Study of the semi-boosted topology in top anti-top quark pair spectra

J. Pácal

Regional Center of Advanced Technologies and Materials
Joint Laboratory of Optics of Palacký University and Institute of Physics AS CR, Faculty of Science, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic

Abstract

The aim of this contribution is to present results of adding the semi-boosted transition region between the resolved and boosted top quark pair reconstruction in the $\ell + \text{jets}$ channel. This region is often omitted due to more complex reconstruction, which usually leads to a smaller efficiency. The reconstruction is performed for all three topologies for comparison and combination purposes. Hypothetical particle $Z'$ was used as a probe to study the increase in the number of events in the studied energy region.

PRESENTED AT

12th International Workshop on Top Quark Physics
Beijing, People’s Republic of China, September 22–27, 2019
1 Introduction

This contribution aims to explore possibilities of adding an almost non-used semi-boosted topology into reconstruction procedures for studies of top anti-top quark pair spectra in the semi-leptonic decay channel ($\ell$+jets). The top anti-top quark pair decay channels are named according to the way the $W$ boson from the decay of top quark decays either to a lepton and a neutrino (leptonic decay), or to a pair of a quark and an anti-quark (hadronic decay). In the semi-leptonic top anti-top decay channel, one $W$ boson decays leptonically and the other one hadronically. Two topologies are commonly used at present, resolved and boosted, but in between them lies a region, where the resolved topology transits to the boosted one and neither of those has a full reconstruction efficiency. This semi-boosted region of energies is often omitted from analyses.

2 Reconstruction of events and topologies

The resolved topology considers all reconstructed elements (jets and leptons) entering the to quark reconstruction angularly separated from each other, in contrary to the boosted topology, which uses collimated products (jets) coming from the hadronically decaying top quark, reconstructed as one object (large jet). The semi-boosted topology is a transition between those two topologies, where the $b$-tagged jet coming from the hadronically decaying top quark is not merged together with the two jets from the $W$ boson decay, which are collimated, or one jet coming from $W$ boson decay is separated while the other is collimated with the $b$-tagged jet (semi-boosted mixed topology). Reconstruction of the leptonically decaying top quark is the same for all topologies by using the lepton with high a transverse momentum, $E_{T,\text{miss}}$ (missing energy in transverse plane, which is usually carried away by neutrinos) and the $b$-tagged jet closest to the lepton by imposing the condition $m_{\ell\nu} = m_W$. The situation for the boosted and semi-boosted topologies is shown in Figure 1. The tagging of large jets as coming either from the $W$ boson or the top quark decay is done by using

\footnote{At Large Hadron Collider is the $b$-tagged jet defined as a jet with a secondary vertex, which was identified by the algorithm as a decay of the $B$-hadron inside the jet. The Delphes package uses an efficiency formula from the real detector for the determination of the flavor of quark, this also simulates the misidentification of the tagging algorithm.}
a substructure variable $\tau_{32}$ \cite{1} for the top quark tagging and $\tau_{21}$ for the $W$ boson tagging, and a mass window for the large jet, the chosen regions are shown in Figure 1(c) and d), respectively.

Migration of events between topologies between particle and detector levels and migration of events for a particular kinematic variable are shown in Figure 2.

The reconstruction algorithm is set to first try the boosted reconstruction, if the event does not pass the criteria, then both semi-boosted reconstruction algorithms are tried. If the event is not selected by any of them, resolved reconstruction algorithm is attempted. The event is discarded if all three reconstruction algorithms fail.

3 Samples and resolution

The simulated samples of the process $pp \to Z' \to t\bar{t}$ at $\sqrt{s} = 14$ TeV were prepared by the MadGraph generator version 2.6.4 \cite{2} interfaced with Pythia8 \cite{3} for the parton showering and hadronization simulation; and then processed by the Delphes package version 3.4.1 \cite{4} for a simplified parameterized detector simulation with different values of $Z'$ mass to cover the effective range of energies, where the semi-boosted topology is relevant. This range begins around 500 GeV and ends around 1000 GeV. The prepared samples were used as input for a private framework based on ROOT 6.16 \cite{5}. Fraction of events as function of the reconstructed mass of the top anti-top quark pair for all the topologies for a sample with $m_{Z'} = 700$ GeV is shown in Figure 3.

