ATLAS NOTE

ATLAS-CONF-2012-066

June 26, 2012

Studies of the impact and mitigation of pile-up on large radius and groomed jets in ATLAS at $\sqrt{s} = 7$ TeV

The ATLAS Collaboration

Abstract

Large radius jets provide one avenue towards efficient reconstruction of massive boosted objects whose decay products are sufficiently collimated so as to make standard reconstruction techniques impractical. The potentially adverse impact of additional proton-proton interactions on such large jets is assessed for a variety of jet types and hadronic final state topologies. The mitigation of these effects via jet grooming algorithms such as trimming, pruning, and filtering is then studied for high transverse momentum jets at $\sqrt{s} = 7$ TeV using an integrated luminosity of 4.7 fb^{-1} collected with the ATLAS detector. A total of 29 jet algorithms and grooming configuration combinations are studied. The application of jet trimming and filtering significantly improves the robustness of large-R jets and reduces their sensitivity to the intense environment of the high-luminosity LHC. The consequence is an overall improvement in the physics potential of searches for heavy boosted objects.

ATLAS-CONF-2012-066
04 July 2012

1 Introduction

Jets have historically been utilized at high energy colliders as proxies for the quarks and gluons produced in the primary hard interaction. The jet definition is the means by which the four-momenta of those partons, stable hadrons, or even experimentally measured calorimeter energy deposits are typically compared. Many studies in ATLAS have demonstrated the performance [1] and physics potential [2–5] of jet measurements based on the relatively new anti-k_t jet algorithm [6, 7] using data from 2010 and 2011 proton-proton (pp) collisions. Typically, comparisons to Monte Carlo (MC) simulations in these analyses are based on the final four-momenta found, and thus the jets serve as merely "surrogates for the individual short-distance energetic parton" [8]. However, much more information is encoded within the structure of the jet itself, along with potentially erroneous additions to, or subtractions from it [9, 10]. Measurements of jet shapes and internal structure seek to extract this additional information in the context of understanding the structure of the underlying QCD parton shower evolution, as well as in searches for new physics in cases where this structure is expected to appear qualitatively different.

A companion study [11] presents a comprehensive comparison and description of the performance of advanced jet reconstruction algorithms in ATLAS, including jet modification techniques that selectively remove or redefine the constituents of a jet – often referred to as jet "grooming" – such as filtering [12], pruning [8, 13], and trimming [14]. This note focuses specifically on the performance of these techniques at high instantaneous luminosities where the impact from multiple simultaneous proton-proton interactions, or pile-up, can have significant negative consequences for large radius jets. In particular, these studies demonstrate that the resilience of jet properties to pile-up is significantly improved through the use of grooming. Various performance measures are considered and techniques for measuring the effects of pile-up *in situ* are presented. With these measures, a subset of the configurations of each grooming algorithm studied, which perform well in context of pile-up, is identified. The stability of the mass reconstruction for both signal and background is given a strong emphasis. The observation is made that the resilience to pile-up provided by certain grooming configurations results in improved discrimination power as well. Recommendations are given for the techniques that provide the best performance in searches for highly boosted particles decaying to a single jet. Studies of pile-up corrections for standard jet algorithms in ATLAS in 2011 are presented elsewhere [15].

2 Data Samples, Event Selection, and Luminosity Profiles

2.1 Event Selection and Data Quality Criteria

The studies presented in this note use the full 2011 dataset, corresponding to an integrated luminosity of (4.7 ± 0.2) fb^{-1} [16]. The data are required to have met some baseline quality criteria. These criteria reject significant contamination from detector noise, non-collision beam backgrounds, and other spurious effects. Some studies shown here focus on the subset of data collected at high instantaneous luminosity near the end of the 2011 data-taking period, corresponding to roughly 1 fb^{-1} of integrated luminosity. These data are characterized by an average instantaneous luminosity of approximately $2 - 3 \times 10^{33}$ cm^{-2}s^{-1}, a mean number of reconstructed primary vertices of $N_{\mathrm{PV}} \approx 7 - 8$, and a peak number of interactions per bunch crossing (μ) near 17.

A three-level trigger system was 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 trigger levels, level-2 and the event filter, which together reduce the event rate to a few hundred Hz. Events were selected if the leading jet in the event passed a single jet trigger at the event filter stage with a transverse momentum defined with respect to the beam

direction of $p_{\mathrm{T}}^{\mathrm{jet}} > 350$ GeV[1].

The ATLAS data quality selection is based on individual assessments for each subdetector, usually separated into barrel, endcap, and forward regions, as well as for the trigger and for each type of reconstructed physics object (jets, electrons, muons, etc.). The primary systems of interest for this study are the electromagnetic and hadronic calorimeters, as well as the inner tracking detector for studies of the properties of tracks associated with jets.

To reject non-collision beam backgrounds, events are required to contain a primary vertex consistent with the LHC beamspot, reconstructed from at least 2 tracks with transverse momenta $p_{\mathrm{T}}^{\mathrm{track}} > 400$ MeV. Jet-specific requirements are also applied. All jets reconstructed with the anti-k_t algorithm [6, 7] using a radius parameter of $R = 0.4$ and a measured $p_{\mathrm{T}}^{\mathrm{jet}} > 20$ GeV are required to satisfy the "looser" requirements discussed in detail in Ref. [17]. 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. Any event containing a jet of this type that fails the selection criteria is rejected.

For the studies shown here, the inputs to jet reconstruction are either stable truth particles with a lifetime of at least 10 ps (excluding muons and neutrinos) in the case of MC "truth jets", three-dimensional topological clusters, or "topo-clusters," in the case of fully reconstructed calorimeter jets, or charged-particle tracks in the case of so-called "track jets" [1]. In the latter case, track quality selection is applied in order to ensure good quality tracks that originate from the reconstructed primary vertex which has the largest $\sum(p_{\mathrm{T}}^{\mathrm{track}})^2$ in the event and contains at least two tracks. For reconstructed 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 [18]. The resulting topo-clusters are then classified as either electromagnetic or hadronic based on their shape, depth and energy density, and are treated as massless inputs to the jet algorithm. Energy corrections are then applied in order to calibrate the clusters to the hadronic scale.

2.2 Monte Carlo Simulation

The data are compared to inclusive jet events generated by two MC simulations: PYTHIA 6.425 [19] and POWHEG-BOX 1.0 [20–22] (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 leading-order proton parton distribution function (PDF) set MRST LO* [23]. In the latter, POWHEG+PYTHIA uses the CTEQ6L1 PDF set [24]. For both cases, PYTHIA is tuned with corresponding AUET2B tune [25, 26]. The comparison between PYTHIA and POWHEG+PYTHIA represents an important juxtaposition, at least at the matrix element (ME) level, between a leading order ME MC (PYTHIA) and an NLO ME generator (POWHEG). After simulation of the parton shower and hadronization, events are passed through the full GEANT4 [27] detector simulation [28]. Following this, the same trigger, event, quality, jet, and track selection criteria are applied to the MC simulation as are applied to the data.

Additional MC samples of events containing boosted hadronic particle decays are used for direct comparisons of the performance of the various reconstruction and jet substructure techniques in signal-like events. For these studies, events containing two collimated jets of three closely-spaced objects from the decay of a single massive particle, sometimes referred to as "three-prong" decays, are obtained from boosted hadronically-decaying $t\bar{t}$ pairs. The boost of the top quarks in these events is due to the decay of

[1]The 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 θ is measured with respect to the LHC beam-line. The azimuthal angle ϕ 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 η is an approximation for rapidity y in the high energy limit, and it is related to the polar angle θ as $\eta = -\ln\tan\frac{\theta}{2}$. Transverse momentum and energy are defined as $p_{\mathrm{T}} = p \times \sin\theta$ and $E_{\mathrm{T}} = E \times \sin\theta$, respectively.

an additional heavy gauge boson, $Z' \to t\bar{t}$ ($M_{Z'} = 1.6$ TeV). These events were generated using PYTHIA 6.425 and also used the MRST LO* PDF set. 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 [26] 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 of pp interactions. 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 the average number of interactions per bunch crossing (μ) agrees with the data, as measured by the luminosity detectors in ATLAS [16].

3 Performance of Grooming in the Context of Pile-up

Detailed descriptions of the trimming [14], pruning [8, 13], mass-drop filtering [12] algorithms, their implementation, and performance in ATLAS are provided in Ref. [11].

In this note, performance measures that describe the impact of pile-up and the extent to which jet grooming helps to ameliorate those effects are categorized into three general types:

1. *In situ* measures of the dependence of the jet kinematics and properties on pile-up;
2. Comparisons of jet mass reconstructed using the calorimeter to those measured via tracks in both data and the MC simulation;
3. MC simulation measures of the differential impact of pile-up on jets containing massive particle decays and those resulting from the QCD parton showers.

