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

The ATLAS Collaboration

DOI
10.1007/JHEP08(2019)033

Publication date
2019

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):
https://doi.org/10.1007/JHEP08(2019)033
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

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A measurement of jet substructure observables is presented using data collected in 2016 by the ATLAS experiment at the LHC with proton-proton collisions at $\sqrt{s} = 13$ TeV. Large-radius jets groomed with the trimming and soft-drop algorithms are studied. Dedicated event selections are used to study jets produced by light quarks or gluons, and hadronically decaying top quarks and $W$ bosons. The observables measured are sensitive to substructure, and therefore are typically used for tagging large-radius jets from boosted massive particles. These include the energy correlation functions and the $N$-subjettiness variables. The number of subjets and the Les Houches angularity are also considered. The distributions of the substructure variables, corrected for detector effects, are compared to the predictions of various Monte Carlo event generators. They are also compared between the large-radius jets originating from light quarks or gluons, and hadronically decaying top quarks and $W$ bosons.

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1903.02942
1 Introduction

Increasing the centre-of-mass energy of proton-proton (pp) collisions from 7 and 8 TeV in Run 1 to 13 TeV in Run 2 of the Large Hadron Collider (LHC) leads to a larger fraction of heavy particles such as top quarks, vector bosons and Higgs bosons being produced with large transverse momenta. This large transverse momentum leads to collimated decay products. They are usually reconstructed in a large-radius jet, whose internal (sub)structure shows interesting features that can be used to identify the particle that initiated the jet formation [1, 2].

This is relevant for a host of measurements and searches, which involve identifying the large-radius jets coming from top quarks [3–7], or Higgs bosons [8–11], for example in Run 2 in ATLAS. Usually a two step procedure is employed. In the first step, termed grooming, the effect of soft, uncorrelated radiation contained in the large-radius jet is reduced. Then jet substructure observables, which describe the spatial energy distribution inside the jets,
are used to classify the jets originating from different particles. This process is called jet tagging and the algorithms are referred to as taggers.

Most of the grooming algorithms and jet substructure observables were developed on the basis of theoretical calculations or Monte Carlo (MC) simulation programs and then they are applied to data. Given that often large differences have been seen between predictions from MC and data, large correction factors need to be applied to simulation results. Additionally, taggers suffer from large systematic uncertainties as the modelling of the substructure observables is not well constrained [2, 12]. Most of these variables have never been measured in data, and performing a proper unfolded measurement is a common request from the theory community. Measuring these observables will help in optimising and developing current and future substructure taggers, as well as tuning hadronization models in the important but still relatively unexplored regime of jet substructure. The choice of variables measured in this paper prioritized jet shapes commonly used in jet tagging, as well as those most useful for model tuning.

The ATLAS Collaboration has performed measurements of jet mass and substructure variables at the $pp$ centre-of-mass energies of $\sqrt{s} = 7, 8$ and 13 TeV [13–19] in inclusive jet events, and the CMS Collaboration has performed measurements of jet mass and substructure in dijet, $W/Z$ boson, and $tt$ events [20–24] at $\sqrt{s} = 7, 8$ and 13 TeV. This paper presents measurements of substructure variables in large-radius jets produced in inclusive multijet events and in $tt$ events at $\sqrt{s} = 13$ TeV using 33 fb$^{-1}$ of data collected in 2016 by the ATLAS experiment. In this analysis, the lepton+jets decay mode of $tt$ events is selected, where one $W$ boson decays into a muon and a neutrino, and the other $W$ boson decays into a pair of quarks. Then the large-radius jets are separated into those that contain all the decay products of a hadronically top quark and those containing only hadronic $W$ boson decay products.

The contents of this paper are organised as follows. First, a description of the ATLAS detector is presented in section 2 and then the MC samples used in the analysis are discussed in section 3. In section 4, event and object selections are summarised. The measured jet substructure observables are defined in section 5. The background estimation is described in section 6 and the systematic uncertainties are assessed in section 7. In section 8, detector-level mass and $p_T$ distributions corresponding to selected large-radii jets are shown, and the unfolding is described in section 9. Finally, the unfolded results are presented in section 10, and the conclusions in section 11.

### 2 ATLAS detector

The ATLAS experiment uses a multipurpose particle detector [25, 26] with a forward-backward symmetric cylindrical geometry and a near 4$\pi$ coverage in solid angle.\footnote{ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upwards. Cylindrical coordinates ($r, \phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. An angular separation between two objects is defined as $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$, where $\Delta \eta$ and $\Delta \phi$ are the separations in $\eta$ and $\phi$. Momentum in the transverse plane is denoted by $p_T$.} It con-
sists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic (EM) and hadron calorimeters, and a muon spectrometer. The ID consists of silicon pixel, silicon microstrip, and straw-tube transition-radiation tracking detectors, covering the pseudorapidity range $|\eta| < 2.5$. The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Electromagnetic calorimetry is performed with barrel and endcap high-granularity lead/liquid-argon (LAr) sampling calorimeters, within the region $|\eta| < 3.2$. There is an additional thin LAr presampler covering $|\eta| < 1.8$, to correct for energy loss in material upstream of the calorimeters. For $|\eta| < 2.5$, the LAr calorimeters are divided into three layers in depth. Hadronic calorimetry is performed with a steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters, which cover the region $1.5 < |\eta| < 3.2$. The forward solid angle up to $|\eta| = 4.9$ is covered by copper/LAr and tungsten/LAr calorimeter modules, which are optimised for energy measurements of electrons/photons and hadrons, respectively. The muon spectrometer consists of separate trigger and high-precision tracking chambers that measure the deflection of muons in a magnetic field generated by superconducting air-core toroids.

