ALL HADRONIC DECAY OF $t\bar{t}$ PAIRS AT LHC

Lj. Simić [1], G. Škoro [2], I. Mendaš [1], D. Popović [1]
and I. Borjanović [1]

[1] Institute of Physics, P.O. Box 68, 11081 Belgrade, Yugoslavia
[2] Faculty of Physics, University of Belgrade, P. O. Box 368, 11001 Belgrade, Yugoslavia

Abstract

An investigation has been made of the potential of the ATLAS detector at the LHC to study the "all hadronic" decay of $t\bar{t}$ pairs in which both W bosons decay to jets, leading to a multi-jet final state. Various kinematic variables for signal and background events have been analyzed. It is demonstrated that performing a kinematic fit of the selected events to the $t\bar{t}$ hypothesis allows extraction of the signal above background, and subsequent study of the kinematic properties of the individual top quarks and of the $t\bar{t}$ pairs. It is also shown that an effective way to further suppress the large QCD multijet background in this decay mode is to select a subsample of high p_T top events after the kinematic fit.

ATL-PHYS-2002-009
10/03/2002

I Introduction

At hadron colliders, the Standard Model (SM) predicts that top quarks are produced primarily in $t\bar{t}$ pairs, through gg fusion and $q\bar{q}$ annihilation processes. The SM further predicts that each top quark decays almost exclusively into a b quark and a W boson, so the resultant $t\bar{t}$ event topology depends on the decay modes of the two W bosons. Approximately 44.4% of $t\bar{t}$ decays lead to the final state where both W bosons decay into quarks. These events are typically dubbed "all hadronic" or "all jets" $t\bar{t}$ events. This final state consists, in the absence of initial or final state radiation, of six jets (including two b-jets), no high p_T leptons, and small missing E_T. With no final state energetic neutrinos, the all hadronic mode is the most kinematically constrained of all the $t\bar{t}$ topologies, but it is also the most challenging to measure due to the large QCD multijet background. Nevertheless, at the Fermilab Tevatron Collider both the CDF and DØ collaborations have shown that it is possible to isolate a $t\bar{t}$ signal in this channel [1, 2]. The CDF collaboration obtained a signal significance over background of better than three standard deviations [1] by applying simple selection cuts and relying on their high b-tagging efficiency ($\approx 46\%$). To compensate for their less efficient b-tagging, the DØ collaboration developed a more sophisticated event selection technique [2], in which ten kinematic variables were used in a neural network to separate signal and background. The output of this network was then combined in a second network with three additional variables designed to best characterize the $t\bar{t}$ events.

The predicted $t\bar{t}$ cross-section at the LHC is 833 pb [3]. With a combined branching ratio of $6/9 \times 6/9 \approx 44.4\%$ for the all hadronic final state, one expects production of 3.7 million multijet $t\bar{t}$ events for an integrated luminosity of 10 fb^{-1}. The enormous sample of $t\bar{t}$ events will allow a detailed scrutiny of the top quark's properties and a comparison with predictions of the Standard Model. However, due to concerns about the large QCD multijet background, most $t\bar{t}$ studies to date have concentrated on final states in which at least one of the W bosons decays leptonically.

In this note, we explore the potential of the ATLAS detector [4] to study the all hadronic decays of $t\bar{t}$ pairs. In the search for an optimal strategy for signal extraction from background, we investigate kinematic properties of both signal and background events, and perform a kinematic fit of selected events. In a latter stage, a clean signal sample is obtained by restricting to a sample of high p_T $t\bar{t}$ events in which both reconstructed top and antitop quarks have $p_T > 200$ GeV. This subsample is then used to investigate the kinematic properties of the top quarks and of the reconstructed $t\bar{t}$ pairs.

II Signal and Background Simulation

For this analysis, both $t\bar{t}$ signal and QCD multijet background events were generated using PYTHIA 6.115 [6]. The generated events were analysed with ATLFAST [7], a fast simulation package which parametrizes the response and performance of the ATLAS detector. According to PYTHIA 6.115, the cross-section for $t\bar{t} \to WWb\bar{b} \to (jj)(jj)b\bar{b}$, with the top quark mass set to 175 GeV, is ≈ 330 pb. In the further analysis we use the NLO cross section of 370 pb instead.