Both the anti-k_t and Cambridge-Aachen (C/A) [29, 30] jet algorithms are used for jet finding. The k_t algorithm [31, 32] is also used to define several substructure observables and in the grooming procedures themselves (e.g. in trimming and pruning). These algorithms are implemented within the framework of the FastJet software [6, 33]. The first two measures are discussed in Section 3.2 and Section 3.3, whereas the third is the subject of Section 3.4.

3.1 Description of the Grooming Algorithms

Each grooming algorithm uses a particular jet definition for jet finding, and a potentially different definition to define the substructure criteria on which the algorithm is based. In each case, the mass of the jet, m^{jet}, is defined by

$$(m^{\mathrm{jet}})^2 = (\sum_i E_i)^2 - (\sum_i p_i)^2, \tag{1}$$

where the sums are taken over either the original constituents of the jet or over the remaining constituents after grooming. A brief description of each of the grooming algorithms follows.

Mass-drop Filtering: The mass-drop filtering procedure seeks to isolate concentrations of energy within a jet by identifying relatively symmetric subjets, j_1 and j_2, each with a significantly smaller mass than that of the original jet with mass m^{jet}. The first requirement in the mass-drop criterion is that there be a significant difference between the original jet mass (m^{jet}) defined by the C/A algorithm

and the highest mass subjet, m^{j_1}, after the splitting such that $m^{j_1}/m^{\text{jet}} < \mu_{\text{frac}}$, where μ_{frac} is a parameter of the algorithm. This splitting is defined by the C/A algorithm itself and consists of reversing the last stage of the 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 splitting into those subjets is required to be relatively symmetric in mass-drop filtering, such that $\frac{\min[(p_T^{j_1})^2,(p_T^{j_2})^2]}{(m^{\text{jet}})^2} \times \Delta R^2_{j_1,j_2} > y_{\text{cut}}$, where $\Delta R_{j_1,j_2} = \sqrt{(y_{j_1} - y_{j_2})^2 + (\phi_{j_1} - \phi_{j_2})^2}$ is the opening angle between j_1 and j_2, and y_{cut} defines the energy sharing between the two highest p_{T} subjets within the original jet.

This technique was developed and optimized using C/A jets in the search for a Higgs boson decaying to two b-quarks, $H \to b\bar{b}$ [12]. 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 [10].

Trimming: The trimming algorithm [14] 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. The ratio of the p_{T} of subjets to that of the jet is used as a selection criterion. The inclusive k_t algorithm is used to create subjets using a radius parameter of R_{sub} from the constituents of a jet, since this algorithm, along with the C/A algorithm, uses the structural information of a group of objects to determine the recombination properties. This is in contrast to the anti-k_t algorithm, which seeks out the highest energy particles in a given region and effectively draws a circle around those objects to define the jet or subjet. Any subjets with $p_{Ti}/p_{\text{T}}^{\text{jet}} < f_{\text{cut}}$ are removed, where f_{cut} is a parameter of the method and p_{Ti} is the transverse momentum of the i^{th} subjet. The remaining constituents form the trimmed jet. Values of f_{cut} ranging from 1% to 5% are tested in these studies.

Completely removing the softer components from the final jet is possible as the majority of energy in large-R jets due to soft additional radiation from pile-up, MPI, and ISR is separated from that due to the hard-scatter. The resulting primary effect of pile-up on the jet mass and substructure of large-R jets is additional low-energy topo-clusters falling within the wide area formed by the jet as opposed to additional energy being added to topo-clusters from hard-scatter particles. This is especially true in the case of jets with very high p_{T} and hard substructure where the fractional impact of overlapping energy deposits from pile-up is reduced simply due to the significantly softer spectrum. 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 [1].

Pruning: The pruning algorithm [8, 13] 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 (in these studies, the k_t algorithm), based on the branching at each point in the jet reconstruction. This approach therefore does not require the explicit 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 re-building the pruned jet. Instead of the "top-down" approach taken by trimming, wherein the choice to remove constituents is driven by a reclustered set of objects with a momentum scale determined by the parent jets, the pruning algorithm attempts to make this choice dynamically. At each recombination step with constituents j_1 and j_2 (where $p_T^{j_1} > p_T^{j_2}$), pruning

requires that either condition $p_T^{j_2}/p_{\mathrm{T}}^{j1+j2} > z_{\mathrm{cut}}$ or $\Delta R_{j_1,j_2} < R_{\mathrm{cut}} \times \frac{2m^{\mathrm{jet}}}{p_{\mathrm{T}}^{\mathrm{jet}}}$ be met in order to retain that proto-jet in the recombination. j_2 is merged with j_1 if these criteria are met, otherwise, j_2 is discarded and the algorithm repeats.

The configurations of the grooming algorithms described above are given in Table 1. Furthermore, six additional pruning configurations closer to the values used in Refs. [8,13] were also tested, but exhibit a negligible impact on the pile-up dependence of the jet mass and properties as discussed in Section 3.2.

Jet finding algorithms used	Grooming algorithm	Configurations considered
C/A	Mass-Drop Filtering	$\mu_{\mathrm{frac}} = 0.20, 0.33, \mathbf{0.67}$
anti-k_t and C/A	Trimming	$f_{\mathrm{cut}} = 0.01, \mathbf{0.03}, 0.05$ $R_{\mathrm{sub}} = \mathbf{0.2}, 0.3$
anti-k_t and C/A	Pruning	$R_{\mathrm{cut}} = 0.1, 0.2, 0.3$ $z_{\mathrm{cut}} = 0.05, 0.1$

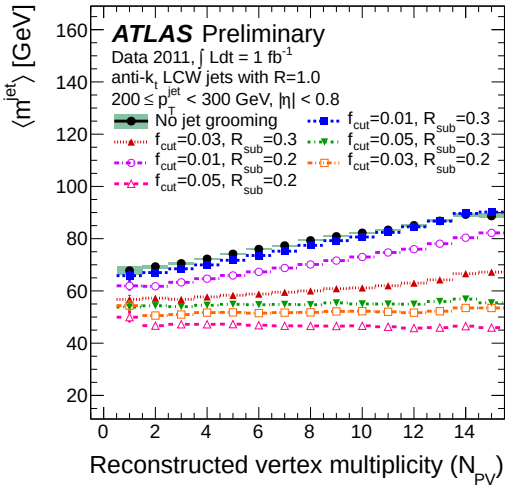
Table 1: Summary of the grooming configurations considered in this study. Values in boldface are optimized configurations reported in Ref. [12] and Ref. [14] for filtering and trimming, respectively.

3.2 Dependence of the Jet Energy Scale and the Jet Mass Scale on pile-up

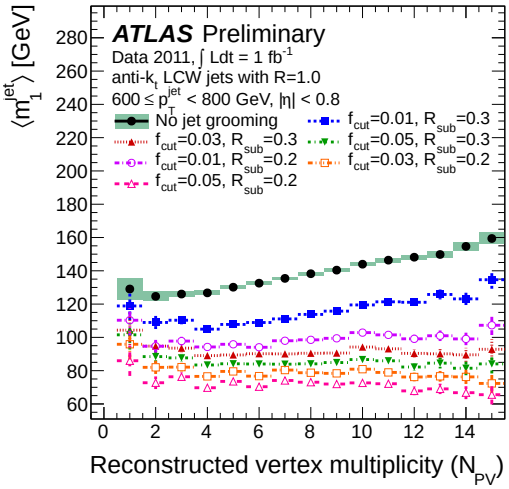
The jet mass may be calibrated using MC-based calibration factors derived as a function of the jet $p_{\mathrm{T}}^{\mathrm{jet}}$ and η [11]. This section elaborates on the magnitude and variation of the impact of pile-up on the jet mass and other observables and the extent to which trimming, filtering, and pruning are able to minimize those effects. In particular, this measure of performance is used as one of the primary figures of merit in determining a subset of groomed jet algorithms on which to focus for physics analysis in ATLAS.

Figure 1 shows the dependence of the mean uncalibrated jet mass, $\langle m^{\mathrm{jet}} \rangle$, on the number of reconstructed primary vertices, N_{PV}, for jets in the central region $|\eta| < 0.8$. This dependence is shown in two $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges of interest for jets after trimming, filtering, and pruning. For these comparisons, only the final period of data collection from 2011 is used, which corresponds to approximately 1 fb^{-1} of integrated luminosity but represents the period with the highest instantaneous luminosity recorded at $\sqrt{s} = 7$ TeV. The lower range, $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV, represents the threshold for most hadronic boosted object measurements and searches, whereas the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV is expected to contain top quarks for which the decay products will be fully merged within an $R = 1.0$ jet nearly 100% of the time [11]. In each figure, the full set of grooming algorithm parameter settings is included for comparison. As noted in Table 1, two values of the subjet radius, R_{sub}, are used for trimming, three R_{cut} factors for pruning are tested, and three μ_{frac} settings are evaluated using the filtering algorithm.

