Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton–proton collisions in the ATLAS detector

ATLAS Collaboration* CERN, 1211 Geneva 23, Switzerland

Received: 30 September 2016 / Accepted: 21 March 2017 © CERN for the benefit of the ATLAS collaboration 2017. This article is an open access publication

Abstract The reconstruction and calibration algorithms used to calculate missing transverse momentum ($E_{\text{T}}^{\text{miss}}$) with the ATLAS detector exploit energy deposits in the calorimeter and tracks reconstructed in the inner detector as well as the muon spectrometer. Various strategies are used to suppress effects arising from additional proton–proton interactions, called pileup, concurrent with the hard-scatter processes. Tracking information is used to distinguish contributions from the pileup interactions using their vertex separation along the beam axis. The performance of the $E_{\text{T}}^{\text{miss}}$ reconstruction algorithms, especially with respect to the amount of pileup, is evaluated using data collected in proton–proton collisions at a centre-of-mass energy of 8 TeV during 2012, and results are shown for a data sample corresponding to an integrated luminosity of 20.3 fb$^{-1}$. The simulation and modelling of $E_{\text{T}}^{\text{miss}}$ in events containing a Z boson decaying to two charged leptons (electrons or muons) or a W boson decaying to a charged lepton and a neutrino are compared to data. The acceptance for different event topologies, with and without high transverse momentum neutrinos, is shown for a range of threshold criteria for $E_{\text{T}}^{\text{miss}}$, and estimates of the systematic uncertainties in the $E_{\text{T}}^{\text{miss}}$ measurements are presented.

Contents
1 Introduction .................................................. 4.1.1 Reconstruction and calibration of the $E_{\text{T}}^{\text{miss}}$ hard terms .................................................. 4.1.2 Reconstruction and calibration of the $E_{\text{T}}^{\text{miss}}$ soft term .................................................. 4.1.3 Jet $p_{\text{T}}$ threshold and JVF selection ...
4.2 Track $E_{\text{T}}^{\text{miss}}$ ........................................... 5 Comparison of $E_{\text{T}}^{\text{miss}}$ distributions in data and MC simulation .......... 5.1 Modelling of $Z \rightarrow \ell\ell$ events ...........
5.2 Modelling of $W \rightarrow \ell\nu$ events ..........
6 Performance of the $E_{\text{T}}^{\text{miss}}$ in data and MC simulation . 6.1 Resolution of $E_{\text{T}}^{\text{miss}}$ ........................................ 6.1.1 Resolution of the $E_{\text{T}}^{\text{miss}}$ as a function of the number of reconstructed vertices ...........................................
6.1.2 Resolution of the $E_{\text{T}}^{\text{miss}}$ as a function of $\Sigma E_{\text{T}}$ .......... 6.2 The $E_{\text{T}}^{\text{miss}}$ response ........................................ 6.2.1 Measuring $E_{\text{T}}^{\text{miss}}$ recoil versus $p_{\text{T}}^{Z}$ .......... 6.2.2 Measuring $E_{\text{T}}^{\text{miss}}$ response in simulated $W \rightarrow \ell\nu$ events ..............
6.3 The $E_{\text{T}}^{\text{miss}}$ angular resolution .......................... 6.4 Transverse mass in $W \rightarrow \ell\nu$ events ............
6.5 Proxy for $E_{\text{T}}^{\text{miss}}$ significance ......................... 6.6 Tails of $E_{\text{T}}^{\text{miss}}$ distributions ......................... 6.7 Correlation of fake $E_{\text{T}}^{\text{miss}}$ between algorithms ..
7 Jet-$p_{\text{T}}$ threshold and vertex association selection ....
8 Systematic uncertainties of the soft term ...........
8.1 Methodology for CST ....................................... 8.1.1 Evaluation of balance between the soft term and the hard term ...........................................
8.1.2 Cross-check method for the CST systematic uncertainties ...........................................
8.2 Methodology for TST and Track $E_{\text{T}}^{\text{miss}}$ ....
8.2.1 Propagation of systematic uncertainties ..
8.2.2 Closure of systematic uncertainties ........
8.2.3 Systematic uncertainties from tracks inside jets ........
9 Conclusions ...............................................
1 Introduction

The Large Hadron Collider (LHC) provided proton–proton (pp) collisions at a centre-of-mass energy of 8 TeV during 2012. Momentum conservation transverse to the beam axis implies that the transverse momenta of all particles in the final state should sum to zero. Any imbalance may indicate the presence of undetectable particles such as neutrinos or new, stable particles escaping detection.

The missing transverse momentum ($E_T^{\text{miss}}$) is reconstructed as the negative vector sum of the transverse momenta of all detected particles, and its magnitude is represented by the symbol $E_T^{\text{miss}}$. The measurement of $E_T^{\text{miss}}$ strongly depends on the energy scale and resolution of the reconstructed “physics objects”. The physics objects considered in the $E_T^{\text{miss}}$ calculation are electrons, photons, muons, $	au$-leptons, and jets. Momentum contributions not attributed to any of the physics objects mentioned above are reconstructed as the $E_T^{\text{miss}}$ “soft term”. Several algorithms for reconstructing the $E_T^{\text{miss}}$ soft term utilizing a combination of calorimeter signals and tracks in the inner detector are considered.

The $E_T^{\text{miss}}$ reconstruction algorithms and calibrations developed by ATLAS for 7 TeV data from 2010 are summarized in Ref. [1]. The 2011 and 2012 datasets are more affected by contributions from additional pp collisions, referred to as “pileup”, concurrent with the hard-scatter process. Various techniques have been developed to suppress such contributions. This paper describes the pileup dependence, calibration, and resolution of the $E_T^{\text{miss}}$ reconstructed with different algorithms and pileup-mitigation techniques.

The performance of $E_T^{\text{miss}}$ reconstruction algorithms, or “$E_T^{\text{miss}}$ performance”, refers to the use of derived quantities like the mean, width, or tail of the $E_T^{\text{miss}}$ distribution to study pileup dependence and calibration. The $E_T^{\text{miss}}$ reconstructed with different algorithms is studied in both data and Monte Carlo (MC) simulation, and the level of agreement between the two is compared using datasets in which events are triggered on leptonically decaying $W$ or $Z$ boson dominate. The $W$ boson sample provides events with intrinsic $E_T^{\text{miss}}$ from non-interacting particles (e.g. neutrinos). Contributions to the $E_T^{\text{miss}}$ due to mismeasurement are referred to as fake $E_T^{\text{miss}}$.

Sources of fake $E_T^{\text{miss}}$ may include $p_T$ mismeasurement, miscalibration, and particles going through un-instrumented regions of the detector. In MC simulations, the $E_T^{\text{miss}}$ from each algorithm is compared to the true $E_T^{\text{miss}}$ ($E_T^{\text{miss, True}}$), which is defined as the magnitude of the vector sum of $p_T$ of stable weakly interacting particles from the hard-scatter collision. Then the selection efficiency after a $E_T^{\text{miss}}$-threshold requirement is studied in simulated events with high-$p_T$ neutrinos (such as top-quark pair production and vector-boson fusion $H \to \tau \tau$) or possible new weakly interacting particles that escape detection (such as the lightest supersymmetric particles).

This paper is organized as follows. Section 2 gives a brief introduction to the ATLAS detector. Section 3 describes the data and MC simulation used as well as the event selections applied. Section 4 outlines how the $E_T^{\text{miss}}$ is reconstructed and calibrated while Sect. 5 presents the level of agreement between data and MC simulation in W and Z boson production events. Performance studies of the $E_T^{\text{miss}}$ algorithms on data and MC simulation are shown for samples with different event topologies in Sect. 6. The choice of jet selection criteria used in the $E_T^{\text{miss}}$ reconstruction is discussed in Sect. 7. Finally, the systematic uncertainty in the absolute scale and resolution of the $E_T^{\text{miss}}$ is discussed in Sect. 8. To provide a reference, Table 1 summarizes the different $E_T^{\text{miss}}$ terms discussed in this paper.

2 ATLAS detector

The ATLAS detector [2] is a multi-purpose particle physics apparatus with a forward-backward symmetric cylindrical geometry and nearly 47r coverage in solid angle. For tracking, the inner detector (ID) covers the pseudorapidity range of $|\eta| < 2.5$, and consists of a silicon-based pixel detector, a semiconductor tracker (SCT) based on microstrip technology, and, for $|\eta| < 2.0$, a transition radiation tracker (TRT). The ID is surrounded by a thin superconducting solenoid providing a 2 T magnetic field, which allows the measurement of the momenta of charged particles. A high-granularity electromagnetic sampling calorimeter based on lead and liquid argon (LAr) technology covers the region of $|\eta| < 3.2$. A hadronic calorimeter based on steel absorbers and plastic-scintillator tiles provides coverage for hadrons, jets, and $\tau$-leptons in the range of $|\eta| < 1.7$. LAr technology using a copper absorber is also used for the hadronic calorimeters in the end-cap region of $1.5 < |\eta| < 3.2$ and for electromagnetic and hadronic measurements with copper and tungsten absorbing materials in the forward region of $3.1 < |\eta| < 4.9$. The muon spectrometer (MS) surrounds the calorimeters. It

---

1 ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates ($r, \phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = - \ln \tan(\theta/2)$.

2 ATLAS defines stable particles as those having a mean lifetime $> 3 \times 10^{-10}$ s.
Table 1 Summary of definitions for $E_T^{\text{miss}}$ terms used in this paper

<table>
<thead>
<tr>
<th>Term</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intrinsic $E_T^{\text{miss}}$</strong></td>
<td>Missing transverse momentum arising from the presence of neutrinos or other non-interacting particles in an event. In case of simulated events the true $E_T^{\text{miss}}$ ($E_T^{\text{miss, True}}$) corresponds to the $E_T^{\text{miss}}$ in such events defined as the magnitude of the vector sum of $p_T$ of non-interacting particles computed from the generator information.</td>
</tr>
<tr>
<td><strong>Fake $E_T^{\text{miss}}$</strong></td>
<td>Missing transverse momentum arising from the miscalibration or misidentification of physics objects in the event. It is typically studied in $Z \rightarrow \mu\mu$ events where the intrinsic $E_T^{\text{miss}}$ is normally expected to be zero.</td>
</tr>
<tr>
<td><strong>Hard terms</strong></td>
<td>The component of the $E_T^{\text{miss}}$ computed from high-$p_T$ physics objects, which includes reconstructed electrons, photons, muons, $\tau$-leptons, and jets.</td>
</tr>
<tr>
<td><strong>Soft terms</strong></td>
<td>Typically low-$p_T$ calorimeter energy deposits or tracks, depending on the soft-term definition, that are not associated to physics objects included in the hard terms.</td>
</tr>
<tr>
<td><strong>Pileup-suppressed $E_T^{\text{miss}}$</strong></td>
<td>All $E_T^{\text{miss}}$ reconstruction algorithms in Sect. 4.1.2 except the Calorimeter Soft Term, which does not apply pileup suppression.</td>
</tr>
<tr>
<td><strong>Object-based</strong></td>
<td>This refers to all reconstruction algorithms in Sect. 4.1.2 except the Track $E_T^{\text{miss}}$, namely the Calorimeter Soft Term, Track Soft Term, Extrapolated Jet Area with Filter, and Soft-Term Vertex-Fraction algorithms. These consider the physics objects such as electrons, photons, muons, $\tau$-leptons, and jets during the $E_T^{\text{miss}}$ reconstruction.</td>
</tr>
</tbody>
</table>

consists of three air-core superconducting toroid magnet systems, precision tracking chambers to provide accurate muon tracking out to $|\eta| = 2.7$, and additional detectors for triggering in the region of $|\eta| < 2.4$. A precision measurement of the track coordinates is provided by layers of drift tubes at three radial positions within $|\eta| < 2.0$. For $2.0 < |\eta| < 2.7$, cathode-strip chambers with high granularity are instead used in the innermost plane. The muon trigger system consists of resistive-plate chambers in the barrel ($|\eta| < 1.05$) and thin-gap chambers in the end-cap regions ($1.05 < |\eta| < 2.4$).

### 3 Data samples and event selection

ATLAS recorded $pp$ collisions at a centre-of-mass energy of 8 TeV with a bunch crossing interval (bunch spacing) of 50 ns in 2012. The resulting integrated luminosity is 20.3 fb$^{-1}$ [3]. Multiple inelastic $pp$ interactions occurred in each bunch crossing, and the mean number of inelastic collisions per bunch crossing ($\langle n_{\text{coll}} \rangle$) over the full dataset is 21 [4], exceptionally reaching as high as about 70.

Data are analysed only if they satisfy the standard ATLAS data-quality assessment criteria [5]. Jet-cleaning cuts [5] are applied to minimize the impact of instrumental noise and out-of-time energy deposits in the calorimeter from cosmic rays or beam-induced backgrounds. This ensures that the residual sources of $E_T^{\text{miss}}$ mismeasurement due to those instrumental effects are suppressed.

#### 3.1 Track and vertex selection

The ATLAS detector measures the momenta of charged particles using the ID [6]. Hits from charged particles are recorded and are used to reconstruct tracks; these are used to reconstruct vertices [7,8].

Each vertex must have at least two tracks with $p_T > 0.4$ GeV; for the primary hard-scatter vertex (PV), the requirement on the number of tracks is raised to three. The PV in each event is selected as the vertex with the largest value of $\Sigma (p_T)^2$, where the scalar sum is taken over all the tracks matched to the vertex. The following track selection criteria$^3$ [7] are used throughout this paper, including the vertex reconstruction:

- $p_T > 0.5$ GeV (0.4 GeV for vertex reconstruction and the calorimeter soft term),
- $|\eta| < 2.5$,
- Number of hits in the pixel detector $\geq 1$,
- Number of hits in the SCT $\geq 6$.

These tracks are then matched to the PV by applying the following selections:

- $|d_0| < 1.5$ mm,
- $|z_0 \sin(\theta)| < 1.5$ mm.

The transverse (longitudinal) impact parameter $d_0$ ($z_0$) is the transverse (longitudinal) distance of the track from the PV and is computed at the point of closest approach to the PV in the plane transverse to the beam axis. The requirements on the number of hits ensures that the track has an

$^3$ The track reconstruction for electrons and for muons does not strictly follow these definitions. For example, a Gaussian Sum Filter [9] algorithm is used for electrons to improve the measurements of its track parameters, which can be degraded due to Bremsstrahlung losses.
accurate $p_T$ measurement. The $|\eta|$ requirement keeps only the tracks within the ID acceptance, and the requirement of $p_T > 0.4$ GeV ensures that the track reaches the outer layers of the ID. Tracks with low $p_T$ have large curvature and are more susceptible to multiple scattering.

The average spread along the beamline direction for $pp$ collisions in ATLAS during 2012 data taking is around 50 mm, and the typical track $z_0$ resolution for those with $|\eta| < 0.2$ and $0.5 < p_T < 0.6$ GeV is 0.34 mm. The typical track $d_0$ resolution is around 0.19 mm for the same $\eta$ and $p_T$ ranges, and both the $z_0$ and $d_0$ resolutions improve with higher track $p_T$.

Pileup effects come from two sources: in-time and out-of-time. In-time pileup is the result of multiple $pp$ interactions in the same LHC bunch crossing. It is possible to distinguish the in-time pileup interactions by using their vertex positions, which are spread along the beam axis. At $\langle \mu \rangle = 21$, the efficiency to reconstruct and select the correct vertex for $Z \rightarrow \mu\mu$ simulated events is around 93.5% and rises to more than 98% when requiring two generated muons with $p_T > 10$ GeV inside the ID acceptance [10]. When vertices are separated along the beam axis by a distance smaller than the resolution, they can be reconstructed as a single vertex. Each track in the reconstructed vertex is assigned a weight based upon its compatibility with the fitted vertex, which depends on the $\chi^2$ of the fit. The fraction of $Z \rightarrow \mu\mu$ reconstructed vertices with more than 50% of the sum of track weights coming from pileup interactions is around 3% at $\langle \mu \rangle = 21$ [7,10]. Out-of-time pileup comes from $pp$ collisions in earlier and later bunch crossings, which leave signals in the calorimeters that can take up to 450 ns for the charge collection time. This is longer than the 50 ns between subsequent collisions and occurs because the integration time of the calorimeters is significantly larger than the time between the bunch crossings. By contrast the charge collection time of the silicon tracker is less than 25 ns.

3.2 Event selection for $Z \rightarrow \ell\ell$

The “standard candle” for evaluation of the $E_T^{\text{miss}}$ performance is $Z \rightarrow \ell\ell$ events ($\ell = e$ or $\mu$). They are produced without neutrinos, apart from a very small number originating from heavy-flavour decays in jets produced in association with the $Z$ boson. The intrinsic $E_T^{\text{miss}}$ is therefore expected to be close to zero, and the $E_T^{\text{miss}}$ distributions are used to evaluate the modelling of the effects that give rise to fake $E_T^{\text{miss}}$.