The major sources of background are QCD multijet events, which arise from $2 \to 2$ parton

processes ($q_iq_j \to q_iq_j$, $q_ig \to q_ig$, $gg \to gg$, $q\bar{q} \to gg$, $q_i\bar{q}_i \to q_j\bar{q}_j$, $gg \to q\bar{q}$) convoluted with parton showers. The heavy-flavour ($c\bar{c}$, $b\bar{b}$, $t\bar{t}$) content in a QCD multijet sample stems from direct production (e.g. $q_i\bar{q}_i \to q_j\bar{q}_j$, $gg \to q\bar{q}$), gluon-splitting (where a final state gluon branches into a heavy quark pair), and flavour excitation (initial state gluon splitting). In the analysis that follows, $t\bar{t}$ production was excluded from the QCD background processes. The QCD background was generated with a p_T cut on the hard scattering process above 100 GeV, resulting in a PYTHIA 6.115 production cross-section of 1.73 μb. Processes involving the production of W and Z bosons (with their subsequent decay into jets) were not included since their contributions are small compared to the QCD multijet background.

As the first step in the selection of the all hadronic $t\bar{t}$ topology, events were required to have six or more reconstructed jets, of which at least two must be tagged as b-jets. Jets were reconstructed using a fixed cone algorithm with ΔR=0.4. Jets were required to have p_T greater than certain thresholds (we present results for cuts of $p_T^{jet} > 15$ GeV, 20 GeV, 25 GeV and 40 GeV), and to satisfy $|\eta| < 3$ ($|\eta| < 2.5$ for b-jet candidates). A b-jet tagging efficiency of 60% was assumed, with mistagging probabilities for c-jets and light quark jets of 10% and 1%, respectively, for the p_T range of interest.

The efficiency for these selection criteria is presented in Table 1 (first column) for both $t\bar{t}$ signal and QCD multijet background. Additionally, the resulting ratio of $t\bar{t}$ signal to QCD background (S/B) (calculated with σ=370 pb for signal and σ = 1.73 μb for QCD multijet background) is 1/56 (1/19) for $p_T^{jet} > 15$ (40) GeV, indicating that these simple selection cuts can already reduce the multijet background to manageable levels.

III Signal and background kinematic properties

Further progress in enhancing the S/B ratio could be sought using variables that provide discrimination between the signal and the QCD background. Therefore, we examined some kinematic variables sensitive to the energy flow in the event, additional radiation and event shape, including several variables used in the neural network analysis of the DØ collaboration [2]. Examples include:
H_T: The sum of the transverse energies of the jets in the event.
H_T^{4j}: H_T without the transverse energy of the two leading jets.
$E_T^{b\ jets\ 1,2}$: The transverse energy of the two leading b-jets.
$\sqrt{\hat{s}}$: The invariant mass of the jets in the final state.
A: The aplanarity, $\frac{3}{2}Q_1$, calculated from the normalized momentum tensor.
S: The sphericity, $\frac{3}{2}(Q_1 + Q_2)$, calculated from the normalized momentum tensor.
C: The centrality, H_T/H_E, where H_E is the sum of all the jet total energies. The centrality characterizes the transverse energy flow.
ΔR_{jj}^{min}: The minimal separation between two jets in η-ϕ space.
The first four of these variables are related to the energy deposition in the event, while the second four are related more to the event shape or topology. The normalized distributions for these variables, for $p_T^{jet} > 40$ GeV, are plotted for $t\bar{t}$ signal and QCD background in Figure 1 (left for the first four variables, and right for the second four). It can be seen that the variables sensitive to the event shape provide a somewhat better discrimination between the signal and background (Figure 1 right). However, it is clear that none of these variables

