Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV

ATLAS Collaboration

DOI
10.1140/epjc/s10052-016-3978-z

Publication date
2016

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
Identification of boosted, hadronically decaying $W$ bosons
and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 22 October 2015 / Accepted: 25 February 2016 / Published online: 17 March 2016
© CERN for the benefit of the ATLAS collaboration. This article is published with open access at Springerlink.com

Abstract This paper reports a detailed study of techniques for identifying boosted, hadronically decaying $W$ bosons using 20.3 fb$^{-1}$ of proton–proton collision data collected by the ATLAS detector at the LHC at a centre-of-mass energy $\sqrt{s} = 8$ TeV. A range of techniques for optimising the signal jet mass resolution are combined with various jet substructure variables. The results of these studies in Monte Carlo simulations show that a simple pairwise combination of groomed jet mass and one substructure variable can provide a 50% efficiency for identifying $W$ bosons with transverse momenta larger than 200 GeV while maintaining multijet background efficiencies of 2–4% for jets with the same transverse momentum. These signal and background efficiencies are confirmed in data for a selection of tagging techniques.

Contents

1 Introduction ............................................. 1
2 A brief introduction to jets, grooming, and substructure variables ............................................. 2
2.1 Jet grooming algorithms ............................................. 2
2.2 Substructure variables ............................................. 3
3 The ATLAS detector ............................................. 4
4 Data and Monte Carlo simulations ............................................. 5
4.1 Monte Carlo samples for the $W$ signal ............................................. 5
4.2 Monte Carlo samples for the multijet background ............................................. 5
5 Object reconstruction and event selection ............................................. 6
6 A comprehensive comparison of techniques in Monte Carlo simulations ............................................. 7
6.1 Performance of grooming algorithms ............................................. 7
6.2 Pileup dependence ............................................. 13
6.3 Performance of substructure variables ............................................. 13
6.4 Summary of grooming and substructure in MC ............................................. 16
7 Detailed studies of selected techniques in data ............................................. 17
7.1 Sample compositions and definitions ............................................. 18
7.2 Event topology effects in Monte Carlo simulations ............................................. 20
7.3 Systematic uncertainties ............................................. 21
7.4 Mass and substructure distributions in $t\bar{t}$ events ............................................. 22
7.5 Signal and background efficiencies and uncertainties ............................................. 24
7.6 Summary of $W$ boson tagging efficiencies in data and MC ............................................. 27
8 Conclusions ............................................. 32
References ............................................. 33

1 Introduction
The high collision energies at the large hadron collider (LHC) can result in the production of particles with transverse momenta, $p_T$, much larger than their mass. Such particles are boosted: their decay products are highly collimated, and for fully hadronic decays they can be reconstructed as a single hadronic jet [1] (a useful rule of thumb is $2M/p_T \sim R$: twice the jet mass divided by the $p_T$ is roughly equal to the maximum opening angle of the two decay products). Heavy new particles as predicted in many theories beyond the Standard Model can be a source of highly boosted particles.

The work presented here is the result of a detailed study of a large number of techniques and substructure variables that have, over recent years, been proposed as effective methods for tagging hadronically decaying boosted particles. In 2012, the ATLAS experiment collected 20.3 fb$^{-1}$ of proton–proton collision data at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, providing an opportunity to determine which of the many available techniques are most useful for identifying boosted, hadronically decaying $W$ bosons. In the studies

\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 $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{\Delta \eta^2 + \Delta \phi^2}$.}

*e-mail: atlas.publications@cern.ch
presented here, jets that contain the W boson decay products are referred to as W-jets.

A brief overview of the existing jet grooming and substructure techniques, along with references to more detailed information, are provided in Sect. 2. The ATLAS detector is described in Sect. 3, and details of Monte Carlo simulations (MC) in Sect. 4. The event selection procedure and object definitions are given in Sect. 5.

The body of the work detailing the W-jet tagging performance studies is divided into a broad study using MC (Sect. 6) and a detailed study of selected techniques in data (Sect. 7).

In Sect. 6 a two-stage optimisation procedure has been adopted: firstly more than 500 jet reconstruction and grooming algorithm configurations are investigated at a basic level, studying the groomed jet mass distributions only. Secondly, 27 configurations that are well-behaved and show potential for W-jet tagging are investigated using pairwise combinations of mass and one substructure variable.

In Sect. 7, one of the four most promising jet grooming algorithms and three substructure variables are selected as a benchmark for more detailed studies of the W-jet tagging performance in data. Jet mass and energy calibrations are discussed. Event topologies are compared, and various sources of systematic uncertainty and their effects on the measurements are investigated. For the first stage, C/A jets with $R = 0.8$ and pruning with $Z_{cut} = 0.10\%$ and $R_{cut} = 1.2$.

**2 A brief introduction to jets, grooming, and substructure variables**

**2.1 Jet grooming algorithms**

The jet grooming algorithms studied here fall into three main categories: trimming [2], pruning [3,4] and split-filtering [5]. Within each category there are several tunable configuration parameters, in addition to the chosen initial jet reconstruction algorithm, Cambridge–Aachen [6] (C/A) or anti-$k_t$ [7], and jet radius parameter $R$. The FastJet [8] package is used for jet reconstruction and grooming. Jet grooming algorithms generally have two uses; (i): to remove contributions from pileup (additional $p\bar{p}$ interactions in the same or adjacent bunch crossings within the detector readout window), and (ii) to reveal hard substructure within jets resulting from massive particle decays by removing the soft component of the radiation.

The three major categories of jet grooming algorithms are described below:

- **Trimming**: Starting with constituents of jets initially reconstructed using the C/A or anti-$k_t$ algorithm, smaller ‘subjets’ are reconstructed using the $k_t$ algorithm [9] with a radius parameter $R = R_{sub}$, and removed if they carry less than a fraction $f_{cut}$ of the original, ungroomed, large-$R$ jet $p_T$. For reference, the recommended trimming configuration from prior ATLAS studies [10] is anti-$k_t$, $R = 1.0$, with $f_{cut} \geq 5\%$ and $R_{sub} = 0.3$.

- **Pruning**: The constituents of jets initially reconstructed with the C/A or anti-$k_t$ algorithms are re-clustered with the C/A algorithm with two parameters: $R_{cut}$ and $Z_{cut}$. The $k_t$ algorithm was used for re-clustering in previous studies [10], but was not found to be as effective. In each pairwise clustering, the secondary constituent is discarded if it is (i) wide-angled: $\Delta R_{12} > R_{cut} \times 2 M/ p_T$, where $\Delta R_{12}$ is the angular separation of the two subjets; or (ii) soft: $f_2 < Z_{cut}$, where $M$ is the jet mass and $f_2$ is the $p_T$ fraction of the softer constituent with respect to the $p_T$ of the pair. A configuration of the pruning algorithm is favoured by the CMS experiment for W-jet tagging [11,12], using C/A jets with $R = 0.8$ and pruning with $Z_{cut} = 10\%$ and $R_{cut} = 1.2$.

- **Split-filtering**: This algorithm has two stages: the first (splitting) is based on the jet substructure, and the second (filtering) is a grooming stage to remove soft radiation. For the first stage, C/A jets are de-clustered through the clustering history of the jet. This declustering is an exact reversal of the C/A clustering procedure, and can be thought of as splitting the jet into two pieces. The momentum balance, $\sqrt{\gamma_{12}}$, is defined as:

$$\sqrt{\gamma_{12}} = \min(p_{T1}, p_{T2}) \frac{m_{12}}{\Delta R_{12}},$$  \hspace{1cm} (1)

where $p_{T1}$ ($p_{T2}$) is the piece with the highest (the lowest) $p_T$, and $m_{12}$ is the invariant mass of the two pieces. The mass-drop fraction $\mu_{12}$ is the fraction of mass carried by the piece with the highest mass:

$$\mu_{12} = \frac{\max(m_{1}, m_{2})}{m_{12}}.$$  \hspace{1cm} (2)

If the requirements on the mass-drop $\mu_{12} < \mu_{max}$ and momentum balance $\sqrt{\gamma_{12}} > \sqrt{\gamma_{min}}$ are met then the jet is accepted and can proceed to the filtering stage. Otherwise the de-clustering procedure continues with the highest mass piece: this is now split into two pieces and the $\mu_{12}$ and $\sqrt{\gamma_{12}}$ requirements are again checked. This process continues iteratively. In the filtering stage, the constituents of the surviving jet are reclustered with a subjet size of $R_{sub} = \min(0.3, \Delta R_{12})$ where $\Delta R_{12}$ is taken from the splitting stage. Any remaining radiation outside the three hardest subjects is discarded. This algorithm differs somewhat from pruning and trimming in....
that it involves both grooming and jet selection. A version of this algorithm is favoured by ATLAS diboson resonance searches [13–15].

2.2 Substructure variables

Substructure variables are a set of jet properties that are designed to uncover hard substructure within jets. An important difference in the substructure variables comes from the choice of distance measure used in their calculation. The various distance measures available are illustrated in Fig. 1. The jet axis is usually defined as the thrust axis (along the jet momentum vector) and can also be defined as the ‘winner-takes-all’ axis which is along the momentum vector of the constituent with the largest momentum.

The many jet substructure techniques can be roughly categorised as follows:

- **Jet shapes** use the relative positions and momenta of jet constituents with respect to each other, rather than defining subjects. The jet mass, $M$, energy correlation ratios $C_{ij}(\beta)$ [16] and $D_{ij}(\beta)$ [17,18], the mass-normalised angularity $a_{3}$ [19], and the planar flow, $P$ [19], all satisfy this description. The calculations of the jet mass and energy correlation ratios are described later in this section.

- **Splitting scales** use the clustering history of the jet to define substructures (‘natural subjects’). The splitting scales studied here are $\sqrt{D_{ij}}$ [20] and its mass-normalised form $\sqrt{\epsilon_{12}}$ [21], and the momentum balance and mass-drop variables $\sqrt{\gamma_{12}}$ and $\mu_{12}$, defined above in the description of the split-filtering algorithm. The soft-drop level $L_{SD}(\beta)$ [22] also belongs in this class of variables.

- **Subjettiness** variables [23,24] force the constituents into substructure templates to see how well they fit (‘synthetic subjets’), and are connected to how likely the corresponding jet is composed of n subjets. The calculations for two forms of 2-subjettiness $\tau_{2}$, $\tau_{2}^{\text{wta}}$, and the corresponding ratios $\tau_{21}$, $\tau_{21}^{\text{wta}}$ are given later in this section. The dipolarity [25], $D$, uses a related method to define hard substructure.

- **Centre-of-mass jet shapes** transform the constituents and then use them with respect to the jet axis. The variables considered are thrust, $T_{\text{min}}$, $T_{\text{maj}}$, sphericity, $S$, and aplannarity, $A$, which have been used in a previous ATLAS measurement [26].

- **Quantum-jet variables** The quantum jets (‘Q-jets’) method [27] is unique in its class, using a non-deterministic approach to jet reconstruction. More information on the use of this method by ATLAS can be found in Ref. [28].

The variables found in the following studies to be most interesting in terms of $W$-jet tagging are described here in more detail.

**Jet mass:**

The mass of a jet is given by the difference between the squared sums of the energy $E_{i}$ and momenta $p_{i}$ of the constituents:

$$M^{2} = \left( \sum_{i} E_{i} \right)^{2} - \left( \sum_{i} p_{i} \right)^{2}. \quad (3)$$

For a two-body decay, the jet mass can be approximated as:

$$M^{2} \approx p_{T1} p_{T2} \Delta R_{12}^{2}. \quad (4)$$

**N-subjettiness:**

The “N-subjettiness” [23,24] jet shape variables describe to what degree the substructure of a given jet $J$ is compatible with being composed of $N$ or fewer subjets. The 0-, 1- and 2-subjettiness are defined as:

$$\tau_{0}(\beta) = \sum_{i \in J} p_{Ti} \Delta R_{i}^{\beta}, \quad (5a)$$

$$\tau_{1}(\beta) = \frac{1}{\tau_{0}(\beta)} \sum_{i \in J} p_{Ti} \Delta R_{a1,i}^{\beta}, \quad (5b)$$

$$\tau_{2}(\beta) = \frac{1}{\tau_{0}(\beta)} \sum_{i \in J} p_{Ti} \min(\Delta R_{a1,i}^{\beta}, \Delta R_{a2,i}^{\beta}), \quad (5c)$$

where the distance $\Delta R$ refers to the distance between constituent $i$ and the jet axis, and the parameter $\beta$ can be used to give a weight to the angular separation of the jet constituents. In the studies presented here, the value of $\beta = 1$ is taken. The calculation of $\tau_{N}$ requires the definition of $N$ axes, such that...
the distance between each constituent and any of these axes is $R_{dN,i}$. In the above functions, the sum is performed over the constituents $i$ in the jet $J$, such that the normalisation factor $t_0$ (Eq. 5a) is equivalent to the magnitude of the jet $p_T$ multiplied by the $\beta$-exponentiated jet radius.

