Measurement of jet substructure observables and jet fragmentation properties using the ATLAS detector*

T. Lenz^a, on behalf of the ATLAS Collaboration

^aPhysikalisches Institut, Universität Bonn, Bonn; Germany.

Abstract

Jet substructure observables and properties of jet fragmentation are sensitive to a wide range of Quantum Chromodynamics (QCD) phenomena. They are essential to probe a new regime of QCD at hadron colliders. Measurements of jet substructure observables are also crucial to improve our understanding of the jet substructure properties used in many searches for physics beyond the Standard Model (BSM). These proceedings present some measurements of jet substructure and jet fragmentation observables performed with 33 fb^{-1} of proton-proton (pp) collision data collected by the ATLAS experiment at the Large Hadron Collider (LHC) at centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$. The measurements are corrected for detector effects, represented as particle-level distributions and are compared with state-of-the-art Monte Carlo (MC) simulations and theoretical calculations.

Keywords:

1. Introduction

The strong interaction is responsible for most of the physics processes occurring in pp collisions at the LHC. The quarks and gluons build collimated sprays of particles called jets. The radiation pattern inside jets, the jet substructure, can be used to identify hadronically decaying massive particles. They are usually reconstructed as large-radius (large- R) jets and the jet internal structure is used to classify their origin. Grooming algorithms are employed to reduce the effect of soft, uncorrelated radiation contained in the large- R jets.

Jet originating from quarks or gluons, from hadronically decaying top quarks or W bosons are studied by the ATLAS experiment [1] using dedicated event selection. These measurements are compared to the predictions of various MC simulation programs and to theoret-

ical calculations. The comparisons are done at leading-logarithm (LL) accuracy for most of the observables, only the large- R jet mass calculations are available at next-to-next-to-leading-logarithm (NNLL) accuracy.

Jet formation from high-energy quarks and gluons is a complex process. It involves the fragmentation of quarks and gluons with net QCD charge into net neutral hadrons. Some important contributions from QCD effects to this process cannot be described by perturbative theory. Thus a measurement of quantities related to fragmentation is essential for further understanding of this process and with it the properties of QCD.

Quark- and gluon-initiated jets have different radiation patterns. These differences in the radiation patterns can be used for jet tagging to differentiate between quark-enriched or gluon-enriched processes. Measurements of jet substructure can be used to calibrate quark-versus-gluon jet taggers. Calibrations that use particle-level measurements are partially independent of the ATLAS and LHC operating conditions and can be updated with the changing detector environment.

Particular attention is paid to gluon-initiated jets from

*Talk given at 22nd International Conference in Quantum Chromodynamics (QCD 19), 2 - 5 July 2019, Montpellier - FR
© 2019 CERN for the benefit of the ATLAS Collaboration. Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

Email address: tlenz@cern.ch (T. Lenz)

the $g \rightarrow b\bar{b}$ splitting. The $g \rightarrow b\bar{b}$ process is the main background for searches with boosted Higgs bosons decaying into $b\bar{b}$ pairs and many other SM measurements and BSM searches by providing an additional source of real b -quarks. On the other hand, the modelling of $g \rightarrow b\bar{b}$ fragmentation provides a useful probe of QCD. The large mass of the b -quark introduces a significant modification to the massless QCD splitting functions. A measurement of the properties of $g \rightarrow b\bar{b}$ at high transverse momenta and low opening angles between the b -jets can be used to validate the MC predictions and to tune the free parameters of the MC models.

2. Jet reconstruction

Various jet reconstruction techniques are utilised by the ATLAS experiment. This section summarises the techniques relevant for the measurements discussed in these proceedings.

Calorimeter-based jets are built from noise-suppressed topological clusters and are reconstructed using FastJet [2] with the anti- k_t algorithm [3] with a radius parameter of $R = 1.0$ or 0.8 (large- R jets) or $R = 0.4$ (small- R jets). The topological clusters of the large- R jets are brought to the hadronic energy scale using the local hadronic cell weighting scheme [4]. The jets are groomed using the trimming technique [5, 6] or the soft-drop procedure [9] to discard the softer components of jets that originate from initial-state radiation, pile-up interactions or underlying event. The trimming is done by reclustering the constituents of the initial jet, using the k_t algorithm [7, 8], into subjets of radius parameter $R = 0.2$ and removing any subjet that has a p_T less than 5% of the parent jet p_T .

