Study of a High Level b-Trigger selection of $t\bar{t}H$
fully hadronic decays

D. Benedetti, L. Fanò

*INFN and University of Perugia, Perugia, Italy*

**Abstract**

We present a High Level Trigger selection of the $t\bar{t}H$ channel decaying into hadronic final states at low luminosity ($2 \times 10^{33}$ cm$^{-2}$s$^{-1}$) where a full detector simulation was performed. The basic idea is to use a fast b-tagging algorithm after the Level-1 trigger selection in order to identify a number of jets in the final state as b-jets. The main source of background considered is direct QCD multijet production. We summarize the efficiency on signal selection and background rejection and finally we give an estimate of the expected event rate.
1 Introduction

While the detection of a directly produced Higgs boson in the $b\bar{b}$ is impossible because of the huge QCD background, one of the most promising channels in which a low mass Higgs search could be successful is the $t\bar{t}H$ associated production where the Higgs decays into a pair of $b$ quarks. The reconstruction of two top quarks in the final state allows the background suppression. At trigger level the signal can be easily selected by means of a high $P_T$ lepton, coming from one of the top decays, together with a large jet multiplicity.

The non-leptonic decays represent a larger fraction of the top decays, but the huge QCD background at trigger level discourage their subsequent use.

Considering the channel $t\bar{t}H \rightarrow t^\pm \nu q g b\bar{b}$ (29% of the total branching ratio) it is possible to achieve for example, with 30 fb$^{-1}$ of integrated luminosity, a signal significance of 5.3 for a Higgs mass of 115 GeV/$c^2$ [1] in the Standard Model scenario. It is also possible to measure the top Higgs Yukawa coupling using the $t\bar{t}H^0$ event rate.

A high discovery potential in the Minimal Supersymmetric extension of the Standard Model (MSSM) could also be investigated in the “maximal $m_H$” scenario[1]. In this case with 60 fb$^{-1}$ of integrated luminosity, most of the $\tan \beta - m_A$ available parameter space can be covered.

For the fully hadronic final state (46% of total branching ratio) has been demonstrated with fast simulation studies [2] that a signal significance of 3.5 can be reached supposing, as for the semileptonic case, a calorimetric L1 trigger requiring at least 8 jets in the final state with $E_T > 20$ GeV in $|\eta| < 2.5$. However, the simulated trigger was not realistic for two reasons: 1) the thresholds considered were low [3] and 2) 100% efficiency was assumed. Also under these condition the collected signal significance was not so high: an improvement is needed for the signal selection.

In this study a high level trigger strategy is introduced and its effect on $t\bar{t}H$ channel in the fully hadronic final state is explored.

![Figure 1: The $t\bar{t}H$ event in the fully hadronic final state](image)

In the case of event reconstruction of fully simulated and reconstructed events, not discussed here, the hadronic final state (Fig. 1) presents two main difficulties: firstly that the two b quarks coming from Higgs decay are soft and secondly the eight jets in the final state is a very complex system to reconstruct.

The relevant COMPHEP [4] cross section for the signal ($m_H = 120$ GeV) and the main QCD interval used as background are listed in Table 1.

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross Section</th>
<th>expected rate at low luminosity (2x10^{33} cm^{-2} s^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}H \rightarrow b\bar{b}b\bar{b}jjjj, m_H = 120$ GeV</td>
<td>0.22 pb</td>
<td>0.46x10^{-3} Hz</td>
</tr>
<tr>
<td>QCD (30 &lt; $P_T$ &lt; 50 GeV)</td>
<td>0.19577 mb</td>
<td>4x10^6 Hz</td>
</tr>
<tr>
<td>QCD (50 &lt; $P_T$ &lt; 80 GeV)</td>
<td>0.02587 mb</td>
<td>4x10^4 Hz</td>
</tr>
<tr>
<td>QCD (80 &lt; $P_T$ &lt; 120 GeV)</td>
<td>3.62904 $\mu$b</td>
<td>7.2x10^3 Hz</td>
</tr>
<tr>
<td>QCD (120 &lt; $P_T$ &lt; 170 GeV)</td>
<td>0.6119 $\mu$b</td>
<td>1.2x10^3 Hz</td>
</tr>
</tbody>
</table>

For this study only QCD multi jet production has been considered as background, while other resonant and non
resonant backgrounds like $Z t \bar{t}$, $j j t \bar{t}$ and $b \bar{b} t \bar{t}$ are expected to contribute according to Table 2.