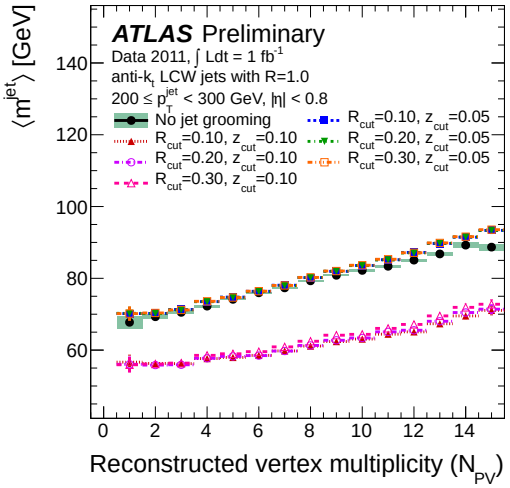
Several observations can be made from Figure 1. Trimming and filtering both have a significant impact on the strong rise for ungroomed jets of $\langle m^{\mathrm{jet}} \rangle$ with pile-up, whereas pruning does not. For at least one of the configurations tested, trimming and filtering are both able to reduce this dependence to approximately zero, as measured by the slope of $\langle m^{\mathrm{jet}} \rangle$ versus N_{PV}. Furthermore, the trimming configurations tested provide a highly tunable set of parameters that allow for a relatively continuous adjustment from small to large reduction of the pile-up dependence of the jet mass. The parameter settings with $R_{\mathrm{sub}} = 0.2$, $f_{\mathrm{cut}} = 0.03$ and $R_{\mathrm{sub}} = 0.3$, $f_{\mathrm{cut}} = 0.05$ exhibit good stability for both low and high $p_{\mathrm{T}}^{\mathrm{jet}}$, with the $f_{\mathrm{cut}} = 0.05$ configuration exhibiting a slightly smaller impact from pile-up at high N_{PV} for low $p_{\mathrm{T}}^{\mathrm{jet}}$. The other parameter settings either do not reduce the pile-up dependence at low $p_{\mathrm{T}}^{\mathrm{jet}}$ (e.g. $f_{\mathrm{cut}} = 0.01$) or



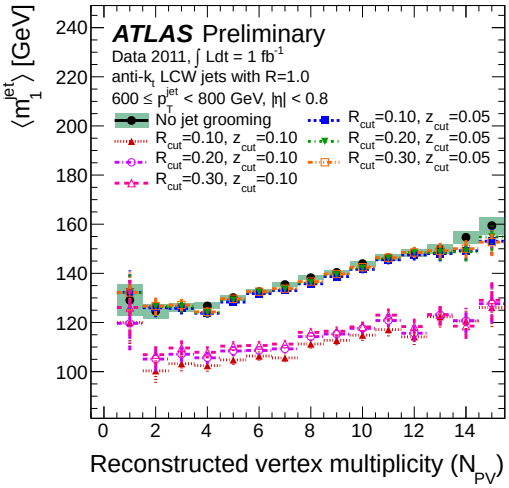
(a) Trimmed anti-k_t: $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV



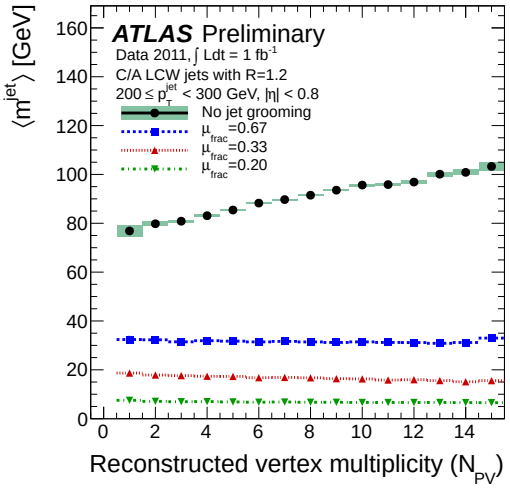
(b) Trimmed anti-k_t: $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV



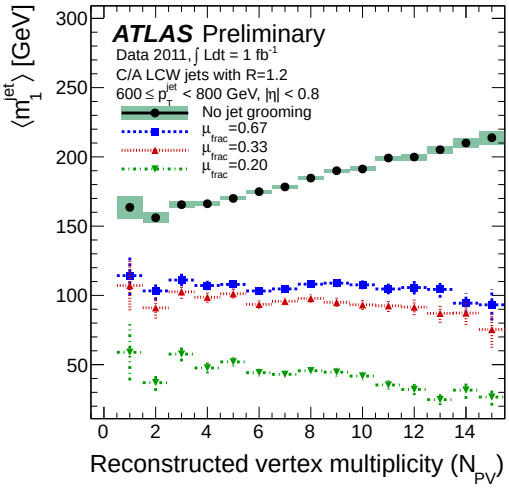
(c) Pruned anti-k_t: $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV



(d) Pruned anti-k_t: $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV



(e) Filtered C/A: $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV



(f) Filtered C/A: $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV

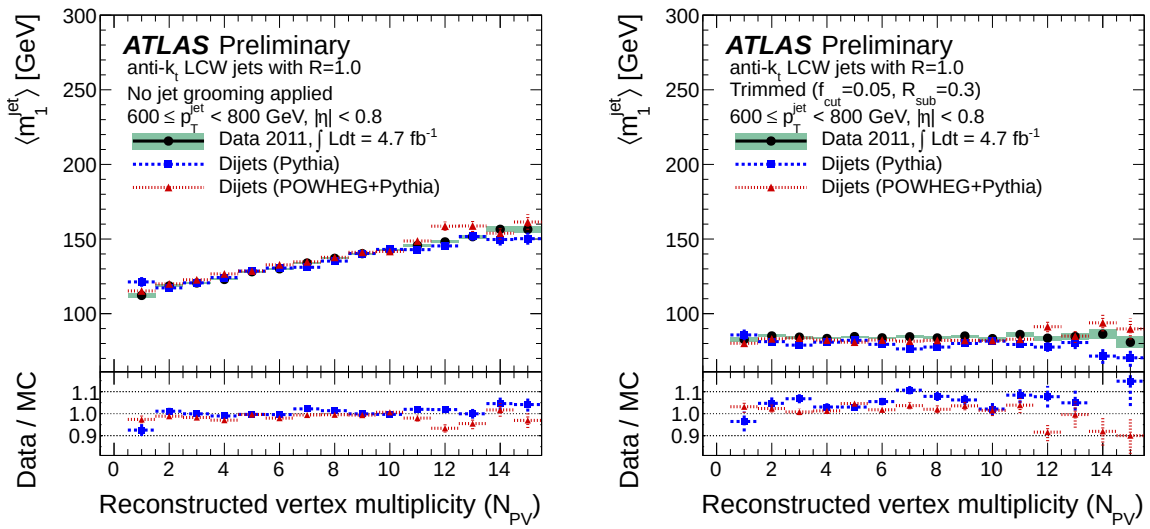
Figure 1: Evolution of the mean jet mass, $\langle m^{\mathrm{jet}} \rangle$, for jets in the central region $|\eta| < 0.8$ as a function of the reconstructed vertex multiplicity, N_{PV} for leading jets in the range $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV (left) and the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV (right). (a)-(b) show trimmed anti-k_t jets with $R = 1.0$, (c)-(d) show pruned anti-k_t jets with $R = 1.0$, and (e)-(f) show split and filtered C/A jets with $R = 1.2$. The error bars indicate the statistical uncertainty on the mean value in each bin.

result in a downward slope of $\langle m^{\text{jet}} \rangle$ as a function of pile-up at high $p_{\text{T}}^{\text{jet}}$ (e.g. $f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.2$ and $\mu_{\text{frac}} = 0.20$).

Pruning, on the other hand, exhibits the smallest impact on the pile-up dependence of the jet mass. Only by increasing the z_{cut} parameter from $z_{\text{cut}} = 0.05$ to $z_{\text{cut}} = 0.10$ can any reduction on the dependence of $\langle m^{\text{jet}} \rangle$ on pile-up be observed. This change is equivalent to reducing the fraction of low p_{T} contributions during the jet recombination; in the language of trimming, this is analogous to raising f_{cut}. This change reduces the magnitude of the variation of the mean jet mass as a function N_{PV} for low $p_{\text{T}}^{\text{jet}}$ slightly. Interestingly, the R_{cut} parameter gives very little impact on the performance, with nearly all of the differences observed being due to the change in z_{cut}. This observation holds for both low and high $p_{\text{T}}^{\text{jet}}$.

The mass-drop filtering algorithm can be made to affect significantly $\langle m^{\text{jet}} \rangle$ solely via the mass-drop criterion, μ_{frac}. A drastic change in $\langle m^{\text{jet}} \rangle$ is observed for all configurations of the jet filtering, with the strictest $\mu_{\text{frac}} = 0.20$ setting rejecting nearly 90% of the jets considered and resulting in a slight negative slope in the mean jet mass versus pile-up. Nevertheless, the other two settings of μ_{frac} tested in these studies exhibit no significant variation as a function of the number of reconstructed vertices, and the optimum value of $\mu_{\text{frac}} = 0.67$ seems to exhibit the greatest stability. Studies from 2010 [10] demonstrate that this reduction in the sensitivity to pile-up is due primarily to the filtering step in the algorithm as opposed to the jet selection itself.