Candidate $Z \rightarrow \ell\ell$ events are required to pass an electron or muon trigger [11,12]. The lowest $p_T$ threshold for the unprescaled single-electron (single-muon) trigger is $p_T > 25$ (24) GeV, and both triggers apply a track-based isolation as well as quality selection criteria for the particle identification. Triggers with higher $p_T$ thresholds, without the isolation requirements, are used to improve acceptance at high $p_T$. These triggers require $p_T > 60$ (36) GeV for electrons (muons). Events are accepted if they pass any of the above trigger criteria. Each event must contain at least one primary vertex with a $z$ displacement from the nominal $p_T$ interaction point of less than 200 mm and with at least three associated tracks.

The offline selection of $Z \rightarrow \mu\mu$ events requires the presence of exactly two identified muons [13]. An identified muon is reconstructed in the MS and is matched to a track in the ID. The combined ID+MS track must have $p_T > 25$ GeV and $|\eta| < 2.5$. The $z$ displacement of the muon track from the primary vertex is required to be less than 10 mm. An isolation criterion is applied to the muon track, where the scalar sum of the $p_T$ of additional tracks within a cone of size $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.2$ around the muon is required to be less than 10% of the muon $p_T$. In addition, the two leptons are required to have opposite charge, and the reconstructed dilepton invariant mass, $m_{\ell\ell}$, is required to be consistent with the $Z$ boson mass: $66 < m_{\ell\ell} < 116$ GeV.

The $E_T^{\text{miss}}$ modelling and performance results obtained in $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ events are very similar. For the sake of brevity, only the $Z \rightarrow \mu\mu$ distributions are shown in all sections except for Sect. 6.6.

3.3 Event selection for $W \rightarrow \ell\nu$

Leptonically decaying $W$ bosons ($W \rightarrow \ell\nu$) provide an important event topology with intrinsic $E_T^{\text{miss}}$; the $E_T^{\text{miss}}$ distribution for such events is presented in Sect. 5.2. Similar to $Z \rightarrow \ell\ell$ events, a sample dominated by leptonically decaying $W$ bosons is used to study the $E_T^{\text{miss}}$ scale in Sect. 6.2.2, the resolution of the $E_T^{\text{miss}}$ direction in Sect. 6.3, and the impact on a reconstructed kinematic observable in Sect. 6.4.

The $E_T^{\text{miss}}$ distributions for $W$ boson events in Sect. 5.2 use the electron final state. These electrons are selected with $|\eta| < 2.47$, are required to meet the “medium” identification criteria [14] and satisfy $p_T > 25$ GeV. Electron candidates in the region $1.37 < |\eta| < 1.52$ suffer from degraded momentum resolution and particle identification due to the transition from the barrel to the end-cap detector and are therefore discarded in these studies. The electrons are required to be isolated, such that the sum of the energy in the calorimeter within a cone of size $\Delta R = 0.3$ around the electron is less than 14% of the electron $p_T$. The summed $p_T$ of other tracks within the same cone is required to be less than 7% of the electron $p_T$. The calorimeter isolation variable [14] is corrected by subtracting estimated contributions from the electron itself, the underlying event [15], and pileup. The

© Springer
electron tracks are then matched to the PV by applying the following selections:

- $|d_0| < 5.0$ mm,
- $|z_0 \sin(\theta)| < 0.5$ mm.

The $W$ boson selection is based on the single-lepton triggers and the same lepton selection criteria as those used in the $Z \rightarrow \ell\ell$ selection. Events are rejected if they contain more than one reconstructed lepton. Selections on the $E_T^{\text{miss}}$ and transverse mass ($m_T$) are applied to reduce the multi-jet backgrounds with one jet misidentified as an isolated lepton. The transverse mass is calculated from the lepton and the $E_T^{\text{miss}}$,

$$m_T = \sqrt{2p_T^\ell E_T^{\text{miss}} (1 - \cos \Delta \phi)},$$

where $p_T^\ell$ is the transverse momentum of the lepton and $\Delta \phi$ is the azimuthal angle between the lepton and $E_T^{\text{miss}}$ directions. Both the $m_T$ and $E_T^{\text{miss}}$ are required to be greater than 50 GeV.

These selections can bias the event topology and its phase space, so they are only used when comparing simulation to data in Sect. 5.2, as they substantially improve the purity of $W$ bosons in data events.

The $E_T^{\text{miss}}$ modelling and performance results obtained in $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ events are very similar. For the sake of brevity, only one of the two is considered in following two sections: $E_T^{\text{miss}}$ distributions in $W \rightarrow e\nu$ events are presented in Sect. 5.2 and the performance studies show $W \rightarrow \mu\nu$ events in Sect. 6. When studying the $E_T^{\text{miss}}$ tails, both final states are considered in Sect. 6.6, because the $\eta$-coverage and reconstruction performance between muons and electrons differ.

### 3.4 Monte Carlo simulation samples

Table 2 summarizes the MC simulation samples used in this paper. The $Z \rightarrow \ell\ell$ and $W \rightarrow \ell\nu$ samples are generated with ALPGEN [16] interfaced with PYTHIA [17] (denoted by ALPGEN+PYTHIA) to model the parton shower and hadronization, and underlying event using the PERUGIA2011C set [18] of tunable parameters. One exception is the $Z \rightarrow \tau\tau$ sample with leptonically decaying $\tau$-leptons, which is generated with ALPGEN interfaced with HERWIG [19] with the underlying event modelled using JIMMY [20] and the AUET2 tunes [21]. ALPGEN is a multi-leg generator that provides tree-level calculations for diagrams with up to five additional partons. The matrix-element MC calculations are matched to a model of the parton shower, underlying event and hadronization. The main processes that are backgrounds to $Z$ bosons in the $Z\rightarrow\ell\ell$ events are presented in Sect. 5.2, as they substantially improve the purity of $W$ bosons in data events. The $E_T^{\text{miss}}$ modelling and performance results obtained in $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ events are very similar. For the sake of brevity, only one of the two is considered in following two sections: $E_T^{\text{miss}}$ distributions in $W \rightarrow e\nu$ events are presented in Sect. 5.2 and the performance studies show $W \rightarrow \mu\nu$ events in Sect. 6. When studying the $E_T^{\text{miss}}$ tails, both final states are considered in Sect. 6.6, because the $\eta$-coverage and reconstruction performance between muons and electrons differ.

### Table 2: Generators, cross-section normalizations, PDF sets, and MC tunes used in this analysis

<table>
<thead>
<tr>
<th>Sample</th>
<th>Generator</th>
<th>Use</th>
<th>Cross-section</th>
<th>PDF set</th>
<th>Tune</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \rightarrow \mu\mu$</td>
<td>ALPGEN+PYTHIA</td>
<td>Signal</td>
<td>NNLO [26]</td>
<td>CTEQ6L1 [27]</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$Z \rightarrow ee$</td>
<td>ALPGEN+PYTHIA</td>
<td>Signal</td>
<td>NNLO [26]</td>
<td>CTEQ6L1 [27]</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>ALPGEN+HERWIG</td>
<td>Signal</td>
<td>NNLO [26]</td>
<td>CTEQ6L1 [27]</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$W \rightarrow \mu\nu$</td>
<td>ALPGEN+PYTHIA</td>
<td>Signal</td>
<td>NNLO [26]</td>
<td>CTEQ6L1 [27]</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$W \rightarrow ev$</td>
<td>ALPGEN+PYTHIA</td>
<td>Signal</td>
<td>NNLO [26]</td>
<td>CTEQ6L1 [27]</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$W \rightarrow tv$</td>
<td>ALPGEN+PYTHIA</td>
<td>Signal</td>
<td>NNLO [26]</td>
<td>CTEQ6L1 [27]</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>POWHEG+PYTHIA</td>
<td>Signal/background</td>
<td>NNLO+NNLL [28,29]</td>
<td>CTEQ6L1</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>VBF $H \rightarrow \tau\tau$</td>
<td>POWHEG+PYTHIA8</td>
<td>Signal</td>
<td>–</td>
<td>NLO CT10 [30]</td>
<td>AU2 [31]</td>
</tr>
<tr>
<td>SUSY 500</td>
<td>HERWIG++</td>
<td>Signal</td>
<td>–</td>
<td>CTEQ6L1</td>
<td>UE EE3 [32]</td>
</tr>
<tr>
<td>$W^\pm Z \rightarrow \ell^\pm\nu^\mp\ell^\mp$</td>
<td>SHERPA</td>
<td>Background</td>
<td>NLO [33,34]</td>
<td>NLO CT10 [30]</td>
<td>SHERPA default</td>
</tr>
<tr>
<td>$ZZ \rightarrow \ell^+\ell^-\ell^+\ell^-$</td>
<td>SHERPA</td>
<td>Background</td>
<td>NLO [33,34]</td>
<td>NLO CT10 [30]</td>
<td>SHERPA default</td>
</tr>
<tr>
<td>$W^+W^- \rightarrow \ell^+\ell^-\nu\nu$</td>
<td>SHERPA</td>
<td>Background</td>
<td>NLO [33,34]</td>
<td>NLO CT10 [30]</td>
<td>SHERPA default</td>
</tr>
<tr>
<td>$tW$</td>
<td>POWHEG+PYTHIA</td>
<td>Background</td>
<td>NNLO+NNLL [35]</td>
<td>CTEQ6L1</td>
<td>PERUGIA2011C [18]</td>
</tr>
<tr>
<td>$Z \rightarrow \mu\mu$</td>
<td>POWHEG+PYTHIA8</td>
<td>Systematic effects</td>
<td>NNLO [36,37]</td>
<td>NLO CT10 [30]</td>
<td>AU2</td>
</tr>
<tr>
<td>$Z \rightarrow \mu\mu$</td>
<td>ALPGEN+HERWIG</td>
<td>Systematic effects</td>
<td>NNLO [36,37]</td>
<td>CTEQ6L1</td>
<td>AUET2</td>
</tr>
<tr>
<td>$Z \rightarrow \mu\mu$</td>
<td>SHERPA</td>
<td>Systematic effects</td>
<td>NNLO [36,37]</td>
<td>NLO CT10 [30]</td>
<td>SHERPA default</td>
</tr>
</tbody>
</table>
with at least one leptonically decaying W boson are considered in Sect. 6.6. The single top quark (tW) production is considered with at least one leptonically decaying W boson. Both the tt and tW processes contribute to the W and Z boson distributions shown in Sect. 5 as well as Z boson distributions in Sects. 4, 6, and 8 that compare data and simulation. A supersymmetric (SUSY) model comprising pair-produced 500 GeV gluinos each decaying to a tt pair and a neutralino is simulated with HERWIG++ [24]. Finally, to study events with forward jets, the vector-boson fusion (VBF) production of \( H \rightarrow \tau \tau \), generated with POWHEG+PYTHIA8 [25], is considered. Both \( \tau \)-leptons are forced to decay leptonically in this sample.

To estimate the systematic uncertainties in the data/MC ratio arising from the modelling of the soft hadronic recoil, \( E_{\text{miss}}^{\tau} \) distributions simulated with different MC generators, parton shower and underlying event models are compared. The estimation of systematic uncertainties is performed using a comparison of data and MC simulation, as shown in Sect. 8.2. The following combinations of generators and parton shower models are considered: SHERPA, ALPGEN+HERWIG, ALPGEN+PYTHIA, and POWHEG+PYTHIA8. The corresponding underlying event tunes are mentioned in Table 2. Parton distribution functions are taken from CT10 [30] for POWHEG and SHERPA samples and CTEQ6L1 [38] for ALPGEN samples.

Generated events are propagated through a GEANT4 simulation [39,40] of the ATLAS detector. Pileup collisions are generated with PYTHIA8 for all samples, and are overlaid on top of simulated hard-scatter events before event reconstruction. Each simulation sample is weighted by its corresponding cross-section and normalized to the integrated luminosity of the data.

### 4 Reconstruction and calibration of the \( E_{\text{T}}^{\text{miss}} \)

Several algorithms have been developed to reconstruct the \( E_{\text{T}}^{\text{miss}} \) in ATLAS. They differ in the information used to reconstruct the \( p_T \) of the particles, using either energy deposits in the calorimeters, tracks reconstructed in the ID, or both. This section describes these various reconstruction algorithms, and the remaining sections discuss the agreement between data and MC simulation as well as performance studies.

#### 4.1 Reconstruction of the \( E_{\text{T}}^{\text{miss}} \)

The \( E_{\text{T}}^{\text{miss}} \) reconstruction uses calibrated physics objects to estimate the amount of missing transverse momentum in the detector. The \( E_{\text{T}}^{\text{miss}} \) is calculated using the components along the \( x \) and \( y \) axes:

\[
E_{\text{T}}^{\text{miss}} = E_{\text{T}}^{\text{miss},\text{e}} + E_{\text{T}}^{\text{miss},\gamma} + E_{\text{T}}^{\text{miss},\tau} + E_{\text{T}}^{\text{miss},\text{jets}} + E_{\text{T}}^{\text{miss},\mu} + E_{\text{T}}^{\text{miss},\text{soft}},
\]

where each term is calculated as the negative vectorial sum of transverse momenta of energy deposits and/or tracks. To avoid double counting, energy deposits in the calorimeters and tracks are matched to reconstructed physics objects in the following order: electrons (e), photons (\( \gamma \)), the visible parts of hadronically decaying \( \tau \)-leptons (\( \tau_{\text{had-vis}} \); labelled as \( \tau \)), jets and muons (\( \mu \)). Each type of physics object is represented by a separate term in Eq. (2). The signals not associated with physics objects form the “soft term”, whereas those associated with the physics objects are collectively referred to as the “hard term”.

The magnitude and azimuthal angle\(^4\) (\( \phi^{\text{miss}} \)) of \( E_{\text{T}}^{\text{miss}} \) are calculated as:

\[
E_{\text{T}}^{\text{miss}} = \sqrt{(E_{\text{miss}}^{\tau})^2 + (E_{\text{miss}}^{\mu})^2},
\]

\[
\phi^{\text{miss}} = \arctan(E_{\text{miss}}^{\tau}/E_{\text{miss}}^{\mu}).
\]

The total transverse energy in the detector, labelled as \( \Sigma E_{\text{T}} \), quantifies the total event activity and is an important observable for understanding the resolution of the \( E_{\text{T}}^{\text{miss}} \), especially with increasing pileup contributions. It is defined as:

\[
\Sigma E_{\text{T}} = \sum p_{\text{T}}^{\tau} + \sum p_{\text{T}}^{\mu} + \sum p_{\text{T}}^{\text{jets}} + \sum p_{\text{T}}^{\text{soft}},
\]

which is the scalar sum of the transverse momenta of reconstructed physics objects and soft-term signals that contribute to the \( E_{\text{T}}^{\text{miss}} \) reconstruction. The physics objects included in \( \Sigma p_{\text{T}}^{\text{soft}} \) depend on the \( E_{\text{T}}^{\text{miss}} \) definition, so both calorimeter objects and track-based objects may be included in the sum, despite differences in \( p_{\text{T}} \) resolution.

#### 4.1.1 Reconstruction and calibration of the \( E_{\text{T}}^{\text{miss}} \) hard terms

The hard term of the \( E_{\text{T}}^{\text{miss}} \), which is computed from the reconstructed electrons, photons, muons, \( \tau \)-leptons, and jets, is described in more detail in this section.

Electrons are reconstructed from clusters in the electromagnetic (EM) calorimeter which are associated with an ID track [14]. Electron identification is restricted to the range of \( |\eta| < 2.47 \), excluding the transition region between the barrel and end-cap EM calorimeters, 1.37 < \( |\eta| < 1.52 \). They are calibrated at the EM scale\(^5\) with the default electron calibration.

---

\(^4\) The \( \arctan \) function returns values from \( [-\pi, +\pi] \) and uses the sign of both coordinates to determine the quadrant.

\(^5\) The EM scale is the basic signal scale for the ATLAS calorimeters. It accounts correctly for the energy deposited by EM showers in the calorimeter, but it does not consider energy losses in the uninstrumented material.
tion, and those satisfying the “medium” selection criteria [14] with $p_T > 10$ GeV are included in the $E_T^{\text{miss}}$ reconstruction.

The photon reconstruction is also seeded from clusters of energy deposited in the EM calorimeter and is designed to separate electrons from photons. Photons are calibrated at the EM scale and are required to satisfy the “tight” photon selection criteria with $p_T > 10$ GeV [14].

Muon candidates are identified by matching an ID track with an MS track or segment [13]. MS tracks are used for $2.5 < |\eta| < 2.7$ to extend the $\eta$ coverage. Muons are required to satisfy $p_T > 5$ GeV to be included in the $E_T^{\text{miss}}$ reconstruction. The contribution of muon energy deposited in the calorimeter is taken into account using either parameterized estimates or direct measurements, to avoid double counting a small fraction of their momenta.

