Hadron Level Definition of top boosted topologies.

Rafael E. Sosa-Ricardo, InSTEC, Cuba

September 6, 2017

Supervisors: Daniela Domínguez Damiani and Hannes Jung.

Abstract

By using PYTHIA and Rivet for QCD background studies, top boosted topologies have been investigated for achieving a Top Jet definition which allows to discriminate Top events from background. After comparing b-Hadrons, fat jet requirements and making use of mass cuts, Top Jets have been defined in Boosted Topologies ($p_{\mathrm{T}}^{Top} > 400 GeV$) with acceptable background ratio and remaining Top events values.

Contents

1 Introduction

The large number of top-antitop quark ($t\bar{t}$) pairs produced at the LHC provide a unique opportunity to improve our understanding of $t\bar{t}$ production and to test the Standard Model (SM) at the TeV scale. New phenomena beyond the Standard Model may distort the top quark transverse momentum (p_T) spectrum, in particular at high p_T [1], and could thus be revealed by a precise measurement. Moreover, due to their high cross-section at the LHC and rich experimental signature, $t\bar{t}$ events constitute a dominant background to a wide range of searches for new massive particles.
The top quark has a very short lifetime:

$$\tau_{top} = O(10^{-24}s)$$

Therefore, it cannot hadronize as the formation of bound states takes about $10^{23}s$ [2]. This fact allows to test the interactions and couplings, as described in the Standard Model, with pure, thus unhadronized samples. According to the Standard Model the top decays almost exclusively into a W boson and a b-quark. Hence, the final state topology of a $t\bar{t}$ event depends on the decay of the W boson. Three different cases have to be distinguished (Fig.1):

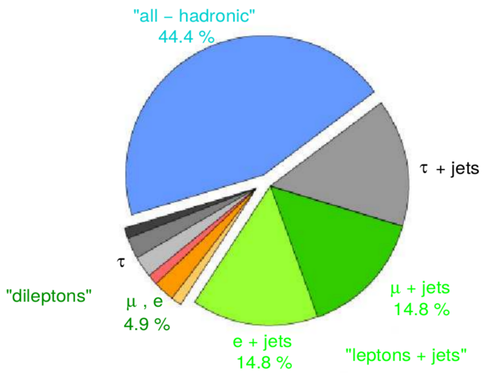

Figure 1: Branching ratios of $t\bar{t}$ decay [3].

1. Leptonic decay($t\bar{t} \longrightarrow W^+b\,W^-\bar{b} \longrightarrow l^+\nu\,l^-\bar{\nu}\,b\bar{b}$):
   The two W bosons decay into charged lepton and neutrino, respectively. Consequently, the final state comprises two jets originating from the two b-quarks, two charged leptons and missing energy arising from the two neutrinos which pass the detector without any interaction and which in consequence cannot be identified directly. The branching ratio of the leptonic (or di-lepton) decay is 4.9% [2].

2. Semileptonic decay: ($t\bar{t} \longrightarrow W^+b\,W^-\bar{b} \longrightarrow l\nu\,q\bar{q}\,b\bar{b}$)
   One W decays into charged lepton and neutrino, the other one into a pair of quark and antiquark. The final state exhibits four jets, one charged lepton and also missing energy.

3. Fully hadronic decay ($t\bar{t} \longrightarrow W^+ b\, W^- \bar{b} \longrightarrow q\bar{q}'\; q\bar{q}'\; b\bar{b}$):
   The hadronic decay is the most frequent $t\bar{t}$ event topology. Its branching ratio corresponds to 44.4% of all $t\bar{t}$ events. The fully hadronic or multijet decay shows six (or more) jets in the final state (Fig.2).

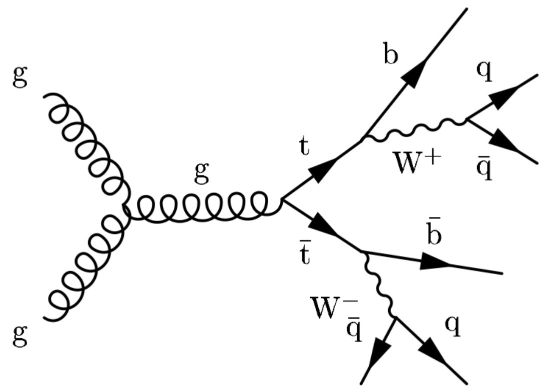

Figure 2: Feynman diagram for the fully hadronic $t\bar{t}$ decay [3].

