technische universität
dortmund

Dissertation zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften

Illuminating the top quark

Observation of single-top-quark production in association with a photon at the ATLAS experiment

Björn Wendland
geboren in Duisburg

Arbeitsgruppe Kröninger
Fakulät Physik
Technische Universität Dortmund

-2025-

Der Fakultät für Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften vorgelegte Dissertation.

Erstgutachter:	Prof. Dr. Kevin Kröninger
Zweitgutachter:	Dr. Dominik Mitzel
Vorsitzender der Prüfungskommision:	Prof. Dr. Marc Aßmann
Vertretung der wiss. Mitarbeiter:	Dr. Tim. Ruhe

Datum des Einreichens der Arbeit:	27.11.2025
Datum der mündlichen Prüfung:	06.03.2026

Abstract

The rich datasets collected by the ATLAS and CMS experiments in proton–proton collisions of the Large Hadron Collider at a center-of-mass energy of 13 TeV allow high-precision measurements of top-quark properties and enable the experimental exploration of several rare top-quark processes for the first time. The t-channel single-top-quark production in association with a photon is such a process. Measurements of this signal process directly probe the top-quark–photon interaction, which is a cornerstone of electroweak physics. In this work, this process is examined using events selected from the data collected by the ATLAS experiment that contain either an electron or a muon, at least one photon, high missing transverse momentum, and a particle jet initiated by a bottom quark.

Studying final states with photons, such as that of the signal process, using proton–proton collision data is challenging, as the majority of reconstructed photon candidates are background photons arising from hadronic activity. This requires excellent rejection of such candidates, which is achieved by applying photon-isolation and photon-identification requirements. Precise measurements of their efficiencies are vital for ensuring consistently high performance and for accurate modeling of the contribution of signal photons in simulation. This work introduces improvements to the method used for measuring the photon-identification efficiency at high photon energies and presents the respective measurement results.

The largest background contributions to the selected data arise from the production of top-quark pairs in association with a photon, from events where an electron is misidentified as a photon, and from the production of a W boson in association with a photon and particle jets. Background contributions arising from events with background leptons and background photons are estimated using data-driven techniques.

Only a small fraction of the selected events originate from the signal process. Deep neural networks are employed to efficiently discriminate between signal and background contributions. This thesis presents the first observation of the signal process, achieved with a statistical significance of 9.3 standard deviations using these neural networks. This observation represents a milestone in electroweak physics. Two fiducial cross sections of the signal process are measured with a precision of 11%. Their values are found to be compatible with Standard Model predictions at next-to-leading order accuracy in quantum chromodynamics at the level of 2.0 and 2.1 standard deviations.

Zusammenfassung

Die umfangreichen Datensätze, die von den ATLAS- und CMS-Experimenten in Proton–Proton-Kollisionen des Large Hadron Colliders bei einer Schwerpunktsenergie von 13 TeV aufgenommen wurden, ermöglichen hochpräzise Messungen der Eigenschaften des Top-Quarks und die erstmalige experimentelle Erforschung einiger seltener Top-Quark-Prozesse. Die t-Kanal Einzelproduktion eines Top-Quarks in Assoziation mit einem Photon stellt einen solchen Prozess dar. Messungen dieses Signalprozesses sind direkte Untersuchungen der Wechselwirkung zwischen Top-Quark und Photon, die ein zentraler Bestandteil der elektroschwachen Physik ist. In dieser Arbeit wird dieser Prozess anhand von Ereignissen untersucht, die aus den Daten des ATLAS-Experiments ausgewählt werden und entweder ein Elektron oder ein Myon, mindestens ein Photon, einen hohen fehlenden transversalen Impuls und einen Teilchenjet, der durch ein Bottom-Quark initiiert wurde, aufweisen.

Die Untersuchung von Endzuständen mit Photonen, wie sie der Signalprozess aufweist, mittels Daten von Proton–Proton-Kollisionen stellt eine Herausforderung dar, da der Großteil der rekonstruierten Photon-Kandidaten Untergrund-Photonen sind, die aus hadronischer Aktivität stammen. Dies erfordert eine starke Unterdrückung solcher Kandidaten, was durch die Anwendung von Photon-Isolations- und Photon-Identifikationskriterien erreicht wird. Präzise Messungen ihrer Effizienzen sind unerlässlich, um eine durchgehend hohe Leistung sicherzustellen und eine genaue Modellierung des Beitrags von Signal-Photonen in Simulationen zu gewährleisten. In dieser Arbeit werden Verbesserungen der Methode zur Messung der Effizienz der Photon-Identifikation bei hohen Photon-Energien samt der zugehörigen Messergebnisse vorgestellt.

Die größten Untergrundbeiträge im untersuchten Datensatz stammen von der Produktion von Top-Quark-Paaren in Assoziation mit einem Photon, von Ereignissen, bei denen ein Elektron fälschlicherweise als Photon identifiziert wurde, und von der Produktion eines W-Bosons in Assoziation mit einem Photon und Teilchenjets. Untergrundbeiträge, die aus Ereignissen mit Untergrund-Leptonen und Untergrund-Photonen stammen, werden datengetrieben abgeschätzt.

Nur ein kleiner Bruchteil der ausgewählten Ereignisse stammt vom Signalprozess. Tiefe neuronale Netze werden eingesetzt, um effizient Signal- und Untergrundbeiträge voneinander zu trennen. In dieser Arbeit wird die erstmalige Beobachtung des Signalprozesses präsentiert, die mithilfe dieser neuronalen Netze mit einer statistischen Signifikanz von 9,3 Standardabweichungen erreicht wird. Diese Beobachtung stellt einen Meilenstein in der elektroschwachen Physik dar. Zwei Wirkungsquerschnitte des Signalprozesses werden in Referenzphasenräumen mit einer Präzision von 11% gemessen. Ihre Werte stimmen innerhalb von 2,0 beziehungsweise 2,1 Standardabweichungen mit den Vorhersagen des Standardmodells auf nächstführender Ordnung in der Quantenchromodynamik überein.

Contents

List of abbreviations

4FS four-flavor scheme

5FS five-flavor scheme

AR analysis region

BPL binned profile likelihood

BSM beyond the Standard Model

CKM Cabibbo-Kobayashi-Maskawa

CR control region

CSC cathode strip chamber

CV conversion vertex

DR diagram removal

DS diagram subtraction

ECAL electromagnetic calorimeter

EDM electric dipole moment

EFT effective field theory

EM electromagnetic

ePSR extended pre-selection region

EW electroweak

EWSB electroweak symmetry breaking

FCAL forward calorimeter

FCNC flavor-changing neutral current

GBDT gradient boosting decision tree

HCAL hadronic calorimeter

HF heavy flavor

HLT high-level trigger

HL-LHC High-Luminosity Large Hadron Collider

IBL Insertable B-Layer

ID inner detector

IP interaction point

IVF iterative vertex finding

JER jet energy resolution

JES jet energy scale

JVT Jet Vertex Tagger

LAr liquid argon

LF light flavor

LH likelihood

LHC Large Hadron Collider

LO leading order

LOFO leave-one-feature-out

MC Monte Carlo

MDM magnetic dipole moment

MDT monitored drift tube

ME matrix element

MET missing transverse energy

ML machine learning

MS muon spectrometer

NLO next-to-leading order

NN neural network

NNLL next-to-next-to-leading logarithm

NNLO next-to-next-to-leading order

NP nuisance parameter

NSV narrow-strip variable

NTR non-tight region

OS opposite sign

PD pseudo-dataset

PDF parton distribution function

p.d.f. probability density function

PID photon identification

PS parton shower

PSR pre-selection region

PV primary vertex

QCD quantum chromodynamics

QED quantum electrodynamics

QFT quantum field theory

ROC receiver operating characteristic

RoI region of interest

RPC resistive plate chamber

RTR relaxed-tight region

RTV relaxed-tight variable

SCT semiconductor tracker

SF scale factor

SM Standard Model of particle physics

SR signal region

SSB spontaneous symmetry breaking

SSV shower shape variable

SV secondary vertex

TGC thin gap chamber

TR transition radiation

TRT transition radiation tracker

TST track soft term

UE underlying event

UV ultraviolet

WP working point

Preamble

The physics analyses that are presented in this thesis were conducted in the context of the physics research program of the ATLAS collaboration. The ATLAS collaboration centrally provides its researchers a comprehensive infrastructure to perform physics analyses. In the following, it is discussed how this infrastructure was used for this thesis. Additionally, the contributions of other researchers of the collaboration are placed into context.

The reconstruction, identification, and isolation algorithms presented in Section 4.3 are developed, maintained, and provided by dedicated subgroups of the ATLAS collaboration along with associated calibrations including their uncertainties. The data of the proton–proton collisions collected with the ATLAS experiment are centrally provided by the collaboration to its researchers in a common digital format, where the algorithms to reconstruct physics objects (cf. Section 4.3) are already applied to the detector signatures.

