Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1

ATLAS Collaboration

DOI
10.1140/epjc/s10052-017-5004-5

Publication date
2017

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date: 13 Apr 2024
Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 10 March 2016 / Accepted: 21 June 2017 / Published online: 24 July 2017

© CERN for the benefit of the ATLAS collaboration 2017. This article is an open access publication

Abstract The reconstruction of the signal from hadrons and jets emerging from the proton–proton collisions at the Large Hadron Collider (LHC) and entering the ATLAS calorimeters is based on a three-dimensional topological clustering of individual calorimeter cell signals. The cluster formation follows cell signal-significance patterns generated by electromagnetic and hadronic showers. In this, the clustering algorithm implicitly performs a topological noise suppression by removing cells with insignificant signals which are not in close proximity to cells with significant signals. The resulting topological cell clusters have shape and location information, which is exploited to apply a local energy calibration and corrections depending on the nature of the cluster. Topological cell clustering is established as a well-performing calorimeter signal definition for jet and missing transverse momentum reconstruction in ATLAS.

Contents

1 Introduction ........................................ 2
2 The ATLAS experiment ............................ 2
  2.1 The ATLAS detector .......................... 2
    2.1.1 The ATLAS detector systems .......... 3
    2.1.2 The ATLAS trigger ...................... 5
  2.2 Dataset ........................................ 5
    2.2.1 Pile-up in data ......................... 5
    2.2.2 Effect on calorimeter noise .......... 7
  2.3 Monte Carlo simulations .................... 8
    2.3.1 Monte Carlo simulations of signal samples .. 8
    2.3.2 Minimum-bias samples and pile-up modelling .... 9
    2.3.3 Minimum-bias overlay samples for 2012 ... 9
    2.3.4 Detector simulation .................... 9
  2.4 Hadronic final-state reconstruction in ATLAS . 9
  3 Topological cluster formation and features ...... 10
    3.1 Topo-cluster formation .................... 10

3.1.1 Collecting cells into topo-clusters ......... 10
3.1.2 Treatment of negative cell signals .......... 11
3.1.3 Cluster splitting .......................... 12
3.1.4 Cluster multiplicities in electromagnetic and hadronic showers .......... 14
3.2 Cluster kinematics ............................ 15
4 Topo-cluster moments ............................ 15
  4.1 Geometrical moments ...................... 16
    4.1.1 Location .............................. 16
    4.1.2 Directions ............................ 16
    4.1.3 Extensions and sizes .................. 17
  4.2 Signal moments ............................. 17
    4.2.1 Signal significance .................... 17
    4.2.2 Signal density ....................... 18
    4.2.3 Signal timing ........................ 18
    4.2.4 Signal composition .................... 18
    4.2.5 Topological isolation .................. 19
5 Local hadronic calibration and signal corrections ... 19
  5.1 General topo-cluster calibration strategy .......... 20
  5.2 Cluster classification ..................... 21
  5.3 Hadronic calibration ....................... 22
  5.4 Correction for out-of-cluster signal losses .......... 24
  5.5 Dead material corrections .................. 27
  5.6 Fully calibrated cluster kinematics ............ 28
6 Performance of the simulation of topo-cluster kinematics and properties .......... 30
  6.1 Single-particle response .................... 30
  6.2 Effect of pile-up on topo-cluster observables .... 31
    6.2.1 Event selection ....................... 32
    6.2.2 Modelling of topo-cluster kinematics in events with pile-up .......... 32
    6.2.3 Transverse momentum flow in the presence of pile-up .......... 32
    6.2.4 Topo-cluster multiplicity in the presence of pile-up .......... 34
    6.2.5 Modelling of the topo-cluster depth location in the presence of pile-up .......... 38
  6.3 Topo-clusters in jets ....................... 38

*e-mail: atlas.publications@cern.ch
1 Introduction

The detectable final state emerging from the proton–proton collisions at the Large Hadron Collider (LHC) consists of particles and jets which are reconstructed with high precision for physics analyses. In the ATLAS experiment [1], clusters of topologically connected calorimeter cell signals (topo-clusters) are employed as a principal signal definition for use in the reconstruction of the (hadronic) part of the final state comprising isolated hadrons, jets and hadronically decaying \( \tau \)-leptons. In addition, topo-clusters are also used to represent the energy flow from softer particles, which is needed for the reconstruction of full-event observables such as the missing transverse momentum.

The algorithm building the topo-clusters explores the spatial distribution of the cell signals in all three dimensions to establish connections between neighbours in an attempt to reconstruct the energy and directions of the incoming particles. The signals from cells determined to be connected are summed, and are used together with the cell locations to calculate direction, location, and shapes of the resulting clusters. Calorimeter cells with insignificant signals found to not be connected to neighbouring cells with significant signals are considered noise and discarded from further jet, particle and missing transverse momentum reconstruction.

The topo-clusters, while well established in deep inelastic scattering experiments such as H1 [2] at HERA and in electron–positron collider experiments such as ALEPH [3] at LEP and BaBar [4] at PEP-II, are used here in an innovative implementation as fully calibrated three-dimensional objects representing the calorimeter signals in the complex final-state environment of hadron–hadron collisions. A similar application in this particular environment, previously developed by the D0 Collaboration, implements the topological clustering in the two dimensions spanned by pseudorapidity and the azimuthal angle, thus applying the noise-suppression strategy inherent in this algorithm for jet reconstruction [5]. Several features and aspects of the ATLAS topo-cluster algorithms and their validations have previously been presented in Refs. [6–9].

Some of the complexity of the final state in hadron–hadron collisions is introduced by particles from the underlying event generated by radiation and multiple parton interactions in the two colliding hadrons producing the hard-scatter final state. Other detector signal contributions from the collision environment, especially important for higher intensity operations at the LHC, arise from pile-up generated by diffuse particle emissions produced by the additional proton–proton collisions occurring in the same bunch crossing as the hard-scatter interaction (in-time pile-up). Further pile-up influences on the signal are from signal remnants from the energy flow in other bunch crossings in the ATLAS calorimeters (out-of-time pile-up).

This paper first describes the ATLAS detector in Sect. 2, together with the datasets used for the performance evaluations. The motivations and basic implementation of the topo-cluster algorithm are presented in Sect. 3. The computation of additional variables associated with topo-clusters including geometric and signal moments is described in Sect. 4. The various signal corrections applied to topo-clusters in the context of the local hadronic calibration are presented in Sect. 5. Section 6 summarises the performance of the topo-cluster signal in the reconstruction of isolated hadrons and jets produced in the proton–proton collisions at LHC. Performance evaluations with and without pile-up are discussed in this section, together with results from the corresponding Monte Carlo (MC) simulations. The paper concludes with a summary and outlook in Sect. 7.
2.1.1 The ATLAS detector systems

The detector closest to the proton–proton collision vertex is the inner tracking detector (ID). It has complete azimuthal coverage and spans the pseudorapidity region $|\eta| < 2.5$. It consists of a silicon pixel detector, a silicon micro-strip detector, and a straw-tube transition radiation tracking detector covering $|\eta| < 2$. The ID is immersed into a uniform axial magnetic field of 2 T provided by a thin superconducting solenoid magnet.

The ATLAS calorimeter system is illustrated in Fig. 1. It comprises several calorimeters with various read-out granularities and with different technologies. The electromagnetic calorimeter (EM) surrounding the ID is a high-granularity liquid-argon sampling calorimeter (LAr), using lead as an absorber. It is divided into one barrel (EMB; $|\eta| < 1.475$) and two end-cap (EMEC; $1.375 < |\eta| < 3.2$) regions.

The barrel and end-cap regions also feature pre-samplers mounted between the cryostat cold wall and the calorimeter modules. The barrel pre-sampler (PreSamplerB) covers $|\eta| < 1.52$, while the end-cap pre-sampler (PreSamplerE) covers $1.5 < |\eta| < 1.8$.

The hadronic calorimeters are divided into three distinct sections. The most central section contains the central barrel region ($|\eta| < 0.8$) and two extended barrel regions ($0.8 < |\eta| < 1.7$). These regions are instrumented with scintillator-tile/steel hadronic calorimeters (Tile). Each barrel region consists of 64 modules with individual azimuthal ($\phi$) coverages of $\pi/32$ rad. The two hadronic end-cap calorimeters (HEC; $1.5 < |\eta| < 3.2$) feature liquid-argon/copper calorimeter modules. The two forward calorimeters (FCAL; $3.1 < |\eta| < 4.9$) are instrumented with liquid-argon/copper and liquid-argon/tungsten modules for electromagnetic and hadronic energy measurements, respectively.

The ATLAS calorimeters have a highly granular lateral and longitudinal segmentation. Including the pre-samplers, there are seven sampling layers in the combined central calorimeters (PreSamplerB, three in EMB and three in Tile) and eight sampling layers in the end-cap region (PreSamplerE, three in EMEC and four in HEC). The three FCal modules provide three sampling layers in the forward region. Altogether, the calorimeter system has about 188 000 read-out channels. The EM calorimeters are between 24 radiation lengths ($X_0$) and 27 $X_0$ deep. The combined depth of the calorimeters for hadronic energy measurements is more than 10 hadronic interaction lengths ($\lambda$) nearly everywhere across the full detector acceptance ($|\eta| \leq 4.9$). The amount of inactive material in front of the calorimeters depends on $\eta$. It varies from about 2 $X_0$ at $\eta = 0$ to about 4 $X_0$ at $|\eta| \approx 1.8$, when measured from the nominal interaction point in ATLAS to the first active sampling layer (including PreSamplerB and PreSamplerE). It can increase to more than 6 $X_0$ in the transition region between central and end-cap calorimeters ($|\eta| \approx 1.45$ and $|\eta| \approx 1.7$). The amount of inactive material for hadrons is approximately $1 \lambda$ across the full covered $\eta$-range, with spikes going up to more than $2 \lambda$ in transition regions and in regions with complex cryostat structures and beam line services ($|\eta| \approx 4$).
The absorption power of the ATLAS calorimeters and their segmentation allow for very precise energy-flow reconstruction based on the topo-clusters described in this paper, with considerable exploitation of the topo-cluster shapes for signal calibration purposes. For more details of the calorimeter read-out structures, absorption characteristics, inactive material distributions, and cell signal formation, see Ref. [1]. The segmentation of the read-out structure in the various calorimeter sampling layers, each named by a dedicated identifier (S_{calo}), is shown in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Calorimeter</th>
<th>Module sampling (S_{calo})</th>
<th>N_{cells}</th>
<th>(\eta)-coverage</th>
<th>(\Delta\eta \times \Delta\phi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electromagnetic calorimeters</td>
<td>EMB</td>
<td>109,568</td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>PreSamplerB</td>
<td>7808</td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EMB1</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EMB2</td>
<td></td>
<td>(1.4 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EMB3</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EMEC</td>
<td>63,744</td>
<td>(1.375 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>PreSamplerE</td>
<td>1536</td>
<td>(1.5 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EME1</td>
<td></td>
<td>(1.375 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EME2</td>
<td></td>
<td>(1.425 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>EME3</td>
<td></td>
<td>(1.425 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td>Hadronic calorimeters</td>
<td>Tile (barrel)</td>
<td>2880</td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>TileBar0/1</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>TileBar2</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>Tile (extended barrel)</td>
<td>2304</td>
<td>(0.8 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>TileExt0/1</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>TileExt2</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>HEC</td>
<td>5632</td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>HEC0/1/2/3</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td>Forward calorimeters</td>
<td>FCAL</td>
<td>3524</td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>FCAL0</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>FCAL1</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>FCAL2</td>
<td></td>
<td>(</td>
<td>\eta</td>
</tr>
</tbody>
</table>

For comparison, the FCAL granularity is approximately \(\Delta\eta \times \Delta\phi = 0.15 \times 0.15(0.3 \times 0.3)\) at \(\eta = 3.5(4.5)\). The total number of read-out cells, including both ends of the calorimeter system, with (without) pre-samplers is 187,652 (178,308).
The muon spectrometer surrounds the ATLAS calorimeters. A system of three large air-core toroids, a barrel and two end-caps with eight coils each, generates a magnetic field in the pseudorapidity range of $|\eta| < 2.7$. The muon spectrometer measures the full momentum of muons based on their tracks reconstructed with three layers of precision tracking chambers in the toroidal field. It is also instrumented with separate trigger chambers.

### 2.1.2 The ATLAS trigger

The trigger system for the ATLAS detector in Run 1 consisted of a hardware-based Level 1 ($L_1$) trigger and a software-based High Level Trigger (HLT) [10]. For the evaluation of the topo-cluster reconstruction performance, samples of minimum-bias (MB) triggered events, samples of events selected by jet triggers, and samples of events with hard objects such as muons, which are not triggered by the calorimeter, are useful.

The ATLAS MB trigger [11] used signals from a dedicated system of scintillators ($MBTS$ [12]; $2.1 < |\eta| < 3.8$) at $L_1$ in 2010 and 2011 data-taking. Depending on the run period, it required one hit in either of the $\eta$ hemispheres, or one hit in each $\eta$ hemisphere. In 2012, the MB samples were triggered by a zero-bias trigger. This trigger unconditionally accepted events from bunch crossings occurring a fixed number of LHC cycles after a high-energy electron or photon was accepted by the $L_1$ trigger. The $L_1$ trigger rate for these hard objects scales linearly with luminosity, thus the collision environment generated by the luminosity-dependent additional proton–proton interactions discussed in Sect. 2.2.1 is well reflected in the MB samples.

For triggering on collision events with jets at $L_1$, jets are first built from coarse-granularity calorimeter towers using a sliding-window algorithm ($L_1$-jets). The events are accepted if they have $L_1$-jets passing triggers based on (1) the transverse momentum ($p_T$) of individual $L_1$-jets (single-jet triggers) or on (2) the detection of several such jets at increasing transverse momenta (multi-jet triggers). Those events accepted by $L_1$ are then subjected to refined jet-trigger decisions based on jet $p_T$ and multi-jet topology in the HLT, now using jets that are reconstructed from calorimeter cell signals with algorithms similar to the ones applied in the offline precision reconstruction [13].

A Z boson sample is collected from muon triggers at $L_1$. Since the trigger rate and the reconstruction of the decay properties of the accepted $Z \rightarrow \mu\mu$ events are basically unaffected by pile-up, this sample is not only unbiased in this respect but also with respect to other possible biases introduced by the ATLAS calorimeter signals.

### 2.2 Dataset

The data used for the evaluation of the topo-cluster reconstruction performance are selected from proton–proton collision events at a centre-of-mass energy of $\sqrt{s} = 7$ TeV, recorded with the ATLAS detector in 2010, and at $\sqrt{s} = 8$ TeV in 2012. The overall amount of high-quality data recorded at those times corresponds to $\sim 45 \text{pb}^{-1}$ in 2010, and $\sim 20.3 \text{fb}^{-1}$ in 2012. Peak instantaneous luminosities reached in the first three years of LHC running (LHC Run 1) are shown in Fig. 2a. Some early data recorded during the very first proton–proton collisions in the LHC in 2009 are considered for the studies of the topo-cluster reconstruction performance as well. The corresponding events are extracted from approximately $540\,000$ proton–proton collisions at $\sqrt{s} = 900$ GeV, recorded during stable beam conditions and corresponding to about $12 \text{mb}^{-1}$. Occasional references to 2011 run conditions, where protons collided in the LHC with $\sqrt{s} = 7$ TeV and ATLAS collected data corresponding to $\sim 5.1 \text{fb}^{-1}$, are provided to illustrate the evolution of the operational conditions during LHC Run 1 relevant to topo-cluster reconstruction. The specific choice of 2010 and 2012 data for the performance evaluations encompasses the most important scenarios with the lowest and highest luminosity operation, respectively.

#### 2.2.1 Pile-up in data

One important aspect of the contribution from additional proton–proton interactions (pile-up) to the calorimeter signal in data is the sensitivity of the ATLAS liquid-argon calorimeters to this pile-up as a function of the instantaneous luminosity, and as a function of the signal history from previous bunch crossings.

In the initial phase of data-taking in 2010 the proton beam intensities at LHC were relatively low. The recorded events contain on average three additional proton–proton interactions, as shown in Fig. 2b. In addition, the initial bunch crossing interval of $t_{BX} = 750 \text{ns}$ was larger than the window of sensitivity of the $LAr$ calorimeter, which is defined by the duration $t_{\text{signal}}$ of the shaped signal, with $t_{\text{signal}} \approx 600 \text{ns}$, as depicted in Fig. 3 for the typical charge collection time of $t_d = 450 \text{ns}$ in this detector. In later data-taking periods in 2010 the bunch crossing interval was reduced to $t_{BX} = 175 \text{ns}$, which is within the sensitivity of the $LAr$ calorimeter signal formation ($t_{BX} < t_{\text{signal}}$). Nevertheless, the still-low instantaneous luminosity reduced the amount of energy scattered into the calorimeter in the other bunch crossings to a negligible contribution with little effect on the signal history.

Throughout operations in 2011 and 2012, the proton beam intensities in the LHC were significantly increased, leading to the corresponding increases in the number of pile-up interac-
The peak luminosities measured by the ATLAS online luminosity monitor system throughout the run years are shown in (a). The mean number of additional proton–proton interactions at the beginning of each LHC fill is shown in (b) for the same period in time.

![Graphs showing peak luminosities and mean number of additional interactions](image)

The bipolar shape of the LAr calorimeter signal shown in Fig. 3 reduces the overall effect of pile-up, because it features a net-zero integral over time. This leads to cancellation on average of in-time pile-up signal contributions by out-of-time pile-up signal residuals in any given calorimeter cell. By design of the shaping amplifier, and the choice of digitally sampling the shaped pulse amplitude in time with a frequency of 40 MHz in the read-out, the most efficient suppression is achieved for 25 ns bunch spacing in the LHC beams. It is fully effective in the limit where for each bunch crossing contributing to out-of-time pile-up about the same amount of energy is deposited in a given calorimeter cell. A small loss of efficiency is observed for 50 ns bunch spacing, due to the less frequent injection of energy by the fewer previous bunch crossings.

Approximately the first ten bunch crossings in each LHC bunch train at 50 ns bunch spacing are characterised by different out-of-time pile-up contributions from the collision history. This history gets filled with signal remnants from an increasing number of past bunch crossings with proton–proton interactions the larger the time difference between the bunch crossing and the beginning of the train becomes. The remaining bunch crossings in a train, about 26 of a total of 36 in 2011 and 62 of a total of 72 in 2012, have an out-of-time pile-up signal contribution which is stable within the bunch-to-bunch fluctuations in the beam intensity. In 2012 data a dedicated cell-by-cell correction is applied in the offline cell signal reconstruction to compensate for the corresponding variations in the out-of-time pile-up. Further details of the
Even with a constant proton bunch intensity and apart from the bunch train effects, the efficiency of pile-up suppression by signal shaping is reduced by the large fluctuations in the number of additional interactions from bunch crossing to bunch crossing, and by the different energy-flow patterns of the individual collisions in the time window of sensitivity $\tau_{\text{signal}}$ in the LAr calorimeters. Consequently, the signal shows a principal sensitivity to pile-up, even after shaping and digital filtering in the read-out. This is evident from the residual event-by-event deviation of the cell-signal baseline, which depends on the specific pile-up condition at the time of the triggered event, from the (average zero) baseline expected from the signal shaping. These baseline fluctuations can lead to relevant signal offsets once the noise suppression is applied, which is an important part of the calorimeter signal extraction strategy using topo-clusters presented in Sect. 3.

The Tile calorimeter shows very little sensitivity to pile-up since most of the associated (soft particle) energy flow is absorbed in the LAr calorimeters in front of it. Moreover, out-of-time pile-up is suppressed by a shorter signal collection time and a short pulse shaping time, reducing the sensitivity of the signal to only about three bunch crossings at 50 ns intervals [12].

2.2.2 Effect on calorimeter noise

In ATLAS operations prior to 2011 the cell noise was dominated by electronic noise. The short bunch crossing interval and higher instantaneous luminosity in 2011 and 2012 LHC running added additional and dominant noise contributions from the cell-signal baseline fluctuations introduced by pile-up, as discussed in Sect. 2.2.1. These fluctuations, even though not perfectly following a Gaussian distribution,\(^2\) can nevertheless be expressed as noise measured by the standard deviation of their distribution, taken from simulated MB events and scaled to the expected number of pile-up interactions. The cell noise thresholds steering the topo-cluster formation described in Sect. 3 thus needed to be increased from those used in 2010 to accommodate this pile-up-induced noise. This is done by adjusting the nominal energy-equivalent noise $\sigma_{\text{noise}}$ according to

$$
\sigma_{\text{noise}} = \begin{cases} 
\sigma_{\text{noise}}^{\text{electronic}} & (2010 \text{ operations}), \\
\sqrt{\left(\sigma_{\text{noise}}^{\text{electronic}}\right)^2 + \left(\sigma_{\text{noise}}^{\text{pile-up}}\right)^2} & (2011 \text{ and } 2012 \text{ operations}).
\end{cases}
$$

Here, $\sigma_{\text{noise}}^{\text{electronic}}$ is the electronic noise, and $\sigma_{\text{noise}}^{\text{pile-up}}$ the noise from pile-up, corresponding to an average of eight additional proton–proton interactions per bunch crossing ($\mu = 8$) in 2011, and $\mu = 30$ in 2012. These configurations are choices based on the expected average ($\langle \mu \rangle$) for the run year. They needed to be made before the respective data-taking started, to allow for a fast turn-around reconstruction of the collected data. As $\mu$ changes with the decrease of the instantaneous luminosity $L_{\text{inst}}$ through-out the LHC proton fill, $\sigma_{\text{noise}}^{\text{pile-up}}$ is only optimal for the small subset of data recorded when $L_{\text{inst}}$ generated the nominal (a priori chosen) $\mu$ pile-up interactions on average. LHC operations at lower $\mu$ lead to slightly reduced calorimeter sensitivity to relevant small signals, as $\sigma_{\text{noise}}^{\text{pile-up}}$ is too large. For data-taking periods with higher than nominal $\mu$ the noise suppression is not optimal, leading to more noise contributions to the topo-cluster signals.

The change of the total nominal noise $\sigma_{\text{noise}}$ and its dependence on the calorimeter region in ATLAS can be seen by comparing Fig. 4a–c. In most calorimeter regions, the total noise rises significantly above the electronic noise with increasing pile-up activity, as expected. This increase is largest in the forward calorimeters, where $\sigma_{\text{noise}}^{\text{pile-up}} \gg \sigma_{\text{noise}}^{\text{electronic}}$ by more than one order of magnitude, already under 2011 run conditions.

\(^2\) Selected examples of the actual distributions taken from data are shown in the context of the topo-cluster formation discussed in Sect. 3.1.1.
2.3 Monte Carlo simulations

The energy and direction of particles produced in proton–proton collisions are simulated using various MC event generators. An overview of these generators for LHC physics can be found in Ref. [17]. The samples for comparisons to 2010 data are produced at $\sqrt{s} = 7$ TeV, while the MC samples for 2012 analyses are generated at $\sqrt{s} = 8$ TeV. Some configuration details for the inclusive jet and inclusive $Z$ boson MC samples and the simulated MB samples are given below.

