Jet Substructure at the Large Hadron Collider: Experimental Review

Lily Asquith, Mario Campanelli, Chris Delitzsch, Andreas Hinzmann, Deepak Kar, Roman Kogler, Christine McLean, Benjamin Nachman, Justin Pilot, Alexander Schmidt, Nhan Tran, Caterina Vernieri, Marcel Vos, and Emma Winkels

University of Sussex, UK
University College London, UK
University of Arizona, USA
Universität Hamburg, Germany
University of Witwatersrand, South Africa
University of California, Davis, USA
Lawrence Berkeley National Laboratory, USA
RWTH Aachen University, Germany
Fermilab, USA
IFIC Valencia, Spain

March 22, 2018

Abstract

Jet substructure has emerged to play a central role at the Large Hadron Collider (LHC), where it has provided numerous innovative new ways to search for new physics and to probe the Standard Model, particularly in extreme regions of phase space. In this article we focus on a review of the development and use of state-of-the-art jet substructure techniques by the ATLAS and CMS experiments. ALICE and LHCb have been probing fragmentation functions since the start of the LHC [1, 2] and have also recently started studying other jet substructure techniques [3]. It is likely that in the near future all LHC collaborations will make significant use of jet substructure and grooming techniques. Much of the work in this field in recent years has been galvanized by the Boost Workshop Series [4, 5, 6], which continues to inspire fruitful collaborations between experimentalists and theorists. We hope that this review will prove a useful introduction and reference to experimental aspects of jet substructure at the LHC. A companion overview of recent progress in theory and machine learning approaches is given in [7]; the complete review will be submitted to Reviews of Modern Physics.

Contents

1 Jet Reconstruction 2
1.1 ATLAS and CMS detectors . . . . . . . . . . 2
1.2 Inputs to Jet Reconstruction . . . . . . . . 3
1.3 Calibration . . . . . . . . . . . . . . . . . . 4
1.4 Jet Grooming . . . . . . . . . . . . . . . . . . 7
1.5 Jet Mass . . . . . . . . . . . . . . . . . . . . 7

2 Pileup Mitigation Techniques 10
2.1 Defining Pileup . . . . . . . . . . . . . . . . 10
2.2 Mitigation Methods . . . . . . . . . . . . . 10
2.3 Performance Studies . . . . . . . . . . . . . 12

3 Boosted Particle Tagging 13
3.1 Quark/Gluon Discrimination . . . . . . . . 13
3.2 Vector Boson Tagging . . . . . . . . . . . . 16
3.3 Top Tagging . . . . . . . . . . . . . . . . . 19
3.4 $H \to b \bar{b}$ Tagging . . . . . . . . . . . . 24

4 Measurements 28
4.1 Measurements with jet substructure . . . . 28


This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.
techniques for using physics jets as proxies for their origin; the following experimental overview describes the final structure that encodes a wealth of information about asproxies for quarks and gluons, but they have a rich inter-

tative corrections. Jets have traditionally been used to identify potentially interesting events for analysis the more coarse objects reconstructed very rapidly in or-

ter (the focus was Higgs mass reconstruction), and ac-

etivities for hard-scattering events at the LHC. Most hadrons leave a signal in the calorimeters of CMS and ATLAS (however only those with sufficient energy, roughly 100-

all samples (PbWO$_4$) crystals, as opposed to the Liquid Argon (LAr) used by ATLAS. The differences in the ATLAS and CMS calorimeter designs are a result of the different ranking of priorities decided by the two collaborations; ATLAS chose a radiation-hard technology with sufficient resolution in a fine sampling LAr calorimeter, while CMS prioritized the excellent resolution of a total absorption crystal calorimeter (the focus was Higgs mass reconstruction), and accepted the accompanying limitations in radiation-hardness associated with this technology. The CMS ECAL crystal response varies under irradiation, which is partially recovered in a few hours at room temperature. The ATLAS ECAL is segmented into three (two) longitudinal layers for $|\eta| < 2.5$ ($|\eta| > 2.5$). The granularity of the ATLAS ECAL in Table 1 refers to its second layer (as most of the electromagnetic energy is deposited there); the first

The ATLAS [10] and CMS [11] detectors are designed to observe leptons, photons, and hadrons resulting from LHC pp collisions. The physics of the hard reaction producing unstable particles and color-confined quarks takes place at the point of collision (the primary vertex) within the beam pipe. Beyond the beam pipe, at 4.4 cm (3.3 cm) in CMS (ATLAS), the first cylindrical layer of detectors encountered are silicon pixels and strips for identification of charged particles. CMS provides a 3.8 T magnetic field via a solenoid positioned outside the silicon tracking detector, the Electromagnetic Calorimeter (ECAL) and most of the Hadronic Calorimeter (HCAL). ATLAS has an additional layer of straw drift tubes (Transition Radiation Tracking: TRT) in the tracking detector, with a 2 T magnetic field encompassing the silicon and TRT, and the ECAL and HCAL situated outside the solenoidal magnet. The calorimeters are surrounded by muon spectrometers which build the outermost part of the ATLAS and CMS detectors.

The energy and momentum ranges and resolutions for the barrel regions of ATLAS and CMS are shown in Table 1 along with the measurement granularity, which limits the angular resolution. The better energy resolution of the CMS ECAL is due to the use of lead tungstate (PbWO$_4$) crystals, as opposed to the Liquid Argon (LAr) used by ATLAS. The differences in the ATLAS and CMS calorimeter designs are a result of the different ranking of priorities decided by the two collaborations; ATLAS chose a radiation-hard technology with sufficient resolution in a fine sampling LAr calorimeter, while CMS prioritized the excellent resolution of a total absorption crystal calorimeter (the focus was Higgs mass reconstruction), and accepted the accompanying limitations in radiation-hardness associated with this technology. The CMS ECAL crystal response varies under irradiation, which is partially recovered in a few hours at room temperature. The ATLAS ECAL is segmented into three (two) longitudinal layers for $|\eta| < 2.5$ ($|\eta| > 2.5$). The granularity of the ATLAS ECAL in Table 1 refers to its second layer (as most of the electromagnetic energy is deposited there); the first

The ATLAS ECAL barrel covers the pseudorapidity range $|\eta| < 1.475$, the end-caps cover $1.375 < |\eta| < 3.2$ and the forward ECAL layer extends the coverage up to $|\eta| < 4.9$. The CMS ECAL barrel covers $|\eta| < 1.48$, the end-caps extend the coverage up to $|\eta| < 3$. 

The majority of particles detected in particle physics experiments are hadrons - bound states of quarks and gluons. Quarks and gluons themselves cannot be observed in isolation due to color confinement, thus observable hadrons are used as a proxy for the quarks and gluons describing perturbative processes. Hadrons are produced in vast quantities by hard-scattering events at the LHC. Most hadrons leave a signal in the calorimeters of CMS and ATLAS (however only those with sufficient energy, roughly 100-500 MeV depending on the noise level, can be detected); charged hadrons additionally leave signals in the tracking detectors.

When hadrons are produced in high energy collisions, the resulting tracks and energy deposits tend to form roughly conical shapes extending from the point of interaction into the detectors. The detectors are arranged cylindrically and symmetrically around the LHC beam. The term ‘jet’ describes a wide range of objects, from the more coarse objects reconstructed very rapidly in order to identify potentially interesting events for analysis (‘trigger-level jets’) to the ‘physics jets’ used at analysis level, which may have gone through a dozen steps of post-processing corrections. Jets have traditionally been used as proxies for quarks and gluons, but they have a rich internal structure that encodes a wealth of information about their origin; the following experimental overview describes techniques for using physics jets as proxies for $b$ quarks, $W \to q\bar{q}$ decays, top quarks and others. For a theoretical introduction to jets, we recommend Ref. [8, 9].
layer has a finer granularity in \( \eta \). The multiple layers allow for a finer granularity than the cell size in any of the individual layers, being advantageous over a laterally segmented calorimeter, and additionally provide pointing information. The difference between ATLAS and CMS for the HCAL resolution is particularly large at higher energies: a 1 TeV jet has \( \frac{\sigma(E)}{E} \sim 2\% \) in ATLAS, in contrast to \( \frac{\sigma(E)}{E} \sim 5\% \) in CMS. To counter this, CMS are using a particle flow approach which includes tracks as discussed in Section 1.2 below.

<table>
<thead>
<tr>
<th></th>
<th>ATLAS</th>
<th>CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_T ) range (GeV)</td>
<td>( p_T &gt; 0.1 )</td>
<td>( p_T &gt; 0.2 )</td>
</tr>
<tr>
<td>( 1/p_T ) resolution (%)</td>
<td>( 5 \times 10^{-4}p_T \oplus 0.015 )</td>
<td>&gt; 0.8%</td>
</tr>
<tr>
<td>( d_0 ) resolution (( \mu )m)</td>
<td>140 – 20</td>
<td>&gt; 9</td>
</tr>
</tbody>
</table>

**ECAL**

\( E \) resolution (%)

<table>
<thead>
<tr>
<th></th>
<th>ATLAS</th>
<th>CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( 0.1/\sqrt{E} \oplus 0.007 )</td>
<td>( 0.03/\sqrt{E} \oplus 0.003 )</td>
</tr>
<tr>
<td>granularity</td>
<td>( 0.025 \times 0.025 )</td>
<td>( 0.017 \times 0.017 )</td>
</tr>
</tbody>
</table>

**HCAL**

\( E \) resolution (%)

<table>
<thead>
<tr>
<th></th>
<th>ATLAS</th>
<th>CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( 0.5/\sqrt{E} \oplus 0.03 )</td>
<td>( 1.0/\sqrt{E} \oplus 0.05 )</td>
</tr>
<tr>
<td>granularity</td>
<td>( 0.1 \times 0.1 )</td>
<td>( 0.087 \times 0.087 )</td>
</tr>
</tbody>
</table>

Table 1: A comparison of ATLAS and CMS detectors in the barrel regions. The transverse momentum is denoted \( p_T \). The granularity is in pseudorapidity and azimuth \( (\eta \times \phi) \) for the second layer (inner three layers) of the ATLAS ECAL (HCAL), and \( d_0 \) is the transverse impact parameter resolution with respect to the beam-line. Information taken from [11, 12, 13, 14, 15].

### 1.2 Inputs to Jet Reconstruction

For general jet reconstruction (in events that do not require jet substructure techniques) both ATLAS and CMS use the anti-\( k_t \) [16] jet clustering algorithm to form jets from their respective inputs. This iterative recombination algorithm takes a single argument: the distance parameter \( R \), which is generally referred to as the jet radius in experimental publications. During the LHC’s Run 1 (2010-2012), ATLAS favored a distance parameter \( R = 0.4 \), while CMS favored \( R = 0.5 \). For Run 2, the two experiments agreed on a common choice of \( R = 0.4 \) for the sake of better comparability and the additional advantage of the applicability of a theoretical calculation to both experiments.

In ATLAS, three dimensional topological clusters [17] are constructed from the cells of the calorimeter based on their significance, defined as the ratio of the cell signal to the average noise. The topological clustering algorithm implicitly suppresses fluctuations due to electronic and pileup noise (pileup-induced noise dominates since 2011 data-taking) by removing isolated cells with a small cell signal resulting in an improved energy and spatial resolution. The topological clusters are calibrated using the Local Hadronic Cell Weighting (LCW) procedure to account for the non-compensating calorimeter response to hadrons, signal losses due to energy deposited in inactive detector material and signal losses on cluster boundaries caused by the topological clustering algorithm. This calibration scheme relies on classification of clusters as hadronic or electromagnetic in origin based on the energy and position of the cluster, the longitudinal depth \( \lambda_{clus} \) and normalized signal energy density, as illustrated in Figure 1; hadronic showers tend to occur deeper in the calorimeter and be less dense.

In recent years, ATLAS has investigated the use of tracks in addition to topoclusters for use in jet reconstruction, the motivation being that the tracker has better transverse momentum \( (p_T) \) resolution for low-energy charged particles, and better angular resolution [18]. The Particle Flow technique has been shown to improve the
jet energy resolution and reduce the pileup dependence for jets with $p_T < 100 \text{ GeV}$, by replacing the calorimeter energy deposits from charged hadrons by their measurement in the inner detector [18]. Jets reconstructed with particle flow have improved angular resolution with respect to those reconstructed only from topoclusters, as shown in Figure 2. Tracks are also used by ATLAS for improving the reconstruction of groomed jet masses, as discussed in Section 1.5, and reducing the topology dependence via the Global Sequential Calibration (GSC) (Section 1.3).

Topological clusters from charged-particles inside dense environments (like the core of high $p_T$ jets) can merge, resulting in a loss in tracking efficiency and degraded cluster position resolution. ATLAS has implemented a stacked neural network (NN) approach to examine pixel clusters to identify multi-particle clusters, estimate the position of the particles passing through the clusters, and also predict the residual resolution of the position estimates [19, 20]. This approach significantly improves the efficiency inside jet cores and for boosted hadronically decaying tau leptons. Furthermore, the NN performance has been extensively studied in data and simulation using a variety of methods [20, 21, 22, 23].

In CMS, a particle flow [24, 25, 26] event algorithm is used to reconstruct and identify each individual particle with an optimized combination of information from various elements of the detector. The energy of photons is directly obtained from the ECAL measurement, corrected for threshold effects in the formation of calorimeter clusters. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for threshold effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy. The momentum for each particle type is calibrated separately based on detector simulation and data measurements. The momentum of charged hadrons is mainly measured with the tracking detector for particle $p_T < 500 \text{ GeV}$, while at higher $p_T$ the energy measurement of these hadrons comes mainly from the ECAL and HCAL detectors, since the resolution of the tracker for high momentum charged hadrons becomes worse than the ECAL and HCAL resolution. Since the angular resolution of the HCAL is not sufficient to separate close-by particles in jets, the granularity of the ECAL is used by splitting clusters according to their lateral energy profile in the ECAL [27]. The particle flow algorithm dramatically improves the energy response and resolution (Fig. 3) as well as the angular resolution of jets before dedicated corrections, particularly for jets with $p_T < 100 \text{ GeV}$ as shown in Fig. 4.

For Run 2, CMS has introduced a dedicated tracking step for the dense cores of high-$p_T$ jets. In the innermost layers of the pixel detector charge deposits from separate tracks can be merged into one cluster. A cluster splitting procedure attempts to split such merged clusters exploiting the information of the jet direction, predicting the expected cluster shape and charge. The jet core tracking step also allows a larger number of hit combinations to be evaluated, compared to the standard tracking algorithm. This results in an improvement of the tracking efficiency of up to 10% in the innermost cores of dense jets [28].

1.3 Calibration

The ratio of the measured energy $E_{\text{reco}}$ to the deposited energy $E_{\text{true}}$ is known as the jet energy response, and is
Figure 3: Jet energy resolution for particle flow (red) and calorimeter-only (blue) jets in the barrel region in CMS simulation, with no pileup, as a function of the $p_T$ of the reference jet. Taken from [26].

Figure 4: Jet angular resolution for particle flow (red) and calorimeter-only (blue) jets in the barrel region in CMS simulation, with no pileup, as a function of the $p_T$ of the reference jet. The $\phi$ resolution is expressed in radians. Taken from [26].

Figure 5: The ATLAS jet energy response prior to calibration, as a function of $\eta$ for a range of different particle-level jet energies. Adapted from [30].

A crucial aspect of any measurement that uses jets reconstructed from calorimeter energy deposits. The average response is strongly dependent on both the energy and $\eta$ of the jet, as shown in the pre-calibration distributions of Figure 5. For this reason, the Jet Energy Scale (JES) is calculated in bins of the particle-level jet energy $E_{\text{true}}$ and $\eta_{\text{det}}$ as the mean of a Gaussian fit to the response distribution and a numerical inversion procedure is used to derive calibration factors in bins of the reconstructed jet energy from $E_{\text{true}}$ [29, 30, 31, 32].

In ATLAS, the calibration of the JES is undertaken in stages, as illustrated in Figure 6 starting from jets either at the electromagnetic (EM) or LCW (built from calibrated inputs) scale. Using calibrated inputs improves the uncalibrated JES by 50% for jets with $E = 30$ GeV and $|\eta| < 0.3$ [29]. The Global Sequential Calibration [30, 33] was introduced for Run 2 and reduces the sensitivity to differences in the responses of quark versus gluon-initiated jets. This additional calibration results in a significant jet $p_T$ resolution improvement of up to 35% depending on the $p_T$ and $\eta$ of the jet [33]. The JES uncertainty varies between 1-6% in the central region with $\eta = 0$ as shown in Figure 7 [30]. The flavor fraction composition for an unknown quark/gluon composition is of the same order as the absolute $\text{in-situ}$ uncertainty. The flavor fraction composition is however analysis-dependent and the uncertainty can be reduced if the quark/gluon composition is known. Single-particle response studies and ATLAS combined test beam measurements [34] are used to derive uncertainties for jets with $p_T > 2$ TeV due to the limited statistics of
collision data for the multijet balance technique.