Jets are reconstructed from three-dimensional topological clusters (topoclusters) [41] of energy deposits in the calorimeter using the anti-$k_t$ algorithm [42] with a distance parameter $R = 0.4$. The topological clustering algorithm suppresses noise by forming contiguous clusters of calorimeter cells with significant energy deposits. The local cluster weighting (LCW) [43,44] calibration is used to account for different calorimeter responses to electrons, photons and hadrons. Each cluster is classified as coming from an EM or hadronic shower, using information from its shape and energy density, and calibrated accordingly. The jets are reconstructed from calibrated topoclusters and then corrected for in-time and out-of-time pileup as well as the position of the PV [4]. Finally, the jet energy scale (JES) corrects for jet-level effects by restoring, on average, the energy of reconstructed jets to that of the MC generator-level jets. The complete procedure is referred to as the LCW+JES scheme [43,44]. Without changing the average calibration, additional corrections are made based upon the internal properties of the jet (global sequential calibration) to reduce the flavour dependence and energy leakage effects [44]. Only jets with calibrated $p_T$ greater than 20 GeV are used to calculate the jet term $E_T^{\text{miss,jets}}$ in Eq. (2), and the optimization of the 20 GeV threshold is discussed in Sect. 7.

To suppress contributions from jets originating from pileup interactions, a requirement on the jet vertex-fraction (JVF) [4] may be applied to selected jet candidates. Tracks matched to jets are extrapolated back to the beamline to ascertain whether they originate from the hard scatter or from a pileup collision. The JVF is then computed as the ratio shown below:

$$\text{JVF} = \frac{\sum_{\text{track, PV, jet}} p_T}{\sum_{\text{track, jet}} p_T}.$$  \hfill (5)

This is the ratio of the scalar sum of transverse momentum of all tracks matched to the jet and the primary vertex to the $p_T$ sum of all tracks matched to the jet, where the sum is performed over all tracks with $p_T > 0.5$ GeV and $|\eta| < 2.5$

and the matching is performed using the “ghost-association” procedure [45,46].

The JVF distribution is peaked toward 1 for hard-scatter jets and toward 0 for pileup jets. No JVF selection requirement is applied to jets that have no associated tracks. Requirements on the JVF are made in the STVF, EJAF, and TST $E_T^{\text{miss}}$ algorithms as described in Table 3 and Sect. 4.1.3.

Hadronically decaying $\tau$-leptons are seeded by calorimeter jets with $|\eta| < 2.5$ and $p_T > 10$ GeV. As described for jets, the LCW calibration is applied, corrections are made to subtract the energy due to pileup interactions, and the energy of the hadronically decaying $\tau$ candidates is calibrated at the $\tau$-lepton energy scale (TES) [47]. The TES is independent of the JES and is determined using an MC-based procedure. Hadronically decaying $\tau$-leptons passing the “medium” requirements [47] and having $p_T > 20$ GeV after TES corrections are considered for the $E_T^{\text{miss}}$ reconstruction.

4.1.2 Reconstruction and calibration of the $E_T^{\text{miss}}$ soft term

The soft term is a necessary but challenging ingredient of the $E_T^{\text{miss}}$ reconstruction. It comprises all the detector signals not matched to the physics objects defined above and can contain contributions from the hard scatter as well as the underlying event and pileup interactions. Several algorithms designed to reconstruct and calibrate the soft term have been developed, as well as methods to suppress the pileup contributions. A summary of the $E_T^{\text{miss}}$ and soft-term reconstruction algorithms is given in Table 3.

Four soft-term reconstruction algorithms are considered in this paper. Below the first two are defined, and then some motivation is given for the remaining two prior to their definition.

- **Calorimeter Soft Term (CST)**
  This reconstruction algorithm [1] uses information mainly from the calorimeter and is widely used by ATLAS. The algorithm also includes corrections based on tracks but does not attempt to resolve the various $pp$ interactions based on the track $z_0$ measurement. The soft term is referred to as the CST, whereas the entire $E_T^{\text{miss}}$ is written as CST $E_T^{\text{miss}}$. Corresponding naming schemes are used for the other reconstruction algorithms. The CST is reconstructed using energy deposits in the calorimeter which are not matched to the high-$p_T$ physics objects used in the $E_T^{\text{miss}}$. To avoid fake signals in the calorimeter, noise suppression is important. This is achieved by calculating the soft term using only cells belonging to topoclusters, which are calibrated at the LCW scale [43,44]. The tracker and calorimeter provide redundant $p_T$ measurements for charged particles, so an energy-flow algorithm is used to determine which measurement to use. Tracks
with $p_T > 0.4$ GeV that are not matched to a high-$p_T$ physics objects are used instead of the calorimeter $p_T$ measurement, if their $p_T$ resolution is better than the expected calorimeter $p_T$ resolution. The calorimeter resolution is estimated as $0.4 \cdot \sqrt{p_T}$ GeV, in which the $p_T$ is the transverse momentum of the reconstructed track. Geometrical matching between tracks and topoclusters (or high-$p_T$ physics objects) is performed using the $\Delta R$ significance defined as $\Delta R/\sigma_{\Delta R}$, where $\sigma_{\Delta R}$ is the $\Delta R$ resolution, parameterized as a function of the track $p_T$. A track is considered to be associated to a topocluster in the soft term when its minimum $\Delta R/\sigma_{\Delta R}$ is less than 4. To veto tracks matched to high-$p_T$ physics objects, tracks are required to have $\Delta R/\sigma_{\Delta R} > 8$. The $E_T^{\text{miss}}$ calculated using the CST algorithm is documented in previous publications such as Ref. [1] and is the standard algorithm in most ATLAS 8 TeV analyses.

- **Track Soft Term (TST)**
  The TST is reconstructed purely from tracks that pass the selections outlined in Sect. 3.1 and are not associated with the high-$p_T$ physics objects defined in Sect. 4.1.1. The detector coverage of the TST is the ID tracking volume ($|\eta| < 2.5$), and no calorimeter topoclusters inside or beyond this region are included. This algorithm allows excellent vertex matching for the soft term, which almost completely removes the in-time pileup dependence, but misses contributions from soft neutral particles. The track-based reconstruction also entirely removes the out-of-time pileup contributions that affect the CST.

To avoid double counting the $p_T$ of particles, the tracks matched to the high-$p_T$ physics objects need to be removed from the soft term. All of the following classes of tracks are excluded from the soft term:

- tracks within a cone of size $\Delta R = 0.05$ around electrons and photons
- tracks within a cone of size $\Delta R = 0.2$ around $\tau_{\text{had-vis}}$
- ID tracks associated with identified muons
- tracks matched to jets using the ghost-association technique described in Sect. 4.1.1
- isolated tracks with $p_T \geq 120$ GeV ($\geq 200$ GeV for $|\eta| < 1.5$) having transverse momentum uncertainties larger than 40% or having no associated calorimeter energy deposit with $p_T$ larger than 65% of the track $p_T$. The $p_T$ thresholds are chosen to ensure that muons not in the coverage of the MS are still included in the soft term. This is a cleaning cut to remove mis-measured tracks.

A deterioration of the CST $E_T^{\text{miss}}$ resolution is observed as the average number of pileup interactions increases [1]. All $E_T^{\text{miss}}$ terms in Eq. (2) are affected by pileup, but the terms which are most affected are the jet term and CST, because their constituents are spread over larger regions in the calorimeters than those of the $E_T^{\text{miss}}$ hard terms. Methods to suppress pileup are therefore needed, which can restore the $E_T^{\text{miss}}$ resolution to values similar to those observed in the absence of pileup.

### Table 3 Summary of $E_T^{\text{miss}}$ and soft-term reconstruction algorithms used in this paper

<table>
<thead>
<tr>
<th>Term</th>
<th>Brief description</th>
<th>Section list</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST $E_T^{\text{miss}}$</td>
<td>The Calorimeter Soft Term (CST) $E_T^{\text{miss}}$ takes its soft term from energy deposits in the calorimeter which are not matched to high-$p_T$ physics objects. Although noise suppression is applied to reduce fake signals, no additional pileup suppression techniques are used.</td>
<td>Section 4.1.2 (definition)</td>
</tr>
<tr>
<td>TST $E_T^{\text{miss}}$</td>
<td>The Track Soft Term (TST) $E_T^{\text{miss}}$ algorithm uses a soft term that is calculated using tracks within the inner detector that are not associated with high-$p_T$ physics objects. The JVF selection requirement is applied to jets.</td>
<td>Section 4.1.2 (definition)</td>
</tr>
<tr>
<td>EJAF $E_T^{\text{miss}}$</td>
<td>The Extrapolated Jet Area with Filter $E_T^{\text{miss}}$ algorithm applies pileup subtraction to the CST based on the idea of jet-area corrections. The JVF selection requirement is applied to jets.</td>
<td>Section 4.1.2 (definition)</td>
</tr>
<tr>
<td>STVF $E_T^{\text{miss}}$</td>
<td>The Soft-Term Vertex-Fraction (STVF) $E_T^{\text{miss}}$ algorithm suppresses pileup effects in the CST by scaling the soft term by a multiplicative factor calculated based on the fraction of scalar-summed track $p_T$ not associated with high-$p_T$ physics objects that can be matched to the primary vertex. The JVF selection requirement is applied to jets.</td>
<td>Section 4.1.2 (definition)</td>
</tr>
<tr>
<td>Track $E_T^{\text{miss}}$</td>
<td>The Track $E_T^{\text{miss}}$ is reconstructed entirely from tracks to avoid pileup contamination that affects the other algorithms.</td>
<td>Section 4.2 (definition)</td>
</tr>
</tbody>
</table>
The TST algorithm is very stable with respect to pileup but does not include neutral particles. Two other pileup-suppressing algorithms were developed, which consider contributions from neutral particles. One uses an \( \eta \)-dependent event-by-event estimator for the transverse momentum density from pileup, using calorimeter information, while the other applies an event-by-event global correction based on the amount of charged-particle \( p_T \) from the hard-scatter vertex, relative to all other \( pp \) collisions. The definitions of these two soft-term algorithms are described in the following:

- Extrapolated Jet Area with Filter (EJAF)
  The jet-area method for the pileup subtraction uses a soft term based on the idea of jet-area corrections \([45]\). This technique uses direct event-by-event measurements of the energy flow throughout the entire ATLAS detector to estimate the \( p_T \) density of pileup energy deposits and was developed from the strategy applied to jets as described in Ref. \([4]\).

The topoclusters belonging to the soft term are used for jet finding with the \( k_T \) algorithm \([48, 49]\) with distance parameter \( R = 0.6 \) and jet \( p_T > 0 \). The catchment areas \([45, 46]\) for these reconstructed jets are labelled \( A_{\text{jet}} \); this provides a measure of the jet’s susceptibility to contamination from pileup. Jets with \( p_T < 20 \text{ GeV} \) are referred to as soft-term jets, and the \( p_T \)-density of each soft-term jet \( i \) is then measured by computing:

\[
\rho_{\text{jet},i} = \frac{p_{\text{jet}}}{A_{\text{jet},i}}. \tag{6}
\]

In a given event, the median \( p_T \)-density \( \rho_{\text{med}} \) for all soft-term \( k_T \) jets in the event \( N_{\text{jets}} \) found within a given range \( -\eta_{\text{max}} < \eta_{\text{jet}} < \eta_{\text{max}} \) can be calculated as

\[
\rho_{\text{med}} = \text{median}\{\rho_{\text{jet},i}\} \quad \text{for } i = 1 \ldots N_{\text{jets}} \text{ in } |\eta_{\text{jet}}| < \eta_{\text{max}}. \tag{7}
\]

This median \( p_T \)-density \( \rho_{\text{med}} \) gives a good estimate of the in-time pileup activity in each detector region. If determined with \( \eta_{\text{max}} = 2 \), it is found to also be an appropriate indicator of out-of-time pileup contributions \([45]\). A lower value for \( \rho_{\text{med}} \) is computed by using jets with \( |\eta_{\text{jet}}| \) larger than 2, which is mostly due to the particular geometry of the ATLAS calorimeters and their cluster reconstruction algorithms.\(^6\)

In order to extrapolate \( \rho_{\text{med}} \) into the forward regions of the detector, the average topocluster \( p_T \) in slices of \( \eta \), \( N_{\text{PV}} \), and \( \langle \mu \rangle \) is converted to an average \( p_T \) density \( \langle \rho \rangle \) for the soft term. As described for the \( \rho_{\text{med}} \), \( \langle \rho \rangle \) is found to be uniform in the central region of the detector with \( |\eta| < \eta_{\text{plateau}} = 1.8 \). The transverse momentum density profile is then computed as

\[
P_{\text{jet}}^\rho(\eta, N_{\text{PV}}, \langle \mu \rangle) = \frac{\langle \rho \rangle(\eta, N_{\text{PV}}, \mu)}{\langle \rho \rangle_{\text{central}}(N_{\text{PV}}, \mu)} \tag{8}
\]

where \( \langle \rho \rangle_{\text{central}}(N_{\text{PV}}, \mu) \) is the average \( \langle \rho \rangle \) for \( |\eta| < \eta_{\text{plateau}} \). The \( P_{\text{jet}}^\rho(\eta, N_{\text{PV}}, \langle \mu \rangle) \) is therefore 1, by definition, for \( |\eta| < \eta_{\text{plateau}} \) and decreases for larger \( |\eta| \).

A functional form of \( P_{\text{jet}}^\rho(\eta, N_{\text{PV}}, \langle \mu \rangle) \) is used to parameterize its dependence on \( \eta \), \( N_{\text{PV}} \), and \( \langle \mu \rangle \) and is defined as

\[
P_{\text{jet}}^\rho(\eta, N_{\text{PV}}, \langle \mu \rangle) = \begin{cases} 
1 & \quad |\eta| < \eta_{\text{plateau}} \\
(1 - G_{\text{base}}(\eta_{\text{plateau}})) \cdot G_{\text{core}}(\eta - \eta_{\text{plateau}}) + G_{\text{base}}(\eta) \quad (|\eta| \geq \eta_{\text{plateau}})
\end{cases} \tag{9}
\]

where the central region \( |\eta| < \eta_{\text{plateau}} = 1.8 \) is plateaued at 1, and then a pair of Gaussian functions \( G_{\text{core}}(\eta - \eta_{\text{plateau}}) \) and \( G_{\text{base}}(\eta) \) are added for the fit in the forward regions of the calorimeter. The value of \( G_{\text{core}}(0) = 1 \) so that Eq. (9) is continuous at \( \eta = \eta_{\text{plateau}} \). Two example fits are shown in Fig. 1 for \( N_{\text{PV}} = 3 \) and 8 with \( \langle \mu \rangle = 7.5-9.5 \) interactions per bunch crossing. For both distributions the value is defined to be unity in the central region \( (|\eta| < \eta_{\text{plateau}}) \), and the sum of two Gaussian functions provides a good description of the change in the amount of in-time pileup beyond \( \eta_{\text{plateau}} \). The baseline Gaussian function \( G_{\text{base}}(\eta) \) has a larger width and is used to describe the larger amount of in-time pileup in the forward region as seen in Fig. 1. Fitting with Eq. (9) provides a parameterized function for in-time and out-of-time pileup which is valid for the whole 2012 dataset.

The soft term for the EJAF \( E_{\text{T}}^{\text{miss}} \) algorithm is calculated as

\[
E_{\text{T}}^{\text{miss,soft}}(x(y)) = -\sum_{i=0}^{N_{\text{after-jet}}} p_{x(y),i}^{\text{jet,corr}}, \tag{10}
\]

which sums the transverse momenta, labelled \( p_{x(y),i}^{\text{jet,corr}} \), of the corrected soft-term jets matched to the primary vertex. The number of these filtered jets, which are selected

\(^6\) The forward ATLAS calorimeters are less granular than those in the central region, which leads to fewer clusters being reconstructed.
The average transverse momentum density shape $P^\rho(\eta, N_{PV}, \langle \mu \rangle)$ for jets in data is compared to the model in Eq. (9) with $\langle \mu \rangle = 7.5–9.5$ and with a three reconstructed vertices and b eight reconstructed vertices. The increase of jet activity in the forward regions coming from more in-time pileup with $N_{PV} = 8$ can be seen by the flatter shape of the Gaussian fit of the forward activity $G_{base}(N_{PV}, \langle \mu \rangle)$ (blue dashed line).

Fig. 1

The mean $\alpha_{STVF}$ weight is shown versus the number of reconstructed vertices ($N_{PV}$) for 0-jet and inclusive events in $Z \to \mu\mu$ data. The inset at the bottom of the figure shows the ratio of the data to the MC predictions with only the statistical uncertainties on the data and MC simulation. The bin boundary always includes the lower edge and not the upper edge.

4.1.3 Jet $p_T$ threshold and JVF selection

The TST, STVF, and EJAF $E_T^{miss}$ algorithms complement the pileup reduction in the soft term with additional requirements on the jets entering the $E_T^{miss}$ hard term, which are also aimed at reducing pileup dependence. These $E_T^{miss}$ reconstruction algorithms apply a requirement of JVF > 0.25 to jets with $p_T < 50$ GeV and $|\eta| < 2.4$ in order to suppress those originating from pileup interactions. The maximum $|\eta|$ value is lowered to 2.4 to ensure that the core of each jet is within the tracking volume ($|\eta| < 2.5$) [4]. Charged particle...
cles from jets below the \( p_T \) threshold are considered in the soft terms for the STVF, TST, and EJAF (see Sect. 4.1.2 for details).

The same JVF requirements are not applied to the CST \( E_T^{\text{miss}} \) because its soft term includes the soft recoil from all interactions, so removing jets not associated with the hard-scatter interaction could create an imbalance. The procedure for choosing the jet \( p_T \) and JVF criteria is summarized in Sect. 7.

