Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking

The ATLAS Collaboration

DOI
10.1007/JHEP08(2020)080

Publication date
2020

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):

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

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: The factor of four increase in the LHC luminosity, from $0.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ to $2.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, and the corresponding increase in pile-up collisions during the 2015–2018 data-taking period, presented a challenge for the ATLAS trigger, particularly for those algorithms that select events with missing transverse momentum. The output data rate at fixed threshold typically increases exponentially with the number of pile-up collisions, so the legacy algorithms from previous LHC data-taking periods had to be tuned and new approaches developed to maintain the high trigger efficiency achieved in earlier operations. A study of the trigger performance and comparisons with simulations show that these changes resulted in event selection efficiencies of $> 98\%$ for this period, meeting and in some cases exceeding the performance of similar triggers in earlier run periods, while at the same time keeping the necessary bandwidth within acceptable limits.

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 2005.09554
1 Introduction

The trigger system [1] of the ATLAS experiment [2] is responsible for deciding which proton-proton ($pp$) bunch-crossing events are kept for later analysis. Storage and processing requirements limit the fraction of events that can be retained to the order of $10^{-5}$, with the rest being discarded and hence unavailable for further physics analysis.
Particles that interact via neither the strong nor the electromagnetic force, and that escape the experiment without decaying, leave no visible signature. Efficient trigger selection of events that contain such invisible particles is nevertheless essential for much of the ATLAS physics programme. Examples include searches for decays of the Higgs boson into invisible final states [3, 4], searches for new charged Higgs bosons decaying into $\tau\nu$ [5], searches for dark matter based on, for example, events in which invisible particles recoil against a single energetic jet [6], supersymmetry searches that involve a stable and invisible neutralino [7, 8], top-quark scalar partner searches [9] and searches for final states with stable long-lived particles [10]. Another recent example is the Standard Model (SM) Higgs boson decay into $b$-quarks [11], a process first observed in events in which the Higgs boson was produced in association with a $Z$ boson which itself decayed into unobserved neutrinos.

Selecting events that contain invisible particles is particularly difficult, precisely because such particles do not register in the detector. The strategy employed is to deduce the presence of these invisible particles from the apparent imbalance of the momentum calculated from the visible particles. In practice the imbalance in the direction parallel to the proton beams is not sensitive since the fraction of each proton’s momentum that participates in the collision is unknown, and much of the outgoing momentum in the beam direction is not observed. Instead, the momentum imbalance in the plane perpendicular to the proton beams is the quantity of most interest; it is known as the missing transverse momentum, and its magnitude is conventionally denoted by $E_{\text{T}}^{\text{miss}}$.

The $E_{\text{T}}^{\text{miss}}$ triggers used by ATLAS are based on transverse momentum imbalance within the calorimeter only. Muons are approximately invisible in the calorimeter [12], and so are treated in these calculations much like neutrinos. Neglecting muons results in a negligible cost in terms of additional trigger rate since events containing muons with large transverse momentum are rare. These calorimeter-only algorithms also have the advantage that they can efficiently select events that contain high-$p_T$ muons. For example, the $E_{\text{T}}^{\text{miss}}$ trigger is also used to select events in which the Higgs boson is produced in association with a $Z$ boson decaying into muons, or events containing a $W$ boson decaying into $\mu\nu$ [11].

Given that the selection of events by the $E_{\text{T}}^{\text{miss}}$ trigger is based on energy deposited throughout the calorimeter, there are particular reconstruction challenges. ATLAS employs a trigger system that uses a region-of-interest trigger strategy [1] where the lowest-level trigger identifies potentially interesting objects in each event and, for those events that satisfy the selection criteria, it provides regions of interest to be further analysed by the higher-level trigger. This technique of reconstructing objects only in particular regions of the detector is useful for simplifying the computational task, but generally unsuited to $E_{\text{T}}^{\text{miss}}$ triggers which must sum momenta over the full solid angle that is instrumented.

The most significant challenge to the $E_{\text{T}}^{\text{miss}}$ triggers during the 13 TeV Run-2 data-taking period (2015–2018) was the factor of four increase in the number of proton-proton collisions occurring within each bunch crossing. The additional collisions, known as pile-up, were a consequence of the corresponding increase in LHC luminosity from $0.5 \times 10^{34}$ cm$^{-2}$s$^{-1}$ in 2015 to $2.0 \times 10^{34}$ cm$^{-2}$s$^{-1}$ in 2017 and 2018. The peak luminosity of $2.0 \times 10^{34}$ cm$^{-2}$s$^{-1}$ was achieved with 2544 bunches of circulating protons, a mean number of $pp$ interactions per bunch crossing ($\mu$) = 56, and a peak pile-up of 70 interactions. The energy from the ad-
ditional pile-up collisions is deposited throughout the detector. Due to the shaping time of the front-end electronics, the calorimeter response is affected by pile-up from several preceding bunch crossings [13]. The overall effect of both forms of pile-up is to degrade the $E_{\text{T}}^{\text{miss}}$ resolution of the detector. With the existing Run-1 algorithms, this rise in pile-up would have led to an unacceptable order-of-magnitude increase in trigger rate unless thresholds were raised, and that would in turn have significantly diminished the signal efficiencies.

This paper describes algorithms introduced during Run 2 that provide greater pile-up resilience and background rejection while maintaining a signal acceptance similar to that in Run 1. These algorithms were able to keep output rates within a tolerable 100 Hz even at $\langle \mu \rangle = 56$. The design of these algorithms is described in detail, and comparative studies of their performance using data and simulation are provided.

The paper is organized as follows. Section 2 describes the ATLAS detector. The $E_{\text{T}}^{\text{miss}}$ trigger algorithms are introduced in section 3. The offline $E_{\text{T}}^{\text{miss}}$ algorithm against which the trigger is compared is defined in section 4. The trigger performance studies and their results are described in section 5. The conclusions are presented in section 6.

2 ATLAS detector

The ATLAS detector [2] at the LHC covers nearly the entire solid angle around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnet systems.

The inner-detector system is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the collision vertex region [14]. It is followed by the silicon microstrip tracker. These silicon detectors are complemented by the transition radiation tracker.

The calorimeter system has approximately 188,000 cells and covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) sampling calorimeters (ECAL), with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. The ECAL is between 24 and 27 radiation lengths ($X_0$) deep, and its granularity in the barrel in terms of $\Delta \eta \times \Delta \phi$ is typically $0.025 \times \pi/128$, with variations in segmentation with layer and $|\eta|$ as described in ref. [13].

Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter (HCAL), segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules (FCAL) optimized for electromagnetic (FCAL1) and hadronic (FCAL2 and FCAL3) measurements respectively. The combined depth of the

---

[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 upwards. Cylindrical coordinates ($r, \phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular separation is measured in units of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

---
calorimeters for hadronic energy measurements is more than 10 nuclear interaction lengths nearly everywhere across the full detector acceptance ($|\eta| < 4.9$). The granularity is as fine as $0.1 \times \pi/32$, again with variations in segmentation with layer and $|\eta|$ as described in ref. [13].

The muon spectrometer comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroids. The field integral of the toroids ranges between $2.0$ and $6.0 \text{T m}$ across most of the detector. A set of precision chambers covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive plate chambers in the barrel and thin gap chambers in the endcap regions.

A two-level trigger system is used to select interesting events [1]. It consists of a hardware-based first-level trigger (Level-1, L1) and a software-based high-level trigger (HLT) running on a farm of approximately 50 k processing units. The L1 trigger decision is formed by the Central Trigger Processor, which receives inputs from the L1 calorimeter (L1Calo) [15] and L1 muon triggers as well as several other subsystems. The L1 trigger decision is formed with a latency of $2.2 \mu$s. The HLT has access to the full event and a decision is made within an average time of 500 ms.

3 Description of the $E_{\text{T}}^{\text{miss}}$ trigger algorithms

The operational demands of the trigger prioritize low latency, rapid processing, and large background rejection while making use of limited detector information. Thus the online $E_{\text{T}}^{\text{miss}}$ trigger algorithms are specifically designed for this purpose and so differ from the offline $E_{\text{T}}^{\text{miss}}$ reconstruction algorithms used in subsequent physics analyses [16, 17].

The ATLAS HLT processes approximately 100 kHz of L1 accepted events, of which about 5 to 10 kHz come from the L1 $E_{\text{T}}^{\text{miss}}$ trigger. The HLT $E_{\text{T}}^{\text{miss}}$ algorithms accept events at a rate of about 1200 Hz averaged over a typical LHC fill [18]. The requirement that the $E_{\text{T}}^{\text{miss}}$ algorithms utilize not more than $O(100 \text{ ms})$ makes the use of inner-detector tracking information generally too computationally expensive, since the corresponding evaluation time can take $O(1-5 \text{ s})$. Thus, all of the algorithms described below use only the calorimeter.\footnote{Here and in what follows it should be understood that the singular ‘calorimeter’ refers to the calorimeter system as a whole.}

For all algorithms the energy measured by the calorimeter is associated with some set of energy depositions, generally referred to as elements. The definition of the set of elements is algorithm-dependent. For example, the set of elements could be all of the calorimeter cells or the reconstructed jets. In each case, the individual elements characterize the local energy deposits, while the complete set captures the overall distribution of energy in the calorimeter. Elements are indexed by the label $i$: the energy $E_i$ deposited in each element is also associated with a polar angle $\theta_i$ (or equivalently a pseudorapidity $\eta_i$) and an azimuthal angle $\phi_i$. 
The components of the missing transverse momentum two-vector $\vec{E}_{\text{miss}}^T$ are calculated from the energy in the elements in the approximation of massless particles

$$E_{\text{miss}}^x = - \sum_{i=1}^{\text{Elements}} E_i \sin \theta_i \cos \phi_i,$$

$$E_{\text{miss}}^y = - \sum_{i=1}^{\text{Elements}} E_i \sin \theta_i \sin \phi_i,$$

(3.1)

where $\text{Elements}$ indicates the number of elements. The magnitude $E_{\text{miss}}^T = \sqrt{(E_{\text{miss}}^x)^2 + (E_{\text{miss}}^y)^2}$ of this two-vector is used in the selection of candidate events for further study. The quantity $E_T = E_i \sin \theta_i$ is conventionally known as the transverse energy, and is useful in characterizing events. The total transverse energy in the calorimeter is given by the scalar sum

$$\Sigma E_T = \sum_{i=1}^{\text{Elements}} E_i \sin \theta_i.$$

The algorithms used are presented in the following sections. They differ in how they select the elements which enter into the sums and in how they make corrections to the energies of the elements.