Simulated samples of physics processes are centrally produced and provided in the common digital format as well. The author was the main developer of the signal "NLO $tq\gamma$" sample and contributed to the production of the signal "LO inclusive" sample (cf. Chapter 6). Both samples are centrally provided by the ATLAS collaboration. The simulation of the detector response is uniformly applied to all centrally provided samples. Additional samples used in Chapter 6 that are not centrally provided were produced by the author. The author did not contribute to the development of any other samples that are used and described in this work.

In this thesis, two physics analyses are presented. The analysis presented in Section 5.3 was conducted within the "Photon Identification" group. This is a subgroup of researchers of the ATLAS collaboration that focuses on developing algorithms to identify detector signatures caused by a photon, to measure their performance, and to provide these algorithms along with the associated calibrations to all researchers of the collaboration. The measurement of the efficiency of the photon identification algorithm at high photon energies is presented and was accompanied by continuous feedback from the group. Data and simulated samples used in these studies were processed through a commonly used software framework for event preselection, and for applying identification and isolation algorithms to the reconstructed physics objects. This software framework was developed and maintained by the "HGam" group, which focuses on studying processes involving at least one Higgs boson decay into photons. The method to measure the efficiencies was developed by the ATLAS collaboration, and this work builds upon previous iterations of the measurement [1–3]. The existing software framework dedicated to this measurement was adapted and improved by the author and provided to the collaboration for future studies and measurements. The presented improvements to the methods were individually developed by the author. The results of the measurement were presented by the author at the "The European Physical Society Conference on High Energy Physics" in 2021 conference as a poster contribution [4] and in the proceedings [5]. The ATLAS collaboration published these measurement results in the "Journal of High Energy Physics" [6].

The second research project presented focuses on the observation and fiducial cross-section measurements of the single production of top quarks in association with a photon. It is presented in Chapters 6–10 and represents the main project of this thesis. It was conducted by an international team of researchers of the ATLAS collaboration consisting of the author, Harish Potti, Johannes Erdmann, and Peter Onyisi within the "Top" group, which is a subgroup dedicated to the exploration of physics processes of the top quark. The related studies and results were compiled by the author and Harish Potti, while Johannes Erdmann and Peter Onyisi supervised this work. A common framework developed and maintained by the Top group was used to process data and simulated samples for preselection, and for applying identification and isolation algorithms to the reconstructed physics objects. For further processing of the samples, Harish Potti and the author developed a shared software framework to create a common minimal data format for the individual aspects of the analysis. Separate software projects were developed for these individual aspects. The analysis strategy, the corresponding event selection, as well as the background

modeling and categorization of the simulated events, presented in Section 7–7.2 were developed jointly by the team members and are based on previous works conducted by other researchers within the Top group. The preselection of events, presented in Section 7.3, was mainly implemented by the author and Harish Potti together, while the optimization of the event selection and the definition of the final selection criteria were carried out primarily by the author. The results discussed in Section 8.2 and related studies were compiled by Harish Potti and are presented in Ref. [7]. A preliminary setup of the statistical analysis presented in Sections 7.5 and 7.6, and Chapter 10 was developed by the author and Harish Potti together. This preliminary setup is presented in Ref. [7]. This work presents the final setup, refined by the author. The strategy for separating signal from background contributions through neural networks presented in Section 9.1 was defined by the author and Harish Potti together. The method to select input features for the neural networks, discussed in Section 9.2, and a preliminary version of the software project for training these neural networks were developed by Harish Potti. The author extended the software project by the optimization steps for the neural networks presented in Sections 9.3–9.6. All other aspects of the analysis were individually conducted by the author. The entire analysis was reviewed by the team, the top group, and designated members of the ATLAS collaboration. A preliminary result of the work was published by the ATLAS collaboration for the Moriond 2022 EW conference [8] and presented by the author at the LHCP 2022 conference [9]. The final result was published in the journal "Physical Review Letters" as an Editors' Suggestion [10].

The author also contributed to the "Search for flavour-changing neutral-current couplings between the top quark and the photon with the ATLAS detector at $\sqrt{s} = 13\,\text{TeV}$" [11]. The author adapted the method for estimating background contributions, where hadronic activity in the detector is misidentified as a photon discussed in Section 8.3 to this analysis and provided the related results. The author further implemented the modeling of the single-top-quark production in association with a photon discussed in Chapter 6 into this analysis.

1 Introduction

The concept of having a limited number of building blocks from which all matter in the universe is constructed is one of the oldest ideas of natural science. Documented theories of atomism, a branch of natural philosophy, date back at least to the fifth century BCE, such as those attributed to Leucippus and Democritus. These concepts were further explored and developed over the following centuries, but they remained abstract and philosophical. However, they gained credibility in physics in the 19th century primarily due to the work of John Dalton on stoichiometry [12]. Dalton concluded that each element is made up of particles of a single type with unique properties, called atoms. Since then, theoretical and experimental studies of these fundamental building blocks, known as elementary particles, and their interactions have become central elements in physics.

In subsequent decades, physicists extensively studied elementary particles and their interactions, leading to a dynamic evolution in the definition of elementary particles and of the models describing their interactions. Initially, it was discovered that atoms consist of electrons and a nucleus. This finding ultimately led to the creation of two new branches of physics: atomic physics and nuclear physics. It was later discovered that the nucleus is not elementary but is composed of nucleons, the protons and neutrons. These were later found to belong to a group of composite particles known as hadrons, which are composed of quarks.

Particle physics is the branch of physics that studies fundamental interactions between particles at the subatomic level, such as hadrons, quarks, and electrons. In contrast, the structure and dynamics of atomic nuclei are the focus of nuclear physics. The main objective of theoretical particle physics is to formulate mathematical models that describe these particles and their interactions. Experimental particle physics focuses on designing and conducting experiments that test the predictions of these models and gather empirical data to guide their development. The ultimate goal of particle physics is to reveal all fundamental building blocks of the universe and to formulate a unified theoretical framework that describes all observable phenomena of the universe at the subatomic level.

The 20th century was characterized by rapid development in both experimental and theoretical particle physics. Several elementary particles and numerous composite subatomic particles were experimentally discovered. These were either predicted by theory, or their discovery guided the evolution of theory. The Standard Model of particle physics (SM) serves as the mathematical foundation and comprises the current knowledge of elementary particles and their interactions. It describes the electromagnetic, weak, and strong interactions between elementary particles, which are mediated by gauge bosons.

The Large Hadron Collider (LHC), located at CERN in Geneva, the European Organization for Nuclear Research, is the most powerful particle collider to date and one of the most advanced facilities for experimental studies in particle physics [13]. In 2012, two of its main experiments, the ATLAS and CMS experiments [14, 15], observed the last predicted elementary particle of the SM, the Higgs boson, in collisions of highly energetic proton beams [16, 17]. This observation marked the completion of the SM.

The SM is one of the most extensively tested and fundamental theories in physics and has proven to be a powerful and successful framework, as no experimental result has significantly contradicted its predictions to date. However, there are conceptual problems within the SM and observed phenomena that cannot be explained by it. Describing them requires a model defined at the subatomic level. These limitations demonstrate that the SM does not satisfy the aforementioned ultimate goal of particle physics and needs to be extended or superseded by a more comprehensive theory. It is not clear yet how to proceed with this task. The predictions of the SM are tested as precisely as possible to deepen the understanding of the fundamental interactions and to find patterns that shed light on this path. This ongoing effort requires advancements in both experimental and theoretical particle physics.

The top quark is the most massive known elementary particle.[1] It was discovered in 1995 by the CDF [18] and DØ [19] experiments located at the Tevatron, a powerful particle collider located at the Fermilab research center. It has unique features compared to other elementary particles and is strongly connected to the Higgs boson due to its large mass. Furthermore, it is linked to many promising theories that extend the SM. These features make the top quark an exceptionally powerful experimental probe. While it has been extensively studied since its discovery, it has only recently become possible to study its properties precisely and to explore rare top-quark processes through the experiments at the LHC.

In proton–proton (pp) collisions at center-of-mass energies ($\sqrt{s}$) typical at the LHC, which are in the range of several TeV, top quarks are predominantly produced in pairs or singly, with pair production occurring at a higher frequency. To study properties related to the electroweak (EW) sector, which describes electromagnetic and weak interactions of elementary particles, top-quark production in association with a boson of this sector, the photon, the Z boson, the W boson, and the Higgs boson, is of special interest. While top-quark pair production has been observed in association with all of these bosons, single-top-quark production has only been observed in association with the Z and W bosons.