Throughout most of this paper the number of jets is computed without a JVF requirement so that the \( E_T^{\text{miss}} \) algorithms are compared on the same subset of events. However, the JVF \( > 0.25 \) requirement is applied in jet counting when 1-jet and \( \geq 2 \)-jet samples are studied using the TST \( E_T^{\text{miss}} \) reconstruction, which includes Figs. 8 and 22. The JVF removes pileup jets that obscure trends in samples with different jet multiplicities.

4.2 Track \( E_T^{\text{miss}} \)

Extending the philosophy of the TST definition to the full event, the \( E_T^{\text{miss}} \) is reconstructed from tracks alone, reducing the pileup contamination that afflicts the other object-based algorithms. While a purely track-based \( E_T^{\text{miss}} \), designated Track \( E_T^{\text{miss}} \), has almost no pileup dependence, it is insensitive to neutral particles, which do not form tracks in the ID. This can degrade the \( E_T^{\text{miss}} \) calibration, especially in event topologies with numerous or highly energetic jets. The \( \eta \) coverage of the Track \( E_T^{\text{miss}} \) is also limited to the ID acceptance of \( |\eta| < 2.5 \), which is substantially smaller than the calorimeter coverage, which extends to \( |\eta| = 4.9 \).

Track \( E_T^{\text{miss}} \) is calculated by taking the negative vectorial sum of \( \vec{p}_T \) of tracks satisfying the same quality criteria as the TST tracks. Similar to the TST, tracks with poor momentum resolution or without corresponding calorimeter deposits are removed. Because of Bremsstrahlung within the ID, the electron \( p_T \) is determined more precisely by the calorimeter than by the ID. Therefore, the Track \( E_T^{\text{miss}} \) algorithm uses the electron \( p_T \) measurement in the calorimeter and removes tracks overlapping its shower. Calorimeter deposits from photons are not added because they cannot be reliably associated to particular \( pp \) interactions. For muons, the ID track \( p_T \) is used and not the fits combining the ID and MS \( p_T \). For events without any reconstructed jets, the Track and TST \( E_T^{\text{miss}} \) would have similar values, but differences could still originate from muon track measurements as well as reconstructed photons or calorimeter deposits from \( \tau_{\text{had-vis}} \), which are only included in the TST.

The soft term for the Track \( E_T^{\text{miss}} \) is defined to be identical to the TST by excluding tracks associated with the high-\( p_T \) physics objects used in Eq. (2).

5 Comparison of \( E_T^{\text{miss}} \) distributions in data and MC simulation

In this section, basic \( E_T^{\text{miss}} \) distributions before and after pileup suppression in \( Z \rightarrow \ell \ell \) and \( W \rightarrow \ell \nu \) data events are compared to the distributions from the MC signal plus relevant background samples. All distributions in this section include the dominant systematic uncertainties on the high-\( p_T \) objects, the \( E_T^{\text{miss,soft}} \) (described in Sect. 8) and pileup modelling [7]. The systematics listed above are the largest systematic uncertainties in the \( E_T^{\text{miss}} \) for \( Z \) and \( W \) samples.

5.1 Modelling of \( Z \rightarrow \ell \ell \) events

The CST, EJAF, TST, STVF, and Track \( E_T^{\text{miss}} \) distributions for \( Z \rightarrow \mu \mu \) data and simulation are shown in Fig. 3. The \( Z \) boson signal region, which is defined in Sect. 3.2, has better than 99% signal purity. The MC simulation agrees with data for all \( E_T^{\text{miss}} \) reconstruction algorithms within the assigned systematic uncertainties. The mean and the standard deviation of the \( E_T^{\text{miss}} \) distribution is shown for all of the \( E_T^{\text{miss}} \) algorithms in \( Z \rightarrow \mu \mu \) inclusive simulation in Table 4. The CST \( E_T^{\text{miss}} \) has the highest mean \( E_T^{\text{miss}} \) and thus the broadest \( E_T^{\text{miss}} \) distribution. All of the \( E_T^{\text{miss}} \) algorithms with pileup suppression have narrower \( E_T^{\text{miss}} \) distributions as shown by their smaller mean \( E_T^{\text{miss}} \) values. However, those algorithms also have non-Gaussian tails in the \( E_T^{\text{miss}} \) and \( E_y^{\text{miss}} \) distributions, which contribute to the region with \( E_T^{\text{miss}} \gtrsim 50 \) GeV. The Track \( E_T^{\text{miss}} \) has the largest tail because it does not include contributions from the neutral particles, and this results in it having the largest standard deviation.

The tails of the \( E_T^{\text{miss}} \) distributions in Fig. 3 for \( Z \rightarrow \mu \mu \) data are observed to be compatible with the sum of expected signal and background contributions, namely \( t\bar{t} \) and the summed diboson (\( VV \)) processes including \( WW, WZ, \) and \( ZZ \), which all have high-\( p_T \) neutrinos in their final states. Instrumental effects can show up in the tails of the \( E_T^{\text{miss}} \), but such effects are small.

The \( E_T^{\text{miss}} \phi \) distribution is not shown in this paper but is very uniform, having less than 4 parts in a thousand difference from positive and negative \( \phi \). Thus the \( \phi \)-asymmetry is greatly reduced from that observed in Ref. [1].

The increase in systematic uncertainties in the range 50–120 GeV in Fig. 3 comes from the tail of the \( E_T^{\text{miss}} \) distribution for the simulated \( Z \rightarrow \mu \mu \) events. The increased width in the uncertainty band is asymmetric because many systematic uncertainties increase the \( E_T^{\text{miss}} \) tail in \( Z \rightarrow \mu \mu \) events by creating an imbalance in the transverse momentum. The largest of these systematic uncertainties are those associated with the jet energy resolution, the jet energy scale, and pileup. The pileup systematic uncertainties affect mostly the CST and EJAF \( E_T^{\text{miss}} \), while the jet energy scale uncertainty
Fig. 3 Distributions of the $E_T^{miss}$ with the a CST, b EJAF, c TST, d STVF, and e Track $E_T^{miss}$ are shown in data and MC simulation events satisfying the $Z \rightarrow \mu\mu$ selection. The lower panel of the figures shows the ratio of data to MC simulation, and the bands correspond to the combined systematic and MC statistical uncertainties. The far right bin includes the integral of all events with $E_T^{miss}$ above 300 GeV.
Table 4 The mean and standard deviation of the $E_{\text{miss}}$ distributions in Z → $\mu\mu$ inclusive simulation

<table>
<thead>
<tr>
<th>$E_{\text{miss}}$ alg.</th>
<th>Mean ± SD [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST $E_{\text{miss}}$</td>
<td>20.4 ± 12.5</td>
</tr>
<tr>
<td>EJAF $E_{\text{miss}}$</td>
<td>16.8 ± 11.5</td>
</tr>
<tr>
<td>TST $E_{\text{miss}}$</td>
<td>13.2 ± 10.3</td>
</tr>
<tr>
<td>STVF $E_{\text{miss}}$</td>
<td>13.8 ± 10.8</td>
</tr>
<tr>
<td>Track $E_{\text{miss}}$</td>
<td>13.9 ± 14.4</td>
</tr>
</tbody>
</table>

causes the larger systematic uncertainty for the TST and STVF $E_{\text{miss}}$. The Track $E_{\text{miss}}$ does not have the same increase in systematic uncertainties because it does not make use of reconstructed jets. Above 120 GeV, most events have a large intrinsic $E_{\text{miss}}$, and the systematic uncertainties on the $E_{\text{miss}}$, especially the soft term, are smaller.

Figure 4 shows the soft-term distributions. The pileup-suppressed $E_{\text{miss}}$ algorithms generally have a smaller mean soft term as well as a sharper peak near zero compared to the CST. Among the $E_{\text{miss}}$ algorithms, the soft term from the EJAF algorithm shows the smallest change relative to the CST. The TST has a sharp peak near zero similar to the STVF but with a longer tail, which mostly comes from individual tracks. These tracks are possibly mismeasured and further studies are planned. The simulation under-predicts the TST relative to the observed data between 60–85 GeV, and the differences exceed the assigned systematic uncertainties. This

![Figure 4](image-url)

Fig. 4 Distributions of the soft term for the a CST, b EJAF, c TST, and d STVF are shown in data and MC simulation events satisfying the Z → $\mu\mu$ selection. The lower panel of the figures show the ratio of data to MC simulation, and the bands correspond to the combined systematic and MC statistical uncertainties. The far right bin includes the integral of all events with $E_{\text{miss,soft}}$ above 160 GeV.
region corresponds to the transition from the narrow core to the tail coming from high-\(p_T\) tracks. The differences between data and simulation could be due to mismodelling of the rate of mismeasured tracks, for which no systematic uncertainty is applied. The mismeasured-track cleaning, as discussed in Sect. 4.1.2, reduces the TST tail starting at 120 GeV, and this region is modelled within the assigned uncertainties. The mismeasured-track cleaning for tracks below 120 GeV and the core of the distribution were used instead. The resolution. This includes important information about the tails, which would be lost if the result of a Gaussian fit over only the central core of the distribution were used instead. The resolution of the \(E_T^{\text{miss}}\) distribution is extracted using the RMS over the central core of the distribution.

5.2 Modelling of \(W \rightarrow \ell \nu\) events

In this section, the selection requirements for the \(m_T\) and \(E_T^{\text{miss}}\) distributions are defined using the same \(E_T^{\text{miss}}\) algorithm as that labelling the distribution (e.g. selection criteria are applied to the CST \(E_T^{\text{miss}}\) for distributions showing the CST \(E_T^{\text{miss}}\)). The intrinsic \(E_T^{\text{miss}}\) in \(W \rightarrow \ell \nu\) events allows a comparison of the \(E_T^{\text{miss}}\) scale between data and simulation. The level of agreement between data and MC simulation for the \(E_T^{\text{miss}}\) reconstruction algorithms is studied using \(W \rightarrow \nu\) events with the selection defined in Sect. 3.3. The CST and TST \(E_T^{\text{miss}}\) distributions in \(W \rightarrow \nu\) events are shown in Fig. 6. The \(W \rightarrow \tau \nu\) contributions are combined with \(W \rightarrow \nu\) events in the figure. The data and MC simulation agree within the assigned systematic uncertainties for both the CST and TST \(E_T^{\text{miss}}\) algorithms. The other \(E_T^{\text{miss}}\) algorithms show similar levels of agreement between data and MC simulation.

6 Performance of the \(E_T^{\text{miss}}\) in data and MC simulation

6.1 Resolution of \(E_T^{\text{miss}}\)

The \(E_x^{\text{miss}}\) and \(E_y^{\text{miss}}\) are expected to be approximately Gaussian distributed for \(Z \rightarrow \ell \ell\) events as discussed in Ref. [1]. However, because of the non-Gaussian tails in these distributions, especially for the pileup-suppressing \(E_T^{\text{miss}}\) algorithms, the root-mean-square (RMS) is used to estimate the resolution. This includes important information about the tails, which would be lost if the result of a Gaussian fit over only the central core of the distribution were used instead. The resolution of the \(E_T^{\text{miss}}\) distribution is extracted using the RMS from the combined distribution of \(E_x^{\text{miss}}\) and \(E_y^{\text{miss}}\), which are determined to be independent from correlation studies.
The previous ATLAS $E_T^{\text{miss}}$ performance paper [1] studied the resolution defined by the width of Gaussian fits in a narrow range of $\pm 2 \text{RMS}$ around the mean and used a separate study to investigate the tails. Therefore, the results of this paper are not directly comparable to those of the previous study. The resolutions presented in this paper are expected to be larger than the width of the Gaussian fitted in this manner because the RMS takes into account the tails.

In this section, the resolution for the $E_T^{\text{miss}}$ is presented for $Z \rightarrow \mu\mu$ events using both data and MC simulation. Unless it is a simulation-only figure (labelled with “Simulation” under the ATLAS label), the MC distribution includes the signal sample (e.g. $Z \rightarrow \mu\mu$) as well as diboson, $t\bar{t}$, and $tW$ samples.

6.1.1 Resolution of the $E_T^{\text{miss}}$ as a function of the number of reconstructed vertices

The stability of the $E_T^{\text{miss}}$ performance as a function of the amount of pileup is estimated by studying the $E_T^{\text{miss}}$ resolution as a function of the number of reconstructed vertices ($N_{\text{PV}}$) for $Z \rightarrow \mu\mu$ events as shown in Fig. 7. The bin edge is always including the lower edge and not the upper. For example, the events with $N_{\text{PV}}$ in the inclusive range $30–39$ are combined because of small sample size. In addition, very few events were collected below $N_{\text{PV}}$ of 2 during 2012 data taking. Events in which there are no reconstructed jets with $p_T > 20$ GeV are referred to collectively as the 0-jet sample. Distributions are shown here for both the 0-jet and inclusive samples. For both samples, the data and MC simulation agree within 2% up to around $N_{\text{PV}} = 15$ but the deviation grows to around 5–10% for $N_{\text{PV}} > 25$, which might be attributed to the decreasing sample size. All of the $E_T^{\text{miss}}$ distributions show a similar level of agreement between data and simulation across the full range of $N_{\text{PV}}$.

For the 0-jet sample in Fig. 7a, the STVF, TST, and Track $E_T^{\text{miss}}$ resolutions all have a small slope with respect to $N_{\text{PV}}$, which implies stability of the resolution against pileup. In addition, their resolutions agree within 1 GeV throughout the $N_{\text{PV}}$ range. In the 0-jet sample, the TST and Track $E_T^{\text{miss}}$ are both primarily reconstructed from tracks; however, small differences arise mostly from accounting for photons in the TST $E_T^{\text{miss}}$ reconstruction algorithm. The CST $E_T^{\text{miss}}$ is directly affected by the pileup as its reconstruction does not apply any pileup suppression techniques. Therefore, the CST $E_T^{\text{miss}}$ has the largest dependence on $N_{\text{PV}}$, with a resolution ranging from 7 GeV at $N_{\text{PV}} = 2$ to around 23 GeV at $N_{\text{PV}} = 25$. The $E_T^{\text{miss}}$ resolution of the EJAF distribution, while better than that of the CST $E_T^{\text{miss}}$, is not as good as that of the other pileup-suppressing algorithms.

For the inclusive sample in Fig. 7b, the Track $E_T^{\text{miss}}$ is the most stable with respect to pileup with almost no dependence on $N_{\text{PV}}$. For $N_{\text{PV}} > 20$, the Track $E_T^{\text{miss}}$ has the best resolution showing that pileup creates a larger degradation in the resolution of the other $E_T^{\text{miss}}$ distributions than excluding neutral particles, as the Track $E_T^{\text{miss}}$ algorithm does. The EJAF $E_T^{\text{miss}}$ algorithm does not reduce the pileup dependence as much as the TST and STVF $E_T^{\text{miss}}$ algorithms, and the CST $E_T^{\text{miss}}$ again has the largest dependence on $N_{\text{PV}}$.

Figure 7 also shows that the pileup dependence of the TST, CST, EJAF and STVF $E_T^{\text{miss}}$ is smaller in the 0-jet sample than in the inclusive sample. Hence, the evolution
The resolution obtained from the combined distribution of $E_T^{\text{miss}}$ and $E_T^{\text{miss}}$ for the CST, STVF, EJAF, TST, and Track $E_T^{\text{miss}}$ algorithms as a function of $N_{PV}$ in a 0-jet and b inclusive $Z \rightarrow \mu\mu$ events.

Fig. 7

The resolution predicted by the simulation is modelled by the simulation, which leads to larger $E_T^{\text{miss}}$ resolution at high $N_{PV}$ than the individual jet categories, which indicates that the behaviour seen in the inclusive sample is driven by an increased number of pileup jets included in the $E_T^{\text{miss}}$ calculation at larger $N_{PV}$.

6.1.2 Resolution of the $E_T^{\text{miss}}$ as a function of $\Sigma E_T$

The resolutions of $E_T^{\text{miss}}$, resulting from the different reconstruction algorithms, are compared as a function of the scalar sum of transverse momentum in the event, as calculated using Eq. (4). The CST $E_T^{\text{miss}}$ resolution is observed to depend linearly on the square root of the $\Sigma E_T$ computed with the CST $E_T^{\text{miss}}$ components in Ref. [1]. However, the $\Sigma E_T$ used in this subsection is calculated with the TST $E_T^{\text{miss}}$ algorithm. This allows studies of the resolution as a function of the momenta of particles from the selected PV without including the amount of pileup activity in the event. Figure 9 shows the resolution as a function of $\Sigma E_T$ (TST) for $Z \rightarrow \mu\mu$ data and MC simulation in the 0-jet and inclusive samples.