Figure 2 presents the pile-up dependence of $\langle m_1^{\text{jet}} \rangle$ in data compared to two MC generators. Here, only the range $600 \leq p_{\text{T}}^{\text{jet}} < 800$ GeV for ungroomed and trimmed anti-k_t jets is shown for brevity, but similar conclusions apply in all $p_{\text{T}}^{\text{jet}}$ ranges. The comparison is made using the full 4.7 fb^{-1} of integrated luminosity collected in 2011. PYTHIA and POWHEG both model the data fairly accurately, with a slight 5%–10% discrepancy appearing in the prediction from PYTHIA for the trimmed jets. Most importantly, the impact of pile-up is very well modeled, with the slopes of the $\langle m_1^{\text{jet}} \rangle$ versus N_{PV} agreeing to within 3% compared to POWHEG +PYTHIA for both ungroomed and trimmed jets.

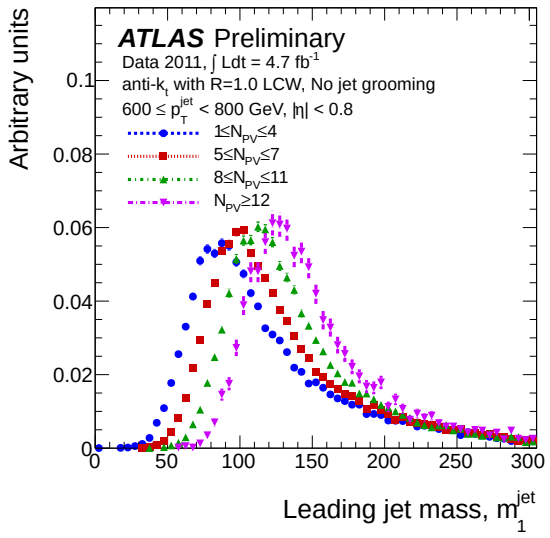


(a) anti-k_t, $R = 1.0$: Ungroomed

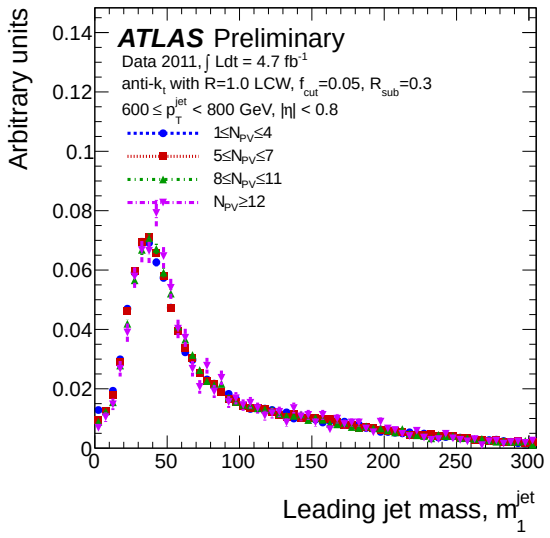
(b) anti-k_t, $R = 1.0$: Trimmed

Figure 2: Dependence of the mean jet mass, $\langle m^{\text{jet}} \rangle$, on the reconstructed vertex multiplicity for anti-k_t jets with $R = 1.0$ in the range $600 \leq p_{\text{T}}^{\text{jet}} < 800$ GeV in the central region $|\eta| < 0.8$. Both (a) untrimmed anti-k_t jets and (b) trimmed anti-k_t jets with $f_{\text{cut}} = 0.05$ show good agreement between data and MC. The error bars indicate the statistical uncertainty on the mean value in each bin.

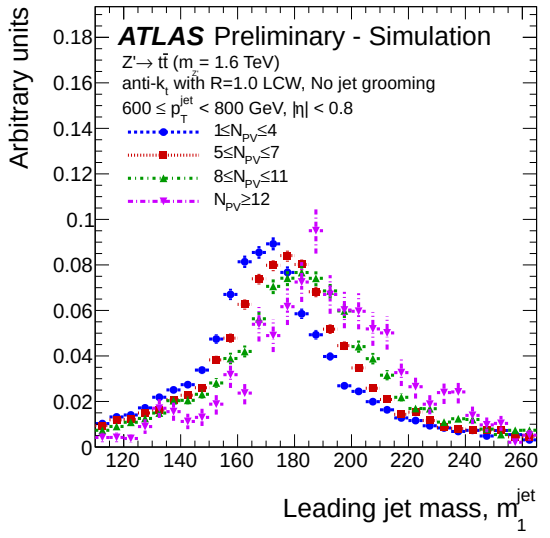
Beyond simply providing a pile-up-independent average jet mass, the optimal grooming configurations render the full jet mass spectrum insensitive to high instantaneous luminosity. Figure 3 demonstrates this by comparing the jet mass spectrum for leading ungroomed and trimmed anti-k_t jets. The comparison is performed both in data and using the $Z' \to t\bar{t}$ MC sample. The result using the inclusive jet sample obtained from data shows that a nearly identical trimmed m^{jet} spectrum is obtained regardless of the level of pile-up. The peak of the leading jet mass distribution for events with $N_{\text{PV}} \geq 12$ is shifted comparatively more due to trimming: from $m^{\text{jet}} \approx 125$ GeV to $m^{\text{jet}} \approx 45$ GeV as compared to an initial peak position of $m^{\text{jet}} \approx 90$ GeV for events with $1 \leq N_{\text{PV}} \leq 4$. Nonetheless, the resulting trimmed jet mass spectra exhibit no dependence on N_{PV}.



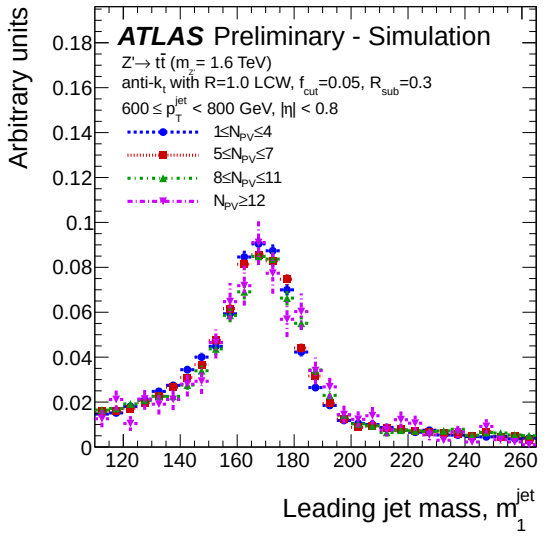
(a) Data: anti-k_t, $R = 1.0$: Ungroomed



(b) Data: anti-k_t, $R = 1.0$: Trimmed



(c) Z': anti-k_t, $R = 1.0$: Ungroomed



(d) Z': anti-k_t, $R = 1.0$: Trimmed

Figure 3: Jet mass spectra for four primary vertex multiplicity ranges for anti-k_t jets with $R = 1.0$ in the range $600 \leq p_{\text{T}}^{\text{jet}} < 800$ GeV. Both untrimmed anti-k_t jets (left) and trimmed anti-k_t jets (right) are compared for the various N_{PV} ranges in data (top) and for a $Z' \to t\bar{t}$ sample (bottom).

Comparisons performed using the $Z' \to t\bar{t}$ sample demonstrate the same performance of the trimming algorithm, but in the context of the reconstruction of highly boosted top quarks. Figures 3(c)-(d) indicate that the ability to render the full jet mass distribution independent of pile-up does not come at the cost of the mass resolution or scale. Prior to jet trimming, a shift in the peak position of the jet mass of nearly 15 GeV is observed between the lowest and the highest ranges of N_{PV} studied. After jet trimming, each of the resulting mass spectra for the various N_{PV} ranges are narrower and lie directly on top of one another, even in the case of a jet containing a highly boosted, and thus very collimated, top quark decay. This observation, combined with that above, demonstrates that the trimming algorithm is working as expected by removing soft, wide-angle contributions to the jet mass while retaining the relevant hard substructure of the jet.

Finally, although not shown explicitly, the same observations made here for the resilience of the trimmed jet mass spectrum can also be made for the filtered jet mass.

The method used to determine the mass scale uncertainty is discussed at length in Ref. [11]. The approach utilizes the ratio of the calorimeter m^{jet} to the track jet mass, which measures the same quantity as in Eq. (1) but using tracks as input to the jet algorithm instead of clusters. The mass ratio, $r^{m}_{\text{track jet}}$, is defined explicitly as

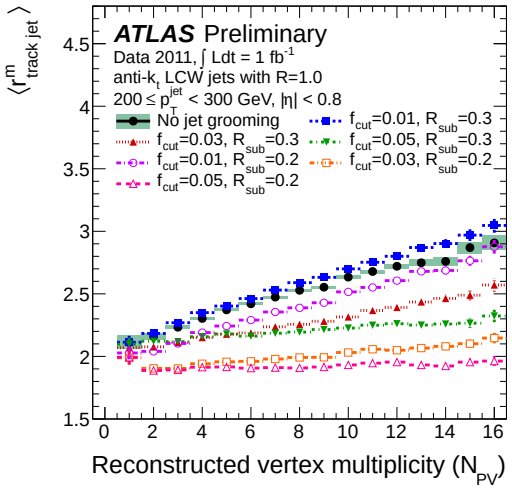
$$r^{m}_{\text{track jet}} = \frac{m^{\mathrm{jet}}}{m^{\text{track jet}}}. \tag{2}$$

The matching between calorimeter and track jets is performed using a matching radius of $\Delta R < 0.3$ in $\eta - \phi$ coordinates. The ratio in Eq. (2) is 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 were unaccounted for by the MC simulation, then as long as this process affects both the track jet mass and the calorimeter jet mass, then the ratio of data to MC should be relatively unaffected. The double ratio of $r^{m}_{\text{track jet}}$ evaluated in the data to that evaluated in the MC simulation is constructed to test this. Systematic effects on the calibration of these jets are estimated via deviations between data and MC for the dependence of $r^{m}_{\text{track jet}}$ on $p_{\mathrm{T}}^{\mathrm{jet}}$ and m^{jet}. In Figure 4, the dependence of the mean value of $r^{m}_{\text{track jet}}$ on N_{PV} is evaluated for trimming, filtering, and pruning for two $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges of interest. This is analogous to Figure 1, but isolates the impact of pile-up on the mass itself on a jet-by-jet basis given that $m^{\text{track jet}}$ for each jet is not affected by pile-up, to first order.

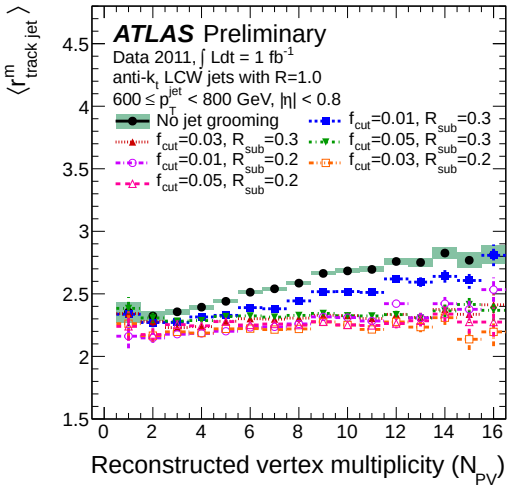
Figure 4 demonstrates that trimming and filtering both provide robust jet definitions in the presence of high luminosity. $r^{m}_{\text{track jet}}$ is observed to be nearly equal among the various trimming configurations in the case of little or no in-time pile-up (i.e. $N_{\mathrm{PV}} \approx 1$) whereas filtering shows a significant, although small, difference between the configurations using $\mu_{\mathrm{frac}} = 0.67, 0.33, 0.20$. This shows that the filtering method does affect the magnitude of m^{jet} and $m^{\text{track jet}}$ slightly differently, resulting in an approximate 12% relative drop in $\langle r^{m}_{\text{track jet}} \rangle$ after filtering. As indicated in Figure 5, these distributions are nonetheless well modeled by the MC simulation, resulting in double ratios of data to MC very close to one (shown in the bottom panels in Figure 5). The trimming configuration with $R_{\mathrm{sub}} = 0.3$, $f_{\mathrm{cut}} = 0.01$ has almost no impact on the dependence of $r^{m}_{\text{track jet}}$ with pile-up.

3.3 Jet Substructure Properties

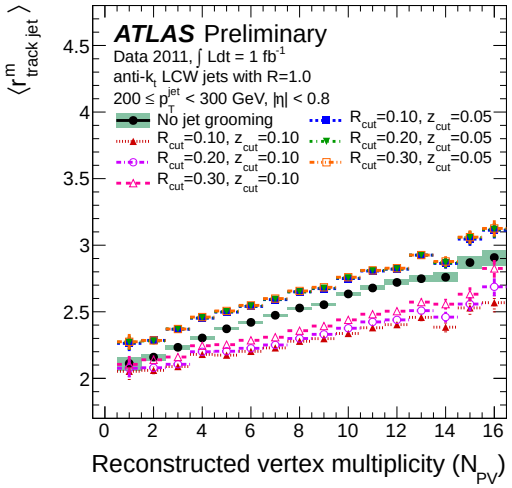
Searches and measurements that utilize the information encoded in the jet mass can often also benefit from more detailed and potentially complex measures of jet substructure. Two such observables are $\sqrt{d_{ij}}$ and τ_N. The family of $\sqrt{d_{ij}}$ variables, most often seen in the form of $\sqrt{d_{12}}$ or $\sqrt{d_{23}}$, classify the massive "scales" of a given jet. For example, the expected value for a two-body heavy particle decay is approximately $\sqrt{d_{12}} \approx m^{\mathrm{jet}}/2$ [11], whereas jets formed from the parton shower of light quarks and gluons will tend to exhibit a steeply falling spectrum of both $\sqrt{d_{12}}$ and $\sqrt{d_{23}}$. These are defined by



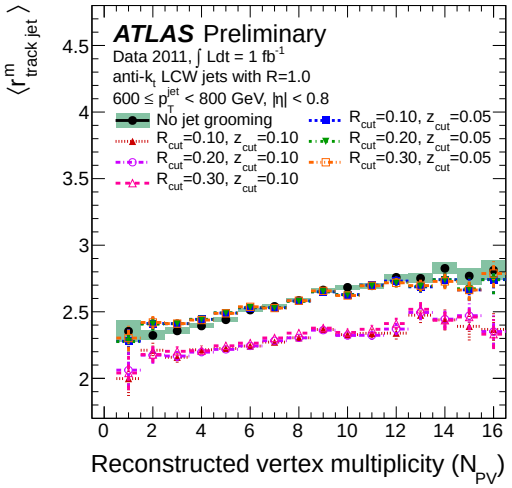
(a) Trimmed anti-k_t: $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV



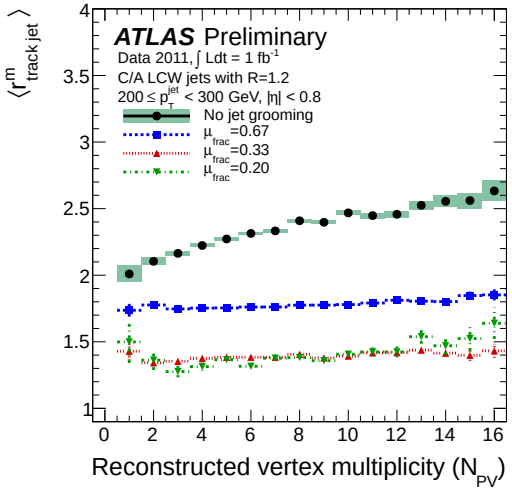
(b) Trimmed anti-k_t: $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV



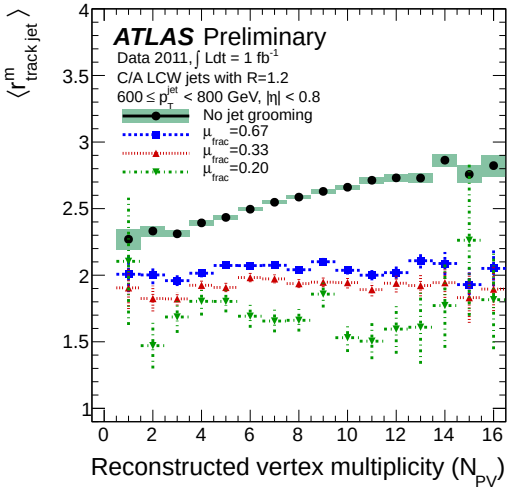
(c) Pruned anti-k_t: $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV



(d) Pruned anti-k_t: $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV

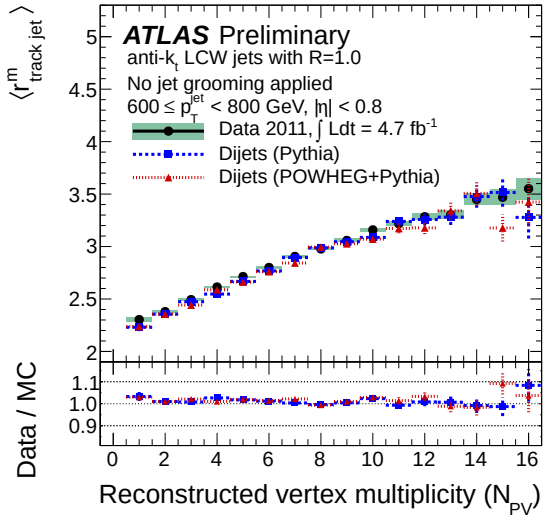


(e) Filtered C/A: $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV

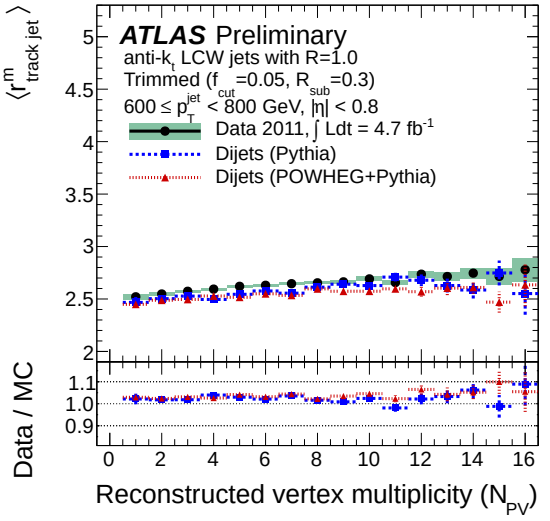


(f) Filtered C/A: $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV

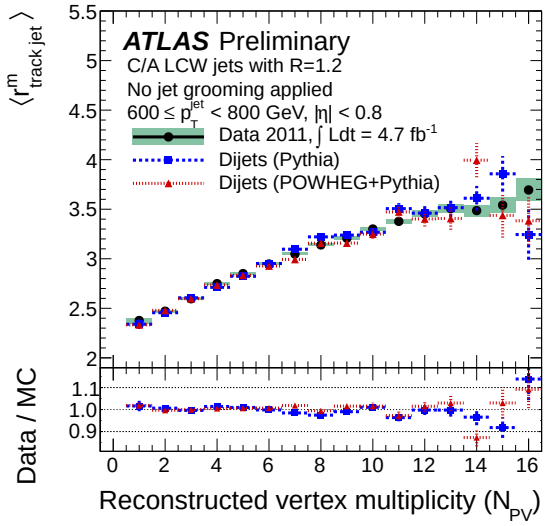
Figure 4: Evolution of the mean jet mass to track jet mass ratio, $\langle r^m_{\text{track jet}} \rangle$, as a function of the reconstructed vertex multiplicity, N_{PV} for the range $200 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 300$ GeV range (left), and the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV (right). (a)-(b) show trimmed anti-k_t jets with $R = 1.0$, (c)-(d) show pruned anti-k_t jets with $R = 1.0$, and (e)-(f) show split and filtered C/A jets with $R = 1.2$. The error bars indicate the statistical uncertainty on the mean value in each bin.



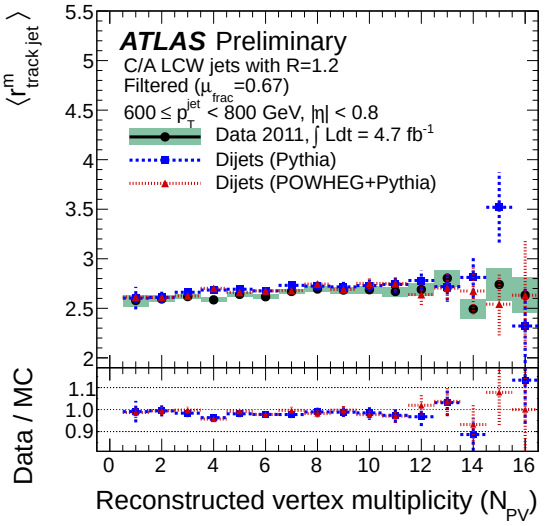
(a) anti-k_t, $R = 1.0$: Ungroomed



(b) anti-k_t, $R = 1.0$: Trimmed



(c) C/A, $R = 1.2$: Ungroomed



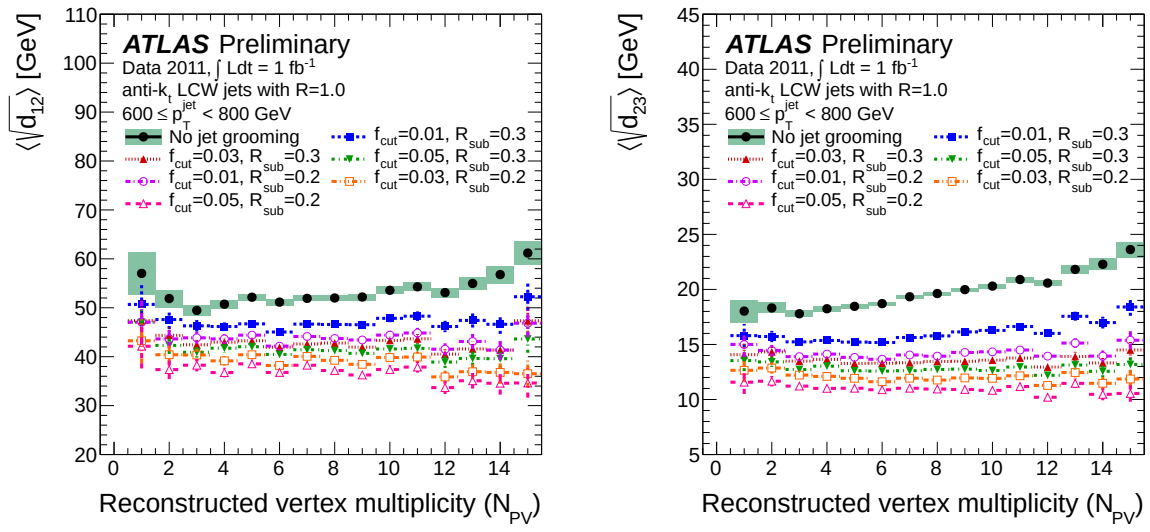
(d) C/A, $R = 1.2$: Filtered

Figure 5: Dependence of the mean jet mass to track jet mass ratio, $\langle r^m_{\text{track jet}} \rangle$, on the reconstructed vertex multiplicity for jets in the range $600 \leq p_{\text{T}}^{\text{jet}} < 800$ GeV. (top) anti-k_t jets with $R = 1.0$ with both (a) ungroomed and (b) trimmed configurations. (bottom) C/A jets with $R = 1.2$ also with (c) ungroomed and (d) filtered configurations. The error bars indicate the statistical uncertainty on the mean value in each bin.

applying the k_t algorithm to the constituents of a jet and calculating

$$\sqrt{d_{ij}} = \min(p_{\text{T}i}, p_{\text{T}j}) \times \Delta R_{ij}, \tag{3}$$

where ΔR_{ij} is the distance in $y - \phi$ 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.



(a) $\sqrt{d_{12}}$ (b) $\sqrt{d_{23}}$

Figure 6: Variation of the mean splitting scales (a) $\langle \sqrt{d_{12}} \rangle$ and (b) $\langle \sqrt{d_{23}} \rangle$ measured in data for anti-k_t jets with $R = 1.0$ in the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV before and after trimming. The error bars indicate the statistical uncertainty on the mean value in each bin.

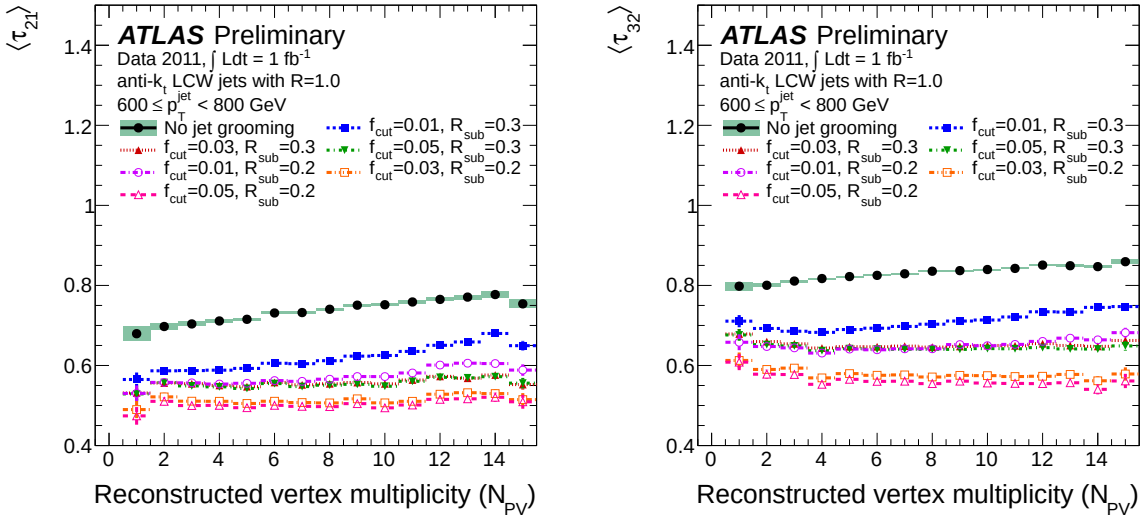
The N-subjettiness variables, τ_N, on the other hand, classify the extent to which a given jet is likely to have N subjets. The variables τ_N are then defined in Eq. (4) as the sum over all constituents k of the jet:

$$\tau_N = \frac{1}{d_0} \sum_k p_{\mathrm{T}k} \times \min(\delta R_{1k}, \delta R_{2k}, ..., \delta R_{Nk}) \,, \text{ with } \; d_0 \equiv \sum_k p_{\mathrm{T}k} \times R \tag{4}$$

where R is the jet radius parameter in the jet algorithm, $p_{\mathrm{T}k}$ is the p_{T} of constituent k and δR_{ik} is the distance in $y-\phi$ coordinates from the subjet i to constituent k. Improved discrimination has been demonstrated [34, 35] when using the ratio of two τ_N variables, as in the case of $\tau_{21} = \tau_2/\tau_1$ and $\tau_{32} = \tau_3/\tau_2$. Most often, τ_{21} is used to seek out and classify two-body decays, whereas τ_{32} is used for three-body decays.