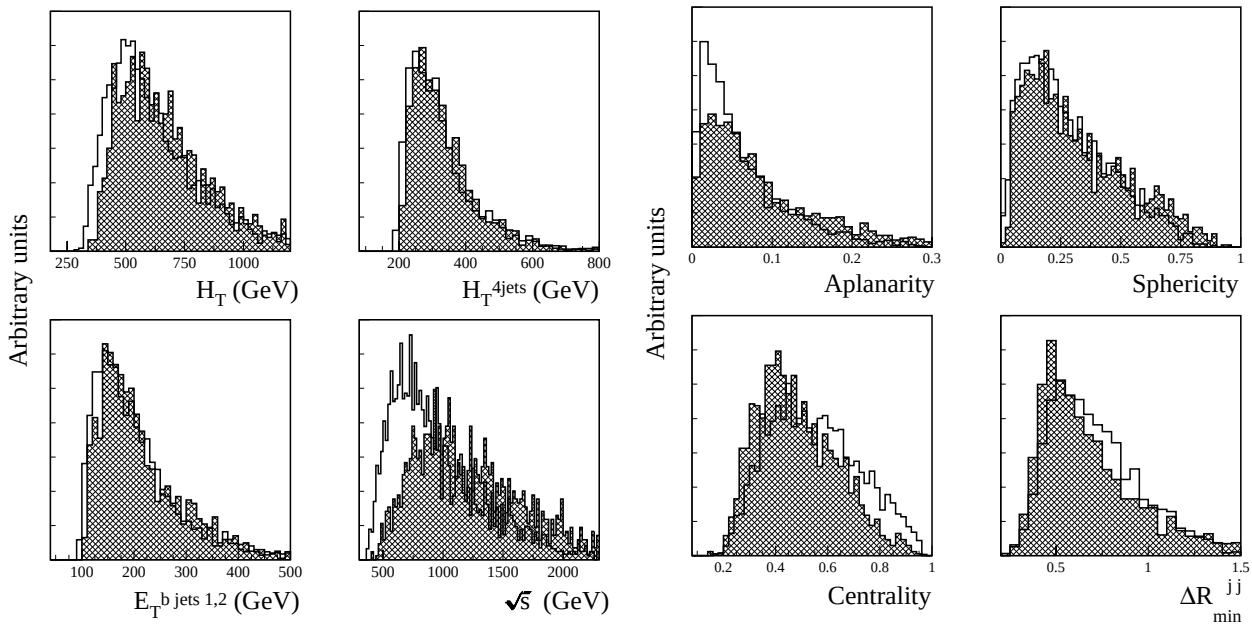

Fig. 1: Distributions for $t\bar{t}$ signal (hatched) and QCD multijet background with $p_T^{jet} > 40$ GeV. Left: the H_T, $H_T^{4\ jets}$, $E_T^{b\ jets\ 1,2}$, and $\sqrt{\hat{s}}$ distributions. Right: aplanarity, sphericity, centrality and and ΔR_{min}^{jj} distributions. For more details, see the text.

provides at the LHC the clear discrimination which was observed at the Tevatron energy [2]. Therefore, it would appear difficult to select a relatively clean signal based on cuts on these variables, or even the use of a more sensitive cut based on a multivariate discriminant, where the variables are treated collectively [2, 8].

IV Kinematic fit

The key feature distinguishing top quark events from QCD multijet background is the fitted mass obtained from the least-squares kinematic fit of the events to the $t\bar{t}$ decay hypothesis [9]. In this procedure we try to match, with specified accuracy, the observed jets with decay products at the parton level. In order to simplify the analysis we assume massless jets and neglect the error on the measured jet direction with respect to the error on the measured jet energy.