The trimming procedure is not analytically calculable and the trimmed jet measurements cannot be compared to calculations beyond the LL accuracy of parton shower (PS) MC programs. The reason for this is the presence of nonglobal logarithms: resummation terms associated with particles that radiate out of a jet and then radiate back into a jet. A new procedure has been developed to systematically remove soft and wide-angle radiation from the jet to make the grooming algorithm insensitive to these effects: the soft-drop grooming algorithm [9]. In the first step, the $R = 1.0$ calorimeter jets are reclustered with the Cambridge-Aachen (C/A) algorithm [10, 11] to combine jet constituents purely according to their angular separation. In the next step the soft-drop algorithm reverses the C/A algorithm clustering history and removes the softer jet constituent at a

specific step of the clustering history unless the following criterion is fulfilled:

$$\frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^\beta, \quad (1)$$

where z_{cut} and β are the parameters of the algorithm, $p_{T,1}$ and $p_{T,2}$ are the transverse momenta of the jet constituents at each step of the clustering history, ΔR_{12} is the angular distance between the jet constituents and R_0 is the jet radius.

A correction to the jet energy, pseudorapidity η and mass is derived from MC simulation to bring the reconstructed jet to the particle jet scale [12]. In addition, residual corrections determined using in situ measurements to bring data in agreement with MC simulation is applied to data.

Track jets are built from charged-particle tracks using the anti- k_t algorithm with a radius parameter of $R = 0.2$. At least two tracks with $p_T > 500$ GeV, $|\eta| < 2.5$ and either associated with the primary vertex or having a longitudinal impact parameter $|z_0 \sin(\theta)| < 3$ mm are required. These requirements reduce the number of tracks from the pile-up vertices, keeping a high selection efficiency for tracks from the hard-scattering vertex. Track jets containing b -hadrons are identified using a multivariate algorithm (MV2c10) [13, 14]. It exploits the information about the jet kinematics, the impact parameter of tracks within the jet and the presence of displaced vertices. A particular requirement on the value of the MV2c10 discriminant corresponds to a particular b -jet selection efficiency calculated from the inclusive p_T and η spectra of jets from an inclusive $t\bar{t}$ sample.

3. Jet substructure observables

The usage of jet substructure observables significantly extended the search program for BSM physics at the LHC. They are also essential to probe the collinear regime of QCD. The jet substructure observables discussed in these proceedings can be categorised as follows:

- Jet shapes use the relative positions and momenta of jet constituents with respect to each other. Some examples are the jet mass and the energy correlation ratios C_2 [15] and D_2 [16, 17], which are useful to identify two-body structures.
- Splitting scales use the clustering history of the jet to define substructures. Commonly used examples are $\sqrt{d_{12}}$ [18] and its mass-normalised form $\sqrt{z_{12}}$ [19].

- Subjettiness observables describe to what degree the substructure of a given jet is compatible with being composed of N or fewer subjets. The subjets are reconstructed from the large- R jet constituents using the k_t algorithm with $R = 0.2$. Often the ratios of N -subjettiness τ_{21}, τ_{32} are used [20].

The soft-drop jet mass is the most precisely calculable jet substructure observable at a hadron collider. The soft-drop mass is calculated at next-to-leading order (NLO) with next-to-leading-logarithms (NLL) and leading order (LO) with NNLL accuracy [21, 22]. The soft-drop mass is the most powerful observable for identifying boosted hadronically decaying massive particles as well. Thus a measurement of the soft-drop jet mass distribution in hadron collisions is of particular importance.

The soft-drop jet mass is measured using 32.9 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ pp data collected in 2016 by the ATLAS detector [23]. For this measurement large- R jets with radius parameter $R = 0.8$ are used. Events were selected online using the jet trigger with the lowest p_T threshold, which is fully efficient for jets with $p_T > 600 \text{ GeV}$. The offline selection requires events with at least two jets, of which the leading- p_T jet has to fulfill the $p_T > 600 \text{ GeV}$ requirement. The dijet topology is ensured by requiring that the leading two jets satisfy $p_{T,1}/p_{T,2} < 1.5$.

