Not reviewed, for internal circulation only

ATLAS Note

ATL-COM-PHYS-2017-673

16th November 2018

Draft version 5.1

Measurement of the $t\bar{t}\gamma$ production cross-section in pp collisions at $\sqrt{s}$ = 13 TeV with ATLAS detector

Binish Batool[c], Julien Caudron[a], Markus Cristinziani[a], Ivor Fleck[c], Sara Ghasemi[c], Mazuza M. A. Ghneimat[a], Sebastian Heer[a], Yichen Li[c], Maria Moreno Llacer[d], John Meshreki[c], Arnulf Quadt[b], Amartya Rej[c], Elizaveta Shabalina[b], Joshua Wyatt Smith[b], Benedikt Völkel[b], Knut Zoch[b]

[a]*University of Bonn*
[d]*CERN*
[b]*University of Göttingen*
[c]*University of Siegen*

The production of a top-quark pair in association with a photon is measured in proton-proton collisions at a center-of-mass energy of 13 TeV. The integrated luminosity is 36.1 fb^{-1}, collected in 2015 and 2016. The measurement is performed in the single and dilepton channels. In the single lepton channels, exactly one lepton and at least four jets are requested, with at least one jet being b-tagged. In the dilepton channels, exactly two opposite charge leptons, at least two jets with at least one jet being b-tagged, and a large missing transverse momentum are requested. A veto is also imposed for events compatible with a Z boson mass reconstruction. For all channels exactly one isolated photon with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.37$ is required. This photon should not be too close to the selected leptons and jets. In addition, an event-level multivariate analysis is performed using event-level variables with the output variable distribution used in a profile likelihood fit, from which the final results are obtained. The fiducial cross-sections in the different channels are extracted as $\sigma_{\mathrm{fid}}^{\mathrm{SL}} = 521.42\,^{+40.93}_{-39.20}(\mathrm{sys.})\,^{+9.21}_{-9.19}(\mathrm{stat.})$ fb and $\sigma_{\mathrm{fid}}^{\mathrm{DL}} = 68.57\,^{+3.95}_{-3.76}(\mathrm{sys.})\,^{+2.64}_{-2.58}(\mathrm{stat.})$ fb for single lepton and dilepton, respectively, with a final combined fit to all 5 channels giving $\sigma_{\mathrm{fid}}^{\mathrm{Incl.}} = 589.43\,^{+33.77}_{-32.05}(\mathrm{sys.})\,^{+9.87}_{-9.73}(\mathrm{stat.})$ fb. Several differential cross-sections are also extracted: as function of the photon p_{T}, the photon $|\eta|$ and the distance between the photon and closest lepton for both single lepton and dilepton channels and as function of the η and ϕ angle between the two leptons in the dilepton channel. All these measurements are in agreement with the theoretically computed numbers.

© 2018 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

Not reviewed, for internal circulation only

Contents

1 Introduction **15**

2 ATLAS detector **18**

3 Data and simulation samples **19**
3.1 Samples description 19
3.2 Removal of the double counting of events in X+jets and X+γ samples 20

4 Objects definition **21**
4.1 Objects selection 21
4.2 Prompt photon tagger 23
4.2.1 PPT training 24
4.2.2 Systematic uncertainties for the PPT 31

5 Signal region definition **43**
5.1 Event selection 43
5.2 Event level MVA 60
5.2.1 Discriminator for single lepton and dilepton channels 60
5.2.2 Variable selection and training 60
5.2.3 Application and evaluation 66

6 Background descriptions **72**
6.1 Background due to mis-reconstructed electrons 72
6.1.1 Control regions 72
6.1.2 Fake sources 73
6.1.3 Fake rates 75
6.1.4 Signal region $e \to \gamma$ fake 89
6.1.5 PPT systematics 90
6.1.6 Application of fake rates 91
6.2 Background due to hadronic fakes 92
6.2.1 Estimation of hadronic fakes in single lepton channel 92
6.2.2 Estimation of hadronic fakes in dilepton channel 105
6.3 Background due to fake leptons 106
6.3.1 Fake leptons in single lepton channel 106
6.3.2 Fake leptons in dilepton channel 110
6.4 Prompt photons background 111
6.4.1 $Z\gamma$ validation region 112
6.4.2 $W\gamma$ validation region 115

7 Cross section definitions **118**
7.1 Fiducial region 118
7.1.1 Particle definition for fiducial region 118
7.1.2 Event selection for fiducial region 118
7.1.3 Acceptance and correction factor 119
7.2 Fiducial cross section 120
7.3 Differential cross section 121

Not reviewed, for internal circulation only

8 Fit strategy **128**
8.1 Likelihood function 128
8.2 Signal strength 129
8.3 Fiducial cross section 129
8.4 Types of fits 129

9 Unfolding **131**
9.1 Iterative Bayesian method 131
9.1.1 The pseudo-data 132
9.1.2 Optimising the number of iterations 132
9.1.3 The closure test 144
9.1.4 Pull study 147
9.1.5 Stress tests 150
9.2 Cross-check methods 153
9.2.1 Bin-by-bin correction method 153
9.2.2 The singular value decomposition method 153
9.2.3 Comparison test 153

10 Systematic uncertainties **156**
10.1 Smoothing, symmetrisation and pruning 156
10.2 Theoretical Uncertainties 162
10.2.1 Signal Modeling 162
10.2.2 Background Modelling 166
10.3 Experimental Uncertainties 169
10.3.1 Sources of experimental uncertainties 169
10.4 Propagation of the systematic uncertainties to the unfolding framework 175

11 Pre-fit distributions with scale factors and systematics applied **176**

12 Theory prediction **180**
12.1 LO jet inclusive cross-section 180
12.2 NLO cross-section 181
12.3 LO cross-section 181
12.4 Kfactor and fiducial cross-section 182

13 Results **184**
13.1 Fit model and results 184
13.1.1 μ, best-fit 184
13.1.2 Post-fit plots and yields 185
13.1.3 Nuisance parameters 192
13.2 Fiducial cross section 203
13.3 Differential cross section 204

14 Conclusion **221**

Not reviewed, for internal circulation only

Appendices **226**

A Studies on removal of the double counting of events between samples **227**
A.1 Definition of true hadronic fakes 227
A.2 Comparison with other methods 233
A.3 Et ratio plots 234

B Additional information for PPT **235**
B.1 Photon kinematics in the training samples 235
B.2 Comparison of different architectures of PPT 235
B.3 Further data/MC plots for PPT systematics 236
B.4 Derivations used for training 239
B.5 Shower shape comparisons in different p_{T} and η regions 239
B.6 Shower shape comparisons for training and test split 242
B.7 Systematic scale factors for the PPT 242

C Cuts optimization **245**
C.1 Cuts optimization in the single lepton channel 245
C.1.1 Results in the single lepton channel 246
C.2 Cuts optimization in the dilepton channel 249
C.2.1 Results in the dilepton channel 249

D Radiative decay versus radiative production **254**

E Photon isolation cut **255**

F Hadronic-fake background appendix **257**
F.1 Study on origin of hadronic fakes 257
F.1.1 Processes contribution to the hadronic fake background 257
F.1.2 Truth matching to the parton level 258
F.2 $t\bar{t}$ MC modelling systematics in hadronic fake control regions 259
F.3 $\mathrm{SF}^{\text{h-fake}}$ for e+jets and μ+jets channels 261
F.4 Data-driven double ratio 268
F.5 Red/blue plots for hadronic fake systematic uncertainties 273
F.5.1 Uncertainty due to the subtraction of $t\bar{t}\gamma$ 273
F.5.2 Uncertainty due to the statistical uncertainties of the data-driven part or of the θ_{MC} factor 274

G $Z\gamma$ validation regions **277**
G.1 $Z\gamma$ heavy jet validation region 277
G.2 $Z\gamma$ light jet validation region 277

H $W\gamma$ validation regions **281**
H.1 $W\gamma$ heavy jet validation region 281
H.2 $W\gamma$ light jet validation region 281

I ELD binning optimisation **285**

Not reviewed, for internal circulation only

J 100 PDF shape uncertainties **286**

K Varying the single top, diboson and $t\bar{t}V$ systematic uncertainty **288**

L Systematic uncertainty breakdown **289**

M Fits with symmetrisation off **291**

N Unfolding checks **294**
N.1 Comparison of different number of iterations 294
N.2 Pre-fit versus post-fit total error for correlated and uncorrelated scenarios 294
N.3 Pre-fit versus post-fit background shape comparisons 300
N.4 Pre-fit unfolding inputs 302
N.5 Pre-/post-fit unfolding comparison 315

O Unfolding using fine binning **322**
O.1 Detector resolution 322
O.2 Inputs 322
O.2.1 Optimising the number of iterations 322
O.2.2 The closure test 333
O.2.3 Pull study 333
O.2.4 Stress tests 333
O.3 Cross-check methods 336
O.3.1 Closure tests 336

P Pull study for unfolding with IBS **343**

Q Results appendix **346**
Q.1 Asimov fits to ELD, number of jets and 1 bin distributions 346
Q.2 Post-fit plots for single lepton and dilepton channels 346
Q.3 Results for e+jets, μ+jets, ee, $\mu\mu$ and $e\mu$ 350
Q.4 Final fits for all scenarios 368
Q.5 Nuisance parameter pull studies 368
Q.5.1 $t\bar{t}\gamma$ ISR/FSR 368
Q.5.2 Jet Energy Resolution studies 370
Q.5.3 Lepfake studies 372
Q.5.4 $Z\gamma$ modelling 373
Q.6 Asimov fit scenarios: $V\gamma$ floating or constrained 373
Q.6.1 single lepton and dilepton combined pull plots 375
Q.6.2 Individual channel pull plots 377

R Systematic uncertainties: RhoTopology studies **382**

S Impact of the PPT on the analysis **383**

T New physics observations using Effective Field Theory **389**
T.1 EFT samples 390
T.2 The study in the fiducial region 390

T.2.1 The fit 391

U First look at EFT 394
U.1 Cross check of first EFT results 394
U.2 The Monte Carlo study in the fiducial region in the single lepton channel 395
U.3 Further variables for the EFT in the fiducial region 397
U.4 EFT validation 400

V Studying larger weights in the $W\gamma$ and $Z\gamma$ MC samples 402

W Estimation of a possible mis-modelling of prompt photons from Pile-Up 406

X Objects overlap removal: loose photons vs. tight photons 408

Y Correction of Jet Flavour Composition Uncertainty 410

Z Monte Carlo Samples 412
Z.1 Nominal samples 412
Z.2 Samples for estimation of systematic uncertainties 416

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

List of contributions

Binish Batool	Fiducial region definition (including calculation of acceptance and correction factors)
Julien Caudron	Internal note editor (until June 2018), analysis coordination (until June 2018), fake lepton background for dilepton channel, ntuple production, samples overlap removal
Markus Cristinziani	Supervision of M. Ghneimat and S. Heer
Ivor Fleck	Paper editor, supervision of B. Batool, S. Ghasemi, J. Meshreki and A. Rej
Sara Ghasemi	Internal note editor, hadronic fake background
Mazuza M. A. Ghneimat	Cut optimisation for dilepton channel, validation and production of the new benchmark signal sample, Zgamma validation region, unfolding, EFT preliminary studies
Sebastian Heer	Hand-over of 8 TeV experience on the dilepton analysis
Yichen Li	Paper editor, analysis coordination, efake background, theoretical cross-section prediction, final differential cross section calculations
Maria Moreno Llacer	Help with defining the analysis strategy and with the profile likelihood fit
John Meshreki	Fake lepton background for single lepton, took over unfolding from M. Ghneimat
Arnulf Quadt	Supervision of J. Smith, B. Völkel and K. Zoch
Amartya Rej	Signal modelling systematics, determination of gluon jets fraction for JES systematics
Elizaveta Shabalina	Supervision of J. Smith, B. Völkel and K. Zoch
Joshua Wyatt Smith	Internal note editor, cut optimisation for single lepton channel, ELD development, Profile likelihood fit, systematics, Wgamma and Zgamma validation regions
Benedikt Völkel	PPT development
Knut Zoch	PPT development (took over from B.Völkel)

Not reviewed, for internal circulation only

Updates / History

- **version 5.1, 16 November 2018**
  - Updated jet b-tagging weight plots (for jets with highest, second highest and third highest weight) in Figure 20.
  - Updated Table 10.
- **version 5.0, 1 October 2018**
  - Add Appendix S: Impact of the PPT on the analysis
  - Updated all the differential cross-section results (plots and tables) to be consistent with the paper draft for ATLAS second circulation
- **version 4.3.5, 13 September 2018**
  - A few edits in Differential cross-section Section, in order to be consistent with the paper draft before the Paper Approval Meeting:
    - Updated the normalised differential cross-section plots, the correlation matrices and the $\chi 2$ tables
    - Added the decomposed systematics plots (Figures 117 and 118)
    - Added the statistical correlations using the bootstrap method (Figures 119 and 120)
- **version 4.3.4, 29 August 2018**
  - Updated contribution and author list.
- **version 4.3.3, 29 August 2018**
  - Only minor changes, to address the additional follow-up comments from TOP WG approval:
    - Updated contribution list
    - Added integral of differential measurements in Chapter 13 for differential measurement.
    - Added Appendix F.5 which contains motivation for Equation 14 and 21 red/blue plots for Table 16 (overall θ_{MC} as well as 10 individual statistical NPs each for θ_{MC} and the SF.)
- **version 4.3.2, 26 August 2018**
  - Only minor changes, to address the additional follow-up comments from TOP WG approval:
    - Chapter 4: Minor bugfix for muon overlap removal. (dR cone was tweaked.)
    - Chapter 3: Add the $t\bar{t}V$ description which was missing.
    - Chapter 11: added a table with prefit yields with all systematics and scale factors.
    - Section 13.3: added signal modelling uncertainties in the χ^2 calculation.
- **version 4.3.1, 19 July 2018**
  - Only minor changes, to address the additional follow-up comments from TOP WG approval:

Not reviewed, for internal circulation only

    * Chapter 8 Fit strategy: Removed subsections 8.4 (Total cross section) and 8.5 (Differential cross section) that we don't need any more.
    * Chapter 13 Results: Added post-fit plots with 5-channel inclusive fit mu applied to all channels.

- **version 4.3, 03 July 2018**
  - Addressing the follow-up comments from TOP WG approval.
  - Change in the order of chapters: Moved Systematics (Chapter 10) and "Pre-fit distributions with scale factors and systematics applied" (Chapter 11) later so the fit is referenced before it (were Chapter 7 and 8 previously).
  - Chapter 10: Systematics
    * Add plots showing which systematic gets symmetrised and smoothed
    * More red/blue plots for various systematics
  - Chapter 13, Results
    * Reinterpret results for pulls and constraints and link to more appendixes with more systematic plots.
    * Include 5 channel inclusive ranking and pull plot
    * Updated all differential cross section results. We now use pre-fit uncertainties and background normalizations for unfolding (before post-fit was being used).
  - Added Appendix Y where the estimation of gluon fraction of jets used for JES flavour composition uncertainty is explained.
  - Appendix M added which shows results of switching off all symmetrisation, as a check.
- **version 4.2.2, 16 June 2018**
  - The few final comments from TOP WG approval. See previous history (4.2) for the previous TOP WG approval changes.
  - Chapter 6.4 Prompt backgrounds
    * Rewrite $W\gamma$ and $Z\gamma$ sections to include fits with all systematics on the jet pt distribution
  - Chapter 11 now shows the variables used to train the respective SL and DL ELD's, with *all* systematic uncertainties and scale factors. Note, the ordering has been switched with Chapter 10.
  - Chapter 13, Results
    * Added Chi-square, p-value tests as well as covariance matrices.
  - Added Q.2: Show's the post-fit plot of Chapter 11.
  - Added N.5: for comparison of unfolding results using pre-/post-fit inputs
- **version 4.2, 12 June 2018**

Not reviewed, for internal circulation only

- The draft version for which the Top WG approval comments have been applied.
- Chapter 4.2 Prompt photon tagger
  - Added isolation dependence plot for PPT.
- Chapter 5 ELD
  - Add loss curves for SL and DL ELD.
- Chapter 6.2.1 Estimation of hadronic fakes in single lepton channel
  - Moved control plots and results of individual e+jets and μ+jets channels to Appendix F.3 , and replaced with control plots of combined single lepton channel.
  - Added ELD plots
  - Added plots and tables for binned θ_{MC}
- Chapter 10 Systematics
  - For signal modelling, include a blurb on how and why we separate shape and acceptance effects, pointing the reader to chapter 7
  - For PDF shape uncertainties add table as well as link to Appendix J which demonstrates that PDF shape uncertainties are completely negligible and not needed in the final fit.
- Chapter 8.4, Types of fits
  - Added a description and figure for the different types of fits we do. This also references the new Appendix Q.4, which shows a plot for all scenarios.
- Chapter 12, Theory prediction
  - Add Section 12.1 to compare MadGraph and theoretical calculation in jet inclusive phase space to address Jay's approval comment.
- Chapter 13, Results
  - Updated all fit results, plots and figures to what is hopefully the final result.
  - Added NP discussion to Section 13.1.3 which references new plots in the systematics chapter (Chapter 10) or the new Appendix Q.5.
  - Added tables of differential cross section numbers, and added the normalised differential cross sections in Tables and figures.
  - Added comparsion of differential cross section with a number of MC predictions (A14 variation, etc.)
- Added Appendix A: Studies on removal of the double counting of events between samples
- Added Appendix F.1 on origin of hadronic fakes, and Appendix F.2 on detail of $t\bar{t}$ MC modelling systematics in hadronic fake control regions
- Added Appendix I "Binning optimisation for the ELD ".
- Added Appendix J "100 PDF shape uncertainties"

Not reviewed, for internal circulation only

- Added Appendix K "Varying the single top, diboson and $t\bar{t}V$ systematic uncertainty" that shows what happens when these uncertainties are doubled.
- Added Appendix N : Unfolding using pre-fit inputs and, to compare 3 and 6 iterations in photon p_{T} bins in the dilepton channel.
  - Added N.2: Pre-fit versus post-fit total error for correlated and uncorrelated scenarios
  - Added N.3: Pre-fit versus post-fit background shape comparisons
- Added Appendix Q.1: Asimov fits to `ELD`, number of jets and 1 bin distributions, and Q.4: Final fits for all scenarios
- Added Appendix X: study of the effects of using tight photons in the object overlap removal procedure

- **version 4.1, 13 April 2018**
  - The draft to be circulated in Top WG.
  - Chapter 6.4 Prompt photons background
    - Updated $W\gamma$ and $Z\gamma$ background section to include some more plots and χ^2 values.
  - Chapter 10, Systematic uncertainties
    - Included information on the ROOT TH1:Smooth method which we now use for some systematic sources.
  - Chapter 13, Results
    - Updated fits using custom jet flavour composition histograms.
    - Added an "inclusive combined" fit result.
- **version 4.0, 3rd April 2018**
  - Chapter 3, Data and simulation samples
    - 3.2 Add more explanations on effect of the removal of double counting events and comparison with other methods
  - Chapter 4, Object definition
    - 4.2 Prompt Photon Tagger (PPT):
      - Systematics now use slicing in η and p_{T}
      - Shape discrepancies from PPT prompt CR now not only used as systematics, but also as corrections for prompt contributions in the signal regions
      - Additional systematic for hadronic fake to account for isolation uncertainty
  - Chapter 6, Background descriptions
    - 6.1 Background due to mis-reconstructed electrons:
      - Updated $e \to \gamma$ fake numbers/plots/tables to their latest version.

Not reviewed, for internal circulation only

        - Added subsection 6.1.4 to explain an additional data-driven correction to $e \rightarrow \gamma$ fake.
    - 6.2: Background due to hadronic fakes
        - 6.2.1: hadronic fake SF's are now separated for converted and unconverted photons and are provided in $\eta - p_{\mathrm{T}}$ 2D bins.
        - 6.2.2: Added extrapolation uncertainty for using hadronic fake SF's of the single lepton channel to the dilepton channel.
- Chapter 8, Systematic uncertainties
    - Added signal and $t\bar{t}$ modelling systematic + red/blue plots.
    - Summarised all systematics and added many plots.
- Chapter 9, Cross section definitions
    - 9.1: Fiducial region: Updated table of acceptance and correction factor
    - 9.4: Differential cross section: Added a new variable (minimum $\Delta R(\text{lepton}, \gamma)$) for the inputs to the unfolding (efficiencies and migration matrices).
- Chapter 11, Unfolding
    - Added new variables minimum $\Delta R(\text{lepton}, \gamma)$ in both channels, $\Delta\eta(\text{lepton}, \text{lepton})$ and $\Delta\phi(\text{lepton}, \text{lepton})$ in dilepton channel.
- Chapter 12, Theory prediction
    - Rewrite almost the whole chapter
    - Document the new k-factor approach and the final predicted fiducial cross-sections
- Chapter 13, Results
    - 13.2 Fiducial cross section:
        - Updated cross section results
    - 13.2 Differential cross section:
        - Updated differential cross section results, plotting and systematic tables.
        - Added missing signal modelling systematics.
- Chapter 14, EFT
    - Moved the entire chapter to appendix (Appendix T).
- Remove appendix B (Cut flow tables), since it was outdated and in general not useful
- Remove the cut optimisations for met + mwt from appendix C.1 (Cut optimisation in the single lepton channel, now Appendix C.1), since it was outdated and buggy.
- Added an appendix for extra material for hadronic fake background (Appendix F)
- Added an appendix for extra material for unfolding (Appendix O).
- Added appendix V: Studying large weights in $W\gamma$ and $Z\gamma$ MC samples

Not reviewed, for internal circulation only

- Added appendix W: Checked for prompt photons from pile-up.

- **version 3.0, 21 December 2017**
  - Chapter 6, Background description
    - 6.1 p_{T}-η 2D SFs are provided as well as their systematics studied. $e \to \gamma$ fake PPT shape systematics is also studied.
    - 6.2 Double ratio factor to correct for photon's isolation and identification correlation is added.
    - 6.3.1 Different parameterisations for Multijet estimation for single lepton channel are studied. Nominal and systematic variations of parameterisation are given.
  - Chapter 7, Pre-fit distributions with SF's
    - Added a chapter to show single lepton after the SF's from background section have been applied.
  - Chapter 8, Systematics
    - 8.2.3 Further section added describing the methods used to prune and smooth nuisance parameters.
    - 8.3 Propagation of uncertainties into the unfolding framework.
  - Chapter 9, Cross section definitions
    - 9.4 The correction and accpetance values are updated with systematics.
    - 9.4 The description of differential cross section is simplified. Inputs for unfolding (efficiencies and migration matrices) are updated.
  - Chapter 10, Fit strategy
    - 10.5 The description of fit strategy for differential cross section is simplified.
  - Chapter 11, Unfolding
    - 11.1 Closure/Pull/Stress tests of the iterative bayesian unfolding are added
  - Chapter 11, Theory prediction
    - 11.3 The NLO fiducial cross section values are updated with systematics.
  - Chapter 13, Results
    - 12.1 Final μ results have been added with previous studies moved to the appendix .
    - Included Post fit plots, yields, and systematic studies.
    - 12.3 Added differential cross section results.

- **version 2.2, 04 August 2017**
  - General: Changed the order of chapters
  - Chapter 4, Object definition

Not reviewed, for internal circulation only

    - Chapter 4.2.1: PPT training: Included selection cuts and trigger information, added plots for PPT applied to $t\bar{t}\gamma$ ntuples
    - Chapter 4.2.2: Systematic uncertainties for PPT from hadronic fake CR.
  - Chapter 5, Signal region definition
    - Included many more plots including ph_pt, jet_pt, nbjets, etc.
    - Included plots showing b-tags vs ELD output.
  - Chapter 6, Background description
    - 6.2 Background due to hadronic fakes: $e \to \gamma$ fake is also subtracted from data now, plots and numbers are updated with new lumi syst and taking into account the $e \to \gamma$ fake SF.
    - 6.4.1 Prompt photons background, $Z\gamma$ Validation region. Redone with NLO samples.
  - Chapter 8, Cross section definitions
    - 8.4 Differential cross section: Added new potential for differential study.
- **version 2.1, 18 August 2017**
  - first draft being given to the EB

Not reviewed, for internal circulation only

1 Introduction

The study of the top quark particle is one of the major fields in modern particle physics. Due to its heavy mass and short life time, the top quark particle plays an important role in the Standard Model, and is a good candidate to interact with new physics in many Beyond Standard Model theories. The precise measurement of its properties is therefore an important topic, and has been the subject to a lot of analyses, as for example in [1]. The top quark coupling measurements can uncover new physics. In particular, the electroweak top quark coupling with the photon, $t\gamma$, can be affected through anomalous dipole moments of the top quark [2–7]. The top-photon coupling can be probed in the top quark pair production associated with a photon ($t\bar{t}\gamma$).

The first evidence of $t\bar{t}\gamma$ production was obtained at the Tevatron [8], and since then, several measurements have been realized at the LHC experiments. The previous results are summarized in Table 1.

Table 1: Previous measurements of the $t\bar{t}\gamma$ cross-section. If not splitted, the uncertainties is the combined uncertainties. Otherwise, the keywords *sy*, *st*, *lu*, *sc* and *pdf* stands for, respectively, the uncertainty due to systematical uncertainties, statistical uncertainties, luminosity, scale and PDF. The last-but-one column is the relative combined uncertainty of the observed result.

Luminosity	Measurement	p_T^γ cut [GeV]	Theory [fb]	Observed result [fb]	Obs. rel. uncert.	Ref.
Tevatron (p-$\bar{p}$), CDF, 1.96 TeV						
6.0fb^{-1}	inclusive	10	170 ± 30	180 ± 70(st) ± 40(sy) ± 10(lu)	8.1%	[8]
LHC (p-p), ATLAS, 7 TeV						
1.04fb^{-1}	inclusive non-all-hadronic	8	2100 ± 400	2000 ± 500(st) ± 700(sy)	43.0%	[9]
4.59fb^{-1}	fiducial ℓ+jets	20	48000 ± 10000	63000 ± 8000(st)$^{+17000}_{-13000}$(sy)	27.0%	[10]
LHC (p-p), CMS, 8 TeV						
19.7fb^{-1}	inclusive μ+jets	20	1800 ± 500	2400 ± 200(st) ± 600(sy)	26.3%	[11]
19.7fb^{-1}	inclusive ℓ+jets	25	591 ± 71(sc) ± 30(pdf)	515 ± 108	21.0%	[12]
19.7fb^{-1}	fiducial ℓ+jets	25	-	127 ± 27	21.2%	[12]
LHC (p-p), ATLAS, 8 TeV						
20.2fb^{-1}	fiducial ℓ+jets	15	151 ± 24	139 ± 7(st) ± 17(sy)	13.2%	[13]
20.2fb^{-1}	differential ℓ+jets	15	-	-	-	[13]

The final state of the $t\bar{t}\gamma$ process is similar to the final state of the usual top quark pair production, but contains an additional prompt energetic photon. In this final state, the photon can have different origin. It can be the result of Initial State Radiation (ISR) of the initial partons, which is mainly obtained when the top quark pair is produced by quark-antiquark annihilation. It can be the result of a radiation directly from the top quark, which is the relevant contribution for the study of the top-photon coupling. Finally, it can also be the result of Final State Radiation (FSR), radiated by the decay products of the top quark: the W-boson, the b-quark or the decay products of the W-boson. Theoretical NLO cross-section calculations have been performed considering the unstable top quark hypothesis, therefore including the latter case [14]. Some representative Feynman diagrams for the photon radiated during top production and radiated during top decay are shown in Figure 1 and Figure 2, respectively. The considered signal in this analysis is originating from any of these sources. However, some selections are intended to reduce the contribution of the photon radiated from the decay products of the top quark, enhancing the contribution originating from the top-photon coupling.

The strategy is to select a single lepton or dilepton final state of $t\bar{t}$, and request the presence of a photon. In previous $t\bar{t}\gamma$ ATLAS analyses, the dominant background was the contribution due to non-prompt photons, and a maximum-likelihood fit using templates on a discriminative variable (the photon isolation

Not reviewed, for internal circulation only

Figure 1: Some representative Feynman diagrams for the $t\bar{t}\gamma$ process where the photon is radiated during top-quark production (a,b,c) or radiated from initial partons (d).

Figure 2: Some representative Feynman diagrams for the $t\bar{t}\gamma$ process where the photon is radiated during top-quark decay.

in the tracker) was performed. In this analysis, an analysis-independent multivariate analysis discriminant optimized to select prompt photons has been introduced. This discriminant enters in a second multivariate analysis discriminant, analysis-dependent and exploiting event-level variables, used as final variable on which the fit is performed. The final states considered are the single lepton channels (e+jets and μ+jets) and dilepton channels (ee, $e\mu$, $\mu\mu$), including electrons and muons from possible leptonic τ lepton decay. The inclusive and differential cross-sections for the different channels are extracted. The differential cross-sections are presented as a function of the photon transverse momentum and as a function of the photon pseudo-rapidity. An unfolding procedure is performed, and the results are compared with the theoretical prediction.

This note is organised as follows. Section 2 provides a brief introduction to the ATLAS detector. Section 3 contains the description of the data and simulated samples used in this analysis. Section 4 describes the object reconstruction and selection, including the description of the prompt-photon discriminant. Section 5 describes the event selection for the signal region, including the description of the event-level discriminant. Section 6 describes the background processes studies, detailing the estimation of the fake photon contribution due to mis-reconstructed hadrons and mis-reconstructed electrons, the estimation of the prompt photon background and the estimation of the contribution from fake leptons. The pre-fit distributions with all the scale factors applied to them are shown in Section 11. Section 10 lists the systematic uncertainties taken into account in this analysis. Section 7 describes the method used to extract the total cross-section from the measured fiducial cross-section, followed by Section 8 which describes the fit procedure and a dedicated chapter, Section 9, explaining the intensive study of the unfolding procedure. Section 12 presents the theory prediction for the signal. Section 13 presents the final results. Finally, the

Not reviewed, for internal circulation only

conclusion is in Section 14.

The appendices include studies on removal of the double counting of events between samples (Appendix A), additional information on the prompt photon tagger (Appendix B), the cuts optimization studies (Appendix C), the study of the signal contributions for photon radiated from top production or radiated from top decay (Appendix D), a comparison of before and after applying the photon isolation cut (Appendix E), the additional studies for hadronic fake background (Appendix F), the additional validation regions for $Z\gamma$ (Appendix **??**) and $W\gamma$ (Appendix **??**), the binning optimisation of the discriminating variable ELD (Appendix I), the study of the impact of the PDF shape uncertainties on the final fit (Appendix J), the study of the impact of varying the single top, diboson and $t\bar{t}V$ systematic uncertainties on the final fit (Appendix K), the extra materials about the unfolding (Appendix N, O , and P), the extra material from result section (Appendix Q), some discussions about new constrains on Effective Field Theory operators and their validation (Appendix T, U and U.4), studies of the issue of large weights in the $V + \gamma$ samples (Appendix V), the study on the modelling of prompt photons from pile-up (Appendix W), and finally the detailed list of the MC samples (Appendix Z).

Not reviewed, for internal circulation only

2 ATLAS detector

The ATLAS detector [15] consists of three main components. The innermost component is the Inner Detector (ID) which surrounds the beam pipe and is used for tracking charged particles. 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 [16, 17]. The ID is surrounded by a superconducting solenoid operating with a magnetic field of 2 T and the calorimeter which is divided into two subsystems. The first subsystem is the electromagnetic calorimeter (EMC) and the second is the hadronic calorimeter (HCAL). The outermost layer is the third main component of the ATLAS detector: the Muon Spectrometer (MS). The ID provides tracking information from pixel and silicon microstrip detectors in the pseudorapidity[1] range $|\eta| < 2.5$ and from a transition radiation tracker (TRT) covering $|\eta| < 2.0$. The 2 T magnetic field of the superconducting solenoid bends the charged particles for the momentum measurement. The EMC uses lead and liquid argon (LAr) and is divided into barrel ($|\eta| < 1.475$) and end-cap ($1.375 < |\eta| < 3.2$) regions. The HCAL is provided by 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 located in toroidal fields of approximately 0.5 T and 1 T in the central and end-cap regions of the ATLAS, respectively. The MS is also instrumented with separate trigger chambers covering $|\eta| < 2.4$.

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

Not reviewed, for internal circulation only

3 Data and simulation samples

3.1 Samples description

This analysis has been realized with proton-proton collision data harvested during 2015 and 2016, at a center-of-mass of energy of $\sqrt{s}$ = 13 TeV. Only luminosity blocks of recorded data satisfying good quality criteria in the detector have been considered. This corresponds to a total integrated luminosity of 3212.96 pb^{-1} in 2015 and 32884.6 pb^{-1} in 2016. The uncertainty for the combined integrated luminosity is 2.1%. It is derived, following a methodology similar to that detailed in [18], from a calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016.

The $t\bar{t}\gamma$ signal sample has been simulated for the semileptonic and dileptonic channels of $t\bar{t}$ decays at leading-order, with the MG5_aMC@NLO generator [19]. The generation is performed at LO, since higher order predictions of $t\bar{t}\gamma$ are quite complicated to compute, consuming high power and CPU, and thus impossible to generate. The NNPDF2.3LO parton distribution function (PDF) set [20] is used. The parton showering is simulated using PYTHIA 8 [21], and the A14 tune [22]. ATLAS recommendations for the SM and EW inputs are used with the particle masses listed in Table 2.

Table 2: Masses of quarks and leptons set for the simulation of the $t\bar{t}\gamma$ sample.

Particle	Quarks			Leptons		
	light-quarks (u, d, c, s)	t-quark	b-quark	electron	muon	tau
Mass [GeV]	0	172.5	0	0	0.105	1.777

In the simulation, the fine structure constant of 1/137 is used. The renormalisation and the factorization scales are dynamic and correspond to half the sum of the transverse masses of all the final state particles:

$$\mu = \frac{1}{2}\sum_{f}\sqrt{m_f^2 + p_{T,f}^2} \tag{1}$$

where f runs over all final state particles, m_f is the rest mass of particle f, and $p_{T,f}$ is the transverse momentum of particle f. In order to avoid infrared and collinear singularities, a set of kinematic cuts listed below have been applied. The photon and the charged lepton transverse momenta and pseudo-rapidities are required to be larger than 15 GeV and smaller than 5.0, respectively. The ΔR distance between the photon and all other charged particles in the final state should be larger than 0.2. The total cross section defined by these cuts and calculated by MadGraph for the $t\bar{t}\gamma$ signal sample is found to be 4.623 pb. It is checked that the MadGraph calculation is consistent with theoretical computation in a slightly smaller phase-space. More details can be found in Section 12.1.

The MC sample for the inclusive $t\bar{t}$ production is generated with POWHEG-BOX v2 [23] interfaced to PYTHIA8, using the A14 tune. This tune uses the NNPDF2.3LO set. The production of W-and Z-bosons $+\gamma$ as well as the other vector boson production samples (W-and Z-bosons + jets) are simulated using SHERPA [24] 2.2.2 and 2.2.1, respectively, with the NNPDF30NNLO pdf set.

The single top-quark t-, s- and Wt-channel samples are produced by POWHEG-Box v1 generator. POWHEG matrix element events are generated with CT10 (NLO) PDF set [25], interfaced with PYTHIA 6 and using the Perugia2012 tune [26]. An additional photon radiation in the fragmentation process is simulated with PHOTOS [27]. The Wt-channel with inclusive diagram removal is used to estimate the

Not reviewed, for internal circulation only

background while the Wt-channel with inclusive diagram subtraction is used for uncertainties evaluation. The WW-, WZ- and ZZ-diboson samples are simulated using SHERPA 2.1 with the CT10(NLO) pdf set.

The $t\bar{t}V$ samples are generated with MadGraph 2.2.1, using NLO matrix elements and the NNPDF3.0NLO PDF set. Showering is performed using Pythia 8.210 and the A14 tune. The samples are normalised to the NLO cross section computed with MadGraph.

A full list of MC samples along with higher order cross-sections and k-factors can be found in Appendix Z.

3.2 Removal of the double counting of events in X+jets and X+γ samples

Because the showering procedure will add photon radiations, the $t\bar{t}$ and V+jets samples can contain events already taking into account by the $t\bar{t}\gamma$ and $V\gamma$ samples. This overlap is removed by vetoing the events in $t\bar{t}$ and V+jets where the selected photon is originating from the hard interaction.

The reconstructed photons are associated to their corresponding truth particle via the *Truth to Cluster* matching algorithm provided by the *MCTruthClassifier* package. This package also provide a classification of the origin of the photon, using two variables: *type* and *origin*. If the reconstructed photon is associated to a truth electron, or if a truth electron (with $p_{\mathrm{T}} > 10$GeVand $|\eta| < 3$) is present in a 0.05 cone around the reconstructed photon, the photon is considered as $e \rightarrow \gamma$ fake. If the reconstructed photon is associated with the type *BkgPhoton*[2], with an origin corresponding to a hadron, the photon is considered as hadronic fake (the exact condition is written $23 \leq \text{origin} \leq 35 or \text{origin} == 42$ && type == 16). All other cases are considered as photon originating from the hard interaction. For future usage, it is recommended to also add the photons with type *Hadron*[3] to the hadronic fake category, as discussed in Appendix A.1.

After this double counting removal, the $t\bar{t}$ and V+jets samples will only contribute to hadronic fake and $e \rightarrow \gamma$ fake backgrounds, while the dedicated $t\bar{t}\gamma$ and $V\gamma$ samples only provide the events containing prompt photons. Furthermore, this allows the classification of each event into the following categories: $t\bar{t}\gamma$ (contributions from only our signal sample), *hadronic fake*, $e \rightarrow \gamma$ *fake*, *Other prompt* and *Lep fake*. This procedure rejects less than 1% of events in $t\bar{t}\gamma$ and $V\gamma$ samples (with the exception of the $Z\gamma$ in the e+jets channel which can also contribute to the $e \rightarrow \gamma$ fake category when the prompt photon is not reconstructed/selected). For $t\bar{t}$ and V+jets samples, the prompt photon contribution (respectively $\sim 75\%$ and $\sim 60\%$ of the events passing the signal selection) is subject to larger statistical uncertainty than the dedicated samples and its modelling is less trusted, it is why the $t\bar{t}\gamma$ and $V\gamma$ samples are used. But the yields and shapes of events rejected by this procedure in those samples are compatible with what is expected from $t\bar{t}\gamma$ and $V\gamma$.

Even if the details of this algorithm has a low impact on the physics measurement, as the $e \rightarrow \gamma$ fake and hadronic fake are normalized from data-driven control region, several check, documented in Appendix A.2, have been done.

[2] This type *BkgPhoton* corresponds to a photon not originating from ISR/FSR, from a boson or from a lepton.
[3] This type *Hadron* corresponds to a reconstructed photon associated to a hadron truth paticle.

Not reviewed, for internal circulation only

4 Objects definition

4.1 Objects selection

The main physics objects considered in this analysis are electrons, muons, jets, b-jets, missing transverse energy and photons. The reconstruction of those objects follows the standard methods in the ATLAS Collaboration. Typical recommended corrections and calibrations are applied, using the top analysis group software AnalysisTop [28] in its version 2.4.33. The object definitions of leptons, jets, b-jets and missing transverse momentum are chosen to be as close as possible to the ones of the usual $t\bar{t}$ analyses. The photon reconstruction, identification and selection have been optimized for this analysis.

The electrons are reconstructed from energy deposits in the central region of the Electromagnetic Calorimeter associated with reconstructed tracks from the Inner Detector, and are required to satisfy the `TightLH` identification criteria and the `Gradient` isolation criteria. Scale factors based on comparison between data and MC are used to correct the simulation. Only electrons with calibrated $E_{\mathrm{T}} > 25$ GeV and $|\eta_{\mathrm{clus}}| < 2.47$, excluding the crack region $1.37 < |\eta_{\mathrm{clus}}| < 1.52$, are considered ($\eta_{\mathrm{clus}}$ is the pseudo-rapidity of the calorimeter cluster associated to the electron) (later, the event selection also requires that the electron candidate matching the electron trigger also has a consistent E_{T} threshold). The recommended requirements on the impact parameter variables are also applied: the transverse impact parameter with respect to the beam line $|d_0|/\sigma(d_0)$ is required to be lower than 5 and the difference in the z-axis between the track origin and the primary vertex when expressed at the beam line $|\Delta z_0 \sin(\theta)|$ is required to be lower than 0.5 mm. The reconstruction, identification and scale factor extraction methods are described in [29], the isolation is described in [30], and the calibration is described in [31].

The muons are reconstructed with the combined algorithm, using the track segments in the various layers of the muon spectrometer and the tracks in the Inner Detector. They are required to have the `Medium` identification quality criteria and the `Gradient` isolation criteria. Scale factors based on comparison between data and MC are used to correct the simulation. Only muons with calibrated $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$ are considered (later, the event selection also requires that the muon candidate matching the muon trigger also has a consistent p_{T} threshold). The recommended requirements on the impact parameter variables are also applied: the transverse impact parameter with respect to the beam line $|d_0|/\sigma(d_0)$ is required to be lower than 3 and the difference in the z-axis between the track origin and the primary vertex when expressed at the beam line $\Delta z_0 \sin(\theta)$ is required to be lower than 0.5 mm. The reconstruction, identification, calibrations and scale factor extraction methods are described in [32], the isolation is described in [30].

The photons are reconstructed from energy deposits in the central region of the Electromagnetic Calorimeter. If the cluster is considered not matched to any reconstructed tracks in the Inner Detector, an unconverted photon candidate is reconstructed. If the cluster is matched with reconstructed tracks that are consistent with originating from a photon conversion and if in addition a conversion vertex is reconstructed, a converted photon candidate is reconstructed. Both kind of photons are considered in this study. The photons are required to satisfy the `Tight` identification criteria. The photon is also required to be isolated, using the `FixedCutTight` working point defined as $p_T^{\mathrm{topoetcone40}} < 0.022\, p_T(\gamma) + 2.45$ GeV and $p_T^{\mathrm{cone20}}/p_T(\gamma) < 0.05$. Scale factors based on comparison between data and MC are used to correct the simulation. Only photons with calibrated $p_{\mathrm{T}} > 20$ GeV and $|\eta_{\mathrm{clus}}| < 2.37$, excluding the crack region $1.37 < |\eta_{\mathrm{clus}}| < 1.52$, are considered ($\eta_{\mathrm{clus}}$ is the pseudo-rapidity of the calorimeter cluster associated to the photon). The reconstruction, identification and scale factor extraction methods are described in [33], the isolation is described in [30], and the calibration is described in [31].

Not reviewed, for internal circulation only

The jets are reconstructed using the anti-k_t algorithm [34] with a distance parameter $R = 0.4$ (in $\eta - \phi$ space) and are reconstructed from topological calorimeter clusters [35]. The jet energy scale and jet energy resolution are calibrated [36]. The jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. In order to reject jets from pile-up or other primary vertices, jets are required to have a Jet Vertex Tagger discriminant [37] larger than 0.59 for jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$. This Jet Vertex Tagger efficiency is corrected by a scale factor extracted from data and MC comparison.

The b-tagging algorithm to identify jets from b quark hadronisation is the MV2c10 algorithm [38], which is based on a boosted decision tree (BDT) using the output weights of the JetFitter, IP3D and SV1 algorithms as input. In the event selection, the 77% working point is used, corresponding to a cut at 0.645925. In addition, the pseudo-continuous b-tagging is used as variables in the multivariate analysis. The efficiency calibration, corresponding to the recommendation for Summer 2017 conferences (`2016-20_7-13TeV-MC15-CDI-2017-06-07_v2.root`), is applied using a scale factor.

The energy carried by the neutrinos is accounted for in the reconstructed missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ [39]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is computed as the vector sum over all topoclusters in the event (considered as a massless particle originating from the center of the detector), and is further refined by applying corrections for identified electrons, muons, photons and jets.

An overlap removal procedure is applied to avoid that a same deposit in the calorimeters or track is used to reconstruct two different objects. The followed overlap removal procedure corresponds to the procedure recommended for Moriond 2017. First electrons that are sharing their track with a muon candidate are removed. Then, all the jets that are close to an electron in the $\eta - \phi$ plane ($\Delta R < 0.2$) are removed. Then, all the electrons that are close ($\Delta R < 0.4$) to a remaining jet are removed. For muons and jets that are close, the jet is removed if it has no more than 2 associated tracks and is within $\Delta R < 0.2$ of the muon. Remaining muons are removed if they are close to a jet ($\Delta R < 0.4$). All the photons that are close ($\Delta R < 0.4$) to a remaining electron or muon are removed. Finally, all the jets that are close ($\Delta R < 0.4$) to a remaining photon are removed. The photon collection used for this last step contains looser photon candidates as defined before, with $p_{\mathrm{T}} > 15$ GeV and no identification and isolation criteria, as discussed in Appendix X.

The object selections are summarised in Table 3.

Object	Definition	p_{T} (E_{T})	$\lvert\eta\rvert$	Removed
Electron	`TightLH`, `Gradient`, $\lvert d_0\rvert/\sigma(d_0) < 5$, $\lvert\Delta z_0 \sin(\theta)\rvert < 0.5mm$	25 GeV	[0, 1.37] & [1.52, 2.47]	(1) if sharing track with muon (3) if $\Delta R < 0.4$ wrt jet
Muon	`Medium`, `Gradient`, $\lvert d_0\rvert/\sigma(d_0) < 3$, $\lvert\Delta z_0 \sin(\theta)\rvert < 0.5mm$	25 GeV	[0, 2.5]	(5) if $\Delta R < 0.4$ wrt jet & jet has ≥ 2 tracks
Photon	`Tight`, `FixedCutTight`	20 GeV	[0, 1.37] & [1.52, 2.37]	(6) if $\Delta R < 0.4$ wrt electron/muon
Jet	anti-k_t, R = 0.4, JVT > 0.59 when $p_{\mathrm{T}} < 60$ GeV and $\lvert\eta\rvert < 2.4$, b-jets: 77% WP	25 GeV	[0, 2.5]	(2) if $\Delta R < 0.2$ wrt electron (4) if $\Delta R < 0.2$ wrt muon & jet has < 2 tracks (7) if $\Delta R < 0.4$ wrt loose photon

Table 3: Summary of the object selection for electrons, muons, photons and jets, and their overlap removal (last column), with their removal order marked by numbers.

Not reviewed, for internal circulation only

4.2 Prompt photon tagger

Photons originating from hadronic decays are uncorrelated to the initial hard interaction and are therefore considered background. Another important background contribution in this analysis are jets/hadrons misidentified as `Tight` photons. Both contributions are called *hadronic fakes* and are further discussed in Section 6. Photons from the hard interaction, on the other hand, are called *prompt photons*. To increase the purity of prompt photons in the set of `Tight` photon candidates and to suppress the contamination of hadronic fakes, an analysis-independent neural network (`NN`) was developed to classify photons, called *prompt-photon-tagger* (`PPT`).

The `NN`'s architecture was developed and trained using Keras [40]. For reconstruction and application of the `NN`, the LightWeight Trained Neural Network library or LWTNN [41] is used. The input variables are chosen to be a subset of shower shape variables also used for photon identification, summarised in Table 4 [42]. The `PPT` classifies photon candidates in such a way that values towards 1 indicate prompt photons, whereas values towards 0 indicate hadronic fakes. Hence, the `PPT` is a binary classification tool. In this analysis, the full continuous `PPT` discriminant variable is used as an input to another event-level `NN`, further described in Section 5.2.

In Section 4.2.1 the `NN` training and results for the `PPT` are presented, and in Section 4.2.2 the procedure to estimate systematic uncertainties of the `PPT` is described.

name	description
Hadronic leakage	
R_{had} or R_{had1}	Transverse energy leakage in the hadronic calorimeter normalised to transverse energy of the photon candidate in the electromagnetic calorimeter. In the region $0.8 \leq \lvert\eta\rvert \leq 1.37$ the entire energy of the photon candidate in the hadronic calorimeter is used (R_{had}), while in the region $\lvert\eta\rvert < 0.8$ and $\lvert\eta\rvert > 1.37$ the energy of the first layer of the hadronic calorimeter is used (R_{had1})
Energy ratios and width in the second layer of electromagnetic calorimeter	
R_η	Energy ratio of 3×7 to 7×7 cells in the $\eta \times \phi$ plane.
R_ϕ	Energy ratio of 3×3 to 3×7 cells in the $\eta \times \phi$ plane.
$w_{\eta 2}$	Lateral width of cluster in $\eta \times \phi = 3 \times 5$: $\sqrt{\frac{\sum_i E_i \eta_i^2}{\sum_i E_i} - \left(\frac{\sum_i E_i \eta_i}{\sum_i E_i}\right)^2}$
Energy ratios and widths in the first (strip) layer of electromagnetic calorimeter	
$w_{\eta 1}(w_{s3})$	Energy weighted width using 3 strips around the maximum: $\sqrt{\frac{\sum_i E_i (i - i_{\max})^2}{\sum_i E_i}}$
$w_{\text{tot,s1}}(w_s)$	Energy weighted width using 20 strips around the maximum, see $w_{\eta,1}$.
f_{side}	Energy within 7 strips without 3 central strips normalised to energy in 3 central strips.
E_{ratio}	Ratio between difference of first 2 energy maxima divided by their sum ($E_{\text{ratio}} = 1$ if there is no second maximum).
ΔE	Difference between the second energy maximum and the minimum between first and second maximum ($\Delta E = 1$ if there is no second maximum).

Table 4: Summary of shower shape variables used for cut-based photon tight identification.

Not reviewed, for internal circulation only

trigger name
HLT_g10_loose
HLT_g15_loose_L1EM7
HLT_g20_loose_L1EM12
HLT_g25_loose_L1EM15
HLT_g35_loose_L1EM15
HLT_g40_loose_L1EM15
HLT_g45_loose_L1EM15
HLT_g50_loose_L1EM15
HLT_g60_loose
HLT_g70_loose
HLT_g80_loose
HLT_g100_loose
HLT_g120_loose
HLT_g140_loose

Table 5: List of triggers to be passed for event selection in the PPT training.

4.2.1 PPT training

The photon candidates used to train the PPT are required to pass $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.37$, excluding the crack region where $1.37 < |\eta| < 1.52$. In addition, the candidates must fulfil the Tight photon identification criteria.[4] To distinguish prompt photons and hadronic fakes, truth information is taken into account using MCTruthClassifier information. Prompt photons are required to originate from the hard interaction, that is, from either the matrix element calculation or the showering, as these types cannot be distinguished in the experiment. Objects identified as photons, but originating from hadronic decays or misidentified hadrons/jets are classified as hadronic fakes. Photons radiated by charged final-state leptons after hadronisation are not taken into account. The MC-generated events used for training the PPT are QCD-Compton processes containing prompt photons, as well as di-jet events with hadronic fakes. These samples are independent of those of the $t\bar{t}\gamma$ analysis and are listed in Table 70. Selected events have to pass triggers listed in Table 5. Good quality criteria in the detector are required and at least one photon has to be present according to the object definitions given above.

Variables were chosen as an input to the PPT based on their separation powers between prompt photons and hadronic fakes in MC events. The separation $\mathcal{S}$ is calculated according to

$$\mathcal{S} = \frac{1}{2} \sum_{i \in \mathrm{bins}} \frac{(s_i - b_i)^2}{(s_i + b_i)}, \tag{2}$$

where s_i and b_i are the number of prompt photons and hadronic fakes in bin i of the variable, respectively.

The calculated separation powers for all shower shape variables used as features, namely R_{had}, R_η, R_ϕ, $w_{\eta,2}$, $w_{\eta,1}$ and f_{side}, are listed in Table 6. These variables were found to have the largest separation powers. Separation plots are shown in Figure 3. One key separation feature, reflected in the chosen input variables,

[4] It is worth noting that, due to the analysis independence of the PPT, photon definitions do not have to match those used in the analysis exactly.

Not reviewed, for internal circulation only

variable	separation
R_{had}	3.33 %
R_{η}	4.83 %
R_{ϕ}	7.01 %
$w_{\eta 2}$	2.01 %
$w_{\eta 1}$	4.14 %
f_{side}	7.21 %

Table 6: Separation power of shower shape variables calculated according to Eq. (2).

is the additional hadronic activity around hadronic fakes which leads to broader energy deposits in the electromagnetic calorimeter and larger hadronic leakage.

Figure 4 shows correlation matrices for the input variables and the other shower-shape variables, both for prompt photons and hadronic fakes. Between some of the input variables, for instance between R_{η} and $w_{\eta 2}$, correlations are as strong as −59.77% and −53.89% for prompt photons and hadronic fakes, respectively. Strong correlations are one of the reasons why simpler techniques, such as rectangular cuts on shower-shape variables, cannot be used to increase the separation between prompt photons and hadronic fakes efficiently. MVA techniques, on the contrary, take correlations between variables into account and are therefore applicable in this case.

Optimisations were done to determine the best NN architecture: the current PPT consists of 6 layers. The first, third and fifth layer contain 64, 40 and 52 neurons, respectively, of which the first uses a *rectified linear unit* activation function while the other two use a *softmax* activation function. The second and fourth layers are *batch normalisation* layers [43]. The final output layer comprises one single neuron with a *sigmoid* activation function.

About one million prompt photons and 200k hadronic fakes were used for training and testing the PPT, splitting the full set of samples into a training set with 80% and a test set with 20% of all samples. Figure 5 shows the visualisation of the receiver operating characteristic (ROC) curves for both training and test samples. The ROC curves show the background rejection and the signal efficiency for different working points of the PPT. The two curves for training and test set overlay, therefore, the PPT can be generalised and applied to unseen data, that is, data that was not used within the training process.[5] Both under- and overtraining could be avoided. The overall classification performance measure is given by the *Area Under Curve* (AUC) which is 0.8693 and 0.8690 for the training and test sets, respectively. The ROC curves and therefore the classification performances for the two sets are very similar. AUC values for different NN architectures and feature sets, all showing poorer performances, can be found in Appendix B.2. Figure 6 shows the classifier output for signal and background, training and test MC. Good agreement can be seen between seen and unseen data and the response to the NN.

In addition, the stability of the PPT training and the dependence of its performance on the training and test sets were evaluated by applying a 5-fold cross-validation. All samples were split into five orthogonal subsets. Each subset i was then once used as test set after the classifier had been trained on the remaining four subsets. Figure 5 shows the five ROC curves of the cross-validation. The largest obtained deviation is

[5] A necessary condition to draw this conclusion from overlaying training and test ROC curves is that the feature distributions in the training and test set agree with each other. That is the case within statistical uncertainties. The corresponding plots can be found in Appendix B.6.

Not reviewed, for internal circulation only

≈ 0.0090, which is small compared to the nominal performance. Hence, the trained PPT can be assumed to be stable with a reliable classification performance.

In addition, the NN can be applied to the $t\bar{t}\gamma$ analysis MC samples described in Section 3 before event-level cuts. Figure 7 shows the PPT output values for different photon truth particle types. Truth type 20 corresponds to $e \rightarrow \gamma$ fake photons. A value of 10 means the photon is classified as hadronic fake. For any value below 10, the photon is classified as signal-like, the exact source of which cannot be determined experimentally. For those signal-like photons a clear trend towards the signal-like region can be seen. The distribution for hadronic fake photons is flatter, with an increasing slope towards the background-like region. $e \rightarrow \gamma$ fake photons are classified to be more signal-like than background-like. This is expected as $e \rightarrow \gamma$ fake photons were not considered at any point in the PPT training. → ?

Figure 8 shows the PPT's dependence on isolation as applied to Tight ID photons. Each cut represents an isolation working point. FixedCutLoose is the loosest requirement while FixedCutTight is the tightest requirement. There is a small deviation from the photons in which a cut is applied and a clear trend in going from non-isolated photons to any type of isolation requirement. Since this tool is trained in an analysis independent way (i.e., trained with no requirement on isolation) this is expected and motivates the 3-D scale factors and systematics derived in the next section.

Not reviewed, for internal circulation only

Figure 3: Distribution of shower-shape variables R_η, R_ϕ, R_{had}, f_{side}, $w_{\eta,1}$ and $w_{\eta,2}$ for prompt photons (red) and hadronic fakes (blue), both normalised to unity.

