Performance of large-\(R\) jets and jet substructure reconstruction with the ATLAS detector

The ATLAS Collaboration

Abstract

This paper presents the application of techniques to study jet substructure. The performance of modified jet algorithms for a variety of jet types and event topologies is investigated. Properties of jets subjected to the mass-drop filtering, trimming and pruning algorithms are found to have a reduced sensitivity to multiple proton-proton interactions and exhibit improved stability at high luminosity. Monte Carlo studies of the signal-background discrimination with jet grooming in new physics searches based on jet invariant mass and jet substructure properties are also presented. The application of jet trimming is shown to improve the robustness of large-\(R\) jet measurements, reduce sensitivity to the superfluous effects due to the intense environment of the high luminosity LHC, and improve the physics potential of searches for heavy boosted objects. The analyses presented in this note use the full 2011 ATLAS dataset, corresponding to an integrated luminosity of 4.7 ± 0.2 fb\(^{-1}\).
1 Introduction

Jets have historically been utilized at high energy colliders as proxies for the quarks and gluons produced in the primary collisions. With the high-luminosity conditions at the LHC, soft particles unrelated to the hard scattering contaminate jets in the detector considerably more than at previous experiments, making it more difficult to resolve the particles originating from the hard interaction. This is especially true for boosted objects whose decay products may be sufficiently collimated such that standard reconstruction techniques begin to fail. In events where such decays are fully contained within individual large-radius jets, a diminished mass resolution due to the high luminosity environment weakens the sensitivity to new physics processes.

One example of a new physics process which may produce heavy objects with a significant Lorentz boost is the decay of a new heavy gauge boson, the $Z'$, to top quark pairs. Figure 1 shows the true angular separation between the $W$ and $b$ decay products of a top quark in simulated $Z' \to t\bar{t}$ ($m_{Z'} = 1.6$ TeV) events, as well as the separation between the light quarks of the subsequent hadronically-decaying $W$. In each case, the angular separation of the decay products is approximately

$$\Delta R \approx \frac{2m}{p_T},$$

where $\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$, and $p_T$ and $m$ are the transverse momentum and mass of the decaying particle, respectively. For $p_T^W > 200$ GeV, the ability to resolve the individual hadronic decay products using standard narrow-cone jet algorithms begins to degrade, and above $p_T^{top} > 350$ GeV, the decay products of the top quark tend to have a separation $\Delta R < 1.0$. Techniques designed to recover sensitivity in such cases focus on large-$R$ jets in order to maximize efficiency. At $\sqrt{s} = 7$ TeV, nearly one thousand Standard Model $t\bar{t}$ events per $fb^{-1}$ are expected in this momentum regime. It is in this region of phase space, which was limited by statistics and the available energy at previous colliders, where new physics may appear.

A single jet that contains all of the decay products of a massive particle will have significantly different properties than a single jet of the same $p_T$ originating from a single light-quark or gluon. The characteristic two-body or three-body decays of a vector boson or top quark result in a hard substructure that is absent from the light-quark and gluon jets, and this can be more clearly resolved by removing soft radiation from jets. This selective removal of soft radiation during the process of iterative recombination in jet reconstruction is generally referred to as jet “grooming”.

Recently many jet grooming algorithms have been designed to remove contributions to a given jet that are irrelevant or detrimental to resolving the hard decay products from a boosted object. The structural differences between jets formed from light quarks or gluons and individual jets originating from a boosted hadronic particle decay form the basis for these tools. Consequently, the characteristic substructure within such a jet is retained while reducing the impact of the fluctuations of the parton shower and underlying event (UE), and mitigating the influence of pile-up (additional $pp$ collisions apart from the primary hard collision in an event). Thus, a groomed jet can also be a powerful tool to discriminate between the dominant multi-jet background and the heavy particle decay, thereby increasing signal sensitivity.

---

1The ATLAS coordinate system is a right-handed system with the x-axis pointing to the center of the LHC ring and the y-axis pointing upwards. The polar angle $\theta$ is measured with respect to the LHC beam-line. The azimuthal angle $\phi$ is measured with respect to the x-axis. The rapidity is defined as $y = 0.5 \times \ln[(E + p_z)/(E - p_z)]$, where $E$ denotes the energy and $p_z$ is the component of the momentum along the beam direction. The pseudorapidity $\eta$ is an approximation for rapidity $y$ in the high energy limit, and it is related to the polar angle $\theta$ as $\eta = -\ln \tan \frac{\theta}{2}$. Transverse momentum and energy are defined as $p_T = p \times \sin \theta$ and $E_T = E \times \sin \theta$, respectively.

2In this paper, “large-$R$” refers to jets with a distance parameter $R \geq 1.0$.
Figure 1: (a) The opening angle between the $W$ and $b$ in top decays, $t \rightarrow Wb$, as a function of the top $p_T$ in simulated PYTHIA $Z' \rightarrow tf$ ($m_{Z'} = 1.6$ TeV) events. (b) The opening angle of the $W \rightarrow q\bar{q}$ system from $t \rightarrow Wb$ decays as a function of the $W$ $p_T$. Both distributions are at the particle level.

This note presents the results of a comprehensive study of the performance of jet grooming algorithms using the 2011 ATLAS dataset corresponding to an integrated luminosity of $(4.7 \pm 0.2) \text{fb}^{-1}$. Three jet grooming techniques are studied: mass-drop filtering, trimming, and pruning. These techniques utilize the internal structure of the jet in order to reduce the sensitivity to pile-up and UE, as well as improve jet mass resolution.

Measurements of groomed jet properties in the presence of pile-up (along with a companion study [1]) are made for jets across a wide range of jet transverse momentum ($p_{\text{jet}}^T$). Comparisons are made to generators incorporating leading-order (LO), e.g. PYTHIA [2] and next-to-leading-order (NLO), e.g. POWHEG [3,4] matrix elements, interfaced to PYTHIA for parton showering and hadronization, as well as to GEANT4 for full detector simulation.

Section 2 describes the design and implementation of the grooming algorithms used in ATLAS, as well as the definitions of the various substructure observables discussed throughout the text. The data samples and Monte Carlo simulation used for comparison are introduced in Section 3. The performance and validation of jet calibrations for large-$R$ and groomed jets described in Section 4 provide the starting point necessary to establish the use of these new jet algorithms in physics analyses. Data to MC comparisons are then discussed in Section 5 for the jet substructure observables introduced in Section 2. Finally, Section 6 presents comparisons of grooming algorithms applied to multiple QCD jet events and a sample selected to contain high $p_T$ hadronically-decaying top quarks in data.

2 Jet Grooming Algorithms and Substructure Observables in ATLAS

This section describes jet reconstruction algorithms and presents three jet algorithm modification procedures studied in ATLAS, referred to as jet “grooming.” Mass-drop filtering, trimming, and pruning are described and performance measures related to each are defined. The different configurations of the grooming algorithms described in this section are summarized in Table 1. Additionally, a technique to tag boosted top quarks using the mass-drop/filtering method is introduced.
2.1 Inputs to jet reconstruction

The inputs to jet reconstruction, or “proto-jets”, are either stable particles with a lifetime of at least 10 ps (excluding muons and neutrinos) in the case of MC “truth jets”, charged particle tracks in the case of so-called “track jets” [5], or three-dimensional topo-clusters [6] in the case of fully reconstructed calorimeter jets. In the reconstruction of track jets, track quality selection is applied in order to ensure good quality tracks that originate from the primary reconstructed vertex which has the largest \( \sum (p_T^{\text{track}})^2 \) in the event and contains at least two tracks. The selection criteria are:

- Transverse momentum: \( p_T^{\text{track}} > 0.5 \) GeV;
- Transverse impact parameter: \(|d_0| < 1.0 \) mm;
- Longitudinal impact parameter: \(|z_0| \times \sin(\theta) < 1.0 \) mm;
- Silicon detector hits on tracks: \( N_{\text{Pixel}} \geq 1, N_{\text{SCT}} \geq 6 \);

where the impact parameters are computed with respect to the primary vertex, and \( \theta \) is the angle between the track and the beam. In the reconstruction of calorimeter jets, calorimeter cells are clustered together using a topological clustering algorithm. These objects provide a three-dimensional representation of energy depositions in the calorimeter with a nearest neighbor noise suppression algorithm. The resulting topo-clusters are then classified as either electromagnetic or hadronic based on their shape, depth and energy density. Energy corrections are then applied in order to calibrate the clusters to the hadronic scale.

2.2 Jet algorithms

Three jet algorithms are studied here: the anti-\( k_t \) algorithm [7], the Cambridge-Aachen (C/A) algorithm [8, 9], and the \( k_t \) algorithm [10, 11]. These algorithms are implemented within the framework of the FastJet software [12, 13]. They represent the most widely used infrared and collinear-safe jet algorithms available for hadron-hadron collider physics today. Furthermore, in the case of the \( k_t \) and C/A algorithms, the clustering history of the algorithm – that is, the ordering and structure of the pair-wise recombinations made during jet reconstruction – provides spatial and kinematic information about the anatomy of that jet. The anti-\( k_t \) algorithm provides stable, robust jets that are defined primarily by the highest-\( p_T \) constituent. The compromise is that the structure of the jet as defined by the anti-\( k_t \) algorithm carries little or no information about the ordering of the shower or wide angular-scale structure. It is, however, possible to exploit this stability and recover meaningful information about the jet substructure: anti-\( k_t \) jets are selected for analysis based on their kinematics (\( \eta \) and \( p_T \)), and then the jet constituents are reclustered with the \( k_t \) algorithm to enable use of the \( k_t \)-ordered splitting scales described in Section 2.3. The four-momentum recombination scheme is used in all cases and the jet finding is performed in rapidity-azimuthal angle (\( y-\phi \)) coordinates. Jet selections and corrections are made in pseudorapidity-azimuthal angle (\( \eta-\phi \)) coordinates.

The iterative recombination procedure works by first creating a list of all objects (either hadrons, topo-clusters or tracks) in an event. The ordering of the list is irrelevant, and proto-jets are formed from these objects. Two distance measures in \( y-\phi \)-space are associated to each member of the list, between the proto-jet and its closest neighbor (as defined in Eq. 2) and between the proto-jet and the beam (Eq. 3).

\[
\rho_{ij} = \min\left(p_{Ti}^{2p}, p_{Tj}^{2p}\right) \frac{(\Delta R_{ij})^2}{R^2}. \tag{2}
\]

\[
\rho_{iB} = p_{Ti}^{2p} \tag{3}
\]
Here, $p_{Ti}$ is the transverse momentum of the proto-jet, $p$ is an integer, $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ is a measure of the opening angle between the two constituents, and $R$ is the jet radius parameter used by the algorithm to define the final size of the jet. The two distances are then compared:

- If $\rho_{iB} < \rho_{ij}$ then the proto-jet is “closer” to the beam than it is to any other proto-jet in the event, so it is defined as a jet and removed from the list.
- If $\rho_{iB} > \rho_{ij}$ then the two proto-jets $i$ and $j$ are combined into one, thereby forming a new proto-jet. This procedure continues through all proto-jets in the event.