In CMS, jets are clustered from calibrated particles, thus the uncalibrated JES is within 6% of the expected value of 1 for central ($|\eta| < 0.7$) jets with $p_T > 30$ GeV [35], as shown in Fig. 8. To account for deviations from unity, factorized JES calibrations are applied in multiple stages [36] including pileup corrections, simulation-based response corrections and small residual corrections for tracking inefficiencies and threshold effects, derived in situ from $\gamma$+jet, $Z$+jet and dijet samples [31]. This additional correction is not used when jet substructure observables are constructed, but dedicated corrections are derived as described in Section 1.5. Figure 9 shows the calibrated JES uncertainty obtained in CMS, which is below 1% for jets with $p_T > 100$ GeV in the central region with $\eta = 0$. Even for jet $p_T$ as low as 10 GeV the uncertainty is below 3%, owing to the excellent performance of the particle flow reconstruction.

A detailed discussion of the different approaches for deriving jet energy scale uncertainties in ATLAS and CMS

Figure 6: ATLAS jet energy scale calibration. Adapted from [30].

Figure 7: ATLAS jet energy scale uncertainty. Adapted from [30].

Figure 8: Simulated jet response versus $\eta$ for jets with $R = 0.5$ in CMS. Taken from [35].

Figure 9: CMS jet energy scale uncertainty, from [35]. JEC means Jet Energy Correction, which has the same meaning as JES.
can be found in Ref. [36].

1.4 Jet Grooming

The use of jets as a proxy for quarks and gluons has become an experimental tool of unreckonable importance, particularly at the LHC experiments where jets are present in nearly every event and thus in nearly every analysis. The machinery of jet reconstruction has been evolving at every level over the last decade, from trigger-level to analysis-level, and jet grooming techniques are an area of work that has seen a particularly high level of interest from the experimental and theoretical communities alike.

A rule of thumb for the decay of a massive object such as a top quark or $W/Z/H$ boson is that the decay products are likely to lie within a cone of radius $R = 2M/p_T$, where $M$ and $p_T$ are the mass and transverse momentum of the object. Using this rule-of-thumb for the example of a top quark decay, a top quark with $p_T = 350$ GeV will have its decay products captured by a jet of radius $R = 1.0$, and the higher the $p_T$ of the object, the more collimated the decay products. Jet grooming is an additional ‘post-processing’ treatment of large radius jets, an extra step used to remove unwanted soft radiation and to allow the underlying hard substructure associated with a three-prong (e.g. top quark) or two-prong (e.g. $W$ boson) decay to be utilized in the analysis.

In Run 1, the role of grooming has traditionally satisfied two purposes in ATLAS, being the mitigation of pileup effects on jets and the removal of soft/wide-angle radiation whereas the particle flow algorithm employed in CMS allowed for an efficient correction for pileup effects, reducing the purpose of grooming to the removal of soft/wide-angle radiation. For higher pileup scenarios as currently achieved in Run 2, also the particle flow algorithm is not able anymore to remove all pileup contributions from large-$R$ jets.

ATLAS performed a broad study of the relative performance of different grooming techniques for boson-tagging [37, 38, 39], top-tagging [40, 41] and SM measurements [42, 43], using the removal of pileup-dependence, the jet mass resolution, and the tagging efficiency versus background rejection as performance metrics. The ‘standard’ grooming procedure adopted by ATLAS is Trimming [44] with $f_{cut} = 0.05$ for boson tagging in both Run 1 ($R_{sub} = 0.3$) and Run 2 ($R_{sub} = 0.2$). The trimming algorithm with the same parameters was adopted for top tagging, along with several other techniques (see Sec. 3.3). Another technique currently in use by ATLAS is the reclustering of small-$R$ jets [45], which uses fully-calibrated anti-$k_t$, $R = 0.4$ jets as inputs to the anti-$k_t$ algorithm with a larger distance parameter (typically $R = 1.0$). This has proven a popular method in ATLAS analyses due to the flexibility of optimizing the jet distance parameter depending on the considered phase-space of the analysis [46, 47, 48].

A recent study of in-situ measurements [49] (including ‘closeby’ effects) confirm that the data / MC differences observed with reclustered jets are indeed covered by simply propagating the uncertainties associated with the input anti-$k_t$, $R = 0.4$ jets.

CMS studied a large number of grooming techniques in the context of boosted boson-tagging [27, 50], top-tagging [51, 52] and SM measurements [53, 54]. During Run 1 the grooming techniques were used together with charged-hadron subtraction for pileup mitigation (see Sec. 2). All groomers studied showed reasonable or good agreement between data and simulation and the Pruning [55] algorithm ($R = 0.8, z = 0.1$ and $d_{cut} = 0.5$) showed the best performance for boson tagging [50]. For Run 2 the modified mass-drop tagger (mMDT) [56] is used for jets with $R = 0.8$ in jet substructure analyses in CMS together with the pileup removal algorithm PUPPI [57] (see Sec. 2). In the remainder of this review, jets groomed with the mMDT are referred to as soft drop [58] jets, a generalization of the mMDT algorithm, with $z_{cut} = 0.1$ and $\beta = 0$. Soft drop jets in combination with PUPPI show a similar performance as pruning when comparing signal efficiency versus background rejection [52, 59], but allow for better theoretical control. While grooming techniques were found to improve the performance (higher background rejection at fixed signal efficiency) of the jet mass, $N$-subjettiness ratios [60, 61] were found to perform better without grooming for boosted boson tagging [50]. For top-tagging applications, however, soft drop groomed $N$-subjettiness ratios improved the performance with respect to ungroomed ones for jets with $p_T < 400$ GeV. For higher $p_T$ jets there was not significant gain observed with grooming for $N$-subjettiness ratios [52].

1.5 Jet Mass

The reconstruction of jet energies mainly relies on the capability of a detector to measure the total energy of all particles deposited in the detector; however, the measurement of jet mass requires detection of the deposited energy with a granularity that is finer than the size of a jet. The
mass of a jet can only be estimated if the energy is deposited in at least two detector elements, as it depends on both the energy and opening angle between the particles (or detected energy deposits) in a jet. For jet substructure techniques that rely on the rejection of soft particles, it is also important to be able to reconstruct particles with low \( p_T \) separately from harder particles in a jet.

The jet mass response distribution \( R_{\text{reco}} \) is constructed from the calibrated, reconstructed jet mass \( M_{\text{reco}} \) divided by the particle-level jet mass \( M_{\text{true}} \). As for the energy response (see 1.3), the mass response distribution is calculated in bins of reconstructed jet \( p_T^{\text{reco}} \) and \( \eta^{\text{reco}} \). In ATLAS, the Jet Mass Scale (JMS) is defined as the mean of this response distribution. The Jet Mass Resolution (JMR) is then defined as half the 68% interquartile range (IQnR) of the response distribution, as

\[
r = 0.5 \times 68\% \text{IQnR}(R_{\text{reco}}).
\]

This is robust to large non-Gaussian tails but, if the distribution is Gaussian, is equal to half its 1\( \sigma \) width. The fractional JMR is expressed as the JMR divided by the median of the response distribution.

ATLAS has recently developed a data-driven approach to extract the JMS and JMR from an enriched samples of boosted \( t\bar{t} \) events, however the method can also be extended to other final states. This forward-folding approach folds the particle-level mass spectra by a modified response function such that the average reconstructed jet mass \( \langle M_{\text{reco}}^{m,p_T} \rangle \) in a given bin of particle-level jet mass and reconstructed jet \( p_T \) is scaled by the scale parameter \( s \) and the resolution scaled by the resolution parameter \( r \):

\[
M_{\text{fold}} = s \times M_{\text{reco}} + (M_{\text{reco}} - \langle M_{\text{reco}}^{m,p_T} \rangle)(r - s).
\]

The values of \( r \) and \( s \) for which the \( M_{\text{fold}} \) distribution best matches the data are extracted from a 2 dimensional \( \chi^2 \) fit as shown in Figure 10 and detailed in Ref. [62, 63].

With the forward-folding approach, the JMS and JMR for hadronically decaying boosted \( W \) bosons with \( p_T > 200 \) GeV are determined with 2–3% and 20% systematic uncertainties, respectively (see Figure 11). As the jet mass and its detector-response depend on kinematics and jet substructure, the measurement was repeated differentially with an increased luminosity for boosted \( W \) and top quarks in Ref. [64]. It will be important to extend the technique (possibly through simulation methods) to other final states in the future. A detailed study of the various contributions to the JMS and JMR has been performed in context of the soft drop mass measurement [42], described in Section 4.2, by propagating experimental uncertainties on the inputs to the jet reconstruction (topological clusters) to the jet mass. The dominating uncertainties are due to the theoretical modeling of jet fragmentation and the cluster energy scale.

As the forward-folding method is currently restricted to jets with \( p_T < 350 (500) \) GeV for boosted \( W \) bosons (top quarks), the results are combined with the so-called \( R_{\text{trk}} \) method which constrains the mass scale by comparing the calorimeter jet mass to the mass calculated from track jets and extends up to \( p_T = 3000 \) GeV [64]. The \( R_{\text{trk}} \) method can also be generalized to other variables and is used in ATLAS to constrain the \( p_T \) scale of large-\( R \) jets as well as to derive systematic uncertainties on jet substructure variables.

The concept of a Track-Assisted Mass for trimmed, large-\( R \) jets has recently been studied in ATLAS [62] to maintain performance for highly boosted particles due to the limited granularity of the calorimeter. The track-
Assisted mass is defined as:

\[ m_{\text{TA}} = \frac{p_{T}^{\text{calo}}}{p_{T}^{\text{track}}} \times m_{\text{track}}, \]  

where \( p_{T}^{\text{calo}} \) is the transverse momentum of the calorimeter jet, \( p_{T}^{\text{track}} \) is the transverse momentum of the four-vector sum of tracks associated to the calorimeter jet, and \( m_{\text{track}} \) is the invariant mass of this four-vector sum, where the track mass is set to the pion mass \( m_{\pi} \). The track-assisted mass exploits the excellent angular resolution of the tracking detector (whereas the momentum resolution of high-\( p_{T} \) tracks is significantly worse than that of the calorimeter) and the ratio \( p_{T}^{\text{calo}} \) to \( p_{T}^{\text{track}} \) corrects for charged-to-neutral fluctuations.

The Combined Mass is defined as:

\[ m_{\text{comb}} = \left( \frac{\sigma_{\text{calo}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}} \right) m_{\text{calo}} + \left( \frac{\sigma_{\text{TA}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}} \right) m_{\text{TA}}, \]

where \( \sigma_{\text{calo}} \) and \( \sigma_{\text{TA}} \) are the calorimeter-based jet mass resolution and the track-assisted mass resolution, respectively. The jet mass resolution for the calorimeter mass, track-assisted mass and combined mass are shown in Figure 12 as a function of jet \( p_{T} \). Similar techniques that take advantage of the excellent angular resolution of the tracking detector at high \( p_{T} \) have been developed to correct topoclusters to improve the resolution of jet substructure variables [65].

It is important to point out that in ATLAS unlike in CMS, the jet energy scale directly impacts the jet mass scale. As opposed to the description of the JES calibration for small-\( R \) jets in Section 1.3, the area subtraction, residual correction and GSC calibration are not applied to large-\( R \) jets.

In CMS, the jet mass is by default reconstructed as a combination of track and calorimeter measurements via the virtues of the particle flow algorithm. Since the particle flow algorithm reconstructs neutral hadrons from a single energy cluster in the less granular HCAL, which can be associated to multiple clusters in the finer segmented ECAL, the HCAL energy is split according to the number and energy distribution of the ECAL clusters, thereby obtaining an optimal estimate of the spatial energy distribution as input to jet substructure algorithms [27]. Since the residual in-situ jet energy corrections are not applied when reconstructing jet masses, dedicated corrections are derived from simulation and data. Firstly, the jet mass response is corrected as a function of \( p_{T} \) and \( \eta \) using simulation of \( W \) jets from boson pair production. Secondly, residual corrections are obtained from a data sample enriched in lepton+jets \( t \bar{t} \) production where the hadronic \( W \) jet can be studied in data [50, 59]. The selection is de-
Figure 13: Jet mass distribution in a sample enriched with lepton+jets $t\bar{t}$ events, where the hadronic $W$ jet with $p_T > 200$ GeV is selected, taken from [59].

finely by requiring exactly one isolated lepton (electron or muon), large missing $p_T$, at least one $b$-tagged $R = 0.4$ jet, and one groomed (pruning or soft drop) $R = 0.8$ jet with $p_T > 200$ GeV in the final state. The contribution from hadronic $W$ jets in this sample shows a peak at the $W$ mass in the jet mass distribution, as shown in Fig. 13 for the soft drop grooming case. This is used to calibrate the JMS and JMR in CMS, and can also be used for dedicated efficiency corrections on other jet substructure observables, such as the $N$-subjettiness ratio $\tau_2/\tau_1$.

The JMS is described within 1-2% by simulation and the JMR differs by about 10%, which is about the same size as the statistical uncertainty of this measurement. Since these measurements are performed in samples of $W$ jets with $p_T \approx 200$ GeV, additional systematic uncertainties apply at higher $p_T$ [66]. A detailed study of the various contributions to the JMS has also been performed for fully merged top-jets in the context of an unfolded top-jet mass measurement [67]. To summarize the impact of the various sources of systematic uncertainty to the measurement of residual corrections for jet substructure observables, we quote here the dominant uncertainties related to the scale factor measurement of an $N$-subjettiness ratio $\tau_2/\tau_1 < 0.4$ selection [59]. The statistical uncertainty of $6\%$ (with $2.3/\text{fb}$ of data) is comparable to the systematic uncertainties related to the simulation of the $t\bar{t}$ topology (nearby jets, $p_T$ spectrum) contributing 4%, the choice of method to derive the scale factors contributing 6% and the modeling of the $p_T$ dependence that rises from 5% at $p_T = 500$ GeV to 13% at $p_T = 2000$ GeV.

2 Pileup Mitigation Techniques

2.1 Defining Pileup

Pileup originates from incidental proton-proton ($pp$) collisions that occur in addition to a hard scattering collision of interest, referred to here as the Primary Vertex (PV). Pileup is uncorrelated with the PV and typically consists of an admixture of inelastic, elastic and diffractive $pp$ processes occurring either in-time or out-of-time with the hard scatter. Most pileup mitigation techniques focus on in-time pileup, which occurs in the same proton bunch crossing as the PV. The effects of out-of-time pileup (a result of leftover signals from neighboring bunch-crossings) can be substantial, but are often sufficiently suppressed through the detector signal processing.

During the LHC Run 1 the mean number of pileup interactions reached $\langle \mu \rangle = 21$, and $\mu$ values up to 60 were achieved in certain runs of 2017 (Run 2) with possibly even higher values in Run 3, and culminating at the high luminosity LHC (HL-LHC) reaching up to $\langle \mu \rangle = 140–200$.

To a good approximation, pileup typically leaves about 0.5 GeV of energy in the detector per unit area ($\eta, \phi$), per pileup vertex: the effects of this are present in all aspects of LHC physics, from detector design and software performance to the final sensitivity of measurements and searches.

2.2 Mitigation Methods

There are a number of basic properties of pileup interactions which can be exploited in order to discriminate pileup particles from particles originating from the primary vertex. The basic strategies for removing pileup at the LHC are based on these properties, listed below, with examples of some of the commonly-used methods falling in each category.

(a) Asymptotic behavior: The collective behavior of pileup can be treated as a uniform field of radiation.
Area Subtraction uses a pileup $p_T$ density per unit area estimator, $\rho$, and defines a jet catchment area, $A$, to remove energy that originate from pileup interactions from a jet. This corrects the jet naively in the following way: $p_T^{\text{corr}} = p_T^{\text{orig}} - \rho \times A$. An example of $\rho$ can be seen in Fig. 14. There are many subtleties in defining both $\rho$ and $A$, which are discussed in e.g. Refs. [68, 69, 70]. Extensions to this method, which corrects only the jet $p_T$, include Shape Subtraction [71] and Constituent Subtraction [57, 72, 73], the latter of which will be discussed in more detail below.

(b) Charged particle vertexing: Charged particles leave tracks in high granularity tracking detectors at the heart of multi-purpose detectors like ATLAS and CMS and can be separated based on their $\hat{z}$ (along the beamline) position in the beam luminous region (see Fig. 15). Charged Hadron Subtraction (CHS) [69] uses the charged particle tracking information to identify pileup by each track individually. Used in concert with particle flow concepts which attempt to identify each particle in the event uniquely, CHS can effectively remove all charged pileup radiation from the event. Identification of Pileup Jets, formed predominantly from the energy of one or many pileup vertices, is another technique for removing pileup using charged particles; by determining the fraction of energy of the jet from the primary vertex, one can distinguish such pileup jets from the PV jets [70, 74].