Not reviewed, for internal circulation only

(a) Shower-shape correlations for prompt photons.

(b) Shower-shape correlations for hadronic fakes.

Figure 4: Correlations between shower shape variables as described in Table 4.

Not reviewed, for internal circulation only

(a) ROC curves of the training and test sets.

(b) ROC curves of the test sets in a 5-fold cross-validation.

Figure 5: ROC curves of training and test sets as well as those obtained in a 5-fold cross-validation.

Figure 6: NN output for training and testing samples.

Not reviewed, for internal circulation only

(a) Photons without isolation requirements.

(b) Photons passing `FixedCutTight` requirements.

Figure 7: PPT output distributions for different photon truth particle types. The distributions are obtained by applying the PPT to the $t\bar{t}\gamma$ analysis MC samples before event-level cuts. All distributions were normalised to unity individually.

Figure 8: PPT output distributions for increasing isolation cuts. The distributions are obtained by applying the PPT to the $t\bar{t}\gamma$ analysis MC samples before event-level cuts, but includes object definitions such as the `Tight` ID photon requirement.

Not reviewed, for internal circulation only

4.2.2 Systematic uncertainties for the PPT

Systematics for the PPT within the $t\bar{t}\gamma$ analysis are derived conservatively from three dedicated PPT control regions: two hadronic fake control regions, one of which targets non-isolated photons, the other isolated photons with inverted ID requirements. In a third region, systematic uncertainties for prompt photons are derived. For all three regions, the analysis MC samples and experimental data introduced in Section 3 as well as the data-driven fake lepton background derived in Section 6.3.1 are used. Scale factors for the $e \rightarrow \gamma$ fake and hadronic fake backgrounds derived in Sections 6.1 and 6.2 are not applied, as those were derived for the signal region only. The observed shape discrepancies between MC and data in the three regions are used as systematic variations.

In a dedicated hadronic fake control region, the PPT shape uncertainty with respect to hadronic fakes is estimated. The selection requirements are identical to those of the e+jets and μ+jets signal regions described in Section 5.1, except for the reversed photon isolation requirements, that is, hadronic fake candidates must fail the `FixedCutTight` isolation working point. To further reduce signal contamination of the control region, track isolations of the photons are required to fulfil $p_T^{\text{cone20}} > 3$ GeV. A conservative method, comparing only shapes of the hadronic fake distributions, would be to remove the remaining signal contribution entirely. Another approach to avoid any sensitivity to the shape of the $t\bar{t}\gamma$ signal is to vary the signal contribution and choose the scenario that maximises shape differences between MC and data. To be conservative, scenarios with up and down variations of the signal of ±50% are tested. The shape differences between MC and data are found to be maximal with the $t\bar{t}\gamma$ signal scaled to 50%, hence, this scenario is chosen for the estimation of the shape systematics. After applying the $t\bar{t}\gamma$ signal scaling, all MC are scaled to data for shape comparison with a global scaling factor of 1.285.[6]

Good agreement between MC and data in the combined e+jets and μ+jets control region can be seen for various observables in Figure 9. The histograms show an enhanced contribution of hadronic fakes due to the reversed photon isolation requirements. In Figure 10(a), the data/MC comparison plot of the PPT output is displayed.[7] As the observed disagreement between data and MC shape is not covered by statistical uncertainties, and because of expected dependencies of the PPT performance on η and p_{T} of the photon, the hadronic fake control region is further divided into three $|\eta|$ and five p_{T} slices, the PPT distributions of which can be found in Figure 11. The used $|\eta|$ slices are $[0, 0.60]$, $[0.60, 1.37]$ and $[1.52, 2.37]$, motivated by the material distributions in front of the Electromagnetic Calorimeter and the excluded crack region for the photon reconstruction. The photon p_{T} is divided into $[20, 27]$, $[27, 35]$, $[35, 50]$, $[50, 80]$ and $[80, 300]$ (in units of GeV) to generate slices with approximately the same data yields. The discrepancies in the data/MC ratios are extracted from all individual slices and then applied as η and p_{T}-dependent systematic variations of the hadronic fake contributions in the signal regions.

As the hadronic fake systematic variations are extracted from a control region with orthogonal isolation requirements compared to the signal regions, an additional hadronic fake isolation control region is defined to extract a second systematic variation for hadronic fake contributions. The selection requirements follow those of the e+jets and μ+jets signal regions, except for the reversed photon ID requirements, that is, hadronic fake candidates must fail the `Tight` ID cuts. In Figure 10(b), the data/MC comparison plot of the PPT output is displayed. The hadronic fake isolation control region is divided into the same $|\eta|$ slices as the previous region, but – due to lower data yields – only into three p_{T} slices, for which the PPT output distributions can be found in Figure 12. The used p_{T} slices are $[20, 27]$, $[27, 50]$, and $[50, 300]$ (in units

[6] Different scaling factors are necessary for the jet p_{T} histogram and distributions specific to the e+jets and μ+jets channels. These are quoted in the caption of Figure 9.

[7] Plots, where the $t\bar{t}\gamma$ signal contribution is increased to 150% can be found in Figures 128 and 129 in Appendix B.3.

Not reviewed, for internal circulation only

of GeV). The discrepancies in the data/MC are extracted from the individual slices and are also applied as systematic variations of the hadronic fake contributions in the signal regions. To scale MC to data, the hadronic fake isolation control region uses a scaling factor of 1.288.

In a third PPT control region, where prompt photons from processes other than $t\bar{t}\gamma$ are enhanced, the PPT shape uncertainty for prompt photons is estimated. The selection requirements are identical to those of the ee and $\mu\mu$ signal regions described in Section 5.1, with removed cuts on the number of jets and b-jets to increase statistics. Additionally, to suppress signal contamination, the invariant mass of the dilepton pair is required to be within $[60, 100]$ GeV. Again, the MC distribution is scaled to data to extract the shape differences with a global scaling factor of 1.153.[8] Control plots with ee and $\mu\mu$ combined are shown in Figure 13 with good agreement between MC and data. The PPT output distribution in the prompt-photon control region in Figure 10(c) shows a clear slope in the data/MC ratio. Again, the control region is divided into three η and five p_{T} slices (using the same slices as the PPT hadronic fake control region) to account for different PPT behaviour for different photon kinematics. The PPT distributions of all slices are shown in Figure 14. As the PPT prompt control region covers photons that are identical to those in the signal regions in terms of ID and isolation requirements, the observed shape discrepancies in the data/MC ratio are extracted and applied as a correction to the PPT values of the prompt constributions in the signal regions. On top of that, the extracted shape discrepancies are additionally applied as η and p_{T}-dependent systematic variations to the prompt contributions in the signal regions.

[8] As in the PPT hadronic fake control region, different scaling factors are necessary for the jet p_{T} histogram and distributions specific to the e+jets and μ+jets channels. These are quoted in the caption of Figure 13.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 9: Control plots in the PPT hadronic fake control region, including distributions of the photon shower shape variables R_{had}, f_{side}, R_η, R_ϕ, $w_{\eta 1}$ and $w_{\eta 2}$ used as input to the NN. The $t\bar{t}\gamma$ signal is decreased to 50% and the total MC is scaled to data for shape comparisons with a global scaling factor of 1.285. The jet p_{T} distribution and the histograms specific for the e+jets and μ+jets channels have scaling factors of 1.321, 1.284 and 1.286 applied, respectively. The displayed uncertainties are statistical only.

Not reviewed, for internal circulation only

(a) (b)

(c)

Figure 10: PPT distributions for (a) the hadronic fake control region, (b) the additional hadronic fake isolation control region and (c) the prompt photon control regions. For the hadronic fake control region shown in (a), the signal contribution is decreased to 50% to maximise shape differences. MC is scaled to data for shape comparison with scaling factors of 1.285, 1.288 and 1.153, respectively. Only statistical uncertainties are included. The observed shape differences in the MC/data ratios of the three control regions are used as systematic variations of the PPT shape.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 11: PPT output distributions for η and p_{T} slices in the PPT hadronic fake control region. The rows correspond to slices of $|\eta|$ and the columns to slices of p_{T} (continued over both pages). $|\eta|$ is divided into $[0, 0.60]$, $[0.60, 1.37]$, and $[1.52, 2.37]$, p_{T} into $[20, 27]$, $[27, 35]$, $[35, 50]$, $[50, 80]$, and $[80, 300]$ (in units of GeV). MC is scaled to data for shape comparisons with a global scaling factor of 1.285 and individual adjustments in the slices to fit the data yields. The displayed uncertainties are statistical only. The observed shape discrepancies in the data/MC ratio are extracted individually and used as systematic variations of the hadronic fake contributions in the signal regions.

Not reviewed, for internal circulation only

Figure 12: PPT output distributions for η and p_T slices in the additional PPT hadronic fake isolation control region. The rows correspond to slices of $|\eta|$ and the columns to slices of p_{T}. $|\eta|$ is divided into $[0, 0.60]$, $[0.60, 1.37]$, and $[1.52, 2.37]$, p_{T} into $[20, 27]$, $[27, 50]$, and $[50, 300]$ (in units of GeV). MC is scaled to data for shape comparisons with a global scaling factor of 1.288 and individual adjustments in the slices to fit the data yields. The displayed uncertainties are statistical only. The observed shape discrepancies in the data/MC ratio are extracted individually and used as systematic variations of the hadronic fake contributions in the signal regions.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 13: Control plots in the PPT prompt-photon control region, including distributions of the photon shower shape variables R_{had}, f_{side}, R_{η}, R_{ϕ}, $w_{\eta 1}$ and $w_{\eta 2}$ used as input to the NN. MC is scaled to data for shape comparisons with a global scaling factor of 1.153. The jet p_{T} distribution and the histograms specific for the e+jets and μ+jets channels have scaling factors of 1.110, 1.137 and 1.159 applied, respectively. The displayed uncertainties are statistical only.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 14: PPT output distributions for η and p_T slices in the PPT prompt control region. The rows correspond to slices of $|\eta|$ and the columns to slices of p_{T} (continued over both pages). $|\eta|$ is divided into $[0, 0.60]$, $[0.60, 1.37]$, and $[1.52, 2.37]$, p_{T} into $[20, 27]$, $[27, 35]$, $[35, 50]$, $[50, 80]$, and $[80, 300]$ (in units of GeV). MC is scaled to data for shape comparisons with a global scaling factor of 1.153 and individual adjustments in the slices to fit the data yields. The displayed uncertainties are statistical only. The observed shape discrepancies in the data/MC ratio are extracted and applied as a correction to the PPT values of the prompt contributions in the signal regions. In addition, the extracted shape discrepancies are applied as systematic variations.

Not reviewed, for internal circulation only

5 Signal region definition

5.1 Event selection

Each event is required to have at least one reconstructed primary vertex, obtained from at least two tracks. The event is required to have fired at least one lepton trigger. The considered triggers are summarised in Table 7. The lepton triggers definition and associated scale factors are described in [29] and [44].

Year	electron trigger	muon trigger
2015	HLT_e24_lhmedium_L1EM20VH or HLT_e60_lhmedium or HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 or HLT_mu50
2016	HLT_e26_lhtight_nod0_ivarloose or HLT_e60_lhmedium_nod0 or HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium or HLT_mu50

Table 7: The lepton triggers considered in the event selections.

The rest of the selection depends on the considered decay channel:

- The single lepton channels are characterized by the presence of only one reconstructed lepton, while the dilepton channels are characterized by the presence of exactly two leptons. For e+jets and ee channels, at least one electron trigger amongst the ones described previously should be fired, while for μ+jets and $\mu\mu$ channels, at least one muon trigger should be fired. For the $e\mu$ channel, at least one trigger should be fired which can be an electron trigger or a muon trigger. In order to be consistent with the trigger requirement, in addition to satisfying the lepton requirements described in Section 4.1, the minimum p_{T} of the highest p_{T} lepton is required to be 27 GeV (27.5 GeV) in the 2016 selection for electron (muon). At least one of the selected leptons in the event should be matched to a fired trigger.
- If the event appears to have badly reconstructed jets (based on criteria on energy fraction variables, using the LooseBad working point [45]) or badly reconstructed muons (inconsistency of the q/p uncertainty between the combined and the local tracks), this event is rejected.
- In the single lepton channels, at least four reconstructed jets are required. In the dilepton channels, at least two reconstructed jets are required.
- In the dilepton channels, the two leptons are required to be of opposite charges, and the invariant mass of the dilepton system should be larger than 15 GeV.
- For all channels, exactly one reconstructed photon in the event is required, satisfying the requirements described in Section 4.1.
- Event double counting removal: In the W+jets, Z+jets and $t\bar{t}$ MC samples, the events with a prompt photon are removed based on the truth information, to avoid a double counting with the dedicated $W\gamma$, $Z\gamma$ and $t\bar{t}\gamma$ MC samples (more detail in Section 3.2).
- At least one b-tagged jet at a working point of 77% [38] is required for all the channels.
- $E_{\mathrm{T}}^{\mathrm{miss}}$: For the ee and $\mu\mu$ dilepton channels, $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV.

Not reviewed, for internal circulation only

- $m(\ell,\ell)$: For the ee and $\mu\mu$ dilepton channels, an invariant mass veto between the two leptons in the window [85,95] GeV.
- $m(\ell,\ell,\gamma)$: For the ee and $\mu\mu$ dilepton channels, an invariant mass veto between the two leptons and the photon in the [85,95] GeV.
- $m(\gamma,e)$: For the e+jets channel, an invariant mass veto between the photon and the electron in the [85,95] GeV.
- $\Delta R(\gamma,\ell)$: In all channels, the distance between the prompt photon and the lepton must be greater than 1.0. This is to limit the contribution from the photons originating from radiative decay. Further studies on this have been done and can be found in Appendix D.

The event selection for all channels are summarised in Table 8.

Channel	e+jets	μ+jets	ee	$\mu\mu$	$e\mu$
Common	Primary vertex, Event cleaning, Run number, etc.				
	1 e	1 μ	2 e, OS	2 μ, OS	1 e + 1 μ, OS
	Trigger match				
	-		$m(\ell,\ell) > 15$ GeV		
Photon	1 γ with $p_{\mathrm{T}} > 20$ GeV				
Jet	≥ 4		≥ 2		
b-jet	≥ 1				
$m(\ell,\ell)$	-		not in [85,95] GeV		-
$m(\ell,\ell,\gamma)$	-		not in [85,95] GeV		-
$E_{\mathrm{T}}^{\mathrm{miss}}$	-		> 30 GeV		-
$m(\gamma,e)$	not in [$m(Z)$-5, $m(Z)$+5] GeV	-			
$\Delta R(\gamma,\ell)$	> 1.0				

Table 8: Summary of the event selections.

Yields for the above mentioned cuts can be found in Table 9 . A selection of pre-fit plots for each channel can be found in Figures 15, 16, 17, 18 and 19. In these plots, only statistical uncertainties are shown. For the $\mu\mu$ and ee channels larger statistical error bars can be seen for the $Z\gamma$ background. This is due to slightly larger than normal weights and is addressed in Appendix V.

	e+jets	μ+jets	ee	$\mu\mu$	$e\mu$
$t\bar{t}\gamma$	3207 ± 14	3200 ± 13	145 ± 3.0	174 ± 3.1	401 ± 5.2
hadronic fake	440.6 ± 13	457.2 ± 14	7.2 ± 1.5	5.7 ± 2.2	17.6 ± 2.5
$e \to \gamma$ fake	773.68 ± 21	603.84 ± 16	0.46 ± 0.27	0.24 ± 0.54	1.07 ± 0.71
Lep fake	293 ± 35	63 ± 22	-	-	-
$W\gamma$	541 ± 29	579 ± 34	-	-	-
$Z\gamma$	-	-	21 ± 3.8	54 ± 14	-
Other prompt	410.5 ± 47	274.6 ± 15	5.1 ± 1.2	4.4 ± 0.84	8.3 ± 1.2
Total	5665 ± 71	5178 ± 50	179 ± 5.2	238 ± 14	428 ± 5.9
Data	6002	5660	196	233	473

Table 9: Pre-fit yields for each background and channel. Only statistical uncertainties are included. Scale factors for the different backgrounds are not included.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 15: A selection of pre-fit distributions for the e+jets channel. Only statistical uncertainties are included.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 16: A selection of pre-fit distributions for the μ+jets channel. Only statistical uncertainties are included.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 17: A selection of pre-fit distributions for the *ee* channel. Only statistical uncertainties are included.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 18: A selection of pre-fit distributions for the $\mu\mu$ channel. Only statistical uncertainties are included.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 19: A selection of pre-fit distributions for the $e\mu$ channel. Only statistical uncertainties are included.

Not reviewed, for internal circulation only

5.2 Event level MVA

5.2.1 Discriminator for single lepton and dilepton channels

After the event selection, a multivariate discriminator at event level (ELD) has been developed separately for the single lepton and dilepton channels, and takes the form of neural networks.

Training was carried out on the full set of MC samples described in Section 3 and the data-driven fake lepton sample described in Section 6.3.1. Keras was used to train the neural networks and LWTNN was used to reconstruct them within analysis code.

For the single lepton training, the full set of selection cuts defined in Section 5.1 were applied. Looser cuts were used to train the dilepton NN. This includes dropping the two mass window cuts between the two leptons and the two leptons and the photon, neglecting the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut and not requiring a cut on the number of jets. This allows for more background contamination, enhances statistics and thus makes training more stable. Training only in the signal region (or close to the signal region) on events that are either signal or background in the final selection means the weights reach optimal values in a shorter amount of time. Each feature (or input variable) has been normalized to give a standard deviation of 1 and a mean of 0.

5.2.2 Variable selection and training

Each of the neural networks had hyper-parameter grid searches performed to achieve the best architecture and thus maximum separation.

The NN final topology is the same for single lepton and dilepton channels: 1 input dense layer consisting of 30 nodes is followed by a dropout layer then a batch normalization layer. A dense layer of 20 nodes follows with a final dense output layer consisting of 1 node with a sigmoid activation function to reduce outputs to values between 0 and 1. All layers' weights are initialized with uniform distributions and input and hidden layers receive a rectified linear unit (ReLU) activation function.

The dropout layer has a 30% chance that any node and its weights will randomly be removed from the epoch. This ensures that the network does not become too dependent on a node, which results in over-training. The batch normalization layer helps limit internal covariate shift in the weights that can occur due to fluctuations in the training set, and as a side effect, speeds up training.

The single lepton NN uses 75% of the raw input events for training with 25% used for testing. This equates to 144491 events used for training and 48164 events used for testing. Accumulatively, there are 160679 signal and 31976 background events. The dilepton NN uses 67% of the raw input events for training with 33% used for testing. This equates to 20258 events used for training and 10160 events used for testing. Accumulatively, there are 21591 signal and 8827 background events. The ratio of the split between training and testing is done to ensure that when separate K-Fold validation is performed the training and test samples can be split consistently into 4 and 3 different tests for single lepton and dilepton, respectively.

The single lepton NN is trained using a batch size of 150 events while the dilepton NN uses 300 events. This describes the number of events used for each gradient update. Optimisation for batch size was performed using a grid search. Event weights are accounted for by scaling the *binary cross-entropy* loss function during training, with negative event weights set to 0.

Not reviewed, for internal circulation only

A range of potentially interesting variables were ranked by their signal/background separation (Equation 2). Their ranking and physics motivation determine the priority in which different trainings were performed. The rankings for each channel are shown in Table 10. A check mark means it was used in the final NN configuration. Of note, the variables p_{T} (γ), $\eta(\gamma)$, $dR(\gamma, \ell)$ and $dR(\gamma, jet)$ were neglected from this optimisation due to being considered for differential measurements. Tables 11 and 12 show the ROC curve results for the training and test batches when performed on an increasing number of variables (added according to separation rank).

In both cases, the NN performs the best with the most number of variables considered. However, for the dilepton NN with more than 7 variables we see larger variations between multiple trainings and testings and so with the limited statistics it's deemed more unstable with over training likely to occur. As a results the network using 7 variables was selected. The distributions for the chosen variables in the combined single lepton and dilepton channels are shown in Figures 20 and 21. These plots show the distributions with the respective full signal region cuts applied.

Table 10: Separation of each variable used in the training of the single lepton and dilepton NN's. A check mark indicates that the variable was used in that NN.

Variable	single lepton separation (%)	dilepton separation (%)
Prompt Photon Tagger	3.46 ✓	6.76
H_{T}	1.79 ✓	1.32
number of jets	0.71 ✓	1.41
number of bjets	3.48 ✓	11.00 ✓
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.78 ✓	12.8 ✓
$m_T(W)$	0.86 ✓	-
p_{T} of first jet	2.05 ✓	11.4 ✓
p_{T} of second jet	1.77 ✓	7.72 ✓
p_{T} of third jet	1.72 ✓	3.29
p_{T} of fourth jet	1.30 ✓	1.13
p_{T} of fifth jet	0.92 ✓	0.262
MV2C10 weight of jet with highest weight	4.31 ✓	10.5 ✓
MV2C10 weight of jet with the second highest weight	2.75 ✓	13.6 ✓
jet with third highest MV2C10 tag weight	0.8 ✓	-
$m(\ell, \ell)$	-	13.4 ✓
$m(\gamma, \ell)$	0.32 ✓	5.79
$m(\gamma, \ell, \ell)$	-	2.32
$dR(\gamma, \ell)$	0.45	13.4
$dR(\gamma, jet)_{leading}$	1.80	3.07
$dR(\gamma, jet)_{subleading}$	1.71	4.3

Figure 22 shows the single lepton and dileptonNN's' response as well as the ROC curves for training on seen and unseen training data.

Training occurs by minimising the loss function for a given neural network. A decreasing loss function means the network is still learning. However, to prevent overtraining, in addition to monitoring the loss of the training set one needs to consider the loss function of the validation test set, i.e. a set of data that the network has never seen before. A "patience" parameter is set ensuring that if the loss function of the

Not reviewed, for internal circulation only

Figure 20: Variables used in training for the single lepton ELD neural network.

Not reviewed, for internal circulation only

Figure 21: Variables used in training for the dilepton ELD neural network.

Not reviewed, for internal circulation only

Table 11: Variable optimisation for the single lepton NN. The starting point is the first row. Each subsequent row includes more variables described on top of the previous row. Variables are chosen according to a combination of separation rank and physics motivation.

Variables	ROC (train)	ROC (test)	Selected NN
(5) jet btag weight leading, jet btag weight subleading, jet btag weight subsubleading, PPT, number of bjets	0.6483	0.6474	
(10) +, p_{T} of first jet, H_{T}, p_{T} of second jet, p_{T} of third jet, p_{T} of fourth jet	0.6744	0.6736	
(14) +,p_{T} of fifth jet ,$m_T(W)$, $E_{\mathrm{T}}^{\mathrm{miss}}$, number of jets,	0.6996	0.6906	
(15) +, $m(\gamma, \ell)$	0.7017	0.6929	✓

Table 12: Variable optimisation for the dilepton NN. The starting point is the first row. Each subsequent row includes more variables described on top of the previous row. Variables are chosen according to a combination of separation rank and physics motivation.

Variables	ROC (train)	ROC (test)	Selected NN
(3) jet btag weight subleading, $m(\ell, \ell)$,$E_{\mathrm{T}}^{\mathrm{miss}}$	0.9289	0.9233	
(7) +, p_{T} of first jet, number of bjets, jet btag weight leading, p_{T} of second jet	0.9313	0.9258	✓
(8) +, PPT	0.9329	0.9261	
(10) +, p_{T} of third jet, $m(\gamma, \ell, \ell)$	0.9365	0.9313	
(13) +,$m_T(W)$,number of jets, H_{T}	0.9428	0.9393	

validation set is constant (or increasing) after a given number of epochs the training is stopped and the model configuration for the minimal value of the loss function is used. The loss functions for the chosen models can be seen in Figure 23.

A further check to ensure over-training has not occurred is to perform k-fold cross-validation on training data. In this test, the data has been split into 4 separate batches for the single lepton case and 3 separate batches for the dilepton case. One batch is used as a validation or test set, the other two or three are used for training the NN. Each batch has a turn to be the test set. All ROC curves should be comparable with each other and only deviate slightly from the initial ROC curve (due to statistical fluctuations). The k-fold cross-validation plots are shown in Figure 24 for the single lepton and dilepton NN's. The agreement is good for both channels and over-training has not occurred.

Not reviewed, for internal circulation only

(a) The ELD response for training and unseen test samples for the single lepton NN.

(b) The ROC curve for the training sample compared to an unseen test sample for the single lepton NN

(c) The ELD response for training and unseen test samples for the dilepton NN.

(d) The ROC curve for the training sample compared to an unseen test sample for the dilepton NN.

Figure 22: Event Level Discriminator training output.

(a) Loss for single lepton ELD.

(b) Loss for dilepton ELD.

Figure 23: The loss curves for the training and validation datasets. The best model (minimal point on the loss curve) according to the validation loss is saved.

Not reviewed, for internal circulation only

(a) The ROC curve for the training samples for the single lepton NN.

(b) The ROC curve for the test samples for the single lepton NN.

(c) The ROC curve for the training samples for the dilepton NN.

(d) The ROC curve for the test samples for the dilepton NN.

Figure 24: The k-fold cross validation plots is performed for each NN.

5.2.3 Application and evaluation

The output from the neural networks for each channel is shown in Figures 25, 26 and 27. Due to the smaller contribution of background in the dilepton channels, only the separation between the signal and the total background is shown. Only statistical uncertainties are included. For the single lepton channels the ELD struggles to separate out the hadronic fake and $e \rightarrow \gamma$ fake background. These have the smallest separation at 13.4% and 15.7% for the e+jets channel, 11.9% and 10.2% for the μ+jets channel, respectively. The dilepton NN performs well where there is a background to separate, while for the $e\mu$ channel it is completely signal dominated in any case. These distributions are used to perform a binned maximum likelihood fit, as described in Section 8.

Since the b-tagged variables play such an import role in the ELD, we can examine the relationship between the b-jet dependence and the ELD output. Figure 28 shows this correlation for the single lepton and dilepton channel NN's, where each row of each plot has been normalised. In general, as more b-tagged jets enter into the event, the more signal-like it is. Background processes with a higher number of b-tagged jets in the event have an increased chance of being tagged as signal by the ELD.

Correlation plots for the ELD and the training input variables are shown in Figure 29 and Figure 30 . As

Not reviewed, for internal circulation only

(a) e+jets ELD

(b) Total background

(c) hadronic fake background

(d) $e \to \gamma$ fake background

(e) $W\gamma$ background

(f) Other prompt background

(g) Lep fake background

Figure 25: Event Level Discriminator variable for the e+jets channel with separation plots shown for signal and the different background components.

Not reviewed, for internal circulation only

(a) μ+jets ELD

(b) Total background

(c) hadronic fake background

(d) $e \rightarrow \gamma$ fake background

(e) $W\gamma$ background

(f) Other prompt background

(g) Lep fake background

Figure 26: Event Level Discriminator variable for the μ+jets channel with separation plots shown for signal and the different background components.

Not reviewed, for internal circulation only

Figure 27: Event Level Discriminator variable for the dilepton channels with the corresponding separation plots for all summed backgrounds.

(a) b-tagged jets versus ELD output for the single lepton channel.

(b) b-tagged jets versus ELD output for the dilepton channel.

Figure 28: The number of b-tagged jets versus the ELD output for the signal, the background and the combination of both for each of the NN's. Each row of each plot is normalised.

Not reviewed, for internal circulation only

expected the highest correlations for the ELD are with respect to the b-tagged variables in both signal and background for both channels. Also shown in these plots are the three and five variables used for differential measurements. These are ($p_T(\gamma)$, $\eta(\gamma)$ and $\Delta R(\gamma, l)$) and ($p_T(\gamma)$, $\eta(\gamma)$, $\Delta R(\gamma, l)$, $\Delta\phi(l, l)$ and $\Delta\eta(l, l)$) for the single lepton and dilepton channels, respectively.

(a) Signal

(b) hadronic fake background

(c) $e \to \gamma$ fake background

(d) Other prompt background

Not reviewed, for internal circulation only

(e) Lep fake background

Figure 29: Correlation plots for the signal and separate backgrounds for the single lepton channel comparing the ELD and training variables. Full event selection and event weights have been applied. The last three variables ($p_T(\gamma)$, $\eta(\gamma)$ and $\Delta R(\gamma, l)$) do not enter the training and are used for differential measurements.

(a) Signal

(b) Total background

Figure 30: Correlation plots for the signal and sum of backgrounds for the dilepton channel comparing the ELD and training variables. Full event selection and event weights have been applied. The last five variables ($p_T(\gamma)$, $\eta(\gamma)$, $\Delta R(\gamma, l)$, $\Delta\phi(l, l)$ and $\Delta\eta(l, l)$) do not enter the training and are used for differential measurements.

Not reviewed, for internal circulation only

6 Background descriptions

Several background processes mimic the $t\bar{t}\gamma$ signature. They can be events with a fake photon, either a hadronic fake (introduced in Section 4.2) or a fake due to an electron misidentification (a.k.a. $e \to \gamma$ fake), or they can be non-$t\bar{t}$ events with a prompt photon. This section describes the studies on the background processes and the detail of their estimation.

6.1 Background due to mis-reconstructed electrons

One of the important background contributions in $t\bar{t}\gamma$ originates from events in which one electron is misidentified as a photon, denoted as $e \to \gamma$ fake. It is an important background source in the single lepton channel. The main processes contributing to this background are the $t\bar{t}$ dileptonic decays (ee and $e\mu$ channels) and $Z \to ee$ decay, where one electron fakes a photon. The strategy of estimating this background is to use a fake rate scale factor, which is the ratio between the data-driven fake rate and the MC predicted fake rate, to correct for the number of fake photons predicted by MC in the signal region.

Section 6.1.1 defines the control regions used for the study. Section 6.1.2 describes the studies on the source of $e \to \gamma$ fakes by matching them to the truth level particles. Section 6.1.3 defines the fake rate and contains the calculation of the fake rate scale factor.

6.1.1 Control regions

To study the $e \to \gamma$ fake rate, a fake enriched control region is defined by selecting events having a pair of back-to-back electron and photon, which will be called $e\gamma$ control region. More specifically, the requirements include:

- Exactly one electron which is trigger matched.
- At least one photon. The leading p_{T} photon is referred simply as photon in the following.
- The opening angle between the electron and the photon to be larger than 150 degrees.
- The invariant mass of the electron and the photon to be within 50 GeV around the Z mass.
- To avoid a trigger bias later when studying the fake rate, the p_{T} of the electron should be larger than the p_{T} of the photon.

In the selected events, the electron is called *tag electron* and the photon is called *probe photon*.

Another control region, that will be called ee control region, is defined in exactly the same way as above, but by replacing the photon in the requirements with an electron which should have opposite charge sign with respect to the tag electron. Thus, this electron is called *probe electron* and is used as a reference to be compared with the probe photon to define the fake rate later.

Not reviewed, for internal circulation only

6.1.2 Fake sources

The $Z \to ee$ MC events selected in the above described $e\gamma$ control region can be used to study the source of $e \to \gamma$ fakes, by matching the probe photon with a truth particle before the detector simulation. The matching is done by extrapolating the track of the truth particle to the calorimeter layer and calculating the angular distance between the truth particle and the EM cluster, from which the photon is reconstructed. If the distance is smaller than a reference value ($\Delta R < 0.3$), the truth particle is considered to be the source of the photon.

After truth matching, the photon can be categorised into four classes (in the order of decreasing population):

- Type (a): denoted as "mis-reco.", where the photon is matched to a true electron. 59% of the selected photons belong to this class.
- Type (b): denoted as "mis-match", where the photon is matched to a true photon, but the photon's p_{T} is larger than that of the true photon by more than 10%, and at the same time, there is a nearby true electron with $\Delta R < 0.05$ w.r.t. the photon. 26% of the selected photons belong to this class.
- Type (c): denoted as "prompt QED", where the photon is matched to a true photon, and there is no nearby true electron with $\Delta R < 0.05$ w.r.t. the photon. 12% of the selected photons belong to this class.
- Type (d): denoted as "non-prompt QED", where the photon is matched to a true photon, and their relative p_{T} difference is smaller than 10%, although there is a nearby true electron with $\Delta R < 0.05$ w.r.t. the probe photon. 3% of the selected photons belong to this class.

The categorization is summarized in Fig. 31.

Figure 31: The categorisation of the selected probe photons in the $e\gamma$ control region via truth particle matching.

In the following, to better understand these four types of photons, their kinematics are shown and compared to that of the probe electron, when available.

The p_{T} and η of the probe photons are compared to those of the probe electron in Fig. 32. It can be seen that the p_{T} spectrum is rather similar between photon type (a), (b), and (d) and the probe electron, which indicates that they are truly $e \to \gamma$ fakes. For the η distribution, type (a) and (b) peak at high absolute η region. Connecting to the fact that there is also larger upstream material in high absolute η region, it implies (a) and (b) are likely to be bremsstrahlung induced fake photons. For type (d), its η spectrum is very similar to that of the probe electron. This could be explained by a very hard non-prompt QED that takes away almost all kinematics of its mother electron.

Not reviewed, for internal circulation only

The p_T of tag electrons in the two control regions and the invariant mass between the tag and probe are compared in Fig. 33. The lower-shifted p_T and mass spectrum for type (c) indicates that it is a true prompt photon from the three body decay of $Z \to ee + \gamma$. Therefore, type (c) is not to be counted as fake.

(a) (b)

Figure 32: The p_T (a) and η (b) distributions of the probe photon, shown for different classes and compared to those of the probe electron.

(a) (b)

Figure 33: The p_T of the tag electron in the two control regions (a) and the invariant mass between the tag and probe (b).

The invariant mass distribution of tag electron and probe photon is shown in Fig. 34, with each type of photons being normalized to their expected yields. It can be seen that the fake is dominated by the type (a) and (b).

Not reviewed, for internal circulation only

Figure 34: The invariant mass between the tag and probe.

6.1.3 Fake rates

Using the $Z \to ee$ MC, the number of $Z \to ee$ events observed with a tag electron in any place and a fake photon in a particular bin can be expressed as follows:

$$N^i_{e,\gamma} = N^i_{true} \times \epsilon^{reco}_e \times \epsilon^{\text{others}}_e \times p^i_{e\to\gamma} \times \epsilon^i_{\gamma(F)} \tag{3}$$

where

- i is the binning index of p_{T} or η or both
- N^i_{true}: true number of generated $Z \to ee$ events where the probe electron is in bin i
- ϵ^{reco}_e and $\epsilon^{\text{others}}_e$: reconstruction and other selection efficiencies of tag electron
- $p^i_{e\to\gamma}$: probability of misidentifying an electron as photon in bin i
- $\epsilon^i_{\gamma(F)}$: selection efficiency of the fake photon in bin i ("F" denotes the fact that it's a fake photon so that the efficiency can be different from the true photon)

Also, the number of $Z \to ee$ events observed with a tag electron in any place and a probe electron in a particular bin can be expressed as follows:

$$N^i_{e,e} = N^i_{true} \times \epsilon^{reco}_{e1} \times \epsilon^{\text{others}}_{e1} \times \epsilon^{reco,i}_{e2} \times \epsilon^{\text{others},i}_{e2} \tag{4}$$

where

- ϵ^{reco}_{e1} and $\epsilon^{\text{others}}_{e1}$: reconstruction and other selection efficiencies of tag electron
- $\epsilon^{reco,i}_{e2}$ and $\epsilon^{\text{others},i}_{e2}$: reconstruction and other selection efficiencies of probe electron in bin i

Not reviewed, for internal circulation only

The fake rate (FR) in a particular bin i is then defined as the ratio between $N^i_{e,\gamma}$ and $N^i_{e,e}$:

$$\mathrm{FR}^i \equiv \frac{N^i_{e,\gamma}}{N^i_{e,e}} = p^i_{e\to\gamma} \times \frac{\epsilon^i_\gamma(F)}{\epsilon^{reco,i}_{e2} \cdot \epsilon^{\mathrm{others},i}_{e2}} = p^i_{e\to\gamma} \times C^i \tag{5}$$

From the above formula, it is known that FR^i is proportional to the $e \to \gamma$ faking probability in that bin. The proportion coefficient C^i could also vary.

When we don't do p_T and η binning, an overall fake rate can be calculated using the same formula and its MC value are shown in Table 13. In Fig. 35, the p_T and η dependencies of the fake rate are shown. In both Table 13 and Fig. 35, the type (c), which is not really a fake photon, is shown for completeness.

Table 13: The overall fake rate calculated from $Z \to ee$ MC for different types of photon. Type (c) is not counted as fake photon and its value is just for shown. Only statistical uncertainties are shown.

	N(probe γ)	N(probe e)	$\mathrm{FR_{MC}}$(%)
Type(a)	78240.80 ± 1096.92	4551436.50 ± 8419.50	1.72 ± 0.02
Type(b)	34156.20 ± 760.62	4551436.50 ± 8419.50	0.75 ± 0.02
Type(c)	16534.88 ± 493.96	4551436.50 ± 8419.50	0.36 ± 0.01
Type(d)	3786.88 ± 248.38	4551436.50 ± 8419.50	0.08 ± 0.01
All	132718.77 ± 1444.81	4551436.50 ± 8419.50	2.92 ± 0.03

(a) (b)

Figure 35: The p_T (a) and η (b) dependency of the fake rates

In order to calculate the data-driven fake rate in bin i, the only difference with respect to the above MC-based fake rate is that:

$$\mathrm{FR}^i_{dd} = \frac{N^{\mathrm{data},i}_{e,\gamma} - N^{non\text{-}Z,i}_{e,\gamma}}{N^{\mathrm{data},i}_{e,e} - N^{non\text{-}Z}_{e,e},i} \tag{6}$$

which means the denominator and numerator are replaced by their data-driven ones.

The subtraction of non-Z events from data in these control regions can't be done with MC samples since from MC study the non-Z background are negligible. Instead, it be done by side band fit, as shown in

Not reviewed, for internal circulation only

Figure 36, where the $Z \to ee$ signal is modelled by double-sided Crystal-ball function and the non-Z background by a Bernstein 4th order polynomial.

Figure 36: The side band fit in the two control regions to subtract non-Z events.

After that the data-driven fake rate in bin i is divided by the corresponding MC fake rate to derive a set of fake rate scale factors:

$$\mathrm{SF}^{i}_{\mathrm{FR}} = \frac{\mathrm{FR}^{i}_{dd}}{\mathrm{FR}^{i}_{MC}} \tag{7}$$

these scale factors will be applied to the $e \to \gamma$ MC samples as a data-driven correction.

The overall scale factor is calculated to be 0.97 ± 0.02, with the error being statistical only. (this overall scale factor is just for show, in practise, the pt/eta binned scale factors are used). To estimate the systematic uncertainty of the scale factor, the following variations are considered:

- signal function shape is changed from double sided Crystal-ball to MC predicted template (if necessary with smoothing to reduce fluctuation), the result of which is shown in Figure 37 (a). The scale factor is calculated to be 1.14, with a difference of 17% with respect to the nominal SF.
- fitting mass range shrunk by 5 or 10 GeV in the low end and 10 GeV in the high end, the result of which is shown in Figure 37 (b). The scale factor is calculated to be 1.08, with a difference of 11% with respect to the nominal SF.
- background function shape is changed from Bernstein to Gaussian, the result of which is shown in Figure 37 (c). The scale factor is calculated to be 1.04, with a difference of 7% with respect to the nominal SF.
- the MC model for the subtraction of prompt QED contribution is changed from $Z \to ee$ sample to a $Z \to ee\gamma$ sample. The scale factor is calculated to be 0.99, with a difference of 2% with respect to the nominal SF.

The total uncertainty is derived from their quadratic sum (including the statistical uncertainty), which results in the final overall SF to be 0.97 ± 0.21.

The SF can also be derived differentially with respect to the photon p_{T} and η. The choice of binning for p_{T} is [25, 35, 45, 60, infinity] (in GeV) and for η is [0, 0.5, 1.0, 1.37] and [1.52, 2.0, 2.37]. So in the end,

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 37: The side band fit in the two control regions to subtract non-Z events, where the signal function is switched from double-sided crystal ball to MC temlate (a), the fit range is shortened by 10 GeV(b), or the background function is switched from 4th order Bernstein polynomial to Gaussian (c).

Not reviewed, for internal circulation only

there will be $4 \times 5 = 20$ bins. The nominal mass peak fits for each of these bins are shown in Figure 38 and Figure 39. The final 2D scale factors are summarized in Figure 40, including statistical uncertainties as well as the above systematic uncertainties. The scale factors are also shown as a function of η in different pt bins in Figure 41, where the total uncertainties are decomposed. To be noted: starting from the 2nd p_{T} bin, which means from 45 GeV, the role of tag/probe is switched to gain statistics in high p_{T} region.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 38: The postfit of the (e,γ) invariant mass distribution in different photon (p_{T},η) bins.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 39: The postfit of the (e,e) invariant mass distribution in different photon (p_T,η) bins.

Not reviewed, for internal circulation only

Figure 40: The final 2D fake rate scale factors with all uncertainties included.

Figure 41: The final 2D fake rate scale factors illustrated in the form of 1D η distribution in each p_{T} bins. The total uncertainties are decomposed.

Not reviewed, for internal circulation only

6.1.4 Signal region $e \to \gamma$ fake

The $e \to \gamma$ fake in the signal region where the fake photon is in bin i could be expressed as following:

$$N^i_{\ell,\gamma} = N^i_{\ell,e} \times \mathrm{FR}^i \tag{8}$$

ℓ could be an electron or a muon, and when it's electron, the formula should be applied twice: one for the case of the leading electron faking the photon and the other for the subleading electron. $N^i_{\ell,e}$ is the number of events in a modified signal region by replacing all the photon related cuts with those on an electron and the electron being in bin i. It plays the role of a pool of $e \to \gamma$ fake candidate. The above section corrected the FR with data-driven method. Here we try to correct $N^i_{\ell,e}$ in this modified signal region.

The p_{T} distributions of the leading lepton and subleading lepton when it's the modified e+jets signal region and when it's the modified μ+jets signal region are shown in Figure 42, where a good data/MC agreement can be found and the ratio of the two are rather flat. The resulting overall data-over-MC SFs are 0.97 ± 0.01 and 0.98 ± 0.01 for the modified e+jets and μ+jets channels, respectively, with the error being statistical uncertainties. These SFs are used as an additional correction, on top of the binned FR SFs, to the $e \to \gamma$ fake MC samaples. To this step, we have corrected all components of the $e \to \gamma$ sample with data.

Figure 42: The p_{T} distributions of the leading lepton and subleading lepton of the modifed e+jets signal region (top) and of the modified μ+jets signal region (bottom). Only statistical uncertainties are shown in the error band.

Not reviewed, for internal circulation only

6.1.5 PPT systematics

Since in the end we do an ELD fit and the single lepton channel ELD takes as input the PPT, it is necessary to estimate the PPT systematics of the $e \rightarrow \gamma$ fake. This is done by comparing the PPT shape in the $Z \rightarrow e\gamma$ fake photon control region defined as above.

To subtract the non-Z background in the PPT distribution, a set of event weights are derived from the invariant mass distribution:

$$w_i = \frac{m_{i,Z}}{m_{i,all}}$$

where $m_{i,Z}$ is the number of estimated $Z \rightarrow ee$ event in the ith bin of the invariant mass distribution, and $m_{i,all}$ is the total number of event in that bin. Then this set of weights are applied to the data and the resulting data PPT distribution is compared with its MC correspondent.

In Figure 43, this PPT shape comparison is shown. The data/MC shape discrepancy looks very similar to that of the prompt photon, thus the prompt photon's PPT systematics are applied to the $e \rightarrow \gamma$ fake. This is not coincidence, as from Section 4.2 it is observed that the PPT shape is very similar between prompt photon and $e \rightarrow \gamma$ fake.

Figure 43: The PPT shape comparison between data and MC in the $Z \rightarrow e\gamma$ fake photon control region. The ratio is very similar to that of the prompt photon PPT.

Not reviewed, for internal circulation only

6.1.6 Application of fake rates

The above fake rate SFs, derived from $Z \to ee$ sample, are applied to MC based $e \to \gamma$ fake photon samples in the signal region. This is based on the belief that the mis-simulation of $e \to \gamma$ fake rates should be universal to all processes, thus can be corrected with the same set of SFs. The correctness of this assumption can be checked by MC to compare the fake rate of $Z \to ee$ process in the above fake rate control region with that of the $t\bar{t}$ process in the signal region.

The result of the test is shown in Figure 44 for the two major fake types and for their p_{T} and η dependencies. As can be seen, these fake rates are very similar, which shows that we could use the same set of SFs to correct the fake rate in the SR. The $\Delta R(\gamma, j)$ and jet multiplicity distributions are also compared for these MC fake rates in Figure 45. The last bin of the $\Delta R(\gamma, j)$ for the Z MC is those events without any jet, thus $t\bar{t}$ MCs have no event there. In general, the fake rate slightly increases as it is getting further away from the jet and both Z and $t\bar{t}$ MCs show this trend. The first 4 bins of the jet multiplicity for the Z MC are those events which have less than 4 jets, thus $t\bar{t}$ MCs, which is selected via signal region event selection, have no event there. In general, the fake rate slightly decreases as the jet multiplicity increases, but then increases again at the high jet multiplicity region. And in the region where most of the $t\bar{t}$ events reside, which is between 4 jets and 6 jets, the fake rates are similar between the Z and $t\bar{t}$ MCs.

Figure 44: The comparison of fake rates predicted by $Z \to ee$, $t\bar{t} \to ee$, and $t\bar{t} \to \mu e$ MCs in different photon p_{T} and η bins, for the two major types of electron fake.

Not reviewed, for internal circulation only

Figure 45: The comparison of fake rates predicted by $Z \to ee$, $t\bar{t} \to ee$, and $t\bar{t} \to \mu e$ MCs as function of $\Delta R(\gamma, j)$ and jet multiplicity, for the two major types of electron fake.

6.2 Background due to hadronic fakes

The background contribution due to hadronic fakes (defined in Section 4.2) is largely reduced by applying the photon isolation requirement, but still is one of the important backgrounds.

To estimate the hadronic fake background, first a data-driven two-dimensional sidebands method (ABCD method) is used. Then, the ratio of the data-driven estimate to the MC prediction of hadronic fake events is considered as the hadronic fake scale factor, and is used to scale the MC derived hadronic fake background in the final fit.

6.2.1 Estimation of hadronic fakes in single lepton channel

Four different regions in data are defined by dividing the photons into four categories, based on their isolation and identification. The rest of the object and event selections remain the same as for the signal region (Sections 4.1 and 5.1). These four regions are illustrated in Figure 46. Region D is identical to the signal region, while regions A, B and C are dominated by hadronic fakes.

Photons in regions B and C are required to be non-isolated. For this, they must fail the isolation working point (`FixedCutTight`, see Section 4.1) and in addition fulfil $p_T^{\text{cone20}} > 3$ GeV. The extra cut on the photon track isolation at 3 GeV, which hereafter is called *isolation gap*, is applied to reduce the prompt photon contamination, specially in region C.

Not reviewed, for internal circulation only

Figure 46: An illustration of ABCD method to estimate the hadronic fake background, using four category of photons based on their identification and isolation.

Photons in regions A and B are required to be identified as *fail-tight*. The fail-tight identification requires the photon to pass all the tight identification cuts except for the cuts on the shower-shape variables f_{side}, w_{s_3}, ΔE and E_{ratio} (see Table 4 for the description of the variables): at least two of these four cuts are explicitly asked to be failed. These four shower-shape variables contain information from the first layer of the electromagnetic calorimeter (strip layer). The strip layer is finely granulated in η, with the purpose of suppressing the background photons which typically have a broader shower profile. While these four variables have strong discrimination power between the prompt and hadronic fake photons, their correlation with photon isolation is very small [46]. The choice of the least number of cuts to fail, is based on a compromise between reducing the prompt photon contamination and keeping enough statistics, specifically in region A.

Control plots for regions A, B and C for single lepton channel are shown in Figures 47, 48 and 49, while control plots for e+jets and μ+jets channels separately can be find in Appendix F.3.

Assuming no correlation between photon isolation and photon identification, the following relation between the hadronic fake contribution in the four regions holds:

$$\frac{N_{\mathrm{A}}^{\text{h-fake}}}{N_{\mathrm{B}}^{\text{h-fake}}} = \frac{N_{\mathrm{D}}^{\text{h-fake}}}{N_{\mathrm{C}}^{\text{h-fake}}} \,. \tag{9}$$

Although the isolation and identification criteria are chosen to minimise any such correlation, but any deviation from this assumption can be quantified by the following double ratio derived from MC:

$$\theta_{\mathrm{MC}} = \frac{N_{\mathrm{D,MC}}^{\text{h-fake}} / N_{\mathrm{C,MC}}^{\text{h-fake}}}{N_{\mathrm{A,MC}}^{\text{h-fake}} / N_{\mathrm{B,MC}}^{\text{h-fake}}} \,. \tag{10}$$

The measured θ_{MC} with its uncertainties are shown in Figures 50 and 51. Figure 50 shows θ_{MC} as function of photon p_{T} and photon η for all photons (50(a) and 50(b)), and separately for converted (50(d) and 50(e)) and unconverted photons (50(f) and 50(g)). Figure 51 shows θ_{MC} as function of photon p_{T} in slices of photon η, separated for converted and unconverted photons. Since the $N_{i,\mathrm{MC}}^{\text{h-fake}}$ ($i \in \mathrm{A, B, C, D}$) are dominated by the $t\bar{t}$ process, the modelling systematics of $t\bar{t}$ MC sample are considered. The uncertainty on hard scattering modelling and parton shower and hadronisation (gen. & PS) are estimated by comparing the $t\bar{t}$ nominal sample produced by Powheg and interfaced by Pythia8, with the $t\bar{t}$ sample produced

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 47: Plots of region A in single lepton channel, where photons are identified as fail-tight and are isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 48: Plots of region B in single lepton channel, where photons are identified as fail-tight and are non-isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 49: Plots of region C in single lepton channel, where photons are identified as tight and are non-isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes.

Not reviewed, for internal circulation only

by SHERPA. The uncertainty on modelling of additional QCD radiation (rad.-up/-down) is estimated by comparing the $t\bar{t}$ nominal sample with two samples that are produced with different PYTHIA8 tunes with higher or lower QCD radiation activity (A14 var3c eigentune), modified factorisation and renormalisation scales (at ×0.5 and ×2) and modified NLO radiation (with *hdamp* = 1.5 and 3 times the mass of the top quark). More information about the effect of $t\bar{t}$ modelling systematics can be find in Appendix F.2. The detector systematics are expected to be much smaller, therefore are not added.

The measured inclusive and binned θ_{MC} are also summarised in Tables 14 and 15, in 1D, 2D and 3D bins. The double ratio is also calculated from a data-driven method as a cross check, the detail of which can be find in Appendix F.4. The deviation of measured θ_{MC} from unity indicates to a residual correlation that should be taken into account.

Channel	conv. type	p_{T} bin [GeV]	η bin	θ_{MC} ± total uncertainty	stat. uncertainty	syst. uncertainty		
						rad.-up	rad.-down	gen. & PS
single lepton	all	inclusive	inclusive	1.44 ± 0.24	±0.06	+0.08	-0.03	±0.22
		$20 < p_{\mathrm{T}} < 30$	inclusive	1.29 ± 0.31	±0.08	-0.06	-0.19	±0.23
		$30 < p_{\mathrm{T}} < 50$		1.33 ± 0.21	±0.10	+0.05	-0.12	±0.15
		$p_{\mathrm{T}} > 50$		1.80 ± 0.62	±0.11	+0.35	-0.51	±0.34
		inclusive	$\lvert\eta\rvert < 0.60$	1.35 ± 0.12	±0.10	+0.00	-0.07	±0.01
			$0.60 < \lvert\eta\rvert < 1.37$	1.25 ± 0.40	±0.07	-0.02	-0.05	±0.39
			$1.52 < \lvert\eta\rvert < 2.37$	1.85 ± 0.44	±0.14	+0.40	-0.18	±0.11
	conv.	inclusive	inclusive	1.51 ± 0.48	±0.08	+0.20	-0.14	±0.43
		$20 < p_{\mathrm{T}} < 30$	inclusive	1.24 ± 0.40	±0.10	+0.00	-0.36	±0.15
		$30 < p_{\mathrm{T}} < 50$		1.45 ± 0.71	±0.17	+0.04	-0.32	±0.60
		$p_{\mathrm{T}} > 50$		1.88 ± 0.90	±0.14	+0.70	-0.61	±0.54
		inclusive	$\lvert\eta\rvert < 0.60$	1.45 ± 0.47	±0.17	+0.25	-0.17	±0.36
			$0.60 < \lvert\eta\rvert < 1.37$	1.34 ± 0.57	±0.08	+0.12	-0.06	±0.55
			$1.52 < \lvert\eta\rvert < 2.37$	1.92 ± 0.50	±0.19	+0.39	-0.38	±0.24
	unconv.	inclusive	inclusive	1.24 ± 0.13	±0.07	-0.06	-0.11	±0.02
		$20 < p_{\mathrm{T}} < 30$	inclusive	1.26 ± 0.34	±0.12	-0.11	-0.06	±0.30
		$30 < p_{\mathrm{T}} < 50$		1.12 ± 0.30	±0.10	-0.01	-0.14	±0.25
		$p_{\mathrm{T}} > 50$		1.45 ± 0.42	±0.15	-0.07	-0.31	±0.25
		inclusive	$\lvert\eta\rvert < 0.60$	1.17 ± 0.28	±0.11	-0.15	-0.00	±0.21
			$0.60 < \lvert\eta\rvert < 1.37$	1.20 ± 0.15	±0.07	-0.12	-0.07	±0.06
			$1.52 < \lvert\eta\rvert < 2.37$	1.45 ± 0.55	±0.17	+0.44	-0.45	±0.28

Table 14: MC-based double ratio θ_{MC} measured for inclusive, 1D bins of photon p_{T}, photon η and photon conversion type, and 2D bins of photon p_{T}-conversion type and photon η-conversion type.

Channel	conv. type	p_{T}-η bin	θ_{MC} ± total uncertainty	stat. uncertainty	syst. uncertainty		
					rad.-up	rad.-down	gen. & PS
single lepton	conv.	$\lvert\eta\rvert < 0.60$	1.45 ± 0.47	±0.17	+0.25	-0.17	±0.36
		$0.60 < \lvert\eta\rvert < 1.37$, $20 < p_{\mathrm{T}} < 30$	1.08 ± 0.55	±0.14	+0.10	-0.29	±0.44
		$0.60 < \lvert\eta\rvert < 1.37$, $p_{\mathrm{T}} > 30$	1.34 ± 0.64	±0.11	+0.00	-0.12	±0.62
		$1.52 < \lvert\eta\rvert < 2.37$, $20 < p_{\mathrm{T}} < 50$	1.73 ± 0.89	±0.23	+0.05	-0.60	±0.61
		$1.52 < \lvert\eta\rvert < 2.37$, $p_{\mathrm{T}} > 50$	2.30 ± 1.82	±0.34	+1.72	-0.75	±0.49
	unconv.	$\lvert\eta\rvert < 0.60$, $20 < p_{\mathrm{T}} < 30$	1.46 ± 0.45	±0.26	-0.30	-0.24	±0.22
		$\lvert\eta\rvert < 0.60$, $p_{\mathrm{T}} > 30$	1.03 ± 0.46	±0.11	-0.10	-0.06	±0.44
		$0.60 < \lvert\eta\rvert < 1.37$, $20 < p_{\mathrm{T}} < 30$	1.18 ± 0.55	±0.15	-0.13	-0.33	±0.41
		$0.60 < \lvert\eta\rvert < 1.37$, $p_{\mathrm{T}} > 30$	1.23 ± 0.74	±0.13	-0.07	-0.67	±0.29
		$1.52 < \lvert\eta\rvert < 2.37$	1.45 ± 0.55	±0.17	+0.44	-0.45	±0.28

Table 15: MC-based double ratio θ_{MC}, in 3D bins of photon p_{T}-η-conversion type, in the same binning that is used for the final 3D binned $\mathrm{SF}^{\text{h-fake}}$.

