



**ATLAS PUB Note**  
ATL-PHYS-PUB-2021-023  
July 20, 2021



# **Expected performance of the ATLAS detector under different High-Luminosity LHC conditions**

The ATLAS Collaboration

A major upgrade of the LHC is planned to be completed by 2027, resulting in the so-called High-Luminosity LHC (HL-LHC) operating at a target luminosity of  $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , with a possibility to reach an ultimate peak luminosity of up to  $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . This corresponds to an average of 140 to 200 proton–proton collisions per bunch crossing. To help the HL-LHC experiments separate particles coming from the different collisions in a single crossing, the HL-LHC can manipulate the beams in various ways to change the density of these collisions in the spatial dimensions, at the expense of potential performance degradation and financial costs. This note documents studies of the performance of the object-level reconstruction and identification with the ATLAS detector at HL-LHC as a function of the number of collisions per bunch crossing and their density.

# 1 Introduction

The luminosity upgrade of the LHC to High-Luminosity (HL-LHC) [1] is planned for the long shutdown in 2025-2027 with major upgrades for the injector chain installed in the 2019-2021 shutdown. The goal of HL-LHC is to accumulate at least  $3000 \text{ fb}^{-1}$  and potentially up to  $4000 \text{ fb}^{-1}$  of proton–proton ( $pp$ ) data at center-of-mass energy  $\sqrt{s} = 14 \text{ TeV}$  over a period of about 10 years. To achieve such high integrated luminosities, peak instantaneous luminosities between  $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  (baseline) and  $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  (ultimate) will be required. Furthermore, such luminosities will need to be maintained for several hours during a fill, achieved through so-called luminosity-leveling. This means the virtual peak luminosity is even higher (up to  $20 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ ), but the actual luminosity is kept roughly constant by not focusing the beams fully at the highest intensity, and then gradually increasing the focus as the beam intensity decreases during the fill.

If we assume an inelastic  $pp$  cross section of 81 mb and 2808 colliding bunch pairs [2], the instantaneous peak luminosities correspond to an average of 128 and 192 inelastic proton-proton collisions per bunch crossing, respectively. Since the bunch-to-bunch variations are significant, this has historically been rounded to averages of 140 and 200 collisions, respectively. Those values are therefore used to model the pileup in the studies presented in this document.

Any event containing a physics signal of interest will also contain particles produced in additional collisions in the same bunch crossing, and we refer to this effect as *in-time pileup*. In addition, the signals from particles produced in previous and subsequent bunch crossings, so-called *out-of-time pileup*, will affect the measurements of particles in a given event. All sub-detectors are not equally sensitive to the latter effect, but it is important to consider for e.g. the calorimeters. If not mitigated, pileup will result in degraded sensitivity in physics studies due to lower reconstruction and identification efficiencies, higher fake rates and degraded resolutions.

The most effective pileup mitigation strategies involve associating particles to different reconstructed proton–proton vertices, through either track impact location along the beam axis ( $z$  coordinate) or a very precise time measurement associated to the particle, or both. Both methods exploit the intrinsic spacial and temporal distribution of the proton collisions. The two-dimensional distribution of the pileup collisions in time and space is determined by the properties of the colliding beams, and is shown in Figure 1 along with the one-dimensional distribution vs.  $z$ . The beam parameters can be varied by the accelerator within certain limitations, producing higher or lower pileup density along the beam axis. However, this has an impact on the integrated luminosity, and in some cases also requires additional accelerator hardware. This impact needs to be compared to the possible gain or loss in the physics performance.

In this document, the performance of the ATLAS detector to reconstruct and correctly identify different physics objects, such as muons and jets, will be studied under different pileup conditions. This is done using simulation of the upgraded ATLAS detector as described in the Phase-II detector TDRs [3–8]. The performance depends on how strongly the reconstruction of such objects depends on the tracker and calorimeter information. The inner tracker (ITk) has a low occupancy thanks to its high granularity, and it provides excellent discrimination for tracks coming from vertices that are separated in space. But when the interaction vertices in a bunch crossing are packed more closely together, the capability to discern them is degraded. So the ITk performance is mostly sensitive to the *pileup density*. In contrast, the calorimeters have very little sensitivity to the collision location and their performance therefore mostly depends on the total number of pileup interactions. Only the pileup density in the spatial dimension along the beam line is studied in this document. There is also work ongoing to use the High-Granularity Timing Detector [8] to

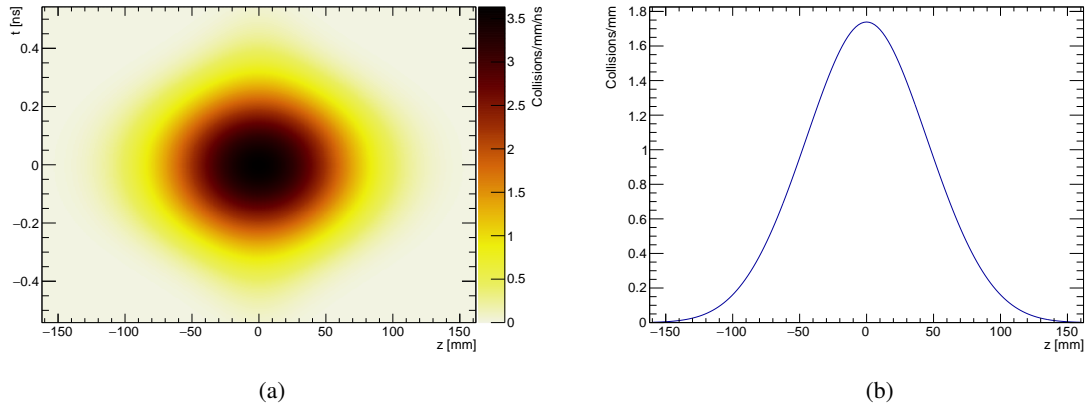


Figure 1: Example of (a) two-dimensional distribution of position along the beam axis and time, and (b) one-dimensional distribution along the beam axis. Both figures represent the scenario with baseline beam parameters and 200 pileup collisions. The distributions result from a full calculation integrating over the analytical expressions for the bunch parameters.

distinguish collisions in the temporal domain. There are however very few possibilities on the accelerator side to vary the temporal pileup density, so these studies are not included in this note.

This document collects both previously available public results and new studies that demonstrate the expected effect of pile-up or pile-up density on the ATLAS reconstruction and combined performance. It is envisaged that some of these effects can be mitigated with dedicated efforts and sophisticated algorithms.

This document is structured as follows. Section 2 introduces different HL-LHC running scenarios and their associated pileup conditions. In Section 3 the simulation used for the performance evaluation is described. The tracking and vertexing performance is presented in Section 4 and higher-level reconstructed objects such as electrons and  $b$ -tagged jets are presented in Section 5. Section 6 discusses how the changes in physics object performance impacts key physics channels.

## 2 HL-LHC luminous region scenarios

The running conditions of the HL-LHC are still evolving. In addition to the baseline configuration, several alternatives are, or have been, under consideration. The alternative configurations are designed to mitigate certain risks or reduce the overall costs of the accelerator upgrade. Most alternatives would also affect the luminous region and thus the pileup conditions in ATLAS. The main beam parameters for a selection of configurations are presented in Table 1, along with the parameters describing the resulting luminous regions. In all cases a bunch intensity of  $2.2 \times 10^{11}$  protons/bunch is assumed. The different beam configurations are briefly summarized in the following. For more details, see Ref. [9] and references therein.

- **Baseline:** In the nominal configuration, round beam optics (same  $\beta^*$  in  $x$  and  $y$  planes) are used as in the LHC, but with a smaller minimum  $\beta^*$  value. This requires a large crossing angle, which is partially ( $\sim 75\%$ ) compensated by two crab cavities (CCs) [10] located around IP1, leading to lower beam overlap and thus luminosity.

Table 1: Beam parameters for different HL-LHC configurations. All scenarios the peak pileup level of 200 interactions per bunch crossing.

Parameter	Baseline	Flat	8b+4e	200 MHz	No CCs
Number of bunches	2760	2760	1972	2760	2760
Number of colliding bunches in ATLAS	2748	2748	1967	2748	2748
RMS bunch length [cm]	7.6	7.6	7.6	15.0	7.6
Minimum $\beta_x^*$ [cm]	15.0	18.0	15.0	18.0	31.5
Minimum $\beta_{  }^*$ [cm]	15.0	7.5	15.0	7.5	7.5
Full crossing angle [ $\mu$ rad]	500	490	470	490	410
Norm. transversal emittance (start) [ $\mu$ m]	2.5	2.5	2.2	2.5	2.5
$z_{\text{RMS}}$ at start of leveling [mm]	47	49	49	54	37
$z_{\text{RMS}}$ at end of leveling [mm]	41	39	42	44	29
Peak instantaneous luminosity [ $\times 10^{34} \text{ cm}^{-1} \text{ s}^{-2}$ ]	7.5	7.5	5.5	7.5	7.5
Peak pileup density at start of leveling [ $\text{mm}^{-1}$ ]	1.62	1.56	1.58	1.48	2.13
Peak pileup density at end of leveling [ $\text{mm}^{-1}$ ]	1.95	2.02	1.88	1.93	2.73
Effective pileup density over fill [ $\text{mm}^{-1}$ ]	1.20	1.20	1.17	1.08	1.58
Yearly integrated luminosity [ $\text{fb}^{-1}/160 \text{ days}$ ]	326	340	243	304	293