This thesis explores t-channel single-top-quark production in association with a photon using data collected by the ATLAS experiment during the years 2015–2018 in pp collisions at $\sqrt{s} = 13\,\text{TeV}$, corresponding to an integrated luminosity of $139\,\text{fb}^{-1}$. Measurements of this rare top-quark process directly probe the interaction of the top quark with the photon. Therefore, they represent important tools to deepen the understanding of EW physics and to find deviations from the SM predictions. The CMS collaboration reported evidence corresponding to a statistical significance of 4.4 standard deviations for this process using a dataset of pp collisions at $\sqrt{s} = 13\,\text{TeV}$ corresponding to $35.9\,\text{fb}^{-1}$ collected by the CMS experiment [20]. A fiducial cross section of this process was measured with a precision of 30% and is compatible with the SM prediction. This thesis presents the first observation and two fiducial cross-section measurements of this process, published by the ATLAS collaboration [10]. This observation represents a milestone in EW physics.

Studying processes with photons in the final state using pp collisions is challenging, as the majority of detected photons are fake photons from hadronic activity. Photon-isolation and photon-identification algorithms are used to suppress these fake contributions. The measurement of the photon-identification efficiency at high energies using this dataset, reported by the ATLAS collaboration [6], is also presented.

This thesis is organized as follows: Chapter 2 provides an overview of the SM and its limitations. It summarizes the current knowledge about top-quark properties and illuminates the interaction between top quarks and photons within the SM and in scenarios beyond it. Chapter 3 starts with a brief overview of the LHC and discusses the production of top quarks in LHC pp collisions, with a specific focus on photon-associated single-top-quark production. Chapter 4 provides an overview of how physics studies are conducted with the ATLAS experiment using pp-collision data, including details about the ATLAS detector, the simulation of pp collisions, and the reconstruction of physics objects from detector signatures. Chapter 5 presents the photon-identification algorithm and the measurement of the photon-identification efficiency for signal photons at high energies. In Chapter 6, the simulation of t-channel single-top-quark production in association with a photon in pp collisions is discussed and comprehensive studies of properties of this process are performed. Chapter 7 outlines the analysis strategy and details the event selection, the modeling of background contributions, the event categorization, and the strategies for the fiducial cross-section measurement and statistical analysis. Chapter 8 focuses on the data-driven estimations of background contributions that include misidentified physics objects. In Chapter 9, the fitting and optimization of deep neural networks for separating signal from background contributions are discussed. Chapter 10 presents the optimization and results of the statistical analysis, including the fiducial cross-section measurements and observation of the photon-associated single-top-quark production. Finally, Chapter 11 provides a summary of the thesis and an outlook for future investigations of t-channel single-top-quark production in association with a photon.

[1] Natural units are used throughout this thesis.

2 Top-quark–photon interaction in the Standard Model and beyond

Particle physics is the branch of physics that studies the fundamental interactions of subatomic particles. It does not include the examination of the structure and dynamics of atomic nuclei, which is the focus of nuclear physics. The ultimate goal of particle physics is to reveal all fundamental building blocks of the universe and to describe all their interactions within a single theoretical framework. The SM is the theoretical framework that comprises the current knowledge of these building blocks, the elementary particles, and their interactions. An overview is presented in the following section. The SM is characterized by its outstanding success in predicting and describing observations in particle physics experiments. However, this chapter also discusses its limitations, which demonstrate that the SM is not the most comprehensive theoretical framework that particle physicists seek to develop. The top quark, one of the elementary particles of the SM, is considered to play an important role in advancing this quest and is a main subject of this thesis. It is introduced by briefly reviewing its main properties according to the SM and their current experimental status. This thesis illuminates the top quark mainly with respect to its interaction with the photon, the mediator of the electromagnetic (EM) interaction. This chapter discusses properties of the top-quark–photon interaction in the SM and highlights its connection to theoretical models describing beyond the Standard Model (BSM) physics.

2.1 Overview of the Standard Model of particle physics

The SM is the theoretical framework used to describe and predict phenomena at the level of elementary particles. It is formulated as a renormalizable [21, 22] gauge quantum field theory (QFT), in which the mathematical structures describing the EM, strong, and weak interactions between elementary particles are induced by gauge groups. The dynamics and kinematics are described by a Lagrange density, or Lagrangian $\mathcal{L}$, which must respect Lorentz and gauge invariance, i.e., the Lagrangian is invariant under local symmetry transformations associated with the gauge groups. The elementary particle content of the SM consists of fermions with spin $1/2$ and bosons with spin 0 or 1. Each of these particles corresponds to a field in the Lagrangian. The gauge interactions between elementary particles are mediated by vector bosons with spin 1. Each of these interactions is related to a charge, which an elementary particle must carry to undergo the corresponding interaction. According to Noether's theorem, these charges are conserved in SM interactions. The properties of the fermion fields under gauge symmetry transformations determine whether the corresponding fermion carries the respective charge. These transformation properties are assigned so that they align with measurements and observed structures of the interactions. Moreover, for each fermion, there exists an anti-fermion carrying charges of the opposite sign.[2] Each interaction is furthermore connected to a coupling constant, which corresponds to the strength of the interaction.

The strong interaction is described by quantum chromodynamics (QCD) with $\mathrm{SU(3)_C}$ as the underlying gauge symmetry group [23–26], where the subscript C stands for color. Only elementary particles that carry color charge, which can take three different values, conventionally labeled red, blue, and green, are allowed to participate in this interaction. Color-charged fermions are referred to as quarks and their fields are $\mathrm{SU(3)_C}$ triplets. Color-neutral fermions are referred to as leptons and their fields are $\mathrm{SU(3)_C}$ singlets. The strong interaction is mediated by eight vector bosons, known as gluons, which carry color charge and therefore interact with one another. The coupling constant of the strong interaction is denoted by g_s.

The weak and EM interactions are described by a unified theory, the EW theory [27–29], in which $\mathrm{SU(2)_L \times U(1)}_Y$ is the underlying gauge group. The subscript L denotes left-handed chirality, and the subscript Y stands for hypercharge (as defined below). Four vector bosons mediate these interactions:

[2] Throughout this thesis, any reference to a particle implicitly includes its corresponding anti-particle. An explicit distinction is made only when it is relevant or cannot be inferred from the context.

W^1, W^2, W^3, and B. The B boson is the gauge boson of the $\mathrm{U}(1)_Y$ group, while the others are the gauge bosons of the $\mathrm{SU}(2)_\mathrm{L}$ group. The corresponding coupling constants are denoted by g' and g, respectively. All elementary particles that carry weak isospin T undergo interactions with the W^1, W^2, and W^3 bosons, while particles that carry weak hypercharge Y interact with the B boson. The weak hypercharge is related to the electric charge Q and the third component of the weak isospin T_3 by $Y = 2\,(Q - T_3)$. Quarks carry either an electric charge of $+\frac{2}{3}$ or $-\frac{1}{3}$, referred to as up-type quarks q^u and down-type quarks q'^d, respectively. The prime denotes that the down-type quark is expressed in the weak-interaction basis. Leptons are categorized into electrically charged leptons ℓ with an electric charge of -1 and into neutrinos ν carrying no electric charge. The left-handed chirality states of the fermions carry a weak isospin of $\frac{1}{2}$ and are $\mathrm{SU}(2)_\mathrm{L}$ doublets, while the right-handed chirality states do not carry weak isospin and are $\mathrm{SU}(2)_\mathrm{L}$ singlets. A weak-isospin quark doublet is formed by the left-handed chirality state of an up-type quark with $T_3 = \frac{1}{2}$ and a down-type quark with $T_3 = -\frac{1}{2}$, $q_\mathrm{L} = (q^u_\mathrm{L}, q'^d_\mathrm{L})^T$, while a weak-isospin lepton doublet is formed by the left-handed chirality states of a neutrino with $T_3 = \frac{1}{2}$ and an electrically charged lepton with $T_3 = -\frac{1}{2}$, $l_\mathrm{L} = (\nu_\mathrm{L}, \ell_\mathrm{L})^T$. As a consequence of coupling only to the left-handed chirality states of fermions, interactions involving weak bosons violate both the charge conjugation (C) and parity (P) symmetries, which is an experimentally established feature of weak interactions. Table 1 provides a summary of the EW quantum numbers for leptons and quarks. The right-handed chirality states of neutrinos do not carry any charge of the SM interactions. Since these states have not been observed and are not required in the formulation of the SM, they are not included in its minimal formulation. For simplicity, electrically charged leptons will be referred to as leptons and neutrinos as neutrinos throughout the thesis.