The ATLAS detector selects events using a tiered trigger system [27]. The first level is implemented in custom electronics. The second level is implemented in software running on a general-purpose processor farm which processes the events and reduces the rate of recorded events to 1 kHz.

3 Monte Carlo samples

Simulated events are used to optimise the event selection, correct the data for detector effects and estimate systematic uncertainties. The predictions of different phenomenological models implemented in the Monte Carlo (MC) generators are compared with the data corrected to the particle level (i.e. observables constructed from final-state particles within the detector acceptance).

The generators used to produce the samples are listed in table 1. The dijet (to obtain multijet events), $t\bar{t}$ and single-top-quark samples are considered to be signal processes in this analysis, corresponding to the dedicated selections. The background is estimated using $Z/W +$jets and diboson samples. The $t\bar{t}$ samples are scaled to next-to-next-to-leading order (NNLO) in perturbative QCD, including soft-gluon resummation to next-to-next-to-leading-log order (NNLL) [28] in cross-section, assuming a top quark mass $m_t = 172.5$ GeV. The POWHEG model [29] resummation damping parameter, $h_{\text{damp}}$, which controls the matching of matrix elements to parton showers and regulates the high-$p_T$ radiation, was set to $1.5m_t$ [30]. The single-top-quark [31–36] and $W/Z$ samples [37] are scaled to the NNLO theoretical cross-sections.

The predicted shape of jet substructure distributions depends on the modelling of final-state radiation (FSR), and fragmentation and hadronisation, as well as on the merging/matching between matrix element (ME) and parton shower (PS) generators. The PYTHIA8 and the SHERPA generators use a dipole shower ordered in transverse momentum, with the Lund string [38] and cluster hadronisation model [39] respectively. The
Table 1. Main features of the Monte Carlo models used to simulate signal and background samples, and to produce predictions to be compared with data. The nominal samples listed are used for comparisons with corrected data at particle level as well. For convenience, MG5_aMC@NLO is referred to as MG5_aMC in figures 3–9.

The MC samples were processed through the full ATLAS detector simulation [54] based on Geant4 [55], and then reconstructed and analysed using the same procedure and software that are used for the data. Additional pp collisions generated by PYTHIA8, with parameter values set to the A2 tune [56] and using the MSTW2008 [57] PDF set, were overlaid to simulate the effects of additional collisions from the same and nearby bunch crossings (pile-up), with a distribution of the number of extra collisions matching that of data.

4 Object and event selection

This analysis uses pp collision data at √s = 13 TeV collected by the ATLAS detector in 2016, that satisfy a number of criteria to ensure that the ATLAS detector was in good operating condition. All selected events must have at least one vertex with at least two associated tracks with pT > 400 MeV. The vertex with the highest ∑pT,track, where pT,track is the transverse momentum of a track associated with the vertex, is chosen as the primary vertex.

Jets are reconstructed from the EM-scale or locally-calibrated topological energy clusters [58] in both the EM and hadronic calorimeters using the anti-kT algorithm [59] with a radius parameter of R = 0.4 or R = 1.0, referred to as small-radius and large-radius jets respectively. These clusters are assumed to be massless when computing the jet four-vectors.
and substructure variables. A trimming algorithm [60] is employed for the large-radius jets to mitigate the impact of initial-state radiation, underlying-event activity, and pile-up. Trimming removes subjets of radius $R_{\text{sub}} = 0.2$ with $p_{T, i}/p_{T,J} < f_{\text{cut}}$, where $p_{T,i}$ is the transverse momentum of the $i$th subjet, $p_{T,J}$ is the transverse momentum of the jet under consideration, and $f_{\text{cut}} = 0.05$. All large-radius jets used in this paper are trimmed before applying the selection criteria. The energies of jets are calibrated by applying $p_T$- and rapidity-dependent corrections derived from Monte Carlo simulation with additional correction factors for residual non-closure in data determined from data [58, 61].

In order to reduce the contamination by small-radius jets originating from pile-up, a requirement is imposed on the output of the Jet Vertex Tagger (JVT) [62]. The JVT algorithm is a multivariate algorithm that uses tracking information to reject jets which do not originate from the primary vertex, and is applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$. Small-radius jets containing $b$-hadrons are tagged using a neural-network-based algorithm [63–65] that combines information from the track impact parameters, secondary vertex location, and decay topology inside the jets. The operating point corresponds to an overall 70% $b$-tagging efficiency in simulated $t\bar{t}$ events, and to a probability of mis-tagging light-flavour jets of approximately 1%.

Muons are reconstructed from high-quality muon spectrometer track segments matched to ID tracks. Muons with a transverse momentum greater than 30 GeV and within $|\eta| < 2.5$ are selected if the associated track has a longitudinal impact parameter $|z_0 \sin(\theta)| < 0.5$ mm and a transverse impact parameter significance $|d_0|/\sigma(d_0) < 3$. The impact parameter $d_0$ is measured relative to the beam line. The muon candidates are also required to be isolated from nearby hadronic activity [66]. The muon isolation criteria remove muons that lie a distance $\Delta R(\mu, \text{jet}) < 0.04 + 10 \text{ GeV}/p_{T,\mu}$ from a small-radius jet axis, where $p_{T,\mu}$ is the $p_T$ of the muon. Since muons deposit energy in the calorimeters, an overlap removal procedure is applied in order to avoid double counting of leptons and small-radius jets.