2.3.1 Monte Carlo simulations of signal samples

Simulated signal samples include inclusive jet-production, which is generated using PYTHIA [18] version 6.425 for 2010 analyses, and PYTHIA8 [19] version 8.160 for 2012 analyses. Both generators model the hard sub-process in the final states of the generated proton–proton collisions using a $2 \to 2$ matrix element at leading order in the strong coupling $\alpha_S$. Additional radiation is modelled in the leading-logarithmic (LL) approximation by $p_T$-ordered parton showers [20]. Multiple parton interactions (MPI) [21], as well as fragmentation and hadronisation based on the Lund string model [22], are also generated.

For comparisons with 2012 data, samples of $Z$ bosons with $Z \to \mu\mu$ are generated. The next-to-leading-order (NLO) POWHEG [23, 24] model is used, with the final-state partons showered by PYTHIA8 using the CT10 NLO parton distribution function (PDF) [25] and the ATLAS AU2 [26] set of tuned parton shower and other soft underlying event generation parameters. PYTHIA8 also provides the MPI, fragmentation and hadronisation for these events.
2.3.2 Minimum-bias samples and pile-up modelling

The MB samples for 2012 running conditions are generated using PYTHIA8 with the ATLAS AM2 [26] set of tuned soft interaction parameters and the MSTW2008LO PDF set [27]. A single, fully simulated event for that run year is built by overlaying a number \( N_{\text{PU}} \) of generated MB events onto one generated hard-scatter event. The actual \( N_{\text{PU}} \) is drawn from a Poisson distribution around the average number \( \langle \mu \rangle \) of additional proton–proton collisions per bunch crossing. The value of \( \langle \mu \rangle \) is measured by the experiment as an average over one luminosity block, which can last as long as two minutes, with its actual duration depending on the central data acquisition configuration at the time of the data-taking. The measurement of \( \langle \mu \rangle \) is mainly based on single \( \eta \)-hemisphere hit counting as well as counting coincidental hits in both \( \eta \)-hemispheres with the fast ATLAS luminosity detectors consisting of two small Cherenkov counter (LUCID; \( 5.6 < |\eta| < 6.0 \)) and two sets of small diamond sensors forming two beam conditions monitors (BCM; \( |\eta| = 4.2 \)). Details of these detectors and the measurement are given in Ref. [28]. The distribution of the measured \( \langle \mu \rangle \) over the whole run period is taken into account in the pile-up simulation.

The LHC bunch train structure with 72 proton bunches per train and 50 ns spacing between the bunches in 2012, is also modelled by organising the simulated collisions into four such trains. This allows the inclusion of out-of-time pile-up effects driven by the distance of the hard-scatter events from the beginning of the bunch train, as discussed in Sect. 2.2.1. A correction depending on the bunch position in the train is applied to data and MC simulations to mitigate these effects. Bunch-to-bunch intensity fluctuations in the LHC are not included in the MC modelling. These are corrected in the data by the correction depending on the position of the bunch in the train.

2.3.3 Minimum-bias overlay samples for 2012

In addition to the fully generated and simulated MC samples described earlier, samples with events mixing data and MC simulations are used to study the topo-cluster reconstruction performance. These samples are produced by overlaying one event from the MB samples collected by the zero-bias trigger described in Sect. 2.1.2 and a hard-scatter interaction from the MC generator [29–31]. The generated hard-scatter event is simulated using the detector simulation described in Sect. 2.1, but without any noise effects included. The recorded and simulated raw electronic signals are then overlaid prior to the digitisation step in the simulation. This results in modelling both the detector noise and the effect of pile-up from data with the correct experimental conditions on top of the simulated event. Theses samples are useful for detailed comparisons of topo-cluster signal features in 2012, as they do not depend on limitations in the soft-event modelling introduced by any of the generators.

2.3.4 Detector simulation

The GEANT4 software toolkit [32] within the ATLAS simulation framework [33] propagates the stable particles\(^3\) produced by the event generators through the ATLAS detector and simulates their interactions with the detector material and the signal formation. Hadronic showers are simulated with the quark–gluon-string-plasma model employing a quark–gluon string model [34] at high energies and the Bertini intra-nuclear cascade model [35–37] at low energies (QGSP_BERT). There are differences between the detector simulation used in 2010 and in 2012. A newer version of GEANT4 (version 9.4) is employed in 2012, together with a more detailed description of the LAr calorimeter absorber structure. These geometry changes introduce an increase of about 2% in the calorimeter response to pions with energies of less than 10 GeV.

2.4 Hadronic final-state reconstruction in ATLAS

The fully reconstructed final state of the proton–proton collisions in ATLAS includes identified individual particles comprising electrons, photons, muons, and \( \tau \)-leptons, in addition to jets and missing transverse momentum (\( E_{\text{T}}^{\text{miss}} \)). Calorimeter signals contribute to all objects, except for muons. The topo-clusters introduced in detail in Sect. 3 are primarily used for the reconstruction of isolated hadrons, jets and \( E_{\text{T}}^{\text{miss}} \).

Jets are reconstructed using topo-clusters, with their energies either reconstructed on the basic (electromagnetic) scale presented in Sect. 3.2, or on the fully calibrated and corrected (hadronic) scale described in Sect. 5.

Additional refinement of the jet energy scale (JES) may include reconstructed charged-particle tracks from the ID. More details of jet reconstruction and calibration can be found in Refs. [16,38].

Jets used in the studies presented here are reconstructed in data and MC simulations using the anti-\( k_T \) jet algorithm [39] as implemented in the FASTJET package [40]. The jet size is defined by the radius parameter \( R \) in the jet algorithm, where both \( R = 0.4 \) and \( R = 0.6 \) are used. Full four-momentum recombination is used, restricting the input topo-cluster signals to be positive for a meaningful jet formation. The jets are fully calibrated and corrected after formation, including a correction for pile-up signal contributions. For 2012, the pile-up correction employs the reconstructed median transverse momentum density in the event and the area of the jet to subtract the \( p_T \) contribution from pile-up, following the sug-

\(^3\) Stable particles are those with laboratory frame lifetimes \( \tau \) defined by \( c\tau > 10 \text{ mm} \).
gestions in Ref. [41]. In addition, an MC simulation-based residual correction is applied [42].

3 Topological cluster formation and features

The collection of the calorimeter signals of a given collision event into clusters of topologically connected cell signals is an attempt to extract the significant signal from a background of electronic noise and other sources of fluctuations such as pile-up. This strategy is most effective in a highly granular calorimeter system such as the one employed by ATLAS. Finely segmented lateral read-out together with longitudinal sampling layers allows the resolution of energy-flow structures generating these spatial signal patterns, thus retaining only signals important for particle and jet reconstruction while efficiently removing insignificant signals induced by noise. The signal extraction is guided by reconstructing three-dimensional “energy blobs” from particle showers in the active calorimeter volume. Individual topo-clusters are not solely expected to contain the entire response to a single particle all of the time. Rather, depending on the incoming particle types, energies, spatial separations and cell signal formation, individual topo-clusters represent the full or fractional response to a single particle (full shower or shower fragment), the merged response of several particles, or a combination of merged full and partial showers.

3.1 Topo-cluster formation

The collection of calorimeter cell signals into topo-clusters follows spatial signal-significance patterns generated by particle showers. The basic observable controlling this cluster formation is the cell signal significance \( \varsigma_{\text{cell}}^{\text{EM}} \), which is defined as the ratio of the cell signal to the average (expected) noise \( \sigma_{\text{noise,cell}}^{\text{EM}} \) in this cell, as estimated for each run year according to Eq. (1) (with \( \sigma_{\text{noise,cell}}^{\text{EM}} = \sigma_{\text{noise}} \)).

\[
\varsigma_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}.
\] (2)

Both the cell signal \( E_{\text{cell}}^{\text{EM}} \) and \( \sigma_{\text{noise,cell}}^{\text{EM}} \) are measured on the electromagnetic (EM) energy scale. This scale reconstructs the energy deposited by electrons and photons correctly but does not include any corrections for the loss of signal for hadrons due to the non-compensating character of the ATLAS calorimeters.

Topo-clusters are formed by a growing-volume algorithm starting from a calorimeter cell with a highly significant seed signal. The seeding, growth, and boundary features of topo-clusters are in this algorithm controlled by the three respective parameters \( S, N, P \), which define signal thresholds in terms of \( \varsigma_{\text{cell}}^{\text{EM}} \) and thus apply selections based on \( \varsigma_{\text{cell}}^{\text{EM}} \) from Eq. (2).

\[
\left| E_{\text{cell}}^{\text{EM}} \right| > S \sigma_{\text{noise,cell}}^{\text{EM}} \Rightarrow \left| \varsigma_{\text{cell}}^{\text{EM}} \right| > S
\]

(primary seed threshold, default \( S = 4 \));

\[
\left| E_{\text{cell}}^{\text{EM}} \right| > N \sigma_{\text{noise,cell}}^{\text{EM}} \Rightarrow \left| \varsigma_{\text{cell}}^{\text{EM}} \right| > N
\]

(threshold for growth control, default \( N = 2 \));

\[
\left| E_{\text{cell}}^{\text{EM}} \right| > P \sigma_{\text{noise,cell}}^{\text{EM}} \Rightarrow \left| \varsigma_{\text{cell}}^{\text{EM}} \right| > P
\]

(principal cell filter, default \( P = 0 \)).

Useful configurations employ a \( S > N \geq P \) rule, as reflected in the default configuration for ATLAS indicated above. The default values are derived from optimisations of the response and the relative energy resolution for charged pions in test-beam experiments using ATLAS calorimeter prototypes [43].

3.1.1 Collecting cells into topo-clusters

Topo-cluster formation is a sequence of seed and collect steps, which are repeated until all topologically connected cells passing the criteria given in Eqs. (3) and (4) and their direct neighbours satisfying the condition in Eq. (5) are found. The algorithm starts by selecting all cells with signal significances \( \varsigma_{\text{cell}}^{\text{EM}} \) passing the threshold defined by \( S \) in Eq. (3) from calorimeter regions which are allowed to seed clusters. These seed cells are then ordered in decreasing \( \varsigma_{\text{cell}}^{\text{EM}} \).

Each seed cell forms a proto-cluster. The cells neighbouring a seed and satisfying Eqs. (4) or (5) are collected into the corresponding proto-cluster. Here neighbouring is generally defined as two calorimeter cells being directly adjacent in a given sampling layer, or, if in adjacent layers, having at least partial overlap in the \((\eta, \phi)\) plane. This means that the cell collection for topo-clusters can span modules within the same calorimeter as well as calorimeter sub-detector transition regions. Should a neighbouring cell have a signal significance passing the threshold defined by the parameter \( N \) in Eq. (4), its neighbours are collected into the proto-cluster as well. If a particular neighbour is a seed cell passing the threshold \( S \) defined in Eq. (3), the two proto-clusters are merged. If a neighbouring cell is attached to two different proto-clusters and its signal significance is above the threshold defined by \( N \), the two proto-clusters are merged. This procedure is iteratively applied to further neighbours until the last set of neighbouring cells with significances passing

\( \varsigma_{\text{cell}}^{\text{EM}} \).  

4 Calorimeter cells marked as having read-out or general signal extraction problems in the actual run conditions are not considered as seeds.
the threshold defined by $P$ in Eq. (5), but not the one in Eq. (4), is collected. At this point the formation stops.

The resulting proto-cluster is characterised by a core of cells with highly significant signals. This core is surrounded by an envelope of cells with less significant signals. The configuration optimised for ATLAS hadronic final-state reconstruction is $S = 4$, $N = 2$, and $P = 0$, as indicated in Eqs. (3) to (5). This particular configuration with $P = 0$ means that any cell neighbouring a cell with signal significance passing the threshold given by $N$ in Eq. (4) is collected into a proto-cluster, independent of its signal. Using the correlations between energies in adjacent cells in this way allows the retention of cells with signals that are close to the noise levels while preserving the noise suppression feature of the clustering algorithm.

The implicit noise suppression implemented by the topo-cluster algorithm discussed above leads to significant improvements in various aspects of the calorimeter performance, such as the energy and spatial resolutions in the presence of pile-up. Contributions from large negative and positive signal fluctuations introduced by pile-up can survive in a given event, though, and thus contribute to the sensitivity to pile-up observed in e.g. the jet response [42], in addition to the cell-level effects mentioned in Sect. 2.2.1. Examples of the effect of this noise suppression on the cells contributing to zero-bias events recorded with ATLAS in 2012 are shown in the cell signal-significance spectra in Fig. 5a–d for four different L$\alpha$T calorimeters in ATLAS.

3.1.2 Treatment of negative cell signals

Negative cell signals in the ATLAS calorimeters are the result of fluctuations introduced predominantly by pile-up and, to a lesser extent, by electronic noise, as discussed in Sects. 2.2.1 and 2.2.2. The thresholds in Eqs. (3)–(5) are applied in terms of the absolute value of $\varsigma_{\text{cell}}^{\text{EM}}$. This means that not only large positive cell signals can seed a cluster, but also those with large negative signals. In addition, cells with neg-
ative signals can also contribute to the cluster growth control and are added to the envelope around the topo-cluster core.

The use of cells with $E_{\text{EM}}^{\text{cell}} < 0$ as topo-cluster seeds provides a diagnostic tool for the amount of noise in the overall calorimeter signal for a given event. At the fixed noise value given in Eq. (1) and used in Eq. (3), the luminosity-dependent actual noise in the event is reflected in the number of topo-clusters reconstructed with negative seeds. This number serves as an estimator mainly for out-of-time pile-up.

Topo-clusters with negative seeds often have a total energy $E_{\text{clus}}^{\text{EM}} < 0$ as well, especially when $|E_{\text{cell}}^{\text{EM}}| \gg P$. This is due to the dominance of the negative seed and the correlation between this seed signal and signals in the neighbouring cells, which likely also have $E_{\text{EM}}^{\text{cell}} < 0$. If a negative seed signal is generated by out-of-time pile-up, it is induced by a particle injected into the calorimeter more than 100 ns before the event. Its residual signal trace is scaled by the negative undershoot of the shaping function shown in Fig. 3. This particle also injected significant energy in the neighbouring cells at the same time, due to its electromagnetic or hadronic shower, which leads to $E_{\text{EM}}^{\text{cell}} < 0$ in these cells at the time of the event. For the same reasons, topo-clusters from out-of-time pile-up seeded by $E_{\text{cell}}^{\text{EM}} > 0$ often yield $E_{\text{clus}}^{\text{EM}} > 0$, because they are typically generated by particles injected in past bunch crossings closer in time (within 100 ns). The topo-clusters with $E_{\text{EM}}^{\text{clus}} < 0$ can be used to provide an average global cancellation of contributions of clusters seeded by positive fluctuations in out-of-time pile-up in full event observables including $E_T^{\text{miss}}$ [44].

Clustering cells with $E_{\text{EM}}^{\text{cell}} < 0$ in any topo-cluster, including those containing and seeded by large positive signals, improves noise suppression due to the local cancellation of random positive (upward) noise fluctuations by negative (downward) fluctuations within this cluster. Allowing only positive signals to contribute introduces a bias in the cluster signal, while the random cancellation partially suppresses this bias.

To reconstruct physics objects such as jets from topo-clusters, only those clusters with a net energy $E_{\text{EM}}^{\text{clus}} > 0$ are considered. The expectation is that clusters with net negative energy have no contribution to the signal of the reconstructed object, as there is no correlation of the corresponding downward fluctuation mainly induced by the energy flow in previous bunch crossings with the final state that is triggered and reconstructed.

3.1.3 Cluster splitting

The proto-clusters built as described in Sect. 3.1.1 can be too large to provide a good measurement of the energy flow from the particles generated in the recorded event. This is true because spatial signal structures inside those clusters are not explicitly taken into account in the formation. In particular, local signal maxima indicate the presence of two or more particles injecting energy into the calorimeter in close proximity.

To avoid biases in jet-finding and to support detailed jet substructure analysis as well as a high-quality $E_T^{\text{miss}}$ reconstruction, proto-clusters with two or more local maxima are split between the corresponding signal peaks in all three spatial dimensions. A local signal maximum is defined by $E_{\text{EM}}^{\text{clus}} > 500 \text{MeV}$, in addition to the topological requirements for this cell to have at least four neighbours and that none of the neighbours has a larger signal. Also, the location of cells providing local maxima is restricted to cells in the EM sampling layers EMB2, EMB3, EME2 and EME3, and to FCAL0. This means that for a proto-cluster located completely inside the electromagnetic calorimeters, or extending from the electromagnetic to the hadronic calorimeters, splitting is guided by the spatial cell signal distributions in the highly granular electromagnetic calorimeters. The cluster splitting is refined in an additional step, where signal maxima can be provided by cells from the thin EM sampling layers EMB1 and EME1 with a highly granular $\eta$-strip read-out geometry, all sampling layers in the hadronic calorimeters (HEC0 to HEC3, Tile0 to Tile2), and the hadronic forward calorimeter modules FCAL1 and FCAL2. The use of EMB1 and EME1 in the topo-cluster splitting improves the photon separation in $\pi^0 \rightarrow \gamma \gamma$.

The cluster splitting algorithm can find cells which are neighbours to two or more signal maxima. In this case, the cell is assigned to the two highest-energy clusters after splitting of the original topo-cluster it is associated with. This means that each cell is only shared once at most, and, even then, is never shared between more than two clusters.

The sharing of its signal between the two clusters with respective energies $E_{\text{EM}}^{\text{clus},1}$ and $E_{\text{EM}}^{\text{clus},2}$ is expressed in terms of two geometrical weights $w_{\text{clus},1}^{\text{geo}}$ and $w_{\text{clus},2}^{\text{geo}}$. These weights are calculated from the distances of the cell to the centres of the two clusters $(d_1, d_2)$, measured in units of a typical electromagnetic shower size scale in the ATLAS calorimeters, and the cluster energies,

$$w_{\text{clus},1}^{\text{geo}} = \frac{E_{\text{EM}}^{\text{clus},1}}{E_{\text{EM}}^{\text{clus},1} + r E_{\text{EM}}^{\text{clus},2}},$$

5 Signals in the pre-samplers and gap scintillators are not considered at all in guiding the topo-cluster splitting (see Ref. [1] for a detailed description of the ATLAS calorimeters).

6 This scale is motivated by the Molière radius of the electromagnetic shower, which in good approximation is set to 5 cm for all calorimeters.
Fig. 6 Stages of topo-cluster formation in the first module (FCAL0) of the FCAL calorimeter for a simulated dijet event with at least one jet entering this calorimeter. Shown in a are cells with signal significance $|\varsigma_{\text{cell}}^{\text{EM}}| > 4$ that can seed topo-clusters, in b cells with $|\varsigma_{\text{cell}}^{\text{EM}}| > 2$ controlling the topo-cluster growth, and in c all clustered cells and the outline of topo-clusters and topo-cluster fragments in this module. All clusters shown in c which do not contain a seed cell from this module are seeded in other modules of the FCAL, or in other calorimeters surrounding it. Pile-up is not included in this simulation, but electronic noise is modelled. Cells not colour coded but inside a topo-cluster have a negative signal, while cells shaded grey are completely surrounded by clustered cells but not part of a topo-cluster themselves. The cell and cluster boundaries are displayed on a dimensionless grid using the polar angle $\theta$ and the azimuthal angle $\phi$. This view maintains the cell shapes and proportions. For the definition of the cell signal significance $\varsigma_{\text{cell}}^{\text{EM}}$ see Eq. (2).

Figure 6 shows an example of topo-clusters generated by an MC simulated jet in the first module of the ATLAS forward calorimeter under 2010 run conditions (no pile-up). Possible seed cells, as defined in Eq. (3), are shown in Fig. 6a. Cells with signal significances above the threshold $N$ specified in Eq. (4) are displayed in Fig. 6b. The cells from this module included in any topo-cluster are shown in Fig. 6c. This display shows the effectiveness of cluster splitting in tracing signal structures. Comparing Figs. 6a and c clearly shows

$$w_{\text{cell},2}^{\text{geo}} = 1 - w_{\text{cell},1}^{\text{geo}},$$

$$r = \exp(d_1 - d_2).$$

The geometrical weights reflect the splitting rule that each cell can only appear in two proto-clusters at most, as $w_{\text{cell},1}^{\text{geo}} + w_{\text{cell},2}^{\text{geo}} = 1$. After splitting, the final proto-clusters are the topo-clusters used for further reconstruction of the recorded or simulated final state.
the survival of cells with $|\frac{S}{\text{cell}}| < 2$ in the vicinity of more significant signals, even if those are not in the same module (or sampling layer).

### 3.1.4 Cluster multiplicities in electromagnetic and hadronic showers

One of the original motivations behind any cell clustering is to reconstruct single-particle showers with the highest possible precision in terms of energy and shape. The immediate expectation is that the clustering algorithm should be very efficient in reconstructing one cluster for each particle entering the calorimeter. While this view is appropriate for dense and highly compact electromagnetic showers with relatively small shower-to-shower fluctuations in their longitudinal (along the direction of flight of the incoming particle) and lateral (perpendicular to the direction of flight) extensions, hadronic showers are subject to much larger intrinsic fluctuations leading to large shower-to-shower variations in their shapes and compactness. Hadrons generated in inelastic interactions in the course of the hadronic shower can even travel significant distances and generate sub-showers outside the direct neighbourhood of the calorimeter cell containing the initial hadronic interaction. This means that topo-clusters can contain only a fraction of the hadronic shower.

The distributions of the topo-cluster multiplicity $N_{\text{clus}}$ for single particles which primarily generate electromagnetic showers ($\pi^0$) and hadronic showers ($\pi^-$) in the central (barrel) calorimeter region are shown in Fig. 7a. The dependence of the average $N_{\text{clus}}$ on the pseudorapidity $\eta$ is displayed in Fig. 7b.

Neutral pions with $E_{\pi^0} = 100$ GeV injected into the detector at a fixed direction often generate only one topo-cluster from largely overlapping electromagnetic showers, as the angular distance between the two photons from $\pi^0 \rightarrow \gamma \gamma$ is small. This is demonstrated by the $N_{\text{clus}}$ distribution for topo-clusters generated by $\pi^0$ at $|\eta| = 0.3$ in ATLAS in Fig. 7a peaking at $N_{\text{clus}} = 1$, with a probability only slightly larger than the one for $N_{\text{clus}} = 2$. In the latter case the two topo-clusters from the $\pi^0$ are generated by (1) resolving the two photon-induced showers, (2) a possible residual imperfect signal collection and proto-cluster splitting in the topo-cluster algorithm, or by (3) accidental inclusion of additional topo-cluster(s) generated by electronic noise. While the particular reason for the second cluster depends on effects introduced by local features including the calorimeter readout granularity and cell noise levels at a given direction $\eta$, hypothesis (1) is found to be least likely as it is observed that the energy sharing between the two topo-clusters is typically very asymmetric. The leading topo-cluster generated by $\pi^0$ at 100 GeV contains very close to 100% of the total energy in this calorimeter region, indicating that the second and any further topo-clusters arise from hypotheses (2) and (3).