Table 1: Summary of the EW quantum numbers for up-type (q^u) and down-type (q'^d) quarks, which carry color charge, and for electrically charged leptons (ℓ) and neutrinos (ν), which are color-neutral particles. The table separately lists these numbers for the right-handed and left-handed chirality states, indicated by subscripts R and L, respectively. The electric charge (Q), the weak isospin (T), the third component of the weak isospin (T_3), and the weak hypercharge (Y) are given in the table.

Fermion	Q	T	T_3	Y
Color-charged particles - Quarks				
q^u_L	$+2/3$	$1/2$	$+1/2$	$+1/3$
q'^d_L	$-1/3$	$1/2$	$-1/2$	$+1/3$
q^u_R	$+2/3$	0	0	$+4/3$
q^d_R	$-1/3$	0	0	$-2/3$
Color-neutral particles - Leptons				
ℓ_L	-1	$1/2$	$-1/2$	-1
ℓ_R	-1	0	0	-2
ν_L	0	$1/2$	$+1/2$	-1

The formulation of the SM Lagrangian as described so far does not allow for direct mass terms for the vector bosons and fermions, as such terms violate gauge invariance. However, this contradicts the experimental observations of massive fermions and vector bosons. Therefore, the introduction of a mechanism into the SM Lagrangian is required to allow for this. This mechanism is the Brout-Englert-Higgs mechanism [30–35]. A scalar field, known as the Higgs field (ϕ), is incorporated into the SM Lagrangian. This field transforms as an $\mathrm{SU}(2)_\mathrm{L}$ doublet with $Y = 1$ and is characterized by a non-vanishing vacuum expectation value $v \approx 246\,\mathrm{GeV}$. As a consequence, the Lagrangian respects the full $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ symmetry, while the vacuum state or ground state does not. The process by which a physical system dynamically transitions from a symmetric state to an asymmetric one (the vacuum state) is referred to as spontaneous symmetry breaking (SSB). Furthermore, the massive Higgs boson (H), which does not carry color and electric charges, is introduced with a mass given by $m_H = \sqrt{2\lambda v^2}$, where λ is the coupling constant for the quartic Higgs-boson self-interactions. The vector bosons and fermions acquire their masses by interacting with the Higgs field. Interactions of fermion

fields with a scalar field are referred to as Yukawa interactions. These are associated with the Yukawa couplings y_f that represent the coupling constants of these interactions. The fermion masses are given by $m_f = y_f v/\sqrt{2}$. The neutrinos remain massless due to the absence of their right-handed chirality states in the SM. After SSB, the gauge symmetry of the SM is given by $SU(3)_C \times U(1)_{\mathrm{EM}}$, where $U(1)_{\mathrm{EM}}$ denotes the gauge group of the EM interactions described by quantum electrodynamics (QED). As the EW symmetry is broken, this specific SSB is referred to as electroweak symmetry breaking (EWSB).

The SM contains twelve different types of fermions, which are referred to as flavors. There are six different quark flavors: three up-type quarks, namely the up (u), charm (c), and top (t) quarks, and three down-type quarks, namely the down (d), strange (s), and bottom (b) quarks. In addition, there are six different lepton flavors: three leptons, namely the electron (e^-), muon (μ^-), and tau lepton (τ^-); and three neutrinos, namely the electron neutrino (ν_e), the muon neutrino (ν_μ), and the tau neutrino (ν_τ). The fermions are ordered in three different generations, which contain one weak isospin quark doublet, $(u_{\mathrm{L}}, d'_{\mathrm{L}})^T$, $(c_L, s'_{\mathrm{L}})^T$, or $(t_{\mathrm{L}}, b'_{\mathrm{L}})^T$, and the corresponding right-handed chirality states, as well as one weak isospin lepton doublet, $(\nu_{e_{\mathrm{L}}}, e_{\mathrm{L}})^T$, $(\nu_{\mu_{\mathrm{L}}}, \mu_{\mathrm{L}})^T$, or $(\nu_{\tau_{\mathrm{L}}}, \tau_{\mathrm{L}})^T$, and the corresponding right-handed chirality states of the leptons. While the properties of the fermions with respect to the SM interactions are identical across generations, they are distinguished by the value of their masses, which fulfill the following relations: $m_u < m_c < m_t$, $m_d < m_s < m_b$, and $m_e < m_\mu < m_\tau$ [36]. The fermion masses are free parameters of the SM, and hence their values, as well as these relations, are determined experimentally. The generations are ordered by the mass values, with the generations containing the least and most massive particles referred to as the first and third generations, respectively.

Observable particles are required to be mass eigenstates and eigenstates of the residual symmetry group after EWSB. The experimentally observed $W^\pm$ bosons, the Z boson, and the photon (γ) are expressed as superpositions of the W^1, W^2, W^3, and B bosons, which are eigenstates of the symmetry group before EWSB. These eigenstates are expressed by Equations (1)–(3), which introduce the weak mixing angle, also referred to as the Weinberg angle θ_W, defined by $\tan(\theta_W) = \frac{g'}{g}$.

$$W^\pm = \frac{1}{\sqrt{2}} \left(W^1 \mp \mathrm{i}\, W^2 \right) \tag{1}$$

$$Z = \cos\theta_W\, W^3 - \sin\theta_W\, B \tag{2}$$

$$\gamma = \sin\theta_W\, W^3 + \cos\theta_W\, B \tag{3}$$

While the $W^\pm$ and Z bosons acquire masses, $m_{W^\pm} = m_W = gv/2$ and $m_Z = gv/(2\cos\theta_W)$, the photon remains massless after EWSB. Fermions that carry weak isospin or electric charge interact with the Z boson, while only those carrying weak isospin interact with the $W^\pm$ bosons. Conversely, the photon only interacts with fermions that carry electric charge. The photon is the gauge boson corresponding to the unbroken $U(1)_{\mathrm{EM}}$ symmetry. Furthermore, fermion interactions with the $W^\pm$ boson alter the third component of the weak isospin T_3 by ∓ 1 and consequently, the flavor of the fermion.[3] For instance, an up-type quark is transformed into a down-type quark, and a neutrino is transformed into a lepton. The $W^\pm$ bosons will be commonly referred to as W bosons in the following.

By the use of a bi-unitary transformation, the weak-isospin quark doublets are transformed from their mass eigenstates to the corresponding weak-interaction eigenstates as shown in Equation (4).[4] The matrix V denotes the Cabibbo-Kobayashi-Maskawa (CKM) matrix [37, 38]. As a consequence of this quark mixing, interactions of quarks with W bosons can change the quark flavor across generations, i.e., a top quark can be transformed into a bottom quark, a strange quark, or a down quark. The interaction vertex of quarks and $W^\pm$ bosons involves the corresponding CKM matrix element V_{ij}, which affect the transition probabilities.[5] Due to its unitary nature, the CKM matrix introduces four free parameters into the SM:

[3] The flavor of the fermion coincides with the mass eigenstate.

[4] The mass eigenstates of the up-type quarks coincide with their weak eigenstates, i.e., $q'_u = q_u$. It should be noted that this is a conventional choice.

[5] In general, the transition probability is smaller for larger values of $|i - j|$ with i, j indicating the generation of the involved quarks.

three mixing angles and one complex phase. This complex phase leads to the violation of CP-symmetry, the breaking of the combined C and P symmetries, in weak interactions when there is interference of at least two different contributing processes that carry distinct complex phases. Due to the lack of neutrino masses in the SM, an analogous mixing in the lepton sector is not realized. As a consequence, lepton flavor is conserved by SM interactions. Flavor-changing transitions and CP-symmetry violation are experimentally established features of interactions mediated by the W boson.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \tag{4}$$

To extract predictions for observables related to SM interactions, the matrix element $\mathcal{M}$, which describes the underlying transition from the initial state to the final state, needs to be calculated. If the coupling constants associated with the interactions are sufficiently small, so that perturbation theory is applicable, $\mathcal{M}$ can be expressed using an asymptotic expansion in those coupling constants. The terms in the expansion are ordered by powers of the coupling constants. The terms with the smallest powers are referred to as leading-order (LO) contributions; the terms with the second-smallest powers are referred to as next-to-leading-order (NLO) contributions; and the n-th terms are referred to as next-to-next-to-...-leading order $\mathrm{N}^{n-1}\mathrm{LO}$ contributions. At any given order, there can be multiple possibilities for how the transitions can be realized according to the SM. Each of these possibilities represents a term that can be calculated using Feynman rules. These rules are encoded in Feynman diagrams, which are pictorial representations of these contributions. In such diagrams, elementary particles are depicted as lines, while interactions between them are represented by vertices. If a contribution does not include any internal loops, it is referred to as a tree-level contribution, which is usually larger than contributions including loops to the same interaction. Resummation techniques and other methods can be applied to improve the convergence or address specific challenges of the series.