The soft-drop algorithm is run on the leading two jets in the selected events. Both jets are used for the measurement. Three different values of the β -parameter (0, 1 and 2) are considered. The value of z_{cut} is set to 0.1. To reduce the dependence of the soft-drop jet mass on the jet p_T , a dimensionless mass $\rho = m^{\text{soft drop}}/p_T^{\text{ungroomed}}$ is used. The ungroomed jet p_T is necessary because the groomed version is not collinear safe for $\beta \neq 0$.

After the event selection, the data are unfolded to correct detector-level (measured outputs of the detector) observables to the particle level (particles that interact with the detector). The particle-level selection is as close as possible to the detector-level selection. MC simulations are used to perform the unfolding applying the iterative Bayesian technique [24]. Various systematic uncertainties impact the soft-drop mass distribution. One of the dominant uncertainties is due to the theoretical modelling of jet fragmentation evaluated by comparing the unfolding procedure performed using the Pythia [25] MC generator to the unfolding performed with the Sherpa [26] and Herwig++ [27, 28] generators. Another dominant uncertainty is the cluster energy scale uncertainty evaluated by matching of isolated calorimeter-cell clusters to tracks and determining the mean of the energy-to-momentum ratio.

Figure 1: The unfolded $\log_{10}(\rho^2)$ distribution in data compared to Pythia, Sherpa, and Herwig++ at the particle level, and NLO+NLL and LO+NNLL theory predictions. The soft-drop algorithm is applied for $\beta = 1$. Open red markers indicate the region, where the nonperturbative (NP) effects are expected to be large but not applied for the LO+NNLL predictions [23].

The unfolded $\log_{10}(\rho^2)$ distribution in data is shown in Figure 1. The data are compared with Pythia, Sherpa, and Herwig++ at the particle level, and NLO+NLL and LO+NNLL theory predictions. The NLO+NLL calculations are available with nonperturbative (NP) corrections based on the average of various MC models with NP effects turned on and off. No NP corrections are available for the LO+NNLO calculations. The distribution is normalised to the integrated cross section measured in the resummation region $-3.7 < \log_{10}(\rho^2) < -1.7$. The measured and predicted shapes agree well in the resummation region, and the data and NLO+NLL prediction agree well at higher values of $\log_{10}(\rho^2)$. At values below -3.7 , nonperturbative effects lead to significantly different predictions between the MC generators and the calculations without NP corrections.

Furthermore, other jet substructure observables briefly introduced in this section are measured in 2016 ATLAS data [33]. Three dedicated event selections are applied to study jets produced by light quarks or gluons, hadronically decaying top quarks and W bosons. The event selection is based on the trimmed large- R jets with $R = 1.0$, but only the leading- p_T jet is considered for this measurement. In addition, the selected large- R jet constructed from the original constituents before the trimming procedure is groomed using the soft-drop algorithm, and the jet substructure observables are constructed from it. The measurement is performed for the trimmed and the soft-drop jet substructure observables.

The dijet topology is ensured by requiring at least

Figure 2: Distribution of N subjects for the dijet selection compared with different MC predictions for soft-drop large- R jets. Sherpa is tested with cluster and string hadronisation models. The shaded bands represent the total uncertainty, the error bars show the statistical uncertainty [33].

two jets with $p_T > 200$ GeV and the leading- p_T jet has to have $p_T > 450$ GeV. For the top-quark and W -boson selections, exactly one muon with $p_T > 30$ GeV is required to target semimuonic $t\bar{t}$ events with geometrically separated decay products of the two top-quark candidates. The hadronically decaying top quark is reconstructed as a large- R jet, which contains the b -tagged small- R jet within the angular distance $\Delta R(\text{large-}R \text{ jet, small-}R \text{ jet}) < 1.0$. Its jet mass has to be above 140 GeV. The hadronically decaying W boson is reconstructed as a large- R jet with the b -tagged small- R jet within $1 < \Delta R(\text{large-}R \text{ jet, small-}R \text{ jet}) < 1.8$ and the jet mass between 60 and 100 GeV. For unfolding, a particle-level selection is defined to be as close as possible to the detector-level selection mentioned above.