(c) Local, topological information: QCD radiation from pileup vertices are often uncorrelated and soft and can thus be removed based on the local energy profile, i.e. if the radiation is not consistent with hard scattering radiation from the PV. This can be done in the transverse plane $\eta, \phi$ and also as a function of radiation depth $r$. Grooming (Sec. 1.4) is a technique to clean the jet of soft and wide-angle radiation which incidentally removes pileup radiation. Topoclustering [17], used by the ATLAS Collaboration, is deployed at the formation of clusters in the calorimeter requiring radiation to have a certain topological profile. In the forward region, where no tracking information is available, jet shapes and topological correlations can be used to identify pileup [75].

(d) Timing information: Detector signals are read out at the LHC bunch frequency (40 MHz) or faster, and can be used to reduce out-of-time pileup. With potential future detectors, precision timing capabilities with $O(10 \, \text{ps})$ could possibly distinguish between pileup vertices within the same $pp$ bunch crossing, as pileup interactions can differ from the PV in time as well as $\hat{z}$ position.

While the above methods have been successfully deployed in the LHC experiments, they each have some deficiencies as well; ideally, one would hope to effectively combine all pileup mitigation handles in order to maximally distinguish pileup from PV radiation and to remove pileup at the most granular level possible, i.e. at the particle or constituent level, in order to be as generic as possible. For example, while area subtraction is very effective for correcting the jet $p_T$, it is not used to mitigate the pileup dependence of jet substructure observables as it is only able to correctly remove pileup contributions on average. In fact, jet substructure variables are among the most difficult to correct for pileup because they are so reliant on radiation profiles. A number of hybrid methods have been proposed operating at the event constituent level, a few examples of which are Constituent Subtraction [73], Soft-
Killer [72], PUMML [77] and PUPPI [57]. Precursor hybrid methods include *jets without jets* [78] and *jet cleansing* [79]. Most hybrid strategies assume perfect tracking information for identifying charged particles from pileup. Constituent Subtraction and SoftKiller use additional information based on the energy density ($\rho$) to further remove neutral particles from the event, while PUPPI uses both event energy density and local topological information incorporated in an event-by-event particle-level discriminator to determine if a particle is from pileup. Particle four-vectors are then weighted proportional to the value of the discriminator value.

### 2.3 Performance Studies

Pileup removal algorithms are commissioned for use in ATLAS and CMS via detailed studies of jet observables in terms of the resolution and absolute scale, pileup dependence, and ultimately the effective background rejection power versus signal efficiency for boosted heavy particle taggers.

For observables like $p_T$, dependencies on the number of reconstructed vertices and $\mu$ are observed even with area subtraction methods for the pileup levels currently observed at the LHC, $\langle \mu \rangle \sim 25$. To correct for these effects, an additional residual correction is applied [30, 35]. Enhancements are also possible from combining area subtraction methods with e.g. CHS.

For jet substructure observables, particle- or constituent-level pileup mitigation strategies have been shown to improve performance, especially in simulation studies going up to $\langle \mu \rangle \approx 40$. An example is given in Fig. 16, where the leading ungroomed jet mass in simulated QCD multijet events is corrected with different pileup removal techniques. The jet mass resolution can be improved further when using grooming algorithm. The effect of different pileup removal techniques on the groomed jet mass depends however strongly on the choice of the grooming algorithm as discussed in detail in Ref. [39, 69].

Generally these techniques, particularly those which operate at particle-level, can also be used to improve performance of non-jet objects such as missing transverse energy and lepton isolation.

Preliminary studies (detector configurations have not yet been finalized) into the application of these advanced hybrid techniques at the higher pileup levels anticipated at the HL-LHC suggest that they are effective in the $\langle \mu \rangle = 140 − 200$ range [80, 81].
3 Boosted Particle Tagging

Particle identification is an experimental challenge that is traditionally faced using custom-designed charged-particle detectors, muon chambers and calorimeters with granularity fine enough to allow shower shape measurements. The ATLAS and CMS detectors were designed with particle identification in mind. The use of jet grooming to reveal the jet mass for top quark and boson decays has provided experimentalists with a potential new way to identify high $p_T$, massive particles - exactly the kind of particles that are most interesting for precision measurements and new physics searches at the LHC. Several substructure variables have been developed by the theoretical community that can be used along with the jet mass for boosted particle identification. The term ‘tagger’ indicates the use of one or more of these variables (often after grooming has been applied) to discriminate between jets coming from different types of particle. A summary of the tagging algorithms currently in use by ATLAS and CMS is tabulated in Table 2.

3.1 Quark/Gluon Discrimination

In the nearly 40 years since the discovery of the gluon at PETRA [82, 83, 84, 85], jet substructure observables have been widely used for quark-initiated (quark) versus gluon-initiated (gluon) jet tagging. The probability for a gluon to radiate a gluon is enhanced by a factor of $C_A/C_F = 9/4 \sim 2$ over the probability for a quark to radiate a gluon of the same energy fraction and opening angle [86]. As a result, gluon jets tend to have more constituents and a broader radiation pattern than quark jets. There are also more subtle differences due to quark and gluon electric charges and spins.

There are three key challenges of quark versus gluon jet (q/g) tagging: (1) quark and gluon labeling schemes are not unique; (2) for a given labeling scheme, quark and gluon jets are not that different; (3) the differences that do exist are sensitive to both perturbative and non-perturbative modeling choices. Since quarks and gluons carry color charge and only colorless hadrons are observed, there is not a unique way to label a jet in simulation as originating from a quark or a gluon. Many labeling conventions exist, ranging in simplicity and model-dependence from matching to out-going matrix element partons to parsing an entire jet clustering history [87, 88]; however, no treatment escapes the problem that the notion of a quark and gluon jet are not universal\(^2\): quark and gluon jet radiation depends on the production mechanism. There is a plethora of jet substructure observables that can be used for q/g tagging; see e.g. Ref. [90] for a large survey. Many of these observables exhibit Casimir scaling which results in nearly the same, limited discrimination power for all the observables [91, 92]. The most powerful single q/g observable is the particle multiplicity inside a jet, which does not exhibit Casimir scaling and recent theoretical advances [93] have shown that its discrimination power can be largely understood from perturbative theory. There is further q/g separation possible when using the full radiation pattern inside a jet, though the combination of multiplicity and a Casimir scaling observable carries a significant fraction of the total discrimination power [94].

The modeling of q/g tagging observables has a long history - see Ref. [95] for a recent and detailed study.

Despite the challenges listed above, both ATLAS and CMS extensively use explicit or implicit quark versus gluon tagging. Explicit taggers are algorithms designed to directly isolate quark and gluon jets while implicit techniques are designed for another purpose that also happens to perform some quark versus gluon jet tagging. The explicit taggers developed by ATLAS [96, 97, 98, 99]

\(^2\)This can be largely mitigated by jet grooming; see e.g. Ref. [89].
and CMS [59, 100, 101, 102] include a variety of observables and data-driven calibration and validation techniques. These and related techniques have been successfully deployed in a variety of physics analyses (see e.g. [27, 103, 104, 105, 106, 107, 108]). Additionally, it has been shown that an improved W tagger can be constructed by utilizing q/g discrimination on subjets [27].

Both ATLAS and CMS have developed likelihood-based discriminants for explicit q/g tagging. The discriminants are constructed from variables sensitive to the radiation pattern of quark and gluon jets, also taking into account differences between light (uds) and heavy flavor (cb) quark jets, where the latter are more similar to gluon jets. ATLAS uses the number of tracks $n_{\text{track}}$ as an approximation for the number of jet constituents and the jet width [99] while CMS utilizes the number of particle-flow constituents $n_{\text{const}}$, the jet axes and fragmentation functions [59]. Since the distributions of these variables depend on $\eta$, $p_T$, and $\rho$, the likelihood discriminators are constructed differentially with respect to these variables. In Run 2, ATLAS also introduced a simple and robust tagger using solely $n_{\text{track}}$ [98], which has the advantage of a much-simplified uncertainty derivation.

Figure 17 shows the CMS q/g tagging performance in simulation. For a 50% gluon (quark) efficiency, the quark (gluon) efficiency is about 10%. This performance depends slightly on the jet $p_T$, in part because the particle multiplicity increases with $p_T$ (and therefore the performance improves). The $p_T$ and $|\eta|$ dependence is illustrated for the same tagger in Fig. 18. Outside the tracking acceptance ($|\eta| > 2.5$), q/g tagging significantly degrades due to the coarse calorimeter granularity and increased pileup sensitivity.

ATLAS [99] and CMS [102] are also actively studying sophisticated approaches based on modern machine learning. While these methods hold great promise for their power and flexibility, Fig. 19 indicates that the deep convolutional neural network (CNN) has similar performance to a 2D binned likelihood ratio (LLH) based on only two variables for lower efficiencies whereas promising performance gains are observed for high efficiencies. Architecture design and input optimization are still an active area of research and development.

The modeling of q/g discriminating observables is a key concern for tagging applications. Typically, Pythia tends to describe quarks better than Herwig, whereas the opposite is observed for gluons. Figure 20 illustrates the typical magnitude of the mis-modeling (the differences between the dashed and dotted lines). Pythia tends to overestimate the q/g tagging performance with respect to data, as illustrated quite strikingly in Figure 21. In contrast, Herwig tends to underestimate the performance (leading to the famous ‘Pythia-data-Herwig sandwich’).

Data-driven reweighting techniques can be used to mitigate the observed mis-modeling (solid and open symbols...
in Fig. 20). In particular, multiple samples with a different (but known) q/g composition can be used to extract the distribution of q/g tagging observables. ATLAS and CMS have both used Z/γ+jets and dijet samples, which are enriched in quark and gluon jets, respectively.

The Run 2 ATLAS tagger is based entirely on dijets, exploiting the rapidity dependence of the q/g fraction to extract the track multiplicity separately for quarks and gluons. A Run 1 measurement is used to constrain the particle-level modeling, and dedicated track reconstruction uncertainties are used to complement the particle-level uncertainty with a Run 2 detector-level uncertainty. The uncertainties on q/g tagging are 2-5% over a wide range of 200 GeV < p_T < 1 TeV at a working point of 60% quark jet efficiency [98]. The template-based calibration can also be used to directly construct the q/g tagger in data, avoiding mis-modeling concerns; however, when more than two observables are used to construct the tagger, it becomes impossible in practice to extract the high-dimensional templates.

Despite its power, the template technique has some residual non-closure because the resulting calibrated tagger applied to another final state may not have the same performance. This is illustrated in Fig. 22, which shows how the average track multiplicities extracted for quark and gluon jets (using high-purity Z/γ+jets and dijets data respectively) differ from the values obtained in the γ+2-jet and trijet samples used for validation.

Explicit tagging is often the focus of modern q/g discrimination, but there is a broad program of implicit tagging as well. One ubiquitous example of this is the ATLAS jet calibration procedure. Since the calorimeter response is non-linear, a jet with a higher particle multiplicity will have a lower response for the same energy. After applying a simulation-based correction to eliminate this inclusive bias in the JES (see Figure 6), a residual calibration is ap-
plied to correct for the dependence of the bias on the number of tracks associated to the jet and the jet width [33]. After applying this residual Global Sequential Calibration, the difference in response between quark and gluon jets is reduced. Implicit q/g tagging also appears in pileup jet identification [74, 75], boson and top tagging [27, 106, 108], and elsewhere.

Despite its long history, quark versus gluon jet tagging is still a very active topic of research. Since most analyses at the LHC target processes with a known and asymmetric q/g jet composition, q/g tagging holds great promise for improving searches and measurements in the future. Further studies are required to understand the limits of q/g tagging performance and to mitigate the sample dependence for universal definitions and calibrations. Interestingly, recent studies have shown how modern machine learning classifiers can be directly trained on data even though there are no per-jet labels [109, 110].

### 3.2 Vector Boson Tagging

The hadronic, two-prong decays of weak vector bosons V have a distinct radiation pattern compared to individual high-\(p_T\) quark or gluons; however, distinguishing V-jets from quark- or gluon-initiated jets requires a high-efficiency tagger across a range of \(p_T\) from 0.2 TeV up to multiple TeV, and \(|\eta|\) from 0 to 2.5, while keeping the misidentification rate for quark and gluon jet backgrounds small. Stability of the misidentification rate with e.g. \(p_T\) and \(\eta\) allows data-driven background estimates to be used. This is necessary in light of the known mis-modeling in multijet Monte Carlo simulations. Good separation power between W and Z bosons is also desirable in a number of analyses, most notably searches for diboson resonances (see Sec. 5.2).

ATLAS and CMS performed a broad range of studies during Run 1 and the beginning of Run 2, systematically identifying the influence of pileup reduction and grooming techniques on jet substructure observables used for V tagging [37, 50]. Simulated samples containing W jets (rather than Z jets) are primarily used for these studies, as W jets are abundant in data thanks to the large quantity of \(t\bar{t}\) events.

The optimization of the V-tagging algorithm is generally based on various factors concerning the tagged jet mass: (i) a sensible JMS (i.e., tagged jet mass close to the W mass), (ii) a narrow jet mass response with an approximate Gaussian lineshape, (iii) stability with respect to pileup and jet \(p_T\), and (iv) good background rejection at a given signal efficiency. Considering all of these factors, ATLAS decided on using the Trimming algorithm with \(f_{\text{cut}} = 0.05\) and \(R_{\text{sub}} = 0.2\) on anti-\(k_t\), \(R = 1.0\) jets in Run 2, while CMS opted for using anti-\(k_t\), \(R = 0.8\) jets, treating the pileup first with PUPPI and then applying soft drop grooming with \(\zeta_{\text{cut}} = 0.1\) and \(\beta = 0\).

In addition to the comprehensive studies of grooming options [27, 37, 38, 39, 111], ATLAS and CMS both investigated the discrimination powers for a plethora of jet substructure variables, including: \(N\)-subjettiness [60, 61], Qjet volatility [112], ratios of energy correlation functions \(C_2^\beta\) [91], \(D_2^\beta\) [113, 114] and \(N_2^\beta\) [115], angularities and planar flow [116], splitting scales [117, 118] the jet and subjet quark/gluon likelihood, and the jet pull angle [119].

Both ATLAS and CMS developed simple taggers that rely on the combination of the jet mass with one other variable that improves the discriminating power between the signal and background. The standard ATLAS V-tagger for Run 2 was chosen to be the trimmed jet (\(R = 1.0, f_{\text{cut}} = 0.05, R_{\text{sub}} = 0.2\)) mass and \(D_2^{\beta=1}\) [37], known as ‘R2D2’, while CMS opted to use the soft drop jet (\(\zeta_{\text{cut}} = 0.1, \beta = 0\)) mass and the \(N\)-subjettiness ratio \(\tau_{21} = \frac{\tau_2}{\tau_1}\). Despite the different choices of tagging observables and detector design, ATLAS and CMS reach a very similar background rejection at a given tagging efficiency. An active field of developments is the usage of multivariate techniques for boosted V identification which have shown to be able to significantly improve the background rejection [50, 120].

In the ATLAS studies the variable \(C_2^{\beta=1}\) in combination with the trimmed jet mass has been shown to be as good a discriminator as \(\tau_{21}\) as shown in Fig. 23. This is in contradiction to the study by CMS, where \(C_2^\beta\) is one of the weaker observables; however, a direct comparison is difficult, since in ATLAS groomed substructure variables are used, calculated for trimmed jets, while in CMS ungroomed variables are used. Also, the particulars of particle reconstruction have a large impact on the performance of individual observables. While a study of the performance of \(D_2^\beta\) at CMS is still pending, the soft drop \(N_2^\beta\) observable was found to give similar performance to \(\tau_{21}\) in CMS [122].

CMS studied the quark/gluon likelihood (QGL) discriminator for its potential in V-tagging applications in

---

\(^3\) A different axis definition for the subjet axes is used in ATLAS when calculating \(\tau_N\), known as the \(\text{the-winner-takes-all}\) axis [121], which is consistently found to perform slightly better than the standard subjet axis definition in tagging bosons.
Run 1 [27], finding that a combination of the groomed jet mass and the QGL achieved a similar discrimination power as the groomed jet mass and $\tau_{21}$. When adding the QGL to the Run 1 V-tagger (pruned jet mass and $\tau_{21}$), the misidentification rate was reduced slightly from 2.6% to 2.3% at a constant signal efficiency of 50%. A similar reduction of the misidentification rate was observed when adding $C_2^{\beta^2}$, showing that $C_2^{\beta^2}$ carries additional information with respect to the groomed jet mass and $\tau_{21}$. However, the QGL and $C_2^{\beta^2}$ exhibit a considerable pileup dependence, resulting in a degradation of their discrimination power with increasing activity. This pileup dependence is expected to be reduced when using PUPPI in place of particle flow + CHS.