Not reviewed, for internal circulation only

Figure 50: MC-based double ratio θ_{MC} measured for all photons as function of photon (a) p_{T}, (b) η and (c) conversion type, as well as separately for the converted photons as function of photon (d) p_{T} and (e) η, and for the unconverted photons as function of photon (f) p_{T} and (g) η. The choice of binning for p_{T} is based on the statistics and the $|\eta|$ bins are chosen according to the detector geometry.

Not reviewed, for internal circulation only

(a)

(b)

Figure 51: MC-based double ratio θ_{MC} measured in different photon η slices as function of photon p_{T}, separately for (a) converted and (b) unconverted photons.

Considering the relation in Equation 9, and using θ_{MC} as a correction factor in order to take into account any residual correlation between photon isolation and identification, the hadronic fake background is estimated as:

$$N_{\mathrm{est.}}^{\mathrm{h\text{-}fake}} = \frac{N_{\mathrm{A,\,data}}^{\mathrm{h\text{-}fake}} \times N_{\mathrm{C,data}}^{\mathrm{h\text{-}fake}}}{N_{\mathrm{B,data}}^{\mathrm{h\text{-}fake}}} \times \theta_{\mathrm{MC}}\,, \tag{11}$$

where $N_{i,\mathrm{data}}^{\mathrm{h\text{-}fake}}$ ($i \in \mathrm{A, B, C}$) is the number of hadronic fake events in data in regions A, B and C. Although these three regions are dominated by hadronic fakes, as it can be seen from the control plots there are contaminations from events with prompt photons and $e \to \gamma$ fakes. Therefore, the contamination from prompt photons are subtracted by using MC, and the $e \to \gamma$ fakes are subtracted by using the MC scaled by the $e \to \gamma$ fake scale factor (Section 6.1). Since in the data-driven fake lepton background estimation (Section 6.3.1), events with prompt or fake photons are not distinguished and the estimated fake lepton background is entering the final fit directly, the fake lepton background is also subtracted:

$$N_{i,\mathrm{data}}^{\mathrm{h\text{-}fake}} = N_{i,\mathrm{data}} - N_{i,\mathrm{MC}}^{\mathrm{prompt}\text{-}\gamma} - \mathrm{SF}^{e\to\gamma} \times N_{i,\mathrm{MC}}^{e\to\gamma} - N_{i,\mathrm{d.d.}}^{\mathrm{fake\text{-}lepton}}, \quad i \in \mathrm{A, B, C}\,. \tag{12}$$

Not reviewed, for internal circulation only

Finally, the hadronic fake scale factor is calculated as:

$$\mathrm{SF}^{\text{h-fake}} = \frac{N_{\text{est.}}^{\text{h-fake}}}{N_{\text{MC}}^{\text{h-fake}}} \,. \tag{13}$$

The $\mathrm{SF}^{\text{h-fake}}$ is calculated for bins of photon p_{T} and η, separately for converted and unconverted photons. This is done by using the θ_{MC} that is measured for the same kinematic range and conversion type. The resulting $\mathrm{SF}^{\text{h-fake}}$'s are shown in Figure 52. The choice of binning for p_{T} is based on the statistics and the $|\eta|$ bins are chosen according to the detector geometry. A clear dependency on p_{T}, η and conversion type can be seen. Furthermore, $\mathrm{SF}^{\text{h-fake}}$ is calculated in 2D bins of p_{T}-η separated for converted and unconverted photons, hence 3D scale factors, using the θ_{MC} derived for the same binning. The neighbouring kinematic bins that resulted to $\mathrm{SF}^{\text{h-fake}}$'s consistent within statistical uncertainties are merged together. This leads to five bins for each of the converted and unconverted category, and the resulting $\mathrm{SF}^{\text{h-fake}}$'s are shown in Figure 53.

The final $\mathrm{SF}^{\text{h-fake}}$ used in the analysis are the 3D scale factors of Figure 53, which are also summarised in Table 16. The systematics uncertainty of $\mathrm{SF}^{\text{h-fake}}$ includes the uncertainties of θ_{MC} as explained before, and the uncertainties of the subtraction of non-hadronic-fakes in Equation 12, for which 100% is considered for the $t\bar{t}\gamma$, 50% for the rest of the prompts, 50% for the $e \to \gamma$ fakes and the uncertainty of the data-driven fake-lepton background as explained in Section 6.3.1. For the $t\bar{t}\gamma$ and other prompt subtractions, conservative uncertainties are assigned because the photon ID and isolation efficiencies are unknown for the fail-tight and reversed isolation reconstructed photons. Similarly, the $e \to \gamma$ fakes fake subtraction systematics is conservative due to the fact that the $e \to \gamma$ fakes fake rates that are used here (Section 6.1) are calculated from photons with tight ID.

In the final fit, the systematics uncertainties of $\mathrm{SF}^{\text{h-fake}}$ due to the $t\bar{t}$ modelling (coming from θ_{MC} factor) are correlated to the other relevant same systematics sources. The systematic uncertainties due to the statistical uncertainty of $\mathrm{SF}^{\text{h-fake}}$ (i.e. statistical uncertainty of data-driven part of $\mathrm{SF}^{\text{h-fake}}$ and statistical uncertainty of θ_{MC} factor) are uncorrelated bin-by-bin, by considering one Nuisance parameter for statistical uncertainty in each bin of $\mathrm{SF}^{\text{h-fake}}$ (see Appendix F.5.2). Also, the anti-correlation of $\mathrm{SF}^{\text{h-fake}}$ due to $t\bar{t}\gamma$ subtraction with the signal strength, μ (see Section 8 for definition), which is the parameter of interest in the fit, is considered. This is done by using the expression:

$$\mathrm{SF}^{\text{h-fake}}(\mu) = \mathrm{SF}^{\text{h-fake}}(\mu = 1) \times (1 - 0.12(\mu - 1)) \,, \tag{14}$$

where the factor -0.12 is the relative change in hadronic fake background normalisation if the $t\bar{t}\gamma$ normalisation goes up by 100% (i.e. if $\mu = 2$). Although $\mathrm{SF}^{\text{h-fake}}$ is 3D binned, since the $t\bar{t}\gamma$ subtraction systematics effect on ELD distribution is more or less flat, and for the simplicity, the overall change (i.e. -0.12) is used. This is motivated by Figure 159 in Appendix F.5.1.

Not reviewed, for internal circulation only

Figure 52: The $SF^{h\text{-}fake}$ in photon p_T bins for (a) converted and (b) unconverted photons, and in photon η bins for (c) converted and (d) unconverted photons.

Figure 53: Final 3D $SF^{h\text{-}fake}$, derived for $p_T - \eta$ bins separately for (a) converted and (b) unconverted photons.

Not reviewed, for internal circulation only

Channel	conv. type	$p_T - \eta$ bin	$SF^{h\text{-fake}}$	stat.	syst.				
					$t\bar{t}\gamma$ subtract.	rest subtract.	θ_{MC} stat.	θ_{MC} rad.-up/-down	θ_{MC} gen. & PS
single lepton	all	inclusive	1.60	±0.07	±0.19	±0.12	±0.07	+0.09/-0.03	±0.24
	conv.	$\lvert\eta\rvert < 0.60$	1.52	±0.17	±0.13	±0.15	±0.18	+0.26/-0.18	±0.38
		$0.60 < \lvert\eta\rvert < 1.37$, $20 < p_T < 30$	1.74	±0.27	±0.14	±0.09	±0.23	+0.16/-0.47	±0.71
		$0.60 < \lvert\eta\rvert < 1.37$, $p_T > 30$	1.00	±0.14	±0.07	±0.09	±0.08	+0.00/-0.09	±0.46
		$1.52 < \lvert\eta\rvert < 2.37$, $20 < p_T < 50$	1.71	±0.26	±0.15	±0.23	±0.23	+0.05/-0.59	±0.60
		$1.52 < \lvert\eta\rvert < 2.37$, $p_T > 50$	0.79	±0.22	±0.06	±0.12	±0.12	+0.59/-0.26	±0.17
	unconv.	$\lvert\eta\rvert < 0.60$, $20 < p_T < 30$	2.36	±0.35	±0.25	±0.13	±0.42	-0.49/-0.39	±0.36
		$\lvert\eta\rvert < 0.60$, $p_T > 30$	1.37	±0.18	±0.24	±0.12	±0.15	-0.13/-0.08	±0.58
		$0.60 < \lvert\eta\rvert < 1.37$, $20 < p_T < 30$	1.72	±0.25	±0.18	±0.05	±0.22	-0.19/-0.48	±0.60
		$0.60 < \lvert\eta\rvert < 1.37$, $p_T > 30$	1.33	±0.23	±0.23	±0.14	±0.14	-0.08/-0.72	±0.31
		$1.52 < \lvert\eta\rvert < 2.37$	3.15	±0.48	±0.57	±0.27	±0.37	+0.95/-0.98	±0.61

Table 16: The hadronic fake scale factor ($SF^{h\text{-fake}}$) for single lepton channel derived for inclusive and 3D bins of photon p_T-η-conversion type. The 3D $SF^{h\text{-fake}}$'s are used in the analysis and the inclusive scale factor is just shown for comparison. All the individual uncertainty components considered for the $SF^{h\text{-fake}}$ are shown: the uncertainty due to the statistics of $N^{h\text{-fake}}_{i,\text{data}}$ with $i \in$ A, B, C (stat.), the uncertainty due to the subtraction of $t\bar{t}\gamma$ ($t\bar{t}\gamma$ subtract.), the uncertainty due to the subtraction of rest of the prompts, $e \to \gamma$ fakes, and fake-lepton background (rest subtract.), the uncertainty due to the statistical uncertainty of the θ_{MC} factor (θ_{MC} stat.), the uncertainty due to modelling of additional QCD radiation in $t\bar{t}$ MC (θ_{MC} rad.-up/-down), and finally the uncertainty due to the hard scattering and parton shower and hadronisation modellings of $t\bar{t}$ MC (θ_{MC} gen. & PS).

6.2.2 Estimation of hadronic fakes in dilepton channel

For the dilepton channel, there are not enough statistics in the regions A and C to use the ABCD method. The contribution of hadronic fake background is much smaller in this channel, the MC estimate shows 3-5% contribution for different dilepton channels (Table 9).

Since the source of hadronic fakes are not depending on the number of leptons, the same scale factor of single lepton channel are used, with additional extrapolation systematics.

The extrapolation systematics are studied by varying the ratio between MC prediced hadronic fakes in the dilepton channel over that of the single lepton channel, since the dilepton hadronic fakes can be expressed as:

$$N_{dilepton} = \text{SF} \times N^{MC}_{dilepton} = N^{data}_{singlelepton} \times \frac{N^{MC}_{dilepton}}{N^{MC}_{singlelepton}} \tag{15}$$

where systematics enter mainly thorough the ratio term.

The considered systematic variations include: the ttbar modelling uncertainties discussed above, which amounts to 10%, dominated by the Sherpa variation. all the experimental uncertainties, which are 8%, dominated by pile-up uncertainty.

Not reviewed, for internal circulation only

6.3 Background due to fake leptons

Another background contribution to $t\bar{t}\gamma$ SR comes from the so called *fake leptons*. They are the non-prompt leptons misidentified as prompt leptons, or jets that are mis-reconstructed as leptons. The non-prompt leptons could be coming from the decay of a heavy hadron (bottom or charm hadrons), or from a photon conversion, or they can be produced from the decay of a pion or a kaon.

The fake lepton background contribution to the single lepton channel is typically from QCD multijets events with associated photon production, and the contribution to the dilepton channel is typically from single lepton $t\bar{t}\gamma$ events or W+jets events with associated photon production, with a fake lepton in addition to the real lepton (events with two fake leptons very rarely happen).

6.3.1 Fake leptons in single lepton channel

The estimation of this background in the single lepton channel is done in two steps: firstly, the *matrix method* in the standard $t\bar{t}$ analysis is applied to estimate fake lepton events satisfying all the signal region cuts, except for the photon related ones; then, the photon cuts are reapplied to the above fake lepton sample to estimate the background of fake lepton plus photon. More details are given below.

The Matrix Method for fake lepton estimation: The method is based on classifying leptons into two categories: *loose* and *tight*. The loose and tight definitions are as follow:

- Tight definition (same as nominal lepton definition in section 4.1):
  For electrons: Tight LH, $p_{\mathrm{T}} > 25$ GeVand gradient isolation.
  For muons: Medium ID, $p_{\mathrm{T}} > 25$GeVand gradient isolation.
- Loose definition:
  For electrons: Medium LH and no isolation.
  For muons: Medium ID and no isolation. Besides, the low p_{T} muon trigger (below 51 GeV) for data of year 2016 is replaced from the nominal one as mentioned in Section 5.1 to the pre-scaled trigger of HLT_mu24, with a pre-scale of 50 to avoid possible trigger bias.

Based on these two lepton identification categories, two $t\bar{t}$ data samples are defined, differing only in the lepton identification while keeping the rest of the selection same as the signal region selection described in Section 5, but neglecting the photon related cuts.

The tight sample contains mostly events with real leptons, while the loose sample is enriched in events with fake leptons. The number of events in loose and tight samples is a linear combination of real and fake leptons:

$$N^{\mathrm{loose}} = N^{\mathrm{loose}}_{\mathrm{real}} + N^{\mathrm{loose}}_{\mathrm{fake}} \tag{16}$$

$$N^{\mathrm{tight}} = N^{\mathrm{tight}}_{\mathrm{real}} + N^{\mathrm{tight}}_{\mathrm{fake}} \tag{17}$$

Not reviewed, for internal circulation only

By defining ϵ_{real} (ϵ_{fake}) as the probability of real (fake) leptons in the loose sample to pass the tight selection, one can write:

$$N^{\text{tight}} = \epsilon_{\text{real}} N^{\text{loose}}_{\text{real}} + \epsilon_{\text{fake}} N^{\text{loose}}_{\text{fake}} \tag{18}$$

From equations 16-18, the number of tight events in data with a fakes leptons can be written as:

$$N^{\text{tight}}_{\text{fake}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (\epsilon_{\text{real}} N^{\text{loose}} - N^{\text{tight}}) \tag{19}$$

Both ϵ_{real} and ϵ_{fake} have dependency on lepton kinematics and event topology (e.g. number of jets or b-jets) and they are estimated as described in [47]. To estimate the fake lepton background, an event weight could be computed using these real and fake efficiencies, where the efficiencies would be functions of various object and event properties:

$$w_i = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (\epsilon_{\text{real}} - \delta_i) \tag{20}$$

where δ_i is unity if loose event i pass the tight selection and 0 otherwise. The sample of N^{loose} together with this set of weights constitute a fake lepton sample in the signal region with the photon related cuts stripped away. What's left to do is just to reapply the photon cuts to this sample, which is trivial.

The parametrisation of the efficiencies: The real and fake efficiencies are measured in a fake lepton control region with respect to different variables, which are called different parametrisation of the efficiencies. They are all centrally measured in the top group and distributed to each analyses. The variables or parametrisations tested in this analysis include:

- "pt": p_{T} of the lepton
- "eta": $|\eta|$ of the lepton
- "jetpt": p_{T} of the leading jet
- "dR": minimum ΔR between the lepton and the closest jet
- "dPhi": $\Delta\phi$ between the lepton and $E_{\text{T}}^{\text{miss}}$
- "nbtag": number of b-jets
- "mtw": transverse mass of W boson candidate [9]

Besides, the double combination and triple combination of the above choices are also tested. Thus in total $7 + C_7^2 + C_7^3 = 63$ parametrisations are studied. From the distribution of event yield of all these parameterisations, some give negative values or values that are very far way from the peak of the distribution, so they are removed in the following. In practise, all "pt" related parametrisations in e+jets channel and all "dPhi" related parametrisations in μ+jets channel are removed.

[9] $m_{\text{T}}(W) = \sqrt{2 p_{\text{T}}(l) \times E_{\text{T}}^{\text{miss}} (1 - \cos\phi)}$, where ϕ is the azimuthal angle between the lepton direction and the $E_{\text{T}}^{\text{miss}}$.

Not reviewed, for internal circulation only

For the e+jets channel, the remaining parametrisations are shown in the left plot of Figure 54. This distribution is projected onto its y axis, as shown in the right plot of Figure 54. It shows a peak with certain width. Thus it is decided to use the parametrisation that is most close to the peak of the distribution as the nominal and the two parametrisations corresponding to the lower and upper edge of this peak as up and down systematic variations.

The same procedure can be performed for the μ+jets channel and the results are summarized in Figure 55. And the finally predicted lepton fake backgrounds, together with its systematic variations, are summarized in table 17.

In Figure 56 and Figure 57, some kinematic distributions are compared between the nominal and the up and down variations for the e+jets and μ+jets channels respectively. We can see the up and down variations enclose the nominal very well. Besides, since the up and down variation already cover a wide range and the fake lepton is a small background, no additional systematics will be considered.

One thing to note is the occurrence of negative weights in the Lep fake samples which are a result of the method above. In general, negative weights are fine. If however the overall yield for the Lep fake sample in a given bin is less than zero, the yield for that bin is set to 1×10^{-6}. This is necessary for RooStats to avoid "NaN"'s. This means that for that bin entering the fit, just the Lep fake background process essentially has 0 events.

Figure 54: Left: the distribution of predicted lepton fake for different parameterisations in the e+jets channel. Right: the projection of the distribution in the left onto its y axis.

Table 17: The nominal and up and down variation of the fake lepton in the single lepton channel.

Variation	Down	Nominal	Up
e+jets para.	eta:dR:jetpt	eta:mtw	eta:nbtag:mtw
e+jets N	159.1 ± 20.9	293.5 ± 38.6	428.8 ± 51.6
μ+jets para.	pt:eta:dR	pt:mtw	jetpt:nbtag
μ+jets N	9.8 ± 25.0	72.1 ± 26.8	135.8 ± 38.4

Not reviewed, for internal circulation only

Figure 55: Left: the distribution of predicted lepton fake for different parameterisations in the μ+jets channel. Right: the projection of the distribution in the left onto its y axis.

Figure 56: The lepton p_{T}, leading jet p_{T}, MET, and MWT distributions for the nominal (black), up variation (blue) and down variation (red) of the fake lepton background in the e+jets channel.

Not reviewed, for internal circulation only

Figure 57: The lepton p_{T}, leading jet p_{T}, MET, and MWT distributions for the nominal (black), up variation (blue), and down variation (red) of the fake lepton background in the e+jets channel.

6.3.2 Fake leptons in dilepton channel

Events with one prompt lepton and one fake lepton are calculated in a control region, defined in the same way as the signal region described in Section 5 except for the opposite-sign two leptons are replaced by same-sign two leptons. The non-prompt (i.e. fake) leptons are identified using the truth MC record.

The same MC samples described in Section 3 are used to estimate this background. It was found that the majority is coming from $t\bar{t}$ single lepton events, with an extra fake lepton. While the most leptons in the same sign regions are from prompt sources as described in Table 18. Other is a lepton coming from the decay of a pion or a kaon or a misidentified lepton.

The final number of fakes can be considered as the same sign data events after substracting the prompt MC events, which is found to be very small and negligible in the same flavour channel (ee and $\mu\mu$) compared to the opposite sign signal region. Thus, a 100% uncertainty is assigned.

Not reviewed, for internal circulation only

Table 18: The number of events in the same sign region in data, and the expected number of events with two prompt leptons or events with at least one fake lepton estimated from MC. The numbers are normalized to the total integrated luminosity of 36.47 fb^{-1} and statistical uncertainties are given.

Component	$e\mu$	ee	$\mu\mu$
Prompt lepton	5.8 ± 1.3	7.6 ± 1.5	0.1 ± 0.1
Heavy flavour	1.9 ± 1.0	0	0.5 ± 0.4
Other	0.5 ± 0.3	0	0
Data	17	3	2

6.4 Prompt photons background

All backgrounds to the $t\bar{t}$ process are also considered when accompanied by prompt photon radiation. More specifically, it includes $W\gamma$ + jets ($W\gamma$), $Z\gamma$ + jets ($Z\gamma$), single top, diboson and $t\bar{t}V$ production with an additional prompt photon.

These backgrounds are estimated from MC simulations with the corresponding MC generator described in Section 3. Events are selected using the $t\bar{t}\gamma$ event selections as described in Section 5, The $W\gamma$ and $Z\gamma$ samples are normalised to the NLO cross sections.

For the single top and diboson samples, the QED radiation is simulated either internally in the parton shower or with PHOTOS [27].

Since the contributions from $W\gamma$ and $Z\gamma$ processes are the dominant ones in this prompt photon background dedicated validation regions are defined to check how well our MC is modelling these processes. The $Z\gamma$ background is modelled in the ee and $\mu\mu$ channels which can be found in Section 6.4.1, while the modelling of the $W\gamma$ background in the e+jets and μ+jets channels is described in Section 6.4.2.

Table 19 summarises the contributions of backgrounds with an emission of a prompt photon after the full selections in the dilepton and single lepton channels. The associated errors are the statistical uncertainties. From here-on, the prompt photon contributions from the single top, diboson and $t\bar{t}V$ samples are grouped into the Other prompt background category. In addition, depending on the channel, the $V\gamma$ process will also be included in Other prompt. I.e., Other prompt will also contain $Z\gamma$ prompt photons for the single lepton channels, and $W\gamma$ prompt photons for the dilepton channels.

Table 19: The expected yields of prompt photon background events from MC. The numbers are normalised to the total integrated luminosity of 36.1 fb^{-1}. Only statistical uncertainties are included.

Process	$e\mu$	$\mu\mu$	ee	e+jets	μ+jets
$W\gamma$	-	-	-	540.80±29.36	579.36±33.57
$Z\gamma$	-	53.66±13.65	21.40±3.85	295.23±46.27	151.98±13.21
Single top	5.81±1.14	2.64±0.83	3.66±1.21	84.61±5.45	92.17±6.20
Diboson	-	0.19±0.30	0.16±0.31	4.02±0.96	5.60±1.21
$t\bar{t}V$	2.50±0.20	1.52 ± 0.14	1.31±0.13	26.54±0.76	24.80±0.71

Not reviewed, for internal circulation only

6.4.1 $Z\gamma$ validation region

The $Z\gamma$ background modelling was checked in different validation regions, orthogonal to the SR, i.e. exactly two leptons and one photon as defined in Section 4.1 is required. The non-$Z\gamma$ contributions are estimated from MC in the same way as in the signal region. Some selections are different with respect to the signal region to have an enriched $Z\gamma$ VR. For the chosen VR the selections are defined as the following: jets ≥ 0 , number of b-tagged jets = 1 and finally the invariant mass of the two leptons is required to be in a mass window of [60, 100] GeV. With the requirement of b-tagged jets = 1 we enter the heavy flavour jet modelling regime which is close to our signal region. A requirement asking for more than one b-tagged jet introduces more $t\bar{t}\gamma$ signal into the region and so will not be considered.

Full systematics described in Chapter 10 are used in these fits[10]. After performing a binned likelihood fit (discussed in Chapter 8) to the jet p_T distribution we obtain scale factors of

$$\text{Heavy flavour } SF_{ee} = 1.033 + 0.350 - 0.243$$
$$\text{Heavy flavour } SF_{\mu\mu} = 1.096 + 0.287 - 0.225.$$

Pre- and post-fit distributions can be found in Figure 58. While difficult to see in these plots due to increased statistics, the $Z\gamma$ Sherpa samples are prone to slightly higher weights in very few events. Further studies have been done to see if this impacts the result. This can be seen in Appendix V.

A further cross-check can be performed to assess the performance of the VR in the b-tagged jets = 0 region, i.e., the light flavour region. The requirement on the cuts is identical to the above with the exception of the b-tagged jet cut. From this fit we obtain scale factors of

$$\text{Light flavour } SF_{ee} = 1.193 + 0.258 - 0.216$$
$$\text{Light flavour } SF_{\mu\mu} = 1.221 + 0.235 - 0.197.$$

These plots can be seen in Section 59. It can be concluded that we achieve decent modelling of the $Z\gamma$ background in the light and heavy flavour jet region

Pre-fit plots with only statistical uncertainty (due to CPU intensiveness for these plots) for both heavy and light jet regions can be seen in Appendix G.1 and G.2.

Important to note are that the results presented here are purely for validation reasons and at no point enter into the final fit or result.

[10] $t\bar{t}\gamma$ parton shower and scale variation shape systematics are not included here.

Not reviewed, for internal circulation only

(a) Heavy flavour *ee* channel

(b) Heavy flavour $\mu\mu$ channel

Figure 58: Heavy flavour validation regions (where number of b-tagged jets is 1) for the dilepton channel. All systematic uncertainties are included.

Not reviewed, for internal circulation only

(a) Light flavour *ee* channel

(b) Light flavour $\mu\mu$ channel

Figure 59: Light flavour validation regions (where number of b-tagged jets is 0) for the dilepton channel. All systematic uncertainties are included.

Not reviewed, for internal circulation only

6.4.2 $W\gamma$ validation region

Prompt photons in addition to $W \rightarrow l\nu$ form a non-negligible part of the single lepton background contribution in the signal region. This validation region is used as a cross-check as the $W\gamma$ background will be a free parameter in the final fit. The results presented here are purely for validation reasons and at no point enter into the final fit or result.

For our chosen validation region, the difference in cuts with respect to the SR are the following: $2 \leq$ number of jets ≤ 3, number of b-tagged jets = 1 and $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV. In addition, there is a further cut on the ELD distribution of < 0.04 to further reduce signal contamination. These cuts select an orthogonal phase space to our SR with the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut reducing contributions mainly from the Other prompt background, but also from Lep fake and $e \rightarrow \gamma$ fake backgrounds. For e+jets we include a cut of $m(\gamma, lep) < 80$ GeV to further reduce the contribution from $e \rightarrow \gamma$ fake[11].

Full systematics described in Chapter 10 are used in these fits[12]. After performing a binned likelihood fit (discussed in Chapter 8) to the jet $\mathrm{p_T}$ distribution we obtain scale factors of

$$\text{Heavy flavour } SF_{e\text{+jets}} = 1.143 + 0.327 - 0.320$$
$$\text{Heavy flavour } SF_{\mu\text{+jets}} = 1.262 + 0.132 - 0.136.$$

Pre- and post-fit distributions can be found in Figure 60. Due to a less pure VR and slightly larger weights found in Sherpa samples a larger uncertainty band can be seen for the e+jets distribution. Appendix V goes into more detail about this type of issue.

As a cross-check, a second VR is defined to encapsulate the data/MC agreement specifically for light jets. The difference in cuts from our SR are as follows: $1 \leq$ number of jets ≤ 3, number of b-tagged jets = 0, $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV, and $dR(\gamma, lep) < 2.8$. For the e+jets channel we exclude events where the mass between the photon and the lepton falls in a window of [60, 100] GeV. The scale factors for e+jets and μ+jets fits are

$$\text{Light flavour } SF_{e\text{+jets}} = 1.148 \pm 0.052$$
$$\text{Light flavour } SF_{\mu\text{+jets}} = 1.177 \pm 0.060.$$

This shows that the normalisation in data and MC are in good agreement for the light and heavy jet flavour phase space. These plots can be seen in Section 61.

Pre-fit plots with only statistical uncertainty (due to CPU intensiveness for these plots) for both heavy and light jet regions can be seen in Appendix H.1 and H.2.

[11] These cuts leave the majority of events with 2 jets, and only 4 events with more than two jets.
[12] $t\bar{t}\gamma$ parton shower and scale variation shape systematics are not included here.

Not reviewed, for internal circulation only

(a) Heavy flavour e+jets channel

(b) Heavy flavour μ+jets channel

Figure 60: Heavy flavour validation regions (where number of b-tagged jets is 1) for the single lepton channel. All systematic uncertainties are included.

Not reviewed, for internal circulation only

(a) Light flavour e+jets channel

(b) Light flavour μ+jets channel

Figure 61: Light flavour validation regions (where number of b-tagged jets is 0) for the single lepton channel. All systematic uncertainties are included.

Not reviewed, for internal circulation only

7 Cross section definitions

This section defines the fiducial region, in which the inclusive and differential cross sections will be measured. The various factors (for correcting signal efficiency, migration, and etc.) needed for the extraction of the cross sections from observation are also explained.

7.1 Fiducial region

The fiducial region is defined to mimic the signal region at reconstruction level, described in Section 5, in terms of both object selection and event selection. Thus, there are five fiducial regions corresponding to the five channels.

7.1.1 Particle definition for fiducial region

- **Leptons:**
  Electrons and muons are selected dressed with close by photons (photons which are not originating from hadrons, in a $\Delta R < 0.1$ cone around the lepton). Leptons are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, and not being originated from hadron decays.
- **Jets:**
  Jets are clustered with the anti-k_t algorithm with a radius of $R = 0.4$. Non-interacting particles and muons are not considered in the clustering. Jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$.
- **b-jets:**
  The flavour of the jet is determined by ghost matching [48].
- **Photons:**
  Photons are required to not originate from a hadron decay. Photons are required to have $E_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.37$.
- **Overlap Removal:**
  The overlap removal is done between the following object, in the following order:
  (i) Muon-jet: The jets with $\Delta R(\mu, j) \leq 0.4$ are removed.
  (ii) Electron-jet: The jets with $\Delta R(e, j) \leq 0.4$ are removed.
  (iii) Photon-jet: The jets with $\Delta R(j, \gamma) \leq 0.4$ to an isolated photon are removed. Isolation computed from charged stable particles in a $R = 0.3$ cone. Photon is considered as isolated if $p_T^{\mathrm{cone30}}/p_T < 0.1$.

7.1.2 Event selection for fiducial region

Stable particles after parton shower but before detector simulation are used for the following object definitions.

- **Leptons:**
  Exactly one good electron (muon) is required in the e+jets (μ+jets) channel. Exactly two good electrons (muons) are required in the ee ($\mu\mu$) channel. Exactly one good electron and one good muon are required in the $e\mu$ channel.

Not reviewed, for internal circulation only

- **Jets:**
  At least four (two) good jets in the single lepton (dilepton) channel are required, among which at least one should be a b-jet.

- **Photon:**
  Exactly one good photon is required. The event is dropped if the photon has a $\Delta R(j, \gamma) < 0.4$ with any good jet, or a $\Delta R(l, \gamma) < 1.0$ with any good lepton.

7.1.3 Acceptance and correction factor

The signal acceptance is defined as the fraction of events falling into the fiducial region, out of the total generated events:

$$A_i = \frac{N_{\text{gen}}^{\text{fid},i}}{N_{\text{gen}}^{\text{all}}}, \tag{21}$$

where $N_{\text{gen}}^{\text{all}}$ is the total number of generated events and $N_{\text{gen},i}^{\text{fid}}$ is the number of events generated inside the fiducial region i, with i running over the five channels.

The correction factor is used to unfold the number of signal events selected at reconstruction level $N_{\text{reco},i}$ in channel i to the number of signal events generated in the fiducial region $N_{\text{gen}}^{\text{fid},i}$ corresponding to channel i:

$$N_{\text{gen}}^{\text{fid},i} = \frac{N_{\text{reco},i}}{C_i}. \tag{22}$$

To do the unfolding, first the non-fiducial signal events in $N_{\text{reco},i}$ should be removed:

$$N_{\text{reco},i}^{\text{fid},i} = N_{\text{reco},i} \times (1 - f_{\text{mig},i}), \tag{23}$$

where $N_{\text{reco},i}^{\text{fid},i}$ is the number of events generated in fiducial region i and selected at reconstruction level in channel i, and $f_{\text{mig},i}$ is the fraction of events selected at reconstruction level in channel i that are not generated from the fiducial region i:

$$f_{\text{mig},i} = \frac{N_{\text{reco},i}^{non\text{-}fid,i}}{N_{reco,i}}. \tag{24}$$

e.g. events generated with true photon $p_{\text{T}} < 20$ GeV can be reconstructed as a photon with $p_{\text{T}} > 20$ GeV due to resolution effects of the detector, or events generated from other fiducial regions can be reconstructed into this fiducial region.

Not all of the generated events in the fiducial region will be reconstructed and selected, because of the event reconstruction and selection efficiencies. Thus, to get the true generated events in the fiducial region, an additional efficiency correction should be applied:

$$N_{\text{gen}}^{\text{fid},i} = \frac{N_{\text{reco},i}^{\text{fid},i}}{\epsilon_i}, \tag{25}$$

where ϵ_i is the efficiency of events generated in fiducial region i being selected at reconstruction level in channel i:

$$\epsilon_i = \frac{N_{\text{reco},i}^{\text{fid},i}}{N_{\text{gen}}^{\text{fid},i}}. \tag{26}$$

Not reviewed, for internal circulation only

Combining the Equation 23 and Equation 25, the correction factor can be expressed as:

$$C_i = \frac{N_{\text{reco},i}}{N_{\text{gen}}^{\text{fid},i}} = \frac{\epsilon_i}{1 - f_{\text{mig},i}} \,. \tag{27}$$

The first equation shows the exact way of calculation of C_i while the second one shows the physical meaning of this factor.

The values of the acceptances and correction factors are summarized in Table 20. The numbers are shown not only for the five channels but also for the merged single lepton and dilepton channels. Also, the signal modelling systematics as described in Section 10.2 are shown.

Table 20: Acceptances and correction factors for the five channels as well as the two merged channels. The systematics are shown as relative uncertainties in %. The systematics of the correction factors enter the final fit, while the systematics of the acceptances are just for shown.

Channel	e+jets	μ+jets	ee	$e\mu$	$\mu\mu$	single lepton	dilepton
Acceptance	0.041	0.041	0.0024	0.0047	0.0024	0.082	0.0096
Scales	0.72%	0.72%	0.71%	0.70%	0.73%	0.72%	0.71%
PDF	0.48%	0.48%	0.55%	0.60%	0.58%	0.48%	0.58%
ISR/FSR Up	4.51 ± 0.02%	4.90 ± 0.02%	-0.79 ± 0.40%	2.41 ± 0.21%	2.12 ± 0.95%	4.71 ± 0.02%	1.53 ± 0.28%
ISR/FSR Dn	-4.42 ± 0.02%	-4.35 ± 0.02%	-0.79 ± 0.40%	0.82 ± 0.20%	-1.84 ± 0.92%	-4.38 ± 0.02%	-0.25 ± 0.27%
Parton Shower	-1.10 ± 0.03%	-0.99 ± 0.03%	-0.60 ± 0.44%	0.38 ± 0.22%	-0.84 ± 1.00%	-1.04 ± 0.02%	-0.17 ± 0.30%
Correction	0.36	0.36	0.24	0.35	0.26	0.36	0.30
Scales	0.23%	0.08%	0.48%	0.23%	0.38%	0.15%	0.29%
PDF	0.09%	0.06%	0.23%	0.10%	0.20%	0.07%	0.13%
ISR/FSR Up	-0.25 ± 0.60%	0.28 ± 0.59%	-6.04 ± 2.85%	-2.76 ± 1.73%	0.89 ± 2.82%	0.01 ± 0.42%	-2.42 ± 1.31%
ISR/FSR Dn	0.27 ± 0.61%	0.76 ± 0.61%	-4.54 ± 2.87%	-1.32 ± 1.77%	5.12 ± 2.97%	0.52 ± 0.43%	-0.40 ± 1.35%
Parton Shower	2.48 ± 0.62%	1.85 ± 0.61%	-3.76 ± 2.94%	-0.63 ± 1.79%	-10.77 ± 2.52%	2.17 ± 0.44%	-3.52 ± 1.31%

7.2 Fiducial cross section

After subtracting background events $N_{\text{bkgs},i}$ from data $N_{\text{data},i}$ and applying the correction factor C_i, the true number of generated signal events in the fiducial region i can be derived. Dividing this number by the integrated luminosity, the fiducial cross section can be measured. These are summarized in the following expression:

$$\sigma_i^{\text{fid}} = \frac{N_{\text{data},i} - N_{\text{bkgs},i}}{L \times C_i} \,. \tag{28}$$

On the other hand, if the total cross section is known from theory prediction, the predicted fiducial cross section can be calculated from:

$$\sigma_i^{\text{fid}} = \sigma^{\text{tot}} \times A_i \,. \tag{29}$$

where A_i is the acceptance defined in the previous section. The calculation of the predicted fiducial cross section will be described in detail in Section 12.

Not reviewed, for internal circulation only

7.3 Differential cross section

In the fiducial region, the cross section can be measured as a function of an observable, e.g. photon p_{T} or η, which means a differential cross section measurement.

For this goal, first the signal spectrum of this observable $N_{\mathrm{reco},j}$ is measured from data spectrum subtracting background spectrum:

$$N_{\mathrm{reco},j} = N_{\mathrm{data},j} - N_{\mathrm{bkgs},j}\,, \tag{30}$$

where j is the binning index of the observable at reconstruction level and the channel index i is left out for simplicity.

Then, the differential cross section σ_k^{diff} can be related to the signal spectrum by:

$$N_{\mathrm{reco},j} = \frac{1}{1 - f_{\mathrm{mig},j}} \times \sum_k (L \times \sigma_k^{\mathrm{diff}} \times \epsilon_k \times M_{kj})\,, \tag{31}$$

- where k is the binning index of the observable at truth level (i.e. particle level), which is not necessarily binned in the same way as at reconstruction level, and
- $f_{\mathrm{mig},j}$ is the fraction of signal events migrated from outside the fiducial region into the bin j of the reconstructed distribution and for these events, the objects that are used to construct the observable to be unfolded have to be matched between the reconstruction and truth levels with a ΔR value smaller than 0.1. The ΔR distributions for the photon and leptons of the events passing both the fiducial region selection and event selection are shown in Figure 62 for the single-lepton and dilepton channels. In all cases, the distributions peak around zero, except for the leptons in the dilepton channel, where a tail at large ΔR value can be seen, which implies a decreased matching efficiency.
- ϵ_k is the efficiency of signal events generated in bin k of the truth distribution but got reconstructed and selected and for these events, the relevant objects have to match as above.
- M_{kj} is the bin-by-bin migration matrix, representing the probability of a signal event generated in bin k of the truth distribution but observed in bin j of the reconstructed distribution. The events used to calculate this probability have to pass both the fiducial region selection and event selectionand and the relevant objects have to match as above.

Figure 62: The distributions of ΔR between the object (photon or lepton) at recontruction and truth levels for the single-lepton and dilepton channels.

Not reviewed, for internal circulation only

Solving the above equation, the differential cross section can be expressed as:

$$\sigma_k^{\text{diff}} = \frac{1}{L} \times \frac{1}{\epsilon_k} \times \sum_j M_{jk}^{-1} \times (N_{\text{data},j} - N_{\text{bkgs},j}) \times (1 - f_{\text{mig},j}), \tag{32}$$

where M_{jk}^{-1} is the inversion of the bin-by-bin migration matrix.

In this analysis, it is the normalized differential cross section instead of the above absolute cross section that is being measured. The predicted normalized differential cross section is directly taken from the MC sample, thus without higher order correction.

Only two channels are defined for the differential cross section measurement, which are the single lepton and dilepton channel. This is due to the similarity of each single lepton (dilepton) channels and for the aim of increasing data statistics in the low populated region of the distributions.

The binning of the distributions to be unfolded are optimized taking into account of the detector's resolution effect as well as the expected data fluctuation. More specifically: 1) the bin width is chosen to be wider than twice of the variable's resolution, so that more than 68% of the events will be in the same bin of the truth and reconstructed distributions, 2) the bin width can not get narrower than a size that has more than 10% (15%) statistical uncertainty of the bin for the single lepton (dilepton) channel. A re-binning can be performed for the final unfolded results, in the case of the statistical uncertainties getting larger. Another choice of binning is shown in Appendix O.

Figure 63 shows the efficiencies and migration fractions (in the form of one minus the fraction) for the single lepton channel, for the observables photon p_{T}, photon η and $\Delta R(\text{lepton}, \gamma)$. and Figure 64 shows them in the dilepton channel, for the observables photon p_{T}, photon η, $\Delta R_{\text{min}}(\text{lepton}, \gamma)$, $\Delta\eta(\text{lepton}, \text{lepton})$ and $\Delta\phi(\text{lepton}, \text{lepton})$. In the efficiency plots, the probabilities of matching all the objects relevant to the observable to be unfolded are shown for the events passing both the fiducial region selection and event selection. In the migration-fraction plots, the same probabilities are shown for the migrated events. The matching probabilities are in general high, except for the observables involving the two leptons in the dilepton channel, The numberical values of these distributions as well as their overall values across all the bins of the relevant observable are shown in Table 21 to 28. Figures 65 and 66 show the corresponding bin-by-bin migration matrices, normalized to the number of events in each row, in the single lepton and dilepton channels, respectively.

Table 21: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the photon p_T of the single lepton channel.The object-matching probabilities defined in the text are also shown.

photon p_T	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Efficiency	0.27 ± 0.16	0.23 ± 0.00	0.27 ± 0.00	0.29 ± 0.00	0.30 ± 0.00	0.32 ± 0.00	0.34 ± 0.00	0.34 ± 0.00	0.36 ± 0.00	0.41 ± 0.00
Match. prob.	1.00 ± 0.01	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
1 - f_{mig}	0.77 ± 0.25	0.72 ± 0.00	0.77 ± 0.00	0.77 ± 0.00	0.80 ± 0.00	0.79 ± 0.01	0.81 ± 0.01	0.81 ± 0.01	0.82 ± 0.01	0.84 ± 0.00
Match. prob.	0.85 ± 0.27	0.72 ± 0.00	0.77 ± 0.00	0.77 ± 0.00	0.80 ± 0.00	0.79 ± 0.01	0.81 ± 0.01	0.81 ± 0.01	0.82 ± 0.01	0.84 ± 0.00

Table 22: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the photon $|\eta|$ of the single lepton channel.The object-matching probabilities defined in the text are also shown.

photon $\lvert\eta\rvert$	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8
Efficiency	0.27 ± 0.12	0.29 ± 0.00	0.29 ± 0.00	0.30 ± 0.00	0.30 ± 0.00	0.30 ± 0.00	0.30 ± 0.00	0.20 ± 0.00	0.27 ± 0.00
Match. prob.	1.00 ± 0.01	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
1 - f_{mig}	0.77 ± 0.19	0.77 ± 0.00	0.77 ± 0.00	0.78 ± 0.00	0.77 ± 0.00	0.76 ± 0.00	0.76 ± 0.00	0.76 ± 0.00	0.75 ± 0.00
Match. prob.	0.85 ± 0.17	0.77 ± 0.00	0.77 ± 0.00	0.78 ± 0.00	0.77 ± 0.00	0.76 ± 0.00	0.76 ± 0.00	0.76 ± 0.00	0.75 ± 0.00

Not reviewed, for internal circulation only

(a)

(b)

Figure 63: The reconstruction and matching efficiencies (a) and one minus the fraction of non-outside migration $(1 - f_{\text{mig}})$ (b) for three different observables in the single lepton channel. In the efficiency plots, the probabilities of matching all the objects relevant to the observable to be unfolded are shown for the events passing both the fiducial region selection and event selection. In the migration-fraction plots, the same probabilities are shown for the migrated events.

Table 23: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the $\Delta R(\text{lepton}, \gamma)$ of the single lepton channel.The object-matching probabilities defined in the text are also shown.

$\Delta R(\text{lepton}, \gamma)$	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Efficiency	0.27 ± 0.14	0.27 ± 0.00	0.28 ± 0.00	0.28 ± 0.00	0.28 ± 0.00	0.28 ± 0.00	0.28 ± 0.00	0.28 ± 0.00	0.27 ± 0.00	0.27 ± 0.00
Match. prob.	1.00 ± 0.01	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
1 - f_{mig}	0.77 ± 0.23	0.76 ± 0.00	0.79 ± 0.00	0.79 ± 0.00	0.78 ± 0.00	0.77 ± 0.00	0.77 ± 0.00	0.77 ± 0.00	0.76 ± 0.00	0.76 ± 0.00
Match. prob.	0.76 ± 0.23	0.76 ± 0.00	0.79 ± 0.00	0.79 ± 0.00	0.78 ± 0.00	0.77 ± 0.00	0.77 ± 0.00	0.77 ± 0.00	0.76 ± 0.00	0.76 ± 0.00

Table 24: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the photon p_T of the dilepton channel.The object-matching probabilities defined in the text are also shown.

photon p_T	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Efficiency	0.27 ± 0.15	0.23 ± 0.00	0.26 ± 0.00	0.28 ± 0.01	0.29 ± 0.01	0.30 ± 0.01	0.33 ± 0.01	0.33 ± 0.01	0.35 ± 0.01	0.38 ± 0.01
Match. prob.	1.00 ± 0.01	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
1 - f_{mig}	0.89 ± 0.20	0.86 ± 0.01	0.89 ± 0.01	0.89 ± 0.01	0.91 ± 0.01	0.91 ± 0.01	0.91 ± 0.01	0.91 ± 0.01	0.91 ± 0.02	0.93 ± 0.01
Match. prob.	0.78 ± 0.30	0.86 ± 0.01	0.89 ± 0.01	0.89 ± 0.01	0.91 ± 0.01	0.91 ± 0.01	0.91 ± 0.01	0.91 ± 0.01	0.91 ± 0.02	0.93 ± 0.01

Not reviewed, for internal circulation only

(a)

(b)

Figure 64: The efficiency (a) and one minus the fraction of outside migration $(1 - f_{\text{mig}})$ (b) for five different observables in the dilepton channel. In the efficiency plots, the probabilities of matching all the objects relevant to the observable to be unfolded are shown for the events passing both the fiducial region selection and event selection. In the migration-fraction plots, the same probabilities are shown for the migrated events.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 65: The bin-by-bin migration matrix of the photon p_{T} (a), the photon $|\eta|$ (b), and $\Delta R(\mathrm{lepton}, \gamma)$ (c), in the single lepton channel.

Table 25: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the photon $|\eta|$ of the dilepton channel.The object-matching probabilities defined in the text are also shown.

photon $\lvert\eta\rvert$	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8
Efficiency	0.27 ± 0.12	0.28 ± 0.01	0.28 ± 0.01	0.29 ± 0.01	0.28 ± 0.01	0.28 ± 0.01	0.30 ± 0.01	0.20 ± 0.00	0.26 ± 0.01
Match. prob.	1.00 ± 0.01	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
1 - f_{mig}	0.89 ± 0.15	0.89 ± 0.01	0.90 ± 0.01	0.88 ± 0.01	0.89 ± 0.01	0.90 ± 0.01	0.88 ± 0.01	0.89 ± 0.01	0.87 ± 0.01
Match. prob.	0.78 ± 0.19	0.89 ± 0.01	0.90 ± 0.01	0.88 ± 0.01	0.89 ± 0.01	0.90 ± 0.01	0.88 ± 0.01	0.89 ± 0.01	0.87 ± 0.01

Table 26: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the min $\Delta R(\mathrm{lepton}, \gamma)$ of the dilepton channel.The object-matching probabilities defined in the text are also shown.

min $\Delta R(\mathrm{lepton}, \gamma)$	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Efficiency	0.22 ± 0.10	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.21 ± 0.01
Match. prob.	0.81 ± 0.18	0.81 ± 0.01	0.80 ± 0.01	0.81 ± 0.01	0.82 ± 0.01	0.81 ± 0.01	0.82 ± 0.01	0.82 ± 0.01	0.81 ± 0.02	0.79 ± 0.01
1 - f_{mig}	0.72 ± 0.19	0.71 ± 0.01	0.71 ± 0.01	0.71 ± 0.01	0.73 ± 0.01	0.72 ± 0.01	0.74 ± 0.01	0.71 ± 0.02	0.75 ± 0.02	0.70 ± 0.01
Match. prob.	0.39 ± 0.16	0.71 ± 0.01	0.71 ± 0.01	0.71 ± 0.01	0.73 ± 0.01	0.72 ± 0.01	0.74 ± 0.01	0.71 ± 0.02	0.75 ± 0.02	0.70 ± 0.01

Not reviewed, for internal circulation only

(a)

(b)

(c)

(d)

(e)

Figure 66: The bin-by-bin migration matrix of the photon p_{T} (a), the photon $|\eta|$ (b), $\Delta R_{\mathrm{min}}(\mathrm{lepton}, \gamma)$ (c), $\Delta\eta(\mathrm{lepton}, \mathrm{lepton})$ (d) and $\Delta\phi(\mathrm{lepton}, \mathrm{lepton})$ (e), in the dilepton channel.

Not reviewed, for internal circulation only

Table 27: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the $\Delta\Phi$(lepton, lepton) of the dilepton channel.The object-matching probabilities defined in the text are also shown.

$\Delta\Phi$(lepton, lepton)	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Efficiency	0.22 ± 0.10	0.21 ± 0.01	0.22 ± 0.01	0.19 ± 0.01	0.20 ± 0.01	0.21 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.23 ± 0.01	0.23 ± 0.01
Match. prob.	0.81 ± 0.18	0.83 ± 0.01	0.82 ± 0.01	0.79 ± 0.02	0.80 ± 0.01	0.82 ± 0.01	0.83 ± 0.01	0.78 ± 0.01	0.81 ± 0.01	0.82 ± 0.01
1 - f_{mig}	0.72 ± 0.19	0.72 ± 0.02	0.74 ± 0.02	0.71 ± 0.02	0.71 ± 0.02	0.73 ± 0.01	0.73 ± 0.01	0.70 ± 0.01	0.71 ± 0.01	0.72 ± 0.01
Match. prob.	0.39 ± 0.16	0.72 ± 0.02	0.74 ± 0.02	0.71 ± 0.02	0.71 ± 0.02	0.73 ± 0.01	0.73 ± 0.01	0.70 ± 0.01	0.71 ± 0.01	0.72 ± 0.01

Table 28: The numerical values of the efficiency and one minus the fraction of outside migration (1-f_{mig}) for the $\Delta\eta$(lepton, lepton) of the dilepton channel.The object-matching probabilities defined in the text are also shown.

$\Delta\eta$(lepton, lepton)	Overall	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8
Efficiency	0.22 ± 0.11	0.21 ± 0.01	0.22 ± 0.01	0.21 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.22 ± 0.01	0.21 ± 0.00	0.22 ± 0.00
Match. prob.	0.81 ± 0.20	0.79 ± 0.01	0.83 ± 0.01	0.81 ± 0.01	0.83 ± 0.01	0.82 ± 0.01	0.79 ± 0.01	0.81 ± 0.01	0.81 ± 0.01
1 - f_{mig}	0.72 ± 0.22	0.69 ± 0.02	0.72 ± 0.01	0.72 ± 0.01	0.72 ± 0.01	0.74 ± 0.01	0.71 ± 0.01	0.72 ± 0.01	0.72 ± 0.01
Match. prob.	0.39 ± 0.17	0.69 ± 0.02	0.72 ± 0.01	0.72 ± 0.01	0.72 ± 0.01	0.74 ± 0.01	0.71 ± 0.01	0.72 ± 0.01	0.72 ± 0.01

Not reviewed, for internal circulation only

8 Fit strategy

The events selected in the signal region, as described in Section 5, subtracting the estimated background events as detailed in Section 6, are the observed signal events, which can be used to:

- be compared with the MC predicted signal events and use their ratio, denoted as μ and called signal strength, to quantify the agreement between observation and SM at the reconstruction level.
- extract the fiducial cross section, after correcting for the signal efficiency loss and migration from outside fiducial region using MC. Furthermore, with the MC predicted signal acceptance, the fiducial phase space can be extrapolated to the total phase space, thus a total cross section measurement is possible.
- with the background subtraction done differentially, perform differential cross section measurement.

In the following, the strategies for the above measurements are described in details.

8.1 Likelihood function

Before introducing the detailed strategies, the general likelihood function, which is the base for all sort of physical interpretation, should be introduced. The most general form of the likelihood is as follows:

$$\mathcal{L} = \mathcal{L}(p.o.i, \boldsymbol{\theta}|\mathbf{ELD}), \tag{33}$$

where ELD is the observed event level discriminator to be fitted as introduced in Section 5.2, $p.o.i$ is the parameter(s) of interest which could be the signal strength, fiducial cross section, total cross section, or differential cross section, and $\boldsymbol{\theta}$ is the modeling of the systematic uncertainties described in section 10.

In practise, a binned maximum likelihood fit is performed and in each bin, the observed number of data follows a Poisson distribution. Thus, the likelihood can also be expressed as:

$$\mathcal{L} = \prod_j \prod_i P(N_{i,j}|N^s_{i,j} + \sum_b N^b_{i,j}) \times \prod_t G(\vartheta|\theta_t, 1), \tag{34}$$

where

- $P(N_{i,j}|N^s_{i,j} + \sum_b N^b_{i,j})$ is the Poisson function,
- $N_{i,j}$, $N^s_{i,j}$ and $N^b_{i,j}$ are the observed number of data, signal events and background events in bin i of the ELD of channel j, respectively,
- j could run over e+jets, μ+jets, ee, $e\mu$, and $\mu\mu$ channels,
- $G(\vartheta|\theta_t, 1)$ is a Gaussian function modeling the systematic uncertainty $\theta_t \in \boldsymbol{\theta}$, with a mean of 0 and a standard deviation of 1. θ_t will enter the expressions for $N^s_{i,j}$ and $N^b_{i,j}$ in a way that when it varies within [-1,+1], the corresponding predictions vary within their estimated standard deviation.

Depending on the definition of the $p.o.i$, it will enter the expressions for $N^s_{i,j}$ differently, thus corresponding to different interpretation of the data, which will be discussed below.

Not reviewed, for internal circulation only

8.2 Signal strength

The signal strength is defined as the ratio between the observed number of signal events over the expected one:

$$\mu = \frac{N_s^{obs}}{N_s^{SM}} \,. \tag{35}$$

If more than one channel is involved in the fit, μ will be the average of ratios in each channel. Alternatively, one can also define multiple μ's, one for each channel j:

$$\mu_j = \frac{N_{s,j}^{obs}}{N_{s,j}^{SM}} \,. \tag{36}$$

Signal strength enters the likelihood according to:

$$N_{i,j}^{s} = \mu \times N_{s,j}^{SM} \times f_{i,j}^{ELD} \tag{37}$$

or

$$N_{i,j}^{s} = \mu_j \times N_{s,j}^{SM} \times f_{i,j}^{ELD} \,, \tag{38}$$

where $N_{s,j}^{SM}$ is the predicted number of signal events in the channel j and $f_{i,j}^{ELD}$ is the fraction of events falling into a particular ELD bin i of channel j, which is in fact the shape information of the ELD.

This way of parametrisation is used most often since it compares between observation and SM prediction without any need of unfolding.

8.3 Fiducial cross section

When the $p.o.i$ is the fiducial cross section σ_j^{fid}, as defined in Section 7, the expression for $N_{i,j}^{s}$ becomes:

$$N_{i,j}^{s} = L \times \sigma_j^{fid} \times C_j \times f_{i,j}^{ELD} \,, \tag{39}$$

where L is the integrated luminosity and C_j is the correction factor (corrects for signal efficiency and migration), as described in Section 7.1.

In this way, the measurement does not depend on the prediction of the normalisation of signal event.

Usually there will be one fiducial cross section for one channel. But sometimes, the fiducial regions of each channel can be combined into one region so that a combined fiducial cross section is measured.

8.4 Types of fits

In the above sections references to fitting the ELD distribution for the different channels are made. However, there are multiple ways this type of fit can be carried out. An overview of the scenarios is shown in Figure 67 and explained below.

For the individual channels this is straightforward and separate likelihood fits in each of the five channels (e+jets, μ+jets, ee, $e\mu$ and $\mu\mu$) are performed. When it comes to carrying out single lepton and dilepton channel fits there are two options considered. The first is a “merged" fit where histograms from e+jets and

Not reviewed, for internal circulation only

μ+jets are essentially added (or stacked) together to form the single lepton channel and histograms from ee, $e\mu$ and $\mu\mu$ are added to form the dilepton channel. The advantage (and ultimately deciding factor) to using this method is that only one signal region is fitted, meaning only one outputted signal strength and set of histograms for each single lepton and dilepton channel. This makes differential measurements more straightforward. These scenarios are shown in purple on the figure.