In the 0-jet sample shown in Fig. 9a, the use of tracking information in the soft term, especially for the STVF, TST, and Track $E_T^{\text{miss}}$, greatly improves the resolution relative to the CST $E_T^{\text{miss}}$. The EJAF $E_T^{\text{miss}}$ has a better resolution than that of the CST $E_T^{\text{miss}}$ but does not perform as well as the other reconstruction algorithms. All of the resolution curves have an approximately linear increase with $\Sigma E_T$ (TST); however, the Track $E_T^{\text{miss}}$ resolution increases sharply starting at $\Sigma E_T$ (TST) $= 200$ GeV due to missed neutral contributions like photons. The resolution predicted by the simulation is about 5% larger than in data for all $E_T^{\text{miss}}$ algorithms at $\Sigma E_T$ (TST) $= 50$ GeV, but agreement improves as $\Sigma E_T$ (TST) increases until around $\Sigma E_T$ (TST) $= 200$ GeV. Events with jets can end up in the 0-jet event selection, for example, if a jet is misidentified as a hadronically decaying $\tau$-lepton. The $\Sigma p_T$ increases with $\Sigma E_T$ (TST), and the rate of jets misreconstructed as hadronically decaying $\tau$-leptons is not well modelled by the simulation, which leads to larger $E_T^{\text{miss}}$ resolution at high $\Sigma E_T$ (TST) than that observed in the data. The Track $E_T^{\text{miss}}$ can be more strongly affected by misidentified jets because neutral particles from the high-$p_T$ jets are not included.

For the inclusive sample in Fig. 9b, the pileup-suppressed $E_T^{\text{miss}}$ distributions have better resolution than the CST $E_T^{\text{miss}}$ for $\Sigma E_T$ (TST) $< 200$ GeV, but these events are mostly those with no associated jets. For higher $\Sigma E_T$ (TST), the
impact from the $\Sigma E_T^{\text{jets}}$ term starts to dominate the resolution as well as the $\Sigma E_T$ (TST). Since the vector sum of jet momenta is mostly common\(^7\) to all $E_T^{\text{miss}}$ algorithms except for the Track $E_T^{\text{miss}}$, those algorithms show similar performance in terms of the resolution. At larger $\Sigma E_T$ (TST), the Track $E_T^{\text{miss}}$ resolution begins to degrade relative to the other algorithms because it does not include the high-$p_T$ neutral particles coming from jets. The ratio of data to MC simulation for the Track $E_T^{\text{miss}}$ distribution is close to one, while for other algorithms the MC simulation is below the data by about 5% at large $\Sigma E_T$ (TST). While the Track $E_T^{\text{miss}}$ appears well modelled for the Alpgen+Pythia simulation used in this figure, the modelling depends strongly on the parton shower model.

6.2 The $E_T^{\text{miss}}$ response

The balance of $E_T^{\text{miss}}$ against the vector boson $\vec{p}_T$ in $W/Z+\text{jets}$ events is used to evaluate the $E_T^{\text{miss}}$ response. A lack of balance is a global indicator of biases in $E_T^{\text{miss}}$ reconstruction and implies a systematic misestimation of at least one of the $E_T^{\text{miss}}$ terms, possibly coming from an imperfect selection or calibration of the reconstructed physics objects. The procedure to evaluate the response differs between $Z+\text{jets}$ events (Sect. 6.2.1) and $W+\text{jets}$ events (Sect. 6.2.2) because of the high-$p_T$ neutrino in the transverse decay of the $W$ boson.

6.2.1 Measuring $E_T^{\text{miss}}$ recoil versus $p_T^Z$

In events with $Z\rightarrow\mu\mu$ decays, the $\vec{p}_T$ of the $Z$ boson defines an axis in the transverse plane of the ATLAS detector, and

\footnote{As defined in Sect. 4.1.3, the CST $E_T^{\text{miss}}$ does not apply a JVF requirement on the jets like the TST, EJAF, and STVF $E_T^{\text{miss}}$. However, large $\Sigma E_T^{\text{jets}}$ tends to come from hard-scatter jets and not from pileup.}

for events with 0-jets, the $E_T^{\text{miss}}$ should balance the $\vec{p}_T$ of the $Z$ boson ($\vec{p}_T^Z$) along this axis. Comparing the response in events with and without jets allows distinction between the jet and soft-term responses. The component of the $E_T^{\text{miss}}$ along the $\vec{p}_T^Z$ axis is sensitive to biases in detector responses [50]. The unit vector of $\vec{p}_T^Z$ is labelled as $\hat{A}_Z$ and is defined as:

$$\hat{A}_Z = \frac{\vec{p}_T^{\ell^+} + \vec{p}_T^{\ell^-}}{|\vec{p}_T^{\ell^+} + \vec{p}_T^{\ell^-}|},$$

(12)

where $\vec{p}_T^{\ell^+}$ and $\vec{p}_T^{\ell^-}$ are the transverse momentum vectors of the leptons from the $Z$ boson decay.

The recoil of the $Z$ boson is measured by removing the $Z$ boson decay products from the $E_T^{\text{miss}}$ and is computed as

$$\vec{R} = \vec{E}_T^{\text{miss}} + \vec{p}_T^Z.$$  

(13)

Since the $E_T^{\text{miss}}$ includes a negative vector sum over the lepton momenta, the addition of $\vec{p}_T^Z$ removes its contribution. With an ideal detector and $E_T^{\text{miss}}$ reconstruction algorithm, $Z\rightarrow\ell\ell$ events have no $E_T^{\text{miss}}$, and the $\vec{R}$ balances with $\vec{p}_T^Z$ exactly. For the real detector and $E_T^{\text{miss}}$ reconstruction algorithm, the degree of balance is measured by projecting the recoil onto $\hat{A}_Z$, and the relative recoil is defined as the projection $\vec{R} \cdot \hat{A}_Z$ divided by $p_T^Z$, which gives a dimensionless estimate that is unity if the $E_T^{\text{miss}}$ is ideally reconstructed and calibrated. Figure 10 shows the mean relative recoil versus $p_T^Z$ for $Z\rightarrow\mu\mu$ events where the average value is indicated by angle brackets. The data and MC simulation agree within around 10% for all $E_T^{\text{miss}}$ algorithms for all $p_T^Z$; however, the agreement is a few percent worse for $p_T^Z > 50$ GeV in the 0-jet sample.

The $Z\rightarrow\mu\mu$ events in the 0-jet sample in Fig. 10a have a relative recoil significantly lower than unity ($\langle \vec{R} \cdot \hat{A}_Z / p_T^Z \rangle < 1$) throughout the $p_T^Z$ range. In the 0-jet sample,
the relative recoil estimates how well the soft term balances the \( \vec{p}_T \) of muons from the \( Z \) decay, which are better measured than the soft term. The relative recoil below one indicates that the soft term is underestimated. The CST \( E_T^{\text{miss}} \) has a relative recoil measurement of \( \langle \vec{R} \cdot \hat{A}_Z / p_T^Z \rangle \sim 0.5 \) throughout the \( p_T^Z \) range, giving it the best recoil performance among the \( E_T^{\text{miss}} \) algorithms. The TST and Track \( E_T^{\text{miss}} \) have slightly larger biases than the CST \( E_T^{\text{miss}} \) because neutral particles are not considered in the soft term. The TST \( E_T^{\text{miss}} \) recoil improves relative to that of the Track \( E_T^{\text{miss}} \) for \( p_T^Z > 40 \) GeV because of the inclusion of photons in its reconstruction.

The relative recoil distribution for the STVF \( E_T^{\text{miss}} \) shows the largest bias for \( p_T^Z < 60 \) GeV. The STVF algorithm scales the recoil down globally by the factor \( \alpha_{\text{STVF}} \) as defined in Eq. (11), and this correction decreases the already underestimated soft term. The \( \alpha_{\text{STVF}} \) does increase with \( p_T^Z \) going from 0.06 at \( p_T^Z = 0 \) GeV to around 0.15 at \( p_T^Z = 50 \) GeV, and this results in a rise in the recoil, which approaches the TST \( E_T^{\text{miss}} \) near \( p_T^Z \sim 70 \) GeV.

In Fig. 10b, the inclusive \( Z \rightarrow \mu \mu \) events have a significantly underestimated relative recoil for \( p_T^Z < 40 \) GeV. The balance between the \( \vec{R} \) and \( \vec{p}_T^Z \) improves with \( p_T^Z \) because of an increase in events having high-\( p_T \) calibrated jets recoiling against the \( Z \) boson. The presence of jets included in the hard term also reduces the sensitivity to the soft term, which is difficult to measure accurately. The difficulty in isolating effects from soft-term contributions from high-\( p_T \) physics objects is one reason why the soft term is not corrected. As with the 0-jet sample, the CST \( E_T^{\text{miss}} \) has a significantly under-calibrated relative recoil in the low-\( p_T^Z \) region, and all of the other \( E_T^{\text{miss}} \) algorithms have a lower relative recoil than the CST \( E_T^{\text{miss}} \). Of the pileup-suppressing \( E_T^{\text{miss}} \) algorithms, the TST \( E_T^{\text{miss}} \) is closest to the relative recoil of the CST \( E_T^{\text{miss}} \). The relative recoil of the Track \( E_T^{\text{miss}} \) is significantly lower than unity because the neutral particles recoiling from the \( Z \) boson are not included in its reconstruction. Finally, the STVF \( E_T^{\text{miss}} \) shows the lowest relative recoil among the object-based \( E_T^{\text{miss}} \) algorithms as discussed above for Fig. 10a, even lower than the Track \( E_T^{\text{miss}} \) for \( p_T^Z < 16 \) GeV.

### 6.2.2 Measuring \( E_T^{\text{miss}} \) response in simulated \( W \rightarrow \ell v \) events

For simulated events with intrinsic \( E_T^{\text{miss}} \), the response is studied by looking at the relative mismeasurement of the reconstructed \( E_T^{\text{miss}} \). This is referred to here as the “linearity”, and is a measure of how consistent the reconstructed \( E_T^{\text{miss}} \) is with the \( E_T^{\text{miss, True}} \). The linearity is defined as the mean value of the ratio, \( (E_T^{\text{miss}} - E_T^{\text{miss, True}})/E_T^{\text{miss, True}} \) and is expected to be zero if the \( E_T^{\text{miss}} \) is reconstructed at the correct scale.

For the linearity studies, no selection on the \( E_T^{\text{miss}} \) or \( m_T \) is applied, in order to avoid biases as these are purely simulation-based studies. In Fig. 11, the linearity for \( W \rightarrow \mu \nu \) simulated events is presented as a function of the \( E_T^{\text{miss, True}} \). Despite the relaxed selection, a positive linearity is evident for \( E_T^{\text{miss, True}} < 40 \) GeV, due to the finite resolution of the \( E_T^{\text{miss}} \) reconstruction and the fact that the reconstructed \( E_T^{\text{miss}} \) is positive by definition. The CST \( E_T^{\text{miss}} \) has the largest deviation from zero at low \( E_T^{\text{miss, True}} \) because it has the largest \( E_T^{\text{miss}} \) resolution.

For the events in the 0-jet sample in Fig. 11a, all \( E_T^{\text{miss}} \) algorithms have a negative linearity for \( E_T^{\text{miss, True}} > 40 \) GeV, which diminishes for \( E_T^{\text{miss, True}} \lesssim 60 \) GeV. The region of \( E_T^{\text{miss, True}} \) between 40 and 60 GeV mostly includes events lying in the Jacobian peak of the \( W \) transverse mass, and these events include mostly on-shell \( W \) bosons. For \( E_T^{\text{miss}} \gtrsim 40 \) GeV, the on-shell \( W \) boson must have non-zero \( p_T \), which typically comes from its recoil against jets. However, no reconstructed or generator-level jets are found in this 0-jet sample. Therefore, most of the events with \( 40 < E_T^{\text{miss, True}} < 60 \) GeV have jets below the 20 GeV threshold contributing to the soft term, and the soft term is not cal-
imbibed. The under-estimation of the soft term, described in Sect. 6.2.1, causes the linearity to deviate further from zero in this region. Events with $E_{\text{T}}^{\text{miss, True}} > 60$ GeV are mostly off-shell $W$ bosons that are produced with very low $p_T$. For these events, the $\vec{p}_T$ contributions to the $E_{\text{T}}^{\text{miss}}$ reconstruction come mostly from the well-measured muon $\vec{p}_T^\mu$, and the soft term plays a much smaller role. Hence, the linearity improves as the impact of the soft term decreases with larger $E_{\text{T}}^{\text{miss, True}}$.

For inclusive events in Fig. 11b with $E_{\text{T}}^{\text{miss, True}} > 40$ GeV, the deviation of the linearity from zero is smaller than 5% for the CST $E_{\text{T}}^{\text{miss}}$. The linearity of the TST $E_{\text{T}}^{\text{miss}}$ is within 10% of unity in the range of 40–60 GeV and improves for higher $E_{\text{T}}^{\text{miss, True}}$ values. The STVF $E_{\text{T}}^{\text{miss}}$ has the most negative bias in the linearity among the object-based $E_{\text{T}}^{\text{miss}}$ algorithms for $E_{\text{T}}^{\text{miss, True}} > 40$ GeV. The TST, CST, STVF, and EJAF $E_{\text{T}}^{\text{miss}}$ algorithms perform similarly for all $E_{\text{T}}^{\text{miss, True}}$ values. As expected, the linearity of the Track $E_{\text{T}}^{\text{miss}}$ settles below zero due to not accounting for neutral particles in jets.

6.3 The $E_{\text{T}}^{\text{miss}}$ angular resolution

The angular resolution is important for the reconstruction of kinematic observables such as the transverse mass of the $W$ boson and the invariant mass in $H \rightarrow \tau\tau$ events [51]. For simulated $W \rightarrow \ell\nu$ events, the direction of the reconstructed $E_{\text{T}}^{\text{miss}}$ is compared to the $E_{\text{T}}^{\text{miss, True}}$ for each $E_{\text{T}}^{\text{miss}}$ reconstruction algorithm using the difference in the azimuthal angles, $\Delta \phi(E_{\text{T}}^{\text{miss}}, E_{\text{T}}^{\text{miss, True}})$, which has a mean value of zero. The RMS of the distribution is taken as the resolution, which is labelled RMS ($\Delta \phi$).

No selection on the $E_{\text{T}}^{\text{miss}}$ or $m_T$ is applied in order to avoid biases. The RMS ($\Delta \phi$) is shown as a function of $E_{\text{T}}^{\text{miss, True}}$ in Fig. 12a for the 0-jet sample in $W \rightarrow \ell\nu$ simulation; the angular resolution generally improves as the $E_{\text{T}}^{\text{miss, True}}$ increases, for all algorithms. For $E_{\text{T}}^{\text{miss, True}} \lesssim 120$ GeV, the pileup-suppressing algorithms improve the resolution over the CST $E_{\text{T}}^{\text{miss}}$ algorithm, but all of the algorithms produce distributions with similar resolutions in the higher $E_{\text{T}}^{\text{miss, True}}$ region. The increase in RMS ($\Delta \phi$) at around 40–60 GeV in the 0-jet sample is due to the larger contribution of jets below 20 GeV entering the soft term as mentioned in Sect. 6.2.2. The distribution from the inclusive sample shown in Fig. 12b has the same pattern as the one from the 0-jet sample, except that the performance of the Track $E_{\text{T}}^{\text{miss}}$ algorithm is again significantly worse. In addition, the transition region near $40 < E_{\text{T}}^{\text{miss, True}} < 60$ GeV is smoother as the underestimate of the soft term becomes less significant due to the presence of events with high-$p_T$ calibrated jets. The TST $E_{\text{T}}^{\text{miss}}$ algorithm has the best angular resolution for both the 0-jet and inclusive topologies throughout the entire range of $E_{\text{T}}^{\text{miss, True}}$.

6.4 Transverse mass in $W \rightarrow \ell\nu$ events

The $W$ boson events are selected using kinematic observables that are computed from the $E_{\text{T}}^{\text{miss}}$ and lepton transverse momentum. This section evaluates the scale of the $m_T$, as defined in Eq. (1), reconstructed with each $E_{\text{T}}^{\text{miss}}$ definition. The $m_T$ computed using the reconstructed $E_{\text{T}}^{\text{miss}}$ is compared to the $m_T^{\text{True}}$, which is calculated using the $E_{\text{T}}^{\text{miss, True}}$ in $W \rightarrow \ell\nu$ MC simulation. The mean of the difference between the reconstructed and generator-level $m_T$, $(m_T - m_T^{\text{True}})$, is shown as a function of $m_T^{\text{True}}$ in Fig. 13 for the 0-jet and inclusive samples. No $E_{\text{T}}^{\text{miss}}$ or $m_T$ selection is made in these figures, to avoid biases. All distributions for the $E_{\text{T}}^{\text{miss}}$ algorithms have a positive bias at low values of $m_T^{\text{True}}$ coming from the positive-definite nature of the $m_T$ and the finite $E_{\text{T}}^{\text{miss}}$ resolution. For the 0-jet sample, the CST algorithm has the smallest bias for $m_T^{\text{True}} \lesssim 60$ GeV because it includes the neutral particles with no corrections for pileup. However, for the inclusive sample the TST $E_{\text{T}}^{\text{miss}}$ has the smallest bias as the $E_{\text{T}}^{\text{miss}}$ resolution plays a larger role. The STVF and Track $E_{\text{T}}^{\text{miss}}$ have the largest bias for $m_T^{\text{True}} < 50$ GeV in the 0-jet and inclusive samples, respectively. This is due to

![Fig. 11](image-url) $E_{\text{T}}^{\text{miss}}$ linearity in $W \rightarrow \mu\nu$ MC simulation is shown versus $E_{\text{T}}^{\text{miss, True}}$ in the a 0-jet and b inclusive events.
the over-correction in the soft term by $\alpha_{\text{STVF}}$ for the former
and from the missing neutral particles in the latter case. For
events with $m_{T} \gtrsim 60$ GeV, all of the $E^{\text{miss}}$ algorithms have
$m_{T} - m_{T}^{\text{True}}$ close to zero, with a spread of less than 3 GeV.