In Fig. 23 the ATLAS measurements of signal efficiencies versus background rejection power are shown for $350 < p_T < 500$ GeV. Taken from Ref. [37].

Figure 23: Signal efficiency versus background rejection power compared with measurements from ATLAS for $350 < p_T < 500$ GeV. Taken from Ref. [37].

Run 1 [27], finding that a combination of the groomed jet mass and the QGL achieved a similar discrimination power as the groomed jet mass and $\tau_{21}$. When adding the QGL to the Run 1 V-tagger (pruned jet mass and $\tau_{21}$), the misidentification rate was reduced slightly from 2.6% to 2.3% at a constant signal efficiency of 50%. A similar reduction of the misidentification rate was observed when adding $C_2^{\beta^2}$, showing that $C_2^{\beta^2}$ carries additional information with respect to the groomed jet mass and $\tau_{21}$. However, the QGL and $C_2^{\beta^2}$ exhibit a considerable pileup dependence, resulting in a degradation of their discrimination power with increasing activity. This pileup dependence is expected to be reduced when using PUPPI in place of particle flow + CHS.

In the ATLAS measurement the dominating systematic uncertainty comes from the jet substructure scale, which has been derived by comparing calorimeter-jets with track-jets. Once again, the distributions in data lie between the ones derived with Pythia and Herwig, leading to large modeling uncertainties [37, 38]. A similar observation is made by CMS [50, 59]. Improving the modeling of jet properties and thereby reducing the differences between different event generators is a major task, but crucial for future precision studies using jet substructure.

A crucial aspect of V-tagging is the derivation of background rates from multijet production in real collision data when performing measurements. A commonly used method is the extrapolation from one or more control regions, which are defined orthogonally to the signal region. Usually, these control regions are defined by inverting the jet mass window selection, see e.g. [123, 124, 125, 126].
Simulation Preliminary

13 TeV

CMS

the mean reconstructed background efficiency shows a residual as a function of \( p_T \). The requirement is thus a flat signal or background efficiency (depending on the needs of the analysis). In order to achieve a flat signal efficiency, ATLAS developed a \( p_T \)-dependent selection on the value of \( D_2 = \frac{1}{2} \), as this distribution shows a strong dependence on \( p_T \) [38]. In contrast to the Run 1 studies described above, no \( p_T \)-dependent selection is made on the trimmed jet mass, as the calibrated jet mass is used to define the V-tagging working point. While the jet mass resolution still increases with \( p_T \), a constant window of \( \pm 15 \text{GeV} \) around the mean reconstructed \( W \) or \( Z \) boson mass is used. This results in a \( p_T \)-dependent signal and background efficiency, which can also be countered with the \( p_T \)-dependent cut on \( D_2 = \frac{1}{2} \). This leads to a constant signal efficiency, while the background efficiency shows a residual \( p_T \) dependence, as shown in Fig. 24.

Another possibility has been explored by CMS. Instead of introducing \( p_T \)-dependent selection criteria, a linear transformation of the ratio \( \tau_{21} \) has been studied [59], given by \( \tau_{21}^{\text{DDT}} = \tau_{21} - M \cdot \log(m^2/p_T/1 \text{ GeV}) \) [129], where \( M \) is a constant determined from simulation. The replacement of \( \tau_{21} \) with \( \tau_{21}^{\text{DDT}} \) does not affect the overall performance of the tagger, but results in an approximately flat misidentification rate as a function of \( p_T \), as shown in Fig. 25 (bottom), while introducing an increasing \( p_T \)-dependence for the V-tagging efficiency, as shown in Fig. 25 (bottom).

The development of decorrelated jet substructure taggers is an active field with new techniques e.g. described in Refs. [130, 131, 132].

A less-studied possibility to lift the \( p_T \)-dependence of substructure observables is the application of variable-\( R \) jets [133]. By shifting the \( p_T \)-dependence to the jet-clustering level with a distance parameter proportional to \( p_T^{-1} \), a stable position of the jet mass and jet substructure variables with respect to changes in \( p_T \) can be achieved [134]. This can lead to a stable tagging performance without the necessity of \( p_T \)-dependent optimization steps, but further experimental studies are needed to commission this strategy for use in analyses.

For some analyses the requirement of \( p_T \gtrsim 200 \text{ GeV} \) is too restrictive, and hadronically decaying \( V \) bosons with lower \( p_T \) need to be selected. This poses a particular challenge due to the abundance of light flavor jets at the LHC and their indistinguishability from jets from \( W/Z \) decays. An attempt was made by CMS to discriminate ‘resolved’ (non-merged) hadronic \( W \) decays from multijet background using the QGL, the sum of the jet charges of the dijet pair and the jet pull angle. Combining these variables into a Boosted Decision Tree, a misidentification rate of about 25% is achieved for a signal efficiency of 50% [27].

While this is a first success, the performance is about an order of magnitude worse than V-tagging for fully merged decays, showing the power of substructure techniques in

Figure 25: Efficiency and misidentification rate of various identification techniques for boosted \( W \) tagging. Taken from Ref. [59].
The boson type-tagger was optimized for a relatively low boson boost, $200 < p_T < 400\text{ GeV}$. The discrimination power of all of the input variables degrades with $p_T$ due to the worsening jet mass resolution, tracking efficiency and momentum resolution, as well $b$-tagging efficiency. However, there are recent developments to address each of these challenges, such as the track-assisted jet mass (Sec. 1.5), pixel-cluster splitting [19], and track-jet $b$-tagging [136].

### 3.3 Top Tagging

The three-prong decays of highly boosted top quarks in the fully hadronic decay channel offer richer phenomenology for their identification than the two-prong decays of $W/Z/H$ bosons. This has been exploited in a number of algorithms, which usually aim at an optimal performance in a particular kinematic regime.

The techniques for tagging boosted top quarks have evolved as fairly complex methods in comparison to the $V$-taggers; these techniques include:

(a) The Johns Hopkins / CMS top tagger (CMSTT) [137] was designed for tagging top quarks with $p_T > 1\text{ TeV}$. The algorithm is based on a decomposition of the primary jet into up to four subjets by reversing the Cambridge/Aachen clustering sequence. It has been adapted by the CMS Collaboration [138, 139], and was adopted as the standard top-tagging algorithm in CMS in Run 1, where it was typically used in the region of $p_T > 400\text{ GeV}$, with an average identification efficiency of 38% at 3% misidentification rate [51].

(b) The HEPTopTagger (HTT) [140, 141] was designed to target $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel. In $t\bar{t}H$ production the top quark $p_T$ distribution peaks around 150 GeV and is steeply falling towards increasing $p_T$, where it is already an order of magnitude smaller at $p_T \sim 400\text{ GeV}$. This results in a requirement of non-zero signal efficiency already at $p_T \approx 200\text{ GeV}$, where the top quark decay is only moderately boosted. The HTT achieves this with a large jet distance parameter of 1.5 and a sequence of declustering, filtering and re-clustering of the original Cambridge/Aachen jet. The performance of the HTT was studied by the ATLAS and CMS Collaborations on data with a center-of-mass energy $\sqrt{s} = 7$ and 8 TeV [40, 51, 142]. Efficiencies of 10% with misidentification rates of 0.5% for jets with...
200 < \pT < 250 \text{ GeV} \) were observed. The efficiency increases with increasing jet \pT, where a plateau is reached for \pT > 400 \text{ GeV}, with efficiencies of approximately 40% at 3\% misidentification rate, very similar to the performance achieved with the CMS TT.

(c) Shower Deconstruction [143, 144] was designed to be analogous to running a parton shower Monte Carlo generator in reverse, where emission and decay probabilities at each vertex, color connections, and kinematic requirements are considered. Small-radius (generally \( R = 0.2 \)) subjets are reconstructed with the Cambridge/Aachen algorithm and all possible shower histories that can lead to the observed leading final state anti-\( k_t \), \( R = 1.0 \) jet are calculated. Each shower history is assigned a probability weight factor based on the aforementioned considerations (to be signal-like or background-like), then a likelihood ratio \( \chi(p_N) \) is constructed, and the log \( \chi(p_N) \) is used as the discriminating substructure variable. For top quark tagging, efficiencies of 80\% with misidentification rates of 50\% for jets with 500 < \pT < 1000 \text{ GeV} were observed. The efficiency increases with increasing jet \pT, where a plateau is reached for \pT > 2000 \text{ GeV}, with efficiencies of \( \sim 80\% \) at 10\% misidentification rate. Recently, the Shower Deconstruction algorithm was optimized for top quarks with \pT > 800 \text{ GeV} in context of the \( W' \) to \( t\bar{b} \) hadronic search [145] by using exclusive \( k_T \) subjets.

In addition to the dedicated techniques described above, simpler algorithms using grooming and substructure as per \( V \)-tagging have been investigated by ATLAS for a number of years. A performance study at 7 \text{ TeV} [142] investigated a variety of performance metrics relating to the usage of groomed jets. Different grooming algorithms were investigated for their resilience to pileup and mass resolution. It was concluded that trimmed anti-\( k_t \) jets with a distance parameter of 1.0 and trimming parameters of \( R_{\text{sub}} = 0.3 \) and \( f_{\text{cut}} = 0.05 \) were a very good candidate for a one-fits-all large-\( R \) jet definition. This jet definition became standard in ATLAS for \( W/Z/H \) and top quark tagging in Run 1. A later ATLAS study [40] investigated the various methods available for tagging hadronic, highly boosted top quarks. The so-called Tagger V has \( M_{\text{jet}} > 100 \text{ GeV}, \sqrt{d_{12}} > 40 \text{ GeV} \) and \( \sqrt{d_{23}} > 20 \text{ GeV} \), where \( \sqrt{d_{ij}} \) is the \( k_T \)-splitting scale [117]. The efficiency versus rejection is shown for various taggers in Fig. 27. The difference between Taggers III and V is the additional requirement on \( \sqrt{d_{23}} \) in Tagger V. At efficiencies smaller than 45\%, the \( W' \) tagger, based on \( \sqrt{d_{12}} \) and the \( N \)-subjettiness ratios \( \tau_{21} \) and \( \tau_{32} \), has better background rejection than Taggers III and V. ATLAS also tested the HTT and Shower Deconstruction [146], which have been found to have good background rejection (larger than 50) for efficiency values smaller than about 35\%. However, similar as for the CMS experiment, the background efficiencies of the two taggers show a significant rise with increasing \pT.

CMS has focused on enhancing the performance of CMS TT and HTT by identifying observables which carry discriminatory power, but have only small or moderate correlations with the observables used in the main algorithm. Typically, correlation coefficients of about 0.3 or less are required for noticeable improvement when augmenting an algorithm with additional variables. Examples for discriminating variables which fulfill this are \( N \)-subjettiness ratios, energy correlation functions and their ratios, and \( b \)-tagging. A study by the CMS Collaboration showed that at 20\% signal efficiency, the background rejection of the CMSTT could be improved by a factor of 5 when adding information from \( \tau_{32} \) and subjet \( b \)-tagging [51]. At higher efficiencies, the improvements become smaller. For the HTT, improvements of similar size were observed for \pT > 200 \text{ GeV}, becoming less significant at higher \pT.

The ATLAS choice of \( R = 1.0 \) jets compared to CMS (\( R = 0.8 \)) results in an earlier rise of the tagging efficiency with increasing jet \pT.

The large difference in performance of the single variable
\( \tau_{32} \) between ATLAS (Fig. 27) and CMS (Fig. 28) is due
to jet grooming. Although the CMS study shows only the
ROC curves for 800 < \( p_T \) < 1000 GeV, the overall picture
does not change when studying top quarks in the region of
\( p_T \approx 400 \) GeV. Instead, in ATLAS \( \tau_{32} \) is calculated from
trimmed jets, which results in less discrimination power
when used as sole tagging variable compared to ungroomed
\( \tau_{32} \). However, groomed \( \tau_{32} \) can still lead to considerable
improvements when combined with other variables.

As with V-tagging discussed above, ATLAS and CMS
took advantage of the LHC shutdown between Run 1 and
Run 2 to perform broad studies of the different top-taggers
available, with emphasis on their stability with respect to
pileup and other detector effects, instead of the utmost
gain in performance \([41, 52]\). Single variables and their
combinations were studied and compared with Shower De-
construction, CMSTT, HTT, and an improved version of
the HTT with shrinking cone size (HTTv2) \([147]\).

Figure 28 shows a comparison based on simulation of
the single variable performance in CMS, where signal jets
were generated through a heavy resonance decaying to \( t\bar{t} \)
and background jets are taken from QCD multijet prod-
duction. Note that for this study reconstructed jets are
matched to a generated parton, and the distance between
the top quark and its decay products must be less than 0.6
(0.8) for a reconstructed \( R = 0.8 \) (1.5) jet, to ensure that
the top quark decay products are fully merged and recon-
structed in a single jet. The best single variable in terms of
efficiency versus background rejection is the discriminator
\( \log \chi \), calculated with Shower Deconstruction. The second
best variables are the \( N \)-subjetiness ratio \( \tau_{32} \) at low effi-
ciency and the jet mass calculated with the HTTv2 at high
efficiency values. The individual groomed jet masses show
similar performance, and the CMS Collaboration moved to
using the soft drop mass due to its beneficial theoretical
properties \([7]\). The default for CMS Run 2 analyses was
chosen to be the soft drop jet mass combined with \( \tau_{32}/\tau_2 \)
for top tagging at high \( p_T \). Generally, at high boost, the
combination of a groomed mass with \( \tau_{32} \) leads to a large
gain in background rejection.

The CMS study also investigated combining single vari-
ables with more complex taggers. Combining Shower De-
construction with the soft drop mass, \( \tau_{32} \), and subjet b-
tagging can lead to improvements, as shown in Fig. 29;
however, the efficiency and misidentification rate for this
combination were found not to be stable as a function of
jet \( p_T \) (the combined algorithms were studied using work-
ing points corresponding to a background efficiency of 0.3).

At low boosts, the dedicated HTTv2 shows the best per-
formance. In this kinematic region, using groomed \( \tau_{32} \),
obtained by using the set of particles from the soft drop
jet instead of the original jet, helps to improve the perform-
ance.

In the shutdown between Run 1 and Run 2, ATLAS
commissioned a single top tagger for use by physics analy-
ses. The rationale behind this approach was the poten-
tial benefit of having an efficient top tagger with well-
understood efficiency and associated systematic uncertain-
ties validated in the Run 1 dataset. Similarly as for Run
1, the supported top-tagger makes use of anti-\( k_t \), \( R = 1.0 \)
trimmed jets, but with a parameter of \( R_{\text{sub}} = 0.2 \)
instead of 0.3 as used in Run 1. Candidate top jets are
required to satisfy a calibrated mass window requirement
122.5 < \( M_{\text{jet}} \) < 222.5 GeV and a \( p_T \)-dependent, one-sided
cut on \( \tau_{32} \) \([41]\). The background rejection for fixed signal
efficiency working points for the pairwise variable combi-
nations considered is shown in Ref. 30. The ratio \( \tau_{32} \) was
used instead of the previously used \( k_t \)-splitting scales or
the minimum pairwise mass \( Q_{W} \) \([118]\) because of a smaller
correlation with \( M_{\text{jet}} \) (shown in Ref. 31), a reduced \( p_T \)-
dependence, and good performance of this combination
across a large range in \( p_T \). The distribution of \( \tau_{32} \) mea-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure28}
\caption{Background versus signal efficiency for the sin-
gle variables studied in the optimization of top tagging for
CMS Run 2 analyses. Taken from Ref. [52]}
\end{figure}
Figure 29: Background efficiency versus signal efficiency for combined variables studied in the optimization of top tagging for CMS Run 2 analyses. Taken from Ref. [52].

A common problem of top-tagging algorithms is the rise of the misidentification rate with increasing $p_T$, which is due to the peak of the mass distribution for quark- and gluon-initiated background jets shifting to higher values. For some taggers, for example the CMSTT, this shift also results in a decrease of the efficiency once a very high $p_T$ threshold is crossed (larger than 1 TeV) [139]. A possible solution to this is offered by the variable-$R$ (VR) algorithm, introduced in Sec. 3.2. The ATLAS Collaboration studied the performance of the VR algorithm for top-tagging and reported a stabilization of the position of the jet mass peak for a large range of $p_T$ [134]. The VR jets are shown to improve the performance of the jet mass, $\sqrt{d_{12}}$ and $\tau_{32}$ for top tagging, when compared to trimmed jets.