The reconstruction algorithm and fit procedure proceed in two steps. First, we seek to reconstruct the two $W \to jj$ decays by selecting, from among the jets not tagged as b-jets, di-jet combinations. This is done by minimizing the χ^2_W function

$$\chi^2_W = \left(\frac{E_u^m - E_u}{\sigma_{E_u}}\right)^2 + \left(\frac{E_d^m - E_d}{\sigma_{E_d}}\right)^2, \tag{1}$$

using the constraint that

$$K_W \equiv \frac{M_W^2}{2(1-\cos\theta_{ud})} = E_u E_d \tag{2}$$

where E_u^m and E_d^m are the measured jet energies, σ_{E_u} and σ_{E_d} are their respective errors. θ_{ud} is the measured angle between the two jets, and M_W is the known W boson mass. For the errors on the measured jet energies the following form is used: $\sigma_E/E = 0.5/\sqrt{E} \oplus 0.03$. The condition

$$\partial\chi^2_W/\partial E_u = 0 \tag{3}$$

leads to a quartic equation in E_u

$$\sigma^2_{E_d}E_u^4 - \sigma^2_{E_d}E_u^m E_u^3 + K_W\sigma^2_{E_u}E_d^m E_u - \sigma^2_{E_u}K_W^2 = 0. \tag{4}$$

In general, this equation has four roots. We choose the real and positive solution which has the minimum value of χ^2_W.

Having determined the minimum value of χ^2_W for each possible di-jet combination, we next choose the pair of di-jet combinations for which the resultant sum of the two χ^2_W's is a minimum. With this procedure we select the most likely assignment of the W boson decay jets, and use those to reconstruct the two W boson four-momenta.

Next, the two $W \to jj$ candidates are combined with the b-tagged jets to form the top and antitop quark candidates (jjb combination). The energies of the b- and $\bar{b}$-jets are constrained by minimizing the χ^2_t function

$$\chi^2_t = \left(\frac{E_b^m - E_b}{\sigma_{E_b}}\right)^2 + \left(\frac{E_{\bar{b}}^m - E_{\bar{b}}}{\sigma_{E_{\bar{b}}}}\right)^2, \tag{5}$$

and by the condition that the top- and antitop-quark masses are equal

$$\begin{aligned} &E_{W_+}E_b - \sqrt{E^2_{W_+} - M^2_W}\sqrt{E_b^2 - M_b^2}\cos\theta_{W_+b} \\ &= E_{W_-}E_{\bar{b}} - \sqrt{E^2_{W_-} - M^2_W}\sqrt{E_{\bar{b}}^2 - M_b^2}\cos\theta_{W_-\bar{b}}. \end{aligned} \tag{6}$$

Here, E_{W_+} and E_{W_-} are the W boson energies, and θ_{W_+b} and $\theta_{W_-\bar{b}}$ are the angles between the W bosons and their associated b-jets. Since we do not know which b-tagged jet is from the b and which from the $\bar{b}$ in the event, we have two ways to pair b-tagged jets with each W candidate. We select the pairing which gives the smallest value of χ^2_t.

After this event reconstruction and fit procedure, additional cuts are applied to the quality of the fit to the $t\bar{t}$ hypothesis, by requiring that the sum of the χ^2_W of the two W candidates be less than 3.5, and that χ^2_t be less than 7. The reconstructed m_{jj} and m_{jjb} distributions for signal events, for $p_T^{jet} > 20$ GeV and $p_T^{jet} > 40$ GeV are shown in Figure 2. The shaded histograms show the same distributions at the parton level. With an increase in the p_T^{jet} threshold, the reconstructed m_{jjb} distribution narrows around the true top mass and becomes similar to that at the parton level, confirming the expectation that effects of gluon radiation become less important with increasing p_T^{jet}.

Table 1 presents the efficiency and S/B ratio for $t\bar{t}$ signal and QCD multijet background after the selection cuts, after kinematic fit procedure, and after additional the requirement that the reconstructed top and antitop quarks masses lie within the window 130-200 GeV. The kinematic fit and limits on the top (antitop) mass significantly improve the value of S/B ratio so that, for top masses within the 130-200 GeV window, the S/B ratio is 1/9 for $p_T^{jet} >$

