PROCEEDINGS OF SCIENCE

Studies of jet shapes and jet substructure in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS

Peter LOCH[*] **(On behalf of the ATLAS Collaboration.)**
University of Arizona
E-mail: loch@physics.arizona.edu

Jet substructure reconstruction techniques are applied in the context of the reconstruction of heavy boosted particles, with a particular emphasis on searches for new physics, and in other analyses sensitive to the flavor structure of the jets. The ATLAS experiment at the LHC conducts extensive performance evaluations for a series of jet substructure techniques using proton-proton collision data taken in 2011 at $\sqrt{s} = 7$ TeV. Selected results for various observables and their sensitivity to experimental conditions like pile-up are presented, together with comparisons to Monte Carlo simulations. It is found that most of the considered jet substructure related mass, distance and splitting scale measures can be understood with precisions of better than 10%.

The European Physical Society Conference on High Energy Physics
18-24 July, 2013
Stockholm, Sweden

[*]Speaker.

PoS(EPS-HEP 2013)442

1. Introduction

The analysis of the substructure of jets produced in hadron collisions is motivated by the attempt to determine the origin of a given jet from its internal transverse momentum flow pattern. In particular, it exploits the differences between the pattern generated by radiation, showering and fragmentation of a light quark or gluon and the one generated by the products of a two- or three-prong decay of a (possibly boosted) heavy particle collected into one jet. This makes the jet substructure analysis a useful tool not only for the reconstruction of boosted heavy Standard Model particles like the W boson and the top quark, but also for searches for even heavier new particles indicating physics beyond the Standard Model.

The ATLAS detector [1] is well suited for the analysis of final states with jets generated by the proton-proton collisions at the LHC. The high center-of-mass energy of $\sqrt{s} = 7$ TeV in these collisions, and the large amount of data recorded in 2011, corresponding to about 4.7 fb^{-1}, for the first time allows an extensive evaluation of the performance of jet substructure reconstruction techniques with ATLAS. The results of this evaluation can be found in Ref. [2], with some selected results presented here.

2. Jets in ATLAS in 2011

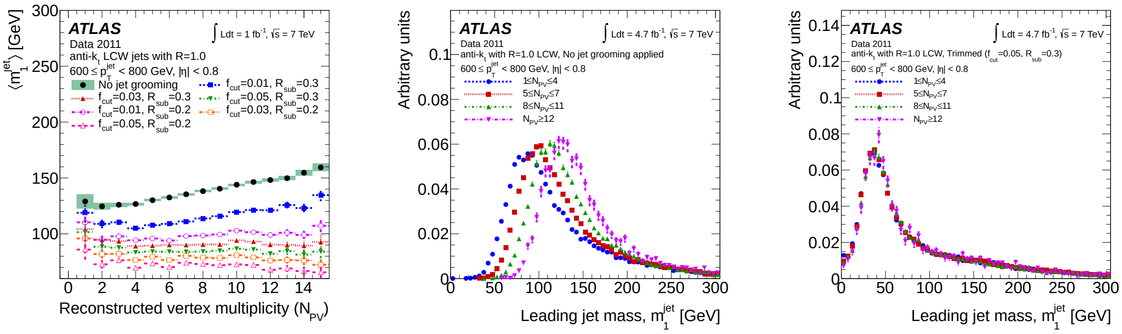
Jets in ATLAS are reconstructed from clusters of cells following signal topologies in the calorimeter [3]. For the resolved analysis of the hadronic final state, the anti-k_t algorithm [4], as provided by the FASTJET [5] package, is used with distance parameters (jet sizes) $R = 0.4$ and $R = 0.6$. All jets are calibrated and corrected to the hadron level, following the procedures outlined in Ref. [6]. For jet substructure analyses, larger jets are typically used to accommodate the decays of heavy particles, as the jet size relates to the particle mass m and its transverse momentum $p_{\rm T}$ as $R \approx 2m/p_{\rm T}$ (see e.g. Ref. [7]).

