Vol. 31 (2000) ACTA PHYSICA POLONICA B No 9

THE $t\bar{t}b\bar{b}$ BACKGROUND TO THE HIGGS SEARCHES: PYTHIA VERSUS CompHEP RATES*

Mariusz Sapiński

Institute of Nuclear Physics
Kawiory 26a, 30-059 Kraków, Poland

(Received August 7, 2000)

The $t\bar{t}b\bar{b}$ production is a dominant background for the SM and MSSM Higgs search in the associated production $t\bar{t}H$ with $H \to b\bar{b}$. This paper presents a comparison for the background rates obtained with the Monte Carlo generators: PYTHIA5.7 and CompHEP. In the case of PYTHIA the $t\bar{t}$ production is generated in the hard process and the remaining two b-quarks are obtained from the partonic cascade. In the case of CompHEP an exact matrix element for $t\bar{t}b\bar{b}$ production is used. A detailed comparison of the cross-sections and the expected acceptances is shown for several steps of the selection used for the extraction of the Higgs signal peak. The final estimates for the expected number of background events are similar in both cases with predictions from PYTHIA being even slightly higher. This comparison gives confidence in the predictions of the PYTHIA generator for the $t\bar{t}b\bar{b}$ background.

PACS numbers: 14.80.Bn, 14.80.Cp, 14.65.Ha

1. Introduction

The $t\bar{t}H$ associated production with the $H \to b\bar{b}$ decay channel is considered as a discovery channel for both SM and MSSM Higgs scenarios for Higgs boson masses in the range 80–130 GeV [1]. The semileptonic decay of one of the top quarks provides an isolated high p_{T} lepton to trigger the experiment. The reconstruction of both top-quarks, required by the selection algorithm, reduces the combinatorial background to the m_{bb} mass spectrum for the signal itself and suppresses, otherwise significant, W+jet background. The only dominant background to this channel remains the $t\bar{t}jj$ background. The true $t\bar{t}b\bar{b}$ events contribute about 60% to the total background. For an integrated luminosity of 30 fb^{-1} about 61 signal and 150 background events,

* Supported in part by the Polish State Committee for Scientific Research grants 2P03B12218, 2P03B11819 and 620/E-77/SPUB/CERN/P-03/DZ 2/99.

giving a significance of $S/\sqrt{B}$ of 5σ, are expected inside the mass window for the Higgs mass of 100 GeV. For collected 100 fb^{-1} (30 fb^{-1} with low luminosity and 70 fb^{-1} with high luminosity) the 5σ can be reached for the Higgs mass of 120 GeV.

The above estimates (see [7] for more details) were obtained with the PYTHIA 5.7 Monte Carlo generator. The $t\bar{t}$ production was generated and additional jets (b-jets) were obtained from the initial and final state radiation (parton shower). It is well known, however, that the parton shower might be giving results on the hard jets multiplicity, being much below, factor 2–3, than these obtained with matrix element calculations. Such comparisons in case of multijet events [5], and W+jets events [6], have been already documented. In the case of multi-b-jet final state this effect might however be compensated by the enhanced gluon splitting into heavy flavours, with respect to what is expected from the lowest order matrix element calculation. In the study presented below for the $t\bar{t}b\bar{b}$ background rates, estimates obtained with PYTHIA Monte Carlo are compared to predictions from the exact matrix element calculations implemented in the CompHEP Monte Carlo.

2. Generation of events

The following procedure is applied for the event generation:

- PYTHIA events: events are generated with the activated production channel for $t\bar{t}$ production (ISUB=81, 82). The initial and final state radiation, hadronisation and decays are activated as well. For consistency with CompHEP, events were filtered on 2 b-quarks, which appear as first ones in the cascade from gluon splitting, with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 3.0$. The default setting of parameters gathering QCD physics in PYTHIA is used.
- CompHEP events: events are generated according to the exact matrix element calculated for the $t\bar{t}b\bar{b}$ process. In the generated a cut is set on the $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 3$ of the two hard b-quarks which do not originate from the top decay. The generated hard process events are read as an external events by PYTHIA. Then initial and final state radiation (ISR and FSR), hadronisation and decays (also $t \to Wb$) are simulated with the PYTHIA Monte Carlo. A fine tuning of the default Q^2 scale for the initial and final state radiation is applied. For more details see [9].
- In both cases the CTEQ4L structure function is used, the top quark is decayed before hadronisation, one W-boson is forced to decay leptonically (e, μ) and the second one to decay in the hadronic channel.

There is a clear ambiguity in tuning different parameters in both generation streams (*e.g.* α_{QCD}, Q^2 evolution scale, *etc.*). They might lead to the difference in the predictions for the total cross section even by factor few. These issues have, to some extend, already been discussed in [8].