Recent studies [29] have shown that an effective alternative axis definition can increase the discrimination power of these variables. The ‘winner-takes-all’ axis uses the direction of the hardest constituent in the exclusive $k_t$ subjet instead of the subjet axis, such that the distance measure $\Delta R_{d1,i}$ changes in the calculation. The ratio of the N-subjettiness functions found with the standard subjet axes, $\tau_{21}$, and with the ‘winner-takes-all’ axes, $\tau_{21}^{wta}$, can be used to generate the dimensionless variables that have been shown in particle-level MC to be particularly useful in identifying two-body structures within jets:

$$\tau_{21} = \frac{\tau_2}{\tau_1}, \quad \tau_{21}^{wta} = \frac{\tau_{21}^{wta}}{\tau_{1}^{wta}}. \quad (6)$$

**Energy correlation ratios:**

The 1-point, 2-point and 3-point energy correlation functions for a jet $J$ are given by:

$$E_{CF0}(\beta) = 1, \quad (7a)$$

$$E_{CF1}(\beta) = \sum_{i \in J} p_{Ti}, \quad (7b)$$

$$E_{CF2}(\beta) = \sum_{i < j \in J} p_{Ti} p_{Tj} (\Delta R_{ij})^\beta, \quad (7c)$$

$$E_{CF3}(\beta) = \sum_{j < k \in J} p_{Ti} p_{Tj} p_{Tk} (\Delta R_{ij} \Delta R_{jk} \Delta R_{jk})^\beta, \quad (7d)$$

where the parameter $\beta$ is used to give weight to the angular separation of the jet constituents. In the above functions, the sum is over the constituents $i$ in the jet $J$, such that the 1-point correlation function Eq. (7b) is approximately the jet $p_T$. Likewise, if one takes $\beta = 2$, it is noted that the 2-point correlation functions are equivalent to the mass of a particle undergoing a two-body decay in collider coordinates.

An abbreviated form of these definitions can be written as:

$$e_2^{(\beta)} = \frac{E_{CF2}(\beta)}{E_{CF1}(\beta)^2}, \quad (8a)$$

$$e_3^{(\beta)} = \frac{E_{CF3}(\beta)}{E_{CF1}(\beta)^3}. \quad (8b)$$

These ratios of the energy correlation functions can be used to generate the dimensionless variable $C_2^{(\beta)}$ [16], and its more recently modified version $D_2^{(\beta)}$ [17,18], that have been shown in particle-level MC to be particularly useful in identifying two-body structures within jets:

$$C_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^2},$$

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}. \quad (9b)$$

Values of $\beta = 1$ and 2 are studied here.

### 3 The ATLAS detector

The ATLAS detector [30] at the LHC covers nearly the entire solid angle around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroid magnets.

The inner-detector system (ID) is immersed in a 2 T axial magnetic field and provides charged particle tracking in the range $|\eta| < 2.5$. A high-granularity silicon pixel detector covers the vertex region and typically provides three measurements per track. It is followed by a silicon microstrip tracker, which usually provides four two-dimensional measurement points per track. These silicon detectors are complemented by a transition radiation tracker, which enables radially extended track reconstruction up to $|\eta| = 2.0$. The transition radiation tracker also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) electromagnetic calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$, to correct for energy loss in material upstream of the calorimeters. For the jets measured here, the transverse granularity ranges from $0.003 \times 0.1$ to $0.1 \times 0.1$ in $\Delta \eta \times \Delta \phi$, depending on depth segment and pseudorapidity. Hadronic calorimetry is provided by a steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. This system enables measurements of the shower energy deposition in three depth segments at a transverse granularity of typically $0.1 \times 0.1$. The solid angle coverage is extended with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements respectively.

A muon spectrometer (MS) comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by superconducting air-core toroids. The precision chamber system covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, com-
implemented with cathode strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

A three-level trigger system is used to select interesting events [31]. The Level-1 trigger is implemented in hardware and uses a subset of detector information to reduce the event rate to a design value of at most 75 kHz. This is followed by two software-based trigger levels which together reduce the event rate to about 400 Hz.

4 Data and Monte Carlo simulations

The data used for this analysis were collected during the $pp$ collision data-taking period in 2012, and correspond to an integrated luminosity of $20.3 \text{ fb}^{-1}$ with a mean number of $pp$ interactions per bunch crossing, $\langle \mu \rangle$, of about 20. The uncertainty on the integrated luminosity, 2.8%, is derived following the same methodology as that detailed in Ref. [32] using beam-separation scans. Data quality and event selection requirements are given in Sect. 5.

Events from Monte Carlo generator are passed through a GEANT4-based [33] simulation of the ATLAS detector [34], and reconstructed using the same algorithm used as for data. All MC samples are produced with the addition of pileup, using hits from minimum-bias events that are produced with PYTHIA (8.160) [35] using the A2M set of tunable parameters (tune) [36] and the MSTW2008LO [37] PDF set. This simulated pileup does not exactly match the distribution of $\langle \mu \rangle$ measured in data. As such, event weights are derived as a function of $\langle \mu \rangle$ for the MC samples used in the data/MC comparisons, making the differences between the data and MC $\langle \mu \rangle$ distributions negligible.

4.1 Monte Carlo samples for the W signal

Samples of the hypothetical process $W' \rightarrow WZ \rightarrow qq\ell\ell$ are produced as a source of signal high-$p_T$ $W$-jets, with the boost in $p_T$ coming from the high mass of the parent $W'$. These samples are produced using PYTHIA (8.165) with the AU2 [36] tune and the MSTW2008LO [37] PDF set. Nine separate signal samples are produced with $W'$ masses ranging from 400 to 2000 GeV in steps of 200 GeV. This ensures good coverage over a wide range of $W$-jet $p_T$. The nine samples are combined and the events are given weights such that when the event weights are applied, the $p_T$ distribution of the combined signal $W$-jets sample matches that of the multijet background sample described in Sect. 4.2. These are used as the signal samples in the preliminary optimisation studies presented in Sect. 6.

The $W$ boson tagging efficiency from top quark decays in data, detailed in Sect. 7, is measured using $t\bar{t}$ samples simulated with the POWHEG-BOX (version 1, r2330) NLO generator [38] interfaced with PYTHIA (6.427). A cross-check is performed with MC@NLO [39] (4.03), with parton showers provided by HERWIG (6.520) [40] + JIMMY (4.31) [41]. In both cases, the next-to-leading order CT10 [42] PDF set is used, and the top quark mass is set to 172.5 GeV. Single-top-quark events in the $s$-, $t$- and $Wt$-channels are simulated with POWHEG-BOX interfaced with PYTHIA (6.426), with the Perugia 2011c [43] tune. The $t$-channel is also generated with POWHEG-BOX in the four-flavour scheme. Background $W+$jet and $Z+$jet events are simulated using ALPGEN [44] (2.14) in the four-flavour scheme ($b$-quarks are treated as massive) followed by PYTHIA (6.426) for the parton shower. Up to five extra partons are considered in the matrix element. The CTEQ6L1 [45] PDF set and the Perugia 2011c tune are used. For diboson events, the SHERPA [46] (1.4.3) generator is used with up to three extra partons in the matrix element and the masses of the $b$- and $c$-quarks are taken into account.

The effects of differences between the $W' \rightarrow WZ$ process used for $W$-jets in the preliminary optimisation studies and the $t\bar{t}$ process used in the detailed comparisons with data are discussed in Sect. 7.2.

4.2 Monte Carlo samples for the multijet background

The background sample used in Sect. 6 is made up of several high-$p_T$ multijet event samples produced using PYTHIA [35] with the AU2 [36] tune and the CT10 [42] PDF set. Eight samples in total are produced according to the leading jet’s $p_T$, four of which are used in this analysis to cover the $p_T$ range 200–2000 GeV. These samples are combined with event weights determined by their relative cross-sections to produce the smoothly falling $p_T$ distribution predicted by PYTHIA. The MC optimisation studies use the leading jets from these events. The jets in these background samples are initiated by light quarks and gluons, the interactions of which are described by Quantum Chromodynamics, QCD.

The $W$-tagging efficiency in multijet background events is studied on the same multijet samples as used for the optimisation studies, using PYTHIA (8.165) with the AU2 tune and the CT10 PDF set, and also a HERWIG++ (2.6.3) sample with the EE3 tune [47] and CTEQ6L1 [45] PDF set. It is these samples that are used for the comparisons with data in Sect. 7.

The effects of differences between these samples due to using the leading jets (for the MC-based optimisation) or both leading and sub-leading jets (for the multijet background efficiency measurement in data) are discussed in Sect. 7.2.
5 Object reconstruction and event selection

In the studies presented here, calorimeter jets are reconstructed from three-dimensional topological clusters (topoclusters) [48] which have been calibrated using the local cluster weighting (LCW) scheme [49]. In MC simulated events, truth jets are built from generator-level particles that have a lifetime longer than 10 ps, excluding muons and neutrinos. Jets are reconstructed using one of the iterative recombination jet reconstruction algorithms [50,51] C/A or anti-kt. The $k_t$ algorithm is also used by the jet trimming algorithm to reconstruct subjets.

In all following discussions, the term constituents means particles in the case of truth jets and LCW topoclusters in the case of calorimeter jets.

For the MC-based optimisation studies discussed in Sect. 6, events are characterised using the leading jet, reconstructed from generator-level particles with the C/A, $R = 1.2$ algorithm.

Objects used to select $t\overline{t}$ events in data and MC for the studies in Sect. 7 include reconstructed leptons (electrons and muons), missing transverse momentum ($E_T^{\text{miss}}$), small-$R$ jets (reconstructed with the anti-$k_t$ algorithm with radius parameter $R = 0.4$), trimmed anti-$k_t$, $R = 1.0$ jets and $b$-tagged jets, defined below.

- **Electrons:** Electron candidates are reconstructed from energy deposits in the EM calorimeter matched to reconstructed tracks in the ID. Candidates are required to be within $|\eta| < 2.47$, excluding the barrel/endcap transition region, $1.37 < |\eta| < 1.52$, of the EM calorimeter, and must have a transverse energy $E_T > 25$ GeV. They are required to satisfy tight identification criteria [52] and to fulfil isolation requirements; excluding its own track, the scalar sum of the $p_T$ of charged tracks within a cone of size $\Delta R = \min(10 \, \text{GeV} / E_T, 0.4)$ around the electron candidate must be less than 5% of the $p_T$ of the electron.

- **Muons:** Muons are reconstructed by matching MS to ID tracks. Muons are required to be within $|\eta| < 2.5$ and have $p_T > 25$ GeV. In order to reject non-prompt muons from hadron decays, the significance of their transverse impact parameter must be $|d_0|/\sigma_{d_0} < 3$, the longitudinal impact parameter must be $|z_0| < 2$ mm, and the scalar sum of $p_T$ of the charged tracks within a cone of size $\Delta R = \min(10 \, \text{GeV} / p_T, 0.4)$ around the muon candidate, excluding its own track, must be less than 5% of the $p_T$ of the muon.

- **Trigger leptons:** Events are selected by requiring an unprescaled single-lepton trigger for the electron and muon channels. Two single-electron triggers, with transverse energy thresholds of $E_T > 24$ GeV for isolated electrons and $E_T > 60$ GeV without isolation criteria, are used in combination with two single-muon triggers, with transverse momentum of $p_T > 24$ GeV for isolated muons and $p_T > 36$ GeV without isolation criteria. The selected muon (electron) must be matched to a trigger and is required to fulfil $p_T > 25(20)$ GeV and $|\eta| < 2.5$. Events are rejected if any other electron or muon satisfying the identification criteria is found in the event.

- **Missing transverse momentum, $E_T^{\text{miss}}$ and transverse mass, $m_W^T$:** The missing transverse momentum is calculated from the vector sum of the transverse energy of topological clusters in the calorimeter [54]. The clusters associated with the reconstructed electrons and small-$R$ jets are replaced by the calibrated energies of these objects. Muon $p_T$ determined from the ID and the muon spectrometer are also included in the calculation. The $E_T^{\text{miss}}$ is required to exceed 20 GeV. The sum of the $E_T^{\text{miss}}$ and the transverse mass, $m_W^T = \sqrt{2 p_T E_T^{\text{miss}} (1 – \cos \Delta \phi)}$, reconstructed from the $E_T^{\text{miss}}$ and the transverse momentum of the lepton, must be $E_T^{\text{miss}} + m_W^T > 60$ GeV.

