984) 323.

[17] E. Richter-Was, D. Froidevaux and L. Poggioli, ATLAS Internal Note ATL-PHYS-98-131 (1998).

[18] J. Goldstein et al. Phys. Rev. Lett. **86** (2001) 1694.

[19] D. Green, K. Maeshima, R. Vidal and W. Wu, *A study of $t\bar{t} + H$ at CMS*, CMS Note 2001/039.

[20] S. Jadach, E. Richter-Was, B.F.L. Ward and Z Was, Comput. Phys. Commun. **70** (1992) 305.

[21] S. Jadach, Comput. Phys. Commun. **130** (2000) 244.

[22] B. Kersevan and E. Richter-Was, *The $q\bar{q} \to W(\to \ell\nu)g^*(\to b\bar{b})$ process: matrix element implementation to PYTHIA 6.1*, ATLAS Physics Communication, ATL-COM-PHYS-2001-013 (2001).

[23] B. Kersevan and E. Richter-Was, *The $gg, q\bar{q} \to Z/\gamma^*(\to \ell\ell)b\bar{b}$ process: matrix element implementation to PYTHIA 6.1*, ATLAS Physics Communication, ATL-COM-PHYS-2001-014 (2001).