In both cases events are processed with the fast simulation package ATLFAST of the ATLAS Detector [7]. Only the true $t\bar{t}b\bar{b}$ events are discussed in the following. Comparison of the background from mistagged jets would require implementations of the additional matrix element calculations within CompHEP.

3. Cross sections and acceptances

A rather delicate point in the comparison presented below is the consistent treatment of the p_{T} threshold for the b-quarks not originating from top-quark decays. In case of events generated with CompHEP, b-quarks not originating from top quarks are produced in the hard process. In case of events generated with PYTHIA these quarks originate from the partonic cascade.

For the consideration of background rates meaningful rates are expected for the b-tagged reconstructed jets. For the analysis presented below the threshold of 15 GeV on the p_{T} of the b-quarks is chosen for b-quarks before FSR and 5 GeV for b-quarks after FSR. The first one corresponds to the threshold set in the analysis on the reconstructed jet transverse momenta, the second one is related to the internal criteria for the labelling of b-jets, see [7]. In Figs. 1, 2 respectively the following distributions are compared: multiplicity of all reconstructed jets and the cone separation between two b-quarks not originating from top quarks.

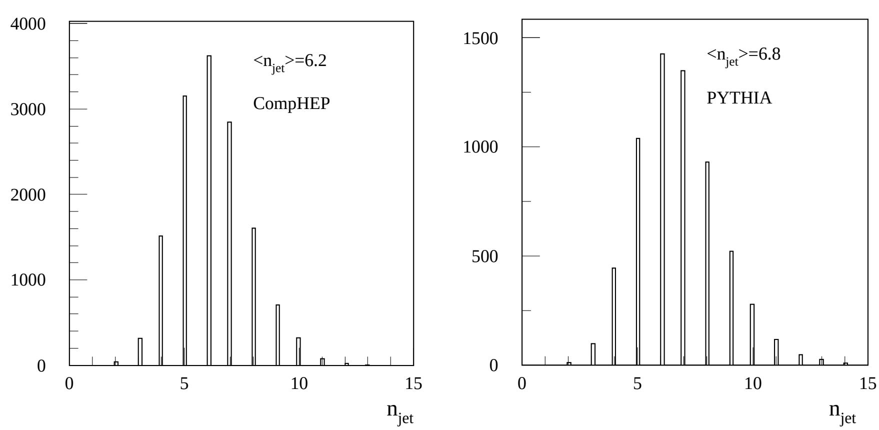

Fig. 1. Multiplicity of all reconstructed jets for events generated with CompHEP (left) and PYTHIA (right).

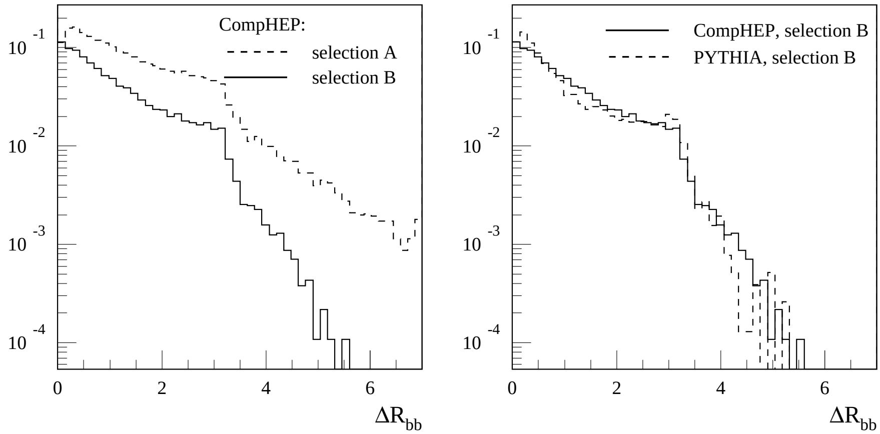

Fig. 2. On the left: cone separation for two b-quarks not originating from top decays from CompHEP events for selection A and B (see text and Table I) On the right: cone separation for two b-quarks not originating from top decays for selection B. Both the results from CompHEP (solid) and PYTHIA (dashed) are shown.

Table I shows the expected cross-section and the acceptance after selection criteria consistent with the analysis foreseen for extraction of the Higgs signal from the $t\bar{t}b\bar{b}$ background, see [4]. As the generation procedures for

TABLE I

Cross-sections and sequential acceptances after different steps of the selection in the analysis.