- **Small-$R$ Jets (anti-$k_t$, $R = 0.4$):** Using locally calibrated topological clusters as input, small-$R$ jets are formed using the anti-$k_t$ algorithm with a radius parameter $R = 0.4$. Small-$R$ jets are required to be within $|\eta| < 2.5$ and to have $p_T > 25$ GeV. To reject jets with significant pileup contributions, the jet vertex fraction [55], defined as the scalar sum of the $p_T$ of tracks associated with the jet that are assigned to the primary vertex divided by the scalar sum of the $p_T$ of all tracks associated to the jet, is required to be greater than 0.5 for jets with $p_T < 50$ GeV. At least one small-$R$ jet must be found. In addition, at least one small-$R$ jet must lie within $\Delta R = 1.5$ of the lepton. The leading small-$R$ jet within $\Delta R = 1.5$ of the lepton is defined as the “leptonic-top jet” and denoted $j_{\ell \ell}$. Jets have to satisfy specific cleaning requirements [56] to remove calorimeter signals coming from non-collision sources or calorimeter noise. Events containing any jets that fail these requirements are rejected.

- **b-jets (anti-$k_t$, $R = 0.4$):** The output of the MV1 [57] algorithm is used to identify small-$R$ jets containing $b$-hadrons. Small-$R$ jets are tagged as $b$-jets if the MV1 weight is larger than the value corresponding to the 70% $b$-tagging efficiency working point of the algorithm. At least one small-$R$ jet must be tagged as a $b$-jet. Loose $b$-jets are defined as having an MV1 weight larger than the value corresponding to the 80% working point. All loose $b$-jets must be separated by $\Delta R > 1.0$ from the $W$-jet candidate.

- **Trimmed $R = 1.0$ Jets:** Using locally calibrated topological clusters as inputs, anti-$k_t$, $R = 1.0$ jets are groomed using the trimming algorithm with parameters $f_{\text{cut}} = 5\%$ and $R_{\text{sub}} = 0.2$. The pseudorapidity, energy and mass of these jets are calibrated using a simulation-based calibration scheme as mentioned in Sect. 6.4. At least one trimmed anti-$k_t$, $R = 1.0$ jet with $p_T > 200$ GeV and
$|\eta| < 1.2$ is required. If more than one jet satisfies these criteria, the leading jet is used to reconstruct the $W$ boson candidate, $J_W$. This candidate, $J_W$, has to be well separated from the leptonic-top jet, $\Delta R(J_W, J_t) > 1.2$.

- **Overlapping jets and leptons**: An overlap removal procedure is applied to avoid double-counting of leptons and anti-$k_t$, $R = 0.4$ jets, along with an electron-in-jet subtraction procedure to recover prompt electrons that are used as constituents of a jet. If an electron lies $\Delta R < 0.4$ from the nearest jet, the electron four-momentum is subtracted from that of the jet. If the subtracted jet fails to meet the small-$R$ jet selection criteria outlined above, the jet is marked for removal. If the subtracted jet satisfies the jet selection criteria, the electron is removed and its four-momentum is added back into the jet. Next, muons are removed if $\Delta R(\text{muon}, \text{jet}) < 0.04 + 10 \text{ GeV}/p_T^{\text{muon}}$ using jets that are not marked for removal after the electron subtraction process.

For the measurement of the multijet background efficiency, a different selection is used to ensure a multijet-enriched sample. The multijet sample is selected using a single, un-prescaled, $R = 1.0$ jet trigger that is 80% efficient for jets with $p_T > 450$ GeV. No grooming is applied to jets at the trigger level. For events with a leading jet above the trigger threshold, both the leading and the sub-leading jets are used for this performance study, making it applicable for jets with $p_T$ down to 200 GeV. At least one anti-$k_t$, $R = 1.0$ jet, trimmed with $f_{\text{cut}} = 5\%$ and $R_{\text{sub}} = 0.2$, is required to have $p_T > 200$ GeV and $|\eta| < 1.2$. Events containing fake jets from noise in the calorimeter or non-collision backgrounds, according to Refs. [58, 59], are rejected.

For the $t\bar{t}$ and multijet background selection, good data quality is required for events in data, meaning that all the detectors of ATLAS as well as the trigger and data acquisition system are required to be fully operational. Events are required to have at least one reconstructed primary vertex with at least five associated tracks, and this vertex must be consistent with the LHC beam spot.

6 A comprehensive comparison of techniques in Monte Carlo simulations

The initial phase of this study evaluates the performance of a large number of grooming and tagging algorithms in MC simulated events.

To account for correlations between the $W$ boson $p_T$ and the resulting jet substructure features, events are categorised by the $p_T$ of the leading (highest $p_T$) jet reconstructed with the C/A [6] algorithm with radius parameter $R = 1.2$, using stable particles as inputs. These ranges in the ungroomed truth jet $p_T$, $p_T^{\text{truth}}$, are: [200, 350] GeV, [350, 500] GeV, [500, 1000] GeV. This large, ungroomed jet is considered a rough proxy for the $W$ boson, and this choice does not introduce a bias towards any particular grooming configuration for the $p_T^{\text{truth}}$ ranges in question. Only events with a $C/A$, $R = 1.2$ truth jet within $|\eta| < 1.2$ are considered, ensuring that jets are within the acceptance of the tracking detector, which is necessary for the derivation of the systematic uncertainties.

First, in Sect. 6.1, more than 500 jet reconstruction and grooming algorithm configurations are selected based on prior studies [10, 11, 60–63]. The leading-groomed-jet mass distributions for $W$-jet signal and multijet background in MC are examined. An ordered list is built rating each configuration based on the background efficiency. The notation for the background efficiency at this grooming stage is $\epsilon^G_{QCD}$, and this is measured within a mass window that provides a signal efficiency of 68%, denoted $\epsilon^G_W = 68\%$. The best performers for each category described in Sect. 2.1 (trimming, pruning, split-filtering) are retained for the next stage: a total of 27 jet collections.

Observations about pileup-dependence are summarised in Sect. 6.2. Jet grooming reduces the pileup-dependence of the jet mass and helps distinguish $W$-jets from those initiated by light quarks and gluons by improving the mass resolution, but does not provide strong background rejection. Further information coming from the distribution of energy deposits within a jet can be used to improve the ratio of signal to background.

In the second stage, 26 substructure variables are studied for all 27 selected jet collections. These studies are detailed in Sect. 6.3. Substructure variables can be calculated using jet constituents before or after grooming; in these studies all variables are calculated from the groomed jet’s constituents, such that the potential sensitivity to pileup conditions is reduced.

The aim of these studies is to find an effective combination of groomed jet mass and one substructure variable. The background efficiency $\epsilon^{G&T}_{QCD}$ (where G&T indicates grooming plus tagging) versus the signal efficiency $\epsilon^{G&T}_W$ is calculated for all variables in each configuration, and background efficiencies for ‘medium’ (50%) and ‘tight’ (25%) signal efficiency working points are determined. Four grooming algorithms and three tagging variables are identified as having a particularly low background efficiency at the medium signal efficiency working point, $\epsilon^{G&T}_W = 50\%$.

In Sect. 6.4 the conclusions of these preliminary studies of combined groomed mass and substructure taggers are presented.

6.1 Performance of grooming algorithms

A set of more than 500 jet reconstruction and grooming algorithm configurations (introduced in Sect. 2.1) are explored within the parameter space summarised in Table 1.
Table 1 Details of the different trimming, pruning and split-filtering configurations that were tried in order to define the best grooming algorithms. All combinations of the grooming parameters are explored in these studies.

<table>
<thead>
<tr>
<th>Trimming configurations</th>
<th>Input jet algorithms</th>
<th>R</th>
<th>Rsub</th>
<th>fcut (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A, anti-k_t</td>
<td>RR</td>
<td>0.6, 0.8, 1.0, 1.2</td>
<td>0.1, 0.2, 0.3</td>
<td>1, 2, 3, 4, 5, 7, 9, 11, 13, 15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pruning configurations</th>
<th>Input jet algorithm</th>
<th>R</th>
<th>Reclust. alg.</th>
<th>Zcut (%)</th>
<th>Rcut</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A, anti-k_t</td>
<td>RR</td>
<td>0.8, 1.0, 1.2</td>
<td>C/A</td>
<td>10, 15, 20, 25, 30</td>
<td>1 / 100, 1 / 10, 1 / 8, 1 / 4, 1 / 2, 1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Split-filtering configurations</th>
<th>Input jet algorithm</th>
<th>R</th>
<th>Rsub</th>
<th>μmax</th>
<th>ycuf</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A</td>
<td>RR</td>
<td>0.8, 1.0, 1.2</td>
<td>0.3, min(0.3, ΔR/2)</td>
<td>67, 78, 89, 100</td>
<td>0.06, 0.07, ..., 0.20</td>
</tr>
</tbody>
</table>

Fig. 2 Uncalibrated mass distributions for various selected grooming configurations: a trimmed with \( R_{\text{sub}} = 0.2 \), b trimmed with \( R_{\text{sub}} = 0.3 \), c pruned, and d split-filtered. The transverse momentum range \( p_T^{\text{truth}} = [200, 350] \text{ GeV} \) is shown for \( W \) signal (solid blue line) and multijet background (dashed red line). The (black) Gaussian fit uses an initial-condition mass set to 80.4 GeV. The dotted vertical lines indicate the 1σ fit interval. The dashed lines contain 68% of the signal and define the mass window. These are examples of grooming algorithms leading to satisfactory mass distributions. Uncertainty bands are statistical only.
The signal and background mass distributions for a selection of grooming configurations in the range $200 < p_T^{\text{Truth}} < 350$ GeV are shown in Fig. 2. A Gaussian fit to the $W$ boson mass peak (with the $W$ mass set as the initial condition) is shown. Two alternative signal mass window definitions are considered:

1. The $\sigma$ boundaries of the Gaussian fit.
2. The smallest interval that contains 68% of the integral.

Comparing the extent of these two mass windows allows an estimation of how closely the signal mass peak resembles a Gaussian distribution. The $W$-jet mass is required to be within the boundaries defined by this latter definition of the signal window; this leads, by definition, to a baseline signal efficiency of $\epsilon_W^{(a)} = 68\%$ for all algorithms.

The groomed jet mass distributions for leading jets are examined for all combinations of grooming configurations for $W$-jet signal and multijet background. The background efficiency, $\epsilon_Q^{G}$, is defined as follows:

- The denominator is the total number of pre-selected events from the multijet background sample, where the pre-selection requires an ungroomed C/A, $R = 1.2$ truth jet with $p_T^{\text{Truth}} > 200$ GeV and $|\eta^{\text{Truth}}| < 1.2$.
- The numerator is the number of pre-selected events where the groomed jet mass falls in the window that contains 68% of the $W$-jet signal, $\epsilon_W^{(a)} = 68\%$.

The minimisation of $\epsilon_Q^{G}$ is the primary criterion for ordering the algorithms according to their performance. In addition, there are a number of possible pathologies revealed in the mass distributions: features that show obviously unsuitable configurations, or may make it impossible to derive a jet mass calibration, or indicate the need for additional pileup removal techniques. These are:

(i) The $\epsilon_W^{(a)} = 68\%$ window does not contain the $W$ boson mass [64]. An example of this is shown in Fig. 3a.
(ii) The signal mass distribution is strongly non-Gaussian. An example of this is shown in Fig. 3b.
(iii) The background mass distribution has an irregular shape (e.g. it has local maxima) in the region of the signal peak. An example of this is also shown in Fig. 3b.
(iv) The jet mass after grooming is strongly affected by pileup. Configurations where the average jet mass increases by $>1$ GeV times the number of primary vertices, NPV, are rejected. This issue is discussed in Sect. 6.2.

Algorithms that are susceptible to any of these pathologies are removed from the list of well-behaved algorithm configurations.

