PROCEEDINGS OF SCIENCE

Top reconstruction and boosted top: experimental overview

Louise Skinnari*, **on behalf of the ATLAS and CMS Collaborations**
Cornell University (US)
E-mail: louise.skinnari@cern.ch

An overview of techniques used to reconstruct resolved and boosted top quarks is presented. Techniques for resolved top quark reconstruction include kinematic likelihood fitters and pseudo-top reconstruction. Many tools and methods are available for the reconstruction of boosted top quarks, such as jet grooming techniques, jet substructure variables, and dedicated top taggers. Different techniques as used by ATLAS and CMS analyses are described and the performance of different variables and top taggers are shown.

8th International Workshop on Top Quark Physics, TOP2015
14-18 September, 2015
Ischia, Italy

---

*Speaker.

1. Introduction

Reconstructing top quarks and top-quark pairs ($t\bar{t}$) involves identifying and correctly assigning observed decay products to the original top quark(s). *Resolved* top decays typically result in well-separated jets and isolated leptons, while for *boosted* tops, the decay products may be overlapping, resulting in merged jets and non-isolated leptons. The separation of the decay products in η–ϕ can be approximated as $\Delta R \approx 2m/p_{\mathrm{T}}$, where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, m the mass of the decaying particle, and p_{T} its transverse momentum. This corresponds to that a top quark with e.g. $p_{\mathrm{T}} = 350$ GeV has its decay products contained in a cone of $\Delta R \approx 1.0$. These proceedings discuss different techniques used by ATLAS [1] and CMS [2] analyses.

2. Resolved top reconstruction

2.1 Kinematic likelihood techniques

Many techniques have been developed to associate the final state decay products (jets) to the partons from the original top quark. One technique for reconstructing resolved top quarks in $t\bar{t} \rightarrow l+$jets final states is a kinematic likelihood fitter, KLFitter [3], which uses a 3D template fit. The method aims to correctly assign jets to partons by maximizing a likelihood, testing each possible permutation of jet-parton association. The likelihood is formed using Breit-Wigner functions which constrain the reconstructed dijet/trijet mass to the mass of the W boson (m_W) or top quark (m_{top}), and transfer functions which map the measured jet energy to the energy of the final state partons. The method is used in numerous ATLAS analyses such as the 7 TeV top mass measurement [4], where the observables are the reconstructed top quark mass ($m_{\mathrm{top}}^{\mathrm{reco}}$), the reconstructed W mass (m_W^{reco}), and a variable sensitive to the relative jet energy scale for b-quark jets vs light-quark jets.

The kinematic likelihood technique is also used by CMS for the 13 TeV differential $t\bar{t}$ cross section measurement in the $l+$jets channel [5]. The neutrino momentum is first reconstructed using m_W and m_{top} constraints. Using the solution best compatible with the measured missing transverse energy (MET) leads to an improved neutrino p_{T} measurement. Next, a likelihood for the most probable quark-to-jet assignment is defined from the 2D probability of the invariant masses of the 2-jets and 3-jets systems tested as the W boson and hadronic top quark decay products, respectively, and the probability of the neutrino reconstruction to correctly select the b-jet. The combined reconstruction efficiency is about 60%.

2.2 Pseudo-top reconstruction

Pseudo-top reconstruction aims to minimize the model dependence of differential cross-section measurements, allowing QCD precision tests in the top quark sector. The "pseudo top" is a top-quark proxy object at stable-particle level. For $l+$jets events, the leptonic W is defined from the electron/muon and the MET, solving for the z-component of the neutrino momentum assuming m_W. The leptonic pseudo-top is formed from the leptonic W and the closest b-jet. The hadronic W is formed from the next two highest-p_{T} jets and the hadronic pseudo-top from this W combined with the remaining b-jet. The usage of pseudo-top reconstruction avoids performing the largely model-dependent extrapolation from detector-level to parton-level [6, 7].

3. Boosted top reconstruction

Boosted (high-p_{T}) top quarks require dedicated reconstruction techniques due to their collimated decay products. Techniques for reconstructing boosted objects are used both to allow testing predictions of high-p_{T} top-quark production within the Standard Model as well as to probe models of physics beyond the Standard Model which result in boosted objects, such as new heavy resonances or vector-like quarks.

3.1 Boosted leptonic tops

For boosted top quarks which decay leptonically, standard isolation variables are suboptimal as the lepton and b-jet are typically not well-separated. Run-I ATLAS analyses used a shrinking isolation cone variable, defined by summing the p_{T} of tracks in a cone around the lepton with radius inversely proportional to the lepton p_{T}. CMS instead relied on kinematic cuts, using a 2D variable requiring the lepton to either be separated in ΔR from the closest jet or have a component of p_{T} transverse to the axis of the closest jet.