6.5 Proxy for $E^{\text{miss}}$ significance

The $E^{\text{miss}}$ significance is a metric defined to quantify how
likely it is that a given event contains intrinsic $E^{\text{miss}}$ and is
computed by dividing the measured $E^{\text{miss}}$ by an estimate of
its uncertainty. Using 7 TeV data, it was shown that the
CST $E^{\text{miss}}_{T}$ resolution follows an approximately stochastic
behaviour as a function of $\Sigma E_{T}$, computed with the CST
components, and is described by

$$\sigma(E^{\text{miss}}_{T}) = a \cdot \sqrt{\Sigma E_{T}},$$

(14)

where $\sigma(E^{\text{miss}}_{T})$ is the CST $E^{\text{miss}}_{T}$ resolution [1]. The typical
value of $a$ in the 8 TeV dataset is around 0.97 GeV$^{1/2}$
for the CST $E^{\text{miss}}_{T}$. The proxy of the $E^{\text{miss}}$ significance
presented in this section is defined as the $\frac{1}{a}E^{\text{miss}}_{T}/\sqrt{\Sigma E_{T}}$. This
choice is motivated by the linear relationship for the CST
$E^{\text{miss}}_{T}$ between its $\sqrt{\Sigma E_{T}}$ and its $E^{\text{miss}}_{T}$ resolution. The same
procedure does not work for the TST $E^{\text{miss}}_{T}$ resolution, so a
value of 2.27 GeV$^{1/2}$ is used to tune the $x$-axis so that integral
of $Z \rightarrow \mu\mu$ simulation fits the multiples of the standard
deviation of a normal distribution at the value of 2. Ideally,
only events with large intrinsic $E^{\text{miss}}_{T}$ have large values of
$\frac{1}{a}E^{\text{miss}}_{T}/\sqrt{\Sigma E_{T}}$, while events with no intrinsic $E^{\text{miss}}_{T}$ such as
$Z \rightarrow \mu\mu$ have low values. It is important to point out that in
general $Z \rightarrow \mu\mu$ is not a process with large $E^{\text{miss}}$ uncertain-
ties or large $\sqrt{\Sigma E_{T}}$. However, when there are many additional
jets (large $\Sigma E_{T}$), there is a significant probability that one of them is mismeasured, which generates fake $E^{\text{miss}}_{T}$.

The distribution of $\frac{1}{a}E^{\text{miss}}_{T}/\sqrt{\Sigma E_{T}}$ is shown for the CST
and TST $E^{\text{miss}}_{T}$ algorithms in Fig. 14 in $Z \rightarrow \mu\mu$ data
and MC simulation. The data and MC simulation agree
within the assigned uncertainties for both algorithms. The
CST $E^{\text{miss}}_{T}$ distribution in Fig. 14a has a very narrow core
for the $Z \rightarrow \mu\mu$ process, having 97% of data events with
$1.03\cdot E^{\text{miss}}_{T}/\sqrt{\Sigma E_{T}} < 2$. The proxy of the $E^{\text{miss}}$ significance,
therefore, provides discrimination power between events

Fig. 12 The resolution of $\Delta \phi(E^{\text{miss}}_{T}, E^{\text{miss}, \text{True}}_{T})$, labelled as $\text{RMS}(\Delta \phi)$, is shown for $W \rightarrow \mu\nu$ MC simulation for the a 0-jet and b inclusive samples

Fig. 13 The $\langle m_{T} - m_{T}^{\text{True}} \rangle$ is shown versus $m_{T}^{\text{True}}$ for $W \rightarrow \mu\nu$ MC simulation in the a 0-jet and b inclusive samples

$\sigma(E^{\text{miss}}_{T}) = a \cdot \sqrt{\Sigma E_{T}}$,
with intrinsic $E_T^{\text{miss}}$ (e.g. $t\bar{t}$ and dibosons) and those with fake $E_T^{\text{miss}}$ (e.g. poorly measured $Z \rightarrow \mu\mu$ events with a large number of jets).

The TST $E_T^{\text{miss}}$ is shown as an example of a pileup-suppressing algorithm. The $\Sigma E_T$ is not always an accurate reflection of the resolution when there are significant contributions from tracking resolution, as discussed in Sect. 5.1. In particular, the performance of the TST reconstruction algorithm is determined by the tracking resolution, which is generally more precise than the calorimeter energy measurements because of the reduced pileup dependence, especially for charged particles with lower $p_T$. Neutral particles are not included in the $\Sigma E_T$ for the Track $E_T^{\text{miss}}$ and TST algorithms, but they affect the resolution. In addition, a very small number of tracks do have very large over-estimated momentum measurements due to multiple scattering or other effects in the detector, and the momentum uncertainties of these tracks are not appropriately accounted for in the $\Sigma E_T$ methodology.

6.6 Tails of $E_T^{\text{miss}}$ distributions

Many analyses require large $E_T^{\text{miss}}$ to select events with high-$p_T$ weakly interacting particles. The selection efficiency, defined as the number of events with $E_T^{\text{miss}}$ above a given threshold divided by the total number of events, is used to compare the performance of various $E_T^{\text{miss}}$ reconstruction algorithms. As $Z \rightarrow \ell\ell$ events very rarely include high-$p_T$ neutrinos, they can be rejected by requiring substantial $E_T^{\text{miss}}$. For events with intrinsic $E_T^{\text{miss}}$ such as $W \rightarrow \ell\nu$, higher selection efficiencies than the $Z \rightarrow \ell\ell$ events are expected when requiring reconstructed $E_T^{\text{miss}}$. For both cases, it is important to evaluate the performance of the reconstructed $E_T^{\text{miss}}$.

The selection efficiencies with various $E_T^{\text{miss}}$ algorithms are compared for simulated $Z \rightarrow \mu\mu$ and $W \rightarrow \mu\nu$ processes as shown in Fig. 15 using the MC simulation. The event selections discussed in Sects. 3.2 and 3.3 are applied except the requirements on $E_T^{\text{miss}}$ and $m_T$ for the $W \rightarrow \mu\nu$ selection.

As shown in Fig. 15a, the selection efficiency for $Z \rightarrow \mu\mu$ events is around 1% for $E_T^{\text{miss}} > 50$ GeV, for all $E_T^{\text{miss}}$ algorithms. Thus a $E_T^{\text{miss}}$ threshold requirement can be used to reject a large number of events without intrinsic $E_T^{\text{miss}}$. However, the $E_T^{\text{miss, True}}$, which does not include detector resolution effects, shows the selection efficiency under ideal conditions, indicating there may be additional potential for improvement of the reconstructed $E_T^{\text{miss}}$. Namely, the selection efficiency with $E_T^{\text{miss, True}}$ provides a benchmark against which to evaluate the performance of different $E_T^{\text{miss}}$ algorithms. The STVF, TST, and Track $E_T^{\text{miss}}$ distributions have narrow cores, so for $E_T^{\text{miss}}$ threshold $\leq 50$ GeV these three $E_T^{\text{miss}}$ definitions have the lowest selection efficiencies for $Z \rightarrow \mu\mu$ events. Above 50 GeV, the Track $E_T^{\text{miss}}$ performance is degraded as a result of missing neutral particles, which gives it a very high selection efficiency. The TST and STVF $E_T^{\text{miss}}$ algorithms continue to have the lowest selection efficiency up to $E_T^{\text{miss}}$ threshold $\approx 110$ GeV. For 110–160 GeV, the TST $E_T^{\text{miss}}$ has a longer tail than the CST $E_T^{\text{miss}}$, which is a result of mismeasured low-$p_T$ particles that scatter and are reconstructed as high-$p_T$
tracks. Such mismeasurements are rare but significant in the $E_T^{\text{miss}}$ tail. The TST, STVF, CST, and EJAF $E_T^{\text{miss}}$ algorithms provide similar selection efficiencies for $E_T^{\text{miss}} \gtrsim 160$ GeV. Above this threshold, the $E_T^{\text{miss}}$ is dominated by mismeasured high-$p_T$ physics objects which are identical in all object-based $E_T^{\text{miss}}$ definitions. Hence, the events with $E_T^{\text{miss}} \gtrsim 160$ GeV are correlated among the TST, STVF, CST, and EJAF $E_T^{\text{miss}}$ distributions.

Figure 15b shows the selection efficiency for the $W \rightarrow \mu \nu$ simulated events passing a $E_T^{\text{miss}}$ threshold for all $E_T^{\text{miss}}$ algorithms. Requiring the $W \rightarrow \mu \nu$ events to pass the $E_T^{\text{miss}}$ threshold should ideally have a high selection efficiency similar to that of the $E_T^{\text{miss}, \text{True}}$. The CST $E_T^{\text{miss}}$ algorithm gives the highest selection efficiency between 30–120 GeV but does not agree as well as that of the other $E_T^{\text{miss}}$ algorithms with the $E_T^{\text{miss}, \text{True}}$ selection efficiency for $E_T^{\text{miss}}$ threshold $\lesssim 110$ GeV. This comes from the positive-definite nature of the $E_T^{\text{miss}}$ and the worse resolution of the CST $E_T^{\text{miss}}$ relative to the other $E_T^{\text{miss}}$ definitions. The Track $E_T^{\text{miss}}$ has the efficiency closest to that of the $E_T^{\text{miss}, \text{True}}$, but for Track $E_T^{\text{miss}} \lesssim 60$ GeV, the amount of jet activity increases, which results in a lower selection efficiency because of missing neutral particles. The EJAF, STVF, and TST $E_T^{\text{miss}}$ distributions are closer than the CST to the $E_T^{\text{miss}, \text{True}}$ selection efficiency for $E_T^{\text{miss}}$ threshold $\lesssim 100$ GeV, but the efficiencies for all the object-based algorithms and $E_T^{\text{miss}, \text{True}}$ converge for $E_T^{\text{miss}}$ threshold $\gtrsim 110$ GeV. Hence, for large $E_T^{\text{miss}}$ all object-based algorithms perform similarly.

In Fig. 16, selection efficiencies are shown as a function of the $E_T^{\text{miss}}$ threshold requirement for various simulated physics processes defined in Sect. 3.4 with no lepton, jet, or $m_T$ threshold requirements. The physics object and event selection criteria are not applied in order to show the selection efficiency resulting from the $E_T^{\text{miss}}$ threshold requirement without biases in the event topology from the ATLAS detector acceptance for leptons or jets. Only the efficiencies for the CST and TST $E_T^{\text{miss}}$ distributions are compared for brevity.

---

8 For the TST and Track $E_T^{\text{miss}}$, mismeasured high-$p_T$ tracks with $p_T > 120$ (200) GeV are removed using the track quality requirements in high (low) $|\eta|$ as defined in Sect. 4.1.2.
In Fig. 16a, the efficiencies with the TST $E_T^{\text{miss}}$ selection are shown. Comparing the physics processes while imposing a moderate $E_T^{\text{miss}}$ threshold requirement of $\sim 100$ GeV results in a selection efficiency of 60% for an ATLAS search for gluino-pair production [52], which is labelled as “SUSY”. The VBF $H \rightarrow \tau \tau$ and $t\bar{t}$ events are also selected with high efficiencies of 14 and 20%, respectively. With the 100 GeV $E_T^{\text{miss}}$ threshold the selection efficiencies for these processes are more than an order of magnitude higher than those for leptonically decaying $W$ bosons and more than two orders of magnitude higher than for $Z$ boson events.

The $Z \rightarrow ee$ events have a lower selection efficiency (around 20 times lower at $E_T^{\text{miss}} = 100$ GeV) than the $Z \rightarrow \mu\mu$ events. This is due to the muon tracking coverage, which is limited to $|\eta| < 2.7$, whereas the calorimeter covers $|\eta| < 4.9$. Muons behave as minimum-ionizing particles in the ATLAS calorimeters, so they are not included in the $E_T^{\text{miss}}$ outside the muon spectrometer acceptance.

The electrons on the other hand are measured by the forward calorimeters. The electron and muon decay modes of the $W$ boson have almost identical selection efficiencies at $E_T^{\text{miss}} = 100$ GeV because there is $E_T^{\text{miss},\text{True}}$ from the neutrino. However, the differences in selection efficiency are around a factor of four higher for $W \rightarrow \ell\nu$ than for $W \rightarrow e\nu$ at $E_T^{\text{miss}} = 350$ GeV. Over the entire $E_T^{\text{miss}}$ spectrum, the differences between the electron and muon final states for $W$ bosons are smaller than that for $Z$ bosons because there is a neutrino in $W \rightarrow \ell\nu$ events as opposed to none in the $Z \rightarrow \ell\ell$ final state.

In Fig. 16b, the selection efficiencies for CST $E_T^{\text{miss}}$ threshold requirements are divided by those obtained using the TST $E_T^{\text{miss}}$. The selection efficiencies resulting from CST $E_T^{\text{miss}}$ thresholds for SUSY, $t\bar{t}$, and VBF $H \rightarrow \tau\tau$ are within 10% of the efficiencies obtained using the TST $E_T^{\text{miss}}$. For $E_T^{\text{miss}}$ thresholds from 40–120 GeV, the selection efficiencies for $W$ and $Z$ boson events are higher by up to 60–160% for CST $E_T^{\text{miss}}$ than TST $E_T^{\text{miss}}$, which come from pileup contributions broadening the CST $E_T^{\text{miss}}$ distribution. The $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ events, which have no $E_T^{\text{miss},\text{True}}$, show an even larger increase of 2.6 times as many $Z \rightarrow ee$ events passing a $E_T^{\text{miss}}$ threshold of 50 GeV. The increase is not as large for $Z \rightarrow \mu\mu$ as $Z \rightarrow ee$ events because neither $E_T^{\text{miss}}$ algorithm accounts for forward muons ($|\eta| > 2.7$) as discussed above. Moving to a higher $E_T^{\text{miss}}$ threshold, mis-measured tracks in the TST algorithm cause it to select more $Z \rightarrow ee$ events with $120 < E_T^{\text{miss}} < 230$ GeV. In addition, the CST $E_T^{\text{miss}}$ also includes electron energy contributions ($p_T < 20$ GeV) in the forward calorimeters ($|\eta| > 3.1$) that the TST does not.

The CST and TST $E_T^{\text{miss}}$ distributions agree within 10% in selection efficiency for $E_T^{\text{miss}} > 250$ GeV for all physics processes shown. This demonstrates a strong correlation between the $E_T^{\text{miss}}$ distributions for events with large $E_T^{\text{miss},\text{True}}$, or a strong correlation between the physics objects that cause a large mismeasurement in $E_T^{\text{miss}}$ for $Z$ events.

6.7 Correlation of fake $E_T^{\text{miss}}$ between algorithms

The tracking and the calorimeters provide almost completely independent estimates of the $E_T^{\text{miss}}$. These two measurements complement each other, and the $E_T^{\text{miss}}$ algorithms discussed in this paper combine that information in different ways. The distribution of the TST $E_T^{\text{miss}}$ versus the CST $E_T^{\text{miss}}$ is shown for the simulated 0-jet $Z \rightarrow \mu\mu + 0$ jet sample in Fig. 17. This figure shows the correlation of fake $E_T^{\text{miss}}$ between the two algorithms, which originates from many sources including incorrect vertex association and miscalibration of high-$p_T$ physics objects.

Vector correlation coefficients [53], shown in Table 5, are used to estimate the correlation between the $E_T^{\text{miss}}$ distributions resulting from different reconstruction algorithms. The value of the vector correlation coefficients ranges from 0 to 2, with 0 being the least correlated and 2 being the most correlated. The coefficients shown are obtained using the simulated 0-jet and inclusive $Z \rightarrow \mu\mu$ MC samples. The least-correlated $E_T^{\text{miss}}$ distributions are the CST and Track $E_T^{\text{miss}}$, which use mostly independent momenta measurements in their reconstructions. The correlations of the other $E_T^{\text{miss}}$ distributions to the CST $E_T^{\text{miss}}$ decrease as more tracking information is used to suppress the pileup dependence of the soft term, with the TST $E_T^{\text{miss}}$ distribution having the second smallest vector correlation coefficient with respect to the CST $E_T^{\text{miss}}$ distribution. Placing requirements on a combination of $E_T^{\text{miss}}$ distributions or requiring the difference in azimuthal direction between two $E_T^{\text{miss}}$ vectors to be small
can greatly reduce fake $E_{T}^{\text{miss}}$ backgrounds, especially using the least-correlated $E_{T}^{\text{miss}}$ distributions. Such strategies are adopted in several Higgs boson analyses in ATLAS [54–56].