The $W$ boson tagging efficiency performance is studied independently for three different ranges in the $p_T$ of the ungroomed truth jet reconstructed with the C/A, $R = 1.2$ algorithm: [200, 350], [350, 500], [500, 1000] GeV. The results for the three grooming categories share some common features:

![Fig. 3 Uncalibrated mass distributions for two problematic grooming configurations in the transverse momentum range $p_T^{\text{Truth}} = [200, 350]$ GeV for $W$ signal and multijet background. The Gaussian fit uses an initial-condition mass set to 80.4 GeV. The dotted vertical lines indicate the $1\sigma$ fit interval. The dashed lines contain 68% of the signal and define the mass window. These plots show examples of unwanted behaviours: in (a) most signal events are reconstructed with a small mass, indicating that the $W$ boson decay products are not fully contained in the jet; and in (b) the signal mass distribution is strongly asymmetric.](image-url)
Trimmed jets, $R = 1.0$

350 < $p_T$ < 500 GeV

$|\eta|$ < 1.2

ATLAS

Pythia8 simulation

$\sqrt{s} = 8$ TeV

ATLAS

Trimmed jets, $R = 1.0$

$350 < p_T < 500$ GeV

$|\eta|$ < 1.2

Truth

$T350 < p_T < 1.2$

Truth

$\eta$

C/A

tanti-k

= subR

0.1

0.2

0.3

-cutf

1%

2%

3%

4%

5%

$M \ [GeV]$

40

60

80

100

120

Shaded bands indicate extent of 68% mass windows

W mass

Bars indicate stat. + baseline syst. uncertainty from JMS+JES

10% background efficiency

Fig. 4 Mass windows and background efficiencies for various configurations of trimming ($R = 1.0$ shown). The baseline systematic uncertainty on the background efficiency for the $p_T$ bin in question (the range 350 < $p_T$ < 500 GeV is shown here) is calculated by varying the jet mass scale (JMS) and jet energy scale (JES) by ±1σ for a representative jet collection. For trimming, this representative configuration is $R_{sub} = 0.2$ and $f_{cut} = 5\%$. The stars indicate the favoured trimming configurations for W-tagging, as detailed in Sect. 6.4.

- The jets reconstructed with $R = 0.6$ and $R = 0.8$ are too small to contain all the decay products of a W-jet for $p_T < 500$ GeV and $p_T < 350$ GeV, respectively. The reconstructed jet mass is often much smaller than 80 GeV, indicating that some of the W boson decay products are not clustered, and the 68% signal mass window is wider, resulting in a higher background efficiency. Small radii jets can, however, have good performance at high $p_T$.
- In the highest $p_T$ bin, 500–1000 GeV, the various configurations result in a similar performance.

The unique features of each grooming category are presented below.

**Trimming:**

Various trimming configurations are studied, varying the algorithm and size of the initial jet (C/A with $R = 0.6–1.2$, anti-$k_t$ with $R = 0.8–1.2$), and the $R_{sub}$ and $f_{cut}$ parameters summarised in Table 1. The background rejection and the boundaries of the 68% signal mass windows obtained with a subset of trimming configurations for the range 350 < $p_T$ < 500 GeV are shown in Fig. 4 for anti-$k_t$, $R = 1.0$ and C/A, $R = 1.0$ jets. The systematic uncertainties resulting from the uncertainty on the jet mass and energy scale (described in detail in Sect. 7.5) are provided to give the reader an idea of the relevance of the differences in performance between the grooming configurations.

The following characteristics are noted:

- C/A and anti-$k_t$ jets have a similar performance under the same configurations.
- The larger values of $f_{cut}$ can lead to significantly lower background efficiency.
- The dependence of the performance on $R_{sub}$ is less significant, but the background efficiency does decrease somewhat for smaller $R_{sub}$ values.

Based on the performance of these algorithms, the trimming implementations considered for further investigation are given in Table 2. Although promising, configurations with $R_{sub} = 0.1$ are not pursued further in these studies, as this size is approaching the limiting granularity of the hadronic

<table>
<thead>
<tr>
<th>Initial algorithm</th>
<th>$R$</th>
<th>$f_{cut}$ (%)</th>
<th>$R_{sub}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-$k_t$</td>
<td>1.2</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>anti-$k_t$</td>
<td>1.0</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>C/A</td>
<td>1.0</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>anti-$k_t$</td>
<td>0.8</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>C/A</td>
<td>0.8</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>C/A</td>
<td>0.6</td>
<td>5</td>
<td>0.2</td>
</tr>
</tbody>
</table>
ATLAS

Pythia8 simulation

= 8 TeVs

Pruned jets, R=1.0

ATLAS

< 500 GeV

| < 1.2

C/A-prun.

Rcut = 1/8

10%

ATLAS

Zcut = 15%

20%

γ5 = 8 TeV

Shaded bands indicate extent of 68% mass windows

Bars indicate stat. + baseline syst. uncertainty from JMS+JES

W mass

ε_{QCD}

10% background efficiency

Fig. 5  Mass windows and background efficiencies for various configurations of pruning (R = 1.0 shown). The baseline systematic uncertainty on the background efficiency for the \( p_T \) bin in question (the range 350 < \( p_T \) < 500 GeV is shown here) is calculated by varying the jet mass scale (JMS) and jet energy scale (JES) by \( \pm 1\sigma \) for a representative jet collection. For pruning, this representative configuration is \( R_{\text{cut}} = \frac{1}{2} \) and \( Z_{\text{cut}} = 15\% \). The star indicates the favoured pruning configuration for \( W \)-tagging, as detailed in Sect. 6.4

Table 3  The best pruning configurations for \( W \)-tagging with each \( R \) based on the first stage of the MC-based optimisation studies

<table>
<thead>
<tr>
<th>Initial algorithm</th>
<th>( R )</th>
<th>( Z_{\text{cut}} ) (%)</th>
<th>( R_{\text{cut}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>10</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>15</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>1.0</td>
<td>10</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>1.0</td>
<td>15</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>0.8</td>
<td>10</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>0.8</td>
<td>15</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>0.6</td>
<td>10</td>
<td>0.5</td>
</tr>
<tr>
<td>C/A</td>
<td>0.6</td>
<td>15</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Split-filtering:

Split-filtering is studied with C/A jets with \( R = 1.2 \) and 1.0, and various values of the parameters \( \sqrt{y_{\text{min}}} \), \( R_{\text{sub}} \) and \( \mu_{\text{max}} \). The background efficiencies and 68% signal mass windows obtained with a subset of split-filtering configurations for the range 350 < \( p_T \) < 500 GeV are shown in Figs. 6 and 7.

Observations from the results of these studies include the following:

- Larger \( \sqrt{y_{\text{min}}} \) values tend to result in lower background efficiencies.
- The performance has a dependence on \( \sqrt{y_{\text{min}}} \) and the optimal requirement varies with jet \( p_T \). For \( y_{\text{cut}} \geq 0.09 \), the background efficiency is relatively stable.
Fig. 6 Mass windows and background efficiencies for various additional configurations of split-filtering ($R = 1.2$ shown). The baseline systematic uncertainty on the background efficiency for the $p_T$ bin in question (the range $350 < p_T < 500$ GeV is shown here) is calculated by varying the jet mass scale (JMS) and jet energy scale (JES) by ±1σ for a representative jet collection. For split-filtering, this representative configuration is $\mu_{\text{max}} = 1$, $R_{\text{sub}} = 0.3$ and $y_{\text{cut}} = 15\%$.

Fig. 7 Mass windows and background efficiencies for various configurations of split-filtering ($R = 1.2$ shown). The baseline systematic uncertainty on the background efficiency for the $p_T$ bin in question (the range $350 < p_T < 500$ GeV is shown here) is calculated by varying the jet mass scale (JMS) and jet energy scale (JES) by ±1σ for a representative jet collection. For split-filtering, this representative configuration is $\mu_{\text{max}} = 1$, $R_{\text{sub}} = 0.3$ and $y_{\text{cut}} = 15\%$. The star indicates the favoured split-filtering configuration for $W$-tagging, as detailed in Sect. 6.4.

- For $\sqrt{y_{\text{min}}} > 0.09$, there is not a strong dependence of the performance on $R_{\text{sub}}$ or $\mu_{\text{max}}$. A total of 11 split-filtering jet collections are considered for further study, all with $\mu_{\text{max}} = 100\%$ and $R_{\text{sub}} = 0.3$. These are given in Table 4.
6.2 Pileup dependence

The influence of pileup on the reconstructed groomed jets is examined during the first stage of algorithm optimisation, and configurations that show large susceptibility to pileup after grooming are discarded. There are a number of methods [61,65–71] available for reducing the effects of pileup, either on their own or combined with grooming; these techniques are not considered in this study. Most grooming configurations almost completely remove the effects of pileup from the mean jet mass as illustrated in Fig. 8 in which the correlation between average jet mass \( \langle M \rangle \) and number of reconstructed primary vertices is shown. The significant correlation between the average ungroomed jet mass and the number of reconstructed primary vertices for a well-behaved trimming configuration is examined during the first stage of algorithm optimisation, and configurations almost completely remove the effects of pileup from the mean jet mass as illustrated in Fig. 8 in which the correlation between average jet mass \( \langle M \rangle \) and number of reconstructed primary vertices is shown. The significant correlation between the average ungroomed jet mass and the number of reconstructed primary vertices is absent for trimmed jets in both signal and background.

The pileup dependence of the mean jet mass obtained with all 27 of the grooming configurations selected for stage two of the optimisation studies is shown in terms of the fitted slope of \( \delta \langle M \rangle / \delta \text{NPV} \) in Fig. 9 for the \( p_T \) range 350–500 GeV. In general, the average masses of jets with larger radii have a more pronounced pileup dependence, and the trimmed jet mass has a weaker pileup dependence than that obtained with the pruning and split-filtering algorithms. For all jet algorithms, the pileup dependence is much reduced with respect to that of ungroomed jets.

<table>
<thead>
<tr>
<th>Initial algorithm</th>
<th>( R )</th>
<th>( \sqrt{\text{NPV}} ) (%)</th>
<th>( \mu \text{max} ) (%)</th>
<th>( R_{\text{sub}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>0</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>4</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>9</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>12</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>1.2</td>
<td>15</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>0.8</td>
<td>0</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>0.8</td>
<td>4</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>0.8</td>
<td>9</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>0.6</td>
<td>0</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>0.6</td>
<td>4</td>
<td>100</td>
<td>0.3</td>
</tr>
<tr>
<td>C/A</td>
<td>0.6</td>
<td>9</td>
<td>100</td>
<td>0.3</td>
</tr>
</tbody>
</table>

6.3 Performance of substructure variables

Substructure variables are introduced in Sect. 2.2. A brief description of the variables studied in this analysis are listed below:

- The energy correlation ratios \( C_2^{(\beta)} \) and \( D_2^{(\beta)} \), described in detail in Sect. 2.2.
- The N-subjettiness ratios \( \tau_2 \), \( \tau_2^{\text{wa}} \), \( \tau_2^{12} \), and \( \tau_2^{12} \) are also described in detail in Sect. 2.2.
- Planar flow [19], \( P \), is a measure of how uniformly distributed the energy of a jet is, perpendicular to its axis.
- The angularity, \( a_3 \), distribution is expected to peak sharply at values close to zero for a balanced two-body decay, such as that of a W boson, while a broader tail is expected for jets initiated by quarks and gluons. The general formula for the mass-normalised angularity can be found in Ref. [19].
- Splitting scales [20] are calculated, within the jet clustering algorithm, and can be calculated for any jet using its constituents. The splitting scale \( \sqrt{d_{12}} \), is calculated for a jet (re)clustered with the \( k_t \)-clustering algorithm, and is the \( k_t \) distance between the two proto-jets of the final clustering step.
- The variable \( \sqrt{z_{12}} \) [21] is a variant on the original splitting scale \( \sqrt{d_{12}} \) which uses the jet mass.
- The momentum balance [5], \( \sqrt{\tau_{12}} \), and mass-drop fraction \( \mu_{12} \), are defined at the first de-clustering step that satisfies a minimum mass-drop and momentum balance requirement, and are only available for those jets that are groomed with the split-filtering algorithm.
- The soft-drop algorithm [22] declusters the jet, following the path of highest \( p_T \) through the clustering history. A condition is defined:
Fig. 9 A summary of the pileup dependence $\delta(M)/\delta\text{NPV}$ for the 27 jet configurations selected for further study. The top panel shows the dependence for signal $W$-jets, the bottom panel for background multijets, and from left to right shows decreasing values of the initial jet radius parameter, $R$. Each value of $\delta(M)/\delta\text{NPV}$ is the slope of a straight line fit of $\langle M \rangle$ versus NPV, an example of which is shown in Fig. 8.