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross Section</th>
<th>expected rate at low luminosity ($2 \times 10^{33}$ cm$^{-2}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}Z$</td>
<td>0.29 pb</td>
<td>0.58x10$^{-3}$ Hz</td>
</tr>
<tr>
<td>$b\bar{b}t\bar{t}$</td>
<td>1.5 pb</td>
<td>3x10$^{-3}$ Hz</td>
</tr>
<tr>
<td>$j j t \bar{t}$</td>
<td>233 pb</td>
<td>4.66x10$^{-1}$ HZ</td>
</tr>
</tbody>
</table>

2 Signal and Background Samples

This study was carried out using fully hadronic $t \bar{t} H$ decays as signal. A total of 6000 events were generated and 1000 events fully simulated using the official CMS 2002 production.

The generation has been done using PYTHIA [5] version 6.158, the simulation using CMSIM [6] 125 and the hit formatting and reconstruction steps using ORCA 6.1.0 [7].

The generator switches for the production and the decay channels were the following:

MSUB(121)=1: $gg \rightarrow t \bar{t} H$

MSUB(122)=1: $q \bar{q} \rightarrow t \bar{t} H$

MDME(214,1)=1: $H \rightarrow b \bar{b}$

MDME(190-205,1)=1: all hadronic W decay modes switched on.

A simple investigation has been done on the sample at generator level to check the sample quality and test the reconstruction strategy that is intended to be adopted. In this case the $PYCELL$ subroutine [5], a UA1-like algorithm, has been used at generator level to define jets using a cone of $\Delta R = 0.4$ ($\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$).

Events have been selected requiring at least eight jets in the final state with $E_T > 10$ GeV. Eight jets are necessary to reconstruct the whole event as will be shown later. The event selection efficiency, in this case, is more than 30%.

Figure 2 shows the number of reconstructed jets per event (left) and the number of jets identified as coming from b quarks (right). A jet is flagged to come from the hadronisation of a b quark if the $\Delta R$ between the jet and quark direction is smaller than 0.3. If more than one jet is matched, then the nearest one is taken. The mean number of reconstructed jets per event is 9.2 while on average 3.1 jets are associated with a b parton.

W bosons have been reconstructed (without any calibration on jet energy) taking the combination of 4 non-b jets that gives the best value for a combined $\chi^2_{combined} = (m_{reconstructed1} - m_W)^2 + (m_{reconstructed2} - m_W)^2$ variable with the W mass set to 80 GeV/c$^2$ (Fig.3). The W boson reconstructed mass is 76.8$\pm$9.0 GeV.

After the two W bosons have been reconstructed, the top quarks are searched for, using two out of the four b jets in the final state (fig.3); the b jet is chosen to give the best combination for the top reconstruction with a supposed
Finally the two remaining b jets can be used to reconstruct the Higgs boson mass. The used QCD samples used are datasets of the official CMS production 2002 [8]. The total number of QCD events used is about 7000.

3 The High Level b-Trigger Algorithm

The algorithms which will run at the HLT stage should be relatively robust and fast. A wide range of algorithms have been developed within CMS to tag b-jets. For this study a b-tagging algorithm was chosen which relies upon the track impact parameter. In order to speed up event reconstruction, only tracks within a jet cone will be used. The performance of the tagger depends crucially on the quality of the tracks (secondary and primary tracks composition) and the jet direction resolution.

Tracks are reconstructed within the jet defined by the direction given by the Level-1 calorimeter jet candidate. The cone size is decided as a function of the reconstructed jet and its apex given by the reconstructed primary vertex of the hard interaction.

The primary vertex location along the beam line is reconstructed using the algorithm presented in reference [9], which makes use of the so-called pixel-lines.

Track reconstruction is based on the partial reconstruction of tracks using a regional approach: starting from the pixel lines as seeds, subsequent hits are sought in a region compatible with the track $p_T$ in a region around the jet axis using a Kalman Filter technique. The reconstruction is stopped after a suitable number of hits is found along the trajectory.