A different method is to fit multiple signal regions at once which allows the possibility for nuisance parameters to be more effectively constrained. For example, a “combined" fit could be carried out on the e+jets and μ+jets workspaces which outputs one single lepton signal strength but two sets of histograms. Hence, a multiple signal region fit. Scenarios that use this strategy are in yellow on the figure.

Finally, the figure shows the chosen scenarios used for the final fits. These have red borders. This includes the five individual channels, the single lepton and dilepton merged fits, and the “inclusive 5 channel fit". The last is a combined fit of the five individual channels, i.e. fitting five signal regions simultaneously. Since differential measurements are only done separately for single lepton and dilepton channels this is not an issue. These results will be shown in Section 13. For completeness and as a cross-check a plot with all scenarios is shown in Appendix Q.4. One can see that there are negligible differences in the scenarios and all results agree with each other.

Figure 67: An overview of the different scenarios for how fits can be carried out. Purple indicates a merged or single region fit, yellow indicates a combined or multiple region fit. The red boxes indicate the final chosen strategy.

Not reviewed, for internal circulation only

9 Unfolding

The procedure of retrieving the truth spectrum of an observable from the measured spectrum is called unfolding [49]. The measured spectrum is unfolded to the particle level, i.e, after the hadronisation and showering processes. The unfolding is performed in order to correct for the detector effects due to limited resolution and acceptance, this enables the measured distributions to be compared to the prediction of an existing theory. Unfolding also enables the comparison with other experiments which could have different migration matrices.

The main problem of unfolding is the inversion of the migration matrix. To derive the inverse of the migration matrix, various methods have been developed. In this section, different methods for unfolding (to particle level) are introduced and compared [49, 50]: the benchmark method which is the iterative Bayesian method (IBS) [51] is introduced in 9.1, and two other methods, the singular value decomposition (SVD) [52] and the bin-by-bin approach [50], which are used as cross-check methods, are introduced in 9.2. Another method which could be used for cross-check is the maximum likelihood method (ML), which was already described in Section **??** and [50]. All of the methods except for the ML are implemented in the RooUnfold package [53].

For each of the unfolding methods, the same inputs are used: efficiency, outside fiducial region migration, and bin-by-bin migration matrix, which are all already introduced in Section 7.3. Some illustrations for these inputs can be found in Figures 64 and 63 for efficiencies and outside migrations, and in Figures 66 and 65 for bin-by-bin migration matrices, in the dilepton and the single lepton channels respectively.

Besides, for simplicity, a simple cut-and-count analysis, instead of ELD template fit, is performed in each bin of the observable.

9.1 Iterative Bayesian method

Before introducing the method, it will be convenient to rewrite the Equation 32 for differential cross section as following:

$$N^{\text{corr}}_{\text{gen},k} = \sum_j M^{-1}_{jk} \times N^{\text{corr}}_{\text{reco},j}\,, \tag{40}$$

where $N^{\text{corr}}_{\text{gen},k}$ is the efficiency corrected truth spectrum:

$$N^{\text{corr}}_{\text{gen},k} = L \times \sigma^{\text{diff}}_k \times \epsilon_k\,, \tag{41}$$

and $N^{\text{corr}}_{\text{reco},j}$ is the observed spectrum with background and outside fiducial migration corrected:

$$N^{\text{corr}}_{\text{reco},j} = (N_{\text{data},j} - N_{\text{bkgs},j}) \times (1 - f_{\text{mig},j})\,. \tag{42}$$

In the above formulas, “gen” means particle level and k is the corresponding binning index, “reco” means reconstruction level and j is the corresponding binning index. The definitions of ϵ_k, M_{jk}, and $f_{\text{mig},j}$ can be found in Section 7.3.

Not reviewed, for internal circulation only

Then the differential cross section measurement is reduced to the problem of converting the corrected observed spectrum $N^{\text{corr}}_{\text{reco},j}$ to the corrected truth spectrum $N^{\text{corr}}_{\text{gen},k}$ by using the inverted migration matrix M^{-1}_{jk}.

To revert the migration matrix, the iterative Bayesian method is proposed by D'Agostini, and explained in more details in his paper [51]. Concepts of cause and effect are introduced where E means effect, corresponding to the measured values at reconstruction level and C means cause, corresponding to the truth ones at particle level. The migration matrix can be expressed as:

$$M_{kj} = P(E_j|C_k)\,, \tag{43}$$

which means the probability for observing an effect in bin j, given a cause in bin k. Then its inversion M^{-1}_{jk} can be expressed as:

$$M^{-1}_{jk} = P(C_k|E_j)\,, \tag{44}$$

which means the probability for a cause in bin k, given the observation of an effect in bin j. According to Bayesian theorem, $P(C_k|E_j)$ can be written in the following equation:

$$P(C_k|E_j) = \frac{P(E_j|C_k)\cdot P_0(C_k)}{\sum_l P(E_j|C_l)\cdot P_0(C_l)}\,, \tag{45}$$

given some prior truth distribution $P_0(C_l)$, from e.g. the MC sample.

Then this inversion matrix can be used to derive an unfolded distribution, which again can be fed into the calculation of the reversion matrix as a prior. Finally, when the unfolded distribution and the prior, which is the unfolded distribution from the last iteration, converge, the iteration stops. The differential cross section can be then easily derived by solving Equation 41.

9.1.1 The pseudo-data

As the unfolding is an iterative procedure, the convergence point is needed to be determined and the point will be know as the number of iterations. The determined number will be used later to perform various checks in order to confirm that the unfolding study is not biased against the bin choice and the specific choice of the MC sample to obtain the truth spectrum and the measured spectrum. To perform these checks, the measured spectrum is replaced by the pseudo-data.

The pseudo-data is generated using the $t\bar{t}\gamma$ MC events which are randomly sampled into two parts, *testing* and *training*, in a way that the testing part has the equivalent statistical power as the observed signal. The training sample is used to build the migration matrix from, and the testing sample is being unfolded. 100 pseudo-data sets are generated in this way.

The following Sections 9.1.2, 9.1.3 and 9.1.4 are using the same pseudo-data.

9.1.2 Optimising the number of iterations

In order to choose the number of iterations to be used for the unfolding, two methods are employed. The final chosen number of iterations is a compromise between having a low difference in the unfolded results between different iterations, having a stable statistical uncertainty in all bins, and having no bias against any number of iterations.

Not reviewed, for internal circulation only

In the first method, the convergence of the unfolded results are checked. 1000 pseudo-data sets are generated by using a Poisson distribution, from the reconstructed-level distribution which is obtained from the testing part. The new pseudo-data is then unfolded using the migration matrix from the training part. The unfolding is done for different number of iterations. The number of pseudo-unfolded events after $i + 1$ iterations, is compared to the previous number of unfolded events obtained after i iterations. The relative difference between the two numbers is calculated and plotted versus the observable. The result of the test is shown in Figure 68 in the dilepton channel and Figure 69 in the single lepton channel. The first number of iterations that results in the relative difference as low as less than 1% in all the bins, is taken into account:

$$\frac{N^{\text{unfold, i+1}} - N^{\text{unfold, i}}}{N^{\text{unfold, i+1}}} < 1\%, \tag{46}$$

where $N^{\text{unfold, i}}$ ($N^{\text{unfold, i+1}}$) is the unfolded number of event at truth level at iteration i ($i + 1$).

From Figures 68 and 69, the number of iterations 3 seems to be preferred. However, the decision should not be taken before checking the RMS.

In the second method, to check the statistical uncertainty on the unfolded pseudo-data, the 100 pseudo-data sets described previously are used. The training pseudo-data is used to get the migration matrix, while the pseudo-data of the testing reconstructed-level is unfolded. The quantity RMS is calculated to compare the unfolded reconstructed distribution to the pseudo-particle level distribution:

$$\text{RMS} = \frac{(\text{truth} - \text{unfold})}{\text{unfold}}. \tag{47}$$

A bias against some number of iterations is observed in the dilepton channel in the fourth bin of photon p_{T}, in particular, and a smaller bias is observed in η and minimum $\Delta R(\text{lepton}, \gamma)$, as shown in Figure 70, Figure 71 and Figure 72, while the other two observables ($\Delta\eta(\text{lepton, lepton})$ and $\Delta\phi(\text{lepton, lepton})$) are stable, as shown in Figure 73 and Figure 74. Furthermore, the uncertainty is mostly stable, only increasing slightly with more number of iterations. For the single lepton channel, most of the bins are stable and no clear bias is observed, as can be seen in Figure 75, Figure 76 and Figure 77 for the photon p_{T}, η, and $\Delta R(\text{lepton}, \gamma)$, respectively.

Since in the first method, the two variables and channels agree for 3 iterations, and in the second method most of bins are also stable after 3 iterations, the 3 iterations are chosen, to be harmonised among all the observables and the both channels. The number of 6 iterations which looks to be preferred in the photon p_{T} in the dilepton channel is checked in Appendix N.1. The check compares all the tests in the next sections and the final results when three or six iterations are used. No visible difference is observed, hence the three iterations are justified.

Not reviewed, for internal circulation only

(a) (b) (c) (d) (e)

Figure 68: The relative difference on the unfolded result after two successive iterations, as a function of five different observables in dilepton channel.

Not reviewed, for internal circulation only

(a) (b)

(c)

Figure 69: The relative difference on the unfolded result after two successive iterations as a function of the photon p_{T} (a), the photon $|\eta|$ (b), and $\Delta R(\mathrm{lepton}, \gamma)$ (c) in the single lepton channel.

Not reviewed, for internal circulation only

Figure 70: The RMS against the number of iterations for the 9 bins in p_{T} in the dilepton channel.

Not reviewed, for internal circulation only

Figure 71: The RMS against the number of iterations for the 8 bins in η in the dilepton channel.

Not reviewed, for internal circulation only

Figure 72: The RMS against the number of iterations for the 9 bins in the minimum $\Delta R(\text{lepton}, \gamma)$ in the dilepton channel.

Not reviewed, for internal circulation only

Figure 73: The RMS against the number of iterations for the 8 bins in $\Delta\eta(\text{lepton}, \text{lepton})$ in the dilepton channel.

Not reviewed, for internal circulation only

Figure 74: The RMS against the number of iterations for the 9 bins in $\Delta\phi(\text{lepton}, \text{lepton})$ in the dilepton channel.

Not reviewed, for internal circulation only

Figure 75: The RMS against the number of iterations for the 9 bins in p_{T} in the single-lepton channel.

Not reviewed, for internal circulation only

Figure 76: The RMS against the number of iterations for the 8 bins in η in the single-lepton channel.

Not reviewed, for internal circulation only

Figure 77: The RMS against the number of iterations for the 9 bins in $\Delta R(\text{lepton}, \gamma)$ in the single-lepton channel.

Not reviewed, for internal circulation only

9.1.3 The closure test

The closure test is performed to check if the the unfolding procedure can recover the truth spectrum. To do that, the 100 pseudo-data distributions of the testing part at the reconstruction-level are unfolded, using the migration matrix from the training part. A good closure is obtained, as can be seen in Figure 78 for the dilepton channel and in Figure 79 for single lepton channel.

Not reviewed, for internal circulation only

(a)

(b)

(c)

(d)

(e)

Figure 78: The ratio of the number of unfolded events for one pseudo-data after 3 iterations to the number of truth events in the dilepton channel.

Not reviewed, for internal circulation only

(a) (b)

(c)

Figure 79: The ratio of the number of unfolded events for one pseudo-data after 3 iterations to the number of truth events, as a function of the photon p_{T} (a), the photon η (b), and $\Delta R(\mathrm{lepton}, \gamma)$ (c) in the single lepton channel.

Not reviewed, for internal circulation only

9.1.4 Pull study

The binning of the observables are initially chosen according to the criteria described earlier in Section 7.3. The stability of the binning choice is checked by performing the pull tests.

The same pseudo-data described in Section 9.1.1 is used and another 1000 pseudo-experiments are built from the reconstructed-level in the testing part using the Poisson distribution, and then unfolded. The unfolded pseudo-results are then compared to the pseudo-particle level. The pulls are calculated per bin as:

$$\text{Pull} = \frac{(\text{truth} - \text{unfold})}{\sigma_{\text{unfold}}}, \tag{48}$$

and then fit by a Gaussian function (see Appendix P). The pull is expected to have a mean value of zero, which indicates no bias, and a width of one, which could mean that the statistical uncertainty is estimated correctly. Figure 80 shows the result of the test for dilepton channel and Figure 81 for single lepton channel. No bias in the bin choice is seen and the width is consistent with one.

Not reviewed, for internal circulation only

(a)

(b)

(c) (d)

(e)

Figure 80: The pull test in each bin of the five observables, obtained after 3 iterations, in the dilepton channel.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 81: The pull test in each bin of the photon p_{T} (a), the photon η (b) and the $\Delta R(\mathrm{lepton}, \gamma)$ (c), obtained after 3 iterations, in the single lepton channel.

Not reviewed, for internal circulation only

9.1.5 Stress tests

The stress test is performed in order to verify that the unfolding procedure is not biased to any specific shape of the particle level distribution. The particle-level and reconstruction-level distributions obtained from the nominal MC sample are reweighed, and then the reweighted reconstructed distribution is unfolded using the nominal inputs from the MC sample, and the unfolded results are compared to the corresponding particle level distribution.

Different weights have been checked, the first one is by taking the observed difference at reconstruction level between data and MC as the following:

$$\text{weight} = 1 + Y \cdot \frac{\text{data}_i - \text{MC}_i}{\text{data}_i} = 1 + Y \cdot \text{Obs}, \tag{49}$$

where i is the bin index and Y= 1, -1. The result of the stress test is shown in Figure 82 and Figure 83 for the dilepton and single lepton channels, respectively. The unfolding is able to retrieve the reweighted particle distribution for all observables in both channels.

A different weight, corresponding to a linear skewness of the shape, is also used. It is defined as the following, in case of the photon p_{T}:

$$\text{weight} = 1 + y \cdot \frac{100 - i}{300} = 1 + y \cdot X, \tag{50}$$

while given by the following for the photon η:

$$\text{weight} = 1 + y \cdot \frac{1.2 - i}{2.37} = 1 + y \cdot X, \tag{51}$$

and for the minimum $\Delta R(\text{lepton}, \gamma)$ by:

$$\text{weight} = 1 + y \cdot \frac{1.8 - i}{6} = 1 + y \cdot X, \tag{52}$$

and for the $\Delta\eta(\text{lepton}, \text{lepton})$ by:

$$\text{weight} = 1 + y \cdot \frac{1.2 - i}{2.5} = 1 + y \cdot X, \tag{53}$$

and for the $\Delta\phi(\text{lepton}, \text{lepton})$ by:

$$\text{weight} = 1 + y \cdot \frac{1.75 - i}{3.14} = 1 + y \cdot X, \tag{54}$$

where y = -1, 1, and i is the bin centre. The results of the second stress test are shown in the same Figures 82 and 83. The reweighted particle level distributions are in different shapes from the nominal ones, and the unfolding procedure is able to retrieve the reweighted particle level distributions in all observables in both channels.

Not reviewed, for internal circulation only

(a)

(b)

(c)

(d)

(e)

Figure 82: The stress test for the five observables in the dilepton channel. Both the dots and lines are ratios made with respect to the nominal particle level. The dots are the ratio of the unfolded reweighted distributions to the nominal particle level distribution, while the solid lines are the ratio of the reweighted particle level distributions to the nominal one. Obs. is defined in Equation 49, and X in Equations 50 – 54.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 83: The stress test for the photon p_{T} (top), the photon η (middle), $\Delta R(\text{lepton}, \gamma)$ (bottom), in the single-lepton channel. Both the dots and lines are ratios made with respect to the nominal particle level. The dots are the ratio of the unfolded reweighted distributions to the nominal particle level distribution, while the solid lines are the ratio of the reweighted particle level distributions to the nominal one. Obs. is defined in Equation 49, and X in Equations 50 – 52.

Not reviewed, for internal circulation only

9.2 Cross-check methods

9.2.1 Bin-by-bin correction method

The method is based on extracting a bin-by-bin correction factor from MC. This method is only favoured with a very low bin-to-bin migration. The correction factor is defined by the ratio of the reconstructed MC distribution to the particle-level distribution. The unfolded spectrum N_k^{unfold} is given by the following equation:

$$N_k^{\text{unfold}} = (N_k^{\text{data}} - N_k^{\text{bkgs}}) \cdot (\frac{N_k^{\text{gen}}}{N_k^{\text{reco}}}) \tag{55}$$

where k is simultaneously the binning index at particle and reconstruction levels, N_k^{gen} is the generated event in bin k at particle level in the fiducial region, and N_k^{reco} is the selected event in bin k at reconstrucion level in the signal region.

9.2.2 The singular value decomposition method

The singular value decomposition method described in [52], simplifies the inversion process of the migration matrix by decomposing the migration matrix of other matrices:

$$M = US^{-1}V^T\,, \tag{56}$$

where U and V are nxn orthogonal matrices and S is nxn diagonal matrix. Then:

$$M^{-1} = VS^{-1}U^T\,. \tag{57}$$

The method also introduces a normalization term to regularize the solution. The term acts as cut-off for the large sensitivity to the small fluctuations. Since the method is used here only for a cross-check, the regularization parameters is set here to the default value, which is the number of bins divided by 2; No optimisation work is done.

9.2.3 Comparison test

The test is done here to compare different methods by unfolding the reconstruction-level distribution, which is estimated from MC. Figure 84 and Figure 85 show the different unfolded results with different methods, compared to the truth distribution in the dilepton and single lepton channels, respectively. The SVD method, show always slightly different central values from the other methods, but agrees within the statistical uncertainties. Here, the SVD method should be taken as first approximation since its regularisation parameter is not optimised.

Not reviewed, for internal circulation only

(a)

(b)

(c)

(d)

(e)

Figure 84: The number of unfolding events in the dilepton channel, obtained by the iterative Bayesian method after 3 iterations in black, the SVD in red and the bin-by-bin in yellow. All are compared to the truth distribution in blue (the ratio plot), and normalized to a luminosity of 36.1 fb^{-1}.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 85: The number of unfolding events as a function of the photon p_{T} (a), the photon η (b), and the $\Delta R(\text{lepton}, \gamma)$ (c) in the single lepton channel, obtained by the iterative Bayesian method after 3 iterations in black, the SVD in red and the bin-by-bin in yellow. All are compared to the truth distribution in blue (the ratio plot), and normalized to a luminosity of 36.1 fb^{-1}.

Not reviewed, for internal circulation only

10 Systematic uncertainties

Various sources of systematic uncertainties are considered for the $t\bar{t}\gamma$ cross section measurement. A discussion on how systematics can be smoothed, symmetrised and pruned is presented in Section 10.1. The modelling uncertainties of the signal and backgrounds are studied in Section 10.2. The experimental uncertainties, which are common to both signal and backgrounds, are explained in Section 10.3. Table 31 details all the systematics considered for this analysis in the pre-pruning stage. In addition to the plots shown in this section, further systematic uncertainty plots are shown in Appendices Q.5.2, Q.5.3 and Q.5.4 for JER, Lep fake and $Z\gamma$ studies.

10.1 Smoothing, symmetrisation and pruning

Smoothing and *symmetrisation* are methods in which statistical fluctuations in various systematic sources are avoided. Furthermore, the systematic is then centred around a mean value. The first step is symmetrisation. *Two-sided* symmetrisation is performed when an up and a down variation is provided for any given systematic. The difference between the two variations is calculated then divided by the mean of the variations. This value is then taken as positive (for up) and negative (for down). This is described by Equation 58.

$$\text{Variation up/down} = \left| \frac{\text{up} - \text{down}}{(\text{up} + \text{down})/2} \right| \tag{58}$$

Thus, the $\pm 1\sigma$ variation is centred around the nominal value. It is important to notice that if the up/down variations are symmetric then applying this procedure has no effect. *One-sided* symmetrisation is when only an up or down variation is provided. An example is the PPT systematics. In this case the variation is simply mirrored to reflect the supplementary variation.

Smoothing algorithms average statistics across bins. This prevents large statistical spikes in many of the systematics that are expected to give small contributions. In certain cases such as for signal, $t\bar{t}$ and $Z\gamma$ modelling, this option is turned off to allow for maximum shape variations.[13] There are two smoothing algorithms used.

- The TRExFitter algorithm which is based on two parameters: the *tolerance* and the threshold for number of slope variations in the systematic histogram. The first step is for the algorithm to ensure that the statistical uncertainty for a group of bins is less than an 8% tolerance. If this isn't the case the histogram is re-binned. The number of slope variations of the histogram is then checked. If the number of variations in slope is less than or equal to the provided threshold of four bins this new histogram is kept. If this isn't the case, then the statistical tolerance is halved, the histogram re-binned, and the slope variation checked once again. This procedure is repeated until the slope variations is four or less. This process only affects the shape of the systematic histogram with the overall normalisation kept fixed to the original histogram.

[13] For the dilepton channels, the $t\bar{t}$ modelling is smoothed due to very small background contributions and large statistical fluctuations.

Not reviewed, for internal circulation only

- The second algorithm is more conservative and makes use of ROOT's TH1::Smooth method. Based on the integral of the histogram and neighbouring bin information it averages bin contents. This method is used when the TRExFitter algorithm over smooths systematic uncertainties in which there are many derivative sign changes as well as higher statistical uncertainties for certain processes. The cases where this has been applied are for *Jet eta intercalibration, jet energy resolution, jet flavour response, jet pileup, pileup* and *MET objects* for all channels.

As a test, all symmetrisation for the systematics explained below is switched off to see the impact on the sensitivity of the measurement. The results can be seen in Appendix M. They show that with symmetrisation turned off the results don't dramatically change.

A summary of which systematics are symmetrised and smoothed is shown in Figures 86 and 87 for the single lepton and dilepton channels[14].

Pruning is the final step in which systematics that are too small to contribute to the final fit result are removed. This increases the speed of the fit and also grooms various plots and tables in order to make them more human-readable. Four pruning scenarios were tested for each of the five channels. Asimov fits were performed with pruning values set at 0.1%, 0.5%, 0.7% and 1.0%. In each case the error in μ is presented for the up and down variation. These results are summarised in Table 29. An overall trend of a lower uncertainty towards higher pruning values can be seen. However, the largest difference between any two values does not exceed 3%. Thus, a conservative value of 0.7% is chosen.

A non-negligible contribution from the JES Pileup (RhoTopology) nuisance parameter is shown as an example of smoothing (making use of the TH1:Smooth algorithm) and symmetrisation in Figure 88. Data-driven backgrounds will receive only a shape contribution, while MC receives shape and normalisation.

Table 29: Different pruning values for Asimov fits and the results for the error of μ for each channel. These tables do not include the hadronic fake normalisation and shape uncertainties (which are expected to be large) or the MC generator weights (which are expected to be very small).

Pruning value [%]	e+jets	μ+jets	ee	$e\mu$	$\mu\mu$
0.1	+0.1139 -0.1037	+0.1141 -0.1051	+0.1215 -0.1126	+0.0730 -0.0690	+0.1119 -0.1030
0.5	+0.1125 -0.1022	+0.1129 -0.1039	+0.1200 -0.1114	+0.0716 -0.0678	+0.1102 -0.1016
0.7	+0.1113 -0.1009	+0.1118 -0.1027	+0.1200 -0.1115	+0.0712 -0.0674	+0.1100 -0.1015
1.0	+0.1108 -0.1005	+0.1123 -0.1030	+0.1201 -0.1115	+0.0711 -0.0674	+0.1100 -0.1015

[14] The respective smoothing and symmetrisation also applies to the individual channels in single lepton and dilepton.

Not reviewed, for internal circulation only

Figure 86: Symmetrisation and smoothing methods applied to the systematic uncertainties for the single lepton, e+jets and μ+jets channels.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 87: Symmetrisation and smoothing methods applied to the systematic uncertainties for the dilepton, ee, $\mu\mu$ and $e\mu$ channels.

Not reviewed, for internal circulation only

(a) single lepton

(b) dilepton

Figure 88: Systematic plots shown for the JES Pileup (RhoTopology) contribution for the single lepton and dilepton channel. The dotted line represents the systematic before smoothing and symmetrisation has been applied. The solid line represents the after affect and the nuisance parameter that will enter the fit.

Not reviewed, for internal circulation only

10.2 Theoretical Uncertainties

10.2.1 Signal Modeling

The following signal modelling systematics are split into shape and normalisation effects and enter the fit as separate (but correlated) systematics. The reason for this is to distinguish between systematics of the overall correction factors (as defined in Section 7.1.3) and those associated with the ELD shape. The systematics of the correction factors are presented in Section 7. Specifically, the results are provided in Table 20. Here we focus on the shape systematics.

A caveat is that the normalisation systematics (from the correction factors) do not enter the differential fit. This is because in differential measurements, we unfold the input at detector level using dedicated unfolding algorithm, thus we do not need to use the correction factors to bridge the detector level quantities to those at particle level. With the exception of $t\bar{t}\gamma$ initial and final state radiation, these systematics are not symmetrised unless they only have one of the up/down variations.

- Renormalisation and Factorisation scales (denoted as "Scales"): The effect of the choice of the renormalisation (μ_r) and factorisation (μ_f) scales is estimated by varying μ_r and μ_f simultaneously or independently up and down by a factor of 2 with respect to the nominal sample value, thus resulting in 3 sets of up-and-down variations. The variation is done in terms of event weights to reduce the effect of statistical fluctuations. The final uncertainty on the normalisation is the quadratic sum of the 3 sets of variations shown in Figure 89. For the shape, the maximum variation is chosen to enter the fit and is shown in Figure 90.

- Uncertainty due to the parton shower and hadronisation (denoted as "Parton Shower"): The uncertainty is estimated by comparing the $t\bar{t}\gamma$ nominal samples produced using MadGraph + Pythia8, with Herwig7 showering the same MadGraph events. The outcome of the shape affecting the nominal signal sample is shown in Figure 91.

- Initial and final state radiation uncertainty (denoted as "ISR/FSR"): The uncertainty is estimated by comparing the $t\bar{t}\gamma$ samples produced using the same MadGraph events as for the nominal signal MadGraph + Pythia8 sample, but using Pythia8 tunes with high or low QCD radiation activity (A14 var3c eigentune). These effects are shown in Figure 92. Under scrutiny was whether to symmetrise this systematic contribution or not. A study was done and is shown in Appendix Q.5.1 where fits are performed with and without this contribution being symmetrised. The results show little difference between the two except for slightly more symmetrical errors. Thus, it was determined to symmetrise this contribution.

- PDF uncertainty: The PDF uncertainty is evaluated by the envelope error of the 100 PDF error sets in the NNPDF23LO, which are stored as weights in the nominal sample. The shape effect is calculated by performing 100 fits (a fit for each variation) and computing the standard deviation with the respect to the central value of μ. This is technically inconvenient to implement in the final fit, and so separate studies were done to see the final impact on μ. Should this shape effect be negligible they can be left out of the final fit. The errors relative to μ for each channel are shown in Table 30. It's clear that the impact on μ is negligible, bearing in mind that from Table 29 the final pruning value is 0.7%. Thus, the 100 PDF shape weights are ignored in the fit. Plots for each channel can be seen in Appendix J.

Not reviewed, for internal circulation only

Figure 89: The quadratic sum of the three variations of the renormalisation and factorisation scales applied to the single lepton and dilepton ELD. Only normalisation is shown.

Figure 90: The largest variation for the renormalisation and factorisation scale systematics (μ_f) applied to the single lepton and dilepton ELD. Only the impact of shape is shown. The solid lines represent the nuisance parameter that will enter the fit after the removal of normalisation.

Table 30: The relative uncertainty on μ for the the 100 PDF shape uncertainty fits.

Channel	Relative error on μ [%]
e+jets	0.066
μ+jets	0.066
ee	0.137
$\mu\mu$	0.109
$e\mu$	0.027
single lepton	0.066
dilepton	0.200

Not reviewed, for internal circulation only

(a) The shape impact of the parton shower variation (HERWIG7) on the nominal signal sample.

(b) The overall impact (shape+acceptance) of the parton shower variation (HERWIG7) on the nominal signal sample.

Figure 91: The shape and overall impact of the parton shower variation (HERWIG7) on the nominal signal sample. The solid lines represent the nuisance parameter that will enter the fit after *onesided* symmetrisation and the removal of normalisation.

Not reviewed, for internal circulation only

(a) The shape impact of the QCD radiation tunes (ISR/FSR) on the nominal signal sample.

(b) The overall impact (shape+acceptance) of the QCD radiation tunes (ISR/FSR) on the nominal signal sample.

Figure 92: The shape and overall impact of the QCD radiation tunes (ISR/FSR) on the nominal signal sample. The solid lines represent the nuisance parameter that will enter the fit after the removal of normalisation.

Not reviewed, for internal circulation only

10.2.2 Background Modelling

In addition to those uncertainties detailed in Section 6.1 and Section 6.2 for the $e \rightarrow \gamma$ fake and hadronic fake backgrounds, the $t\bar{t}$ MC modelling systematics which affect the shape for these two backgrounds are included and applied to the respective background events. These are:

- The uncertainty on hard scattering generator and parton shower and hadronisation: This is estimated by comparing the $t\bar{t}$ nominal sample produced by PowhegBox+Pythia8 with the $t\bar{t}$ sample produced by Sherpa. The outcome of the shape effect on $e \rightarrow \gamma$ fake and hadronic fake backgrounds for the single lepton channel are shown in Figure 93. The dilepton contribution is negligible. These contributions are *onesided* symmetrised where needed (where only one up/down variation exists), otherwise they are not symmetrised.
- The uncertainty on initial and final state radiation: This is estimated by comparing the nominal $t\bar{t}$ sample, with two samples produced with different Pythia8 tunes with higher or lower QCD radiation activity (A14 var3c eigentune), modified factorisation and renormalisation scales (at $\times 0.5$ and $\times 2$) and modified NLO radiation with *hdamp* = 1.5 and 3 times the mass of the top quark. The outcome of the shape and overall effect on $e \rightarrow \gamma$ fake and hadronic fake backgrounds for single lepton are shown in Figure 94. This also includes the ISR/FSR modelling systematic derived for η and p_{T} bins in Section 6.2. The dilepton contribution is negligible. These contributions are not symmetrised.

The $W\gamma$ background is floated in the single lepton fit while the $Z\gamma$ shape is modelled with MG5_aMC@NLO and showered with Pythia8. The $Z\gamma$ background also has a separate 50% normalisation uncertainty correlated to the shape. This is due to the large differences between the systematic variation sample and the large generator weights in the nominal sample. Additionally, the $W\gamma$ and $Z\gamma$ backgrounds receive QCD renormalisation and factorisation scale variation systematics. The largest contribution is chosen (μ_r for both cases). Since the $W\gamma$ background is a free-floating parameter, it is assigned only shape. The $Z\gamma$ scale variation receives both shape and normalisation.

Other backgrounds with a prompt γ are very small, thus assigned a conservative systematic of 50%. As a cross-check, this uncertainty is doubled and leads to negligible changes. See Appendix K for more information.

The estimation on the Lep fake background systematics can be seen in Section 6.3.

Not reviewed, for internal circulation only

(a) $t\bar{t}$ shape systematics for the modelling of the parton shower for hadronic fake backgrounds.

(b) $t\bar{t}$ shape systematics for the modelling of the parton shower for $e \rightarrow \gamma$ fake backgrounds.

(c) hadronic fake systematic for the modelling of the parton shower derived in Section 6.2.

Figure 93: The components of the $t\bar{t}$ parton shower modelling for the single lepton channel. Since the $t\bar{t}$ samples are embedded in the hadronic fake and $e \rightarrow \gamma$ fake background groups, a systematic needs to be defined for the $t\bar{t}$ part of each, which are then correlated in the fit.

Not reviewed, for internal circulation only

(a) $t\bar{t}$ shape systematics for the modelling of extra radiation for hadronic fake backgrounds.

(b) $t\bar{t}$ shape systematics for the modelling of extra radiation for $e \to \gamma$ fake backgrounds.

(c) hadronic fake systematic for the modelling of extra radiation derived in Section 6.2.

Figure 94: The components of the $t\bar{t}$ ISR/FSR modelling for the single lepton channel. Since the $t\bar{t}$ samples are embedded in the hadronic fake and $e \to \gamma$ fake background groups, a systematic needs to be defined for the $t\bar{t}$ part of each, which are then correlated in the fit.

Not reviewed, for internal circulation only

10.3 Experimental Uncertainties

Experimental uncertainties common to signal and background processes are discussed in this section. These are mainly the uncertainties of the reconstructed physical objects used in the analysis, such as the lepton, photon, jet, and $E_{\mathrm{T}}^{\mathrm{miss}}$ in terms of their reconstruction and identification efficiencies, momentum and energy scales as well as momentum and energy resolutions. In addition, the uncertainties of the jet flavour tagging, the jet vertex fraction cut and the integrated luminosity as well as the pile-up simulation are also described in this section. In the following, the sources of experimental uncertainties are introduced.

10.3.1 Sources of experimental uncertainties

Leptons

- **Lepton Efficiency -** The lepton efficiency refers to the reconstruction and identification efficiencies, isolation efficiency as well as the trigger efficiency. For Monte Carlo (MC) samples, these efficiencies are corrected by applying scale factors, which are calculated as the ratio between the efficiencies measured in data and those in simulation using tag-and-probe method on $Z \to ee$ and $J/\Psi \to ee$ ($Z \to \mu\mu$) decays for electrons (muons). These scale factors, which are E_{T} and η dependent, are varied up and down by one standard deviation to study the impact of lepton efficiency uncertainties on the analysis result [29][32].
- **Lepton energy (momentum) scale and resolution -** The measured lepton energy (momentum) is calibrated using MC-based techniques. Correction factors, which are derived from the study of dileptonic decays of the Z-boson are applied to correct possible detector mis-modelings in the calibration. These correction factors, are varied up and down by one standard deviation to study the lepton energy (momentum) scale uncertainty. For electrons, the energy scale and resolution are calculated together with photons as EGamma scale and resolution.

Photons

- **Photon efficiency -** The photon identification efficiency is measured with samples of photons from the radiative decays of the Z-boson and electrons and positrons from Z decays exploiting the similarity between electron and photon electromagnetic (EM) showers. Scale factors are defined as the ratio between efficiencies measured in data and those in MC samples and are applied to MC samples to correct possible detector mis-modelings. The scale factors for photon isolation have been measured [33]. These scale factors are varied up and down by one standard deviation to study their impacts on the analysis.
- **Photon energy scale and resolution -** The photon energy calibration and the study of its scale and resolution uncertainties are calculated together with the electrons as EGamma scale and resolution.
- **Prompt Photon Tagger (PPT) -** The strategy of deriving systematic uncertainties of the PPT is described in Section 4.2.2. The 6 types of systematics as applied to the ELD are shown in Figure 95.

Jets

Not reviewed, for internal circulation only

(a) Prompt systematic applied to $t\bar{t}\gamma$ signal, and $W\gamma$ and Other prompt backgrounds.

(b) Hadronic fake systematic.

(c) $e \rightarrow \gamma$ fake systematic.

(d) Hadronic fake isolation systematic.

Figure 95: PPT shape systematics for the single lepton channel.

- **Jet Energy Scale -** The jet energy calibration, or jet energy scale (JES) used in this analysis follows the globally reduced parameter configurations of 21NPs [54, 55]. The calibration is done in several steps, combining Monte Carlo simulation and *in situ* measurements, to correct for topo-cluster energy, pile-up effect, etc. Its uncertainty is split into several independent categories: modelling and statistical uncertainties on the extrapolation of the jet calibration from the central region, jet flavour compostion(see Appendix Y for gluon fraction estimation), high-p_{T} jet behaviour, b-jet energy scale uncertainties, uncertainties due to pile-up, uncertainties on *in situ* jet energy corrections, etc. In one category, there are usually more than one physical source of the uncertainty. To study the JES uncertainty, each source is varied up and down independently by its corresponding uncertainty.

- **Jet Energy Resolution -** The jet energy resolution (JER) is measured using the balance between jets and well measured objects like photons or Z bosons, and it is found to be in agreement between data and MC. There is a single source of uncertainty for the JER, which is varied by one sigma to study its impacts on the analysis.

- **Jet Vertex Tagging -** The systematic uncertainty associated to the jet vertex tagging(JVT) is obtained by varying up and down the JVT cut using the *JetVertexTaggerTool*[37]

b-tagging

Jets coming from a b-quark is called a b-jet. It has its own topological feature, e.g. long lived b-hadrons, and can be distinguished from the other lighter flavour jets, with a procedure called b-tagging. The b-tagging in this analysis is performed with the MV2c10 algorithm for few different working points. Each jet can be assigned to a different working point (a procedure called "pseudo-continuous b-tagging"), and the b-tagging uncertainties on this jet are derived for this specific working point. Those uncertainties are accounted for by varying the calibration scale factors provided by Flavour Tagging group for b-,c-, and

Not reviewed, for internal circulation only

light-flavour jets up and down by their corresponding systematic uncertainties independently. For each jet category, the uncertainties are decomposed into several uncorrelated components using eigenvector method, for example, there are 30, 15 and 80 eigenvectors for *b*-,*c*-, and light-flavour jets uncertainties, respectively [56].

Missing Transverse Energy

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed[39] from the vector sum of several terms corresponding to different types of reconstructed objects. The estimated uncertainties for electrons, muons, photons and jets are propagated into the uncertainty of $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus, the only new contribution is the systematic uncertainty of the soft terms $E_{\mathrm{x,y}}^{\mathrm{RefSoftJet}}$ and $E_{\mathrm{x,y}}^{\mathrm{CellOut}}$.

- **Soft term scale and resolution -** The systematic uncertainty of the soft-term scale is estimated by comparing the ratio of MC simulation to data. The average deviation of the ratio from unity is taken as a flat uncertainty on the absolute scale. The systematic uncertainty of the soft-term resolution is estimated by evaluating the level of agreement between data and MC in the $E_{\mathrm{x}}^{\mathrm{miss}}$ and $E_{\mathrm{y}}^{\mathrm{miss}}$ resolution. Both the scale and resolution of the soft term are varied up and down by one standard deviation to study their impact on the analysis.

Pile Up

The systematics due to pile-up has been studied by variation of the μ rescaling from the nominal value of 1.09 to 1.0(up) or 1.18(down) [57]. Due to the larger $V\gamma$ weights it's prudent to smooth this systematic otherwise constraints and pulls are seen in the single lepton channels. Figure 96 and 97 shows the pileup NP for each of the backgrounds for the single lepton and dilepton channels, respectively. We can see that smoothing this systematic is reasonable, especially for the Other prompt background process which has a large spike in the third bin of the single lepton distribution. This spike is caused by the large $V\gamma$ weights. More information can be seen in Appendix V.

Luminosity

The total integrated luminosity has an uncertainty of 2.1%. It is derived, following a methodology similar to that detailed in [18], from a calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016.

Not reviewed, for internal circulation only

Figure 96: The pileup systematic contribution for the single lepton channel split into signal and background contributions. The dotted line represents the systematic before smoothing and symmetrisation has been applied. The solid line represents the after affect and the nuisance parameter that will enter the fit.

Not reviewed, for internal circulation only

Figure 97: The pileup systematic contribution for the dilepton channel split into signal and background contributions. The dotted line represents the systematic before smoothing and symmetrisation has been applied. The solid line represents the after affect and the nuisance parameter that will enter the fit.

Not reviewed, for internal circulation only

Table 31: Complete list of systematics considered before pruning is applied. A “N" indicates that only normalisation (or acceptance) has been considered as a systematic. A “S" indicates that only the shape is considered as a systematic. “SN" indicates that both shape and normalisation have been considered as a systematic. For PDF systematics, see Appendix J on why the negligible shape contribution is not included.

Systematic name	Type	Components	Notes
Miscellaneous			
Luminosity	N	1	
Pileup reweighting	SN	1	
Background modelling			
$W\gamma$ (single lepton)	free floating	1	
$W\gamma$ (dilepton)	N	1	50%
$Z\gamma$ (single lepton)	N	1	50%
VV	N	1	50%
ST	N	1	50%
$t\bar{t}V$	N	1	50%
Lep fake	SN	1	
$e \to \gamma$ fake	SN	11	
hadronic fake d.d.	SN	25	SL hadronic fake $t\bar{t}$ modelling from θ_{MC} is anti-correlated with the POI when fitting, Equation 14
DL $hadronicfake$ ISR/FSR extrap.	N	1	
DL $hadronicfake$ Scales extrap.	N	1	
ISR/FSR	SN*, S**	1	(hadronic fake)*, ($e \to \gamma$ fake)**
Parton Shower	SN*, S**	1	(hadronic fake)*, ($e \to \gamma$ fake, $Z\gamma$)**
Scales	S	1	hadronic fake, $e \to \gamma$ fake, $W\gamma$, $Z\gamma$
Prompt Photon Tagger			
Prompt photons	S	3	Split into $t\bar{t}\gamma$, $W\gamma$ and Other prompt contributions and correlated in fit
$e \to \gamma$ fake	S	1	
hadronic fake	S	1	
hadronic fake isolation	S	1	
Object reconstruction			
Electrons (trigger, reco, ID, isolation)	SN	5	
Egamma (resolution, scale)	SN	2	
Muons (trigger, reco, ID, isolation)	SN	15	
MET (resolution, scale)	SN	3	
Photons (efficiency, isolation)	SN	3	
Jet energy scale	SN	21	
Jet energy resolution	SN	1	
Jet vertex tagger	SN	1	
b-tagging efficiency	SN	30	
c-tagging efficiency	SN	15	
Light-jet tagging efficiency	SN	80	
b-tagging extrapolation	SN	1	
Signal modelling			
ISR/FSR	SN	1	
Parton Shower	SN	1	
Scales	SN	1	
PDF	N	1	

Not reviewed, for internal circulation only

10.4 Propagation of the systematic uncertainties to the unfolding framework

The same systematic uncertainties described above are propagated to the unfolding results. Prefit systematic variations are used for the study. In the single lepton channel, the $W\gamma$ background is floated so that a stat-only fit without any systematics considered is performed to derive its normalization as well as uncertainty.

In practise, nominal signal+backgrounds histogram is varied. For systematics related only to backgrounds, the nominal histogram is shifted by the amount of shift on the backgrounds, then unfolded and compared with the nominal unfolded distribution. For experimental systematics, the nominal histogram is shifted by the amount of shift on both the signal and backgrounds, then unfolded and compared with the nominal unfolded distribution. In the above cases, the unfolding core, which includes the migration matrix, efficiency and migration, are kept to be nominal. However, for systematics related to signal modelling, the unfolding core is varied, while the input is kept to be nominal.

Besides, the MC statistical uncertainty of the signal is considered. Each of the bin of the migration matrix is fluctuated while keeping the integral of the matrix constant. The magnitude of fluctuation corresponds to the Poisson fluctuation of the raw event of that bin. 200 toys are produced in this way for each bin of the migration matrix. Then the nominal input is used and unfolded with these 200 migration matrices. The standard deviation of the 200 unfolded distributions are assigned as the MC statistical uncertainty of a certain bin of the migration matrix. The total MC statistical uncertainty is a quadratic sum of the MC statistical uncertainty of all the bins.

The total uncertainty is the quadrature sum of all these variations, for each bin of the unfolded distribution. The covariance matrix for each systematic variation is implemented in a simple way:

$$C_{ij} = \sigma_i \times \sigma_j \tag{59}$$

where σ_i and σ_j are the uncertainty for bin i and bin j of the distribution. For comparison, the data statistical covariance matrix is directly taken from the output of the unfolding algorithm.

Not reviewed, for internal circulation only

11 Pre-fit distributions with scale factors and systematics applied

This section shows the pre-fit plots used for training the ELD (Table 10) with all systematics applied as discussed in Section 10. All respective scale factors have also been applied. Table 32 gives the yields for the signal, backgrounds and data.

Figure 98 shows this for the single lepton channel and Figure 99 for the dilepton channel.

	single lepton	dilepton
$t\bar{t}\gamma$	6488 ± 420	720 ± 34
hadronic fake	1435 ± 290	49 ± 27
$e \to \gamma$ fake	1653 ± 170	2 ± 1
Lep fake	356 ± 200	-
$W\gamma$	1126 ± 45	-
$Z\gamma$	-	75 ± 52
Other prompt	691 ± 260	18 ± 7
Total	11747 ± 710	863 ± 78
Data	11662	902

Table 32: Pre-fit yields for signal and backgrounds for the single lepton and dilepton channels. All scale factors and systematic uncertainties are included.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 98: Pre-fit distributions used for training the single lepton ELD. All scale factors and systematic uncertainties included.

Not reviewed, for internal circulation only

Figure 99: Pre-fit distributions used for training the dilepton ELD. All scale factors and systematic uncertainties included.

Not reviewed, for internal circulation only

12 Theory prediction

Next-to-leading order (NLO) QCD corrections to the production of a $t\bar{t}$ pair in association with a hard photon are available by K. Melnikov et al. [14], extending the results of Duan et. al. [58] performed in the approximation of stable top quarks. The calculation of [14] allows for decays of top quarks and it includes photon radiation in both the production and decay stages of the $t\bar{t}$ pair. The method of generalised D-dimensional unitarity extended to massive particles and the dipole formalism are used respectively to calculate one-loop virtual amplitudes and real emission corrections. Top quarks are treated in the narrow width approximation with all spin correlations retained. The hadronic decays of W-bosons are considered into two families of light quarks, always treated as massless. The W-bosons are considered in their mass-shells and no QCD radiative corrections to the hadronic decays are considered. The strong coupling constant is evaluated using one- and two-loop running with five massless flavours.

While the results presented in [14] are calculated at the centre-of-mass energy of $\sqrt{s}$ = 14 TeV, a dedicated calculation at $\sqrt{s}$ = 13 TeV has been performed in the single lepton and dilepton channels, in a phase space very close to the fiducial region of the measurement, as described in Section 7.1. The NNPDF set was used, and a fine structure constant of $\alpha_{QED} = 1/137$ is used. The renormalisation and factorisation scales $\mu_R = \mu_f = \mu = m_t$ are used.

A comparison of the leading order (LO) jet inclusive cross-section calculated by MadGraph and by the above theoretical framework is performed in Section 12.1. The LO cross-section of MadGraph is corrected to NLO precision in a phase space which is quite close to the fiducial region defined in Section 7.1. The reason to define a fiducial kfactor instead of an inclusive kfactor is that we find in the dilepton channel, the inclusive kfactor gives a very bad data/MC agreement, which implies there is a phase space dependent of the kfactor. This phase space as well as the theoretical computation of its NLO cross-section are described in Section 12.2. The corresponding LO prediction by MadGraph in the same phase space is described in Section 12.3. The resulting NLO kfactor in this phase space as well as the final predicted cross-section in the fiducial region are presented in Section 12.4.

12.1 LO jet inclusive cross-section

As a cross-check, the MadGraph calculated LO jet inclusive cross-section is compared with that of the theoretical computation.

The jet inclusive phase space is defined by the following cuts. At least one lepton is required to have $p_{\mathrm{T}} > 15$ GeV. The photon is required to have $p_{\mathrm{T}} > 15$ GeV, $\mid \eta \mid < 5$. Jets are required to have $p_{\mathrm{T}} > 10$ GeV and $\mid \eta \mid < 5$. The photon has to satisfy $\Delta R(\gamma, l) > 0.5$ and $\Delta R(\gamma, \mathrm{jet}) > 0.2$, otherwise the event is rejected. This jet inclusive phase space differs from the total phase space used by MadGraph only in terms of the cut on $\Delta R(\gamma, l)$.

There is a difference which should be corrected for making a fair comparison. The values of the electroweak parameters, like the top and W boson width, used by MadGraph and the theoretical computation are different. This difference is corrected by adjusting the MadGraph parameter values to be the same as the theoretical computation. In this way, the total cross-section reported in Section 3.1 changes from 4.623 pb to 3.356pb.

The jet inclusive cross-section calculated by MadGraph is compared with the one by theoretical calculation at different scales, a fixed scale of $2m_t$ and a dynamic scale defined as the the sum of the transverse

Not reviewed, for internal circulation only

Table 33: The comparison of LO total cross-section calculated by MadGraph (after correcting for some electroweak parameter values) and that of the theory calculation in the single lepton and dilepton channels at fixed or dynamic scales.

Scale		Single lepton	Dilepton
$\mu = 2m_t$	Theory	1170.4	227.7
	MadGraph	1170.9	232.3
Dynamic scale	Theory	1398.4	276.02
	MadGraph	1396.8	279.9

mass of all the final state particles, and in different channels, the single lepton channel and the dilepton channel. The results are summarised in Table 33, where a good agreement can be seen between the two calculations.

12.2 NLO cross-section

The NLO calculation is performed in a phase space very close to the fiducial region of the measurement. It can be done in the exact fiducial region. But due to a mis-communiation with the theorist, a small difference is introduced between the phase space where the NLO calculation is performed and the fiducial phase space. However, it is found that the resulting NLO kfactor is not sensitive to this difference.

This phase space is defined with the following cuts. At least one lepton is required to have $p_{\mathrm{T}} > 25$ GeVand $\mid \eta \mid < 2.5$ The photon is required to have $p_{\mathrm{T}} > 20$ GeV, $\mid \eta \mid < 2.37$, Jets are reconstructed from stable particles using the anti-k_T algorithm ($R = 0.4$), and are considered if they have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. For single lepton (dilepton) channel, it is required to have at least four (two) jets. There are several overlap removal requirements: jets are required to be separated from photon with $\Delta R(\gamma, \mathrm{jet}) > 0.4$; leptons are required to be separated from photon with $\Delta R(\gamma, \mathrm{lepton}) > 1.0$; if it's dilepton channel, leptons are required to be separated from jets with $\Delta R(\mathrm{jet}, \mathrm{lepton}) > 0.4$.

The calculated NLO cross sections are 112 fb and 28.3 fb for the single lepton and dilepton channels, respectively. These numbers are then scaled up by 7% to take into account the difference of top and W widths between those used for theory calculation and those from ATLAS recommendation.

12.3 LO cross-section

The NLO calculation is performed at parton level. In principle, the LO cross-section should also be done at parton level. But since there is jet multiplicity cut involved in the definition of the phase space, using a parton level LO cross-section will over-estimate the kfactor. More specifically, the NLO parton calculation has added an additional jet to the LO parton calculation and the LO MC particle level calculation has added $\geq$1 additional jets to the LO parton calculation. So if the kfactor is defined by NLO(parton)/LO(parton) and applied to the LO MC sample, the 1st hardest additional emission will be corrected for twice, thus over-estimating the results.

To use particle level sample to calculate the LO cross-section, one has to carefully define the objects to mimic those used in the NLO calculation for defining the phase space. The leptons and photon can be

Not reviewed, for internal circulation only

easily taken as their matrix element correspondents, but now with slighly changed kinematics due to the recoil caused effect caused by the additional QCD radiation from the parton shower. Besides, the leptons have all their PS-added QED radiations reabsorbed since the NLO calculation considers only 1 QED radiation. Jets are defined using the same anti-k_T algorithm of cone size 0.4 and use as input all the final state particles, except for the above photon, leptons and their corresponding lepton neutrinos.

After defing the objects as above, the same phase space cuts as in the previous section are applied. The resulting LO cross-sections are 66.8 fb and 15.4 fb for the single lepton and dilepton channels, respectively.

12.4 Kfactor and fiducial cross-section

The NLO kfactor is defined as:

$$k = \frac{\sigma_{NLO}^{parton}}{\sigma_{LO}^{particle}}. \tag{60}$$

σ_{NLO}^{parton} is the parton level NLO cross-section from theory calculation described in Section 12.2. and $\sigma_{LO}^{particle}$ is the particle level LO cross-section from MadGraph described in Section 12.3. The resulting kfactors are 1.80 and 1.97 for the single lepton and dilepton channels, respectively.

From the equation, it can be seen that theoretical uncertainties could enter via both of its numerator and denominator.

To estimate the uncertainty of σ_{NLO}^{parton}, the renormalization and factorization scales are varied up-and-down by a factor of 2 simultaneously, resulting in an uncertainty of 13%. The PDF is also varied from NNPDF to MMHT2014 or CT14, resulting in an uncertainty of 5% (2%) for the single lepton (dilepton) channel.

To estimate the uncertainty of $\sigma_{LO}^{particle}$, it's mainly the non-perturbative effect and the arbitrary jet cone size that should be considered. The non-perturbative effect brought by the parton shower includes the multiple parton interaction (MPI) and the hadronization of the partons. The MPI systematics is derived by tuning the parton shower to have rather low MPI effect. The hadronization systematics is evaluated by turning off the hadronization of the parton shower. The jet cone size is varied from 0.4 to 0.3 and 0.5 as up and down variation. The results of these systematics are summarized in Table 34, together with the systematics of the σ_{NLO}^{parton}. The final systematics for the kfactors are also given.

Channel	Kfactor	$\delta\sigma_{LO}^{particle}$				$\delta\sigma_{NLO}^{parton}$	δKfactor
		Jet cone	MPI	Hadron.	Sum	Scale/Pdf	
single lepton	1.80	-11%/+4%	-7%	+4%	14%	14%	20%
dilepton	1.97	-6%/+3%	-2%	+3%	7%	13%	15%

Table 34: The systematics of the NLO kfactor, decomposed into systematics of σ_{NLO}^{parton} and that of $\sigma_{LO}^{particle}$.

Using these kfactor and their uncertainties, the fiducial cross-sections can be calculated by $\sigma_{LO} \times A_i \times k_i$, where σ_{LO} is the total cross-section by MadGraph, A_i is the accpetance calculated in Section 7.1.3, and k_i is the kfactor calculated as above, of channel i. The results are summarized in Table 35.

Not reviewed, for internal circulation only

Table 35: The theoretical fiducial NLO cross sections as well as their uncertainties in single lepton and dilepton channels taking into account the branching ratio for each channel.

	e+jets	μ+jets	ee	$e\mu$	$\mu\mu$	single lepton	dilepton
Xsec. (fb)	247.14 ± 49.43	247.95 ± 49.59	15.88 ± 2.38	31.37 ± 4.71	15.93 ± 2.39	495.10 ± 99.02	63.18 ± 9.48

Not reviewed, for internal circulation only

13 Results

13.1 Fit model and results

The unblinded fit results for the analysis are presented in this section for the single lepton, dilepton and inclusive combined channels[15]. The best μ is presented along with post-fit plots, yields and nuisance parameter rankings relative to μ. In the case of the single lepton channel the $W\gamma$ background is an extra floating parameter in the fit. Asimov fit results shown in Table 91 justify the choice to use the ELD in all channels and Appendix I justifies the choice of bin sizes.

13.1.1 μ, best-fit

Figure 100 shows the observed $t\bar{t}\gamma$ signal strength with the total uncertainty. The total uncertainty is further split into a statistical and systematic contribution. μ is shown for each channel, for single lepton and dilepton and for a 5 channel inclusive fit. The vertical dotted line represents the standard model prediction of 1. The shaded region represents the theoretical uncertainty.

The observed best fit value for the single lepton channel is $\mu^{\mathrm{SL}}_{t\bar{t}\gamma} = 1.053\ ^{+0.083}_{-0.079}(\mathrm{sys.})\ ^{+0.019}_{-0.019}(\mathrm{stat.})$ while the observed best fit value for the dilepton channel is $\mu^{\mathrm{DL}}_{t\bar{t}\gamma} = 1.085\ ^{+0.063}_{-0.059}(\mathrm{sys.})\ ^{+0.042}_{-0.041}(\mathrm{stat.})$. The observed 5 channel inclusive fit has a best fit value of $\mu^{\mathrm{Incl.}}_{t\bar{t}\gamma} = 1.056\ ^{+0.060}_{-0.057}(\mathrm{sys.})\ ^{+0.018}_{-0.017}(\mathrm{stat.})$. All measurements agree with theory within uncertainties.

In the case of the single lepton channel, $W\gamma\ SF = 0.796 \pm 0.335$ (tot.), while for the inclusive combined fit $W\gamma\ SF = 0.856 \pm 0.255$ (tot.).