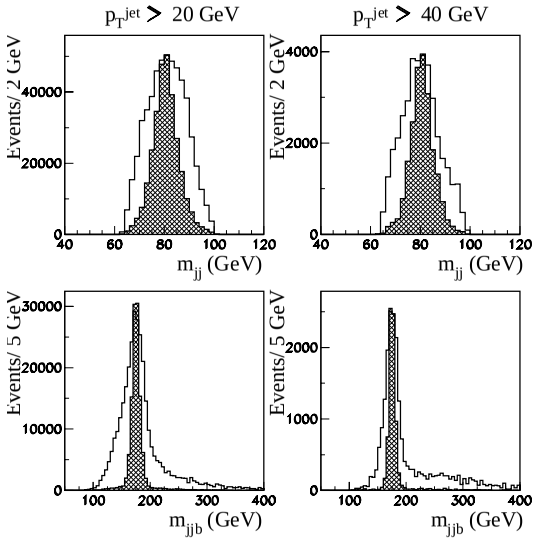

Fig. 2: Invariant mass distributions of the selected jj and jjb combinations, normalized to an integrated luminosity of 10 fb^{-}1. The shaded areas show the corresponding distributions at the parton level.

15 GeV and $S/B \approx 5$ for $p_T^{jet} > 40$ GeV. This last result is obtained with a large statistical error on the remaining background (40%), and a more accurate determination would require generation of significantly larger Monte Carlo background samples.

The results presented in Table 1 are somewhat different from the results reported previously in the Physics TDR [5], which were obtained using PYTHIA 5.7 as the event generator. In addition to this difference, the present analysis is improved in a number of ways; for example, a higher E_T^{jet} cut (> 40 GeV) has been introduced, the expression used in the fit for the errors of the measured jet energies is changed, and finally $t\bar{t}$ production has been

Table 1: Efficiency for $t\bar{t}$ signal and QCD multijet background after applying various level of cuts, and for different p_T^{jet} thresholds. The last column shows the resulting signal-to-background ratio.

p_T^{jet} threshold	after selection cuts		kinematic fit and χ^2 cuts		mass window 130-200 GeV		
	$t\bar{t}$(%)	QCD(%)	$t\bar{t}$(%)	QCD(%)	$t\bar{t}$(%)	QCD(%)	S/B
15 GeV	20.9	0.25	7.1	0.023	4.5	0.0086	1/9
20 GeV	15.5	0.12	4.3	0.008	2.7	0.0025	1/4.4
25 GeV	10.4	0.064	2.4	0.0031	1.5	0.0008	1/2.5
40 GeV	2.8	0.011	0.25	0.00017	0.16	0.000007	$\approx$ 5/1

excluded from the generation of the QCD multijet background.

V High transverse momentum $t\bar{t}$ events

The isolation of a top signal can be further improved by restricting the analysis to a sample of high p_T $t\bar{t}$ events where both reconstructed top and anti-top quarks have $p_T > 200$ GeV. To study this subsample, $t\bar{t}$ signal and QCD background events were generated using PYTHIA 6.115 with a p_T cut on the hard scattering processes above 200 GeV. The corresponding cross-sections in this case are 47.5 pb for signal (or 14.4% of the total $t\bar{t} \to all\ jets$ cross-section) and 86.1 nb for the QCD multijet background. In the further analysis we used for signal the expected NLO cross-section of 53.3 pb instead of 47.5 pb. For the high p_T sample, the same selection cuts were applied as for the inclusive sample of events, but the analysis was performed only for $p_T^{jet} > 40$ GeV. After the kinematic fit and the requirement that both reconstructed top and anti-top quarks have $p_T > 200$ GeV, the efficiency, cross-section and S/B ratio are given in Table 2.

The invariant mass distribution of the accepted jjb combinations for the high p_T $t\bar{t}$ events and QCD background (the shaded histogram) is shown in Figure 3. This figure and the S/B ratio from Table 2 clearly indicate that the signal from the fully reconstructed top and antitop quarks with $p_T > 200$ GeV, can be isolated from the QCD background. Within the window 130-200 GeV the signal to background ratio is ≈ 18. By fitting the bins around the peak of the mass distribution with a Gaussian, we find a top mass consistent with the generated value of 175 GeV. Furthermore, a width of m_{jjb} mass peak of 10.1 GeV is obtained. For an integrated luminosity of 10 fb $^{-1}$, a sample of 3300 events would be collected with fully reconstructed top and antitop quarks with $p_T > 200$ GeV. This number of events would lead to a statistical error of δm_t(stat)=$\pm$0.18 GeV and provide a clean sample for the study of differential distributions for both top (antitop) quarks, and also for $t\bar{t}$ pairs. As examples, the distributions of p_T(top), the invariant mass distribution of $t\bar{t}$, and the pseudorapidity difference between t and $\bar{t}$ are plotted in Figure 4. A kinematic property of $t\bar{t}$ pairs is that,