The 2016 ATLAS data provide a discrimination between the various MC models probed for this measurement. In general, Pythia8 [29] for light-quark and gluon large- R jet observables, and Powheg+Pythia8 [30], Sherpa and MG5_aMC@NLO+Pythia8 [31] for top quark and W -boson large- R jet observables describe the data better than other MC models. All studied distributions can be found in Ref. [33]. Only three examples for the soft-drop jet substructure observables are shown here. Figure 2 shows the subjet multiplicity inside the large- R jet for the dijet selection. About 62% of events have only one subjet. A semi-hard gluon radiation leads

Figure 3: Distributions of D_2 for the top-quark selections compared with different MC predictions for soft-drop large- R jets. The shaded bands represent the total uncertainty, the error bars show the statistical uncertainty [33].

to the presence of more subjects in other cases. Pythia8 and Sherpa describe the data best. The Herwig7 [32] predictions show significant differences from data, this trend holds for all observables. The conclusions are similar for trimmed and soft-drop jet observables.

Figures 3 and 4 compare the data and the MC predictions for the D_2 observable for the top-quark and the W -boson selections, respectively. The value of D_2 increases as the number of subjects inside the large- R jet increases. Especially in the W -boson selection, the MC models exhibit a shift in the peak of the D_2 distribution compared to data relative to data, suggesting that the models are overestimate the gluon radiation.

The inclusive jet spectrum is measured at $\sqrt{s} = 13$ TeV and predictions of various MC generators are compared with the measurement and to each other in distributions of global event observables and jet substructure observables [34]. Good agreement among the different predictions is observed in the region dominated by the hard-scattering matrix element. Different variations are provided to evaluate uncertainties stemming from the modelling of the perturbative (matrix element and shower scales) and the nonperturbative (hadronisation, parton density functions, multiparton interaction) parts.

Figure 4: Distributions of D_2 for the W-boson selections compared with different MC predictions for soft-drop large- R jets. The shaded bands represent the total uncertainty, the error bars show the statistical uncertainty [33].

4. Jet fragmentation properties

Measurements of qualities related to the formation of jets from high-energy quarks and gluons are important for the tuning of MC models. The charged-particle tracks are used to probe the detailed structure of jets. Jets with transverse momentum between 100 GeV and 2500 GeV are used for this measurement [35]. The fragmentation properties of forward and central jets are studied. The data are unfolded and comparisons are made to predictions of various PS generators.

The observables used for the measurement are briefly described below, a full description can be found in Ref. [35]:

- the charged-particle multiplicity is the number of charged particles inside a jet with p_T above some threshold. The multiplicity is not calculable in perturbative theory, but the ratio of the multiplicity for quark-initiated jets to that for gluon-initiated jets is equal to the ratio of colour factors $C_A/C_F = 9/4$ to first order in $\sqrt{\alpha_S}$. The multiplicity increases with p_T for both quark- and gluon-initiated jets.
- the summed fragmentation function is the momentum fraction of the particle transverse momentum divided by the jet transverse momentum summed over charged particles.
- the relative transverse momentum is

$$p_T^{\text{rel}} = p_T^{\text{charged particle}} \sin \Delta\Phi, \quad (2)$$

Figure 5: The unfolded measured spectrum of charged-particle multiplicity for all selected jets as a function of the jet p_T . The lower panel shows the ratio of various MC predictions to the data, with the total uncertainty band centred on the data at unity [35].

where $\Delta\Phi$ is the angle between the momentum of the charged particle and the jet axis in the transverse plane.

- radial profiles give the number of charged particles in various annuli around the jet axis.

Events were selected using jet triggers. The trigger for a particular jet p_T is chosen to ensure that the trigger is 100% efficient and has the lowest prescale factor. Events in data are weighted by the prescale. The offline selection requires at least two jets with $|\eta| < 2.1$ and the two leading- p_T jets must satisfy $p_T^{\text{lead}}/p_T^{\text{sublead}} < 1.5$. The data are corrected for resolution and acceptance effects using the iterative Bayesian unfolding technique. The jet substructure observables are simultaneously unfolded with the jet p_T and for the central and the forward jets.