[15] Pre- and post-fit plots are not shown for the inclusive combined fit since these would be redundant to what is shown in the Appendix Q.3 for each channel.

Not reviewed, for internal circulation only

Figure 100: The observed signal strength for individual as well as combined channels. The standard model prediction of 1 is represented by the vertical line. The theoretical uncertainty for each fit is represented by the shaded region.

13.1.2 Post-fit plots and yields

Figure 101 and 102 show the pre-fit and post-fit distributions for the single lepton channel, while Figure 103 and 104 show pre-fit and post-fit distributions for the dilepton channel, respectively. The ELD shown is used as the discriminating variable in the maximum likelihood fit.

Figure 105 and 106 shows the fit results from the 5 channel inclusive fit applied to the single lepton and dilepton channels, respectively. The difference with the previous plots being that uncertainties are smaller.

Individual channel distributions can be seen in Appendix Q.3. More post-fit variables (used to train the single lepton and dilepton ELD) can be seen in Appendix Q.2.

Final yields are shown in Table 36.

Not reviewed, for internal circulation only

(a) Pre-fit ELD

(b) Pre-fit $|\eta(\gamma)|$

(c) Pre-fit $p_T(\gamma)$

(d) Pre-fit $\Delta R(\text{lepton}, \gamma)$

Figure 101: Pre-fit plots for the single lepton channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Post-fit ELD

(b) Post-fit $|\eta(\gamma)|$

(c) Post-fit $p_T(\gamma)$

(d) Post-fit ΔR(lepton, γ)

Figure 102: Post-fit plots for the single lepton channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Pre-fit ELD (b) Pre-fit $|\eta(\gamma)|$ (c) Pre-fit $p_T(\gamma)$

(d) Pre-fit $\Delta\phi$(lepton, lepton) (e) Pre-fit $\Delta\eta$(lepton, lepton) (f) Pre-fit min ΔR(lepton, γ)

Figure 103: Pre-fit plots for the dilepton channel where this ELD distribution is used as the discriminating variable in the fit.

	single lepton	dilepton
$t\bar{t}\gamma$	7044 ± 350	782 ± 44
hadronic fake	1473 ± 180	49 ± 26
$e \to \gamma$ fake	1621 ± 160	2 ± 1
Lep fake	186 ± 68	-
$W\gamma$	896 ± 370	-
$Z\gamma$	-	55 ± 29
Other prompt	569 ± 180	18 ± 7
Total	11787 ± 180	906 ± 38
Data	11662	902

Table 36: Post-fit yields for each background for the single lepton and dilepton channels. All uncertainties are included.

Not reviewed, for internal circulation only

(a) Post-fit ELD (b) Post-fit $|\eta(\gamma)|$ (c) Post-fit $p_T(\gamma)$

(d) Post-fit $\Delta\phi$(lepton, lepton) (e) Post-fit $\Delta\eta$(lepton, lepton) (f) Post-fit min ΔR(lepton, γ)

Figure 104: Post-fit plots for the dilepton channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Post-fit ELD

(b) Post-fit $|\eta(\gamma)|$

(c) Post-fit $p_T(\gamma)$

(d) Post-fit ΔR(lepton, γ)

Figure 105: Post-fit plots for the 5 channel inclusive fit applied to the single lepton channel.

Not reviewed, for internal circulation only

(a) Post-fit ELD (b) Post-fit $|\eta(\gamma)|$ (c) Post-fit $p_T(\gamma)$

(d) Post-fit $\Delta\phi$(lepton, lepton) (e) Post-fit $\Delta\eta$(lepton, lepton) (f) Post-fit min ΔR(lepton, γ)

Figure 106: Post-fit plots for the 5 channel inclusive fit applied to the dilepton channel.

Not reviewed, for internal circulation only

13.1.3 Nuisance parameters

The relative uncertainties in relation to the best μ for combined, single lepton and dilepton channels are shown in Table 37. To create this table, individual systematics are grouped into categories, for example "Jets", "Background modelling" and "$E\gamma$". All nuisance parameters in the group are kept fixed while a fit is performed with the rest of the nuisance parameters floating. This resulting uncertainty is then subtracted in quadrature from the total uncertainty to obtain $\Delta\mu$. The greatest contribution comes from jet objects, which can be further broken down into individual subcategories. This is shown in Table 38. Background modelling is further broken down into components from hadronic fake, $e \rightarrow \gamma$ fake, $t\bar{t}$ and Other prompt contributions. This is shown in Table 39. The PPT shape systematic uncertainties are broken down into their separate contributions, as is shown in Table 40.

The total contributions relative to μ are also summarised in Figure 107. The JES contribution includes *Jet flavour composition*, *Jet flavour response*, *Jet eta intercalibration*, *Jet pileup* (with the largest contribution coming from *RhoTopology*), *Jet effective NP*, *Jet single particle HightPt*, *Jet Punchthrough MC15* and *Jet BJES response*.

Individual channel nuisance parameters can be seen in Appendix Q.3.

	Inclusive		Single lepton		Dilepton	
	+ [%]	- [%]	+ [%]	- [%]	+ [%]	- [%]
Signal Modelling	2.50	2.36	1.59	1.47	2.87	2.75
Jets	3.14	2.83	5.40	4.80	2.02	1.89
Luminosity	2.28	2.06	2.32	2.01	2.29	2.04
Pileup	1.99	1.86	2.04	1.80	2.32	2.12
Photon Efficiencies	1.06	0.97	1.08	0.95	1.07	0.96
b-Tagging	0.41	0.40	0.76	0.86	0.36	0.40
Background modelling	2.81	2.74	4.80	4.80	2.91	2.86
Leptons	0.96	0.88	0.26	0.27	1.30	1.20
Prompt photon tagger (shape)	1.45	1.45	3.80	4.00	-	-
Eγ	0.06	0.08	0.02	0.03	0.16	0.20
Template Statistics	1.50	1.38	1.89	1.78	1.65	1.56
Total systematic	5.8	5.5	7.9	7.6	5.8	5.4
Total statistical	1.4	1.4	1.5	1.5	3.8	3.8
Total	6.0	5.7	8.1	7.7	7.0	6.6

Table 37: Relative difference to μ due to the up and down variation of the systematic sources. Fits are performed with nuisance parameters in each group held constant with the rest floating. This new uncertainty is subtracted in quadrature from the total uncertainty to obtain $\Delta\mu$

	Inclusive		Single lepton		Dilepton	
	+ [%]	- [%]	+ [%]	- [%]	+ [%]	- [%]
Jet Pileup (RhoTopology)	2.31	2.10	3.80	3.38	1.66	1.52
Jet Flavour Composition	0.05	0.05	0.13	0.14	0.10	0.10
Jet EtaInterCalibration	0.13	0.12	1.10	1.01	0.21	0.18
Jet Effective NP	1.79	1.61	2.91	2.53	0.97	0.89
Jet Flavour Resp.	0.41	0.39	1.94	1.79	0.02	0.00

Table 38: Breakdown of the larger jet systematic components for showing the total contribution to the relative difference in μ.

Not reviewed, for internal circulation only

	Inclusive		Single lepton		Dilepton	
	+ [%]	- [%]	+ [%]	- [%]	+ [%]	- [%]
$e \to \gamma$ fake modelling	1.23	1.19	2.51	2.54	0.05	0.05
$t\bar{t}$ modelling	1.10	1.06	2.02	2.05	0.98	1.13
hadronic fake modelling	1.93	1.87	1.41	1.42	2.13	2.17
Other prompt bkg modelling	1.57	1.55	3.55	3.70	1.26	1.25

Table 39: Breakdown of the background modelling components showing the total contribution to the relative difference in μ.

	Single lepton	
	+ [%]	- [%]
Prompt sources	3.17	3.30
$e \to \gamma$ fake sources	1.74	1.92
hadronic fake sources	0.60	0.60
hadronic fake isolation	0.25	0.23

Table 40: Breakdown of the PPT shape systematics for single lepton channel showing the total contribution to the relative difference in μ.

Figure 107: Select grouped uncertainties for the inclusive, single lepton and dilepton channels relative to the signal strength.

Figure 108 shows the pruning map for which different systematics get pruned before the fit for the single lepton and dilepton channels. As mentioned in Chapter 10 a value of 0.7% is used. Importantly, no extra systematic uncertainties are pruned for the inclusive combined fit. The results from the individual channel fits (shown in Appendix Q.3) are used along with their already pruned workspaces.

Ranking plots with finer granularity showing the largest 40 contributions for the single lepton and dilepton

Not reviewed, for internal circulation only

channels are shown in Figure 109. Figure 110 shows the ranking for the inclusive combined fit. These plots include the statistical uncertainties of each bin which enter the fit as uncorrelated nuisance parameters (the γ's).

The full nuisance parameter pull plots for the single lepton and dilepton channels are shown in Figure 111. Figure 112 shows this for the inclusive fit.

The pulled and constrained nuisance parameters are further discussed.

- $t\bar{t}\gamma$ parton shower: The shape and overall effect can be seen in Figure 91. A pull and constraint occurs in the single lepton (and marginally in the dilepton) channel simply because the most signal rich bins form a ELD distribution shape that favours the HERWIG7 sample.
- $t\bar{t}$ parton shower (ttbar_sherpa): This systematic is constrained in the single lepton channel. The different contributions and overall impact of these systematics are shown in Figure 93. In general, the data slightly prefers the shape of the SHERPA variation samples.
- $t\bar{t}$ ISR/FSR: This systematic is constrained in the single lepton channel. The different contributions and overall impact of these systematics are shown in Figure 94. These systematics are not symmetrised. Large shape differences to the nominal samples (often above or below 20% per bin) ensures that this systematic is constrained in the fit.
- Jet Energy Resolution: The single lepton channels suffer from a constrained JER. By looking at Figure 235 the culprit is found to be the third bin from the left in which a larger data/MC disagreement is observed. Although the constraint is in the negative direction, this is due to a sign flip of the up/down variation. Appendix Q.5.2 demonstrates this by masking the guilty bin.
- Lepfake: The Lepfake systematic uncertainty is constrained due to the data preferring a different parameterisation. The up/down variations for the Lep fake background are from the extremes of the parameterisation “envelope" used in the data driven method. Figure 237 shows the down variation (or parameterisation) being preferred.
- $Z\gamma$ modelling: The dilepton channel pull plot shows constrained $Z\gamma$ systematic uncertainties. This is in part due to larger generator weights as explored in Appendix V, specifically in the first and second bins of the dilepton channels, and the third bin of the single lepton channels (in the Other prompt category). For the dilepton channels this is also due to large differences between the nominal and systematic variations as can be seen in Figure 238.

Finally, Figure 113 and 114 show the correlations between the nuisance parameters and the signal strength, μ, for given thresholds above 15%. For the single lepton channel, large correlations between PPT prompt, Jet Pileup Rhotopology and μ are unsurprising given their very conservative nature and thus their impact on the measurement. These are large NP's that predominantly affect the signal. For the dilepton there are no alarming or surprisingly large correlations.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 108: Pruning plots showing the contributions of each systematic that survive the pruning process for the single lepton and dilepton channels. Grey means the systematic does not apply. Green means both shape and normalisation contributions are kept. Yellow means the shape contribution has been dropped but the normalisation has been kept, while orange is the reverse. Red indicates the whole systematic was pruned.

Not reviewed, for internal circulation only

(a) single lepton (b) dilepton

Figure 109: Systematic uncertainty ranking plots for the largest 40 contributions in the single lepton and dilepton channels.

Not reviewed, for internal circulation only

Figure 110: Systematic uncertainty ranking plots for the largest 40 contributions in the inclusive fit.

Not reviewed, for internal circulation only

(a) single lepton

(b) dilepton

Figure 111: Pull plots for all nuisance parameters in the single lepton and dilepton channels after a 0.7% pruning has been applied.

Not reviewed, for internal circulation only

Figure 112: Pull plots for all nuisance parameters in the inclusive fit. Pruning has already occurred when fitting the individual channels, for which the workspaces are used in the inclusive fit.

Not reviewed, for internal circulation only

Figure 113: Correlation matrix plot for the single lepton channel showing the correlation between each nuisance parameter and the signal strength for values greater than 15%.

Figure 114: Correlation matrix plot for the dilepton channel showing the correlation between each nuisance parameter and the signal strength for values greater than 15%.

Not reviewed, for internal circulation only

13.2 Fiducial cross section

From the theoretical calculations from Chapter 12 and the results in Section 13.1, we obtain $\sigma_{\text{fid}}^{\text{SL}} = 521.42\,^{+40.93}_{-39.20}(\text{sys.})\,^{+9.21}_{-9.19}(\text{stat.})$ fb and $\sigma_{\text{fid}}^{\text{DL}} = 68.57\,^{+3.95}_{-3.76}(\text{sys.})\,^{+2.64}_{-2.58}(\text{stat.})$ fb for the single lepton and dilepton channels, respectively, and $\sigma_{\text{fid}}^{\text{Incl.}} = 589.43\,^{+33.77}_{-32.05}(\text{sys.})\,^{+9.87}_{-9.73}(\text{stat.})$ fb for the combined inclusive fit to all channels. The cross sections are summarised in Table 41. All measurements agree with theoretical predictions within uncertainty.

	σ_{fid} [fb]	Systematic uncertainty		Statistical uncertainty	
		+ [fb]	- [fb]	+ [fb]	- [fb]
e+jets	264.80	20.80	19.69	6.36	6.33
μ+jets	250.16	22.31	20.98	6.63	6.77
ee	15.94	1.63	1.45	1.52	1.39
$\mu\mu$	17.62	1.52	1.35	1.39	1.32
$e\mu$	34.30	1.93	1.81	1.73	1.67
single lepton	521.42	40.93	39.20	9.21	9.19
dilepton	68.57	3.95	3.76	2.64	2.58
Inclusive (5 channels)	589.43	33.77	32.05	9.87	9.73

Table 41: Fiducial cross section summary for all channels.

Not reviewed, for internal circulation only

13.3 Differential cross section

The normalised measured differential cross sections are shown in Figures 115 and 116 in the single-lepton and dilepton channels, respectively. The measured distributions are obtained from data after subtracting the pref-fit background distributions. The unfolded results are compared to the particle-level distribution obtained from the nominal MC signal sample. The impact of the statistical uncertainty on the measured cross section is dominant in the dilepton channel, but nevertheless the systematic uncertainty has also a sizeable impact. For the single-lepton channel, it is the other way around where systematic uncertainty is dominant. The leading systematic uncertainties in the single lepton and dilepton channels are due to the background modelling, the jet identification and reconstruction, the pile-up, and the signal modeling which gets large in some bins due to the particular choice of the generator for parton showering.

The decomposed uncertainties of the unfolded distributions are illustrated in Figure 117 and Figure 118 for the single-lepton and dilepton channels respectively.

The numerical values of the normalised differential cross-sections results are summarised in Tables 42 – 44 for the single-lepton channel while for the dilepton channel they are summarised in Tables 45 – 49. The effect of all systematics on the differential cross section are shown in Tables 50 – 52 in the single-lepton channel and Tables 53 – 57 in the dilepton channel.

To quantify the agreement between the theory and the unfolded distributions, χ^2's and their corresponding p-values are calculated using the total covariance matrix. The formula for χ^2 is:

$$\chi^2 = (\sigma^{\text{diff,norm}}_{j,\text{unfold}} - \sigma^{\text{diff,norm}}_{j,\text{pred}}) \cdot C^{-1}_{jk} \cdot (\sigma^{\text{diff,norm}}_{k,\text{unfold}} - \sigma^{\text{diff,norm}}_{k,\text{pred}}) \tag{61}$$

where C^{-1} is the inverse of the covariance matrix, and $(\sigma^{\text{diff,norm}}_{j,\text{unfold}} - \sigma^{\text{diff,norm}}_{j,\text{pred}})$ is the difference between the unfolded and predicted distributions. For the normalized distribution, the last bin is not included in the above formula to take into account the fact that there is one additional redundant degree of freedom. The χ^2 tests for the normalised cross-sections compared with a number of MC predictions are summarised in Table 58 for the single lepton channel while for the dilepton channel they are summarized in Table 59. In general, data to MC agreements are observed, which can be seen from the χ^2 tables, except that the $\Delta\phi(\text{lepton}, \text{lepton})$ is not well modelled by any of the signal MC.

The correlation matrices in the single-lepton channel can be seen in Tables 60 – 62 while for the dilepton channel they are summarised in Tables 63 – 67.

A closure test is performed where the un-normalised differential cross-sections are integrated. This results in $\sigma_{\text{SL}} = 527 \pm 65$ fb and $\sigma_{\text{DL}} = 70 \pm 9$ fb. Uncertainties are the total uncertainty. Central values are comparable to fitted fiducial measurements, but uncertainties are much larger due to using pre-fit input for differential measurements. The results from the integration of each observable is slightly different and so the average is taken.

The statistical correlations are calcualted using the bootstrab method in the single-lepton and dilepton channels which can be seen in Figures 119 and 120, respectively.

Not reviewed, for internal circulation only

(a) (b)

(c)

Figure 115: The normalised differential cross section in (a) p_{T} bins, (b) $|\eta|$ bins and (c) min $\Delta R(\text{lepton}, \gamma)$ (bottom), in the single lepton channel. The unfolded distribution to the particle-level (dots) is compared to the particle-level distributions of nominal A14 (red), and variations up (cyan) and down (green) and also compared to the replacement of the Pythia8 with Herwig7 (blue).

Not reviewed, for internal circulation only

(a) (b) (c) (d) (e)

Figure 116: The normalised differential cross section in (a) p_{T} bins , (b) $|\eta|$ bins, (c) minimum $\Delta R(\text{lepton}, \gamma)$, (d) $\Delta\eta(\text{lepton}, \text{lepton})$, and (e) $\Delta\phi(\text{lepton}, \text{lepton})$ in the dilepton channel. The unfolded distribution to the particle-level (dots) is compared to the particle-level distributions of nominal A14 (red), and variations up (cyan) and down (green) and also compared to the replacement of the PYTHIA8 with HERWIG7 (blue).

Not reviewed, for internal circulation only

(a) (b)

(c)

Figure 117: The decomposed systematic uncertainties for the normalised differential cross-sections as a function of the (a) photon p_{T}, (b) photon $|\eta|$ and (c) $\Delta R(\gamma, \ell)$ in the single lepton channel.

Not reviewed, for internal circulation only

(a) (b) (c) (d) (e)

Figure 118: The decomposed systematic uncertainties for the normalised 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.

Not reviewed, for internal circulation only

Table 42: The normalised differential cross section in $p_{\mathrm{T}}(\gamma)$ bins, obtained using the iterative Bayesian method, in the single lepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

$p_{\mathrm{T}}(\gamma)$ [Gev]	normalised cross section
20.0 - 35.0	0.0285 ± 0.0013 ± 0.0019
35.0 - 50.0	0.0137 ± 0.0008 ± 0.001
50.0 - 65.0	0.0072 ± 0.0006 ± 0.0005
65.0 - 80.0	0.0044 ± 0.0004 ± 0.0008
80.0 - 95.0	0.0029 ± 0.0003 ± 0.0005
95.0 - 110.0	0.0022 ± 0.0003 ± 0.0003
110.0 - 140.0	0.0013 ± 0.0001 ± 0.0003
140.0 - 180.0	0.0009 ± 0.0001 ± 0.0001
180.0 - 300.0	0.0003 ± 0.0 ± 0.0

Table 43: The normalised differential cross section in $|\eta|(\gamma)$ bins, obtained using the iterative Bayesian method, in the single lepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

$\|\eta\|(\gamma)$	normalised cross section
0.0 - 0.2	0.64 ± 0.04 ± 0.08
0.2 - 0.4	0.67 ± 0.04 ± 0.08
0.4 - 0.6	0.62 ± 0.04 ± 0.07
0.6 - 0.8	0.54 ± 0.05 ± 0.04
0.8 - 1.0	0.53 ± 0.04 ± 0.04
1.0 - 1.2	0.47 ± 0.04 ± 0.06
1.2 - 1.7	0.36 ± 0.02 ± 0.04
1.7 - 2.37	0.18 ± 0.02 ± 0.04

Table 44: The normalised differential cross section in min $\Delta R(\gamma, \mathrm{lep})$ bins, obtained using the iterative Bayesian method, in the single lepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

min $\Delta R(\gamma, \mathrm{lep})$	normalised cross section
1.0 - 1.2	0.395 ± 0.034 ± 0.03
1.2 - 1.4	0.373 ± 0.032 ± 0.015
1.4 - 1.6	0.403 ± 0.035 ± 0.033
1.6 - 1.8	0.346 ± 0.035 ± 0.039
1.8 - 2.0	0.331 ± 0.032 ± 0.052
2.0 - 2.2	0.412 ± 0.036 ± 0.024
2.2 - 2.4	0.363 ± 0.052 ± 0.021
2.4 - 2.6	0.474 ± 0.038 ± 0.042
2.6 - 6.0	0.112 ± 0.004 ± 0.005

Not reviewed, for internal circulation only

Table 45: The normalised differential cross section in $p_{\mathrm{T}}(\gamma)$ bins, obtained using the iterative Bayesian method, in the dilepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

$p_{\mathrm{T}}(\gamma)$ [Gev]	normalised cross section
20.0 - 35.0	0.0285 ± 0.0022 ± 0.0011
35.0 - 50.0	0.0127 ± 0.0018 ± 0.0012
50.0 - 65.0	0.0068 ± 0.0013 ± 0.0006
65.0 - 80.0	0.0047 ± 0.0008 ± 0.0007
80.0 - 95.0	0.0026 ± 0.0006 ± 0.0006
95.0 - 110.0	0.0025 ± 0.0005 ± 0.0005
110.0 - 140.0	0.0012 ± 0.0003 ± 0.0001
140.0 - 180.0	0.0009 ± 0.0002 ± 0.0001
180.0 - 300.0	0.0005 ± 0.0001 ± 0.0001

Table 46: The normalised differential cross section in $|\eta|(\gamma)$ bins, obtained using the iterative Bayesian method, in the dilepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

$\|\eta\|(\gamma)$	normalised cross section
0.0 - 0.2	0.54 ± 0.08 ± 0.03
0.2 - 0.4	0.74 ± 0.09 ± 0.07
0.4 - 0.6	0.53 ± 0.09 ± 0.05
0.6 - 0.8	0.66 ± 0.08 ± 0.04
0.8 - 1.0	0.44 ± 0.1 ± 0.06
1.0 - 1.2	0.36 ± 0.06 ± 0.02
1.2 - 1.7	0.32 ± 0.05 ± 0.04
1.7 - 2.37	0.27 ± 0.03 ± 0.05

Table 47: The normalised differential cross section in min $\Delta R(\gamma, \mathrm{lep})$ bins, obtained using the iterative Bayesian method, in the dilepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

min $\Delta R(\gamma, \mathrm{lep})$	normalised cross section
1.0 - 1.2	0.871 ± 0.105 ± 0.09
1.2 - 1.4	0.758 ± 0.103 ± 0.067
1.4 - 1.6	0.513 ± 0.108 ± 0.085
1.6 - 1.8	0.7 ± 0.104 ± 0.058
1.8 - 2.0	0.504 ± 0.088 ± 0.036
2.0 - 2.2	0.537 ± 0.09 ± 0.073
2.2 - 2.4	0.377 ± 0.081 ± 0.038
2.4 - 2.6	0.152 ± 0.058 ± 0.024
2.6 - 6.0	0.035 ± 0.009 ± 0.011

Table 48: The normalised differential cross section in $\Delta\eta(\mathrm{lep}, \mathrm{lep})$ bins, obtained using the iterative Bayesian method, in the dilepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

$\Delta\eta$(lep, lep)	normalised cross section
0.0 - 0.2	0.47 ± 0.09 ± 0.06
0.2 - 0.4	0.35 ± 0.12 ± 0.1
0.4 - 0.6	0.41 ± 0.09 ± 0.05
0.6 - 0.8	0.52 ± 0.09 ± 0.07
0.8 - 1.0	0.43 ± 0.08 ± 0.04
1.0 - 1.2	0.53 ± 0.09 ± 0.07
1.2 - 1.7	0.34 ± 0.05 ± 0.05
1.7 - 2.5	0.36 ± 0.03 ± 0.03

Not reviewed, for internal circulation only

Table 49: The normalised differential cross section in $\Delta\phi(\text{lep}, \text{lep})$ bins, obtained using the iterative Bayesian method, in the dilepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

$\Delta\phi$(lep, lep)	normalised cross section
0.0 - 0.35	0.19 ± 0.04 ± 0.05
0.35 - 0.7	0.17 ± 0.05 ± 0.04
0.7 - 1.05	0.38 ± 0.08 ± 0.08
1.05 - 1.4	0.23 ± 0.05 ± 0.04
1.4 - 1.75	0.45 ± 0.06 ± 0.05
1.75 - 2.1	0.37 ± 0.06 ± 0.04
2.1 - 2.45	0.36 ± 0.06 ± 0.04
2.45 - 2.8	0.36 ± 0.05 ± 0.02
2.8 - 3.14	0.36 ± 0.06 ± 0.03

Not reviewed, for internal circulation only

Table 50: Summary of the sources of uncertainty on the normalised differential cross section for $p_{\mathrm{T}}(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_{\mathrm{T}}(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
Signal modelling	±0.63	±1.07	±0.52	±4.52	±0.7	±9.75	±3.96	±5.07	±1.9
Exp. systematics	±1.48	±1.2	±1.35	±2.13	±2.9	±2.57	±2.94	±3.22	±2.92
$t\bar{t}$ modelling	±5.6	±6.33	±5.15	±16.14	±14.09	±10.07	±16.64	±12.65	±6.65
Other bkg. modelling	±3.97	±3.79	±3.27	±7.95	±7.35	±6.84	±5.23	±5.99	±3.85
Total systematics	±7	±8	±6	±18	±16	±15	±24	±16	±9
Data statistics	±4	±6	±8	±10	±12	±13	±11	±10	±8
Total uncertainty	±8	±9	±10	±21	±20	±20	±27	±18	±12

Table 51: Summary of the sources of uncertainty on the normalised differential cross section for $|\eta|(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$\|\eta\|(\gamma)$	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.37
Source	systematic uncertainty (%)							
Signal modelling	±2.1	±1.83	±2.46	±1.19	±2.07	±4.39	±0.83	±2.84
Exp. systematics	±1.78	±1.93	±1.59	±2.38	±1.52	±1.34	±1.15	±2.86
$t\bar{t}$ modelling	±11.24	±10.47	±10.76	±5.4	±7.05	±10.86	±8.94	±18.96
Other bkg. modelling	±4.08	±4.43	±3.8	±4.12	±3.36	±3.86	±3.84	±10.03
Total systematics	±12	±12	±12	±7	±8	±12	±10	±22
Data statistics	±6	±6	±7	±10	±7	±8	±7	±9
Total uncertainty	±14	±13	±13	±12	±11	±15	±12	±23

Table 52: Summary of the sources of uncertainty on the normalised differential cross section for min $\Delta R(\gamma, \mathrm{lep})$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

min $\Delta R(\gamma, \mathrm{lep})$	1.0 - 1.2	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.2	2.2 - 2.4	2.4 - 2.6	2.6 - 6.0
Source	systematic uncertainty (%)								
Signal modelling	±1.0	±2.74	±2.2	±1.89	±1.49	±3.49	±0.95	±6.0	±1.19
Exp. systematics	±1.75	±1.56	±1.85	±1.8	±1.72	±1.77	±2.73	±1.52	±1.12
$t\bar{t}$ modelling	±6.83	±2.08	±7.1	±10.56	±15.53	±3.86	±1.38	±6.06	±2.97
Other bkg. modelling	±2.8	±1.53	±2.77	±3.23	±2.15	±1.83	±5.04	±1.74	±2.65
Total systematics	±8	±4	±8	±11	±16	±6	±6	±9	±4
Data statistics	±9	±9	±9	±10	±10	±9	±14	±8	±4
Total uncertainty	±12	±10	±12	±15	±19	±10	±16	±12	±6

Not reviewed, for internal circulation only

Table 53: Summary of the sources of uncertainty on the normalised differential cross section for $p_{\mathrm{T}}(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_{\mathrm{T}}(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
Signal modelling	±0.56	±5.73	±5.93	±7.48	±11.99	±15.67	±5.66	±7.55	±5.54
Exp. systematics	±2.79	±4.11	±6.28	±3.18	±2.87	±2.29	±3.64	±2.81	±2.66
$Z\gamma$ modelling	±1.71	±12.18	±5.71	±3.68	±3.68	±1.81	±4.14	±3.35	±2.82
Other bkg. modelling	±1.61	±0.82	±2.13	±0.96	±2.98	±1.58	±3.4	±1.12	±1.67
Total systematics	±4	±9	±9	±16	±22	±20	±9	±12	±15
Data statistics	±8	±14	±19	±18	±22	±22	±20	±21	±14
Total uncertainty	±9	±17	±21	±23	±31	±30	±22	±24	±21

Table 54: Summary of the sources of uncertainty on the normalised differential cross section for $|\eta|(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$\lvert\eta\rvert(\gamma)$	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.37
Source	systematic uncertainty (%)							
Signal modelling	±2.44	±6.37	±1.73	±3.13	±9.29	±4.93	±5.38	±8.28
Exp. systematics	±3.01	±4.26	±6.53	±2.98	±6.97	±3.27	±8.48	±3.37
$Z\gamma$ modelling	±4.65	±3.02	±11.27	±6.06	±4.12	±1.64	±5.72	±4.11
Other bkg. modelling	±1.75	±3.82	±2.41	±1.27	±1.11	±0.96	±2.58	±5.89
Total systematics	±5	±10	±9	±5	±13	±7	±11	±18
Data statistics	±14	±12	±18	±13	±22	±18	±17	±11
Total uncertainty	±15	±16	±20	±14	±26	±19	±20	±21

Table 55: Summary of the sources of uncertainty on the normalised differential cross section for min $\Delta R(\gamma, \mathrm{lep})$ at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

min $\Delta R(\gamma, \mathrm{lep})$	1.0 - 1.2	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.2	2.2 - 2.4	2.4 - 2.6	2.6 - 6.0
Source	systematic uncertainty (%)								
Signal modelling	±2.17	±4.87	±8.8	±3.07	±3.39	±13.14	±3.22	±5.03	±13.66
Exp. systematics	±3.4	±2.84	±12.49	±3.71	±4.11	±2.48	±2.74	±4.31	±17.69
$Z\gamma$ modelling	±6.05	±3.84	±13.35	±2.97	±9.4	±5.46	±1.49	±2.25	±24.39
Other bkg. modelling	±3.21	±1.46	±3.17	±2.93	±0.51	±1.14	±4.17	±7.64	±9.63
Total systematics	±10	±9	±16	±8	±7	±14	±10	±16	±32
Data statistics	±12	±14	±21	±15	±17	±17	±21	±40	±25
Total uncertainty	±16	±16	±27	±17	±19	±22	±24	±40	±40

Table 56: Summary of the sources of uncertainty on the normalised differential cross section for $\Delta\eta(\mathrm{lep}, \mathrm{lep})$ at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$\Delta\eta(\mathrm{lep}, \mathrm{lep})$	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.5
Source	systematic uncertainty (%)							
Signal modelling	±6.29	±8.34	±10.65	±5.68	±5.49	±5.65	±10.7	±4.26
Exp. systematics	±5.46	±18.47	±4.89	±4.14	±4.68	±4.77	±7.66	±2.88
$Z\gamma$ modelling	±0.97	±21.03	±6.06	±1.46	±2.69	±3.39	±16.25	±6.52
Other bkg. modelling	±3.15	±2.37	±3.77	±1.76	±1.08	±1.31	±1.19	±2.16
Total systematics	±12	±28	±13	±14	±9	±14	±13	±9
Data statistics	±19	±34	±22	±17	±18	±16	±16	±9
Total uncertainty	±22	±40	±26	±22	±20	±21	±21	±13

Not reviewed, for internal circulation only

Table 57: Summary of the sources of uncertainty on the normalised differential cross section for $\Delta\phi$(lepton, lepton) at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$\Delta\phi$(lepton, lepton)	0.0 - 0.35	0.35 - 0.7	0.7 - 1.05	1.05 - 1.4	1.4 - 1.75	1.75 - 2.1	2.1 - 2.45	2.45 - 2.8	2.8 - 3.14
Source	systematic uncertainty (%)								
Signal Modelling	±10.08	±16.19	±12.06	±3.1	±3.3	±5.59	±5.71	±1.79	±8.0
Ttbar Modelling	±0.67	±0.44	±0.64	±0.64	±1.87	±0.59	±0.22	±0.29	±0.04
Hfake dedicated	±1.4	±0.89	±1.35	±1.35	±3.81	±1.24	±0.46	±0.61	±0.08
Efake dedicated	±0.04	±0.17	±0.5	±0.18	±0.17	±0.25	±0.1	±0.17	±0.17
Other Background Modelling	±6.56	±4.57	±2.7	±5.92	±2.11	±1.26	±2.23	±1.52	±3.31
Luminosity	±0.18	±0.31	±0.24	±0.05	±0.09	±0.07	±0.05	±0.04	±0.07
Pileup Effects	±1.84	±0.32	±8.52	±1.65	±2.1	±1.96	±1.88	±1.0	±0.22
Photon Scale Factors	±0.2	±0.14	±0.1	±0.04	±0.08	±0.06	±0.06	±0.03	±0.06
Lepton Identification and Reconstruction	±0.71	±0.99	±0.77	±0.25	±0.46	±0.39	±0.18	±0.2	±0.23
E/Gamma Resolution and Scale	±0.16	±0.18	±0.17	±0.01	±0.21	±0.07	±0.08	±0.04	±0.05
Jet Identification and Reconstruction	±3.62	±5.21	±9.26	±2.07	±1.85	±4.38	±1.36	±1.76	±1.2
b-tagging	±0.79	±2.07	±4.51	±1.1	±0.79	±0.61	±1.44	±0.6	±0.91
Missing Transverse Momentum Reconstruction	±1.5	±2.26	±2.0	±0.46	±0.81	±0.9	±0.45	±0.56	±0.22
Total systematics	±13	±18	±18	±7	±7	±8	±7	±3.3	±9
Data statistics	±25	±26	±33	±18	±20	±15	±15	±11	±11
Total uncertainty	±28	±32	±40	±19	±21	±17	±16	±12	±14

Not reviewed, for internal circulation only

Table 58: χ^2/n.d.f. values and p-values between the measured normalised differential cross-sections and predictions from several generators in the single lepton channel.

Predictions	$p_T(\gamma)$		$\lvert\eta(\gamma)\rvert$		$\Delta R(\gamma,\ell)$	
	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
MadGraph + Pythia8	3.2/8	0.92	0.7/7	1.0	5.0/8	0.76
MadGraph + Herwig7	2.3/8	0.97	0.9/7	1.0	4.8/8	0.78
MadGraph + Pythia8 (A14 Up)	3.3/8	0.91	0.8/7	1.0	4.9/8	0.77
MadGraph + Pythia8 (A14 Down)	2.6/8	0.96	0.9/7	1.0	4.6/8	0.8
POWHEG + Pythia8 $t\bar{t}$	25.4/8	<0.01	2.8/7	0.9	8.7/8	0.37

Table 59: χ^2/n.d.f. values and p-values between the measured normalised differential cross-sections and predictions from several generators in the dilepton channel.

Predictions	$p_T(\gamma)$		$\eta(\gamma)$		$\Delta R(\gamma,\ell)$		$\lvert\Delta\eta(\ell,\ell)\rvert$		$\Delta\phi(\ell,\ell)$	
	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
MadGraph + 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
MadGraph + 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
MadGraph + Pythia8 (A14 Up)	1.6/8	0.99	8.4/7	0.3	7.4/8	0.49	3.4/7	0.85	14.0/8	0.08
MadGraph + 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

Not reviewed, for internal circulation only

Table 60: The correlation matrix for the normalised differential cross-section as a function of photon p_{T} in the single lepton channel, accounting for the statistical and systematic uncertainties.

Bin GeV	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
20.0 - 35.0	1.00	0.32	0.16	0.47	0.34	0.24	0.47	0.19	0.18
35.0 - 50.0	0.32	1.00	0.40	0.60	0.51	0.41	0.59	0.43	0.42
50.0 - 65.0	0.16	0.40	1.00	0.43	0.36	0.28	0.40	0.41	0.39
65.0 - 80.0	0.47	0.60	0.43	1.00	0.67	0.57	0.75	0.44	0.50
80.0 - 95.0	0.34	0.51	0.36	0.67	1.00	0.47	0.69	0.32	0.41
95.0 - 110.0	0.24	0.41	0.28	0.57	0.47	1.00	0.53	0.34	0.36
110.0 - 140.0	0.47	0.59	0.40	0.75	0.69	0.53	1.00	0.40	0.50
140.0 - 180.0	0.19	0.43	0.41	0.44	0.32	0.34	0.40	1.00	0.49
180.0 - 300.0	0.18	0.42	0.39	0.50	0.41	0.36	0.50	0.49	1.00

Table 61: The correlation matrix for the normalised differential cross-section as a function of photon $|\eta|$ in the single lepton channel, accounting for the statistical and systematic uncertainties.

Bin	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.37
0.0 - 0.2	1.00	0.64	0.43	-0.03	0.29	-0.35	0.15	0.34
0.2 - 0.4	0.64	1.00	0.25	-0.13	0.16	-0.40	-0.05	0.15
0.4 - 0.6	0.43	0.25	1.00	0.25	0.56	0.25	0.59	0.71
0.6 - 0.8	-0.03	-0.13	0.25	1.00	0.26	0.36	0.34	0.36
0.8 - 1.0	0.29	0.16	0.56	0.26	1.00	0.25	0.50	0.59
1.0 - 1.2	-0.35	-0.40	0.25	0.36	0.25	1.00	0.42	0.36
1.2 - 1.7	0.15	-0.05	0.59	0.34	0.50	0.42	1.00	0.68
1.7 - 2.37	0.34	0.15	0.71	0.36	0.59	0.36	0.68	1.00

Table 62: The correlation matrix for the normalised differential cross-section as a function of $\Delta R(\gamma, \ell)$ in the single lepton channel, accounting for the statistical and systematic uncertainties.

Bin	1.0 - 1.2	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.2	2.2 - 2.4	2.4 - 2.6	2.6 - 6.0
1.0 - 1.2	1.00	0.06	0.27	0.24	0.45	0.16	-0.02	0.13	0.21
1.2 - 1.4	0.06	1.00	0.15	0.15	0.18	0.15	0.04	0.18	0.12
1.4 - 1.6	0.27	0.15	1.00	0.39	0.50	0.27	0.10	0.35	0.32
1.6 - 1.8	0.24	0.15	0.39	1.00	0.53	0.27	0.12	0.39	0.28
1.8 - 2.0	0.45	0.18	0.50	0.53	1.00	0.33	0.09	0.37	0.44
2.0 - 2.2	0.16	0.15	0.27	0.27	0.33	1.00	0.05	0.29	0.26
2.2 - 2.4	-0.02	0.04	0.10	0.12	0.09	0.05	1.00	0.08	0.15
2.4 - 2.6	0.13	0.18	0.35	0.39	0.37	0.29	0.08	1.00	0.25
2.6 - 6.0	0.21	0.12	0.32	0.28	0.44	0.26	0.15	0.25	1.00

Not reviewed, for internal circulation only

Table 63: The correlation matrix for the normalised differential cross-section as a function of photon p_{T} in the dilepton channel, accounting for the statistical and systematic uncertainties.

Bin GeV	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
20.0 - 35.0	1.00	0.00	0.04	0.03	0.02	-0.03	0.04	-0.01	0.00
35.0 - 50.0	0.00	1.00	0.29	0.49	0.54	0.38	0.26	0.34	0.44
50.0 - 65.0	0.04	0.29	1.00	0.28	0.36	0.25	0.21	0.26	0.26
65.0 - 80.0	0.03	0.49	0.28	1.00	0.50	0.35	0.21	0.33	0.41
80.0 - 95.0	0.02	0.54	0.36	0.50	1.00	0.47	0.30	0.39	0.50
95.0 - 110.0	-0.03	0.38	0.25	0.35	0.47	1.00	0.25	0.33	0.42
110.0 - 140.0	0.04	0.26	0.21	0.21	0.30	0.25	1.00	0.18	0.23
140.0 - 180.0	-0.01	0.34	0.26	0.33	0.39	0.33	0.18	1.00	0.34
180.0 - 300.0	0.00	0.44	0.26	0.41	0.50	0.42	0.23	0.34	1.00

Table 64: The correlation matrix for the normalised differential cross-section as a function of photon $|\eta|$ in the dilepton channel, accounting for the statistical and systematic uncertainties.

Bin	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.37
0.0 - 0.2	1.00	0.05	0.20	0.18	0.07	-0.01	-0.09	-0.02
0.2 - 0.4	0.05	1.00	0.13	0.16	0.22	0.16	0.12	0.20
0.4 - 0.6	0.20	0.13	1.00	0.24	0.22	0.15	0.14	0.27
0.6 - 0.8	0.18	0.16	0.24	1.00	0.16	0.13	0.10	0.22
0.8 - 1.0	0.07	0.22	0.22	0.16	1.00	0.13	0.16	0.28
1.0 - 1.2	-0.01	0.16	0.15	0.13	0.13	1.00	0.12	0.24
1.2 - 1.7	-0.09	0.12	0.14	0.10	0.16	0.12	1.00	0.37
1.7 - 2.37	-0.02	0.20	0.27	0.22	0.28	0.24	0.37	1.00

Table 65: The correlation matrix for the normalised differential cross-section as a function of $\Delta R(\gamma, \ell)$ in the dilepton channel, accounting for the statistical and systematic uncertainties.

Bin	1.0 - 1.2	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.2	2.2 - 2.4	2.4 - 2.6	2.6 - 6.0
1.0 - 1.2	1.00	0.24	-0.12	-0.01	-0.07	-0.04	-0.15	-0.16	0.11
1.2 - 1.4	0.24	1.00	-0.19	-0.04	-0.09	-0.01	-0.12	-0.12	0.06
1.4 - 1.6	-0.12	-0.19	1.00	0.22	0.28	0.23	0.17	0.22	0.24
1.6 - 1.8	-0.01	-0.04	0.22	1.00	0.15	0.14	0.07	0.12	0.16
1.8 - 2.0	-0.07	-0.09	0.28	0.15	1.00	0.16	0.13	0.19	0.23
2.0 - 2.2	-0.04	-0.01	0.23	0.14	0.16	1.00	0.11	0.18	0.22
2.2 - 2.4	-0.15	-0.12	0.17	0.07	0.13	0.11	1.00	0.17	0.13
2.4 - 2.6	-0.16	-0.12	0.22	0.12	0.19	0.18	0.17	1.00	0.19
2.6 - 6.0	0.11	0.06	0.24	0.16	0.23	0.22	0.13	0.19	1.00

Table 66: The correlation matrix for the normalised differential cross-section as a function of $|\Delta\eta(\ell, \ell)|$ in the dilepton channel, accounting for the statistical and systematic uncertainties.

Bin	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.5
0.0 - 0.2	1.00	-0.02	-0.06	-0.17	-0.11	-0.16	-0.22	-0.08
0.2 - 0.4	-0.02	1.00	0.16	0.14	0.11	0.18	0.16	0.36
0.4 - 0.6	-0.06	0.16	1.00	0.17	0.16	0.21	0.29	0.27
0.6 - 0.8	-0.17	0.14	0.17	1.00	0.20	0.28	0.41	0.21
0.8 - 1.0	-0.11	0.11	0.16	0.20	1.00	0.16	0.27	0.20
1.0 - 1.2	-0.16	0.18	0.21	0.28	0.16	1.00	0.39	0.24
1.2 - 1.7	-0.22	0.16	0.29	0.41	0.27	0.39	1.00	0.30
1.7 - 2.5	-0.08	0.36	0.27	0.21	0.20	0.24	0.30	1.00

Not reviewed, for internal circulation only

Table 67: The correlation matrix for the normalised differential cross-section as a function of $\Delta\phi(\ell,\ell)$ in the dilepton channel, accounting for the statistical and systematic uncertainties.

Bin	0.0 - 0.35	0.35 - 0.7	0.7 - 1.05	1.05 - 1.4	1.4 - 1.75	1.75 - 2.1	2.1 - 2.45	2.45 - 2.8	2.8 - 3.14
0.0 - 0.35	1.00	-0.07	0.05	-0.36	-0.18	-0.08	-0.19	-0.13	-0.24
0.35 - 0.7	-0.07	1.00	0.29	0.21	0.09	0.11	0.21	0.02	0.32
0.7 - 1.05	0.05	0.29	1.00	0.12	0.05	0.10	0.14	0.05	0.23
1.05 - 1.4	-0.36	0.21	0.12	1.00	0.26	0.18	0.29	0.20	0.41
1.4 - 1.75	-0.18	0.09	0.05	0.26	1.00	0.15	0.16	0.14	0.21
1.75 - 2.1	-0.08	0.11	0.10	0.18	0.15	1.00	0.17	0.15	0.23
2.1 - 2.45	-0.19	0.21	0.14	0.29	0.16	0.17	1.00	0.12	0.33
2.45 - 2.8	-0.13	0.02	0.05	0.20	0.14	0.15	0.12	1.00	0.18
2.8 - 3.14	-0.24	0.32	0.23	0.41	0.21	0.23	0.33	0.18	1.00

Not reviewed, for internal circulation only

(a)

Figure 119: The statistical correlations among the normalised differential cross-sections of the single-lepton channel. The covariance is evaluated by unfolding 1000 statistically coupled (co-varied) replicas of individual spectra in data using the bootstrap method.

Not reviewed, for internal circulation only

(a)

Figure 120: The statistical correlations among the normalised differential cross-sections of the dilepton channel. The covariance is evaluated by unfolding 1000 statistically coupled (co-varied) replicas of individual spectra in data using the bootstrap method.

Not reviewed, for internal circulation only

14 Conclusion

Not reviewed, for internal circulation only

References

[1] U. Husemann, *Top-Quark Physics: Status and Prospects*, Prog. Part. Nucl. Phys. **95** (2017) 48, arXiv: 1704.01356 [hep-ex].

[2] U. Baur, A. Juste, L. H. Orr and D. Rainwater, *Probing electroweak top quark couplings at hadron colliders*, Phys. Rev. **D71** (2005) 054013, arXiv: hep-ph/0412021 [hep-ph].

[3] A. O. Bouzas and F. Larios, *Electromagnetic dipole moments of the Top quark*, Phys. Rev. **D87** (2013) 074015, arXiv: 1212.6575 [hep-ph].

[4] R. Röntsch and M. Schulze, *Probing top-Z dipole moments at the LHC and ILC*, JHEP **08** (2015) 044, arXiv: 1501.05939 [hep-ph].

[5] M. Schulze and Y. Soreq, *Pinning down electroweak dipole operators of the top quark*, Eur. Phys. J. **C76** (2016) 466, arXiv: 1603.08911 [hep-ph].

[6] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou and C. Zhang, *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, JHEP **05** (2016) 052, arXiv: 1601.08193 [hep-ph].

[7] P.-F. Duan, Y. Zhang, Y. Wang, M. Song and G. Li, *Electroweak corrections to top quark pair production in association with a hard photon at hadron colliders*, Phys. Lett. **B766** (2017) 102, arXiv: 1612.00248 [hep-ph].

[8] T. Aaltonen et al., *Evidence for $t\bar{t}\gamma$ Production and Measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$*, Phys. Rev. **D84** (2011) 031104, arXiv: 1106.3970 [hep-ex].

[9] ATLAS Collaboration, *Measurement of the inclusive $t\bar{t}\gamma$ cross section at $\sqrt{s} = 7$ TeV with the ATLAS detector*, ATLAS-CONF-2011-153, 2011, URL: https://cds.cern.ch/record/1398197.

[10] ATLAS Collaboration, *Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. **D91** (2015) 072007, arXiv: 1502.00586 [hep-ex].

[11] *Measurement of the inclusive top-quark pair + photon production cross section in the muon + jets channel in pp collisions at 8 TeV*, tech. rep. CMS-PAS-TOP-13-011, CERN, 2014, URL: https://cds.cern.ch/record/1644573.

[12] *Measurement of the $t\bar{t}+\gamma$ production cross-section in pp collisions at $\sqrt{s} = 8$ TeV*, tech. rep. CMS-PAS-TOP-14-008, CERN, 1900, URL: https://cds.cern.ch/record/2216566.

[13] M. Aaboud et al., *Measurement of the $t\bar{t}\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, (2017), arXiv: 1706.03046 [hep-ex].

[14] K. Melnikov, M. Schulze and A. Scharf, *QCD corrections to top quark production in association whith a photon at hadron colliders*, Phys.Rev. **D 83** (2011) (), eprint: arXiv:1102.1967[hep-ph], URL: http://arxiv.org/abs/1102.1967.

[15] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

Not reviewed, for internal circulation only

[16] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, URL: https://cds.cern.ch/record/1291633.

[17] *ATLAS Insertable B-Layer Technical Design Report Addendum*, tech. rep. CERN-LHCC-2012-009. ATLAS-TDR-19-ADD-1, Addendum to CERN-LHCC-2010-013, ATLAS-TDR-019, 2012, URL: https://cds.cern.ch/record/1451888.

[18] M. Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** (2016) 653, arXiv: 1608.03953 [hep-ex].

[19] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[20] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195 [hep-ph].

[21] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[22] *ATLAS Run 1 Pythia8 tunes*, tech. rep. ATL-PHYS-PUB-2014-021, CERN, 2014, URL: https://cds.cern.ch/record/1966419.

[23] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[24] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007.

[25] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[26] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv: 1005.3457 [hep-ph].

[27] E. Barberio, B. van Eijk, and Z. Was, *Photos- a universal Monte Carlo for QED radiative corrections in decays*, Comput. Phys. Commun. 66 (1991) 155, 1991.

[28] ATLAS Collaboration, *AnalysisTop framework (restricted)*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisTop.

[29] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016, URL: https://cds.cern.ch/record/2157687.

[30] ATLAS Collaboration, *ATLAS electron, photon and muon isolation in Run 2: The isolation variables reconstruction and performance for Run 2 are documented.*, tech. rep. ATL-COM-PHYS-2017-290, This note contains the Moriond 2017 recommendations. It will be updated when new recommendations become available.: CERN, 2017, URL: https://cds.cern.ch/record/2256658.

Not reviewed, for internal circulation only

[31] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2016-015, 2016, URL: https://cds.cern.ch/record/2203514.

[32] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex].

[33] ATLAS Collaboration, *Photon identification in 2015 ATLAS data*, ATL-PHYS-PUB-2016-014, 2016, URL: https://cds.cern.ch/record/2203125.

[34] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[35] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-036, 2015, URL: https://cds.cern.ch/record/2044564.

[36] M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, (2017), arXiv: 1703.09665 [hep-ex].

[37] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[38] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, 2015, URL: https://cds.cern.ch/record/2037697.

[39] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, tech. rep. ATL-COM-PHYS-2016-407, CERN, 2016, URL: https://cds.cern.ch/record/2149445.

[40] F. Chollet et al., *Keras*, https://github.com/fchollet/keras, 2015.

[41] D. H. Guest, M. Paganini, M. Kagan, J. W. Smith and M. Lanfermann, *lwtnn/lwtnn: Release for Athena v21*, 2017, URL: https://doi.org/10.5281/zenodo.290682.

[42] C. Anastopoulos et al., *Photon identification pre-recommendations for run 2*, tech. rep. ATL-COM-PHYS-2015-496, CERN, 2015, URL: https://cds.cern.ch/record/2022342.

[43] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, CoRR **abs/1502.03167** (2015), URL: http://arxiv.org/abs/1502.03167.

[44] M. Aaboud et al., *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[45] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: https://cds.cern.ch/record/2037702.

[46] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, (2016), arXiv: 1606.01813 [hep-ex].

Not reviewed, for internal circulation only

[47] ATLAS Collaboration, *Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s}$ = 13 TeV with AnalysisTop-2.3.41*, tech. rep. ATL-COM-PHYS-2016-198, CERN, 2016, URL: https://cds.cern.ch/record/2135116.

[48] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph].

[49] V. Blobel, 'An Unfolding method for high-energy physics experiments', *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002*, 2002 258, arXiv: hep-ex/0208022 [hep-ex], URL: http://www.ippp.dur.ac.uk/Workshops/02/statistics/proceedings//blobel2.pdf.

[50] F. Spano, 'Unfolding in particle physics: a window on solving inverse problems', *Proceedings, 6th International Workshop on Top Quark Physics (TOP2013): Durbach, Germany, September 14-19, 2013*, 2014 256, URL: http://inspirehep.net/record/1344095/files/P52.pdf.

[51] G. D'Agostini, *Improved iterative Bayesian unfolding*, (), eprint: arXiv:1010.0632v1[physics.data-an], URL: https://arxiv.org/abs/1010.0632.

[52] A. Hocker and V. Kartvelishvili, *SVD approach to data unfolding*, Nucl. Instrum. Meth. **A372** (1996) 469, arXiv: hep-ph/9509307 [hep-ph].

[53] *RooUnfold package*, URL: http://hepunx.rl.ac.uk/~adye/software/unfold/RooUnfold.html.

[54] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665v2 [hep-ex].

[55] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s}$=7 TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 17, arXiv: 1406.0076v3 [hep-ex].

[56] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: http://cdsweb.cern.ch/record/2160731.

[57] ATLAS Collaboration, *Pile up reweighting tool (restricted)*, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ExtendedPileupReweighting.

[58] D. Peng-Fei et al., *QCD corrections to associated production of $t\bar{t}\gamma$ at hadron colliders*, (2009), arXiv: 0907.1324 [hep-ph].

[59] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, Phys. Lett. **B761** (2016) 136, arXiv: 1606.02699 [hep-ex].

Appendices

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

A Studies on removal of the double counting of events between samples

A.1 Definition of true hadronic fakes

For the removal of the double counting of events between samples, as explained in Section 3.2, if the reconstructed photon is associated with the type *BkgPhoton*, with an origin corresponding to a hadron, the photon is considered as hadronic fake. An study shown here revealed that this is only accounting for the non-prompt photons decayed from hadrons. But another possible source of hadronic fakes are the hadrons (jets) that are mis-reconstructed as photons. To include this missing category, the reconstructed photons that are associated with the type *Hadron* are also added to the hadronic fake truth definition, and the effect is studied. The study is done only in single lepton channel, since hadronic fake background contribution in dilepton channel is very small. Also, this study is done by using $t\bar{t}$ MC sample as $t\bar{t}$ forms more than 90% of hadronic fake events in single lepton channel (Section F.1).

Figure 121 shows the origin, type and PDG ID of the truth particles corresponding to the reconstructed photons in the $t\bar{t}$ MC sample, before performing any truth matching, and Figure 122 shows the same after the hadronic fake truth matching of Section 3.2 is performed. 11% of the events survive the hadronic fake truth matching, and all of the true hadronic fakes have the PDG ID of a photon. Figure 123 shows the same, this time with the category of type *Hadron* added to the nominal hadronic fake truth definition. This adds 5% more events, where all the newly added true hadronic fakes have a PDG ID >100, which is of a hadron, as was checked from 2D plots (and all belong to photon origin *NonDefined*).

Since the hadronic fake background in this analysis is scaled by the data driven scale factors, this 5% underestimation of MC is covered. But the data driven scale factor which is 3D binned of photon p_{T}-η-conversion type, uses a MC driven correction factor called θ_{MC}, also 3D binned (see Section 6.2.1). Therefore, the θ_{MC} factor in bins of photon p_{T}-η-conversion type before and after adding the missing category of hadronic fakes is compared in Figure 124. It can be seen that within the statistical uncertainties the θ_{MC} factor does not change, therefore the missing category of hadronic fakes does not have any effect on the shape and normalisation of the estimated hadronic fake background.

In addition, the shape of ELD, which is used as the discriminating variable in the fit, for the hadronic fake background is compared before and after adding the missing category of hadronic fakes. As Figure 125 shows, the ELD shape of hadronic fake background does not change within the statistical uncertainties.