Selection	CompHEP	PYTHIA
A: $t\bar{t}$; 2 b-quarks ($p_{\mathrm{T}} > 15$ GeV and $\vert\eta\vert < 3.0$ before FSR)	11.9 pb 100%	6.0 pb 100%
B: both $t \to Wb$; 4 b-quarks ($p_{\mathrm{T}} > 5$ GeV and $\vert\eta\vert < 2.5$ after FSR)	5.47 pb 46%	4.44 pb 74%
both $t \to Wb$; 4 b-labelled jets ($p_{\mathrm{T}} > 15$ GeV and $\vert\eta\vert < 2.5$)	0.82 pb 15%	1.10 pb 25%
both W decayed: $W_1 \to jj, W_2 \to l\nu$	0.23 pb 28%	0.31 pb 28%
one isolated lepton with $p_{\mathrm{T}} > 20$ GeV within $\vert\eta\vert < 2.5$	0.19 pb 84%	0.24 pb 80%
solution on p_z^ν exists	0.14 pb 72%	0.18 pb 74%
at least 6 jets with $\vert\eta\vert < 2.5$ and $p_{\mathrm{T}} > 20$ GeV	0.12 pb 82%	0.15 pb 85%

PYTHIA and for CompHEP differs substantially in the first steps, two selections schemes: **A** and **B** are discussed: selection **A** is the selection in which CompHEP sample of events were generated and it is difficult to reproduce exactly the same selection for PYTHIA (FSR must be switched on and b-quarks might appear at any step of the cascade decay, top decay must be switched off).

For the purpose of this analysis selection **B** is more meaningful as b-quarks with $p_{\rm T}$ >5 GeV and $|\eta| < 2.5$ after FSR are used for jets b-labelling, and this selection is easy to reproduce in both: CompHEP and PYTHIA generation streams.

Good agreement between the predictions given by both generating streams is found for the expected 4 b-jets rates.

4. Top quarks reconstruction

The efficiencies for the reconstruction of both top quarks (see [4]), are compared for events generated with CompHEP and PYTHIA.

- The top quark decaying into lepton, neutrino and b-quark is reconstructed with the assumption that the missing transverse energy is equal to the transverse momentum of the neutrino. Up to two solutions for the neutrino longitudinal momentum are possible from the W-mass equation, and the best one minimising χ^2 (see later) is chosen. In about 26% events solution is not found (equation has no real solution).
- The top quark decaying into jets is reconstructed from two non-b-jets with $m_{jj} = m_W \pm 25$ GeV and $|\eta_j| < 2.5$ and from one b-jet.
- The best combination of the $(l\nu b, jjb)$ is found by minimising value of the $\chi^2 = (m_{l\nu b} - m_{\rm top})^2 + (m_{jjb} - m_{\rm top})^2$.

Table II compares the expected efficiencies for the reconstruction of events simulated with CompHEP and PYTHIA. Since in both cases top-quark decays are simulated with PYTHIA one expects rather good agreement between both simulation streams.

TABLE II

Top quarks reconstruction with prefect b-tagging.

Selection	CompHEP	PYTHIA
events with 4 b-jets, 2 extra jets with $p_{\rm T} > 20$ GeV and isolated lepton with $p_{\rm T} > 20$ GeV	100%	100%
both tops reconstructed	59%	58%
both tops with correct b-jets matching	35%	35%

5. Reconstruction of the m_{bb} spectrum

In the analysis for the $t\bar{t}H$ searches, the mass spectrum of the two b-jets which are not used for the top-quarks reconstruction gives the background for extracting the Higgs mass peak. In Fig. 3 the invariant mass of the b-quarks, first in the cascade and not originating from top-quarks decay is shown for events generated with CompHEP and PYTHIA.

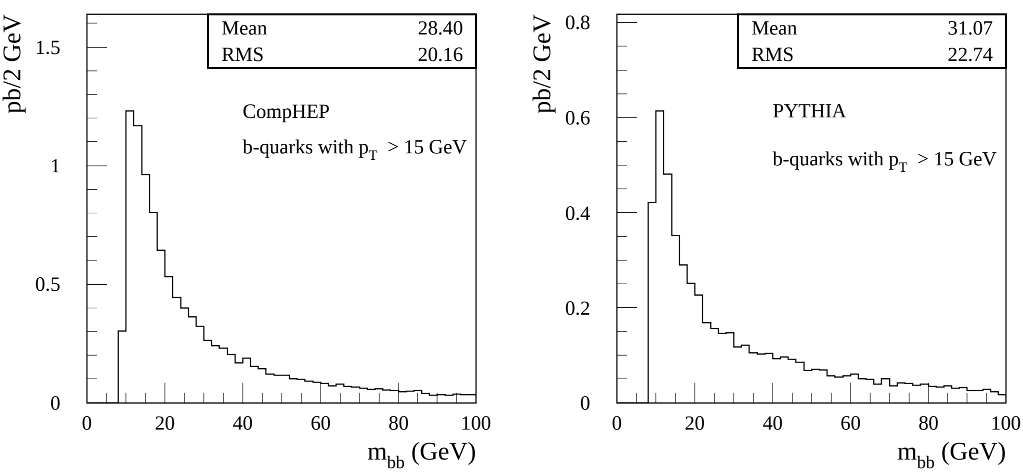

Fig. 3. Comparison of CompHEP (left) and PYTHIA (right). Invariant mass of the two b-quarks not originating from top, before FSR (or first in cascade in case of PYTHIA) and transverse momenta $p_{\mathrm{T}} > 15$ GeV.