Figure 4 shows the error on the track $p_T$ resolution and transverse impact parameter resolution as a function of the number of hits along the track, compared with the full tracker resolution. With only 5 hits the precision in terms of $p_T$ resolution and transverse impact parameter resolution is comparable with the full reconstruction.

A smaller number of hits, however, increases the fraction of ghost tracks. The fake rate decreases to below the 1% level for tracks with at least 5 hits.

Tracks that have been reconstructed within the jet cone can be used to better determine the Level-1 jet direction. In fact, due to the large granularity of the Calorimeter trigger cells, the resolution of the Level-1 jets is rather poor, both in $\eta$ as well as in $\phi$. The effect of a reduced angular resolution of the jet direction causes a sign flip of the track impact parameter, causing a worse performance of the tagger. The new direction is determined by a $p_T$ weighted sum of the tracks’ $\eta$ and $\phi$ and gives, on average, a factor of two improvement in resolution.
4 Event Selection

For this study we assume a baseline pixel detector (three barrel and two endcap layers on each side). In this section the effect of two different trigger strategies on signal selection and background rejection will be analysed. The first strategy is based on a jet energy threshold applied to Level-1 triggers. In the second strategy, the effect of the b-jet identification is explored, then the results are combined in order to achieve the best performance.

4.1 Level-1 Trigger Selection

At Level-1 events are selected using a calorimetric jet trigger. Fig. 5a shows the signal selection efficiency at Level-1 as a function of the cut on the calibrated transverse jet energy while Fig. 5b shows the corresponding QCD rate. The transverse jet energy has been calibrated with the formula [10] [3]:

$$E_{T\text{corrected}} = aE_{T\text{reconstructed}}^2 + bE_{T\text{reconstructed}} + c$$

The coefficients $a$, $b$ and $c$ are functions of the L1 transverse jet energy ($E_{T\text{reconstructed}}$) and $\eta$. The correction used is valid up to $|\eta| < 5$. 

Figure 5: The first and second plot show respectively the signal selection efficiency and the QCD rate as a function of the calibrated transverse energy cut. In both cases rates for 1, 2, 3 and 4 jets in the final state have been plotted. The jets are preselected with reconstructed transverse energy $> 10$ GeV corresponding to 95% of signal event and $5 \times 10^4$ Hz of 4-jets QCD rate.

The Level-1 jet selection could be performed by requiring at least four jets in the final state. It is possible to achieve 80% of signal efficiency by requiring 4 jets with $E_{T\text{corrected}} > 40$ GeV, and a rate of 1 KHz at low luminosity. If
instead the Jet $E_T$ threshold is set increased to 50 GeV the signal selection efficiency drops to 65% with a QCD rate of $\sim 200$ Hz.

### 4.2 HLT Selection

The HLT selection requires a number of jets to be identified as b-jets. In this section we will show how a b-jet identification could lower the rate while keeping a sufficiently large efficiency for the signal.

The b-tagging algorithm used is the track counting method as described in [11]. As stated in section 3 the tracks have been reconstructed by using a conditional reconstruction. In this paper we have stopped the track reconstruction when 5 hits are reached. These results are compared with those obtained by using maximum 10 hits along the track.

The HLT is now applied in order to select different signal samples as a function of the b-tag request. Pixel Lines and Tracks are reconstructed in a cone of $\Delta R = 0.4$ around the Jets direction where the original Level-1 direction is recomputed using the reconstructed tracks as explained in Section 3.

Jets are flagged as b-jets if at least two tracks with impact parameter significance above a threshold were found: from 0.5 to 5, for 1 tagged jet, and from 0.5 to 2.5 (with 0.5 steps) for 2, 3 or 4 tagged jets. The results are shown in fig.6. These results (box) have been compared with the 10 hits (circle) reconstruction scenario.