### 3.1 Level-1 trigger

The ATLAS L1 trigger is implemented in firmware running on custom-made electronics [15]. Analogue sums of the input signals from calorimeter cells forming projective towers are digitized, with the granularity in the projective coordinates $\eta$ and $\phi$ being approximately $\Delta \eta \times \Delta \phi = 0.1 \times 0.1$ for the detector region $|\eta| < 2.5$ and both larger and less regular for $|\eta| > 2.5$, as described in ref. [15]. The digitization results in counts that nominally correspond to 1 GeV in $E_T$. A fixed threshold that depends on $\eta$ is then applied per tower: the energy $E_i$ of any tower which is below this threshold is set to zero in the subsequent calculations. The threshold is adjusted to provide a fixed occupancy of 0.5–1% based on data unbiased by a trigger selection. This occupancy threshold is optimized to give an acceptable rate for a trigger that efficiently selects events with $E_{\text{miss}}^T > 150$ GeV. As the LHC luminosity increased, the occupancy tended to grow, leading to higher thresholds as described in ref. [19].

The calorimeter noise thresholds vary from 1 to 9 GeV depending on the pseudorapidity and whether the calorimeter layer is electromagnetic or hadronic. The noise thresholds were periodically reoptimized during the period under study, particularly when the collider parameters, and as a result the pile-up, were varied. In the performance studies that follow, particular attention is paid to three periods during 2017 that have different pile-up distributions; these periods are labelled with the symbols $\alpha$, $\beta$ and $\gamma$. The pile-up at the start of the LHC fill increased from around $\langle \mu \rangle = 40$ for period $\alpha$ to $\langle \mu \rangle = 60$ for period $\gamma$. The largest changes in threshold occurred for the towers with $4.0 < |\eta| < 4.9$ in the electromagnetic layer, and the thresholds were 6, 7 and 9 GeV for periods $\alpha$, $\beta$ and $\gamma$ respectively. After the threshold is applied, the towers are summed into larger projective towers which have an approximate granularity of $\Delta \eta \times \Delta \phi = 0.2 \times 0.2$ and are referred
to as jet elements. The $\vec{E}_T^{\text{miss}}$ is then computed by summing the $x$ and $y$ projections of the jet elements using eq. (3.1).

Events that are accepted by the L1 trigger are transferred to the HLT where the $E_T^{\text{miss}}$ is recalculated using one or more of the algorithms described in sections 3.2 to 3.5.

### 3.2 Trigger using calorimeter cell signals (cell)

The most basic HLT algorithm, cell, determines $\vec{E}_T^{\text{miss}}$ from a sum over the full set of 188 k calorimeter cells to determine $E_x$ and $E_y$, without adjusting for hadronic vs electromagnetic calibrations or for pile-up corrections. To reduce the effect of noise from electronics and pile-up, only cells satisfying $|E_i| > 2\sigma_i$ are included in this sum. Here $\sigma_i$ is the expected energy-equivalent noise in cell $i$ described in ref. [13]. Its value is based on expectations for electronic noise and pile-up prior to data taking. Negative energy cells are included because the LAr electronics are designed so that signals from pile-up in later bunch crossings appear as negative energy and so tend to cancel energy deposits from earlier pile-up signals [20]. For the 2015 and 2016 run periods, the noise thresholds were configured for an average number of interactions $\langle \mu \rangle = 30$. For 2017 and 2018 they were configured for $\langle \mu \rangle = 40$. In addition, the requirement $E_i > -5\sigma_i$ is used to protect against spurious large negative cell signals.

### 3.3 Trigger using topological clusters of calorimeter cells (tc_lcw)

The topological clustering [13] of calorimeter cells forms an early stage of many ATLAS reconstruction algorithms. It offers the possibility of identifying clusters as either electromagnetic or hadronic in origin, and thus allows appropriate calibration (‘local cell weighting’) before using them as inputs for jet reconstruction and calculation of $E_T^{\text{miss}}$.

Topological clusters are formed in a multistage process. First the algorithm identifies calorimeter seed cells each with $|E_i| > 4\sigma_i$. All cells neighbouring a seed cell are collected in all three spatial dimensions and added to the cluster. If any of those neighbouring cells satisfy $|E_i| > 2\sigma_i$, then their neighbours are collected as well, and the process continues iteratively until no further neighbours satisfying the requirement can be identified. Finally, all neighbouring cells are added to the cluster, regardless of their energy. After this initial cluster formation, an algorithm is run which splits clusters between local signal maxima (again, in three dimensions). The energies of these clusters are corrected for the type of energy deposit after each one has been classified as being either electromagnetic or hadronic in origin.

These energy-calibrated clusters can be used directly in an $E_T^{\text{miss}}$ calculation, which is denoted tc_lcw. These topological clusters also form the inputs to all of the following algorithms.

### 3.4 Trigger based on jets (mht)

In most events of interest, hadronic jets tend to dominate the visible momentum. Since these jets can be calibrated accurately [21], there is good motivation to use them as the basis of an $E_T^{\text{miss}}$ calculation. In addition, the calculation of the $E_T^{\text{miss}}$ from the calorimeter signals described previously includes energy from pile-up, while jets are corrected on-average for pile-up effects. Using only calibrated jets for $E_T^{\text{miss}}$ reconstruction yields a representation that is referred to as mht.
The \texttt{mht} algorithm calculates $E_{\text{T}}^{\text{miss}}$ from the negative transverse momentum vector sum of all jets above a threshold of 7 GeV before calibration. The HLT jets are reconstructed from calibrated topological clusters (defined in section 3.3) using the anti-$k_t$ jet algorithm with a radius parameter of $R = 0.4$ [22] implemented in the \texttt{FastJet} toolkit [23].

These jets are calibrated in a procedure similar to that used for offline physics analysis [21]. First, the estimated pile-up contribution to jets is removed using the jet-area-based pile-up suppression method [24, 25]. After pile-up subtraction, jets are calibrated using the simulation-based calibration described in ref. [26]. The energy deposits that arise from photons, electrons or hadronically decaying $\tau$-leptons, are included in the jet reconstruction.

3.5 Trigger implementing local pile-up suppression (\texttt{pufit})

The \texttt{pufit} algorithm corrects for pile-up effects on high-$E_{\text{T}}$ calorimeter signals contributing to $E_{\text{T}}^{\text{miss}}$. It employs a pile-up estimate obtained from a fit to lower-$E_{\text{T}}$ signals. It takes as inputs the topological clusters defined in section 3.3 and combines them into $\eta$-$\phi$ patches that correspond approximately to the size of a jet with $R = 0.4$. A fit is then performed which estimates the energy contribution to each patch from pile-up, based on the energy deposited and its spatial fluctuations across the calorimeter. Finally, the pile-up-subtracted patches are used to determine the $E_{\text{T}}^{\text{miss}}$.

The strategy is based on the assumption that high-$E_{\text{T}}$ energy deposits are associated with a hard-scatter collision of interest whereas the low-$E_{\text{T}}$ deposits are the result of pile-up. The \texttt{pufit} algorithm proceeds by performing a fit that constrains to zero (within fluctuations) the summed transverse momentum components $E_x$ and $E_y$ from the pile-up energy deposits. The $E_{\text{T}}^{\text{miss}}$ vector is then determined by summing the $E_x$ and $E_y$ of the high-$E_{\text{T}}$ deposits after subtracting the estimated pile-up contributions.

The \texttt{pufit} algorithm uses the measured structure of the energy deposition in each event. This contrasts with other approaches such as that of ref. [27] which estimate pile-up contributions by defining a median transverse energy density $\langle \rho \rangle$ that is then used in subtracting pile-up from high-$E_{\text{T}}$ deposits. The \texttt{pufit} algorithm is observed to outperform the standard pile-up-density algorithms in the context of the HLT, so these other algorithms are not described further. The full definition of the algorithm, and the event-by-event fit performed, can be found in appendix A.

4 Offline object and $E_{\text{T}}^{\text{miss}}$ reconstruction

When defining selections of events for which performance characteristics are desired, standard ATLAS offline algorithms are used to reconstruct and identify electrons, muons, $\tau$-leptons, jets and $b$-tagged jets, as described in appendix B.

The offline $E_{\text{T}}^{\text{miss}}$ is also computed using these reconstructed objects since they tend to have better resolution than individual tracks or clusters in the calorimeter. First, the contributions from high-$p_{\text{T}}$ electrons, photons, $\tau$-leptons and jets are summed, following the procedure described in ref. [16]. To account for the activity from the underlying event, tracks not associated with one of the above objects are also included in the $E_{\text{T}}^{\text{miss}}$ calculation. The $E_{\text{T}}^{\text{miss}}$ definition described above is referred to as ‘tight’ in the following.
In some cases, the so-called ‘tenacious’ offline $E_\text{miss}^\text{T}$ definition is used in order to make the jet selections less sensitive to pile-up. With this algorithm, jets that have $|\eta| > 2.4$ and $p_T < 35 \text{ GeV}$ are vetoed, along with jets with $p_T < 120 \text{ GeV}$ that fail the forward jet vertex tagger (JVT) requirement that utilizes jet correlations to reject pile-up jets in a region without a tracking detector [28]. The working point used corresponds to an efficiency of 92% for hard-scatter jets. Jets with $|\eta| < 2.4$ and $p_T$ within 20–40 GeV are used only if they satisfy a JVT requirement that yields an 85% efficiency for hard-scatter jets. Jets with $p_T$ within 40–60 GeV and 60–120 GeV are used only if they satisfy a similar requirement with an efficiency of 92% and 97% respectively.

For all purposes considered in this paper, the offline $E_\text{miss}^\text{T}$ is computed without any contribution to the visible momentum from any muon(s). This method of computing $E_\text{miss}^\text{T}$ facilitates comparison with the $E_\text{miss}^\text{T}$ trigger algorithms which use calorimeter information only.

5 $E_\text{T}^\text{miss}$ trigger performance

The figures of merit used to characterize the performance of the $E_\text{T}^\text{miss}$ trigger algorithm include: CPU time, trigger rate, efficiencies with respect to well-defined references, stability of the efficiencies for several different kinds of events, and the instantaneous luminosity dependence of these characteristics. Depending on the characteristics under study, the L1 and HLT algorithm performances, both individually and when used consecutively, are of interest. Good performance is characterized by a trigger which has stable high efficiency for signal events of interest, and, at the same time, a stable low output rate.