Table 2: Efficiency for high p_T $t\bar{t}$ signal and QCD multijet background where reconstructed top and antitop quark both have $p_T > 200$ GeV, for different cuts applied and for p_T^{jet} threshold > 40 GeV. The last row shows the resulting signal-to-background ratio.

	kinematic fit and χ^2 cuts	$m_{top} < 250$ GeV	mass window 130-200 GeV
$t\bar{t}$(%)	0.68	0.64	0.63
QCD(%)	0.00041	0.00005	0.000021
$\sigma_{t\bar{t}}$ (pb)	0.36	0.34	0.33
σ_{QCD} (pb)	0.36	0.043	0.018
S/B	1/1	$\approx 8/1$	$\approx 18/1$

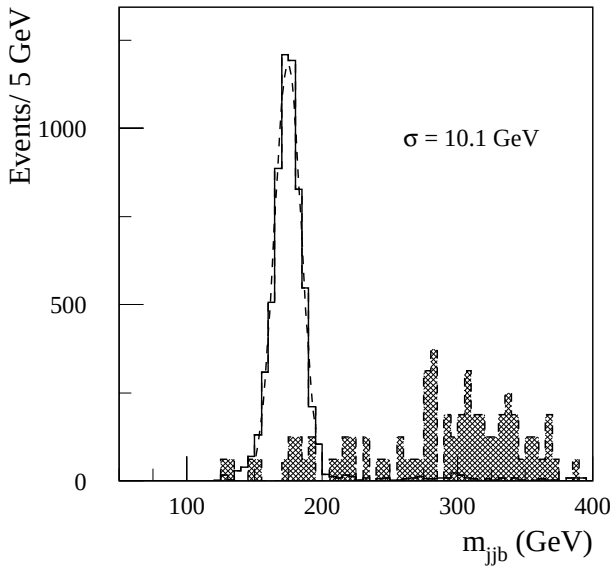

Fig. 3: Invariant mass distribution of the accepted jjb combinations for the high p_T (top) sample, normalized to an integrated luminosity of 10 fb^{-}1. The shaded area shows the QCD multijet background.

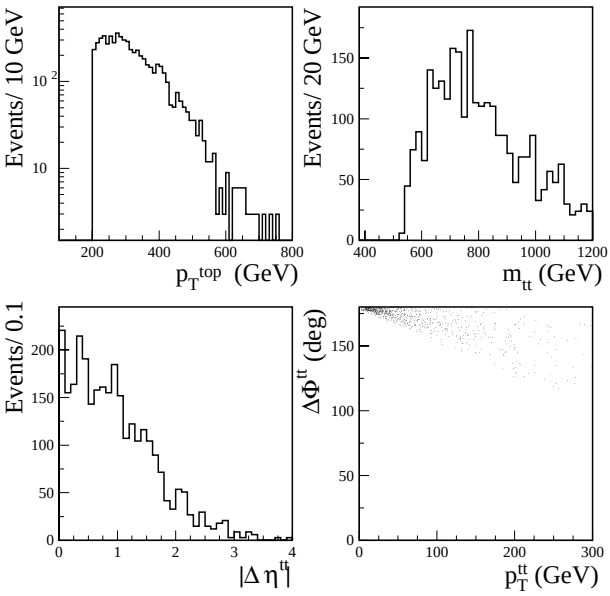

Fig. 4: Selected distributions for high p_T $t\bar{t}$ events: top p_T, the invariant mass of the $t\bar{t}$ pair, pseudorapidity difference between t and $\bar{t}$, and the difference in azimuthal angle between t and $\bar{t}$ as a function of the net transverse momentum of the $t\bar{t}$ pair.

at lowest order, they are produced with $p_T^{t\bar{t}} \approx 0$ and $\Delta\Phi^{t\bar{t}} \approx 180$ deg.