Another check was done to see if the missing category of hadronic fakes is wrongly ending up in the prompt photons category, since as mention in Section 3.2 any reconstructed photon that is not categories under $e \rightarrow \gamma$ fake or hadronic fake after truth matching, is considered as a true prompt photon. By adding the type *Hadron* to the hadronic fake truth definition, the yield of $t\bar{t}\gamma$ and $V\gamma$ samples do not change, and the change in the *other prompts* yield is negligible (0.4%).

The conclusion of this study is that missing hadrons mis-reconstructed as photons in the hadronic fake truth definition did not have any effect on the final result of this analysis, but for future analyses it is recommended to add the type *Hadron* to the hadronic fakes truth definition.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 121: Origin (a), type (b) and PDG ID (c) of the truth particle corresponding to the reconstructed photon in the $t\bar{t}$ MC sample, before performing any truth matching. In (c) the last bin contains the overflow.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 122: Origin (a), type (b) and PDG ID (c) of the truth particle corresponding to the reconstructed photon in the $t\bar{t}$ MC sample, after performing the nominal truth matching to select hadronic fakes as explained in Section 3.2. In (c) the last bin contains the overflow.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 123: Origin (a), type (b) and PDG ID (c) of the truth particle corresponding to the reconstructed photon in the $t\bar{t}$ MC sample, after performing truth matching to select hadronic fakes as explained in Section 3.2 with an additional group of reconstructed photons associated with the truth type *Hadron*. In (c) the last bin contains the overflow.

Not reviewed, for internal circulation only

(a)

(b)

Figure 124: Comparison of θ_{MC} factor with hadronic fake truth definition of Section 3.2 (Nominal) and after adding the type *Hadron* to the hadronic fakes truth definition (Alternative), in different photon η slices as function of photon p_{T}, separated for converted (a) and unconverted (b) photons. The error bars show the statistical uncertainties.

Not reviewed, for internal circulation only

(a) (b)

Figure 125: Comparison of ELD shape of hadronic fake background with hadronic fake truth definition of Section 3.2 (Nominal) and after adding the type *Hadron* to the hadronic fakes truth definition (Alternative), in e+jets (a) and μ+jets (b) channels. The error bars show the statistical uncertainties.

Not reviewed, for internal circulation only

A.2 Comparison with other methods

Other analyses are using $t\bar{t}\gamma$ samples and $t\bar{t}$ samples and need to remove the overlap. The usual method used in those analyses is described here and compared with the method used in the $t\bar{t}\gamma$ measurement. This method considers that a $t\bar{t}$ event is overlaping with $t\bar{t}\gamma$ when, amongst the list of truth particle, exists a truth particle photon with p_{T} higher than the generation threshold and not coming from a hadron. If it is the case, the event is considered as possessing a photon from matrix-element. By construction, the $t\bar{t}\gamma$ sample events all possess a photon from matrix-element, and the events of the $t\bar{t}$ sample that possess a photon from matrix-element are discarded to avoid double counting. For all truth particle with pdgId equals to 22, the parent particle is retrieved (parents with pdgId equals to 22 are skipped, the parent of the parent is then used). If the parent absolute value of pdgId is lower than 100 (or if the particle has no parent), that the barcode of the particle is lower than 200000 and that the p_{T} of the particle is higher than the generator p_{T} threshold, the event is considered as possessing a matrix-element photon. This method is usually used to remove the overlap with $t\bar{t}\gamma$ samples with high photon p_{T} threshold (e.g., $p_{\mathrm{T}} > 100$GeV), and this algotihm has not been optimized or intensively studied in the case of a low threshold (e.g., $p_{\mathrm{T}} > 15$GeV like in the $t\bar{t}\gamma$ measurement). In the case of the $t\bar{t}$ sample, this alternative method agrees with the classification used in this analysis in 95% of the cases. The difference is due to the difference in approach of the two methods. The one on the $t\bar{t}\gamma$ analysis needs to separate the $e \rightarrow \gamma$ fake and hadronic fake that are estimated from data-driven methods, while this method is separating based on what has been generated. In this method, events that are considered as possessing a photon from matrix-element will still be double counted in the $e \rightarrow \gamma$ fake and hadronic fake method if the reconstructed photon does not correspond to this matrix-element photon. In particular, for this method, a photon emitted colinearly by an electron and where the electron deposit is included in the reconstructed photon will be considered as a photon from matrix-element, while in the $t\bar{t}\gamma$ analysis, this contribution is included in the $e \rightarrow \gamma$ fake, as it corresponds to what is extracted from the data-driven estimation. Additionnally, for a really small fraction of events, an artificial intermediate parent particle with high pdgId value can be added by the generator for technical reason, and this method will therefore classify the event has not possessing a photon from matrix-element.

Another method consists into exploring the decay chain from which the truth particle associated to the reconstructed photon is originating. In this case, the algorithm is the following. For the truth particle associated to the photon, the parent is retrieved. If the parent has the same pdgId as the child particle, the parent is ignored. If not, the pdgId value is added to a vector. If a particle has two parents, the current vector is duplicated and the procedure is resumed in parallel for the two parents. If there is no more parents or the parent was already accounted (in case of loop), the procedure is stopped. For the reconstructed photon, if the obtained vectors contain at least one pdgId with an absolute value higher than 100, this reconstructed photon is considered as hadronic fake. The condition for being $e \rightarrow \gamma$ fake are identical to the one used in the analysis, and has priority over the hadronic fake classification. And if the reconstructed photon is not classified as hadronic fake or $e \rightarrow \gamma$ fake, it is therefore considered as prompt. Those two methods provide very similar results, with less than 0.1% disagreement on the classification. Disagreement usually corresponds to a case where the truth particle type and origin is not accounted in the condition of the algorithm used in the $t\bar{t}\gamma$ analysis (for example type *Hadron*).

Not reviewed, for internal circulation only

A.3 Et ratio plots

The quality of truth matching is studied by drawing the 2D plots of $\Delta R(reco, true)$ v.s. $Et(reco)/Et(true)$, for each of the 3 types of truth photons using the $t\bar{t}$ MC in the SR of μ+jets, as shown in Figure 126. The hadron fake has the expected behavior of much more smeared than the others.

Figure 126

Not reviewed, for internal circulation only

B Additional information for PPT

B.1 Photon kinematics in the training samples

Figure 127 shows comparisons of the p_{T} and $|\eta|$ spectra of the photons used as signal in the training of the PPT and of $t\bar{t}\gamma$ photons in the e+jets and μ+jets signal regions. The spectra for the signal regions are taken from the $t\bar{t}\gamma$ Monte-Carlo simulation described in 3, and are identical to the distributions marked as $t\bar{t}\gamma$ in Fig. 15 and Fig. 16.

B.2 Comparison of different architectures of PPT

Tab. 68 summarises the AUC values and the information whether the training and test ROC curves are in agreement of all PPT configurations tested. Tab. 69 contains information on the three different architectures considered.

(a) (b)

Figure 127: p_{T} and $|\eta|$ spectra of the photons used as signal during the training of the PPT in comparison with $t\bar{t}\gamma$ photons in the signal regions of the μ+jets and e+jets channels.

Not reviewed, for internal circulation only

archi-tecture	**features**	**AUC (training)**	**AUC (test)**	**ROC agreement**
1	R_η, R_ϕ, f_{side}, $w_{\eta 1}$	0.8249	0.8266	✓
	R_η, R_ϕ, f_{side}, $w_{\eta 1}$, R_{had}	0.8414	0.8401	✓
	R_η, R_ϕ, f_{side}, $w_{\eta 1}$, R_{had}, $w_{\eta 2}$	0.8486	0.8477	✓
2	R_η, R_ϕ, f_{side}, $w_{\eta 1}$	0.8411	0.8419	✓
	R_η, R_ϕ, f_{side}, $w_{\eta 1}$, R_{had}	0.8668	0.8657	✓
	R_η, R_ϕ, f_{side}, $w_{\eta 1}$, R_{had}, $w_{\eta 2}$	0.8654	0.8649	✓
3	R_η, R_ϕ, f_{side}, $w_{\eta 1}$	0.8460	0.8452	✓
	R_η, R_ϕ, f_{side}, $w_{\eta 1}$, R_{had}	0.8631	0.8621	✓
	R_η, R_ϕ, f_{side}, $w_{\eta 1}$, R_{had}, $w_{\eta 2}$	0.8693	0.8690	✓

Table 68: Three different architectures tested for the PPT. Starting with one layer, further layers are added successively. The output layer is included in all cases.

architectures	
1 (input layer)	`relu`, 64 neurons
2	+ `batchnormalisation` + `softmax`, 40 neurons
3	+ `batchnormalisation` + `softmax`, 52 neurons
output layer	`sigmoid`, 1 neuron

Table 69: Visualisation of the NN architecture used for the PPT.

B.3 Further data/MC plots for PPT systematics

Figure 128 shows control plots for various observables in the PPT hadronic fake control region, where the signal contribution is increased to 150%. Figure 129 shows the corresponding PPT output distribution.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 128: Control plots in the PPT hadronic fake control region including distributions of the photon shower shape variables R_{had}, f_{side}, R_η, R_ϕ, $w_{\eta 1}$ and $w_{\eta 2}$ used as input to the NN. The signal contribution is increased to 150% and MC is scaled to data for shape comparisons. Only statistical uncertainties are included.

Figure 129: PPT distributions in hadronic fake control regions where the signal contribution is scaled to 150%. MC is scaled to data for shape comparison and only statistical uncertainties are included.

Not reviewed, for internal circulation only

B.4 Derivations used for training

Table 70 summarises the derivations used for training and testing the PPT.

single photons, QCD compton (training)
mc15_13TeV.423099.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP8_17.merge.DAOD_STDM2.e4453_s2726_r7725_r7676_p2669/
mc15_13TeV.423100.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP17_35.merge.DAOD_STDM2.e3791_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423101.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP35_50.merge.DAOD_STDM2.e3904_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423102.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP50_70.merge.DAOD_STDM2.e3791_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423103.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP70_140.merge.DAOD_STDM2.e3791_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423104.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP140_280.merge.DAOD_STDM2.e3791_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423105.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP280_500.merge.DAOD_STDM2.e3791_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423106.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP500_800.merge.DAOD_STDM2.e3791_s2608_s2183_r7725_r7676_p2669/
mc15_13TeV.423107.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP800_1000.merge.DAOD_STDM2.e4453_s2726_r7725_r7676_p2669/
mc15_13TeV.423108.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP1000_1500.merge.DAOD_STDM2.e4453_s2726_r7725_r7676_p2669/
mc15_13TeV.423109.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP1500_2000.merge.DAOD_STDM2.e4453_s2726_r7725_r7676_p2669/
mc15_13TeV.423110.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP2000_2500.merge.DAOD_STDM2.e4453_s2726_r7725_r7676_p2669/
mc15_13TeV.423111.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP2500_3000.merge.DAOD_STDM2.e4453_s2726_r7772_r7676_p2669/
mc15_13TeV.423112.Pythia8EvtGen_A14NNPDF23LO_gammajet_DP3000_inf.merge.DAOD_STDM2.e4453_s2726_r7772_r7676_p2669/
QCD di-jet (training)
mc15_13TeV.423300.Pythia8EvtGen_A14NNPDF23LO_perf_JF17.merge.DAOD_STDM2.e3848_s2608_s2183_r7725_r7676_p2666/
mc15_13TeV.423301.Pythia8EvtGen_A14NNPDF23LO_perf_JF23.merge.DAOD_STDM2.e3848_s2608_s2183_r7725_r7676_p2666/
mc15_13TeV.423302.Pythia8EvtGen_A14NNPDF23LO_perf_JF35.merge.DAOD_STDM2.e3848_s2608_s2183_r7725_r7676_p2666/
mc15_13TeV.423303.Pythia8EvtGen_A14NNPDF23LO_perf_JF50.merge.DAOD_STDM2.e3848_s2608_s2183_r7725_r7676_p2666/
Radiative Z, $Z \rightarrow e^+e^-\gamma$ (evaluation, MC)
mc15_13TeV.301535.Sherpa_CT10_eegammaPt10_35.merge.DAOD_EGAM3.e3952_s2608_s2183_r7725_r7676_p2613
mc15_13TeV.301899.Sherpa_CT10_eegammaPt35_70.merge.DAOD_EGAM3.e3952_s2608_s2183_r7725_r7676_p2613
mc15_13TeV.301900.Sherpa_CT10_eegammaPt70_140.merge.DAOD_EGAM3.e3952_s2608_s2183_r7725_r7676_p2613

Table 70: Summary of derivations used for training and evaluation of the PPT.

B.5 Shower shape comparisons in different p_{T} and η regions

The topology of the detector is not homogeneous. For instance, the granularity in the calorimeters is a function of η and different sub-components cover slightly different kinematic regions. Furthermore, material upstream is present in the detector. Since shower shapes evolve due to the photon interacting with the detector material, they consequently depend on the kinematic region they are measured in. Especially the granularity of the strip layer in the end-caps of the electromagnetic calorimeter gets coarser with increasing $|\eta|$ [15]. Figs. 130 and 131 shows the shower shape distributions of R_{had}, R_{η}, R_{ϕ}, $w_{\eta 1}$, $w_{\eta 2}$ and f_{side} in 3 different bins of p_{T} and η, respectively.

The classification performance of the PPT can hence expected to be different between the training on QCD-Compton and di-jet events and the application to $t\bar{t}\gamma$ signal and background events. For instance, Fig. 132 shows the comparison of the PPT distribution of prompt photons in QCD-Compton events and $t\bar{t}\gamma$ signal events.

Not reviewed, for internal circulation only

Figure 130: Photon shower shapes R_{had}, R_η, R_ϕ, $w_{\eta 1}$, $w_{\eta 2}$ and f_{side} in three different bins of p_{T}.

Not reviewed, for internal circulation only

Figure 131: Photon shower shapes R_{had}, R_η, R_ϕ, $w_{\eta 1}$, $w_{\eta 2}$ and f_{side} in three different bins of η.

Not reviewed, for internal circulation only

Figure 132: PPT distribution of prompt photons in QCD-Compton events and $t\bar{t}\gamma$ signal events.

B.6 Shower shape comparisons for training and test split

Figure 133 shows shower shape distributions for the six observables used as input to the NN, randomly split into 80% training and 20% training sets.

B.7 Systematic scale factors for the PPT

As described in section 4.2.2, shape uncertainties for the PPT are extracted from three different control regions and applied as systematic variations to the signal regions. Figure 134 shows all extracted scale factors bin by bin for the three shape systematics: hadronic fake, hadronic fake isolation and prompt.

Not reviewed, for internal circulation only

Figure 133: Shower shape distributions R_{had}, R_{η}, R_{ϕ}, $w_{\eta 1}$, $w_{\eta 2}$ and f_{side} of prompt photons where all samples are randomly split into orthogonal 80 % training and 20 % test sets.

Not reviewed, for internal circulation only

(a) PPT hadronic fake scale factors

(b) PPT hadronic fake isolation scale factors

(c) PPT prompt scale factors

Figure 134: Bin-by-bin scale factors for the three systematic uncertainties of the PPT. The displayed error bars show statistical uncertainties on the derived values. The scale factors are derived from the p_{T} and $|\eta|$ slices of the three control regions: hadronic fake, hadronic fake isolation and prompt, as described in 4.2.2. The scale factors shown in 134(c) are applied as corrections to the prompt photon contributions to the signal regions, and all three sets of scale factors are used as systematic variations.

Not reviewed, for internal circulation only

C Cuts optimization

The cut optimization is done based on reducing the total relative uncertainty on the expected cross section. The expected cross section is given by the following:

$$\hat{\sigma} = \frac{N_{expected} - N_{background}}{\epsilon \cdot \mathcal{L}} = \frac{(N_{signal} + N_{background}) - N_{background}}{\epsilon \cdot \mathcal{L}} = \frac{N_{signal}}{\epsilon \cdot \mathcal{L}} \tag{62}$$

The cross section in Equation 62 includes the Luminosity $\mathcal{L}$, the efficiency ϵ and the expected number of events in data $N_{expected}$. The efficiency ϵ is given by:

$$\epsilon = \frac{N_{signal}}{N_{signal}^{initial}} \tag{63}$$

Where N_{signal} is the number of events in the signal sample after the full selection and $N_{signal}^{initial}$ is the number of events generated initially in the MC signal sample.

The uncertainty on the cross section can be given by:

$$\Delta\hat{\sigma} = \sqrt{(\Delta\hat{\sigma}_{stat})^2 + \left(\Delta\hat{\sigma}_{syst}\right)^2}\,. \tag{64}$$

$$\Delta\hat{\sigma}_{stat} = \sqrt{\left(\frac{\Delta\left(N_{signal} + N_{background}\right)}{\epsilon \cdot \mathcal{L}}\right)^2 + \left(\frac{\Delta N_{background}}{\epsilon \cdot \mathcal{L}}\right)^2} = \frac{\sqrt{N_{signal} + N_{background}}}{\epsilon \cdot \mathcal{L}} \tag{65}$$

$$\Delta\hat{\sigma}_{syst} = \sqrt{\left(\frac{\Delta N_{background}^{syst}}{\epsilon \cdot \mathcal{L}}\right)^2 + \left(\hat{\sigma} \cdot \frac{\Delta\epsilon}{\epsilon}\right)^2 + \left(\hat{\sigma} \cdot \frac{\Delta\mathcal{L}}{\mathcal{L}}\right)^2} \tag{66}$$

C.1 Cuts optimization in the single lepton channel

The simulated MC samples described in Section 3 are used here, normalized to their cross section and an integrated luminosity of 36.47 fb^{-1} (Note: The cut optimisation was done with an older calculation of the total luminosity, different from the current number which is 36.1 fb^{-1}.).

The following systematic uncertainties are considered (Note: The cut optimisation was done with an older conservative set of systematics, different from the current systematics numbers.):

- Luminosity uncertainty: 5.5%
- hadronic fakes background modelling: 5%
- $e \rightarrow \gamma$ fakes background modelling: 10%
- other backgrounds modelling: 50%

Not reviewed, for internal circulation only

The optimization is done for a cut on each variable independently from the other, so it is a one dimensional optimization. The selection cuts are optimized for the following set of variables:

- $E_{\mathrm{T}}^{\mathrm{miss}}$ in a range of [0,100] GeV
- $m_T(W)$ in a range of [0,100] GeV

For the optimization of a variable, the other variable is set to the minimum.

C.1.1 Results in the single lepton channel

The missing transverse momentum distributions in both single lepton channels and the expected relative uncertainties on the cross section for a shorter range of the variable are shown in Figure 135. The same is shown for $m_T(W)$ in Figure 136. For both variables systematic uncertainty is dominating.

For both variables, the relative total uncertainty on the expected cross section shows a rather flat behaviour in the smaller values and increases for the larger values. Therefore, it is decided to apply no cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$, $m_T(W)$ for the signal region selection.

Not reviewed, for internal circulation only

(a) (b)

(c) (d)

Figure 135: The missing transverse momentum distributions in e+jets (a) and μ+jets (b), and the relative uncertainty on the expected cross section as a function of the variable in e+jets (c) and μ+jets (d), with the systematic uncertainty in red, the statistical uncertainty in blue and the total uncertainty in green.

Not reviewed, for internal circulation only

(a) (b)

(c) (d)

Figure 136: The $m_T(W)$ distributions in e+jets (a) and μ+jets (b), and the relative uncertainty on the expected cross section as a function of the variable in e+jets (c) and μ+jets (d), with the systematic uncertainty in red, the statistical uncertainty in blue and the total uncertainty in green.

Not reviewed, for internal circulation only

C.2 Cuts optimization in the dilepton channel

The simulated MC samples described in Section 3 are used here, normalized to their cross section and an integrated luminosity of 36.47 fb^{-1} (Note: The cut optimisation was done with an older calculation of the total luminosity, different from the current number which is 36.1 fb^{-1}.).

The following systematic uncertainties are considered (Note: The cut optimisation was done with an older conservative set of systematics, different from the current systematics numbers.):

- Luminosity uncertainty: 2.9%
- Background modelling: 10%
- b-tagging, jet, leptons scales uncertainties are taken from the published $t\bar{t}$ differential cross section measurement in the $e\mu$ channel [59]
- Photon identification uncertainty

The optimization is done for a cut on each variable independently from the other, so it is a one dimensional optimization. The selection cuts are optimized for the following set of variables:

- The distance between the the photon and the jet in a range of [0.4, 1.0]
- The distance between the photon and any lepton is a range of [0.4, 1.0]
- The missing transverse momentum in the same flavour channels in a range of [0 , 70] GeV
- The number of b-tagged jets in a range of [0, 2] b-tagged jets

For the optimization of any of the first three variables, their values are set to the minimum, while other selections described in Section 5 are used. Similarly, for optimizing the number of b-tagged jets, no b-tag requirement was asked, a cut on missing transverse momentum of 30 GeV(see Figure 137) was applied to reduce background from fake leptons and hence reducing its uncertainty on the cross section, while the two distances variables are set to minimum.

C.2.1 Results in the dilepton channel

The result of optimizing the missing transverse momentum is shown in Figure 137, including the distributions in the three channels and the expected uncertainty on the cross section for a shorter range of the variable. The systematic uncertainty is dominant for this variable. A minimum of the total uncertainty at 30 GeVis observed for ee and $\mu\mu$ channels and hence a cut at this value is preferred, while applying no cut in the $e\mu$ channel is preferred.

The result of optimizing the distance between the lepton and the photon is shown in Figure 138. The statistical uncertainty is dominant for this variable. A minimum of the total uncertainty at 0.5 is observed for ee and $\mu\mu$ channels, while no preferred value for $e\mu$ channel. However a cut of 1.0 is applied in all the three channels in order to enhance the radiative photons production, as a result of the study which is documented in Appendix D).

The result of optimizing the distance between the jet and the photon is shown in Figure 139. The statistical uncertainty is dominant for this variable. No preferred value for minimizing the total uncertainty is observed in any of three channels. Hence no cut at event selection will be applied (the cut at the value 0.4

Not reviewed, for internal circulation only

which can be seen in the distributions comes from the object level selection, from the jet-photon overlap removal described in Section 4.1).

The result of optimizing the number of b-tagged jets is shown in Figure 140. The statistical uncertainty is dominant for this variable. A minimum of the total uncertainty at a cut of at least one b-tagged jet is observed in the three channels, hence this cut value is included in the signal selection.

(a) (b) (c)

(d) (e) (f)

Figure 137: The transverse momentum distributions in $e\mu$(a), ee (b) and $\mu\mu$ (c). The uncertainty on the expected cross section as a function of the variable in $e\mu$(d), ee (e) and $\mu\mu$ (f). The systematic uncertainty in red, the statistical uncertainty on blue and the total uncertainty in black.

Not reviewed, for internal circulation only

Figure 138: The distance between the photon and the lepton distributions in $e\mu$(a), ee (b) and $\mu\mu$ (c). The uncertainty on the expected cross section as a function of the variable in $e\mu$(d), ee (e) and $\mu\mu$ (f). The systematic uncertainty in red, the statistical uncertainty on blue and the total uncertainty in black.

Not reviewed, for internal circulation only

Figure 139: The distance between the photon and the jet distributions in $e\mu$(a), ee (b) and $\mu\mu$ (c). The uncertainty on the expected cross section as a function of the variable in $e\mu$(d), ee (e) and $\mu\mu$ (f). The systematic uncertainty in red, the statistical uncertainty on blue and the total uncertainty in black.

Not reviewed, for internal circulation only

Figure 140: The number of b-tagged jets distributions in $e\mu$(a), ee (b) and $\mu\mu$ (c). The uncertainty on the expected cross section as a function of the variable in $e\mu$(d), ee (e) and $\mu\mu$ (f). The systematic uncertainty in red, the statistical uncertainty on blue and the total uncertainty in black.

Not reviewed, for internal circulation only

D Radiative decay versus radiative production

A brief study is done here, to check what is the fraction of photons that are radiated from initial state or off-shell top quark (called radiative $t\bar{t}\gamma$ production) out of all photons survived some selections in the $t\bar{t}\gamma$ signal sample.

The results are summarized in Table 71. The first column is using the basic $t\bar{t}$ selections with requiring a photon in addition, the second column is using the selections described in Section 5 (the photon isolation cut is not used here) but keeping the distance between the leptons and the photon at 0.7, the third column is using the same selections but changing the distance variable to 1.0, where the fraction of the radiative photon production is enhanced, specially in the dilepton channel. The fourth column is using the selections described in Section 5, but changing the photon transverse momentum to 40 GeV.

Table 71: The fraction of radiative photon production after different sets of selections. The photon isolation requirement is not applied here.

Channel	$t\bar{t}$ sel. + photons	$t\bar{t}\gamma$ sel. with $\Delta R(\text{lepton}, \gamma) > 0.7$	$t\bar{t}\gamma$ sel. with $\Delta R(\text{lepton}, \gamma) > 1.0$	$t\bar{t}\gamma$ sel. with $\Delta R(\text{lepton}, \gamma) > 0.1$ and $p_{\text{T}}(\gamma) > 40$ GeV
e + jets	35%	44%	45%	55%
μ + jets	35%	44%	45%	55%
$e\mu$	34%	48%	55%	67%
$\mu\mu$	33%	40%	54%	63%
ee	34%	39%	59%	68%

Similar study is done after the photon isolation cut was introduced to the analysis, to check again the effect of $\Delta R(\text{lepton}, \gamma)$ cut, by changing its valued from 0.4 to 0.7 and to 1.0, while keeping all the rest of the $t\bar{t}\gamma$ selection as is described in Section 5. The results which are summarised in Table 72, show that $\Delta R(\text{lepton}, \gamma) > 1.0$ enhances the radiative production by 20% in dilepton channel and 12% in single lepton channel.

Table 72: The fraction of radiative photon production after different cut values for $\Delta R(\text{lepton}, \gamma)$.

Channel	$t\bar{t}$ sel. with $\Delta R(\text{lepton}, \gamma) > 0.4$	$t\bar{t}\gamma$ sel. with $\Delta R(\text{lepton}, \gamma) > 0.7$	$t\bar{t}\gamma$ sel. with $\Delta R(\text{lepton}, \gamma) > 1.0$
dilepton	45%	56%	65%
single lepton	45%	54%	57%

After fixing the $\Delta R(\text{lepton}, \gamma)$ cut value to 1.0, the same practice is repeated for different sets of $\Delta R(\text{jet}, \gamma)$ cuts. The result are summarised in Table 73, which shows introducing a cut on $\Delta R(\text{jet}, \gamma)$ beyond what is imposed in the object overlap removal is not as effective as the cut on $\Delta R(\text{lepton}, \gamma)$.

Table 73: The fraction of radiative photon production after different cut values for $\Delta R(\text{jet}, \gamma)$.

Channel	$t\bar{t}$ sel. with $\Delta R(\text{jet}, \gamma) > 0.4$	$t\bar{t}\gamma$ sel. with $\Delta R(\text{jet}, \gamma) > 0.7$	$t\bar{t}\gamma$ sel. with $\Delta R(\text{jet}, \gamma) > 1.0$
dilepton	65%	67%	69%
single lepton	57%	61%	65%

Not reviewed, for internal circulation only

E Photon isolation cut

A comparison before and after applying the photon isolation cut. This cut has a large effect in reducing the main backgorund form hadronic fakes as shown in the tables. The optimised cuts which are described in Section 5 were used.

Table 74: The number of selected data events and the expected yields from MC after applying the photon isolation cut. Normalized to an integrated luminosity of $3.2 fb^{-1}$

Process	$e\mu$	$\mu\mu$	ee
Data	85	35	35
$t\bar{t}\gamma$	38.53 ± 3.19	21.62 ± 2.41	14.96± 1.94
Hadron fake	24.09 ± 2.25	9.36 ± 1.37	8.95 ± 1.34
egamma fake	2.06 ± 0.67	1.32± 0.54	0.01± 0.004
$Z\gamma$ + jets	0.003 ± 0.002	2.16 ± 0.67	2.93± 0.93
Single top + γ	1.70 ± 0.62	1.05± 0.47	0.71 ± 0.33
Diboson + γ	0.12 ±0.12	0.06± 0.03	0.007 ± 0.004

Table 75: The number of selected data events and the expected yields from MC after applying the photon isolation cut, normalised to an integrated luminosity of $3.2 fb^{-1}$. Note: In these numbers table the ph_iso weight is not included.

Process	$e\mu$	$\mu\mu$	ee
Data	50	19	18
$t\bar{t}\gamma$	32.94 ± 2.95	17.12 ± 2.15	12.58 ± 1.84
Hadron fake	1.48 ± 2.25	0.78 ± 0.41	0.94± 0.43
egamma fake	0.36 ± 0.27	0.33± 0.33	0.001 ± 0.0007
$Z\gamma$ + jets	0.003 ± 0.002	1.99± 0.67	2.02± 0.70
Single top + γ	1.22 ± 0.67	0.64± 0.37	0.49± 0.28
Diboson + γ	0.12 ±0.12	0.02± 0.02	0.004 ± 0.004

Not reviewed, for internal circulation only

(a) (b) (c)

(d) (e) (f)

Figure 141: The $p_T cone20$ variable after appying all the optimised selecten which are described in Section 5 and before appyling the photon isolation cut in $e\mu$(a), ee (b) and $\mu\mu$ (c). The $p_T cone20$ variable after appying the same selections described in Section 5 after appyling the phton isolation cut in $e\mu$(a), ee (b) and $\mu\mu$ (c).

Not reviewed, for internal circulation only

F Hadronic-fake background appendix

F.1 Study on origin of hadronic fakes

F.1.1 Processes contribution to the hadronic fake background

In Figures 142 and 143 the hadronic fake origin are shown in e+jets and μ+jets channels, respectively, for the four regions of A, B, C and D (i.e. signal region) and for all the processes that their contribution to hadronic fake background are considered. Here, the truth origin information comes from the `MCTruthClassifier` package, which is the same truth matching package that is used in the definition of true hadronic fake explained in Section 3.2.

One can see that the dominant process is $t\bar{t}$ across all the regions and in both channels, with a contribution of 91%(87%) in region A, 85%(85%) in region B, 89%(89%) in region C and 91%(90%) in signal region in e+jets (μ+jets) channel. The second and third contribution ranks switch between W+jets and single lepton processes across the regions and channels, but with small difference, and together their contributions are 7%(12%) in region A, 13%(13%) in region B, 9%(10%) in region C and 7%(9%) in signal region in e+jets (μ+jets) channel.

The hadronic fakes are dominantly originated from π^0 decays, and a smaller contribution comes from other light mesons decays, and this is consistent across the four regions and the two channels.

Figure 142: Photon truth origin according to the `MCTruthClassifier` package, in the four regions in e+jets channel.

Not reviewed, for internal circulation only

Figure 143: Photon truth origin according to the `MCTruthClassifier` package, in the four regions in μ+jets channel.

F.1.2 Truth matching to the parton level

In this section a study on the origin of hadronic fakes is shown, where the reconstructed photon that is categorised as hadronic fakes is matched to a parton, using the $t\bar{t}$ MC sample. For this, after the full event selection (Section 5.1) is applied and the sample overlap removal to reject the prompt photons and $e \rightarrow \gamma$ fakes (Section 3.2) is performed, the reconstructed photon is matched to the partons within a $\Delta R < 0.2$ distance.

The result of this study are shown in Figure 144 for e+jets and μ+jets channels. The high fraction of hadronic fakes originating from quarks from W boson explains why hadronic fake background in single lepton channel is larger than in dilepton channel. Note that the first bin ("non.") shows the fraction of events where no parton matched to hadronic fakes within $\Delta R < 0.2$, therefore it should contain the events where hadronic fakes are originated from extra gluon jets. [16]

[16] The exact number of gluon jets per event is not known here, since this study is done using a parton-level ntuple

Not reviewed, for internal circulation only

(a) (b)

Figure 144: Partons matched to a hadronic fake in (a) e+jets and (b) μ+jets channels, in $t\bar{t}$ sample. The first bin shows the fraction of events with no parton match is found (within a $\Delta R < 0.2$ distance), therefore it includes the events with hadronic fakes originated from extra jets.

F.2 $t\bar{t}$ MC modelling systematics in hadronic fake control regions

In the calculation of θ_{MC} (Equation10) in 3D bins of photon $p_{\mathrm{T}} - \eta -$ conversion type, the $t\bar{t}$ MC modelling systematics are considered. Figures 145 and 146 show the effect of these modelling uncertainties on the photon p_{T} distributions in the four regions of A, B, C and D (i.e. signal region) in e+jets and μ+jets channels as example, with the relative difference in number of events w.r.t. nominal sample printed on each plot.

Not reviewed, for internal circulation only

Figure 145: $t\bar{t}$ MC modelling systematics for regions A, B and C in e+jets and μ+jets channels. The nominal sample is compared to the Sherpa sample, to estimate the modelling uncertainty of hard scattering generators and parton showering.

Figure 146: $t\bar{t}$ MC modelling systematics for regions A, B and C in e+jets and μ+jets channels. The nominal sample is compared to the samples with higher or lower A14 tunes and different scales and NLO radiation settings, to estimate the extra radiation modelling uncertainty.

Not reviewed, for internal circulation only

F.3 $SF^{h\text{-}fake}$ for e+jets and μ+jets channels

Control plots for regions A, B and C for e+jets and μ+jets channels are shown in Figures 147, 148, 149, 150, 151 and 152.

The measured MC-based double ratio and the resulting inclusive hadronic fake scale factor for e+jets and μ+jets channels are summaries in Tables 76 and 77, respectively. In the analysis, these $SF^{h\text{-}fake}$'s are not used, but the 3D $SF^{h\text{-}fake}$'s calculated for the combined single lepton (see Table 16) are used.

Channel	conv. type	p_T-η bin	θ_{MC} ± total uncertainty	stat. uncertainty	syst. uncertainty		
					rad.-up	rad.-down	gen. & PS
e+jets	all	inclusive	1.41±0.18	±0.07	+0.08(±0.01)	-0.15(±0.02)	-0.07(±0.01)
μ+jets	all	inclusive	1.47±0.40	±0.08	+0.07(±0.01)	-0.12(±0.01)	-0.38(±0.06)

Table 76: Inclusive MC-based double ratio correction factor (θ_{MC}) for e+jets and μ+jets channels. The statistical error of the systematics uncertainties are given in the round brackets in front of them.

Channel	conv. type	$p_T - \eta$ bin	$SF^{h\text{-}fake}$	stat.	syst.				
					$t\bar{t}\gamma$ subtract.	rest subtract.	θ_{MC} stat.	θ_{MC} rad.-up/-down	θ_{MC} gen. & PS
e+jets	all	inclusive	1.63	±0.11	±0.18	±0.15	±0.08	+0.09/-0.17	±0.08
μ+jets	all	inclusive	1.57	±0.10	±0.19	±0.10	±0.09	+0.07/-0.13	±0.41

Table 77: Inclusive hadronic fake scale factor for e+jets and μ+jets channels.

Not reviewed, for internal circulation only

Figure 147: Plots of region A in e+jets channel, where photons are identified as fail-tight and are isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes in Section 6.2.1.

Not reviewed, for internal circulation only

Figure 148: Plots of region A in μ+jets channel, where photons are identified as fail-tight and are isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes in Section 6.2.1.

Not reviewed, for internal circulation only

Figure 149: Plots of region B in e+jets channel, where photons are identified as fail-tight and are non-isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes in Section 6.2.1.

Not reviewed, for internal circulation only

Figure 150: Plots of region B in μ+jets channel, where photons are identified as fail-tight and are non-isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes in Section 6.2.1.

Not reviewed, for internal circulation only

Figure 151: Plots of region C in e+jets channel, where photons are identified as tight and are non-isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes in Section 6.2.1.

Not reviewed, for internal circulation only

Figure 152: Plots of region C in μ+jets channel, where photons are identified as tight and are non-isolated. In each plot, the dashed area shows the total uncertainty of the total expectation and the solid band in the ratio panels show the statistical uncertainty of the total expectation. The systematic uncertainty only includes the $t\bar{t}$ MC modelling systematics and those systematics that are explained in the part about subtraction of non-hadronic-fakes in Section 6.2.1.

Not reviewed, for internal circulation only

F.4 Data-driven double ratio

The double ratio from Equation 10 is additionally calculated by a data-driven method.

One can not use the same A, B, C and D regions in Equation 10 to calculate the data-driven double ratio, as the number of hadronic fakes in region D (signal region) in data is the unknown. However, an equivalent double ratio, $\theta'_{\text{d.d.}}$, can be calculated from four new control regions, A′, B′, C′ and D′, which are derived by dividing regions B and C after the isolation gap $p_T^{\text{cone20}} > 3$ GeV is removed. A schematic representation from these control regions are illustrated in Figure 153.

Figure 153: An illustration of the regions A′, B′, C′ and D′ which are used to calculate the $\theta'_{\text{d.d.}}$ and are derived from dividing regions B and C after the isolation gap is removed.

Then, the data-driven double ratio is calculated as:

$$\theta'_{\text{d.d.}} = \frac{N^{\text{h-fake}}_{\text{D}',\text{data}} / N^{\text{h-fake}}_{\text{C}',\text{data}}}{N^{\text{h-fake}}_{\text{A}',\text{data}} / N^{\text{h-fake}}_{\text{B}',\text{data}}}, \tag{67}$$

where same as in Equation 12, the non-hadronic-fakes and QCD background are subtracted from data.

To obtain the A′, B′, C′ and D′ regions, the p_T^{cone20} distribution of regions B and C after the removal of the isolation gap is used. These distributions can be seen in Figure 154 for inclusive photon p_{T}, η and conversion type.

The boundaries of A′, B′, C′ and D′ are varied, by setting different cuts on the p_T^{cone20}. The lower boundary of A′ and D′ are varied by three different cuts on p_T^{cone20} at 1, 2 and 3 GeV. The upper boundary of A′ and D′, which is the same as the lower boundary of B′ and C′, are varied by nine different cuts on p_T^{cone20} at 7, 8, 9, 10, 11, 12, 13 , 14 and 15 GeV. Regions B′ and C′ don't have an upper boundary.

The resulting $\theta'_{\text{d.d.}}$ are shown in Figure 155(a) for inclusive photon p_{T}, η and conversion type, in Figure 155(b) for converted and in Figure 155(c) for unconverted photons. Also, it is shown in Figures 156 and 157 for different photon p_{T} and η ranges. In all the above figures the $\theta'_{\text{d.d.}}$ is compared to the equivalent MC derived θ'_{MC} which is calculated from same A′, B′, C′ and D′ regions, only by replacing the data by MC in Equation 67. The systematic uncertainty of $\theta'_{\text{d.d.}}$ is due to the subtraction of non-hadronic-fakes and QCD background from data and the systematic uncertainty of θ'_{MC} comes from the $t\bar{t}$ MC sample modelling.

Not reviewed, for internal circulation only

(a) (b)

Figure 154: The p_T^{cone20} distribution of (a) region B without the isolation gap and (b) region C without the isolation gap.

The measured $\theta'_{\text{d.d.}}$ is not distributed flat against the different choices boundaries of A′, B′, C′ and D′ regions. Therefore, the nominal $\theta'_{\text{d.d.}}$ is calculated from $1 < p_T^{\text{cone20}} < 12$ GeV as A′ and D′ and $p_T^{\text{cone20}} > 12$ GeV as B′ and C′, and then the difference of minimum and maximum of $\theta'_{\text{d.d.}}$ in the 27 different choices of region definitions is taken as an extra systematic uncertainty.

Using $\theta'_{\text{d.d.}}$, the estimated number of hadronic fake in signal region is calculated as

$$N_{\text{est.}}^{\text{h-fake}} = \frac{N_{\text{A, data}}^{\text{h-fake}} \cdot N_{\text{C,data}}^{\text{h-fake}}}{N_{\text{B,data}}^{\text{h-fake}}} \times \theta'_{\text{d.d.}} \tag{68}$$

and then the $\text{SF}^{\text{h-fake}}$ is calculated from Equation 13. The resulting $\text{SF}^{\text{h-fake}}$ using $\theta'_{\text{d.d.}}$ is compared with the calculated $\text{SF}^{\text{h-fake}}$ using θ_{MC}, which is the baseline method, in Figure 158.

Within the uncertainties, the two methods agree in most bins. However, the θ_{MC} is chosen as the baseline method, for several reasons: 1) Extending the A′ and D′ are to the low isolation values ($p_T^{\text{cone20}} > 1$ GeV) is needed to make a fair extrapolation of isolated vs non-isolated regions (A′ and D′ vs B′ and C′) but in the other hand it introduces prompt photon contamination. 2) In the final fit the systematics due to background subtraction in $\theta'_{\text{d.d.}}$ and in $N_{\text{est.}}^{\text{h-fake}}$ must be correlated and this is technically difficult.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 155: $\theta'_{\text{d.d.}}$ for different choices of A′, B′, C′ and D′ regions for (a) inclusive (b) converted and (c) unconverted photons, compared to MC. The x axis shows the boundaries of A′ and D′ region which are defined by cuts on p_T^{cone20} in GeV. The lower boundaries of B′ and C′ are identical as upper boundaries of A′ and D′ and they go up to infinity in p_T^{cone20}.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 156: $\theta'_{\text{d.d.}}$ for different choices of A′, B′, C′ and D′ regions for different photon p_{T} ranges, compared to MC. The x axis shows the boundaries of A′ and D′ region which are defined by cuts on p_T^{cone20} in GeV. The lower boundaries of B′ and C′ are identical as upper boundaries of A′ and D′ and they go up to infinity in p_T^{cone20}.

Not reviewed, for internal circulation only

(a)

(b)

(c)

Figure 157: $\theta'_{\text{d.d.}}$ for different choices of A′, B′, C′ and D′ regions for different photon η ranges, compared to MC. The x axis shows the boundaries of A′ and D′ region which are defined by cuts on p_T^{cone20} in GeV. The lower boundaries of B′ and C′ are identical as upper boundaries of A′ and D′ and they go up to infinity in p_T^{cone20}.

Not reviewed, for internal circulation only

Figure 158: $\mathrm{SF}^{\text{h-fake}}$ calculated by using $\theta'_{\text{d.d.}}$ (cross-check method) compared to $\mathrm{SF}^{\text{h-fake}}$ calculated by using θ_{MC} (baseline method). The error bars show the total uncertainty where the thick lines are only statistical uncertainty.

F.5 Red/blue plots for hadronic fake systematic uncertainties

F.5.1 Uncertainty due to the subtraction of $t\bar{t}\gamma$

As referenced in Section 6.2, Equation 14, a 12% anti-correlation between the hadronic fake background and the signal is considered. The number is motivated by Figure 159.

Not reviewed, for internal circulation only

Figure 159: The NP for the the subtraction of signal in the calculation of $SF^{\text{h-fake}}$ ("$t\bar{t}\gamma$ subtract." in Table 16) and the 1σ up and down variations.

F.5.2 Uncertainty due to the statistical uncertainties of the data-driven part or of the θ_{MC} factor

The systematic uncertainties due to the statistical uncertainty of $SF^{\text{h-fake}}$ are presented here. This includes two components: statistical uncertainty of data-driven part of $SF^{\text{h-fake}}$ ("stat." in Table 16), and statistical uncertainty of the θ_{MC} factor ("θ_{MC} stat." in Table 16).

Figure 160 shows the red/blue plot for the case that only 1 NP for the "θ_{MC} stat." is considered. This is not used in the final fit and is just shown here for the comparison, since this way the correlation between the bins are not considered.

Figure 160: The 1 NP for the statistical uncertainty of the θ_{MC} factor ("θ_{MC} stat." in Table 16) and the 1σ up and down variations. This is not used in the fit, instead it is replaced by the 10 NP's shown in Figures 161 and162.

As mentioned in in Section 6.2, actually one NP is considered for "θ_{MC} stat." of each bin of $SF^{\text{h-fake}}$, resulting to 10 NP's in total, which are shown in Figure 161 for converted photons, and in Figure 162 for unconverted photons.

Not reviewed, for internal circulation only

Figure 161: The 5 NP's for the θ_{MC} statistical components of the $\mathrm{SF}^{\text{h-fake}}$ uncertainty for converted photons ("θ_{MC} stat." in Table 16) and their 1σ up and down variations.

Figure 162: The 5 NP's for the θ_{MC} statistical components of the $\mathrm{SF}^{\text{h-fake}}$ uncertainty for unconverted photons ("θ_{MC} stat." in Table 16) and their 1σ up and down variations.

Not reviewed, for internal circulation only

The statistical uncertainties associated with the data-driven part of $SF^{\text{h-fake}}$ ("stat." in Table 16) are shown in Figure 163 for converted photons and in Figure 164 for unconverted photons. These 10 NP's are used in the final fit as well.

Figure 163: The 5 NP's associated with the statistical uncertainty of the data-driven part of $SF^{\text{h-fake}}$ for the converted photons ("stat." in Table 16) and their 1σ up and down variations.

Figure 164: The 5 NP's associated with the statistical uncertainty of the data-driven part of $SF^{\text{h-fake}}$ for the unconverted photons ("stat." in Table 16) and their 1σ up and down variations.

Not reviewed, for internal circulation only

G $Z\gamma$ validation regions

G.1 $Z\gamma$ heavy jet validation region

Figure 165: Pre and post-fit distributions for the *ee* $Z\gamma$ validation region, where number of b-tagged jets is 1. Only statistical uncertainties are shown.

G.2 $Z\gamma$ light jet validation region

Not reviewed, for internal circulation only

Figure 166: Pre and post-fit distributions for the $\mu\mu$ $Z\gamma$ validation region, where number of b-tagged jets is 1. Only statistical uncertainties are shown.

Not reviewed, for internal circulation only

Figure 167: Pre and post-fit distributions for the ee $Z\gamma$ validation region, where number of b-tagged jets is 0. Only statistical uncertainties are shown.

Not reviewed, for internal circulation only

Figure 168: Pre and post-fit distributions for the $\mu\mu$ $Z\gamma$ validation region, where number of b-tagged jets is 0. Only statistical uncertainties are shown.

Not reviewed, for internal circulation only

H $W\gamma$ validation regions

H.1 $W\gamma$ heavy jet validation region

Figure 169: Pre-fit distributions for the e+jets $W\gamma$ validation region, where number of b-tagged jets is 1. Only statistical uncertainties are shown.

H.2 $W\gamma$ light jet validation region

Not reviewed, for internal circulation only

Figure 170: Pre-fit distributions for the μ+jets $W\gamma$ validation region, where number of b-tagged jets is 1. Only statistical uncertainties are shown.

Not reviewed, for internal circulation only

Figure 171: Pre-fit distributions for the e+jets $W\gamma$ validation region, where number of b-tagged jets is 0. Only statistical uncertainties are shown.

Not reviewed, for internal circulation only

Figure 172: Pre-fit distributions for the μ+jets $W\gamma$ validation region, where number of b-tagged jets is 0. Only statistical uncertainties are shown.

Not reviewed, for internal circulation only

I ELD binning optimisation

Different Asimov fit scenarios are tested on the single lepton and dilepton channels to determine the best bin sizes for the ELD. All systematics are included in these scenarios. The first scenario uses a uniform binning distribution with ten and 6 bins for the single lepton and dilepton channels, respectively. Another scenario uses an automatic binning algorithm called “TransformD" (available in TRExFitter). The algorithm scans the distribution and merges bins from the right until a certain fraction of signal and background remains in the bin. For each bin, a threshold for this merging is defined as

$$Z = z_b \frac{n_b}{N_b} + z_s \frac{n_s}{N_s}, \tag{69}$$

where n_s and n_b is the signal and background content respectively. N_s and N_b are the total number of signal and background events. This function takes two inputs, z_s and z_b, which define the maximum allowed fraction of signal and background events in each bin for $z_s + z_b$ = Number of bins. The algorithm stops iterating when $Z > 1$.

The final scenario is a “by-eye" correction to the AutoBin result with the main motivation to improve aesthetics.

For the single lepton channel small differences are seen between each scenario. However, the AutoBin 2 binning algorithm shows the most sensitivity and so will be used. This consists of 15 bins arranged as [0.00,0.16,0.25,0.30,0.36,0.41,0.46,0.50,0.55,0.60,0.64,0.68,0.73,0.77,0.81,1.00]. For simplicity, this binning is also applied to the individual single lepton and μ+jets channels. For the dilepton channel, similarly, small differences are seen. The AutoBin 2 algorithm performs slightly better, however with fewer events as is seen in the ee, $\mu\mu$ and $e\mu$ channels it seems prudent to limit the number of bins. For simplicity sake, the uniform binning with 6 bins is chosen for all dilepton channels. The results for each of the scenarios are shown in Table 78.

Channel	Binning	“+" error	“-" error
single lepton	Uniform	0.0934	0.0915
	AutoBin 1 ($z_s = 5, z_b = 5$)	0.0923	0.0897
	AutoBin 2 ($z_s = 12, z_b = 3$)	**0.0911**	**0.0888**
	AutoBin 2 tweaked	0.0929	0.0914
dilepton	**Uniform**	0.0678	0.0629
	AutoBin 1 ($z_s = 5, z_b = 5$)	0.0663	0.0626
	AutoBin 2 ($z_s = 12, z_b = 3$)	0.0652	0.0606
	AutoBin 2 tweaked	**0.0652**	**0.0615**

Table 78: The different Asimov fit scenarios for the single lepton and dilepton channels. All systematics are included in the fit. The error represents the absolute up/down error of μ. The chosen binning is highlighted in bold text.

Not reviewed, for internal circulation only

J 100 PDF shape uncertainties

As detailed in Chapter 10, a PDF shape uncertainty on the signal needs to be accounted for. The acceptance effect is not explored here (see Section 7). Rather, under consideration is the impact of the shapes in the final fit. For each fit, the only systematic contribution is from 1 of the 100 PDF variations. A full fit is then performed in each of the 7 channels; single lepton, dilepton, e+jets, μ+jets, ee, $\mu\mu$ and $e\mu$. The results of μ are calculated and plotted in Figures 173 and 174, with the standard deviation printed on each plot. The relative error on μ is shown in Table 30. These show that the impact from the 100 PDF shape variations are negligible in all channels and can be disregarded in the fit.

(a) single lepton channel (b) dilepton channel

Figure 173: The 100 fit results for each PDF variation for the signal sample. Shown for the single lepton and dilepton channels.

Not reviewed, for internal circulation only

(a) e+jets channel

(b) μ+jets channel

(c) ee channel

(d) $\mu\mu$ channel

(e) $e\mu$ channel

Figure 174: The 100 fit results for each PDF variation for the signal sample. Shown for the five channels.

Not reviewed, for internal circulation only

K Varying the single top, diboson and $t\bar{t}V$ systematic uncertainty

There is currently a 50% normalisation uncertainty applied to each of the single top, diboson and $t\bar{t}V$ backgrounds This section checks the impact the systematic uncertainty has for these prompt backgrounds, grouped as Other prompt. For the single lepton and dilepton channel two Asimov fits are performed with all systematic sources included. In the first a 50% uncertainty is assigned to each of these three backgrounds. In the second fit the systematic uncertainty is doubled, i.e. a 100% uncertainty is assigned. The results for the fits are shown in Table 79. Doubling the uncertainty on the three backgrounds leads to an absolute increase in the error of $\approx 0.2\%$ for both the single lepton and dilepton channels. This impact is negligible and means these backgrounds do not play an important role in the fit. For this reason, the normalisation uncertainty is left at 50%.

Channel	Normalisation uncertainty	“+" error	“-" error
single lepton	50%	0.0929	0.0914
	100%	0.0945	0.0927
dilepton	50%	0.0678	0.0629
	100%	0.0697	0.0647

Table 79: Two different Asimov fit scenarios for the single lepton and dilepton channels where the normalisation uncertainty on the Other prompt backgrounds has been varied. All systematics are included in the fit. The error represents the absolute up/down error of μ.

Not reviewed, for internal circulation only

L Systematic uncertainty breakdown

The size of the shifted $t\bar{t}\gamma$ reconstructed distribution due to the luminosity estimation, the choice of the parton shower generator of the signal sample and the pile-up scale factor, is compared to the nominal distribution and shown in Figure 175 and Figure 176 in bins of the reconstructed object. It can be seen that the number of events in the shifted samples is mostly the same as the nominal sample, and therefore the majority of differences due to systematic shifts are expected to be significant.

Figure 175: Reconstructed-level distributions of the size of the relative systematic shifts due to signal modelling (left), luminosity (centre) and pile-up (right), in bins of p_{T} (first row), $|\eta|$ (second row), minimum $\Delta R(\ell, \ell)$ (third row) in the single-lepton channel.

Not reviewed, for internal circulation only

Figure 176: Reconstructed-level distributions of the size of the relative systematic shifts due to signal modelling (left), luminosity (centre) and pile-up (right), in bins of p_{T} (first row), $|\eta|$ (second row), minimum $\Delta R(\ell,\ell)$ (third row), $\Delta\eta(\ell,\ell)$ (fourth row) and $\Delta\phi(\ell,\ell)$ in the dilepton channel.

Not reviewed, for internal circulation only

M Fits with symmetrisation off

In a hypothetical scenario, what happens if we turn off all the symmetrisation options for the various systematics that get symmetrised? Table 80 shows the asimov and full fit results for such a situation.

Channel		μ	+	-
single lepton Asimov	Symm. On	1.000	0.096	0.093
	Symm. Off	1.000	0.096	0.094
single lepton Full	Symm. On	1.067	0.093	0.089
	Symm. Off	1.068	0.093	0.089
dilepton Asimov	Symm. On	1.000	0.072	0.068
	Symm. Off	1.000	0.072	0.067
dilepton Full	Symm. On	1.097	0.079	0.072
	Symm. Off	1.093	0.081	0.074

Table 80: Asimov and full fit results comparing the nominal scenario (symmetrisation on in most cases), to when symmetrisation is switched off for all systematic uncertainties.

We can see that there is a small and negligible difference between symmetrisation on and off. Figure 177 and 178 shows the pull plots for the dilepton and single lepton fits, respectively. There are a few systematics which are constrained or pulled, but not alarmingly so (with the exception of three JET NP21 NPs). Given that this is a hypothetical study these will not be examined further. Thus, the majority of systematic uncertainties are symmetrised. The exceptions occurs for most modelling systematics which are is explained more in Chapter 10.

Not reviewed, for internal circulation only

(a) Asimov Symm. Off (b) Full Symm. Off (c) Full Symm. On

Figure 177: Pull plots for the dilepton channel for the different scenarios when switching symmetrisation on and off.

Not reviewed, for internal circulation only

(a) Asimov Symm. Off (b) Full Symm. Off (c) Full Symm. On

Figure 178: Pull plots for the single lepton channel for the different scenarios when switching symmetrisation on and off.

Not reviewed, for internal circulation only

N Unfolding checks

N.1 Comparison of different number of iterations

The closure, pull and stress tests are repeated using 6 iterations and compared to three iterations in Figure 179. No significant difference is observed. The measured differential cross sections are obtained after three and six iterations in the photon p_{T} bins are compared in Table 81

Table 81: The differential cross section in $p_{\mathrm{T}}(\gamma)$ bins, obtained using the iterative Bayesian method, in the dilepton channel. The first and second associated uncertainties are the statistical and the systematic uncertainty, respectively.