Figure 7b shows the average $N_{\text{clus}}$ as a function of the generated particle direction $\eta = \eta_{\text{gen}}$. Especially around transition regions at $|\eta| \approx 1.4$ (central to end-cap calorimeters) and $|\eta| \approx 3.2$ (end-cap to forward calorimeters), which both have reduced calorimetric coverage, $N_{\text{clus}}$ can significantly increase due to reduction or loss of the core signal of the showers.

The number of clusters generated by $\pi^-$ with $E = 100$ GeV injected at $\eta = 0.3$ peaks at $N_{\text{clus}} = 3$ and has a more significant tail to higher multiplicities, as shown in Fig. 7a. This is expected for hadronic showers, where the distance between two inelastic interactions with significant energy release is of the order of the nuclear interaction length
\( \lambda_{\text{nuclei}} \), typically \( O(10\, \text{cm}) \). This can lead to several well-separated topo-clusters. For example, at 100 GeV incident energy the leading topo-cluster generated by \( \pi^- \) contains on average 85 GeV, while the next-to-leading topo-cluster contains about 10 GeV on average. The remaining energy is distributed among one or more low-energy topo-clusters.

The wider hadronic shower spread introduces a higher sensitivity of \( N_{\text{clus}} \) to the calorimeter read-out granularities and transition regions, as can be seen in Fig. 7b. The transition regions at \(| \eta | \approx 0.8–1.0, | \phi | \approx 1.4 \) and \(| \eta | \approx 3.2 \) affect the topo-cluster formation more than in the case of electromagnetic showers, not only in terms of the peak \( N_{\text{clus}} \) but also in terms of the range in \( \eta \). In particular the region around \(| \eta | \approx 0.8–1.0 \) has a larger effect on \( N_{\text{clus}} \) for hadrons than for electromagnetic interacting particles, as this is the transition from the central to the extended Tile calorimeter introducing reduced calorimetric coverage for hadrons. The central electromagnetic calorimeter provides hermetic coverage here, without any effect on \( N_{\text{clus}} \). The sharp drop of \( N_{\text{clus}} \) for \( \pi^- \) at \(| \eta | = 2.5 \) corresponds to the reduction in calorimeter cell granularity by a factor of approximately four.

### 3.2 Cluster kinematics

The cluster kinematics are the result of the recombination of cell energies and directions. The presence of cells with \( E_{\text{cell}} < 0 \) requires a special recombination scheme to avoid directional biases.

The cluster directions are calculated as signal-weighted barycentres \((\eta_{\text{clus}}, \phi_{\text{clus}})\). Using \( E_{\text{cell}} < 0 \) in this scheme leads to distortion of these directions, even projecting them into the wrong hemispheres. Ignoring the contribution of cells with negative signals, on the other hand, biases the cluster directions with contributions from upward noise fluctuations. To avoid both effects, the cluster directions are calculated with absolute signal weights \( |E_{\text{clus}}| \),

\[
\eta_{\text{clus}} = \frac{\sum_{i=1}^{N_{\text{cell}}} u_{\text{cell,}i}^{\text{geo}} \cdot |E_{\text{cell,}i}^{\text{EM}}| \cdot \eta_{\text{cell,}i}}{\sum_{i=1}^{N_{\text{cell}}} u_{\text{cell,}i}^{\text{geo}} \cdot |E_{\text{cell,}i}^{\text{EM}}|}
\]

\[
\phi_{\text{clus}} = \frac{\sum_{i=1}^{N_{\text{cell}}} u_{\text{cell,}i}^{\text{geo}} \cdot |E_{\text{cell,}i}^{\text{EM}}| \cdot \phi_{\text{cell,}i}}{\sum_{i=1}^{N_{\text{cell}}} u_{\text{cell,}i}^{\text{geo}} \cdot |E_{\text{cell,}i}^{\text{EM}}|}
\]

Here \( N_{\text{cell}} \) is the number of cells in the cluster, and \( u_{\text{cell,}i}^{\text{geo}} \) are the geometrical signal weights introduced by cluster splitting, as given in Eqs. (6)–(8) in Sect. 3.1.3. The direction of each cell is given by \((\eta_{\text{cell}}, \phi_{\text{cell}})\), calculated from its location with respect to the centre of ATLAS at \((x = 0, y = 0, z = 0)\) in the detector reference frame. The cluster directions are therefore reconstructed with respect to this nominal detector centre.

The total cluster signal amplitude \( E_{\text{clus}}^{\text{EM}} \) reflects the correct signal contributions from all cells,

\[
E_{\text{clus}}^{\text{EM}} = \sum_{i=1}^{N_{\text{cell}}} u_{\text{cell,}i}^{\text{geo}} E_{\text{cell,}i}^{\text{EM}},
\]

and is calculated using the signed cell signals \( E_{\text{cell,}i}^{\text{EM}} \) and taking into account the geometrical signal weights. In general, all clusters with \( E_{\text{clus}}^{\text{EM}} > 0 \) are used for the reconstruction of physics objects in the ATLAS calorimeters, including the very few ones seeded by cell signals \( E_{\text{cell}}^{\text{EM}} \).

Each topo-cluster is interpreted as a massless pseudo-particle in physics object reconstruction. The energy and momentum components on the EM scale are calculated from the basic reconstructed kinematic variables \((E_{\text{clus}}^{\text{EM}}, \eta_{\text{clus}}, \phi_{\text{clus}})\) as

\[
f_{\text{clus}}^{\text{EM}} = E_{\text{clus}}^{\text{EM}} \cdot (1, \sin \theta_{\text{clus}} \cos \phi_{\text{clus}}, \sin \theta_{\text{clus}} \sin \phi_{\text{clus}}, \cos \theta_{\text{clus}})
\]

\[
= \left( E_{\text{clus}}^{\text{EM}}, p_{\text{clus}}^{\text{EM}} \right)
\]

with terms involving \( \theta_{\text{clus}} \), the polar angle calculated from \( \eta_{\text{clus}} \), and \( \phi_{\text{clus}} \).

The massless pseudo-particle interpretation is appropriate as there is no physically meaningful cluster mass without a specific and valid particle hypothesis for the origin of the signal. Such a hypothesis seems to be impossible to obtain from the calorimeter signals alone, especially for hadrons or hadronically decaying particles, where particle identification often requires a measurement of the charge. A topo-cluster mass could in principle be reconstructed from the cell signals and their spatial distribution, but this observable is dominated by lateral shower spreading, which does not represent a physically meaningful mass. It is also highly affected by the settings for the noise thresholds, which control the lateral and longitudinal spread of the cluster in a given pile-up environment (see Sect. 3.1.1).

In addition, hadronic showers tend to be split more often into two or more topo-clusters, as discussed in Sect. 3.1.4 for single particles. Also, it is very likely in the proton–proton collision environment at the LHC that a given topo-cluster contains signals from several particles, especially when located inside a jet, as a mix of electromagnetic and hadronic showers or shower fragments. These issues make a physical particle hypothesis very unlikely, and any cluster mass measurement would be very hard to interpret or validate in relation to a “real” particle.

### 4 Topo-cluster moments

The shape of a topo-cluster and its internal signal distribution contain valuable information for signal characterisation with
4.1 Geometrical moments

Each topo-cluster with at least three cells with $E_{\text{cell}}^{\text{EM}} > 0$ has a full set of geometrical moments. Simple directional moments (barycentres in $(\eta, \phi)$ space) and locations (centres of gravity) are available for all clusters. Not all geometrical moments can be evaluated in a meaningful way for all topo-clusters, mostly due to lack of relevant information in clusters with few cells. In this case, a default value specific to each of these moments is provided.

4.1.1 Location

The location of a topo-cluster is defined by its centre of gravity $\vec{c}$ in three-dimensional space, as shown in Fig. 8. This centre is calculated from the first moments of the three Cartesian coordinates specifying the calorimeter cell centres, following the definition given in Eq. (13). These locations are provided in the nominal detector frame of reference defined by the interaction point (IP) being located at $(x = 0, y = 0, z = 0)$.

In addition to the absolute location measured by the centre of gravity, the distance $\lambda_{\text{clus}}$ of the centre of gravity from the calorimeter front face, determined along the shower axis (see below and Fig. 8), is calculated for each topo-cluster.

4.1.2 Directions

The direction of a topo-cluster is given by $(\eta_{\text{clus}}, \phi_{\text{clus}})$, reconstructed as given in Eqs. (9) and (10). In addition, the first- and second-order directional moments using $\eta_{\text{cell}}$ and $\phi_{\text{cell}}$ are calculated using Eq. (13) with $n = 1$ and $n = 2$, respectively. The reference for these direction measures is the IP discussed above.

The shower axis is a measure of the direction of flight of the incoming particle. It is defined by a principal value analysis of the energy-weighted spatial correlations between cells with $E_{\text{cell}}^{\text{EM}} > 0$ with respect to the cluster centre in Cartesian coordinates,

$$C_{uv} = \frac{1}{\mathcal{W}} \sum_{i \mid E_{\text{cell},i}^{\text{EM}} > 0} \left( u_{\text{geo},i}^{\text{EM}} E_{\text{cell},i}^{\text{EM}} \right)^2 (u_i - \langle u \rangle) (v_i - \langle v \rangle).$$

(14)

with all permutations of $u, v \in \{x, y, z\}$. The normalisation $\mathcal{W}$ is given by

$$\mathcal{W} = \sum_{i \mid E_{\text{cell},i}^{\text{EM}} > 0} \left( u_{\text{geo},i}^{\text{EM}} E_{\text{cell},i}^{\text{EM}} \right)^2.$$

(15)

7 The first directional moment in $\eta(\phi)$ is only identical to $\eta_{\text{clus}}(\phi_{\text{clus}})$ for topo-clusters without negative signal cells, because negative signal cells are omitted from its calculation while they contribute to the $\eta_{\text{clus}}(\phi_{\text{clus}})$ reconstruction.
The $C_{uv}$ fill a symmetric $3 \times 3$ matrix $C = [C_{uv}]$. The eigenvector of $C$ closest to the direction $\vec{c}$ from the IP to the centre of gravity of the topo-cluster is taken to be the shower axis $\vec{s}$. If the angular distance $\Delta \alpha$ between $\vec{c}$ and $\vec{s}$ is $\Delta \alpha > 20^\circ$, $\vec{s}$ is used as the shower axis. Figure 8 depicts the geometry of the two axis definitions for topo-clusters.

4.1.3 Extensions and sizes

The size of the topo-cluster is calculated with respect to the shower axis $\vec{s}$ and the centre of gravity $\vec{c}$. For this, cells are first located with reference to $\vec{s}$ and $\vec{c}$. The distances of a cell at $\vec{x}_i$ to the shower axis and the centre of gravity are then given by

$$r_i \equiv |(\vec{x}_i - \vec{c}) \times \vec{s}|$$
(radial distance to shower axis); 

$$\lambda_i = (\vec{x}_i - \vec{c}) \cdot \vec{s}$$
(longitudinal distance from shower centre of gravity). 

The first moment $\langle \lambda \rangle$ calculated according to Eq. (13) with $\nu_{cell,i} = \lambda_i$ and $n = 1$ is $\langle \lambda \rangle = 0$ by definition. The same equation is used for the first moment ($r$) of $r_i$ ($\nu_{cell,i} = r_i$, $n = 1$). The longitudinal and lateral extensions of a topo-cluster can then respectively be measured in terms of the second moments $\langle \lambda^2 \rangle$ and $\langle r^2 \rangle^2$, again using Eq. (13), but with $n = 2$. Specifying cluster dimensions in this fashion describes a spheroid with two semi-axes of respective lengths $\sqrt{\langle \lambda^2 \rangle}$ and $\sqrt{\langle r^2 \rangle}$.

As calorimeter technologies and granularities change as function of $\eta$ in ATLAS, measures representing the lateral and longitudinal extension of topo-clusters in a more universal and normalised fashion are constructed. These measures are defined in terms of second moments with value ranges from 0 to 1,

$$m_{\text{lat}}^2 = \frac{\langle r^2 \rangle_{\text{out}}}{\langle r^2 \rangle_{\text{out}} + \langle r^2 \rangle_{\text{core}}} \quad \text{normalised lateral energy dispersion (width measure);}$$

$$m_{\text{long}}^2 = \frac{\langle \lambda^2 \rangle_{\text{out}}}{\langle \lambda^2 \rangle_{\text{out}} + \langle \lambda^2 \rangle_{\text{core}}} \quad \text{normalised longitudinal energy dispersion (length measure).}$$

The $\langle r^2 \rangle_{\text{out}}$ term is calculated using Eq. (13) with $n = 2$ and $\nu_{cell,i} = r_i$, but with $r_i = 0$ for the two most energetic cells in the cluster. The term $\langle r^2 \rangle_{\text{core}}$ is calculated with the same equation, but now with a fixed $r_i = r_{\text{core}}$ for the two most energetic cells, and $r_i = 0$ for the rest. The calculation of the corresponding terms $\langle \lambda^2 \rangle_{\text{out}}$ and $\langle \lambda^2 \rangle_{\text{core}}$ for $m_{\text{long}}^2$ follows the same respective rules, now with $\nu_{cell,i} = \lambda_i$ in Eq. (13) and $\lambda_{\text{core}}$ for the most energetic cells in $\langle \lambda^2 \rangle_{\text{core}}$.

The normalised moments $m_{\text{long}}^2$ and $m_{\text{lat}}^2$ do not directly provide a measure of spatial topo-cluster dimensions, rather they measure the energy dispersion in the cells belonging to the topo-cluster along the two principal cluster axes. Characteristic values are $m_{\text{long}}^2 \to 0$ ($m_{\text{lat}}^2 \to 0$) indicating few highly energetic cells distributed in close proximity along the longitudinal (lateral) cluster extension, and $m_{\text{long}}^2 \to 1$ ($m_{\text{lat}}^2 \to 1$) indicating a longitudinal (lateral) distribution of cells with more similar energies. Small values of $m_{\text{long}}^2$ ($m_{\text{lat}}^2$) therefore mean short (narrow) topo-clusters, while larger values are indicative of long (wide) clusters.

The effective size of the topo-cluster in $(\eta, \phi)$ space can in good approximation be estimated as

$$\sigma_{\eta} \simeq \sigma_{\phi} \simeq \sqrt{\frac{\sqrt{\langle r^2 \rangle}}{|\vec{c}|}} \times \cosh(\eta).$$

The fact that this approximation holds for both the cluster size in $\eta (\sigma_{\eta})$ and $\phi (\sigma_{\phi})$ is due to the particular granularity of the ATLAS calorimeters.

4.2 Signal moments

Topo-cluster related moments to the distribution of the cell signals inside the cluster are useful in determining the density and compactness of the underlying shower, the significance of the cluster signal itself, and the quality of the cluster reconstruction. These moments thus not only provide an important input to the calibrations and corrections discussed in Sect. 5, but also support data quality driven selections in the reconstruction of physics objects. Additional topo-cluster signal quality moments related to instantaneous, short term, and long term detector defects introducing signal efficiency losses are available but very technical in nature, and very specific to the ATLAS calorimeters. Their discussion is outside of the scope of this paper.

4.2.1 Signal significance

The significance of the topo-cluster signal is an important measure of the relevance of a given cluster contribution to
the reconstruction of physics objects. Similar to the cell signal
significance $\varsigma_{\text{cell}}^{\text{EM}}$ given in Eq. (2) in Sect. 3.1, it is measured
with respect to the total noise $\sigma_{\text{noise,clus}}^{\text{EM}}$ in the topo-cluster.

The definition of $\sigma_{\text{noise,clus}}^{\text{EM}}$ assumes incoherent noise in the
cells contributing to the topo-cluster:

$$\sigma_{\text{noise,clus}}^{\text{EM}} = \sqrt{\sum_{i=1}^{N_{\text{cell}}} (\sigma_{\text{noise,cell},i}^{\text{EM}})^2}.$$  (21)

Here $N_{\text{cell}}$ is the number of cells forming the cluster, including
the ones with $E_{\text{cell}}^{\text{EM}} < 0$. As discussed in Sect. 2.2.2, the
individual overall cell noise $\sigma_{\text{noise,cell},i}^{\text{EM}}$ is set according to the
nominal pile-up condition for a given data taking period. The
topo-cluster signal significance $\varsigma_{\text{clus}}^{\text{EM}}$ is then measured using
$\sigma_{\text{noise,clus}}^{\text{EM}}$ and $E_{\text{clus}}^{\text{EM}}$:

$$\varsigma_{\text{clus}}^{\text{EM}} = \frac{E_{\text{clus}}^{\text{EM}}}{\sigma_{\text{noise,clus}}^{\text{EM}}}.$$  (22)

In addition to $\varsigma_{\text{clus}}^{\text{EM}}$, $\varsigma_{\text{cell}}^{\text{EM}}$ of the cell with the highest signif-
ificant signal (the original cluster seed) is available to further evaluate the topo-cluster. A highly significant seed is a strong
indication of an important cluster signal, even if $\varsigma_{\text{clus}}^{\text{EM}}$ may be reduced by inclusion of a larger number of less significant
cell signals.

4.2.2 Signal density

The signal density of the topo-cluster is indicative of the
nature of the underlying particle shower. It can be evaluated in two different approaches. First, $E_{\text{clus}}^{\text{EM}}$ can be divided by the
volume the cluster occupies in the calorimeter. This volume is the sum of volumes of all cells contributing to the cluster. The signal density reconstructed this way is subject to con-
siderable instabilities introduced by signal fluctuations from noise, as large volume cells can be added with a very small signal due to those fluctuations.

The default for topo-cluster calibration is the second and more stable estimate of the topo-cluster signal density measured by the cell-energy-weighted first moment $\rho_{\text{clus}} = \langle \rho_{\text{cell}} \rangle$ of the signal densities $\rho_{\text{cell},i} = E_{\text{cell},i}^{\text{EM}} / V_{\text{cell},i}$ of cells
$i = 1 \ldots N_{\text{cell}}$ forming the cluster. Here $V_{\text{cell},i}$ is the volume of cell $i$. The $\rho_{\text{clus}}$ variable is calculated using Eq. (13) with $\nu_{\text{cell},i} = \rho_{\text{cell},i}^{\text{EM}}$ and $n = 1$. It is much less sensitive to the accidental inclusion of large volume cells with small signals

10 Out-of-time pile-up introduces a coherent component into the
calorimeter cell noise due to the correlation of signals in adjacent cells in showers generated by past energy flow. This contribution is reflected on average in the value for $\sigma_{\text{noise,cell}}^{\text{EM}}$, but cannot explicitly be evaluated for any given cell due to its highly stochastic and beam-conditions dependent nature.

4.2.3 Signal timing

The topo-cluster signal timing is a sensitive estimator of its
signal quality. It is particularly affected by large signal rem-
nants from previous bunch crossings contributing to the cluster,
or even exclusively forming it, and can thus be employed as a tag for topo-clusters indicating pile-up activity.

The reconstructed signal $E_{\text{clus}}^{\text{EM}}$ in all calorimeter cells in ATLAS is derived from the reconstruction of the peak amplitude of the time-sampled analogue signal from the calorimeter shaping amplifiers. In the course of this reconstruction the signal peaking time $t_{\text{clus}}^{\text{EM}}$ with respect to the 40 MHz LHC bunch crossing clock is determined as well. The timing $t_{\text{clus}}^{\text{EM}}$ of a topo-cluster is then calculated from $t_{\text{cell},i}$ of the clustered cells $i = 1 \ldots N_{\text{cell}}$ according to

$$t_{\text{clus}}^{\text{EM}} = \frac{\sum_{i=1}^{N_{\text{clus}}} (t_{\text{cell},i}^{\text{EM}} - \langle \nu_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}} \rangle)^2}{\sum_{i=1}^{N_{\text{clus}}} (t_{\text{cell},i}^{\text{EM}} - \langle \nu_{\text{cell},i}^{\text{geo}} E_{\text{cell},i}^{\text{EM}} \rangle)^2},$$  (23)

where only cells with a signal significance $\varsigma_{\text{cell},i}^{\text{EM}}$ sufficient to reconstruct $E_{\text{cell},i}^{\text{EM}}$ and $t_{\text{cell},i}$ are used ($\varsigma_{\text{cell},i}^{\text{EM}} > 2$). The particular weight of the contribution of $t_{\text{cell},i}$ to $t_{\text{clus}}^{\text{EM}}$ in Eq. (23) is found to optimise the cluster timing resolution [6].

4.2.4 Signal composition

The signal distribution inside a topo-cluster is measured in
terms of the energy sharing between the calorimeters con-
tributing cells to the cluster, and other variables measuring
the cell signal sharing. The energy sharing between the elec-
 tromagnetic and hadronic calorimeters is expressed in terms
of the signal ratio $f_{\text{em}}$, and can be used as one of the character-
istic observables indicating an underlying electromagnetic
shower. The signal fraction $f_{\text{max}}$ carried by the most ener-
gectic cell in the cluster is a measure of its compactness. The signal fraction $f_{\text{core}}$ of the summed signals from the high-
est energetic cell in each longitudinal calorimeter sampling
layer contributing to the topo-cluster can be considered as a
measure of its core signal strength. It is sensitive not only to
the shower nature but also to specific features of individual
hadronic showers. These fractions are calculated for each
topo-cluster with $E_{\text{clus}}^{\text{EM}} > 0$ as follows (EMC denotes the
electromagnetic calorimeters11 in ATLAS).

11 For the purpose of this calculation, the EMC consists of sampling
layers EMB1 to EMB3, EME1 to EME3, and FCAL0.
\[ f_{\text{emc}} = \frac{1}{N_{\text{clus}} \sum_{i \in \text{EM}}} w_{\text{cell,i}} E_{\text{EM},i} \]  

(EMC signal fraction in cluster);  

\[ f_{\max} = \frac{1}{N_{\text{clus}} \sum_{s \in \text{samplings}}} \max \left\{ w_{\text{cell,i}} E_{\text{EM},i} \right\} \]  

(most energetic cell signal fraction in cluster);  

\[ f_{\text{core}} = \frac{1}{N_{\text{clus}} \sum_{s \in \text{samplings}} \max \left\{ w_{\text{cell,i}} E_{\text{EM},i} \right\}} \]  

(core signal fraction in cluster).  

(26)

The index \( s \) steps through the set of calorimeter sampling layers, with cells contributing to the topo-cluster. Only cells with \( E_{\text{cell}} > 0 \) are used in the calculation of these fractions. Correspondingly, they are normalised to \( E_{\text{clus},\text{pos}} \) given by

\[ E_{\text{clus},\text{pos}} = \sum_{i \mid E_{\text{cell,i}} > 0} w_{\text{cell,i}} E_{\text{EM},i} \]  

(27)

All these moments have a value range of \([0,1]\).