The experimental conditions during the 2011 data taking at LHC are characterized by the presence of pile-up introduced by the high instantaneous luminosity and high total inelastic proton-proton cross section of about 71.5 mb [8]. This pile-up is generated by additional proton-proton interactions in the same bunch crossing as the triggered hard scattering interaction. It manifests itself in ATLAS by additional signals arising from the diffuse and stochastic scattering of particles from these interactions, and in signal modifications introduced by the bunch collision history [9]. The corresponding transverse momentum flow affects not only the reconstruction of the global jet kinematics, but also the measurement of the jet mass and structural jet features. Techniques to mitigate the pile-up effects are discussed in Ref. [9] for the global jet energy scale, and in Ref. [10] for substructure related observables before and after application of substructure analysis.

Most performance studies use inclusive high $p_{\rm T}$ large-R jet samples ($p_{\rm T} > 600$ GeV, $R = 1$), which are typical backgrounds for signal objects on mass scales of order 300 GeV decaying into a single jet of this size. The corresponding collision events are extracted mostly from single- and multi-jet triggers, and all jets are required to pass high quality reconstruction criteria (see Ref. [2] for more details).

In Monte Carlo (MC) simulations, the hard scattering process is generated using PYTHIA 6.425 [11] or, alternatively, HERWIG++ [12]. Both generators employ leading order matrix ele-

PoS(EPS-HEP 2013)442



(a) Variation of $\langle m^{\rm jet}\rangle$ with pile-up (b) Ungroomed jets (c) Trimmed jets

Figure 1: The variation of the average leading jet mass $\langle m_1^{\rm jet}\rangle$ as function of the number of reconstructed collision vertices $N_{\rm PV}$ in data, for anti-k_t jets with $R = 1.0$ in the central detector region of ATLAS ($|\eta| < 0.8$) (a). Shown are ungroomed jets and jets trimmed with various configurations $(R_{\rm sub}, f_{\rm cut})$. The distributions of $m_1^{\rm jet}$ for ungroomed jets and jets trimmed with $R_{\rm cut} = 0.3$ and $f_{\rm cut} = 0.05$ are shown in (b) and (c), respectively All jets considered in these figures are required to have $600 \leq p_{\rm T} < 800$ GeV (plots from Ref. [2]).

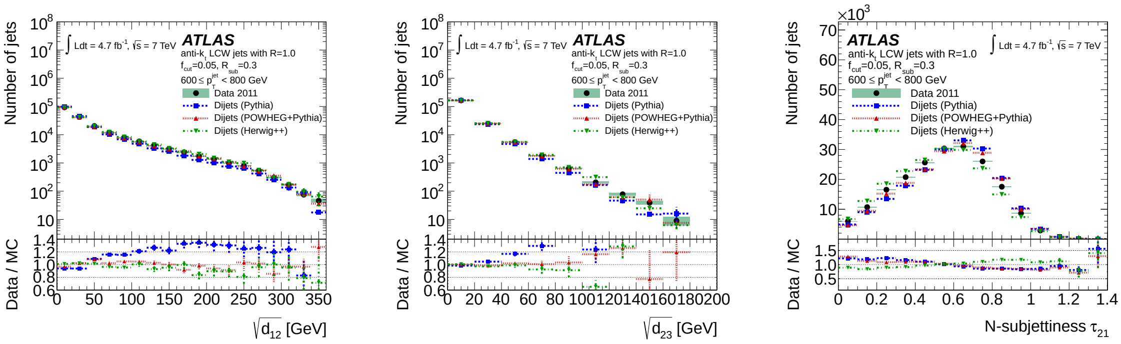
ments for $2 \to 2$ inclusive jet production. In addition, POWHEG [13], which employs a next-to-leading order (NLO) matrix element, is used together with PYTHIA parton shower, hadronization, and multiple parton interaction modeling. Pile-up is added at hadron level by overlaying the appropriate number μ of single minimum bias collisions generated by PYTHIA and tuned to ATLAS data. Here μ follows a Poisson distribution around the average number of pile-up interactions $\langle\mu\rangle$, which itself is distributed as measured for the corresponding event samples in data.

