Supplemental Material for: Observation of single-top-quark production in association with a photon using the ATLAS detector

The ATLAS Collaboration

Technical description of the $e \to \gamma$ rate estimate

Events are selected in two regions: a region enriched in $Z \to e^+e^-$ events with an electron or positron misidentified as a photon ($Z \to e\gamma$ region) and a region enriched in $Z \to e^+e^-$ events with the electron and positron correctly identified ($Z \to ee$ region). Events in the $Z \to e\gamma$ region are selected by requiring exactly one reconstructed electron (or positron) and exactly one reconstructed photon to fulfill the same identification criteria that are used in the definition of the SRs and CRs in the analysis. Events in the $Z \to e^+e^-$ region are selected by requiring exactly two electrons with opposite electric charge and the absence of a photon. For orthogonality with the SRs and CRs and to enhance the fraction of Z-boson events, $E_{\mathrm{T}}^{\mathrm{miss}}$ must be less than 30 GeV in both regions and no b-tagged jet (with the nominal, tight criteria) may be present. The invariant mass of the electron–photon pair ($m_{e\gamma}$) or electron–positron pair (m_{ee}) must be in the range 70–110 GeV.

A template fit to the m_{ee} distribution is used in the $Z \to ee$ region to determine the number of $Z \to e^+e^-$ events, $N(Z \to e^+e^-)$, from data. Templates for $Z \to e^+e^-$ production and for background processes are taken from MC simulations, where the background processes are scaled to their expected numbers of events from MC simulations, normalized to their respective theoretical cross sections. Similarly, a template fit to the $m_{e\gamma}$ distribution is used in the $Z \to e\gamma$ region to determine the number of $Z \to e\gamma$ events, $N(Z \to e\gamma)$, in data. Templates are again taken from MC simulations, except for several small backgrounds (such as diboson or γ+jets production), which are modeled inclusively with a third-order Bernstein polynomial with its parameters determined from data. The number of $Z \to e\gamma$ events is determined in six bins in photon η (motivated by the geometry of the electromagnetic calorimeter) and six bins that correspond to six different types of photon reconstruction: one reconstruction type for unconverted photons with no associated track in the inner detector and five types for converted photons with associated tracks from $\gamma \to e^+e^-$ pair production in the material in front of the electromagnetic calorimeter. Converted photons can have two associated tracks or one associated track (if one track was not reconstructed or not matched to the photon candidate) and the tracks can have hits in the silicon (Si) tracking detectors or only in the transition radiation tracker (TRT), so that five types of converted photons are reconstructed: one track with Si hits (SingleSi), one track only with hits in the TRT (SingleTRT), two tracks both with Si hits (DoubleSi), two tracks both with only hits in the TRT (DoubleTRT), and two tracks with one having Si hits and one having only TRT hits (DoubleSiTRT). The quality of the reconstruction varies across these different reconstruction types and so does the modeling of the $e \to \gamma$ probability in the MC simulation.

Scale factors (SF) are evaluated to correct the $e \to \gamma$ probability in the MC simulations to correspond to data. These SFs are evaluated as a double ratio by dividing $N_i(Z \to e\gamma)/N(Z \to e^+e^-)$ in data by the same ratio in MC simulation, where i is one of the 36 bins in photon η and reconstruction type. While statistical uncertainties are mostly below 5%, and below 10% in all bins (with the exception of DoubleSiTRT in the central part of the detector, which happen rarely), systematic uncertainties dominate. These systematic uncertainties originate from uncertainties in the background subtraction, the relative electron–photon energy scale and the choice of invariant-mass windows for $m_{e\gamma}$ and m_{ee}. Uncertainties in the background modeling are assessed by conservatively removing the polynomial corresponding to the rare backgrounds from the fit, as well as removing the template for $W\gamma$ and $Z\gamma$ production from the fit. The modeling of the Z+jets process is alternatively assessed with MC samples generated with Powheg+Pythia 8. The photon energy scale is varied by ±1% with the electron energy scale fixed and the $m_{e\gamma}$ and m_{ee} windows are varied from 70–110 GeV to 80–100 GeV. The dominant systematic uncertainties vary between the different bins, with the most important sources being the variation of the

Z+jets MC programs, the removal of the background polynomial and the variation of the invariant-mass range. The measured values of the SFs with their total uncertainties are given in Table 1.

Table 1: Measured values of the scale factors to correct the $e \to \gamma$ probability in MC simulations, with their uncertainties, in bins of photon $|\eta|$ and photon reconstruction type.

Photon $\|\eta\|$ range Reco. type	0–0.3	0.3–0.6	0.6–1.0	1.0–1.37	1.52–1.81	1.81–2.37
Unconverted	1.10 ± 0.11	0.97 ± 0.11	0.99 ± 0.11	1.03 ± 0.08	1.05 ± 0.09	1.06 ± 0.05
SingleSi	2.51 ± 0.26	1.79 ± 0.27	1.73 ± 0.18	1.33 ± 0.11	1.02 ± 0.04	1.11 ± 0.06
SingleTRT	1.81 ± 0.16	1.47 ± 0.16	1.04 ± 0.09	0.94 ± 0.10	0.90 ± 0.07	0.92 ± 0.04
DoubleSi	0.99 ± 0.05	0.94 ± 0.04	0.94 ± 0.03	0.90 ± 0.02	0.84 ± 0.05	0.95 ± 0.02
DoubleTRT	1.28 ± 0.31	1.09 ± 0.19	0.90 ± 0.18	0.96 ± 0.10	0.91 ± 0.07	0.80 ± 0.09
DoubleSiTRT	2.00 ± 0.74	1.57 ± 0.49	1.21 ± 0.16	1.20 ± 0.08	0.86 ± 0.06	0.78 ± 0.07