The variation between the jet algorithms comes from the value of $p$ in the exponent of $p_T$ in Eq. (2) and Eq. (3):

$p = +1$: the $k_t$ algorithm. Proto-jets with the smallest $p_T$ tend to be clustered first, so that the highest $p_T$ proto-jets are combined last.

$p = 0$: the C/A algorithm. In this case no $p_T$ information is used at all. Proto-jets are combined based purely on their angular separation from one another and from the beam.

$p = -1$: the anti-$k_t$ algorithm. Proto-jets with the largest $p_T$ are clustered first. A consequence of this is that isolated anti-$k_t$ jets tend to be very close to circular in $\eta$-$\phi$ space, because the axis of the jet is relatively fixed after the first few steps of recombination. This stability makes anti-$k_t$ jets more robust than $k_t$ jets in high multiplicity environments. ATLAS has adopted anti-$k_t$ jets as standard in physics analyses.

2.3 Jet properties and substructure observables

Jet mass: The jet mass is calculated from the energies and momenta of its constituents, as given in Eq. (4):

$$ (m_{\text{jet}}^\text{jet})^2 = (\sum_i E_i)^2 - (\sum_i p_i)^2, \quad (4) $$

where $E_i$ and $p_i$ are the energy and three-momentum of the $i^{th}$ jet constituent. At the detector level, the jet mass is calculated by summing over the topo-clusters or tracks used for jet reconstruction. At the particle level, the jet constituents are stable particles. The jet mass is an essential measurement in the search for boosted, high-mass particles, and can be used as a powerful discriminant between signal and background.

Splitting scales: The $k_t$ splitting scales are defined by reclustering the constituents of a jet with the $k_t$ recombination algorithm. The $k_t$-distance of the final step in combining two proto-jets, referred to as subjets in this case, into the final jet can be used to define a “splitting scale” variable given in Eq. (5) as:

$$ \sqrt{d_{ij}} = \text{min}(p_{Ti}, p_{Tj}) \times \Delta R_{ij}, \quad (5) $$

where $\Delta R_{ij}$ is the distance between the two subjets. With this definition, the subjets identified at the last step of the reclustering in the $k_t$ algorithm provide the $\sqrt{d_{12}}$ observable. Similarly, $\sqrt{d_{23}}$ defines the splitting scale in the second to the last step of the reclustering. These definitions are equivalent to the square root of the distance parameter $\rho_{ij}$ given in Eq. (2) multiplied by the jet radius parameter $R$ in order to remove the explicit dependence on the nominal jet radius. As described in Section 2.2, the $k_t$ algorithm combines the harder constituents last. Because of this, and the fact that $\sqrt{d_{ij}}$ uses the minimum $p_T$ between the $i$ and $j$ subjets, the parameters $\sqrt{d_{12}}$ and $\sqrt{d_{23}}$
\( \sqrt{d_{23}} \) can be used to distinguish heavy particle decays, which tend to be reasonably symmetric, from largely asymmetric splittings in light quark or gluon jets. The expected value for a two-body heavy particle decay is approximately \( \sqrt{d_{12}} \approx m^{\text{jet}}/2 \), whereas jets from the parton shower of light quarks and gluons will tend to exhibit a steeply falling spectrum for both \( \sqrt{d_{12}} \) and \( \sqrt{d_{23}} \) (see Figure 23 in Section 5.3).

\textbf{N-subjettiness:} The \( N \)-subjettiness variables \( \tau_N \) \cite{14,15} are observables related to the subject multiplicity. The \( \tau_N \) variable is calculated by clustering the constituents of the jet with the \( k_t \) algorithm and requiring exactly \( N \) subjets to be found. This is done using the exclusive \( k_t \) algorithm \cite{11} and is based on reconstructing clusters of particles in the jet using all of the jet constituents. The exclusive mode of the \( k_t \) algorithm discards constituents for which \( \rho_{iB} < \rho_{ij} \) and stops clustering either when all \( \rho_{iB} \) and \( \rho_{ij} \) are above some specified \( \rho_{\text{cut}} \) or when there are \( N \) proto-jets remaining. The latter case is used here. These \( N \) final subjets define axes within the jet, around which the jet constituents may be concentrated. The variables \( \tau_N \) are then defined in Eq. (6) as the sum over all constituents \( k \) of the jet:

\[
\tau_N = \frac{1}{d_0} \sum_k p_{T_k} \times \min(\delta R_{1k}, \delta R_{2k}, \ldots, \delta R_{Nk}) , \quad \text{with} \quad d_0 \equiv \sum_k p_{T_k} \times R \quad (6)
\]

where \( R \) is the jet radius parameter in the jet algorithm, \( p_{T_k} \) is the \( p_T \) of constituent \( k \) and \( \delta R_{ik} \) is the distance from the subject \( i \) to constituent \( k \). Using this definition, \( \tau_N \) describes how well jets can be described as containing \( N \) or fewer \( k_t \) subjets by assessing the degree to which constituents are localized near the axes of these subjets. The ratios \( \tau_2/\tau_1 \) and \( \tau_3/\tau_2 \) can be used to provide discrimination between jets formed from the parton shower of light quarks or gluons and jets containing two hadronic decay products (from Z-bosons, for example) or three hadronic decay products from boosted top quarks. These ratios will herein be referred to as \( \tau_{21} \) and \( \tau_{32} \) respectively. For example, \( \tau_{21} \approx 1 \) corresponds to a jet that is very well described by a single subjet whereas a lower value implies a jet that is much better described by two subjets than one.

2.4 Jet grooming algorithms

\textbf{Mass-drop Filtering:} The mass-drop filtering procedure seeks to isolate concentrations of energy within a jet by identifying relatively symmetric subjets, each with a significantly smaller mass than that of their sum. This technique was developed and optimized using C/A jets in the search for a Higgs boson decaying to two \( b \)-quarks: \( H \rightarrow b \bar{b} \) \cite{16}. The procedure is applied only to C/A jets since each clustering step of the algorithm combines the two widest angle proto-jets at that point in the shower history. Therefore, the structure of the C/A jet provides an angular-ordered description of substructure, which tends to be one of the most useful properties when searching for hard splittings within a jet. Although the mass-drop criterion and subsequent filtering procedure are not specifically based on soft-\( p_T \) or wide-angle selections, the algorithm does retain the hard components of the jet through the requirements placed on its internal structure. The first measurements of the jet mass of these filtered jets was performed using 35 pb\(^{-1}\) of data collected in 2010 by the ATLAS experiment \cite{17}.

The mass-drop/filtering procedure has two stages:

- \textbf{Mass-drop and symmetry} Undo the last stage of the C/A clustering so that the jet “splits” into two subjets, \( j_1 \) and \( j_2 \), ordered such that the mass of \( j_1 \) is larger: \( m^{j_1} > m^{j_2} \). The mass-drop criterion requires that there be a significant difference between the original jet mass \( (m^{\text{jet}}) \) and \( m^{j_1} \) after the splitting:

\[
m^{j_1}/m^{\text{jet}} < \mu_{\text{frac}} \quad (7)
\]
where $\mu_{\text{frac}}$ is a parameter of the algorithm. The splitting is also required to be relatively symmetric:
\[
\frac{\min[(p_{T}^{j_{1}})^{2}, (p^{j_{2}}_{T})^{2}]}{(m_{\text{jet}})^{2}} \times \Delta R_{j_{1},j_{2}}^{2} > y_{\text{cut}},
\]
where $\Delta R_{j_{1},j_{2}}$ is the opening angle between $j_{2}$ and $j_{1}$, and $y_{\text{cut}}$ defines the energy sharing between the two highest $p_{T}$ subjets within the original jet. For the analyses presented here, $y_{\text{cut}}$ is set to 0.09, the optimal value obtained in previous studies [16]. To give a sense of the kinematic requirements that this places on a given decay, consider a hadronically decaying $W$ boson with $p_{T}^{W} \approx 200$ GeV. According to the approximation given by Eq. (1), the average angular separation of the two daughter quarks is $\Delta R_{j_{1},j_{2}} \sim 0.8$. The symmetry requirement determined by $y_{\text{cut}}$ in Eq. (8) thereby implies that the transverse momentum of the softer (in $p_{T}$) of the two subjets be greater than approximately 30 GeV. Generally, the requirement entails a minimum $p_{T}$ of the softer subjet of $p_{T}^{\text{subj}} / p_{T}^{\text{jet}} > 0.15$, thus forcing both subjets to carry some significant fraction of the momentum of the original jet. This procedure is illustrated in Figure 2(a). If the mass-drop and symmetry criteria are not satisfied, the jet is discarded.

- **Filtering** The constituents of $j_{1}$ and $j_{2}$ are reclustered using the C/A algorithm with $R_{\text{filt}} < \Delta R_{j_{1},j_{2}}$, where $R_{\text{filt}} = \min[0.3, \frac{\Delta R_{j_{1},j_{2}}}{2}]$. The jet is then filtered; all constituents outside the three hardest subjets are discarded. In isolating $j_{1}$ and $j_{2}$ with the C/A algorithm, the angular scale of any potential massive particle decay is known. By dynamically reclusterking the jet at an appropriate angular scale able to resolve that structure, the sensitivity to highly collimated decays is maximized. This is illustrated in Figure 2(b).

In this analysis, three values of the mass-drop parameter $\mu_{\text{frac}}$ are studied, as summarized in Table 1. The values chosen for $\mu_{\text{frac}}$ are based on a previous study [16] which has shown that $\mu_{\text{frac}} = 0.67$ is optimal in discriminating $H \rightarrow b\bar{b}$ from background. A subsequent study regarding the factorization properties of several groomed jet algorithms [18] found that smaller values of $\mu_{\text{frac}}$ (0.20 and 0.33) are similarly effective at reducing backgrounds, and yet they remain factorizable within the soft collinear effective theory studied in that analysis. Future studies may seek to evaluate the impact of variations of $y_{\text{cut}}$ as well as $\mu_{\text{frac}}$, in particular to assess the extent to which mass-drop filtering may be used for three-body decays, for example from the top quark.

**Trimming:** The trimming algorithm [19] takes advantage of the fact that contamination from pile-up, multiple parton interactions (MPI), and initial-state radiation (ISR) in the reconstructed jet is often much softer than the outgoing partons associated with the hard-scatter and their final-state radiation (FSR). The ratio of the $p_{T}$ of the constituents to that of the jet is used as a selection criterion. Completely removing the softer components from the final jet is possible as there is generally minimal spatial overlap of the soft additional radiation from pile-up, MPI, and ISR with the hard-scatter decay products. As the primary effect of pile-up, for example, is additional low-energy topo-clusters as opposed to additional energy being added to topo-clusters from hard-scatter particles, this allows a relatively simple jet energy offset correction for smaller radius jets ($R = 0.4, 0.6$) as a function of the number of primary reconstructed vertices [5].