Most top-taggers target either the region of low to intermediate boosts, or the highly boosted regime. However, in typical searches for new physics at the LHC non-vanishing efficiency for the full kinematic reach is crucial. Several
attempts of combining different reconstruction and identification algorithms have been made. A search for resonances decaying to $t\bar{t}$ by the ATLAS Collaboration uses a cascading selection from boosted to resolved [150], where the resolved topology is reconstructed and identified using a $\chi^2$-sorting algorithm. To efficiently identify top quarks over a broad $p_T$ range in the search for top squark pair production, reclustered variable-$R$ jets are used with $R = 0.4$ jets as inputs to the jet reclustering algorithm [47, 151].

A search for supersymmetry in CMS [152] uses three distinct topologies: fully-merged top quark decays (Monojet), merged $W$ boson decays (Dijet) and resolved decays (Trijet). The efficiency of the three categories is shown in Fig. 33, where the turn-on of the combined efficiency starts at values as low as $p_T \approx 100$ GeV. The resolved trijet category is identified using three anti-$k_T$ jets with a distance parameter of 0.4, where the large combinatorial background is suppressed through a multivariate analysis, which achieves a misidentification rate of approximately 20%. There exist other approaches to cover the transition from low to high Lorentz boosts, using a single algorithm.

In the HTTv2 algorithm, the jet size is reduced until an optimal size $R_{\text{opt}}$ is found, defined by the fractional jet mass contained in the smaller jet. This results in better performance at high $p_T$, while keeping a low misidentification rate at low $p_T$. The distribution in the variable $R_{\text{opt}}$ is shown in Fig. 34 for jets with $p_T > 200$ GeV, where $R_{\text{opt}}$ is observed to vary between 0.4 and 1.5.

An important step towards the commissioning of top taggers within an experiment are measurements of the efficiency and misidentification rate in real collision data. Generally, high-purity samples of top-jets in data are obtained using a tight signal selection (an electron or muon, well-separated from a high-$_{\text{PT}}$ large-$R$ jet, and an additional $b$-tagged jet) to ensure that events contain a fully-merged top quark decay in a single large-$R$ jet. This can never be fully achieved, as no requirements on the substructure of the large-$R$ jet can be imposed without biasing the efficiency measurement. This results in an efficiency measurement that will be based on a sample also containing partially-merged or even non-merged top quark decays. These can be subtracted from the efficiency measurement by using simulated events, as done in a study by the ATLAS Collaboration [40], with the drawback of relying on a specific simulation and the ambiguous definition of a fully-merged top quark decay. By not correcting for non-merged top quark decays, efficiency values are obtained smaller than the ones suggested by ROC curve studies, see for example [52]. The optimal way of obtain-
An efficiency measurement that relies only minimally on simulation and does not rely on the subtraction of processes with similar features to the signal is to fit the fully-and partially-merged categories simultaneously, as done in a recent measurement by CMS [153].

Measurements of the misidentification rate can be carried out by selecting a dijet sample, which is dominated by light-flavor jets. The disadvantage of this approach is the high $p_T$ threshold of unprescaled jet triggers, which results in measurements starting from $p_T > 400\text{ GeV}$ or higher. A solution to this is the tag-and-probe method, in which the tagged jet can be required to fail top-tagging selection criteria, resulting in a sample with negligible contamination of $t\bar{t}$ production, even after requiring the probe jet to be top-tagged [52]. Another approach is to use a non-isolated electron trigger, where the electron fails offline identification criteria. This yields events mainly from light-flavor multijet production, where a jet is misidentified as an electron at the trigger level. While the top-tag misidentification rate can be measured starting from smaller values of $p_T$ with this strategy, a non-negligible amount of $t\bar{t}$ contamination has to be subtracted after requiring a top-tagged jet [40].

As an example, the efficiency and misidentification rate of Shower Deconstruction with the requirement $\log(\chi) > 2.5$, as measured in ATLAS, are shown in Fig. 35. The efficiency of 30% with a misidentification rate of 1% for $350 < p_T < 400\text{ GeV}$ agrees well with the values obtained from Fig. 27. Note that the largest uncertainty of the efficiency measurement stems from the choice of the MC generator used to simulate $t\bar{t}$ production. The uncertainty of the misidentification rate measurement is dominated by the energy scales and resolutions of the subjets and large-$R$ jets.

### 3.4 $H \rightarrow b\bar{b}$ Tagging

The ability to identify jets originating from $b$ quarks produced in the decay of boosted heavy particles, such as top quarks or Higgs bosons, is critical for measurements and searches of new physics signatures, especially as the LHC moves to higher energies.
ATLAS and CMS both use dedicated $b$-tagging algorithms that have been developed and optimized over more than a decade. Both experiments use multivariate techniques making use of the long lifetime of $b$ hadrons. For Run 2 analyses, CMS uses the CSVv2 algorithm [154] and ATLAS uses the MV2c10 algorithm [155].

The boosted $H \rightarrow b\bar{b}$ signal represents a good signature for many different scenarios beyond the Standard Model: resonant $HH$ and $VH$ production, searches for boosted mono-$H$, or vector-like quark searches in the $tH$ and $bH$ final states. Because of the large predicted branching fraction for the $H \rightarrow b\bar{b}$ decay of about 58%, its coupling to $b$ quarks is one of the most interesting to study. For a large fraction of Higgs bosons with $p_T > 300$ GeV, the two $b$ quark jets merge into a single jet for a jet distance parameter of $R = 0.8$ or 1.0, as used in CMS and ATLAS respectively. The approach to reconstruct the Higgs boson in this topology is different from considering the individual smaller-sized jets separately, as the composite nature of such a jet is revealed by analyzing its substructure. Several phenomenological studies have explored $H \rightarrow b\bar{b}$ tagging algorithms using jet substructure, though ultimately the optimal performance comes from using a combination of substructure information and the track and vertex information related to the $b$ hadron lifetime.

The approaches to identify boosted $H \rightarrow b\bar{b}$ candidates that have been explored (and used) at CMS and ATLAS include:

(a) Subjet $b$-tagging [136, 156, 157, 158, 159, 160], where ‘standard’ $b$-tagging is applied to each of the subjets (the standard for CMS is the CSVv2 algorithm [154], and for ATLAS is MV2c20 [155]). Tagging $b$-jets in dense environments is of particular importance here, and was studied by ATLAS in Ref. [161]. In CMS subjets with $R = 0.4$ are clustered with the $k_T$ algorithm using the constituents of the large-$R$ jet, while for ATLAS track jets with a radius of 0.2 are matched to the large-$R$ jet using the ghost-association technique. At high $p_T$ the subjets start to overlap causing the standard $b$-tagging techniques to break down due to double-counting of tracks and secondary vertices when computing the subjet $b$-tag discriminants.

(b) Double-$b$ tagging [160, 162, 163]. In ATLAS, the term double-$b$ tagging means that the two leading $p_T$ track jets must pass the same $b$-tagging requirement. In CMS, the double-$b$ tagger is a novel approach detailed in [162, 163], which uses the pruned anti-$k_T$, $R = 0.8$

![Figure 36: The rejection of inclusive multijets versus Higgs-jet efficiency using all large-$R$ jets with $p_T > 250$ GeV for single, double, asymmetric, and leading subjet $b$-tagging requirements. Taken from Ref. [160].](image)

jet mass (a window of $50 < M < 200$ GeV is required to reduce the multijet background) and the $N$-subjettiness axes.

The Higgs-jet efficiency versus the inclusive multijet rejection are shown in Fig. 36 for ATLAS subjet $b$-tagging, where the performance curves are shown for double-$b$ tagging, leading subjet $b$-tagging, and asymmetric $b$-tagging (among the two leading $p_T$ track jets, the track jet with the largest $b$-tagging weight must pass the fixed 70% $b$-tagging working point threshold, while the $b$-tagging requirement of the other jet is varied) requirements. None of the curves reach a Higgs-jet efficiency of 100% due to the imperfect efficiency to reconstruct the track jets needed for $b$-tagging and, in the case of asymmetric $b$-tagging, also due to the 70% $b$-tagging working point requirement on one of the track jets.

The CMS double-$b$ tagging approach has been applied to identify $H \rightarrow b\bar{b}$ candidates [162, 163]. The algorithm attempts to fully exploit the strong correlations between the $b$ hadron flight directions and the energy flows of the two subjets, while adapting the variables used in the CSVv2 algorithm. The flexibility of the double-$b$ tagger is ensured by avoiding a strong performance dependency on the jet $p_T$ and mass.

With the double-$b$ tagger, at the same signal efficiency, the misidentification rate is uniformly lower by about a factor of two compared to the subjet $b$-tagging approach. Given the different kinematic properties expected for a
$b\bar{b}$ pair originating from the decay of a massive resonance compared to gluon splitting, the misidentification rate for the gluon splitting background reduces from 60% to 50% at 80% signal efficiency and from 20% to 10% at 35% signal efficiency. At high $p_T$, even larger performance improvements are observed, which is an important gain for searches for heavy resonances, where very high $p_T$ jets are expected. In Fig. 37 the signal efficiencies and misidentification rates for the double-$b$ tagger are shown as a function of jet $p_T$ for three operating points: loose, medium and tight, which correspond to 80%, 70% and 35% signal efficiency, respectively, for a jet $p_T$ of about 1000 GeV. The misidentification rate is mostly flat across the $p_T$ range considered while the signal efficiency decreases with increasing $p_T$, as expected from the degradation of the tracking performance inside high $p_T$ jets.

![Figure 37: The misidentification efficiency for inclusive multijets versus Higgs-jet efficiency using jets with 300 < $p_T$ < 500 GeV and pruned jet mass 70 < $m$ < 200 GeV for three different $b$-tagging requirements. Taken from Ref. [162].](image)

Due to the small cross section of producing events with boosted $H \rightarrow b\bar{b}$ or $Z \rightarrow b\bar{b}$ jets, the efficiency of the ATLAS and CMS Higgs identification algorithms is measured using QCD multijet events enriched in jets from gluon splitting, $g \rightarrow b\bar{b}$ with a topology similar to that of boosted $H \rightarrow b\bar{b}$ jets.

CMS selects topologies as similar as possible to a signal jet by requiring the jet $p_T > 300$ GeV and pruned mass > 50 GeV [162, 163]. Each jet has to contain at least two muons, each with $p_T > 7$ GeV and $|\eta| < 2.4$. Each pruned subjet is required to have at least one muon among its constituents and within $\Delta R < 0.4$ from the subjet axis ("double-muon tagged"). An alternative selection that requires at least one muon is also examined as a cross-check for the measurement ("single-muon tagged"). While this single-muon selection allows for a larger dataset in which to perform the tagger efficiency measurement, the gluon splitting topology in this inclusive phase space is less signal-like relative to the double-muon selection. Thus, to maximize the similarity between the $g \rightarrow b\bar{b}$ and the $H \rightarrow b\bar{b}$ topologies, the measurement is performed requiring double-muon tagged jets. It is worth noting however that the jet mass depends on the number of muons and a large fraction of the signal will not contain two muons.

ATLAS performed a similar measurement selecting events with at least one anti-$k_t$, $R = 1.0$ jets with $p_T > 250$ GeV that has two ghost-associated $R = 0.2$ track jets [160]. As opposed to the measurement from CMS, only one of the subjets is required to have a muon associated to it. Kinematic and substructure variables are compared in data and MC after correcting for flavor composition differences of the large-$R$ jet observed between data and MC simulation and are found to be in good agreement.

One of the major backgrounds for analyses selecting boosted $H$ or $Z$ bosons decaying to $b\bar{b}$ is $t\bar{t}$ production. The misidentification rate for boosted top quark jets faking $H$ jets was measured in data by CMS [162, 163] in enriched data samples of lepton+jets $t\bar{t}$ events.

As previously discussed, for high $p_T$ of the Higgs boson, the two subjets from $b$ quarks start overlapping and the performance of identifying the subjets as fixed-radius track jets decreases significantly. To improve the performance of the ATLAS standard $H \rightarrow b\bar{b}$ identification algorithm for searches that are sensitive to high $p_T$ Higgs bosons, the ATLAS Collaboration studied alternative methods: the use of variable-radius track jets, exclusive $k_T$ subjets, calorimeter subjets reconstructed in the center-of-mass frame of the Higgs jet candidate [164] and the combination of three jet shape and jet substructure variables into a multivariate discriminator [165]. For highly boosted Higgs bosons, these reconstruction techniques significantly outperform the usage of fixed-radius track jets.
<table>
<thead>
<tr>
<th>Taggers in use by ATLAS and CMS for Run 2 of the LHC.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Q/G Taggers</strong></td>
</tr>
<tr>
<td>$LLH(n_{\text{track}}, W)$</td>
</tr>
<tr>
<td>CNN</td>
</tr>
<tr>
<td>$n_{\text{track}}$</td>
</tr>
<tr>
<td>$LD(n_{\text{cons}}, \sigma_2, p_T D)$</td>
</tr>
<tr>
<td>$BDT(n_{\text{ch}}, \sigma_1, \sigma_2, p_T D, W x)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Vector Boson Taggers</strong></th>
<th><strong>Description</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>R2D2</td>
<td>Trimmed ($R = 1.0$, $f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.2$) mass and $D_2^\beta$</td>
</tr>
<tr>
<td>CMS V-tagger</td>
<td>PUPPI soft drop ($z_{\text{cut}} = 0.1$ and $\beta = 0$) mass and ungroomed $\tau_{21}$</td>
</tr>
<tr>
<td>$N_2^\beta$</td>
<td>PUPPI soft drop ($z_{\text{cut}} = 0.1$ and $\beta = 0$) mass and soft drop $N_2^\beta$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Top Taggers</strong></th>
<th><strong>Description</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>R2 $\tau_{32}$</td>
<td>Trimmed ($R = 1.0$, $f_{\text{cut}} = 0.05$, $R_{\text{sub}} = 0.2$) mass and $\tau_{32}$</td>
</tr>
<tr>
<td>SD</td>
<td>Shower Deconstruction</td>
</tr>
<tr>
<td>HTT</td>
<td>HEPTopTagger</td>
</tr>
<tr>
<td>HTTv2</td>
<td>HEPTopTagger with shrinking cone</td>
</tr>
<tr>
<td>CMSTT</td>
<td>mMDT mass and $\tau_{32}$ and subjet b-tagging</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Higgs Taggers</strong></th>
<th><strong>Description</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>CSVv2</td>
<td>Subjet $b$-tagging with the CSVv2 algorithm and $R_{\text{sub}} = 0.4$ jets</td>
</tr>
<tr>
<td>MV2c20</td>
<td>Subjet $b$-tagging with the MV2c20 algorithm and $R_{\text{sub}} = 0.2$ track jets</td>
</tr>
<tr>
<td>Double-$b$ tagging</td>
<td>With the CSVv2 algorithm and pruned $R = 0.8$ jets, using $N$-subjettiness axes</td>
</tr>
</tbody>
</table>
4 Measurements

The measurement of jet properties is crucial to constrain the Standard Model in new energy regimes and constitutes and important test of perturbative calculations of jet structure over a wide region of phase space. Moreover jet cross section measurements provide constraints on the parton distribution functions and the strong coupling constant, \( \alpha_s \). The precise knowledge of jet properties also improves the precision of other measurements due to an improved modeling of important background processes and is necessary to probe for physics beyond the Standard Model. Jet substructure variable measurements are challenging as they require a precise measurement of the radiation pattern within the jet and thus the jet constituents.

4.1 Measurements with jet substructure

4.1.1 \( \bar{t}t \) Cross Section

The selection cuts applied in the traditional \( \bar{t}t \) cross section measurements [166, 167, 168, 169, 170, 171, 172] are chosen to maximize the acceptance and minimize the associated uncertainties on the fiducial cross section measurement. The resulting fiducial region is such that events with top \( p_T \) below 100 GeV and above 600 GeV are under-represented, with the latter being due to collimated decays resulting from the large Lorentz boost. This is evident from Fig. 38, where a drop in selection efficiency below 100 GeV and above 600 GeV is apparent. This results in a small number of events being selected with high top quark \( p_T \), as seen in the ATLAS Run 1 (7 TeV) measurement shown in Fig. 39; unfortunately, a very interesting region in terms of new physics is the least well-measured. Despite often having similar signal efficiencies to resolved reconstruction, boosted top tagging techniques allow for more precise measurements at high \( p_T \) due to their higher background rejection.