The higher-order terms include loop contributions, which typically yield divergences in the large-momentum regime, referred to as ultraviolet divergences. By employing renormalization techniques, these divergences can be absorbed into the parameters of the SM, such as couplings and masses. In contrast to the unrenormalized (bare) parameters of the theory, these parameters are expressed in terms of the renormalization scale μ_R and can be used to compute observable quantities, such as pole masses. Equation (5) shows the dependence of SM coupling constants defined as $\alpha_i = \frac{g_i^2}{4\pi}$ with $g_i = \{g_s, g, g'\}$, which represents the solution to the renormalization group equation at one-loop order.

$$\alpha_i(\mu_\mathrm{R}^2) = \frac{\alpha_i(\mu_0^2)}{1 + \beta_{0,i}\,\alpha_i(\mu_0^2)\,\ln\left(\frac{\mu_\mathrm{R}^2}{\mu_0^2}\right)} \tag{5}$$

Measuring α_i at a reference scale μ_0 fixes the "running" of the parameter. The value and sign of the coefficient $\beta_{0,i}$ are specific to each interaction. The $\beta_{0,i}$ coefficient is negative for the coupling strength associated with the $\mathrm{U}(1)_Y$ gauge group and positive for those associated with the $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{SU}(3)_\mathrm{C}$ gauge groups. As a consequence, the strong and weak coupling strengths converge toward zero for large μ_R, known as asymptotic freedom, and diverge toward low μ_R, while the opposite behavior is observed for g'. The weak interaction is mediated by massive gauge bosons, which ensures that its effective coupling, the Fermi constant $G_\mathrm{F} = \sqrt{2}\,g^2/(8\,m_W^2)$, remains small also for low μ_R. In contrast, the strong interaction is mediated by massless gluons. As a consequence, the coupling constant grows with decreasing energy, and perturbation theory with g_s as expansion parameter cannot be applied in the low-μ_R regime, necessitating the use of non-perturbative or effective models at low energy scales. Furthermore, neither quarks nor gluons have been observed as isolated particles, a phenomenon explained by the confinement theorem.[6] Only composite particles made up of color-charged particles

[6] The aforementioned behavior of g_s already hints at this phenomenology, but it is not the direct cause, as the divergent behavior of g_s cannot be extrapolated to the non-perturbative regime.

can be observed. These composite particles, known as hadrons, are color neutral, i.e., they carry a net color charge of zero. The constituents of hadrons are referred to as partons. There are two main categories of hadrons: mesons, which are bosons characterized by an even number of valence quarks (at least two), and baryons, which are fermions characterized by an odd number of valence quarks (at least three). Valence quarks are the primary quark constituents of mesons and baryons and mainly determine the properties and quantum numbers of these particles. A hadron is called a b hadron if at least one of its valence quarks is a b quark. If this is not the case and at least one of its valence quarks is a c quark, it is called a c hadron. The b and c hadrons are commonly referred to as heavy-flavor (HF) hadrons, while all other hadrons are referred to as light-flavor (LF) hadrons.

2.2 Limitations of the Standard Model

The compatibility of observed data with a model prediction is usually expressed in terms of the statistical significance Z in units of the standard deviation σ. The larger the value, the more the validity of the model in the probed phase space is challenged. Particle physicists demand a statistical significance exceeding 5σ as one of the mandatory requirements to reject the validity of the SM. Such a deviation from the predictions of the SM has not been established in the thousands of measurements performed in particle physics experiments. However, there are conceptual shortcomings and observed phenomena at energy scales not probed by these experiments that cannot be explained by the SM. These lead to the conclusion that the SM is inherently incomplete and not the theory that describes all fundamental interactions among subatomic particles. Therefore, searching for BSM physics and identifying and interpreting the underlying mechanisms of observed phenomena are main objectives of particle physics. Selected examples of these limitations are discussed in the following.

One of the most apparent phenomena not included in the SM is the gravitational interaction, which is described by general relativity. As of now, no formulation of a valid gauge QFT of general relativity has been achieved that is renormalizable and can be unified with the SM [39–44].

As discussed in the previous section, neutrinos are massless in the SM. However, the well-established experimental observation of neutrino oscillations [45, 46], i.e., neutrinos changing their flavor during propagation, reveals that neutrinos must have nonzero masses. An upper bound of 0.45 eV on the neutrino masses has been reported by the KATRIN collaboration [47], but the masses have not been directly measured to date. The SM can be extended to include Dirac mass terms for neutrinos, as for the other fermions, by introducing right-handed neutrino chirality states. These states are singlets with respect to the SM gauge groups and are commonly referred to as sterile neutrinos. They solely couple to the Higgs field via Yukawa interaction. Due to the small expected mass values, this interaction is extremely weak, so that an experimental verification of this extension is not possible in the foreseeable future. Alternatively, Majorana mass terms for neutrinos can be generated by mechanisms such as the see-saw mechanism [36], which postulate new heavy neutrino states that contain right-handed chirality states. These states enable mass mixing with the established light neutrino states and provide a natural explanation for the smallness of the observed mass values. However, due to the lack of empirical evidence, none of these extensions has been universally accepted.

Cosmological observations, such as those related to the rotation curves of galaxies [48] and the cosmic microwave background (CMB), which is a relic microwave radiation that permeates the entire observable universe, reveal that only a small fraction of the energy content of the universe stems from ordinary matter described by the SM. The most common interpretation is provided by the ΛCDM model [49, 50], which describes the evolution of the universe since the Big Bang using six parameters in its basic form. Λ denotes the cosmological constant that is associated with the vacuum energy, also known as dark energy, and CDM refers to an unknown form of matter, the so-called cold dark matter. While the ΛCDM model incorporates the SM in its description, it fundamentally relies on these phenomena, which are not described by the SM. From the most recent and precise measurements of cosmological parameters of

the ΛCDM model by the Planck microwave telescope [51], which was used for extensive studies of the CMB, it follows that around 69%, 26%, and 5% of the energy content stems from dark energy, CDM, and ordinary matter, respectively. Although the SM is not designed to describe the cosmological evolution, the absence of a description of dark matter and dark energy within the SM raises the question about the completeness of the current understanding of particle interactions. This motivates studies of BSM theories that extend the SM to incorporate these missing concepts, for instance, theories proposing new elementary particles as dark matter candidates, like weakly interacting massive particles (WIMPs) [52] or axions [53]. It should be noted that while the ΛCDM model represents a robust framework for explaining cosmological observations and is often referred to as the Standard Model of cosmology for that reason, there are experimental results that challenge the validity of the ΛCDM model, most notably the "Hubble tension". This tension refers to the discrepancy between independent measurements of the Hubble constant, a key parameter of the ΛCDM model, with a statistical significance of $> 4\sigma$ [51, 54, 55].

Moreover, the $\approx 5\%$ observable matter content is characterized by an enormous asymmetry of matter and antimatter particles. In fact, only the production of antimatter is observed, for instance, in cosmic rays, in the laboratory, e.g., in particle physics experiments, or in radioactive decays, but not its permanent existence in isolated regions or stable states. A widely accepted hypothesis is that this asymmetry was caused by a baryon asymmetry in the early stages of the universe, i.e., that the density of baryonic matter was greater than the density of anti-baryonic matter. This imbalance is tiny ($\approx 6 \cdot 10^{-10}$). One interpretation is that this asymmetry is primordial. However, the most common cosmological models assume a symmetric primordial baryon state. In such a scenario, the baryon asymmetry must be generated dynamically, a process referred to as baryogenesis [56]. This dynamic process must satisfy the Sakharov conditions [57], which require mechanisms that violate the baryon number as well as the C and CP symmetries, and lead to a departure from thermal equilibrium. The baryon number (B) is a quantum number defined as $B = +1$ for baryons, $B = -1$ for antibaryons, and $B = 0$ for other particles. The EW interaction offers a mechanism for baryon number violation in the non-perturbative regime [58, 59]. Furthermore, the EW interaction violates the C symmetry and introduces a source of CP symmetry violation. However, these mechanisms yield a baryon asymmetry that is several orders of magnitude smaller than the observed value and therefore provide no valid explanation for a dynamic generation of the baryon asymmetry. This is still the case even when the SM parameters are adjusted to maximize the CP violation (neglecting experimental constraints), so that substantial extensions of the SM are needed to provide such an explanation.

The formulation of the SM, as discussed in the previous section, is characterized by 19 free theory parameters whose values cannot be inferred from first-principle calculations and must instead be measured experimentally. Most of these parameters are associated with the EW sector of the SM and are characterized by a hierarchical structure. As of now, there is no established explanation of this hierarchical structure. It is therefore intriguing to study extensions of the SM through processes that generate the relations and hierarchical structure of the parameters dynamically, as proposed by Froggatt and Nielsen in Ref. [60], to reduce the number of free parameters and uncover underlying mechanisms.