### Table 5

The soft-drop levels $L_{\text{SD}}(\beta)$ are defined as the highest level of balance in the jet history

$$z_g > z_{\text{cut}} \times r_g^{\beta}, \quad (10)$$

where the fractional momentum of the softest of the two branches is $z_g = \frac{\text{min}(p_T1, p_T2)}{p_T1 + p_T2}$, and the fractional angular separation of the two branches (with respect to the $R$ parameter of the initial jet algorithm, $R_0$) is $r_g = \frac{\Delta R_1}{R_0}$. Nine values of the $z_{\text{cut}}$ parameter between 4 and 20% are explored here, given in Table 5. The $\beta$ values chosen here are $-1.0, -0.75, -0.5$. The starting condition of Eq. 10 with $z_{\text{cut}} = 4\%$ is applied to the first step in the declustering. If this condition is not satisfied, the algorithm continues to the next step in the jet’s clustering history, and so on, checking if the condition is satisfied at any point. If it is not, the ‘soft-drop-level’, $L_{\text{SD}}(\beta)$ is zero. If this condition is satisfied, $L_{\text{SD}}(\beta) = 1$. The algorithm then remains at this point in the clustering history and asks for the same condition with the harder momentum condition, $z_{\text{cut}} = 6\%$. If this condition is not satisfied, the algorithm continues to the next step in the jet’s clustering history, and so on.

- The dipolarity [25], $D$, is a measure of the colour flow between two hard centres within a jet.
- Jet shape variables are computed in the centre-of-mass frame of a jet, which can increase the separation power between $W$-jets and jets in multijet events. Sphericity, $S$, aplanarity, $A$, and thrust minor and major, $T_{\text{min}}, T_{\text{maj}}$, already used in a previous ATLAS measurement [26], as well as the ratio of the second to zeroth order Fox–Wolfram moments, $R_2^{\text{FW}}$ [72] are considered.
- For a jet clustered with a given recombination jet clustering algorithm, the Q-jets technique [27] reclusters the jet many times for each step in the clustering. Following this, any jet observable, such as the mass, will have a distribution for a given jet. The Q-jets configuration optimised in Ref. [28] is adopted in this study. The high mass in $W$-jets tends to persist during the re-clustering while the mass of QCD jets fluctuates. A sensitive observable to this trend is the coefficient of variation of the mass distribution for a single jet, called the volatility [27,28], $\nu^2$. The superscript $\alpha$ denotes the rigidity, which controls the sensitivity of the pair selection to the random number generation used in the clustering.

For all 27 jet collections and grooming algorithms described in Sect. 6.1, the full list of substructure variables described above are computed. The distributions of the three variables $\tau^{\text{wa}}_{21}$, $C_2^{(\beta=1)}$ and $D_2^{(\beta=1)}$ are shown in Figs. 10, 11, 12 for anti-$k_t$, $R = 1.0$ jets trimmed with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$, after applying the 68% signal efficiency mass window requirement. This grooming algorithm is referred to in the remainder of this paper as ‘R2-trimming’. At this stage no jet mass calibrations have been applied for any of the grooming configurations. Also shown are the correlations between the
Fig. 10 The $C_2^{(β=1)}$ variable, for R2-trimmed jets: a distributions in signal (blue solid line) and background (red dashed) in MC in the range $350 < p_T < 500$ GeV, obtained after applying the 68% signal efficiency mass window requirement (discussed in Sect. 6.1; b correlation with the leading jet’s mass in (left) multijet background and (right) $W$-jet signal events. No truth-matching requirements are made, so the signal events can contain background jets as well as $W$-jets. The vertical line corresponds to the value of the cut providing a combined 50% efficiency for grooming and tagging (corresponding to a tagging-only efficiency of 50%/68% = 73.5%).

Jet mass and each of these variables, shown separately for the $W$-jet signal and multijet background, in both cases before applying the 68% signal efficiency mass window requirement. No truth-matching between the subjets and the quarks from the $W$ decay is required, such that the signal sample contains both full $W$-jets and jets made of fragments of the $W$-decay, generally because the $W$-decay is not completely captured in the $R = 1.0$ jet. The background jets within the signal sample are particularly visible in the low-mass region of Fig. 10b, where the distributions echo those seen in the background sample.

The background rejection power (1/background efficiency) is shown in Fig. 13 for the $\epsilon_{W}^{G&T} = 50\%$ efficiency working point for each substructure variable inside the mass window determined by the grooming, and for each of the 27 grooming configurations, for the range $350 < p_T < 500$ GeV.

In addition to calculating the background rejection power at a particular signal efficiency working point, full rejection versus efficiency curves (so-called Receiver Operating Characteristic ‘ROC’ curves) are produced for each combination. An example showing the relationship between the $W$-jet signal efficiency and the multijet background rejection for the range $350 < p_T < 500$ GeV is shown in Fig. 14. The maximal efficiency value for each algorithm is by definition 68%, since the tagging criteria are applied after requiring
the jet mass to be within the mass window defined by the grooming.

6.4 Summary of grooming and substructure in MC

Four grooming configurations, given in Table 6, show consistently high performance in all $p_T$ bins. The jet $\eta$, mass and energy calibrations are derived for these four using a simulation-based calibration scheme, used as the standard one by ATLAS in previous studies [10]. The mass window sizes for calibrated jets, the background efficiencies for $\epsilon_W = 68\%$ and the $\delta(M)/\delta$NPV in the range $200 < p_T < 350$ GeV are also given in Table 6.

Since the first algorithm in Table 6 is the only one of the four with negligible pileup dependence across all $p_T$ ranges (the central $p_T$ range only is shown in Fig. 9), it is adopted for all successive studies.

The best substructure variables for use with R2-trimmed jets at the $\epsilon_{W}^{G&T} = 50\%$ working point, providing background efficiencies $\epsilon_{QCD}^{G&T} \sim 2\%$ (background rejection power $\sim 50$, in terms of Fig. 13) for jets with $p_T > 350$ GeV, are given in Table 7. Studies of the R2-trimmed grooming configuration and the three preferred substructure variables are described in the next section, where the results obtained from Monte Carlo simulations are compared to data.
Fig. 12 The $\tau_{21}^{\text{wta}}$ variable, for R2-trimmed jets: a distributions in signal (blue solid line) and background (red dashed) in MC in the range $350 < p_T < 500$ GeV, obtained after applying the 68 % signal efficiency mass window requirement (discussed in Sect. 6.1); b correlation with the leading jet’s mass in (left) multijet background and (right) W-jet signal events. No truth-matching requirements are made, so the signal events can contain background jets as well as W-jets. The vertical line corresponds to the value of the cut providing a combined 50 % efficiency for grooming and tagging (corresponding to a tagging-only efficiency of 50%/68% = 73.5 %)

7 Detailed studies of selected techniques in data

This section describes a comparison of the W-jet and multijet tagging efficiencies measured using three tagging variables $C_2^{(\beta=1)}$, $D_2^{(\beta=1)}$ and $\tau_{21}^{\text{wta}}$, computed for the leading R2-trimmed jet in data and MC.

In data, a relatively pure sample of boosted, hadronically decaying W bosons can be obtained from decays of top quark pairs in the lepton-plus-jets decay channel: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell\nu q\bar{q}bb$. The selection requirements detailed in Sect. 5 are applied to events in data and MC, where relevant. The composition of the data and MC samples introduced in Sect. 4 is discussed in Sect. 7.1. Details of the event topology differences between the $t\bar{t}$ final state examined in this section and the $W'$ final state used in the preliminary optimisation studies are given in Sect. 7.2. The systematic uncertainties are discussed in Sect. 7.3, and the distributions of mass and substructure variables in data and MC are presented in Sect. 7.4. The signal and background efficiency estimation procedures and their uncertainties are detailed in Sect. 7.5. A summary of the signal and background tagging efficiencies measured in data and compared to MC is given in Sect. 7.6.

In all the following studies, events are categorised according to the leading, reconstructed R2-trimmed jet $p_T$ in three ranges: [200, 250], [250, 350], and [350, 500] GeV. This
Fig. 13 For jets with $350 < p_T^{\text{Truth}} < 500\text{GeV}$, the background rejection factors corresponding to a 50 % efficiency are shown for all possible combinations between the 27 grooming configurations and 26 substructure variables, after applying the uncalibrated groomed mass window requirement that provides a 68 % signal efficiency. The error shown are the result of the finite Monte Carlo sample size.

characterisation differs from that used in the first stage of the optimisation in Sect. 6, which uses ungroomed C/A, $R = 1.2$ truth jets and different ranges; the selection is extended only to 500 GeV here because there are insufficient data above 500 GeV in the 2012 dataset. The lowest $p_T$ range used in the preliminary optimisation stage, [200, 350] GeV, is now divided in two, since the 2012 dataset has an abundance of top-decay events in this range.

7.1 Sample compositions and definitions

Signal W-jets are extracted from $\tilde{t}\tilde{t}$ events in data and in the MC samples detailed in Sect. 4. The $\tilde{t}\tilde{t}$ production cross-
cases where the 
function from the 
section is scaled to match the value obtained from NNLO calculations [73]. An additional reweighting is then applied to the \( t\bar{t} \) MC using the generator-level \( p_T \) of the top quark and the \( p_T \) of the \( t\bar{t} \) system to reproduce the \( p_T \)-dependence of the measured cross-section [74].

The dominant backgrounds to the \( t\bar{t} \) event topology come from \( t\bar{t} \) production where there is only partial reconstruction of the \( W \) boson decay, with or without contamination from radiation outside of the top quark decay (such as hard gluon emission, non-tagged \( b \)-jets). Generator-level information from the \( t\bar{t} \) and \( Wt \) samples is used to distinguish the cases where the \( W \) candidate jet is matched to a genuine \( W \) boson or to other jets (referred to as top quark background events). An event is categorised as belonging to the \( W \) signal when both partons from the \( W \) boson decay are within \( \Delta R = 1.0 \) of the jet axis; otherwise, the event is labelled as non-\( W \) background.

The leading non-top background process is production of \( W \) bosons in association with jets. The \( W \) background.

Table 6 The four favoured grooming configurations along with their mass windows (derived using calibrated jets), background efficiencies, and pileup dependence for \( \epsilon_W^G = 68 \% \) in the range 200 < \( p_T \) < 350 GeV

<table>
<thead>
<tr>
<th>Grooming configuration</th>
<th>( \epsilon_W^G ) = 68 % mass range (GeV)</th>
<th>( \epsilon_QCD ) (%)</th>
<th>( \delta(M)/\delta NPV ) (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-( k_t ), ( R = 1.0 ) trimmed ( f_{cut} = 0.05 ), ( R_{sub} = 0.2 )</td>
<td>61–93</td>
<td>11</td>
<td>0.1–0.2</td>
</tr>
<tr>
<td>anti-( k_t ), ( R = 1.0 ) trimmed ( f_{cut} = 0.05 ), ( R_{sub} = 0.3 )</td>
<td>65–99</td>
<td>16</td>
<td>0.5–0.6</td>
</tr>
<tr>
<td>C/A, ( R = 1.0 ) pruned ( Z_{cut} = 0.15 ), ( R_{cut} = 0.5 )</td>
<td>59–111</td>
<td>16</td>
<td>0.9–1.1</td>
</tr>
<tr>
<td>C/A, ( R = 1.2 ) split-filt ( \sqrt{2} ) split-filt = 0.15, ( R_{sub} = 0.3 )</td>
<td>63–103</td>
<td>13</td>
<td>0.1–0.3</td>
</tr>
</tbody>
</table>

Table 7 The masses windows for calibrated \( R_2 \)-trimmed jets that provide \( \epsilon_W^G = 68 \% \), and the requirements on the three substructure variables that result in the lowest background efficiencies \( \epsilon_QCD \), when combined with the mass windows to provide \( \epsilon_W^{G\&T} = 50 \% \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tagging criteria in ( p_T ) range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \epsilon_W^G ) = 68 % mass range</td>
<td>200–350 GeV</td>
</tr>
<tr>
<td>( \epsilon_W^{G&amp;T} = 50 % C_2^{(beta=1)} )</td>
<td>61–93 GeV</td>
</tr>
<tr>
<td>( \epsilon_W^{G&amp;T} = 50 % D_2^{(beta=1)} )</td>
<td>&lt;0.18</td>
</tr>
<tr>
<td>( \epsilon_W^{G&amp;T} = 50 % \tau_{21} )</td>
<td>&lt;1.14</td>
</tr>
<tr>
<td>( \epsilon_W^{G&amp;T} = 50 % \tau_{21} )</td>
<td>&lt;0.32</td>
</tr>
</tbody>
</table>

Fig. 14 For jets with 350 < \( p_T^{\text{Truth}} < 500 \) GeV, the signal efficiency versus background rejection power “ROC” curve for selected tagging variables (combined with the uncalibrated groomed mass window) on a subset of high-performance algorithms is shown. The inset at 68 % signal efficiency is a result of the 68 % mass window. The inset enlarges the high-efficiency region
relative contributions from each jet flavour are found using the charge asymmetry and the flavour fractions are fixed for W+jets events in the signal region before the b-tagged jet requirement is applied. Finally, an overall normalisation is obtained by scaling the simulated W+jets charge asymmetry to match the charge asymmetry in data, after other charge-asymmetric backgrounds are accounted for using MC.