One of the variables that can be considered for further evaluation of the relevance of the cluster signal in the presence of pile-up is the ratio of \( E_{\text{clus},\text{pos}} \) to \( E_{\text{clus}} \). It is sensitive to the negative energy content of a given topo-cluster which is largely injected by out-of-time pile-up dominated by the negative tail of the bipolar signal shaping function discussed in Sect. 3.1.2.

4.2.5 Topological isolation

The implicit noise suppression in the topological clustering algorithms leads to signal losses affecting the calorimeter response to particles, as further discussed in Sect. 5.4. As these signal losses appear at the boundary of the topo-cluster, corresponding corrections need to be sensitive to whether the lost signals may be included in another close-by cluster or if they are lost for good. This is particularly important for jets, where the topo-cluster density can be very high.

The degree of isolation is measured by the isolation moment \( f_{\text{iso}} \), with \( 0 \leq f_{\text{iso}} \leq 1 \). A topo-cluster with \( f_{\text{iso}} = 1 \) is completely isolated, while a cluster with \( f_{\text{iso}} = 0 \) is completely surrounded by others. The isolation measures the sampling layer energy \( (E_{s}^{\text{EM}}) \)-weighted fraction of non-clustered neighbour cells on the outer perimeter of the topo-cluster. Here \( E_{s}^{\text{EM}} \) is defined as the sum of the energies \( E_{\text{EM}} \) of all cells in a topo-cluster located in a given sampling layer \( s \) of the calorimeter.

The isolation moment is reconstructed by first counting the number of calorimeter cells \( N_{\text{clus},s} \) in sampling layer \( s \) neighbouring a topo-cluster but not collected into one themselves. Second, the ratio \( N_{\text{clus},s} / N_{\text{cell},s} \) of this number to the number of all neighbouring cells \( (N_{\text{cell},s})_{\text{neighbour}} \) for each \( s \) contributing to the cluster is calculated. The per-cluster \( E_{s}^{\text{EM}} \)-weighted average of these ratios from all included \( s \) is the isolation moment \( f_{\text{iso}} \).

\[ f_{\text{iso}} = \frac{\sum_{s \in \{ \text{samplings with } E_{s}^{\text{EM}} > 0 \}} E_{s}^{\text{EM}} N_{\text{clus},s} / N_{\text{cell},s}}{\sum_{s \in \{ \text{samplings with } E_{s}^{\text{EM}} > 0 \}} E_{s}^{\text{EM}}} \]  

(28)

5 Local hadronic calibration and signal corrections

The motivation for the calibration scheme described in this section arises from the intention to provide a calorimeter signal for physics object reconstruction in ATLAS which is calibrated outside any particular assumption about the kind of object. This is of particular importance for final-state objects with a significant hadronic signal content, such as jets and, to a lesser degree, \( \tau \)-leptons. In addition to these discrete objects, the precise reconstruction of the missing transverse momentum requires well-calibrated hadronic signals even outside hard final-state objects, to e.g. avoid deterioration of the \( E_{\text{miss}}^{\text{T}} \) resolution due to highly fluctuating (fake) \( p_{T} \)-imbalances introduced by the non-linear hadronic response on the EM scale.

The topo-cluster moments provide information sensitive to the nature of the shower generating the cluster signal. This information can be explored to apply moment-dependent calibrations cluster-by-cluster, and thus correct for the effects of the non-compensating calorimeter response to hadrons, accidental signal losses due to the clustering strategy, and energy lost in inactive material in the vicinity of the topo-cluster. The calibration strategy discussed in some detail in the following is local because it attempts to calibrate highly localised and relatively small (in transverse momentum flow space) topo-clusters.\(^{12} \) As the local hadronic calibration includes cell signal weighting, the calibration based on topo-clusters is referred to as “local hadronic cell weighting” (LCW) calibration.

All calibrations and corrections are derived using MC simulations of single pions (charged and neutral) at various energies in all ATLAS calorimeter regions. This fully simulation-based approach requires good agreement between data and these MC simulations for the topo-cluster signals and moments used for any of the applied corrections in terms of distribution shapes and averages. Reconstructed observables which are not well-modelled by simulation are not considered. The data/MC comparisons for most used observables

\(^{12} \) As cells and clusters are localised in the calorimeters, the preferred variables for this space are the azimuth \( \phi \) and the pseudorapidity \( \eta \), rather than the rapidity \( y \). As topo-clusters are reconstructed as massless pseudo-particles (see Sect. 3.2), \( y = \eta_{\text{clus}} \) for the complete object.
are shown in the context of the discussion of the methods using them.

5.1 General topo-cluster calibration strategy

The LCW calibration aims at the cluster-by-cluster reconstruction of the calorimeter signal on the appropriate (electromagnetic or hadronic) energy scale. In this, the cluster energy resolution is expected to improve by using other information in addition to the cluster signal in the calibration. The basic calorimeter signal inefficiencies that this calibration must address are given below.

Non-compensating calorimeter response: All calorimeters employed in ATLAS are non-compensating, meaning their signal for hadrons is smaller than the one for electrons and photons depositing the same energy \( (e/\pi > 1) \). Applying corrections to the signal locally so that \( e/\pi \) approaches unity on average improves the linearity of the response as well as the resolution for jets built from a mix of electromagnetic and hadronic signals. It also improves the reconstruction of full event observables such as \( E_T^{\text{miss}} \), which combines signals from the whole calorimeter system and requires balanced electromagnetic and hadronic responses in and outside signals from (hard) particles and jets.

Signal losses due to clustering: The topo-cluster formation applies an intrinsic noise suppression, as discussed in detail in Sect. 3.1. Depending on the pile-up conditions and the corresponding noise thresholds, a significant amount of true signal can be lost this way, in particular at the margins of the topo-cluster. This requires corrections to allow for a more uniform and linear calorimeter response.

Signal losses due to energy lost in inactive material: This correction is needed to address the limitations in the signal acceptance in active calorimeter regions due to energy losses in nearby inactive material in front, between, and inside the calorimeter modules.

The corrections collected in the LCW calibration address these three main sources of signal inefficiency. The specifics of the calibrations and corrections applied to correct for these signal inefficiencies depend on the nature of the energy deposit – hadronic (HAD) or electromagnetic (EM). Therefore, the first step of the topo-cluster calibration procedure is to determine the probability \( 0 \leq P_{\text{EM}}^{\text{clus}} \leq 1 \) that a given topo-cluster is generated by an electromagnetic shower. This approach provides straightforward dynamic scales (cluster-by-cluster) for the application of specific electromagnetic \( P_{\text{EM}}^{\text{clus}} \) and hadronic \( (1 - P_{\text{EM}}^{\text{clus}}) \) calibrations and corrections. For topo-clusters with \( P_{\text{EM}}^{\text{clus}} = 1 \), it suppresses the application of a hadronic calibration mostly addressing the non-compensating response to hadrons, and applies the electromagnetic-signal-specific corrections for the losses introduced by clustering and inactive material mentioned above. Reversely, very hadronic topo-clusters with \( P_{\text{EM}}^{\text{clus}} = 0 \) receive the appropriate hadronic calibration and hadronic-signal-specific signal loss corrections.

The main differences in the hadronic and electromagnetic calibration of topo-clusters are the magnitudes of the applied corrections, which in the EM case are significantly smaller than for HAD. Applying an exclusive categorisation based on the probability distributions described in Sect. 5.2 can lead to inconsistent calibrations especially for low-energy or small (few cells only) clusters, as misclassification for these kinds of topo-clusters is more likely than for clusters with higher energies or larger sizes. To allow for smooth transitions and reduce the dependency on the classification, the signal weights \( w_{\text{cell}}^{\text{cal}} \) applied to cell signals in the topo-cluster at any of the calibration and correction steps are calculated as

\[
w_{\text{cell}}^{\text{cal}} = P_{\text{EM}}^{\text{clus}} \cdot w_{\text{cell}}^{\text{em-cal}} + (1 - P_{\text{EM}}^{\text{clus}}) \cdot w_{\text{cell}}^{\text{had-cal}}.
\] (29)

The weights \( w_{\text{cell}}^{\text{em-cal}} \) and \( w_{\text{cell}}^{\text{had-cal}} \) represent the factors applied by the EM or HAD calibration to the cell signal. The effective representation of all calibration steps in terms of these cell-level signal weights implements a consistent approach independent of the nature of the actual correction applied at any given step. As detailed in Sects. 5.3–5.5, the weights can depend on the cell signal itself, thus yielding a different weight for each cell. They can also represent cluster-level corrections generating the same weight for all cells, or a subset of cells, of the topo-cluster. This cell weighting scheme therefore provides not only the corrected overall cluster energy after each calibration step by weighted cell signal re-summation, but also the corresponding (possibly modified) cluster barycentre. Thus the cumulative effect on the topo-cluster energy and direction can be validated after each step. The steps of the general LCW calibration are schematically summarised in Fig. 9, and the individual steps are described in more detail below.

The EM calibrations and corrections and their respective parameters are determined with single-particle MC simulations of neutral pions for a large set of energies distributed uniformly in terms of \( \log(E) \) between 200 MeV and 2 TeV, at various directions \( \eta \). The same energy and \( \eta \) phase space is used for the corresponding simulations of charged pions to determine the HAD calibrations and corrections. The signals in these simulations are reconstructed with thresholds corresponding to the nominal \( \sigma_{\text{noise}}^{\text{EM}} \) for a given run period, which reflects the pile-up conditions according to Eq. (1) in Sect. 2.2.2. Only electronic noise is added into the signal formation in the MC simulation, so that the derived calibrations and corrections effectively correct for signal losses.
introduced by the clustering itself. In particular, additional signal from pile-up and modifications of the true signal by out-of-time pile-up are not considered, as these are expected to cancel on average.

5.2 Cluster classification

As discussed in Sect. 4, most topo-clusters provide geometrical and signal moments sensitive to the nature of the shower producing the cluster signal. In particular, electromagnetic showers with their compact shower development, early starting point and relatively small intrinsic fluctuations can generate cluster characteristics very different from those generated by hadronic showers. The latter are in general subjected to larger shower-by-shower fluctuations in their development and can be located deeper into the calorimeter. In addition, the hadronic showers show larger variations of their starting point in the calorimeter. A classification of each topo-cluster according to its likely origin determines the most appropriate mix of EM and HAD calibration and correction functions to be applied.

The depth $\lambda_{\text{clus}}$ of the topo-cluster (Sect. 4.1.1) and its average cell signal density $\rho_{\text{clus}}$ (Sect. 4.2.2), both determined in bins of the cluster energy $E_{\text{clus}}^\text{EM}$ and the cluster direction $\eta_{\text{clus}}$, are found to be most efficient in classifying the topo-clusters. Using the MC simulations of single charged and neutral pions entering the calorimeters at various pseudorapidities and at various momenta, the probability for a cluster to be of electromagnetic origin ($P_{\text{EM}}^\text{clus}$) is then determined by measuring the efficiency for detecting an EM-like cluster in bins of four topo-cluster observables,

$$\mathcal{O}_{\text{clus}}^\text{class} = \left\{ E_{\text{clus}}^\text{EM}, \eta_{\text{clus}}, \log_{10}(\rho_{\text{clus}}/\rho_0) \right. \left. - \log_{10}(E_{\text{clus}}^\text{EM}/E_0), \log_{10}(\lambda_{\text{clus}}/\lambda_0) \right\}, \quad (30)$$

in this sequence mapped to bin indices $ijkl$ in the full accessible phase space. The density scale is $\rho_0 = 1 \text{ MeV mm}^{-3}$, the signal normalisation is $E_0 = 1 \text{ MeV}$, and longitudinal depth is measured in terms of $\lambda_0 = 1 \text{ mm}$. Here the density $\rho_{\text{clus}}$ is divided by the cluster signal $E_{\text{clus}}^\text{EM}$. This provides a necessary reference scale for its evaluation. As an absolute measure, $\rho_{\text{clus}}$ is less powerful in separating electromagnetic from hadronic energy deposits, as the same densities can be generated by electromagnetically and hadronically interacting particles of different incident energies.

The likelihood $P_{\text{EM}}^\text{clus}$ is defined in each bin $ijkl$ as

$$P_{\text{EM}}^\text{clus}(E_{\text{clus}}^\text{EM}, \eta_{\text{clus}}, \rho_{\text{clus}}/E_{\text{clus}}^\text{EM}, \lambda_{\text{clus}}) \mapsto P_{\text{clus},ijkl}^\text{EM} \quad (31)$$

with $0 \leq P_{\text{EM}}^\text{clus,ijkl} \leq 1$. The efficiencies $\epsilon_{ijkl}^{\pi^0(\pi^\pm)}$ are calculated as

$$\epsilon_{ijkl}^{\pi^0(\pi^\pm)} = \frac{N_{ijkl}^{\pi^0(\pi^\pm)}}{N_{ij}^{\pi^0(\pi^\pm)}}. \quad (32)$$

Here $N_{ijkl}^{\pi^0(\pi^\pm)}$ is the number of topo-clusters from $\pi^0 (\pi^\pm)$ in a given bin $ijkl$, while $N_{ij}^{\pi^0(\pi^\pm)}$ is the number of $\pi^0 (\pi^\pm)$ found in bin $ij$ of the $(E_{\text{clus}}^\text{EM}, \eta_{\text{clus}})$ phase space. On average there is no detectable difference in the development of $\pi^+$ and $\pi^-$ initiated hadronic showers affecting the topo-cluster formation. The distributions of the observables in $\mathcal{O}_{\text{clus}}^\text{class}$ as well as the correlations between them are the same. Therefore topo-clusters from $\pi^+$ and $\pi^-$ showers occupy the same bins in the $\mathcal{O}_{\text{clus}}^\text{class}$ phase space, yielding $N_{ijkl}^{\pi^+} = N_{ijkl}^{\pi^-} = N_{ijkl}^{\pi^0}$. In the definition of $P_{\text{EM}}^\text{clus}$ in Eq. (31). This normalisation reflects the use of all three pion charges at equal probability in MC...
magnetic and hadronic calorimeters in ATLAS. The average reflects the longitudinal segmentation of the electromagnetic in Figs. 11a and b. The specific structure of each distribution of the likelihood for clusters respectively classified as electromagnetic or hadronic, in 2010 data and MC simulations (no pile-up) shows a similar shape up to $E_{\text{clus}} \approx 2$ GeV. This energy range is dominated by topo-clusters from low-energy hadrons, in addition to clusters from less-energetic hadronic shower fragments created by the splitting algorithm described in Sect. 3.1.3. For $E_{\text{clus}} > 2$ GeV the average $\lambda_{\text{clus}}$ is increasingly dominated by higher-energy clusters produced by splitting and located in the electromagnetic calorimeter, thus pulling it to lower values. The rise of $\langle \lambda_{\text{clus}} \rangle$ for topo-clusters with $E_{\text{clus}} \gtrsim 10$ GeV reflects increasing contributions from energetic hadrons with dense showers generating high-energy clusters deeper in the hadronic calorimeter. The good agreement between data and MC simulations for both classes of topo-clusters supports the use of $\lambda_{\text{clus}}$ for the cluster classification derived from MC simulations for data [38].

5.3 Hadronic calibration

The hadronic calibration for topo-clusters attempts to correct for non-compensating calorimeter response, meaning to establish an average $e/\pi = 1$ for the cluster signal. The calibration reference is the locally deposited energy in the cells of a given topo-cluster, which is defined as the sum of all energies released by various shower processes in these cells. In each of the cells, the signal $E_{\text{clus}}^{\text{EM}}$ from this deposited energy $E_{\text{clus}}^{\text{dep}}$ is reconstructed on the electromagnetic energy scale. This yields cell signal weights defined as

$$w_{\text{cell}} = \frac{E_{\text{clus}}^{\text{dep}}}{E_{\text{clus}}^{\text{EM}}}.$$  (33)

In the case of electromagnetic signals, $w_{\text{cell}} = w_{\text{cell}}^{\text{em-cal}} \equiv 1$ by construction of the electromagnetic scale. In hadronic showers, $E_{\text{clus}}^{\text{dep}}$ has contributions from energy loss mechanisms which do not contribute to the signal, including nuclear binding energies and escaping energy carried by neutrinos. In this case, $w_{\text{cell}} = w_{\text{cell}}^{\text{had-cal}} \neq 1$ with $w_{\text{cell}}^{\text{had-cal}} > 1$ for hadronic inelastic interactions within the cell volume, and $w_{\text{cell}}^{\text{had-cal}} < 1$ for deposits by ionisations.\textsuperscript{13} The appropriate value of $w_{\text{cell}}^{\text{had-cal}}$ reflecting on average the energy loss mechanism generating $E_{\text{clus}}^{\text{EM}}$ in a given cell is determined by the hadronic calibration as a function of a set of observables $\rho_{\text{clus}}$ associated with the cell and the topo-cluster it belongs to. It is then applied to $E_{\text{clus}}^{\text{EM}}$ according to Eq. (29) in the signal reconstruction.

Simultaneously using all simulations of charged single pions for all energies and directions, lookup tables are considered, thus maintaining the correct isospin-preserving ratio.

For performance evaluation purposes, any topo-cluster with the set of observables $\rho_{\text{clus}}$ from Eq. (30) located in a bin $ijkl$ with $p_{\text{clus}}^{\text{EM}} \geq 0.5$ is classified as EM and with $p_{\text{clus}}^{\text{EM}} < 0.5$ is classified as HAD. In the rare case where a topo-cluster has too few cells or too little signal to meaningfully reconstruct the observables in $\rho_{\text{clus}}$, the cluster is likely generated by noise or insignificant energy deposits and is thus neither classified nor further corrected nor calibrated. An example of a $p_{\text{clus}}^{\text{EM}}$ distribution in a given phase space bin $ij$ is shown in Fig. 10. All distributions and their bin contents are accessed as lookup tables to find $p_{\text{clus}}^{\text{EM}}$ for a given cluster.

The distributions of $\lambda_{\text{clus}}$ for topo-clusters in jets reconstructed with the anti-$k_t$ algorithm with $R = 0.6$ are shown for clusters respectively classified as electromagnetic or hadronic, in 2010 data and MC simulations (no pile-up) in Figs. 11a and b. The specific structure of each distribution reflects the longitudinal segmentation of the electromagnetic and hadronic calorimeters in ATLAS. The average cluster depth $\langle \lambda_{\text{clus}} \rangle$ as a function of the cluster energy is shown in Figs. 11c and d for the same EM and HAD topo-clusters, respectively. The EM topo-clusters show the expected linear dependence of $\langle \lambda_{\text{clus}} \rangle$ on $E_{\text{clus}}^{\text{EM}}$ in Fig. 11c, with some modulations introduced by the read-out granularity of the EMC. The $\langle \lambda_{\text{clus}} \rangle$ dependence on $E_{\text{clus}}^{\text{EM}}$ shown for HAD topo-clusters in Fig. 11d features a similar shape up to $E_{\text{clus}} \approx 2$ GeV. This energy range is dominated by topo-clusters from low-energy hadrons, in addition to clusters from less-energetic hadronic shower fragments created by the splitting algorithm described in Sect. 3.1.3. For $E_{\text{clus}} > 2$ GeV the average $\lambda_{\text{clus}}$ is increasingly dominated by higher-energy clusters produced by splitting and located in the electromagnetic calorimeter, thus pulling it to lower values. The rise of $\langle \lambda_{\text{clus}} \rangle$ for topo-clusters with $E_{\text{clus}}^{\text{EM}} \gtrsim 10$ GeV reflects increasing contributions from energetic hadrons with dense showers generating high-energy clusters deeper in the hadronic calorimeter. The good agreement between data and MC simulations for both classes of topo-clusters supports the use of $\lambda_{\text{clus}}$ for the cluster classification derived from MC simulations for data [38].

\textsuperscript{13} This is because the electromagnetic energy scale reconstructs a signal larger than expected for the deposited energy in case of pure ionisation, due to the lack of showering.
structured from binned distributions relating $\Omega_{\text{had-cal}}$, defined as

$$
\Omega_{\text{had-cal}} = \{ S_{\text{calo}}, \eta_{\text{cell}}, \log_{10}(\rho_{\text{cell}}/\rho_0), \log_{10}(E_{\text{EM}}/E_0) \},
$$

to the hadronic signal calibration weight $w_{\text{had-cal}}$. The cell location is defined by one of the sampling layer identifiers $S_{\text{calo}}$ listed in Table 1 in Sect. 2.1 and the direction of the cell centre $\eta_{\text{cell}}$ extrapolated from the nominal detector centre of ATLAS. The cell signal density $\rho_{\text{cell}}$ is measured as discussed in Sect. 4.2.2, and $E_{\text{EM}}$ is the signal of the topo-cluster to which the cell contributes to. The lookup tables are binned in terms of $\Omega_{\text{had-cal}}$ such that $w_{\text{had-cal}}$ in each bin in the filled table is the average over all cells with observables fitting into this bin, with each contributing weight calculated as given in Eq. (33). These average weights are then retrieved for any cell in a topo-cluster as a function of $\Omega_{\text{had-cal}}$. The cluster signal and directions are re-summed as discussed in Sect. 5.6. The scales $\rho_0$ and $E_0$ in Eq. (34) are the same as the ones used in Eq. (30).

The $E_{\text{EM}}$ distribution in the PreSampler and the $\rho_{\text{cell}}$ distribution in the the EMB2 sampling of the central electromagnetic calorimeter are shown in Fig. 12 for cells in topo-clusters inside jets reconstructed with the anti-$k_t$ algorithm using a distance parameter $R = 0.6$. Discrepancies between data and MC simulations mostly in the high-end tails of the distributions indicate more compact electromagnetic showers in the simulation. This is also seen in Fig. 13a for the $\rho_{\text{cell}}$ distribution for the same kind of jets in the EME2 sampling of the electromagnetic end-cap calorimeter. Better agreement between data and MC simulations over the whole spectrum is observed for the $\rho_{\text{cell}}$ distributions in the first module (FCAL0) of the forward calorimeter shown in Fig. 13b, and in the second sampling of the central hadronic (Tile1) and the first sampling of the hadronic end-cap (HEC0) calorimeters shown in Fig. 14. Overall, the quality of the modelling of the cell signal densities is sufficient for topo-cluster calibration purposes. The figures are taken from Ref. [38].
5.4 Correction for out-of-cluster signal losses

In the process of applying the noise suppression described in Sect. 3.1, cells with small true deposited energy generated by EM or HAD showers may not be collected into a topo-cluster, either due to lack of significance of their small signal, or due to the absence of a neighbouring cell with a significant signal. The energy losses introduced by this effect are estimated using single-particle MC simulations. A corresponding out-of-cluster correction is determined and applied to nearby topo-clusters. The cells with true energy not included into clusters are referred to as lost cells.