This work is focused only on fully hadronic events, illustrated in Fig.2. The signature of these events shows six or more jets with high transverse jet-momenta $p_{T,jet}$ [4]. Two of these jets originate from the b-quarks. Four jets originate from the quarks coming from the decay of the W bosons. The $t\bar{t}$ multijet events suffer from a huge background of events which can also have six or more jets, especially the QCD multijet background [2]. In contrast to the jets in fully hadronic $t\bar{t}$ events the jets in QCD background events originate predominantly from gluon radiation.

2 Techniques.

For event generation of proton protons collisions at $\sqrt{s} = 13\text{TeV}$, PYTHIA main42 code have been used, and for analysis: Rivet tools and libraries.

2.1 PYTHIA and Rivet.

PYTHIA is a standard tool for the generation of events in high energy collisions, comprising a coherent set of physics models for the evolution from a hard process to a complex multiparticle final state. The program is based on a combination of analytical results and various QCD-based models. The different steps in the simulation include the hard subprocesses, the initial and final-state parton showers, the underlying events, beam remnants and finally hadronization and decays.
The RIVET toolkit (Robust Independent Validation of Experiment and Theory) is a system for comparison of Monte Carlo event generators. It provides a large (and ever growing) set of experimental analyses useful for comparison, as well as a convenient infrastructure for adding your own analyses. RIVET is one of the most widespread way by which analysis code from the LHC and other high-energy collider experiments is preserved for comparison and development of future theory models. It is used by phenomenologists, MC generator developers, and experimentalists on the LHC and other facilities.

2.2 average.cpp and eraseErrYodas.cpp developed Codes.

During several hours of submitting jobs, two specific annoying problems appear:

1. Compute average values for all important variables contained in always more than 100 output files (Cross section, integrals, selected events, . . .).

2. To find and delete corrupted YODA output between all generated files, which is needed for an accurate histogram construction.

For solving these problems and save time, codes average.cpp and eraseErrYodas.cpp have been developed and tested during this work.

3 Definition of Top Jets in Boosted Topologies. QCD background studies.

At very high p_T the $t\bar{t}$ system should have a boosted topology. That means instead of having several jets due the top products decays, we can cluster all in a unique jet (fat jet) (Fig.3). We will consider as high p_T jets those with $p_T > 400\text{GeV}$.

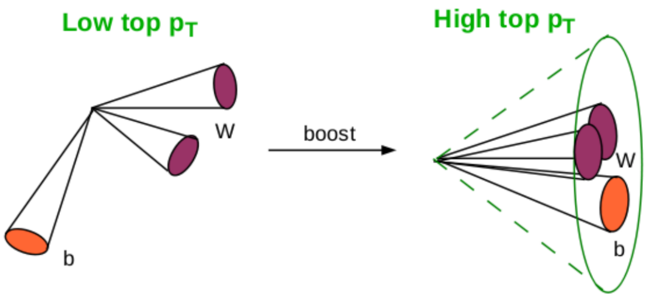

Figure 3: Boosted topology at high p_T

3.1 Boosted Top Scenarios

The fat jet is clustered with the *ANTI-KT* algorithm [5] using $\Delta R = 0.8$. Two jets with $p_T > 400\text{GeV}$ (Fig.4) are required and the top candidates will be the leading (jet with the higher p_T) and sub-leading (jet with the second higher p_T) jets. Then we can cluster these fat jets in two sub-jets (Fig.5) using the Soft Drop Mass Mechanism [6] where ideally we can identify the W and the $bHadron$ separately and removing soft radiation.

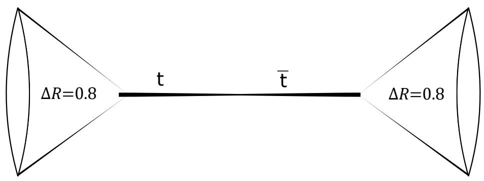

Figure 4: Fat Jets.