The contribution from multijet events to the sample composition is estimated by using loose lepton identification criteria and deriving the contribution of non-prompt leptons using the method matrix [76,77]. This method relies on the fact that the tight lepton identification criteria selects primarily prompt leptons, while loose leptons that do not satisfy the tight criteria are primarily from backgrounds. The probabilities for a non-prompt lepton from multijet production which satisfies the loose/tight identification criteria are measured from data in control regions dominated by multijet events, with prompt-lepton contributions subtracted based on MC. The corresponding probabilities for a lepton from prompt sources (such as W bosons) which satisfies the loose/tight identification criteria are derived from MC samples, corrected using data-to-MC correction factors derived from Z → ℓℓ events. Once the fraction of events satisfying the different identification criteria is known, an event weight is calculated and applied to data events with the loosened lepton identification criteria to provide an estimate of the multijet contribution.

7.2 Event topology effects in Monte Carlo simulations

The preliminary MC-based optimisation studies in Sect. 6 use a signal composed of well-isolated W-jets from the hypothetical process W′ → WZ → qqℓℓ provided by PYTHIA and a background sample of jets initiated by light quarks or gluons, also provided by PYTHIA. In the following sections, efficiencies are measured in data, so the t¯t final state is used as a source of W-jets. As described in Sect. 4, the main t¯t signal processes are provided by either POWHEG-BOX + PYTHIA or MC@NLO + HERWIG and the multijet background is provided by PYTHIA or by HERWIG++.

Despite the backgrounds in both event topologies being PYTHIA multijets, they are different in that the background efficiencies obtained in data include a leading-jet minimum pT requirement of 450 GeV in order to ensure full efficiency with respect to the trigger used. With this selection, the lower pT ranges, [200, 250] and [250, 350] GeV, are composed entirely of sub-leading jets, and the highest pT bin, [350, 500] GeV, is a mixture of leading and sub-leading jets. Jets softer than the sub-leading jet are not considered. In the background sample used for the studies in Sect. 6 there is no comparison with data, thus there are no trigger requirements and the leading jet is always shown. A higher average jet mass is observed in the leading + sub-leading jet selection with the leading-jet selection. This in turn leads to a higher background efficiency for the studies summarised in Sect. 7 than for those in Sect. 6.4. These differences are relevant in that leading and sub-leading jets have different flavour compositions (light-quark versus gluon). Gluon-initiated jets have higher average mass than quark-initiated jets [78].

The signal event topologies are more obviously different, with the W′ process producing potentially more isolated W-jets than those found in the t¯t final state. The W bosons produced in the W′ decay are also generally longitudinally polarised, making them potentially easier to distinguish from multijet background than W-jets from top decays, which are produced in both the longitudinal and transverse modes [11, 63].

The signal efficiency versus background rejection curves in the two different event topologies, including the differences in both signal and background, are shown in Fig. 15. The curves for tagging W-jets from the W′ against a leading-jet background indicate better performance in this event topology, with the magnitude of the difference depending on the substructure variable used for tagging. Figure Fig. 16 shows the curves again, but this time the leading jet from the PYTHIA multijet background is used in both cases, thus removing the differences in background efficiencies, and isolating the differences resulting from the different signal event topologies. With identical background compositions, the performance is generally slightly better in the POWHEG-BOX t¯t sample.
The mass distributions for the different signal and background samples are compared in Fig. 17 for the lowest and highest \( p_T \) ranges. The signal distributions also include the R2-trimmed leading-jet mass from \( tt \) events provided by MC@NLO + HERWIG. The mass shape differences are less pronounced at higher \( p_T \), although the difference in \( \epsilon_W \) for the different signal event topologies is still a non-negligible 10% even in the highest \( p_T \) range.

7.3 Systematic uncertainties

The sources of systematic uncertainty that are common to both the signal and background efficiency measurements include the jet mass scale (JMS), jet energy scale (JES), jet energy resolution (JER) and jet substructure variable (JSS).

The uncertainty on the JER is taken from previous studies [79] and is parameterised as a function of \( p_T \). The size of JER uncertainty is approximately 10% for the \( p_T \) ranges presented here. The uncertainty on JMR is also taken from previous studies [10], where it was determined from the data/MC variations in the widths of the \( W \)-jet mass peaks in \( tt \) events, and is fixed at 20%. The JMS, JES and JSS are varied up and down by \( \pm 1\sigma \), using the standard deviation derived from the double-ratio method; this is described in detail below using the JSS as an example.

The systematic uncertainty on the JSS is needed in order to derive the full systematic uncertainties on the signal and background efficiencies. Uncertainties are derived using in-situ methods by comparing the measured calorimeter jet energy, mass and substructure variables to the same quantities measured by well-calibrated and completely independent detectors in both data and MC, using the double ratio:

\[
\frac{X^{\text{jet}}/X^{\text{ref}}}_W / \frac{X^{\text{jet}}/X^{\text{ref}}}_\text{MC},
\]

where \( X \) denotes a jet variable. In this case, track-jets are used as reference objects, since tracks from charged hadrons are well-measured and are independent of the calorimeter. In addition, the use of track-jets, where tracks are required to come from the hard scattering vertex, suppresses pileup effects. A geometrical matching in the \( \eta-\phi \) plane is applied to associate track-jets with calorimeter-jets. This approach was widely used in the measurement of the jet mass and substructure properties of jets in the 2011 data [10]. Performance studies have also shown that there is excellent agreement between the measured positions of clusters and tracks in data, indicating no systematic misalignment between the calorimeter and the inner detector. This technique achieves a precision of around 3–7% in the central detector region, which is dominated by systematic uncertainties arising from the inner-detector tracking efficiency and MC modelling uncertainties of the charged and neutral components of jets.

The double ratio of Eq. (11) is computed for two different MC generators, PYTHIA and HERWIG++, and the largest disagreement between data and each of the MC generators is taken as a modelling uncertainty. The total uncertainty is then obtained by adding in quadrature this modelling uncertainty to the tracking efficiency uncertainty. Specific uncertainties for tracks inside the core of dense jets are not needed here, because only jets with \( p_T < 1 \) TeV are considered. The scale uncertainties for the jet energy, mass and substructure variables are derived in ranges of the \( p_T, \eta, \) and \( M/p_T \) of the reconstructed calorimeter jet.

Figure 18 shows a set of six representative distributions for \( c_2^{(\beta=1)}, D_2^{(\beta=1)} \), and \( r_{21}^{\text{GTA}} \) in the range 350 < \( p_T < 500 \) GeV. The mean values of the single-ratio \( X^{\text{jet}}/X^{\text{ref}} \) distributions are shown as a function of the jet mass, along with the distributions of \( X^{\text{jet}}/X^{\text{ref}} \) themselves within the relevant \( \epsilon_W \) ~ 68% mass window.

Large discrepancies between data and MC are observed for low-mass jets, while for masses around 80 GeV the data/MC agreement is within 5%. In the distributions of \( X^{\text{jet}}/X^{\text{ref}} \) it is noticed that while the tails of the ratio distributions show discrepancies between data and the MC, the agreement is good for values of the ratio close to one, which represents the large majority of events. In summary, the scale uncertainty of the three jet substructure variables ranges between 1 and 5% in the different kinematic regions.
Additional, sub-dominant systematic uncertainties come from MC sources listed in Table 9 and described in Sect. 7.5 in terms of the uncertainty on the final measured signal and background efficiencies. The full systematic uncertainty on the mass and substructure variables are obtained by adding each of the scale, resolution, statistical and MC uncertainties in quadrature.

7.4 Mass and substructure distributions in $t\bar{t}$ events

The jet mass distribution for the leading R2-trimmed jets in events satisfying the pre-selection criteria in Sect. 5 are shown in Fig. 19. The data and events in POWHEG-BOX + PYTHIA and MC@NLO + HERWIG simulations agree within the uncertainties detailed in Sect. 7.3. Distributions of the three tagging variables $C_2^{\beta=1}$, $D_2^{\beta=1}$ and $D_2^{\alpha_{W}}$ are shown for the same pre-selection criteria, before and after making the relevant $\epsilon_W^{G&T} = 68\%$ mass window requirements for the $p_T$ range in question, in Fig. 20. These variables are used to define medium and tight tagging criteria, where the medium working point provides a signal efficiency of $\epsilon_W^{\alpha_{W}} = 50\%$ and the tight working point provides $\epsilon_W^{G&T} = 25\%$.

The jet mass distributions of the $W$ boson candidates satisfying or failing to satisfy the medium signal efficiency requirement for each of the three substructure variables are shown for the same pre-selection criteria, before and after the medium signal efficiency requirement for each of the three substructure variables are
Fig. 18  **Left** distributions of the mean calorimeter-jet / track-jet ratios as a function of the R2-trimmed jet mass for three tagging variables. **Right** distribution of these ratios for the three variables in data compared to the PYTHIA and HERWIG++ models. a, b C\((\beta=1)\)\(_2\), c, d \(D_{\tau}^{(\beta=1)}\) and e, f \(\tau_{wta}^{(\beta=1)}\). The distributions are shown for R2-trimmed jets in the central calorimeter region, \(|\eta| < 1.2\) and in the range 350 < \(p_T\) < 500 GeV. The data/MC comparisons (the ‘double-ratios’) for PYTHIA (blue dashed) and HERWIG++ (red dotted) are shown in the lower panel of each plot.

shown in Fig. 21. The mass distribution for jets failing the \(C_{\tau}^{(\beta=1)}\) tagger (Fig. 21a) is notably different from the mass distributions for jets that fail the \(D_{\tau}^{(\beta=1)}\) and/or \(\tau_{wta}^{(\beta=1)}\), with a significantly higher mass peak and a low-mass tail that is conspicuous in its absence. This effect can be understood by referring back to Fig. 10b: the correlation between
the mass and $C_2^{(β=1)}$ is strong for background jets with low masses, while there is no clear correlation in the signal mass region. This means that the $C_2^{(β=1)}$ variable performs well when combined with a mass window, but is not very effective without the mass constraint.

7.5 Signal and background efficiencies and uncertainties

Background efficiencies are measured in a multijet-enriched sample of data, using the large-$R$ trigger and event selection described in Sect. 5.

The systematic uncertainties on the background efficiency measurements in multijet events are summarised in Table 8. The uncertainties are propagated coherently through to the measurement and then added together in quadrature. The background efficiency uncertainty due the JSS uncertainty can be as large as $\sim$25% for jets with $p_T > 500$ GeV and is about 15–20% in the lower $p_T$ ranges for the scale uncertainty on $D_2^{(β=1)}$. The background efficiency uncertainties from the JMS are, in general, larger than those from the JES and are of the order of 6–10 and 2–9%, respectively. The impact of JER and JMR uncertainties is much smaller than that of the scale uncertainties.

Signal efficiencies are extracted from data by performing a template fit to the mass distributions of jets that satisfy

or fail to satisfy the requirement on the given tagging variable. The signal template is constructed using the POWHEG-BOX + PYTHIA $t\bar{t}$ events, requiring that both partons from the $W$ boson decay in the event record are within $ΔR = 1.0$ of the jet axis. The mass templates for the background are composed of decays of $W$ bosons from top quarks, where not all the decay products fall inside the jet cone, and the other non-$W$ backgrounds are also estimated using POWHEG-BOX + PYTHIA. The normalisations of both templates are allowed to float.

The statistical uncertainty on the efficiency measurement in data includes the statistical uncertainty of the templates. For most sources of systematic uncertainty, a variation of the fit is performed with templates modified by ±1σ. In the case of the JMS, this variation is between ±0.5σ and ±1.0σ; this reduction in the uncertainty with respect to that obtained with the standard double-ratio technique is made possible by fitting the mass distributions in data to a number of different templates. The templates are obtained by shifting the jet mass up and down by fractions (0.25–1.0) of σ. The $χ^2/ndf$ quality of each template is calculated, and a parabolic fit performed to the $χ^2/ndf$ as a function of the fraction of σ. The fraction of σ that results in a one unit shift from that which minimises $χ^2/ndf$ is used as the uncertainty on the JMS for the signal efficiency calculation.