The challenge in determining this correction is the assignment of the energy deposited in a lost cell to a certain cluster. As discussed in Sect. 3.1.3 and seen in Fig. 7, hadronic showers in particular can generate more than one topo-cluster. An algorithm defining an out-of-cluster neighbourhood to search for the lost cells has been developed for this assignment. This is depicted schematically in Fig. 15. The actual size of the neighbourhood for a given topo-cluster is determined by the maximum angular distance between the cluster and the lost cells. This distance depends on $\eta_{clus}$, and thus reflects granularity changes and shower size variations. It varies from approximately $\pi/3$ rad (60°) at $\eta_{clus} = 0$ to $7\pi/90$ rad (14°)
for $\eta_{\text{clus}} > 3.2$. The energy $E_{\text{clus}}^{\text{ooc}}$ deposited in all lost cells associated with a given topo-cluster is then used to derive the out-of-cluster correction factor $w_{\text{clus}, \text{ooc}}$:

$$w_{\text{clus}, \text{ooc}} = \frac{E_{\text{clus}}^{\text{ooc}} + E_{\text{clus}}^{\text{dep}}}{E_{\text{clus}}^{\text{dep}}}$$

and

$$E_{\text{clus}}^{\text{ooc}} = \sum_{i \in \{\text{lost cells}\}} E_{\text{cell, lost, } i}^{\text{dep}}.$$  

(35)

Here $E_{\text{clus}}^{\text{dep}}$ is the summed deposited energy of all cells inside the cluster. The out-of-cluster correction is a cluster-level correction featuring $w_{\text{clus}, \text{ooc}} > 1$.

Figure 15 shows that a lost cell can be located in the two overlapping out-of-cluster neighbourhoods of two close-by topo-clusters. In this case $E_{\text{cell, lost}}$ of this lost cell is assigned to both clusters, with a weight proportional to their respective deposited energies $E_{\text{clus}}^{\text{dep}}$. The out-of-cluster correction takes into account shared and non-shared lost cells and is derived for each of the two clusters separately using Eq. (35) with

$$E_{\text{clus}, 1(2)}^{\text{ooc}} = \sum_{i \in \{\text{lost cells}\}} E_{\text{cell, lost, } i}^{\text{dep}}$$

non-shared lost cells

$$+ \frac{E_{\text{clus}, 1(2)}^{\text{dep}}}{E_{\text{clus}, 1}^{\text{dep}} + E_{\text{clus}, 2}^{\text{dep}}} \sum_{j \in \{\text{lost cells}\}} E_{\text{cell, lost, } j}^{\text{dep}}.$$  

(36)

There are no spatial distance criteria applied to the sharing.

The scheme for the out-of-cluster correction ignores lost energy deposited in inactive areas of the detector, outside calorimeter cells. This effect is corrected for later in the calibration sequence (see Sect. 5.5) such that this component is not double-counted.

The out-of-cluster correction is different for electromagnetic and hadronic showers and is therefore separately determined with neutral and charged pion single-particle simulations. The three-dimensional set of observables $\Delta_{\text{clus}}^{\text{ooc}}$...
Fig. 16 The distribution of the isolation moment \( f_{\text{iso}} \) in a clusters classified as electromagnetic, and b clusters classified as hadronic. The average isolation \( \langle f_{\text{iso}} \rangle \) as a function of the cluster signal \( E_{\text{clus}}^{\text{EM}} \) is shown in c for electromagnetic and in d for hadronic topo-clusters. The figures are taken from Ref. [38].

\[
Q_{\text{clus}}^{\text{ooc}} = \left\{ \eta_{\text{clus}}, \log_{10}(E_{\text{clus}}^{\text{EM}}/E_0), \log_{10}(\lambda_{\text{clus}}/\lambda_0) \right\} \quad (37)
\]

is used to bin \( w_{\text{clus}}^{\text{ooc}} \). The weight is applied to the signal of nearly all cells of the topo-cluster receiving the out-of-cluster correction such that \( w_{\text{clus}}^{\text{ooc}} = w_{\text{clus}}^{\text{ooc}} \). The exceptions are cells located in the LAr pre-samplers PreSamplerB and PreSamplerE, and the Tile scintillators located between the barrel and end-cap cryostats, where \( w_{\text{clus}}^{\text{ooc}} = 1 \) always. The normalisations \( E_0 \) and \( \lambda_0 \) in Eq. (37) are the same as used in Eq. (30).

While the determination of the out-of-cluster correction depends on this assignment algorithm, the application of the correction is context dependent. A topo-cluster in a jet is likely to have directly neighbouring clusters which can capture its out-of-cluster signal loss. It is therefore expected that topo-clusters in jets need less out-of-cluster corrections than isolated topo-clusters away from other clusters. The degree of isolation is measured by the isolation moment \( f_{\text{iso}} \) introduced in Sect. 4.2.5. The out-of-cluster correction is effectively \( f_{\text{iso}} w_{\text{clus}}^{\text{ooc}} (Q_{\text{clus}}^{\text{ooc}}) \). This correction can change the barycentre and centre of gravity of topo-clusters containing cells from the LAr pre-samplers or the Tile scintillators.

Figure 16 shows \( f_{\text{iso}} \) for topo-clusters classified as either electromagnetic or hadronic in jets reconstructed with the anti-\( k_t \) algorithm and \( R = 0.6 \) [38]. A good agreement between data and MC simulations is observed, both for the details of the respective \( f_{\text{iso}} \) in Figs. 16a and b and the average as a function of \( E_{\text{clus}}^{\text{EM}} \) in Figs. 16c and d. The \( E_{\text{clus}}^{\text{EM}} \) dependence of \( f_{\text{iso}} \) is very similar for both kinds of topo-clusters.

The peak structure in the \( f_{\text{iso}} \) distributions shown in Figs. 16a and b is indicative of topo-clusters which have a large fraction of their energy in one sampling layer in the (regular) ATLAS calorimeter read-out segmentation with at least 16 cells around the perimeter of clustered cells in a sampling layer. The isolation of this layer then dominates the overall \( f_{\text{iso}} \), as given by Eq. (28) in Sect. 4.2.5. This dominance of just one sampling layer with the minimal number of cells is typical for topo-clusters seeded by a cell barely above the seed threshold defined in Eq. (3) and too little energy in the neighbouring samplings to further expand the cluster. Neighbouring cells then limit \( f_{\text{iso}} \) to the multiples...
of 1/16 visible in Figs. 16a and b. Even multiples of 1/16 occur more often than odd multiples since they can be produced more easily by topo-clusters with a different number of neighbours. The fact that clusters close to the noise threshold are mainly responsible for the peaks explains the mismatch between data and MC simulations observed in the peak heights, and points to non-perfect modelling of noise and very small signals. The overall structure of the $f_{iso}$ spectrum in data is well reproduced in terms of the peak locations by MC simulations.

5.5 Dead material corrections

Particles traversing the inactive (dead) material in front of or between calorimeter modules can deposit energy in it, thus reducing the measurable energy. This energy loss is addressed on average by the dead material correction. It is derived with single-particle MC simulations, where the deposited energy in the dead material outside of the active calorimeter can be calculated. This material is divided into virtual cells with a pointing geometry in ($\eta$, $\phi$). These cells are similar to the ATLAS calorimeter cells, but typically larger in size. Depending on the particle’s direction of flight, eight distinct regions are mapped out, as summarised in Table 2. The energy deposited in the dead material cells is determined for charged and neutral pions at various energies and directions, and almost everywhere correlated with measurable signals.

Table 2 Overview of the signals used to correct for dead material losses in the various regions around the ATLAS calorimeters. The numbered regions are shown in Fig. 17. The parameter values used for the dead material correction are extracted from lookup tables. Region 8 comprises all dead material volumes with energy loss outside regions 1–7. These are mostly small volumes located between and behind the active calorimeters.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Description</th>
<th>Cluster signals for dead material correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>In front of EMB</td>
<td>Energy in PreSampler B</td>
</tr>
<tr>
<td>2</td>
<td>Between EMB and Tile</td>
<td>Energies in last layer of EMB and first layer of Tile</td>
</tr>
<tr>
<td>3</td>
<td>In front of Tile gap scintillators</td>
<td>Energy in Tile gap scintillators</td>
</tr>
<tr>
<td>4</td>
<td>In front of EMEC</td>
<td>Energy in PreSampler E</td>
</tr>
<tr>
<td>5</td>
<td>Between EMEC and HEC</td>
<td>Energies in last layer of EMEC and first layer of HEC</td>
</tr>
<tr>
<td>6</td>
<td>In front of FCAL</td>
<td>Energy in first FCAL module</td>
</tr>
<tr>
<td>7</td>
<td>Behind calorimeters</td>
<td>Energy in last layer of hadronic calorimeters and $O_{clus}^{dm}$ given in Eq. (39)</td>
</tr>
<tr>
<td>8</td>
<td>Everywhere else</td>
<td></td>
</tr>
</tbody>
</table>

Figure 17 The average energy loss in the virtual dead material cells for charged 100 GeV pions. The numbers 1–7 indicate the different regions, with region 8 (not displayed) being everywhere outside regions 1–7. The dead material cells are superimposed on a schematic ($r$, $z$) view showing a quarter of the ATLAS calorimeter system with its read-out segmentation sampling layer $s$ is used to assign the dead material energy to topo-clusters. For a given cluster $k$ out of $N_{clus}$ topo-clusters which have cells from $s$ included, the assignment weight $w$ is calculated using

$$w = \sqrt{\frac{E_{clus,k}^{dep}}{\sum_{i=1}^{N_{clus}} E_{clus,i}^{dep}}} \times \exp(-\Delta R_{k}/R_0),$$

with

$$\Delta R_{k} = \sqrt{(\Delta \eta_{k})^2 + (\Delta \phi_{k})^2} \quad \text{and} \quad R_0 = 0.2.$$  

(38)

The choice of $s$ depends on the dead material regions indicated in Fig. 17. The distances $\Delta \eta$ and $\Delta \phi$ are measured between the topo-cluster direction and the dead material cell direction. The normalisation of $w$ is calculated using all $N_{clus}$
clusters such that $0 \leq w \leq 1$. It is rare that two clusters are close to the same dead material cell, most often $w = 1$ is found for the closest topo-cluster, and $w = 0$ for the next closest ones.

This weighted energy loss is collected as a function of observables of the associated topo-cluster given in Table 2. Lost energy deposited in front of the calorimeter is compensated for by applying a correction proportional to the pre-sampler signals in topo-clusters which contain these signals. In the forward region the signal in the first module FCAL0 of the FCAL is used for this purpose.

Energy lost between an electromagnetic and a hadronic calorimeter module (regions 2 and 5 in Table 2; Fig. 17) is found to be proportional to $E_{\text{EM}}^{\text{EM}} \cdot E_{\text{EM}}^{\text{EM}}$, where $E_{\text{EM}}^{\text{EM}}$ is the energy in the last sampling layer of the electromagnetic calorimeter, and $E_{\text{EM}}^{\text{EM}}$ is the energy in the first sampling layer of the hadronic calorimeter. Both $E_{\text{EM}}^{\text{EM}}$ and $E_{\text{EM}}^{\text{EM}}$ are reconstructed on the electromagnetic energy scale. This correction is only applied to topo-clusters which span the material between the two calorimeters.

Dead material corrections for longitudinal leakage (region 7 in Table 2; Fig. 17) are applied to topo-clusters that contain cells from the very last (hadronic) calorimeter sampling layer. These corrections are calculated in three-dimensional bins of a set of observables $\mathcal{O}_{\text{clus}}$, with

$$\mathcal{O}_{\text{clus}} = \{ \eta_{\text{clus}}, \log_{10}(E_{\text{clus}}/E_0), \lambda_{\text{clus}} \}, \quad (39)$$

and $E_0$ from Eq. (30) in Sect. 5.2. The same set of observables is used as input to correct dead material energy losses in topo-clusters that are located in the direct neighbourhood of inactive material categorised as region 8 and that have no other dead material correction applied.

Like the out-of-cluster correction, the dead material correction is a cluster-based correction. It is expressed in terms of a weight $w_{\text{clus}}^{\text{dm}}$, which is determined from the various correction functions or lookup tables. The corresponding cell signal weight is the same for all cells of the given cluster $(w_{\text{clus}}^{\text{dm}} = w_{\text{clus}}^{\text{dm}})$. This correction therefore does not affect the topo-cluster barycentre or centre of gravity.

### 5.6 Fully calibrated cluster kinematics

The reconstructed and fully calibrated topo-cluster energy $E_{\text{clus}}^{\text{LCW}}$ depends on the EM likelihood of the cluster, as discussed in Sect. 5.2, and is characterised by $E_{\text{clus}}^{\text{LCW}} \geq E_{\text{clus}}^{\text{EM}}$. The cluster direction changes due to the calibration, because it is calculated from energy-weighted cell directions using Eqs. (9) and (10) with $w_{\text{clus}}^{\text{geo}} \rightarrow w_{\text{clus}}^{\text{geo-cal}}$.

The effective cell calibration weight $w_{\text{clus}}^{\text{cal}}$ from Eq. (29) after any of the calibrations or corrections are applied yields the cluster energy $E_{\text{clus}}^{\text{clus}}$ after the calibration

$$E_{\text{clus}}^{\text{clus}} = \sum_{i \in \text{cluster}} w_{\text{clus},i}^{\text{clus}} E_{\text{clus},i}^{\text{EM}}. \quad (40)$$

While the signal weights determined for each calibration and correction are independently derived, the overall effect of the calibration sequence leads to a factorised accumulation of $w_{\text{clus}}^{\text{clus}}$ in the reconstruction of the cell energies. This is summarised in Table 3. The overall weight $w_{\text{clus}}^{\text{clus}}$ given in item (5) of the table is used cell-by-cell in Eq. (40) to calculate the final cluster energy $E_{\text{clus}}^{\text{clus}}$ by setting $w_{\text{clus}}^{\text{clus}} = w_{\text{clus},i}^{\text{clus}}$ and thus yielding $E_{\text{clus}}^{\text{clus}} = E_{\text{clus}}^{\text{clus}}$. As discussed earlier, $w_{\text{clus}}^{\text{clus}}$ is also used to recalculate the cluster directions $\eta_{\text{clus}}$ and $\phi_{\text{clus}}$. The final fully calibrated four-momentum reconstructed for any topo-cluster is given by replacing $E_{\text{clus}}^{\text{clus}}$ in Eq. (12) in Sect. 3.2 with $E_{\text{clus}}^{\text{clus}}$.

All input parameter values used in the LCW calibration are derived from dedicated single-particle MC simulations. The validity of this calibration is confirmed with data, where the cumulative effect of the hadronic calibration and the out-of-cluster and dead material corrections on the signal of topo-clusters found in jets is analysed and compared to corresponding MC simulations. Figure 18 summarises the quality of the LCW calibration for these clusters, both as a function.
Fig. 18 The average ratio of reconstructed to EM-scale energy after each calibration step, as a function of the cluster energy $E_{\text{clus}}^{\text{EM}}$ (a, c, e) for topo-clusters in anti-$k_t$ jets with $R = 0.6$ and $p_T > 20$ GeV and with rapidities $|y_{\text{jet}}| < 0.3$. The corresponding average ratios as a function of $\eta_{\text{clus}}$ are shown in b, d, and f. Data recorded in 2010 is compared to the corresponding MC simulations. The figures are adapted from Ref. [38].

The basic cluster signal $E_{\text{clus}}^{\text{EM}}$ and the cluster direction $\eta_{\text{clus}}$ [38]. Data are compared to MC simulations after the application of the hadronic cell weights ($E_{\text{clus}}^{\text{EM}}$ in Fig. 18a, b), followed by the out-of-cluster correction ($E_{\text{clus}}^{\text{cal}+\text{ooc}}$ in Fig. 18c, d), and at the LCW scale after applying the dead material correction ($E_{\text{clus}}^{\text{LCW}}$ in Fig. 18e, f). The differences between data and MC simulations are determined from these results as functions of $E_{\text{clus}}^{\text{EM}}$ and $\eta_{\text{clus}}$ using the respective double-ratio

$$
\frac{\langle E_{\text{clus}}^{\text{cal}} / E_{\text{clus}}^{\text{EM}} \rangle_{\text{data}}}{\langle E_{\text{clus}}^{\text{cal}} / E_{\text{clus}}^{\text{EM}} \rangle_{\text{MC}}}, \quad \frac{\langle E_{\text{clus}}^{\text{cal}+\text{ooc}} / E_{\text{clus}}^{\text{EM}} \rangle_{\text{data}}}{\langle E_{\text{clus}}^{\text{cal}+\text{ooc}} / E_{\text{clus}}^{\text{EM}} \rangle_{\text{MC}}},
$$

and

$$
\frac{\langle E_{\text{clus}}^{\text{LCW}} / E_{\text{clus}}^{\text{EM}} \rangle_{\text{data}}}{\langle E_{\text{clus}}^{\text{LCW}} / E_{\text{clus}}^{\text{EM}} \rangle_{\text{MC}}},
$$

where $E_{\text{clus}}^{\text{EM}}$, $E_{\text{clus}}^{\text{cal}}$, and $E_{\text{clus}}^{\text{cal}+\text{ooc}}$ are the EM, calibration, and calibration plus out-of-cluster corrected energies, respectively, and $E_{\text{clus}}^{\text{LCW}}$ is the energy at the LCW scale after dead material correction.
These double-ratios are shown in Fig. 18 as well, and indicate generally good agreement between data and MC simulations. The particular structures shown in the $q_{\text{clus}}$ dependence of the magnitude of the various calibration steps indicate the cumulative effects of transition regions between calorimeters in ATLAS, due to not only technology changes but also to changes in the read-out granularity. Especially Fig. 18f shows the large correction factors applied by the LCW calibration in the attempt to recover signal losses introduced by (1) the transition between the central and the endcap calorimeters at $|\eta| \approx 1.45$, (2) the transition between end-cap and forward calorimeters at $|\eta| \approx 3.2$, and (3) the upper limit of the ATLAS calorimeter acceptance at $|\eta| \approx 4.9$.

6 Performance of the simulation of topo-cluster kinematics and properties

The reconstruction performance of the topological cell clustering algorithm in ATLAS can be evaluated in the context of reconstructed physics objects such as jets or (isolated) single particles. In addition, features of the topo-cluster signal outside these physics objects can be studied with exclusive samples of low-multiplicity final states without jets. These are preferably selected by muons as those leave only small signals in the calorimeter, nearly independent of their $p_T$ ($W \rightarrow \ell \nu$ or $Z \rightarrow \mu \mu$ without jets). The topo-clusters not used in reconstructing hard physics objects reflect the calorimeter sensitivity to small and dispersed energy flows generated by the proton–proton collisions in the LHC, including pile-up. The level of agreement between data and MC simulations is used in all cases as a metric for the reconstruction performance.

6.1 Single-particle response

The calorimeter response to single isolated charged hadrons with well-measured momentum in the ID was determined using proton–proton collision data at $\sqrt{s} = 900$ GeV in 2009 [45]. The single-hadron response at higher centre-of-mass energies was determined in 2010 at $\sqrt{s} = 7$ TeV and in 2012 at $\sqrt{s} = 8$ TeV [46]. Due to the relatively low luminosities in the 2009 and 2010 run periods, pile-up contributions are insignificant in the corresponding data. These measurements provide important validations of the topo-cluster algorithm and the calorimeter acceptance in general.

The principal observable is the energy-to-momentum ratio $E/p$. The calorimeter energy $E$ is reconstructed using the topo-clusters located around the direction of the track of the incoming charged particle with momentum $p$, including the ones with $E_{\text{clus}} < 0$. The effect of the axial magnetic field is taken into account by extrapolating the reconstructed tracks into the calorimeter. The energy $E$ is then calculated by summing the EM-scale energies from all sampling layers $s$ of topo-clusters which have a barycentre $(\eta_s, \phi_s)$ within $\Delta R = 0.2$ of the track direction extrapolated to each $s$, as described in more detail in Ref. [45]. The sampling layer energies are summed irrespective of their sign, i.e. $E < 0$ is possible.

The results of the measurement of $E/p$ are shown in Figs. 19a and b for reconstructed isolated tracks in proton–proton collisions at $\sqrt{s} = 900$ GeV. The distributions reflect the acceptance of the calorimeter for charged particles in the given momentum ranges. Entries for $E/p < 0$ indicate that the incoming track is matched with a topo-cluster generated by significant electronic noise. The number of tracks with no matching calorimeter signal ($E = 0 \Rightarrow E/p = 0$) is indicative of none or only a small fraction of the particle energy reaching the calorimeter, and the signal generated by
this energy fraction is not sufficiently significant to survive the implicit noise suppression in the topo-cluster formation described in Sect. 3.1.

The likelihood $P_{E=0}(d_{\text{in}})$ to find $E = 0$ for a charged particle passing through inactive material of various thicknesses $d_{\text{in}}$, measured in terms of the nuclear interaction length $\lambda_{\text{nucl}}$, is shown in Fig. 20a for isolated tracks within $|\eta| < 1.0$ in proton–proton collisions at $\sqrt{s} = 900$ GeV. The various values of $d_{\text{in}}$ are extracted from the detector description in the MC simulation using the direction $|\eta|$ of the incoming tracks. The data and MC simulations agree well, indicating an appropriate description of the actual detector geometry in the MC simulation. The likelihood to have no matching signal in the calorimeter shows the expected increase with increasing inactive material.

The dependence of $P_{E=0}$ on the track momentum is shown in Fig. 20b for isolated tracks with $|\eta| < 0.6$. Good agreement between data and MC simulations is observed, which together with the results displayed in Figs. 19 and 20a indicates a good description of the data by the QGSP_BERT hadronic shower model used by the MC simulation.

The dependence of $E/p$ on the track momentum has been evaluated for two different hadronic shower models in GEANT4. In addition to the default QGSP_BERT model introduced in Sect. 2.3.4, the FRITIOF model [47,48] is considered together with the Bertini intra-nuclear cascade to simulate hadronic showers (FTFP_BERT). The results for 2012 data from a dedicated sample with insignificant pile-up ($\mu \approx 0$) are presented in Fig. 21 and show good agreement between data and MC simulations without indicating a strong preference for one of the hadronic shower models. More results of the full systematic evaluation of the topo-cluster response to single charged hadron tracks, including for selected tracks from identified charged mesons and baryons, are available in Ref. [45].

6.2 Effect of pile-up on topo-cluster observables

The topo-cluster reconstruction performance is affected by in-time and out-of-time pile-up. While in-time pile-up is expected to usually increase the number of topo-clusters with increasing number of reconstructed vertices ($N_{\text{PV}}$), the out-of-time pile-up leads to cluster signal and shape modifications introduced by the calorimeter signal shaping functions described in Sect. 2.2.1.

The high density of very significant cell signals generated inside jets in the calorimeter increases the likelihood of low-energy pile-up signals to survive in the topo-cluster formation, according to the formation rules given in Sect. 3.1. Cell signals generated by the energy flow of relatively isolated particles entering the calorimeter outside jets or (stochastic) jet-like flow structures\(^{14}\) often have less significant neighbouring cells and thus contribute less often to topo-clusters. Consequently, the acceptance of the calorimeter for these particles, many of which are produced by pile-up, is lower than for particles in or around a jet.

In this section the modelling of the pile-up effects on the kinematics and moments used for the LCW calibration is compared to data for topo-clusters formed inside and outside jets for the conditions during 2012 running. The effect of pile-up on jets reconstructed from topo-clusters is discussed in Sect. 6.3, together with the stability of topo-cluster-based observables associated with the jet and its composition.