3.1.1 Jets and sub-jets mass distribution. Cross sections.

In Fig.6 it can be seen how the top mass can be reconstructed in the fat jet for leading and sub-leading jets, while for sub-jet0 a peak exists in the W mass region. There is also a contribution of lower masses suggesting that the W is not always clustered in this sub-jet. For sub-jet1 no peak is observed suggesting that b-Hadron is inside this jet. From this analysis we can say that Leading and Sub-leading jets are supposed to be the Top-Jets, and the first Sub-Jet in both leading jets could contain the W decay products.

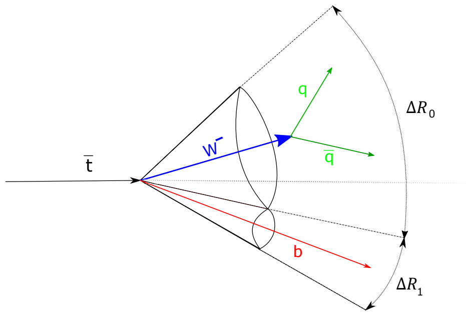

Figure 5: Sub-jets definition and topology.

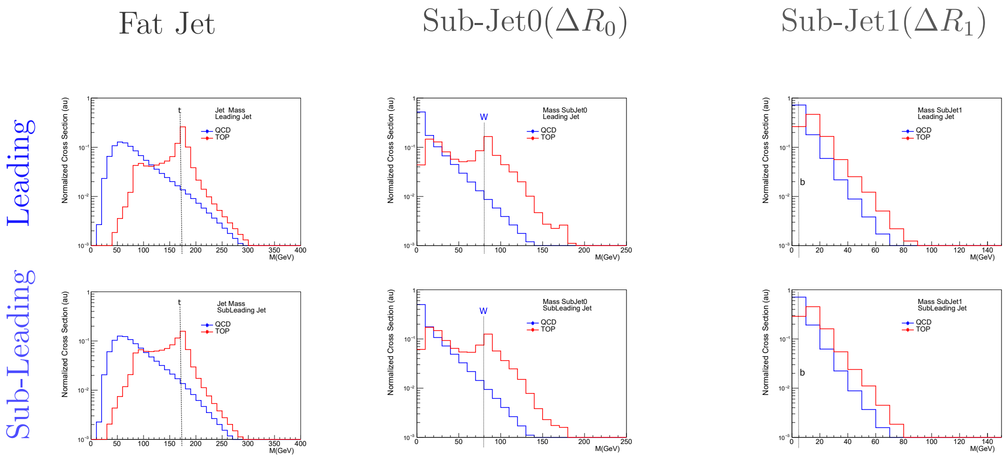

Figure 6: Mass distribution for Leading and sub-leading Fat, Sub0, and Sub1 jets.

If we require that at least two jets with $p_t > 400$GeV and in $\eta < 2.4$ we obtain an initial signal-background ratio for our analysis: $\sigma_{QCD}/\sigma_{TOP} = 1082$, remaining $\sim 60\%$ of top generated events. In Fig.7 the cross sections for leading and sub-leading jets are shown, confirming that the QCD background is over Top events in the entire p_t range.

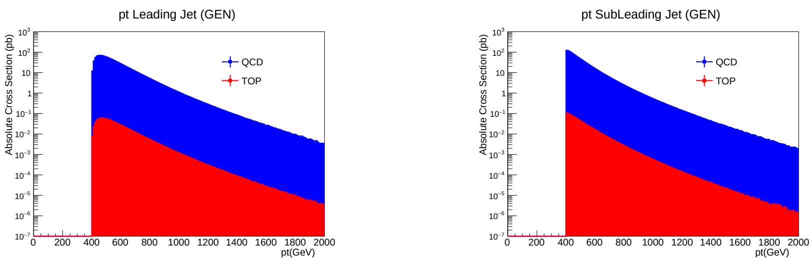

Figure 7: Cross Sections: at least two jets with $p_t > 400$GeV and $\eta < 2.4$

4 Particle Level Studies

To understand better the scenario which we are dealing with, we will look at particle level distributions. Ideally we should be able to cluster top decay products inside the fat jet, but this is not really true, because the possibility exists that the angle between b and W exceeds the value of $2\Delta R$ making impossible the construction of a fat jet containing b and W simultaneously (Fig.8). First we analyzed the angle between the W and b particle distribution. In Fig.8 is shown how top often decay products can be clustered in the same fat jet ($\Delta R < 0.8$), specially at the lowest p_T range.