The full set of contributions to the systematic uncertainty on the signal efficiency is summarised in Table 9, after applying the mass and $D_2^{(β=1)}$ medium tagging requirements. As in the background efficiency uncertainty estimate, the JSS contributes the largest uncertainty on this efficiency, varying between 3 and 5% for the $D_2^{(β=1)}$ scale. The contribution from the JMR is $\sim$3%. The contribution from JER is less significant than JMR, being negligible in the lowest $p_T$ bin and $\sim$1% for jets with $250 < p_T < 500$ GeV. The contribution from JMS variations is also $\sim$1% (symmetrised as a result of the profiling technique) and increases to $\sim$10% in the highest $p_T$ range (350 < $p_T$ < 500 GeV). The uncertainty from the JES is around 2–4%.

In addition to the scale and resolution uncertainties, two other types of uncertainty are considered for the signal efficiency measurement: (a) $t\bar{t}$ modelling—initial-state radiation (ISR), final-state radiation (FSR), and generator uncertainty; (b) the normalisation of the main background sources—multijet, $W$+jets, partial-$W$ and non-$W$ in single top and $t\bar{t}$.

The generator uncertainty is taken into account as the difference between the signal efficiency measurement using the MC@NLO + HERWIG mass templates for the signal instead of the default POWHEG-BOX + PYTHIA ones. These uncertainties are between 1 and 3%. The modelling uncertainty of the QCD radiation is estimated using ACERMC [80] v3.8 plus PYTHIA v6.426 MC samples by varying the parame-
Fig. 20 Distributions of the W candidate jet substructure variables before (left) and after (right) the $\epsilon_{GW} = 68\%$ mass window for selected lepton+jets $t\bar{t}$ events in data and POWHEG-BOX + PYTHIA MC for the combined electron and muon channel. a, b, C$^2_{(=1)}$, c, d, D$^2_{(=1)}$ and e, f. Data points are shown with statistical uncertainties, and the combined MC is shown with full systematic and statistical uncertainties. The lower panels show the data/MC ratios, with the statistical uncertainty on the MC given in black forward-slashed bands, and the full systematic uncertainty given in the blue, back-slashed bands.
Fig. 21 The distribution of the $W$ candidate mass for R2-trimmed jets failing (left) and passing (right) the selection corresponding to $\epsilon_{G&T} = 50\%$ for the combined electron and muon channel in the $p_T$ range 200–250 GeV, without application of the mass cut. In a, b the variable used for selection is $C_2^{(\beta=1)}$, in c and d it is $D_2^{(\beta=1)}$, and in e, f it is $\tau_2^{WS}$. Data points are shown with statistical uncertainties, and the combined MC is shown with full systematic and statistical uncertainties. The lower panels show the data/MC ratios, with the statistical uncertainty on the MC given in black forward-slashed bands, and the full systematic uncertainty given in the blue, back-slashed bands.
ters controlling the ISR and FSR in a range consistent with a previous ATLAS measurement [81]. The resulting uncertainties on the signal efficiency increase with jet $p_T$ and are 2–6%. The normalisation uncertainties for the main background sources are evaluated using a ±1σ variation of the cross-section. The normalisation uncertainties are negligible with respect to the scale and resolution uncertainties, and for the $t\bar{t}$ signal and W+jets background they are <1%.

7.6 Summary of W boson tagging efficiencies in data and MC

The $W$-jet tagging efficiency in $t\bar{t}$ events using the R2-trimmed jet mass window and the medium and tight $C_2^{(β=1)}$ selections is measured in top-enriched data and in MC provided by POWHEG-BOX + PYTHIA and MC@NLO + HERWIG. The background efficiency with the same selection is measured in multijet-enriched data and in PYTHIA and HERWIG++ simulations. The results of these measurements are shown in Fig. 22. In both the signal and background efficiency distributions, the ratio of data to each of the two MC models is shown in the lower panels. The corresponding signal and background efficiency distributions for $D_2^{(β=1)}$ and $t\bar{t}^{(wt)}$ are shown in Figs. 23 and 24 respectively. Systematic errors from background modelling are added for the signal data points, while no background modelling is involved in the derivation of background efficiencies, whose points only show statistical error. Good agreement is observed between data and predictions.

The signal efficiency at the medium working point is not exactly 50% because the selection requirements for the $e_W^{G\&T} = 50\%$ working point are calculated using $W$-jets from $W^+ → WZ → qq\ell\ell$ events, and are applied here to $W$-jets in $t\bar{t}$ events.

The data points are the result of fits using templates extracted from POWHEG-BOX + PYTHIA; the difference with respect to the results that would be obtained using templates from MC@NLO + HERWIG is added in quadrature as an additional source of systematic uncertainty.

The $D_2^{(β=1)}$ tagger has the smallest background efficiency for the medium and tight working points in all $p_T$ ranges except for the lowest, 200 <$p_T<$ 250 GeV. The background efficiencies decrease with increasing $p_T$, with the exception of the $C_2^{(β=1)}$ tagger, for which the background efficiency increases for jets in the range 250 <$p_T<$ 350 GeV. This behaviour can be explained by the stronger

Table 8 Relative systematic uncertainties (in %) on the background efficiency from different sources, for jets in the PYTHIA multijet sample after tagging with the R2-trimmed mass and medium $D_2^{(β=1)}$ requirement that results in a signal efficiency $e_W^{G\&T} ≈ 50\%$. Uncertainties on scales (JMS, JES and JSS indicate the mass, energy and substructure scale uncertainties) can be in both directions, and so result in pairs of efficiency uncertainties. The mass and energy resolution uncertainties are denoted JMR and JER respectively, and ISR/FSR indicate the uncertainties from the modeling of the initial/final state radiation.

<table>
<thead>
<tr>
<th>Source</th>
<th>$p_T$ range (GeV)</th>
<th>200–250</th>
<th>250–350</th>
<th>350–500</th>
<th>500–1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>JES</td>
<td>+3.0/−5.6</td>
<td>+7.9/−8.3</td>
<td>+8.8/−5.0</td>
<td>+2.5/−4.3</td>
<td></td>
</tr>
<tr>
<td>JMS</td>
<td>−9.3/+8.6</td>
<td>−9.7/−9.6</td>
<td>−6.0/−6.7</td>
<td>−8.0/+5.7</td>
<td></td>
</tr>
<tr>
<td>JER</td>
<td>+1.0</td>
<td>−1.6</td>
<td>+0.5</td>
<td>+0.8</td>
<td></td>
</tr>
<tr>
<td>JMR</td>
<td>−2.0</td>
<td>+1.8</td>
<td>+1.0</td>
<td>+0.1</td>
<td></td>
</tr>
<tr>
<td>JSS ($D_2^{(β=1)}$)</td>
<td>−13.2/+15.9</td>
<td>−15.7/+19.7</td>
<td>−17.6/+22.7</td>
<td>−19.4/+23.8</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>+16.7/−19.1</td>
<td>+20.3/−23.5</td>
<td>+20.6/−24.2</td>
<td>+21.2/−24.8</td>
<td></td>
</tr>
</tbody>
</table>

Table 9 Relative systematic uncertainties (in %) on the W-jet tagging efficiency from different sources after tagging with the R2-trimmed mass and medium $D_2^{(β=1)}$ requirement that results in a signal efficiency $e_W^{G\&T} ≈ 50\%$. The uncertainties on scales (JMS, JES and JSS indicate the mass, energy and substructure scale uncertainties) can be in both directions, and so result in pairs of efficiency uncertainties. The mass and energy resolution uncertainties are denoted JMR and JER respectively, and ISR/FSR indicate the uncertainties from the modeling of the initial/finale state radiation.

<table>
<thead>
<tr>
<th>Source</th>
<th>$p_T$ range (GeV)</th>
<th>200–250</th>
<th>250–350</th>
<th>350–500</th>
</tr>
</thead>
<tbody>
<tr>
<td>JMS</td>
<td>+1.1</td>
<td>+1.1</td>
<td>+9.6</td>
<td></td>
</tr>
<tr>
<td>JES</td>
<td>−3.5/+3.6</td>
<td>−1.7/+2.5</td>
<td>+1.6/−2.3</td>
<td></td>
</tr>
<tr>
<td>JER</td>
<td>−0.1</td>
<td>+1.0</td>
<td>+1.0</td>
<td></td>
</tr>
<tr>
<td>JMR</td>
<td>+2.7</td>
<td>+3.7</td>
<td>+4.3</td>
<td></td>
</tr>
<tr>
<td>JSS ($D_2^{(β=1)}$)</td>
<td>+4.3/−2.9</td>
<td>+4.2/−4.5</td>
<td>+5.1/−4.8</td>
<td></td>
</tr>
<tr>
<td>MC generator</td>
<td>−0.9</td>
<td>+1.9</td>
<td>−3.2</td>
<td></td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>+1.6/−2.2</td>
<td>+2.7/−4.0</td>
<td>+4.4/−5.6</td>
<td></td>
</tr>
<tr>
<td>Multijet normalisation</td>
<td>−0.4/+0.4</td>
<td>−0.3/+0.3</td>
<td>+0.1/−0.1</td>
<td></td>
</tr>
<tr>
<td>Single-top normalisation</td>
<td>−0.1/+0.1</td>
<td>−0.1/+0.1</td>
<td>−0.1/0.1</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ normalisation</td>
<td>0.6/−0.5</td>
<td>0.6/−0.6</td>
<td>+0.5/−0.5</td>
<td></td>
</tr>
<tr>
<td>W+jets normalisation</td>
<td>−0.3/+0.3</td>
<td>−0.4/+0.4</td>
<td>−0.5/0.4</td>
<td></td>
</tr>
<tr>
<td>MC statistics</td>
<td>−1.0</td>
<td>−1.5</td>
<td>−3.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>+6.6/−5.4</td>
<td>+7.3/−6.6</td>
<td>+13.1/−13.2</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 22  $W$ boson tagging efficiencies in ranges of jet $p_T$ for (left) signal $W$-jets in $t\bar{t}$ events and (right) multijet background. The $\epsilon_{W}^{G&T} \sim 50\%$ working points obtained with the combined mass window and $c_{2}^{(\beta=1)}$ requirements are shown in a and b, and the $\sim 25\%$ working points are shown in c, d. The deviations from 50 and 25% in a and e respectively are due to the optimisations being based on $W$-jets in a different $W' \rightarrow WZ$ topology, as discussed in the text. The lower panels show ratios of the efficiency measured in data to the efficiency in two different MC simulations.
Fig. 23  $W$ boson tagging efficiencies in ranges of jet $p_T$ for (left) signal $W$-jets in $t\bar{t}$ events and (right) multijet background. The $\epsilon_W^{\rm G&T} \sim 50\%$ working points obtained with the combined mass window and $D_2^{(\beta=1)}$ requirements are shown in a and b, and the $\sim 25\%$ working points are shown in c, d. The deviations from 50 and 25 in a and c respectively are due to the optimisations being based on $W$-jets in a different $W' \rightarrow WZ$ topology, as discussed in the text. The lower panels show ratios of the efficiency measured in data to the efficiency in two different MC simulations.
Fig. 24  W boson tagging efficiencies in ranges of jet \( p_T \) for (left) signal \( W \)-jets in \( t\bar{t} \) events and (right) multijet background. The \( \epsilon_{\text{G&T}} \frac{W}{W} \sim 50 \% \) working points obtained with the combined mass window and \( \tau_{21}^{\text{Tight}} \) requirements are shown in a and b, and the \( \sim 25 \% \) working points are shown in c, d. The deviations from 50 and 25 \% in a and c respectively are due to the optimisations being based on \( W \)-jets in a different \( W' \to WZ \) topology, as discussed in the text. The lower panels show ratios of the efficiency measured in data to the efficiency in two different MC simulations.
Fig. 25 Signal efficiency versus background rejection power (1/background efficiency) curves derived using POWHEG-BOX + PYTHIA signal efficiencies and PYTHIA background efficiencies compared with points from data. Three $p_T$ ranges are shown: a 200–250 GeV, b 250–350 GeV, and c 350–500 GeV. The data points include systematic uncertainties on the signal efficiency measurement in $t\bar{t}$ events and the uncertainties on the PYTHIA background efficiency predictions.

$\rho_T$ dependence of the $C_2^{(\beta=1)}$ tagger compared to the $D_2^{(\beta=1)}$ and $\tau_{wta}^{21}$ taggers.