The ATLAS Collaboration performed a measurement of the boosted \( \bar{t}t \) differential cross section as a function of the \( p_T \) in the lepton+jets channel [173]. A least one anti-\( k_t \) jet, trimmed with \( R_{\text{sub}} = 0.3 \) and \( f_{\text{cut}} = 0.05 \) is required with \( |\eta| < 2 \) and \( p_T > 300 \) GeV. To select events with boosted top quarks, the large-\( R \) jet is required to have a mass larger than 100 GeV and \( \sqrt{d_{T2}} > 40 \) GeV (Tagger III, see Sec. 3.3). The reconstructed \( p_T \) distribution of the anti-\( k_t \), \( R = 1.0 \) trimmed jet is unfolded to the parton and particle-level. The measured particle-level differential cross section is compared in Fig. 40 to the predictions of several MC generators normalized to the NNLO+NNLL inclusive cross section. Overall good agreement is observed, but a harder \( p_T \) spectrum is predicted by the simulation than observed in data with larger discrepancies at high \( p_T \). The differential cross section measurement is also compared to predictions from Powheg+Pythia in Fig. 40 using either the HERAPDF or CT10 PDF set and two different values for \( h_{\text{damp}} \), \( h_{\text{damp}} = m_{\text{top}} \) and \( h_{\text{damp}} = \infty \). The best data/MC agreement is observed when using the HERAPDF set and \( h_{\text{damp}} = m_{\text{top}} \). For each of the settings, the trend of a harder \( p_T \) spectrum in simulation compared to data persists.

A similar measurement by the CMS Collaboration based on 8 TeV data [174] uses the CMSTT algorithm to reconstruct boosted top quarks. The unfolded results are in agreement with the ATLAS measurement and show a similar trend between data and simulation, as shown in Fig. 41.

These measurements extend up to a top quark \( p_T \) of 1.2 TeV, allowing for higher precision thanks to the usage of jet substructure techniques. The largest uncertainties at the highest values of \( p_T \) in ATLAS and CMS come from the large-\( R \) jet energy scale and the extrapolation of the \( b \)-jet calibration to high \( p_T \).
The parton-level differential cross section in top quark \( p_T \) has also been measured in the all-hadronic final state by the CMS Collaboration using 8 TeV data [175]. This measurement relies on pruned jets with an \( N \)-subjettiness and subjet-\( b \) tagging requirement to suppress the huge amount of background from QCD dijet production. The cross section is determined from a maximum likelihood fit to the jet mass distributions for signal-enriched and signal depleted regions. This allows for a simultaneous extraction of the \( t\bar{t} \) cross section while constraining the QCD background. The measurement is in agreement with the results from the CMS Collaboration using 8 TeV data [175]. This measurement relies on pruned jets with an \( N \)-subjettiness and subjet-\( b \) tagging requirement to suppress the huge amount of background from QCD dijet production. The measurement is in agreement with the results from the CMS Collaboration using 8 TeV data [175].

A first measurement based on 3.2 fb\(^{-1}\) of 13 TeV data in the lepton+jets channel has been performed by ATLAS [172]. The measurement extends to \( p_T \) of 1.5 TeV and a similar trend as at 8 TeV is observed between the data and the simulation at high \( p_T \). A newer measurement of the \( t\bar{t} \) differential cross section in the all-hadronic channel is performed by the ATLAS Collaboration with 36.1 fb\(^{-1}\) of 13 TeV data [176]. The measurement uses trimmed anti-\( k_t \) \( R = 1.0 \) jets with \( R_{\text{sub}} = 0.2 \) and \( f_{\text{cut}} = 0.05 \). To obtain a flat signal efficiency of 50%
and a quark/gluon rejection of approximately 17 (10) for \( p_T = 500 \) (1000) GeV. \( p_T \) dependent criteria are applied on the jet mass and \( \tau_23 \). Furthermore the two top-tagged large-\( R \) jets are required to have a b-tagged small-\( R \) jet within \( \Delta R < 1.0 \). The event selection results in a signal-to-background ratio of 3-to-1. The measured fiducial phase-space cross section is \( \sigma = 292 \pm 7 \) (stat) \( \pm 76 \) (sys) fb compared to the Powheg+Pythia8 prediction of 384\( \pm 36 \) fb at NNLO+NNLL. The measured normalized differential cross section as a function of the top jet mass within \( \Delta R < 1.0 \). The measurement is dominated by the systematic uncertainties on the jet energy, mass and substructure scale of the large-\( R \) jets, alternative parton shower model and the uncertainties on the \( b \)-jet identification.

4.1.2 W/Z/H Cross Sections

The cross section of boosted \( W \) and \( Z \) bosons was measured by ATLAS in 4.6 fb\(^{-1} \) of 7 TeV pp collisions [43]. The hadronically decaying \( W \) and \( Z \) bosons are reconstructed as one single ungroomed anti-\( k_t \) \( R = 0.6 \) jet with \( p_T > 320 \) GeV, \( |\eta| < 1.9 \) and masses ranging between 50 and 140 GeV. The \( W \) and \( Z \) signal is enhanced over the dominating QCD background by constructing a likelihood discriminant from three substructure variables; thrust minor [177, 178], sphericity [179] and aplanarity [179], resulting in a signal efficiency of 56\% and a background rejection of 89\%. The jet mass distribution after subtracting the expected background from \( t\bar{t} \) events is shown in Fig. 43. A binned maximum likelihood fit to the jet mass distribution is used to extract the \( W/Z \) jet signal yield and to calculate the inclusive cross section. Only the combined \( W + Z \) cross section measurement is performed in this analysis due to the limited jet mass resolution. The combined \( W + Z \) cross section is measured to be \( \sigma_{W + Z} = 8.5 \pm 0.8 \) (stat.) \( \pm 1.5 \) (syst.) pb and is in agreement with the Standard Model prediction of \( \sigma_{W + Z} = 5.1 \pm 0.5 \) pb within 2 standard deviations. The dominating systematic uncertainties are the jet mass resolution and the choice of the QCD background pdf. The signal significance was furthermore studied when using groomed jets instead of ungroomed jets. Without an optimization of the analysis for groomed jets, similar significances were observed for groomed and ungroomed jets as expected due to the low number of pileup vertices in the 7 TeV dataset.

![Figure 42](image1.png)

Figure 42: The normalized differential cross section as a function of the \( t\bar{t} \) \( p_T \) as measured by ATLAS in the all-hadronic channel at 13 TeV. Taken from Ref. [176].

![Figure 43](image2.png)

Figure 43: Binned maximum likelihood fit to the jet mass distribution in data for selected \( W/Z \) events reconstructed as one single ungroomed anti-\( k_t \) \( R = 0.6 \) jet. Taken from Ref. [43].
$Z \rightarrow b\bar{b}$ cross section using two $b$-tagged anti-$k_t$ $R = 0.4$ jets in 19.5 fb$^{-1}$ of 8 TeV $pp$ collisions instead of one single large-$R$ jet [180]. The measured fiducial cross section was determined to be $\sigma_{Z \rightarrow b\bar{b}} = 2.02 \pm 0.33$ pb which is in excellent agreement with the next-to-leading-order theoretical predictions.

As discussed in Section 3.4 the SM Higgs boson decays with approximately 58$\%$ into $b\bar{b}$. However the $H \rightarrow b\bar{b}$ decay in the resolved channel can only be studied in the associated production with either a vector boson ($W/Z$) [181, 182] or top quark or via the vector-boson-fusion production mechanism due to the overwhelming multijet background. To search directly for $H \rightarrow b\bar{b}$ production mechanism due to the overwhelming multijet background. The CMS Collaboration performed a search for the SM Higgs boson using a single-jet topology with 35.9 fb$^{-1}$ of 13 TeV $pp$ collisions [183]. The analysis uses anti-$k_t$ $R = 0.8$ jets that are modified with the soft drop algorithm ($\beta = 0$, $z_{cut} = 0.1$) to mitigate the effects from pileup and to remove wide-angle radiation. At least one large-$R$ jet with $p_T > 450$ GeV and $|\eta| < 2.5$ is required. To distinguish the two prong structure of a jet containing the full $H \rightarrow b\bar{b}$ decay from quark- or gluon-initiated jets, the $N^1_2$ variable, calculated from the generalized energy correlation functions, is exploited. To ensure a flat QCD background rejection of 26$\%$ over the considered mass and $p_T$ range, a decorrelation procedure [129] is applied to $N^1_2$. The multijet background is further suppressed by utilizing the double-$b$ tagger which exploits not only the presence of two $b$ quarks but also uses the fact that the flight direction of the $b$ hadron is strongly correlated to the energy flow of the two subjets. The $W/Z +$jets background are estimated from MC simulation and the shape of the multijet background is determined in a validation region in data with lower values of the double-$b$ tagger discriminator. The soft drop mass distribution of the leading jet is shown in Fig. 44 with a clear resonant structure at the mass of the $W$ and $Z$ boson. The SM background processes and the potential signal from SM $H \rightarrow b\bar{b}$ production are estimated simultaneously. The measured cross section for the $Z +$jets process is $0.85 \pm 0.16$ (stat.) $^{+0.0}_{-0.4}$ (syst.) pb which is in agreement with the SM prediction of $1.09 \pm 0.11$ pb. This is the first observation of $Z \rightarrow b\bar{b}$ in the single jet topology. The observed (expected) significance for the $H \rightarrow b\bar{b}$ process is 1.5(0.7)$\sigma$.

Figure 44: Soft drop jet mass $m_{SD}$ of anti-$k_t$, $R = 0.8$ jets in data and for the dominating background processes; multijet production and $W/Z +$jets events. Jets are required to pass criteria on $N^1_2$ and to be identified as double-$b$ jets by the double-$b$ tagger introduced in Section 3.4. Taken from Ref. [183].

4.2 Measurements of jet substructure

The ATLAS and CMS Collaborations have performed a series of precision measurements of hadronic jet substructure in $pp$ collisions, correcting for acceptance and resolution such as jet and event shapes [184, 185, 186, 187], charged particle multiplicities [187, 188, 189], the jet fragmentation functions [190], color flow [191] and $k_T$ splitting scales, $N$-subjettiness ratios as well as further substructure variables such as Planar Flow and angularity [192, 193, 193]. This section will focus on measurements of the jet mass and the jet charge [194, 195].

4.2.1 Jet mass

The particle-level $t\bar{t}$ cross section was measured differentially as a function of the top-jet mass by the CMS Collaboration in lepton+jets events using 8 TeV data [67]. Large-$R$ jets are reconstructed with the C/A algorithm using a distance parameter of 1.2 and $|\eta| < 2.5$. The larger value of $R$ in this measurement compared to the default $R = 0.8$ applied for top tagging applications in CMS is due to an optimization of statistical precision due to the
number fully-merged top quarks and the width of the jet mass distribution, which grows with increasing $R$. The leading jet $p_T$ is required to be above 400 GeV to ensure the hadronic top quark decay to be fully captured within the large-$R$ jet. A requirement of $p_T > 150$ GeV is imposed on the subleading jet to select the $b$ quark from the leptonically decaying top quark. The particle-level differential $t\bar{t}$ cross section as a function of the leading jet mass is shown in Fig. 45. The shown MC predictions exceed the measurement, consistent with the cross section measurements from the ATLAS and CMS Collaborations at high $p_T$. The total cross section in the fiducial region is $\sigma_{t\bar{t}} = 101 \pm 11$ (stat) $\pm 13$ (sys) $\pm 9$ (model) for $140 < M_{\text{jet}} < 350$ GeV. The experimental systematic uncertainties are dominated by the uncertainties on the jet mass and energy scale, but are smaller than the uncertainties due to the signal modeling, coming from the choice of the top quark mass, the parton showering and the choice of the factorization and renormalization scales.

For the first time, the top quark mass was determined in boosted $t\bar{t}$ events by the CMS Collaboration. The top quark mass was extracted from the normalized differential cross section measurement, shown in Fig. 46, since only the shape of the top-jet mass distribution can be reliably calculated. The top quark is measured to be $m_t = 170.8 \pm 6.0$ (stat) $\pm 2.8$ (sys) $\pm 4.6$ (model) $\pm 4.0$ (theo) GeV which is in agreement with top quark mass measurement in resolved $t\bar{t}$ events (see e.g. Refs. [196, 197]). The usage of boosted techniques is of particular theoretical interest to perform precision measurements of the top quark mass. The most precise top quark mass measurements [196, 198, 199] are currently based on template and matrix element fits to the reconstructed top quark decay products and thus measure the Monte Carlo top quark mass $m_{t}^{MC}$ which is often assumed to be the same as the pole mass but depends in fact on the parton shower and hadronization model. The systematic uncertainties of these measurements are dominated by MC modeling uncertainties. The reconstruction of the top quark decay products as one single jet allows for the extraction of the top quark mass without ambiguities that arise from the translation of the MC mass into a short-distance mass scheme [200]. Furthermore it has been shown that the usage of grooming techniques can significantly reduce the MC modeling systematic uncertainties [201].

The first measurement of the normalized dijet differential cross section as a function of the jet mass was performed by the ATLAS Collaboration with a dataset corresponding to 35 pb$^{-1}$ of 7 TeV $pp$ collisions [192]. Both the
cross section for groomed and ungroomed C/A $R = 1.2$ jets was measured separately to gain sensitivity to both the hard and soft jet physics and to gain a deeper understanding of the various effects involved in QCD radiation. For the ungroomed jet mass, large discrepancies were observed in the tails of the mass distribution between the MC predictions from Pythia and Herwig++ and the data whereas the core of the mass distribution agreed within approximately 20% over the considered $p_T$ range. The largest discrepancies occur at low jet masses which is sensitive to the underlying event description, hadronization model and pileup effects. The normalized cross section after applying the split filtering algorithm with $\mu_{\text{frac}} = 0.67$ and $y_{\text{filt}} = 0.09$ is shown in Fig. 47. After removing soft radiation from the jet which is difficult to model, the MC prediction is in excellent agreement with the data within statistical precision. The CMS Collaboration performed a similar measurement with anti-$k_t$ $R = 0.7$ jets using various grooming techniques in selected dijet events using 5 fb$^{-1}$ of $\sqrt{s} = 7$ TeV data and found as well that the agreement between data and the MC prediction improves significantly after grooming techniques are applied [53]. Furthermore a measurement of the cross section was performed in $V + \text{jet}$ final states which overall show a slightly better data/MC agreement than that observed in dijet events suggesting that the simulation of quark jets is better than for gluon jets.

The CMS (ATLAS) Collaboration measured the double-differential jet cross section in balanced dijet events at $\sqrt{s} = 13$ TeV for groomed anti-$k_t$ $R = 0.8$ jets with the soft drop algorithm with $z_{\text{cut}} = 0.1$ and $\beta = 0$ ($\beta = 0, 1, 2$) [42, 54]. The soft drop algorithm was chosen as it allows to compare the unfolded measurement directly to theoretical calculations which significantly exceed the precision of parton shower MC simulations. The jet energy of the ungroomed jets used in the ATLAS measurement are corrected for pileup effects and calibrated to the generator-level while no explicit mass calibration is applied to the groomed jets as the unfolding procedure accounts for differences between the reconstructed and generator-level mass. The CMS Collaboration applied calibration factors derived from simulation and using in-situ techniques to correct the jet energy and mass scale. Furthermore the jet energy and mass are smeared in MC simulation to match the resolution measurements in data. Various sources of systematic uncertainties, categorized as experimental and theoretical uncertainties, that impact the jet mass measurement are taken into account.

While CMS evaluated the effect of the jet energy and mass scale uncertainties on the measurement by varying the energy and mass by their respective uncertainties, ATLAS evaluated the experimental uncertainties based on the accuracy of the modelling of the topological cluster energies and positions as well their reconstruction efficiency. Theoretical uncertainties on the physics model are taken into account by comparing the response matrix for various MC generators.

The comparison of the normalized cross section with two analytical calculations as measured by CMS is shown in Fig. 48. ATLAS measured instead the unfolded $\log_{10} \rho^2$ distribution, shown in Fig. 49, where $\rho$ is the ratio of the soft drop jet mass to the ungroomed jet $p_T$. Both measurements are compared to calculations at next-to-leading order with next-to-leading-logarithm and leading order with next-to-next-to-leading-logarithm accuracy. Good agreement between the data and the predictions is observed in resummation regime $-3.7 < \log_{10} \rho^2 < -1.7$. For higher jet masses, where non-perturbative effects play an important role, the NLO+NNLL calculation, which includes non-perturbative effects, provides a better description than the LO+NNLL calculation.


2.3 fb\(^{-1}\) (13 TeV)

Figure 48: Normalized differential cross section as a function of the mass for jets groomed with the soft drop algorithm in data and for two theoretical calculations. Taken from Ref. [54].

Figure 49: Comparison of the unfolded \(\log_{10} \rho^2\) distribution for \(z_{\text{cut}} = 0.1, \beta = 1\) in data to various Monte Carlo particle-level predictions and theory predictions, normalized to the integrated cross section measured in the resummation regime \(-3.7 < \log_{10} \rho^2 < -1.7\). Taken from Ref. [42].