The formulation of the SM allows for terms related to the strong interaction that violate the CP-symmetry in SM interactions. However, this violation has not been observed and only tiny values of the CP violating parameter that do not contradict observations are allowed. This puzzling aspect of the SM is referred to as the strong CP-problem. It is therefore intriguing to search for a mechanism that provides a natural and dynamic explanation for this observation. One of the most prominent theoretical solutions is the Peccei-Quinn mechanism that predicts the existence of new elementary particles, called axions [61].

Apart from the aforementioned shortcomings, there are measurements probing observables that challenge their SM prediction. Large tensions must be scrutinized both theoretically and experimentally to rule out systematic errors and to ensure that no effects significantly impacting the SM prediction have been

neglected. Furthermore, it is typically investigated whether these tensions are observed in independent datasets, ideally from different experiments, or manifest themselves in different, but related processes through a common pattern. This provides information on whether the observed tensions arise from statistical fluctuations, insufficient theoretical knowledge in describing established mechanisms, or potential indications of BSM physics.

Several measurements of angular observables and branching fractions [62–73] in semileptonic decays of b hadrons involving $b \to s\ell\ell$ transitions are in tension with SM predictions. These tensions follow a common pattern. Therefore, exploring BSM models explaining these "B anomalies" is intriguing, even though the result of the test of lepton-flavor universality reported in Ref. [74], initially found to be in tension with the SM prediction and to follow the common pattern, was later found to agree with the SM prediction in an updated result [75]. It remains an open question whether these tensions genuinely form a consistent pattern that indicates BSM physics.

Measurements of the W-boson mass are in good agreement with the SM prediction, including the latest and most precise one reported in Ref. [76], except for the result presented in Ref. [77]. This result deviates from the SM prediction by a statistical significance of 7σ and strongly disagrees with other experimental results. This conflict has not yet been resolved.

Numerous BSM models that extend the SM and offer solutions to one or multiple of the aforementioned limitations, conceptual problems, or experimental tensions have been proposed over the last decades. Many of these models predict the existence of new elementary particles, which play a crucial role in resolving these issues. These new particles can either be observed directly, e.g., as a resonance, and/or alter the properties of observable particles, allowing the BSM effects to be observed indirectly. However, due to the absence of conclusive experimental evidence, no BSM model has superseded or extended the SM so far. Assuming that a specific BSM model is realized in nature, there are two main reasons why its experimental manifestation has not yet been observed in current measurements:

- The couplings of the BSM particles to SM particles are sufficiently tiny, so that the resolution and precision of measurements, which are limited by statistical, experimental, and theoretical uncertainties, are insufficient to resolve the BSM effects.
- The new particles emerge and noticeably affect particle properties at energy scales not accessible by current experiments.

As a consequence, it is paramount to deepen the understanding of the SM by improving the precision of experimental results and theory predictions. This is especially important for those observations that are in tension with the SM, and by exhaustively studying as many aspects of the SM as possible to find direct evidence for, or patterns of, new physics. In this context, the top quark is a particularly interesting research subject. Its unique features, which are discussed in Section 2.3, make it especially sensitive to promising BSM scenarios and allow for its thorough experimental investigation in collider experiments. The study of the top-quark–photon interaction represents a cornerstone of EW physics and an intriguing probe for BSM physics, as discussed in Section 2.5.

2.3 Top-quark properties in the Standard Model

In 1964, Ref. [78] reported the experimental observation of CP-symmetry violation. This was an unexpected finding, as no established mechanism for this violation existed at the time. Furthermore, quark models were not widely accepted and only contained two quark generations consisting of the up, down, and strange quarks, which did not allow for a theoretical description of this violation. In 1973, Kobayashi and Maskawa showed in Ref. [37] that this observation can be explained by the quark model if an additional, third quark generation exists. This allows quark mixing to be parameterized using a unitary 3×3 matrix, which, in contrast to the 2×2 Cabibbo matrix [38] used for two generations, introduces a complex phase necessary for the theoretical description of CP-symmetry violation. The

existence of the third generation of quarks was first inferred from the discovery of the bottom quark in 1977 [79]. The top quark was discovered in 1995 as the second third-generation quark and most massive elementary particle, with a mass of around 173 GeV, by the CDF [18] and DØ [19] experiments at the Tevatron at Fermilab in proton–antiproton collisions. Since its discovery, the properties of the top quark have been studied exhaustively in hadron collider experiments, while high statistical precision in these studies has been achieved only in recent years. The main properties of the top quark in the context of the SM will be summarized in the following.

Measurements [80–83] revealed that the top-quark charge Q_t agrees well with the SM expectation of $+\frac{2}{3}$. Measurements of the top-quark mass m_t have been performed at the CDF and DØ experiments at the Tevatron in proton–antiproton collisions and by the ATLAS and CMS experiments at the LHC in pp collisions. Several of these measurements have been combined into a world average value of $m_t = 173.34 \pm 0.76\,\text{GeV}$ [84]. The most recent combination of results from the ATLAS and CMS collaborations is more precise, yielding a value of $m_t = 172.52 \pm 0.33\,\text{GeV}$. The most recent single result released by the CMS collaboration reports a value of $m_t = 171.77 \pm 0.37\,\text{GeV}$ [85], representing an improvement in precision of 25% relative to the next most precise single result [86]. The top quark is not only the most massive elementary particle, but its mass is more than ≈ 40 times larger than that of the second-to-most massive fermion, the bottom quark. The top quark is also considered to be the only particle with a natural mass value, as its Yukawa coupling y_t is approximately one. This suggests that the top quark is strongly connected to the EWSB and is therefore considered a key particle in theories that aim to solve the limitations occurring in the EW sector discussed in the previous section.

Another fundamental property of the top quark is its decay width Γ_t, which can be expressed by Equation (6) at LO precision. This assumes $|V_{tb}| = 1.0$, so that Γ_t^0 corresponds to the $t \to Wb$ decay, and neglects terms proportional to $\frac{m_b^2}{m_t^2}$.

$$\Gamma_t^0 = \frac{G_\text{F} m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_t^2}\right) \tag{6}$$

Using the world average results of the top-quark mass and the W-boson mass[7] reported by the Particle Data Group of $m_W = 80.377 \pm 0.012\,\text{GeV}$ [36], as well as $G_\text{F} = 1.1663788 \pm 0.0000006 \cdot 10^{-5}\,\text{GeV}^{-2}$ [36], a value of $\Gamma_t^0 = 1.51\,\text{GeV}$ is computed. Corrections to Equation (6) arising at NLO QCD ($\delta_\text{QCD}^{(1)}$), next-to-next-to-leading order (NNLO) QCD ($\delta_\text{QCD}^{(2)}$), and NLO EW (δ_EW) as well as corrections from the finite bottom-quark mass (δ_f^b) and W-boson decay width (δ_f^W) have been calculated in Refs. [87–89]. These have been found to lower the value of the decay width by $\approx 10.7\%$ [90]. Differences of 0.5 GeV in the top-quark mass result in a variation of the top-quark decay width of $\Delta\Gamma_t \approx 0.2\,\text{GeV}$, while uncertainties in other SM parameters entering the calculation contribute an uncertainty at the level of $\approx 6\%$. The impact of other uncertainty sources on this parameter is small ($< 1\%$).

The top-quark decay width can be measured experimentally either in an indirect or a direct fashion. The indirect measurements provide a higher precision with the most precise result of $\Gamma_t = 1.36 \pm 0.02\,(\text{stat.})^{+0.014}_{-0.011}\,(\text{syst.})\,\text{GeV}$ [91] reported by the CMS collaboration. The direct measurements are less model-dependent, and the ATLAS collaboration has provided a preliminary result of $\Gamma_t = 1.9 \pm 0.5\,\text{GeV}$ [92], while the most precise result is $\Gamma_t = 1.28 \pm 0.30\,\text{GeV}$ [93].

The corresponding mean lifetime of the top quark is $\tau_t = 1/\Gamma_t \approx 5 \cdot 10^{-25}\,\text{s}$, which is one order of magnitude shorter than the timescale of hadronization processes, $1/\Lambda_\text{QCD} \approx 3 \cdot 10^{-24}\,\text{s}$ with Λ_QCD being the energy scale below which QCD interactions become non-perturbative.[8] Hence, the top quark decays in the vast majority of all cases before it undergoes hadronization. This is a unique feature of the top quark, as all other quarks undergo hadronization before they decay. In the SM, the top

[7] This world average neglects the most recent measurement by the CDF collaboration discussed in the previous section.
[8] Assuming a typical value of $\Lambda_\text{QCD} \approx 250\,\text{MeV}$.

quark predominantly decays via the weak interaction into a W boson and a quark. The probability of the quark flavor is proportional to the corresponding CKM matrix elements. With the relevant CKM matrix element in $t \to b$ transitions being $|V_{tb}| \approx 1$, it follows that the top quark decays almost exclusively into a W boson and a bottom quark. Assuming the unitarity of the CKM matrix, i.e., only three quark generations, the most precise SM value extracted from fits to experimental results is $|V_{tb}| = 0.999118^{+0.000031}_{-0.000036}$ [36], while the combination of experimental results of $|V_{tb}|$ without this assumption yields $|V_{tb}| = 1.014 \pm 0.029$ [36].