For the signal efficiencies, the uncertainty bands of the ratios account for the correlations in the systematic uncertainties between data and MC. In general, data and POWHEG-BOX + PYTHIA agree better than data and MC@NLO + HERWIG. For the medium working point, there is agreement between the two MC models within 1σ except in the range $200 < \rho_T < 250$ GeV, while for the tight working point ($\epsilon_{G&T} \sim 25\%$) the efficiency of MC@NLO + HERWIG is 1.5σ to 2σ higher than both the efficiency predicted by POWHEG-BOX + PYTHIA and the measurements in data. There is a potential bias towards POWHEG-BOX + PYTHIA, as this generator provides the signal template used in determining the background subtraction that is necessary to define the signal efficiency in data. However, even when using MC@NLO + HERWIG for the templates in the subtraction, POWHEG + PYTHIA gives a better description of the signal efficiency measured in data. The differences in the MC signal efficiencies stem from the differences in the signal mass distributions between models; the mass peak has a different width, so the fraction of signal in the mass window (which is the same for both Monte Carlo samples) is already significantly different after the requirement on the groomed jet mass is applied (see for example Fig. 17).

Figure 25 shows the $t\bar{t}$ MC efficiency versus rejection curves with data measurements at the medium and tight working points, including systematic uncertainties on the sig-
nal and background efficiencies. Generally good agreement between data and MC simulation is observed in all $p_T$ ranges for these measurements.

8 Conclusions

Several combinations of jet grooming algorithms and tagging variables have been studied to find an optimal $W$-jet tagger in terms of (a) maximising multijet background rejection power for given values of $W$-jet signal efficiency; (b) minimising systematic uncertainties and the effects of pileup; and (c) the modelling of the jet mass and substructure variables in Monte Carlo simulations.

The signal efficiency working point $\epsilon^G_W = 68\%$ is chosen as a suitable baseline for the comparison of grooming algorithms. The performances of the best few configurations of trimming, pruning and split-filtering are similar at this working point, and the anti-$k_t$, $R = 1.0$ jet trimmed with $f_{\text{cut}} = 5\%$ and $R_{\text{sub}} = 0.2$ (‘R2-trimming’) does particularly well in terms of removing pileup-dependence. Cambridge-Aachen pruning also provides significant discrimination for $W$-jet tagging, as does split-filtering without the mass-drop requirement. The irrelevance of the mass-drop requirement was shown previously in phenomenological studies [82], and is verified here in MC samples with a full ATLAS detector simulation. Trimming with $R_{\text{sub}} = 0.1$ shows promise in terms of the jet mass; it is not pursued further in these studies because it is challenging in terms of systematic uncertainties, as one is entering the arena of single-cluster jet, but it may well be considered in future extensions of these studies (for example in tagging $W$ bosons with $p_T > 1\ TeV$).

The energy correlation ratios $D_2^{(\beta=1)}$, $C_2^{(\beta=1)}$ are found to be particularly good variables for tagging $W$-jets, as shown for the first time here in data. However, there is some evidence of the $C_2^{(\beta=1)}$ variable having a higher background efficiency for low-$p_T$ jets. Similarly good is the $N$-subjettiness ratio $\tau_{21}^{\text{wta}}$, which performs better than its predecessor $\tau_{21}$.

The signal and background efficiencies obtained using pairwise combinations of the R2-trimmed mass and three different substructure variables are measured in $t\bar{t}$ and multijet events from 20.3 fb$^{-1}$ of 8 TeV $pp$ collisions recorded by ATLAS at the LHC. These are compared to various MC predictions which show in general good agreement within the uncertainties with the data measurements of signal efficiencies around 50% for background efficiencies around 2%.

In some configurations, significant differences are observed in both the signal and background efficiencies from different Monte Carlo predictions. This can provide important information to improve the Monte Carlo simulations for searches for physics beyond the Standard Model. It further highlights the potential for data measurements such as these to be utilised for tuning Monte Carlo simulations.

These studies are necessarily limited in scope to comparing simple two-variable taggers, made up of a groomed mass window and a substructure variable requirement, both of which are sensitive to $p_T$ and therefore optimised for three different $p_T$ ranges. Extensions to these studies could include combining three or more variables and using multivariate techniques to further boost the signal efficiency and/or reduce the background; investigating how these conclusions change if dedicated pileup-removal techniques are used alongside grooming; and varying the $\epsilon^G_W$ baseline at which the grooming algorithms are compared.

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; COCIC, Colombia; MESMT CR, MPO CR and VSC CR, Czech Republic; DNRF, DNRC, and Lundbeck Foundation, Denmark; IN2P3-CNRS, CEA, DSM/IRFU, France; GNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RCG, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNISW 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, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, FP7, Horizon 2020 and Marie Sklodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Region Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and 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 (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP3.

1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany, NY, USA
3 Department of Physics, University of Alberta, Edmonton, AB, Canada
4 (a)Department of Physics, Ankara University, Ankara, Turkey; (b)Istanbul Aydin University, Istanbul, Turkey;
   (c)Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5 LAPP, CNRS/IN2P3 and Université Savoie Mont Blanc, Annecy-le-Vieux, France
6 High Energy Physics Division, Argonne National Laboratory, Argonne, IL, USA
7 Department of Physics, University of Arizona, Tucson, AZ, USA
8 Department of Physics, The University of Texas at Arlington, Arlington, TX, USA
9 Physics Department, University of Athens, Athens, Greece
10 Physics Department, National Technical University of Athens, Zografou, Greece
11 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
12 Institut de Física d’Altes Energies and Departament de Física de la Universitat Autònoma de Barcelona, Barcelona, Spain
13 Institute of Physics, University of Belgrade, Belgrade, Serbia
14 Department for Physics and Technology, University of Bergen, Bergen, Norway
15 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, USA
16 Department of Physics, Humboldt University, Berlin, Germany
17 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
18 School of Physics and Astronomy, University of Birmingham, Birmingham, UK
19 (a)Department of Physics, Bogazici University, Istanbul, Turkey; (b)Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey;
   (c)Department of Physics, Dogo University, Istanbul, Turkey
20 (a)INFN Sezione di Bologna, Bologna, Italy; (b)Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy
21 Physikalisches Institut, University of Bonn, Bonn, Germany
22 Department of Physics, Boston University, Boston, MA, USA
23 Department of Physics, Brandeis University, Waltham, MA, USA
24 (a)Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; (b)Electrical Circuits Department,
   Federal University of Juiz de Fora (UFJJF), Juiz de Fora, Brazil; (c)Federal University of Sao Joao del Rei (UFSJ), Sao
   Joao del Rei, Brazil; (d)Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil
25 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
University of Iowa, Iowa City, IA, USA
Department of Physics and Astronomy, Iowa State University, Ames, IA, USA
Joint Institute for Nuclear Research, JINR Dubna, Dubna, Russia
KEK, High Energy Accelerator Research Organization, Tsukuba, Japan
Graduate School of Science, Kobe University, Kobe, Japan
Faculty of Science, Kyoto University, Kyoto, Japan
Kyoto University of Education, Kyoto, Japan
Department of Physics, Kyushu University, Fukuoka, Japan
Instituto de Fisica La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
Physics Department, Lancaster University, Lancaster, UK
(a)INFN Sezione di Lecce, Lecce, Italy; (b)Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy
Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK
Department of Physics, Jožef Stefan Institute and University of Ljubljana, Ljubljana, Slovenia
School of Physics and Astronomy, Queen Mary University of London, London, UK
Department of Physics, Royal Holloway University of London, Surrey, UK
Department of Physics and Astronomy, University College London, London, UK
Louisiana Tech University, Ruston, LA, USA
Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France
Fysiska institutionen, Lunds universitet, Lund, Sweden
Departamento de Física Teórica C-15, Universidad Autonoma de Madrid, Madrid, Spain
Institut für Physik, Universität Mainz, Mainz, Germany
School of Physics and Astronomy, University of Manchester, Manchester, UK
CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
Department of Physics, University of Massachusetts, Amherst, MA, USA
Department of Physics, McGill University, Montreal, QC, Canada
School of Physics, University of Melbourne, Melbourne, VIC, Australia
Department of Physics, The University of Michigan, Ann Arbor, MI, USA
Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
(a)INFN Sezione di Milano, Milan, Italy; (b)Dipartimento di Fisica, Università di Milano, Milan, Italy
B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus
National Scientific and Educational Centre for Particle and High Energy Physics, Minsk, Republic of Belarus
Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA
Group of Particle Physics, University of Montreal, Montreal, QC, Canada
P.N. Lebedev Institute of Physics, Academy of Sciences, Moscow, Russia
Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
National Research Nuclear University MEPhI, Moscow, Russia
D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
Nagasaki Institute of Applied Science, Nagasaki, Japan
Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
(a)INFN Sezione di Napoli, Naples, Italy; (b)Dipartimento di Fisica, Università di Napoli, Naples, Italy
Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA
Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands
Department of Physics, Northern Illinois University, De Kalb, IL, USA
Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
Department of Physics, New York University, New York, NY, USA
Ohio State University, Columbus, OH, USA
Faculty of Science, Okayama University, Okayama, Japan
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, OK, USA
Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel

Department of Physics, Aristotle University of Thessaloniki, Thessaloníki, Greece

International Center for Elementary Particle Physics and Department of Physics, The University of Tokyo, Tokyo, Japan

Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan

Department of Physics, Tokyo Institute of Technology, Tokyo, Japan

Department of Physics, University of Toronto, Toronto, ON, Canada

(a) TRIUMF, Vancouver, BC, Canada; (b) Department of Physics and Astronomy, York University, Toronto, ON, Canada

Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan

Department of Physics and Astronomy, Tufts University, Medford, MA, USA

Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia

Department of Physics and Astronomy, University of California Irvine, Irvine, CA, USA

(a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy; (b) ICTP, Trieste, Italy; (c) Dipartimento di Chimica Fisica e Ambiente, Università di Udine, Udine, Italy

Department of Physics, University of Illinois, Urbana, IL, USA

Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden

Instituto de Física Corpuscular (IFIC) and Departamento de Física Atómica, Molecular y Nuclear and Departamento de Ingeniería Electrónica and Instituto de Microelectrónica de Barcelona (IMB-CNMB), University of Valencia and CSIC, Valencia, Spain

Department of Physics, University of British Columbia, Vancouver, BC, Canada

Department of Physics and Astronomy, University of Victoria, Victoria, BC, Canada

Department of Physics, University of Warwick, Coventry, UK

Waseda University, Tokyo, Japan

Department of Particle Physics, The Weizmann Institute of Science, Rehovot, Israel

Department of Physics, University of Wisconsin, Madison, WI, USA

Fakultät für Physik und Astronomie, Julius-Maximilians-Universität, Würzburg, Germany

Fachbereich C Physik, Bergische Universität Wuppertal, Wuppertal, Germany

Department of Physics, Yale University, New Haven, CT, USA

Yerevan Physics Institute, Yerevan, Armenia

Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France

a Also at Department of Physics, King’s College London, London, UK
b Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
c Also at Novosibirsk State University, Novosibirsk, Russia
d Also at TRIUMF, Vancouver, BC, Canada
e Also at Department of Physics, California State University, Fresno, CA, USA
f Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
g Also at Departamento de Física e Astronomia, Faculdade de Ciencias, Universidade do Porto, Porto, Portugal
h Also at Tomsk State University, Tomsk, Russia
i Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
j Also at Universita di Napoli Parthenope, Napoli, Italy
k Also at Institute of Particle Physics (IPP), Canada
l Also at Particle Physics Department, Rutherford Appleton Laboratory, Didcot, UK
m Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
n Also at Louisiana Tech University, Ruston, LA, USA
o Also at Institució Catalana de Recerca i Estudis Avançats, ICREA, Barcelona, Spain
p Also at Department of Physics, The University of Michigan, Ann Arbor MI, United States of America
q Also at Graduate School of Science, Osaka University, Osaka, Japan
r Also at Department of Physics, National Tsing Hua University, Taiwan
s Also at Department of Physics, The University of Texas at Austin, Austin, TX, USA
t Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia
u Also at CERN, Geneva, Switzerland
v Also at Georgian Technical University (GTU), Tbilisi, Georgia
w Also at Manhattan College, New York, NY, USA
x Also at Hellenic Open University, Patras, Greece
y Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
z Also at LAL, Université Paris-Sud and CNRS/IN2P3, Orsay, France
aa Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
ab Also at School of Physics, Shandong University, Shandong, China
ac Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
ad Also at Section de Physique, Université de Genève, Geneva, Switzerland
ae Also at International School for Advanced Studies (SISSA), Trieste, Italy
af Also at Department of Physics and Astronomy, University of South Carolina, Columbia, SC, USA
ag Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China
ah Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
ai Also at National Research Nuclear University MEPhI, Moscow, Russia
aj Also at Department of Physics, Stanford University, Stanford CA, USA
ak Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
al Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
a Deceased