The trimming procedure uses a $k_{T}$ algorithm to create subjets of size $R_{\text{sub}}$ from the constituents of a jet. Any subjets with $p_{T}/p_{T}^{\text{jet}} < f_{\text{cut}}$ are removed, where $p_{T}$ is the transverse momentum of the $j^{th}$ subjet, and $f_{\text{cut}}$ is a parameter of the method, which is typically a few percent. The remaining constituents form the trimmed jet. This procedure is illustrated in Figure 3. Low-mass jets (with $m_{\text{jet}} < 100$ GeV) from a light quark or gluon typically lose 30-50% of their mass, while jets containing the decay products of a boosted object will lose only a few percent of their mass,
most of which is due to the removal pileup or the UE (see, for example, Figures 22 and 25 in Section 5.3). The fraction removed increases with the number of interactions in the event [1].

Six configurations of trimmed jets are studied here, arising from combinations of $f_{\text{cut}}$ and $R_{\text{sub}}$, given in Table 1. They are based on the optimized parameters in Ref. [19] ($f_{\text{cut}} = 0.03, R_{\text{sub}} = 0.2$) and variations suggested by the authors of the algorithm. This set represents a wide range of phase space for trimming and is somewhat broader than considered in the original paper on the subject.
Pruning: The pruning algorithm [20, 21] is similar to trimming in that it removes constituents with a small relative $p_T$, but additionally utilizes a wide-angle radiation veto. The pruning procedure is invoked at each successive recombination of the jet algorithm used (either C/A or $k_t$), based on the branching at each point in the jet reconstruction, and as such does not require the reconstruction of subjets. This results in definitions of the terms “wide-angle” or “soft” that are not directly related to the original jet but rather to the proto-jets formed in the process of rebuilding the pruned jet.

Figure 4: A cartoon illustrating the pruning procedure.

The procedure is as follows:

- Run either the C/A or $k_t$ recombination jet algorithm on the constituents found by any jet finding algorithm.
- At each recombination step with constituents $j_1$ and $j_2$ (where $p_{T1}^j > p_{T2}^j$), require that $p_{T2}^j / p_{T1}^{j1+j2} > z_{\text{cut}}$ or $\Delta R_{j_1, j_2} < R_{\text{cut}} \times \frac{2m_{\text{cut}}}{p_{T1}}$.
- Merge $j_2$ with $j_1$ if the above criteria are met, otherwise, discard $j_2$ and continue with the algorithm.

The pruning procedure is illustrated in Figure 4. Six configurations, given in Table 1, based on combinations of $z_{\text{cut}}$ and $R_{\text{cut}}$ are studied here. They are not configurations that have been studied before in Refs. [20, 21] but are chosen based on discussion with the authors of the pruning algorithm [22]. This set of parameters also represents a relatively wide range of possible configurations.

2.5 HEPTopTagger

The HEPTopTagger [23] is an example of how jet grooming techniques may be used to optimize the selection of boosted objects (in this case, top quarks with a hadronically-decaying $W$ boson daughter) over a large multi-jet background. The method uses the C/A jet algorithm and a variant on the mass-drop filtering technique described in Section 2.4 in order to utilize information about the recombination history of the jet. The algorithm proceeds as follows:

- Decomposition into substructure objects: The mass-drop criterion defined in Eq. (7) is applied to a large-$R$ C/A jet, where $j_1$ and $j_2$ are the two subjets from the last stage of clustering. If the criterion is satisfied, the same prescription is followed iteratively on both $j_1$ and $j_2$ until $N_i$ subjets are left, where the subjets either have masses $m_i \leq m_{\text{cut}}$ or represent individual constituents, such
Jet finding algorithms used | Grooming algorithm | Configurations considered
---|---|---
C/A | Mass-Drop Filtering | $\mu_{trc} = 0.20, 0.33, 0.67$
anti-$k_t$ and C/A | Trimming | $f_{cut} = 0.01, 0.03, 0.05$
anti-$k_t$ and C/A | Pruning | $R_{cut} = 0.1, 0.2, 0.3$
C/A | HEPTopTagger | (see Table 2)

Table 1: Summary of the grooming configurations considered in this study. Values in boldface are optimized configurations reported in Ref. [16] and Ref. [19] for filtering and trimming, respectively.

as topo-clusters, tracks, or truth particles (i.e. no clustering history). If at any stage $m_{j1} > m_{j2}\mu_{trc}$, the mass-drop criterion and subsequent iterative de-clustering is not applied to $j_2$. The values of $m_{cut}$ and $R$ studied in this note are summarized in Table 2. $R$ values of 1.5 and 1.8, somewhat larger than used generally in mass-drop filtering, are chosen based on previous studies [23]. When the iterative process of de-clustering the jet is complete, there must exist at least three substructure objects, otherwise the jet is discarded.

- **Filtering:** Combinations of three substructure objects are filtered at a time. The constituents of the substructure objects in a given triplet are reclustered into $N_i$ subjets using the C/A algorithm with a distance parameter $R_{filt} = \min[0.3, \Delta R_{j1,j2}^2]$, where $\Delta R_{j1,j2}$ is the minimum separation between all possible pairs in the current triplet.

- **Top mass window requirement:** If the invariant mass of the four-vector determined by summing the constituents of the $N_i$ subjets is not in the range $140 \leq m_{j1} < 200$ GeV then the triplet combination is ignored. If more than one triplet satisfies the criteria, only the one with mass closest to the top quark mass, $m_{top}$, is used.

- **Reclustering of subjets:** From the $N_i$ subjets formed from the chosen top candidate triplet, a number of leading-$p_T$ subjets ($N_{subjet}$) are chosen, where $3 \leq N_{subjet} \leq N_i$. Of these chosen subjets, exactly three jets are built by applying the C/A algorithm to the constituents of the $N_{subjet}$ subjets (exclusive clustering using a distance parameter $R_{jet}$ listed in Table 2). These subjets are calibrated as described in [24].

- **$W$ boson mass requirements:** Relations listed in A1 of [23] are defined using the total invariant mass of the three subjets ($m_{123}$) and the invariant mass $m_{ij}$ formed from combinations of two of the three C/A jets ordered in $p_T$. These relations include:

  \[ R_- < \frac{m_{23}}{m_{123}} < R_+ \]  
  \[ 0.2 < \arctan \frac{m_{13}}{m_{12}} < 1.3 \]

Here, $R_{\pm} = (1 \pm f_W) \frac{m_{top}}{m_W}$, $f_W$ is a resolution variable (given in Table 2), and the quantities $m_W$ and $m_{top}$ denote the $W$ boson and top quark masses, respectively. If at least one of the criteria in A1 of [23] is met, the four-momentum addition of the three subjets is considered a candidate top quark.
The procedure is illustrated in Figure 5.

<table>
<thead>
<tr>
<th></th>
<th>Default</th>
<th>Tight</th>
<th>Loose</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.8</td>
</tr>
<tr>
<td>$m_{\text{cut}}$[GeV]</td>
<td>30</td>
<td>30</td>
<td>70</td>
<td>30</td>
</tr>
<tr>
<td>$R_{\text{jet}}$</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>$N_y$</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>$f_W$[%]</td>
<td>15</td>
<td>10</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 2: The settings used for studying the performance of the HEPTopTagger.
(a) Every object encountered in the de-clustering process is considered a ‘substructure object’ if it is of sufficiently low mass or has no clustering history.

(b) The mass-drop criterion is applied iteratively, following the highest subjet-mass line through the clustering history, resulting in $N_i$ substructure objects.

(c) For every triplet-wise combination of the substructure objects, recluster into subjets and select the $N_{\text{subjet}}$ leading-$p_T$ subjets, with $3 \leq N_{\text{subjet}} \leq N_i$ (here, $N_{\text{subjet}} = 5$).

(d) Recluster the constituents of the $N_{\text{subjet}}$ subjets into exactly three subjets to make the top candidate for this triplet-wise combination of substructure objects.

Figure 5: The HEPTopTagger procedure.
3 Data and Monte Carlo Samples

3.1 Data quality criteria and event selection

The data used in the analyses presented in this note, corresponding to \((4.7 \pm 0.2) \text{ fb}^{-1}\) of integrated luminosity, are required to have met baseline quality criteria and were taken during periods in which the detector was fully operational. The ATLAS data quality (DQ) criteria reject data with significant contamination from detector noise or issues in the read-out based upon individual assessments for each subdetector. These criteria are established separately for the barrel, endcap and forward regions, and differ depending on the trigger conditions and reconstruction of each type of physics object, such as jets, electrons, muons, etc. The primary systems of interest in these studies are the electromagnetic and hadronic calorimeters and the inner tracking detector (the latter specifically for studies of the properties of tracks associated with jets).

To reject non-collision backgrounds, events are required to contain a primary vertex consistent with the LHC beam spot, reconstructed from at least 2 tracks each with transverse momentum \(p_{\text{track}} > 400\text{ MeV}\). All jets in the event reconstructed with the anti-\(k_t\) algorithm with \(R = 0.4\) and a measured \(p_{\text{jet}}^T > 20\text{ GeV}\) are required to satisfy the “looser” requirements discussed in detail in Ref. [25]. These selections are designed to provide an efficiency to retain good quality jets of greater than 99.8% with as high a fake jet rejection as possible. In particular, this selection is very efficient at rejecting fake jets that arise due to calorimeter noise.

A three-level trigger system is used to select interesting events. The level-1 trigger is implemented in hardware and uses a subset of detector information to reduce the event rate to a design value of at most 75 kHz. This is followed by two software-based triggers, level-2 and the event filter, which together reduce the event rate to a few hundred Hz. Events used in this analysis were selected if the leading jet in the event passed a single jet trigger at the event filter stage with \(p_{\text{jet}}^T > 350\text{ GeV}\). This trigger threshold was un-prescaled for the entire 2011 data-taking period and thus represents the full integrated luminosity with negligible inefficiency.

3.2 Monte Carlo simulation

The data are compared to inclusive jet events generated by two MC simulations: PYTHIA 6.425 [2] and POWHEG-BOX 1.0 [3,4,26] (patch 4) interfaced to PYTHIA 6.425 for the parton shower, hadronization, and underlying event (UE) models. In the former case, standalone PYTHIA uses the modified-LO parton distribution function (PDF) set MRST LO* [27]. In the latter, POWHEG+PYTHIA uses the CTEQ6L1 PDF set [28]. For both cases, PYTHIA is tuned with the corresponding AUET2B tune [29,30]. The comparison between PYTHIA and POWHEG+PYTHIA represents an important juxtaposition, at least at the matrix element (ME) level, between an LO (PYTHIA) and an NLO (POWHEG) ME generator. After simulation of the parton shower and hadronization, events are passed through the full Geant4 [31] detector simulation [32]. Following this, the same trigger, event, quality, jet, and track selection criteria are applied to the MC simulation as are applied to the data.