The trigger efficiency is defined by:

$$\varepsilon(S_i) = \frac{N(\text{trigger}|S_i)}{N(S_i)},$$

where $N(S_i)$ is the size of the sample of events satisfying some selection $S_i$ which is typically designed to isolate events within a narrow range of $E_\text{T}^\text{miss}$. To assess the efficiency of the trigger, $S_i$ is relaxed to capture events that satisfy some lower $E_\text{T}^\text{miss}$ threshold. In either case, the numerator $N(\text{trigger}|S_i)$ is the size of the subset of events that also satisfies the $E_\text{T}^\text{miss}$ trigger requirement.

5.1 Background model based on detector resolution

The $E_\text{T}^\text{miss}$ trigger rate behaviour in the absence of pile-up corrections can be studied with the cell $E_\text{T}^\text{miss}$ algorithm. This algorithm does not attempt to correct for the effects of pileup, other than via adjustments to the cell noise thresholds. A model has been constructed that captures the dependencies of the unbiased event acceptance (and hence trigger rate) of the cell $E_\text{T}^\text{miss}$ trigger algorithm on pile-up. The model is sufficient for the purpose of understanding the behaviour of the rate and demonstrates the need for more-sophisticated algorithms to deal with the large increase in pile-up through the period under study.
The $E_{T}^{\text{miss}}$ distribution is modelled with two components. The first is due to calorimeter energy resolution effects. This resolution is assumed to depend on the instantaneous number $\mu$ of $pp$ interactions per bunch crossing only through their combined contribution to the total calorimeter transverse energy $\Sigma E_{T}$, upon which the resolution in turn depends. The second component is the high $E_{T}^{\text{miss}}$ tail of the distribution, which is assumed to arise from events with rarer measurement fluctuations and events containing non-interacting particles (such as semileptonic decays of $b$- or $c$-hadrons). The probability of the second class of fluctuations is assumed to scale linearly with instantaneous luminosity. The two components are combined to form the vector sum of the two $E_{T}^{\text{miss}}$ values, with the azimuthal angle difference between the two components randomly oriented with respect to each other. By modelling the dependencies in this way, and by measuring the parameters of the model at low luminosity (and hence low pile-up), predictions of the $E_{T}^{\text{miss}}$ distribution and trigger rates can be obtained for higher $\mu$. The detailed description of the model may be found in appendix C.

To compare the calculation with measurements, data are selected by combining events obtained with several triggers, shown in table 1. For low $E_{T}^{\text{miss}}$ values, events obtained with an unbiased random trigger (zero bias) are used. The background events, which dominate the rate, were selected using a zero bias trigger, weighted to the instantaneous luminosity per bunch by requiring that an electron trigger fired in the previous LHC orbit of this bunch. Such triggers are prescaled, meaning that only one in $N$ events is accepted for some number $N$. Since the prescale factor $N$ for random triggers is high, there are not enough recorded events at high $E_{T}^{\text{miss}}$ for the study of the trigger background. These events are therefore supplemented with samples collected by a suite of triggers which require $L1 E_{T}^{\text{miss}}$ to be greater than a set of thresholds in the range 30 GeV to 50 GeV, as shown in table 1. The efficiencies for selecting events with higher $L1 E_{T}^{\text{miss}}$ thresholds (and lowest prescales) are found successively from those selected at lower thresholds (and correspondingly higher prescale), until those with the lowest $E_{T}^{\text{miss}}$ have their efficiencies determined using the random trigger.

Figure 1 compares the two-component model and its individual components with the full $E_{T}^{\text{miss}}$ distribution measured in data. When comparing data with the model, it is assumed that the instantaneous mean number $\mu$ of interactions per bunch crossing in the model is equal to its time-average $\langle \mu \rangle$ as measured over short periods in data. The data are also expected to have sensitivity to details that are not modelled, such as changes of calorimeter settings and the LHC bunch structure. The lower-luminosity data from earlier years of Run 2 were recorded under conditions different from those for the higher-luminosity data recorded in later years, giving differences of up to an order of magnitude in rates depending on threshold and luminosity. As is described in appendix C, the model pa-

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Random triggers</th>
<th>$L1 E_{T}^{\text{miss}} &gt; 30 \text{ GeV}$</th>
<th>$L1 E_{T}^{\text{miss}} &gt; 50 \text{ GeV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescale</td>
<td>$\mathcal{O}(10^6)$</td>
<td>$\mathcal{O}(10^5)$</td>
<td>$\mathcal{O}(10^3)$</td>
</tr>
</tbody>
</table>

Table 1. Triggers used for comparing the cell $E_{T}^{\text{miss}}$ acceptance model with data.
rameters were extracted from the full data set, and therefore are averaged over these effects. Nonetheless, as can be seen in figure 1, the model reproduces the key features of the data over the approximately nine orders of magnitude range of each distribution. Comparisons performed for values of average pile-up in the range $15 \lesssim \langle \mu \rangle \lesssim 60$ show that the model accounts for all qualitative features of the data in this range. Beyond these values it is found to somewhat underestimate (overestimate) the $E_T^{\text{miss}}$ tail for higher (lower) values of $\langle \mu \rangle$.

Three regions can be seen in figure 1. For low $E_T^{\text{miss}}$, the resolution term dominates, and the rate grows exponentially with increasing $\mu$. At high $E_T^{\text{miss}}$, the tail term dominates, and the rate is linear in $\mu$. Both of these terms contribute at intermediate $E_T^{\text{miss}}$ values. In this region there is a transition from exponential to linear behaviour with increasing $E_T^{\text{miss}}$ threshold. As $\mu$ increases, this transition region moves to higher values of $E_T^{\text{miss}}$. For a fixed $E_T^{\text{miss}}$ threshold trigger, the rate dependence on $\mu$ varies from linear to exponential with increasing $\mu$. The value of $\mu$ at which this transition occurs will vary according to the $E_T^{\text{miss}}$ threshold applied.

Figure 2 shows the prediction for the cell $E_T^{\text{miss}}$ algorithm pass-fraction at fixed threshold as a function of $\mu$. LHC Run-2 luminosities produced instantaneous $\mu$ as high as about 70, although the figure also shows extrapolated predictions up to $\mu = 200$. If the cell $E_T^{\text{miss}}$ algorithm had been the primary $E_T^{\text{miss}}$ trigger during Run 2, the threshold would have been raised considerably to keep the trigger rate within affordable limits. This increase in thresh-
Figure 2. The $E_{\text{T}}^{\text{miss}}$ model predicted trigger rate as a function of $\mu$ for the cell $E_{\text{T}}^{\text{miss}}$ algorithm with a threshold of 80 GeV and 120 GeV, assuming no additional pile-up mitigation.

Figure 3. (a) The L1 $E_{\text{T}}^{\text{miss}}$ trigger efficiency, shown as a function of $p_T(\mu\mu)$ in $Z \rightarrow \mu\mu$ events. The muons have little interaction with the calorimeter, so the transverse momentum $p_T(\mu\mu)$ of the dimuon system provides a good estimate of the $E_{\text{T}}^{\text{miss}}$ expected in the trigger calculations.

(b) The efficiencies in the plot are shown for events satisfying a $Z \rightarrow \mu\mu$ selection and with $p_T(\mu\mu)$ larger than 150 GeV vs pile-up for each of the four years of data taking. The uncertainties are statistical.

old would have significantly decreased the efficiency for signal events. Algorithms which better correct for pile-up were therefore introduced for Run 2 and used either in conjunction with or in place of the cell $E_{\text{T}}^{\text{miss}}$ algorithm.

5.2 Level-1 trigger performance

The efficiency of the L1 $E_{\text{T}}^{\text{miss}}$ trigger is determined using a $Z \rightarrow \mu\mu$ events. The muons have little interaction with the calorimeter, so the transverse momentum $p_T(\mu\mu)$ of the dimuon system provides a good estimate of the $E_{\text{T}}^{\text{miss}}$ expected in the trigger calculations.

To select events with two muons, a trigger requiring either two muon candidates each with $p_T > 14$ GeV, or an asymmetric threshold of 22 GeV for the leading muon and 8 GeV
for the sub-leading muon was used. The offline selection then requires that each of the two muons has $p_T > 25$ GeV, and that the dimuon invariant mass be in the range $66.6 \text{ GeV} < m(\mu\mu) < 116.6 \text{ GeV}$.

The efficiency is shown as a function of $p_T(\mu\mu)$ in figure 3a for an L1 nominal threshold of 50 GeV, at which the algorithm was generally run without prescaling. It can be observed that the algorithm achieves an efficiency of approximately 90% for a dimuon $p_T$ of 150 GeV.

The L1 $E_T^{\text{miss}}$ trigger efficiency for a $Z \rightarrow \mu\mu$ selection is shown as a function of the mean pile-up $\langle \mu \rangle$ for different years in figure 3b. A threshold of $p_T(\mu\mu) > 150$ GeV is used for the efficiency calculation since for $E_T^{\text{miss}}$ values in the range 150–175 GeV, the L1 trigger is sufficiently close to fully efficient to be interesting for many physics analyses. It is observed that the same efficiency was maintained to within a few percent as the pile-up increased.

Figure 4a shows the corresponding typical trigger rate as a function of the mean pile-up $\langle \mu \rangle$, which rises with increasing luminosity. Each of the three periods shown has its own set of values of the L1 calorimeter noise thresholds, which increase with increasing $\langle \mu \rangle$ as the period changes from $\alpha$ to $\beta$ to $\gamma$. The effect of the different noise thresholds used during the periods in 2017 (labelled $\alpha$, $\beta$ and $\gamma$), can be observed. As anticipated, higher calorimeter noise thresholds lead to much reduced trigger rates, particularly at higher $\langle \mu \rangle$.

The L1 $E_T^{\text{miss}}$ trigger efficiency for a $Z \rightarrow \mu\mu$ selection is shown as a function of $\langle \mu \rangle$ for the three periods with different noise thresholds during 2017 in figure 4b. Even though the calorimeter noise thresholds increase to moderate the trigger rate, the efficiency remains stable.

5.3 High-level trigger performance

The HLT background acceptance, which is proportional to the trigger rate, is defined as the fraction of events that have $E_T^{\text{miss}}$ computed by the HLT algorithm above a given threshold.
It is determined using events collected by a dedicated set of L1 triggers, unbiased by the HLT, as described in section 5.1.

The signal efficiency is determined by events collected using the $Z \rightarrow \mu\mu$ selection described in section 5.2. A subsample is selected with an additional requirement that the L1 trigger satisfy $E_{T}^{\text{miss}} > 50 \text{ GeV}$, in order to determine the efficiency of the HLT algorithms alone.