3. Jet substructure analysis

The common feature of all jet substructure reconstruction techniques is the attempt to provide a structural analysis of the internal $p_{\rm T}$ flow in a jet and enhance patterns in it arising from pronged particle decays. The following suite of jet substructure reconstruction techniques is evaluated, with the values in braces indicating configurations considered in all combinations by ATLAS for this *jet grooming*:

Trimming [14] removes small sub-jets with size $R_{\rm sub} = \{0.2, 0.3\}$, clustered with the k_t algorithm [18, 19] inside a large-R jet, which have less than a fraction $f_{\rm cut} = \{0.01, 0.03, 0.05\}$ of the original jet $p_{\rm T}$.

Mass drop filtering [16] undoes the last recombination of a large-R jet built with the Cambridge/Aachen (C/A) algorithm [15] until the leading sub-jet mass is smaller by a factor of $\mu_{\rm frac} = \{0.20, 0.33, 0.67\}$ of the original jet mass, and the splitting is not too asymmetric ($y > y_{\rm cut} = 0.09$). If this criteria is met, the accepted jet is filtered such that only the three leading C/A sub-jets reconstructed with $R_{\rm filt} = \max(0.3, \Delta R(\mathrm{j}_1, \mathrm{j}_2))$ are used to build the groomed jet.



(a) $\sqrt{d_{12}}$, trimmed jets (b) $\sqrt{d_{23}}$, trimmed jets (c) $\tau_{21} = \tau_2/\tau_1$, trimmed jets

Figure 2: The distributions of the splitting scales $\sqrt{d_{12}}$ (a) and $\sqrt{d_{23}}$ (b) for anti-k_t jets with $R = 1.0$, in data and for various MC simulations. The distributions of the N-subjettiness ratio τ_{21} for the same jet samples is shown in (c). Jets are required to have $600 \leq p_{\mathrm{T}} < 800$ GeV (plots from Ref. [2]).

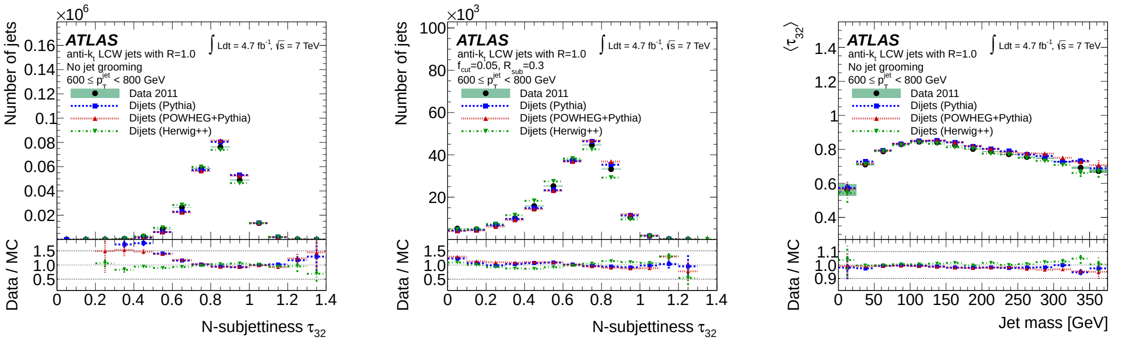
Pruning [17] suppresses large distance recombinations $\Delta R > R_{\mathrm{cut}} = \{0.1, 0.2, 0.3\}$, and recombinations where the lower p_{T} constituent has only a small fraction $z_{\mathrm{cut}} = \{0.05, 0.10\}$ of the combined p_{T}, in the C/A or k_t jet cluster sequence.