Technical description of the $h \to \gamma$ rate estimate

Events are selected in three regions (A, B, C) with definitions based on whether the reconstructed photon fulfills the track and calorimeter isolation criteria and whether it fulfills the requirements on a subset of the photon identification criteria. This subset is based only on variables that describe the inner core of the photon shower, so that the requirements on this subset are largely uncorrelated with whether the photon fulfills the isolation criteria. The variables in the subset are [**?**] F_{side}, $w_{s,3}$, ΔE and E_{ratio}. All other photon identification criteria have to be fulfilled. The three regions are defined as follows:

- The photon fulfills track and calorimeter isolation criteria but fails to fulfill at least one of the criteria for the identification subset variables (region A).
- The photon fails to fulfill both the track and calorimeter isolation criteria but fulfills all criteria for the identification subset variables (region C).
- The photon fails to fulfill both the track and calorimeter isolation criteria as well as at least one of the criteria for the identification subset variables (region B).

All other event-selection criteria correspond to those in the SR or CR for which the number of $h \to \gamma$ events is to be estimated.

The events in regions A, B and C are largely due to events with a $h \to \gamma$ fake, and the number of $h \to \gamma$ events, $N_{h\to\gamma}$, is obtained by subtracting the expected number of events with prompt photons (including those with fake leptons) as well as events with $e \to \gamma$ fakes. The contribution of $h \to \gamma$ events is about 52% (48%) in region A, about 94% (96%) in region B and about 75% (72%) in region C in case of converted (unconverted) photon candidates. The number of $h \to \gamma$ events in a SR or CR (region D) is then estimated by:

$$N_{h\to\gamma}(D) = N_{h\to\gamma}(C) \cdot \frac{N_{h\to\gamma}(A)}{N_{h\to\gamma}(B)} \cdot \theta$$

where the correlation factor θ accounts for the residual correlations for fake photons between the isolation criteria and the identification criteria for the selected subset of variables and is determined from MC simulations:

$$\theta = \frac{N^{\text{MC}}_{h\to\gamma}(B)}{N^{\text{MC}}_{h\to\gamma}(A)} \cdot \frac{N^{\text{MC}}_{h\to\gamma}(D)}{N^{\text{MC}}_{h\to\gamma}(C)}$$

Scale factors to correct the MC prediction for $h \to \gamma$ are derived separately for unconverted and converted photons, in the same six bins in photon η as for the $e \to \gamma$ rate estimate, motivated by the detector geometry, and two bins in photon p_{T} ($<$ 40 GeV, $\geq$ 40 GeV), to capture the kinematic dependence of the modeling of $h \to \gamma$ in the MC simulations and considering the limited number of events in regions A, B and C. Due to the limited number of events, the two last η bins are combined into one bin for unconverted photons. In these 22 bins, θ takes values ranging from 0.94 to 1.30 with statistical uncertainties in the range from ±0.05 to ±0.23.

Scale-factor uncertainties from various sources are evaluated and comprise statistical uncertainties in the data and MC samples as well as systematic uncertainties in the background subtraction and uncertainties in the correlation factor θ. Uncertainties due to the limited number of events in data and in the MC samples are evaluated using 1000 pseudoexperiments and are in the range ±11% to ±28%. Uncertainties in the background subtraction are evaluated by varying the $e \to \gamma$ SFs within their uncertainties, by varying the $t\bar{t}$ cross section by ±6% and by varying the factorization and renormalization scales in $t\bar{t}\gamma$ and $t\bar{t}$ production by factors of 0.5 and 2, as well as by varying the $t\bar{t}\gamma$ uncertainty normalization by ±15% simultaneously with normalization variations of other processes with prompt photons (in particular $W\gamma$ and $Z\gamma$ production) by ±30%. The correlation factor θ is re-evaluated in the MC samples by varying the number of identification variables that enter the definitions of regions A, B and C. Three alternative subsets of identification variables are defined that (1) exclude ΔE and E_{ratio} from the original subset, (2) exclude only E_{ratio} from the original subset, or (3) add $w_{s,\text{tot}}$ to the original subset. The largest variations seen in θ across all bins and subset definitions are used to estimate conservative uncertainties in θ. These uncertainties dominate over all other sources of uncertainty and vary between ±29% and ±53%. The measured values of the SFs with their total uncertainties are given in Table 2.

Table 2: Measured values of the scale factors to correct the $h \to \gamma$ probability in MC simulations, with their uncertainties, separately for unconverted and converted photons and in bins of photon $|\eta|$ and p_{T}.

Photon $\|\eta\|$ range		0–0.3	0.3–0.6	0.6–1.0	1.0–1.37	1.52–1.81	1.81–2.37
(Un)converted	p_{T} range						
Unconverted	20–40 GeV	1.3 ± 0.6	1.2 ± 0.6	1.7 ± 0.9	1.4 ± 0.9	1.4 ± 0.9	
Unconverted	≥40 GeV	0.7 ± 0.4	1.0 ± 0.5	1.2 ± 0.6	1.3 ± 0.6	1.9 ± 1.0	
Converted	20–40 GeV	1.5 ± 0.8	1.7 ± 0.8	1.6 ± 0.8	1.0 ± 0.6	1.1 ± 0.6	1.2 ± 0.5
Converted	≥40 GeV	1.0 ± 0.4	0.9 ± 0.4	1.0 ± 0.4	1.0 ± 0.4	0.9 ± 0.5	0.7 ± 0.4