Electrons are reconstructed from energy deposits measured in the EM calorimeter which are matched to ID tracks. They are required to be isolated from nearby hadronic activity by using a set of $p_T$- and $\eta$-dependent criteria based on calorimeter and track information as described in ref. [67]. Their selection also requires $p_T > 30$ GeV and $|\eta| < 2.5$, excluding the region $1.37 < |\eta| < 1.52$ which corresponds to the transition region between the barrel and end-cap calorimeters. Photon candidates are reconstructed from clusters of energy deposited in the EM calorimeter, and must have $p_T > 30$ GeV and $|\eta| < 2.5$. Photon identification is based primarily on shower shapes in the calorimeter [68].

The missing transverse momentum, with magnitude $E_{\text{miss}}^T$, is calculated as the negative vectorial sum of the transverse momenta of calibrated photons, electrons, muons and jets associated with the primary vertex [69]. The transverse mass of the leptonically decaying $W$ boson, $m_W^T$, is defined using the absolute value of $E_{T}^\text{miss}$ as $m_W^T = \sqrt{2p_{T,\mu}E_{T}^\text{miss}(1 - \cos \Delta \phi(\mu, E_{T}^\text{miss}))}$.

In order to examine large-radius jets originating from light quarks and gluons, from top quarks and from $W$ bosons, three event selections are defined. These are referred to as dijet, top and $W$ selections, and are indicative of the origin of the large-radius jet.
In the dijet selection, the events are accepted by a single-large-radius-jet trigger that becomes fully efficient for jets with $p_T > 400$ GeV. The offline dijet selection requires a leading trimmed large-radius jet with $p_T > 450$ GeV and $|\eta| < 1.5$, and at least one other trimmed large-radius jet with $p_T > 200$ GeV and $|\eta| < 2.5$, and rejects the event if an electron or muon is present.

For both the top and $W$ selections, events are collected with a set of single-muon triggers that become fully efficient for muon $p_T > 28$ GeV. The top quarks and the $W$ bosons are identified from their decay products. A geometrical separation between the decay products of the two top quark candidates is required. Additional requirements are applied to separate large-radius jets containing all decay products of the top quark from those where the large-radius jet only contains the hadronic $W$ boson decays, with the $b$-tagged small-radius jet reconstructed independently. These form the top selection and the $W$ selection respectively. The selections are described in table 2. After these requirements the data sample contains about $3.2 \times 10^7$ events in the dijet selection, and roughly 6800 and 4500 events in the top and $W$ selection respectively.

Particle-level observables in Monte Carlo simulation are constructed from stable particles, defined as those with proper lifetimes $c\tau \gtrsim 10$ mm. Muons at particle level are dressed by including contributions from photons with an angular distance $\Delta R < 0.1$ from the muon. Particle-level jets do not include muons or neutrinos. Particle-level $b$-tagging is performed by requiring a prompt $b$-hadron to be ghost-associated [70] with the jet.

5 Definition of the jet observables

All large-radius jets are trimmed before being used in the selections, and subsequently only the leading trimmed large-radius is considered in the analysis. Then the large-radius jet constructed from the original constituents of the selected jet before the trimming step is groomed using the soft-drop algorithm, and the jet substructure observables studies are constructed from that soft-dropped large-radius jet.

Soft-drop [71, 72] is an extension of the original split-filtering technique [73] and relies on reclustering the jet constituents using the angle-ordered Cambridge-Aachen jet algorithm and then sequentially considering each splitting in order to remove soft and wide-angle radiation. At each step the jet is split into two proto-jets. The removal of proto-jets in a splitting is controlled by two parameters: a measure of the energy balance of the pair, $z_{cut}$, and the significance of the angular separation of the proto-jets, $\beta_{SD}$. These are used to define the soft-drop condition:

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \left( \frac{\Delta R_{12}}{R} \right)^{\beta_{SD}}$$

where $R_{12}$ is the angular distance between the two proto-jets and $R$ is the radius of the large jet. In this analysis, values of $z_{cut} = 0.1$ and $\beta_{SD} = 0.0$ are used, based on previous ATLAS studies [18], which is equivalent to modified mass drop tagger [74]. An important feature of soft-drop is that groomed observables are analytically calculable to high-order resummation accuracy [75–77].
Table 2. Summary of object event selections for detector-level and particle-level dijet and ℓt events. “Leptonic top” refers to the top quark that decays into a leptonically decaying W boson, while “b-tagged jet” refers to small-radius jets that pass a b-tagging requirement. The top and W selections are common up to the requirement on the leptonic top, then they differ on the requirements on the leading-p_T trimmed large-radius jet. All selections are inclusive, unless otherwise mentioned.

The following substructure variables are measured in this analysis:

- Number of subjets with p_T > 10 GeV, reconstructed from the selected large-radius jet constituents using the k_t algorithm [78] with R = 0.2.

