Top Tagging with the ATLAS Experiment

P. Berta
Charles University in Prague, Faculty of Mathematics and Physics, Prague, Czech Republic.

Abstract. Top tagging is an approach to identify boosted hadronic top quarks. Studies of top tagging using jet shape variables are presented. The impact of multiple proton–proton interactions (pile-up) on jet shape variables is considered and a possibility to correct the jet shape variables for pile-up is presented.

Introduction

The top quark is the heaviest elementary particle known-to-date. It is the only quark which does not form bound states, because of its extremely short lifetime. The top quark was discovered at the particle collider Tevatron, FNAL, in 1995, where its basic properties were studied. Nowadays, the only place where the top quarks are produced and detected is the Large Hadron Collider (LHC). At the LHC, three types of beams are collided: proton–proton ($p+p$), proton–nucleus ($p+Pb$) and nucleus–nucleus ($Pb+Pb$). For top quark physics, only the $p+p$ collisions are relevant. The center of mass energy of pairs of colliding protons was $\sqrt{s} = 7$ TeV in the year 2011 and $\sqrt{s} = 8$ TeV in the year 2012. At the LHC, there are two Standard Model (SM) modes of the top quark production: top–antitop ($t \bar{t}$) pair production or single top production.

The ATLAS (A Toroidal LHC ApparatuS) experiment [The ATLAS Collaboration, 2008] is one of the main experiments at the LHC which effectively detects the top quark decay products. According to the SM and the measured CKM matrix elements, the top quark decays in $\approx 100\%$ to a $W$ boson and a $b$ quark. The $b$ quark hadronizes and it creates a narrow cone of hadrons (jet). The exact signature of a top quark in the ATLAS detector depends on the decay mode of the $W$ boson originating from the top quark. The $W$ boson can decay hadronically (in $\approx 68\%$) with signature of two jets or leptonically (in $\approx 32\%$) with signature of one lepton and missing energy from a neutrino. The top quark is called hadronic in case of hadronic decay of top’s $W$ and the top quark is called leptonic in case of leptonic decay of top’s $W$. This article focuses only on the hadronic top quarks.

Depending on the magnitude of the transverse momentum of the top quark, $p_T^T$, there are two main types of top quark: resolved top and boosted top. For low $p_T^T$, the top decay products are well separated (resolved top) while for high $p_T^T$, the top decay products are more collimated due to the Lorentz boost (boosted top), see Figure 1. There is no clear boundary between resolved and boosted tops. The resolved tops are reconstructed standardly using a kinematic fit. The invariant mass of certain pair of jets in an event containing one or more hadronic top quarks should be the $W$ boson mass and the invariant mass of this pair plus one jet should be the top quark mass (within energy measurement uncertainties). With this information, the correct combination of jets can be found in the events containing one or more resolved tops. At the ATLAS experiment, the jets are reconstructed primarily using the anti-$k_t$ jet algorithm [Cacciari et al., 2008]. The free parameter of this jet algorithm called radius is often chosen to be 0.4 to effectively contain the particles originating from one quark. When the top decay products are not well separated, the jet algorithm can reconstruct jets which contain mixture of particles originating from different quarks and then the reconstruction using the kinematic fit is not efficient. One of the variables which can indicate how well are the top decay products separated is the angular separation $\Delta R(W,b)$ defined as $\Delta R(W,b) = \sqrt{(y_W - y_b)^2 + (\phi_W - \phi_b)^2}$ where $y_x$ and $\phi_x$ is the rapidity and the azimuth of the particle $x$, respectively. The Figure 2 shows the distribution of $\Delta R(W,b)$ vs. $p_T$ for top quarks from a simulation of a hypothetical
Figure 1. Illustration of the difference between the resolved top quark on the left and the boosted top quark on the right. The three jets from a boosted top quark are overlapping each other.

Figure 2. Distribution of the angular separation of the top decay products $W$ and $b$ versus the transversal momentum of the top quark in $Z' \rightarrow t\bar{t}$ events \cite{The ATLAS Collaboration, 2012}. The mass of the simulated $Z'$ boson is 1.6 TeV and the events are simulated with PYTHIA \cite{Mrenna et al., 2006} at $\sqrt{s} = 7$ TeV.

new gauge boson $Z'$ decaying to a $t\bar{t}$ pair. Thanks to the large mass of the $Z'$ boson, the two top quarks are obtaining large values of $p_T^{ij}$. The ability to resolve the three jets from a hadronic top quark using the anti-$k_T$ jet algorithm with radius of 0.4 begins to degrade for $p_T^{ij} \gtrsim 300$ GeV. Therefore, novel reconstruction techniques are necessary to identify boosted tops.