Figures 6 and 7 show the variation observed with pile-up for each of the splitting scales and N-subjettiness observables for jets in the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV as measured in data. In this case, the focus is on the jet trimming algorithm since the mass-drop filtering algorithm makes a predefined choice to search for properties of a jet characteristic of a two-body decay. The constraints placed on subjet multiplicity by the filtering procedure are not appropriate for calculating generic jet shapes given the strict substructure requirements they place on a jet. Furthermore, pruning does not seem to mitigate the effects of pile-up. The trimming configurations with $R_{\mathrm{sub}} = 0.3$ and $f_{\mathrm{cut}} = 0.03, 0.05$ yield the most stable jet substructure properties with the smallest deviation in their observed mean values at low N_{PV}. This conclusions holds for all other jet $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges as well, with larger differences between $f_{\mathrm{cut}} = 0.03, 0.05$ appearing at low $p_{\mathrm{T}}^{\mathrm{jet}}$.

Figure 8 presents a comparison of the $\langle \sqrt{d_{12}} \rangle$ and $\langle \tau_{32} \rangle$ for ungroomed and trimmed ($R_{\mathrm{sub}} = 0.3$, $f_{\mathrm{cut}} = 0.05$) anti-k_t jets with $R = 1.0$. The description of the dependence of these observables on N_{PV} provided by the MC is observed to be very close to that observed in the data.



(a) τ_{21} (b) τ_{32}

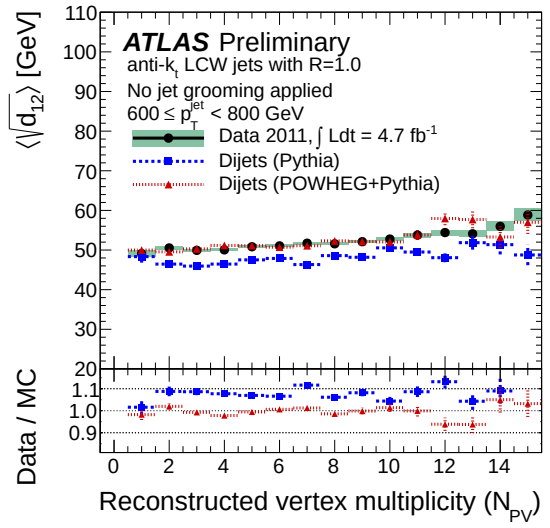
Figure 7: Variation of the mean N-subjettiness ratios (a) $\langle\tau_{21}\rangle$ and (b) $\langle\tau_{32}\rangle$ measured in data for anti-k_t jets with $R = 1.0$ in the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV before and after trimming. The error bars indicate the statistical uncertainty on the mean value in each bin.

3.4 Impact of Pile-up on Signal and Background in the Monte Carlo

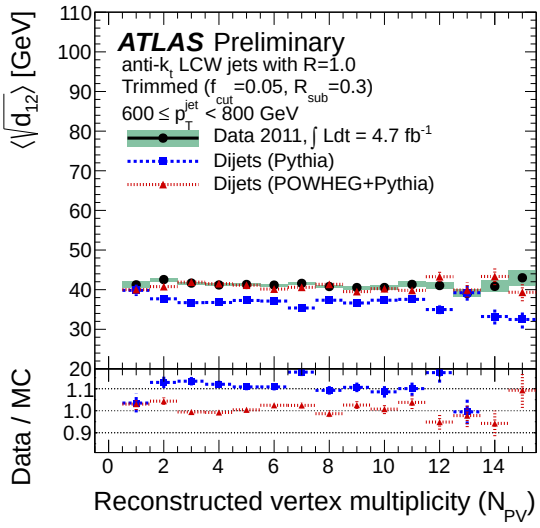
In addition to the comparisons between data and MC, and between the various grooming configurations, a comparison of how grooming impacts signal-like events versus background-like events is crucial.

Figure 9 shows the variation of the average leading jet mass, $\langle m_1^{\mathrm{jet}}\rangle$, in the range $600 \leq p_{\mathrm{T}}^{\mathrm{jet}} < 800$ GeV for ungroomed and trimmed anti-k_t, $R = 1.0$ jets, for both the $Z' \to t\bar{t}$ sample and the POWHEG dijet sample. The average ungroomed leading jet mass in the sample of light quarks and gluons in the inclusive POWHEG dijet events exhibits a slope of approximately $d\langle m_1^{\mathrm{jet}}\rangle/dN_{\mathrm{PV}} \approx 3.00$ GeV/N_{PV}. The leading jets in the Z' sample are typically entirely composed of fully hadronic boosted top quark decays contained in a single jet. The mass reconstruction in this case proceeds as usual (four-momentum recombination) but the mass distribution is highly peaked near the top quark mass of approximately 175 GeV. Jets in this peak but without grooming exhibit a slope of roughly $d\langle m_1^{\mathrm{jet}}\rangle/dN_{\mathrm{PV}} \approx 2.15$ GeV/N_{PV}, or about 30% smaller than in the inclusive jet sample. In the case of trimmed jets, the slopes as a function of N_{PV} for both signal-like jets and jets in dijet events are consistent with zero.

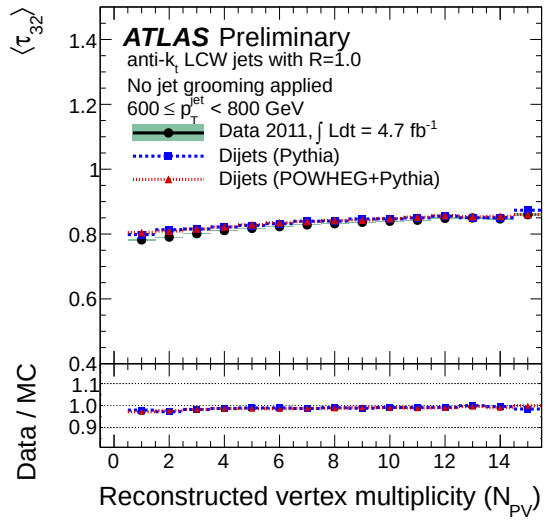
Most importantly, the average separation in the mean jet mass for signal-like jets in the Z' sample and the background represented by the POWHEG QCD dijet sample increases by nearly 50% after trimming and remains stable across the full range of N_{PV}. The separation shown here is significant since the widths of the peaks of each of the distributions are also simultaneously narrowed by the grooming algorithm, as shown in Figure 3. This differential impact of trimming is again due to the design of the algorithm [11]: soft, wide angle contributions to the jet mass that are ubiquitous in QCD jets are suppressed whereas the hard components present in a jet with true substructure – as in the case of the top-quark jets here – are preserved.