- Generalised angularities defined as:

  \[ \lambda_{\beta}^{\kappa,\text{LHA}} = \sum_{i \in J} z_i^{\kappa} \theta_i^{\beta \text{LHA}}, \]

where z_i is the transverse momentum of jet constituent i as a fraction of the scalar sum of the p_T of all constituents and \( \theta_i \) is the angle of the i^{th} constituent relative to the jet axis, normalised by the jet radius. The exponents \( \kappa \) and \( \beta \text{LHA} \) probe different aspects of the jet fragmentation. The (\( \kappa = 1, \beta \text{LHA} = 0.5 \)) variant is termed the Les Houches angularity (LHA) [79] and used in this analysis. It is an infrared-safe version of the jet-shape angularity, and provides a measure of the broadness of a jet.
Energy correlation functions ECF2 and ECF3 \cite{80}, and related ratios $C_2$, $D_2$ \cite{81}. The 1-point, 2-point and 3-point energy correlation functions for a jet $J$ are given by:

$$ECF1 = \sum_{i \in J} p_T_i,$$

$$ECF2(\beta_{ECF}) = \sum_{i < j \in J} p_T_i p_T_j (\Delta R_{ij})^{\beta_{ECF}},$$

$$ECF3(\beta_{ECF}) = \sum_{i < j < k \in J} p_T_i p_T_j p_T_k (\Delta R_{ij} \Delta R_{ik} \Delta R_{jk})^{\beta_{ECF}},$$

where the parameter $\beta_{ECF}$ weights the angular separation of the jet constituents. In the above functions, the sum is over the $i$ constituents in the jet $J$, such that the 1-point correlation function ECF1 is approximately the jet $p_T$. Likewise, if one takes $\beta_{ECF} = 2$, the 2-point correlation functions scale as the mass of a particle undergoing a two-body decay in collider coordinates. In this analysis, $\beta_{ECF} = 1$ is used, and for brevity, $\beta_{ECF}$ is not explicitly mentioned hereafter.

The ratios of some of these quantities (written in an abbreviated form) are defined as:

$$e_2 = \frac{ECF2}{(ECF1)^2}, \quad e_3 = \frac{ECF3}{(ECF1)^3},$$

The observables $e_2$ and $e_3$ are measured, and are later referred to as $ECF2^{\text{norm}}$ and $ECF3^{\text{norm}}$. These ratios are then used to generate the variable $C_2$ \cite{80}, and its modified version $D_2$ \cite{79,81}, which have been shown to be particularly useful in identifying two-body structures within jets \cite{82}. The $C_2$ and $D_2$ variables as defined below are measured in this analysis:

$$C_2 = \frac{e_3}{(e_2)^2}, \quad D_2 = \frac{e_3}{(e_2)^3}.$$

Ratios of $N$-subjettiness \cite{83}, $\tau_{21}$ and $\tau_{32}$. The $N$-subjettiness describes to what degree the substructure of a given jet is compatible with being composed of $N$ or fewer subjets.

In order to calculate $\tau_N$, first $N$ subjet axes are defined within the jet by using the exclusive $k_t$ algorithm, where the jet reconstruction continues until a desired number of jets are found. The 0-, 1-, 2-, and 3-subjettiness are defined as:

$$\tau_0(\beta_{NS}) = \sum_{i \in J} p_T_i R^{\beta_{NS}}, \quad (5.1a)$$

$$\tau_1(\beta_{NS}) = \frac{1}{\tau_0(\beta_{NS})} \sum_{i \in J} p_T_i \Delta R^{\beta_{NS}}_{a_1,i}, \quad (5.1b)$$

$$\tau_2(\beta_{NS}) = \frac{1}{\tau_0(\beta_{NS})} \sum_{i \in J} p_T_i \min(\Delta R^{\beta_{NS}}_{a_1,i}, \Delta R^{\beta_{NS}}_{a_2,i}), \quad (5.1c)$$

$$\tau_3(\beta_{NS}) = \frac{1}{\tau_0(\beta_{NS})} \sum_{i \in J} p_T_i \min(\Delta R^{\beta_{NS}}_{a_1,i}, \Delta R^{\beta_{NS}}_{a_2,i}, \Delta R^{\beta_{NS}}_{a_3,i}), \quad (5.1d)$$
where $\Delta R$ is the angular distance between constituent $i$ and the jet axis, $a_i$, and $\Delta R_{a,n}$ is the angular distance between constituent $i$ and the axis of the $n^{th}$ subjet. The term $R$ in equation (5.1a) is the radius parameter of the jet. The parameter $\beta^{NS}$ gives a weight to the angular separation of the jet constituents. In the studies presented here, the value of $\beta^{NS} = 1$ is used. In the above functions, the sum is performed over the constituents $i$ in the jet $J$, and a normalisation factor $\tau_0$ (eq. (5.1a)) is used. The ratios of the $N$-subjettiness functions, $\tau_{21} = \tau_2/\tau_1$ and $\tau_{32} = \tau_3/\tau_2$ have been shown to be particularly useful in identifying two-body and three-body structures within jets.

Studies presented in ref. [84] have shown that an alternative axis definition can increase the discrimination power of these variables. The winner-takes-all (WTA) axis uses the direction of the hardest constituent in the subjet obtained from the exclusive $k_t$ algorithm instead of the subjet axis, such that the distance measure $\Delta R_{a,1;i}$ changes in the calculation. In this analysis, the same observables calculated with the WTA axis definition, $\tau_{21}^{WTA}$ and $\tau_{32}^{WTA}$, are used.

6 Data-driven background estimation

The largest non-$t\bar{t}$ contributions to the $W$ and top selections come from the $W$+jets and single-top processes. Additionally non-prompt and mis-reconstructed muons are a separate source of background for the top and $W$ selections. Contributions from other processes were considered and found to be negligible. A data-driven method, following ref. [85], is used to estimate the contribution from the $W$+jets process while the single-top process is considered part of the signal.