The difference in azimuthal angle between the t and $\bar{t}$ as a function of the transverse momentum of the $t\bar{t}$ pairs is also plotted in Figure 4. As can be seen, most of the pairs are indeed produced nearly back-to-back, with low $p_T^{t\bar{t}}$. Incidentally, we note that cuts on these two variables could be used as an additional criterion indicating whether the $t\bar{t}$ pairs are correctly identified.

VI Conclusions

In this note we explored, using the PYTHIA 6.115 Monte Carlo generator and the ATLFAST simulation package, possibilities for study of the all hadronic decay mode of $t\bar{t}$ pairs with the ATLAS detector at the LHC. Various kinematic variables were analyzed for both signal and QCD multijet background, and a kinematic fit of selected events was performed to the $t\bar{t}$ decay hypothesis. After this fit, isolation of a top (antitop) signal was possible with a signal to background ratio ≈ 5 within the mass window 130-200 GeV.

Restricting the analysis to a subsample of events in which both the reconstructed top and antitop quarks have $p_T > 200$ GeV, the signal-to-background ratio was improved to $\approx$ 18, with a signal of 3300 $t\bar{t}$ events for an integrated luminosity of 10 fb^{-1}. We demonstrated

that this clean sample could be used to measure differential spectra of the top (antitop) quarks, and also of the $t\bar{t}$ pairs [11, 12].

This analysis is limited by the PYTHIA modeling of the QCD multijet background production of which involves higher-order processes that may not be well modeled in this Monte Carlo simulation. The multijet rates and topologies as generated by PYTHIA suffer from very large uncertainties [10]. Consequently, reliable extraction of the signal from the background in the all hadronic decay mode could be more difficult. As an additional check, it is planned to repeat this analysis with the HERWIG Monte Carlo generator.

An additional concern is that the all hadronic final state poses difficulties for triggering at the LHC. For example, the trigger menus examined thus far by ATLAS [5] consider multi-jet trigger thresholds only up to four jets, for which a jet p_T threshold of 55 GeV is applied at low luminosity. Further study is required to determine appropriate thresholds for a six-jet topology.

Acknowledgements

We thank M. Cobal from CERN and J. Parsons from Nevis Labs, N.Y. for their contribution to this work. We also thank our colleagues D. Salihagić, A. P. Cheplakov and V. V. Kukhtin from JINR Dubna for their continued support.

References

[1] F. Abe *et al.*, CDF Collaboration, Phys. Rev. Lett. **79** (1997) 1992.

[2] B. Abbott *et al.*, DØ Collaboration, Phys. Rev. D **60** (1999) 012001; B. Abbot *et al.*, DØ Collaboration, Phys. Rev. Lett. **83** (1999) 1908; hep-ex/9901023.

[3] Bonciani *et al.*, Nucl. Phys. **B529** (1998) 424.

[4] ATLAS Technical Design Report CERN/LHCC/96-41 (1996).

[5] ATLAS Detector and Physics Performance, Technical Design Report, ATLAS TDR 14, CERN/LHCC 99-14 (1999).

[6] T. Sjöstrand, Comp. Phys. Comm., **135** (2001) 238; hep-ph/0010017.

[7] E. Richter-Was, D. Froidevaux, L. Poggioli, ATLAS Internal Note, ATL-PHYS-No-079 (1996).

[8] B. Abbott *et al.*, DØ Collaboration, Phys. Rev. D **58** (1998) 052001.

[9] J.M. Benlloch, N.Wainer, W.T. Giele, Phys. Rev. D **48** (1993) 5226.

[10] E. Richter-Was and D. Froidevaux, ATLAS Internal Note, ATL-PHYS-No-104 (1997).

[11] J. Kühn, G. Rodrigo Phys. Rev. Lett. **81** (1998) 49.

[12] J. Kühn, G. Rodrigo Phys. Rev. D **59** (1999) 054017.