3.2 Boosted hadronic tops

Boosted top quarks decaying hadronically are reconstructed by clustering the decay products into a single large-radius (R) jet. Jet grooming techniques are applied to reduce contamination from pileup (PU), initial-state radiation (ISR), and underlying event (UE) to improve the separation in jet mass for signal vs background processes. In addition to the groomed jet mass, both jet substructure variables and dedicated top taggers are used to reconstruct boosted hadronic tops.

Jet grooming: Jet trimming [8], common in Run-I ATLAS analyses, reclusters jets into subjets, discarding subjets with small p_{T} (typically $\sim$5% of the original jet p_{T}). Jet pruning [9], used in Run-I CMS analyses with boosted W bosons, removes soft and wide-angle radiation as part of the jet algorithm. A new technique, soft drop [10], also removes wide-angle soft radiation through recursively declustering the jet. Soft drop is becoming the default in Run-2 for CMS as a powerful method in pushing the QCD jet mass to low values while preserving the hard scale of top quarks or heavy resonances. The soft drop mass for simulated $t\bar{t}$ and QCD multijet events is shown in Figure 1.

HEP Top Tagger: The HEP Top Tagger [11] starts from $R = 1.5$ Cambridge-Aachen jets and declusters the jet into hard subjets using a “mass drop” criteria. All combinations of three subjets are tested for compatibility with a hadronic top decay. PU and UE contributions are removed by reclustering and filtering. Finally, the combination with filtered jet mass closest to the top mass is chosen. Kinematic cuts are applied to differentiate top quarks from background jets, requiring one subjet pair to have an invariant mass around m_W and the combined jet to have a mass around m_{top}. The HEP Top Tagger was used by both ATLAS and CMS for moderately boosted top quarks ($p_{\mathrm{T}} > 200$ GeV).

CMS Top Tagger: The CMS Top Tagger algorithm [12], used extensively in CMS Run-I analyses, declusters jets into up to four subjets, removing soft and wide-angle clusters. $R = 0.8$ jets are used, resulting in an efficient reconstruction for $p_{\mathrm{T}} > 400$ GeV. Three subjets, minimum pairwise mass > 50 GeV, and a jet mass within an m_{top} window are required.

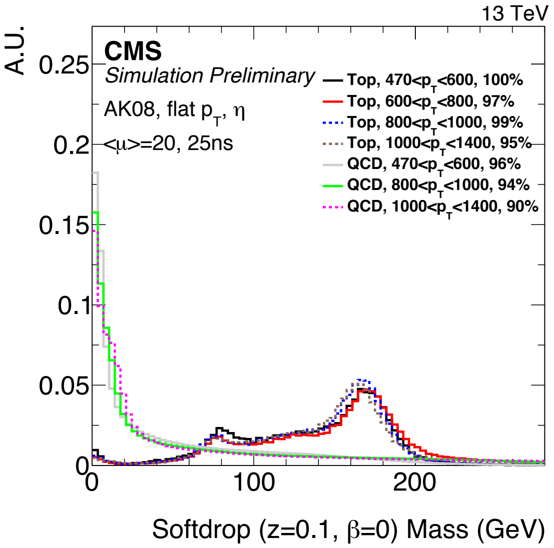


Figure 1: Distribution of soft drop mass for anti-k_t $R = 0.8$ jets for different p_{T} ranges, comparing simulated $t\bar{t}$ and QCD multijet events at 13 TeV from CMS [15].

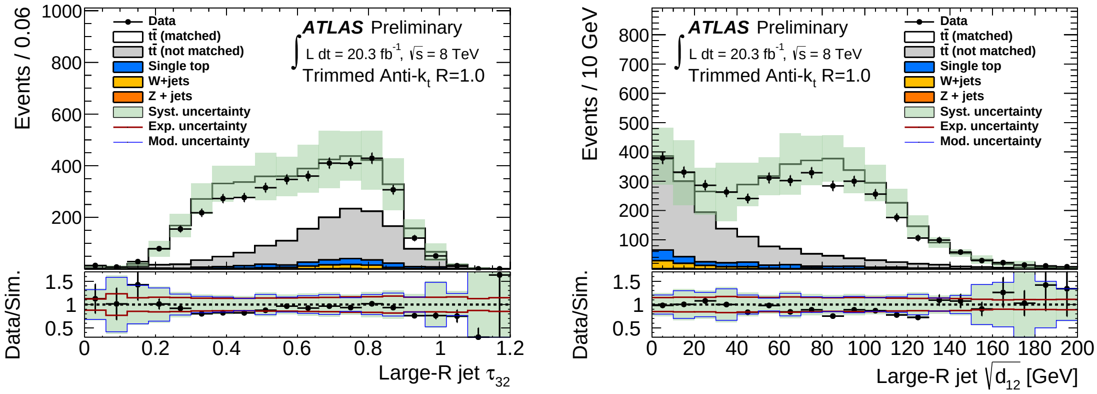
Shower deconstruction: Shower deconstruction [13] defines a likelihood that a large-R jet is due to a signal or background process. Using as input $R = 0.1 - 0.2$ subjets, it is assumed that each subjet comes from a certain source of radiation such as a top decay or ISR. The probability that a given subjet configuration originates from a particular decay chain is then calculated. Shower deconstruction allows for a very pure signal selection.

N-subjettiness: N-subjettiness [14] is a jet shape variable used to measure the consistency of a jet to have N subjets. It is defined as the p_{T}-weighted sum of distances of the constituent particles from the subjet axes: τ_3/τ_2 discriminates between a three-prong subjet structure (top) vs a non-top jet. This variable is shown in Figure 2(a).

k_t splitting scale: Several Run-I ATLAS analyses use the k_t splitting scale, a measure of the scale of the last recombination step in the k_t jet clustering algorithm which clusters high-p_{T} and large-angle components last. The observable corresponding to the splitting scale of the last step, $\sqrt{d_{12}}$, is shown in Figure 2(b). For large-R jets in $t\bar{t}$ signal, it peaks around $m_{\mathrm{top}}/2$.

Subjet b-tagging: Since top decays involve a b-quark, applying subjet b-tagging can significantly increase the QCD multijet background rejection. ATLAS defines an improved b-tagger for additional discrimination in the boosted regime (MVb) [17] and CMS uses an improved tagger which uses as input an inclusive vertex finder algorithm independent, by construction, of the jet algorithm used [18].

The performance of different top taggers and jet substructure variables are studied using simulation. Figure 3(a) shows the background rejection as a function of the tagging efficiency at 8 TeV for ATLAS. The best performance in terms of large background rejection is achieved for shower deconstruction. A common working point in Run-I included trimmed jet mass > 100 GeV with cuts on the k_t splitting scale. Figures 3(b) and 3(c) show similar performance comparisons from CMS at 13 TeV. Shower deconstruction is again the single most discriminating variable, though



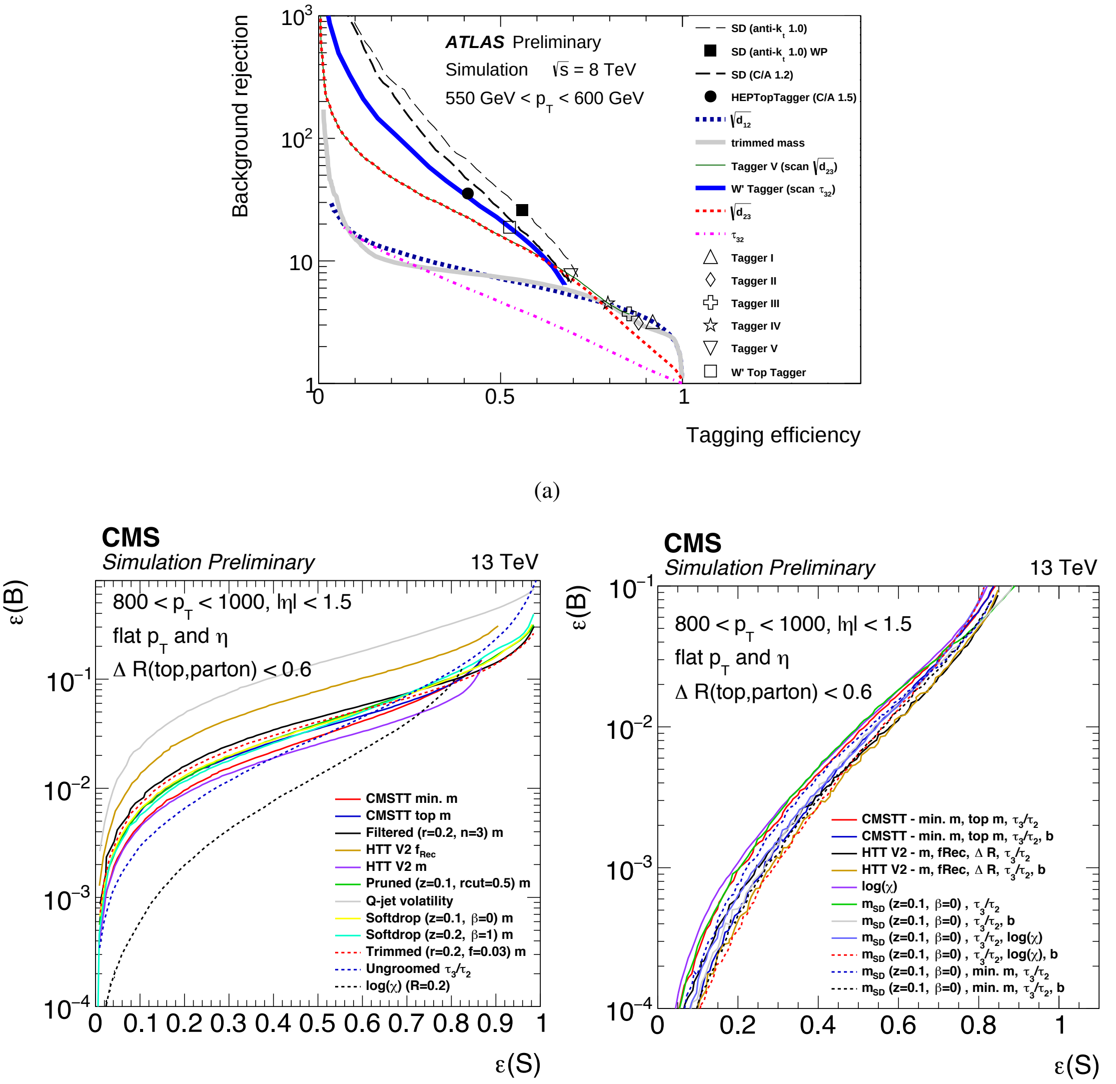
(a) (b)

Figure 2: Distribution of (a) the N-subjettiness ratio τ_3/τ_2 and (b) the splitting scale $\sqrt{d_{12}}$ for trimmed $R = 1.0$ jets in events passing a $t\bar{t}$ signal selection at 8 TeV from ATLAS [16].

it does not reach as high efficiency. For combined taggers, similar performance is achieved for a variety of variables.

4. Conclusions

In conclusion, many different methods are used for reconstructing resolved and boosted top quarks. Jet substructure techniques are used extensively in identifying boosted hadronic top quarks in searches for beyond the Standard Model physics and more recently also in top quark cross-section measurements targeting the high-p_T regime. With higher center-of-mass energy, boosted objects will become even more important in Run-II.



(a)

(b)

(c)

Figure 3: Background rejection as a function of tagging efficiency shown for large-R jets with $550 < p_T < 600$ GeV for different top taggers from ATLAS at 8 TeV in (a) [16]. Taggers I-V correspond to different working points using N-subjettiness, k_t splitting scale, and trimmed jet mass. Background efficiency as a function of tagging efficiency shown for large-R jets with $800 < p_T < 1000$ GeV for (b) single variables and (c) combined variables at 13 TeV from CMS [15].

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.

[2] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST 3 (2008) S08004.

[3] J. Erdmann et al, *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, Nucl. Instrum. Meth. A748 (2014).

[4] ATLAS Collaboration, *Measurement of the top quark mass in the* $t\bar{t} \to$ *lepton+jets and* $t\bar{t} \to$ *dilepton+jets channels using* $\sqrt{s} = 7$ *TeV ATLAS data*, Eur. Phys. J. C (2015) 75:330.

[5] CMS Collaboration, *Measurement of the inclusive and differential top quark pair production cross sections in lepton+jets final states at* $\sqrt{s} = 13$ *TeV*, CMS-PAS-TOP-15-005 (2015).

[6] ATLAS Collaboration, *Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using* pp *collisions at* $\sqrt{s} = 7$ *TeV in the ATLAS detector*, JHEP 06 (2015) 100.

[7] CMS Collaboration, *Measurement of the differential cross section for top quark pair production in* pp *collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C 75 (2015) 542.

[8] D. Krohn, J. Thaler, L. T. Wang, *Jet Trimming*, JHEP 1002 (2010) 084.

[9] S. D. Ellis, C. K. Vermilion, J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, Phys. Rev. D 81 (2010) 094023.

[10] A. J. Larkoski, S. Marzani, G. Soyez, J. Thaler, *Soft Drop*, JHEP 1405 (2014) 146.

[11] T. Plehn, M. Spannowsky, M. Takeuchi, D. Zerwas, *Stop Reconstruction with Tagged Tops*, JHEP 1010 (2010) 078.

[12] CMS Collaboration, *Boosted top jet tagging at CMS*, CMS-PAS-JME-13-007 (2014).

[13] D. E. Soper, M. Spannowsky, *Finding top quarks with shower deconstruction*, Phys. Rev. D 87 (2013) 054012.

[14] J. Thaler, K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP 1103 (2011) 015.

[15] CMS Collaboration, *W/Z/Top tagging for Run 2 in CMS*, CMS DP-2015/043 (2015).

[16] ATLAS Collaboration, *Identification of high transverse momentum top quarks in* pp *collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, ATLAS-CONF-2015-036 (2015).

[17] ATLAS Collaboration, *b-tagging in dense environments*, ATL-PHYS-PUB-2014-014 (2014).

[18] CMS Collaboration, *Performance of b-tagging in boosted topology events*, CMS DP-2014/031 (2014).