The results are presented in two ways: inclusive spectra for both jets together and separate spectra for forward and central jets to determine features of quark-initiated and gluon-initiated jets. The inclusive spectrum of charged-particle multiplicity is shown in Figure 5. The charged-particle multiplicity increases from about 10 at a jet p_T of 100 GeV to over 20 at a jet p_T of 2500 GeV. Pythia 8 and Sherpa predictions bracket the data and are accurate to about 5%. The Herwig++ [36] prediction is between the two and closer to the data except for jets with $p_T < 500$ GeV.

Figure 6: The jet topics extracted using charged-particle multiplicity for jets with p_T between 900 and 1000 GeV together with the topics and quark and gluon distributions from Pythia. The uncertainty bands on the data include both statistical and systematic uncertainties while the error bars indicate statistical uncertainties only [35].

A new data-based approach is used to extract underlying distributions for gluon-initiated and quark-initiated jets without using inputs from MC simulation or theory, the “jet topics” approach. The definition of the first and the second topics can be found in Refs. [37, 38]. The extracted topics using charged-particle multiplicity are shown in Figure 6 for jets with p_T between 900 and 1000 GeV. The very low charged-particle multiplicity region is dominated by quarks and the very high charged-particle multiplicity region is dominated by gluons. Therefore, the first topic is well aligned with quarks and the second topic is more gluon-like. The agreement is better for quarks than for gluons, but the second topic does converge to the gluon distribution at high jet p_T .

Figure 7 shows the charged-particle multiplicity as a function of jet p_T for both extracted quark-like and gluon-like jets as well as the topic extraction along with the prediction for the pure quark and gluon case. Gluon jets from data deviate significantly from predictions at high jet p_T . Quark jets are better described at high p_T , but have a different slope, most prominent at low p_T .

Another measurement of the gluon fragmentation is carried out using the gluon splitting to b -quark pairs [39]. The identified b -tagged small- R jets within large- R jets provide a proxy for the quark daughters of the initial gluon. Events are selected using a single-

Figure 7: The charged-particle multiplicity as a function of jet p_T for quark and gluon jets in data and in predictions from Pythia and perturbative QCD calculations. The predictions are normalised to the data in the sixth p_T bin, since the calculation cannot predict the overall normalisation [35].

jet trigger with p_T threshold of 420 GeV. In the offline selection, jets with $p_T > 450$ GeV and $|\eta| < 2.0$ are required to ensure that the trigger selection is fully efficient. The highest- p_T jet is required to have at least two associated track-jets with $p_T > 10$ GeV and $|\eta| < 2.5$. Only the leading two track-jets are used for subsequent analysis. The leading track-jet must be b -tagged by the MV2c10 algorithm at the 60% efficiency working point.

The kinematic properties of the $g \rightarrow b\bar{b}$ process are characterised by three quantities:

- opening angle between the b -quarks: $\Delta R(b_1, b_2) = \sqrt{\Delta\phi(b_1, b_2)^2 + \Delta\eta(b_1, b_2)^2}$ measured between the two track-jets;
- the momentum sharing between the b -quarks: $z(p_T) = p_{T,2}/(p_{T,1} + p_{T,2})$, where $p_{T,1}$ and $p_{T,2}$ are the transverse momenta of the leading and subleading track-jets, respectively;
- the orientation of the gluon splitting relative to the gluon production plane: $\Delta\theta_{pg,gb}$, which is the angle between the plane spanned by the beam line and the vector sum of the two track-jets.

In addition to these quantities, the dimensionless mass $\log(m_{bb}/p_T)$ is also measured, where the mass and p_T are computed from the four-vector sum of the two track-jets.

Figure 8: The unfolded distribution of $\Delta\theta_{ppg,gb\bar{b}}$ compared with predictions from the Pythia and Sherpa simulations. Error bands represent the sum in quadrature of statistical and systematic uncertainties. The bands for the Pythia prediction represented by a square indicate the Var2 variations (dominated by a 10% variation in the final state shower α_s). The additional Pythia prediction with the gluon polarisation turned off uses $m_{b\bar{b}}^2/4$ for the renormalisation scale [39].

After the event selection, the contribution from large- R jets that do not have two associated track-jets containing b -hadrons is subtracted from data before correcting for detector effects. The normalisation of this background is determined using data template fits to the impact parameter distribution. In each bin of the target observable distributions, the distribution of the signed impact parameter significance is fitted to data using templates from simulation while letting the fraction of each jet flavour component float in the fit. Due to the long lifetime of b -hadrons, the values of the signed impact parameter significance tend to be larger than those for tracks in c -jets and light-flavour jets. Therefore, this distribution can be used to extract the fractions of b -jets, c -jets and light-flavour jets.

After subtracting the background from the detector-level distributions, the data are corrected for resolution and acceptance effects using the iterative Bayesian unfolding technique and compared to MC simulation predictions. In general, the Sherpa predictions are more accurate than those from Pythia. Although significant differences between both generators and the data are observed at low mass, low $z(p_T)$ and for all $\Delta\theta_{ppg,gb\bar{b}}$ distributions. The $\Delta\theta_{ppg,gb\bar{b}}$ observable is shown in Figure 8. The data distribution is inverted with respect to the one

from Pythia, showing a minimum instead of a maximum at $\pi/2$. Sherpa predicts a relatively uniform distribution. For comparison, a version of Pythia with the gluon polarisation turned off is also shown in this figure. This version is closer to Sherpa and to the data. In general, the properties of gluon polarisation inside unpolarised hadrons is largely unconstrained by experimental data. This and future measurements of $\Delta\theta_{ppg,gb\bar{b}}$ may provide a new way to extract p_T -dependent parton distributions in order to better understand the proton structure.

5. Conclusions

These proceedings summarise some measurements of jet substructure observables and jet fragmentation properties performed with 33 fb^{-1} of proton-proton LHC collision data collected with the ATLAS detector at $\sqrt{s} = 13 \text{ TeV}$.

The measurement of the soft-drop jet mass shows that the theoretical calculations agree well with the data in the perturbative regions. In the regions where the non-perturbative effects are expected to be significant, the MC simulation predictions agree better with the data than the predictions from the calculations. The soft-drop jet mass is the first jet substructure observable for which calculations are available at the NNLL accuracy.

The measurements of other jet substructure variables are compared with MC predictions as well. Pythia8 for the light-quark- and gluon-jet observables and Powheg+Pythia8, Sherpa and MG5_aMC@NLO+Pythia8 for top-quark and W -boson jet observables describe the data better than other models. For most of the observables Herwig7 in the dijet selection and Powheg+Herwig7 in the top-quark and W -boson selections show the largest disagreement between the data and the predictions.

The track-based jet fragmentation properties including the charged-particle multiplicity, the momentum fraction carried by charged particles, and angular properties of the radiation pattern inside jets are measured. The radial profile is systematically broader in the data than in the simulation. Other variables are well modelled in general within the precision of the measurement. A first measurement of jet topics for the charged-particle multiplicity provides a promising way to extract the underlying distributions for gluon-initiated and quark-initiated jets without using inputs from MC simulation or theory.

Various properties of the $g \rightarrow b\bar{b}$ process have been measured and compared with MC predictions. A flavour-fraction fit is used to remove contributions from

processes other than $g \rightarrow b\bar{b}$. The measured properties are unfolded and compared to the Sherpa and Pythia predictions. Sherpa simulations provide a better model than Pythia, especially for the $\Delta\theta_{pg,gb}$ observable sensitive to the gluon polarisation.

The measured particle-level spectra are publicly available and can be used for further interpretations to validate MC predictions and tune the MC models.

References

- [1] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, 2008 JINST 3 S08003.
- [2] M. Cacciari, G. P. Salam and G. Soyez, FastJet User Manual, Eur. Phys. J. C, 72 (2012) 1896.
- [3] M. Cacciari, G. P. Salam and G. Soyez, The anti- k_t jet clustering algorithm, JHEP 04 (2008) 063.
- [4] ATLAS Collaboration, Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1, Eur. Phys. J. C 77 (2017) 490.
- [5] ATLAS Collaboration, Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector, JHEP 09 (2013) 076.
- [6] D. Krohn, J. Thaler and L.-T. Wang, Jet Trimming, JHEP 02 (2010) 084.
- [7] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Longitudinally invariant k_t clustering algorithms for hadron hadron collisions, Nucl. Phys. B 406 (1993) 187.
- [8] S. D. Ellis and D. E. Soper, Successive combination jet algorithm for hadron collisions, Phys. Rev. D 48 (1993) 3160.
- [9] A. J. Larkoski, S. Marzani, G. Soyez, J. Thaler, Soft Drop, JHEP 05 (2014) 146.
- [10] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, Better jet clustering algorithms, J. High Energy Phys. 08 (1997) 001.
- [11] M. Wobisch and T. Wengler, Hadronization corrections to jet cross-sections in deep inelastic scattering, arXiv:hep-ph/9907280.
- [12] ATLAS Collaboration, In situ calibration of large-radius jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector, Eur. Phys. J. C 79 (2019) 135.
- [13] ATLAS Collaboration, Performance of b-jet identification in the ATLAS experiment, JINST 11 (2016) P04008.
- [14] ATLAS Collaboration, Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run, ATL-PHYS-PUB-2016-012, 2016.
- [15] A. J. Larkoski, G. P. Salam and J. Thaler, Energy correlation functions for jet substructure, JHEP 06 (2013) 108.
- [16] A. J. Larkoski, J. Thaler and W. J. Waalewijn, Gaining (mutual) information about quark/gluon discrimination, JHEP 11 (2014) 129.
- [17] A. J. Larkoski, I. Moult and D. Neill, Analytic boosted boson discrimination, JHEP 05 (2016) 117.
- [18] J. M. Butterworth, B. E. Cox and J. R. Forshaw, WW scattering at the CERN LHC, Phys. Rev. D 65 (2002) 096014.
- [19] J. Thaler and L.-T. Wang, Strategies to Identify Boosted Tops, JHEP 07 (2008) 092.
- [20] J. Thaler and K. Van Tilburg, Identifying boosted objects with N-subjettiness, JHEP 03 (2011) 015.
- [21] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, Factorization for groomed jet substructure beyond the next-to-leading logarithm, J. High Energy Phys. 07 (2016) 064.
- [22] C. Frye et al., Precision physics with pile-up insensitive observables, arXiv:1603.06375.
- [23] ATLAS Collaboration, Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector, Phys. Rev. Lett. 121 (2018) 092001.
- [24] G. D'Agostini, A multidimensional unfolding method based on Bayes theorem, Nucl. Instrum. Meth. A 362 (1995) 487.
- [25] T. Sjöstrand, S. Mrenna, and P. Z. Skands, A brief introduction to PYTHIA 8.1, Comput. Phys. Commun. 178 (2008) 852.
- [26] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, Event generation with SHERPA 1.1, J. High Energy Phys. 02 (2009) 007.
- [27] M. Bahr et al., HERWIG++ physics and manual, Eur. Phys. J. C 58, 639 (2008).
- [28] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes), J. High Energy Phys. 01 (2001) 010.
- [29] T. Sjöstrand et al., An introduction to PYTHIA 8.2, Comput. Phys. Commun. 191 (2015) 159.
- [30] S. Alioli, P. Nason, C. Oleari and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, JHEP 06 (2010) 043.
- [31] J. Alwall et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, JHEP 07 (2014) 079.
- [32] J. Bellm et al., Herwig 7.0/Herwig++ 3.0 release note, Eur. Phys. J. C 76 (2016) 196.
- [33] ATLAS Collaboration, Measurement of jet-substructure observables in top quark, W boson and light jet production in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, JHEP 08 (2019) 033.
- [34] ATLAS Collaboration, Multijet simulation for 13 TeV ATLAS Analyses, ATL-PHYS-PUB-2019-017 (2019), <http://cdsweb.cern.ch/record/2672252>.
- [35] ATLAS Collaboration, Properties of jet fragmentation using charged particles measured with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV, submitted to Phys. Rev. D, arXiv:1906.09254 (2019).
- [36] M. Bähr et al., Herwig++ physics and manual, Eur. Phys. J. C 58 (2008) 639.
- [37] E. M. Metodiev and J. Thaler, Jet Topics: Disentangling Quarks and Gluons at Colliders, Phys. Rev. Lett. 120 (2018) 241602.
- [38] P. T. Komiske, E. M. Metodiev and J. Thaler, An operational definition of quark and gluon jets, JHEP 11 (2018) 059.
- [39] Properties of $g \rightarrow b\bar{b}$ at small opening angles in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV, Phys. Rev. D 99 (2019) 052004.