20.0 - 35.0	3 iter.	30.14 ± 2.33 ± 1.58
	6 iter.	30.15 ± 2.34 ± 1.58
35.0 - 50.0	3 iter	13.63 ± 1.68 ± 0.93
	6 iter.	13.63 ± 1.71 ± 0.94
50.0 - 65.0	3 iter.	7.15 ± 1.19 ± 0.44
	6 iter.	7.15 ± 1.21 ± 0.44
65.0 - 80.0	3 iter.	4.61 ± 0.77 ± 0.26
	6 iter.	4.62 ± 0.79 ± 0.26
80.0 - 95.0	3 iter.	2.52 ± 0.40 ± 0.23
	6 iter.	2.50 ± 0.40 ± 0.23
95.0 - 110.0	3 iter.	2.38 ± 0.62 ± 0.30
	6 iter.	2.39 ± 0.63 ± 0.31
110.0 - 140.0	3 iter.	2.57 ± 0.45 ± 0.15
	6 iter.	2.56 ± 0.46 ± 0.15
140.0 - 180.0	3 iter.	2.42 ± 0.52 ± 0.17
	6 iter.	2.42 ± 0.53 ± 0.17
180.0 - 300.0	3 iter.	4.16 ± 0.83 ± 0.23
	6 iter.	4.16 ± 0.83 ± 0.23

N.2 Pre-fit versus post-fit total error for correlated and uncorrelated scenarios

A study is carried out to examine the size of the systematic error in three scenarios. The first scenario looks at the background and signal distributions for pre-fit total errors. This should give the highest contribution to the total error as the individual contributions are simply added in quadrature before any optimisations have been done by the fit. This refers to the *uncorrelated pre-fit*. The second two scenarios looks at post-fit histograms. The first of which sums the contributions in quadrature as was done for the pre-fit histograms, this is *uncorrelated post-fit*. The second takes the error as calculated by the fit, which includes the correlation. Thus, this is *correlated post-fit*.

The plots for these scenarios are shown in Figure 180 and 181 for single lepton and dilepton, respectively. It is expected that the uncorrelated pre-fit uncertainties will be largest. What is surprising is that they are not much larger than uncorrelated post-fit histograms. As expected, uncorrelated methods show significantly higher uncertainties compared to the correlated post-fit uncertainties.

Not reviewed, for internal circulation only

Figure 179: The closure (top), pull (centre) and the stress (bottom) test obtained after three iterations (left) and six iterations (right).

Not reviewed, for internal circulation only

(a) Background and signal for $\Delta R(\gamma, l)$.

(b) Background and signal for the ELD.

(c) Background and signal for $\eta(\gamma)$.

Not reviewed, for internal circulation only

(d) Background and signal for $p_T(\gamma)$.

Figure 180: Plots comparing pre-fit uncorrelated errors, post-fit uncorrelated and post-fit correlated errors for the single lepton channel.

Not reviewed, for internal circulation only

(a) Background and signal for $\Delta\eta(l,l)$.

(b) Background and signal for $\Delta\phi(l,l)$.

(c) Background and signal for $\Delta R(\gamma,l)$.

Not reviewed, for internal circulation only

(d) Background and signal for the ELD

(e) Background and signal for $\eta(\gamma)$.

(f) Background and signal for $p_T(\gamma)$.

Figure 181: Plots comparing pre-fit uncorrelated errors, post-fit uncorrelated and post-fit correlated errors for the dilepton channel.

Not reviewed, for internal circulation only

N.3 Pre-fit versus post-fit background shape comparisons

A study is carried out to examine the difference in the shape of backgrounds from pre-fit versus post-fit histograms. This is to determine the effect the fit has on the overall shape of backgrounds once optimisations have been done. For the single lepton channel three “types" of backgrounds are considered: the pre-fit background, the post-fit background where a StatOnly fit has been performed but with the $W\gamma$ scale factor floating, and the post-fit background. For the dilepton channel, two “types" of backgrounds are considered: the pre-fit background, and the post-fit background.

The results are shown in Figure 182 and 183 for the single lepton and dilepton channels, respectively. Of note is that the pre-fit background uncertainty band only includes statistical uncertainty while the post-fit bands include the total uncertainties. Since only the shape is being considered in these plots, this is of little importance. One can see that in all distributions and for both channels the overall shape varies negligibly and well within total post-fit uncertainties.

Figure 182: Plots comparing pre-fit and post-fit background shapes for the ELD and differential variables in the single lepton channel.

Not reviewed, for internal circulation only

Figure 183: Plots comparing pre-fit and post-fit background shapes for the ELD and differential variables in the dilepton channel.

Not reviewed, for internal circulation only

N.4 Pre-fit unfolding inputs

The measured signal distribution, which is obtained from data after subtracting the pre-fit backgrounds is unfolded. The $W+\gamma$ background is obtained from the statistical only fit, in the single-lepton channel, since its a floating background. The pre-fit shifted systematic uncertainties are unfolded and shown as uncertainty band. The unfolded results are shown in Figure 190 and Figure 192 in the single-lepton and dilepton channels, respectively. These results can be compared to the results in Figure 115 and Figure 116, where post-fit inputs are used and unfolded. It can be seen that no significant differences are observed in terms of the shapes, and only larger systematic uncertainty band is observed in the pre-fit distributions. This is mainly due to the constrained systematics by the fit, in particular the signal parton showering and $t\bar{t}$ Sherpa modelling, and the $Z+\gamma$ modelling in the dilepton channel.

The effect of pre-fit systematics on the measured cross sections are shown in Table 82 and Table 83 in the single-lepton and dilepton channels, in photon p_{T} bins, which can be compared to the post-fit systematics in Table **??** and Table **??**. The breakdown of the main pref-fit background uncertainties is shown in Table 84 and Table 85, in the single-lepton and dilepton channels, respectively. The shape of the leading pre-fit uncertainty after unfolding, are compared to the nominal unfolded result, and shown in Figure 186 and Figure 187 for all observables. More leading uncertainties in photon p_{T} bins are shown in Figure 188 and Figure 189, in the single-lepton and dilepton channels, respectively. The same comparison is performed for the post-fit inputs and shown in the same Figure. It can be seen that the same deviations in the pre-fit and post-fit results are obtained.

The impact of pre-fit systematic uncertainties on the measured cross section for the remaining observables are shown in Tables 86 – 90 in the single-lepton channel.

Not reviewed, for internal circulation only

Figure 184: The absolute (right) and normalised (left) differential cross section, using pre-fit inputs, in p_{T} bins (top), $|\eta|$ bins (centre) and min $\Delta R(\mathrm{lepton}, \gamma)$ (bottom), in the single-lepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 185: The absolute (right) and normalised (left) differential cross section, using pre-fit inputs, in p_{T} bins (first row), $|\eta|$ bins (second row), minimum $\Delta R(\mathrm{lepton}, \gamma)$ (third row), $\Delta\eta(\mathrm{lepton}, \mathrm{lepton})$ (fourth row), and $\Delta\phi(\mathrm{lepton}, \mathrm{lepton})$ (fifth row) in the dilepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Table 82: Summary of the sources of uncertainty on the absolute differential cross section for $p_T(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Pre-fit inputs are used. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_T(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
Signal Scale Variation	±1.15	±0.64	±0.24	±0.26	±0.7	±0.95	±1.69	±2.5	±4.17
Signal ISR/FSR Setting	±0.41	±1.44	±0.43	±1.25	±1.7	±2.03	±2.21	±3.82	±0.41
Signal Parton Showering	±6.21	±0.43	±2.51	±2.74	±3.04	±11.7	±4.58	±4.24	±14.04
Background Modelling	±16.17	±14.61	±16.36	±25.37	±29.39	±27.23	±45.38	±26.38	±10.13
Luminosity	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97
Pileup Effects	±2.52	±1.9	±1.7	±2.46	±0.95	±1.79	±1.96	±1.79	±1.22
Photon Scale Factors	±2.13	±1.47	±1.16	±0.92	±0.88	±0.89	±0.97	±1.83	±2.12
Lepton Identification and Reconstruction	±1.13	±1.16	±1.15	±1.13	±1.16	±1.12	±1.13	±1.17	±1.19
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±5.88	±5.43	±5.51	±5.11	±4.99	±4.72	±4.68	±3.92	±2.97
b-tagging	±1.67	±1.7	±1.69	±1.62	±1.7	±1.52	±1.57	±1.64	±1.74
Missing Transverse Momentum Reconstruction	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
PPT	±1.71	±0.5	±5.04	±0.69	±7.32	±2.14	±1.63	±2.31	±3.28
Total systematics	±19	±16	±19	±27	±31	±31	±33	±28	±20
Data statistics	±5	±6	±7	±9	±11	±14	±11	±12	±8
Total uncertainty	±20	±17	±20	±28	±33	±34	±34	±30	±21

Table 83: Summary of the sources of uncertainty on the absolute differential cross section for $p_T(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Pre-fit inputs are used. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_T(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
Signal Scale Variation	±1.16	±0.57	±0.18	±0.24	±0.42	±1.13	±1.52	±2.11	±3.65
Signal ISR/FSR Setting	±1.04	±5.34	±5.41	±2.99	±7.46	±3.6	±5.12	±6.13	±0.0
Signal Parton Showering	±1.79	±3.9	±0.97	±2.5	±13.93	±27.15	±0.89	±7.96	±6.82
Background Modelling	±10.2	±23.65	±19.86	±4.22	±1.67	±6.54	±6.2	±4.32	±3.85
Luminosity	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97
Pileup Effects	±3.52	±1.55	±0.86	±4.0	±2.06	±5.35	±3.81	±3.57	±1.36
Photon Scale Factors	±2.04	±1.47	±1.2	±0.97	±0.96	±0.91	±1.03	±1.8	±2.07
Lepton Identification and Reconstruction	±1.65	±1.65	±1.66	±1.64	±1.74	±1.74	±1.74	±1.68	±1.89
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±2.2	±2.06	±2.44	±2.29	±0.99	±1.43	±1.62	±1.05	±0.61
b-tagging	±1.43	±1.55	±1.38	±1.61	±1.62	±1.98	±1.6	±1.33	±1.25
Missing Transverse Momentum Reconstruction	±0.06	±0.35	±0.29	±0.35	±0.36	±0.51	±0.53	±0.55	±0.54
Total systematics	±12	±25	±21	±8	±17	±28	±10	±12	±10
Data statistics	±8	±14	±18	±18	±18	±26	±19	±24	±20
Total uncertainty	±15	±29	±28	±20	±25	±40	±22	±27	±22

Table 84: Summary of the main sources of uncertainty on the absolute differential cross section for $p_T(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Pre-fit inputs are used. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_T(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
ttbar Sherpa Modelling/Hadronic Fake	±7.58	±2.82	±0.28	±1.96	±0.91	±0.18	±1.86	±0.76	±1.12
ttbar Sherpa/Electron Fake	±0.01	±4.81	±7.02	±10.82	±16.62	±14.35	±5.08	±12.81	±0.6
ttbar Sherpa/Hadronic Fake	±0.46	±3.73	±4.04	±5.57	±0.8	±3.61	±21.87	±6.36	±3.84
Hadronic-fake SF Background Subtraction	±5.36	±4.5	±3.2	±2.78	±2.39	±2.76	±1.59	±0.68	±0.85
Hadronic-fake SF thetaMC	±6.42	±3.43	±2.19	±2.24	±1.78	±2.02	±1.33	±0.35	±0.68
Elecrton-fake signal function	±3.81	±0.8	±0.14	±8.48	±7.53	±6.69	±4.15	±3.32	±1.45
Z+photon Modelling	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0

Table 85: Summary of the sources of uncertainty on the absolute differential cross section for $p_T(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the dilepton channel. Pre-fit inputs are used. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_T(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
Z+photon Modelling	±9.57	±23.56	±19.67	±2.94	±1.3	±5.88	±6.14	±3.79	±3.5
ttbar Sherpa Modelling/Hadronic Fake	±1.39	±0.85	±0.39	±1.12	±0.15	±1.26	±0.03	±0.9	±0.6
Hadronic-fake SF Background Subtraction	±1.86	±1.13	±0.53	±1.5	±0.2	±1.68	±0.04	±1.2	±0.81
Hadronic-fake SF thetaMC	±1.6	±0.98	±0.46	±1.29	±0.17	±1.45	±0.04	±1.04	±0.7

Not reviewed, for internal circulation only

Figure 186: The unfolded $t\bar{t}$ Sherpa modelling, affecting the electron-fake background, compared to the nominal unfolded, using post-fit (left) inputs and pre-fit (right) inputs, p_{T} bins (top), $|\eta|$ bins (middle), and min $\Delta R(lepton, \gamma)$ (bottom), in the single-lepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 187: The unfolded $Z + \gamma$ modelling systematic uncertainty compared to the nominal unfolded, using post-fit (left) inputs and pre-fit (right) inputs, p_{T} bins (first row), in $|\eta|$ bins (second row), minimum $\Delta R(\mathrm{lepton}, \gamma)$ (third row), $\Delta\eta(\mathrm{lepton}, \mathrm{lepton})$ (fourth row), and $\Delta\phi(\mathrm{lepton}, \mathrm{lepton})$ bins in the single-lepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 188: The unfolded up-shifted systematics compared to the nominal unfolded, using post-fit (left) inputs and pre-fit (right) inputs, in p_{T} bins in the single-lepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 189: The unfolded up-shifted systematics compared to the nominal unfolded, using post-fit (left) inputs and pre-fit (right) inputs, in p_{T} bins in the single-lepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Table 86: Summary of the sources of uncertainty on the normalised differential cross section for $p_{\mathrm{T}}(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$p_{\mathrm{T}}(\gamma)$ [Gev]	20.0 - 35.0	35.0 - 50.0	50.0 - 65.0	65.0 - 80.0	80.0 - 95.0	95.0 - 110.0	110.0 - 140.0	140.0 - 180.0	180.0 - 300.0
Source	systematic uncertainty (%)								
Signal Scale Variation	±0.99	±0.48	±0.12	±0.38	±0.84	±1.08	±1.83	±2.65	±4.32
Signal ISR/FSR Setting	±0.42	±1.43	±0.42	±1.27	±1.68	±2.02	±2.23	±3.83	±0.43
Signal Parton Showering	±5.56	±0.23	±3.16	±3.38	±3.68	±12.32	±5.21	±4.88	±14.64
Background Modelling	±7.49	±12.55	±15.19	±22.99	±28.34	±25.35	±45.49	±27.5	±10.47
Luminosity	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Pileup Effects	±0.43	±0.2	±0.38	±0.38	±1.13	±0.3	±0.13	±0.3	±0.86
Photon Scale Factors	±0.78	±0.54	±0.61	±0.85	±0.87	±0.8	±0.72	±0.6	±0.81
Lepton Identification and Reconstruction	±0.04	±0.04	±0.02	±0.04	±0.04	±0.09	±0.1	±0.07	±0.21
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±0.52	±0.06	±0.14	±0.26	±0.38	±0.63	±0.67	±1.41	±2.32
b-tagging	±0.11	±0.21	±0.21	±0.28	±0.22	±0.41	±0.42	±0.42	±0.54
Missing Transverse Momentum Reconstruction	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
PPT	±1.62	±0.41	±5.15	±0.59	±7.42	±2.24	±1.74	±2.22	±3.19
Total systematics	±9	±13	±17	±24	±30	±29	±33	±29	±20
Data statistics	±4	±5	±7	±9	±10	±13	±10	±11	±8
Total uncertainty	±10	±14	±18	±25	±31	±32	±34	±31	±21

Table 87: Summary of the sources of uncertainty on the absolute differential cross section for $|\eta|(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$\|\eta\|(\gamma)$	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.37
Source	systematic uncertainty (%)							
Signal Scale Variation	±0.11	±0.12	±0.16	±0.18	±0.07	±0.16	±0.3	±0.42
Signal ISR/FSR Setting	±0.4	±0.79	±0.93	±1.19	±1.43	±1.24	±0.21	±0.28
Signal Parton Showering	±1.75	±0.65	±1.57	±0.62	±0.41	±7.0	±1.19	±5.76
Background Modelling	±15.53	±17.61	±9.46	±13.24	±10.49	±19.66	±17.97	±22.41
Luminosity	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97
Pileup Effects	±1.68	±3.2	±2.4	±0.6	±2.06	±1.0	±3.49	±1.51
Photon Scale Factors	±1.01	±1.04	±1.05	±1.68	±1.62	±1.63	±1.94	±2.35
Lepton Identification and Reconstruction	±1.15	±1.13	±1.13	±1.14	±1.14	±1.16	±1.15	±1.15
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±5.08	±4.91	±5.27	±5.15	±5.52	±5.19	±5.49	±5.82
b-tagging	±1.62	±1.7	±1.66	±1.66	±1.71	±1.73	±1.63	±1.61
Missing Transverse Momentum Reconstruction	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
PPT	±0.44	±0.9	±0.97	±4.67	±0.77	±0.89	±8.62	±2.38
Total systematics	±17	±19	±12	±16	±13	±22	±21	±24
Data statistics	±6	±7	±7	±10	±8	±8	±7	±9
Total uncertainty	±18	±20	±14	±19	±15	±23	±22	±26

Table 88: Summary of the sources of uncertainty on the normalised differential cross section for $|\eta|(\gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

$\|\eta\|(\gamma)$	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2	1.2 - 1.7	1.7 - 2.37
Source	systematic uncertainty (%)							
Signal Scale Variation	±0.12	±0.13	±0.17	±0.2	±0.08	±0.14	±0.28	±0.4
Signal ISR/FSR Setting	±0.4	±0.79	±0.94	±1.18	±1.43	±1.25	±0.2	±0.29
Signal Parton Showering	±1.59	±0.82	±1.75	±0.8	±0.58	±7.17	±1.03	±5.6
Background Modelling	±14.79	±17.7	±7.29	±7.26	±6.54	±16.21	±12.57	±14.73
Luminosity	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Pileup Effects	±0.44	±1.09	±0.27	±1.51	±0.07	±1.11	±1.38	±0.62
Photon Scale Factors	±0.54	±0.52	±0.51	±0.23	±0.15	±0.14	±0.4	±0.83
Lepton Identification and Reconstruction	±0.03	±0.02	±0.02	±0.01	±0.02	±0.02	±0.02	±0.02
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±0.23	±0.39	±0.04	±0.16	±0.21	±0.12	±0.18	±0.53
b-tagging	±0.14	±0.22	±0.23	±0.29	±0.22	±0.27	±0.12	±0.15
Missing Transverse Momentum Reconstruction	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
PPT	±0.02	±1.31	±1.39	±5.07	±0.32	±1.3	±8.19	±2.79
Total systematics	±15	±18	±8	±9	±7	±18	±15	±16
Data statistics	±6	±6	±6	±10	±7	±8	±7	±9
Total uncertainty	±16	±19	±10	±13	±10	±19	±16	±18

Not reviewed, for internal circulation only

Table 89: Summary of the sources of uncertainty on the absolute differential cross section for min $\Delta R(\text{lepton}, \gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

min $\Delta R(\text{lepton}, \gamma)$	1.0 - 1.2	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.2	2.2 - 2.4	2.4 - 2.6	2.6 - 6.0
Source	systematic uncertainty (%)								
Signal Scale Variation	±0.37	±0.41	±0.33	±0.48	±0.35	±0.1	±0.19	±0.14	±0.51
Signal ISR/FSR Setting	±0.98	±1.2	±2.51	±0.77	±0.16	±0.28	±1.14	±1.28	±0.21
Signal Parton Showering	±4.16	±2.71	±0.16	±5.51	±0.77	±0.61	±1.39	±9.26	±0.83
Background Modelling	±12.63	±10.14	±14.89	±29.59	±18.6	±22.73	±13.44	±18.69	±12.29
Luminosity	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97	±2.97
Pileup Effects	±2.08	±1.73	±2.52	±2.62	±3.71	±2.8	±1.21	±2.55	±1.47
Photon Scale Factors	±1.51	±1.53	±1.52	±1.52	±1.5	±1.48	±1.51	±1.49	±1.54
Lepton Identification and Reconstruction	±1.1	±1.12	±1.13	±1.12	±1.15	±1.15	±1.14	±1.16	±1.16
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±4.95	±4.67	±5.02	±4.66	±4.93	±4.85	±5.17	±5.61	±5.81
b-tagging	±1.83	±1.62	±1.67	±1.68	±1.56	±1.62	±1.65	±1.7	±1.67
Missing Transverse Momentum Reconstruction	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
PPT	±0.53	±0.65	±0.21	±0.48	±0.64	±0.06	±0.39	±0.33	±0.01
Total systematics	±15	±12	±17	±31	±20	±24	±15	±22	±14
Data statistics	±9	±9	±9	±10	±9	±9	±13	±9	±4
Total uncertainty	±18	±15	±19	±33	±22	±25	±20	±24	±15

Table 90: Summary of the sources of uncertainty on the normalised differential cross section for min $\Delta R(\text{lepton}, \gamma)$ at particle level, presented as a percentage of the measured cross section in each bin in the single lepton channel. Entries of 0.0 are uncertainties that are less than 0.005 in magnitude. Any asymmetric systematic uncertainties have been symmetrised.

min $\Delta R(\text{lepton}, \gamma)$	1.0 - 1.2	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.2	2.2 - 2.4	2.4 - 2.6	2.6 - 6.0
Source	systematic uncertainty (%)								
Signal Scale Variation	±0.37	±0.41	±0.33	±0.48	±0.35	±0.1	±0.19	±0.14	±0.51
Signal ISR/FSR Setting	±0.97	±1.19	±2.51	±0.78	±0.16	±0.28	±1.15	±1.27	±0.21
Signal Parton Showering	±4.14	±2.74	±0.19	±5.53	±0.79	±0.59	±1.41	±9.24	±0.85
Background Modelling	±9.49	±5.42	±13.18	±27.47	±16.48	±20.96	±7.28	±16.37	±4.76
Luminosity	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Pileup Effects	±0.03	±0.32	±0.48	±0.58	±1.69	±0.76	±0.83	±0.51	±0.57
Photon Scale Factors	±0.07	±0.07	±0.03	±0.03	±0.04	±0.04	±0.01	±0.03	±0.05
Lepton Identification and Reconstruction	±0.09	±0.06	±0.06	±0.06	±0.05	±0.03	±0.02	±0.04	±0.06
E/Gamma Resolution and Scale	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
Jet Identification and Reconstruction	±0.35	±0.61	±0.28	±0.62	±0.36	±0.43	±0.13	±0.33	±0.53
b-tagging	±0.54	±0.17	±0.32	±0.19	±0.21	±0.15	±0.22	±0.41	±0.1
Missing Transverse Momentum Reconstruction	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0	±0.0
PPT	±0.54	±0.65	±0.2	±0.48	±0.65	±0.06	±0.39	±0.33	±0.0
Total systematics	±11	±7	±13	±28	±17	±21	±7	±19	±5
Data statistics	±9	±8	±9	±9	±8	±9	±12	±9	±4
Total uncertainty	±14	±11	±16	±30	±18	±23	±14	±21	±6

N.5 Pre-/post-fit unfolding comparison

A comparison between the unfolding results in the single lepton channel using the pre-fit and post-fit inputs is shown in Figure 190 and Figure 191, respectively. The same comparison in the dilepton channel can be seen in Figure 192 and Figure 193, repectively. The post-fit results are produced by an older version of the unfolding framework which had a few bugs(they're fixed in the new version) in the event selection. This will consequently give a slight difference when compared with the post-fit results in Section 13.3 after the fix.

97,0-1 Bot

Not reviewed, for internal circulation only

Figure 190: The absolute (right) and normalised (left) differential cross section, using pre-fit inputs, in p_{T} bins (top), $|\eta|$ bins (centre) and min $\Delta R(\mathrm{lepton}, \gamma)$ (bottom), in the single-lepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Figure 191: The absolute (left) and normalised (right) differential cross section, using post-fit inputs, in p_{T} bins (top), $|\eta|$ bins (centre) and min $\Delta R(\mathrm{lepton}, \gamma)$ (bottom), in the single-lepton channel. The unfolded distribution to the particle-level (dots) is compared to the particle-level distribution (red)

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 192: The absolute (right) and normalised (left) differential cross section, using pre-fit inputs, in p_{T} bins (first row), $|\eta|$ bins (second row), minimum $\Delta R(\mathrm{lepton}, \gamma)$ (third row), $\Delta\eta(\mathrm{lepton}, \mathrm{lepton})$ (fourth row), and $\Delta\phi(\mathrm{lepton}, \mathrm{lepton})$ (fifth row) in the dilepton channel. The unfolded distribution (dots) is compared to the particle level distribution (red line).

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 193: The absolute (left) and normalised (right) differential cross section, using post-fit, in p_{T} bins (first row), $|\eta|$ bins (second row), minimum $\Delta R(\mathrm{lepton}, \gamma)$ (third row), $\Delta\eta(\mathrm{lepton}, \mathrm{lepton})$ (fourth row), and $\Delta\phi(\mathrm{lepton}, \mathrm{lepton})$ (fifth row) in the dilepton channel. The unfolded distribution to the particle-level (dots) is compared to the particle-level distribution(red)..

Not reviewed, for internal circulation only

O Unfolding using fine binning

The binning used in this Appendix is optimised based only on the resolution and expected data statistics checks. However, the final binning used in the nominal analysis and shown in the main body of the note is further optimised based on the final statistical uncertainty on the unfolded results. It turned out that after unfolding the statistical uncertainty was magnified, thus a few bins are merged by eyes to have a rough statistical uncertainty in the same level as the other bins. All checks are repeated for finer and wider binning.

O.1 Detector resolution

The resolution of the three variables, photon p_{T}, photon η, and $\Delta R(\ell, \gamma)$, are studied with single lepton channel signal MC (resolution should be universal to all channels) and shown in Figure 194. The binning is chosen in a way that: the bin width is larger than twice of the resolution of the observable, and the expected statistical uncertainty of the bin should be smaller than 10% (15%) for the single lepton (dilepton) channel. The comparison between bin width and resolution, as well as the expected statistical uncertainties are shown in Figure **??** for the photon p_{T}, Figure **??** for the photon η, and Figure **??** for the $\Delta R(\ell, \gamma)$. To be mentioned, these are not the final binning used for the nominal analysis, since the unfolded statistical uncertainty could be different from the estimation performed here, which is based on the Poisson formula. In fact, it is found that the statistical uncertainty is underestimated by the Poisson formula, thus a few bins are merged based on the binning defined here.

O.2 Inputs

Figure 198 shows $1/\epsilon$ and $(1 - f_{\mathrm{mig}})$ in the dilepton channel, while Figure 199 in the single lepton channel. The migrations are shown in Figure 200 and Figure 201 in the dilepton and single lepton channels.

O.2.1 Optimising the number of iterations

In order to choose the number of iterations that to be used for unfolding, two methods are employed. The final chosen number of iterations is a compromise between having a low difference in the unfolded results between different iterations, having a stable statistical uncertainty in all bins, and having no bias against any number of iterations.

For the first method, the convergence of the unfolded results, is shown in Figure 202.

$$\frac{N^{\mathrm{unf,\,i+1}} - N^{\mathrm{unf,\,i}}}{N^{\mathrm{unf,\,i+1}}} < 1\% \tag{70}$$

The results of the RMS test are shown in Figure 203 for p_{T} and Figure 204 for η in the dilepton channel. While for single lepton channel, are shown in Figure 205 for the photon's p_{T}, and Figure 206 for the photon η.

Not reviewed, for internal circulation only

(a) (b)

(c)

Figure 194: The resolution of the three variables to be unfolded: photon p_{T}, photon η, and $\Delta R(\ell, \gamma)$, studied by using the signal MC sample in the single lepton channel.

Not reviewed, for internal circulation only

Figure 195: The proposed binning for the photon p_{T} and its comparison to the resolution, as well as its expected statistical uncertainty for the single lepton and dilepton channels.

Not reviewed, for internal circulation only

Figure 196: The proposed binning for the photon η and its comparison to the resolution, as well as its expected statistical uncertainty for the single lepton and dilepton channels.

Not reviewed, for internal circulation only

Figure 197: The proposed binning for the $\Delta R(\ell, \gamma)$ and its comparison to the resolution, as well as its expected statistical uncertainty for the single lepton and dilepton channels.

Not reviewed, for internal circulation only

Figure 198: The inverse of signal efficiency $1/\epsilon$ (left) and fraction of non-outside migration $(1 - f_{\mathrm{mig}})$ (right) as a function of (a), (b) the photon p_{T}, (c), (d) the photon $|\eta|$ in the dilepton channel.

Not reviewed, for internal circulation only

Figure 199: The inverse of signal efficiency $1/\epsilon$ (left) and fraction of non-outside migration $(1 - f_{\mathrm{mig}})$ (right) as a function of (a), (b) the photon p_{T}, (c), (d) the photon $|\eta|$ in the single lepton channel.

Not reviewed, for internal circulation only

(a) (b)

Figure 200: The bin-by-bin migration matrix of (a) the photon p_T and (b) the photon $|\eta|$ in the dilepton channel.

(a) (b)

Figure 201: The bin-by-bin migration matrix of (a) the photon p_T and (b) the photon $|\eta|$ in the single lepton channel.

Not reviewed, for internal circulation only

Figure 202: The relative difference on the unfolded result after two successive iterations as a function of the photon p_{T} (left) and the photon $|\eta|$ (right), (a) and (b) in dilepton channel, (c) and (d) in the single lepton channel.

Not reviewed, for internal circulation only

From the first methods, the two variables and channels agree for 6 iterations, and most of bins are stable after 6 iterations from the second method.

$$\text{RMS} = \frac{(\text{truth} - \text{unfolded})}{\text{unfolded}} \tag{71}$$

Figure 203: The RMS against the number of iterations for the 12 Bins in p_{T} in the dilepton channel.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 204: The RMS against the number of iterations for the 14 Bins in η in the dilepton channel.

O.2.2 The closure test

The closure test is performed to check if the the unfolding procedure can recover the reconstructed spectrum. To do that, the 100 pseudo data distributions of the testing par at the reconstruction-level are unfolded using the migration matrix from the training part. A perfect closure is obtained, and shown in Figure 207 for the dilepton and single lepton channels.

O.2.3 Pull study

The binning was chosen initially to have more diagonal matrices, with more than 50% of particle-level events are reconstructed, as shown earlier in Section 7. The stability of the bin choice is checked by performing pull tests. The same pseudo-data described in Section 9.1.1 is used and another 1000 pseudo-experiments are built from the reconstructed-level in the testing part using the Poisson distribution, and then unfolded. The unfolded pseudo-results are then compared to the pseudo-particle level. The pulls are calculated per bin (Equation 72) and fit by a Gaussian function (see Appendix P). The pull is expected to have a mean value of zero, which indicates no bias, while a width of one which could mean that the statistical uncertainty is estimated correctly. Figure 208 shows no bias in the bin choice, and the width is consistent with one.

$$\text{Pull} = \frac{(\text{truth} - \text{unfolded})}{\sigma_{\text{unfold}}} \tag{72}$$

O.2.4 Stress tests

The stress test is performed in order to verify that the unfolding procedure isn't biased to any specific shape of the particle level distribution. The particle level and reconstructed level distributions obtained from the nominal MC sample are reweighed, the reweighted reconstructed distribution is then unfolded using the nominal inputs from the MC sample, and the unfolded results are compared to the corresponding particle

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 205: The RMS against the number of iterations for the 27 Bins in p_{T} in the single lepton channel.

level distribution. Different weights have been checked, the first is by taking the observed difference between data and MC as the following:

$$\text{weight} = 1 + Y \cdot \frac{\text{data}_i - \text{MC}_i}{\text{data}_i} = 1 + Y \cdot \text{Obs}, \tag{73}$$

where i is the bin index and Y= 1, -1. The result of the stress test is shown in Figure 209. The unfolding is able to retrieve the reweighted particle distribution for both variables η or p_{T}.

A different weight is tried, and defined as the following in case of the photon's p_{T}:

$$\text{weight} = 1 + y \cdot \frac{100 - i}{300} = 1 + y \cdot X, \tag{74}$$

while for the photon's η is given by:

$$\text{weight} = 1 + y \cdot \frac{1.2 - i}{2.37} = 1 + y \cdot X, \tag{75}$$

where y = -1, 1 and i is the bin centre. The result of the second test is shown in the same Figure 209. The reweighted particle level distribution are in different shapes from the nominal ones, and the unfolding procedure is able to retrieve the reweighted particle level distributions in both channels.

O.3 Cross-check methods

O.3.1 Closure tests

The test is done here to compare different methods by unfolding the reconstructed-level distribution, which is estimated from MC. Figures 210 and 210 show the different unfolded results with different methods, compared to the truth distribution in the dilepton and single lepton channels. The SVD method, show always slightly different central values from the other methods, but agrees within the statistical uncertainty, very good closure is achieved in photon's p_{T}.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 206: The RMS against the number of iterations for the 20 Bins in η in the single lepton channel.

Not reviewed, for internal circulation only

(a)

(b)

(c)

(d)

Figure 207: The ratio of the number of unfolded events for one pseudo-data after 6 iterations to the number of truth events, as a function of the photon p_{T} (left) and the photon η (right) in (a) and (b) the dilepton channel, (c) and (d) the single lepton channel.

Not reviewed, for internal circulation only

Figure 208: The pull test in each bin of the photon p_T (left), the photon η (right), after 6 iterations in (a) and (b) the dilepton channel, (c) and (d) the single lepton channel.

Not reviewed, for internal circulation only

(a) (b) (c) (d)

Figure 209: The stress test for the photon p_{T} (left), the photon η (right), in (a) and (b) the dilepton channel, (c) and (d) the single lepton channel. Both the dots and lines are ratios made with respect to the nominal particle level. The dots are the ratio of the unfolded reweighted distributions to the nominal particle level distribution, while the solid lines are the ratio of the reweighted particle level distributions to the nominal one. Obs. is defined in Equation 73, and X in Equations 74 and 75.

Not reviewed, for internal circulation only

(a)

(b)

(c) (d)

Figure 210: The number of unfolding events as a function of the photon p_T (left) and the photon η (right) in (a) and (b) the dilepton channel, (c) and (d) in the single lepton channel, obtained by the iterative method in black, after 6 iterations, the SVD in red and the bin-by-bin in magenta, in the dilepton channel. All are compared to the truth distribution in blue (the ratio plot). The yellow distribution is the original reconstructed distribution before unfolding. The distribution are normalized to a luminosity of 36.1 fb^{-1}.

Not reviewed, for internal circulation only

P Pull study for unfolding with IBS

Example distributions of the pulls calculated per photon p_{T} and photon $|\eta|$ bins in the dilepton channel (see Section 9.1.4) are shown in Figure 211 and Figure 212, respectively. The Gaussian fit is shown in red.

Not reviewed, for internal circulation only

Figure 211: The distributions of the pull in each bin of p_{T} (black dots) in the dilepton channel. The Gaussian fit is shown in red.

Not reviewed, for internal circulation only

Figure 212: The distributions of the pull in each bin of η (black dots) in the dilepton channel. The Gaussian fit is shown in red.

Not reviewed, for internal circulation only

Q Results appendix

Q.1 Asimov fits to ELD, number of jets and 1 bin distributions

The topic of whether the ELD is needed in the final dilepton fit has emerged a few times. Table 91 compares three Asimov fits performed (with all systematics) on the ELD, the number of jets and single bin distributions. See Chapter 5 for the number of jets distributions, bearing in mind that these are shown for separate channels. From the table we can see large gains in sensitivity when fitting the ELD distribution in both the single lepton and dilepton channels.

	ELD		number of jets		1 bin	
	+	-	+	-	+	-
single lepton	0.096	0.093	0.419	0.330	1.204	0.779
dilepton	0.072	0.068	0.092	0.106	0.080	0.076

Table 91: Asimov fit results showing error of μ for the ELD, number of jets and single bin distributions for the single lepton and dilepton channels.

Q.2 Post-fit plots for single lepton and dilepton channels

The section shows the post-fit plots for the variables presented in Chapter 11. All systematic uncertainties and scale factors are included. Figure 213 shows the single lepton channel, while Figure 214 shows the dilepton channel.

Not reviewed, for internal circulation only

Not reviewed, for internal circulation only

Figure 213: Post-fit distributions used for training the single lepton ELD. All scale factors and systematic uncertainties included.

Not reviewed, for internal circulation only

Figure 214: Post-fit distributions used for training the dilepton ELD. All scale factors and systematic uncertainties included.

Not reviewed, for internal circulation only

Q.3 Results for e+jets, μ+jets, ee, $\mu\mu$ and $e\mu$

This section contains the results for each of the individual channels that make up the single lepton and dilepton channels. Pre- and post-fit plots can be seen in Figures 215, 216, 217, 218 and 219 where the shown ELD is the distribution used in the fit.

Figures 220, 221, 222, 223 and 224 show the 5 channel inclusive fit results as applied to the e+jets, μ+jets, ee, $\mu\mu$ and $e\mu$ channels, respectively.

The total contributions relative to μ are summarised in Figure 225. The JES contribution includes *Jet flavour composition*, *Jet flavour response*, *Jet eta intercalibration*, *Jet pileup* (with the largest contribution coming from *RhoTopology*), *Jet effective NP*, *Jet single particle HightPt*, *Jet Punchthrough MC15* and *Jet BJES response*.

Nuisance parameter pull plots are shown in Figure 226 and Figure 227.

Pruning plots (at 0.7%) are shown in Figures 228, 229, 230, 231 and 232. The general trend is that the shape components of the signal modelling do not get pruned but the normalisation components do (with some exceptions).

(a) Pre-fit plots

(b) Post-fit plots

Figure 215: Pre- and post-fit plots for the e+jets channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Pre-fit plots

(b) Post-fit plots

Figure 216: Pre- and post-fit plots for the μ+jets channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Pre-fit plots

(b) Pre-fit plots

(c) Post-fit plots

(d) Post-fit plots

Figure 217: Pre- and post-fit plots for the *ee* channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Pre-fit plots

(b) Pre-fit plots

(c) Post-fit plots

(d) Post-fit plots

Figure 218: Pre- and post-fit plots for the $\mu\mu$ channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Pre-fit plots

(b) Pre-fit plots

(c) Post-fit plots

(d) Post-fit plots

Figure 219: Pre- and post-fit plots for the *ee* channel where this ELD distribution is used as the discriminating variable in the fit.

Not reviewed, for internal circulation only

(a) Post-fit ELD

(b) Post-fit $|\eta(\gamma)|$

(c) Post-fit $p_T(\gamma)$

(d) Post-fit ΔR(lepton, γ)

Figure 220: Post-fit plots for the 5 channel inclusive fit applied to the e+jets channel.

Not reviewed, for internal circulation only

(a) Post-fit ELD

(b) Post-fit $|\eta(\gamma)|$

(c) Post-fit $p_T(\gamma)$

(d) Post-fit ΔR(lepton, γ)

Figure 221: Post-fit plots for the 5 channel inclusive fit applied to the μ+jets channel.

Not reviewed, for internal circulation only

(a) Post-fit ELD (b) Post-fit $|\eta(\gamma)|$ (c) Post-fit $p_T(\gamma)$

(d) Post-fit $\Delta\phi$(lepton, lepton) (e) Post-fit $\Delta\eta$(lepton, lepton) (f) Post-fit min ΔR(lepton, γ)

Figure 222: Post-fit plots for the 5 channel inclusive fit applied to the *ee* channel.

Not reviewed, for internal circulation only

(a) Post-fit ELD

(b) Post-fit $|\eta(\gamma)|$

(c) Post-fit $p_T(\gamma)$

(d) Post-fit $\Delta\phi$(lepton, lepton)

(e) Post-fit $\Delta\eta$(lepton, lepton)

(f) Post-fit min ΔR(lepton, γ)

Figure 223: Post-fit plots for the 5 channel inclusive fit applied to the $\mu\mu$ channel.

Not reviewed, for internal circulation only

(a) Post-fit ELD (b) Post-fit $|\eta(\gamma)|$ (c) Post-fit $p_T(\gamma)$

(d) Post-fit $\Delta\phi$(lepton, lepton) (e) Post-fit $\Delta\eta$(lepton, lepton) (f) Post-fit min ΔR(lepton, γ)

Figure 224: Post-fit plots for the 5 channel inclusive fit applied to the $e\mu$ channel.

Not reviewed, for internal circulation only

Figure 225: Grouped uncertainties for the individual channels relative to the signal strength.

Not reviewed, for internal circulation only

(a) e+jets (b) μ+jets

Figure 226: Pull plots for all nuisance parameters in the single lepton channels after a 0.7% pruning has been applied.

Not reviewed, for internal circulation only

(a) ee (b) $\mu\mu$ (c) $e\mu$

Figure 227: Pull plots for all nuisance parameters in the dilepton channels after a 0.7% pruning has been applied.

Not reviewed, for internal circulation only

Figure 228: Pruning plots showing the contributions of each systematic that survive the pruning process for the e+jets channel. Grey means the systematic does not apply. Green means both shape and normalisation contributions are kept. Yellow means the shape contribution has been dropped but the normalisation has been kept, while orange is the reverse. Red indicates the whole systematic was pruned.

Not reviewed, for internal circulation only

Figure 229: Pruning plots showing the contributions of each systematic that survive the pruning process for the μ+jets channel. Grey means the systematic does not apply. Green means both shape and normalisation contributions are kept. Yellow means the shape contribution has been dropped but the normalisation has been kept, while orange is the reverse. Red indicates the whole systematic was pruned.

Not reviewed, for internal circulation only

Figure 230: Pruning plots showing the contributions of each systematic that survive the pruning process for the *ee* channel. Grey means the systematic does not apply. Green means both shape and normalisation contributions are kept. Yellow means the shape contribution has been dropped but the normalisation has been kept, while orange is the reverse. Red indicates the whole systematic was pruned.

Not reviewed, for internal circulation only

Figure 231: Pruning plots showing the contributions of each systematic that survive the pruning process for the $e\mu$ channel. Grey means the systematic does not apply. Green means both shape and normalisation contributions are kept. Yellow means the shape contribution has been dropped but the normalisation has been kept, while orange is the reverse. Red indicates the whole systematic was pruned.

Not reviewed, for internal circulation only

Figure 232: Pruning plots showing the contributions of each systematic that survive the pruning process for the $\mu\mu$ channel. Grey means the systematic does not apply. Green means both shape and normalisation contributions are kept. Yellow means the shape contribution has been dropped but the normalisation has been kept, while orange is the reverse. Red indicates the whole systematic was pruned.

Not reviewed, for internal circulation only

Q.4 Final fits for all scenarios

It is possible to perform a fit in different ways to extract the parameter of interest. Figure 100 shows the results for all the scenarios. These are further explained in Section 8.4. As a cross check, the results of all possible scenarios are shown in Figure 233. All results are within agreement of each other.

Figure 233: The observed signal strength for all the fit scenarios. The standard model prediction of 1 is represented by the vertical line. The theoretical uncertainty for each fit is represented by the shaded region.

Q.5 Nuisance parameter pull studies

This section shows the extra studies and investigations for those nuisance parameters that are constrained in the final fit.

Q.5.1 $t\bar{t}\gamma$ ISR/FSR

Figure 234 and Table 92 highlight the differences between symmetrising the $t\bar{t}\gamma$ ISR/FSR systematic versus leaving it asymmetric. The impact is that a slightly larger up variation in some of the fits is seen, however central values of μ are not changed.

Not reviewed, for internal circulation only

(a) single lepton (b) dilepton

Figure 234: Pull plots for all nuisance parameters in the single lepton and dilepton channels after symmetrising the $t\bar{t}\gamma$ ISR/FSR systematic.

Not reviewed, for internal circulation only

	$t\bar{t}\gamma$ ISR/FSR non-symm.			$t\bar{t}\gamma$ ISR/FSR symm.		
	μ	+	-	μ	+	-
ee	0.999	0.151	0.120	1.002	0.139	0.125
e+jets	1.064	0.089	0.085	1.061	0.088	0.084
$e\mu$	1.094	0.090	0.085	1.095	0.104	0.093
μ+jets	1.021	0.092	0.088	1.020	0.095	0.090
$\mu\mu$	1.109	0.128	0.117	1.108	0.129	0.117
single lepton	1.040	0.103	0.084	1.044	0.087	0.084
dilepton	1.083	1.083	0.078	1.085	0.074	0.071

Table 92: The effect of smoothing the $t\bar{t}\gamma$ ISR/FSR systematics. Small differences can be seen mainly in the “up" error, however central values remain the same.

Q.5.2 Jet Energy Resolution studies

The Jet Energy Resolution plots are shown in Figure 235. In the data and MC combined plot we can see the third bin from the right experiences some tension. The data prefers the “down" variation of the systematic (the blue line which changes signs). Thus a negative constraint is experienced. To demonstrate this a fit is performed with just the JER NP and with masking the third bin from the right. Figure 236 shows the effect of fitting with and without the masked bin[17]. A clear improvement can be seen when masking the bin.

[17] This initial NP is less constrained due to being the only systematic in the toy fit.

Not reviewed, for internal circulation only

Figure 235: The JER systematic contribution for the single lepton channel split into signal and background contributions.

Not reviewed, for internal circulation only

(a) Fitting the full ELD distribution.

(b) Fitting the ELD distribution with one bin masked.

(c) The ELD distribution with one bin masked.

Figure 236: The JER systematic contribution for the single lepton channel when fitting the full distributions vs masking the offending bin.

Q.5.3 Lepfake studies

The Lepfake systematic uncertainty is constrained due to the data preferring a different parameterisation. The up/down variations for the Lep fake background are from the extremes of the parameterisation “envelope" used in the data driven method. Figure 237 shows the down variation (or parameterisation) being preferred.

Figure 237: The Lepfake systematic contribution for the single lepton channel.

Not reviewed, for internal circulation only

Q.5.4 $Z\gamma$ modelling

In the dilepton channels (specifically the $\mu\mu$) channel heavily constrained $Z\gamma$ modelling systematics can be seen. This in part due to the larger $V\gamma$ weights, but also the large differences in the modelling of SHERPA and the AMCATNLOPYTHIA8 sample. Given our validation region is sane, we take this shape difference as a conservative estimate and apply a 50% normalisation uncertainty on the $Z\gamma$ cross section.

Figure 238: The $Z\gamma$ parton shower systematic contribution for the dilepton channel.

Q.6 Asimov fit scenarios: $V\gamma$ floating or constrained

Two Asimov fit scenarios are considered: either fixing $V\gamma$ backgrounds to MC prediction with a 50% uncertainty or floating them.

The fits results for are summarised in Table 93, from which it can be found that the difference in terms of signal strength between these two scenarios is very small. In the same table, it is attempted to split one signal region into two signal regions according to the number of b-jets, of which the results are very similar to the case without splitting, except that a tighter constraint can be imposed on the floated $V\gamma$ normalisation.

These fits make use of an older k-factor and also contain the k-factor theoretical uncertainty. However, since they are self-contained tests on Asimov data the behaviour of the fit is still relevant.

Not reviewed, for internal circulation only

Table 93: 36.1 fb^{-1} Asimov fit results for each channel and a variety of scenarios. $V\gamma$ refers to $W\gamma$ for single lepton and and $Z\gamma$ for dilepton. A dash infers the that $V\gamma$ background has been constrained with a systematic uncertainty of 50%, opposed to being left floating in the fit. Systematics described in 10 have been included in the fit with the exception of the PPT and $e \rightarrow \gamma$ fake systematics. Also included for these scenarios is the signal k-factor uncertainty.

channel	ELD template(s)	± error (μ)	± error ($V\gamma$)	S/B	$S/Sqrt(S+B)$
$e + jets$	≥1 btag	0.2693	-	1.36 ± 0.06	43.82 ± 2.09
	≥1 btag	0.2689	0.5680		
	=1 btag, ≥2 btag	0.2686	-		
	=1 btag, ≥2 btag	0.2691	0.3603		
$\mu + jets$	≥1 btag	0.2671	-	1.62 ± 0.07	44.11 2.06
	≥1 btag	0.2684	0.2879		
	=1 btag, ≥2 btag	0.2688	-		
	=1 btag, ≥2 btag	0.2696	0.2435		
single lepton	≥1 btag	0.2633	-	1.47 ± 0.09	62.13 ± 4.09
	≥1 btag	0.2658	0.2714		
	=1 btag, ≥2 btag	0.2653	-		
	=1 btag, ≥2 btag	0.2659	0.2139		
ee	≥1 btag	0.3045	-	3.15 ± 0.41	10.24 ± 1.70
	≥1 btag	0.3054	0.3443		
	=1 btag, ≥2 btag	0.3025	-		
	=1 btag, ≥2 btag	0.3029	0.3495		
$\mu\mu$	≥1 btag	0.2987	-	1.93 ± 0.29	10.46 ± 1.67
	≥1 btag	0.2996	0.2915		
	=1 btag, ≥2 btag	0.2987	-		
	=1 btag, ≥2 btag	0.2998	0.2840		
$e\mu$	≥1 btag	0.2810	-	11.34 ± 0.87	18.71 ± 1.85
	=1 btag, ≥2 btag	0.2856	-		
dilepton	≥1 btag	0.2771	-	4.20 ± 0.50	23.53 ± 3.13
	≥1 btag	0.2779	0.2362		
	=1 btag, ≥2 btag	0.2767	-		
	=1 btag, ≥2 btag	0.2778	0.2337		

Not reviewed, for internal circulation only

Q.6.1 single lepton and dilepton combined pull plots

The pull plots for the single lepton and dilepton combined fits are shown in Figure 239 and 240.

From the nuisance parameter plots we can see that in both the separated and combined single lepton cases the Lep fake background is constrained. The $W\gamma$ background, when not a floating parameter in the fit, is also constrained. Similarity for the separate and combined dilepton, there is tension for the $Z\gamma$ background when it is constrained in the fit, hinting that 50% assigned uncertainty is perhaps too conservative. These fits make use of the systematics defined in Section 10 with the exception of the PPT and $e \rightarrow \gamma$ fake systematics. The theoretical uncertainty on the k-factor is also included. Scale factors and uncertainties derived for the $e \rightarrow \gamma$ fake and hadronic fake background are also included.

(a) ≥ 1 btag, $W\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $W\gamma$ constrained

(c) ≥ 1 btag, $W\gamma$ floating

(d) = 1 btag, ≥ 2 btag, $W\gamma$ floating

Figure 239: Pull plots for the combined single lepton channels showing $W\gamma$ constrained and floating. See Table 93 for reference. They are also shown for the case of splitting the signal region according to the number of b-jets.

Not reviewed, for internal circulation only

(a) ≥ 1 btag, $Z\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $Z\gamma$ constrained

(c) ≥ 1 btag, $Z\gamma$ floating

(d) = 1 btag, ≥ 2 btag, $Z\gamma$ floating

Figure 240: Pull plots for the combined dilepton channels showing $Z\gamma$ constrained and floating. See Table 93 for reference. They are also shown for the case of splitting the signal region according to the number of b-jets.

Not reviewed, for internal circulation only

Q.6.2 Individual channel pull plots

The pull plots for individual channels are shown in Figure 241, 242, 243, 244 and 245.

(a) ≥ 1 btag, $W\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $W\gamma$ constrained

(c) ≥ 1 btag, $W\gamma$ floating

(d) = 1 btag, ≥ 2 btag, $W\gamma$ floating

Figure 241: Pull plots for the e+jets channels showing $W\gamma$ constrained and floating. See Table 93 for reference.

Not reviewed, for internal circulation only

(a) ≥ 1 btag, $W\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $W\gamma$ constrained

(c) ≥ 1 btag, $W\gamma$ floating

(d) = 1 btag, ≥ 2 btag, $W\gamma$ floating

Figure 242: Pull plots for the μ+jets channels showing $W\gamma$ constrained and floating. See Table 93 for reference.

Not reviewed, for internal circulation only

(a) ≥ 1 btag, $Z\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $Z\gamma$ constrained

(c) ≥ 1 btag, $Z\gamma$ floating

(d) = 1 btag, ≥ 2 btag, $Z\gamma$ floating

Figure 243: Pull plots for the ee channels showing $Z\gamma$ constrained and floating. See Table 93 for reference.

Not reviewed, for internal circulation only

(a) ≥ 1 btag, $Z\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $Z\gamma$ constrained

(c) ≥ 1 btag, $Z\gamma$ floating

(d) = 1 btag, ≥ 2 btag, $Z\gamma$ floating

Figure 244: Pull plots for the $\mu\mu$ channels showing $Z\gamma$ constrained and floating. See Table 93 for reference.

Not reviewed, for internal circulation only

(a) ≥ 1 btag, $Z\gamma$ constrained

(b) = 1 btag, ≥ 2 btag, $Z\gamma$ constrained

Figure 245: Pull plots for the $e\mu$ channels showing $Z\gamma$ constrained. See Table 93 for reference.

Not reviewed, for internal circulation only

R Systematic uncertainties: RhoTopology studies

It was recently discovered that the systematic uncertainty for jet pileup RhoTopology has been incorrectly estimated for release 20.7. A fix on the timeline for this publication is not foreseen. The impact on this systematic is estimated to be reduced by up to a factor of 4. For this reason, three Asimov fits are presented. The first shows the “nominal" scenario in which no adjustments to this systematic are made. The next two scenarios are identical except for scaling the up and down contributions of the RhoTopology systematic by a factor of 0.5 and 0.25. Table 94 shows the impact this has for the single lepton channel and Figure 246 shows the affect each scaling has for signal sample.

Given that this systematic is ranked the highest, this is not surprising. These fits are not done in the dilepton channel as it enters as the fifth highest contribution. The main gain will be in the single lepton channel.

	nominal		RhoTop/0.5		RhoTop/0.25	
	+ [%]	- [%]	+ [%]	- [%]	+ [%]	- [%]
Error in μ	0.096	0.093	0.089	0.088	0.087	0.086

Table 94: The error of μ for the “nominal" fit versus a fit done with scaling the RhoTopology systematic by 0.5 and 0.25 for the single lepton channel.

(a) Nominal (b) RhoTop/0.50 (c) RhoTop/0.25

Figure 246: The affect of scaling the RhoTopology systematic by 0.5 and 0.25 for the $t\bar{t}\gamma$ sample in the single lepton channel.

Not reviewed, for internal circulation only

S Impact of the PPT on the analysis

This section explores the impact that the PPT has on the final ELD and thus the overall sensitivity. This brief study is provided purely to supply useful information for groups or analyses looking to repeat the process of creating a novel tool such as the PPT.

The re-training of the ELD has been performed in the exact same way, using the exact same samples, architectures and optimisations presented in Chapter 5.2. The *only* difference is the removal of the PPT. Thus, this new ELD variable only has 14 input variables. The number of epochs, layers, patience parameters in the training etc. remain the same. This enables us to see *exactly* the difference the PPT brings to the analysis. Figure 247 shows the training process including the output on the new ELD, the ROC curves and the loss functions. These plots can be directly compared to Figure 22 and 23 in the main part of the text. For example, an AUC of 0.69 can be compared to the AUC of 0.67 when training without the PPT.

Figure 248 shows the stacked distributions of the signal and various backgrounds. Clearly, the hadronic fake and $e \rightarrow \gamma$ fake backgrounds sit directly under the signal. Both pre-fit and post-fit distributions are shown, where the exact same fitting infrastructure (minus any PPT related systematics and scale factors) as the main part of this note has been used. Figure 249 shows that without the PPT in the ELD training, we learn nothing about the hadronic fake background (and even the $e \rightarrow \gamma$ fake background).

The NPs and correlation matrix for the fit to the ELD without the PPT is shown in Figure 250 and 251. The constraints and pulls that are seen are very similar to the results shown in the main part of this note.

Clearly, there is an improvement when using the PPT. However, due to the novelty of the tool, conservative systematic uncertainties have been derived in conjunction with egamma/PhotonID groups within ATLAS. The question arises, after the systematics have been included, do we still see a gain in overall performance when including the PPT?

Table 95 shows the sensitivity of the ELD with and without the PPT for Asimov fits. This shows that using our current method (with PPT and conservative systematics) gives the worst expected sensitivity. A clear gain can be seen between the cases of the ELD with no PPT versus the ELD with the PPT but no PPT systematics. This confirms that we can be more sensitive up to about 14% (7%) for the “up" (“down") variation when using the PPT.

Table 96 shows the sensitivities for the final full fit, which follow the same trend as the Asimov fit results. We do better with the PPT, but we wash-out our result with systematics. The relative “up" error increases by about 5% while the “down" error increases by about 8% compared to the results with no PPT. If the size of the PPT uncertainties were reduced we could expect improvements up to about 8% for the “up" variation and up to 7% for the “down" variation compared to using no PPT.

	Signal strength error		$W\gamma$ Norm error
	+	-	±
ELD with PPT	0.091	0.089	0.438
ELD with PPT, no PPT systs.	0.074	0.069	0.364
ELD without PPT	0.084	0.074	0.321

Table 95: Asimov fit results to the ELD with the PPT including all systematics, with all systematics but excluding the PPT systematics, and the ELD trained without the PPT.