The W boson itself has an even larger decay width of $\Gamma_W = 2.091 \pm 0.001$ GeV, as determined by the most recent global fit to experimental results testing the parameters of the EW sector [94], leading to a shorter lifetime than the top quark. Therefore, it also decays almost instantaneously. It decays into a charged lepton and a neutrino in $\approx 33\%$ or into a quark-antiquark pair involving first- or second-generation quarks in $\approx 66\%$ of all cases [36]. The former is known as the leptonic channel, while the latter is referred to as the hadronic channel. Top-quark decays with a subsequent leptonic W-boson decay are referred to as semileptonic top-quark decays, while those with a subsequent hadronic W-boson decay are referred to as hadronic top-quark decays.

2.4 Top-quark–photon interaction in the Standard Model

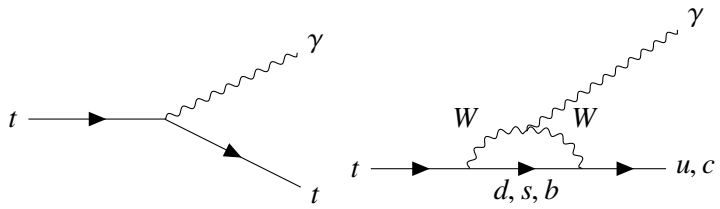


Figure 1: Leading-order Feynman diagrams for the flavor-conserving neutral top-quark–photon interaction on the left and for the flavor-changing interaction on the right.

In addition to studying the top-quark mass, width, and decay, it is intriguing to study its direct coupling to the gauge bosons of the SM. This work focuses on examining the interaction between the top quark (t) and the photon. Figure 1 shows LO Feynman diagrams for the flavor-conserving neutral and flavor-changing top-quark–photon interactions according to the SM. Without loss of generality, the top-quark–photon interaction can be parameterized in terms of four form factors, κ^{L}_{tU}, κ^{R}_{tU}, $\lambda^{\mathrm{L}}_{tU}$, and $\lambda^{\mathrm{R}}_{tU}$, in the form of a Lagrangian [95] given by Equation (7).[9] The photon field is denoted by A^{μ}, while q^U represents the field of an up-type quark ($U = \{u, c, t\}$) and q^t represents the field of the top quark. The tensor $\sigma^{\mu\nu}$ is related to the commutator of the Dirac matrices γ^{μ} and is given by $\frac{i}{2}[\gamma^{\mu}, \gamma^{\nu}]$. P^{L} and P^{R} are operators, which project fermion fields onto their left-handed and right-handed chiral states, respectively. The four-momentum of the photon is given by $p = p_t - p_U$ with p_q and p_U denoting the four-momenta of the top quark and the up-type quark, respectively.

$$\mathcal{L}_{\gamma tU} = -e\bar{q}^U \gamma^{\mu} \left(\kappa^{\mathrm{L}}_{tU} P^{\mathrm{L}} + \kappa^{\mathrm{R}}_{tU} P^{\mathrm{R}}\right) q^t A_{\mu} - e\bar{q}^U \frac{i\sigma^{\mu\nu} p_{\nu}}{m_t} \left(\lambda^{\mathrm{L}}_{tU} P^{\mathrm{L}} + \lambda^{\mathrm{R}}_{tU} P^{\mathrm{R}}\right) q^t A_{\mu} + h.c. \tag{7}$$

For flavor-conserving neutral top-quark–photon interactions ($q^U = t$), κ^{L}_{tU} equals κ^{R}_{tU}, and Equation (7) can be transformed into Equation (8). For flavor-changing top-quark–photon interactions ($q^U = \{q^u, q^c\}$), κ^{L}_{tU} and κ^{R}_{tU} equal zero and Equation (7) can be transformed into Equation (9). These relations are consequences of the gauge symmetries.

$$\mathcal{L}_{\gamma tt} = -eF_{1V} \bar{q}^t \gamma^{\mu} q^t A_{\mu} - e\bar{q}^t \frac{i\sigma^{\mu\nu} p_{\nu}}{2\, m_t} \left(F^{\gamma}_{2V} + iF^{\gamma}_{2A}\gamma^5\right) q^t A_{\mu} \tag{8}$$

[9] The Einstein summation convention is used

$$F_{2V}^{\gamma} = \lambda_{tU}^{\mathrm{L}} + \lambda_{tU}^{\mathrm{R}} \quad F_{2A}^{\gamma} = -i(\lambda_{tU}^{\mathrm{L}} - \lambda_{tU}^{\mathrm{R}})$$

According to the SM, F_{1V} coincides with Q_t. The form factors F_{2V} and F_{2A} are related to the top quark's magnetic dipole moment (MDM), a_t, and electric dipole moment (EDM), d_t, with $F_{2V} = Q_t\, a_t$ and $F_{2A} = \frac{2\, m_t}{e} d_t$. The latter introduces a source of CP-symmetry violation. At tree level, the values of a_t and d_t are both zero. At one-loop order, a_t acquires a nonzero value. It has been calculated to be $a_t = 0.02$ [96] when considering terms up to the two-loop order. In contrast, d_t only acquires a nonzero value at three-loop order. Consequently, the SM prediction for d_t is tiny, estimated to be $|d_t| < 10^{-30} e$ cm [97, 98].

$$\mathcal{L}_{\gamma tU} = -e\bar{q}^{U} \frac{i\sigma^{\mu\nu} p_\nu}{m_t} \left(\lambda_{tU}^{\mathrm{L}} P^{\mathrm{L}} + \lambda_{tU}^{\mathrm{R}} P^{\mathrm{R}}\right) q^t A_\mu + h.c. \tag{9}$$

According to the SM, transitions of a top quark into another up-type quark via interactions with bosons carrying no electric charge, referred to as flavor-changing neutral currents (FCNCs), are forbidden at tree level. Thus, LO contributions first occur at the one-loop level and involve a W boson. Such processes are heavily suppressed in the SM by the Glashow-Iliopoulos-Maiani (GIM) mechanism [99] and the small CKM matrix elements for $t \to U$ transitions, with $U = \{u, c\}$. The rate of $t \to U\gamma$ transitions is expressed through a partial decay width, which is given by Equation (10) [100].

$$\Gamma_t(t \to U\gamma) = \frac{\alpha}{2} m_t \left(|\lambda_{tU}^{\mathrm{L}}|^2 + |\lambda_{tU}^{\mathrm{R}}|^2\right) \tag{10}$$

The SM predicts that the branching ratio ($\mathcal{BR}(t \to U\gamma) = \Gamma_t(t \to U\gamma)/\Gamma_t$) is $\approx 5 \cdot 10^{-14}$ for $t \to c\gamma$ and $\approx 4 \cdot 10^{-16}$ for $t \to u\gamma$ transitions [100].

While the measurements of Q_t have ruled out significant deviations from the SM value [80–83], studying the top quark's MDM and EDM as well as the rate of FCNC top-quark interactions with a photon remains an intriguing task. Any significant deviation from the small predicted values of these properties would be a direct indication of physical mechanisms that are not described by the SM.

2.5 Top-quark–photon interaction beyond the Standard Model

Properties of the top-quark–photon interaction are affected in various BSM scenarios. Special attention is paid to scenarios that enhance the EDM of the top quark, as it represents a new source of CP-symmetry violation. Exploring these sources is compelling because they are essential for satisfying the Sakharov conditions for baryogenesis. Experimental data directly probing this quantity only allow values of $|d_t| < O(1) \cdot 10^{-17}\, e$ cm [101, 102]. Additionally, data obtained from studies of the neutron [103] and the mercury atom [104–106] can be translated into a constraint for the top-quark EDM corresponding to $|d_t| < 5 \cdot 10^{-20}\, e$ cm [107]. Contributions to the top-quark EDM in specific BSM models have been extensively studied, e.g., in Refs. [108–111], showing potential values up to 10^{-19}–$10^{-18}\, e$ cm, which represents enhancements of more than ten orders of magnitude. BSM models that enhance the MDM have also been examined. For example, Ref. [112] studied such a scenario and found BSM contributions in the range 10^{-6}–10^{-5}.