- **Flat beams:** Uses non-round beam optics, i.e. with larger  $\beta^*$  in one direction. This allows a slightly smaller crossing angle and higher peak virtual luminosity, leading to favorable experimental conditions with reduced pileup density and higher integrated luminosity. However, this running mode has not yet been fully demonstrated in LHC operations.
- **8b+4e beams:** Effects due to electron clouds [11] remain a major concern for HL-LHC which uses higher bunch intensities than the LHC. The effect should be partially mitigated with the planned beam-screen coating, but additional mitigation might still be needed. The insertion of four empty bunches after each eight filled bunches has been demonstrated in Run 2 to effectively reduce electron-cloud effects. However, this reduces the number of colliding bunches by almost 30%, and running at nominal or ultimate luminosity would therefore result in much higher pileup than in the baseline scenario. Instead it will be assumed that the luminosity is leveled at the nominal and ultimate pileup values of 140 and 200 interactions per crossing, respectively.
- **200 MHz RF:** An alternative electron-cloud mitigation strategy is to install a second 200 MHz RF system in addition to the main 400 MHz one. This would allow using longer bunches which would reduce the electron clouds and also give a significant reduction in the pileup density. Flat optics is assumed in this configuration to minimize the peak luminosity loss from the increased bunch length.
- **No crab cavities:** The use of crab cavities has not been demonstrated to work at the LHC and is not guaranteed to be available for HL-LHC, at least during the early phase. Without the CCs compensating for the crossing angle, the peak virtual luminosity is significantly reduced as well as the luminous region length. This can be partially compensated by using more flat beams allowing the crossing angle to be somewhat reduced, but this still represents a more extreme scenario.

In the past an additional scheme called *crab-kissing* [12] was considered, which required a second set of crab cavities to be added in the non-crossing plane. This would allow an additional rotation of the beams that can be used to lengthen the luminous region and thus reduce pileup density. This scheme was explicitly

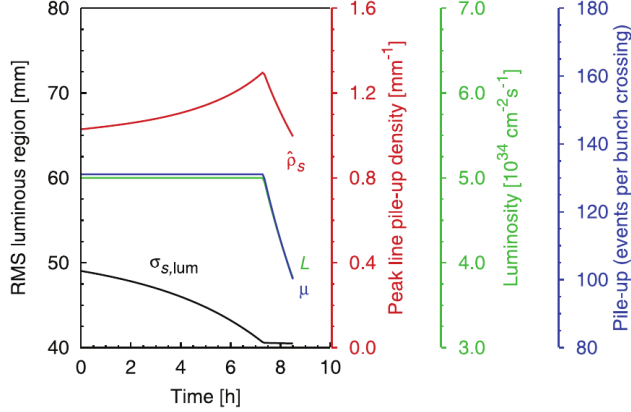


Figure 2: Example evolution of the luminous region size and pileup density as a function of time in an LHC fill [9].

designed to minimize the pileup density at the cost of requiring additional hardware, but is no longer under consideration given the high cost of additional crab cavities.

The longitudinal pileup density,  $\rho(z)$ , will depend on the detailed shape of the luminous region and the overall pileup level. In most scenarios, a Gaussian shape approximates the shape of the luminous region well and  $\rho(z)$  can be expressed as:

$$\rho(z) = \frac{\mu}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{z^2}{2\sigma_z^2}\right), \quad (1)$$

where  $\sigma_z$  is the Gaussian width of the luminous region along  $z$ , and  $\mu$  is the average number of collisions per bunch-crossing (overall pileup). This results in a peak pileup density of  $\frac{\mu}{\sqrt{2\pi}\sigma_z}$  at  $z = 0$ , and an average pileup density of  $\frac{\mu}{2\sqrt{\pi}\sigma_z}$ . As part of the luminosity-leveling procedure, the width of the luminous region will change with time during a fill. Typically, the additional focusing and the changes in crossing angle will make the width smaller with time, resulting in a higher pileup density at the end of the leveling period than at the start as shown in Table 1. Once the leveling period ends, the luminous region size stays constant, but the luminosity and therefore the pileup and pileup density drops off rapidly. This is illustrated in Figure 2. Integrating the pileup density over the collision position along  $z$  and over the duration of a fill, the effective average pileup density is determined for each of the beam configurations. To the extent that the performance impact depends only linearly on the pileup density, this gives a better measure of the overall difference in pileup density between each scenario. For example the ‘flat beams’ scenario has a higher peak pileup density than the ‘baseline’ scenario, but the effective pileup density remains the same and so a 4% gain in integrated luminosity should result in a net physics gain. In the ‘no crab cavities’ scenario, the peak pileup density increases by 40%, whereas the effective pileup density only increases by 30%, so comparing performance only at peak pileup density might overestimate the impact somewhat.

In this document, the time evolution of the luminous region size and the overall pileup are not simulated. Instead a fixed width is used and the performance is mostly measured as a function of pileup density by using the pileup density variation as a function of  $z$  given by Eq. 1. In reality the physics impact will depend on how many pileup collisions (and how many charged particles they produce) are near the hard-scatter collision of interest. For some studies, the performance has therefore also been studied as a function of the

actual pileup density, defined as:

$$\rho_{\text{actual}}(z) = \frac{N(PV_i : |z_i - z| < \Delta z)}{2\Delta z}, \quad (2)$$

where  $N$  is the number of collisions within a  $\pm\Delta z$  window around the hard-scatter collision. In this document,  $\Delta z = 2$  mm is used, though some measurements will also be sensitive to vertices further away.

### 3 Simulation

The studies presented in this document were performed using Monte Carlo samples of both single particles and several types of simulated  $pp$  collisions at  $\sqrt{s} = 14$  TeV. Di-jet events are generated using PYTHIA8 [13] with the ATLAS AU2-CT10 set of tuned parameters [14] for the underlying event (UE). For the generation of  $t\bar{t}$  events, matrix elements were calculated at NLO in QCD using a POWHEG kernel [15, 16] interfaced with PYTHIA6 [17] to simulate the parton shower, fragmentation and UE, employing the Perugia 2011C set of tuned parameters [18]. POWHEG interfaced with PYTHIA8 was used to generate  $Z \rightarrow \mu\mu/\tau\tau$  events, also here using the AU2-CT10 tune.  $H \rightarrow \gamma\gamma$  via gluon-gluon fusion and  $H \rightarrow ZZ \rightarrow 4\nu$  via vector-boson fusion were generated at NLO using POWHEG with the CT10 parton distribution function set interfaced with PYTHIA8 with parameters set according to the AZNLOCTEQ6L1 tune [14].

The effect of multiple interactions in the same and neighboring bunch crossings is modeled by overlaying simulated inelastic  $pp$  interactions generated with PYTHIA8 [13] with the A2 set of tuned parameters [14]. At fixed beam conditions, the number of pileup interactions follow a poisson distribution with average value  $\mu$ . As the bunches vary in intensity giving a spread in luminosity for different bunch crossings, the simulated samples are generated with a spread of  $\mu$  values. Samples with *average* pileup  $\langle\mu\rangle$  of 80, 140 and 200 therefore contain events with  $\mu$  in the ranges 70–90, 130–150 or 190–210, respectively.

To model the size of the luminous region, both the hard-scatter and the pileup overlay interactions were all generated with vertex distribution in  $z$  which follow a Gaussian with a width of 50 mm, centered at  $z = 0$ . Besides these ‘nominal  $\rho$ ’ samples, additional ‘high  $\rho$ ’ samples were generated by filtering both the hard-scatter and pileup overlay events to have a  $z$  distribution with a width of just 23 mm before the overlay is done. This increases the average pileup density from 1.13 events/mm to 2.45 events/mm.

All simulated event samples were processed with the GEANT4 simulation of the ATLAS detector [19, 20], including the upgraded tracker ITk [3, 4] with further updates as described in Refs. [8, 21]. The simulated events are reconstructed with the track and vertex reconstruction adjusted to ITk as described in Refs. [4, 21]. Globally, various detector geometries have been used for the various studies presented in this document. While the geometries affect the overall performance, the relative impact due to pile-up dependence is not expected to be different for them.

### 4 Tracking and vertexing performance

Highly efficient reconstruction of the trajectories of charged particles and precise determination of their momenta and points of origin are essential prerequisites for achieving a performant reconstruction of higher-level physics objects such as electron and muon candidates, jets, hadronically decaying tau leptons

and converted photons. One aim of the upgrade to the ITk tracking detector [3, 4] is to ensure that this is possible even under the most extreme pileup conditions expected at the HL-LHC.

The robustness of the ITk track reconstruction performance was extensively investigated in Ref. [4]. Figure 3 compares the track reconstruction efficiency for simulated single muons for  $\langle\mu\rangle = 0$  and 200. The presence of pileup degrades the efficiency by less than 1% for  $p_T = 1$  GeV and no degradation is observed for  $p_T = 10$  GeV. In addition, the inclusive track reconstruction efficiency is studied in simulated  $\sqrt{s} = 14$  TeV  $t\bar{t}$  events explicitly as a function of the number of overlaid  $pp$  interactions in the event, as shown in Figure 4(a). The figure shows that the track reconstruction efficiency is robust against pileup, also for complex final-state topologies. Figure 4(b) shows the ratio between the number of reconstructed tracks and the number of generated charged particles as a function of the number of overlaid  $pp$  interactions for the same simulated  $t\bar{t}$  events. The ratio is generally higher than the efficiency, indicating that additional tracks above the 1 GeV threshold are found, corresponding either to particles produced in material interactions or fake tracks due to random coincidental arrangements of hits that resulted in the reconstruction of a track. Given the nearly constant efficiency, a rising slope in this distribution would be indicative of an increase in the number of fake tracks. No such slope is observed, indicating that the fake track contribution does not grow significantly at high pileup.

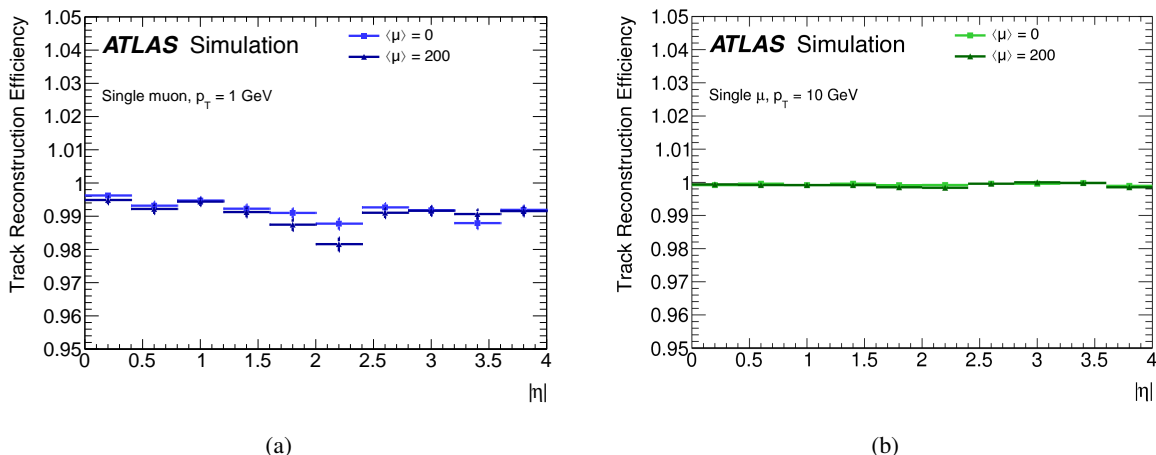


Figure 3: Track reconstruction efficiency for single muons in the ITk as a function of the pseudorapidity  $\eta$ , with a fixed transverse momentum  $p_T$  of (a) 1 GeV, or (b) 10 GeV for  $\langle\mu\rangle = 0$  and  $\langle\mu\rangle = 200$  [4].

In addition to the track reconstruction, the assignment of the reconstructed tracks to vertex candidates, and the identification of the hard-scatter vertex candidate among the reconstructed vertices, are both important for ensuring effective pileup mitigation for high-level physics objects. In what follows, the hard scatter-vertex is considered to be reconstructed if a vertex is found to be within 0.1 mm of the true hard-scatter position. Figure 5(a) shows that at higher pileup density, the number of reconstructed vertices decreases as vertices are merged. For  $t\bar{t}$ , a process with significant transverse activity, the efficiency to reconstruct the primary  $t\bar{t}$  vertex remains close to 100% and drops by less than 1% at the highest pileup density, as shown in Figure 5(b). In the case of processes with less visible activity, such as  $H \rightarrow 4\nu$  production via vector boson fusion where most of the charged particle activity is in the forward direction, the efficiency still remains stable within 3%. Identifying the correct hard-scatter vertex in such events is more difficult as selecting the vertex with the highest  $\sum p_T^2$  will select the wrong vertex more often, as the overall pileup increases. Further studies are documented in Refs. [4, 21, 22].

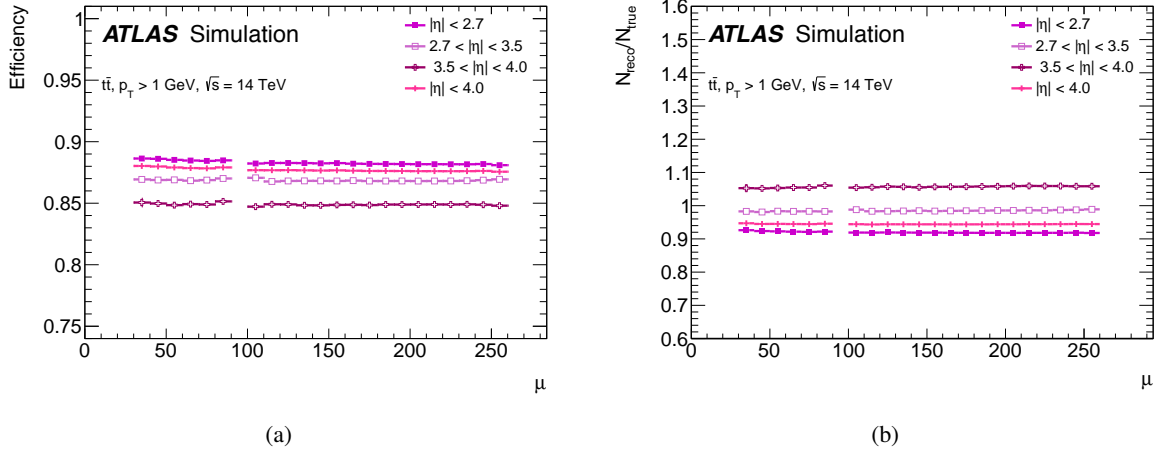


Figure 4: (a) Efficiency and (b) number of reconstructed tracks divided by the number of generated charged particles, measured in simulated  $t\bar{t}$  events as a function of  $\langle\mu\rangle$  [4].

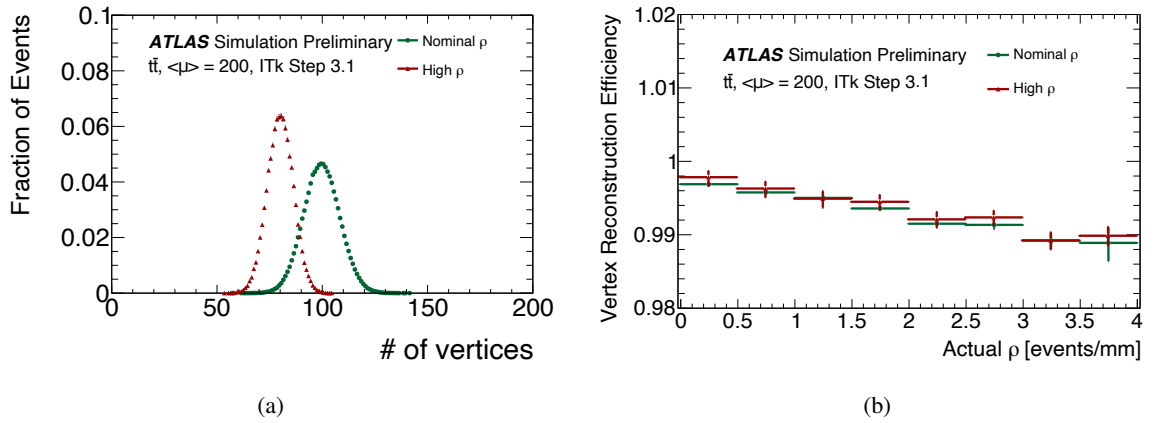


Figure 5: (a) Number of reconstructed vertices, and (b) reconstruction efficiency for the primary vertex of the hard-scatter process, for  $t\bar{t}$  events in samples with nominal and high pileup densities.

## 5 Physics object performance

High levels of pileup and high pileup densities give rise to challenging experimental conditions, but as seen above, they do not significantly affect track and vertex reconstruction. This section contains studies on the impact of these challenging conditions on the object reconstruction performance, where track and vertex requirements are combined with selections on based on quantities measured in the calorimeters and muon spectrometer (MS). These high-level objects (electrons, muons and jets) are used in physics analyses and are therefore important to study further. The impact is observed to be more significant than in pure tracking and vertexing studies.



## 5.1 Electron and Photon performance

The reconstruction of both electrons and photons starts by forming clusters of deposited energy in the electromagnetic calorimeter. In the case of electrons and converted photons, reconstructed tracks pointing to these clusters are matched to form candidates [23, 24]. The electron and photon identification is done using a cut-based selection on multiple variables related to the cluster shape, track quality, and track–cluster matching. More advanced reconstruction and identification methods have been developed for Run 2 [25], but have not yet been studied in detail for HL-LHC simulated data.

The reconstruction and identification steps depend strongly on the occupancy of the electromagnetic calorimeter and are consequently sensitive to the average pileup  $\langle\mu\rangle$ . This is illustrated in Figure 6 which shows the electron identification efficiency dropping by about 1% as  $\langle\mu\rangle$  increases from 140 to 200, while the mis-identification probability stays approximately constant. In contrast, very little dependence on pileup density  $\rho$  is expected.

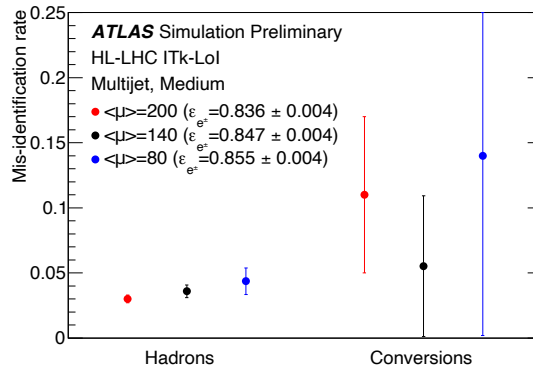


Figure 6: The probability to mis-identify hadrons or photon conversions as electrons estimated using simulated multijet samples is given for the ‘Medium’ working point at different pileup levels. The efficiencies are estimated with respect to electron candidates passing the track-quality requirements. The efficiencies for isolated electrons quoted in the legend ( $\epsilon_{e^\pm}$ ) were estimated using  $Z \rightarrow ee$  decays. In all cases,  $E_T > 7$  GeV and  $|\eta| \leq 2.47$  are required, the transition region between the barrel and endcap electromagnetic calorimeters,  $1.37 < |\eta| \leq 1.52$ , is included, and the reconstructed candidates are matched to their respective generator-level counterparts. [26]

To suppress backgrounds from mis-identified jets and hadron decays, isolation criteria are applied to electrons and photon candidates. Track-based isolation is sensitive to pileup density and is discussed in Section 5.3. Calorimeter-based isolation variables are instead sensitive to the overall pileup level. This is illustrated in Figure 7, which shows the calorimeter cone isolation variable  $E_{T,\text{topo}}^{R < R_c}$  [27] for simulated photons for two different cone radii.  $E_{T,\text{topo}}^{R < R_c}$  is computed as the sum of transverse energies of positive-energy topological clusters in the calorimeter within a cone of  $R_c = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$  centered around the photon candidate. The transverse energy of the candidate is removed and the contributions of the underlying event and pileup in  $R < R_c$  are subtracted based on the method suggested in Ref. [28]. Despite this so-called *pileup subtraction*, some pileup dependence is clearly visible.

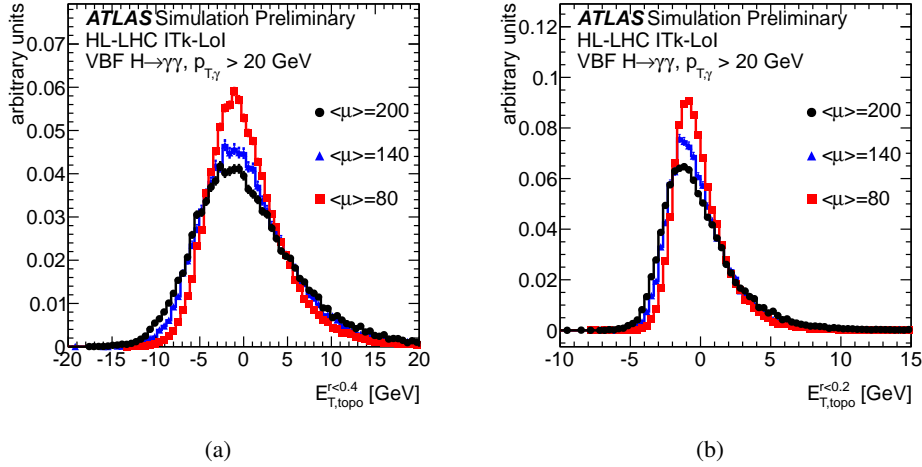


Figure 7: Calorimeter-based photon isolation variables (a)  $E_{T,topo}^{R<0.4}$  and (b)  $E_{T,topo}^{R<0.2}$  in simulated  $H \rightarrow \gamma\gamma$  events [26].

## 5.2 Muon performance

Muon candidates are formed by combining tracks in ITk and the MS [5]. Given that the MS is the outermost detector in ATLAS, its performance is mainly sensitive to  $\langle\mu\rangle$ , rather than the pileup density  $\rho$ . Performance degradation in the MS with increased  $\langle\mu\rangle$  is expected to occur primarily due to increasing hit rates from secondary photons and neutrons originating from nuclear decays in activated detector material.

As described in Ref. [5], the innermost set of endcap muon chambers, covering  $1.3 < |\eta| < 2.7$ , will be replaced by more radiation hard technology. Since this upgrade is not yet available in simulation, studies related to the muon spectrometer are limited to the region  $0.1 < |\eta| < 1.3$ , where the muon spectrometer is fully instrumented and expected to perform similarly to Run 2.

Figure 8 shows the muon reconstruction and identification efficiency measured in simulated  $\sqrt{s} = 14$  TeV  $Z \rightarrow \mu\mu$  events as a function of the muon  $p_T$  and of the local pileup density for  $\langle\mu\rangle = 140$  and 200. *Medium* muon identification criteria [29], which represent the combination of high efficiency and strong background rejection most commonly used in ATLAS physics analyses, are required to be satisfied in the efficiency definition. A slight efficiency degradation of about 0.5% at the highest pileup considered is observed. As expected, this is largely independent of the pileup density. The degradation is slightly more pronounced at low transverse momenta, where the reconstruction is more sensitive to the innermost muon chambers which have the highest background occupancy.

In addition to the efficiency, the momentum resolution is another important aspect of muon performance. For the best-reconstructed muons, the so-called *combined muons*, the momentum measurement uses information from both the ITk and the muon spectrometer. Below  $p_T \sim 100$  GeV, the ITk information contributes most to the precision of the measurement, while for high momenta ( $p_T \gtrsim 500$  GeV) and high pseudorapidities, the contribution from the MS measurement becomes dominant [5]. To explore both regimes, the relative momentum resolution is studied for simulated single muons of  $p_T = 10$  GeV, populating the ITk-dominated region, and  $p_T = 1000$  GeV, where the MS measurement is dominant. The resulting resolutions are shown in Figure 9 as a function of the local pileup density.

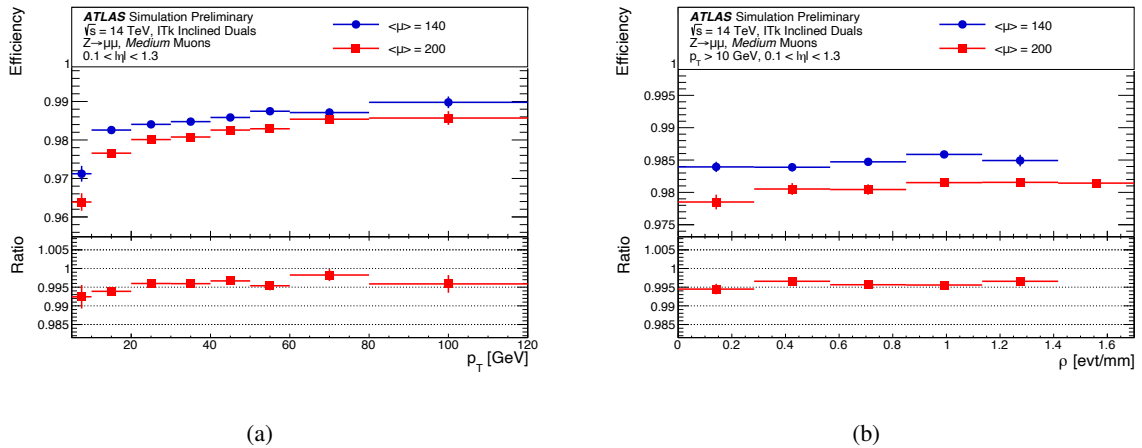


Figure 8: Muon reconstruction and identification efficiency for simulated  $\sqrt{s} = 14$  TeV  $Z \rightarrow \mu\mu$  events at  $\langle\mu\rangle=140$  (blue) / 200 (red), as a function of (a) the muon transverse momentum, and (b) the local pileup density.

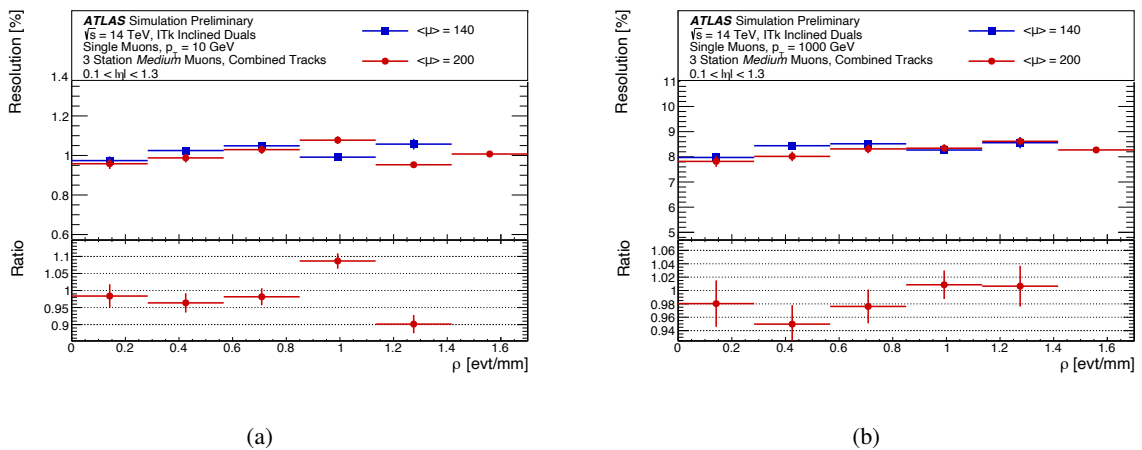


Figure 9: Muon transverse momentum resolution for simulated single muons of  $p_T = 10$  GeV (a) or 1 TeV (b) at  $\langle\mu\rangle=140$  (blue) and 200 (red), as a function of the local pileup density.

For both cases, the resolution is found to be independent of both the total amount of pileup and the pileup density. Together, these findings indicate that key aspects of muon performance are largely unaffected by the choice of HL-LHC running scenario.

### 5.3 Track isolation performance

Leptons from heavy particles will normally be isolated from other parts of the hard-scatter process and can therefore be efficiently separated from background using calorimeter and track isolation variables.

The track isolation variable ( $p_T^{\text{cone}}$ ) is computed by summing the transverse momenta of selected tracks within a cone centered around the electron track or the photon cluster. Tracks matched to the leptons are excluded. Since for leptons produced in the decay of heavy high-momentum particles, other decay products

can be very close to the lepton direction, the track isolation is typically defined with a variable cone size ( $p_T^{\text{varcone}}$ ) that shrinks for larger transverse momenta. The track selection requires well-reconstructed tracks that are consistent with the same interaction vertex as the lepton. A simple selection inspired by Run-2 studies requiring  $|\Delta z_0| \sin \theta < 3$  mm, where  $\Delta z_0$  is the longitudinal impact parameter of the track with respect to the primary vertex, was found to lead to a large dependence of the isolation efficiency on the pileup density. The  $|\Delta z_0| \sin \theta$  requirement was therefore optimized as a function of track  $p_T$  and  $\eta$ , taking advantage of the improved tracking capabilities of the ITk.

The resulting improved track isolation is demonstrated in Figure 10. The selection was adjusted to provide fixed background suppression and the efficiency is seen to decrease linearly on the pileup density. The efficiency loss at nominal peak pileup density compared to no pileup is less than 1% for central muons, while it reaches 5% for forward muons. Comparing the samples for nominal and high pileup density, the same efficiency is observed at a fixed actual pileup density, but as the average density is twice as high in the high-density sample, a 3% efficiency loss is expected for muons in  $t\bar{t}$  events as shown in Figure 10(b).

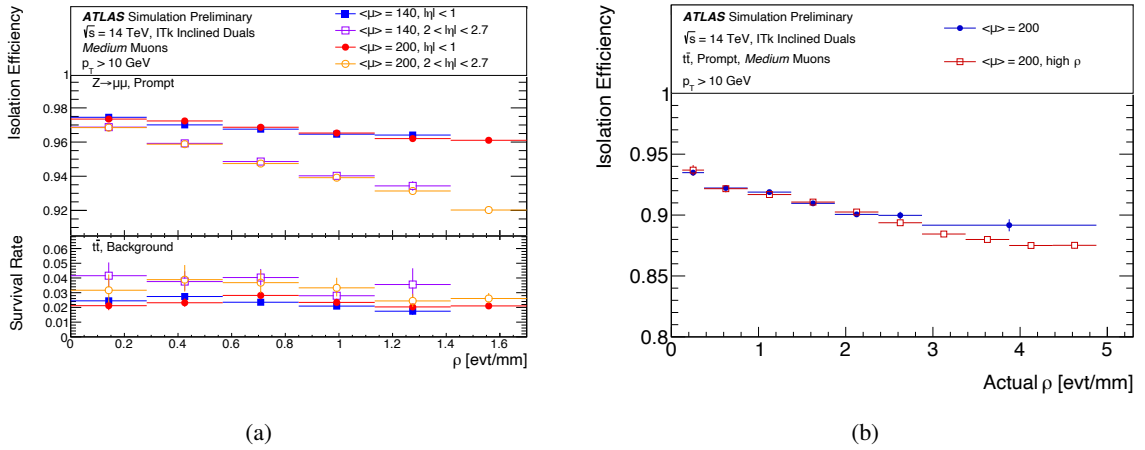


Figure 10: (a) Isolation efficiency and background survival rate of track-based isolation selection criteria for  $\langle\mu\rangle = 140$  and 200. (b) Comparison of efficiency vs actual pileup density for samples with nominal and very high average pileup density.

## 5.4 Tau performance

Visible hadronic decays of  $\tau$ -leptons (denoted as  $\tau_{\text{had-vis}}$ ) are reconstructed in ATLAS using jets formed from clusters of energy in the calorimeters using the anti- $k_r$  algorithm [30, 31] with radius parameter  $R = 0.4$  and calibrated using a local hadronic calibration [32]. These  $\tau_{\text{had-vis}}$  candidates are required to have  $p_T > 10$  GeV and be within the geometrical acceptance of the ITk,  $|\eta| < 4.0$ , and outside the transition between the barrel and endcap calorimeters ( $1.37 < |\eta| < 1.52$ ). A set of boosted decision trees (BDTs) is used to select tracks consistent with those found in  $\tau$  decays and reject those originating from secondary interactions, the underlying event or pileup. Each  $\tau_{\text{had-vis}}$  candidate is required to contain exactly one or three of the tracks selected by the BDTs (prongs), and have a total charge of  $\pm 1$ .

To separate  $\tau_{\text{had-vis}}$  from other jets, a recurrent neural network (RNN) algorithm is trained separately for one- and three-prong taus using tracking and calorimetric shower information. The track-selection BDTs and the tau identification RNN are optimized for the ATLAS detector in a scenario with  $\langle\mu\rangle = 200$ . The

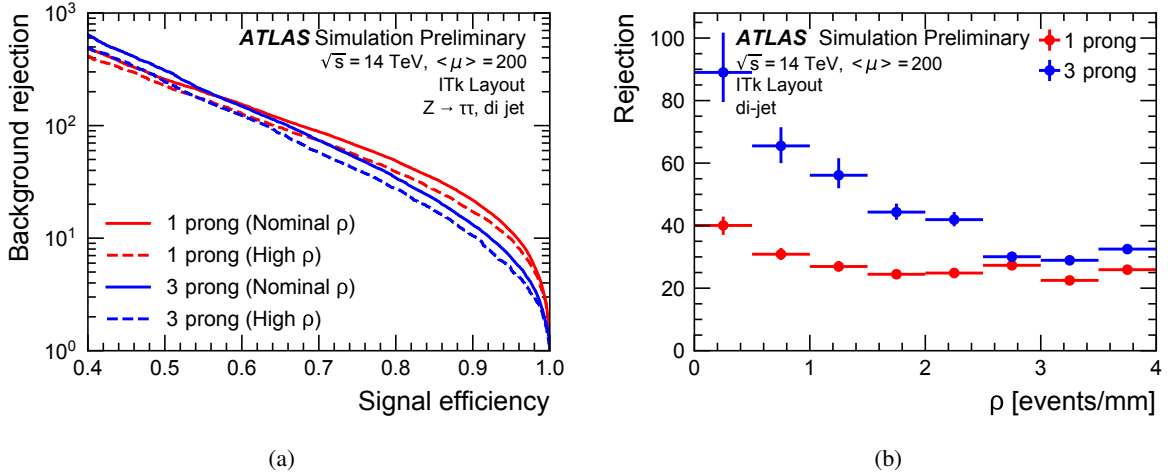


Figure 11: Jet rejection as a function of (a)  $\tau_{\text{had-vis}}$  efficiency for the scenario with nominal pileup density compared to a scenario with high pileup density, and (b) average pileup density for the loose  $\tau$  identification RNN working point, with corresponding efficiencies of 85% (75%) for one(three)-prong candidates, evaluated in the sample with high pileup density. One-prong (red) and three-prong (blue) performance is shown separately.

identification RNN is trained on a Drell-Yan ( $Z/\gamma^* \rightarrow \tau^+\tau^-$ ) signal sample, and was designed to provide  $\tau$  identification efficiencies independent of the total amount of pileup in the sample. The  $\tau_{\text{had-vis}}$  identification efficiency is defined as the ratio of reconstructed  $\tau_{\text{had-vis}}$  candidates passing a given RNN selection cut to the total amount of reconstructed  $\tau_{\text{had-vis}}$  candidates. Three efficiency working points are used for  $\tau_{\text{had-vis}}$  identification. They are evaluated for one-prong (three-prong)  $\tau_{\text{had-vis}}$  candidates, with efficiencies of 85% (75%) for the loose, 75% (60%) for the medium and 60% (45%) for the tight working point.

A plot of jet rejection versus  $\tau_{\text{had-vis}}$  efficiency is shown in Figure 11(a). Here, the jet rejection is defined as the inverse of the fake rate, where the fake rate represents the ratio of reconstructed jets with one (three) tracks that pass the RNN selection requirement to the total amount of reconstructed jets with one (three) tracks. The performance is compared in  $Z \rightarrow \tau\tau$  and di-jet events for scenarios with lower and higher pileup density. A slight decrease in performance is seen for the scenario with higher pileup density, of the order of 2-3% for all working points. This degradation impacts both one- and three-prong  $\tau_{\text{had-vis}}$  candidates. Figure 11(b) shows the degradation of the jet rejection for the loose identification working point as a function of pileup density.

## 5.5 Jet performance

So far, jets in ATLAS at the HL-LHC have primarily been studied with reconstruction starting from topological clusters of energy depositions in the calorimeters that are then formed into jets using for instance the anti- $k_t$  algorithm [30, 31] with different radii parameters. As the calorimeters cannot discriminate between different collisions within the same bunch crossing, a non-negligible amount of ‘pileup jets’ will arise from the pileup interactions, either from real jets from another interaction than the hard-scatter process of interest, or from combinations of deposits from multiple low-energy interactions.

Efficient identification and rejection of pileup jets is therefore essential to enhance the physics potential of the HL-LHC. A simple discriminant is the  $R_{p_T}$  jet variable [33] defined as the scalar sum of the  $p_T$  of all

tracks that are inside the jet cone and originate from the selected primary vertex, divided by the jet  $p_T$ , i.e.  $R_{p_T} = \frac{\sum p_T^{\text{trk}}}{p_T^{\text{jet}}}$ . For this study, the transverse momentum of the jet is calculated at the constituent-level scale, with jet-area-based subtraction of pileup energy. The tracks are associated with the jets via *ghost association* [34–36], and with the primary vertex (PV) according to the distance between the longitudinal impact parameter  $z_0$  and the  $z$  position of the PV. This distance is required to be less than 2.5 times the  $z_0$  resolution parameterized as a function of  $\eta$  and  $p_T$  of the track. In this study, jets are reconstructed from clusters of calorimeter energy deposits using the anti- $k_t$  algorithm with radius parameter  $R = 0.4$ . Jets in the simulated events are classified as hard-scatter or pileup jets by matching to truth jets reconstructed from stable and detector-interacting final-state particles emanating from the hard-scatter vertex. Jets within  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$  of a hard-scatter truth jet with  $p_T > 10$  GeV are classified as hard-scatter jets. Jets that are at least  $\Delta R > 0.6$  away from any true hard-scatter jet with  $p_T > 4$  GeV are classified as pileup jets.

Figure 12 shows the pileup-jet rejection factor as a function of the efficiency to keep real hard-scatter jets with  $30 < p_T < 50$  GeV. A substantial degradation in pileup-jet rejection is seen when going from the nominal pileup density to the high-density sample, particularly for jets in the forward region where the  $\Delta z_0$  resolution is worse. Typically an analysis will choose a predetermined working point to achieve the needed rejection factor for pileup jets. A higher pileup density would therefore lead to a lower efficiency for selecting the hard-scatter jets. Figure 13 shows the hard-scatter efficiency when requiring a fixed rejection factor of 50 for pileup jets in each pileup-density bin. The efficiency is seen to drop roughly linearly with pileup density with about a 3% (11%) drop when going from peak nominal pileup density to high pileup density for central (forward) jets.

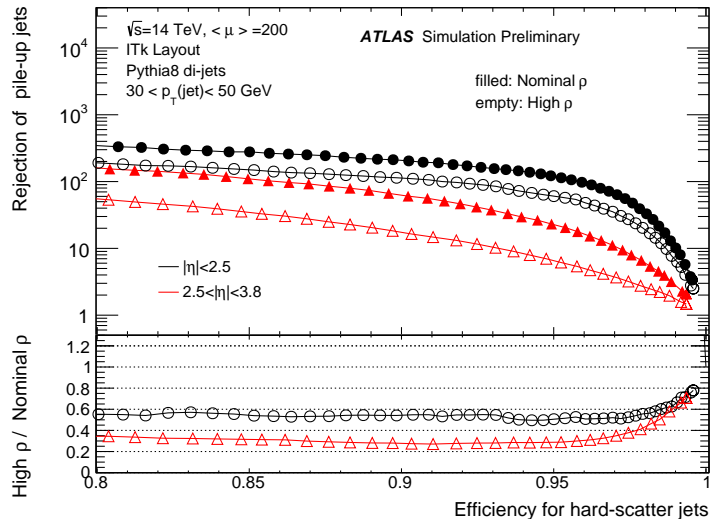


Figure 12: Pileup-jet rejection as a function of efficiency for hard-scatter jets, for  $|\eta| < 2.5$  and  $2.5 < |\eta| < 3.8$ , in events with nominal (filled symbols) and high pileup density (empty symbols).

## 5.6 Flavor-tagging performance

Identification of jets originating from  $b$ -quarks,  $b$ -tagging, is a crucial component of various HL-LHC analyses, including searches for di-Higgs production. The  $b$ -tagging performance is characterized by

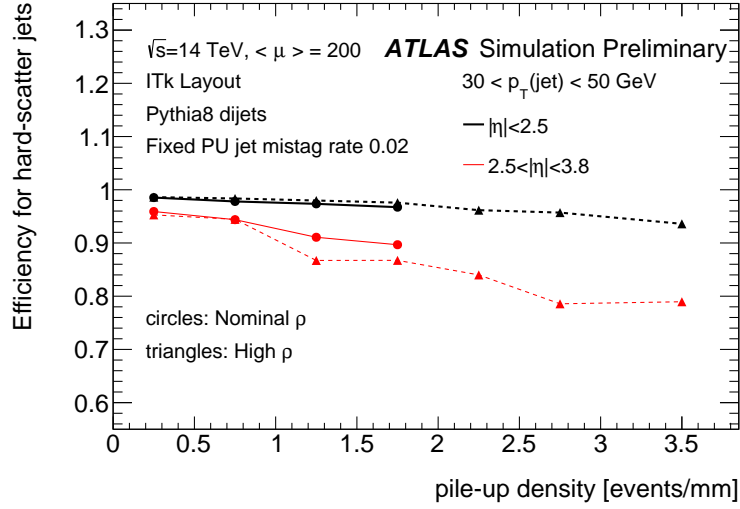


Figure 13: Efficiency for keeping hard-scatter jets, for  $|\eta| < 2.5$  and  $2.5 < |\eta| < 3.8$ , as a function of pileup density measured in events with nominal (circles) and high pileup density (triangles).

rejection (inverse of probability to misidentify as a  $b$ -jet) of light-flavor jets and  $c$ -jets as a function of  $b$ -tagging efficiency (probability to correctly identify a  $b$ -jet). The  $b$ -tagging algorithm used in the studies is MV2c10 that was developed in Run 2 and adapted for the ITk layout [37, 38]. It relies on a combination of track-based algorithms that make use of large impact parameters of tracks originating from  $b$ -decays, and algorithms based on secondary vertices. As both approaches are affected by the presence of additional tracks, pileup is expected to deteriorate the algorithm performance, especially in the forward region where the impact-parameter resolution is worse. Furthermore, since the algorithm efficiently suppresses the tracks originating far away from the hard-scatter PV, the  $b$ -tagging performance is expected to depend mainly on the presence of pileup vertices near the PV, i.e. be more sensitive to pileup density than the overall amount of pileup. The  $b$ -tagging performance has been evaluated for events with correctly reconstructed PV (where the distance between the true and reconstructed PV is less than 0.1 mm), so that the PV reconstruction and selection efficiencies described in Section 4 have been factored out.

Figure 14 shows the overall performance curve (light or  $c$ -jet rejection as a function of  $b$ -tagging efficiency) for samples with nominal and high pileup density. As expected, the performance deteriorates in the latter case, especially in the forward region.

To study the dependence of the performance as a function of the pileup density, the  $b$ -tagging operating point (the selection cut on the algorithm output) was adjusted in pileup density bins such that corresponding light-jet rejection remained the same for all bins. The resulting  $b$ -tagging efficiency is shown in Figure 15 as a function of pileup density. The  $b$ -tagging efficiency at fixed light-jet rejection decreases as the pileup density increases. For a given pileup density, the  $b$ -tagging efficiency is similar for both the nominal (red points) and high-density (blue points) samples.

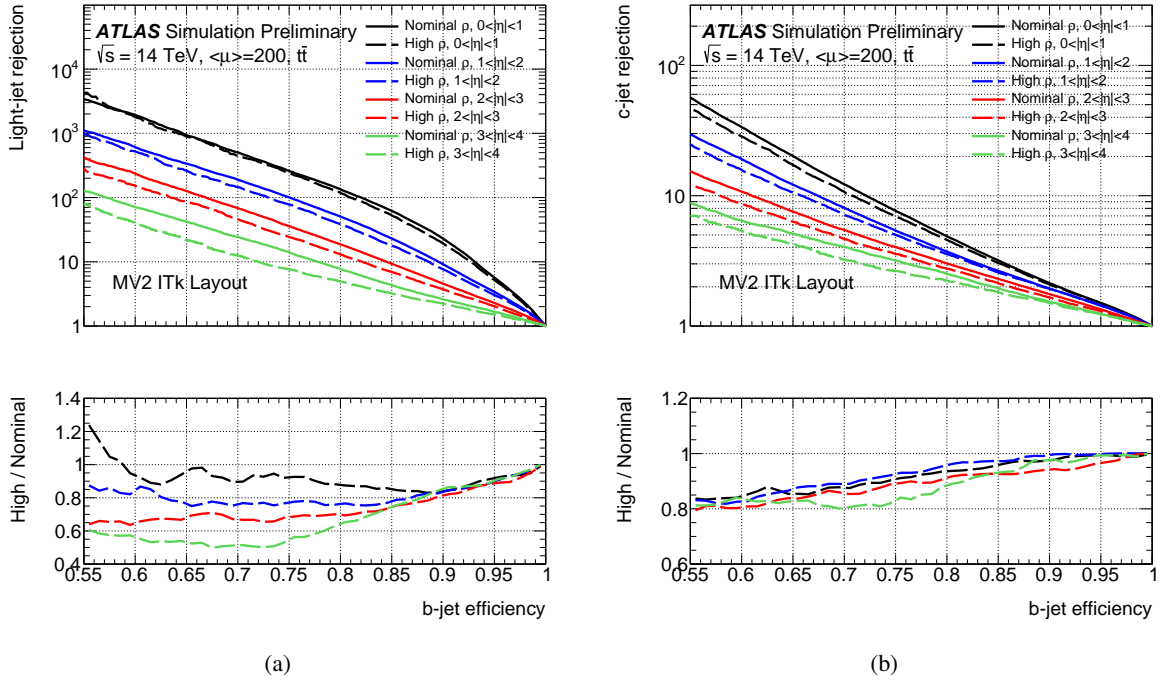


Figure 14: Comparison of (a) light-jet and (b)  $c$ -jet rejection as a function of  $b$ -tagging efficiency for samples with nominal and high pileup density in different jet pseudo rapidity regions.