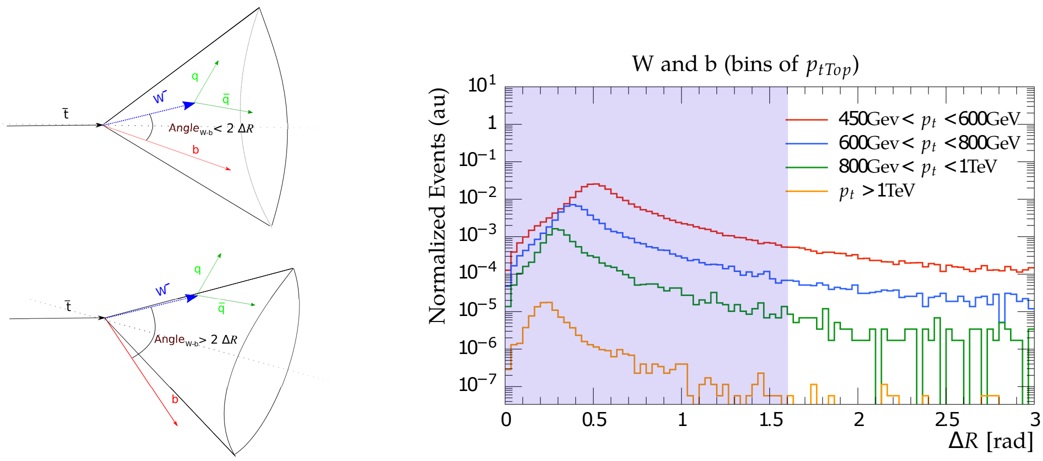

Figure 8: Angle between W and b analysis.

In order to understand which of the particles are most important this effect, the angle between W,b and the top direction can be studied. From Fig.9 we see that at lower p_t the possibility of the b-Hadron being outside the fat jet can't be neglected. ($\Delta R_{bHadronTop} < 0.8$) while the W is almost always inside the fat jet.

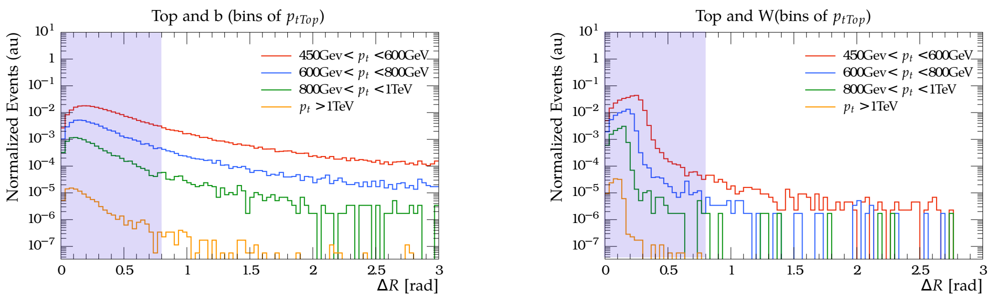

Figure 9: Angle between W,b and Top analysis.

4.1 Top Jet bHadron requirement.

The probability to have a certain quantity of b-Hadrons inside the jet can be obtained for the leading and sub-leading jets (Fig.10). The difference in probability between having 1 b-Hadron and any other number higher than 2 is of the order of ~ 100. Also its important to remark this change does not occur in QCD until reaching the value of 3. Taken this into account, we decided to follow and test the following b-Hadron requirements:

1. The Jet has to contain at least 1 b-Hadron.
    - Notice: The probability in top events to find two b-Hadrons inside the same fat jet is negligible.
2. The Jet has to contain just 1 bHadron.
    - Rejecting $\sim$ half of the background.
    - Not affecting efficiency for selecting top events.