At the LHC the production rate of $W^+\text{+jets}$ events is larger than that of $W^-\text{+jets}$ due to the higher density of $u$-quarks than $d$-quarks in the proton. This results in more events with positively charged leptons. Other processes do not contribute significantly to this charge asymmetry. The data are used to derive scale factors that correct the normalisation and flavour fraction given by the MC simulation [86].

Normalisation scale factors are determined by comparing the charge asymmetry in data with the asymmetry estimated by simulation. Contributions to the asymmetry from other processes are estimated by simulation and subtracted. A selection that contains the full top and $W$ selection criteria without any $b$-tagging requirements is initially used. The total number of $W$+jets events in data, $N_{W+} + N_{W-}$, is given by

$$N_{W+} + N_{W-} = \left( \frac{r_{MC} + 1}{r_{MC} - 1} \right) (D_+ - D_-)$$

where $r_{MC}$ is the ratio of the number of events with positive muons to the number of events with negative muons obtained from the MC simulation while $D_+$ and $D_-$ are the number of events with positive and negative muons in data, respectively, after using simulation to subtract the estimated background contribution of all processes other than $W$+jets. From the above equation the scale factor $C_A$ is extracted which is defined as the ratio of $W$+jets
events evaluated from data to the number predicted by the simulation

\[ C_A = \left( \frac{r_{MC} + 1}{r_{MC} - 1} \right) (D_+ - D_-) \cdot \frac{1}{N_{MC}^W} \]

where \( N_{MC}^W \) is the predicted number of \( W+\text{jets} \) events.

Scale factors correcting the relative fractions of \( W \) bosons produced in association with jets of different flavour are also estimated using data. The fractions of \( W+b\bar{b}, W+c\bar{c}, W+c \) and \( W+\text{light-quark} \) events are initially estimated from simulation in a selection without the \( b\)-tagging requirements, which corresponds to the selection mentioned in table 2 without the \( \Delta R \) requirement imposed during the top and \( W \) selections. A system of three equations is used to fit the fractions estimated from simulation to the selection with full \( b\)-tagging requirements:

\[
\begin{pmatrix}
(C_A(N_{bb}^- + N_{cc}^+) - C_A N_c^- & C_A N_{light}^-)
\end{pmatrix} \cdot \begin{pmatrix}
K_{bb,cc} \\
K_c \\
K_{light}
\end{pmatrix} = \begin{pmatrix}
D_{W^-} \\
1 \\
D_{W^+}
\end{pmatrix},
\tag{6.1}
\]

where \( f_{bb}, f_{cc}, f_c \) and \( f_{light} \) are flavour factors estimated from simulation while \( K_{bb}, K_{cc}, K_c \) and \( K_{light} \) are the respective correction factors. The corresponding number of events estimated by simulation with positive (negative) leptons are given by \( N_{bb}^{+,(-)}, N_{cc}^{+,(-)}, N_c^{+,(-)} \) and \( N_{light}^{+,(-)} \). The terms \( D_{W^\pm} \) are the expected numbers of \( W+\text{jets} \) events with positively or negatively charged leptons in the data. An iterative process is used to find the \( K_{\text{flavour}} \) correction factors which are used to correct the associated \( f_{\text{flavour}} \) fractions used in the calculation of \( C_A \). The correction factors are determined by inverting eq. (6.1) and then the process is repeated with a new \( C_A \) calculated using the corrected flavour fractions. This process is repeated 10 times and further iterations produce negligible changes in \( C_A \).

This process is repeated individually for all variables in the top and \( W \) selections since, depending on the substructure of the selected large-radius jet, events can fall out of the acceptance for a subset of the variables. The final calculated scale factors are, however, consistent across both selections and all variables. These scale factors are \( 0.84\pm0.02 \), where the uncertainty is statistical, and the overall contribution to the final selections is shown in table 3. In order to determine the uncertainty in the shape of the subtracted \( W+\text{jets} \) distribution, the contribution from an alternative MC generator (MG5_aMC@NLO+PYTHIA8 as opposed to default SHERPA) was used. Both MC samples were scaled to the estimated number of events and the envelope of the shape difference was taken as an uncertainty.

There is also a contribution from events where a jet is misreconstructed as a muon or when a non-prompt muon is misidentified as a prompt muon which satisfies the selection criteria. This contribution is estimated using the matrix method, comparing the yields of muons and non-prompt muons that pass a loose selection with the yields of those that pass a tight selection. The efficiency for real muon selection \( (\varepsilon_{\text{real}}) \) is measured using a tag-and-probe method with muons from \( Z \rightarrow \mu\mu \) events. The efficiency for misreconstructed muon selection \( (\varepsilon_{\text{fake}}) \) is measured in control regions dominated by background from multijet processes, after using simulation to subtract the contribution of other processes. Event weights
<table>
<thead>
<tr>
<th>Background</th>
<th>Top selection</th>
<th>W selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Percent contributions)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.0 ± 0.1</td>
<td>2.6 ± 0.1</td>
</tr>
<tr>
<td>Misreconstructed and non-prompt muons</td>
<td>6.6 ± 0.1</td>
<td>5.5 ± 0.1</td>
</tr>
</tbody>
</table>

Table 3. Contributions from background processes which are subtracted in the top and $W$ selections. The uncertainties are statistical only.

are computed using the above efficiencies, which are parameterised in the kinematics of the event. The weight for event $i$, where the muons satisfy the loose criteria, is given by

$$w_i = \frac{\varepsilon_{\text{fake}}}{\varepsilon_{\text{real}} - \varepsilon_{\text{fake}}}(\varepsilon_{\text{real}} - \delta_i)$$

where $\delta_i$ equals unity if the muon in event $i$ satisfies the tight criteria and zero otherwise. The background estimate in a given bin is therefore the total sum of weights in that bin. The estimated contributions to the yield from misreconstructed or non-prompt muons for the top and $W$ selections are shown in table 3. These corrections have very little effect on the shape of the distributions considered.