![Figure 6: Signal selection versus QCD rate after b-trigger request.](image)

Also in this case the jets are preselected with $E_T > 10$ GeV

In this case the jets are preselected with $E_T > 10$ GeV and the efficiency and rates shown if fig.6. Now one can try to optimize the signal efficiency versus QCD background rate. A 50% efficiency corresponds to a QCD rate of 2 KHz for 1 b-tagged jet. The QCD rate can be further decreased requiring, for example, 3 tagged jets: in this case a $\sim 500$ Hz rate with a signal efficiency of 30% can be obtained. Fig. 7 shows the results of using both the energetic threshold and the b-trigger criteria: events are selected requiring at least 4 jets with calibrated $E_T > 50$ GeV and, separately, for 1, 2, 3 or 4 tagged jets.

In this case it is possible, by requiring one tagged jet, to select the 55% of the signal events and reduce the QCD rate.
Figure 7: Signal selection versus QCD rate after b-trigger request. Different plots are for 1, 2, 3 or 4 tagged jets using partial (box) and full (circle) track reconstruction. Jets are preselected with $E_T > 50$ GeV to 20 Hz (at low luminosity). In table 3 are summarized signal efficiencies and QCD rates for analysed triggers.

Table 3: Summary of signal efficiency and QCD rates for different trigger criteria.

<table>
<thead>
<tr>
<th>Trigger type</th>
<th>Signal Efficiency</th>
<th>QCD rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1: 4 jets with $E_T &gt; 10$ GeV</td>
<td>95%</td>
<td>$5 \times 10^4$ Hz</td>
</tr>
<tr>
<td>L1: 4 jets with $E_T &gt; 50$ GeV</td>
<td>65%</td>
<td>200 Hz</td>
</tr>
<tr>
<td>L1: 4 jets with $E_T &gt; 10$ GeV + HLT: at least 1 b-jet (3 b-tagged jets)</td>
<td>50% (30%)</td>
<td>$2 \times 10^3$ Hz (500 Hz)</td>
</tr>
<tr>
<td>L1: 4 jets with $E_T &gt; 50$ GeV + HLT: at least 1 b-jet</td>
<td>55%</td>
<td>20 Hz</td>
</tr>
</tbody>
</table>
4.3 Signal Estimation

The event rate for signal and QCD background can be explicitly calculated over the relevant $P_T$ range for the Low Luminosity. Table 4 summarises the number of events expected for 30 $fb^{-1}$ of integrated luminosity for different trigger criteria.

<table>
<thead>
<tr>
<th>Trigger type</th>
<th># Signal events</th>
<th># QCD events</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1: 4 jets with $E_T &gt; 10$ GeV</td>
<td>6512</td>
<td>$7.5 \times 10^{11}$</td>
</tr>
<tr>
<td>L1: 4 jets with $E_T &gt; 50$ GeV</td>
<td>4456</td>
<td>$3 \times 10^9$</td>
</tr>
<tr>
<td>L1: 4 jets with $E_T &gt; 10$ GeV + HLT: 1 b-tagged jet (3 b-tagged jets)</td>
<td>3428 (2057)</td>
<td>$3 \times 10^{10}$ ($7.5 \times 10^9$)</td>
</tr>
<tr>
<td>L1: 4 jets with $E_T &gt; 50$ GeV + HLT: 1 b-tagged jet</td>
<td>3770</td>
<td>$3 \times 10^8$</td>
</tr>
</tbody>
</table>

5 Conclusions and perspectives

In this work a high level trigger based on b quark selection has been deeply analysed. In particular the Higgs boson with $t\bar{t}$ associated production in a fully hadronic final state has been investigated. It has been shown that the QCD background in the low luminosity scenario can be reduced from $5 \times 10^4$ Hz to 20 Hz while selecting the 55% of the signal events. This is done by requiring 4 jets in the final state with $E_T > 50$ GeV and at least one b-triggered jet.

The next steps will be to inspect the full signal reconstruction capabilities considering all background sources and study the High Luminosity scenario.
References

[1] CMS NOTE 2001/054 - Searching for Higgs Bosons in Association with Top Quark Pairs in the $H^0 \rightarrow b\bar{b}$ Decay Mode - V. Drollinger, Th. Muller and D. Denegri

[2] PhD Thesis - Reconstruction and Analysis Methods for Searches of Higgs Bosons in the Decay Mode $H^0 \rightarrow b\bar{b}$ at Hadron Colliders - V. Drollinger


[4] HEP-PH/0211352v1 - NLO QCD corrections to $t\bar{t}H$ production in hadron collisions