Curves of background rejection versus signal efficiency are obtained by varying the HLT trigger threshold. Figure 5 compares such curves for the four $E_{T}^{\text{miss}}$ algorithms defined in section 3, for different amounts of pile-up. For low pile-up ($\langle \mu \rangle < 20$) the efficiencies at which the $\text{tc}_{\text{lscw}}$ and the $\text{mht}$ $E_{T}^{\text{miss}}$ algorithms have equal-efficiency rejection power within a factor of three to that of $\text{pufit} E_{T}^{\text{miss}}$. The $\text{cell} E_{T}^{\text{miss}}$ algorithm has lower corresponding efficiency. As pile-up increases, $\text{tc}_{\text{lscw}}$ and $\text{mht}$ suffer the most degradation in their performance, whereas the $\text{pufit} E_{T}^{\text{miss}}$ trigger, which was designed to be robust against increasing pile-up, continues to simultaneously achieve good signal efficiency and large background rejection.

By combining different high-level triggers it was found to be possible to further improve the overall HLT performance. The simplest way to achieve this is by demanding that more than one $E_{T}^{\text{miss}}$ algorithm indicates that the event has high $E_{T}^{\text{miss}}$. The rationale for such a combination is as follows. The trigger rate of each algorithm for $E_{T}^{\text{miss}}$ greater than about 50 GeV is typically dominated by contributions from the resolution tails of poorly measured events which often contain little true $E_{T}^{\text{miss}}$. Since these tails depend on the details of the algorithm, populations of poorly reconstructed events in the high $E_{T}^{\text{miss}}$ tails differ between algorithms. By contrast, events with large true $E_{T}^{\text{miss}}$ caused by invisible particles tend to produce a large $E_{T}^{\text{miss}}$ with all algorithms. Therefore, requiring events to have large $E_{T}^{\text{miss}}$ in more than one algorithm, with appropriate thresholds for each, can result in reduced trigger rates for a similar overall efficiency.

The joint use of two $E_{T}^{\text{miss}}$ algorithms was found to be particularly useful when combining the $\text{pufit}$ and $\text{cell}$ algorithms. Figure 6 shows the relative signal acceptance and background rejection curves of the combined $\text{pufit-cell}$ algorithm compared with those of $\text{pufit}$ alone or $\text{cell}$ alone. With suitable thresholds, the combinations can have a higher rejection at the same efficiency than does either algorithm used alone.

The efficiencies of the $\text{cell}$, $\text{pufit}$ and combined $\text{pufit-cell}$ algorithms are shown as a function of $p_T(\mu\mu)$ and as a function of the offline $E_{T}^{\text{miss}}$ in figure 7. In order to have a fair comparison between the algorithms, each algorithm’s trigger threshold has been set such that their background rejections (and hence trigger acceptance rates) are equal. The combined $\text{pufit-cell}$ algorithm is again observed to have higher efficiency for signal events throughout the turn-on region than does either of the individual algorithms. The behaviour is consistent regardless of whether the efficiency is calculated as a function of $p_T(\mu\mu)$ or the offline $E_{T}^{\text{miss}}$ (with muons treated as invisible).

The improved acceptance for physics analyses that results from using the new algorithms can be considerable. Figure 7 shows that the new $\text{pufit-cell}$ trigger reaches its plateau efficiency at a $Z$ boson $p_T$ about 25 GeV below that of the $\text{cell}$ trigger operating at the same rate. This permits a correspondingly 25 GeV lower offline $E_{T}^{\text{miss}}$ threshold to
Figure 5. Background acceptance vs signal efficiency for each of four individual HLT $E_T^{\text{miss}}$ algorithms for a $Z \rightarrow \mu\mu$ selection with $p_T(\mu\mu) > 175$ GeV for data recorded in the year 2017. The diamond indicates the performance of the $p^{\text{fit}}$ $E_T^{\text{miss}} > 110$ GeV trigger. Each of the four lower panels shows a different range of $\langle \mu \rangle$: (a) $0 \leq \langle \mu \rangle < 20$, (b) $20 \leq \langle \mu \rangle < 30$, (c) $30 \leq \langle \mu \rangle < 40$ and (d) $40 \leq \langle \mu \rangle$.

be used for any analysis using an $E_T^{\text{miss}}$ trigger. This has a particularly important effect for physics channels in which the $E_T^{\text{miss}}$ distribution falls rapidly. For those analyses that select events on the $E_T^{\text{miss}}$ turn-on region where the trigger is not fully efficient, the $p^{\text{fit}}+\text{cell}$ algorithm recovers up to double the number of events of interest compared to cell alone. An example is for the search for Higgs to $b$-quarks [11] associated with the decay $Z \rightarrow \nu\nu$. For this analysis the acceptance decreases from 12% with $E_T^{\text{miss}} > 150$ GeV to only 5% with $E_T^{\text{miss}} > 200$ GeV [29]. If the offline threshold were to increase to 225 GeV the acceptance would have been only 3.5%.

To further examine the efficiency of the trigger algorithms with respect to the offline $E_T^{\text{miss}}$, the $E_T^{\text{miss}}$ trigger efficiency is calculated after applying either an additional offline $E_T^{\text{miss}} > 150 (175)$ GeV requirement or an offline $p_T(\mu\mu) > 150 (175)$ GeV requirement. Figure 8 (left) shows efficiencies for both the L1 trigger and the full (L1+HLT) trigger chain for data recorded at the end of 2018. The trigger efficiencies for a fixed $p_T(\mu\mu)$
Figure 6. Relative background acceptance fraction vs. relative efficiency for two different $p_T(\mu\mu)$ thresholds: (a) $p_T(\mu\mu) > 150 \text{ GeV}$ and (b) $p_T(\mu\mu) > 175 \text{ GeV}$ for data recorded in the year 2018. Two of the curves show the performance of the stand-alone cell algorithm and the stand-alone pufit algorithm. The other two show combined algorithms each formed by requiring that the event satisfy both a fixed threshold (either 65 GeV or 70 GeV as shown in the legend) for the cell algorithm and a pufit $E_T^{\text{miss}}$ threshold which varies along the curve. In each plot the background acceptance fractions and the efficiencies are relative to those of the pufit $E_T^{\text{miss}} > 110 \text{ GeV}$ trigger and thus can be greater than one. The diamond indicates the performance of the pufit $E_T^{\text{miss}} > 110 \text{ GeV}$ trigger while the cross indicates the performance of the combined (pufit $E_T^{\text{miss}} > 110 \text{ GeV}$ and cell $E_T^{\text{miss}} > 50 \text{ GeV}$) trigger.

Figure 7. Turn-on efficiency curves are shown for $Z \rightarrow \mu\mu$ events for three algorithms: the cell algorithm alone, the pufit algorithm alone and the combined cell+pufit algorithm. The thresholds are set such that the algorithms have equal rates, and the data were recorded in the year 2018. (a) The trigger efficiency with respect to $p_T(\mu\mu)$. (b) The trigger efficiency with respect to the offline $E_T^{\text{miss}}$ calculation with muons treated as being invisible.
Figure 8. Efficiencies for $Z \rightarrow \mu\mu$ events are shown for the L1 $E_T^{\text{miss}} > 50$ GeV trigger (square) and for the complete L1+HLT trigger chain (circle) that also requires $\text{pufit } E_T^{\text{miss}} > 110$ GeV. The uncertainties are statistical. Each is shown as a function of $(\mu_i)$, either for a $p_T(\mu\mu)$ threshold as in the left plots: (a) and (c) or for an offline $E_T^{\text{miss}}$ threshold as shown in the right plots: (b) and (d). The upper two plots (a) and (b) show thresholds of 150 GeV, while the lower two plots (c) and (d) correspond to thresholds of 175 GeV.

threshold show no significant decrease, even for the highest values of $(\mu)$. However, when compared with an offline $E_T^{\text{miss}}$ threshold in figure 8 (right), an apparent degradation of the trigger efficiency is observed for high $(\mu)$. This indicates that the difference between the online and offline $E_T^{\text{miss}}$ definitions has a larger effect at higher $(\mu)$.

5.4 Trigger menu evolution and performance

Due to the dependence of the algorithm efficiencies and trigger rates upon luminosity, it was necessary to update the primary physics triggers to cope with the increasing pile-up levels. Since the L1 rate was reduced by adjusting calorimeter noise thresholds, only small adjustments needed to be made to the overall L1 threshold. Table 2 summarizes the algorithms and trigger thresholds used during Run-2 data taking. In 2015–2016, the $\text{mht } E_T^{\text{miss}}$ was used. From 2016, the $\text{pufit } E_T^{\text{miss}}$ was combined with $\text{cell } E_T^{\text{miss}}$, thereby mitigating the effect of pile-up.
<table>
<thead>
<tr>
<th>Year</th>
<th>Trigger name</th>
<th>HLT algorithm</th>
<th>L1 threshold [GeV]</th>
<th>HLT threshold [GeV]</th>
<th>$\int \mathcal{L} , dt$ [fb$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>HLT_xe70_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>70</td>
<td>3.5</td>
</tr>
<tr>
<td>2016</td>
<td>HLT_xe90_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>90</td>
<td>12.7</td>
</tr>
<tr>
<td>2016</td>
<td>HLT_xe110_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>110</td>
<td>30.0</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe90_pufit_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>90,50</td>
<td>21.8</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe100_pufit_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>100,50</td>
<td>33.0</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe110_pufit_L1XE50(55)</td>
<td>pufit,cell</td>
<td>50(55)</td>
<td>110,50</td>
<td>47.7</td>
</tr>
<tr>
<td>2018</td>
<td>HLT_xe110_pufit_xe65_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>110,65</td>
<td>57.0</td>
</tr>
<tr>
<td>2018</td>
<td>HLT_xe110_pufit_xe70_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>110,70</td>
<td>62.6</td>
</tr>
</tbody>
</table>

**Table 2.** The evolution of the primary $E_T^{\text{miss}}$ physics triggers through the years of the LHC physics Run 2 from 2015 to 2018. For each year the table shows the algorithms used, the L1 and HLT thresholds applied and the integrated luminosity collected. Where two HLT thresholds are given, the first corresponds to the pufit algorithm and the second to the cell algorithm. In 2017, the pufit algorithm was used in conjunction with an additional requirement that cell $E_T^{\text{miss}} > 50$ GeV, which is not explicit in its name. The integrated luminosities are not exclusive and cannot be summed to obtain a total integrated luminosity.