7 Systematic uncertainties

7.1 Large-radius jet uncertainties

As jets are built from topological clusters reconstructed in the calorimeter, systematic uncertainties in the jet substructure observables are calculated using a bottom-up approach applied to the clusters forming each jet [18]. The following components of the uncertainty are considered:

- Cluster reconstruction efficiency (CE): accounts for low energy particles that fail to seed a cluster based on the fraction of inner-detector tracks matched to no clusters in low $\mu$ data. The uncertainty is the observed difference between simulation and data. Since the efficiency reaches 100% for cluster energy above 2.5 GeV, no uncertainty is assumed above this value.

- Cluster energy scale variation (CESu/CESd): the cluster energy scale is determined by studying clusters matched to isolated tracks in data events with low pile-up. A fit of the $E/p$ distribution is used to extract an overall energy scale. The uncertainty in the scale is given by taking the difference of the ratio of the scales calculated in data and simulation from unity. Clusters are independently scaled up and down and the resulting variations in observables are added in quadrature.

- Cluster energy smearing (CES): the difference in quadrature of the width of the $E/p$ distribution measured in data and given by simulation is defined as the uncertainty in the energy resolution. The cluster energies are smeared by this value and the effect on the observables is taken as an uncertainty.
Cluster angular resolution (CAR): the radial distance between clusters and their matched tracks (extrapolated to the corresponding calorimeter layer) is measured in bins of $\eta$ and as a function of $E$, to account for the resolution in various regions of the calorimeter. A conservative uncertainty of 5 mrad is used to smear cluster positions.

Uncertainties in the jet $p_T$ and mass are derived by the R$_{trk}$ method [87], comparing the variables calculated using the energy deposited in the calorimeter with those using the momenta of charged-particle tracks. The largest effect on the majority of measured distributions comes from cluster energy smearing for the top and $W$ selections, typically around 8\% but can be as high as 16\% in some regions. The other cluster uncertainty components contribute between 1\% and 6\% in the statistically significant part of the distributions for the top and $W$ selections. For the dijet selection, the typical values are between 2\% and 4\% for all observables, but reach 10\% in some bins. The dominant large-radius jet uncertainties for a subset of variables are shown in figure 1.

In addition to the above uncertainties the sensitivity of the measured distributions to other detector effects was considered. This is summarised as follows:

- Energy scaling correlation scheme: applying the variations to clusters with different kinematics and with different properties, assuming them to be uncorrelated.

- Since the cluster energy calibration is based on pion energy deposition, additional tests are carried out to account for the different energy deposited by non-pion hadrons, such as $K_L$, and the impact on the distributions under study.

- Cluster merging and splitting: topo-clusters can be split or merged during the clustering procedure and this process can be sensitive to noise fluctuations.

In all cases, very conservative variations were applied in order to ensure that the distributions considered were not sensitive to the above effects. For the majority of the distributions the observed variations due to other detector effects were smaller than the cluster uncertainties. However, it was found that $N$-subjettiness variables in the dijet selection had shifts of about 50\% when some of the cluster merging and splitting variations were applied. Using a different axis definition, rather than the WTA variant, did not sufficiently reduce the sensitivity of the variables to this effect. While these variations were conservative, in order to ensure that no systematic uncertainties are being underestimated the $N$-subjettiness variables and their ratios were not used in the dijet selection.

7.2 Other sources of uncertainties

Systematic uncertainties are also derived for other reconstructed objects which are considered in the top and the $W$ selections [88]. Uncertainties associated with small-radius jets, $b$-tagged jets, reconstructed muons and $E_T^{\text{miss}}$ are all considered and are found to be subdominant. The theory normalisation uncertainties are also found to be negligible.

Finally, uncertainties in the shape of the subtracted $W$+jets component are derived by comparing, for each variable, the shapes obtained using the nominal MC sample and
an alternative sample, as listed in table 1. The envelope is taken as an uncertainty in the subtracted shape, and results in uncertainties which are smaller than 1%. The uncertainties due to signal modelling in MC generators are accounted for in unfolding, as described in section 9.

8 Detector-level results

The distributions of the trimmed large-radius jet mass and $p_T$ at detector level are shown in figure 2 for dijet, top and $W$ selections. The peaks in the distributions due to the top and $W$ masses are clearly visible. In general, good agreement is observed between data and simulation for the distribution of transverse momenta, while a shift is observed for the distributions of mass. This is a known effect [2], due to the lack of in situ calibrations of jet mass, and to jet mass scale uncertainties in the detector-level plots.