4.2.2 Jet Charge

This section summarizes the measurements of the jet charge [194, 195], defined as the energy weighted sum of the electric charges of the jet constituents

\[
Q_{\kappa} = \sum_{i \in J} \left( \frac{p_{T,i}}{p_{T,J}} \right)^{\kappa} q_i ,
\]

where \(q_i\) is the electric charge of particle \(i\) and the free parameter \(\kappa\) that controls the sensitivity to soft particles within the jet. The ATLAS (CMS) Collaboration measured the jet charge for different values of \(\kappa\) using anti-\(k_t\) jets with a radius parameter of \(R = 0.4\) (\(R = 0.5\)) in a sample of dijet events. The ATLAS Collaboration distinguishes between the two leading jets using the pseudorapidity instead of the \(p_T\) to avoid cases where the leading particle-level jet is reconstructed as the sub-leading detector-level jet due to the jet energy resolution. The average jet charge at detector- and particle-level for the more forward of the leading jets and for \(\kappa = 0.5\) is shown in Fig. 50. Due to the increasing fraction of scattering valence up quark jets (up quark charge \(> 0\)), the average jet charge increases with \(p_T\). The difference of the average jet charge distribution at detector-level and particle-level in Fig. 50 shows that the unfolding corrections are large and growing at high \(p_T\), due to the loss of charged-particle tracks inside jets as a result of track merging. The average jet charge as predicted by Pythia 8 [202] using the Perugia tunes [203] is smaller than that observed in data due to a well-known over-estimation of the multiplicity inside jets.

The dominating systematic uncertainties are the track \(p_T\) resolution and the choice of MC generator used to construct the response matrix (Pythia 6 versus Herwig++) for the CMS Collaboration whereas the uncertainties on the unfolding procedure, the jet energy resolution at low \(p_T\) and uncertainties on the tracking at high \(p_T\) dominate the measurement of the ATLAS Collaboration. The unfolded jet charge distribution (\(\kappa = 0.6\)) of the leading jet in data is compared to the prediction from Powheg+Pythia8 (PH+P8) and Powheg+Herwig++ (PH+HPP) in Fig. 51. The different hadronization and fragmentation model used by Pythia8 and Herwig++ have the largest impact on the jet charge distribution. Variations of the jet charge can also be observed for different PDF sets however the effect of the relative flavor fraction in the dijet samples is significantly smaller than the choice of the showering and fragmentation model. It was further found that the predicted jet charge distribution has a significant dependency on the chosen value of \(\alpha_s\) that describes final state radiation whereas it is insensitive to NLO QCD effect in the matrix element calculation, color-reconnection and multiple parton interactions. These findings are consistent between the ATLAS and CMS Collaboration.

In addition to studying the sensitivity to various non-perturbative aspects of hadronization and parton distribution functions, the jet charge measurement by ATLAS...
includes the first direct comparison of a jet substructure quantity with a perturbative calculation at a hadron collider. As it is not collinear safe, the average jet charge is not calculable. However, the $p_T$ dependence for a particular jet type has been calculated [204, 205]. A new technique was introduced in Ref. [194] to separately extract the average up and down quark jet charge. For a fixed $p_T$, the more forward of the two dijets has a higher energy and is therefore more likely to be the scattering parton with a higher momentum fraction of the proton. In turn, the higher momentum fraction parton is most likely to be a valence quark. Therefore, the fraction of up quark jets is higher for the more forward dijet than the more central dijet. Assuming further that the jet charge is entirely determined by the jet $p_T$ and parton origin, one can then solve a system of equations to extract the average up and down quark jet charge in each bin of jet $p_T$:

\[
\langle Q_f^j \rangle = f_u^f \langle Q_u^j \rangle + f_d^f \langle Q_d^j \rangle \\
\langle Q_c^j \rangle = f_u^c \langle Q_u^j \rangle + f_d^c \langle Q_d^j \rangle,
\]

where $f = \text{forward}$, $c = \text{central}$, $u = \text{up}$ and $d = \text{down}$. As expected (though not an input), the average up quark charge is positive and the average down quark charge is negative; furthermore, the latter is roughly half the former in absolute value. The $p_T$ dependence of $\langle Q_{u,d}^j \rangle$ are fit with a logarithmic scale violating term $c$: $\langle Q \rangle_i = \langle Q \rangle_0(1 + c \kappa \ln(p_T_i/p_T_0))$, where $i$ represents the $p_T$ bin. Figure 52 shows the measured and predicted values of $c_\kappa$. The uncertainties are large, but there is an indication that $c < 0$ and $\partial c/\partial \kappa < 0$, as predicted.

5 Searches for New Physics

Jet substructure methods have been successfully applied in a large variety of searches for physics beyond the SM. The respective exclusion limits are substantially improved through the application of these methods. In some cases the decay signature of heavy BSM particles would not be accessible without the application of jet substructure methods.

As the number of such BSM searches is very large, only a small subset of the published results can be discussed here. The following sections give an overview of a selection of searches for $t\bar{t}$ resonances [150, 206, 207, 208, 209, 210, 211], diboson resonances [66, 123, 124, 125, 126, 127, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227], vector-like quarks [228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243] and leptophobic $Z'$ [122, 244]. Further searches using jet substructure techniques can be found in Refs. [47, 106, 145, 151, 245, 246, 247, 248, 249, 250].

Figure 50: The detector- and particle-level average jet charge as a function of jet $p_T$. Reproduced from Ref. [194].

Figure 51: Unfolded jet charge distribution for $\kappa = 0.6$ in data and MC prediction. Taken from Ref. [195].
Figure 52: The measured and predicted value of the average jet charge scale violation parameter $c_\kappa$. Reproduced from Ref. [194].


5.1 $tt' $ Resonances

Several models of new physics predict resonances coupling strongly to the top quark to play a role in the cancellation of large corrections to the Higgs mass. These models include extensions of the SM Higgs doublet, where the simplest realizations are two-Higgs-doublet models [271] with heavy, neutral Higgs bosons, which can have large branching fractions to top quarks. Alternative mechanisms of electroweak symmetry breaking often include new heavy particles with large couplings to top quarks. Examples for such models are the topcolor model containing a $Z'$ boson [272] as well as composite Higgs models [273, 274, 275, 276, 277, 278, 279, 280]. The Randall-Sundrum Kaluza-Klein model predicts a massive gluon decaying to top quarks [281, 282, 283, 284]. While small resonance masses are mostly excluded, the searches are generally focusing at high masses with $m_{t\bar{t}} > 1$ TeV such that the top quarks receive high Lorentz boosts and must be reconstructed with jet substructure methods. In the following sections the analysis strategies and results from CMS and ATLAS using proton-proton collisions at 8 TeV are discussed, followed by a summary of the first results with 13 TeV data.

**Lepton+jets Channel** Due to the high boost of the top quarks from a heavy resonance decay, the lepton from the $W$ boson decay may overlap with the associated $b$ quark jet. Therefore, the lepton-isolation criteria, which are used to mitigate the contamination with QCD multijet background, are relaxed. The CMS and ATLAS Collaborations follow different strategies for this purpose. In the CMS analysis [207], the lepton must have a large angular separation from the associated $b$-jet candidate of $\Delta R(\text{lepton, jet}) > 0.5$ or it must have a transverse momentum relative to the jet axis $p_T^{\text{rel}} > 25$ GeV. This selection requirement removes background contributions from semi-leptonic $B$ hadron decays. In the ATLAS analysis [150], the lepton isolation is achieved by a so-called mini-isolation requirement which applies a variable isolation cone that changes as a function of the lepton energy [285]. Interestingly, studies performed in CMS for 13 TeV show that the CMS implementation of mini-isolation is not as powerful as their own method for $t\bar{t}$ resonance searches.

To reconstruct the boosted hadronic top decay, the presence of a single high-momentum, large-$R$, top-tagged jet is required. In CMS (ATLAS) the large-$R$ jet is reconstructed with the Cambridge/Aachen (anti-$k_t$) algorithm with a size parameter of $R = 0.8$ (1.0). The selection requirement on the transverse momentum is $p_T > 400(300)$ GeV. ATLAS applies trimming to the large-$R$ jets with the parameters $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ and the jets are required to have a mass $m_{\text{jet}} > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV. The strategy followed by CMS is to apply the “CMS top-tagging” algorithm (as defined in Section 3.3), where the mass of the jet has to satisfy $140 < m_{\text{jet}} < 250$ GeV. In addition, the $N$-subjetness ratio $\tau_{32}$ must be smaller than 0.7.

The variable of interest is the invariant mass $m_{t\bar{t}}$ of the $t\bar{t}$ system. It is reconstructed from the top-tagged large-$R$ jet, a $b$-tagged small-$R$ jet as well as the lepton and the missing energy. Once the top pair system is reconstructed, events are further divided into categories based on the lepton flavor and the number of $b$-tagged and top-tagged jets. This gives several analysis categories with different background compositions: the top-tagged and $b$-tagged events are dominated by the SM $t\bar{t}$ background, while events without top-tags and $b$-tags are mostly composed of $W$+jets events. No excess of data events is seen over expected SM backgrounds, and exclusion limits are
set on the $Z'$ and RS-KK gluon physics models, which are discussed in more detail below.

**Hadronic channel** An analogous search is performed by CMS assuming hadronic decays for both top quarks [207]. The same CMS top-tagging algorithm as in the lepton+jets channel is applied. In addition, the region with jets of lower momenta with $200 < p_T < 400$ GeV is covered with a dedicated algorithm using a larger jet size parameter of $R = 1.5$ (CA15 jets). The larger jet size allows to cover topologies with larger angles between the decay products of the top quark in case of intermediate or smaller Lorentz boosts. These low-$p_T$ jets are required to be identified as top-jets by the HEPTopTagger algorithm (as described in Section 3.3). This approach allows to improve the sensitivity for smaller masses of the hypothetical $t\bar{t}$ resonance.

Additionally, events are divided based on the rapidity separation $|\Delta y|$ of the two jets. Jets from a high-mass particle will in general have a lower $|\Delta y|$ compared to QCD dijet events, where the high mass usually comes from a larger rapidity separation. Two regions are used to divide events, defined by $|\Delta y|$ greater or less than 1.0. This results in six event categories for the hadronic analysis.

Even with the requirement of two top-tagged jets, the selected sample is dominated by QCD dijet events. This background is estimated using a data-driven technique, where an anti-tag and probe method is used. The $\tau_{32}$ requirement is reversed on one jet to select a sample dominated by QCD events. The opposite jet is then used to measure the misidentification rate for the top-tagging requirements. The measured misidentification rate ranges from 5 to 10%, depending on the jet momentum, $\tau_{32}$ and the $b$-tagging requirements applied. This differential rate is used in a sample of single top-tagged events to predict the double top-tagged event contribution from QCD processes in each individual event category. Closure tests performed in data and simulation are performed to validate the background estimation for each of the signal regions.

No significant excess above the predicted background is observed in the measured $t\bar{t}$ invariant mass spectrum. The exclusion limits obtained on the $Z'$ and RS-KK gluon models are discussed below.

**Results** The analysis of the ATLAS Collaboration [150] finds upper limits on the cross section times branching ratio of a narrow $Z$ boson decaying to top quark pairs in the range from 4.2 pb to 0.03 pb for resonance masses ranging from 0.4 TeV to 3.0 TeV. The limits for larger widths (15%) are slightly worse.

The CMS Collaboration performs a combination of the lepton+jets, hadronic and di-leptonic channels. The obtained limits on the production cross sections range from 0.86 pb to 0.0059 pb for narrow resonances with masses from 0.75 TeV to 3.0 TeV, excluding a topcolor narrow $Z'$ boson with a mass below 2.4 TeV. The limit for a $Z'$ boson with a width of 10% is 2.9 TeV, while Kaluza-Klein gluon excitations predicted in Randall-Sundrum models are excluded below 2.8 TeV.

**First results with 13 TeV data** A similar analysis strategy as for 8 TeV is adopted by the ATLAS and CMS Collaborations for the first data at the LHC collision energy of 13 TeV [210, 211]. Figure 53 shows the $m_{t\bar{t}}$ spectrum in the category with the highest S/B fraction in the CMS analysis which does not exhibit any deviations from the SM background. With an integrated luminosity of 3.2 and 2.6 fb$^{-1}$, Topcolor $Z'$ bosons with relative widths of 1%, 10%, and 30% are excluded for the mass ranges of 0.6–2.5, 0.5–3.9, and 0.5–4.0 TeV, respectively. Kaluza-Klein gluon excitations predicted in Randall-Sundrum models are excluded below 3.3 TeV. These limits are significantly more stringent than those obtained with 8 TeV data.
5.2 Diboson Resonances

The new physics models discussed in the previous section also predict new resonances decaying to pairs of $W$, $Z$ or Higgs bosons. In more than 60% of the cases, $W/Z/H$ bosons decay into a quark anti-quark pair, which makes the reconstruction of such decays an essential ingredient for these searches.

The search for diboson resonances is performed at $\sqrt{s} = 13$ TeV in semi-leptonic [220, 225?] and fully hadronic final states [66, 125, 126, 127] by the ATLAS and CMS Collaborations. As the methods of jet substructure analyses exhibit their full strength in hadronic final states, the following discussion gives a summary and comparison of the ATLAS and CMS results in the search for $W/Z$ resonances in hadronic final states only.

In the analysis performed by the CMS Collaboration [66] events with two anti-$k_t$ jets with $R = 0.8$ and $65 < m_{\text{soft drop}} < 105$ GeV are selected, where the jet is called a $W$ boson candidate if the mass is in the range 65–85 GeV; while it is called a $Z$ boson candidate if the mass is in the range 85–105 GeV. This leads to the three signal categories $WW$, $ZZ$ and $WZ$. The jets are further categorized according to $\tau_{21}$ into high purity (HP, $\tau_{21} < 0.35$) and low purity (LP, $0.35 < \tau_{21} < 0.75$). Events are always required to have one HP $V$ jet, and are divided into HP and LP events, depending on whether the other $V$ jet is of high or low purity. To further suppress the large QCD multijet background a requirement on the dijet kinematics $|\eta_1 - \eta_2| < 1.3$ is applied.

The background is estimated from a signal + background fit with the function $\frac{dN}{dm_{jj}} = \frac{P_0}{(m_{jj}/\sqrt{s})^{\beta_1}}$, where $P_0$ is a normalization parameter and $P_1$ is a parameter describing the shape. As shown in Fig. 54 the data is well described by the fit function. Figure 54 also shows that no excess over the background-only hypothesis is observed. Exclusion limits are set on new physics resonances under different hypotheses of spin and charge. $W'$ and $Z'$ resonances with masses below 3.6 and 2.7 TeV are excluded for the heavy vector triplet model B [286].

A similar analysis has been performed by the ATLAS Collaboration [127]. In this analysis events are required to have at least two large-$R$ jets with $p_T > 200$ GeV in the pseudo-rapidity range $|\eta| < 2.0$. These jets are reconstructed with the anti-$k_t$ algorithm with a radius parameter $R = 1.0$. The trimming algorithm is applied using $k_T$ subjets with $R = 0.2$. The rapidity separation between the two leading jets has to satisfy $|\Delta y_{12}| < 1.2$. The large-$R$ jet mass is computed from the Combined Mass (see Sec. 1.5), and is required to be within a window of the expected $W$ or $Z$ mass value. The window width varies from 22 to 40 GeV depending on the jet $p_T$. In addition, the $D_2^{\beta=1}$ variable is used to select jets with a two-prong structure.

Similar as in CMS analysis, the background is estimated by fitting the dijet invariant mass distribution with the parametric form $\frac{dn}{dx} = p_1(1 - x)^{p_2 + \xi p_3 x p_3}$, where $n$ is the number of events, $x$ is a dimensionless variable related to the dijet mass $m_{jj}$, $x = m_{jj}/\sqrt{s}$, $p_1$ is a normalization factor, $p_2$ and $p_3$ are dimensionless shape parameters, and $\xi$ is a constant chosen to remove the correlation between $p_2$ and $p_3$ in the fit.

The dijet invariant mass distributions for these events are shown in Figs. 55, where good agreement is found between data and the expectations from the background fit. Limits are set on the cross section times branching ratio for diboson resonances. In the heavy vector triplet model B a spin-1 vector triplet is excluded for masses between 1.2 and 3.5 TeV.

The analyses described above investigate final states with boosted $W$ and $Z$ bosons. In case of boosted $H$
bosons, different reconstruction methods have to be used to benefit from the presence of $b$ quarks in $H \rightarrow b \bar{b}$ decays (see Sec. 3.4). Results have been published on the search for for $WH/ZH$ final states [126, 221, 226] as well as for $HH$ final states [222, 223, 227] using proton-proton collisions at $\sqrt{s} = 13$ TeV.

5.3 Vector-like Quarks

Vector-like quarks (VLQs) are predicted by a variety of theories introducing a mechanism that stabilizes the mass of the Higgs particle. Such theories include little Higgs models [287, 288], models with extra dimensions [289, 290], and composite Higgs models [289, 290, 291]. As VLQs are expected to have large masses, jet substructure analyses have been applied in many searches for VLQs.

The first search for VLQs using jet substructure methods was an inclusive search for pair-produced $T$ quarks [228]. This search was conducted by the CMS Collaboration using $pp$ collisions at $\sqrt{s} = 8$ TeV corresponding to an integrated luminosity of 19.7 fb$^{-1}$. In this search the Cambridge/Aachen algorithm was used with a distance parameter $R = 0.8$ (CA8 jets). Boosted $W$ jets are identified based on the mass of the CA8 jet while boosted top jets are identified with the CMSTT, described in Sec. 5.1. The limits on the hypothetical mass of a VLQ are between 687 and 782 GeV, depending on the decay modes.

The first search for VLQs in the all-hadronic final state has been performed by CMS and is described in Ref. [229].

The strategy of this analysis is to apply the C/A algorithm with a large size parameter of $R = 1.5$ to cluster the decay products from top quarks and Higgs bosons into single large jets. To identify the origin of the large C/A jets a top tagging algorithm (HEPTopTagger) and a Higgs tagging algorithm are used. This was the first time the HEPTopTagger algorithm has been applied in a data analysis by the CMS Collaboration. This was also the first time a Higgs tagging algorithm based on subjet $b$-tagging (see Sec. 3.4) has been used. Two subjets must be $b$-tagged and their invariant mass must be greater than 60 GeV to fulfill the Higgs tagging requirement. The multiplicity of these Higgs tags is shown in Fig. 56 which demonstrates that both the QCD multijet and the $t\bar{t}$ backgrounds can be suppressed by several orders of magnitudes. The efficiencies of these algorithms were also measured, since these are needed for modelling the signal efficiencies. As a result of this analysis VLQ masses of less than 745 GeV have been excluded at 95% confidence level, assuming a 100% branching fraction into the $tH$ final state.

Extensive use of substructure methods has also been made by the ATLAS Collaboration, in particular for the search for single production of VLQs. The single pro-

Figure 55: The observed data in the signal region of the WZ category. Also shown is the fitted background prediction. The gray region represents the uncertainty in the background estimate. Taken from Ref. [127].

Figure 56: Multiplicity of CA15 jets which fulfill the Higgs tagging criteria. The solid histograms represent the simulated background processes ($t\bar{t}$ and QCD multijet). The hatched error bands show the statistical uncertainty of the simulated events. Taken from Ref. [229].
duction modes may have higher cross sections than pair production depending on the VLQ mass and the coupling parameters\cite{292}. ATLAS performed an analysis \cite{240} where the VLQ is searched for in the decay mode with a $W$ boson and a top quark ($B \to tW$). Final states with at least one lepton are considered, where either the $W$ boson or the top quark appear in a boosted configuration. They are identified by the application of a jet mass requirement ($m > 50\,\text{GeV}$) on a trimmed large-$R$ anti-$k_t$ jet with a distance parameter $R = 1.0$.

A different strategy is followed in another ATLAS search \cite{241}, where the decay into the $bW$ final state is investigated ($T/Y \to bW$). As the $W$ boson is assumed to decay leptonically, no boosted hadronic $W$ or top quark decays are present. Nevertheless, the analysis uses a veto on the presence of massive ($m > 70\,\text{GeV}$), trimmed large-$R$ anti-$k_t$ jets with $R = 1.0$, to suppress the dominant $tt$ background.

In the meantime jet substructure methods are widely employed in almost all VLQ searches at $\sqrt{s} = 13\,\text{TeV}$ published by the LHC Collaborations, see e.g. Refs. \cite{232, 233, 234, 235, 236, 237, 242}. Note that this list is not exhaustive by far and many more analyses are being finalized presently. The excluded VLQ masses are exceeding 1 TeV for all branching fractions. The sensitivity of these searches to VLQ masses up to and beyond 1 TeV is a huge success, which has been made possible thanks to jet substructure techniques.

### 5.4 Leptophobic $Z'$

Many extensions of the Standard Model predict resonances that couple to quarks and gluons \cite{293, 294, 295, 296}, including simplified Dark Matter (DM) models in which resonances couple only to quarks and DM particles \cite{297, 298, 299}.

The ATLAS and CMS Collaboration searched for low mass leptophobic $Z'$ decaying into $q\bar{q}$ in the full 2015+2016 dataset collected at 13 TeV \cite{122, 244}. The search for the $Z'$ production is performed in association with a high $p_T$ jet from initial-state radiation (or also a photon in the case of ATLAS) to be able to cover resonance masses of $m_{Z'} \ll 1\,\text{TeV}$. The sensitivity to low resonance masses is usually degraded due to high trigger thresholds and the enormous QCD multijet background. The decay products of the $Z'$ are collimated and captured within one single large-$R$ jet. Jets in the CMS analysis are reconstructed with the anti-$k_t$ algorithm with $R = 0.8$ and corrected for pileup effects with the soft drop algorithm ($\beta = 0$, $z_{\text{cut}} = 0.1$) whereas anti-$k_t$ $R = 1.0$ jets, trimmed with $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 5\%$ are used in ATLAS. To suppress the dominating QCD multijet background, CMS applies criteria on $N_2^1$ \cite{115} and ATLAS chooses $\tau_{21}$ as discriminator. To avoid distortions of the jet mass spectrum due to large correlation between the jet mass and substructure variables, the substructure variables are decorrelated from the jet mass with the designed decorrelated tagger method. Data-driven techniques are used to determine the most-dominating backgrounds, namely the QCD multijet production whereas sub-domination processes such as $W/Z+\text{jets}$ events are estimated from MC simulation. The jet mass distributions of the large-$R$ jet is shown in Fig. 57 and 58 for the CMS and ATLAS analyses, respectively. No evidence for a resonant structure on top of the SM background is observed and the results are interpreted by CMS (ATLAS) as 95% confidence level limits on the production cross section and the coupling of the $Z'$ to quarks, $g_q$, for $Z'$ masses in the range from 50 to 300 GeV (100 to 220 GeV). The CMS Collaboration excludes coupling values of $g_q > 0.25$ over the full mass range, with stronger constraints for low jet masses. After combining the ISR jet and ISR photon channels, ATLAS sets upper limits on $g_q$ of 0.17 for $m_{Z'} = 100\,\text{GeV}$ and 0.21 for $m_{Z'} = 220\,\text{GeV}$.

![Figure 57: Soft drop jet mass of anti-$k_t$ $R = 0.8$ jets in data and for the dominating background processes; multijet production and $W/Z+\text{jets}$ events. Taken from Ref. \cite{122}](image)
### 6 Future Collider Projects

The LHC guarantees a rich scientific output in the next decade, stretching well into the 2030s with the planned luminosity upgrade. A diverse range of possible energy frontier machines beyond the LHC has been proposed. A summary is given in Table 3.

Lepton colliders come in two varieties: the linear collider projects promise scalable machines [300, 301, 302], that can explore high-energy processes such as vector-boson-fusion Higgs production, associated production of a top quark pair and a Higgs boson and di-Higgs boson production, and can ultimately reach the multi-TeV regime [303]. Large circular colliders (with a circumference of 100 km) can provide superior luminosity at the maximum of the $ZH$ production cross section ($\sqrt{s} \sim 250$ GeV [304, 305]) and may reach the top pair production threshold [306].

Ambitious projects exist in Europe and China for a new $pp$ collider with a size several times that of the LHC. The FCC$\text{hh}$ [308, 309, 310, 311] and SPPC [305] projects are set to reach a center-of-mass energy of 100 TeV. An intermediate step could be found in the upgrade of the LHC dipoles with 16 T bending magnets, which could roughly double the energy reach. Conceptual Design Reports are being prepared for the machines that have not yet presented detailed design, such that a complete assessment of scientific merits and financial cost can be made for example for the update of the European Strategy in 2019.

#### 6.1 Hadron Colliders

A future high-energy hadron collider could reach an energy of up 100 TeV, with a luminosity that significantly exceeds that of the LHC [312]. Cross sections increase with center-of-mass energy. Combined, this yields a factor 100 increase in the sample of $WW$, $ZZ$, $t\bar{t}$ and Higgs events that can be collected. The increase is even larger for rare processes, such as $tH$ or di-Higgs boson production [313, 314] and is dramatic in the high-energy tails. A 100 TeV collider is a boosted-object factory, extending the mass reach for searches [309] to several tens of TeV. The study of boosted-object production rates also brings a much greater indirect sensitivity to new physics at even higher scale [315, 316, 317, 318].

With higher energies come new challenges. During LHC Run 1 and Run 2, a top quark is generally described as being boosted if it has $p_T$ greater than twice its mass, e.g. $p_T > 350$ GeV, as this would indicate a high probability of capturing the hadronic decay products in a single jet with $R = 1.0$. In future ultra-high energy colliders, boosted object taggers must evolve to exploit different features for objects with $p_T$ up to 10 or 20 TeV. At these energies, the decay products of multi-TeV top quarks, and especially $W$, $Z$ and Higgs bosons can be contained by jets with $\Delta R < 0.1$.

The analysis of the substructure in jets formed by hyper-boosted objects pose stringent requirements on the granularity and size of the experiment. To resolve the substructure of jets at a very small scale the detector granularity and size must exceed that of current LHC hadron calorimeter significantly [319, 320]. To contain high-energy hadrons a deep hadronic calorimeter is required, with up to 12 interaction lengths [321] (cf. the CMS barrel calorimeter has a depth of 7 $\lambda_{\text{int}}$, including 1 $\lambda_{\text{int}}$ in the EM system). The tracking efficiency in the core of jets, already under some pressure at the LHC, is another area where requirements related to jet substructure may condition the detector design. Detailed full-simulation studies are required to assess the experimental limitations.
Table 3: Projects for energy-frontier facilities. Integrated luminosities correspond to different running times (typically ten years per energy point) and are subject to large uncertainties. The last column provides a reference to the most recent design report. References to detailed running scenarios are given where available. All $e^+e^-$ colliders envisage a brief period of running at energies close to the top quark pair production threshold, that is not listed here.

### Energy-frontier lepton colliders

<table>
<thead>
<tr>
<th>Project</th>
<th>Host</th>
<th>Type</th>
<th>$\sqrt{s}$ (TeV)</th>
<th>$\int L$ ($ab^{-1}$)</th>
<th>Status</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILC</td>
<td>Japan</td>
<td>linear $e^+e^-$</td>
<td>0.25</td>
<td>0.5 (2)</td>
<td>TDR 2013 [301]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td>staging [300, 307]</td>
<td></td>
</tr>
<tr>
<td>CLIC</td>
<td>CERN</td>
<td>linear $e^+e^-$</td>
<td>0.38</td>
<td>0.5</td>
<td>CDR 2012 [303]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.5,3</td>
<td>staging [302]</td>
<td></td>
</tr>
<tr>
<td>CEPC</td>
<td>China</td>
<td>circular $e^+e^-$</td>
<td>0.25</td>
<td>5</td>
<td>CDR 2017 [305]</td>
<td></td>
</tr>
<tr>
<td>FCCee</td>
<td>CERN</td>
<td>circular $e^+e^-$</td>
<td>0.25/0.36</td>
<td>10/2.6</td>
<td>CDR &lt; 2019 [304]</td>
<td></td>
</tr>
<tr>
<td>$\mu$ collider</td>
<td>FNAL</td>
<td>racetrack $\mu^+\mu^-$</td>
<td>0.125-3</td>
<td>-</td>
<td>R&amp;D</td>
<td></td>
</tr>
</tbody>
</table>

### Very-high-energy pp colliders

<table>
<thead>
<tr>
<th>Project</th>
<th>Host</th>
<th>Type</th>
<th>$\sqrt{s}$ (TeV)</th>
<th>$\int L$ ($ab^{-1}$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE-LHC</td>
<td>CERN</td>
<td>LEP/LHC</td>
<td>25</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>FCCChh</td>
<td>CERN</td>
<td>new</td>
<td>100</td>
<td>20</td>
<td>[308, 309, 310, 311]</td>
</tr>
<tr>
<td>SPPC</td>
<td>China</td>
<td>new</td>
<td>70-140</td>
<td>3</td>
<td>[305]</td>
</tr>
</tbody>
</table>
Several authors have explored different approaches to tag highly boosted top quarks. A "traditional" substructure analysis may be feasible [309] well into the multi-TeV regime. Larkoski et al. show that a substructure analysis limited to charged particle tracks provides significant separation power, thus avoiding strong constraints on the calorimeter granularity [322]. For top quarks, the lepton-in-jet signature [316] may be sufficient to isolate the signal.

Ref. [323] studies the prospect to identify highly boosted $W, Z$ or Higgs bosons in 100 TeV collisions at the FCC. As these are color singlets the energy of the boson is contained fully inside a critical radius $R \sim 2M/p_T$. This radiation pattern differs significantly from QCD jets. This suggests that isolation criteria such as those used for hadronic $\tau$-lepton decays may be used to isolate a sample of hadronic gauge boson decays. The same report [323] finds that a more traditional substructure analysis yields a separation between a boosted gauge boson sample and quark or gluon jet background for $p_T$ up to several tens of TeV.

The reconstruction and isolation of hyper-boosted objects with $p_T$ of 10-20 TeV poses important requirements on the detector design. This unexplored kinematic regime requires the development of new algorithms and techniques.

### 6.2 Lepton Colliders

A high-energy $e^+e^-$ collider can provide precise measurements of the interactions of the Higgs boson, SM gauge bosons, the top quark and possible new states [300, 324, 325]. An accurate reconstruction of hadronic final states is a prerequisite for a precise measurements of Higgs boson and top quark couplings [326, 327, 328].

Excellent jet clustering is required to take full advantage of the potential of the machine and detectors. This is considerably more challenging than at the previous generation of lepton colliders, especially in complex multijet final states with multiple energy scales [329], where the contribution of jet clustering to the mass resolution may be sizeable or even dominant [328].

Detailed detector designs [330] have been developed for the linear collider projects. They achieve excellent single-particle reconstruction with highly granular calorimeters [331, 332], and particle-flow algorithms [333]. This approach yields a superb response to the jet energy and jet substructure observables [328].

The collection of a clean sample of energetic jets offers excellent opportunities to study jet formation in an environment that is more readily calculable, less affected by pileup, and much less prone to biases in trigger and selection. The detailed understanding of fragmentation would benefit analyses employing jet substructure at hadron colliders [334].

At linear colliders $\gamma\gamma \rightarrow$ hadrons production yields a diffuse background superposed on the signal [335]. Longitudinally invariant algorithms developed for hadron colliders are found to be significantly more robust than traditional $e^+e^-$ jet reconstruction algorithms [303]. The same background resilience can be achieved with a classical $e^+e^-$ inter-particle distance criterion by modifying the beam distance [327, 328].

Work in this direction has only just started [328] and only a few studies exist in the literature on the use of boosted object tagging and jet substructure analysis at future lepton colliders. The separation between quark and gluon jets that can be achieved with jet substructure is likely to play an important role in the isolation of a $H \rightarrow gg$ signal at a Higgs factory with $\sqrt{s} = 250$ GeV. The reconstruction of boosted gauge bosons, Higgs bosons and top quarks in $e^+e^-$ colliders with a center-of-mass above 1 TeV can benefit from reconstruction, grooming and tagging algorithms developed for the LHC [328].

### 7 Conclusions

Jet substructure is a term used to describe the calculations, algorithms, and analysis techniques developed over the last decade and reviewed in this article. These methods are used to exploit the details of hadronic activity detectable by modern particle detectors such as ATLAS and CMS, and measurements and searches at both these experiments increasingly rely on one or more of the tools developed by the jet substructure community for boosted particle tagging and pileup suppression. Our hope is that this review serves as a useful reference, facilitating the community to continue tackling the challenges of high-energy experimental particle physics.

### Acknowledgments

The editors thank CERN and the ATLAS and CMS collaborations, the participants and organizers of the Boost Workshops held in Zurich 2016 [336] and Buffalo 2017 [337] for discussions and input, and Jon Butterworth for sug-
gesting this jet substructure review article. We also thank Andrew Larkoski and Ian Moul for the collaboration on the theoretical section of the review.

References


[54] CMS Collaboration. Measurement of the differential jet production cross section with respect to jet mass and transverse momentum in dijet events from


[73] Peter Berta, Martin Spousta, David W. Miller, and Rupert Leitner. Particle-level pileup subtraction for


[181] ATLAS Collaboration. Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector. JHEP, 12:024,


[198] ATLAS Collaboration. Measurement of the top quark mass in the \( t\bar{t} \rightarrow \text{dilepton} \) channel from \( \sqrt{s} = 8 \) TeV pp collisions at \( \sqrt{s} = 7 \) TeV with the ATLAS detector.


[249] CMS Collaboration. Search for supersymmetry in proton-proton collisions at 13 TeV using identified


59