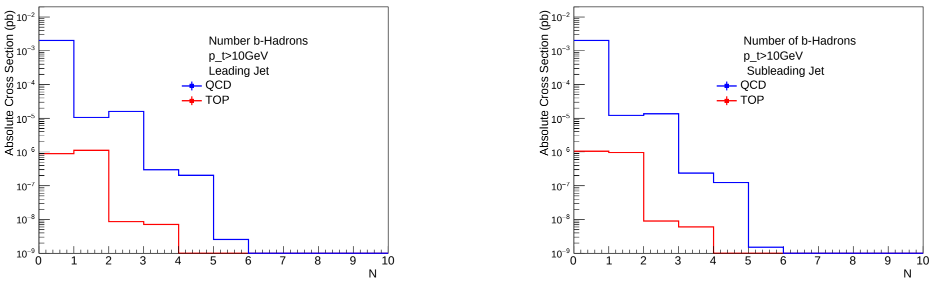

Figure 10: probability to have a certain quantity of bHadrons for leading and sub-leading jets.

5 Top Selection Criteria. Comparison between bHadron requirements.

To select top events, we start with criteria I using b-Hadron requirement 1:

- $p_t > 400$GeV, $\eta < 2.4$.
- at least one bHadron inside fat jet ($\Delta R_{bj} < 0.8$) and $SD_{Mass} > 50$GeV.

With SD_{Mass} is the tested value for Soft Drop Mass and ΔR_{bj} the angle between b and top. In Fig.11 QCD and Top cross sections are compared; the QCD cross section is larger than the Top cross section: $\sigma_{QCD}/\sigma_{TOP} = 3.3$. Applying above conditions we keep only $\sim 27.0\%$ of the generated Top events.

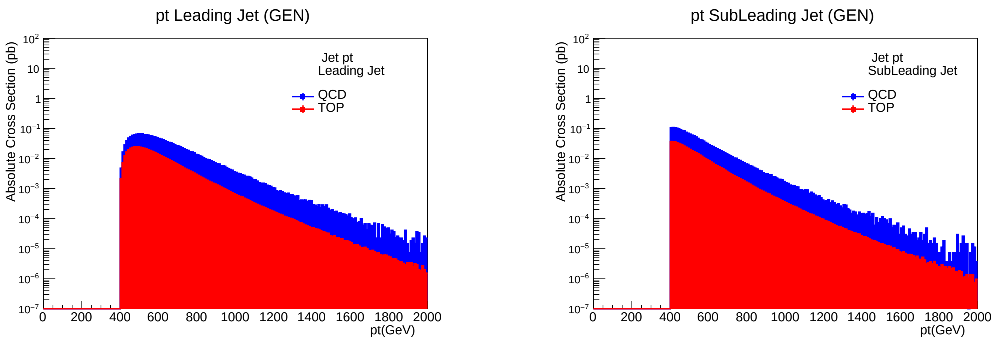

Figure 11: QCD and Top cross sections after Pre-selection criteria I.

After Pre-selection criteria I, the mass distributions for Leading and sub-leading Fat Jets (Fig.12) allows to define mass regions (Mass Cuts) to obtain QCD background discrimination: 150GeV$< Jet\ Mass <$ 200GeV and $Mass_{1stSubJet} > 30$GeV.

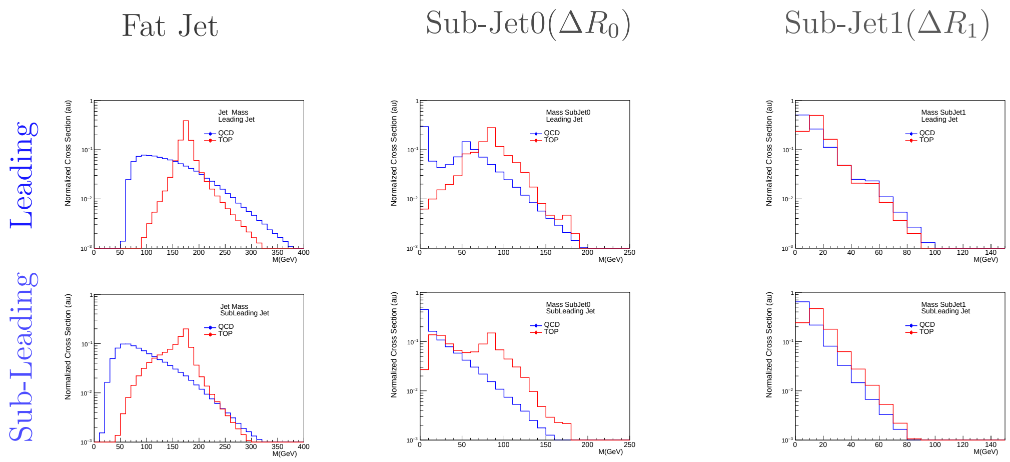

Figure 12: Mass distribution after Pre-selection criteria I.