9 Unfolding

The measured distributions are unfolded to correct for detector effects. The Iterative Bayesian (IB) unfolding method [89] with three iterations (as implemented in RooUnfold [90]) is used to correct detector-level data to particle level, as defined in section 4.
Figure 2. Comparison of detector-level distributions in data and MC simulation for trimmed large-radius jets for dijet (top row), top (middle row), and $W$ (bottom row) selections. For the top and $W$ selections, jet mass requirements have not been applied. The mass is shown in the left column, while the transverse momentum is in the right column. The shaded bands represent the combined statistical and systematic uncertainty. Contributions from dominant backgrounds are shown for the top and $W$ selections, while the smaller contributions from other processes are grouped under other backgrounds.
Response matrices \( (a_{ji}) \) for each distribution are derived from MC simulation and used in order to estimate the probability for a given event at particle level \( (T) \), contributing to bin \( i \), to be reconstructed in a given detector-level \( (D) \) bin \( j \), also defined as \( P(D_j|T_i) \). Rather than using a simple matrix inversion, IB unfolding uses a probabilistic approach. In order to do this, the unfolding matrix \( (\theta_{ij}) \) is defined such that the number of events in a particle-level bin, \( T_i \), is given by

\[
T_i = \sum_j \theta_{ij} d_j
\]  

(9.1)

where \( d_j \) is the number of data events measured in bin \( j \). Using Bayes’ theorem, one can define the unfolding matrix as:

\[
\theta_{ij} = P(T_i|D_j) = \frac{P(D_j|T_i) \cdot P(T_i)}{\sum_i P(D_j|T_i) \cdot P(T_i)} = \frac{a_{ji} \cdot P(T_i)}{\sum_i a_{ji} \cdot P(T_i)}.
\]

where \( P(T_i) \) is the input prior. The unfolding matrix can therefore be constructed using the response matrix obtained from simulation. After corrections are applied for detector acceptance and reconstruction efficiency, eq. (9.1) can be used to perform the unfolding. To ensure that the final distributions are not biased by the shape predicted by simulation the process is iterated, each subsequent iteration using the previous estimate for the final corrected distribution as \( P(T_i) \). The number of iterations is chosen such that differences between multiple subsequent iterations are smaller than data-driven cross-closure uncertainties, described below.

The consistency of the unfolding procedure was tested using several closure and cross-closure tests.

- **MC closure**: a test where the distributions from the nominal MC generator are unfolded using the nominal method. Uncertainties are found to be negligible.

- **Cross-closure**: accounts for modelling differences between two different MC generators. The distributions from an alternative generator are unfolded using the nominal method and the differences account for differences in the predicted shape. These result in the largest uncertainties and are typically around 5% in the dijet selection and around 14% in the top and \( W \) selections, depending on the observable and the bin.

- **Data-driven cross-closure**: accounts for the sensitivity of the unfolding method to differences between the shape of the observable seen in data and in simulation. The particle-level substructure distributions are reweighted such that the corresponding detector-level distributions match the data. These reweighted distributions are unfolded using the nominal method and uncertainties are estimated as the differences between the reweighted particle-level and unfolded distributions.

The binning of variables in the dijet selection was chosen to reduce uncertainties from the above effects by increasing the bin purity. For the top and \( W \) selections binning was determined based on the statistical uncertainty of the dominant systematic uncertainties.
10 Particle-level results

The results are presented in two sets of distributions: substructure observables in data are compared with MC predictions, and distributions measured in data corresponding to different selections are compared with each other. For the latter, it must be noted that the comparisons are performed in different large-radius jet $p_T$ ranges; however, in each instance the most inclusive selection is used. They are indicative of different substructures of the large-radius jets according to their origin even with somewhat different kinematic ranges. All plots with soft-drop grooming are shown; the trimmed versions have very similar characteristics [91]. The dominant systematic uncertainties in the measurement are the large-radius jet uncertainties resulting from the bottom-up approach using clusters, and modelling uncertainties affecting the unfolding closure and cross-closure.

In figure 3, the subjet multiplicity inside the large-radius jets from the three different selections is compared with different MC predictions, and the data are compared between the three selections. While for the dijet selection most events have one subjet, for the top selection and $W$ selection the distributions peak at three and two subjets respectively, as expected. In both cases a non-negligible fraction of events have more subjets, indicating the
Figure 4. Les Houches angularity is compared with different MC predictions for soft-dropped large-radius jets from dijet (top left), top (top right), and $W$ (bottom left) selections. For the dijet selection, SHERPA is tested with two different hadronisation models. Data are compared between the soft-dropped large-radius jets for the three selections mentioned above (bottom right). The shaded bands represent the total uncertainty, while the error bars show the statistical uncertainty, except in the bottom right plot, where the shaded areas represent the total uncertainty.

presence of semi-hard gluon radiation. In the $W$ selection, the instances with one subjet are few, while for the top selection, some fraction of events have two subjets, indicating either non-containment of the top quark decay products, or overlapping subjets that get reconstructed as a single subjet. For the dijet selection, PYTHIA8 and SHERPA describe the data the best, while for the top selection and $W$ selection, there is more spread among MC predictions. Predictions from HERWIG7 are very different from data for the dijet selection, a trend which is consistent across all observables. The difference between the different hadronisation models used in SHERPA is negligible. Although these observables depend on hadronisation modelling, it can be inferred that both models can be tuned to give a good description of data.

In figure 4, the Les Houches angularity (LHA) is compared between large-radius jets for the three selections and with MC model predictions. For the dijet selection, all models except HERWIG7 describe the data, while for the top and $W$ selections, the level of agreement between all models and data is worse, and the peaks of the distributions in the models are shifted relative to those in data. While in the case of the top and $W$ selections the shapes are similar, the distribution for the dijet selection peaks at the lowest value.
This indicates that the additional radiation in quark/gluon jets is soft, with little activity away from the large-radius jet axis, while for the large-radius jets from top quarks and W bosons, there are hard emissions separated by appreciable angles.