Several samples of events containing boosted hadronic particle decays are used for direct comparisons of the performance of the various reconstruction and jet substructure techniques. For two-prong decays, a sample of hadronically decaying Z bosons was generated using the HERWIG 6.5.10 [33] event generator interfaced with JIMMY 4.3.1 [34] for MPI. In order to test the performance of techniques designed for three-prong decays, \(t\bar{t}\) events from an additional heavy gauge boson (\(Z'\) with \(m_{Z'} = 1.6\) TeV) were generated using the same PYTHIA 6.425 tune as stated above. This model provides a relatively narrow \(t\bar{t}\) resonance and very high \(p_T\) top quarks.

Pile-up is simulated by overlaying additional soft \(pp\) collisions, or minimum bias events, which are generated with PYTHIA 6.425 using the ATLAS MC11 AUET2B tune [30] and the CTEQ6L1 PDF
set. The minimum bias events are overlaid onto the hard scattering events according to the measured distribution of the average number ($\mu$) of $pp$ interactions per bunch crossing. The proton bunches were organized in four trains of 36 bunches with a 50 ns spacing between the bunches. Therefore, the simulation also contains effects from out-of-time pile-up, i.e. contributions from the collision of neighboring bunches to that where the event of interest occurred. Simulated events are reweighted such that the MC distribution of $\langle \mu \rangle$ agrees with the data, as measured by the luminosity detectors in ATLAS [35].

4 Calibration of Large-$R$ and Groomed Jets

This section describes the procedure used for determining the energy and mass scale calibrations for large-$R$ and groomed jets using Monte Carlo simulation. In addition, the impact of detector simulation and the partonic origin of a jet on the mass calibration and response are studied.

4.1 Monte Carlo based calibration

The MC calibration scheme starts from the measured calorimeter energy at the electromagnetic (EM) energy scale [36–44], which correctly measures the energy deposited by electromagnetic showers. A local cluster weighting (LCW) calibration method first clusters together topologically connected calorimeter cells and classifies these clusters as either electromagnetic or hadronic. Based on this classification energy corrections are derived from single pion MC simulations. Dedicated hadronic corrections are derived for the effects of non-compensation, signal losses due to noise suppression threshold effects, and energy lost in non-instrumented regions. The results shown here use LCW clusters as input to the jet algorithm.

The final jet energy calibration is derived as a correction relating the calorimeter’s response to the true jet energy. It can be applied to EM scale jets, with the resulting calibrated jets referred to as EM+JES, or to LCW calibrated jets, with the resulting jets referred to as LCW+JES jets. More details regarding the evaluation and validation of this approach for standard anti-$k_t$, $R = 0.4$, 0.6 jets can be found in Ref. [5].

The JES correction is derived from a PYTHIA MC sample including pile-up events, as in the standard JES determination procedure [5]. For standard jet algorithms, the dependence of the jet response on the number of primary vertices ($N_{PV}$) and the average number of interactions ($\langle \mu \rangle$) is removed by applying a pile-up offset correction on the EM or LCW scale before applying the JES correction. However, for large-$R$ jets and jets with the various grooming algorithms applied, no explicit pile-up correction is applied.

4.2 Jet mass scale determination

Since one of the primary goals of the use of large-$R$ and groomed jet algorithms is to reconstruct the masses of jets accurately and precisely, a last step is added to the calibration procedure of large-$R$ jets wherein the mass of the jet is calibrated based on multi-jets MC simulation. Explicit jet mass calibration is important for using the individual invariant jet mass in physics analyses since it is particularly susceptible to soft, wide-angle contributions that do not otherwise significantly impact the jet energy scale. The procedure measures the jet mass response for jets built from LCW clusters after the standard JES calibration. The mass response is determined from the mean of a Gaussian fit to the core of the distribution of the reconstructed jet mass divided by the corresponding truth jet mass.

Figure 6 shows the jet mass response in a number of jet energy bins as a function of $\eta$, before and after calibration to the true jet mass for anti-$k_t$, $R = 1.0$ jets. One can see from this figure that even very high $p_T$ jets near the central part of the detector can have a mean mass scale (or JMS) up to 20% different than that of the particle level true jet mass. In particular, the reconstructed mass is, on average, greater
than that of the particle-level jet due in part to noise and pile-up in the detectors. Additionally, the finite resolution of the detector has a differential impact on the mass response as a function of $\eta$. Following the jet mass calibration, performed also as function of $\eta$, a uniform mass response can be restored to within 3\% across the full energy and $\eta$ range.

4.3 In-situ validation based on track jets

In order to validate the jet mass measurement made by the calorimeter, calorimeter-based jets are compared to track jets reconstructed from charged particle tracks. Track jets have a very different set of systematic uncertainties and allow for a reliable determination of the relative systematic uncertainties associated with the calorimeter-based measurement. Performance studies [45] have shown that there is excellent agreement between the measured positions of clusters and tracks in data, indicating no systematic misalignment between the calorimeter and the inner detector.

The use of track jets reduces or eliminates the impact of additional $pp$ collisions by requiring the jet inputs (tracks) to come from the hard-scattering vertex. It also provides a reference with which to compare the calorimeter jet measurement. The inner detector and calorimeter have largely uncorrelated systematic effects, and so comparison of variables such as jet mass and energy between the two systems allows for a separation of physics and detector effects. It is therefore possible to validate the JES and JMS and also to estimate the pile-up energy contribution to jets directly. This approach was used extensively in the measurement of the jet mass and substructure properties of jets in the 2010 data [17] where pile-up was significantly less important and the statistical reach of the measurement was smaller than with the full integrated luminosity of 4.7 $fb^{-1}$ for the 2011 dataset.

The method to determine the relative uncertainty uses the ratio of the calorimeter $p_T^\text{jet}$ ($m^\text{jet}$) to the track jet transverse momentum, $p_T^\text{track jet}$($m^\text{track jet}$). The ratios are defined explicitly as

$$r^p_{\text{track jet}} = \frac{p_T^\text{jet}}{p_T^\text{track jet}}, \quad r^m_{\text{track jet}} = \frac{m^\text{jet}}{m^\text{track jet}},$$

(11)

where the matching between calorimeter and track jets is performed using a matching radius of $\Delta R < 0.3$. These ratios are expected to be well described by the detector simulation in the case that detector effects are well modeled. That is to say, even if some underlying physics process was unaccounted for by the simulation, as long as this process affects both the track jet and calorimeter jet $p_T$ or masses in a similar way, then the ratio of data to simulation should be relatively unaffected.

Double ratios of $r^m_{\text{track jet}}$ and $r^p_{\text{track jet}}$ are constructed in order to evaluate this agreement. These double ratios, $R^p_{r \text{ track jet}}$ and $R^m_{r \text{ track jet}}$, are defined as:

$$R^p_{r \text{ track jet}} = \frac{r^p_{\text{track jet, data}}}{r^p_{\text{track jet, MC}}}, \quad R^m_{r \text{ track jet}} = \frac{r^m_{\text{track jet, data}}}{r^m_{\text{track jet, MC}}},$$

(12)

The dependence of $R^p_{r \text{ track jet}}$ and $R^m_{r \text{ track jet}}$ on $p_T^\text{jet}$ and $m^\text{jet}$ provides a test for the deviation between data and simulation, thus allowing for an estimation of the calibration uncertainty.

Figure 7 shows the distribution of $r^m_{\text{track jet}}$ for four jet algorithms and trimming configurations, and for jets in the range $500 \leq p_T^\text{jet} < 600$ GeV in the central calorimeter region, $|\eta| < 0.8$. This $p_T^\text{jet}$ range is fairly typical and was chosen for illustrative purposes because of its relevance to boosted vector bosons and boosted tops, as the decay products of both are expected to be fully merged into a large-$R$ jet. The peak position near $r^m_{\text{track jet}} \approx 2$ and the shape of the distribution are both generally well-described. Both the ungroomed and the trimmed anti-$k_t$, $R = 1.0$ distributions show some discrepancies at very low $p_T^\text{track jet}$.
where the description of very soft radiation and hadronization is important, and at high values of \( r_{\text{track jet}} \) above \( r_{\text{track jet}} \gtrsim 4 \). The differences are of order 20%. However, these spectra are used primarily to test the overall scale, so that the important comparison is of the mean value of the distributions. These are quite well described, as discussed in Section 4.4.

### 4.4 Evaluation of jet mass scale systematics

The relative systematic uncertainty on the jet kinematics is first estimated for each MC generator sample as the weighted average absolute deviation of the double ratio, \( R_{m_{\text{track jet}}} \), from unity. Measurements of \( R_{m_{\text{track jet}}} \) are performed in exclusive \( p_T \) and \( \eta \) ranges. The statistical uncertainty is used as the weight in this case. The final relative uncertainty is then determined by the maximum of the weighted average relative uncertainties considered. Comparisons are made using PYTHIA and POWHEG +PYTHIA.

Figure 8 presents the distributions of both \( r_{\text{track jet}} \) and the double ratio with respect to MC, \( R_{m_{\text{track jet}}} \), for the same four jet algorithms and grooming configurations as shown in Figure 7. In the peak of the jet mass distribution, logarithmic soft terms dominate [46] and lower \( p_T \) particles constitute a large fraction of the calorimeter jet mass. These particles are often bent by the magnetic field or not reconstructed as charged tracks and thus contribute less to the track jet mass, resulting in the shape observed in the \( r_{\text{track jet}} \) distribution in this region. Higher mass jets tend to be composed of higher \( p_T \) particles that contribute more similarly to the calorimeter and track-based mass reconstruction, resulting in a flatter and fairly stable \( r_{\text{track jet}} \) ratio. This flat \( r_{\text{track jet}} \) distribution is present across the mass range for both filtered and trimmed jet masses, as both these algorithms are designed to remove softer particles as compared to the original jet \( p_T \). Although there is a difference in the phase space of emissions probed at low mass and high mass, the calorimeter response relative to the tracker response is well modeled by both PYTHIA and POWHEG MCs.

The weighted average deviation of \( R_{m_{\text{track jet}}} \) from unity ranges from approximately 2% to 4% for the set of jet algorithms and grooming configurations tested for jets in the range \( 500 \leq p_T^{\text{jet}} < 600 \) GeV and in the central calorimeter, \( |\eta| < 0.8 \). The results are fairly stable for the slightly less central \( \eta \) range
Figure 7: $r_{\text{track jet}}$ distributions for (a) anti-$k_t$, $R = 1.0$, (b) anti-$k_t$, $R = 1.0$ trimmed jets ($f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.03$), (c) C/A, $R = 1.2$ jets, and (d) C/A, $R = 1.2$ filtered jets ($\mu_{\text{frac}} = 0.67$) in the range $500 \leq p_T^{\text{jet}} < 600$ GeV and in the central calorimeter, $|\eta| < 0.8$. The ratios between data and MC distributions are shown in the lower section of each figure.
0.8 ≤ |η| < 1.2.