7 Jet-$p_T$ threshold and vertex association selection

Jets can originate from pileup interactions, so tracks matched to the jets are extrapolated back to the beamline to ascertain whether they are consistent with originating from the hard scatter or a pileup collision. The JVF defined in Sect. 4.1.1 is used to separate pileup jets and jets from the hard scatter. The STVF, EJAF, and TST $E_{T}^{\text{miss}}$ algorithms improve their jet identification by removing jets associated with pileup vertices or jets that have a large degradation in momentum resolution due to pileup activity. Energy contributions from jets not associated with the hard-scatter vertex are included in the soft term. For the TST, this means that charged particles from jets not associated with the hard-scatter vertex may then enter the soft term if their position along the beamline is consistent with the $z$-position of the hard-scatter vertex.

Applying a JVF cut is a trade-off between removing jets from pileup interactions and losing jets from the hard scatter.

Therefore, several values of the JVF selection criterion are considered in $Z \rightarrow \ell\ell$ events with jets having $p_T > 20$ GeV; their impact on the $E_{T}^{\text{miss}}$ resolution and scale is investigated in Fig. 18. Larger JVF thresholds on jets reduce the pileup dependence of the $E_{T}^{\text{miss}}$ resolution, but they simultaneously worsen the $E_{T}^{\text{miss}}$ scale. Thus the best compromise for the value of the JVT threshold is chosen. Requiring JVF > 0.25 greatly improves the stability of the $E_{T}^{\text{miss}}$ resolution with respect to pileup by reducing the dependence of the $E_{T}^{\text{miss}}$ resolution on the number of reconstructed vertices as shown in Fig. 18a. The $E_{T}^{\text{miss}}$ in $Z \rightarrow \ell\ell$ events ideally has a magnitude of zero, apart from some relatively infrequent neutrino contributions in jets. So its magnitude should be consistently zero along any direction. The $p_T^{\ell\ell}$ remains unchanged for different JVF requirements, which makes it a useful reference to check the calibration of the $E_{T}^{\text{miss}}$. The difference from zero of the average value of the reconstructed $E_{T}^{\text{miss}}$ along $p_T^{\ell\ell}$ increases as tighter JVF selections are applied as shown in Fig. 18b. Requiring a JVF threshold of 0.25 or higher slightly improves the stability of the resolution with respect to pileup, whereas it visibly degrades the $E_{T}^{\text{miss}}$ response by removing too many hard-scatter jets. Lastly, pileup jets with $p_T > 50$ GeV are very rare [4], so applying the JVF requirement above this $p_T$ threshold is not useful. Therefore, requiring JVF to be larger than 0.25 for jets with $p_T < 50$ GeV within the tracking volume ($|\eta| < 2.4$) is the preferred threshold for the $E_{T}^{\text{miss}}$ reconstruction.

In addition, the $p_T$ threshold, which defines the boundary between the jet and soft terms, is optimized. For these studies, the jets with $p_T > 20$ GeV and $|\eta| < 2.4$ are required to have JVF > 0.25. A procedure similar to that used for the JVF optimization is used for the jet-$p_T$ threshold using the same two metrics as shown in Fig. 19. While applying a higher $p_T$ threshold improves the $E_{T}^{\text{miss}}$ resolution versus the number of pileup vertices, by decreasing the slope, the
The systematic uncertainties of the soft term

The soft term, $E_T^{\text{miss}}$, becomes strongly biased in the direction opposite to $\vec{p}_T^Z$. Therefore, the $p_T$ threshold of 20 GeV is preferred.

8 Systematic uncertainties of the soft term

The $E_T^{\text{miss}}$ is reconstructed from the vector sum of several terms corresponding to different types of contributions from reconstructed physics objects, as defined in Eq. (2). The estimated uncertainties in the energy scale and momentum resolution for the electrons [14], muons [13], jets [44], $\tau_{\text{had-vis}}$ [47], and photons [14] are propagated into the $E_T^{\text{miss}}$. This section describes the estimation of the systematic uncertainties for the $E_T^{\text{miss}}$ soft term. These uncertainties take into account the impact of the generator and underlying-event modelling used by the ATLAS Collaboration, as well as effects from pileup.

The balance of the soft term with the calibrated physics objects is used to estimate the soft-term systematic uncertainties in $Z \rightarrow \mu\mu$ events, which have very small $E_T^{\text{miss, True}}$. The transverse momenta of the calibrated physics objects, $\vec{p}_T^{\text{hard}}$, is defined as

$$\vec{p}_T^{\text{hard}} = \sum \vec{p}_T^\mu + \sum \vec{p}_T^\tau + \sum \vec{p}_T^\gamma + \sum \vec{p}_T^{\text{jet}},$$

(15)

which is the vector sum of the transverse momenta of the high-$p_T$ physics objects. It defines an axis (with unit vector $\hat{p}_T^{\text{hard}}$) in the transverse plane of the ATLAS detector along which the $E_T^{\text{miss}}$ soft term is expected to balance $p_T^{\text{hard}}$ in $Z \rightarrow \mu\mu$ events. This balance is sensitive to the differences in calibration and reconstruction of the $E_T^{\text{miss,soft}}$ between data and MC simulation and thus is sensitive to the uncertainty in the soft term. This discussion is similar to the one in Sect. 6.2.

The $E_T^{\text{miss}}$ is decomposed along the $\hat{p}_T^{\text{hard}}$ direction. The direction orthogonal to $\hat{p}_T^{\text{hard}}$ is referred to as the per-
pendicular direction while the component parallel to \( \vec{p}_{\text{hard}} \) is labelled as the longitudinal direction. The projections of \( E_{\text{miss}}^{\text{soft}} \) along those directions are defined as:

\[
\begin{align*}
E_{\parallel}^{\text{miss,soft}} &= E_{\text{miss,soft}} \cos(\phi_{E_{\text{miss,soft}}^{\text{soft}}, \vec{p}_{\text{hard}}}), \\
E_{\perp}^{\text{miss,soft}} &= E_{\text{miss,soft}} \sin(\phi_{E_{\text{miss,soft}}^{\text{soft}}, \vec{p}_{\text{hard}}}),
\end{align*}
\]

The \( E_{\parallel}^{\text{miss,soft}} \) is sensitive to scale and resolution differences between the data and simulation because the soft term should balance the \( \vec{p}_{\text{hard}} \) in \( Z \rightarrow \mu\mu \) events. For a narrow range of \( p_{T}^{\text{hard}} \) values, the mean and width of the \( E_{\parallel}^{\text{miss,soft}} \) are compared between data and MC simulation. On the other hand, the perpendicular component, \( E_{\perp}^{\text{miss,soft}} \), is only sensitive to differences in resolution. A Gaussian function is fit to the \( E_{\text{miss}}^{\text{soft}} \) projected onto \( \vec{p}_{\text{hard}} \) in bins of \( p_{T}^{\text{hard}} \), and the resulting Gaussian mean and width are shown in Fig. 20. The mean increases linearly with \( p_{T}^{\text{hard}} \), because the soft term is not calibrated to the correct energy scale. On the other hand, the width is relatively independent of \( p_{T}^{\text{hard}} \), because the width is mostly coming from pileup contributions.

The small discrepancies in mean and width between data and simulation are taken as the systematic uncertainties for the scale and resolution, respectively. A small dependence on the average number of collisions per bunch crossing is observed for the scale and resolution uncertainties for high \( p_{T}^{\text{hard}} \), so the uncertainties are computed in three ranges of pileup and three ranges of \( p_{T}^{\text{hard}} \). The scale uncertainty varies from \(-0.4\) to \(0.3\) GeV depending on the bin, which reduces the uncertainties from the 5% shown in Fig. 20 for \( p_{T}^{\text{hard}} > 10\) GeV. A small difference in the uncertainties for the resolution along the longitudinal and perpendicular directions is observed, so they are considered separately. The average uncertainty is about 2.1% (1.8%) for the longitudinal (perpendicular) direction.

8.1.2 Cross-check method for the CST systematic uncertainties

As a cross-check of the method used to estimate the CST uncertainties, the sample of \( Z \rightarrow \mu\mu +0\)-jet events is also used to evaluate the level of agreement between data and simulation. The projection of the \( E_{\text{miss}}^{\text{soft}} \) onto \( \vec{p}_{\text{hard}} \) provides a test for potential biases in the \( E_{\text{miss}}^{\text{soft}} \) scale. The systematic uncertainty in the soft-term scale is estimated by comparing the ratio of data to MC simulation for \( E_{\text{miss}}^{\text{soft}} \cdot \vec{p}_{\text{had}} \) versus \( \Sigma E_{\text{T}} \) (CST) as shown in Fig. 21a. The average deviation from unity in the ratio of data to MC simulation is about 8%, which is taken as a flat uncertainty in the absolute scale. The systematic uncertainty in the soft-term resolution is estimated by evaluating the level of agreement between data and MC simulation in the \( E_{\text{x}}^{\text{miss}} \) and \( E_{\text{y}}^{\text{miss}} \) resolution as a function of the \( \Sigma E_{\text{T}} \) (CST) (Fig. 21b). The uncertainty on the soft-term resolution is about 2.5% and is shown as the band in the data/MC ratio.

Even though the distributions appear similar, the results in this section are derived by projecting the full \( E_{T}^{\text{miss}} \) onto the \( \vec{p}_{\text{hard}} \) in the 0-jet events, and are not directly comparable to the ones in Sect. 8.1.1, in which only the soft term is projected onto \( \vec{p}_{\text{hard}} \).

8.2 Methodology for TST and Track \( E_{T}^{\text{miss}} \)

A slightly different data-driven methodology is used to evaluate the systematic uncertainties in the TST and Track \( E_{T}^{\text{miss}} \). Tracks matched to jets that are included in the hard term are removed from the Track \( E_{T}^{\text{miss}} \) and are treated separately, as described in Sect. 8.2.3.

The method exploits the balance between the soft track term and \( \vec{p}_{\text{hard}} \) and is similar to the balance method for the CST. The systematic uncertainties are split into two compo-
The longitudinal ($E_{\parallel}^{\text{miss,soft}}$) and transverse ($E_{\perp}^{\text{miss,soft}}$) projections onto $p_T^\text{hard}$ as defined in Eq. (16).

The $E_{\parallel}^{\text{miss,soft}}$ in data is fit with the MC simulation convolved with a Gaussian function, and the fitted Gaussian mean and width are used to extract the differences between simulation and data. The fitted mean and width define the systematic uncertainties. For the perpendicular component, the simulation is only smeared by a Gaussian function of width $\sigma_\perp$ to match the data. The mean, which is set to zero in the fit, is very small in data and MC simulation because the hadronic recoil only affects $E_{\parallel}^{\text{miss,soft}}$.

The fitting is done in 5 or 10 GeV bins of $p_T^\text{hard}$ from 0–50 GeV, and a single bin for $p_T^\text{hard} > 50$ GeV.

An example fit is shown in Fig. 22 for illustration. The 1-jet selection with the JVF requirement is used to show that the differences between data and simulation, from the jet-related systematic uncertainties, are small relative to the differences in the soft-term modelling. The impact of the jet-related systematic uncertainties is less than 0.1% in the Gaussian smearing ($\sigma = 1.61$ GeV), indicating that the jet-related systematic uncertainties do not affect the extraction of the TST systematic uncertainties.

The Gaussian width squared of $E_{\parallel}^{\text{miss,soft}}$ and $E_{\perp}^{\text{miss,soft}}$ components and the fitted mean of $E_{\parallel}^{\text{miss,soft}}$ for data and MC simulation are shown versus $p_T^\text{hard}$ in Fig. 23. The systematic uncertainty squared of the convolved Gaussian width and the systematic uncertainty of the offset for the longitudinal component are shown in the bands. While the systematic uncertainties are applied to the MC simulation, the band is shown centred around the data to show that all MC generators plus parton shower models agree with the data within the assigned uncertainties. Similarly for the $E_{\perp}^{\text{miss,soft}}$, the width of the convolved Gaussian function for the perpendicular component is shown in the band. The ALPGEN+HERWIG simulation has the largest disagreement with data, so the Gaussian smearing parameters and offsets applied to the simulation are used as the systematic uncertainties in the soft term. The $p_T^\text{hard} > 50$ GeV bin has the smallest number of data entries; therefore, it has the largest uncertainties in the fitted mean and width. In this bin of the distribution shown in Fig. 23(a), the statistical uncertainty from the ALPGEN+HERWIG simulation, which is not the most discrepant from data, is added to the uncertainty band, and this results in a systematic uncertainty band that spans the differences in MC generators for $\sigma^2(E_{\parallel}^{\text{miss,soft}})$ for events with $p_T^\text{hard} > 50$ GeV.
The impact of uncertainties coming from the parton shower model, the number of jets, $\mu$ dependence, JER/JES uncertainties, and forward versus central jet differences was evaluated. Among the uncertainties, the differences between the generator and parton shower models have the most dominant effects. The total TST systematic uncertainty is summarized in Table 6.

### 8.2.1 Propagation of systematic uncertainties

The CST systematic uncertainties from the balance method defined in Sect. 8.1.1 are propagated to the nominal $E_{T}^{\text{miss,soft}}$ as follows:

\[
E_{\parallel,\text{reso}}^{\text{miss,soft}} = (1 \pm R_{\perp,\text{reso}})(E_{\parallel}^{\text{miss,soft}} - \langle E_{\parallel}^{\text{miss,soft}} \rangle) + \langle E_{\parallel}^{\text{miss,soft}} \rangle \tag{17a}
\]

\[
E_{\parallel,\text{scale}}^{\text{miss,soft}} = E_{\parallel}^{\text{miss,soft}} \pm \Delta_{\text{CST}} \tag{17b}
\]

where $E_{\parallel,\text{reso}}^{\text{miss,soft}}$ and $E_{\parallel,\text{scale}}^{\text{miss,soft}}$ are the values after propagating the resolution and scale uncertainties, respectively, in the longitudinal (perpendicular) directions. The mean values of parameters are denoted using angled brackets. The $\Delta_{\text{CST}}$ is the scale uncertainty, and the $R_{\parallel,\text{reso}}$ is the fractional resolution uncertainty taken from the lower portion of Fig. 20b. Both depend on the $p_{T}^{\text{hard}}$ and the average number of pileup

<table>
<thead>
<tr>
<th>$p_{T}^{\text{hard}}$ range (GeV)</th>
<th>$\Delta_{\text{TST}}$ (GeV)</th>
<th>$\sigma_{\parallel}$ (GeV)</th>
<th>$\sigma_{\perp}$ (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>0.3</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>10–15</td>
<td>0.4</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>15–20</td>
<td>0.6</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>20–25</td>
<td>0.7</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>25–30</td>
<td>0.8</td>
<td>1.9</td>
<td>1.7</td>
</tr>
<tr>
<td>30–35</td>
<td>1.0</td>
<td>2.1</td>
<td>1.8</td>
</tr>
<tr>
<td>35–40</td>
<td>1.1</td>
<td>2.4</td>
<td>2.1</td>
</tr>
<tr>
<td>40–50</td>
<td>1.2</td>
<td>2.6</td>
<td>2.2</td>
</tr>
<tr>
<td>&gt;50</td>
<td>1.4</td>
<td>5.2</td>
<td>2.7</td>
</tr>
</tbody>
</table>

---

Fig. 23 The fitted TST a $\sigma^{2}(E_{T}^{\text{miss,soft}})$, b $\sigma^{2}(E_{\perp}^{\text{miss,soft}})$, and c $\langle E_{\parallel}^{\text{miss,soft}} \rangle$ in each case versus $p_{T}^{\text{hard}}$ are shown in data and Alpgen+Herwig, Alpgen+Pythia6, Powheg+Pythia8, and Alpgen+Pythia $Z \rightarrow \mu \mu$ simulation. The error bars on the data and MC simulation points are the errors from the Gaussian fits. The solid band, which is centred on the data, shows the parameter’s systematic uncertainties from Table 6. The insets at the bottom of the figures show the ratios of the MC predictions to the data.
interactions per bunch crossing. Each propagation of the systematic uncertainties in Eq. (17b) is called a variation, and all of the variations are used in ATLAS analyses.

The systematic uncertainties in the resolution and scale for the CST using the cross-check method defined in Sect. 8.1.2 are propagated to the nominal $E_T^{\text{miss},\text{soft}}$ as follows:

\begin{align}
E_{x(y),\text{reso}}^{\text{miss},\text{soft}} &= E_{x(y)}^{\text{miss},\text{soft}} \cdot \text{Gaus}(1, \delta_{\text{CST}}), \\
E_{x(y),\text{scale}}^{\text{miss},\text{soft}} &= E_{x(y)}^{\text{miss},\text{soft}} \cdot (1 + \delta),
\end{align}

where $E_{x(y),\text{reso}}^{\text{miss},\text{soft}}$ and $E_{x(y),\text{scale}}^{\text{miss},\text{soft}}$ are the values after propagating the resolution and scale uncertainties, respectively, in the $x$ ($y$) directions. Here, $\delta$ is the fractional scale uncertainty, and $\delta_{\text{CST}}$ corrects for the differences in resolution between the data and simulation.