5.1 Selection Criteria I and II

The Selection criteria I can be defined as:

- $p_t > 400\text{GeV}$, $\eta < 2.4$.
- at least one bHadron inside fat jet ($\Delta R_{bj} < 0.8$) and $SD_{Mass} > 50\text{GeV}$.
- $150\text{GeV} < Jet\ Mass < 200\text{GeV}$ and $Mass_{1stSubJet} > 30\text{GeV}$.

With this selection the cross section of Top events are is larger than the one from QCD (Fig.13) obtaining a $\sigma_{TOP}/\sigma_{QCD} = 6.5$ with $\sim 10.6\%$ of remaining Top events.

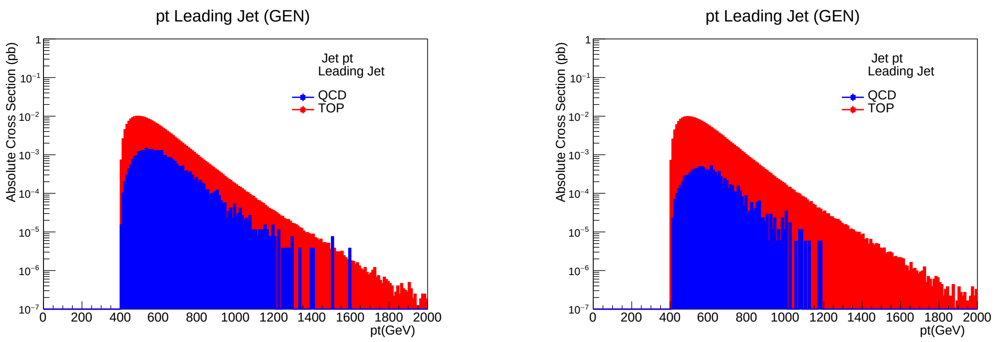

Figure 13: QCD and Top Leading Jets cross sections: (**left**) Selection criteria I. (**right**) Selection criteria II.

In addition to Criteria I, using the second bHadron requirement, Selection Criteria II can be defined:

- just one bHadron inside fat jet ($\Delta R_{bj} < 0.8$) and $SD_{Mass} > 50\text{GeV}$.

With this criteria, the QCD contribution can be proper reduced: $\sigma_{TOP}/\sigma_{QCD} = 19.8$ (Fig.13) and without considerable changes in Top efficiency ($\sim 10.4\%$).

6 Conclusions

- Top Jets have been defined for Boosted Topologies ($p_{\mathrm{T}}^{Top} > 400GeV$) with fat jets:

  - just one b-Hadron inside the fat jet.
  - top mass window (150GeV$< Mass <$ 200GeV).
  - the mass of the first sub-jet above 30GeV.

- Signal background ratio: $\sigma_{TOP}/\sigma_{QCD} = 19.8$
- Remaining Top events: $\sim 10.4\%$

References

[1] C. Englert et al. Constraining the intrinsic structure of top-quarks. *Phys. Lett. B*, 721:261, 2013.

[2] M. Benke, I. Efthymiopoulos, M. L. Mangano, and J. Womersley. Top quark physics. *hep-ph/0003033*, 2000.

[3] Marion Lambacher. Study of fully hadronic $t\bar{t}$ decays and their separation from qcd multijet background events in the first year of the atlas experiment. Technical report, Ludwig-Maximilians-Universität München, 2007.

[4] B. Abott, M. Abolins, V. Abramov, B. S. Acharya, I. Adam, M. Adams, S. Ahn, H. Aihara, and G. A. Alves. Measurement of the top quark pair production cross section in $p\bar{p}$ collisions using multijet final states. *Phys. Rev. D*, 60:012001, 2001.

[5] M. Cacciari and G. P. Salam. The anti-k_t jet clustering algorithm. *LPTHE-07-03*, 2008.

[6] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler. Soft drop. *JHEP*, 2014.