Finally, a study can be made to see the different significances (s/b and $s/\sqrt{b}$) for various cuts on the ELD distribution. Figure 252 shows this for curves with and without the PPT used in the training of the

Not reviewed, for internal circulation only

(a) The ELD distribution for training and test samples.

(b) The ROC curves for the training and test samples.

(c) The loss for the training of the ELD. The validation curve has a "patience" parameter where if no improvement is seen after a certain number of epochs, the best model is saved (around 50 epochs).

Figure 247: The figures related to training of the ELD for the single lepton channel.

	Signal strength error			$W\gamma$ norm	
	μ	+	-	$W\gamma$SF	$\pm$
ELD with PPT	1.053	0.085	0.081	0.796	0.325
ELD with PPT, no PPT systs.	1.058	0.075	0.070	0.797	0.293
ELD without PPT	1.057	0.081	0.075	0.807	0.305

Table 96: Full fit results to the ELD with the PPT including all systematics, with all systematics but excluding the PPT systematics, and the ELD trained without the PPT.

ELD. For each bin of the ELD distribution from left to right a cut is performed and the remaining signal and background events used to calculate s/b and $s/\sqrt{b}$. The uncertainty bands include all systematic and statistical uncertainties. When a cut is made on a bin, the remaining bins uncertainties are assumed uncorrelated and added in quadrature. The figures corroborate that using the PPT is effective at increasing the sensitivity, but at a price of increasing the uncertainties.

Not reviewed, for internal circulation only

Figure 248: The pre- and post-fit distributions for the ELD training without the PPT. All scale factors and systematic uncertainties are included.

Figure 249: The normalised separation for the signal and hadronic fake contribution in the ELD trained without the PPT.

Not reviewed, for internal circulation only

Figure 250: The nuisance parameters after a fit is performed on the ELD which has been trained without the PPT as an input variable.

Not reviewed, for internal circulation only

Figure 251: The correlation matrix between NPs after a fit is performed on the ELD, which has been trained without the PPT as an input variable.

Not reviewed, for internal circulation only

(a) No uncentainties included

(b) Systematic and statistical uncertainties included

Figure 252: Curves showing the s/b and $s/\sqrt{b}$ for different cuts on the ELD distributions trained with and without the PPT.

Not reviewed, for internal circulation only

T New physics observations using Effective Field Theory

With the increase of signal events obtained at 100 fb^{-1}of integrated luminosity, possible deviations with respect to the SM could become visible. The model-independent Effective Field Theory (EFT) approach has been chosen in order to explore the possibility of observation of deviations with respect to the SM. In this approach, any new model can be modelized by the addition of new higher-dimensional operators to the SM Lagrangian, with corresponding coefficient. A study in the context of top quark pair production [6] has shown that three operators can have a visible impact on the $t\bar{t}\gamma$ production: the dipole operators O_{tB}, O_{tW} and O_{tG}, coupling to respectively weak hypercharge and isospin gauge bosons and gluons and creating electroweak and chromomagnetic dipole moments. The operators are described by Equation 76 and shown in Figure 253.

$$\begin{aligned} O_{tW} &= y_t g_w (\bar{Q}\sigma^{\mu\nu}\tau^I t)\bar{\phi}W^I_{\mu\nu}, \\ O_{tB} &= y_t g_Y (\bar{Q}\sigma^{\mu\nu} t)\bar{\phi}B_{\mu\nu}, \\ O_{tG} &= y_t g_s (\bar{Q}\sigma^{\mu\nu}T^A t)\bar{\phi}G^A_{\mu\nu}, \end{aligned} \tag{76}$$

where ϕ is the Higgs field, Q is the left-handed quark doublet of the third generation, g_w, g_t and g_s are the SM coupling constants, and y_t is the top-Yukawa coupling constant.

Figure 253: Feynman diagrams for the EFT vertices in top-quark production and decay.

The effect of the above operators on the inclusive cross-section is given by:

$$\sigma = \sigma_{\mathrm{SM}} + \frac{C_i}{\Lambda^2}\sigma_i^{(1)} + \frac{C_i^2}{\Lambda^4}\sigma_i^{(2)}, \tag{77}$$

Not reviewed, for internal circulation only

where σ_{SM} is the SM cross-section, $\sigma_i^{(1)}$ is the cross-section of the interference between diagrams with operator O_i vertex and SM diagrams and $\sigma_i^{(2)}$ is the cross-section of the quadratic term of the operator O_i. C_i is the coefficient associated to the operator O_i, and Λ is the cut-off in TeVunits.

T.1 EFT samples

The chosen strategy is to generate several samples, using MadGraph5_aMC@NLO with a customized Lagrangian as done in the reference paper[6], with all coefficients set to zero except one. The samples are generated at the next-to-leading order and including only photons emitted from initial state or off-shell top quark. For each operator, two samples are generated, with positive or negative coefficient value. The list of those samples is given in Table 97. The positive and negative variations of the coefficients are used to extract $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ in Equation 77.

By considering only one operator at a time, this strategy neglects the possible cancellation from the contributions of different EFT operators and the interferences between different EFT operators. The SM sample is generated by setting all the coefficients very close to zero. The samples are generated using the MadGraph5_aMC@NLO, and Pythia for parton showering. The selections at the generator level are as the following, the leptons, photons, and jet transverse momenta are asked to be larger than 15 GeVand $\eta < 5.0$. The NNPDF30 and the dynamic scale of the sum of transverse mass divided by two are used. The cross section value for each coefficient variation is given in Table 97. $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ compared to the SM cross section in Table 98. The validation of the generated samples is shown in the Appendix U.4.

Table 97: Expected NLO cross section for each different C_i generated at a time. Only statistical uncertainty is mentioned.

EFT operator	Cross section [pb]
σ_{SM}	2.652 ± 0.013
$C_{tB}/\Lambda^2 = +5$	3.005 ± 0.013
$C_{tB}/\Lambda^2 = -5$	2.807 ± 0.013
$C_{tG}/\Lambda^2 = +0.3$	2.914 ± 0.014
$C_{tG}/\Lambda^2 = -0.3$	2.469 ± 0.012
$C_{tW}/\Lambda^2 = +2$	2.757 ± 0.015
$C_{tW}/\Lambda^2 = -2$	2.670 ± 0.013

T.2 The study in the fiducial region

The same object definitions introduced in the fiducial region in Section 7.1.1 are implemented in rivet routine and used for the study. The event level selections described in Section 5 are also applied. The result of the inclusive cross sections are summarized in Table 99. The quadratic term has a larger impact for the O_{tB} operator, while the interference term for the other two operators, O_{tG} and O_{tW} in both the single lepton and dilepton channels. The differential cross sections in photon p_{T}, and the invariant mass of the photon and the hardest jet, are compared to the SM cross section and shown in Figure 254 and Figure 255, in the single lepton and dilepton channels respectively. The residual due to the O_{tB} operator is increasing with higher photons p_{T} or higher values of the invariant mass variable, this is due to the amount

Not reviewed, for internal circulation only

Table 98: The interference and quadratic cross section terms due to the new operators at the generator level in dilepton channel, the results are in pb.

Cross section/Operator	O_{tB}	O_{tG}	O_{tW}
σ_{SM}	2.652	2.652	2.652
$\sigma^{(1)}$	0.010	0.022	0.004
$\sigma^{(2)}$	0.025	0.004	0.006
$\sigma^{(1)}/\sigma_{SM}$	$< 10^{-3}$	10^{-3}	10^{-3}
$\sigma^{(2)}/\sigma_{SM}$	$< 10^{-3}$	10^{-3}	10^{-3}

of energy enters the EFT vertex, the behaviour is similar in both channels. However, for O_{tW} operator, the effect on the SM cross section is getting smaller with higher invariant mass values in the single lepton channel, while the interference and quadratic terms behave differently in the dilepton channel. For the photon p_{T}, it shows a similar behaviour as for the O_{tB} but with less impact on the SM cross section in both of the single lepton and dilepton channels. For the last operator, O_{tG}, the impact is stable over p_{T} and the invariant mass in the two channels.

Further variables have been checked in the Appendix U.3, since no visible effect of the operators has been observed for those variables, they will not be considered.

Table 99: The interference and quadratic cross section terms due to the new operators and their effect on the fiducial cross section in the single lepton and dilepton channels, the results are in pb.

Channel		O_{tB}	O_{tG}	O_{tW}
σ_{SM}	single lepton	0.286	0.286	0.286
	dilepton	0.059	0.059	0.059
$C\sigma^{(1)}$	single lepton	0.007	0.022	0.007
	dilepton	0.003	0.005	0.001
$C^2\sigma^{(2)}$	single lepton	0.036	0.007	0.001
	dilepton	0.007	0.001	0.001
$C\sigma^{(1)}/\sigma_{SM}$	single lepton	0.024	0.077	0.024
	dilepton	0.051	0.085	0.017
$C^2\sigma^{(2)}/\sigma_{SM}$	single lepton	0.126	0.024	$< 10^{-3}$
	dilepton	0.119	0.017	0.017

T.2.1 The fit

The SM cross section is fixed to the nominal one from the $t\bar{t}\gamma$ sample. The weights; $\sigma^{(1)}$ and $\sigma^{(2)}$, described in Table 99, are used as inputs to the fit (currently inclusive values, will be done later differentially). Table 100 shows a very preliminary Asimov fit results done in the single lepton and dilepton channels separately and also for their combinations. The results are in units of the corresponding coefficient for each

Not reviewed, for internal circulation only

Figure 254: The effect of the operators O_{tB} (left) O_{tG} (right), and O_{tW} (middle) on the cross section as a function of the photon p_{T} (top) and invariant mass of the photon and the hardest jet (bottom), in the single channel, for the chosen value of $C_{tB} = \pm 5$, $C_{tG} = \pm 0.3$ and $C_{tW} = \pm 2$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red.

operator ($C_{tB} = 5$, $C_{tG} = 0.3$ and $C_{tW} = 2$). It can be seen that the coefficient are not will constrained. The limits get improved by correcting the weights with fraction of photons from production to photons from decay, as shown in Table 101. The correction are predicted from theory at $\sqrt{s} = 14$ TeV [14]. A stronger constrain is observed in single lepton channel, and gets even stronger with the combination of the two channels.

Table 100: The constraints on the EFT coefficients by doing inclusive asimov fit in the single lepton and dilepton channels, and considering only photons from production.

		O_{tB}	O_{tG}	O_{tW}
Channel	single lepton	3.51754	1.19667	1.84423
	dilepton	5.33926	1.90162	1.89772
	combined	3.02848	1.01512	1.56445

Not reviewed, for internal circulation only

Figure 255: The effect of the operators O_{tB} (left) O_{tG} (right), and O_{tW} (middle) on the cross section as a function of the photon p_{T} (top) and invariant mass of the photon and the hardest jet (bottom), in the dilepton channel, for the chosen value of $C_{tB} = \pm 5$, $C_{tG} = \pm 0.3$ and $C_{tW} = \pm 2$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red.

Table 101: The constraints on the EFT coefficients by doing inclusive Asimov fit in the single lepton and dilepton channels, and considering photons from top production and decay. (Note: In this table small statistics is used.)

		O_{tB}	O_{tG}	O_{tW}
Channel	single lepton	2.55576	0.553034	7.05593
	dilepton	2.91095	0.837023	1.24251
	combined	1.89281	0.463836	1.09634

Not reviewed, for internal circulation only

U First look at EFT

U.1 Cross check of first EFT results

The results of the effect on the inclusive cross-section as a function of the photon transverse momentum are shown in Figure 256 for the two operators O_{tG} and O_{tB}, the valuse of C_{tB} and C_{tG} are set to ±4 and ±1 respictively, the two valuse are chosen only to be compared to the results in [6]. The results are obtained for all the decay modes of $t\bar{t}\gamma$, and with selections applied only at the generator level; the leptons and jet transverse momenta should be greater than 15 GeVand $\eta < 5.0$. The photon p_{T} should be greater than 20 GeV. As shown in the figure, the squared term is increasing fast with the transverse momentum for the O_{tB} while it is mostly constant for the O_{tG}, this is due to the amount of the momentum passing through the EFT vertex.

Table 102: The interference and quadratic NLO cross section terms and their effect on the inclusive cross section, the results are in pb.

EFT operator	O_{tB}	O_{tG}	O_{tW}
σ_{SM}	2.668	2.668	2.668
$\sigma^{(1)}$	0.0135	0.833	0.007
$\sigma^{(2)}$	0.0077	0.289	0.002
$\sigma^{(1)}/\sigma_{SM}$	0.005	0.312	0.003
$\sigma^{(2)}/\sigma_{SM}$	0.003	0.108	0.0007
$\sigma^{(2)}/\sigma^{(1)}$	0.571	0.347	0.286

Not reviewed, for internal circulation only

Figure 256: The effect of the operators O_{tB} (left) and O_{tG} (right) on the incluseive cross section as a function of the photon transverse momentum for the chosen value of $C_{tB} = \pm 4$ and $C_{tG} = \pm 1$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red. All decay channels are included.

U.2 The Monte Carlo study in the fiducial region in the single lepton channel

A fiducial space region is defined in order to enhance the contributions from photons from production, and also to focus on regions where the effect of the new EFT operators can be more visible. The event selections used to define this region are originally optimised for the SM fiducial cross section measurement. In the single lepton channel, the event is required to have exactly one lepton (either electron or muon), at least for jets with a transverse momentum $p_{\mathrm{T}} > 25$ GeVand $\eta < 2.5$, and at least one of them should be b-tagged is also required. Exactly one photon of $p_{\mathrm{T}} > 20$ GeVand $\eta < 2.37$ should exist, the distance between the photon and the jet should be greater than 0.4, and the distance between the photon and the lepton should be greater than 0.7.

The results in Table 102 and Figure 256 are shown for all the decay modes of $t\bar{t}\gamma$, and with loose selections applied at the generator level. The previous results can be compared to Table 103 and Figure 257 after applying the fiducial selections in the single lepton channel. The results in the fiducial region shows that the effect of quadratic term is further enhanced especially for the O_{tB}. Another interesting variable is the invariant mass between the photon and the leading jet which shows a rising effect at high value of the variable as shown in Figure 258. The results for separated single muon and single electron channels are aummarized in Table 104

Not reviewed, for internal circulation only

Table 103: The inerference and quadratic cross section terms due to the new operators and their effect on the fiducial cross section in the single lepton channel, the results are in pb.

EFT operator	O_{tB}	O_{tG}	O_{tW}
σ_{SM}	0.299	0.299	0.299
$\sigma^{(1)}$	0.002	0.078	0.003
$\sigma^{(2)}$	0.002	0.094	0.0006
$\sigma^{(1)}/\sigma_{SM}$	0.007	0.261	0.010
$\sigma^{(2)}/\sigma_{SM}$	0.007	0.314	0.002
$\sigma^{(2)}/\sigma^{(1)}$	1.	0.347	0.2

Table 104: The inerference and quadratic cross section terms due to the new operators and their effect on the fiducial cross section in the single muon and single electron channels, the results are in pb.

Channel		O_{tB}	O_{tG}	O_{tW}
σ_{SM}	μ+jets	0.156	0.156	0.156
	e+jets	0.159	0.159	0.159
$\sigma^{(1)}$	μ+jets	0.0012	0.0433	0.001
	e +jets	0.0011	0.0383	0.0025
$\sigma^{(2)}$	μ+jets	0.0007	0.0011	0
	e+jets	0.0007	0.0055	0.0012
$\sigma^{(1)}/\sigma_{SM}$	μ+jets	0.0075	0.2776	0.0064
	e+jets	0.0069	0.2408	0.0013
$\sigma^{(2)}/\sigma_{SM}$	μ+jets	0.0070	0.0047	0
	e+jets	0.0047	0.0353	0.0079
$\sigma^{(2)}/\sigma^{(1)}$	μ+jets	0.5833	0.0254	0
	e+jets	0.6727	0.1464	0.5

Figure 257: The effect of the operators O_{tB} (left) O_{tG} (right), and O_{tW} (middle) on the cross section as a function of the photon transverse momentum for the chosen value of $C_{tB} = \pm 5$, $C_{tG} = \pm 0.3$ and $C_{tW} = \pm 2$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red.

Not reviewed, for internal circulation only

Figure 258: The effect of the operators O_{tB} (left) O_{tG} (right), and O_{tW} (middle) on the cross section as a function of invariant mass of the photon and the leading jet for the chosen value of $C_{tB} = \pm 5$, $C_{tG} = \pm 0.3$ and $C_{tW} = \pm 2$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red.

U.3 Further variables for the EFT in the fiducial region

Not reviewed, for internal circulation only

Figure 259: The effect of the operators O_{tB} (left) O_{tG} (right), and O_{tW} (middle) on the cross section as a function of the distance between the photon and the jet (top), between the photon and the lepton (middle), and the photon η (bottom), in the single lepton channel, for the chosen value of $C_{tB} = \pm 5$, $C_{tG} = \pm 0.3$ and $C_{tW} = \pm 2$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red.

Not reviewed, for internal circulation only

Figure 260: The effect of the operators O_{tB} (left) O_{tG} (right), and O_{tW} (middle) on the cross section as a function of the distance between the photon and the jet (top), between the photon and the lepton (middle), and the photon η (bottom), in the dilepton channel, for the chosen value of $C_{tB} = \pm 5$, $C_{tG} = \pm 0.3$ and $C_{tW} = \pm 2$. The SM differential cross section in black, the effect of the interference term in blue, the effect of both interference and quadratic terms in red.

Not reviewed, for internal circulation only

U.4 EFT validation

The validation is performed in order to confirm that the EFT model is performing similarly to the SM, if the EFT coefficients are set very close to zero; in this case the process is referred as a SM-like. The coefficients are not set exactly to zero to avoid possible divergences in the calculations of the amplitudes in the EFT model. The validation is done by generating two samples, one is using the SM, and the other is using the EFT model with coefficients ≈ 0, both at LO and NLO. The differential cross section as a function of the photon's observables are checked and found to behave very similarly in the two samples, as and shown in Figure 261 for the LO process, and in Figure 262 for the NLO process.

Figure 261: A comparison between generating a small sample using the SM, shown in blue, and using the EFT model and setting its coefficients very close to zero, and shown in red. Both are generated at LO

Not reviewed, for internal circulation only

Figure 262: A comparison between generating a small sample using the SM, shown in blue, and using the EFT model and setting its coefficients very close to zero, and shown in red. Both are generated at NLO

Not reviewed, for internal circulation only

V Studying larger weights in the $W\gamma$ and $Z\gamma$ MC samples

It is apparent (especially so in the dilepton channel) that some $V + \gamma$ samples have higher weights than expected. This is shown by the large blue uncertainty band in the ELD distribution. Studies have been made in both the single lepton and dilepton channels to determine the origin of these higher-than-normal weights.

Figure 263 shows the individual distributions for *weight_mc* and *weight_pileup* as well as the product of the two versus the ELD for the $\mu^+\mu^-$ channel. Figure 264 shows the pileup dependence on the pileup weight. Evidently, there is one. Figure 265 shows the product of the two weights versus the ELD distribution for the e+jets channel.

(a) *weight_mc* (b) *weight_pileup* (c) *weight_mc* × *weight_pileup* vs ELD

Figure 263: *weight_mc* and *weight_pileup* and the product of both versus the ELD for the $\mu^+\mu^-$ channel.

Figure 264: Pileup dependence of the weight for the $Z\gamma$ samples in the $\mu^+\mu^-$ channel. It was suggested if there was not a dependance we could drop this weight for these samples. This isn't the case.

Various strategies have been discussed. One possibility is to place an arbitrary cut at some value of *weight_mc* × *weight_pileup*. While this solves the problem of these outlying large weights, it's quite hard to justify since no weight is individually very large. Thus, a study is made to see if cutting these events away or simply ignoring them actually changes our sensitivity.

Three (two) Asimov tests are performed for the single lepton (dilepton) channel. In each channel, a fit is performed without placing any cut on *weight_mc* × *weight_pileup* for the $V + \gamma$ samples. A second fit is performed placing a cut at *weight_mc* × *weight_pileup* < 5. For the single lepton channel a third fit is performed with *weight_mc* × *weight_pileup* < 12. The summary of the error in μ for each respective fit is shown in Table 105. It's apparent that neglecting the larger weights in the $V + \gamma$ samples has a

Not reviewed, for internal circulation only

Figure 265: *weight_mc* × *weight_pileup* vs ELD for the e+jets channel.

negligible affect. For the single lepton channel a cut of < 5 is in any case too aggressive and so wouldn't be considered.

Figure 266 shows the respective pull plots for the single lepton fits, while Figure 267 shows the same for the dilepton fits. Very little difference (if any) is seen between the different scenarios. For the single lepton channels the $W\gamma$ background is floating and so one can look at the associated errors in μ. These are 1 ± 0.43379, 1 ± 0.42762 and 1 ± 0.40686 for no cut, <12 and < 5, respectively. Thus, all acceptable for a floating background. For the dilepton channel, a 50% prior is placed on the $Z\gamma$ background since it is not floating. In both pull plots it can be seen to be constrained. This is because a cut at < 5 reduces weights which are uncommonly high, but this doesn't remove all of them. However, this is understood and so is acceptable.

Table 105: Summary of Asimov fits performed for different cuts on *weight_mc* × *weight_pileup*.

channel	*weight_mc* × *weight_pileup* cut	$\pm\Delta\mu$
single lepton	no cut	+0.10357 -0.09558
	<12	+0.10280 -0.09504
	<5	+0.09967 -0.09234
dilepton	no cut	+0.07337 -0.06773
	< 5	+0.07361 -0.06806

In summary, we abstain from cutting the few larger weights in the $V + \gamma$ SHERPA samples knowing that they will not drastically impact our final result.

Not reviewed, for internal circulation only

(a) No cut (b) *weight_mc* × *weight_pileup* < 12 (c) *weight_mc* × *weight_pileup* < 5

Figure 266: Pull plots for all nuisance parameters in the single lepton channel for various Asimov fits with or without cuts on *weight_mc* × *weight_pileup* for $V + \gamma$ samples.

Not reviewed, for internal circulation only

(a) No cut

(b) *weight_mc* × *weight_pileup* < 5

Figure 267: Pull plots for all nuisance parameters in the dilepton channel for various Asimov fits with or without cuts on *weight_mc* × *weight_pileup* for $V + \gamma$ samples.

Not reviewed, for internal circulation only

W Estimation of a possible mis-modelling of prompt photons from Pile-Up

The possibility of mismodelling and under-estimating the presence of prompt photons due to a production in pile-up events has been evaluated. To do so, the z position of the primary vertex of the event, corresponding to the hard interaction producing the particles of the signal final state is compared to the z position of the extrapolated vertex of the selected photon. If this photon is the result of pile-up activity, this z position should be uncorrelated to the primary vertex z position. The extrapolated vertex of the selected photon is obtained from the PhotonPointingTool tool from the PhotonVertexSelection package provided by the EGamma group. For converted photons, this vertex z position is obtained from the vertex of the conversion tracks. For unconverted photon, the z position is extrapolated from the shower shape of the photon in the calorimeter. The discriminating variable is $(z(PV) - z(\gamma))/\Delta z(\gamma)$ where $z(PV)$ and $z(\gamma)$ is the z position of the primary vertex and the vertex of the selected photon respectively, and $\Delta z(\gamma)$ is the uncertainty on the vertex position of the selected photon. In order to test if the simulation is missing some pile-up prompt photon, three different distribution are created from the signal sample: the nominal one and two where in a fraction of the events (10% and 30%) the z position of the selected photon has been randomized. This randomization is done by taking the z position of the previous selected event, and therefore correspond to a distribution corresponding to the typical expected z position. As visible in Figure 268, the addition of randomized position will not affect significantly the bulk of the event, but will have effects on the tails, as visible in the underflow and overflow bins. Comparing with the data distribution, we can conclude that there is no need to add a percentage of randomized photon vertex position, and therefore that the simulation accounts correctly for the prompt photon not associated to the primary vertex. Those distributions are normalized to unity, and done for the dilepton channels after the full signal selection. The distributions of the single variables ($z(PV)$, $z(\gamma)$ and $\Delta z(\gamma)$) also show a good shape agreement between data and simulation.

Not reviewed, for internal circulation only

Figure 268: Distribution of The bin-by-bin migration matrix of (a) the photon p_{T} and (b) the photon $|\eta|$ in the dilepton channel.

Not reviewed, for internal circulation only

X Objects overlap removal: loose photons vs. tight photons

In this analysis, the overlap removal between objects is applied with loose photon. Those photons have a $p_{\mathrm{T}} > 15$ GeV, and no identification or isolation criteria. This can only affect the last step of the overlap removal procedure, when jets close to a remaining loose photon are removed.

This choice is justified by the need of having a consistent definition in the signal selection and the fake photon control regions, in particular hadronic fake. If only the tight photons, such as the ones used in the signal selection, are used in this procedure, the looser photon candidates used in the hadronic fake control regions will be biased by the double counting of the loose photon and the possible jet corresponding to the same deposit. A dedicated production for the hadronic fake control regions with a different object overlap removal is also not acceptable, because this modification affects the event selection efficiency, which biases the ABCD method.

In order to understand the impact of the loose photon in the object overlap removal, the signal sample was reprocessed with an object overlap removal using tight photons. With respect to the procedure applied to the analysis, those photons have $p_{\mathrm{T}} > 20$ GeV, `Tight` identification and `FixedCutTight` working point isolation criteria. Additionally, the `FixedCutTight` working point used in this test corresponds to an updated value, and is slightly different with respect to the one used in the analysis. With respect to the analysis selection, the selection of this procedure has an increase efficiency of $\sim 5\%$ in the single lepton channel and $\sim 2\%$ in the dilepton channel. This increase is mainly due to events that are rejected in the analysis procedure and not in this procedure: only $< 0.5\%$ of the events passing the selection with loose photon are failing the selection with tight photon, which is due to the slightly different `FixedCutTight` working point. This increase of efficiency is due to the impact on the jet collection. With the loose photon used, more jets are vetoed, which affects both the jet selection cut and the b-jet selection cut. When the cuts on the jet multiplicity and b-jet multiplicity are removed, the difference between the two methods is indeed reduced to $< 1\%$ (which is the result of the slightly different `FixedCutTight` working point). This also has effects on the selected topology, are illustrated by Figure 269, which show the distributions of the jet and b-jet multiplicity when the jet and b-jet cuts are not applied. In the jet multiplicity plots, it can be seen that the most affected region is the low number of jets which is not passing selection. The remaining slope corresponds to a migration of events from one bin to another of a few percents. But as the object overlap removal procedure is applied identically in simulation and data, this change of selection efficiency does not impact the final measurement, and a similar effect would have been observed with a slightly different jet definition.

Not reviewed, for internal circulation only

Figure 269: Distribution of the jet and *b*-jet multiplicity for single lepton and dilepton channels for the object overlap removal when using the loose photon and the tight photon.

Not reviewed, for internal circulation only

Y Correction of Jet Flavour Composition Uncertainty

There are different uncertainties due to the flavour dependence of the jet energy scale(JES). One of them is the flavour composition uncertainty which depends on the gluon fraction of the jets and its uncertainty: $f_g \pm \Delta f_g$

For a particular p_T/η bin of the jets, f_g can be calculated from the signal MC sample as:

$$f_g = \frac{N_g}{N_g + N_{lq} + N_{cq}} \tag{78}$$

and its uncertainty,

$$\Delta f_g = \sqrt{(f_{g,nominal} - f_{g,Herwig7})^2 + (\frac{f_{g,radUp} - f_{g,radDown}}{2})^2} \tag{79}$$

which gives a better estimation of the flavour composition.

Not reviewed, for internal circulation only

Figure 270: Gluon fraction of jets for different p_T/η bins of the jets extracted from the signal(ttγ) MC sample

Not reviewed, for internal circulation only

Z Monte Carlo Samples

Z.1 Nominal samples

Table 106: Nominal MC samples.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
$t\bar{t}\gamma$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
410389	MadGraphPythia8EvtGen A14NNPDF23 ttgamma nonallhadronic	4926000	MadGraph +Pythia8 +EvtGen	A14 NNPDF23LO	1065.47271429	4.6233	1.7	1.0	7.860
$t\bar{t}$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
410501	PowhegPythia8EvtGen A14 hdamp258p75 ttbar nonallhad	58439000	Powheg +Pythia8 +EvtGen	A14 NNPDF23LO	147.178260062	730.19	1.1390	0.54378	452.25443603
$Z\gamma$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364500	Sherpa 222 NNPDF30NNLO eegamma_pty_7_15	3968000	Sherpa	NNPDF3.0 NNLO	68.8243660457	57.654	1.0	1.0	57.654
364501	Sherpa 222 NNPDF30NNLO eegamma_pty_15_35	2950000	Sherpa	NNPDF3.0 NNLO	85.3390418885	34.568	1.0	1.0	34.568
364502	Sherpa 222 NNPDF30NNLO eegamma_pty_35_70	989000	Sherpa	NNPDF3.0 NNLO	157.148759017	6.2934	1.0	1.0	6.2934
364503	Sherpa 222 NNPDF30NNLO eegamma_pty_70_140	484000	Sherpa	NNPDF3.0 NNLO	984.560304319	0.49159	1.0	1.0	0.49159
364504	Sherpa 222 NNPDF30NNLO eegamma_pty_140_E_CMS	99000	Sherpa	NNPDF3.0 NNLO	1570.13258897	0.063052	1.0	1.0	0.063052
364505	Sherpa 222 NNPDF30NNLO mumugamma_pty_7_15	3954000	Sherpa	NNPDF3.0 NNLO	68.530426192	57.697	1.0	1.0	57.697
364506	Sherpa 222 NNPDF30NNLO mumugamma_pty_15_35	2957000	Sherpa	NNPDF3.0 NNLO	85.4500794683	34.605	1.0	1.0	34.605
364507	Sherpa 222 NNPDF30NNLO mumugamma_pty_35_70	992000	Sherpa	NNPDF3.0 NNLO	157.790927022	6.2868	1.0	1.0	6.2868
364508	Sherpa 222 NNPDF30NNLO mumugamma_pty_70_140	489500	Sherpa	NNPDF3.0 NNLO	991.231800417	0.49383	1.0	1.0	0.49383
364509	Sherpa 222 NNPDF30NNLO mumugamma_pty_140_E_CMS	100000	Sherpa	NNPDF3.0 NNLO	1587.10005079	0.063008	1.0	1.0	0.063008
364510	Sherpa 222 NNPDF30NNLO tautaugamma_pty_7_15	3957000	Sherpa	NNPDF3.0 NNLO	68.6335726923	57.654	1.0	1.0	57.654
364511	Sherpa 222 NNPDF30NNLO tautaugamma_pty_15_35	2958000	Sherpa	NNPDF3.0 NNLO	85.3802857555	34.645	1.0	1.0	34.645
364512	Sherpa 222 NNPDF30NNLO tautaugamma_pty_35_70	980000	Sherpa	NNPDF3.0 NNLO	155.523463412	6.3013	1.0	1.0	6.3013
364514	Sherpa 222 NNPDF30NNLO tautaugamma_pty_140_E_CMS	98000	Sherpa	NNPDF3.0 NNLO	1555.0865612	0.063019	1.0	1.0	0.063019
$W\gamma$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364521	Sherpa 222 NNPDF30NNLO enugamma_pty_7_15	12374000	Sherpa	NNPDF3.0 NNLO	62.0748469951	199.34	1.0	1.0	199.34
364522	Sherpa 222 NNPDF30NNLO enugamma_pty_15_35	12855000	Sherpa	NNPDF3.0 NNLO	95.6402053419	134.41	1.0	1.0	134.41
364523	Sherpa 222 NNPDF30NNLO enugamma_pty_35_70	1992000	Sherpa	NNPDF3.0 NNLO	104.424407633	19.076	1.0	1.0	19.076
364524	Sherpa 222 NNPDF30NNLO enugamma_pty_70_140	248000	Sherpa	NNPDF3.0 NNLO	128.817785165	1.9252	1.0	1.0	1.9252
364525	Sherpa 222 NNPDF30NNLO enugamma_pty_140_E_CMS	97500	Sherpa	NNPDF3.0 NNLO	325.347036839	0.29968	1.0	1.0	0.29968
364526	Sherpa 222 NNPDF30NNLO munugamma_pty_7_15	12386000	Sherpa	NNPDF3.0 NNLO	62.0852130326	199.5	1.0	1.0	199.5
364527	Sherpa 222 NNPDF30NNLO munugamma_pty_15_35	12853000	Sherpa	NNPDF3.0 NNLO	95.56133829	134.5	1.0	1.0	134.5
364528	Sherpa 222 NNPDF30NNLO munugamma_pty_35_70	1986000	Sherpa	NNPDF3.0 NNLO	103.664265581	19.158	1.0	1.0	19.158
364529	Sherpa 222 NNPDF30NNLO munugamma_pty_70_140	248000	Sherpa	NNPDF3.0 NNLO	129.362057274	1.9171	1.0	1.0	1.9171
364530	Sherpa 222 NNPDF30NNLO munugamma_pty_140_E_CMS	97500	Sherpa	NNPDF3.0 NNLO	330.284552846	0.2952	1.0	1.0	0.2952
364531	Sherpa 222 NNPDF30NNLO taunugamma_pty_7_15	12352000	Sherpa	NNPDF3.0 NNLO	61.9489442801	199.39	1.0	1.0	199.39
364532	Sherpa 222 NNPDF30NNLO taunugamma_pty_15_35	12841000	Sherpa	NNPDF3.0 NNLO	95.4721189591	134.5	1.0	1.0	134.5
364533	Sherpa 222 NNPDF30NNLO taunugamma_pty_35_70	1987000	Sherpa	NNPDF3.0 NNLO	103.928029709	19.119	1.0	1.0	19.119
364534	Sherpa 222 NNPDF30NNLO taunugamma_pty_70_140	245500	Sherpa	NNPDF3.0 NNLO	126.748928701	1.9369	1.0	1.0	1.9369
364535	Sherpa 222 NNPDF30NNLO taunugamma_pty_140_E_CMS	99500	Sherpa	NNPDF3.0 NNLO	331.854717673	0.29983	1.0	1.0	0.29983
$ST + \gamma$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]

Not reviewed, for internal circulation only

Table 106: Nominal MC samples.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
410011	PowhegPythiaEvtGen P2012 singletop_tchan_lept_top	4986200	Powheg +Pythia +EvtGen	Perugia2012	113.998948307	43.739	1.0094	1.0	44.1501466
410012	PowhegPythiaEvtGen P2012 singletop_tchan_lept_antitop	4989800	Powheg +Pythia +EvtGen	Perugia2012	193.568158895	25.778	1.0193	1.0	26.2755154
410013	PowhegPythiaEvtGen P2012 Wt_inclusive_top	4985800	Powheg +Pythia +EvtGen	Perugia2012	146.602369961	34.009	1.054	1.0	35.845486
410014	PowhegPythiaEvtGen P2012 Wt_inclusive_antitop	4985600	Powheg +Pythia +EvtGen	Perugia2012	146.682750302	33.989	1.054	1.0	35.824406
410025	PowhegPythiaEvtGen P2012 SingleTopSchan_noAllHad_top	997800	Powheg +Pythia +EvtGen	Perugia2012	486.328410586	2.0517	1.005	1.0	2.0619585
410026	PowhegPythiaEvtGen P2012 SingleTopSchan_noAllHad_antitop	995400	Powheg +Pythia +EvtGen	Perugia2012	789.060642093	1.2615	1.022	1.0	1.289253
$VV+\gamma$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
361063	Sherpa CT10 llll	2964000	Sherpa	CT10	230.679430306	12.849	0.91	1.0	11.69259
361064	Sherpa CT10 lllvSFMinus	448800	Sherpa	CT10	243.357553411	1.8442	0.91	1.0	1.678222
361065	Sherpa CT10 lllvOFMinus	898000	Sherpa	CT10	247.69680587	3.6254	0.91	1.0	3.299114
361066	Sherpa CT10 lllvSFPlus	596600	Sherpa	CT10	232.88312905	2.5618	0.91	1.0	2.331238
361067	Sherpa CT10 lllvOFPlus	1197000	Sherpa	CT10	238.218436555	5.0248	0.91	1.0	4.572568
361068	Sherpa CT10 llvv	5929600	Sherpa	CT10	423.542857143	14.0	0.91	1.0	12.74
361070	Sherpa CT10 llvvjj ss EW6	10000	Sherpa	CT10	232.536508232	0.043004	0.91	1.0	0.03913364
361071	Sherpa CT10 lllvjj EW6	9800	Sherpa	CT10	233.238927101	0.042017	0.91	1.0	0.03823547
361072	Sherpa CT10 llljj EW6	60000	Sherpa	CT10	1905.00381001	0.031496	0.91	1.0	0.02866136
361077	Sherpa CT10 ggllvv	6289000	Sherpa	CT10	7364.08238779	0.85401	0.91	1.0	0.7771491
361091	Sherpa CT10 WplvWmqq SHv21 improved	3993900	Sherpa	CT10	160.494273659	24.885	0.91	1.0	22.64535
361092	Sherpa CT10 WpqqWmlv SHv21 improved	3993700	Sherpa	CT10	160.667015328	24.857	0.91	1.0	22.61987
361093	Sherpa CT10 WlvZqq SHv21 improved	3993600	Sherpa	CT10	347.450843919	11.494	0.91	1.0	10.45954
361094	Sherpa CT10 WqqZll SHv21 improved	3990500	Sherpa	CT10	1165.65402816	3.4234	0.91	1.0	3.115294
361095	Sherpa CT10 WqqZvv SHv21 improved	4962400	Sherpa	CT10	732.241404751	6.777	0.91	1.0	6.16707
361096	Sherpa CT10 ZqqZll SHv21 improved	3988900	Sherpa	CT10	1690.07837686	16.445	0.91	0.14352	2.147769624
361097	Sherpa CT10 ZqqZvv SHv21 improved	4483500	Sherpa	CT10	966.017888751	16.432	0.91	0.28245	4.223508744
$Z+jets$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364100	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV0 70 CVetoBVeto	7891000	Sherpa	NNPDF3.0 NNLO	4.84043821455	1983.0	0.9751	0.8221	1589.63171493
364101	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV0 70 CFilterBVeto	4917000	Sherpa	NNPDF3.0 NNLO	21.9786141692	1978.4	0.9751	0.11308	218.146906947
364102	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV0 70 BFilter	7902000	Sherpa	NNPDF3.0 NNLO	62.1324429023	1982.2	0.9751	0.064161	124.013153838
364103	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV70 140 CVetoBVeto	5917000	Sherpa	NNPDF3.0 NNLO	78.8760104787	108.92	0.9751	0.68873	73.1485614572
364104	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV70 140 CFilterBVeto	1969800	Sherpa	NNPDF3.0 NNLO	96.8068045993	109.42	0.9751	0.18596	19.8410843943
364105	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV70 140 BFilter	5900600	Sherpa	NNPDF3.0 NNLO	476.29608478	108.91	0.9751	0.11375	12.0800385388
364106	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV140 280 CVetoBVeto	4943000	Sherpa	NNPDF3.0 NNLO	203.538739262	39.878	0.9751	0.60899	23.6805991698
364107	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV140 280 CFilterBVeto	2954400	Sherpa	NNPDF3.0 NNLO	318.519317285	39.795	0.9751	0.23308	9.04446067686
364108	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV140 280 BFilter	12339300	Sherpa	NNPDF3.0 NNLO	1932.72316485	43.675	0.9751	0.14618	6.22543965365

Not reviewed, for internal circulation only

Table 106: Nominal MC samples.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364109	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV280 500 CVetoBVeto	1973000	Sherpa	NNPDF3.0 NNLO	413.369042021	8.5375	0.9751	0.55906	4.65412767872
364110	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV280 500 CFilterBVeto	986000	Sherpa	NNPDF3.0 NNLO	435.210414507	8.5403	0.9751	0.26528	2.20915807148
364111	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV280 500 BFilter	1971400	Sherpa	NNPDF3.0 NNLO	1321.91527905	8.4932	0.9751	0.17559	1.4541870954
364112	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV500 1000	2960500	Sherpa	NNPDF3.0 NNLO	1655.66802752	1.7881	0.9751	1.0	1.74357631
364113	Sherpa 221 NNPDF30NNLO Zmumu MAXHTPTV1000 E CMS	988000	Sherpa	NNPDF3.0 NNLO	6689.68785971	0.14769	0.9751	1.0	0.144012519
364114	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV0 70 CVetoBVeto	7900000	Sherpa	NNPDF3.0 NNLO	4.8550350808	1981.8	0.9751	0.82106	1586.66000797
364115	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV0 70 CFilterBVeto	4940500	Sherpa	NNPDF3.0 NNLO	22.0822865422	1980.8	0.9751	0.11295	218.160449136
364116	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV0 70 BFilter	7883600	Sherpa	NNPDF3.0 NNLO	62.3454455468	1981.7	0.9751	0.063809	123.301682947
364117	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV70 140 CVetoBVeto	5925000	Sherpa	NNPDF3.0 NNLO	77.6616159528	110.5	0.9751	0.69043	74.3928313765
364118	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV70 140 CFilterBVeto	1972600	Sherpa	NNPDF3.0 NNLO	97.0003619098	110.63	0.9751	0.18382	19.8296400357
364119	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV70 140 BFilter	5855000	Sherpa	NNPDF3.0 NNLO	463.844185493	110.31	0.9751	0.11443	12.3084662448
364120	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV140 280 CVetoBVeto	4949000	Sherpa	NNPDF3.0 NNLO	197.722621181	40.731	0.9751	0.61452	24.4067667684
364121	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV140 280 CFilterBVeto	2962600	Sherpa	NNPDF3.0 NNLO	316.111997843	40.67	0.9751	0.23044	9.13863212948
364122	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV140 280 BFilter	12330900	Sherpa	NNPDF3.0 NNLO	2032.5277594	40.643	0.9751	0.14927	5.91571777281
364123	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV280 500 CVetoBVeto	1932800	Sherpa	NNPDF3.0 NNLO	396.941483556	8.6743	0.9751	0.56134	4.74798769611
364124	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV280 500 CFilterBVeto	988900	Sherpa	NNPDF3.0 NNLO	433.732058608	8.6711	0.9751	0.26294	2.22320755605
364125	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV280 500 BFilter	1976850	Sherpa	NNPDF3.0 NNLO	1322.86442976	8.6766	0.9751	0.17223	1.45716098463
364126	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV500 1000	2973000	Sherpa	NNPDF3.0 NNLO	1644.26746308	1.8081	0.9751	1.0	1.76307831
364127	Sherpa 221 NNPDF30NNLO Zee MAXHTPTV1000 E CMS	988000	Sherpa	NNPDF3.0 NNLO	6650.06394292	0.14857	0.9751	1.0	0.144870607
364128	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV0 70 CVetoBVeto	7907000	Sherpa	NNPDF3.0 NNLO	4.85769756199	1981.6	0.9751	0.82142	1587.19549779
364129	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV0 70 CFilterBVeto	4941000	Sherpa	NNPDF3.0 NNLO	22.0697176888	1978.8	0.9751	0.11314	218.306784343
364130	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV0 70 BFilter	7890600	Sherpa	NNPDF3.0 NNLO	61.7741911262	1981.8	0.9751	0.064453	124.552404811
364131	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV70 140 CVetoBVeto	5935500	Sherpa	NNPDF3.0 NNLO	78.0718037803	110.37	0.9751	0.68883	74.1331155392
364132	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV70 140 CFilterBVeto	1961200	Sherpa	NNPDF3.0 NNLO	97.0301270827	110.51	0.9751	0.1829	19.7089932529
364133	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV70 140 BFilter	5912550	Sherpa	NNPDF3.0 NNLO	480.932428087	110.87	0.9751	0.110886	11.9878119426
364134	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV140 280 CVetoBVeto	4956000	Sherpa	NNPDF3.0 NNLO	199.81121944	40.781	0.9751	0.60821	24.185807051
364135	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV140 280 CFilterBVeto	2973000	Sherpa	NNPDF3.0 NNLO	318.709713854	40.74	0.9751	0.22897	9.09596467878
364136	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV140 280 BFilter	4932950	Sherpa	NNPDF3.0 NNLO	900.322268996	40.761	0.9751	0.13442	5.34266418886
364137	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV280 500 CVetoBVeto	1973000	Sherpa	NNPDF3.0 NNLO	411.797480449	8.5502	0.9751	0.56036	4.67188943921

Not reviewed, for internal circulation only

Table 106: Nominal MC samples.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364138	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV280 500 CFilterBVeto	986000	Sherpa	NNPDF3.0 NNLO	433.287517426	8.6707	0.9751	0.26245	2.21896214715
364139	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV280 500 BFilter	1974950	Sherpa	NNPDF3.0 NNLO	1314.14727158	8.6804	0.9751	0.17313	1.46541699447
364140	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV500 1000	2944800	Sherpa	NNPDF3.0 NNLO	1627.32095491	1.8096	0.9751	1.0	1.76454096
364141	Sherpa 221 NNPDF30NNLO Ztautau MAXHTPTV1000 E CMS	980000	Sherpa	NNPDF3.0 NNLO	6606.44465417	0.14834	0.9751	1.0	0.144646334
$W + jets$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364156	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV0 70 CVetoBVeto	24723000	Sherpa	NNPDF3.0 NNLO	1.56772318768	19143.0	0.9702	0.8238	15300.0572987
364157	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV0 70 CFilterBVeto	9847000	Sherpa	NNPDF3.0 NNLO	3.94926016845	19121.0	0.9702	0.1304	2419.07572368
364158	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV0 70 BFilter	17226200	Sherpa	NNPDF3.0 NNLO	20.405404216	19135.0	0.9702	0.044118	819.040831686
364159	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV70 140 CVetoBVeto	14788000	Sherpa	NNPDF3.0 NNLO	23.1996228452	944.85	0.9702	0.67463	618.428915666
364160	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV70 140 CFilterBVeto	9853800	Sherpa	NNPDF3.0 NNLO	44.7969889819	937.78	0.9702	0.23456	213.410699631
364161	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV70 140 BFilter	19639000	Sherpa	NNPDF3.0 NNLO	274.827498956	944.63	0.9702	0.075648	69.3298810068
364162	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV140 280 CVetoBVeto	9882000	Sherpa	NNPDF3.0 NNLO	46.4914010851	339.54	0.9702	0.62601	206.221283425
364163	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV140 280 CFilterBVeto	7408000	Sherpa	NNPDF3.0 NNLO	75.2561266792	340.06	0.9702	0.28947	95.5037405876
364164	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV140 280 BFilter	9826000	Sherpa	NNPDF3.0 NNLO	28.9391529717	339.54	0.9702	0.11229	36.9147888
364165	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV280 500 CVetoBVeto	4940000	Sherpa	NNPDF3.0 NNLO	125.436572946	72.067	0.9702	0.54647	38.208856376
364166	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV280 500 CFilterBVeto	2958000	Sherpa	NNPDF3.0 NNLO	129.0699178	72.198	0.9702	0.31743	22.234860368
364167	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV280 500 BFilter	2959500	Sherpa	NNPDF3.0 NNLO	308.003993468	72.045	0.9702	0.13337	9.32230412883
364168	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV500 1000	5910500	Sherpa	NNPDF3.0 NNLO	393.770819454	15.01	0.9702	1.0	14.562702
364169	Sherpa 221 NNPDF30NNLO Wmunu MAXHTPTV1000 E CMS	3959000	Sherpa	NNPDF3.0 NNLO	3207.22618276	1.2344	0.9702	1.0	1.19761488
364170	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV0 70 CVetoBVeto	24740000	Sherpa	NNPDF3.0 NNLO	1.56883756535	19127.0	0.9702	0.82447	15299.7024868
364171	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV0 70 CFilterBVeto	9853500	Sherpa	NNPDF3.0 NNLO	3.953039329	19130.0	0.9702	0.1303	2418.3583578
364172	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV0 70 BFilter	17242400	Sherpa	NNPDF3.0 NNLO	20.4139516402	19135.0	0.9702	0.044141	819.467821557
364173	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV70 140 CVetoBVeto	14660500	Sherpa	NNPDF3.0 NNLO	23.2587435151	942.58	0.9702	0.66872	611.538499092
364174	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV70 140 CFilterBVeto	9818400	Sherpa	NNPDF3.0 NNLO	45.5631726263	945.67	0.9702	0.22787	209.068226178
364175	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV70 140 BFilter	9801900	Sherpa	NNPDF3.0 NNLO	100.28754283	945.15	0.9702	0.10341	94.8253702473
364176	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV140 280 CVetoBVeto	9879000	Sherpa	NNPDF3.0 NNLO	48.7043751025	339.81	0.9702	0.59691	196.791474684
364177	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV140 280 CFilterBVeto	7410000	Sherpa	NNPDF3.0 NNLO	75.2717206264	339.87	0.9702	0.28965	95.5097338041
364178	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV140 280 BFilter	9880900	Sherpa	NNPDF3.0 NNLO	29.1059856251	339.48	0.9702	0.11229	36.9965304

Not reviewed, for internal circulation only

Table 106: Nominal MC samples.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
364179	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV280 500 CVetoBVeto	4923800	Sherpa	NNPDF3.0 NNLO	125.468709773	72.084	0.9702	0.54441	38.0738015769
364180	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV280 500 CFilterBVeto	2963400	Sherpa	NNPDF3.0 NNLO	129.708896015	72.128	0.9702	0.31675	22.1657169888
364181	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV280 500 BFilter	2958000	Sherpa	NNPDF3.0 NNLO	306.317350162	72.113	0.9702	0.13391	9.36888360547
364182	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV500 1000	5916800	Sherpa	NNPDF3.0 NNLO	388.649500788	15.224	0.9702	1.0	14.7703248
364183	Sherpa 221 NNPDF30NNLO Wenu MAXHTPTV1000 E CMS	3947000	Sherpa	NNPDF3.0 NNLO	3200.09729204	1.2334	0.9702	1.0	1.19664468
364184	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV0 70 CVetoBVeto	24784000	Sherpa	NNPDF3.0 NNLO	1.56866295484	19152.0	0.9702	0.82495	15328.6190165
364185	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV0 70 CFilterBVeto	9865600	Sherpa	NNPDF3.0 NNLO	3.98248214869	19153.0	0.9702	0.12934	2403.4269992
364186	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV0 70 BFilter	17273200	Sherpa	NNPDF3.0 NNLO	20.2130975747	19163.0	0.9702	0.044594	829.089088304
364187	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV70 140 CVetoBVeto	14808500	Sherpa	NNPDF3.0 NNLO	23.1909855548	947.65	0.9702	0.67382	619.516866415
364188	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV70 140 CFilterBVeto	9860000	Sherpa	NNPDF3.0 NNLO	46.86705154	946.73	0.9702	0.22222	204.11294685
364189	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV70 140 BFilter	9857000	Sherpa	NNPDF3.0 NNLO	100.562851001	943.3	0.9702	0.10391	95.0973575706
364190	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV140 280 CVetoBVeto	9899000	Sherpa	NNPDF3.0 NNLO	48.9242450604	339.36	0.9702	0.59622	196.303689268
364191	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV140 280 CFilterBVeto	7365000	Sherpa	NNPDF3.0 NNLO	74.7127079545	339.63	0.9702	0.29025	95.6399947965
364192	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV140 280 BFilter	9834000	Sherpa	NNPDF3.0 NNLO	257.927814235	339.54	0.9702	0.11229	36.9907635913
364193	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV280 500 CVetoBVeto	4931200	Sherpa	NNPDF3.0 NNLO	125.39558066	72.065	0.9702	0.54569	38.1532603845
364194	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV280 500 CFilterBVeto	2956400	Sherpa	NNPDF3.0 NNLO	129.786408974	71.976	0.9702	0.31648	22.1001513385
364195	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV280 500 BFilter	2954100	Sherpa	NNPDF3.0 NNLO	305.484551818	72.026	0.9702	0.13426	9.38203847935
364196	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV500 1000	5945000	Sherpa	NNPDF3.0 NNLO	395.121627011	15.046	0.9702	1.0	14.5976292
364197	Sherpa 221 NNPDF30NNLO Wtaunu MAXHTPTV1000 E CMS	3946000	Sherpa	NNPDF3.0 NNLO	3197.99011265	1.2339	0.9702	1.0	1.19712978

Z.2 Samples for estimation of systematic uncertainties

Table 107: MC samples used for modelling systematics.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
$t\bar{t}\gamma$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
410404	MadGraphPythia8EvtGen A14NNPDF23Var3cUp ttgamma nonallhad	4899000	MadGraph +Pythia8 +EvtGen	A14 NNPDF23LO	1059.632729868	4.6233	1.0	1.0	4.6233
410405	MadGraphPythia8EvtGen A14NNPDF23Var3cDown ttgamma nonallhad	4927000	MadGraph +Pythia8 +EvtGen	A14 NNPDF23LO	1065.689010014	4.6233	1.0	1.0	4.6233
410395	MGH7EG H7UE ttgamma nonallhadronic	4804000	MadGraph +Herwig7 +EvtGen	MMHT2014	1039.084636515	4.6233	1.0	1.0	4.6233
$t\bar{t}$ MC15c									
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
410250	Sherpa 221 NNPDF30NNLO ttbar SingleLeptonP MEPS NLO	13587000	Sherpa	NNPDF3.0NNLO	85.582010582	158.760	1.1484	1.0	182.319984

Not reviewed, for internal circulation only

Table 107: MC samples used for modelling systematics.

DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
410251	Sherpa 221 NNPDF30NNLO ttbar SingleLeptonM MEPS NLO	14923000	Sherpa	NNPDF3.0NNLO	93.87305781	158.970	1.1484	1.0	182.561148
410252	Sherpa 221 NNPDF30NNLO ttbar dilepton MEPS NLO	9944000	Sherpa	NNPDF3.0NNLO	130.36695203	76.277	1.1484	1.0	87.596507
410511	PowhegPythia8EvtGen A14v3cUp hdamp517p5 ttbar nonallhad		Powheg +Pythia8 +EvtGen	A14 NNPDF23LO		824.51	1.0088	0.54383	452.339134105
410512	PowhegPythia8EvtGen A14v3cDo hdamp258p75 ttbar nonallhad		Powheg +Pythia8 +EvtGen	A14 NNPDF23LO		640.4	1.2988	0.54388	452.377629
				$Z\gamma$ MC15c					
DSID	explanation	nevents	genName	genTune	Sample Lumi [fb-1]	AMIXsec [pb]	K-factor	FiltEff	σ×Kfactor×FiltEff [pb]
345775	aMcAtNloPythia8 NNPDF30NLO FxFx ee-gamma HT0 125	1948000	aMcAtNlo Pythia8	A14 NNPDF23LO	93.187906621	20.904	1.0	1.0	20.904
345777	aMcAtNloPythia8 NNPDF30NLO FxFx ee-gamma HTGT125 MjjLT500	391000	aMcAtNlo Pythia8	A14 NNPDF23LO	332.241609497	10.176	1.0	0.11565	1.1768544
345778	aMcAtNloPythia8 NNPDF30NLO FxFx ee-gamma HTGT125 MjjGT500	99000	aMcAtNlo Pythia8	A14 NNPDF23LO	749.068567285	10.179	1.0	0.012984	0.132164136
345779	aMcAtNloPythia8 NNPDF30NLO FxFx mu-mugamma HT0 125	1953000	aMcAtNlo Pythia8	A14 NNPDF23LO	93.213058419	20.952	1.0	1.0	20.952
345781	aMcAtNloPythia8 NNPDF30NLO FxFx mumugamma HTGT125 MjjLT500	393000	aMcAtNlo Pythia8	A14 NNPDF23LO	328.843168141	10.191	1.0	0.11727	1.19509857
345782	aMcAtNloPythia8 NNPDF30NLO FxFx mumugamma HTGT125 MjjGT500	96000	aMcAtNlo Pythia8	A14 NNPDF23LO	730.468599335	10.187	1.0	0.012901	0.131422487
345783	aMcAtNloPythia8 NNPDF30NLO FxFx tautaugamma pty 10	2244000	aMcAtNlo Pythia8	A14 NNPDF23LO	80.360979802	27.924	1.0	1.0	27.924