\(^{14}\) These can be generated by particles from different pile-up collisions going in the same direction in a given event.
6.2.1 Event selection

The data used for the evaluation of the pile-up effects on topo-cluster kinematics and moments are collected from \(Z \rightarrow \mu\mu\) events recorded in 2012. As indicated in Sect. 2.1, the corresponding sample is defined by a muon-based trigger. The additional event selection, applied to both data and the corresponding MC simulations, requires two muons with \(p_T > 25\) GeV within \(|\eta| < 2.4\) and an invariant mass \(m_{\mu\mu}\) of the muon pair of 80 GeV < \(m_{\mu\mu}\) < 100 GeV for the inclusive sample. For the analysis of an exclusive sample with softer hadronic recoil against the \(Z\) boson transverse momentum \((p_T^Z)\), events with at least one jet reconstructed with the anti-\(k_t\) algorithm and a distance parameter \(R = 0.4\) and \(p_T > 20\) GeV are removed. This sample is characterised by a final state dominated by in-time pile-up signal contributions, with only a small number of topo-clusters associated with the hadronic recoil.

Another exclusive sample for the analysis of topo-cluster features in jets is selected by requiring at least one anti-\(k_t\) jet with \(p_T > 20\) GeV in the event. Like in the selection applied to collect the exclusive sample without jets, all jets are fully calibrated and corrected, including a correction for pile-up (see Sect. 2.4). All inclusive and exclusive samples are thus characterised by their stability against pile-up.

6.2.2 Modelling of topo-cluster kinematics in events with pile-up

Detailed data/MC comparisons of topo-cluster kinematics yield significant differences between the measured and the modelled spectra. The transverse momentum spectra of topo-clusters reconstructed on the EM scale \((p_{T,clus}^{EM})\) for the final state of an inclusive \(Z \rightarrow \mu\mu\) sample, are shown in Figs. 22a and b for the central, in Figs. 22c and d for the end-cap, and in Figs. 22e and f for the forward detector region. The comparison between the \(p_{T,clus}^{EM}\) spectra from MC simulations with fully modelled pile-up and data in the various \(\eta_{clus}\) ranges shows significant disagreements. Possible sources are an imperfect detector simulation or the modelling of the underlying soft physics processes in the MC generator.

Using the data overlay method described in Sect. 2.3.3 improves the data/MC comparison of the \(p_{T,clus}^{EM}\) spectra significantly, especially in the low-\(p_T\) regime, where pile-up is expected to have a large effect. This improvement can be seen in Fig. 22b, d and f for the respective \(\eta_{clus}\) ranges.

6.2.3 Transverse momentum flow in the presence of pile-up

The transverse momentum flow in the \(Z \rightarrow \mu\mu\) sample without jets with \(p_T > 20\) GeV is reconstructed using the exclusive selection defined in Sect. 6.2.1. Topo-clusters are selected by \(p_{T,clus}^{EM} > p_{T,clus,min}\), where \(p_{T,clus,min} \in \{0, 100\,\text{MeV}, 250\,\text{MeV}, 500\,\text{MeV}, 1\,\text{GeV}, 2\,\text{GeV}\}\). The flow is measured by the average total transverse momentum \((\Sigma p_{T,clus}^{EM})\), carried by all or selected topo-clusters in any given direction \(\eta_k \leq \eta_{clus} < \eta_{k+1}\), and averaged over a given number of events \(N_{evts}\):

\[
(\Sigma p_{T,clus}^{EM})(\eta_{clus}) = \frac{1}{N_{evts}} \sum_{i=1}^{N_{evts}} \sum_{|j| < \eta_{clus,j} < \eta_{k+1}} p_{T,clus,j}^{EM} 
\]

(41)
Fig. 22 The distribution of the transverse momentum of topo-clusters reconstructed on the EM scale ($p^E_{\text{clus}}$) for an inclusive $Z \rightarrow \mu\mu$ event sample recorded in 2012. Data are compared to distributions from MC simulations (a, c, e) including fully simulated pile-up and (b, d, f) with pile-up overlaid from data for all topo-clusters within (a, b) $|\eta_{\text{clus}}| < 0.2$, (c, d) $2.0 < |\eta_{\text{clus}}| < 2.2$, (e, f) $3.8 < |\eta_{\text{clus}}| < 4.0$. The ratio of the distribution from data to the one from MC simulation is evaluated bin-by-bin and shown below the respective distribution. The shaded bands indicate the statistical uncertainties from MC simulations for both the spectra and the ratios.
Here \( \eta_k \) denotes the lower boundary of the \( k \)-th \( \eta \)-bin used to sum the transverse momentum of the selected topo-clusters in each event. Figure 23 shows \( \langle \Sigma p_{T,\text{clus}}^{EM} \rangle(\mu) \) as a function of \( \eta_{\text{clus}} \) for the various topo-cluster selections for this \( Z \rightarrow \mu\mu \) data sample and the corresponding MC simulations.

The pile-up dependence of the average transverse momentum flow in various detector regions, as expressed by \( \langle \Sigma p_{T,\text{clus}}^{EM} \rangle(\mu) \), is shown in Fig. 24 for an inclusive \( \langle p_{T,\text{clus}}^{EM} > 0 \rangle \) and a exclusive \( p_{T,\text{clus}}^{EM} > 1 \text{ GeV} \) topo-cluster selection. The MC simulations predict the flow in the detector regions \( |\eta| < 0.2 \) and \( 2.0 < |\eta| < 2.2 \) well, in particular for the more pile-up-sensitive cluster selection shown in Figs. 24a and c. Larger deviations are observed for these two regions with the exclusive selection in Figs. 24b and d. In the forward region, MC simulations predict higher \( p_T \)-flow for both topo-cluster selections, as can be seen in Figs. 24e and f. The slope of the \( \langle \Sigma p_{T,\text{clus}}^{EM} \rangle(\mu) \) dependence in this region is very similar for data and MC simulations.

The observations in Figs. 22, 23, 24 indicate that in the case of the fully simulated pile-up the simulation of the topo-cluster response to the underlying transverse energy flow outside jets suffers from MC simulation deficiencies. The use of overlaid pile-up from data, while not demonstrated here in all details, promises significant improvements for the modelling of the soft-event signals.

### 6.2.4 Topo-cluster multiplicity in the presence of pile-up

The calorimeter signal occupancy in the exclusive \( Z \rightarrow \mu\mu \) sample is determined using selected topo-clusters with \( p_{T,\text{clus}}^{EM} > p_{T,\text{min}} \) and the \( p_{T,\text{min}} \) values used in Sect. 6.2.3. The relevant observable is the cluster number density, which is given by the number of topo-clusters per unit \( \eta \) \( (\partial N_{\text{clus}} / \partial \eta) \). Figure 25 shows the average \( \langle \partial N_{\text{clus}} / \partial \eta \rangle(\eta_{\text{clus}}) \) for these topo-cluster selections, for data and MC simulations with fully simulated pile-up. The shape observed especially for the less restrictive selections with \( p_{T,\text{min}} \leq 500 \) MeV in Fig. 25a-d, reflects the variations of the calorimeter segmentation and the effect of sub-detector transition regions on the topo-cluster formation across the full ATLAS acceptance \( |\eta_{\text{clus}}| < 4.9 \). Generally, MC simulations describe the \( p_T \)-flow better than the number of clusters. This is expected as the description of the summed \( p_T \)-flow is constrained with more weight in the numerical fits for the ATLAS tunes than the particle number density.

The topo-cluster number density changes rapidly at \( |\eta_{\text{clus}}| = 2.5 \). This is a consequence of the reduction of the calorimeter cell granularity by about a factor of four in terms of pseudorapidity and azimuth \( (\Delta \eta \times \Delta \phi) \), which reduces the number of potential topo-cluster seeds. The granularity change also introduces more signal overlap between individual particles in any given cell and thus less spatial resolution for the reconstruction of the corresponding energy flow due to this merging of particle signals. In addition, the larger cells increase the noise thresholds, as shown in Figs. 4b and c, which changes the calorimeter sensitivity. This change of sensitivity can be evaluated by comparing \( \langle \Sigma p_{T,\text{clus}}^{EM} \rangle \) with the corresponding quantity

\[
\langle \Sigma p_{T,\text{clus}}^{EM} \rangle(\eta_{\text{clus}}) = \frac{1}{N_{\text{clus}}} \sum_{i=1}^{N_{\text{clus}}} \left[ \sum_{j \mid |\eta_i| < \eta_{\text{clus}}, j < \eta_{i+1}} p_{T,\text{clus},j}^{EM} \right],
\]

reconstructed from all calorimeter cell signals in each \( \eta \) bin, similar to Eq. (41) for clusters. The cell-based \( p_T \)-flow expressed by \( \langle \Sigma p_{T,\text{clus}}^{EM} \rangle(\eta_{\text{clus}}) \) is unbiased with respect to noise suppression as none is applied. Consequently, it is subject to larger fluctuations. Figure 26 shows this measurement for a 2012 MB data sample with pile-up close to the nominal \( \mu = 30 \) used for the noise thresholds (see Sect. 2.2.2). It indicates signal losses due to clustering up to about 50 % for \( 2.5 < |\eta| < 4.5 \), and some signal increase due to suppression of cells with \( E < 0 \), in particular in the end-cap region \( 1.5 < |\eta| < 2.5 \). All topo-clusters and calorimeter cell signals are accepted for this study.

The geometry effect yields the steep drop in topo-cluster number density at this boundary. Raising the transverse momentum threshold for accepted topo-clusters increasingly mitigates the geometrical and noise effects on the cluster number density. The data/MC comparison shows larger deficiencies for more inclusive topo-cluster selections, which capture more signals from pile-up. It improves as the \( p_{T,\text{min}} \) threshold increases, when the selections are dominated by clusters that are generated by harder emissions than those due to pile-up.

The dependence of the average number of topo-clusters in a given calorimeter region on the pile-up activity, expressed in terms of \( \mu_i \), is shown for clusters with \( p_{T,\text{clus}}^{EM} > 0 \) and \( p_{T,\text{clus}}^{EM} > 1 \text{ GeV} \) in Fig. 27. Applying the (inclusive) \( p_{T,\text{clus}}^{EM} > 0 \) selection yields more topo-clusters in MC simulations than in data in the selected central (\( |\eta| < 0.2 \)) and end-cap (\( 2.0 < |\eta| < 2.2 \)) regions, with the difference rising with increasing \( \mu_i \) in Figs. 27a and c. In the forward region the number of topo-clusters in MC simulations is closer to the number in data for low \( \mu_i \), but tends to be lower than data at higher \( \mu_i \), as seen in Fig. 27e. These qualitative differences between the observations for the central and end-cap regions and the forward region can arise from the modelling of soft physics, which is tuned with reconstructed charged tracks in the detector region \( |\eta| < 2.5 \) but is not experimentally constrained in the forward region.

In addition, imperfections in the description of the inactive material in front of the calorimeter in the detector simulation.
Fig. 23 The average $\langle \Sigma p_{T,\text{clus}}^{\text{EM}} \rangle$ of clusters at the EM scale, calculated as function of $\eta$ using Eq. (41), for clusters with $p_{T,\text{clus}}^{\text{EM}} > 0$, $b$ $p_{T,\text{clus}}^{\text{EM}} > 100$ MeV, $c$ $p_{T,\text{clus}}^{\text{EM}} > 250$ MeV, $d$ $p_{T,\text{clus}}^{\text{EM}} > 500$ MeV, $e$ $p_{T,\text{clus}}^{\text{EM}} > 1$ GeV, and $f$ $p_{T,\text{clus}}^{\text{EM}} > 2$ GeV. Results are obtained from a 2012 $Z \rightarrow \mu \mu$ sample without jets with $p_T > 20$ GeV in data and MC simulation. The ratios of $\langle \Sigma p_{T,\text{clus}}^{\text{EM}} \rangle(\eta)$ from data and MC simulations are shown below each plot.
Fig. 24 The average transverse momentum flow ($\langle \Sigma_{\text{clus}} p_{\text{EM}}^T \rangle$) evaluated as function of the pile-up activity measured by the number of proton–proton interactions per bunch crossing $\mu$, in several calorimeter regions. In a, c, and e, $\langle \Sigma_{\text{clus}} p_{\text{EM}}^T \rangle(\mu)$ is shown in the central ($|\eta| < 0.2$), end-cap ($2.0 < |\eta| < 2.2$), and the forward ($3.8 < |\eta| < 4.0$) region, respectively, using topo-clusters with $p_{\text{EM}}^T > 0$. The corresponding results using topo-clusters with $p_{\text{EM}}^T > 1$ GeV are presented in b, d, and f. Results are obtained from a 2012 $Z \rightarrow \mu\mu$ sample without jets with $p_T > 20$ GeV in data and MC simulations. The narrow shaded bands around the results for MC simulations indicate statistical uncertainties, both for $\langle \Sigma_{\text{clus}} p_{\text{EM}}^T \rangle(\mu)$ and the corresponding data-to-MC simulation ratios shown below each plot.
Fig. 25  Average topo-cluster number density $\langle \partial N/\partial \eta \rangle$ as a function of $\eta_{\text{clus}}$, for clusters with $p_{\text{T,clus}} > p_{\text{T,min}}$, for various $p_{\text{T,min}}$ values. Results are obtained from a 2012 $Z \rightarrow \mu\mu$ sample without jets with $p_{\text{T}} > 20 \text{ GeV}$ in data and MC simulations. The corresponding data-to-MC simulation ratios are shown below each figure.
The forward region is suspected to introduce a higher sensitivity, particularly the increased signal overlap between particles in the different calorimeter regions. In some cases, pile-up can actually increase the cluster splitting behaviour in the various calorimeter regions. The reconstructed average transverse momentum flow on EM scale, measured with topo-clusters in bins of $\eta$ using $\langle \Sigma p_{T,\text{clus}}^{\text{EM}}(\eta) \rangle$ in Eq. (41) and with all calorimeter cells in the same $\eta$-bins using $\Sigma p_{T,\text{cell}}^{\text{EM}}(\eta)$ given in Eq. (42), in 2012 MB data can change the acceptance for low-energy particles significantly in different ways in the various $\eta$-regions. Also, mismodelling in the simulation of the (mostly hadronic) lateral and longitudinal shower spreads in the calorimeters, as e.g. documented in Refs. [49,50], can lead to different topo-cluster splitting behaviour in the different calorimeter regions. In particular the increased signal overlap between particles in the forward region is suspected to introduce a higher sensitivity of the cluster splitting to the detector simulation.

As can be seen in Figs. 27b and f, counting only topo-clusters with $p_{T,\text{clus}}^{\text{EM}} > 1$ GeV introduces a more similar slope in the cluster number density as a function of $\mu$. The qualitative behaviour of $\langle \partial N_{\text{clus}}/\partial \eta \rangle(\eta_{\text{clus}})$ in the various detector regions is different than for the more inclusive topo-cluster selection, with MC simulation predicting fewer clusters in the central and end-cap regions shown in Figs. 27b and d. In the forward region, data shows overall fewer clusters than MC simulation, as can be seen in Fig. 27f, with larger differences at any given $\mu$, but a very similar number of additional clusters per additional proton–proton interaction. Pile-up is expected to affect cluster moments as well as the overall cluster kinematics. Its diffuse energy emission can not only produce additional topo-clusters, but also change the centre of gravity, the barycentre, and other cluster shapes. In some cases, pile-up can actually increase the cluster splitting, as additional local signal maxima can be inserted into a topo-cluster by pile-up. In addition, the increased cell noise can produce additional signal minima in groups of previously connected cells in the topo-cluster. This last effect can be more important for topo-clusters in jets and is discussed in Sect. 6.3. The topo-cluster depth location $\lambda_{\text{clus}}$ discussed here serves as an example for the quality of modelling cluster moments in the presence of pile-up. Other moments are investigated in the context of jets.

The modelling of $\lambda_{\text{clus}}$ in the calorimeter is compared to data in Fig. 28 for the inclusive $Z \rightarrow \mu\mu$ sample in the same bins of $\eta_{\text{clus}}$ used for the study of $p_{T,\text{clus}}$ in Fig. 22. The fully simulated events with pile-up from the minimum-bias simulations show significant differences from the data, while the MC simulations overlaid with pile-up from data show good agreement with respect to all features of these distributions. The complex structure of the distributions reflects the longitudinal calorimeter segmentation in the various regions defined by $\eta_{\text{clus}}$. For example, the forward direction $3.8 < |\eta_{\text{clus}}| < 4.0$ is covered by the FCal, which has three coarse and deep longitudinal segments (approximately 2.5/3.5/3.5 $\lambda_{\text{amu}}$). This structure generates topo-clusters preferably in the depth centre of each module, as can be seen in Figs. 28e and f. These distributions are dominated by low-energy clusters associated with pile-up interactions such that the improvement seen by using data overlay is expected.

Similarly to the studies of the kinematic and flow properties of topo-clusters discussed in Sects. 6.2.2 and 6.2.3, more exclusive topo-cluster selections are also investigated. Figure 29 shows data/MC comparisons of the $\lambda_{\text{clus}}$ distributions for clusters within $|\eta_{\text{clus}}| < 0.2$ for $p_{T,\text{clus}}^{\text{EM}} > p_{T,\text{min}}$ with $p_{T,\text{min}} \in \{1, 2, 5\}$ GeV, for MC simulations with fully simulated pile-up and for MC simulations with pile-up from data overlaid. The MC simulation with overlaid pile-up agrees better with data than the one with fully simulated pile-up, particularly in the case of the least restrictive $p_{T,\text{min}} = 1$ GeV topo-cluster selection.

6.3 Topo-clusters in jets

Jets are important in many analyses at the LHC. Therefore, the performance of the simulation of their constituents is important, in particular for analyses employing jet substructure techniques or relying on the jet mass. In order to study the topo-cluster features in jets and the jet topo-cluster composition, exclusive jet samples are extracted from data and MC simulation using the $Z \rightarrow \mu\mu$ and jet selection described in Sect. 6.2.1. As the jets are globally corrected for pile-up [16], they form a stable kinematic reference for the evaluation of pile-up effects on the topo-clusters used to reconstruct them. Jets include only topo-clusters with $E > 0$, as required by the kinematic recombination.

The full evaluation of the reconstruction performance for jets formed with topo-clusters on both EM and LCW scale is presented in Refs. [16,38]. The evaluation of the jet energy resolution can be found in Ref. [51].
Fig. 27 The dependence of the average number of positive-energy topo-clusters on the pile-up activity measured by the number of proton–proton collisions per bunch crossings \( \mu \) in several regions of the detector is shown in a, b for \( |\eta| < 0.2 \), in c, d for \( 2.0 < |\eta| < 2.2 \), and in e, f for \( 3.8 < |\eta| < 4.0 \). Plots a, c and e show the results for counting all clusters with \( p_{T,\text{clus}}^{\text{EM}} > 0 \), while b, d and f show the results for only counting clusters with \( p_{T,\text{clus}}^{\text{EM}} > 1 \) GeV. The corresponding ratio of data to MC simulations is shown below each plot. All results are obtained from a 2012 \( Z \rightarrow \mu\mu \) sample without jets with \( p_T > 20 \) GeV in data and MC simulations. The narrow shaded bands indicate the statistical uncertainties associated with the results from MC simulations for the mean values and the ratios.
Fig. 28 The distribution of the topo-cluster depth location, measured in terms of $\log_{10}(\lambda_{\text{clus}}/\lambda_0)$, for clusters in various bins of $|\eta_{\text{clus}}|$ for an inclusive $Z \rightarrow \mu\mu$ event sample recorded in 2012. Data is compared to distributions from MC simulations including fully simulated pile-up for all topo-clusters within a $|\eta_{\text{clus}}| < 0.2$, c $2.0 < |\eta_{\text{clus}}| < 2.2$, and e $3.8 < |\eta_{\text{clus}}| < 4.0$. The corresponding distributions for MC simulations with pile-up from data overlaid are depicted in b, d, and f. The ratios of the distributions for data and MC simulations are shown below the respective distributions. The shaded bands indicate the statistical uncertainties for MC simulations.
Fig. 29 The distribution of the topo-cluster depth location, measured in terms of $\log_{10}(\lambda_{\text{clus}}/\lambda_0)$, for selected topo-clusters within $|\eta_{\text{clus}}| < 0.2$ and with a transverse momentum $p_{\text{T,clus}}^\text{EM}$, evaluated on the EM scale, larger than various thresholds. Results are shown for an inclusive $Z \rightarrow \mu\mu$ event sample recorded in 2012. Data are compared to distributions from MC simulations including fully simulated pile-up for all topo-clusters with $p_{\text{T,clus}}^\text{EM} > 1 \text{ GeV}$, $p_{\text{T,clus}}^\text{EM} > 2 \text{ GeV}$, and $p_{\text{T,clus}}^\text{EM} > 5 \text{ GeV}$. The corresponding distributions for MC simulations with pile-up from data overlaid are depicted in (b), (d), and (f). The shaded bands indicate the statistical uncertainties for the distributions obtained from MC simulations and the corresponding uncertainties in the ratio plots.
6.3.1 Jet energy scale and topo-cluster-based response in pile-up

As mentioned above, the fully calibrated four-momentum $P_{\text{jet}}$ of jets reconstructed from topo-clusters is corrected for pile-up effects. Therefore, the corresponding transverse momentum $p_{\text{T,jet}}$ provides a stable signal for event selections and the kinematics of the true particle flow. The basic jet four-momentum is reconstructed on two different scales, the EM scale and the LCW scale using locally calibrated topo-clusters with $E > 0$:

$$P_{\text{EM,clus}}^\text{jet} = \sum_{i=1}^{N_{\text{clus}}^\text{jet}} p_{\text{clus}}^\text{EM}$$

$$P_{\text{LCW,clus}}^\text{jet} = \sum_{i=1}^{N_{\text{clus}}^\text{jet}} p_{\text{clus}}^\text{LCW}$$ \hspace{1cm} (43)

The sum runs over the number $N_{\text{clus}}^\text{jet}$ of topo-clusters in a given jet. Both $P_{\text{EM,clus}}^\text{jet}$ and $P_{\text{LCW,clus}}^\text{jet}$ are not corrected further. The corresponding $p_T$ responses $p_{T,jet}^\text{EM}$ and $p_{T,jet}^\text{LCW}$ are therefore affected by pile-up. A full jet energy scale (JES) calibration is applied to both scales, yielding $p_{T,jet}^\text{EM+JES}$ and $p_{T,jet}^\text{LCW+JES}$, respectively. This JES calibration includes pile-up corrections, response calibration, direction corrections and refinements from in situ transverse momentum balances, similar to those outlined for 2011 data in Ref. [16]. The respective fully calibrated transverse momentum is then $p_{T,jet}^\text{EM+JES}$ and $p_{T,jet}^\text{LCW+JES}$.

Figure 30 shows the pile-up dependence of the fully calibrated $p_{T,jet}^\text{LCW+JES}$ and the uncorrected $p_{T,jet}^\text{LCW}$ on the pile-up activity in the event, measured by $\mu$. Results are obtained from a $Z \rightarrow \mu\mu$ sample of events with one jet with 30 GeV < $p_{T,jet}^\text{LCW+JES}$ < 40 GeV in data and MC simulations. While Fig. 30a shows the stability of the corrected jet $p_T$ scale, Fig. 30b indicates the different sensitivities of the uncorrected response to pile-up in the various detector regions. The different shapes seen in this figure are mostly related to the calorimeter granularity and the specific shaping functions in the different LAr calorimeters. While the general expectation that every pile-up interaction adds energy to the jet is indicated in the rise of $p_{T,jet}^\text{LCW+JES}$ with increasing $\mu$, the dependence of this ratio on $\mu$ is less pronounced for jets with $3.5 < |\eta_{\text{jet}}| < 4.5$ in the FCAL calorimeter. This observation is related to the much coarser calorimeter geometry in this region, in addition to the different (faster) shaping function in the FCAL, yielding a better average online in-time pile-up suppression by the out-of-time pile-up signal history in 2012 running conditions (50 ns bunch crossings).

6.3.2 Topo-cluster multiplicity in jets

Figure 31 shows the distributions of the number of topo-clusters ($N_{\text{clus}}^\text{jet}$) in central, end-cap and forward jets. Distributions are shown using fully simulated pile-up and using data overlay. The discrepancies between MC simulations and data, while slightly reduced in the simulations employing the pile-up overlaid from data, generally persist.
Fig. 31 The distribution of the number of topo-clusters inside anti-$k_T$ jets formed with $R = 0.4$ in the (a, b) central ($|\eta| < 0.6$), the (c, d) end-cap ($2.0 < |\eta| < 2.5$), and the (e, f) forward detector region ($3.5 < |\eta| < 4.5$) of ATLAS. Shown are results from the analysis of $Z \to \mu\mu$ events with at least one jet with $30 \text{ GeV} < p_T < 40 \text{ GeV}$ in 2012 data and MC simulations with fully simulated pile-up in a, c and e, and with pile-up from data overlaid in b, d and f. The ratios of results for data and MC simulations are shown below the distributions. The shaded bands show the statistical uncertainties for the distributions obtained from MC simulations and the corresponding uncertainty bands in the ratio plots.
The data/MC comparisons of the cluster multiplicity distributions counting only topo-clusters with $p_{T,\text{clus}}^\text{EM} > 1 \text{ GeV}$ for the same $Z \rightarrow \mu\mu$ data and MC simulations are shown in Fig. 32. This comparison is significantly improved with respect to Fig. 31, indicating that the number of low-energy topo-clusters in jets is poorly simulated. The comparison of data to MC simulations with fully simulated pile-up and with pile-up overlaid from data for the more inclusive cluster multiplicities in Fig. 31 indicates that pile-up is likely not the main source for the deficiencies in the MC simulation, as the comparison does not improve significantly when pile-up is taken from the data. This observation, together with the insensitivity of the data/MC comparison of the multiplicity of harder topo-clusters to the choice of pile-up modelling in MC simulations shown in Fig. 32, suggests that the deficiencies in the simulation of the low-energy topo-cluster multiplicity arise from imperfections in the detector model, response or tuning of the parton shower and other sources of soft emissions, including multiple parton interactions in the underlying event, rather than from the modelling of pile-up or electronic noise. Further investigations concerning the distribution of the topo-cluster location in jets confirm this interpretation and are presented in Sect. 6.3.3.

The dependence of the number of clusters $N_{\text{clus}}^\text{jet}$ forming the anti-$k_T$ jets of size $R = 0.4$ and with $30 \text{ GeV} < p_{T,\text{jet}}^\text{LCW+JES} < 40 \text{ GeV}$ as a function of the pile-up activity, measured by $\mu$, is shown in Fig. 33. As indicated in Fig. 33a, $N_{\text{clus}}^\text{jet}$ rises approximately linearly with increasing $\mu$ in the central and end-cap detector regions. The gradient of this rise is much smaller in the forward region, where the coarser read-out geometry and the signal shaping effects already discussed in Sect. 6.3.1 in the context of Fig. 30b lead to merging and suppression of pile-up signals. Figure 33a also confirms the already mentioned deficiencies in the MC simulation of the absolute values of the most inclusive $\langle N_{\text{clus}}^\text{jet} \rangle$ in any given $\mu$ range, except for the forward detector region. The slope of $\langle N_{\text{clus}}^\text{jet} \rangle(\mu)$, on the other hand, compares well with data.

The number of topo-clusters in the core of the jet ($N_{\text{clus}}^\text{core}$) is defined by counting the clusters at distances $\Delta R < 0.3$ around the jet axis. Figure 33b shows a residual dependence of the average $\langle N_{\text{clus}}^\text{core} \rangle$ on $\mu$ in the central and end-cap regions, with significant differences between data and the predictions from MC simulations. The figure shows good data/MC agreement for $\langle N_{\text{clus}}^\text{core} \rangle$ in the forward region. Comparing $\langle N_{\text{clus}}^\text{jet} \rangle(\mu)$ in Fig. 33a with $\langle N_{\text{clus}}^\text{core} \rangle(\mu)$ in Fig. 33b shows a steeper slope for $\langle N_{\text{clus}}^\text{jet} \rangle(\mu)$ than for $\langle N_{\text{clus}}^\text{core} \rangle(\mu)$ in the central and end-cap calorimeter regions. Pile-up interactions add more topo-clusters at the margin of the jet than in the core. For forward jets, $\langle N_{\text{clus}}^\text{jet} \rangle$ rises only slightly with increasing $\mu$, while $\langle N_{\text{clus}}^\text{core} \rangle$ shows no observable dependency on pile-up.

Calculating $N_{\text{clus}}^\text{jet}$ and $N_{\text{clus}}^\text{core}$ with only considering topo-clusters with $p_{T,\text{clus}}^\text{EM} > 2 \text{ GeV}$ yields the result for the pile-up dependence of $\langle N_{\text{clus}}^\text{jet} \rangle$ and $\langle N_{\text{clus}}^\text{core} \rangle$ displayed in Figs. 33c and d, respectively. While both $\langle N_{\text{clus}}^\text{jet} \rangle$ and $\langle N_{\text{clus}}^\text{core} \rangle$ are nearly independent of $\mu$ in the central detector region, they show more complex dependencies on the pile-up activity in the end-cap region. The loss of hard topo-clusters in both the overall jet and in its core with increasing $\mu$ reflects additional cluster splitting induced by the diffuse energy flow from pile-up in the end-cap calorimeters. The observations in both the central and the end-cap regions are well described by MC simulations.

A good quality of the MC predictions is also achieved when comparing the number of hard topo-clusters above the $p_{T,\text{clus}}^\text{EM}$ threshold in forward jets. This number shows only a small increase with increasing $\mu$, as shown in Fig. 33c. This is due to the fact that the cluster splitting behaviour does not change with increasing pile-up in the coarse granularity of the FCAL. In this module, the residual signal contribution from pile-up shifts a small number of additional clusters above the $2 \text{ GeV}$ threshold, yielding an increase of about 10% for both $\langle N_{\text{clus}}^\text{jet} \rangle$ and $\langle N_{\text{clus}}^\text{core} \rangle$ for $\mu < 10$ to $\mu > 30$. A comparison of $\langle N_{\text{clus}}^\text{jet} \rangle(\mu)$ with and without the $p_{T,\text{clus}}^\text{EM} > 2 \text{ GeV}$ selection shows that the cut occasionally removes a topo-cluster from a forward jet such that $\langle N_{\text{clus}}^\text{jet} \rangle$ is reduced by more than 15% for any given $\mu$. The selection affects $\langle N_{\text{clus}}^\text{core} \rangle(\mu)$ in a different way. While $\langle N_{\text{clus}}^\text{core} \rangle(\mu) \approx \text{const}$ without the cut, the average number of topo-clusters in the jet core passing the $p_{T,\text{clus}}^\text{EM}$ selection is smaller by approximately 15% in the region of lower pile-up activity, where $\langle N_{\text{clus}}^\text{jet} \rangle(\mu < 10) \approx \langle N_{\text{clus}}^\text{core} \rangle(\mu < 10)$ both with and without the selection. It is only about 5% smaller for higher pile-up, where $\langle N_{\text{clus}}^\text{jet} \rangle(\mu > 30) \approx \langle N_{\text{clus}}^\text{core} \rangle(\mu > 30)$ independent of the cut, as can be seen by comparing Fig. 33b with d for forward jets.

### 6.3.3 Topo-cluster location in jets

The distribution of the depth location of all topo-clusters inside anti-$k_T$ jets reconstructed with $R = 0.4$ and with $30 \text{ GeV} < p_{T,\text{jet}}^\text{LCW+JES} < 40 \text{ GeV}$ in $Z \rightarrow \mu\mu$ events in 2012 data and MC simulations is shown in Fig. 34. Like for the depth distribution of topo-clusters in the inclusive $Z \rightarrow \mu\mu$ sample presented in Fig. 28, the MC simulations with overlaid pile-up data show better agreement with data than the ones with fully simulated pile-up. The differences in the jet context are significantly smaller than observed for the inclusive selection.

Applying a $p_{T,\text{clus}}^\text{EM} > 1 \text{ GeV}$ cut to the topo-clusters in the jets results in the depth distributions shown in Fig. 35. This selection also shows better data/MC agreement for the sample with fully simulated pile-up, an indicator consistent with the better simulation of harder signals observed in e.g. Fig. 29. A noticeable difference from the depth distributions obtained from the inclusive sample in Fig. 29a is that for topo-
Fig. 32 The distribution of the number of topo-clusters with $p_{T,\text{clus}}^\text{EM} > 1$ GeV inside anti-$k_t$ jets with $R = 0.4$ in the (a, b) central ($|\eta| < 0.6$), the (c, d) end-cap ($2.0 < |\eta| < 2.5$), and the (e, f) forward detector region ($3.5 < |\eta| < 4.5$) of ATLAS. Shown are results from the analysis of $Z \rightarrow \mu\mu$ events with at least one jet with $30 \text{ GeV} < p_T < 40 \text{ GeV}$ in 2012 data and MC simulations with fully simulated pile-up in a, c and e, and with pile-up from data overlaid in b, d and f. The data-to-MC simulation ratios are shown below the respective distributions. The shaded bands indicate statistical uncertainties for the distributions from MC simulations and the corresponding statistical uncertainty bands for the ratios.
clusters in jets the data/MC agreement in the case of the fully simulated pile-up is already better for the $p_{T,\text{clus}} > 1 \text{ GeV}$ selection, as can be seen in Fig. 35a. In addition, comparing Figs. 34 and 35 shows that the $p_{T,\text{clus}} > 1 \text{ GeV}$ selection predominantly removes topo-clusters at small depth $\lambda_{\text{clus}}$, as the distributions are depopulated more for smaller values of $\lambda_{\text{clus}}$ than for larger ones. This means that mostly topo-clusters generated by soft particles with little penetration depth into the calorimeters, including those consistent with pile-up, are removed. The data/MC comparisons are thus less sensitive to pile-up modelling issues, and therefore show better agreement.

6.3.4 Calibration and signal features of the leading topo-cluster

The leading topo-cluster in a jet is defined as the one with the highest $p_{T,\text{clus}}$. Its moments and its signal contribution to the jet provide a good reference for the dependence of important topo-cluster calibration inputs on pile-up. The leading cluster is found in the anti-$k_t$ jets reconstructed with $R = 0.4$ and with $30 \text{ GeV} < p_{T,\text{clus}} < 40 \text{ GeV}$ in the 2012 $Z \rightarrow \mu\mu$ sample in data and MC simulations with full pile-up simulation. The distributions of the topo-cluster moments relevant to the LCW calibration for the leading cluster in the
Fig. 34 The distribution of the depth location, measured in terms of $\log_{10}(\lambda_{\text{clus}}/\lambda_0)$ with $\lambda_0 = 1$ mm, of all topo-clusters in jets reconstructed with the anti-$k_t$ algorithm with $R = 0.4$ and with $30 \text{ GeV} < p_{T,jet} < 40 \text{ GeV}$ in $Z \rightarrow \mu \mu$ events in 2012 data and MC simulations with (a, c, e) fully simulated pile-up and with (b, d, f) overlaid pile-up from data. Distributions are shown for jets in the (a, b) central ($|\eta| < 0.6$), the (c, d) end-cap ($2.0 < |\eta| < 2.5$), and the (e, f) forward detector region ($3.5 < |\eta| < 4.5$). The bin-by-bin ratios of the distributions from data and MC simulations are shown below the plots. The shaded bands indicate statistical uncertainties for the distributions from MC simulations and the corresponding uncertainty bands in the ratio plots.
Fig. 35 The distribution of the depth location, measured in terms of $\log_{10}(\lambda_{\text{clus}}/\lambda_0)$ with $\lambda_0 = 1$ mm, of topo-clusters with $p_{T,\text{clus}}^{\text{EM}} > 1$ GeV in jets reconstructed with the anti-$k_t$ algorithm with $R = 0.4$ and with $30 < p_{T,jet}^{\text{JES}} < 40$ GeV in $Z \rightarrow \mu\mu$ events in 2012 data and MC simulations with (a, c, e) fully simulated pile-up and with (b, d, f) overlaid pile-up from data. Distributions are shown for jets in the (a, b) central ($|\eta| < 0.6$), the (c, d) end-cap ($2.0 < |\eta| < 2.5$), and the (e, f) forward detector region ($3.5 < |\eta| < 4.5$). The data-to-MC simulation ratios are shown below the distributions. The shaded bands shown for the distributions obtained from MC simulations indicate statistical uncertainties and the corresponding uncertainty bands in the ratio plots.
Counts / Jet

0.02
0.04
0.06
0.08
0.1
0.12
0.14

L dt ~ 20 fb

\int | \eta | < 0.6

Powheg-Pythia Z \rightarrow \mu \mu

ATLAS

Counts / Jet

0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2

L dt ~ 20 fb

\int | \eta | < 2.5

Powheg-Pythia Z \rightarrow \mu \mu

ATLAS

Counts / Jet

0.02
0.04
0.06
0.08
0.1
0.12
0.14

L dt ~ 20 fb

\int | \eta | < 4.5

Powheg-Pythia Z \rightarrow \mu \mu

ATLAS

(a)

(b)

(c)

Fig. 36 The distribution of the leading topo-cluster depth location measure \( \log_{10}(\lambda_{\text{clus}}/\lambda_0) \) in fully calibrated jets reconstructed with the anti- \( k_T \) algorithm with \( R = 0.4 \) and \( 30 \text{ GeV} < p_T^{\text{jet}} < 40 \text{ GeV} \) in regions of (a) the central \((|\eta_{\text{jet}}| < 0.6)\), (b) the end-cap \((2.0 < |\eta_{\text{jet}}| < 2.5)\), and (c) forward \((3.5 < |\eta_{\text{jet}}| < 4.5)\) calorimeters in ATLAS. Data is compared to MC simulations with fully simulated pile-up for \( Z \rightarrow \mu \mu \) events recorded in 2012. The ratio of the distribution from data to the one from MC simulations is shown below each plot. The shaded bands show statistical uncertainties for the distributions from MC simulations and the corresponding uncertainty bands in the ratio plots. The reference scale for \( \lambda_{\text{clus}} \) is \( \lambda_0 = 1 \text{ mm} \).

The distribution of the overall LCW calibration weight described in Sect. 5.6 is shown in Fig. 38.

The distribution of the depth location of the leading topo-cluster, already discussed for all and selected topo-clusters in the inclusive \( Z \rightarrow \mu \mu \) sample in Sect. 6.2.5 and the \( Z \rightarrow \mu \mu \) sample with jets in Sect. 6.3.3, is shown in Fig. 36a–c for jets reconstructed in the central, end-cap, and the forward detector region, respectively. As expected from the previous observations, MC simulations agree reasonably well with data. It is also observed that the leading cluster in the central and end-cap detector regions is most often located either in the electromagnetic or in the hadronic calorimeters, and rarely between the modules. In the forward region, the hardest cluster is most often located in the first FCAL module.

The signal density \( \rho_{\text{clus}} \) of topo-clusters is defined in Sect. 4.2.2. Figure 37 shows the \( \rho_{\text{clus}} \) distributions for the leading topo-cluster in the jet. The complex structures of these distributions are well modelled. Their shape in the central
Fig. 37 The distribution of the leading topo-cluster signal density measure \( \log_{10}(\rho_{\text{clus}}/\rho_0) - \log_{10}(E_{\text{clus}}/E_0) \) in fully calibrated jets reconstructed with the anti-\( k_t \) algorithm with \( R = 0.4 \) and 30 GeV \( < p_{\text{T}}^\text{LCW+HEG} < 40 \) GeV in regions of a the central \( (|\eta_{\text{jet}}| < 0.6) \), b the end-cap \( (2.0 < |\eta_{\text{jet}}| < 2.5) \), and c the forward \( (3.5 < |\eta_{\text{jet}}| < 4.5) \) calorimeters in ATLAS. Data is compared to MC simulations with fully simulated pile-up for \( Z \rightarrow \mu \mu \) events recorded in 2012. The ratio of the distribution from data to the one from MC simulations is shown below each plot. The shaded bands show statistical uncertainties for the distributions from MC simulations and the corresponding uncertainty bands in the ratio plots. The reference scale for \( \rho_{\text{clus}} \) is \( \rho_0 = 1 \) MeV/mm\(^3\), and for the energy \( E_0 = 1 \) MeV.

and end-cap regions is driven by the jet fragmentation. Jets with a leading photon, or two nearby photons from a neutral pion decay, can produce the leading topo-cluster with a high signal density, reflecting the single or the two largely overlapping compact electromagnetic shower(s) reconstructed in the highly granular electromagnetic calorimeters. Jets with a leading hadron that reaches the detector typically produce less dense topo-cluster signals in the corresponding hadronic calorimeters. For these jets an additional geometric effect is introduced, as the leading topo-cluster is more likely located in the hadronic calorimeters in ATLAS.\(^{15}\) The typically larger cell sizes in these detectors introduce lower density signals even for compact showers.

The forward detector region has a coarser longitudinal segmentation, with the first module FCAL0 closest to the

\(^{15}\) In the case of a leading (stable) hadron in the jet, the leading topo-cluster may still arise from a photon, as the selection of this cluster is performed on the EM scale. This introduces a bias due to \( e/\pi > 1 \), which is nevertheless well modelled in MC simulations, according to Fig. 37.
Fig. 38 The distribution of the ratio of the cluster signal reconstructed on EM scale $E_{\text{EM}}^{\text{clus}}$ to the fully calibrated signal $E_{\text{LCW}}^{\text{clus}}$ for the leading topo-cluster in fully calibrated jets reconstructed with the anti-$k_T$ algorithm with $R = 0.4$ and $30 \text{ GeV} < p_{\text{T,jet}}^{\text{LCW}+\text{JES}} < 40 \text{ GeV}$ in regions of
a the central ($|\eta_{\text{jet}}| < 0.6$), b the end-cap ($2.0 < |\eta_{\text{jet}}| < 2.5$), and c the forward ($3.5 < |\eta_{\text{jet}}| < 4.5$) calorimeters in ATLAS. Data is compared to MC simulations with fully simulated pile-up for $Z \rightarrow \mu\mu$ events recorded in 2012. The ratio of the distribution from data to the one from MC simulations is shown below each plot. The shaded bands show statistical uncertainties for the distributions from MC simulations and the corresponding uncertainty bands in the ratio plots. The reference scale for $\rho_{\text{clus}}$ is $\rho_0 = 1 \text{ MeV/mm}^3$, and for the energy $E_0 = 1 \text{ MeV}$

The overall effect of the LCW calibration described in Sect. 5 on the signal scale of the leading topo-cluster can be measured by the ratio of the basic EM scale signal $E_{\text{EM}}^{\text{clus}}$ to the fully calibrated cluster signal $E_{\text{LCW}}^{\text{clus}}$. The distribution of this ratio is shown for the three detector regions in Fig. 38a. These distributions are inclusive with respect to the topo-cluster classification described in Sect. 5.2. The shapes observed in the central and end-cap detector regions reflect this classification of the leading topo-cluster. The rightmost peak is mostly produced by topo-clusters that are generated by elec-

collision vertex being about $30 X_0$ and $2.5 \lambda_{\text{nucl}}$ deep [52]. Consequently, most leading topo-clusters in jets going in this direction are located in FCAL0, as can be seen in the $\lambda_{\text{clus}}$ distribution in Fig. 36c. The $\rho_{\text{clus}}$ distribution in Fig. 37c therefore does not show the features seen in Figs. 37a and b, because the calorimeter read-out granularity changes smoothly within this module. The hard transitions between calorimeter modules with very different granularity affecting the $\rho_{\text{clus}}$ distributions in the central and end-cap regions are absent.
tromagnetic showers and predominantly calibrated as such. In this case the calibration corrections consist of relatively small out-of-cluster and dead material corrections only, as outlined in Sect. 5. As a consequence, \( E_{\text{EM}}^{\text{clus}} / E_{\text{clus}}^{\text{LCW}} \) is closer to unity. Topo-clusters classified as hadronic receive much larger corrections, and are more likely to populate the lower side of the \( E_{\text{EM}}^{\text{clus}} / E_{\text{clus}}^{\text{LCW}} \) spectrum.

The \( E_{\text{EM}}^{\text{clus}} / E_{\text{clus}}^{\text{LCW}} \) distribution in the forward detector region shown in Fig. 38c does not display these shapes. This is due to a lack of classification power in the coarse geometry of the FCAL. Here most topo-clusters are classified as hadronic and receive relatively large corrections. The populated ranges of \( E_{\text{EM}}^{\text{clus}} / E_{\text{clus}}^{\text{LCW}} \) in Figs. 38a and b indicate that the magnitude of the total correction scaling the basic cluster signal \( E_{\text{EM}}^{\text{clus}} \) up to the locally calibrated signal \( E_{\text{clus}}^{\text{LCW}} \) reaches considerably higher values in the central region than in the end-cap detector regions. This reflects the fact that the incoming particle energies are higher at larger \( |\eta| \) for a given range in jet \( p_T \). Therefore, the calorimeter response to hadrons relative to the response to electrons and photons \( (e/\pi) \) rises with increasing \( |\eta| \), and reduces the amount of corrections needed. This effect is initially expected to be observed when comparing the end-cap with the forward region displayed in Fig. 38c as well, yet in the FCAL the out-of-cluster and dead material corrections are larger than the hadronic calibration addressing \( e/\pi > 1 \) and thus dominate the overall LCW calibration.

The signal fraction carried by the leading topo-cluster in the jet is calculated relative to the fully corrected and calibrated \( p_T^{\text{clus,lead}} \) (which provides a stable signal reference in the presence of pile-up (see Fig. 30a),

\[
\begin{align*}
\mu_{\text{clus,lead}} &= \frac{p_T^{\text{EM}}}{p_T^{\text{clus,lead}} + p_T^{\text{clus,JES}}}.
\end{align*}
\]

This means that \( \mu_{\text{clus,lead}} \) is expected to satisfy 0 < \( \mu_{\text{clus,lead}} < 1 \). Figure 39 shows the distribution of \( \mu_{\text{clus,lead}} \) in the three detector regions. The \( \mu_{\text{clus,lead}} \) distributions in the central region shown in Fig. 39a and the end-cap region shown in Fig. 39b display very similar features and indicate the most probable value is \( \mu_{\text{clus,lead}} \approx 12–15 \% \). The distribution of \( \mu_{\text{clus,lead}} \) in the forward detector region shown in Fig. 39c displays a significantly different shape introduced by the relatively low topo-cluster multiplicity in jets in this region, as shown in Figs. 31e and f. The peak at \( \mu_{\text{clus,lead}} \approx 60 \% \) in this distribution is consistent with jets with \( N_{\text{clus}}^{\text{clus}} = 1 \), and the leftmost shoulder indicates contributions from jets with \( N_{\text{clus}}^{\text{clus}} = 2 \), with the region in between populated by jets with \( N_{\text{clus}}^{\text{clus}} > 2 \). All distributions of \( \mu_{\text{clus,lead}} \) are modelled well in the MC simulations with fully simulated pile-up.

\[\text{Springer}\]

6.3.5 Pile-up dependence of leading topo-cluster signal features

The pile-up dependence of the average leading cluster signal fraction \( \langle f_{\text{clus,lead}} \rangle \), the average \( \langle E_{\text{EM}}^{\text{clus}} / E_{\text{clus}}^{\text{LCW}} \rangle \), and the average depth location of the leading topo-cluster are displayed in Fig. 40. The pile-up activity is measured in terms of \( \mu \) for this evaluation. A small linear drop of \( \langle f_{\text{clus,lead}} \rangle (\mu) \) is observed for increasing \( \mu \) in all three detector regions in Fig. 40a. This signal loss of the leading topo-cluster can arise from two effects. First, the increase of the out-of-time pile-up contributions due to the rising \( \mu \) reduces the signal due to the bipolar signal shaping function employed in the ATLAS calorimeters (see discussion in Sect. 2.2.1). Second, the increasing in-time pile-up contributions at higher \( \mu \) and the increased noise introduced by more out-of-time pile-up leads to additional splitting in the topo-cluster formation, which can take signal away from the leading cluster in the jets.

Figure 40b shows that the overall LCW calibration applied to the leading topo-cluster, measured by the average ratio \( \langle E_{\text{EM}}^{\text{clus}} / E_{\text{clus}}^{\text{LCW}} \rangle \), in the end-cap and forward detector regions is stable against increasing pile-up activity. A slight drop can be observed with increasing \( \mu \) in the central detector region, which indicates changes in the topo-cluster properties relevant to the LCW calibration introduced by increasing pile-up. One possible reason for that may be effects on the topo-cluster splitting in this region, as pile-up can induce spatial energy distributions leading to modifications in the splitting even for hard signal clusters.\(^\text{17}\) The depth location \( \lambda_{\text{clus}} \), which enters the LCW calibration in the classification step discussed in Sect. 5.2, is found to be rather stable against pile-up, as shown in Fig. 40c. The pile-up dependence of the leading topo-cluster features discussed here are found to be well modelled in MC simulations with fully simulated pile-up.

6.3.6 Leading topo-cluster geometry and shapes

The spatial extensions of the leading topo-cluster in a jet are calculated as described in Sect. 4.1. The distributions of the normalised lateral energy dispersion \( m_{\text{lat}}^{\text{clus}} \) given in Eq. (18) and the normalised longitudinal energy dispersion \( m_{\text{long}}^{\text{clus}} \) given in Eq. (19) are shown in Fig. 41 for the leading topo-cluster in jets reconstructed with the anti-\( k_T \) algorithm with \( R = 0.4 \) and 30 GeV < \( p_T^{\text{clus,JES}} \) < 40 GeV, in \( Z \rightarrow \mu \mu \) events in 2012 data and MC simulations with fully simulated pile-up. The lateral extensions represented by \( m_{\text{lat}}^{\text{clus}} \) are

\[\text{Springer}\]

\(^{17}\) In particular, pile-up can introduce an additional signal maximum at the boundary of a relatively dense leading topo-cluster, which can have a significant effect on e.g. \( \rho_{\text{clus}} \) and other cluster properties pertinent to the LCW calibration.

\[\text{Springer}\]
Fig. 39 The distribution of the signal fraction \( f_{\text{lead}} \) carried by the leading topo-cluster in jets, as defined in Eq. (45), in (a) the central, (b) the end-cap, and (c) the forward detector region. The jets are reconstructed using the anti-\( k_t \) algorithm with \( R = 0.4 \) and with \( 30 \, \text{GeV} < p_{T,\text{jet}} < 40 \, \text{GeV} \) in \( Z \to \mu\mu \) events in 2012 data and MC simulations with fully simulated pile-up. The data-to-MC simulation ratios are shown below the distributions. The shaded bands shown for the distributions obtained from MC simulations indicate statistical uncertainties and the corresponding uncertainty bands in the ratio plots reasonably well modelled in all three detector regions, with some residual discrepancies in particular in the low-value tails and upper edges of the spectra in the end-cap and forward regions. The longitudinal extensions measured by \( m_{\text{long}}^{2} \) are modelled well in the central and forward detector regions, but their modelling shows some deficiencies in the end-cap region.

The distribution of the leading topo-cluster length measure \( \sqrt{\lambda^2} \) defined in Sect. 4.1.3 in the three detector regions is shown in Fig. 42a–c. The MC simulations reproduce the shape of the \( \sqrt{\lambda^2} \) distributions from data well in the central and forward regions, with some deficiencies observed in the end-cap region. The shapes in the central and end-cap region are due to leading topo-clusters contained in the electromagnetic calorimeters populating the left peak of the distribution (short clusters) and leading topo-clusters in the hadronic calorimeters populating the right peak with longer clusters. The shape of the length distribution in the forward region shown in Fig. 42c is characterised by a sharp drop on the right of the spectrum, which corresponds to the half-depth of cells (225 mm) in the FCAL modules. This shows that in this detector region the leading topo-cluster rarely extends into all three.
Fig. 40 The pile-up dependence of a $f_{\text{lead}}$ defined in Eq. (45), b $p_{\text{EM}}/p_{\text{LCW}}$ and c the depth location $\lambda_{\text{clus}}$ of the leading topo-cluster in fully calibrated anti-$k_t$ jets reconstructed with $R = 0.4$ and with $30 \text{ GeV} < p_{\text{T,jet}} < 40 \text{ GeV}$ in $Z \rightarrow \mu\mu$ events in 2012 data and MC simulations with fully simulated pile-up. The reference scale for $\lambda_{\text{clus}}$ is $\lambda_0 = 1$ mm. The pile-up activity is measured in terms of the number of pile-up interactions $\mu$. The shaded bands shown for the results obtained from MC simulations indicate statistical uncertainties.

The pile-up dependence of leading topo-cluster geometry and shapes

The dependence of the geometry and shape of the leading topo-cluster in a jet on the pile-up activity measured by $\mu$ is shown in Fig. 44. No significant dependence is observed.

18 The location of this peak is consistent with the change of the cell size in sampling layers $\text{EME1}$ and $\text{EME2}$ of the electromagnetic end-cap calorimeter at $|\eta| = 2.5$, see Table 1.
Fig. 41 The distribution of the normalised (a, c, e) lateral ($m_{2\text{lat}}$) and (b, d, f) longitudinal ($m_{2\text{long}}$) extension measures of the leading topo-cluster in fully calibrated anti-$k_t$ jets with $R = 0.4$ and $30\text{ GeV} < p_{T,\text{jet}}^{\text{LCW+JES}} < 40\text{ GeV}$ in $Z \rightarrow \mu\mu$ events in 2012 data and MC simulations with fully simulated pile-up, for jets in the (a, b) central ($|\eta| < 0.6$), the (c, d) end-cap ($2.0 < |\eta| < 2.5$), and the (e, f) forward detector region ($3.5 < |\eta| < 4.5$) of ATLAS. The ratios of data and MC simulation distributions are shown below the plots. The shaded bands shown for the distributions obtained from MC simulations indicate statistical uncertainties and the corresponding uncertainty bands in the ratio plots.
for the average longitudinal extension of this cluster shown in Fig. 44a, the average size of this cluster in (η, φ) space in Fig. 44b, and its average lateral energy dispersion, defined in Eq. (18) and displayed in Fig. 44c.

The data/MC comparison of the average pile-up dependences shows generally acceptable agreement, but also suggests some residual deficiencies likely related to the simulation of the longitudinal and lateral (hadronic) shower shapes. Corresponding observations are reported in Refs. [49,50,53,54] in the context of detailed comparisons of ATLAS test-beam data with simulations.

7 Conclusion

Topological cell signal clusters (topo-clusters) provide a well-understood and calibrated signal definition for hadronic final-state reconstruction in the ATLAS calorimeters. The
primal algorithm generating these topo-clusters includes a noise-suppression scheme based on signal-significance patterns which is similar to applications in previous experiments. The innovative approach developed for the ATLAS calorimeters not only employs a highly refined implementation of this algorithm in a high-energy, high-luminosity hadron collider environment characterised by significant collision backgrounds introduced by pile-up, but also uses the topo-clusters as a signal base for a local hadronic calibration (LCW) in a non-compensating calorimeter.

Both the topo-cluster formation and the LCW calibration have been validated in collisions without pile-up recorded in 2010, and in the more active pile-up environments observed in 2011 and 2012 operations. The residual effects of pile-up on cluster kinematics and observables in data are well controlled in that they can be reproduced with sufficient precision in MC simulations for topo-clusters either inside or outside jets. The largest observed data–MC differences mainly arise from imperfect modelling of the soft collision physics affecting pile-up. Overlaying pile-up from data on gener-
Fig. 44 The average pile-up dependence of various geometric observables reconstructed from the leading topo-cluster in anti-$k_t$ jets reconstructed with $R = 0.4$ and $30 \text{ GeV} < p_L^\text{LCW+JES} < 40 \text{ GeV}$ in $Z \rightarrow \mu\mu$ events in 2012 data and MC simulations with fully simulated pile-up. The average cluster length, represented by $\log_{10}(\langle \lambda^2/\lambda_0^2 \rangle^{1/2})$ with the reference scale $\lambda_0 = 1 \text{ mm}$, is shown as a function of $\mu$ in (a), for three detector regions. The average size $\langle R^\text{lead}_{\eta\phi} \rangle$ of the leading topo-cluster in $(\eta, \phi)$ space is displayed for the same detector regions and as a function of $\mu$ in (b). The average normalised lateral energy dispersion $\langle m^2_{\text{lat}} \rangle$ of the cluster, as a function of $\mu$ for the three detector regions, is shown in (c). The shaded bands shown around the results obtained from MC simulations indicate statistical uncertainties.

From the LHC Run 1 experience, topo-clusters are now established as a well-performing signal base for jet and transverse missing momentum ($E_{\text{miss}}^T$) reconstruction in ATLAS. They provide noise suppression important for a high-quality calorimeter signal, and in this reduce the amount of data needed to represent the final state in the detector. Their spatial resolution allows not only detailed analysis of the energy flow in the proton–proton collision events as needed for $E_{\text{miss}}^T$ reconstruction but also analysis of more localised energy-flow structures inside jets. This is done routinely in boosted-object reconstruction techniques applied in jet substructure analysis, with recent examples from ATLAS discussed inRefs. [55–58].

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; SRNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Rus-
sian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SBC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, UK; DOE and NSF, USA. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partagé le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [59].

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

References


1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany, NY, USA
3 Department of Physics, University of Alberta, Edmonton, AB, Canada
4 (a)Department of Physics, Ankara University, Ankara, Turkey; (b)Istanbul Aydin University, Istanbul, Turkey; (c)Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5 LAPP, CNRS/IN2P3 and Université Savoie Mont Blanc, Annecy-le-Vieux, France
6 High Energy Physics Division, Argonne National Laboratory, Argonne, IL, USA
7 Department of Physics, University of Arizona, Tucson, AZ, USA
8 Department of Physics, The University of Texas at Arlington, Arlington, TX, USA
9 Physics Department, University of Athens, Athens, Greece
10 Physics Department, National Technical University of Athens, Zografou, Greece
11 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
12 Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain
13 Institute of Physics, University of Belgrade, Belgrade, Serbia
14 Department for Physics and Technology, University of Bergen, Bergen, Norway
15 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, USA
16 Department of Physics, Humboldt University, Berlin, Germany
17 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
18 School of Physics and Astronomy, University of Birmingham, Birmingham, UK
19 (a)Department of Physics, Bogazici University, Istanbul, Turkey; (b)Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey; (c)Department of Physics, Dogus University, Istanbul, Turkey
20 Centro de Investigaciones, Universidad Antonio Narino, Bogotá, Colombia
21 (a)INFN Sezione di Bologna, Bologna, Italy; (b)Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy
22 Physikalisches Institut, University of Bonn, Bonn, Germany
23 Department of Physics, Boston University, Boston, MA, USA
24 Department of Physics, Brandeis University, Waltham, MA, USA
Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil; Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei, Brazil; Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil

Physics Department, Brookhaven National Laboratory, Upton, NY, USA

Transilvania University of Brasov, Brasov, Romania; National Institute of Physics and Nuclear Engineering, Bucharest, Romania; Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj Napoca, Romania; University Politehnica Bucharest, Bucharest, Romania; West University in Timisoara, Timisoara, Romania

Departamento de Fisica, Universidad de Buenos Aires, Buenos Aires, Argentina

Cavendish Laboratory, University of Cambridge, Cambridge, UK

Department of Physics, Carleton University, Ottawa, ON, Canada

CERN, Geneva, Switzerland

Enrico Fermi Institute, University of Chicago, Chicago, IL, USA

Departamento de Fisica, Pontificia Universidad Católica de Chile, Santiago, Chile; Departamento de Fisica, Universidad Técnica Federico Santa María, Valparaíso, Chile

Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui, China; Department of Physics, Nanjing University, Nanjing, Jiangsu, China; School of Physics, Shandong University, Jinan, Shandong, China; Department of Physics and Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology (also affiliated with PKU-CHEP), Shanghai Jiao Tong University, Shanghai, China; Physics Department, Tsinghua University, Beijing 100084, China

Laboratoire de Physique Corpusculaire, Clermont Université and Université Blaise Pascal and CNRS/IN2P3, Clermont-Ferrand, France

Nevis Laboratory, Columbia University, Irvington, NY, USA

Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy; Dipartimento di Fisica, Università della Calabria, Rende, Italy

Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Kraków, Poland; Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland

Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, Poland

Physics Department, Southern Methodist University, Dallas, TX, USA

Physics Department, University of Texas at Dallas, Richardson, TX, USA

DESY, Hamburg and Zeuthen, Germany

Institut für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany

Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany

Department of Physics, Duke University, Durham, NC, USA

SUPA-School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK

INFN Laboratori Nazionali di Frascati, Frascati, Italy

Fakultät für Mathematik und Physik, Albert-Ludwigs-Universität, Freiburg, Germany

Section de Physique, Université de Genève, Geneva, Switzerland

INFN Sezione di Genova, Genoa, Italy; Dipartimento di Fisica, Università di Genova, Genoa, Italy

E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia; High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia

II Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany

SUPA-School of Physics and Astronomy, University of Glasgow, Glasgow, UK

II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany

Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France

Department of Physics, Hampton University, Hampton, VA, USA

Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA, USA

Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; ZITI Institut für technische Informatik, Ruprecht-Karls-Universität Heidelberg, Mannheim, Germany

Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima, Japan
(a) Department of Physics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong; (b) Department of Physics, The University of Hong Kong, Hong Kong, China; (c) Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

62 Department of Physics, Indiana University, Bloomington, IN, USA
63 Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria
64 University of Iowa, Iowa City, IA, USA
65 Department of Physics and Astronomy, Iowa State University, Ames, IA, USA
66 Joint Institute for Nuclear Research, JINR Dubna, Dubna, Russia
67 KEK, High Energy Accelerator Research Organization, Tsukuba, Japan
68 Graduate School of Science, Kobe University, Kobe, Japan
69 Faculty of Science, Kyoto University, Kyoto, Japan
70 Kyoto University of Education, Kyoto, Japan
71 Department of Physics, Kyushu University, Fukuoka, Japan
72 Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
73 Physics Department, Lancaster University, Lancaster, UK
74 (a) INFN Sezione di Lecce, Lecce, Italy; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy
75 Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK
76 Department of Physics, Jožef Stefan Institute and University of Ljubljana, Ljubljana, Slovenia
77 School of Physics and Astronomy, Queen Mary University of London, London, UK
78 Department of Physics, Royal Holloway University of London, Surrey, UK
79 Department of Physics and Astronomy, University College London, London, UK
80 Louisiana Tech University, Ruston, LA, USA
81 Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France
82 Fysiska institutionen, Lunds universitet, Lund, Sweden
83 Departamento de Física Teorica C-15, Universidad Autonoma de Madrid, Madrid, Spain
84 Institut für Physik, Universität Mainz, Mainz, Germany
85 School of Physics and Astronomy, University of Manchester, Manchester, UK
86 CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
87 Department of Physics, University of Massachusetts, Amherst, MA, USA
88 Department of Physics, McGill University, Montreal, QC, Canada
89 School of Physics, University of Melbourne, Melbourne VIC, Australia
90 Department of Physics, The University of Michigan, Ann Arbor, MI, USA
91 Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
92 (a) INFN Sezione di Milano, Milan, Italy; (b) Dipartimento di Fisica, Università di Milano, Milan, Italy
93 B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus
94 National Scientific and Educational Centre for Particle and High Energy Physics, Minsk, Republic of Belarus
95 Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA
96 Group of Particle Physics, University of Montreal, Montreal, QC, Canada
97 P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
98 Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
99 National Research Nuclear University MEPhI, Moscow, Russia
100 D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
101 Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany
102 Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Munich, Germany
103 Nagasaki Institute of Applied Science, Nagasaki, Japan
104 Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
105 (a) INFN Sezione di Napoli, Naples, Italy; (b) Dipartimento di Fisica, Università di Napoli, Naples, Italy
106 Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA
107 Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
108 Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands
109 Department of Physics, Northern Illinois University, DeKalb, IL, USA
110 Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
111 Department of Physics, New York University, New York, NY, USA
112 Ohio State University, Columbus, OH, USA
113 Faculty of Science, Okayama University, Okayama, Japan
114 Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, OK, USA
115 Department of Physics, Oklahoma State University, Stillwater, OK, USA
116 Palacký University, RCPTM, Olomouc, Czech Republic
117 Center for High Energy Physics, University of Oregon, Eugene, OR, USA
118 LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
119 Graduate School of Science, Osaka University, Osaka, Japan
120 Department of Physics, University of Oslo, Oslo, Norway
121 Department of Physics, Oxford University, Oxford, UK
122 (a)INFN Sezione di Pavia, Pavia, Italy; (b) Dipartimento di Fisica, Università di Pavia, Pavia, Italy
123 Department of Physics, University of Pennsylvania, Philadelphia, PA, USA
124 National Research Centre “Kurchatov Institute” B.P. Konstantinov Petersburg Nuclear Physics Institute, St. Petersburg, Russia
125 (a) INFN Sezione di Pisa, Pisa, Italy; (b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa, Italy
126 Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, USA
127 (a) Laboratório de Instrumentação e Física Experimental de Partículas-LIP, Lisbon, Portugal; (b) Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal; (c) Departamento de Física, Universidade de Coimbra, Coimbra, Portugal; (d) Centro de Física Nuclear da Universidade de Lisboa, Lisbon, Portugal; (e) Departamento de Física, Universidade do Minho, Braga, Portugal; (f) Departamento de Física Teorica y del Cosmos and CAFPE, Universidad de Granada, Granada, Spain; (g) Dep Fisica and CEFITEC of Faculdade de Ciencias e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
128 Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
129 Czech Technical University in Prague, Prague, Czech Republic
130 Faculty of Mathematics and Physics, Charles University in Prague, Prague, Czech Republic
131 State Research Center Institute for High Energy Physics (Protvino), NRC KI, Moscow, Russia
132 Particle Physics Department, Rutherford Appleton Laboratory, Didcot, UK
133 (a) INFN Sezione di Roma, Rome, Italy; (b) Dipartimento di Fisica, Sapienza Università di Roma, Rome, Italy
134 (a) INFN Sezione di Roma Tor Vergata, Rome, Italy; (b) Dipartimento di Fisica, Università di Roma Tor Vergata, Rome, Italy
135 (a) INFN Sezione di Roma Tre, Rome, Italy; (b) Dipartimento di Matematica e Fisica, Università Roma Tre, Rome, Italy
136 (a) Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies-Université Hassan II, Casablanca, Morocco; (b) Centre National de l’Energie des Sciences Techniques Nucleaires, Rabat, Morocco; (c) Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech, Marrakech, Morocco; (d) Faculté des Sciences, Université Mohammed Premier and LPTPM, Oujda, Morocco; (e) Faculté des Sciences, Université Mohammed V, Rabat, Morocco
137 DSM/IRFU (Institut de Recherches sur les Lois Fondamentales de l’Univers), CEA Saclay (Commissariat à l’Energie Atomique et aux Energies Alternatives), Gif-sur-Yvette, France
138 Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz, CA, USA
139 Department of Physics, University of Washington, Seattle, WA, USA
140 Department of Physics and Astronomy, University of Sheffield, Sheffield, UK
141 Department of Physics, Shinshu University, Nagano, Japan
142 Fachbereich Physik, Universität Siegen, Siegen, Germany
143 Department of Physics, Simon Fraser University, Burnaby, BC, Canada
144 SLAC National Accelerator Laboratory, Stanford, CA, USA
145 (a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
146 (a) Department of Physics, University of Cape Town, Cape Town, South Africa; (b) Department of Physics, University of Johannesburg, Johannesburg, South Africa; (c) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
Also at Louisiana Tech University, Ruston LA, USA
Also at Institutio Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
Also at Graduate School of Science, Osaka University, Osaka, Japan
Also at Department of Physics, National Tsing Hua University, Hsinchu, Taiwan
Also at Department of Physics, The University of Texas at Austin, Austin TX, USA
Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia
Also at CERN, Geneva, Switzerland
Also at Georgian Technical University (GTU), Tbilisi, Georgia
Also at Ochadai Academic Production, Ochanomizu University, Tokyo, Japan
Also at Manhattan College, New York NY, USA
Also at Hellenic Open University, Patras, Greece
Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
Also at at LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
Also at Academy Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
Also at School of Physics, Shandong University, Shandong, China
Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
Also at Section de Physique, Université de Genève, Geneva, Switzerland
Also at International School for Advanced Studies (SISSA), Trieste, Italy
Also at Department of Physics and Astronomy, University of South Carolina, Columbia SC, USA
Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China
Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria
Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
Also at National Research Nuclear University MEPhI, Moscow, Russia
Also at Department of Physics, Stanford University, Stanford CA, USA
Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
Also at Flensburg University of Applied Sciences, Flensburg, Germany
Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
*Deceased