The systematic uncertainties in the resolution and scale for the TST $E_T^{\text{miss},\text{soft}}$ are propagated to the nominal $E_T^{\text{miss},\text{soft}}$ as follows:

\begin{align}
E_{\parallel(\perp),\text{reso}}^{\text{miss},\text{soft}} &= E_{\parallel(\perp)}^{\text{miss},\text{soft}} + \text{Gaus}(\Delta_{\text{TST}}, \sigma_{\parallel(\perp)}), \\
E_{\parallel,\text{scale}}^{\text{miss},\text{soft}} &= E_{\parallel}^{\text{miss},\text{soft}} \pm \Delta_{\text{TST}}.
\end{align}

The symbol $\text{Gaus}(\Delta_{\text{TST}}, \sigma_{\parallel(\perp)})$ represents a random number sampled from a Gaussian distribution with mean $\Delta_{\text{TST}}$ and width $\sigma_{\parallel(\perp)}$. The shift $\Delta_{\text{TST}}$ is zero for the perpendicular component. All of the TST systematic-uncertainty variations have a wider distribution than the nominal MC simulation, when the Gaussian smearing is applied. To cover cases in which the data have a smaller resolution (narrower distribution) than MC simulation, a downward variation is computed using Eq. (20). To compute the yield of predicted events in the variation, $Y_{\text{down}}(X)$, for a given value $X$ of the $E_T^{\text{miss}}$, the yield is defined as the

$$Y_{\text{down}}(X) = \frac{[Y(X)]^2}{Y_{\text{smeared}}(X)},$$

(20)

where the square of the yield of the nominal distribution, $Y(X)$, is divided by the yield of events after applying the variation with Gaussian smearing to the kinematic variable, $Y_{\text{smeared}}(X)$. In practice, the yields are typically the content of histogram bins before ($Y(X)$) and after ($Y_{\text{smeared}}(X)$) the systematic uncertainty variations. This procedure can be applied to any kinematic observable by propagating only the smeared soft-term variation to the calculation of the kinematic observable $X$ and then computing the yield $Y_{\text{down}}(X)$ as defined in Eq. (20).

There are six total systematic uncertainties associated with the TST:

- Increase scale ($E_{\parallel,\text{scale}}^{\text{miss},\text{soft}}$)
- Decrease scale ($E_{\parallel,\text{scale}}^{\text{miss},\text{soft}}$)
- Gaussian smearing of $E_{\parallel}^{\text{miss},\text{soft}}$ ($E_{\parallel,\text{reso}}^{\text{miss},\text{soft}}$)
- The downward variation of the above $E_{\parallel,\text{reso}}^{\text{miss},\text{soft}}$ computed using Eq. (20)
- Gaussian smearing of $E_{\perp}^{\text{miss},\text{soft}}$ ($E_{\perp,\text{reso}}^{\text{miss},\text{soft}}$)
- The downward variation of the above $E_{\perp,\text{reso}}^{\text{miss},\text{soft}}$ computed using Eq. (20)

8.2.2 Closure of systematic uncertainties

The systematic uncertainties derived in this section for the CST and TST $E_T^{\text{miss}}$ are validated by applying them to the $Z \rightarrow \mu\mu$ sample to confirm that the differences between data and MC simulation are covered.
Fig. 25 Distributions of a $E_{\text{T}}^{\text{miss, soft}}$ and b $E_{\text{T}}^{\text{miss}}$ with the CST algorithm. Data are compared to the nominal simulation distribution as well as those resulting from applying the shifts/smearing according to the scale and resolution systematic uncertainties on the $E_{\text{T}}^{\text{miss, soft}}$. The resulting changes from the variations are added in quadrature, and the insets at the bottom of the figures show the ratios of the data to the MC predictions. The uncertainties are estimated from the data/simulation ratio in Sect. 8.1.2.

Fig. 26 Distributions of a $E_{\text{T}}^{\text{miss, soft}}$ and b $E_{\text{T}}^{\text{miss}}$ with the TST algorithm. Data are compared to the nominal simulation distribution as well as those resulting from applying the scale and resolution systematic uncertainties to the $E_{\text{T}}^{\text{miss, soft}}$ and adding the variations in quadrature, and the insets at the bottom of the figures show the ratios of the data to the MC predictions. The uncertainties are estimated from the method in Sect. 8.2.

The effects of these systematic uncertainty variations on the CST $E_{\text{T}}^{\text{miss}}$ are shown for the $Z \rightarrow \mu\mu$ events in Figs. 24 and 25 for the primary (Sect. 8.1.1) and the cross-check (Sect. 8.1.2) methods, respectively. The uncertainties are larger for the cross-check method, reaching around 50% for $E_{\text{T}}^{\text{miss, soft}} > 60$ GeV in Fig. 25a.

The corresponding plots for the TST $E_{\text{T}}^{\text{miss}}$ are shown in Fig. 26 using the $Z \rightarrow \mu\mu + 0$-jet control sample, where the uncertainty band is the quadratic sum of the variations with the MC statistical uncertainty. The systematic uncertainty band for the TST is larger in Fig. 26a than the one for the primary CST algorithm. In all the distributions, the systematic uncertainties in the soft term alone cover the disagreement between data and MC simulation.

8.2.3 Systematic uncertainties from tracks inside jets

A separate systematic uncertainty is applied to the scalar summed $p_{\text{T}}$ of tracks associated with high-$p_{\text{T}}$ jets in the Track $E_{\text{T}}^{\text{miss}}$ because these tracks are not included in the TST. The fraction of the momentum carried by charged particles within jets was studied in ATLAS [57], and its uncertainty varies from 3 to 5% depending on the jet $\eta$ and $p_{\text{T}}$. These uncertainties affect the azimuthal angle between the Track $E_{\text{T}}^{\text{miss}}$ and the TST $E_{\text{T}}^{\text{miss}}$, so the modelling is checked with $Z \rightarrow \mu\mu$ events produced with one jet. The azimuthal angle between the Track $E_{\text{T}}^{\text{miss}}$ and the TST $E_{\text{T}}^{\text{miss}}$ directions is well modelled, and the differences between data and MC simulation are within the systematic uncertainties.
Weakly interacting particles, which leave the ATLAS detector undetected, give rise to a momentum imbalance in the plane transverse to the beamline. An accurate measurement of the missing transverse momentum ($E_T^{miss}$) is thus important in many physics analyses to infer the momentum of these particles. However, additional interactions occurring in a given bunch crossing as well as residual signatures from nearby bunch crossings make it difficult to reconstruct the $E_T^{miss}$ from the hard-scattering process alone.

The $E_T^{miss}$ is computed as the negative vector sum of the reconstructed physics objects including electrons, photons, muons, $\tau$-leptons, and jets. The remaining energy deposits not associated with those high-$p_T$ physics objects are also considered in the $E_T^{miss}$. They collectively form the so-called soft term, which is the $E_T^{miss}$ component most affected by pileup. The calorimeter and the tracker in the ATLAS detector provide complementary information to the reconstruction of the high-$p_T$ physics objects as well as the $E_T^{miss}$ soft term. Charged particles are matched to a particular collision point or vertex, and this information is used to determine which charged particles originated from the hard-scatter collision. Thus tracking information can be used to greatly reduce the pileup dependence of the $E_T^{miss}$ reconstruction. This has resulted in the development of $E_T^{miss}$ reconstruction algorithms that combine the information from the tracker and the calorimeter. The performance of these reconstruction algorithms is evaluated using data from 8 TeV proton–proton collisions collected with the ATLAS detector at the LHC corresponding to an integrated luminosity of 20.3 fb$^{-1}$.

The Calorimeter Soft Term (CST) is computed from the sum of calorimeter topological clusters not associated with any hard object. No distinction can be made between energy contributions from pileup and hard-scatter interactions, which makes the resolution on the $E_T^{miss}$ energy contributions from pileup and hard-scatter interactions. Thus tracking information can be used to greatly reduce the pileup dependence of the $E_T^{miss}$ reconstruction. This has resulted in the development of $E_T^{miss}$ reconstruction algorithms that combine the information from the tracker and the calorimeter. The performance of these reconstruction algorithms is evaluated using data from 8 TeV proton–proton collisions collected with the ATLAS detector at the LHC corresponding to an integrated luminosity of 20.3 fb$^{-1}$.

The Calorimeter Soft Term (CST) is computed from the sum of calorimeter topological clusters not associated with any hard object. No distinction can be made between energy contributions from pileup and hard-scatter interactions, which makes the resolution on the $E_T^{miss}$ energy contributions from pileup and hard-scatter interactions. Thus tracking information can be used to greatly reduce the pileup dependence of the $E_T^{miss}$ reconstruction. This has resulted in the development of $E_T^{miss}$ reconstruction algorithms that combine the information from the tracker and the calorimeter. The performance of these reconstruction algorithms is evaluated using data from 8 TeV proton–proton collisions collected with the ATLAS detector at the LHC corresponding to an integrated luminosity of 20.3 fb$^{-1}$.

The Track Soft Term (TST) algorithm does not use calorimeter energy deposits in the soft term and uses only the inner detector (ID) tracks. It has stable $E_T^{miss}$ resolution with respect to the amount of pileup; however, it does not have as good a response as the CST $E_T^{miss}$ due mainly to missing neutral particles in the soft term. Nevertheless, its response is better than that of the other reconstruction algorithms that aim to combine the tracking and calorimeter information. For large values of $E_T^{miss}$, the CST and TST $E_T^{miss}$ algorithms all perform similarly. This is because contributions from jets dominate the $E_T^{miss}$ performance, making the differences in soft-term reconstruction less important.

The Extrapolated Jet Area with Filter (EJAF) and Soft-Term Vertex-Fraction (STVF) $E_T^{miss}$ reconstruction algorithms correct for pileup effects in the CST $E_T^{miss}$ by utilizing a combination of the ATLAS tracker and calorimeter measurements. Both apply a vertex association to the jets used in the $E_T^{miss}$ calculation. The EJAF soft-term reconstruction subtracts the pileup contributions to the soft term using a procedure similar to jet area-based pileup corrections, and the EJAF $E_T^{miss}$ resolution has a reduced dependence on the amount of pileup, relative to the CST algorithm. The STVF reconstruction algorithm uses an event-level correction of the CST, which is the scalar sum of charged-particle $p_T$ from the hard-scatter vertex divided by the scalar sum of all charged-particle $p_T$. The STVF correction to the soft term greatly decreases the dependence of the $E_T^{miss}$ resolution on the amount of pileup but causes the largest under-estimation of all the soft-term algorithms.

Finally, the Track $E_T^{miss}$ reconstruction uses only the inner detector tracks with the exception of the reconstructed electron objects, which use the calorimeter $E_T$ measurement. The resolutions on the Track $E_T^{miss}$ magnitude and direction are very stable against pileup, but the limited $|\eta|$ coverage of the tracker degrades the $E_T^{miss}$ response, as does not accounting for high-$p_T$ neutral particles, especially in events with many jets.

The different $E_T^{miss}$ algorithms have their own advantages and disadvantages, which need to be considered in the context of each analysis. For example, removing large backgrounds with low $E_T^{miss}$, such as Drell–Yan events, may require the use of more than one $E_T^{miss}$ definition. The tails of the track and calorimeter $E_T^{miss}$ distributions remain uncorrelated, and exploiting both definitions in parallel allows one to suppress such backgrounds even under increasing pileup conditions.

The systematic uncertainties in the $E_T^{miss}$ are estimated with $Z \rightarrow \mu\mu$ events for each reconstruction algorithm, and are found to be small.

**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 CF, 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; CNRST, Georgia; BMBF, GfK, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRSST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; PME,Standard Model (SM) predictions, ATLAS Collaboration; Steven B. Goto, Université de Lausanne; the Kavli Institute for Cosmology, University of Cambridge. The results presented here contribute to the following projects: CMS Collaboration; ATLAS Collaboration; CERN; the Kavli Institute for Theoretical Physics, University of California, Santa Barbara; the Kavli Institute for Cosmology, University of Cambridge.
Appendix

A. Calculation of EJAF

A jet-level \( \eta \)-dependent pileup correction of the form

\[
\rho_{\eta}^{\text{med}}(\eta) = \rho_{\text{evt}}^{\text{med}} \cdot P_f(\eta, N_{\text{PV}}, \langle \mu \rangle),
\]

is used, where the \( N_{\text{PV}} \) and \( \langle \mu \rangle \) are determined from the event properties. This multiplies the median soft-term jet \( p_T \)-density, \( \rho_{\text{evt}}^{\text{med}} \), from Eq. (7) by the functional form, \( P_f(\eta, N_{\text{PV}}, \langle \mu \rangle) \) defined in Eq. (9), which was fit to the average transverse momentum density. The median transverse momentum density \( \rho_{\text{evt}}^{\text{med}} \) is determined from soft-term jets with \( |\eta| < 2 \) and extrapolated to higher \( |\eta| \) as discussed in Sect. 4.1.2 using the fitted \( P_f(\eta, N_{\text{PV}}, \langle \mu \rangle) \).

The pileup correction \( \rho_{\eta}^{\text{med}}(\eta) \) from Eq. (21) is applied to the transverse momenta of the soft-term jets passing a JVF selection. The pileup-corrected jet \( p_T \) is labelled \( p_{T,i}^{\text{filter-jet,corr}} \), and it is computed as

\[
p_{T,i}^{\text{filter-jet,corr}} = \begin{cases} 
0 & \text{if } p_{T,i}^{\text{filter-jet}} < \rho_{\eta}^{\text{med}}(\eta_i^{\text{filter-jet}}) \cdot A_i^{\text{filter-jet}} \cdot \rho_{\eta}^{\text{med}}(\eta_i^{\text{filter-jet}}) \cdot A_i^{\text{filter-jet}}, \text{ and to the jet } \eta (\eta_i^{\text{filter-jet}}) \text{ variables.}
\end{cases}
\]

While all other jets used in this paper use an \( R = 0.4 \) reconstruction, the larger value of \( R = 0.6 \) is used to reduce the number of \( k_t \) soft-term jets with \( p_T = 0 \) (see Eq. (22)) in the central detector region. While negative energy deposits are possible in the ATLAS calorimeters, their contributions cannot be matched to the soft-term jets by ghost-association. Studies that modify the cluster-to-jet matching to include negative-\( p_T \) clusters indicate no change in the \( E_T^{\text{miss}} \) performance, so negative-\( p_T \) clusters are excluded from the soft-term jets. Finally, only filter-jets with \( p_{T,i}^{\text{filter-jet}} \) larger than the pileup correction contribute to the EJAF soft term.

References

5. ATLAS Collaboration, Data-quality requirements and event cleaning for jets and missing transverse energy reconstruction with the ATLAS detector in proton–proton collisions at a center of mass energy \( \sqrt{s} = 7 \) TeV. ATLAS-CONF-2010-038 (2010). http://cds.cern.ch/record/1277678
ATLAS Collaboration


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

Department of Physics, Indiana University, Bloomington, IN, USA

Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria

University of Iowa, Iowa City, IA, USA

Department of Physics and Astronomy, Iowa State University, Ames, IA, USA

Joint Institute for Nuclear Research, JINR Dubna, Dubna, Russia

KEK, High Energy Accelerator Research Organization, Tsukuba, Japan

Graduate School of Science, Kobe University, Kobe, Japan

Faculty of Science, Kyoto University, Kyoto, Japan

Kyoto University of Education, Kyoto, Japan

Department of Physics, Kyushu University, Fukuoka, Japan

Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina

Physics Department, Lancaster University, Lancaster, UK

(a) INFN Sezione di Lecce, Lecce, Italy; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy

Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK

Department of Physics, Jožef Stefan Institute, University of Ljubljana, Ljubljana, Slovenia

School of Physics and Astronomy, Queen Mary University of London, London, UK

Department of Physics, Royal Holloway University of London, Surrey, UK

Department of Physics and Astronomy, University College London, London, UK

Louisiana Tech University, Ruston, LA, USA

Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France

Fysiska institutionen, Lunds universitet, Lund, Sweden

Departamento de Fisica Teorica C-15, Universidad Autonoma de Madrid, Madrid, Spain

Institut für Physik, Universität Mainz, Mainz, Germany

School of Physics and Astronomy, University of Manchester, Manchester, UK

CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France

Department of Physics, University of Massachusetts, Amherst, MA, USA

Department of Physics, McGill University, Montreal, QC, Canada

School of Physics, University of Melbourne, Victoria, Australia

Department of Physics, The University of Michigan, Ann Arbor, MI, USA

Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA

(a) INFN Sezione di Milano, Milan, Italy; (b) Dipartimento di Fisica, Università di Milano, Milan, Italy

B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus

National Scientific and Educational Centre for Particle and High Energy Physics, Minsk, Republic of Belarus

Group of Particle Physics, University of Montreal, Montreal, QC, Canada

P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia

Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia

National Research Nuclear University MEPhI, Moscow, Russia

D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia

Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany

Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Munich, Germany

Nagasaki Institute of Applied Science, Nagasaki, Japan

Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan

(a) INFN Sezione di Napoli, Naples, Italy; (b) Dipartimento di Fisica, Università di Napoli, Naples, Italy

Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA

Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands

Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands

Department of Physics, Northern Illinois University, DeKalb, IL, USA

Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia