Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s} = 13$ TeV in ATLAS

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Abstract   Inclusive and differential cross-sections for the production of a top-quark pair in association with a photon are measured with proton-proton collision data corresponding to an integrated luminosity of 36.1 fb$^{-1}$, collected by the ATLAS detector at the LHC in 2015 and 2016 at a centre-of-mass energy of 13 TeV. The measurements are performed in single-lepton and dilepton final states in a fiducial volume. Events with exactly one photon, one or two leptons, a channel-dependent minimum number of jets, and at least one $b$-jet are selected. Neural network algorithms are used to separate the signal from the backgrounds. The fiducial cross-sections are measured to be $521 \pm 9$ (stat.) $\pm 41$ (sys.) fb and $69 \pm 3$ (stat.) $\pm 4$ (sys.) fb for the single-lepton and dilepton channels, respectively. The differential cross-sections are measured as a function of photon transverse momentum, photon absolute pseudorapidity, and angular distance between the photon and its closest lepton in both channels, as well as azimuthal opening angle and absolute pseudorapidity difference between the two leptons in the dilepton channel. All measurements are in agreement with the theoretical predictions.

1 Introduction

Measurements of top-quark properties play an important role in testing the Standard Model (SM) and its possible extensions. Studies of the production and kinematic properties of a top-quark pair in association with a photon ($t\bar{t}\gamma$) probe the $t\gamma$ electroweak coupling. For instance, deviations in the transverse momentum ($p_T$) spectrum of the photon from the SM prediction could point to new physics through anomalous dipole moments of the top quark [1–3]. A precision measurement of the $t\bar{t}\gamma$ production cross-section could effectively constrain some of the Wilson coefficients in top-quark effective field theories [4]. Furthermore, differential distributions of photon production in $t\bar{t}$ events could provide insight on the $t\bar{t}$ production mechanism, in particular about the $t\bar{t}$ spin correlation and the charge asymmetry [5].

Evidence for the production of a top-quark pair in association with an energetic, isolated photon was found in proton-antiproton ($p\bar{p}$) collisions at the Tevatron collider at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV by the CDF Collaboration [6]. Observation of the $t\bar{t}\gamma$ process was reported by the ATLAS Collaboration in proton-proton ($pp$) collisions at $\sqrt{s} = 7$ TeV [7]. Recently, both the ATLAS and CMS Collaborations measured the $t\bar{t}\gamma$ cross-section at $\sqrt{s} = 8$ TeV [8,9]. In the ATLAS measurement, the differential cross-sections with respect to the transverse momentum $p_T$ and absolute pseudorapidity $|\eta|$ of the photon were reported. In the CMS measurement, the ratio of the $t\bar{t}\gamma$ fiducial cross-section to the $t\bar{t}$ total cross-section was measured.

This paper describes a measurement of the $t\bar{t}\gamma$ production cross-section in final states with one or two leptons, electron or muons, referred to as the single-lepton or dilepton channel, based on a data set recorded at the LHC in 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and corresponding to an integrated luminosity of 36.1 fb$^{-1}$. The photon can originate not only from a top quark, but also from its charged decay products, including a charged fermion (quark or lepton) from the decay of the $W$-boson. In addition, it can be radiated from an incoming charged parton. In this analysis, no attempt is made to separate these different sources of photons, but criteria are applied to suppress those radiated from top-quark decay products: e.g. by requiring the photon to have a large angular distance from the lepton(s). In each channel, the fiducial inclusive cross-section, referred to as fiducial cross-section in the following for simplicity, is measured with a likelihood fit to the output of a neural network trained to

\[ \eta = -\ln \tan(\theta/2). \]
differentiate between signal and background events. In both channels, differential cross-sections, normalized to unity, are measured in the same fiducial region without performing the likelihood fit, as a function of the photon \( p_T \), the photon \( |\eta| \), and the distance \( \Delta R \) between the photon and its closest lepton. The distance \( \Delta R \) between two objects is defined as the quadratic sum of their pseudorapidity difference \( \Delta \eta \) and azimuthal opening angle \( \Delta \phi \). In the dilepton channel, the normalized differential cross-sections are also measured as a function of the absolute pseudorapidity difference \( |\Delta \eta| \) and \( \Delta \phi \) between the two leptons, the latter being sensitive to the spin correlation of the \( t \bar{t} \) pair. The measured cross-sections are compared to predictions from leading order (LO) generators. The predictions for the inclusive cross-sections are corrected by next-to-leading order (NLO) \( k \)-factors [10] in the strong interaction, calculated at parton level.

This paper is organized as follows. The ATLAS detector is briefly introduced in Sect. 2. The data and simulation samples used are listed in Sect. 3. The derivation of the NLO correction to the LO cross section is described in Sect. 4. The object and event selection, and the neural-network algorithms are presented in Sect. 5. The estimation of the backgrounds are introduced in Sect. 6. The definition of the fiducial region and the strategies to extract the fiducial and differential cross-sections are described in Sect. 7. The evaluation of the systematics uncertainties are discussed in Sect. 8. Section 9 gives the final results, and Sect. 10 presents the conclusion.

2 ATLAS detector

The ATLAS detector [11] consists of three main components. The innermost component is the Inner Detector (ID), which is used for tracking charged particles. It surrounds the beam pipe and is located inside a superconducting solenoid, operating with a magnetic field of 2 T. An additional silicon pixel layer, the insertable B-layer, was added between 3 and 4 cm from the beam line to improve \( b \)-hadron tagging [12,13] for Run 2. The calorimeter outside the ID is divided into two subsystems. The inner subsystem is the electromagnetic calorimeter (ECAL) and the second is the hadronic calorimeter (HCAL). The outermost layer is the third main component of the ATLAS detector: the muon spectrometer (MS), which is within a magnetic field provided by air-core toroid magnets with a bending integral of about 2.5 Tm in the barrel and up to 6 Tm in the endcaps. The ID provides tracking information from silicon pixel and silicon microstrip detectors in the pseudorapidity range \( |\eta| < 2.5 \) and from a transition radiation tracker (TRT) covering \( |\eta| < 2.0 \). The magnetic field of the superconducting solenoid bends charged particles for the momentum measurement. The ECAL uses lead absorbers and liquid argon (LAr) as active medium and is divided into barrel (\( |\eta| < 1.475 \)) and end-cap (\( 1.375 < |\eta| < 3.2 \)) regions. The HCAL is composed of a steel/scintillating-tile calorimeter, segmented into three barrel structures within \( |\eta| < 1.7 \), and two copper/LAr hadronic endcap calorimeters, that cover the region \( 1.5 < |\eta| < 3.2 \). The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules, optimised for electromagnetic and hadronic measurements respectively, and covering the region \( 3.1 < |\eta| < 4.9 \). The MS measures the deflection of muon tracks within \( |\eta| < 2.7 \) using multiple layers of high-precision tracking chambers in toroidal fields of approximately 0.5 T and 1 T in the central and end-cap regions, respectively. The MS is instrumented with separate trigger chambers covering \( |\eta| < 2.4 \).

Data are selected from inclusive pp interactions using a two-level trigger system [14]. A hardware-based trigger uses custom hardware and coarser-granularity detector data to initially reduce the trigger rate to approximately 100 kHz from the original 40 MHz LHC proton bunch crossing rate. Next, a software-based high-level trigger, which has access to the full detector granularity, is applied to further reduce the event rate to 1 kHz.

3 Data and simulation samples

The data used for this analysis were recorded by the ATLAS detector in 2015 and 2016 at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 36.1 fb\(^{-1}\). Only the data-taking periods in which all detector systems were operating normally are considered. Candidate events were collected using single-lepton triggers, designed to select events with at least one isolated high-\( p_T \) electron or muon.

The signal and background processes were modelled using Monte Carlo (MC) generators and passed through a detector simulation using GEANT 4 [15,16]. The simulated events were reconstructed with the same software algorithms as data. To account for overlapping pp collisions (pile-up), multiple interactions were simulated with the soft QCD processes of PYTHIA v8.186 [17] using the set of tuned parameters called A2 [18] and the MSTW2008LO parton distribution functions (PDF) set [19].

The \( t \bar{t} \gamma \) signal sample was simulated as a \( 2 \rightarrow 7 \) process for the semileptonic and dileptonic decay channels of the \( t \bar{t} \) system at LO by MadGraph5_aMC@NLO v2.33 [20] (denoted as MG5_aMC in the following) interfaced with PYTHIA v8.212 [21], using the A14 set of tuned parameters [22] and the NNPDF2.3LO PDF set [23]. The photon could be radiated from an initial charged parton, an intermediate top quark, or any of the charged final state particles. The top-quark mass, top-quark decay width, W-boson decay width, and fine structure constant were set to 172.5 GeV, 1.320 GeV, 2.085 GeV, and 1/137, respectively. The five-flavour scheme
was used where all the quark masses are set to zero, except for the top quark. The renormalization and the factorization scales were set to $0.5 \times \sum \sqrt{m_i^2 + \left( p_T^i \right)^2}$, where the sum runs over all the particles generated from the matrix element calculation. The photon was requested to have $p_T > 15$ GeV and $|\eta| < 5.0$. At least one lepton with $p_T > 15$ GeV was required, with all the leptons satisfying $|\eta| < 5.0$. The $\Delta R$ between the photon and any of the charged particles among the seven final-state particles were required to be greater than 0.2. The resulting total cross-section of the sample was calculated to be 4.62 pb. The NLO $k$-factors, introduced in Sect. 4, were applied to correct the fiducial cross-sections and acceptances to NLO.

The inclusive $t\bar{t}$ sample [24] was generated with POWHEG-Box v2 [25] using the NNPDF3.0NLO PDF set [26], and interfaced with PYTHIA v8.210 using the A14 tune set and the NNPDF2.3LO PDF set. The $h_{\text{damp}}$ parameter, which controls the $p_T$ of the first additional parton emission beyond Born level in POWHEG, was set to 1.5 times the top-quark mass. The production of a vector boson ($V = W, Z$) in association with a photon ($V\gamma$) was simulated with SHERPA v2.2.2 [27], and the inclusive production of $V$+jets [28] was simulated with SHERPA v2.2.1, both using the NNPDF3.0NLO PDF set. The $s$-channel single top quark and $tW$ samples [24] were produced with POWHEG-Box v1 using the CT10 (NLO) PDF set [29], interfaced with PYTHIA v6.428 using the Perugia 2012 tune set [30] and the CTEQ6L1 PDF set [23]. The $t$-channel single top quark was produced with the same generator and parton shower, with the four-flavour scheme and the corresponding CT104fs PDF set [23]. The diboson samples of $WW$, $WZ$ and $ZZ$ [31] were generated by SHERPA v2.1, using the CT10 (NLO) PDF set. The $t\bar{t}V$ samples [32] were generated with MG5\_aMC v2.2 using the NNPDF3.0NLO PDF set, interfaced with PYTHIA v8.210 using the NNPDF2.3LO PDF set. For all samples without photon radiation in the matrix element calculation, the radiation was simulated by the corresponding parton shower. The EvtGen program [33] was used to simulate the decay of bottom and charm hadrons, except for the SHERPA samples. All these samples were generated with NLO precision in QCD and, in the case of SHERPA samples, the NLO calculations were performed for up to one or two additional partons.

To assess the effects of initial- and final-state radiation (ISR and FSR), alternative signal samples were produced with the relevant PYTHIA 8 A14 Var3c tune parameters [22] varied to increase or decrease the parton radiation. The effect of the choice of parton shower algorithm for the signal is evaluated with a sample generated using Herwig v7.0.1 [34] instead of PYTHIA v8.212. An alternative $t\bar{t}$ sample was generated to enhance the parton shower radiation, with the renormalization and factorization scales varied down by a factor of two, the high radiation variation of the A14 Var3c tune parameter and the $h_{\text{damp}}$ value increased by a factor of two. A corresponding $t\bar{t}$ sample with reduced parton shower radiation was generated with the renormalization and factorization scales multiplied by a factor of two and the low radiation variation of the A14 Var3c tune parameter. The uncertainty arising from the choice of $t\bar{t}$ generator is evaluated using a SHERPA v2.2 sample. An alternative $Z\gamma$ sample, generated by MG5\_aMC v2.33 interfaced with PYTHIA v8.212, is used to evaluate the modelling uncertainty of the $Z\gamma$ background estimate.

The $t\bar{t}$ and $V$+jets samples contain events already accounted for by the $t\bar{t}\gamma$ and $V\gamma$ samples. Based on truth information, the overlap is removed by vetoing the events where the selected photon originates from the hard interaction in the $t\bar{t}$ and $V$+jets samples.

4 Next-to-leading order $k$-factor for $t\bar{t}\gamma$

Calculations at NLO precision in QCD are available for the $t\bar{t}\gamma$ process at a centre-of-mass energy of $\sqrt{s} = 14$ TeV [10], extending results performed using the approximation of stable top quarks [35]. A dedicated calculation at $\sqrt{s} = 13$ TeV has been performed for both single-lepton and dilepton channels, by the authors of Ref. [10]. The renormalization and factorization scales are both set to the top-quark mass, while the rest of the parameters are set to the same values used by the MG5\_aMC $t\bar{t}\gamma$ MC sample, as described in Sect. 3. These calculations are used to derive corrections at parton level to the normalization of the LO $t\bar{t}\gamma$ MC sample.

The $t\bar{t}\gamma$ calculation is performed at parton level in a phase space very close to the fiducial region defined in Sect. 7.1. The lepton (at least one lepton) is required to have $p_T > 25$ GeV for the single-lepton (dilepton) channel, and all leptons must have $|\eta| < 2.5$. The photon is required to have $p_T > 20$ GeV and $|\eta| < 2.37$. Jets are reconstructed from quarks and gluons using the anti-$k_t$ algorithm [36] with a radius parameter of $R = 0.4$, and they are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. At least four (two) jets are required for the single-lepton (dilepton) channel. All jets are required to be separated from the photon by $\Delta R(\gamma, \text{jet}) > 0.4$. Leptons are required to be separated from the photon by $\Delta R(\gamma, \ell) > 1.0$. The leptons in the dilepton channel are required to be separated from the jets by $\Delta R(\ell, \text{jet}) > 0.4$.

The LO cross-sections are calculated using the MG5\_aMC LO sample at particle level in the same phase space as above to derive the NLO $k$-factors for the single-lepton and dilepton channels. Since the kinematic properties of all the objects used in the NLO theoretical calculation are taken from parton level, the photons, leptons, and jets of the LO MC sample must be defined carefully to correspond to those at the parton level. This is achieved by requiring the photon and the leptons to be produced from the matrix element rather than
from the parton shower and adding the QED radiation simulated by the parton shower back to the leptons. The anti-$k_t$ algorithm with $R = 0.4$ is used for jet clustering, using all the final state particles, excluding the above photon, leptons, and their corresponding neutrinos.

The calculated NLO cross-sections are 120 fb and 31 fb for the $e+$jets and $e\mu$ channels, while the calculated LO cross-sections are 92 fb and 21 fb for the same channels, respectively, resulting in NLO $k$-factors of 1.30 and 1.44. These $k$-factors are applied to other single-lepton or dilepton channels. Statistical uncertainties are negligible. Systematic uncertainties of the $k$-factors receive contributions from two sources. For the NLO theoretical cross-section, the relative uncertainty due to the QCD scale and PDF choices are 14% (13%) for the single-lepton (dilepton) channel, with the QCD scale uncertainty dominating. For the LO MC cross-section, the non-perturbative effects in the parton shower model are studied by turning off the multiple parton interaction and hadronization of PYTHIA 8 separately, resulting in an uncertainty of 8% (4%) for the single-lepton (dilepton) channel. In addition, the jet cone size is varied from 0.4 to 0.3 or 0.5 separately to evaluate the impact of additional QCD radiation on the reconstruction of the particle-level jet. The resulting uncertainties are 11% (6%) for the single-lepton (dilepton) channel. Summing up the components in quadrature, the total relative uncertainty on the $k$-factor is 20% (15%) for the single-lepton (dilepton) channel.

5 Object and event selection

The object and event selection at the detector level are introduced in Sects. 5.1 and 5.2 respectively. The neural-network algorithms used in the analysis are described in Sect. 5.3. In Sect. 7.1, the fiducial region at particle level is defined.

5.1 Object selection

Electron candidates are reconstructed from energy deposits in the central region of the ECAL associated with reconstructed tracks from the ID [37] and are required to have a $p_T > 25$ GeV and an absolute calorimeter cluster pseudorapidity $|\eta_{\text{cluster}}| < 2.47$, excluding the transition region between the barrel and endcap calorimeters ($|\eta_{\text{cluster}}| \notin [1.37, 1.52]$). “Tight” likelihood-based identification criteria are applied, which correspond to an efficiency between 80% and 90% for electrons in different $p_T$ and $\eta$ ranges measured in $Z \rightarrow e\ell\gamma$ events [37]. Muon candidates are reconstructed by an algorithm that combines the track segments in the various layers of the MS with the tracks in the ID [38] and are required to have a $p_T > 25$ GeV and $|\eta| < 2.5$. “Medium” cut-based identification criteria are required, which correspond to an average efficiency around 96% in $t\bar{t}$ events for muons in different $p_T$ and $\eta$ ranges [38]. Isolation criteria are applied to both the electron and muon candidates using calorimeter- and track-based information to obtain 90% efficiency for leptons with $p_T = 25$ GeV, rising to 99% efficiency at $p_T = 60$ GeV in $Z \rightarrow e\ell\gamma$ events. The transverse impact parameter divided by its estimated uncertainty $|d_0|/\sigma(d_0)$ is required to be lower than five for electron candidates and three for muon candidates. The longitudinal impact parameter must satisfy $|z_0 \sin(\theta)| < 0.5$ mm for both. The lepton reconstruction and identification efficiencies in simulation are corrected to match the corresponding values in data [37,38].

A photon could convert into an electron positron pair when it traverses the material before entering the active volume of the ECAL. Photon candidates are reconstructed from energy deposits in the central region of the ECAL [39] and classified as unconverted if there is no matching track or reconstructed conversion vertex or as converted if there is a matching reconstructed conversion vertex or a matching track consistent with originating from a photon conversion. They must have a $p_T > 20$ GeV and $|\eta_{\text{cluster}}| < 2.37$, excluding the transition region between the barrel and endcap. “Tight” cut-based identification criteria, based on discriminating variables and corresponding to an efficiency around 85% at 40 GeV, are applied [39]. Cut-based $p_T$-dependent isolation criteria are applied using calorimeter- and track-based information and correspond to an efficiency between 75% and 90% for prompt photons (photons not from hadron decays) in $Z \rightarrow e\ell\gamma$ events. The photon reconstruction and identification efficiencies in simulation are corrected to match the corresponding values in data [39].

Jets are reconstructed using the anti-$k_t$ algorithm with a radius parameter of $R = 0.4$ from topological clusters of energy deposits in the calorimeter [40]. The jet energy scale and jet energy resolution are calibrated using energy- and $\eta$-dependent calibration schemes resulting from simulation and in situ corrections based on data [41]. The jets are required to have a $p_T > 25$ GeV and $|\eta| < 2.5$. Jets likely to originate from pile-up are suppressed by using the output of a multivariate jet-vertex-tagger (JVT) [42]. Scale factors are used to correct the selection efficiency in simulation to match data. Jets containing $b$-hadrons ($b$-jets) are identified with a $b$-tagging algorithm using a multivariate discriminant that combines information about secondary vertices and track impact parameters (MV2c10) [43,44]. The operating point used corresponds to an overall 77% $b$-tagging efficiency in $t\bar{t}$ events, with a corresponding rejection of $c$-jets (light-jets) by a factor of 6 (134). Efficiencies to tag $b$, $c$, and light-jets in the simulation are scaled by $p_T$- and $\eta$-dependent factors [43] to match the efficiencies in data.

The transverse energy carried by the neutrinos is accounted for in the reconstructed missing transverse momentum $E_T^{\text{miss}}$ [45], which is computed as the transverse component of the negative vector sum of all the selected electrons, muons, pho-
tons, and jets, as well as ID tracks associated with the primary vertex but not with any of the above objects, which is called track-based soft term.

An overlap removal procedure is applied to avoid the same calorimeter energy deposit or the same track being reconstructed as two different objects. Electrons sharing their track with a muon candidate are removed. Jets within a \( \Delta R = 0.2 \) cone of an electron are removed. After that, electrons within a \( \Delta R = 0.4 \) cone of a remaining jet are removed. When a muon and a jet are close, the jet is removed if it has no more than two associated tracks and is within \( \Delta R < 0.2 \) of the muon, otherwise the muon is removed if it is within \( \Delta R < 0.4 \) of the jet and the jet has more than two associated tracks. Photons within a \( \Delta R = 0.4 \) cone of a remaining electron or muon are removed. Finally, the jets within a \( \Delta R = 0.4 \) cone of a remaining photon are removed.

5.2 Event selection

The events must have at least one primary vertex with at least two associated tracks, each with \( p_T > 400 \) MeV. Primary vertices are formed from reconstructed tracks spatially compatible with the interaction region. The primary vertex with the highest sum of \( p_T^2 \) over all associated tracks is chosen. Events are categorized into the single-lepton channel if their final state contains exactly one lepton (electron or muon), and into the dilepton channel if they contain two leptons, two muons, or one electron and one muon, with each pair required to be of opposite charge. The lepton (at least one of the leptons) must be matched to a fired single-lepton trigger for the single-lepton (dilepton) channel. The \( p_T \) of the electron (muon) that fired the trigger has to be larger than 27 (27.5) GeV in order to match the higher lepton \( p_T \) trigger threshold in 2016. The selected events must have at least four (two) jets in the single-lepton (dilepton) channel, at least one of which is \( b \)-tagged, and exactly one photon. A Z-boson veto is applied in the single electron channel by excluding events with invariant mass of the system of the electron and the photon around the Z-boson mass \( (m(e, \gamma) - m(Z)) < 5 \) GeV), where \( m(Z) = 91.188 \) GeV. In the dilepton channel when the two leptons have the same flavour, events are excluded if the dilepton invariant mass or the invariant mass of the system of the two leptons and the photon is between 85 and 95 GeV, and \( E_T^{miss} \) is required to be larger than 30 GeV. The dilepton invariant mass is required to be higher than 15 GeV to suppress events from \( J/\psi, \Upsilon \) and \( \gamma^* \) decays. Finally, to suppress photons radiated from lepton(s), the \( \Delta R \) between the selected photon and lepton(s) must be greater than 1.0. The event selection is summarized in Table 1.

There are four types of backgrounds to the selected \( t\bar{t} \gamma \) candidates, three of which are events with a misidentified object. The contribution from events in which the selected photon candidate originates from a jet or a non-prompt photon from hadron decays, referred to as hadronic-fake background, is estimated following the method outlined in Sect. 6.1. The contribution from events in which the selected photon candidate originates from an electron, referred to as electron-fake background, is estimated following the method outlined in Sect. 6.2. The contribution from events in which the selected lepton candidate originates from a jet or a non-prompt lepton from heavy-flavour decays, referred to as fake-lepton background, is estimated following the method outlined in Sect. 6.3. Finally, the contribution from events with a prompt photon (excluding the \( t\bar{t} \gamma \) signal and the fake-lepton background with prompt photon radiation), referred to as prompt-photon background, is estimated following the method outlined in Sect. 6.4. In the single-lepton channel, the main backgrounds are from events with a hadronic-fake or electron-fake photon and \( W_\gamma \) production, while in the dilepton channel, \( Z \gamma \) production and events with a hadronic-fake photon are the dominant backgrounds.

### Table 1

Summary of the event selection. “OS” means the charges of the two leptons must have opposite signs.

<table>
<thead>
<tr>
<th>( e+\text{jets} )</th>
<th>( \mu+\text{jets} )</th>
<th>( ee )</th>
<th>( \mu\mu )</th>
<th>( e\mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary vertex</td>
<td>( 1 \ e )</td>
<td>( 1 \ \mu )</td>
<td>( 2 \ e, \text{OS} )</td>
<td>( 2 \ \mu, \text{OS} )</td>
</tr>
<tr>
<td>Trigger match</td>
<td>( \geq 4 \ fiber \text{jets} )</td>
<td>( \geq 2 \ fiber \text{jets} )</td>
<td>( \geq 1 \ b\text{-jet} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( m(e, \gamma) &gt; 5 ) GeV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( m(\ell, \gamma) \not\in [85,95] ) GeV</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( m(\ell, \gamma) \notin [85,95] ) GeV</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( E_T^{miss} &gt; 30 ) GeV</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( m(\ell, \gamma) &gt; 15 ) GeV</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \Delta R(\gamma, \ell) &gt; 1.0 )</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2  The observed data and the expected event yields for the signal and backgrounds in the single-lepton and dilepton channels. All data-driven corrections and systematic uncertainties are included. The signals are scaled by the NLO \(k\)-factors. The fake-lepton background in the dilepton channel is negligible, represented by a “-“. The Z\(\gamma\) (\(W\gamma\)) background in the single-lepton (dilepton) channel is included in “Other prompt”. The uncertainty of the \(W\gamma\) background in the single-lepton channel is not given since the normalization of this background is a free parameter in the likelihood fit.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Single lepton</th>
<th>Dilepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t\bar{t}\gamma)</td>
<td>(6,490 \pm 420)</td>
<td>(720 \pm 34)</td>
</tr>
<tr>
<td>Hadronic-fake</td>
<td>(1,440 \pm 290)</td>
<td>(49 \pm 27)</td>
</tr>
<tr>
<td>Electron-fake</td>
<td>(1,650 \pm 170)</td>
<td>(2 \pm 1)</td>
</tr>
<tr>
<td>Fake lepton</td>
<td>(360 \pm 200)</td>
<td>–</td>
</tr>
<tr>
<td>(W\gamma)</td>
<td>(1,130)</td>
<td>–</td>
</tr>
<tr>
<td>(Z\gamma)</td>
<td>–</td>
<td>(75 \pm 52)</td>
</tr>
<tr>
<td>Other prompt</td>
<td>(690 \pm 260)</td>
<td>(18 \pm 7)</td>
</tr>
<tr>
<td>Total</td>
<td>(11,750 \pm 710)</td>
<td>(863 \pm 78)</td>
</tr>
<tr>
<td>Data</td>
<td>(11,662)</td>
<td>(902)</td>
</tr>
</tbody>
</table>

A total number of 11,662 and 902 candidate events are selected for the single-lepton and dilepton channels, respectively, with expected numbers of 6490 ± 420 and 720 ± 34 signal events, where the corresponding NLO \(k\)-factors are applied and the uncertainties include the simulation statistical uncertainty and all systematic uncertainties introduced in Sect. 8. Table 2 summarizes the observed data and the expected event yields for the signal and background processes. Figures 1 and 2 show comparisons of the data with the expected simulated distributions. The simulation is corrected with data-driven corrections. The statistical uncertainty of data and systematic uncertainties are included. The signals are scaled by the NLO \(k\)-factors.

5.3 Multivariate analysis

To discriminate the \(t\bar{t}\gamma\) signal from backgrounds, a neural-network algorithm, called the event-level discriminator (ELD), is trained separately for the single-lepton and dilepton channels. Given the significant contribution of hadronic-fake photons in the single-lepton channel, a dedicated neural network, referred to as the prompt-photon tagger (PPT) in the following, is trained to discriminate between prompt photons and hadronic-fake photons. The PPT is used as one of the inputs to the ELD in the single-lepton channel.

Both neural-network algorithms are feedforward binary classifiers that have been trained using Keras [46] and evaluated using LWtnn [47]. Theano [48] is used as backend. The input variables are normalized to have a standard deviation of 1 and a mean of 0. To reduce the risk of over-training, regularization methods such as dropout [49] and batch normalization [50] layers are used. Additionally, \(k\)-fold cross-validation is performed.

Five variables which characterize the photon candidate shower shape in the transverse and lateral directions utilizing the energy deposits in the first and second layer of the ECAL, \(R_{\eta}\), \(R_{\phi}\), \(w_{22}\), \(w_{13}\), and \(F_{side}\), and one variable which characterizes the energy leakage fraction into the HCAL, \(R_{\text{had}}\), are used in the PPT. These are the standard discriminating variables used in ATLAS for photon identification [39] and their definitions are given in the Appendix. Prompt photons from simulated QCD-Compton processes and hadronic-fake photons from simulated dijet events are used as signal and background photons in the training and testing of the PPT. Photons are required to pass the Tight identification and have \(p_T > 25\) GeV and \(|\eta_{\text{clus}}| < 2.37\), excluding the calorimeter transition region. The PPT shape of the prompt photons in simulation is corrected to match data in photon \(p_T-\eta\) bins. The correction factors for each bin are extracted from the ratio between the PPT output distribution in data and that of simulation, using photons in a \(Z \rightarrow \ell\ell\gamma\) control region. The control region is defined by requiring exactly one photon and two opposite-sign leptons, with the invariant mass of the lepton pair between 60 and 100 GeV. The resulting correction factors range from 0.5 to 2.0 and are in general around unity. PPT systematic uncertainties are evaluated separately for prompt photons and fake photons and are discussed in Sect. 8.3. The PPT output distribution after event selection in the single-lepton channel is shown in Fig. 3. The shape difference between data and prediction of the PPT is caused by the shape difference between data and simulation of the input discriminating variables and is covered by the assigned systematic uncertainties.

Simulated signal and background events passing the event selection are used for the training and testing of the ELD, except for the fake-lepton background in the single-lepton channel which is taken from data as described in Sect. 6.3. In the dilepton channel, the selection criteria on the \(E_T^{\text{miss}}\), the invariant masses, and the jet multiplicity are removed to increase the sample size for training. The training takes 15 (7) variables as input for the single-lepton (dilepton) channel, which are summarized in Table 3. The \(b\)-tagging related variables are important for the ELD training in both channels, because of their discriminating power against background without real heavy flavour jets. which have significant contributions. The use of the PTT as input to ELD improves the discrimination power against hadronic-fake background.

Variables like the dilepton invariant mass and missing transverse energy are useful for the ELD training in the dilepton channel, due to the dominant background of \(Z\gamma\). The distributions of the ELD after event selection are shown in Fig. 4 for the single-lepton and dilepton channels. The shapes of the ELD are compared between signal and total background in Fig. 5 for the single-lepton and dilepton channels.
In the single-lepton channel, the kinematic properties and jet flavour compositions are similar between the $t\bar{t}\gamma$ signal and the dominating background, which is $t\bar{t}$ production with a hadronic-fake or electron-fake photon. In the dilepton channel, this is not the case since $Z\gamma$ production is dominant. Thus the ELD is more discriminating in the dilepton channel than in the single-lepton channel. The ELD is used in the likelihood fit to data to extract the fiducial cross-sections.

6 Background estimation

6.1 Hadronic-fake background

The hadronic-fake background is an important background in this analysis. Its main source is the $t\bar{t}$ process, where one of the final state jets is reconstructed and identified as a photon. In addition, there are small contributions from $W+$jets and single top processes to the single-lepton channel and from $Z+$jets events to the dilepton channel.
Fig. 2  Distributions of the a photon $p_T$, b photon $|\eta|$, c minimum $\Delta R(\gamma, \ell)$, d $|\Delta \eta(\ell, \ell)|$, and e $\Delta \phi(\ell, \ell)$ in the dilepton channel after event selection and before likelihood fit. All data-driven corrections and systematic uncertainties are included. Overflow events are included in the last bin. In particular, events with $|\Delta \eta(\ell, \ell)| > 2.5$ are included in the last bin of d.
The hadronic-fake background is estimated using all the simulation samples by requiring that the selected photon is a hadron or from a hadronic decay at generator level. A data-driven method, called the ABCD method, is applied to derive a set of scale factors, based on the ratio of hadronic-fake back-

ground estimated by the method over the one from simulation. This set of scale factors is derived in the single-lepton channel and applied to both the single-lepton and dilepton channels to calibrate the simulation to match data.

In the ABCD method, the isolation selection and part of the Tight identification criteria of the photon, which are assumed to be uncorrelated, are inverted to define three regions enriched with hadronic-fake photons. These regions are orthogonal to one another, and to the signal region. Region A uses photons that pass the isolation selection defined in Sect. 5.1 but fail at least two out of the four identification requirements on the discriminating variables $F_{side}$, $w_3$, $\Delta E$, and $E_{ratio}$ (defined in the Appendix), while passing all other Tight identification criteria. These four variables describe the shower shape in the first layer of ECAL and are chosen for their small correlation with the photon isolation but strong discrimination power between prompt and hadronic-fake photons. Region B uses photons that fail the identification criteria as in region A and do not pass the isolation selection. Additionally, the sum of the $p_T$ of all tracks within $\Delta R = 0.2$ around the photon is required to be larger than 3 GeV to further suppress the prompt-photon contribution. Region C selects photons that fail the isolation requirements as in region B but pass the Tight identification. Region D is the signal region.

The hadronic-fake background in the signal region can be expressed as:

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

Table 3 Input variables for the event-level discriminator for the single-lepton and dilepton channels. For events without the 5th jet, the $p_T(j_5)$ is set to zero

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Single lepton</th>
<th>Dilepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPT</td>
<td>Prompt-photon tagger output</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$H_T$</td>
<td>Scalar sum of the $p_T$ of the leptons and jets</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$m(\gamma, \ell)$</td>
<td>Invariant mass of the system of the photon and the lepton</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$E_T^{miss}$</td>
<td>Missing transverse energy</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_T^W$</td>
<td>Reconstructed transverse mass of the leptonically decaying $W$-boson</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$N_{jets}$</td>
<td>Jet multiplicity</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$p_T(j_1)$</td>
<td>$p_T$ of the leading jet (ordered in $p_T$)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$p_T(j_2)$</td>
<td>$p_T$ of the sub-leading jet</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$p_T(j_3)$</td>
<td>$p_T$ of the third jet</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$p_T(j_4)$</td>
<td>$p_T$ of the fourth jet</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$p_T(j_5)$</td>
<td>$p_T$ of the fifth jet</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$N_{B-jets}$</td>
<td>$b$-jet multiplicity</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$b_1(j)$</td>
<td>highest $b$-tagging score of all jets</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$b_2(j)$</td>
<td>second highest $b$-tagging score of all jets</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$b_3(j)$</td>
<td>third highest $b$-tagging score of all jets</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$m(\ell, \ell)$</td>
<td>Invariant mass of the system of the two leptons</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Fig. 4 Distributions of the ELD for the a single-lepton and b dilepton channels after event selection and before likelihood fit. All data-driven corrections and systematic uncertainties are included.

Fig. 5 Comparison of the shape of the ELD between signal and total background in the a single-lepton and b dilepton channels after event selection. All data-driven corrections are included.

\[
N_{h\text{-fake}}^{D, \text{est.}} = \frac{N_{h\text{-fake}}^{A, \text{data}} \times N_{h\text{-fake}}^{C, \text{data}}}{N_{h\text{-fake}}^{B, \text{data}}} \times \theta_{MC},
\]

\[
\theta_{MC} = \frac{N_{h\text{-fake}}^{D, \text{MC}}/N_{h\text{-fake}}^{A, \text{MC}}}{N_{h\text{-fake}}^{D, \text{MC}}/N_{h\text{-fake}}^{B, \text{MC}}},
\]

where \(N_{h\text{-fake}}^{A, \text{data}}, N_{h\text{-fake}}^{B, \text{data}}, \) and \(N_{h\text{-fake}}^{C, \text{data}}\) are the numbers of hadronic-fake events in regions A, B, and C, estimated by subtracting the events with prompt photons and other backgrounds from the number of data events in these regions, and \(N_{h\text{-fake}}^{A, \text{MC}}, N_{h\text{-fake}}^{B, \text{MC}}, N_{h\text{-fake}}^{C, \text{MC}}\) and \(N_{h\text{-fake}}^{D, \text{MC}}\) are the numbers of hadronic-fake events predicted by simulation in regions A, B, C, and D. The factor \(\theta_{MC}\) corrects for possible bias caused by residual correlation between the isolation variables and the four discriminating variables used to define the regions.

The ABCD method is applied to photons in different \(p_T-\eta\) bins, separately for converted and unconverted photons. The resulting scale factors range from 0.8 to 3.2, with typical
values around 1.5 and large statistical and systematic uncertainties of more than 0.5. These scale factors are applied to the MC-based hadronic-fake background prediction.

6.2 Electron-fake background

Dilepton events where an electron is mis-identified as a photon contribute to the electron-fake background in the single-lepton channel. Its main source is the $t \bar{t}$ dileptonic decay process. When the lepton is an electron, there is also some contribution from the $Z \rightarrow ee$ process. A data-driven method is applied to derive a set of scale factors to correct the electron-to-photon fake rate in simulation to match data.

The electron-to-photon fake rate is measured with a tag-and-probe method using two control regions (CR), exploiting the $Z \rightarrow ee$ process. For the first CR, $Z \rightarrow ee$ event candidates are selected in which one of the electrons fulfills the photon selection criteria, i.e. contributes to the electron-fake background, and the electron-photon pair should have an invariant mass in the range [40, 140] GeV and an opening angle greater than 2.62 rad. The electron is called the tag electron and the electron-fake photon candidate is referred to as probe photon. Non-$Z$-boson backgrounds are subtracted with a sideband fit of the invariant mass distribution. The fitted signal contains $Z \rightarrow ee\gamma$ contributions with one of the electrons not reconstructed or identified which is subtracted using simulation. For the second CR, events with an electron-positron pair satisfying the same requirements as in the first CR are selected and the same procedure is applied. The fake rate is calculated as the ratio of the number of probe photons over the number of probe electrons. To avoid a trigger bias, the tag electron in both CRs must match the single-lepton selection. The fake-lepton efficiency is measured using data and simulation. The fake-lepton efficiency is extracted from a tag-and-probe technique using leptons from $Z$ boson decays. The efficiencies are parametrized as a function of the lepton $\eta$ and $m_{T\gamma}$ (the lepton $p_T$ and $m_{T\gamma}$ when the lepton is an electron (muon)).

In the dilepton channel, the contamination of background processes with at least one fake lepton is estimated by selecting same-sign dilepton events in data, after subtracting events with two same-sign prompt leptons, using simulation. The fake-lepton background in the dilepton channel is found to be negligible.

6.3 Fake-lepton background

In the single-lepton channel, the fake-lepton background is dominated by multi-jet processes with an additional photon which could either be a prompt or a fake photon. It is estimated directly from data by using a matrix method [51]. The number of background events in the signal region is evaluated by applying efficiency factors (fake lepton and real lepton efficiencies) to the number of events satisfying a tight (identical to the signal selection) as well as a looser lepton selection. The fake-lepton efficiency is measured using data in control regions dominated by multi-jet background with the real lepton contribution subtracted using simulation. The real lepton efficiency is extracted from a tag-and-probe technique using leptons from $Z$ boson decays. The efficiencies are parametrized as a function of the lepton $\eta$ and $m_{T\gamma}$ (the lepton $p_T$ and $m_{T\gamma}$ when the lepton is an electron (muon)).

6.4 Prompt-photon background

All background processes to $t\bar{t}$ production are also background to $t\bar{t}\gamma$ production when accompanied by prompt-photon radiation. These processes include $W\gamma$, $Z\gamma$, and associated production of a photon in single top, diboson, and $t\bar{t}V$ productions. In the single-lepton channel, $W\gamma$ is the dominant prompt-photon background, and $Z\gamma$ and the others are grouped together as “Other prompt”. In the dilepton channel, $Z\gamma$ is dominant, with all others grouped as “Other prompt”. Background from $t\bar{t}$ production with a photon produced in an additional $pp$ interaction in the same bunch crossing has been studied and is found to be negligible.

Validation regions are selected to check the modelling of $W\gamma$ in the single-lepton channel and $Z\gamma$ in the dilepton channel. The $Z\gamma$ validation region is selected by requiring exactly one $b$-tagged jet and the invariant mass of the system of the two leptons in a mass window of [60, 100] GeV. The $W\gamma$ validation region is selected with the same event selection as for the signal region of the single-lepton channel, with the following modifications: the number of jets must be either two or three; exactly one $b$-jet is required; the $E_T^{miss}$ is required to be larger than 40 GeV; the ELD value must be smaller than 0.04; and the invariant mass of the system of the lepton and photon is required to be smaller than 80 GeV if the lepton is an electron. The modelling of $W\gamma$ is also checked in a light-flavour validation region requiring zero $b$-jet and without the ELD cut.
The normalization of the $W\gamma$ background is treated as a free parameter of the likelihood fit in the single-lepton channel, since this background is well separated from the $t\bar{t}\gamma$ signal by the ELD and the uncertainty of its theoretical prediction is large. The shape of $W\gamma$ is taken from simulation and checked in the validation region to ensure good modelling. The normalization and shape of the $Z\gamma$ background in the dilepton channel as well as other prompt backgrounds in both channels are predicted by simulation.

7 Analysis strategy

The analysis is performed in two parts, one being the measurement of the fiducial cross-section and the other the measurement of normalized differential cross-sections in the same fiducial region. Both parts share the same strategy for the estimation of backgrounds and systematic uncertainties. In the fiducial cross-section measurement, the ELD is fitted and the post-fit background yields and systematic uncertainties are used. In the normalized differential cross-section measurements, no fit is performed, except for the determination of the $W\gamma$ contribution in the single-lepton channel, where a systematics-free ELD fit is performed.

7.1 Fiducial region definition

The fiducial region of the analysis is defined at particle level in a way that mimics the event selection in Sect. 5.2. Leptons must have $p_T > 25$ GeV and $|\eta| < 2.5$ and must not originate from hadron decays. Photons not from hadron decays and in a $\Delta R = 0.1$ cone around a lepton are added to the lepton before the lepton selection. Photons are required to have $p_T > 20$ GeV and $|\eta| < 2.37$ and must not originate from hadron decays or be used for lepton dressing. The photon isolation computed from the ratio of the scalar sum of the transverse momentum of all stable\(^2\) charged particles around the photon over its transverse momentum must be smaller than 0.1. Jets are clustered using the anti-$k_T$ algorithm with $R = 0.4$ using all final state particles excluding non-interacting particles and muons that are not from hadron decays. Jets must have $p_T > 25$ GeV and $|\eta| < 2.5$. A ghost matching method [52] is used to determine the flavour of the jets, with those matched to $b$-hadrons tagged as $b$-jets. A simple overlap removal is performed: jets within $\Delta R < 0.4$ of a selected lepton or photon are removed. For events in the single-lepton (dilepton) channel, exactly one photon and exactly one lepton (two leptons) are required. At least four (two) jets are required with at least one of them $b$-tagged. Events are rejected if there is any lepton and photon pair satisfying $\Delta R(\gamma, \ell) < 1.0$. The acceptance for the generated signal events to pass the fiducial selection of the single-lepton (dilepton) channel is 8.2% (0.96%).

7.2 Fiducial cross-section

The fiducial cross-section is extracted using a profile likelihood fit to the ELD distribution. The parameter of interest, the fiducial cross-section $\sigma_{\text{fid}}$, is related to the number of signal events in bin $i$ of the ELD as

$$N_i^f = L \times \sigma_{\text{fid}} \times C \times f_i^{\text{ELD}},$$

where $L$ is the integrated luminosity, $C$ is the correction factor for the signal efficiency and for migration into the fiducial region, and $f_i^{\text{ELD}}$ is the fraction of signal events falling into bin $i$ of the ELD. The correction factor $C$ is defined as $N_i^{\text{sel,MC}} / N_i^{\text{sel,MC},\gamma}$, where $N_i^{\text{sel,MC}}$ is the simulated number of signal events passing the event selection described in Sect. 5.2 and $N_i^{\text{sel,MC},\gamma}$ is the corresponding number of signal events generated in the fiducial region defined in Sect. 7.1. The value of $C$ is 0.36 (0.30) for the single-lepton (dilepton) channel, with negligible statistical uncertainty.

A likelihood function is defined from the product over all bins of the ELD distribution:

$$L = \prod_i P\left(N_i^{\text{obs}}|N_i^f(\theta) + \sum_b N_i^b(\theta)\right) \times \prod_t G(0|\theta_t, 1),$$

where $N_i^{\text{obs}}$, $N_i^f$, and $N_i^b$ are the observed number of events in data, the predicted number of signal events, and the estimated number of background events in bin $i$ of the ELD, which form a Poisson term $P$ in that bin. Nuisance parameter $\theta_t$ is to parameterize a systematic uncertainty $t$, which is constrained by a Gaussian $G(0|\theta_t, 1)$, so that when it changes from zero to ±1, the quantities affected by this systematics in the likelihood change by ±1 standard deviation. The collection of all the systematic uncertainties is denoted as $\theta$. For systematic uncertainties related to the finite number of MC events, the Gaussian terms in the likelihood are replaced by Poisson terms. Each systematic uncertainty affects $N_i^f$ and $N_i^b$ in each bin of the ELD. The cross-section is measured by profiling the nuisance parameters and maximizing this likelihood.

7.3 Normalized differential cross-sections

An unfolding procedure is applied to the observed detector-level distribution of a given observable, with backgrounds subtracted, to derive the true distribution of the signal at particle level, from which the differential cross-section as a function of the observable is calculated. The differential cross-section is normalized to unity.
The differential cross-section is given by

\[ \sigma_k = \frac{1}{L} \frac{1}{\epsilon_k} \sum_j M^{-1}_{jk} \times (N^\text{obs}_j - N^b_j) \times (1 - f_{\text{out},j}) . \]

The indices \( j \) and \( k \) indicate the bin of the observable at detector and particle levels, respectively. The variables \( N^\text{obs}_j \) and \( N^b_j \) are the number of observed events and of estimated background events in bin \( j \) at detector level, respectively. The efficiency \( \epsilon_k \) is the fraction of signal events generated at particle level in bin \( k \) of the fiducial region to be reconstructed and selected at detector level and have the objects, that are used to define the observable to be unfolded, matched between reconstruction and particle-levels with \( \Delta R < 0.1 \).

The migration matrix \( M_{kj} \) expresses the probability for an event in bin \( k \) at particle level to end up in bin \( j \) at detector level, calculated from events passing both the fiducial region selection and the event selection, as well as the above matching procedure. The outside-migration fraction \( f_{\text{out},j} \) is the fraction of signal events generated outside the fiducial region but reconstructed and selected in bin \( j \) at detector level or events failing the above matching. The signal MC sample is used to determine \( \epsilon_k \), \( f_{\text{out},j} \), and \( M_{kj} \), the values of which are illustrated in Fig. 6, using the photon \( p_T \) in the single-lepton channel as an example. The normalization and the corresponding uncertainty of the \( W\gamma \) contribution in the single-lepton channel are taken from the likelihood fit introduced in Sect. 7.2 but without systematic uncertainties included. The normalized differential cross-section is

\[ \sigma_k^{\text{norm}} = \frac{\sigma_k}{\sum_k \sigma_k} , \]

where the sum is over all the bins of the observable.

The inversion of the migration matrix \( M_{kj} \) is approximated using the iterative Bayesian method [53] implemented in the \textsc{RooUnfold} package [54]. The method relies on the Bayesian probability formula to invert the migration matrix, starting from a given prior of the particle-level distribution and iteratively updating it with the posterior distribution. The binning choices of the unfolded observables take into account the detector resolution and the expected statistical uncertainty, with the latter being the dominating factor. Three iterations are chosen which give a good convergence of the unfolded distribution and a statistically stable result. Tests are performed, using simulation, to verify that the unfolding procedure does not bias the results while the estimated uncertainties are still reasonable. The results are cross-checked with other unfolding methods, which give consistent results.

The chosen observables to unfold are the photon \( p_T \) and \(|\eta|\) and the \( \Delta R \) between the photon and the closest lepton for both single-lepton and dilepton channels and the \( \Delta \phi \) and \(|\Delta \eta|\) between the two leptons for the dilepton channel. These are all lepton or photon observables, therefore the migration matrices are almost diagonal, making the unfolding simple and converging fast. The kinematic properties of the photon are sensitive to the \( t\gamma \) coupling, while the dilepton \( \Delta \phi \) is sensitive to the \( tt \) spin correlation. The normalized differential cross-sections are measured, since the overall signal normalization is given by the measured fiducial cross-section.

### 8 Systematic uncertainties

Signal and background modelling and experimental uncertainties in the analysis are described in this section, as well as the PPT systematic uncertainty. They affect the normaliza-
tion of signal and background and/or the shape of their corresponding distributions, such as the ELD and the observables to be unfolded. Each of the signal and background modelling uncertainties is correlated between different channels for the relevant signal or background process. Each of the experimental uncertainties is correlated between signal and simulated backgrounds and between different channels. The PPT systematic uncertainty is separately studied for the prompt, electron-fake, and hadronic-fake photons. Table 5 gives a summary of these uncertainties and their impact to the fiducial cross-section measurements.

8.1 Signal modelling uncertainties

The signal modelling uncertainties include the uncertainties due to the choice of the QCD scales, the parton shower, the amount of ISR and FSR, and the PDF set. Their effects on the corrections defined in Sect. 7 (both for the fiducial and normalized differential cross-section measurements) as well as on the shape of the ELD distributions are evaluated.

To study the QCD scale uncertainty, the renormalization and factorization scales are varied up and down by a factor of two from their nominal choices independently or simultaneously. The largest variation of the corrections or the shapes is assigned as the uncertainty. To evaluate the parton shower uncertainty, PYTHIA 8 and HERWIG 7 both interfaced to MG5_aMC are compared. The ISR/FSR uncertainty is studied by comparing the variations of the A14 tune parameters of PYTHIA 8 with its nominal values. The PDF uncertainty is evaluated using the standard deviation of the distribution formed by the 100 eigenvector set of the NNPDF set [23].

8.2 Background modelling uncertainties

The systematic uncertainties on the hadronic-fake background due to background subtraction in the hadronic-fake control regions A, B, and C are estimated by varying up and down the signal by 100%, the other MC-based backgrounds by 50%, and the other data-driven backgrounds by their estimated uncertainties, separately. The statistical uncertainties in the three data control regions are also considered. Systematic uncertainties arising from the correction factor $\theta_{MC}$, the extrapolation of the hadronic-fake scale factors, and the shapes of the distributions of the ELD and the observables to be unfolded are estimated using $t\bar{t}$ samples rather than all of the simulated hadronic-fake samples, since $t\bar{t}$ is the dominant source of the hadronic-fake background. The uncertainty due to the rate of additional QCD radiation is estimated by comparing the samples with enhanced/reduced parton shower radiation as described in Sect. 3 with the nominal sample, and the uncertainty due to the modelling of the generator and parton shower is estimated by comparing POWHEG+PYTHIA 8 with SHERPA.

The systematic uncertainties on the electron-fake background mainly come from the sideband fit when measuring the fake rate, which is estimated by varying the fit parameters within their uncertainties. The uncertainty due to $Z \to ee\gamma$ subtraction is also considered by replacing the $Z\gamma$ MC sample by the $Z$+jets sample, where the photon radiation is described by the parton shower. Uncertainties of the shapes of the distributions are evaluated using $t\bar{t}$ systematic variations MC samples as for the hadronic-fake background. In the dilepton channel where the electron-fake background is very small, a 50% uncertainty is assumed to cover a possible mis-modelling in the estimate.

For evaluating the systematic uncertainty of the fake-lepton background in the single-lepton channel, several alternative parameterizations of the real and fake efficiencies of the matrix method are studied and two predicting larger and smaller yields are selected as up and down variations. The lepton $\eta$, $b$-jet multiplicity, and $m_T^W$ (the lepton $\eta$, minimum $\Delta R$ between the lepton and the closest jet, and the jet $p_T$ parameterization is used as up (down) variation when the lepton is an electron. The jet $p_T$ and $b$-jet multiplicity (the lepton $p_T$, $\eta$, and minimum $\Delta R$ between the lepton and the closest jet) parameterization is used as up (down) variation when the lepton is a muon. The resulting uncertainties are around 50%, and given that this background contribution is relatively small, no additional systematic uncertainties are considered.

The uncertainty on the $V\gamma$ background shape is studied by varying the renormalization and factorization scales up and down by a factor of two from their nominal values independently or simultaneously and then choosing the maximum shape distortions as the QCD scale uncertainties. For the $Z\gamma$ background in the dilepton channel, SHERPA is compared with MG5_aMC+PYTHIA 8 to evaluate the shape uncertainty due to the choice of generator and parton shower. No shape modelling uncertainty is assigned to the other small prompt backgrounds. Apart from the $W\gamma$ background in the single-lepton channel whose normalization is a free parameter in the likelihood fit, a normalization uncertainty of 50% is assigned to each source of the prompt-photon background, included in Table 2 as “$Z\gamma$” and “Other prompt”.

8.3 Experimental uncertainties

Experimental systematic uncertainties affect the normalization and shape of the simulated signal and background samples. For MC-based backgrounds calibrated to data using data-driven techniques, only the shape variations are considered.

The photon identification and isolation efficiencies as well as the efficiencies of the lepton reconstruction, identification,
Fig. 7 The post-fit ELD distributions for the a single-lepton and b dilepton channels. All the systematic uncertainties are included.

Table 4 The observed data and post-fit event yields for the signal and backgrounds in the single-lepton and dilepton channels. All data-driven corrections and systematic uncertainties are included. The fake-lepton background in the dilepton channel is negligible, represented by a "-". The $Z\gamma$ ($W\gamma$) background in the single-lepton (dilepton) channel is included in "Other prompt".

<table>
<thead>
<tr>
<th>Channel</th>
<th>Single lepton</th>
<th>Dilepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}\gamma$</td>
<td>7040 ± 350</td>
<td>780 ± 44</td>
</tr>
<tr>
<td>Hadronic-fake</td>
<td>1470 ± 180</td>
<td>49 ± 26</td>
</tr>
<tr>
<td>Electron-fake</td>
<td>1620 ± 160</td>
<td>2 ± 1</td>
</tr>
<tr>
<td>Fake lepton</td>
<td>186 ± 68</td>
<td>–</td>
</tr>
<tr>
<td>$W\gamma$</td>
<td>900 ± 370</td>
<td></td>
</tr>
<tr>
<td>$Z\gamma$</td>
<td>55 ± 29</td>
<td></td>
</tr>
<tr>
<td>Other prompt</td>
<td>570 ± 180</td>
<td>18 ± 7</td>
</tr>
<tr>
<td>Total</td>
<td>11 790 ± 180</td>
<td>906 ± 38</td>
</tr>
<tr>
<td>Data</td>
<td>11 662</td>
<td>902</td>
</tr>
</tbody>
</table>
The measured fiducial cross-sections normalized to their corresponding NLO SM predictions [10] for the five individual channels and for the single-lepton and dilepton channels, as well as the combination of all channels. The statistical uncertainties are the inner error bars, while the total uncertainties are the outer error bars. The NLO prediction for the inclusive fiducial cross-section is represented by the dashed vertical line, and the theoretical uncertainties are represented by the shaded bands.

The uncertainty in the combined 2015+2016 integrated luminosity is 2.1%. It is derived, following a methodology similar to that detailed in Ref. [62], and using the LUCID-2 detector for the baseline luminosity measurements [63], from calibration of the luminosity scale using x-y beam-separation scans.

The uncertainty associated to the modelling of pile-up in the simulation is assessed by varying the reweighting of the pile-up in the simulation within its uncertainties.

8.4 Systematic uncertainties of the measured differential cross-section

Systematic uncertainties for unfolding arise from the detector response description, signal modelling, and background modelling. The systematic uncertainties due to background modelling and the detector response are evaluated by varying the input detector-level pre-fit distributions, unfolding them with corrections based on the nominal signal sample, and calculating the difference of the resulting unfolded distributions with respect to the nominal one. The systematic uncertainties due to signal modelling are evaluated by varying the signal corrections, i.e. the migration matrix $M_{kj}$, the efficiency $\epsilon_k$ and the fraction $f_{\text{out}, j}$ as defined in Sect. 7.3, with which the nominal input detector-level pre-fit distributions are unfolded, and calculating the difference of the resulting unfolded distributions with respect to the nominal one. The statistical uncertainties of the signal and background MC samples are also considered. The covariance matrix $C_{ij}$ for each of these systematic uncertainties is estimated as $\sigma_i \times \sigma_j$, where $\sigma_i$ and $\sigma_j$ are the symmetrized uncertainties for bin $i$ and bin $j$ of the unfolded distribution. The covariance matrix for the statistical uncertainty of data is calculated by the unfolding algorithm [54].

### Table 5

Summary of the effects of the groups of systematic uncertainties on the fiducial cross-section in the single-lepton and dilepton channels. Due to rounding effects and small correlations between the different sources of uncertainty, the total systematic uncertainty is different from the sum in quadrature of the individual sources.

<table>
<thead>
<tr>
<th>Source</th>
<th>Single lepton (%)</th>
<th>Dilepton (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal modelling</td>
<td>± 1.6</td>
<td>± 2.9</td>
</tr>
<tr>
<td>Background modelling</td>
<td>± 4.8</td>
<td>± 2.9</td>
</tr>
<tr>
<td>Photon</td>
<td>± 1.1</td>
<td>± 1.1</td>
</tr>
<tr>
<td>Prompt-photon tagger</td>
<td>± 4.0</td>
<td>–</td>
</tr>
<tr>
<td>Leptons</td>
<td>± 0.3</td>
<td>± 1.3</td>
</tr>
<tr>
<td>Jets</td>
<td>± 5.4</td>
<td>± 2.0</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>± 0.9</td>
<td>± 0.4</td>
</tr>
<tr>
<td>Pile-up</td>
<td>± 2.0</td>
<td>± 2.3</td>
</tr>
<tr>
<td>Luminosity</td>
<td>± 2.3</td>
<td>± 2.3</td>
</tr>
<tr>
<td>MC sample size</td>
<td>± 1.9</td>
<td>± 1.7</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>± 7.9</td>
<td>± 5.8</td>
</tr>
<tr>
<td>Data sample size</td>
<td>± 1.5</td>
<td>± 3.8</td>
</tr>
<tr>
<td>Total uncertainty</td>
<td>± 8.1</td>
<td>± 7.0</td>
</tr>
</tbody>
</table>
PDF

9 Results

9.1 Fiducial cross-sections

The fiducial cross-section is extracted via a binned maximum likelihood fit to the ELD distribution in data as described in Sect. 7.2. The measured cross-sections are

\[ \sigma_{\text{fid}}^{\text{SL}} = 521 \pm 9 \, (\text{stat.}) \pm 41 \, (\text{sys.}) \, \text{fb} \]

\[ \sigma_{\text{fid}}^{\text{DL}} = 69 \pm 3 \, (\text{stat.}) \pm 4 \, (\text{sys.}) \, \text{fb} \]

for the single-lepton and dilepton channels, respectively, and agree well within uncertainties with the corresponding predicted cross-sections of 495 \pm 99 \, \text{fb} and 63 \pm 9 \, \text{fb}. The ELD distributions after the fit (post-fit) are shown in Fig. 7 for the single-lepton and dilepton channels. The corresponding event yields are summarized in Table 4, including all the
Fig. 10 The normalized differential cross-sections as a function of the a photon $p_T$, b photon $|\eta|$, c minimum $\Delta R(\gamma, \ell)$, d $|\Delta \eta(\ell, \ell)|$, and e $\Delta \phi(\ell, \ell)$ in the dilepton channel. The unfolded distributions are compared to the predictions of the MG5_aMC+PYTHIA 8 together with the up and down variations of the PYTHIA 8 A14 tune parameters, the MG5_aMC+Herwig 7, and the POWHEG+PYTHIA 8 $t\bar{t}$ where photon radiation is modelled in the parton shower. The top ratio-panel shows the ratios of all the predictions over data. The bottom ratio-panel shows the ratios of the alternative predictions and data over the nominal prediction. Overflows are included in the last bin.
systematic uncertainties. Compared to pre-fit (Table 2), the signal event yields are higher, while some of the background event yields are lower. Some of the systematic uncertainties are moderately constrained, e.g. the parton shower uncertainty of the $t\bar{t}\gamma$ and $t\bar{t}$ modelling and the PPT shape uncertainty of prompt photons.

The fiducial cross-sections in each of the individual channels ($e+\text{jets}$, $\mu+\text{jets}$, ee, $e\mu$, and $\mu\mu$) as well as a combined single-lepton and dilepton cross-section are also measured. The former are measured by fitting the ELD distribution in each individual channel separately, while the latter is measured by fitting them simultaneously, sharing the same signal strength parameter $\mu = \sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}^{\text{NLO}}$, which scales coherently the fiducial cross-sections of each channel. A comparison of all measurements with the predictions is shown in Fig. 8.

All the systematic uncertainties introduced in Sect. 8 are grouped into a smaller set of classes and summarized in Table 5 for the single-lepton and dilepton channels. The effect of each group of uncertainties is calculated from the quadratic difference between the relative uncertainty in the measured fiducial cross-section with this group of uncertainties included or excluded from the fit with corresponding nuisance parameters fixed to their fitted values. In the single-lepton channel, the jet-related and background modelling systematic uncertainties are dominant, followed by the PPT and signal modelling systematic uncertainties. In the dilepton channel, the data statistical uncertainty is the leading contribution, followed by the signal and background modelling systematic uncertainties. The luminosity and pile-up uncertainties are also important in this channel.

9.2 Normalized differential cross-sections

The normalized differential cross-sections are shown in Figs. 9 and 10, for the single-lepton and dilepton channels, respectively. They are compared to the nominal $t\bar{t}\gamma$ sample (MG5_AMC+Pythia 8) and the samples with variations of the Pythia 8 A14 tune parameters and the alternative parton shower model of MG5_AMC+Herwig 7. In addition, a comparison with the nominal $t\bar{t}$ Powheg+Pythia 8 MC sample where prompt-photon radiation is modelled in the parton shower is included. All $t\bar{t}\gamma$ samples predict very similar shapes and describe the data well. A small deviation from the prediction is observed in the dilepton $\Delta R(\gamma, \ell)$ distribution, where the leptons in the prediction are more back-to-back than in data. The deviation of data from the prediction is 1.5 standard
deviations, based on the $\chi^2$ calculated according to the procedure described in what follows. The POWHEG+PYTHIA 8 $t\bar{t}$ sample gives an improved agreement with data compared to the nominal and varied $t\bar{t}\gamma$ samples, although the overall agreement is still poor. It can also be seen from Figs. 9(a) and 10(a) that the photons generated by PYTHIA 8 have a softer $p_T$ spectrum than in data.

The systematic uncertainties of the unfolded distributions are decomposed into the signal modelling uncertainty, experimental uncertainty, and background modelling uncertainty in both channels. In the single-lepton (dilepton) channel, the background modelling uncertainty is split into $t\bar{t}$ ($Z\gamma$) and the others. These decomposed uncertainties are illustrated in Figs. 11 and 12 for the single-lepton and dilepton channels, respectively. For the single-lepton channel, the systematic uncertainty is dominated by the $Z\gamma$ modelling, mostly from the comparison between SHERPA and MG5_AMC+PYTHIA 8 generators. Because the unfolding is performed with the distributions before the fit, the background modelling uncertain-

---

Fig. 12 The decomposed systematic uncertainties for the normalized differential cross-sections as a function of the a photon $p_T$, b photon $|\eta|$, c minimum $\Delta R(\gamma, \ell)$, d $|\Delta \eta(\ell, \ell)|$, and e $\Delta \phi(\ell, \ell)$ in the dilepton channel
ties are not constrained as in the case of fiducial cross-section measurement where a fit to ELD distribution is performed, and thus they have a much larger impact on the result.

The differences between the unfolded and the predicted distributions are quantified by the chi-squared per degree of freedom $\chi^2$/ndf, where the $\chi^2$ is

$$\chi^2 = \left( \frac{\sigma_{\text{data}} - \sigma_{\text{pred}}}{\sigma_{\text{data}}^\text{norm}} \right) \cdot C_{jk}^{-1} \cdot \left( \frac{\sigma_{\text{data}}^\text{norm} - \sigma_{\text{pred}}^\text{norm}}{\sigma_{\text{data}}^\text{norm}} \right),$$

where $\sigma_{\text{data}}^\text{norm}$ and $\sigma_{\text{pred}}^\text{norm}$ are the unfolded and predicted normalized differential cross-sections, $C_{jk}$ is the covariance matrix of $\sigma_{\text{data}}^\text{norm}$, and $j$ and $k$ are the binning indices of the distribution. For normalized differential cross-sections, the last bin of the above formula is removed from the $\chi^2$ calculation and ndf is reduced by one since this bin gives redundant information. The total correlation matrix is shown in Table 6, taking the $\Delta\phi(\ell, \ell)$ in the dilepton channel as an example. There is moderate correlation, either positive or negative, between different bins of the unfolded $\Delta\phi(\ell, \ell)$ distribution. The calculated $\chi^2$/ndf values and their corresponding $p$-values are summarized in Tables 7 and 8, quantifying the compatibility between data and each of the predictions.

---

**Table 6** The correlation matrix for the normalized differential cross-section as a function of $\Delta\phi(\ell, \ell)$ in the dilepton channel, accounting for the statistical and systematic uncertainties

<table>
<thead>
<tr>
<th>Bin</th>
<th>0.0–0.35</th>
<th>0.35–0.7</th>
<th>0.7–1.05</th>
<th>1.05–1.4</th>
<th>1.4–1.75</th>
<th>1.75–2.1</th>
<th>2.1–2.45</th>
<th>2.45–2.8</th>
<th>2.8–3.14</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0–0.35</td>
<td>1.00</td>
<td>−0.07</td>
<td>0.05</td>
<td>−0.36</td>
<td>−0.18</td>
<td>−0.08</td>
<td>−0.19</td>
<td>−0.13</td>
<td>−0.24</td>
</tr>
<tr>
<td>0.35–0.7</td>
<td>1.00</td>
<td>0.29</td>
<td>0.21</td>
<td>0.09</td>
<td>0.11</td>
<td>0.21</td>
<td>0.02</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>0.7–1.05</td>
<td>1.00</td>
<td>0.12</td>
<td>0.05</td>
<td>0.10</td>
<td>0.14</td>
<td>0.05</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.05–1.4</td>
<td>1.00</td>
<td>0.26</td>
<td>0.18</td>
<td>0.29</td>
<td>0.20</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4–1.75</td>
<td>1.00</td>
<td>0.15</td>
<td>0.16</td>
<td>0.14</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.75–2.1</td>
<td>1.00</td>
<td>0.17</td>
<td>0.15</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1–2.45</td>
<td>1.00</td>
<td>0.12</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.45–2.8</td>
<td>1.00</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.8–3.14</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7** $\chi^2$/ndf values and $p$-values between the measured normalized differential cross-sections and predictions from several generators in the single-lepton channel

| Predictions                      | $p_T(\gamma)$ | $\eta(\gamma)$ | $\Delta R(\gamma, \ell)$ | $|\Delta \eta(\ell, \ell)|$ | $\Delta \phi(\ell, \ell)$ |
|----------------------------------|----------------|-----------------|---------------------------|----------------------------|---------------------------|
|                                  | $\chi^2$/ndf   | $p$-value       | $\chi^2$/ndf              | $p$-value                   | $\chi^2$/ndf              | $p$-value                   |
| MG5_AMC + PYTHIA8                | 3.2/8          | 0.92            | 0.7/7                     | 1.0                        | 5.0/8                     | 0.76                        |
| MG5_AMC + HERWIG7                | 2.3/8          | 0.97            | 0.9/7                     | 1.0                        | 4.8/8                     | 0.78                        |
| MG5_AMC + PYTHIA8 (A14 Up)       | 3.3/8          | 0.91            | 0.8/7                     | 1.0                        | 4.9/8                     | 0.77                        |
| MG5_AMC + PYTHIA8 (A14 Down)     | 2.6/8          | 0.96            | 0.9/7                     | 1.0                        | 4.6/8                     | 0.80                        |
| POWHEG + PYTHIA8 $t\bar{t}$      | 25.4/8         | $< 0.01$        | 2.8/7                     | 0.9                        | 8.7/8                     | 0.37                        |

**Table 8** $\chi^2$/ndf values and $p$-values between the measured normalized differential cross-sections and predictions from several generators in the dilepton channel

| Predictions                      | $p_T(\gamma)$ | $\eta(\gamma)$ | $\Delta R(\gamma, \ell)$ | $|\Delta \eta(\ell, \ell)|$ | $\Delta \phi(\ell, \ell)$ |
|----------------------------------|----------------|-----------------|---------------------------|----------------------------|---------------------------|
|                                  | $\chi^2$/ndf   | $p$-value       | $\chi^2$/ndf              | $p$-value                   | $\chi^2$/ndf              | $p$-value                   |
| MG5_AMC + PYTHIA8                | 1.7/8          | 0.99            | 7.4/7                     | 0.39                       | 6.9/8                     | 0.55                        | 3.0/7                     | 0.89                        | 14.4/8 | 0.07                        |
| MG5_AMC + HERWIG7                | 2.0/8          | 0.98            | 7.4/7                     | 0.39                       | 6.6/8                     | 0.58                        | 3.1/7                     | 0.88                        | 14.4/8 | 0.07                        |
| MG5_AMC + PYTHIA8 (A14 Up)       | 1.6/8          | 0.99            | 8.4/7                     | 0.30                       | 7.4/8                     | 0.49                        | 3.4/7                     | 0.85                        | 14.0/8 | 0.08                        |
| MG5_AMC + PYTHIA8 (A14 Down)     | 1.6/8          | 0.99            | 7.9/7                     | 0.34                       | 7.5/8                     | 0.48                        | 3.2/7                     | 0.87                        | 14.4/8 | 0.07                        |
| POWHEG + PYTHIA8 $t\bar{t}$      | 20.1/8         | 0.01            | 10.8/7                    | 0.15                       | 8.6/8                     | 0.38                        | 4.5/7                     | 0.72                        | 9.8/8  | 0.28                        |
10 Conclusions

Fiducial cross-sections of top-quark pair production in association with a photon are measured in the single-lepton and dilepton decay channels of the top-quark pair using 36.1 fb$^{-1}$ of 13 TeV $pp$ collision data collected in 2015 and 2016 by the ATLAS detector at the LHC. The normalized differential cross-sections are measured as a function of the photon $p_T$ and $|\eta|$, and the $\Delta R$ between the photon and the closest lepton for both channels, and the $|\Delta \eta|$ and $\Delta \phi$ between the two leptons for the dilepton channel.

In both channels, the measured fiducial cross-sections agree well with the NLO SM predictions within uncertainties. The measured normalized differential cross-sections also agree well with the LO $t\bar{t} \gamma$ prediction and the NLO $t\bar{t}$ prediction, where the photon comes from the parton shower. The largest disagreement between data and LO $t\bar{t} \gamma$ prediction is observed in the distribution of the azimuthal opening angle between the two leptons in the dilepton channel, which is sensitive to $t\bar{t}$ spin correlation, while the NLO $t\bar{t}$ sample provides an improved agreement with data in this variable.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, 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, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; 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, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, Spain; The Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [64].

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restrictions on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk/). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/). This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413 [opendata.cern.ch].]

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

Appendix

The definitions of the discriminating variables used in ATLAS photon identification are given in Table 9.

The shapes of the PPT of the prompt, hadronic-fake, and electron-fake photons are compared in Fig. 13.

Figure 14 shows the post-fit distributions of four important variables for the training of the ELD in the single-lepton channel.

Figure 15 shows the post-fit distributions of four important variables for the training of the ELD in the dilepton channel.

Figure 16 shows the distributions of the ELD and $H_T$ in the hadronic-fake control region B, as defined in Sect. 6.1.

Figure 17 shows the distributions of the efficiency, outside fraction, and migration matrix for the dilepton azimuthal angle in the dilepton channel, which are the inputs for unfolding, as described in Sect. 7.3.
Table 9 Definitions of photon discriminating variables

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadronic leakage</td>
<td>Transverse energy leakage in the hadronic calorimeter normalized to transverse energy of the photon candidate in the ECAL. In the region $0.8 \leq</td>
</tr>
<tr>
<td>$R_{\text{had}}$ or $R_{\text{had1}}$</td>
<td>Energy ratios and width in the second layer of ECAL</td>
</tr>
<tr>
<td>$R_\eta$</td>
<td>Energy ratio of $3 \times 7$ to $7 \times 7$ cells in the $\eta \times \phi$ plane</td>
</tr>
<tr>
<td>$R_\phi$</td>
<td>Energy ratio of $3 \times 3$ to $3 \times 7$ cells in the $\eta \times \phi$ plane</td>
</tr>
<tr>
<td>$w_{\eta 2}$</td>
<td>Lateral width of the shower, using a window of $\eta \times \phi = 3 \times 5$ cells</td>
</tr>
<tr>
<td>Energy ratios and widths in the first (strip) layer of ECAL</td>
<td></td>
</tr>
<tr>
<td>$w_{\eta 3}$</td>
<td>Shower width along $\eta$, using 3 strips around the largest energy deposit</td>
</tr>
<tr>
<td>$w_{\eta \text{tot}}$</td>
<td>Shower width along $\eta$, using $20 \times 2$ strip cells in the $\eta \times \phi$ plane</td>
</tr>
<tr>
<td>$F_{\text{side}}$</td>
<td>Energy outside the 3 central strips but within 7 strips, normalized to the energy within the 3 central strips</td>
</tr>
<tr>
<td>$E_{\text{ratio}}$</td>
<td>Ratio between difference of the first and second energy maximum divided by their sum ($E_{\text{ratio}} = 1$ if there is no second maximum)</td>
</tr>
<tr>
<td>$\Delta E$</td>
<td>Difference between the second energy maximum and the minimum found between first and second maximum ($\Delta E = 0$ if there is no second maximum)</td>
</tr>
</tbody>
</table>

Fig. 13 The simulated shapes of the output of the prompt-photon tagger for the prompt, hadronic-fake, and electron-fake photons after applying Tight identification and isolation criteria.
Fig. 14 Distributions of the a highest and b second highest $b$-tagging scores for all jets, c total number of $b$-tagged jets (at a 77% efficiency), and d scalar sum of the transverse energies of all the selected physical objects in the single-lepton channel after event selection and likelihood fit. Each bin of a and b, from one to five, represents the $b$-tagging efficiencies of $>85\%$, $77–85\%$, $70–77\%$, $60–70\%$, and $<60\%$. All data-driven corrections and systematic uncertainties are included.
Fig. 15  Distributions of the a highest and b second highest $b$-tagging scores for all jets, c total number of $b$-tagged jets (at a 77% efficiency), and d invariant mass between the two leptons in the dilepton channel after event selection and likelihood fit. Each bin of a and b, from one to five, represents the $b$-tagging efficiencies of $> 85\%$, 77–85%, 70–77%, 60–70%, and < 60%. All data-driven corrections and systematic uncertainties are included.
Fig. 16 Distributions of the a ELD and b $H_T$ in the hadronic-fake control region B of the single-lepton channel. Data-driven correction to the electron-fake background and all systematic uncertainties are included. The contributions from $t\bar{t}\gamma$ and $W\gamma +$ other prompts are very small (0.2% and 0.3%, respectively), therefore not visible.

Fig. 17 The a efficiency and outside fraction and b migration matrix for the dilepton azimuthal angle in the dilepton channel.

References

6. CDF Collaboration, Evidence for $t\bar{t}\gamma$ production and measurement of $\sigma_{\gamma} / \sigma_{t\bar{t}}$. Phys. Rev. D 84, 031104(R) (2011). arXiv:1106.3970 [hep-ex]
LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand, France
38 Nevis Laboratory, Columbia University, Irvington, NY, USA
39 Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
40 (a) Dipartimento di Fisica, Università della Calabria, Rende, Italy; (b) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy
41 Physics Department, Southern Methodist University, Dallas, TX, USA
42 Physics Department, University of Texas at Dallas, Richardson, TX, USA
43 (a) Department of Physics, Stockholm University, Stockholm, Sweden; (b) Oskar Klein Centre, Stockholm, Sweden
44 Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen, Germany
45 Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany
46 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany
47 Department of Physics, Duke University, Durham, NC, USA
48 SUPA-School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK
49 INFN e Laboratori Nazionali di Frascati, Frascati, Italy
50 II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
51 Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
52 II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany
53 Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
54 (a) Dipartimento di Fisica, Università di Genova, Genoa, Italy; (b) INFN Sezione di Genova, Genoa, Italy
55 II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
56 SUPA-School of Physics and Astronomy, University of Glasgow, Glasgow, UK
57 LPSC, Université Grenoble Alpes, CNRS/IN2P3, Grenoble INP, Grenoble, France
58 Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA, USA
59 (a) Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei, China; (b) Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao, China; (c) School of Physics and Astronomy, Shanghai Jiao Tong University, KLPPAC-MoE, SKLPPC, Shanghai, China; (d) Tsung-Dao Lee Institute, Shanghai, China
60 (a) Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; (b) Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
61 Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima, Japan
62 (a) Department of Physics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong; (b) Department of Physics, University of Hong Kong, Hong Kong, China; (c) Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
63 Department of Physics, National Tsing Hua University, Hsinchu, Taiwan
64 Department of Physics, Indiana University, Bloomington, IN, USA
65 (a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy; (b) ICTP, Trieste, Italy; (c) Dipartimento Politecnico di Ingegneria e Architettura, Università di Udine, Udine, Italy
66 (a) INFN Sezione di Lecce, Lecce, Italy; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy
67 (a) INFN Sezione di Milano, Milan, Italy; (b) Dipartimento di Fisica, Università di Milano, Milan, Italy
68 (a) INFN Sezione di Napoli, Naples, Italy; (b) Dipartimento di Fisica, Università di Napoli, Naples, Italy
69 (a) INFN Sezione di Pavia, Pavia, Italy; (b) Dipartimento di Fisica, Università di Pavia, Pavia, Italy
70 (a) INFN Sezione di Pisa, Pisa, Italy; (b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa, Italy
71 (a) INFN Sezione di Roma, Rome, Italy; (b) Dipartimento di Fisica, Sapienza Università di Roma, Rome, Italy
72 (a) INFN Sezione di Roma Tor Vergata, Rome, Italy; (b) Dipartimento di Fisica, Università di Roma Tor Vergata, Rome, Italy
73 (a) INFN Sezione di Roma Tre, Rome, Italy; (b) Dipartimento di Matematica e Fisica, Università Roma Tre, Rome, Italy
74 (a) INFN-TIFPA, Povo, Italy; (b) Università degli Studi di Trento, Trento, Italy
75 Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria
76 University of Iowa, Iowa City, IA, USA
77 Department of Physics and Astronomy, Iowa State University, Ames, IA, USA
78 Joint Institute for Nuclear Research, Dubna, Russia
79 (a) Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Brazil

38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79
Brazil; (b) Universidade Federal do Rio de Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; (c) Universidade Federal de São João del Rei (UFSJ), São João del Rei, Brazil; (d) Instituto de Física, Universidade de São Paulo, São Paulo, Brazil
KEK, High Energy Accelerator Research Organization, Tsukuba, Japan
Graduate School of Science, Kobe University, Kobe, Japan
(a) Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Krakow, Poland; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland
Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland
Faculty of Science, Kyoto University, Kyoto, Japan
Kyoto University of Education, Kyoto, Japan
Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka, Japan
Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
Physics Department, Lancaster University, Lancaster, UK
Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK
Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana, Slovenia
School of Physics and Astronomy, Queen Mary University of London, London, UK
Department of Physics, Royal Holloway University of London, Egham, UK
Department of Physics and Astronomy, University College London, London, UK
Louisiana Tech University, Ruston, LA, USA
Fysiska institutionen, Lunds universitet, Lund, Sweden
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France
Departamento de Física Teorica C-15 and CIAFF, Universidad Autónoma de Madrid, Madrid, Spain
Institut für Physik, Universität Mainz, Mainz, Germany
School of Physics and Astronomy, University of Manchester, Manchester, UK
CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
Department of Physics, University of Massachusetts, Amherst, MA, USA
Department of Physics, McGill University, Montreal, QC, Canada
School of Physics, University of Melbourne, Victoria, Australia
Department of Physics, University of Michigan, Ann Arbor, MI, USA
Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Belarus
Research Institute for Nuclear Problems of Byelorussian State University, Minsk, Belarus
Group of Particle Physics, University of Montreal, Montreal, QC, Canada
P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
Institute for Theoretical and Experimental Physics of the National Research Centre Kurchatov Institute, Moscow, Russia
National Research Nuclear University MEPhI, Moscow, Russia
D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Munich, Germany
Nagasaki Institute of Applied Science, Nagasaki, Japan
Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA
Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands
Department of Physics, Northern Illinois University, DeKalb, IL, USA
(a) Budker Institute of Nuclear Physics and NSU, SB RAS, Novosibirsk, Russia; (b) Novosibirsk State University, Novosibirsk, Russia
Institute for High Energy Physics of the National Research Centre Kurchatov Institute, Protvino, Russia
Department of Physics, New York University, New York, NY, USA
Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo, Japan
Ohio State University, Columbus, OH, USA
also at LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France
also at Manhattan College, New York, NY, USA
also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
also at National Research Nuclear University MEPhI, Moscow, Russia
also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
also at School of Physics, Sun Yat-sen University, Guangzhou, China
also at The City College of New York, New York, NY, USA
also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing, China
also at Tomsk State University, Tomsk, and Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
also at TRIUMF, Vancouver, BC, Canada
also at Universidad de Granada, Granada (Spain), Spain
also at Universita di Napoli Parthenope, Napoli, Italy
* Deceased