From theoretical point of view, the motivation to study boosted top quarks is the new physics. A new heavy resonance decaying to $t\bar{t}$ pair can have signature of two boosted top quarks. From experimental point of view, the boosted top quarks have no combinatoric background as in the case of the resolved top quarks which originates from the inefficiency of choosing the correct combination of jets belonging to a hadronic top quark. With increasing center-of-mass energy at the LHC, the ratio of the number of produced boosted tops to the number of produced resolved tops is increasing. All of these reasons contribute to the importance of identifying the boosted hadronic top quarks.

Top Tagging

The top tagging is an approach to identify boosted hadronic top quarks. The Figure 2 shows that the angular separation of the top decay products $W$ and $b$ for the boosted tops is $\Delta R(W,b) \ll 1$. The general idea for top tagging is to use the anti-$k_T$ jet algorithm with large
radius, e.g., 1.0, to merge all three jets coming from a boosted top into one “fat” jet, called top jet. Then by using the internal structure of jets, one can tag the fat jets to distinguish between top jets and background jets originating from other processes, e.g., QCD production of two jets in $p + p$ collisions (QCD jets). One jet shape variable which has large potential to identify the top jets is the observable called $N$-subjettiness.

$N$-subjettiness

The definition and usefulness of the jet shape variable $N$-subjettiness, $\tau_N$, was firstly introduced in [Thaler and Van Tilburg, 2010]. It “counts” the number of subjets in a fat jet and it is defined as:

$$
\tau_N = \frac{\sum_{i \in \text{const}} p_{T,i} \min (\Delta R_{i,1}, \Delta R_{i,2}, \ldots, \Delta R_{i,N})}{R \sum_{i \in \text{const}} p_{T,i}} \quad (1)
$$

where the sums run over all constituent of the jet, $R$ is the jet radius parameter of the jet algorithm, $p_{T,i}$ is the transverse momentum of the constituent $i$ and $\Delta R_{i,k}$ is the angular distance between constituent $i$ and the subjet axis $k$. The subjets are defined by re-clustering the constituents of the jet with the $k_t$ algorithm and requiring that exactly $N$ subjets have to be found. The constituents of a jet at ATLAS are 3D topological calorimeter clusters [Lampl et al., 2008].

The lower the $\tau_N$, the higher the probability that the jet contains exactly $N$ subjets. As one can expect that a top jet contains three subjets, the variables $\tau_1$, $\tau_2$ and $\tau_3$ are relevant for top tagging. Especially a new variable $\tau_{32}$ seems to have large potential:

$$
\tau_{32} = \frac{\tau_3}{\tau_2} \quad (2)
$$

as it is visible in Figure 3 where the $\tau_{32}$ distributions for QCD jets and top jets are plotted for fat jets in mass window around the top quark mass. The subjettiness $\tau_{32}$ has discriminating power which can be used for top tagging. A cut value $\tau_{32}^{\text{cut}}$ can be chosen and one can select events with at least one jet fulfilling $\tau_{32} < \tau_{32}^{\text{cut}}$ and obtain a sample of events containing top jets with higher purity.

One can also use the combination of variables $\tau_1$, $\tau_2$, $\tau_3$, jet mass and other jet shape variables in a multivariate analysis for the top tagging. There is an alternative way of choosing the $N$ subjets by minimizing the expression (1) over all possible subjet directions which can also lead to a better top tagging performance [Thaler and Van Tilburg, 2011].

Figure 3. Distributions of subjettiness $\tau_{32}$ for QCD jets and top jets with mass $m_j \in [145, 205]$ GeV [Thaler and Van Tilburg, 2011].
Figure 4. Mis-tagging efficiency vs. top-tagging efficiency for several top tagging methods, as tested on Herwig++ $t\bar{t}$ and dijet sample for jets with $p_T \in [500, 600]$ GeV [Altheimer et al., 2012].

Other techniques

There are several other possibilities for top tagging. A brief list of them follows with their denotations used in Figure 4 in parenthesis:

- ATLAS top tagger (ATLAS) — based on $k_t$ splitting scales [Brooijmans and The ATLAS Collaboration, 2008],
- Thaler and Wang top tagger (TW) — using similar kinematic variables as ATLAS top tagger [Thaler and Wang, 2008],
- Johns Hopkins top tagger (JH) — using Cambridge-Aachen jet algorithm [Kaplan et al., 2008],
- CMS top tagger (CMS) — variant of Johns Hopkins tagger [Rappoccio, 2009],
- HEP top tagger (HEP) — motivated by Higgs search channel $p + p \rightarrow t\bar{t}H$ [Plehn et al., 2009],
- top tagger using pruned jets (Pruned) — selectively removing constituents of the jet originating from pile-up, underlying event and soft radiation [Ellis et al., 2009],
- top tagger using trimmed jets (Trimmed) — with the same goal as the top tagger using pruned jets but with other technique [Krohn et al., 2009].

The comparison of the performance of these taggers together with the tagger using the jet shape variable $r_{32}$ (NSub) is shown in Figure 4. One can see that the simple top tagging method using the variable $r_{32}$ has one of the best performances, mainly if one requires high signal efficiency.

Pile-up subtraction for jet shapes

At the LHC, the protons in the two colliding beams are grouped in bunches with the spacing of 50 ns. Due to the high instantaneous luminosity, there are simultaneous $p + p$ interactions per one bunch crossing (pile-up). The pile-up affects the jets in the detector.

The jet shape variables used for top tagging are sensitive to pile-up. In the following, a method for pile-up correction for jet shapes is introduced according to [Soyez et al., 2012]. The expected pile-up deposition in a small region of size $\Delta y \cdot \Delta \phi$ can be expressed by the 4-momentum:

$$\rho \cdot [\cos \phi, \sin \phi, \sinh y, \cosh y] \cdot \Delta y \cdot \Delta \phi$$  (3)
Berta: Top Tagging with the ATLAS Experiment

where $\rho$ is the pile-up $p_T$ density weakly dependent on the azimuthal angle $\phi$ and rapidity $y$ (according to [Soyez et al., 2012]). The basic ingredient for the description of the variable $\rho$ are the so-called “ghosts” which are very low $p_T$ particles. Each ghost resides at an area of the same size, $A_g$, in the $y-\phi$ plane where they are uniformly distributed. Each ghost has 4-momentum

$$ p_{T,g} \cdot [\cos \phi_i, \sin \phi_i, \sinh y_i, \cosh y_i] $$

where $\phi_i$ is the azimuthal angle of the ghost $i$, $y_i$ is the rapidity of the ghost $i$ and the $p_{T,g}$ is the transverse momentum (all ghosts have the same transverse momentum). The area $A_g$ is inversely proportional to the number of ghosts. By clustering with the anti-$k_t$ algorithm, some ghosts are clustered into the jets. These ghosts mimic the pile-up constituents of the jet. In practice, the addition of ghosts to the $y-\phi$ plane represents a discrete form of the continuous variable $\rho$. The smaller the ghost area $A_g$, the better is the description of the variable $\rho$ via ghosts.

Jet shape $V$ is an arbitrary function of constituents of the jet. In the next part, only the dependence on pile-up through variable $\rho$ and the ghosts will be assumed: $V = V(\rho, p_{T,g})$. The measured shape is given by $V(\rho = \rho_0, p_{T,g} = 0)$, where $\rho_0$ is the pile-up $p_T$ density measured in the absence of any additional ghosts added to the event. One can express the correspondence between the ghosts and the variable $\rho$ as

$$ V(\rho + x, p_{T,g}) = V(\rho, p_{T,g} + xA_g), $$

where $x$ represents an arbitrary shift in $\rho$. This property can be used for finding the corrected jet shape in the absence of pile-up:

$$ V_{corr} = V(\rho = 0, p_{T,g} = 0) = V(\rho = \rho_0, p_{T,g} = -\rho_0A_g). $$

It is often hard to express the function $V$ explicitly through variable $p_{T,g}$. In such case, the Taylor expansion of the function $V(\rho = \rho_0, p_{T,g})$ at the point $p_{T,g} = 0$ can be used. Then the corrected jet shape can be written as

$$ V_{corr} = \sum_{k=0}^{\infty} (-\rho_0A_g)^k \frac{\partial^k V(\rho = \rho_0, p_{T,g})}{\partial p_{T,g}^k} \bigg|_{p_{T,g}=0}. $$

The derivatives can be computed numerically by alternating the variable $p_{T,g}$ around $p_{T,g} = 0$ and evaluating the change of the shape variable. In practice, only the right derivatives can be computed as the particles with negative energies have no meaning. The pile-up $p_T$ density, $\rho_0$, in the actual event can be obtained using the method described in [Cacciari and Salam, 2007].

The presented correction method was tested on the measured data at the ATLAS experiment and simulated events reconstructed after ATLAS detector simulation [The ATLAS Collaboration, 2013]. In Figure 5, the distribution of shape variable $\tau_{32}$ is presented before and after the correction as well as the distribution of $\tau_{32}$ without detector and pile-up effects. The applied correction recovered the original distribution quite well. One can also observe a reasonable agreement between the measured data and simulated events.

Conclusion

The top tagging is an important field at the ATLAS experiment. There are several top tagging methods using fat jets and one of the most promising is the use of jet shape variable $\tau_{32}$. Jet shape variables are sensitive to pile-up. The negative effect of pile-up can be corrected with the estimation of pile-up $p_T$ density and usage of the ghost technique which determine the sensitivity of a particular jet shape on the pile-up.

Acknowledgments. The presented work was supported by the Grant Agency of Charles University in Prague (GAUK), under Contract 320813.
Figure 5. Comparisons of the uncorrected (blue) and corrected (red) distributions of $\tau_{32}$ for measured data (points) and the simulated QCD jets overlaid by pile-up events reconstructed after detector simulation (solid histogram) at the ATLAS experiment. The distribution of $\tau_{32}$ computed using simulated particles without pile-up effects and without detector simulation (green) is also included [The ATLAS Collaboration, 2013].

References