Figure 9 presents the full set of jet mass scale systematic uncertainties for various jet algorithms estimated using the calorimeter-to-track jet double ratios. The total relative uncertainty includes the 3% uncertainty of the precision of the jet mass scale calibration (see 6(b), for example) as well as the uncertainty on the track measurements themselves. The latter uncertainty takes into account the knowledge of tracking inefficiencies and their impact on the $p_T^{\text{jet}}$ and $m^{\text{jet}}$ measurements using track jets. Each of these two additional components is assumed to be uncorrelated and added in quadrature with the uncertainty determined solely from the calorimeter-to-track jet double ratios.

The impact of the tracking efficiency systematic uncertainty on $r^{m}$_{\text{track jet}} is evaluated by randomly rejecting tracks used to construct track jets according to the efficiency uncertainty. This is evaluated as a function of $\eta$ and $m^{\text{jet}}$ for various $p_T^{\text{jet}}$ ranges. Typically, this results in a 2-3% shift in the measured track jet kinematics (both $p_T$ and mass) and thus a roughly 1% shift on the resulting total uncertainty.

The total systematic uncertainty on the jet mass scale is fairly stable near 4–5% for all jet algorithms up to $p_T^{\text{jet}} ≈ 800$ GeV. At low $p_T^{\text{jet}}$, in the range $200 ≤ p_T^{\text{jet}} < 300$ GeV, the average uncertainty for some jet algorithms rises to approximately 5–7%. The estimated uncertainty is similar for both the ungroomed and the trimmed or filtered jets, except for trimmed anti-$k_t$ jets (see Figure 9(b)) for which the uncertainty in the range $900 ≤ p_T < 1000$ GeV is approximately 8%.
Figure 8: Mean values of $r_{\text{track jet}}^m$ as a function of jet mass for (a) anti-$k_t$, $R = 1.0$, (b) anti-$k_t$, $R = 1.0$ trimmed jets ($f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.03$), (c) C/A, $R = 1.2$ jets, and (d) C/A, $R = 1.2$ filtered jets ($\mu_{\text{frac}} = 0.67$) in the range $500 \leq p_T^{\text{jet}} < 600$ GeV and in the central calorimeter, $|\eta| < 0.8$. The mean ratios between the data and MC distributions (the double ratios $R_{r \text{track jet}}^m$) are shown in the lower section of each figure.
Figure 9: Summary of the jet mass scale (JMS) relative systematic uncertainties as a function of $p_T^{\text{jet}}$. These uncertainties are determined from track jet double ratios. (a) anti-$k_t$, $R = 1.0$ without trimming and (b) anti-$k_t$, $R = 1.0$ with trimming ($f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.03$). (c) C/A, $R = 1.2$ without filtering and (d) C/A, $R = 1.2$ with filtering. These estimates include a 3% relative non-closure uncertainty on the MC-based mass scale calibration factors discussed above, as well as systematic uncertainties due to the uncertainty on the impact of the tracking efficiency and fake rate on the track measurements.
5 The Impact of Grooming on Jet Properties and Signal Discrimination

Using the inclusive jet selection described in Section 3, various generators are compared to the full 2011 dataset for large-\(R\) jets before and after grooming. Furthermore, MC simulations are used to compare boosted jets from various physics processes of interest to the multi-jet background, thereby demonstrating their utility and the effect of grooming on substructure observables.

5.1 Inclusive groomed jet resolution in simulation

The fractional resolution is defined as the width of a Gaussian fit to the central part of the resolution distribution, which is generated by taking the difference between the truth jet mass and the reconstructed jet mass, divided by the same truth jet mass. Here, the truth jet is the simulated particle shower that has been groomed according to the same grooming algorithms used after jet reconstruction. Large-\(R\) truth and reconstructed jets before grooming are matched if they are within \(\Delta R < 0.7\). The matching of ungroomed jets is retained in comparing groomed versions of the truth and reconstructed jets. Note that the mass-drop/\(\text{filtering}\) method was not applied to anti-\(k_t\) jets, as discussed in Section 2.4.

<table>
<thead>
<tr>
<th>Label</th>
<th>Algorithm and Parameters</th>
<th>Label</th>
<th>Algorithm and Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trm PtF1 R30</td>
<td>Trimmed, (f_{\text{cut}} = 1%), (R_{\text{sub}} = 0.3)</td>
<td>Prn Rc10 Zc5</td>
<td>Pruned, (R_{\text{cut}} = 0.1), (z_{\text{cut}} = 5%)</td>
</tr>
<tr>
<td>Trm PtF3 R30(^\dagger)</td>
<td>Trimmed, (f_{\text{cut}} = 3%), (R_{\text{sub}} = 0.3)</td>
<td>Prn Rc10 Zc10</td>
<td>Pruned, (R_{\text{cut}} = 0.1), (z_{\text{cut}} = 10%)</td>
</tr>
<tr>
<td>Trm PtF5 R30(^\ddagger)</td>
<td>Trimmed, (f_{\text{cut}} = 5%), (R_{\text{sub}} = 0.3)</td>
<td>Prn Rc20 Zc5</td>
<td>Pruned, (R_{\text{cut}} = 0.2), (z_{\text{cut}} = 5%)</td>
</tr>
<tr>
<td>Trm PtF1 R20</td>
<td>Trimmed, (f_{\text{cut}} = 1%), (R_{\text{sub}} = 0.2)</td>
<td>Prn Rc20 Zc10</td>
<td>Pruned, (R_{\text{cut}} = 0.2), (z_{\text{cut}} = 10%)</td>
</tr>
<tr>
<td>Trm PtF3 R20</td>
<td>Trimmed, (f_{\text{cut}} = 3%), (R_{\text{sub}} = 0.2)</td>
<td>Prn Rc30 Zc5</td>
<td>Pruned, (R_{\text{cut}} = 0.3), (z_{\text{cut}} = 5%)</td>
</tr>
<tr>
<td>Trm PtF5 R20</td>
<td>Trimmed, (f_{\text{cut}} = 5%), (R_{\text{sub}} = 0.2)</td>
<td>Prn Rc30 Zc10</td>
<td>Pruned, (R_{\text{cut}} = 0.3), (z_{\text{cut}} = 10%)</td>
</tr>
<tr>
<td>MD Filt mf20</td>
<td>Mass-drop/Filtered, (\mu_{\text{frac}} = 0.2)</td>
<td>MD Filt mf67(^\S)</td>
<td>Mass-drop/Filtered, (\mu_{\text{frac}} = 0.67)</td>
</tr>
<tr>
<td>MD Filt mf33</td>
<td>Mass-drop/Filtered, (\mu_{\text{frac}} = 0.33)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Labels used in figures to represent the various configurations of the grooming algorithms. \(^\dagger\)Groomed jets have been calibrated for both anti-\(k_t\) and C/A jets. \(^\ddagger\)Groomed jets have been calibrated for anti-\(k_t\) only. \(^\S\)Groomed jets have been calibrated for C/A only.

Figure 10 shows the fractional mass resolution for an inclusive multi-jet sample with the same pile-up conditions as were observed in the data. Abbreviated versions of the groomed algorithm names used to label the figures are listed in Table 3. In general, the groomed jets have better resolution than the ungroomed large-\(R\) jets, with improvements of up to \(\sim 10\%\) (absolute) in some cases. The trimmed and pruned jet resolution improves with higher \(p_T\), where the calibrated jet collections gain \(\sim 3-5\%\) over the range \(300 \leq p_T^\text{jet} < 800\) GeV. The pruning algorithm, especially with C/A jets, produces larger tails in the resolution distribution than for the trimmed algorithm, worsening the overall fractional resolution by comparison. In the case of the mass-drop/\(\text{filtering}\) algorithm, the resolution is fairly stable over a large \(p_T\) range. However it should be noted that the efficiency is considerably lower for these jets than the other grooming algorithms (\(\sim 30\%\)) due to the strict mass-drop requirement, which is often not met for jets without boosted object substructure.

The specific performance of jets under pile-up conditions is presented in greater detail elsewhere [1], but a summary of the fractional mass resolution for jets before and after grooming in the presence of
Figure 10: Fractional mass resolution comparing the various grooming algorithms (with labels defined in Table 3) for the leading-\(p_T^{\text{jet}}\) jet in POWHEG with PYTHIA dijet simulated events. Here, “nominal” refers to jets before grooming is applied. Three ranges of the nominal jet \(p_T^{\text{jet}}\) are shown. The uncertainty on the width of the Gaussian fit is indicated by the error bars.

Various pile-up conditions is shown in Figure 11. Trimming in both anti-\(k_t\) and C/A jets reduces the dependence of the jet mass on pile-up compared to the ungroomed jet, as does the mass-drop/filtering procedure in the case of C/A jets, while pruning does not seem to have a large impact. In particular, the trimming parameters \(f_{\text{cut}} = 0.03\) and 0.05 not only outperform the looser \(f_{\text{cut}} = 0.01\) setting in events with more than 12 average interactions, they also exhibit a significantly reduced overall variation between different instantaneous luminosities.

Based on the above comparisons of mass resolution in different \(p_T^{\text{jet}}\) ranges and under varying pile-up conditions, two configurations, trimmed anti-\(k_t\) jets \((f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3)\) with \(R = 1.0\) and filtered C/A jets \((\mu_{\text{frac}} = 0.67)\) with \(R = 1.2\), were chosen for detailed comparisons between data and simulation and are presented in the next section.

5.2 Large-\(R\) jets and the effect of grooming observed in data

Previous studies conducted by ATLAS [17] and CMS [47] suggest that even complex observables of jet substructure are fairly well modeled by the MC simulations used by the LHC experiments. This section reviews the description provided by PYTHIA and POWHEG of the jet grooming techniques introduced above, and of the substructure of the un groomed and groomed jets themselves.

Figure 12 presents a comparison of the jet invariant mass for un groomed, trimmed, and filtered (for C/A, \(R = 1.2\)) jets in the range \(600 \leq p_T^{\text{jet}} < 800\) GeV and in the central calorimeter, \(|\eta| < 0.8\). A similar performance is observed in all other \(p_T\) regions in the range \(p_T > 300\) GeV. The description of both the ungroomed and trimmed anti-\(k_t\) jets with \(R = 1.0\) provided by PYTHIA is decent, but is significantly more accurate with the NLO generator POWHEG interfaced to PYTHIA for the parton showering and hadronization. PYTHIA tends to underestimate the fraction of high mass large-\(R\) anti-\(k_t\) jets, whereas POWHEG is accurate to within a few percent, even for very massive jets. The un groomed anti-\(k_t\), \(R = 1.0\) jets are poorly described at low mass due to non-perturbative effects, which increase the jet mass.
thereby reducing the overall catchment area. As discussed in Ref. [1], this has the desirable e
dramatic changes.

jet mass for both ungroomed and trimmed anti-
is not shown here). Significant variation is also observed among the configurations tested, with the large
compared to the original jet is apparent for both grooming algorithms shown (and also for pruning, which
the average over the whole 2011 data-taking period. The significant spectral shift and shape di

Figure 12 also shows that the shape of the jet mass distribution is significantly a

Figure 11 demonstrates the e

A similarly poor description of the low mass region is observed for C/A jets with R = 1.2. In this
case, however, both PYTHIA and POWHEG provide fairly good descriptions of the high mass regime of
the jet mass spectrum. This suggests that there is an angular scale dependence, and the slightly smaller
radius used for the large-R anti-\( k_t \) jets in these studies could play a role in the observed discrepancy with
PYTHIA. Figure 12 also shows that the shape of the jet mass distribution is significantly affected by
the mass-drop and filtering technique. This change is well described by both PYTHIA and POWHEG,
although the accuracy of the POWHEG prediction is again observed to be slightly better.

Figure 13 presents an overview of the shape of the jet mass spectrum for several configurations of the
jet trimming algorithm for anti-\( k_t \) jets and for C/A jets with mass-drop filtering applied. These spectra are
measured using approximately 1 fb\(^{-1}\) from the last data-taking period of 2011 where \( \langle \mu \rangle = 12 \), higher than
the average over the whole 2011 data-taking period. The significant spectral shift and shape difference
compared to the original jet is apparent for both grooming algorithms shown (and also for pruning, which
is not shown here). Significant variation is also observed among the configurations tested, with the large
\( f_{\text{cut}} \), small \( R_{\text{sub}} \) setting for trimming and the small \( \mu_{\text{frac}} \) setting for mass-drop filtering exhibiting the most
dramatic changes.

The significant change observed in the jet mass distribution is due primarily to a reduction in the
effective area of each jet\(^3\). Soft and typically wide angle jet constituents are removed from the jet,
thereby reducing the overall catchment area. As discussed in Ref. [1], this has the desirable effect of also
reducing the impact of pile-up on the jet properties.

Figure 14 demonstrates the effect of grooming by presenting the average jet area as a function of the
jet mass for both ungroomed and trimmed anti-\( k_t \), R = 1.0 jets. Prior to jet trimming, the anti-\( k_t \), R = 1.0

\(^3\)In order to measure the effective area of a jet, the so-called “ghost area” approach is used [48, 49]. Infinitesimally soft,
low-\( p_T \) particles are distributed within the event at a fixed density (one per \( \sqrt{\Delta y^2 + (\Delta \phi)^2} = 0.01 \times 0.01 \)) and are allowed to
participate in the jet clustering algorithm. The number of such particles present in the jet after full reconstruction thus defines
the area of that jet.
Figure 12: Mass of jets within the range $600 \leq p_T^{\text{jett}} < 800$ GeV and in the central calorimeter ($|\eta| < 0.08$). Shown are (a) ungroomed and (b) trimmed ($f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.3$) anti-$k_t$ jets with $R = 1.0$; and (c) ungroomed and (d) filtered ($\mu_{\text{frac}} = 0.67$) C/A jets with $R = 1.2$. The ratios between data and MC distributions are shown in the lower section of each figure.
area is very close to π. A small rise in the ungroomed jet area is observed for jets with very large mass, characterized by small additional clusters near the edge of the jet. For trimmed jets at low mass, the average jet area is reduced by a factor of 3 to 5 and continuously rises as a function of the jet mass to a maximum of approximately 1/2 of the original jet area for high mass ungroomed jets. These features are very well described by the MC simulations across the entire spectrum of jet mass, both before and after pruning and filtering. Similar observations are made with respect to lower and higher \( p_T \) components of the jet, does not significantly affect \( \sqrt{d_{12}} \), shown in Figure 15. It is also not surprising that POWHEG describes this variable better, especially at large values. For the inclusive jet distribution, a second hard splitting is not highly probable, as also evidenced by the spectrum of \( \sqrt{d_{23}} \) as shown in Figure 16 falling more steeply compared to \( \sqrt{d_{12}} \). As a result, the trimming tends to affect this splitting scale slightly more when a third leading subjet within the parent jet is modified during the trimming procedure.

\( N \)-subjettiness helps to discriminate between jets that have well-formed substructure and those that do not. Figures 17 and 18 demonstrate that the MC simulations model the distributions of \( \tau_{21} \) and \( \tau_{32} \) observed in the data within about 20%. The jets in this case are selected to have \( 600 \leq p_T^{\text{jet}} < 800 \text{ GeV} \). Both PYTHIA and POWHEG show a roughly 10-20\% shift with respect to the data. The fraction of events with a low value of \( \tau_{ij} \) is predicted to be slightly lower in data, while it is the opposite for high values of \( \tau_{ij} \). In addition, the distributions of both \( \tau_{21} \) and \( \tau_{32} \) are broadened for trimmed jets (Figures 17(b) and 18(b)). These observations suggest that using jet shape observables in ungroomed jets, and that jet mass and substructure observables (like \( \sqrt{d_{12}} \)) in groomed jets might offer superior

Figure 13: Mass of jets in the range \( 600 \leq p_T^{\text{jet}} < 800 \text{ GeV} \) and in the central calorimeter (|\( \eta \)| < 0.8). In (a) anti-\( k_t \) jets with \( R = 1.0 \) are compared before (ungroomed) and after trimming with several configurations of \( p_T \) fraction (\( f_{\text{cut}} \)) and and subjet size (\( R_{\text{sub}} \)). In (b) C/A jets with \( R = 1.2 \) are compared before and after filtering with three different values of mass-drop fraction (\( \mu_{\text{frag}} \)).
Figure 14: Average area of jets within the range $500 \leq p_T^{\text{jet}} < 600$ GeV as a function of jet mass. Shown are (a) ungroomed and (b) trimmed ($f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.3$) anti-$k_t$ jets with $R = 1.0$. The ratios between data and MC distributions are shown in the lower section of each figure.

Figure 15: Splitting scale $\sqrt{d_{12}}$ of jets within the range $600 \leq p_T^{\text{jet}} < 800$ GeV. Shown are (a) ungroomed and (b) trimmed ($f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.3$) anti-$k_t$ jets with $R = 1.0$. The ratios between data and MC distributions are shown in the lower section of each figure.
discrimination against backgrounds.

For $\tau_{32}$, the MC distributions slightly underestimate the fraction of jets with $0.3 < \tau_{32} < 0.7$, which is the signal range for boosted top quark candidates. Since these observables are intended to be used as discriminants between boosted object signal events and the inclusive jet background, such differences are important for the resulting estimation of signal efficiency compared to background rejection. However, the variations observed in the distribution of each observable translate into much smaller differences in efficiency and rejection.

The modeling of the background with respect to massive boosted objects can be tested by evaluating, for example, the evolution of the mean of substructure variables as a function of jet mass. This is shown for $\tau_{32}$ in Figure 19 for the same $600 \leq p_T^{\text{jet}} < 800$ GeV range as used above. Three important observations can be made. Values of $\tau_{32}$ are slightly lower in data than those predicted by the MC simulations, and the trimmed values are lower than for ungroomed jets. Furthermore, $\tau_{32}$ is a slowly varying function of the jet mass for both ungroomed and trimmed jets. This variation is slightly reduced for trimmed jets.

### 5.3 Comparison of jet performance on the signal before and after grooming

Comparisons between jets containing “signal-like” boosted objects and a light quark or gluon jet background are presented here. Boosted objects are divided into two categories depending on the event topology: “two-pronged”, such as hadronically decaying $W$ or $Z$ bosons, and “three-pronged”, such as the top decaying into a $b$-jet and a hadronically decaying $W$. Performance measures are shown for a sample of $Z \rightarrow q\bar{q}$ and a sample of top quarks (from $Z' \rightarrow t\bar{t}$). Along with the event and object selection listed in Section 3.1, the large-$R$ ungroomed leading-$p_T$ jet is required to be within $|\eta| < 2.0$. For the signal distributions, a $\Delta R < 1.5$ match between the four-vector of the hadronically-decaying boosted object in the truth record and the reconstructed ungroomed leading-$p_T$ jet is made to ensure the non-contamination of light quark or gluon jets (or leptonically-decaying tops in the $Z'$ sample).
Figure 17: $N$-subjettiness $\tau_{21}$ of jets within the range $600 \leq p_T^{\text{jet}} < 800$ GeV. Shown are (a) ungroomed and (b) trimmed ($f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.3$) anti-$k_t$ jets with $R = 1.0$. The ratios between data and MC distributions are shown in the lower section of each figure.

Figure 18: $N$-subjettiness $\tau_{32}$ of jets within the range $600 \leq p_T^{\text{jet}} < 800$ GeV. Shown are (a) ungroomed and (b) trimmed ($f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.3$) anti-$k_t$ jets with $R = 1.0$. The ratios between data and MC distributions are shown in the lower section of each figure.
Figure 19: Mean $\tau_{32}$ as a function of jet mass for (a) ungroomed and (b) trimmed ($f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3$) anti-$k_T$ jets with $R = 1.0$ in the range $600 \leq p_T^{\text{jet}} < 800$ GeV.

Figure 20: Fractional mass resolution of the leading-$p_T$ jet in $Z \rightarrow q\bar{q}$ simulated events comparing the various grooming algorithms. Here, “nominal” refers to jets before grooming is applied. Three ranges of the nominal jet $p_T^{\text{jet}}$ are shown. The uncertainty on the width of the Gaussian fit is indicated by the error bars.

Figures 20 and 21 show the signal fractional mass resolution for the two-pronged and three-pronged cases, respectively. As explained in section 5.1, the mass-drop/filtering algorithm is shown only for the two-pronged signal with C/A jets. In the two-pronged case, as for the multi-jet background shown in Figure 10, the C/A mass-drop/filtering algorithm performs the best, but with a signal reconstruction efficiency of $\sim 45\%$ in $Z \rightarrow q\bar{q}$ events for $\mu_{\text{frac}} = 0.67$. In both two-pronged and three-pronged configu-
are compared in the range $600 \leq p_T < 800$ GeV. As seen in Figures 22–24 showing distributions for the two-pronged decay case, and in Figures 25–28 showing comparisons for the three-pronged decay case, better signal to background discrimination is made after grooming. In these figures, the ungroomed distributions are normalized to unit area, while the groomed distributions have the $e$-subjettiness variable $\tau_{21}$ selected on the trimmed jets, where the $N$-subjettiness variables $\tau_{21}$ is observed to have a much better signal to background discrimination with anti-$k_t$ trimmed jets than for C/A mass-drop/filtered jets, where the discrimination is not improved after applying the mass-drop/filtering criteria.

The three-pronged hadronic top jet mass distributions from $Z' \rightarrow t\bar{t}$ events are shown in Figure 25, where the signal peak remains fairly stable between groomed and ungroomed jets, especially with anti-$k_t$ trimming. The small excess of signal events below 50 GeV is the result of one of the two quarks from the decay of the $Z$ boson falling outside the large-$R$ jet radius, thus leaving only one quark reconstructed as the jet and making it indistinguishable from the background.

Figure 21: Fractional mass resolution of the leading-$p_T$ jet in $Z' \rightarrow t\bar{t}$ ($m_{Z'} = 1.6$ TeV) simulated events comparing the various grooming algorithms. Here, “nominal” refers to the ungroomed jet collection. Three ranges of the ungroomed jet $p_T^{\text{jet}}$ are shown. The uncertainty on the width of the Gaussian fit is indicated by the error bars.
jets. Again the resolution of the signal mass peak improves after grooming, where the \( W \)-mass peak can also be seen after trimming is applied. The enhancement of the \( W \)-mass peak is especially present in jets with lower \( p_T^{\text{jet}} \), as the jet from the \( b \) quark decay falls outside the radius of the large-\( R \) jet. Figures 26 and 27 show the variables \( \sqrt{d_{12}} \) and \( \sqrt{d_{23}} \), respectively. Similarly as in the two-pronged case, signal discrimination with the splitting scales is enhanced after jet trimming. Figure 28 shows the \( \tau_{32} \) distribution before and after trimming. Here, trimming of anti-\( k_t \) and C/A jets results in similar signal/background discrimination.

Figure 22: Leading-\( p_T \) jet mass comparing \( Z \to q\bar{q} \) signal to POWHEG dijets background for jets in the range \( 600 \leq p_T^{\text{jet}} < 800 \) GeV. The dotted lines show the ungroomed jet distributions, while the solid lines show the trimmed and mass-drop/filtered jets for (a) and (b), respectively. The trimmed parameters are \( f_{\text{cut}} = 0.05 \) and \( R_{\text{sub}} = 0.3 \) and the mass-drop/filtering parameter is \( \mu_{\text{frac}} = 0.67 \). The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity. After trimming or mass-drop/filtering, the mass peak corresponding to the \( Z \) is clearly seen at the correct mass.

5.4 Using \( N \)-subjettiness as a top-tagger

One of the primary applications of \( N \)-subjettiness is as a discriminating variable in searches for highly boosted top quarks [14]. A common method of comparing the performance of such discriminating variables or tagging algorithms is to compare the rate at which light quark or gluon jets are selected (the “mis-tag” rate) to the efficiency for retaining jets containing the hadronic particle decay of interest [50, 51].

In order to understand the utility of the \( \tau_{32} \) selections and the potential impact of jet grooming, trimmed anti-\( k_t \), \( R = 1.0 \) jets are compared to their ungroomed counterparts in a boosted top selection using two jet momentum ranges. In each case, the POWHEG inclusive jet sample is used to simulate the multi-jet background, while a PYTHIA \( Z' \) model which decays to two top quarks and has a mass \( m_{Z'} = 1.6 \) TeV is used to simulate boosted hadronic top quark production. The signal mass range is defined as that which contains a large fraction of the boosted top signal. For ungroomed jets this fraction is set to 90%, and the mass range that satisfies this requirement is \( 100 \leq m^{\text{jet}} < 250 \) GeV. A slightly
Figure 23: Leading-$p_T$ jet splitting scale $\sqrt{d_{12}}$ comparing $Z \rightarrow q\bar{q}$ signal to POWHEG dijets background for jets in the range $600 \leq p_T^{\text{jet}} < 800$ GeV. The dotted lines show the ungroomed jet distributions, while the solid lines show the trimmed and mass-drop/filtered jets for (a) and (b), respectively. The trimmed parameters are $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ and the mass-drop/filtering parameter is $\mu_{\text{tac}} = 0.67$. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.

Figure 24: Leading-$p_T$ jet $N$-subjettiness $\tau_{21}$ comparing $Z \rightarrow q\bar{q}$ signal to POWHEG multi-jet background for jets in the range $600 \leq p_T^{\text{jet}} < 800$ GeV. The dotted lines show the ungroomed jet distributions, while the solid lines show the trimmed and mass-drop/filtered jets for (a) and (b), respectively. The trimmed parameters are $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ and the mass-drop/filtering parameter is $\mu_{\text{tac}} = 0.67$. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.
jets. The groomed distributions are normalized with respect to the ungroomed distributions, which are leading-\(\leq\) multi-jet background for jets in the range 600 \(< p_T^{\text{jet}} \leq 800\) GeV. The dotted lines show the ungroomed leading-\(p_T\) jet distribution, while the solid lines show the corresponding trimmed \((f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3)\) jets. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.

Figure 25: Leading-\(p_T\) jet mass comparing \(Z' \rightarrow t\bar{t} (m_{Z'} = 1.6\) TeV\) signal to POWHEG multi-jet background for jets in the range 600 \(< p_T^{\text{jet}} < 800\) GeV. The dotted lines show the ungroomed leading-\(p_T\) jet distribution, while the solid lines show the corresponding trimmed \((f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3)\) jets. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.

Figure 26: Leading-\(p_T\) jet splitting scale \(\sqrt{d_{12}}\) comparing \(Z' \rightarrow t\bar{t} (m_{Z'} = 1.6\) TeV\) signal to POWHEG multi-jet background for jets in the range 600 \(< p_T^{\text{jet}} < 800\) GeV. The dotted lines show the ungroomed leading-\(p_T\) jet distribution, while the solid lines show the corresponding trimmed \((f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3)\) jets. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.
themselves normalized to unity. The groomed distributions are normalized with respect to the ungroomed distributions, which are leading-multiplet background for jets in the range $600 \leq p_T < 800$ GeV. The dotted lines show the ungroomed leading-$p_T$ jet distribution, while the solid lines show the corresponding trimmed ($f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3$) jets. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.

Figure 28: Leading-$p_T$ jet $N$-subjettiness $\tau_{32}$ comparing $Z' \rightarrow t\bar{t} (m_{Z'} = 1.6$ TeV) signal to POWHEG multi-jet background for jets in the range $600 \leq p_T < 800$ GeV. The dotted lines show the ungroomed leading-$p_T$ jet distribution, while the solid lines show the corresponding trimmed ($f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.3$) jets. The groomed distributions are normalized with respect to the ungroomed distributions, which are themselves normalized to unity.
lower signal fraction of 80% for the same mass range is required for groomed jets; this is motivated
by the tendency for trimmed jets to populate an additional small peak around the \( W \) mass, as shown in
Figure 25.

The mis-tag rate is defined as the fraction of the POWHEG multi-jet sample that remains in the mass
window after a simple selection based on \( \tau_{32} \). The signal “top-jet” efficiency is defined as the fraction of
top-jets selected in the \( Z' \) sample with the inverse \( \tau_{32} \) selection. Figure 29 uses this definition to depict the
performance of the \( N \)-subjettiness tagger for jets with \( 600 \leq p_T^{\text{jet}} < 800 \) GeV and \( 800 \leq p_T^{\text{jet}} < 1000 \) GeV.
In both cases, the reduction in high-invariant-mass jets due to trimming results in a relative gain of several
percent in the mis-tag rate for a fixed top-jet signal efficiency. Moreover, in the case of very high \( p_T^{\text{jet}} \), as
in Figure 29(b), the slightly more aggressive trimming configuration results in a slight performance gain
as well.

![Efficiency vs. mis-tag curves using \( \tau_{32} \) as a top tagger for (a) 600 \( \leq p_T^{\text{jet}} < 800 \) GeV and (b) 800 \( \leq p_T^{\text{jet}} < 1000 \) GeV with masses in the range 100 \( \leq m^{\text{jet}} < 250 \) GeV.](image)

Figure 29: Efficiency vs. mis-tag curves using \( \tau_{32} \) as a top tagger for (a) 600 \( \leq p_T^{\text{jet}} < 800 \) GeV and (b) 800 \( \leq p_T^{\text{jet}} < 1000 \) GeV with masses in the range 100 \( \leq m^{\text{jet}} < 250 \) GeV.

### 5.5 HEPTopTagger efficiency results

The efficiency of the HEPTopTagger is measured as a function of the transverse momentum of the generated
top quark, and is the product of the large-\( R \) jet finding efficiency and the efficiency to tag the jet
correctly: \( \varepsilon(\text{total}) = \varepsilon(\text{large-}\ R \ \text{jet}) \cdot \varepsilon(\text{tag}) \).

Figure 30 shows \( \varepsilon(\text{total}) \) for four different filtering configurations of the HEPTopTagger as a function
of the generator level true-top \( p_T \) on the \( t\bar{t} \) MC sample. The efficiency for the default settings is 20% at
200 GeV and reaches a plateau of 40% at 400 GeV. Below 400 GeV the efficiency can be improved by
5% by using a larger radius parameter of \( R = 1.8 \). The maximal efficiency for the tight filtering settings
is 30%.

The fake efficiency evaluated using the PYTHIA multi-jet sample is shown in Figure 31(a). The \( p_T \)
of the leading anti-\( k_T \) jet with \( R = 0.4 \) has been chosen to compute the efficiency as it provides a measure
for the energy available in the event and is easily comparable between different tagging approaches. The
fake efficiency shows a sharp turn-on around 200 GeV with efficiencies below 0.5% below and a plateau of 4% (2.5%) for the default and loose (tight) filtering settings.

Figure 30: Per-top tagging efficiency as a function of the generator level top quark $p_T$ for different filtering settings of the HEPTopTagger, evaluated using the semi-leptonic $t\bar{t}$ MC sample.

Figure 31: Per-event fake efficiency as a function of the leading anti-$k_T$, $R = 0.4$ jet $p_T$ in the event for different filtering settings, measured using (a) the dijet MC sample and (b) a comparison of the per-jet fake-efficiency between the dijet and $W \rightarrow q\bar{q}$ MC samples using the default filtering.
Figure 31(b) shows the fake tagging efficiency as a function of the large-$R$ jet $p_T$ in a multi-jet background sample and for events with a hadronically-decaying $W$ boson. The fake efficiency rises sharply at $300$ GeV and reaches a plateau of $2.5\%$ at a large-$R$ jet $p_T$ of $400$ GeV.

The efficiency of the HEPTopTagger to select jets from boosted top quarks can be increased by varying the filtering parameters. Since this will also increase the fake efficiency, the optimal working-point will depend on the analysis scenario. The HEPTopTagger is shown to be robust against fakes arising from both a multi-jets background and $W$ plus jets events.

6 Effect of Grooming in $t\bar{t}$ Events

6.1 Semi-leptonic $t\bar{t}$ Selection

A selection of $t\bar{t}\rightarrow (Wb)(Wb)\rightarrow (\mu\nu b)(q\bar{q}b)$ events is used to demonstrate in data the effect of grooming on large-$R$ jets with substructure. The semi-leptonic $t\bar{t}$ decay-mode in which one $W$ decays into a neutrino and a muon is chosen in order to tag the $t\bar{t}$ event and reduce the overwhelming multi-jet background so that the top quark signal is visible. The following event-level and object selection criteria are applied on data and simulation:

**Event-level Trigger and Data Quality Selection** The standard data quality and vertex requirements described in Section 3.1 are applied. Events are selected if they satisfy the single muon event-filter trigger with muon $p_T > 18$ GeV.

**Event-level Jet Selection** Events are required to have at least four anti-$k_t$ jets with $R = 0.4$ having $p_T > 25$ GeV and jet-vertex fraction $|JVF| > 0.75$. The jet vertex fraction is a discriminant which contains information regarding the probability that a jet originated from a particular vertex in an event [52].

**Lepton Selection** Muons must be reconstructed in both the inner detector and muon spectrometer with $p_T > 20$ GeV and $|\eta| < 2.5$. The opening angle between the muon and the jet must be greater than $R = 0.4$, where only jets with $p_T > 25$ GeV and jet vertex fraction $|JVF| > 0.75$ are considered. Events with one or more electrons passing standard quality criteria are rejected.

**Event-level neutrino and leptonic-$W$ requirement** To tag events with a leptonically decaying $W$ from a top quark decay, events are required to have missing energy $E_T^{miss} > 20$ GeV. Additionally, the scalar sum of $E_T^{miss}$ and transverse mass of the leptonic $W$ boson must have $E_T^{miss} + m_W^T > 60$ GeV, where $m_W^T$ is the combination of the muon $p_T$ and $E_T^{miss}$ in the event.

To simulate the signal, a $t\bar{t}$ sample generated with MC@NLO v4.01 [53] and interfaced to HERWIG v6.520 [54] and JIMMY v4.31 [34] was used. Various backgrounds to $t\bar{t}$ production were also simulated, including $W \rightarrow \mu\nu$ and $Z \rightarrow \mu\mu$ produced in association with jets (modeled with the ALPGEN v2.13 [55] generator interfaced to HERWIG), single top (modeled with MC@NLO, interfaced with JIMMY for the parton showering in the s-channel and $Wt$ channel, and ACERMC v3.8 [56] in the t-channel), dibosons ($WW$, $WZ$, and $ZZ$, generated with HERWIG) and dijets (PYTHIA). However, due to the selection requirements, $W$ plus jets constitutes the largest background, with smaller contributions from $Z$ plus jets and single top.

6.2 Observed effects of trimming in $t\bar{t}$ events

After the above selection, Figure 32 shows the leading-$p_T$ jet mass of anti-$k_t$ jets with $R = 1.0$ having $p_T > 350$ GeV before and after trimming ($f_{cut} = 0.05$, $R_{sub} = 0.3$). The data and simulation agree
within statistical uncertainty. The $W \rightarrow \mu \nu$ events produced in association with jets form the largest background. Since large-$R$ jets in $W$ events are formed from one or more random light-quark or gluon jets, trimming causes the mass spectrum to fall more steeply and the peak of the distribution to lie at smaller masses, in a similar way as was shown for the multi-jet background in Figure 13(a). However, trimming does not alter the top mass spectrum drastically, and any signal loss near the top mass peak is due to events in which the top quark is not boosted enough to have all three hadronic decay products fall within $R = 1.0$. Figures 33 and 34 show the distributions for $\sqrt{d_{12}}$ and $\sqrt{d_{23}}$, respectively. Again, the top distribution remains relatively unaffected by trimming, while the $W$ plus jets background is pushed lower; however, the effect is smaller for these variables. Figures 35 and 36 show distributions for $\tau_{21}$ and $\tau_{32}$, respectively. Here, there is less discrimination between the $W$ plus jets background and the top quark signal. This is due to the fact that multiple jets in $W$ events can be reconstructed as a $R = 1.0$ anti-$k_T$ jet and mimic the “subjettiness” signature of the large-$R$ jet containing the hadronic decay products of the top quark.

In order to look at events with a reduced $W$ plus jets background, a b-tagging requirement on at least one anti-$k_T$ jet in the event with $R = 0.4$ is applied in addition to the selection in Section 6.1. Figure 37 shows the effect of trimming on the mass of the leading-$p_T$ anti-$k_T$ jet with $R = 1.0$ in a sample of nearly pure $t\bar{t}$ events. Trimming clearly enhances the mass resolution compared to the ungroomed case, with a peak at low mass corresponding to large-$R$ jets containing one quark or gluon (likely from a fully-leptonic $t\bar{t}$ event) and a peak around the top mass, where all three top decay products from the $b$-jet and hadronic-$W$ daughters fall inside the large-$R$ jet radius.

![Figure 32: Leading-$p_T$ jet mass of anti-$k_T$ jets with $R = 1.0$ for (a) unngroomed jets and (b) trimmed jets ($f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$).](image)

6.3 Kinematics for HEPTopTagger candidates

Basic kinematic distributions for large-$R$ jets before and after applying the HEPTopTagger algorithm are shown in this section. Top candidates in data and simulation are compared after selecting events according to the criteria listed in Section 6.1. Figure 38 shows the jet mass distributions of C/A jets with $R = 1.5$ and $R = 1.8$ before applying the HEPTopTagger, for jets having $p_T > 200$ GeV. The selection consists of around 50% $t\bar{t}$-pair events, with other contributions coming mainly from $W$ plus jets and multi-jet events. A larger contribution from multi-jet events compared to what was observed in the previous section is expected, due to the larger jet radius and lower $p_T$ threshold. The large-$R$ jet mass is generally well described by the simulation.
Figure 33: Leading-\( p_T \) jet \( \sqrt{d_{12}} \) of anti-\( k_t \) jets with \( R = 1.0 \) for (a) ungroomed jets and (b) trimmed jets (\( f_{\text{cut}} = 0.05 \) and \( R_{\text{sub}} = 0.3 \)).

Figure 34: Leading-\( p_T \) jet \( \sqrt{d_{13}} \) of anti-\( k_t \) jets with \( R = 1.0 \) for (a) ungroomed jets and (b) trimmed jets (\( f_{\text{cut}} = 0.05 \) and \( R_{\text{sub}} = 0.3 \)).

Figure 39 shows the top candidate mass distribution after applying the HEPTopTagger with four different filtering/large-\( R \) jet configurations. For all settings, the top mass peak is well described by MC and a relatively pure \( \bar{t}t \) selection is obtained for \( m_{\text{top}} > 120 \) GeV.

The good agreement between data and simulation both before and after applying the HEPTopTagger shows that the exploited substructure is modeled well, even for jets with a very large radius.

Figure 40 shows the requirements imposed by the HEPTopTagger on the subjet mass ratios for top-quark identification, as defined in Eq. (9) and Eq. (10). For example in Figure 40(b): if the sub-leading \( p_T \) and the sub-sub-leading \( p_T \) subjet are the decay products of a W boson then \( m_{23}/m_{123} \) peaks at \( m_W/m_{\text{top}} \sim 0.46 \). The \( m_W \) distribution in Figure 40(c) is obtained by taking the subjet pair with the invariant mass closest to the true \( m_W \). All distributions are well modeled by the simulation.
Figure 35: Leading-\(p_T\) jet \(\tau_{21}\) of anti-\(k_t\) jets with \(R = 1.0\) for (a) ungroomed jets and (b) trimmed jets \((f_{\text{cut}} = 0.05\) and \(R_{\text{sub}} = 0.3\)).

Figure 36: Leading-\(p_T\) jet \(\tau_{32}\) of anti-\(k_t\) jets with \(R = 1.0\) for (a) ungroomed jets and (b) trimmed jets \((f_{\text{cut}} = 0.05\) and \(R_{\text{sub}} = 0.3\)).

Figure 37: Leading-\(p_T\) jet mass of anti-\(k_t\) jets with \(R = 1.0\) for (a) ungroomed jets and (b) trimmed jets \((f_{\text{cut}} = 0.05\) and \(R_{\text{sub}} = 0.3\)), where one anti-\(k_t\) jet with \(R = 0.4\) was tagged as a \(b\)-jet.
Figure 38: Mass distribution for C/A jets with (a) $R = 1.5$ and (b) $R = 1.8$ before running the HEPTopTagger.

Figure 39: Top candidate mass distribution after the HEPTopTagger procedure (before applying a $m_{\text{top}}$ window). The distance parameter for large-$R$ jet finding is $R=1.5$ in (a), (c), and (d), and $R=1.8$ in (b). Distributions (a) and (b) are with default filtering settings, while (c) and (d) have been optimized for high purity and high signal efficiency, respectively.
Figure 40: Substructure variables $\text{arctan}(m_{13}/m_{12})$ (a), $m_{23}/m_{123}$ (b), and $m_W$ (c) for HEPTopTagger-tagged top candidates using the default filtering parameters and a jet size of $R = 1.5$. 
7 Conclusions

We have demonstrated that grooming algorithms increase sensitivity to new physics processes with boosted objects while reducing the dependence on pile-up and underlying event.

Calibrations have been derived in simulation for various large-\( R \) jet collections, as well as for jets with grooming applied. These have been validated \textit{in situ} using calorimeter-jet versus track-jet double ratios, and uncertainties on the jet energy scale and jet mass scale have been provided over a wide range of large-\( R \) jet momentum.

Using an inclusive jet sample, ATLAS simulation has shown agreement between mass distributions observed in data using large-\( R \) jets before and after grooming, especially for the POWHEG NLO generator. The substructure variables presented here also show good agreement between data and simulation.

Comparing various performance measures, the parameters of trimmed, pruned and mass-drop/filtered jet algorithms have been optimized for searches in ATLAS. The trimming algorithm exhibits the best performance with superior mass resolution and reduced dependence on pile-up compared to the pruning algorithm. In particular, the anti-\( k_t \) with \( R = 1.0 \) trimmed jets collection using parameters \( f_{\text{cut}} = 0.05 \) and \( R_{\text{sub}} = 0.3 \) is recommended for boosted top physics analyses in ATLAS. Additionally, C/A jets with \( R = 1.2 \) using the mass-drop/filtering parameter \( \mu_{\text{frac}} = 0.67 \) are recommended for boosted two-pronged analyses such as \( H \rightarrow b\bar{b} \) or searches using \( W \rightarrow q\bar{q} \).

The benefit of using these grooming algorithms along with substructure variables is demonstrated in top-tagging studies, where the efficiency of finding a boosted top is greatly increased after grooming is applied. Grooming has been shown to leave the boosted signal mass peak relatively unaffected while systematically shifting the light quark and gluon jet background lower in mass, thus increasing the discrimination of signal over background. The HEPTopTagger was demonstrated to be a robust and versatile tool to reconstruct hadronically-decaying top quarks in the presence of UE and pile-up using jet grooming and substructure techniques. The algorithm is well modeled in simulation when compared to data. The HEPTopTagger performance (efficiency, rejection, mass resolution) can be optimized for a given analysis by varying the algorithm parameters.

From the studies presented here, groomed jets and substructure variables are ready to be used in further ATLAS analyses. These techniques will become extremely beneficial tools in upcoming searches for boosted objects in Higgs, supersymmetric and exotic models, as well as in other Standard Model measurements.
References