## 6 Physics impact

The ATLAS Collaboration has charted an ambitious physics program for HL-LHC in which it intends to make measurements of the production and decay modes, and properties of the Higgs boson with high precision, access its rare decays, and carry out a broad program of other SM measurements and searches for beyond-SM physics. One of the most important pursuits will be the measurement of pair-production of Higgs bosons. This is a challenging measurement that requires sensitivity from as many decay channels as possible to become feasible. Given its large branching ratio,  $HH \rightarrow b\bar{b}b\bar{b}$  is a crucial channel, despite the experimental challenges associated with it. The impact of the HL-LHC running conditions on the sensitivity of this channel is discussed in what follows, as an example physics case.

### 6.1 $HH \rightarrow 4b$

The impact of the degraded  $b$ -tagging performance observed under high pileup densities presented in Section 5.6 has been evaluated for the SM  $pp \rightarrow HH \rightarrow b\bar{b}b\bar{b}$  HL-LHC analysis, re-assessing the expected sensitivity of this channel with respect to the previous published projections [38, 39].

Projections for this analysis were made by extrapolating from the ATLAS Run-2 results with  $24.3 \text{ fb}^{-1}$  of 13 TeV data. Four central jets with  $p_T > 40 \text{ GeV}$  are paired to construct two Higgs boson candidates. The acceptance times efficiency of the full event selection for the SM signal is of 1.6%, and around 95% of the background consists of multi-jet events, which is modeled with data-driven techniques, while the remaining 5% of the background originates from  $t\bar{t}$  processes. The largest source of systematic uncertainty



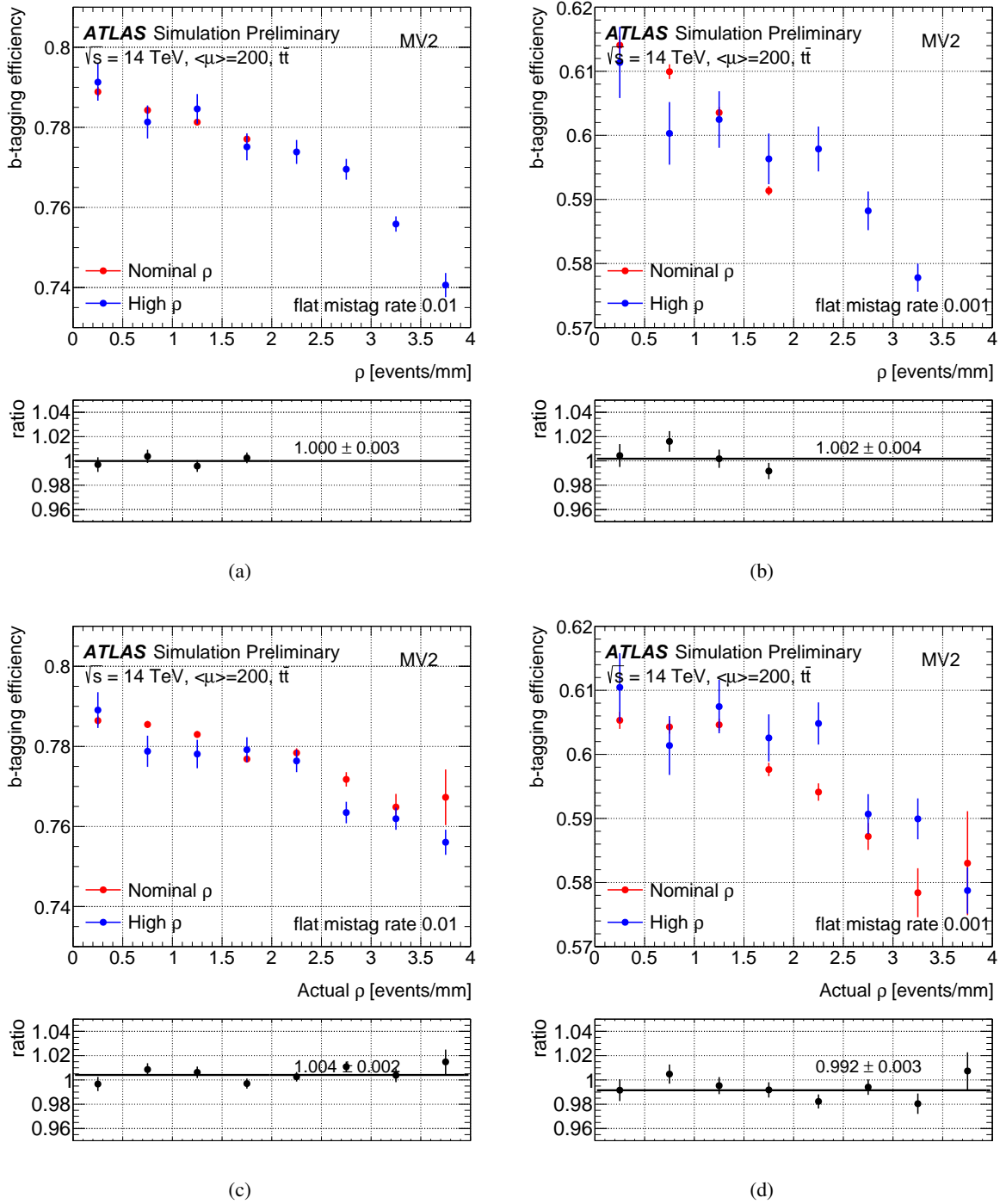


Figure 15: Efficiency for correctly tagging  $b$ -jets as a function of average (top) and local (bottom) pileup density at fixed light-jet rejection of 100 (left) and 1000 (right) for samples with nominal and high average pileup density.

is the ability to model the QCD multi-jet background using control regions in data, and this systematic uncertainty is left unchanged in the extrapolation performed here. The high number of pileup events at the HL-LHC cause difficulties in maintaining high acceptance when triggering on multi-jet final states. The

effect of the high-pileup conditions on the trigger efficiency has not been evaluated here, and the same trigger performance as in the early Run-2 analysis is used in the results presented in this section.

The  $b$ -tagging working point used in the analysis has an efficiency to identify  $b$ -jets of 70%, which corresponds to a rejection factor of 283 for light-quark and gluon-initiated jets, and a factor of 12 for charm-initiated jets [40]. The equivalent mis-tag rates are 0.35% and 8.3% respectively.

As shown in Section 5.6, the higher pileup density results in degraded flavor-tagging performance: at fixed mistag rate, the  $b$ -tagging efficiency decreases by a relative factor of 2.4%. To evaluate the effect of this degradation on the sensitivity of the  $HH \rightarrow b\bar{b}b\bar{b}$  measurement at HL-LHC, an event-by-event scale factor for the yields for both the signal and background processes has been computed to account for the lower  $b$ -tagging efficiency. The dominant multi-jet background, contributing to  $\sim 95\%$  of the total background, has been scaled according to its actual flavor composition derived from simulation. The overall correction factor is 0.9151 for the dominant multi-jet component, and 0.9063 for the signal and the sub-dominant background, primarily composed of  $t\bar{t}b\bar{b}$  events.

Figure 16 shows the flavor composition of the QCD background as a function of the mass of the system of the two Higgs-boson candidates ( $m_{HH}$ ), used as the final discriminant in the  $HH \rightarrow b\bar{b}b\bar{b}$  analysis. The characterisation of the flavor composition has been performed using a di-jet MC sample and does not represent the nominal background estimation of the  $HH \rightarrow b\bar{b}b\bar{b}$  analysis. The jets used for the event selection and the  $m_{HH}$  calculation are solely reconstructed jets, even if the flavor information shown here is based on the MC truth.

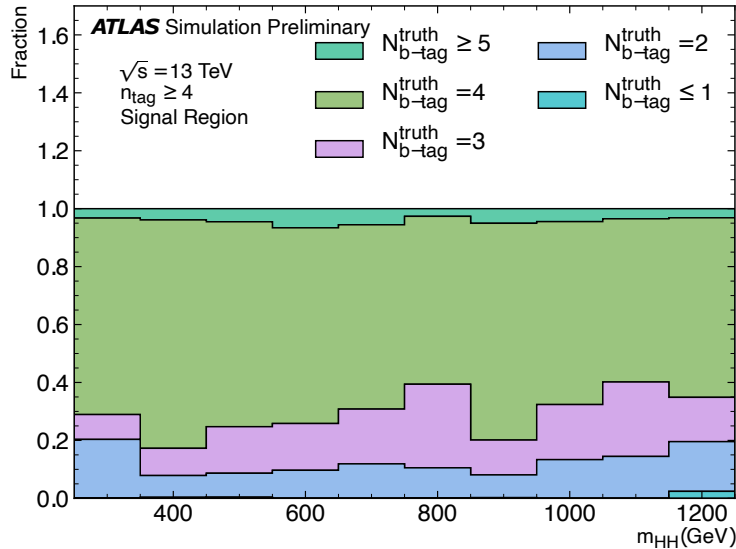


Figure 16: Composition of the QCD background split into categories based on the multiplicity of true  $b$ -jets, as a function of the mass of the system of the two Higgs-boson candidates ( $m_{HH}$ ).

In order to extract the expected significances, an Asimov dataset has been generated. The expected significance of the SM  $pp \rightarrow HH \rightarrow b\bar{b}b\bar{b}$  process relative to the background-only hypothesis, including both statistical and systematic uncertainties, are reported in Table 2.

In particular, the second column reports the expected significance taking into account the latest updated performance results including a more realistic description of the detector material [38], while the third column reports the re-assessed significance at higher pileup density. The degradation in  $b$ -tagging efficiency results in a lower significance by  $\sim 4\%$ . The loss in significance due to  $b$ -tagging in high pileup density environments, can be compensated for by collecting  $340 \text{ fb}^{-1}$  of additional data.

	Nominal $\rho$	High $\rho$
$HH \rightarrow b\bar{b}b\bar{b}$	0.547	0.525

Table 2: SM significance of  $HH \rightarrow b\bar{b}b\bar{b}$  including both statistical and systematic uncertainties, and the  $b$ -tagging performance expected at high pileup density. The study assumes the same jet- $p_T$  thresholds and systematic uncertainties as those used in the Run-2 analysis.

## 7 Summary

The performance of the ATLAS detector, including its planned Phase-II upgrades, has been studied for various HL-LHC running scenarios. The tracking and vertexing capabilities have been found to be quite robust against the challenging pileup conditions, and many important performance metrics are not degraded substantially by increased pileup density. At the level of physics objects, the impact is significant in several places. Generally, objects that rely heavily on track-to-vertex association depend on the pileup density, while objects based primarily on calorimeter and muon-spectrometer signatures are more sensitive to the overall level of pileup rather than the density. A small but visible degradation is seen for electrons, photons and muons, but the impact is expected to be limited on the physics program. The negative impact on the capability to identify of  $\tau$ -leptons and  $b$ -jets is larger when considering scenarios with substantially increased pileup density, yielding a typical degradation in efficiency of a few percent at similar background rejection. For analyses that target final states with high multiplicities of affected objects, the sensitivity can be degraded significantly. As an example, the expected degradation for flavor tagging in the high-pileup-density scenario will reduce the sensitivity of the  $HH \rightarrow 4b$  analysis at a level that would require an additional  $340 \text{ fb}^{-1}$  of HL-LHC data to compensate for it.

## References

- [1] G. Apollinari et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*, CERN Yellow Reports: Monographs, Geneva: CERN, 2017, URL: <https://cds.cern.ch/record/2284929> (cit. on p. 2).
- [2] ATLAS Collaboration, *Expected pile-up values at the HL-LHC*, ATL-UPGRADE-PUB-2013-014, 2013, URL: <https://cds.cern.ch/record/1604492> (cit. on p. 2).
- [3] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, (2017), URL: <https://cds.cern.ch/record/2257755> (cit. on pp. 2, 6, 7).
- [4] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, (2017), URL: <https://cds.cern.ch/record/2285585> (cit. on pp. 2, 6–8).

- [5] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, (2017), URL: <http://cds.cern.ch/record/2285580> (cit. on pp. 2, 10).
- [6] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS LAr Calorimeter*, (2017), URL: <https://cds.cern.ch/record/2285582> (cit. on p. 2).
- [7] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, (2017), URL: <https://cds.cern.ch/record/2285583> (cit. on p. 2).
- [8] ATLAS Collaboration, *Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade*, (2020), URL: <http://cds.cern.ch/record/2719855> (cit. on pp. 2, 6).
- [9] L. Medina, R. Tomás, G. Arduini, and M. Napsuciale, *Assessment of the performance of High-Luminosity LHC operational scenarios: integrated luminosity and effective pile-up density*, *Can. J. Phys.* **97** (2019) 498 (cit. on pp. 3, 5).
- [10] R. Calaga, R. Tomas, and F. Zimmermann, *3rd CARE-HHH-APD Workshop: Towards a Roadmap for the Upgrade of the LHC and GSI Accelerator Complex*, CERN-2007-002, Geneva: CERN, 2007 77, URL: <https://cds.cern.ch/record/974075> (cit. on p. 3).
- [11] G. Rumolo et al., *Electron Cloud Effects at the LHC and LHC Injectors*, (2017) **MOZA1. 7 p**, URL: <http://cds.cern.ch/record/2289719> (cit. on p. 4).
- [12] S. Fartoukh, *Pile up management at the high-luminosity LHC and introduction to the crab-kissing concept*, *Phys. Rev. ST Accel. Beams* **17** (11 2014) 111001, URL: <https://link.aps.org/doi/10.1103/PhysRevSTAB.17.111001> (cit. on p. 4).
- [13] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, arXiv: [0710.3820](https://arxiv.org/abs/0710.3820) [hep-ph] (cit. on p. 6).
- [14] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: <https://cds.cern.ch/record/1474107> (cit. on p. 6).
- [15] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **2007** (2007) 070, arXiv: [0709.2092](https://arxiv.org/abs/0709.2092) [hep-ph] (cit. on p. 6).
- [16] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **2010** (2010), arXiv: [1002.2581](https://arxiv.org/abs/1002.2581) [hep-ph] (cit. on p. 6).
- [17] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, arXiv: [hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175) (cit. on p. 6).
- [18] P. Z. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, *Physical Review D* **82** (2010), arXiv: [1005.3457](https://arxiv.org/abs/1005.3457) [hep-ph] (cit. on p. 6).
- [19] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: [1005.4568](https://arxiv.org/abs/1005.4568) [physics.ins-det] (cit. on p. 6).
- [20] S. Agostinelli et al., *GEANT4: A simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250 (cit. on p. 6).

- [21] ATLAS Collaboration, *Expected Tracking Performance of the ATLAS Inner Tracker at the HL-LHC*, ATL-PHYS-PUB-2019-014, 2019, URL: <https://cds.cern.ch/record/2669540> (cit. on pp. 6, 7).
- [22] M. Safdari et al., *Rescuing VBF Higgs Invisible Events with Novel Vertex Selection*, ATL-PHYS-PROC-2020-042, 2020, URL: <https://cds.cern.ch/record/2719814> (cit. on p. 7).
- [23] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **79** (2019) 639, arXiv: [1902.04655](https://arxiv.org/abs/1902.04655) [hep-ex] (cit. on p. 9).
- [24] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, *Eur. Phys. J. C* **79** (2019) 205, arXiv: [1810.05087](https://arxiv.org/abs/1810.05087) [hep-ex] (cit. on p. 9).
- [25] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, arXiv: [1908.00005](https://arxiv.org/abs/1908.00005) [hep-ex] (cit. on p. 9).
- [26] ATLAS Collaboration, *Expected performance for an upgraded ATLAS detector at High-Luminosity LHC*, ATL-PHYS-PUB-2016-026, 2016, URL: <https://cds.cern.ch/record/2223839> (cit. on pp. 9, 10).
- [27] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) 490, arXiv: [1603.02934](https://arxiv.org/abs/1603.02934) [hep-ex] (cit. on p. 9).
- [28] M. Cacciari, G. P. Salam, and S. Sapeta, *On the characterisation of the underlying event*, *JHEP* **04** (2010) 065, arXiv: [0912.4926](https://arxiv.org/abs/0912.4926) [hep-ph] (cit. on p. 9).
- [29] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **76** (2016) 292, arXiv: [1603.05598](https://arxiv.org/abs/1603.05598) [hep-ex] (cit. on p. 10).
- [30] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph] (cit. on pp. 12, 13).
- [31] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, arXiv: [1111.6097](https://arxiv.org/abs/1111.6097) [hep-ph] (cit. on pp. 12, 13).
- [32] ATLAS Collaboration, *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001-2, 2008, URL: <https://cds.cern.ch/record/1112035> (cit. on p. 12).
- [33] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: <https://cds.cern.ch/record/1700870> (cit. on p. 13).
- [34] ATLAS Collaboration, *Performance of jet substructure techniques for large- $R$  jets in proton–proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, *JHEP* **09** (2013) 076, arXiv: [1306.4945](https://arxiv.org/abs/1306.4945) [hep-ex] (cit. on p. 14).
- [35] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005, arXiv: [0802.1188](https://arxiv.org/abs/0802.1188) [hep-ph] (cit. on p. 14).
- [36] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett. B* **659** (2008) 119, arXiv: [0707.1378](https://arxiv.org/abs/0707.1378) [hep-ph] (cit. on p. 14).

- [37] ATLAS Collaboration, *Measurements of  $b$ -jet tagging efficiency with the ATLAS detector using  $t\bar{t}$  events at  $\sqrt{s} = 13$  TeV*, *JHEP* **08** (2018) 089, arXiv: [1805.01845](https://arxiv.org/abs/1805.01845) [[hep-ex](#)] (cit. on p. 15).
- [38] ATLAS Collaboration, *Expected  $b$ -tagging Performance with the upgraded ATLAS Inner Tracker Detector at the High-Luminosity LHC*, ATL-PHYS-PUB-2020-005, 2020, URL: <https://cds.cern.ch/record/2713377> (cit. on pp. 15, 16, 19).
- [39] M. Cepeda et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2018) 221, arXiv: [1902.00134](https://arxiv.org/abs/1902.00134), URL: <https://cds.cern.ch/record/2650162> (cit. on p. 16).
- [40] ATLAS Collaboration, *Optimisation of the ATLAS  $b$ -tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: <https://cds.cern.ch/record/2160731> (cit. on p. 18).