![Figure 9](https://example.com/figure9.png)

**Figure 9.** High-level trigger output rates, as a function of $\langle \mu \rangle$, shown separately for example runs in each year 2015–2018, for triggers HLT_xe70_mht (2015), HLT_xe90_mht and HLT_xe110_mht (2016), HLT_xe110_pufit (2017), HLT_xe110_pufit_xe65 and HLT_xe110_pufit_xe70 (2018). The HLT_xe110_pufit trigger used during 2017 also included an implicit requirement of cell $E_T^{\text{miss}} > 50$ GeV.
The trigger names carry information about the algorithms and thresholds used. For example, L1\textsubscript{XE50}, denotes that the requirement is placed upon the first-level trigger (L1), that the requirement is on the value of $E_{\text{miss}}^T$ (XE), and provides the value of the L1 trigger threshold (50 GeV). The naming convention for full trigger paths can be parsed to give the trigger algorithms and their thresholds. For example, for the trigger path named HLT\_xe110\_pufit\_L1XE50, the prefix HLT indicates that the event must satisfy the high-level trigger requirement; xe110 indicates that the HLT threshold used was 110 GeV; pufit refers to the HLT algorithm used (except in the special case of the cell $E_{\text{miss}}^T$ algorithm, where the additional algorithm name is omitted), and L1XE50 refers to the L1 item used and its threshold.

Typical output rates for various HLT algorithms are shown year-by-year in figure 9. The reduction in rate obtained by using the pufit-based algorithms is a factor of ten or more for higher values of $\langle \mu \rangle$.

The overall (L1 +HLT) $E_T^{\text{miss}}$ trigger efficiency is shown year-by-year in figure 10. The efficiency is shown both as a function of $p_T(\mu \mu)$ and as a function of pile-up. The latter demonstrates that the efficiency remained stable within a few percent even at the highest pile-up values recorded.

5.5 Algorithm computation times

Average CPU times for the various steps used in the HLT $E_T^{\text{miss}}$ algorithms are given in table 3. For all algorithms except cell, the fraction of the computation time needed for evaluating the final $E_T^{\text{miss}}$ from previously determined input elements is negligible, and most of the CPU time is spent reconstructing cells and topological clusters. All steps satisfy the requirement described in section 3 that the CPU time does not exceed $\mathcal{O}(100 \text{ ms})$. 
### Table 3.
The average execution time of each step in computing $E_T^{\text{miss}}$ in the HLT online farm. The dot (•) indicates the required steps for each algorithm. The time to evaluate the $E_T^{\text{miss}}$ is shown as well, with the total time per step added to the evaluation time.

<table>
<thead>
<tr>
<th>Algorithm step</th>
<th>Time per step [ms]</th>
<th>Algorithm</th>
<th>tc</th>
<th>lcw</th>
<th>pufit</th>
<th>cell</th>
<th>mht</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calorimeter cell reconstruction</td>
<td>20</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>Topological cluster reconstruction</td>
<td>75</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Jet reconstruction</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ evaluation time [ms]</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>40</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Total time (ms)</td>
<td>95</td>
<td>95</td>
<td>60</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 5.6 Dependence on event characteristics

All of the previous efficiencies are computed using the clean reference sample selected with two muons in the final state. This sample is dominated by $Z \to \mu\mu$ events produced with additional jets. Because the detector response is not identical for events selected according to different criteria, the computation of $E_T^{\text{miss}}$ also depends on the event characteristics, for example whether jets or electrons are required to be present. In this section the trigger efficiency is evaluated and compared for a variety of offline event selections.

To complement the $Z \to \mu\mu$ events, four other selections are defined, as shown in table 4. The $t\bar{t}$ selections target the pair production of top quarks, and are particularly useful for examining performance in events with a large number of jets. These selections require that the event contain either (i) exactly one electron and no muons or (ii) exactly one muon and no electrons or (iii) exactly one electron and one muon. The vector boson fusion (VBF) selection targets events characterized by two energetic jets, where typically at least one jet is in the forward calorimeter. The selection requires two jets separated by a large pseudorapidity difference and not back-to-back in azimuth. The $W$ boson sample targets the $W \to e\nu$ process, and samples events with electromagnetic energy deposits that can be larger than 50 GeV.

<table>
<thead>
<tr>
<th>Physics process</th>
<th>Offline $E_T^{\text{miss}}$ definition</th>
<th>Lepton(s)</th>
<th>Kinematics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \to \mu\mu$</td>
<td>$p_T(\mu\mu)$</td>
<td>$\mu\mu$</td>
<td>$66.6 &lt; m(\mu\mu) &lt; 116.6 \text{GeV}$</td>
</tr>
<tr>
<td>$W \to e\nu$</td>
<td>Tight</td>
<td>$e$</td>
<td>—</td>
</tr>
<tr>
<td>VBF</td>
<td>Tight</td>
<td>$\mu, p_T &gt; 30 \text{GeV}$</td>
<td>Exactly two jets, $p_T &gt; 80 (50) \text{GeV}$, $</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>Tight</td>
<td>$e, \mu, e^\pm \mu^\mp$</td>
<td>$\geq 2$ $b$-tagged jets</td>
</tr>
</tbody>
</table>

Table 4. Definition of offline analysis selections used for efficiency measurements, labelled by the physics process being examined. All indicated lepton requirements implicitly require $p_T(\ell) > 25 \text{GeV}$ unless specified. The offline $E_T^{\text{miss}}$ definitions correspond to different working points. When multiple jets are required with different $p_T$ thresholds, the threshold for the subleading jet is listed in parentheses, e.g. 80 (50) GeV.
The W boson, VBF and t\bar{t} samples were each collected using a trigger that selects events containing a single isolated electron or muon. The required lepton transverse momentum thresholds were in the range 24–26 GeV, where the p_T value corresponds to the lowest threshold single lepton trigger available for a given luminosity.

Figure 11 shows the stability of the efficiencies with respect to pile-up after requiring that the offline E_T^{miss} be larger than 150 (175) GeV for these four different physics selections. In general, the efficiency for \langle \mu \rangle > 50 tends to be approximately 10–20% lower for events containing an electron rather than a muon. A difference in behaviour is not unexpected given that electrons are included in the calculation of the visible momentum, whereas muons are not. The offline E_T^{miss} calculation uses offline electrons that have better resolution compared to the trigger algorithm. It can also be seen that events containing forward jets in the VBF selections, or containing jets from top quark decays (right) have a somewhat different behaviour than do those events selected without jet requirements (left). Such variations

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure11.png}
\caption{Efficiencies for the first-level trigger L1XE50 and the combined L1+HLT trigger chain HLT_xe110_pufit_xe65_L1XE50 in data recorded in the year 2018 are shown as a function of \langle \mu \rangle for two different offline E_T^{miss} thresholds and four different physics selections: (a) W \rightarrow e\nu and Z \rightarrow \mu\mu selections with offline E_T^{miss} > 150 GeV (b) t\bar{t} and vector boson fusion selections with offline E_T^{miss} > 150 GeV (c) W \rightarrow e\nu and Z \rightarrow \mu\mu selections with offline E_T^{miss} > 175 GeV (d) t\bar{t} and vector boson fusion selections with offline E_T^{miss} > 175 GeV. The uncertainties are statistical.}
\end{figure}
are also to be expected, since the $E_{T}^{\text{miss}}$ resolution and scale change if any jets are present and depend on their energies and the region(s) of the calorimeter in which they are found.

5.7 Comparison with Monte Carlo simulation

It is important for most physics analyses to quantify the efficiency with which the trigger selects the events of interest. Depending on the details of the analysis, the $E_{T}^{\text{miss}}$ trigger efficiency may be determined from data alone, from a Monte Carlo simulation, or from a combination of the two. A concern with using Monte Carlo simulation to derive the $E_{T}^{\text{miss}}$ trigger efficiency is the effect of any residual difference between data and simulation. For example, as described in section 3.1, the noise thresholds of the L1 $E_{T}^{\text{miss}}$ trigger algorithm are adjusted periodically during data taking, but it is generally impractical to include such changes in the simulation. Because the $E_{T}^{\text{miss}}$ triggers use information from the full calorimeter, the efficiency determined using $E_{T}^{\text{miss}}$ triggers is more sensitive to changes in noise thresholds than the efficiency of other triggers.

Many ATLAS physics analyses render residual $E_{T}^{\text{miss}}$ trigger inefficiencies largely immaterial by requiring that the offline $E_{T}^{\text{miss}}$ be larger than 200 GeV. This requirement means that selected events are in a region in which the trigger efficiency is greater than 99% and therefore inefficiencies are negligible. However, some analyses, particularly those in which the number of events falls rapidly with increasing $E_{T}^{\text{miss}}$ (such as those in refs. [4, 8, 11]) motivate the use of $E_{T}^{\text{miss}}$ thresholds below the trigger plateau in order to maintain high signal efficiency. For these and similar cases the $E_{T}^{\text{miss}}$ trigger efficiency needs to be determined, often by using Monte Carlo simulations. Any differences between the simulation and the data may therefore lead to an incorrect calculation of the efficiency if the simulation alone were to be relied upon.

To account for residual differences between data and simulation, corrections referred to as scale factors are determined by measuring the ratio of the trigger efficiency using data to that expected from simulation. These are subsequently applied to correct the signal and background simulation. In the case of the $E_{T}^{\text{miss}}$ trigger the values of the scale factors vary with properties that include, e.g., the value of the trigger threshold, the cell noise thresholds, the definition of offline $E_{T}^{\text{miss}}$ and the details of the offline selection. Given that the $E_{T}^{\text{miss}}$ trigger is used for a large range of offline selections with widely varying final states, no single scale factor suitable for all cases can be found. Instead, analysis-specific corrections must be employed.

Comparisons between the trigger efficiency as predicted by Monte Carlo simulation and as measured in data were performed for both the L1 and the combined L1+HLT trigger chain, using data recorded during 2018. The trigger employed for those data is the combined $\text{pufit+cell}$ algorithm with thresholds as indicated in table 2. The variant of the offline $E_{T}^{\text{miss}}$ used is that referred to as ‘tenacious’, and is described in section 4. The efficiency is measured using events containing a single muon, which, as elsewhere, is treated as being invisible in the $E_{T}^{\text{miss}}$ calculation. The selection requirements are otherwise similar to the row labelled ‘VBF’ in table 4, except that: the requirements on the VBF jets are changed such that $|\Delta\phi(jj)| < 2.0$ and $m(jj) > 200 \text{ GeV}$. In addition, for figure 12a, exactly two jets are required and $|\Delta\eta(jj)| > 5.0$, while figure 12b is binned in jet multiplicity and $|\Delta\eta(jj)| > 3.5$. This selection is similar to that used in the VBF Higgs-to-invisible analysis [4].
Figure 12. Efficiencies from data and from simulation for L1 (L1XE50) and the combined L1+HLT chain (HLT_xe110_pufit_xe65_L1XE50) triggers are shown for a VBF selection which requires at least two jets that are well separated in rapidity. The uncertainties are statistical only. The lower panels show the ratios between efficiencies in data and simulation. (a) Efficiencies as a function of the offline $E_{T}^{\text{miss}}$. In this plot the jet rapidity difference requirement has been tightened to $\Delta y(jj) > 5$ (b) efficiencies as a function of $\langle \mu \rangle$ for an offline requirement of $E_{T}^{\text{miss}} > 150$ GeV using the ‘tenacious’ working point; the efficiencies in data and MC simulation are compared for two different selections, one requiring exactly two and the other $\geq 3$ jets.

Monte Carlo simulated samples of the $Z \rightarrow \nu \nu$, $W \rightarrow \ell \nu$ and $Z \rightarrow \ell \ell$ processes were generated at next-to-leading order (NLO) in strong coupling constant $\alpha_s$ using SHERPA 2.2.1 [30]. These calculations use the Comix [31] and OpenLoops [32] matrix element generators, and merging was done with the SHERPA parton shower [33] using the ME+PS@NLO prescription [34]. The NNPDF3.0 parton distribution function (PDF) set [35] computed at next-to-next-to-leading order in $\alpha_s$ was used, along with dedicated parton shower tuning parameters developed for SHERPA 2.2.1 [36]. After the events were generated, the response of the detector [37] is simulated using GEANT4 [38]. Each $W$ or $Z$ boson event was overlaid with 15–70 pile-up collisions to match the distribution in data. The pile-up events were simulated using PYTHIA 8.1 [39] with the MSTW2008 PDF set [35] and A3 set of parameters tuned to data [40, 41]. The efficiency as a function of offline $E_{T}^{\text{miss}}$ is shown in figure 12a. In the trigger turn-on region at lower values of offline $E_{T}^{\text{miss}}$, a difference can be observed between data and simulation, an effect largely attributable to the L1 trigger. The lower panel shows the ratio of data-determined to simulated efficiency as a function of the offline $E_{T}^{\text{miss}}$. This ratio is an example of the scale factor that can be applied to simulated Monte Carlo events to correct their $E_{T}^{\text{miss}}$ trigger efficiency. Figure 12b shows the efficiency with respect to $\langle \mu \rangle$ for a selection that requires offline $E_{T}^{\text{miss}} > 150$ GeV and either exactly 2 or $\geq 3$ jets. It can be seen that the Monte Carlo simulations overestimate the efficiency by a few percent. Such considerations show the need both to correct for the differences between data and simulation when working in the turn-on region, and to understand the behaviour of the resulting scale factors for appropriate selections.
6 Conclusion

Despite the considerable increase in luminosity during Run 2 of the LHC (2015–2018), it was possible to maintain the excellent performance of the ATLAS $E_T^{\text{miss}}$ trigger. This was achieved through a dedicated programme of developing, testing, evaluating and optimizing various pile-up mitigation algorithms.

For triggers without any pile-up correction, a steep increase in trigger rate with pile-up is observed. This behaviour is consistent with expectations from a two-component background $E_T^{\text{miss}}$ distribution model.

Several $E_T^{\text{miss}}$ trigger algorithms were introduced in ATLAS during LHC Run 2. Both the first-level and high-level trigger algorithms were improved to maintain a similar level of efficiency throughout the data-taking period. These included a new high-level trigger algorithm which uses a fit to determine pile-up-induced local energy deposits in individual events to reduce the impact of increasing luminosity on the $E_T^{\text{miss}}$ trigger rate. In addition, it was found that combining algorithms related to different sources of high-$E_T^{\text{miss}}$ tails could, with an appropriate choice of thresholds, help maintain high efficiency while keeping trigger rates under control. The new algorithms enabled the offline $E_T^{\text{miss}}$ selection to become efficient about 25 GeV earlier than would have previously been possible. In the trigger turn-on region the new algorithms recovered up to double the number of events that would have been possible previously.

A study of the $E_T^{\text{miss}}$ trigger performance for different signal samples shows only a small degradation of the efficiency despite the factor of four increase in instantaneous luminosity during the four-year LHC Run-2 period. The $E_T^{\text{miss}}$ trigger behaviour agrees in general with predictions from Monte Carlo simulations. However, the low-$E_T^{\text{miss}}$ region, where the trigger is not fully efficient, is more difficult to model precisely. Since different $E_T^{\text{miss}}$ algorithms respond differently depending on the characteristics of the event, it is necessary for analyses working in this low-$E_T^{\text{miss}}$ region to determine the specific corrections appropriate to their particular event selection.

The predictions from the background model extend up to $\mu = 200$, the value anticipated for the HL-LHC [29]. The predicted rates are those that would be obtained if no changes were made to the algorithm to mitigate pile-up. The equal-threshold acceptance fractions predicted for this level of pile-up are about two orders of magnitude higher than in Run 2, making that environment even more challenging for the $E_T^{\text{miss}}$ trigger. The upgrades planned for ATLAS from 2021 to 2025 will enable a factor of ten increase in trigger rate and a first-level trigger that takes advantage of the full calorimeter granularity. As found in the results presented in this paper, mitigation of pile-up is possible, by use of specifically-designed algorithms which can achieve lower rates for the same signal efficiency. Use of tracking information can also mitigate the pile-up effect, as can increasing the cell noise thresholds. However, from the rate predictions it is clear that further development of the high-level $E_T^{\text{miss}}$ trigger algorithms will also be required.
Acknowledgments

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

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

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [50].

A Full definition of the trigger implementing local pile-up suppression

The details of the \texttt{pufit} algorithm, introduced in section 3.5, are as follows.

To estimate the contribution from high-$E_T$ energy deposits, the calorimeter is divided into equal area ‘patches’. Using 8 divisions in $\phi$ and 14 in $\eta$ yields a total of 112 patches, each with an area similar to that of an $R = 0.4$ jet. The energy for each patch is computed from the sum of all the clusters that fall within that patch. To prevent a jet from being split into two low-$E_T$ patches, four sets of patches are constructed in parallel, with each set shifted by half of a patch size along $\eta$ and/or $\phi$. From the four candidate sets of patches, the algorithm selects the set that yields the largest scalar sum of $E_T$ from the high-$E_T$ patches. The transverse energy in each patch is then used to select the hard-scatter and
pile-up patches. Finally, a fit is performed to determine the pile-up contribution to each hard-scatter patch.

The fit is constrained by two assumptions: first that the vector sum of $\vec{E}_{\text{miss}}^T$ over all pile-up contributions should be zero, and second that the pile-up should be relatively evenly distributed throughout the detector.

**Stage 1: selection of hard-scatter patches.** The hard-scatter patch threshold is determined for each event from the trimmed mean and variance of the distribution of the patch $E_T$ values in the event. The patch $E_T$ values are sorted in ascending order and a trimmed mean is determined by discarding the 5% of patches with the lowest $E_T$ and the 5% of patches with the highest $E_T$. The trimmed mean is given by:

$$\langle E_T \rangle_{\text{patch}} = \frac{1}{N} \sum_{j=0.05N}^{0.95N} E_T^j,$$

where $N$ is the total number of patches. The variance of the patch $E_T$ values is estimated using the sample variance, calculated as

$$V_{\text{patch}} = \frac{1}{N} \left\{ \sum_{j=1}^{0.95N} (E_T^j - \langle E_T \rangle_{\text{patch}})^2 + \sum_{j=1}^{0.05N} (E_T^j - \langle E_T \rangle_{\text{patch}})^2 \right\},$$

where the second term uses the lowest 5% of patches (already included in the first term) to estimate the contribution to the variance from pile-up in the highest 5%. The highest 5% are excluded because the sample is biased by hard-scatter jets. The patch threshold is then set to

$$E_{T, \text{thres}} = \langle E_T \rangle_{\text{patch}} + n_{\sigma} \sqrt{V_{\text{patch}}},$$

with $n_{\sigma}$ chosen to be 5, based on rate considerations. The typical value for $E_{T, \text{thres}}$ after the L1 requirement is 30 GeV.

With the threshold determined, patches are categorized as high-$E_T$ or low-$E_T$ if they have transverse energy above or below $E_{T, \text{thres}}$, respectively. Any event that has no patch above threshold has $\text{pufit } E_{\text{miss}}^T$ set to zero and is rejected by the trigger. Approximately 97% of background events that pass the $\text{L1XE50}$ trigger have at least one hard-scatter patch.

**Stage 2: performing the fit.** The fit of the pile-up contribution to each of the high-$E_T$ patches is performed by minimizing the $\chi^2$ function

$$\chi^2(E_{T1}, \ldots, E_{Tm}) = \Delta^T V^{-1} \Delta.$$

In this expression $V$ is the associated covariance matrix, and the variables $E_{Tk}$ are the pile-up contributions to the transverse energies of the $m$ high-$E_T$ patches. The values of $E_{Tk}$ are determined by minimizing this $\chi^2$ function. The $(m + 2)$-dimensional vector $\Delta$ is
given by

\[ \Delta = \begin{pmatrix}
\sum_{i=1}^{N_{\text{low}}} E_{x_i} + \sum_{k=1}^{m} \mathcal{E}_{T_k} \cos \phi_k \\
\sum_{i=1}^{N_{\text{low}}} E_{y_i} + \sum_{k=1}^{m} \mathcal{E}_{T_k} \sin \phi_k \\
(A_1/A_{\text{low}} \times \sum_{i=1}^{N_{\text{low}}} \mathcal{E}_{T_i}) - \mathcal{E}_{T_1} \\
: \\
(A_m/A_{\text{low}} \times \sum_{i=1}^{N_{\text{low}}} \mathcal{E}_{T_i}) - \mathcal{E}_{T_m}
\end{pmatrix} \]

where the \(A_k\) are the areas of each of the \(m\) high-\(E_T\) patches, and \(A_{\text{low}}\) is the summed area of the low-\(E_T\) patches. The first two lines impose the transverse momentum conservation constraints \(\sum E_x = 0\) and \(\sum E_y = 0\) on the pile-up contributions. On each of those two lines, the first sum (for which \(N_{\text{low}} = N - m\)) runs over the low-\(E_T\) patches, and the second over the estimated contributions of pile-up in the high-\(E_T\) patches. The quantities \(\cos \phi_k\) and \(\sin \phi_k\) are given by \(E_{x_k}/E_{T_k}\) and \(E_{y_k}/E_{T_k}\), respectively. Each of the remaining \(m\) lines corresponds to one of the high-\(E_T\) patches, and penalizes any difference between the to-be-fitted pile-up estimate in that patch \(\mathcal{E}_{T_k}\) and the amount of pile-up transverse energy that would be expected in it, based on the event-wide average transverse energy density (as calculated from the low-\(E_T\) patches). The event-wide transverse energy is given by the term \(A_k/A_{\text{low}} \times \sum_{i=1}^{N_{\text{low}}} \mathcal{E}_{T_i}\), where the ratio \((\sum_{i=1}^{N_{\text{low}}} \mathcal{E}_{T_i})/A_{\text{low}}\) is the average energy density from the low-\(E_T\) patches.

For the determination of the covariance matrix, the fluctuations of all of the low-\(E_T\) patches are calculated from the contribution to the calorimeter energies due to detector resolution. The uncertainty in each measured patch transverse energy \(\mathcal{E}_{T_i}\) is taken as

\[ \sigma_{E_T}^2 = r_0^2 + r^2 \mathcal{E}_{T_i}, \]

where \(r\) is the resolution scale, set to 0.5 GeV\(^{1/2}\) and determined from calorimeter energy resolutions. The \(r_0\) term is a resolution floor (0.05 GeV), introduced to avoid numerical problems for events where most of the patches have no deposited \(E_T\). The patch sample variance \(V_{\text{patch}}\) calculated above is used for the lower \(m\) diagonal elements. The full covariance matrix is

\[ V = \begin{pmatrix}
V_{11} & V_{12} & 0 & 0 & \ldots & 0 \\
V_{12} & V_{22} & 0 & 0 & \ldots & 0 \\
0 & 0 & sV_{\text{patch}} & 0 & \ldots & 0 \\
0 & 0 & 0 & \ddots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \ldots & sV_{\text{patch}}
\end{pmatrix} \]
where the upper $2 \times 2$ block is given by

$$V_{\text{cov}} = \begin{pmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{pmatrix} = \begin{pmatrix} \sum_{i}^{N_{\text{low}}} \sigma_i^2 \cos^2 \phi_i & \sum_{i}^{N_{\text{low}}} \sigma_i^2 \cos \phi_i \sin \phi_i \\ \sum_{i}^{N_{\text{low}}} \sigma_i^2 \cos \phi_i \sin \phi_i & \sum_{i}^{N_{\text{low}}} \sigma_i^2 \sin^2 \phi_i \end{pmatrix},$$

and $s = 1$ is determined by optimizing the fit.

The minimization of the $\chi^2$ function with respect to the $E_{T_k}$ is performed analytically, avoiding computationally expensive numerical minimization. The solution involves solving an $m \times m$ linear system, where $m$, the number of above-threshold patches, is typically 4 or 5.

Minimizing the function gives

$$0 = \frac{1}{2} \frac{\partial \chi^2}{\partial E_{T_k}} = (\cos \phi_k, \sin \phi_k) V_{\text{cov}}^{-1} \left( \sum_{i=1}^{N_{\text{low}}} E_{x_i} + \sum_{j=1}^{m} E_{T_j} \cos \phi_j \right) \frac{1}{s V_{\text{patch}}} \left( \frac{A_k}{A_{\text{low}}} \sum_{i=1}^{N_{\text{low}}} E_{x_i} - E_{T_k} \right) + \frac{\delta_{ik}}{s V_{\text{patch}}}.$$

The solution to this equation is given by the matrix equation

$$E_{T_k} = [X^{-1}]_{ki} c_i,$$

where the $m \times m$ matrix

$$X_{ik} = (\cos \phi_i, \sin \phi_i) V_{\text{cov}}^{-1} \left( \frac{\cos \phi_k}{\sin \phi_k} + \frac{\delta_{ik}}{s V_{\text{patch}}} \right)$$

and

$$c_i = \frac{A_i}{A_{\text{low}}} \frac{\sum_{j=1}^{N_{\text{low}}} E_{T_j}}{s V_{\text{patch}}} - (\cos \phi_i, \sin \phi_i) V_{\text{cov}}^{-1} \left( \frac{\sum_{j=1}^{N_{\text{low}}} E_{x_j}}{\sum_{j=1}^{N_{\text{low}}} E_{y_j}} \right).$$

The $E^\text{miss}_x$ for the event is then calculated from the $m$ high-$E_T$ patches and the fit results for the pile-up contribution to the energy in these patches, where now

$$E_{T_k}^\text{miss} = - \sum_{k=1}^{m} (E_{x_k} - E_{T_k} \cos \phi_k),$$

$$E_{T_k}^\text{miss} = - \sum_{k=1}^{m} (E_{y_k} - E_{T_k} \sin \phi_k),$$

and finally $E_T^\text{miss} = \sqrt{(E_x^\text{miss})^2 + (E_y^\text{miss})^2}$.

B Details of the offline reconstruction algorithms

In the offline analyses, tracks with $p_T > 400$ MeV are reconstructed to identify a common vertex [42] and an event is required to have at least one such vertex. If more than one vertex is found in the event, the primary vertex is defined as the one with the largest $\sum_{i} p_{T,i}^2$, where the sum runs over all tracks associated with the vertex.
Electrons are reconstructed by matching tracks that originate from the primary vertex with clusters of energy deposits in the calorimeter \cite{43}. They are required to have $p_T > 25$ GeV and $|\eta| < 2.47$. To suppress electrons that originate from the hadronic showers of jets, the electron is also required to be isolated using information from the calorimeter and tracks within a cone of size $\Delta R = 0.2$ around the electron direction. The total efficiency of electron reconstruction and isolation is 85\% for an electron with $p_T \approx 30$ GeV.

Muons are reconstructed by matching tracks from the muon spectrometer with inner-detector tracks that are associated with the primary vertex \cite{44}. Muons are required to have $p_T > 25$ GeV and $|\eta| < 2.47$, and they are required to be isolated. For the isolation working point used, the efficiency for a muon with $p_T$ in the range 20–100 GeV is 96\%.

Jets are reconstructed starting with information from the calorimeter. Cells are grouped into topological clusters \cite{13}, described in section 3.3, which in turn are used to reconstruct jets using the anti-$k_t$ algorithm with a radius parameter of $R = 0.4$. Jets are required to have $p_T > 20$ GeV and $|\eta| < 4.5$. To suppress jets that come from pile-up, those that have $|\eta| < 2.4$ are required to exceed the jet vertex tagger (JVT) threshold \cite{25}. The JVT uses tracks and vertices to assign a likelihood for a jet to be associated with the primary vertex. Jets that fail the JVT requirement are removed, as they are likely to be due to pile-up. The efficiency to select a jet from a signal process is 92\%, while jets originating from pile-up collisions are rejected 99\% of the time in the $p_T$ range 20–50 GeV. Each jet with $|\eta| < 2.5$ is evaluated for the likelihood that it originated from a $b$-hadron \cite{45}. The efficiency for tagging a jet containing a $b$-hadron is 77\% for jets originating from $t\bar{t}$ events \cite{46}.

When selecting events based on the number of jets and leptons (electrons or muons), it is necessary to resolve the ambiguity for cases where a reconstructed jet and a lepton both result from the same detector signals. To remove such ambiguities, a sequential overlap removal procedure is defined. If the jet is within a cone of size $\Delta R = 0.2$ around any electron, the jet is removed. If a jet is found within a conical annulus defined by $0.2 < \Delta R < 0.4$, then the electron is removed. If there is a muon within $\Delta R = 0.4$ of a jet, and the jet has at least three charged tracks with $p_T > 500$ MeV, then the muon is removed; otherwise the jet is removed.

C The cell $E_T^{\text{miss}}$ background distribution model

The details of the background cell $E_T^{\text{miss}}$ distribution model discussed in section 5.1 are as follows. The model has two components. The first covers the bulk of the distribution, which is assumed to originate from resolution effects. The second component, which dominates in the tail, comes from rare occurrences which scale linearly with the luminosity, and hence, for fixed beam parameters, linearly with the instantaneous average number $\mu$ of interactions per bunch crossing.

**Component 1: $\Sigma E_T$ resolution-dependent bulk distribution.** The bulk resolution effects are assumed to depend on the total transverse energy deposited in the calorimeter, $\Sigma E_T$. It is assumed that the two perpendicular components of $E_T^{\text{miss}}$ are independently identically Gaussian distributed with a width $\sigma(\Sigma E_T)$ that depends on $\Sigma E_T$, such that
Figure 13. Data and fits to data for $E_T^{\text{miss}}$ model. (a) A comparison of the measured cell-algorithm $E_T^{\text{miss}}$ distribution with a Rayleigh distribution fit, for a narrow range of pile-up centred on $\langle \mu \rangle \approx 50$. The black histogram is the distribution from the unbiased data. The red curve is the result of a fit to a Rayleigh distribution. (b) Values of $\sigma$ are shown from Rayleigh fits to $E_T^{\text{miss}}$ distributions for 6 values of $\langle \mu \rangle$ and 50 values of $\sqrt{\Sigma E_T}$ per $\langle \mu \rangle$ value, as well as the data from which the fits are obtained. Also shown are straight-line fits to the 50 $\sqrt{\Sigma E_T}$ bins for each $\langle \mu \rangle$.

events with a given $\Sigma E_T$ will obey a Rayleigh distribution in $E_T^{\text{miss}}$,

$$R( E_T^{\text{miss}} | \sigma(\Sigma E_T) ) = (E_T^{\text{miss}} / \sigma^2) \exp(-[E_T^{\text{miss}}]^2/2\sigma^2).$$

In turn the probability density function of $\Sigma E_T$ is assumed to be a function $P(\Sigma E_T|\mu, \gamma)$, where $\mu$ is the instantaneous average number of interactions per bunch crossing and $\gamma$ is used to parameterize any other dependencies. For a given $\mu$, $E_T^{\text{miss}}$ is distributed according to the probability density function

$$P_{\text{res}}(E_T^{\text{miss}}|\mu) = \int_0^\infty R(E_T^{\text{miss}}|\sigma(\Sigma E_T)) P(\Sigma E_T|\mu, \gamma) \, d(\Sigma E_T). \quad (C.1)$$

In order to sample from the distribution $P_{\text{res}}(E_T^{\text{miss}}|\mu)$, expressions for both $\sigma(\Sigma E_T)$ and $P(\Sigma E_T|\mu, \gamma)$ are needed.

Fits to find $\sigma(\Sigma E_T)$ are discussed first. Figure 13a shows an example cell $E_T^{\text{miss}}$ distribution for a narrow range around $\langle \mu \rangle \approx 50$ and $\Sigma E_T \approx 400$ GeV. The plot also shows a Rayleigh fit to the $E_T^{\text{miss}}$ distribution; the fit characterizes it well except at the highest $E_T^{\text{miss}}$ values. The points in figure 13b show the values of $\sigma$ obtained from fits to the $E_T^{\text{miss}}$ distribution for 6 values of $\langle \mu \rangle$ and 50 bins of $(\Sigma E_T)^{1/2}$ for each $\langle \mu \rangle$. Although some $\langle \mu \rangle$ dependence is visible, the model provides a reasonable description of the data using a parameterization of the resolution given by $\sigma(\Sigma E_T) = \alpha_1 + \alpha_2 \sqrt{\Sigma E_T}$ with $\alpha_1 = 3$ GeV and $\alpha_2 = 0.465 \text{GeV}^{1/2}$. Plots like those in figure 13b show that this linear fit is a good model for the bulk of events and performs poorly only in ranges of the total transverse energy where the Rayleigh fits do not provide a good description of the $E_T^{\text{miss}}$.

To find the $\Sigma E_T$ probability density function $P(\Sigma E_T|\mu, \gamma)$, one may first consider the distribution $S_1(\Sigma E_T|\gamma)$ of the $\Sigma E_T$ for a single proton-proton interaction. The corresponding distribution $S_n(\Sigma E_T|\gamma)$ for $n$ simultaneous proton-proton interactions is given by the
$n$-fold convolution of $S_1(\Sigma E_T|\vec{\gamma})$. The values of $n$ that are sampled for any $\mu$ are assumed to be Poisson distributed with mean $\mu$, so that the resulting distribution $P(\Sigma E_T|\mu, \vec{\gamma})$ is given by a Poisson-weighted sum of terms, each term being the product of the Poisson probability for $n$ interactions given $\mu$ with the corresponding $S_n(\Sigma E_T|\vec{\gamma})$:

$$P(\Sigma E_T|\mu, \vec{\gamma}) = \sum_{n=1}^{\infty} \text{Poi}(n|\mu)S_n(\Sigma E_T|\vec{\gamma}) .$$

In the absence of data taken at low $\langle \mu \rangle$ under the same beam conditions, $S_1(\Sigma E_T|\vec{\gamma})$ cannot be determined directly. Instead, various models of $S_1(\Sigma E_T|\vec{\gamma})$ are explored by varying their parameters until the calculation gives a reasonable match to the measured high-luminosity $P(\Sigma E_T|\mu, \vec{\gamma})$ distributions. It is found that a single-interaction $\Sigma E_T$ distribution of the form

$$S_1(\Sigma E_T|\vec{\gamma}) = \gamma_0 \exp(-\gamma_1 \Sigma') + (1 - \gamma_3) \gamma_2 \exp(-\gamma_2 \Sigma')$$

gives a result that matches the data well for $\mu$ of about 20 or greater. The parameters of this expression are $\gamma_1 = 0.070$ GeV$^{-1}$, $\gamma_2 = 0.009$ GeV$^{-1}$, $\gamma_3 = 0.995$, and for the $\mu$-dependent substitution $\Sigma' = \Sigma E_T - (\gamma_4 + \gamma_5 \times \mu)$ with $\gamma_4 = 25$ GeV and $\gamma_5 = -5$ GeV. Figure 14 shows an example of the $\Sigma E_T$ modelling for $\mu = 50$.

The $n$-fold convolution can be carried out analytically only for a limited number of functions. For example, the single-interaction $\Sigma E_T$ distribution for the LHC Run 1 was approximated by a single exponential function, and the convolution was then performed as described in ref. [47]. For the more complicated functions used in the present calculation, the convolution and Poisson sum is computed using a fast Fourier transform and the method described in refs. [48] and [49].

The Fourier transform $\hat{P}(\omega|\mu, \vec{\gamma})$ of the weighted convolution sum $P(\Sigma E_T|\mu, \vec{\gamma})$ is rewritten as the infinite sum of Fourier transforms over $S_1(\Sigma E_T|\vec{\gamma})$, denoted by $\hat{S}_1(\omega|\vec{\gamma})$. This facilitates calculation, as the infinite sum can be performed using the Fourier transform:

$$\hat{P}(\omega|\mu, \vec{\gamma}) = \sum_{n=1}^{\infty} \text{Poi}(n|\mu)\hat{S}_n(\omega|\vec{\gamma})$$

$$= \sum_{n=1}^{\infty} \text{Poi}(n|\mu)(2\pi)^{n-1}(\hat{S}_1(\omega|\vec{\gamma}))^n$$

$$= \frac{1}{2\pi} e^{-\mu} \sum_{n=1}^{\infty} (2\pi)^n \mu^n (\hat{S}_1(\omega|\vec{\gamma}))^n / n!$$

$$= \frac{1}{2\pi} e^{-\mu} \left( \sum_{n=0}^{\infty} [2\pi \mu \hat{S}_1(\omega|\vec{\gamma})]^n / n! - 1 \right)$$

$$= \frac{1}{2\pi} e^{-\mu} \left( \exp \left( 2\pi \mu \hat{S}_1(\omega|\vec{\gamma}) \right) - 1 \right) .$$

(C.2)

In principle, one could also take a logarithm of the exponential term in the final expression above to directly solve for $\hat{S}_1(\omega|\vec{\gamma})$ from $\hat{P}(\omega|\mu, \vec{\gamma})$. However, the uncertainties in the $P(\Sigma E_T|\mu, \vec{\gamma})$ distribution are amplified in this process, making it impossible in practice to obtain useful results in this way.
Component 2: luminosity-dependent tail from rare events. The Rayleigh distribution resolution model described above is found to describe the observed cell $E_T^{\text{miss}}$ distribution well for $E_T^{\text{miss}}$ up to about 50 GeV, but not beyond. A second component, which contributes more at larger $E_T^{\text{miss}}$ values, is assumed to come from a combination of events with rare calorimeter measurement fluctuations, for example very poor measurement of a single jet in a two-jet event, and from events containing non-interacting particles, such as those in which semileptonic $b$-hadron decays produce neutrinos. The probability of such an event is expected to scale linearly with luminosity, and hence for fixed bunch-crossing rate, to scale linearly with $\mu$. Indeed, a Fréchet function that is linearly dependent on $\mu$,

$$
\mu F(E_T^{\text{miss}}|\mu, \alpha, s) = \mu A \frac{\alpha}{s} \left( \frac{E_T^{\text{miss}} - m}{s} \right)^{-1-\alpha} \exp \left[ - \left( \frac{E_T^{\text{miss}} - m}{s} \right)^{-\alpha} \right]
$$

is found to describe the high $E_T^{\text{miss}}$ distribution, with parameters $\alpha = 3.54$, $s = 8$ GeV, and $m = 40$ GeV, and with the probability per single $pp$ interaction of such a tail event being $A = 10^{-5}$.

Full $E_T^{\text{miss}}$ distribution. The complete model $E_T^{\text{miss}}$ distribution is then obtained by vector addition of the two $E_T^{\text{miss}}$ sources. A random azimuthal angle $\phi_{\text{rand}}$ is assumed between the two vectors and the magnitude of the $E_T^{\text{miss}}$ is sampled from two sources. These are (i) the $E_T^{\text{miss}}$ distribution in eq. (C.1) from the resolution function of the total transverse energy in eq. (C.2), and (ii) a tail modelled with the Fréchet distribution, depending linearly upon $\mu$, in eq. (C.3). The combination yields

$$
\text{Event } E_T^{\text{miss}} = \left| E_T^{\text{miss}}(\text{resolution}) \hat{n}_T(\phi = 0) + E_T^{\text{miss}}(\text{tail}) \hat{n}_T(\phi = \phi_{\text{rand}}) \right|.
$$

It is the distribution of this quantity that is compared with data in section 5.1.

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


The ATLAS collaboration

Tomsk State University, Tomsk; Russia
Department of Physics, University of Toronto, Toronto ON; Canada
\((a)\) TRIUMF, Vancouver BC;\((b)\) Department of Physics and Astronomy, York University, Toronto ON; Canada
Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba; Japan
Department of Physics and Astronomy, Tufts University, Medford MA; United States of America
Department of Physics and Astronomy, University of California Irvine, Irvine CA; United States of America
Department of Physics and Astronomy, University of Uppsala, Uppsala; Sweden
Department of Physics, University of Illinois, Urbana IL; United States of America
Istituto di Fisica Corpuscolare (IFIC), Centro Mixto Universidad de Valencia — CSIC, Valencia; Spain
Department of Physics, University of British Columbia, Vancouver BC; Canada
Department of Physics and Astronomy, University of Victoria, Victoria BC; Canada
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg; Germany
Department of Physics, University of Warwick, Coventry; United Kingdom
Waseda University, Tokyo; Japan
Department of Particle Physics, Weizmann Institute of Science, Rehovot; Israel
Department of Physics, University of Wisconsin, Madison WI; United States of America
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal; Germany
Department of Physics, Yale University, New Haven CT; United States of America
\(a\) Also at Borough of Manhattan Community College, City University of New York, New York NY; United States of America
\(b\) Also at Centro Studi e Ricerche Enrico Fermi; Italy
\(c\) Also at CERN, Geneva; Switzerland
\(d\) Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille; France
\(e\) Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
\(f\) Also at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona; Spain
\(g\) Also at Department of Applied Physics and Astronomy, University of Sharjah, Sharjah; United Arab Emirates
\(h\) Also at Department of Financial and Management Engineering, University of the Aegean, Chios; Greece
\(i\) Also at Department of Physics and Astronomy, Michigan State University, East Lansing MI; United States of America
\(j\) Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY; United States of America
\(k\) Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva; Israel
\(l\) Also at Department of Physics, California State University, East Bay; United States of America
\(m\) Also at Department of Physics, California State University, Fresno; United States of America
\(n\) Also at Department of Physics, California State University, Sacramento; United States of America
\(o\) Also at Department of Physics, King’s College London, London; United Kingdom
\(p\) Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg; Russia
\(q\) Also at Department of Physics, University of Adelaide, Adelaide; Australia
\(r\) Also at Department of Physics, University of Fribourg, Fribourg; Switzerland
\(s\) Also at Dipartimento di Matematica, Informatica e Fisica, Università di Udine, Udine; Italy
\(t\) Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow; Russia