

## Identification of boosted hadronically decaying particles with jet substructure in ATLAS Run-2

---

**Jason Veatch\***

*on behalf of the ATLAS Collaboration*

*Georg-August-Universität Göttingen*

*E-mail: [jason.veatch@cern.ch](mailto:jason.veatch@cern.ch)*

In order to exploit the abundance of high momentum heavy particles produced at LHC (top,  $W/Z$  and Higgs) decaying hadronically, the study of jet substructure has become increasingly important throughout a wide array of searches and measurements. The latest ATLAS results in terms of jet substructure measurements and calibrations are presented including studies of performance sensitivity to pileup, with several grooming algorithms and recently developed constituent level pile-up mitigation techniques.

*XXVI International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2018)  
16-20 April 2018  
Kobe, Japan*

---

\*Speaker.

## 1. Introduction

Large- $R$  jets play a vital role in the physics program at ATLAS [1] using high energy LHC  $pp$  collisions. As searches for Beyond Standard Model physics probes higher mass scales, boosted hadronic final states become increasingly important. In order to better exploit this tool, many advancements have been made to improve the usage of large- $R$  jets to search for new physics. This note presents the work of several published ATLAS analyses and is meant to provide an overview of the methodologies and results.

## 2. Alternative jet definitions

The standard large- $R$  jet definition used in ATLAS to date uses locally-calibrated topological (LCTopo) clusters of energy deposited in calorimeters clustered into jets using the anti- $k_t$  algorithm [2] with  $R = 1.0$  and trimmed [3] with  $R_{\text{sub}} = 0.2$  and  $f_{\text{cut}} = 0.05$ . An analysis [4] is presented that compares the standard large- $R$  jet definition to additional definitions that make use of modified inputs and alternative grooming techniques. The alternative inputs used in the analysis are LCTopo clusters modified with combinations of Voronoi Subtraction [5], Constituent Subtraction (CS) [6], and SoftKiller (SK) [7]. The alternate grooming algorithms include pruning [8], the Modified Mass-Drop Tagger [9], SoftDrop [10], and reclustered jets.

The pile-up dependence of the reconstructed jet mass with each type of input is shown in Figure 1a. Additionally, a comparison of many different inputs and grooming algorithms with a range of parameter values is performed and the four configurations with the best QCD rejection using a 68% efficient mass window for  $W$  tagging are selected. Mass and energy calibrations are derived and applied for each of these configurations and the resulting QCD rejection as a function of  $W$  tagging efficiency for each configuration is shown in Figure 1b using  $\text{mass} + D_2^{\beta=1.0}$ , where  $D_2^{\beta=1.0}$  [11, 12] is an energy correlator ratio that is sensitive to the angular distribution of energy inside a jet.

## 3. Track CaloClusters

At high  $p_T$ , the coarse angular resolution of calorimeter clusters causes a degradation in the resolution of jet substructure observables as clusters merge. This can be mitigated through the use of novel jet inputs that combine charged particle tracks with calorimeter clusters into objects called TrackCaloClusters (TCC) [13]. Tracks are matched to clusters and, in the simplest case of a single track matched to a single cluster, a resulting TCC's energy is that of the cluster and its position is defined as that of the track. This use of tracking information results in improved jet substructure resolutions as shown in Figure 2, especially for observables that are very sensitive to angular distributions such as  $D_2^{\beta=1.0}$ . Additionally, the excellent primary vertex association of tracks causes TCC jets to be more robust against the effects of pile-up than LCTopo jets.

## 4. Reclustered jets

In addition to jets built from calorimeter topological clusters, large- $R$  jets built from calibrated anti- $k_t$   $R = 0.4$  jets are studied [14]. This reclustering (RC) approach makes use of the excellent

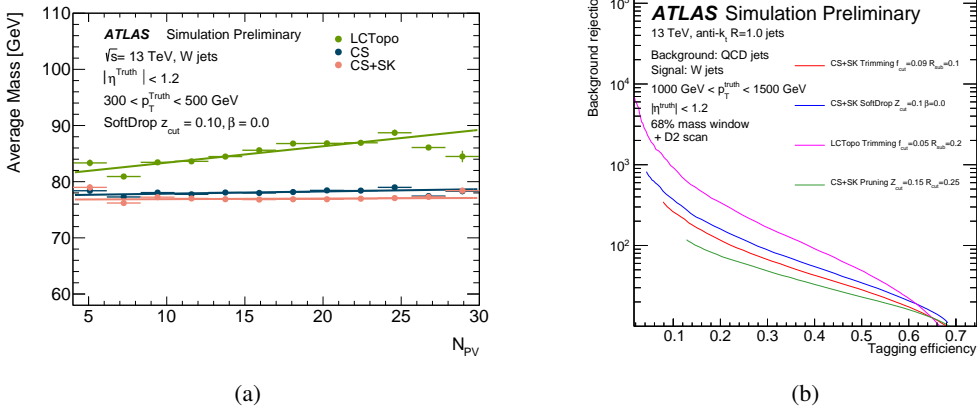


Figure 1: The jet mass of W jets as a function of  $N_{pV}$  (a) shows a strong pile-up dependence for LCTopo jet inputs, which can be mitigated with CS and SK. A comparison of mass+ $D_2^{\beta=1.0}$  W tagging ROC curves for three select LCTopo+CS+SK jet configurations and the standard large-R jet definition (b) shows the superior performance of the standard definition. These are both example plots from the extensive parameter space studied in the Ref. [4].

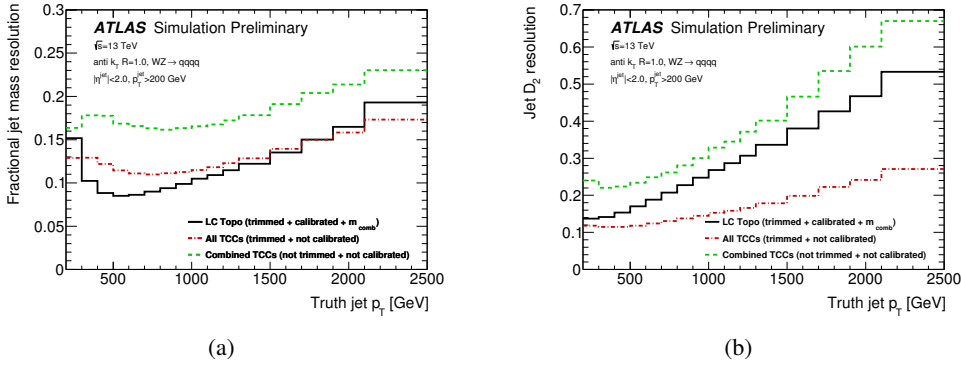


Figure 2: Resolution of jet mass (a) and  $D_2^{\beta=1.0}$  (b) for LCTopo jets and TCC jets [13]. Significant improvements are seen for TCC jets, especially for  $D_2^{\beta=1.0}$ , which is very sensitive to angular distributions.

calibrations and uncertainties [15] as well as the pile-up suppression [16] available for  $R = 0.4$  jets. Figure 4 shows the improved jet mass resolution and the reduced jet energy scale uncertainties from RC jets.

### 5. Measuring large-R jet response with *in situ* techniques

Several *in situ* techniques are used to measure large-R jet energy and mass response in MC and data [17]. The jet energy response is measured using momentum balance in events in which a large-R jet recoils against a well-measured object, either a photon, multiple small-R jets, or another large-R jet. As an example, the uncertainty on the jet energy scale resolution in events using recoiled

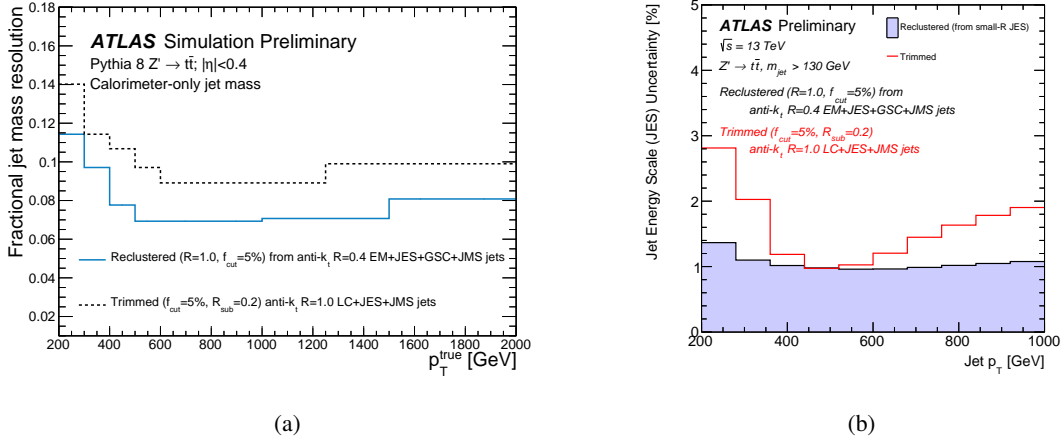


Figure 3: RC jets provide better jet mass resolution (a) and jet energy scale uncertainties (b) than standard large- $R$  jets [14]. Top jets are shown here, but similar performance gains are seen for  $W$  jets.

small- $R$  jets is shown in Figure 4a. The jet mass response is measured using the  $R_{\text{trk}}$  method [18, 19] in which the ratio of cluster-based to track-based observables is used and the forward folding method in which a truth-level jet mass distribution is shifted and smeared to match that in data. The different methods for measuring the jet energy response are combined for a final comparison to data, as are the methods for measuring the jet mass, which is shown in Figure 4b. The *in situ* methods result in jet energy and mass scale uncertainties of  $< 5\%$  and agreement between data and MC within a few percent.

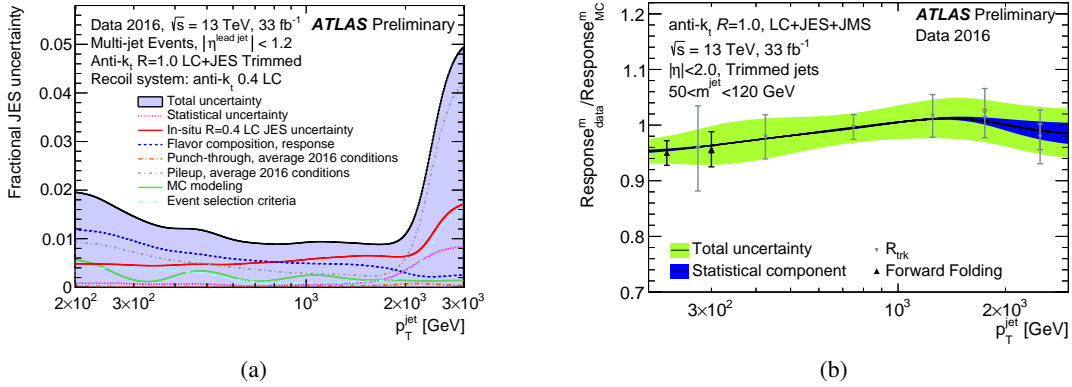


Figure 4: A breakdown of the total uncertainty on the fractional jet energy scale using events in which a large- $R$  jet recoils against multiple small- $R$  jets (a) and the relative jet mass response between data a MC using a combination of the  $R_{\text{trk}}$  and forward folding methods (b) [17].

## 6. $W$ boson and top quark tagging

Improved techniques to identify large- $R$  jets as boosted hadronically decaying  $W$  bosons or

top quarks are developed and validated. Simple tagging techniques were developed at ATLAS in Run 1 using mass and addition jet substructure observables for  $W$  [20] and top [21] jets. Furthermore, HEPTopTagger [22, 23] was shown to be effective at identifying top jets. An analysis using Run 2 data and MC [24] compares the use of such taggers to taggers based on machine learning techniques, BDT and DNN. The performance of the BDT and DNN taggers are found to be well-modeled in MC as shown, as an example, by the similar data and MC jet mass spectra (shape and normalization) after the DNN tagger is applied as shown in Figure 5a. Furthermore, both machine learning techniques provide superior background rejection compared to the other techniques as shown in Figure 5b.

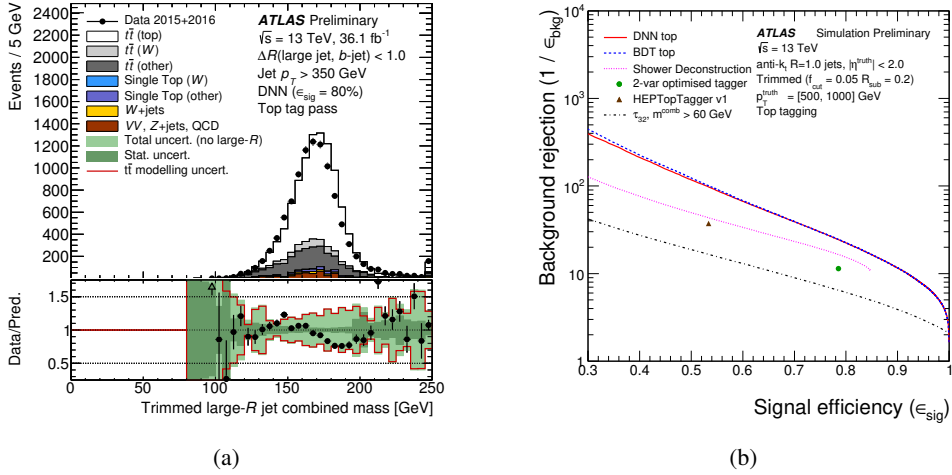


Figure 5: A comparison of jet mass with the 80% efficient DNN top tagger applied to both MC and data (a) shows that the tagger is well-modeled in MC within systematic uncertainties. A comparison of the top tagging ROC curves (a) shows that the BDT and DNN taggers have similar performances and are superior to non-machine-learning techniques. The results for  $W$  and top tagging are similar [24].

### 7. $H \rightarrow b\bar{b}$ tagging

Techniques for identifying boosted  $H \rightarrow b\bar{b}$  decays based on the presence of  $b$ -tagged jets within a large- $R$  jet and jet substructure observables are studied. It is shown [25] that the identification of one or two  $b$ -tagged anti- $k_t R = 0.2$  track-jets within a large- $R$  jet in addition to a jet mass cut provides an effective way of identifying  $H \rightarrow b\bar{b}$  jets, as shown in Figure 6a. The performance of this method degrades for high- $p_T$   $H \rightarrow b\bar{b}$  decays due to the merging of the  $b$ -jets, resulting in a lower  $b$ -tagging efficiency. This can be mitigated using any of three techniques studied in Ref. [26]. Variable- $R$  track-jets, exclusive  $k_t$  calorimeter subjets, and center-of-mass subjet reconstruction all offer improved performance at high  $p_T$  as demonstrated by the double  $b$ -tagging efficiency as a function of jet  $p_T$  using these three techniques is shown in Figure 6b.

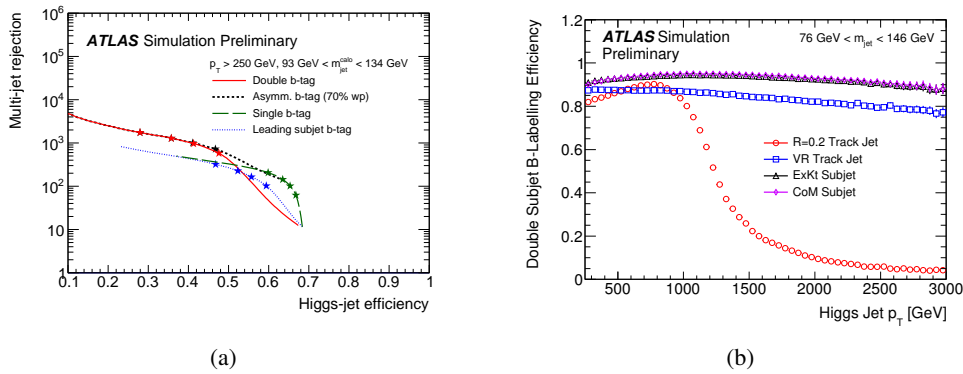


Figure 6: The QCD background rejection vs.  $H \rightarrow b\bar{b}$  tagging efficiency using a jet mass cut and associated  $b$ -tagged small- $R$  track jets [25] is shown in (a). The curves are a result of scanning the  $b$ -tagging efficiency and the 60%, 70%, 77%, and 80% efficiency working points are indicated by stars. The double  $b$ -tagging performance increase at high- $p_T$  using different algorithms [26] is shown in (b).

## References

- [1] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.
- [2] M. Cacciari, G. P. Salam, and G. Soyez. “The Anti- $k(t)$  jet clustering algorithm”. In: *JHEP* 0804 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].
- [3] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jet Trimming”. In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].
- [4] ATLAS Collaboration. *Impact of Alternative Inputs and Grooming Methods on Large- $R$  Jet Reconstruction in ATLAS*. ATL-PHYS-PUB-2017-020. 2017. URL: <https://cdsweb.cern.ch/record/2297485>.
- [5] ATLAS Collaboration. *Constituent-level pile-up mitigation techniques in ATLAS*. ATLAS-CONF-2017-065. 2017. URL: <https://cds.cern.ch/record/2281055>.
- [6] Peter Berta et al. “Particle-level pileup subtraction for jets and jet shapes”. In: *JHEP* 06 (2014), p. 092. DOI: 10.1007/JHEP06(2014)092. arXiv: 1403.3108 [hep-ex].
- [7] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “SoftKiller, a particle-level pileup removal method”. In: *Eur. Phys. J. C* 75.2 (2015), p. 59. DOI: 10.1140/epjc/s10052-015-3267-2. arXiv: 1407.0408 [hep-ph].
- [8] Stephen D. Ellis, Christopher K. Vermilion, and Jonathan R. Walsh. “Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches”. In: *Phys. Rev. D* 81 (2010), p. 094023. DOI: 10.1103/PhysRevD.81.094023. arXiv: 0912.0033 [hep-ph].
- [9] Mrinal Dasgupta et al. “Towards an understanding of jet substructure”. In: *JHEP* 09 (2013), p. 029. DOI: 10.1007/JHEP09(2013)029. arXiv: 1307.0007 [hep-ph].

- [10] Andrew J. Larkoski et al. “Soft Drop”. In: *JHEP* 05 (2014), p. 146. DOI: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph].
- [11] Andrew J. Larkoski, Ian Moult, and Duff Neill. “Analytic Boosted Boson Discrimination”. In: *JHEP* 05 (2016), p. 117. DOI: 10.1007/JHEP05(2016)117. arXiv: 1507.03018 [hep-ph].
- [12] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. “Energy Correlation Functions for Jet Substructure”. In: *JHEP* 06 (2013), p. 108. DOI: 10.1007/JHEP06(2013)108. arXiv: 1305.0007 [hep-ph].
- [13] ATLAS Collaboration. *Improving jet substructure performance in ATLAS using Track-CaloClusters*. ATL-PHYS-PUB-2017-015. 2017. URL: <https://cds.cern.ch/record/2275636>.
- [14] ATLAS Collaboration. *Jet reclustering and close-by effects in ATLAS Run 2*. ATLAS-CONF-2017-062. 2017. URL: <https://cds.cern.ch/record/2275649>.
- [15] ATLAS Collaboration. “Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *Phys. Rev. D* 96 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].
- [16] Matteo Cacciari and Gavin P. Salam. “Pileup subtraction using jet areas”. In: *Phys. Lett. B* 659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].
- [17] ATLAS Collaboration. *In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions*. ATLAS-CONF-2017-063. 2017. URL: <https://cds.cern.ch/record/2275655>.
- [18] ATLAS Collaboration. “Performance of jet substructure techniques for large- $R$  jets in proton–proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector”. In: *JHEP* 09 (2013), p. 076. DOI: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].
- [19] ATLAS Collaboration. “Jet mass and substructure of inclusive jets in  $\sqrt{s} = 7$  TeV  $pp$  collisions with the ATLAS experiment”. In: *JHEP* 05 (2012), p. 128. DOI: 10.1007/JHEP05(2012)128. arXiv: 1203.4606 [hep-ex].
- [20] ATLAS Collaboration. “Identification of Boosted, Hadronically Decaying  $W$  Bosons and Comparisons with ATLAS Data Taken at  $\sqrt{s} = 8$  TeV”. In: *Eur. Phys. J. C* 76 (2016), p. 154. DOI: 10.1140/epjc/s10052-016-3978-z. arXiv: 1510.05821 [hep-ex].
- [21] ATLAS Collaboration. “Identification of high transverse momentum top quarks in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector”. In: *JHEP* 06 (2016), p. 093. DOI: 10.1007/JHEP06(2016)093. arXiv: 1603.03127 [hep-ex].
- [22] Tilman Plehn, Gavin P. Salam, and Michael Spannowsky. “Fat Jets for a Light Higgs”. In: *Phys. Rev. Lett.* 104 (2010), p. 111801. DOI: 10.1103/PhysRevLett.104.111801. arXiv: 0910.5472 [hep-ph].
- [23] Tilman Plehn et al. “Stop Reconstruction with Tagged Tops”. In: *JHEP* 10 (2010), p. 078. DOI: 10.1007/JHEP10(2010)078. arXiv: 1006.2833 [hep-ph].

- [24] ATLAS Collaboration. *Performance of Top Quark and W Boson Tagging in Run 2 with ATLAS*. ATLAS-CONF-2017-064. 2017. URL: <https://cds.cern.ch/record/2281054>.
- [25] ATLAS Collaboration. *Boosted Higgs ( $\rightarrow b\bar{b}$ ) Boson Identification with the ATLAS Detector at  $\sqrt{s} = 13$  TeV*. ATLAS-CONF-2016-039. 2016. URL: <https://cds.cern.ch/record/2206038>.
- [26] ATLAS Collaboration. *Variable Radius, Exclusive- $k_T$ , and Center-of-Mass Subjet Reconstruction for Higgs ( $\rightarrow b\bar{b}$ ) Tagging in ATLAS*. ATL-PHYS-PUB-2017-010. 2017. URL: <https://cds.cern.ch/record/2268678>.