The groomed jet is reconstructed from the surviving sub-jets in case of trimming and mass drop filtering, or by the modified recombination sequence in case of pruning. The variables evaluated for the groomed and the corresponding original jet are the jet mass $m_{\mathrm{jet}} = \sqrt{E_{\mathrm{jet}}^2 - p_{\mathrm{jet}}^2}$, the k_t splitting scales $d_{ij} = \min[p_{\mathrm{T},i}, p_{\mathrm{T},j}] \times \Delta R_{ij}$, and the N-subjettiness τ_N [20], which measures how well a given jet can be described assuming N sub-jets. While the splitting scales d_{ij} have expectation values for pronged decays, e.g. $\langle\sqrt{d_{12}}\rangle \approx m/2$ for jets from two-pronged decays of a particle with mass m, N-subjettiness ratios τ_{N+1}/τ_N can be efficient in separating multi-pronged decay jets from quark or gluon initiated jets.

4. Results

The largest effects of pile-up are expected for the jet mass reconstruction. The additional soft signals, which can appear at large distances from the jet axis, add mass to jets. In ATLAS this additional mass contribution is found to be on average about $d\langle m_{\mathrm{jet}}\rangle/dN_{\mathrm{PV}} \approx 2.6$ GeV[1] for anti-k_t jets with $R = 1.0$, and with $600 \leq p_{\mathrm{T}} < 800$ GeV, see Figure 1(a) . N_{PV} is the number of reconstructed collision vertices in the event, which is a good measure of the in-time pile-up activity. The mass contribution from pile-up to lower p_{T} jets is smaller, at about 1.7 GeV per reconstructed vertex for $200 \leq p_{\mathrm{T}} < 300$ GeV. Applying e.g. trimming with $R_{\mathrm{sub}} = 0.3$ and $f_{\mathrm{cut}} = 0.05$ restores $\langle m^{\mathrm{jet}}\rangle$ at any N_{PV} to the value observed under non-pile-up ($N_{\mathrm{PV}} = 1$) conditions. In addition, trimming with this configuration can also mitigate the effect of pile-up on the m^{jet} distribution, as can be seen in the comparison of Figures 1(b) and (c).

[1]Here, jet masses are reconstructed without dedicated mass calibration, see Ref. [2] for details.



(a) τ_{32}, ungroomed jets (b) τ_{32}, trimmed jets (c) $\langle\tau_{32}\rangle$ versus jet mass, ungroomed jets

Figure 3: The N-subjettiness ratio τ_{32} for ungroomed (a) and trimmed (b) anti-k_t jets with $R = 1.0$ in ATLAS data and MC simulations. The correlation between the mean $\langle\tau_{32}\rangle$ and the jet mass is shown in (c) for ungroomed jets. All jets considered have $600 \leq p_{\mathrm{T}} < 800$ GeV (plots from Ref. [2]).

The understanding of substructure related observables reconstructed in data is based on the ability of MC simulations to reproduce not only average behaviors but also features of the distributions of these variables. Figure 2 shows the distributions of the splitting scales $\sqrt{d_{12}}$ and $\sqrt{d_{23}}$ together with the distributions for the N-subjettiness ratio $\tau_{21} = \tau_2/\tau_1$, for trimmed anti-$k_t$ jets with $R = 1.0$ in ATLAS data and various MC simulations. The spectra for the splitting scales are better described by NLO based simulations, an observation confirmed as expected in similar comparisons of jet mass spectra, due to the correlations between $\sqrt{d_{12}}$ and m^{jet}. The τ_{21} distributions shown in Figure 2(c) show less sensitivity to the order of calculation of the matrix element, as all three MC simulations agree to the same level, with some qualitative differences between PYTHIA and HERWIG++. Those indicate sensitivities of τ_{21} to the soft physics modeling, an observation supported by the good agreement between PYTHIA (LO) and POWHEG (NLO), where in both cases the soft physics modeling is provided by PYTHIA.

The application of trimming does not affect the data-to-MC differences observed for the splitting scale distributions from ungroomed jets. It improves the data-to-MC agreement, though, for the N-subjettiness ratio distributions, as soft physics contributions to τ_N, including those from pile-up, are suppressed by the jet grooming. This can be seen in the comparison of Figures 3(a) and (b). The correlation between the means $\langle\tau_{32}\rangle$ and the jet mass is described to significantly better than 5% over nearly the whole mass range for ungroomed jets by all considered MC models, see Figure 3(c). After grooming, the HERWIG++ based MC simulations show a slightly larger disagreement in the order of 5%, while both PYTHIA and POWHEG based simulations maintain their level of agreement with data.

5. Conclusions

ATLAS is studying the performance of jet trimming, mass drop filtering, and pruning in great detail. Ref [2] contains a full summary for all considered configurations and jets in different final states. In general it is observed that the effect of pile-up on the reconstruction of the jet mass, internal jet splitting scales, and more complex variables like N-subjettiness ratios can be mitigated with high efficiency by applying jet grooming. In particular trimming configured with sub-jet sizes $R_{\mathrm{sub}} = 0.3$ and a sub-jet p_{T} threshold of $f_{\mathrm{cut}} = 5\%$ of the large-R jet p_{T} shows a good performance in the reconstruction of jet mass, jet shapes, and jet substructure.

The comparisons of internal transverse momentum flow patterns in jets in data and MC simulations indicates a preference for NLO matrix element calculations for the modeling of splitting scales. N-subjettiness ratios τ_{N+1}/τ_N at least up to $N = 2$, on the other hand, do not seem to be very sensitive to the order of calculation in the matrix element, but show more dependence on the soft physics modeling.

The generally good agreement between data and MC, and first attempts for the full evaluation of systematic uncertainties in particular for the jet mass calibration described in Ref. [2], suggest systematic uncertainties of 10% or less (about 5% for the jet mass) for the reconstruction of the complex jet substructure observables in data. This allows first application of these techniques in searches for new physics, and leads to extensions of some exclusion limits with respect to a fully resolved analysis for e.g. new heavy particles decaying into three quarks [21].

Acknowledgments

We like to thank CERN and the LHC for providing the proton-proton collisions at high efficiency and quality, and all funding agencies for their continuing support of our experiment.

References

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

[2] ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, arXiv:1306.4945 [hep-ex] (accepted for publication in *JHEP*).

[3] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan and D. Rousseau *et al.*, *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002 (2008).

[4] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **0804**, 063 (2008) [arXiv:0802.1189 [hep-ph]].

[5] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].

[6] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2011 data*, ATLAS-CONF-2013-004 (2013).

[7] 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. D* **81**, 094023 (2010) [arXiv:0912.0033 [hep-ph]].

[8] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross-Section at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, *Nature Commun.* **2**, 463 (2011) [arXiv:1104.0326 [hep-ex]].

[9] ATLAS Collaboration, *Pile-up corrections for jets from proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS in 2011*, ATLAS-CONF-2012-064 (2012).

[10] ATLAS Collaboration, *Studies of the impact and mitigation of pile-up on large-R and groomed jets in ATLAS at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2012-066 (2012).

[11] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605**, 026 (2006) [hep-ph/0603175].

[12] M. Bahr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Platzer and P. Richardson *et al.*, *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58**, 639 (2008) [arXiv:0803.0883 [hep-ph]].

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

[14] D. Krohn, J. Thaler and L. -T. Wang, *Jet Trimming*, *JHEP* **1002**, 084 (2010) [arXiv:0912.1342 [hep-ph]].

[15] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in proceedings of *Monte Carlo generators for HERA physics. Hamburg, Germany, 1998-1999*, DESY-PROC-1999-02 (1999) 714 [hep-ph/9907280].

[16] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100**, 242001 (2008) [arXiv:0802.2470 [hep-ph]].

[17] S. D. Ellis, C. K. Vermilion and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, *Phys. Rev. D* **80**, 051501 (2009) [arXiv:0903.5081 [hep-ph]].

[18] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, *New clustering algorithm for multi - jet cross-sections in e+ e- annihilation*, *Phys. Lett. B* **269**, 432 (1991).

[19] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48**, 3160 (1993) [hep-ph/9305266].

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

[21] ATLAS Collaboration, *Search for pair production of massive particles decaying into three quarks with the ATLAS detector in* $\sqrt{s} = 7$ *TeV pp collisions at the LHC*, *JHEP* **1212**, 086 (2012) [arXiv:1210.4813 [hep-ex]].

PoS(EPS-HEP 2013)442