In figure 5, a comparison of $C_2$ among the three different selections with MC is presented, as well as a comparisons of data and MC predictions for each selection. For the dijet selection, all models except HERWIG7 describe the data well, while for the top and W selections, the models predict shapes that differ from data, with POWHEG+HERWIG7 performing somewhat worse than the rest. The three distributions have distinct peaks, corresponding to their substructure. The value of $C_2$ increases as the number of subjets inside the large-radius jets increases.

In figure 6, comparisons of the data with MC predictions for $D_2$ reveal some interesting features. For the dijet selection, most of the models describe the data well, and for the top selection the some differences can be seen. For the W selection, all MC predictions have a peak shifted relative to data, suggesting that the models are overestimating gluon radiation. The distributions in data for the three selections are also compared in figure 6 (bottom right), where peaks at different values are observed.
The distributions of $E_{CF}^2$ norm, as shown in figure 7 for the different selections, can discriminate between events with two and three prong decays as opposed to one prong decay. Similarly to $C_2$, for the dijet selection, all models except HERWIG 7 describe the data well, while for the top and W selections, the models predict shapes that differ somewhat from data, with agreement being worse for the W selection case.

The modelling of $E_{CF}^3$ norm in the dijet selection is better for PYTHIA 8 than for the other generators, as shown in figure 8. For the top and W selections, none of the models describe the shape of the data distribution well, with noticeable differences at low values. The three different selections again show distinct shapes.

Finally, in figure 9, a comparison of $\tau_{WTA}^{21}$ and $\tau_{WTA}^{32}$ among top quark and W selections is presented. The distribution of $\tau_{WTA}^{21}$ peaks at lower values for the W selection than for the top selection, indicating the two-prong decay of the former. In general, $\tau_{WTA}^{21}$ distributions are modelled well by the MC models, except POWHEG + HERWIG 7. Although most of the models also describe the $\tau_{WTA}^{32}$ distributions well, differences can be observed between them, especially in the W selection.

Figure 6. The distributions of $D_2$ compared with different MC predictions for soft-dropped large-radius jets from dijet (top left), top (top right), and W (bottom left) selections. For the dijet selection, SHERPA is tested with two different hadronisation models. Data are compared between the soft-dropped large-radius jets for the three selections mentioned above (bottom right). The shaded bands represent the total uncertainty, while the error bars show the statistical uncertainty, except in the bottom right plot, where the shaded areas represent the total uncertainty.
Figure 7. The distributions of $ECF_2^{\text{norm}}$ compared with different MC predictions for soft-dropped large-radius jets from dijet (top left), top (top right), and $W$ (bottom left) selections. For the dijet selection, SHERPA is tested with two different hadronisation models. Data are compared between the soft-dropped large-radius jets for the three selections mentioned above (bottom right). The subscript norm indicates that normalised versions of $ECF_2^{\text{norm}}$ are used. The shaded bands represent the total uncertainty, while the error bars show the statistical uncertainty, except in the bottom right plot, where the shaded areas represent the total uncertainty.
Figure 8. The distributions of $E_{\text{CF}}^3\text{Norm}$ are compared with different MC predictions for soft-dropped large-radius jets from dijet (top left), top (top right), and $W$ (bottom left) selections. For the dijet selection, SHERPA is tested with two different hadronisation models. Data are compared between the soft-dropped large-radius jets for the three selections mentioned above (bottom right). The superscript “norm” indicates that normalised versions of $E_{\text{CF}}^3\text{Norm}$ are used. The shaded bands represent the total uncertainty, while the error bars show the statistical uncertainty, except in the bottom right plot, where the shaded areas represent the total uncertainty.
Figure 9. The distributions of $\tau^{WTA}_{21}$ (left) and $\tau^{WTA}_{32}$ (right) are compared with different MC predictions for large-radius jets from top (top row) and $W$ (bottom row) selections. The distributions of $\tau^{WTA}_{21}$ (bottom left) and $\tau^{WTA}_{32}$ (bottom right) in data are compared between the soft-dropped large-radius jets for the two selections mentioned above. The subscript WTA indicates that WTA axis was used in calculating these observables. The shaded bands represent the total uncertainty, while the error bars show the statistical uncertainty, except in the bottom plots, where the shaded areas represent the total uncertainty.
11 Conclusions

A measurement of jet substructure observables using groomed large-radius jets from light quarks or gluons, hadronically decaying top quarks and $W$ bosons is presented using 33 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton-proton collision data taken with the ATLAS detector at the LHC. The data discriminate between the various MC models probed. In general, Pythia8 for light-quark/gluon large-radius jet observables, and Powheg+Pythia8, Sherpa as well as MG5_aMC@NLO+Pythia8 for top quark and $W$ boson large-radius jet observables, describe the data better than other models. The different hadronisation models in Sherpa in the dijet selection result in similar predictions. For most observables, Herwig7 in the dijet selection, and Powheg+Herwig7 in the top and $W$ selections do not describe the data well. These measurements will be useful in improving the modelling of these substructure variables in MC generators. Since searches that utilise boosted topologies use these observables, or combinations of them, in tagging large-radius jets, a better modelling of them will help to increase the sensitivity of such searches.

Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRS, Morocco; NWO, Netherlands; RCN, Norway; MINaW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d’ Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristea programmes co-financed by EU-ESF and the Greek NSRF, Greece; ISF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, Spain; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL.
(U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [92].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


M.H. Seymour and A. Siodmok, *Constraining MPI models using $\sigma_{\text{eff}}$ and recent Tevatron and LHC underlying event data*, JHEP **10** (2013) 113 [arXiv:1307.5015] [INSPIRE].


[91] https://doi.org/10.17182/hepdata.89324