There is an obvious difference between the resolution at the particle and detector levels. The invariant mass of the reconstructed top anti-top quark pair mass was fitted by a Gaussian function and the resolution, in terms of fitted $Z'$ mass width, for different samples and topologies is shown as both absolute and relative values with respect to $Z'$ mass of the studied sample in Figure 4. The relative resolution in the semi-boosted case, especially at the detector level, is similar to the resolution of resolved and boosted topologies.
Figure 1: Schematic of the studied boosted (a) and semi-boosted (b) topologies. Tagging regions of large jets as top quark candidates in the $\tau_{32}$ and large jet mass plane (c); and $W$ boson candidates in the $\tau_{21}$ and large jet mass plane (d).
Figure 2: The migration of events between the resolved (R), semi-boosted mixed (SB-mix, SBM), semi-boosted (SB) and boosted (B) topologies between particle and detector levels (left), and the migration matrix for the transverse momentum of the reconstructed top anti-top quark pair for sample with $m_{Z'} = 700$ GeV (right).

Figure 3: Fraction of events in resolved, semi-boosted mixed, semi-boosted and boosted topologies at particle level (left) and detector level (right) for the sample with $m_{Z'} = 700$ GeV.
4 Unfolding, closure test and the migration matrix properties

Reconstruction was also put to test through the unfolding procedure. The Fully Bayesian Unfolding (FBU) procedure [6] was used to correct for influences of the detector. Unfolding can be described by the following formula

$$\hat{T}_i = f_i,\text{eff} M_{ij}^{-1} f_j,\text{acc}(D_j - B_j),$$

where $\hat{T}_i$ is the unfolded particle level spectrum, $D_j$ is the detector level spectrum, $B_j$ is the background spectrum contributing to the detector level spectrum, $M_{ij}$ is the corresponding migration matrix between detector and particle levels and the inversion symbolizes the unfolding procedure; and $f_i,\text{acc}$ and $f_i,\text{eff}$ are acceptance and efficiency correction factors. The example of the unfolded posterior, migration matrix and example of the detector, particle and unfolded spectra are shown in Figure 5.

The unfolding procedure is working well, this can be seen from the achieved closure test in the Figure 5, where the deviation of the unfolded spectrum from the particle level spectrum is under 3%.

5 Conclusion

This contribution presented a study of possible efficiency enhancement of top anti-top quark pair analyses by adding semi-boosted topologies in
Figure 5: An example of the posterior in a selected bin (top left), the migration matrix (top right) between the detector and the particle level; and the unfolded spectrum for $p_{T,\text{toph}}$ (bottom) for the sample with $m_{Z'} = 700$ GeV for the semi-boosted topology, where the ratio is with respect to the particle level spectrum and the yellow band is RMS of the posterior in the given bin bin.
consideration. The semi-boosted resolution is similar to other topologies in given $Z'$ mass region and it helps to increase the number of events by 60 − 70%, this could be beneficial especially to boosted analyses. The next step will be to find out the way to combine results from different topologies.

6 Acknowledgements

This article was supported by MSMT projects GACR 19-21484S and LTT-17018 and by Palacky University project IGA_PrF_2019_008.

References

[1] Jesse Thaler and Ken Van Tilburg. Identifying boosted objects with n-subjettiness. *Journal of High Energy Physics*, 2011(3), Mar 2011.

[2] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[3] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, Jun 2015.

[4] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:057, 2014.

[5] Rene Brun and Fons Rademakers. ROOT - An Object Oriented Data Analysis Framework. *Nucl. Inst. & Meth. in Phys. Res. A*, 389:81–86, 1997.

[6] Georgios Choudalakis. Fully bayesian unfolding. [https://arxiv.org/abs/1201.4612v4](https://arxiv.org/abs/1201.4612v4), 2012.