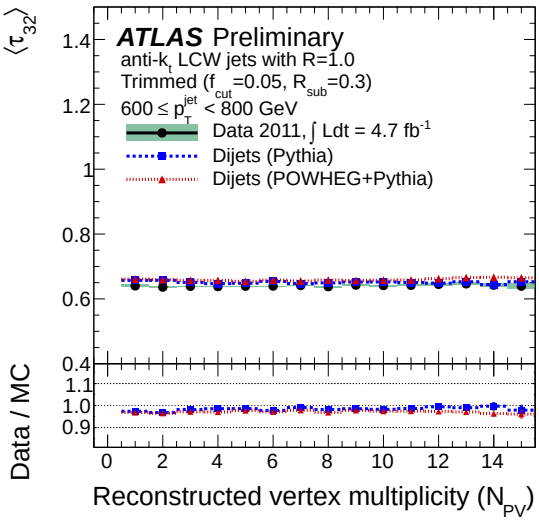
(a) anti-k_t, R = 1.0: Ungroomed



(b) anti-k_t, R = 1.0: Trimmed



(c) anti-k_t, R = 1.0: Ungroomed

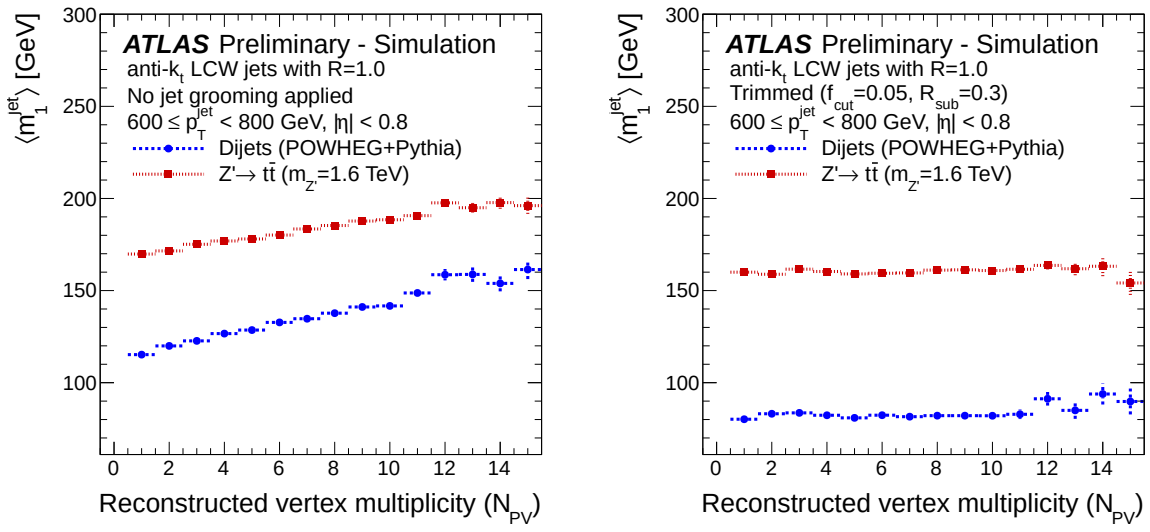


(d) anti-k_t, R = 1.0: Trimmed

Figure 8: Variation as a function of N_{PV} of the mean splitting scales $\langle\sqrt{d_{12}}\rangle$ for in data and MC for (a) ungroomed anti-k_t jets with $R = 1.0$ and (b) trimmed ($R_{\mathrm{sub}} = 0.3$, $f_{\mathrm{cut}} = 0.05$) anti-k_t jets. Variation as a function of N_{PV} of the mean N-subjettiness $\langle\tau_{32}\rangle$ ratios for (c) ungroomed anti-k_t jets with $R = 1.0$ and (d) trimmed ($R_{\mathrm{sub}} = 0.3$, $f_{\mathrm{cut}} = 0.05$) anti-k_t jets. The error bars indicate the statistical uncertainty on the mean value in each bin. The panels below the main figures show the ratio of the mean values measured in data to MC.

4 Conclusions

A study of the effect of pile-up on the properties of jets with and without grooming is conducted in both data and MC simulation. A total of 29 jet grooming algorithm and parameter configurations are studied. The impact of pile-up on the jet mass, calibration validations from *in situ* measurements using tracks, and jet properties, is studied for ungroomed jets as well for the trimming, pruning, and mass-drop filtering



(a) Ungroomed

(b) Trimmed

Figure 9: Variation of the average leading jet mass, $\langle m_1^{\text{jet}} \rangle$, in the range $600 \leq p_{\text{T}}^{\text{jet}} < 800$ GeV for (a) ungroomed and (b) trimmed jets. The error bars indicate the statistical uncertainty on the mean value in each bin.

algorithms. Prior to jet grooming, the average jet mass can be shifted by 2–4 GeV per additional reconstructed vertex in the event which significantly degrades the shape and width of hadronic resonance peaks reconstructed with a single jet. Taking into account only effects due to pile-up but not reconstruction or theoretical uncertainties, trimmed or filtered jets are observed to exhibit the greatest stability among the grooming algorithms in terms of the mass and substructure properties at high luminosity. Furthermore, the trimming configuration utilizing $R_{\text{sub}} = 0.3$ and $f_{\text{cut}} = 0.05$ is found to be the most effective in providing this resilience and is close to that suggested by the authors of the algorithm. This setting also increases the average separation between the mean jet mass for signal-like jets containing a boosted top quark from the $Z' \rightarrow t\bar{t}$ sample and a background-like jet sample represented by the POWHEG QCD dijet sample. Similar conclusions are obtained from a study of the overall performance of these grooming algorithms [11] wherein the trimming configuration with $R_{\text{sub}} = 0.3$ and $f_{\text{cut}} = 0.05$ is found to have superior mass resolution for boosted massive objects and stable jet mass scale for both QCD-like jets and jets containing hard substructure. Of course, the theoretical implications of these procedures still need to be explored, but the performance and resilience to pile-up of trimming and mass-drop filtering look very promising.

References

[1] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Submitted to EPJC (2011) , arXiv:1112.6426 [hep-ex].

[2] ATLAS Collaboration, *Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector*, Eur. Phys. J. **C71** (2011) 1795, arXiv:1109.5816 [hep-ex].

[3] ATLAS Collaboration, *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, Eur. Phys. J. **C71** (2011) 1512, arXiv:1009.5908 [hep-ex].

[4] ATLAS Collaboration, *Study of Jet Shapes in Inclusive Jet Production in pp Collisions at $\sqrt{s}$ = 7 TeV using the ATLAS Detector*, Phys. Rev. **D83** (2011) 052003, arXiv:1101.0070 [hep-ex].

[5] ATLAS Collaboration, *Measurements of multi-jet production cross sections in proton-proton collisions at 7 TeV center-of-mass energy with the ATLAS Detector*, ATLAS-CONF-2010-084. http://cdsweb.cern.ch/record/1298854.

[6] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. **B641** (2006) 57, arXiv:0802.1189 [hep-ph].

[7] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[8] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, Phys. Rev. **D81** (2010) 094023, arXiv:0912.0033 [hep-ph].

[9] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[10] ATLAS Collaboration, *Jet mass and substructure of inclusive jets in $\sqrt{s}$ = 7 TeV pp collisions with the ATLAS experiment*, JHEP **1205** (2012) 128, arXiv:1203.4606 [hep-ex].

[11] ATLAS Collaboration, *Performance of large-R jets and jet substructure reconstruction with the ATLAS detector*, ATLAS-CONF-2011-065. https://cdsweb.cern.ch/record/1455115.

[12] J. M. Butterworth et al., *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, arXiv:0802.2470 [hep-ph].

[13] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, Phys. Rev. **D80** (2009) 051501, arXiv:0903.5081 [hep-ph].

[14] D. Krohn, J. Thaler, and L.-T. Wang, *Jet trimming*, JHEP **2010** (2010) 20, arXiv:0912.1342 [hep-ph].

[15] ATLAS Collaboration, *Pile-up corrections for jets from proton-proton collisions at $\sqrt{s}$ = 7 TeV in ATLAS in 2011*, ATLAS-CONF-2011-064. https://cdsweb.cern.ch/record/1448667.

[16] ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s}$ = 7 TeV using the ATLAS Detector in 2011*, ATLAS-CONF-2011-116. https://cdsweb.cern.ch/record/1376384.

[17] ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, ATLAS-CONF-2012-020. https://cdsweb.cern.ch/record/1430034.

[18] W. Lampl et al., *Calorimeter clustering algorithms: description and performance*, ATL-LARG-PUB-2008-002. https://cdsweb.cern.ch/record/1099735.

[19] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 physics and manual*, JHEP **0605** (2006) 026, arXiv:0603175 [hep-ph].

[20] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

[21] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, arXiv:0709.2092 [hep-ph].

[22] S. Alioli, K. Hamilton, P. Nason, C. Oleari, and E. Re, *Jet pair production in POWHEG*, JHEP **1104** (2011) 081, arXiv:1012.3380 [hep-ph].

[23] A. Sherstnev and R. S. Thorne, *Parton distributions for LO generators*, Eur. Phys. J. **C55** (2008) 553–575, arXiv:0711.2473 [hep-ph].

[24] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:0201195 [hep-ph].

[25] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, ATL-PHYS-PUB-2011-008. https://cdsweb.cern.ch/record/1345343.

[26] ATLAS Collaboration, *ATLAS tunes for Pythia6 and Pythia8 for MC11*, ATLAS-PHYS-PUB-2011-009. https://cdsweb.cern.ch/record/1363300.

[27] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[28] ATLAS Collaboration, *The ATLAS simulation infrastructure*, Eur. Phys. J. **C70** (2010) 823, arXiv:1005.4568 [physics.ins-det].

[29] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Better jet clustering algorithms*, JHEP **9708** (1997) 001, arXiv:hep-ph/9707323 [hep-ph].

[30] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross sections in deep- inelastic scattering*, arXiv:hep-ph/9907280 [hep-ph].

[31] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993) 3160, arXiv:hep-ph/9305266 [hep-ph].

[32] S. Catani, Y. L. Dokshitzer, M. Seymour, and B. Webber, *Longitudinally invariant $k_\perp$ clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993) 187.

[33] M. Cacciari, G. P. Salam, and G. Soyez. http://fastjet.fr/.

[34] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **1103** (2011) 015, arXiv:1011.2268 [hep-ph].

[35] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, JHEP **1202** (2012) 093, arXiv:1108.2701 [hep-ph].