The same distributions but for the reconstructed b-labelled jets are shown in Fig. 4. In the second case the rates are normalised to what is expected for an integrated luminosity of 30 fb^{-1} and a b-tagging efficiency of 60% (contribution from only true b-jets is shown). A more favourable shape of the m_{bb} distributions than what has been shown in [1] and [4] is obtained thanks to the optimised procedure for the jet energy recalibration [10]. Also the distribution presented here, contains only true $t\bar{t}b\bar{b}$ events, while in Fig. 19 of [4] the total background from all $t\bar{t}jj$ events was shown.

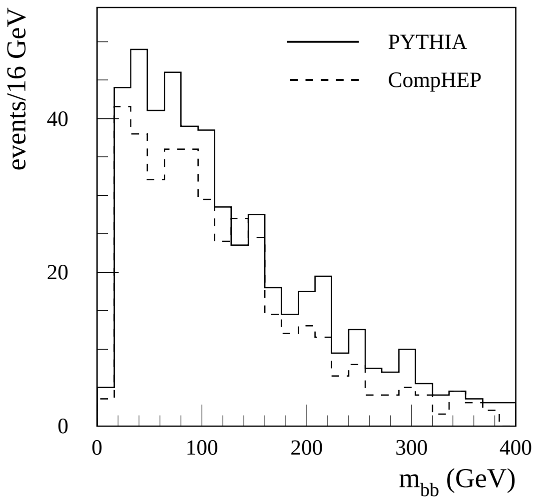

Fig. 4. Comparison of CompHEP (dashed line) and PYTHIA (solid line). Invariant mass of the two b-jets not used for the top reconstruction. Rates are normalised to these expected for an integrated luminosity of 30 fb^{-1} and b-tagging efficiency of 60%.

6. Conclusions

Both generation streams give similar results on the expected background rates. The results obtained from PYTHIA seem to give even higher predictions which is rather unexpected. One should not forget, however, that there is a large degree of ambiguity in setting consistently the parameters for the QCD activity in both simulations streams. The important conclusion from this study is however, that a credible predictivity of the PYTHIA generator for this background channel has been confirmed. The results obtained do not degradate the discovery potential of the $t\bar{t}H$ channel as presented in [1] and [4].

The author would like to thank the CompHEP team, especially Sasha Belyaev, Viatcheslav Ilyin, and Volker Drollinger from the CMS Collaboration for providing sample of events generated with CompHEP and several interactions and discussions which were crucial for performing this comparison. This note would not have been created without inspiration and help of Elżbieta Richter-Wąs, to whom the author is greatly indebted. Author would also like to thank Karl Jakobs for carefull reading of the note and several valuable comments.

REFERENCES

[1] ATLAS Collaboration, ATLAS Detector and Physics Performance Technical Design Report, ATLAS TDR 15, CERN/LHCC/99-15, 25 May 1999.

[2] T. Sjostrand, *Comput. Phys. Commun.* **82**, 74 (1994).

[3] A. Pukhov *et al.*, "CompHEP — a package for evaluation of Feynman diagrams and integration over multi-particle phase space", `hep-ph/9908288`.

[4] E. Richter-Wąs, M. Sapiński, *Acta Phys. Pol.* **B30**, 1001 (1999).

[5] E. Richter-Wąs, D. Froidevaux, ATL-PHYS-97-104 (1997).

[6] M. Cobal *et al.*, ATL-PHYS-96-084 (1996).

[7] E. Richter-Wąs, D. Froidevaux, L. Poggioli, ATL-PHYS-98-131 (1998).

[8] M. Beneke *et al.* "Top quark physics" CERN-2000-004, and `hep-ph/0003033`.

[9] CompHEP web page `http://theory.npi.smu.su/comphep`

[10] E. Richter-Was, D. Froidevaux, L. Poggioli, "ATLFAST 3.0 a fast simulation package for ATLAS", ATLAS Internal Note in preparation. Web page: `http://atlasinfo.cern.ch/Atlas/GROUPS/PHYSICS/HIGGS/Atlfast.html`