Furthermore, various BSM models suggest significantly enhanced rates of flavor-changing top-quark–photon interactions. Studies, such as those reported in Refs. [113–116] and [117, 118], have examined these effects in different BSM scenarios. These studies found that the branching fractions for $t \to c\gamma$ and $t \to u\gamma$ transitions could be as high as 10^{-7} and 10^{-8}, respectively. This represents an increase of seven to eight orders of magnitude relative to the SM predictions.

In the absence of any direct observation of new elementary particles, indirect searches for new physics contributions have become increasingly attractive. The emergence of tensions between SM predictions and experimental results, such as the B anomalies, has accelerated this development. However, individual

indirect observations of BSM physics may be explained by multiple BSM models. This ambiguity can be resolved by complementary direct or indirect observations that reveal an unambiguous pattern. Many BSM models that were once considered to be the most likely candidates have already been extensively studied, and experimental data have either constrained their allowed parameter space or ruled them out. Therefore, model-independent approaches have gained more attention in recent years. One such approach is the effective field theory (EFT), which will be briefly discussed in the following.

The guiding principle of EFTs is that they are not tools for studying how an interaction works, but how it affects observables. EFTs serve as a valid approximation of the full model, which includes the detailed description of the underlying mechanism, if the interaction involves heavy degrees of freedom, such as a new elementary particle with a mass corresponding to a mass scale Λ well above the probed energy scale. This is known as the decoupling theorem [119]. Contributions from those heavy degrees of freedom are suppressed by inverse powers of the associated mass scale. The advantage of the EFT approach is that it allows for the efficient study of large sets of experimental data for different observables, even from different experiments, and for the exploration of interaction patterns. In this way, EFTs provide valuable insight and constraints for potential BSM scenarios.

In the context of high-energy particle physics, these interactions involving heavy degrees of freedom are described by EFT operators. These operators, $O_k^{(d)}$, extend the SM by adding corresponding terms to the Lagrangian of the SM ($\mathcal{L}_{\text{SM}}$), as depicted by Equation (11). The operators are labeled by their mass dimension d (parentheses in the superscript indicate that d is a fixed index and not a variable). Like SM interactions, each of these operators is associated with a coupling strength, $C_i^{(d)}$, known as the Wilson coefficient.

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \sum_{d\geq 5} \sum_{k} \frac{C_k^{(d)}}{\Lambda^{d-4}} O_k^{(d)} \tag{11}$$

The Lagrangian of the SM contains only terms with $d = 4$ and the additional terms must be of the same mass dimension. This is achieved by including scaling with a factor of $1/\Lambda^{d-4}$ in the additional terms. Here, Λ is an energy scale with a mass dimension of one. Furthermore, the EFT operators must respect the symmetries of the SM, such as gauge invariance. To ensure model independence, only known structures can appear in the EFT operators, and, therefore, only SM fields are used in their construction.

Numerous operators that align with the aforementioned principles can be constructed. However, for specific observables, typically only a few of these operators are of interest. In the context of top-quark–photon interactions, two classes of operators are typically studied, which are illustrated in Equations (12) and (13) using the parameterization of the Warsaw basis [120]. Both classes have a mass dimension of $d = 6$, and, for simplicity, the corresponding superscript indicating the mass dimension is neglected. The generation of the quark that interacts with the top quark is denoted by l, for instance, O_{uB}^{23} corresponds to the interaction between c and top quarks. $B_{\mu\nu}$ and $W_{\mu\nu}^I$ denote the field strength tensors of the B and W gauge bosons of the EW interactions before SSB, respectively, while τ^I denotes the corresponding Pauli matrix and φ the Higgs field with $\tilde{\varphi} = i\tau^2\varphi^*$.

$$O_{uB}^{l3} = \left(\bar{q}_l^u \sigma^{\mu\nu} t\right) \tilde{\varphi}\, B_{\mu\nu} \tag{12}$$

$$O_{uW}^{l3} = \left(\bar{q}_l^u \sigma^{\mu\nu} \tau^I t\right) \tilde{\varphi}\, W_{\mu\nu}^I \tag{13}$$

The EFT approach can be viewed as a modern version of the form-factor approach discussed in the previous section. The Wilson coefficients and the form factors of Equation (7) can be related as demonstrated in Ref. [95].

3 Photon-associated single-top-quark production in proton–proton collisions

The extremely short lifetime of the top quark prohibits precise experimental studies of top quarks from natural sources, such as high-energy particle collisions in cosmic rays. Due to its large mass, top-quark production is possible only in high-energy inelastic particle interactions. To enable detailed studies of the interaction between top quarks and photons, this production must be abundant. The LHC [13] is currently the only facility providing such conditions in a controlled experimental environment, which is briefly reviewed in Section 3.1.

The LHC can provide high-energy inelastic interactions via collisions of different types of particles. Among these, pp collisions are best suited for top-quark physics, as they provide high collision energies and interaction rates, while offering simpler initial states than other collision types. Top-quark production in pp collisions under LHC conditions is discussed in Section 3.2. Furthermore, Section 3.3 sheds light on single-top-quark production in association with a photon. This process is the main subject of this thesis and is examined using pp collisions provided by the LHC.

3.1 Brief overview of the Large Hadron Collider

The LHC is a circular particle accelerator designed to provide high-energy particle collisions. It is located beneath the CERN laboratories in Geneva and has a circumference of approximately 26.7 km. Two particle beams are accelerated in opposite directions in separate beam pipes and brought to collision at four interaction points (IPs). Nine detectors designed to examine these collisions are installed at the LHC. Four of them, the detectors of the ATLAS [14], CMS [15], LHCb [121], and ALICE [122] experiments, which are the main LHC experiments, are located at the IPs. The other five detectors are each located in the vicinity of one of the IPs.

The physics program of the LHC is carried out in runs, which are periods of LHC operation without long interruptions or shutdowns. In total, the LHC program foresees five runs, interspersed with long shutdowns lasting two to three years that are used to upgrade or repair components of the LHC and its experiments. The first and second runs (Run 1 and Run 2) were carried out in the years 2010–2013 and 2015–2018, respectively. The third run (Run 3) started in 2022 and is expected to end in July 2026. The LHC was operated with beams of either protons, lead ions, or xenon ions during Run 1 and Run 2. The pp beam configuration, which is the main operation mode of the LHC and therefore offers the largest datasets, is best suited for precision measurements of top-quark properties and studies of rare top-quark processes, as well as searches for BSM phenomena involving top quarks.

The LHC is designed to provide pp collisions up to a center-of-mass energy ($\sqrt{s}$) of 14 TeV. During Run 1, the proton beams were brought to collision at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The proton beam energy was increased to 6.5 TeV per beam, resulting in $\sqrt{s} = 13$ TeV for Run 2. Run 3 is being operated at $\sqrt{s} = 13.6$ TeV, while the subsequent runs are expected to deliver pp collisions at $\sqrt{s} = 14$ TeV. The research presented in this thesis uses data from pp collisions taken during Run 2 of the LHC program.

The LHC is the final stage of the accelerator complex for the proton beams. The journey through this complex and the origin of the proton beams during Run 2 are briefly described in the following. Hydrogen gas is filled into a duoplasmatron, where electrons are stripped from hydrogen atoms to produce protons. These protons are then passed to the first stage of the accelerator complex, the linear accelerator LINAC2, where they acquire a kinetic energy of 50 MeV. The accelerated protons are then transferred to the Proton Synchrotron Booster, followed by the Proton Synchrotron, and finally to the Super Proton Synchrotron, all of which are circular accelerators that further increase the proton energy to 1.4 GeV, 25 GeV, and 450 GeV, respectively. The Super Proton Synchrotron is the last pre-acceleration stage before the proton beams are injected into the LHC.

The LHC proton beams during Run 2 consisted of up to 2808 bunches, each containing $\approx 10^{11}$ protons, separated by 25 ns. One of the major characteristics of a particle accelerator is the instantaneous luminosity $\mathcal{L}^{\text{ins}}$. It quantifies how many particle interactions take place at the IPs per unit of time. The expected rate, $\frac{\mathrm{d}N}{\mathrm{d}t}$, of a physics process is given by Equation (14), where σ^{tot} denotes its total cross section.

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L}^{\text{ins}} \cdot \sigma^{\text{tot}} \tag{14}$$

While enhancing $\mathcal{L}^{\text{ins}}$ has the advantage of increasing the rate of physics processes, which is especially advantageous for the study of rare physics processes with small cross sections, it also has the disadvantage of a larger average total number of interactions, μ, per bunch crossing, which is linearly dependent on $\mathcal{L}^{\text{ins}}$. Interesting physics processes that particle physicists seek to explore are typically characterized by large energy transfers and are referred to as hard-scattering processes. However, such processes occur at much lower rates than processes with smaller energy transfers. Hence, each hard-scattering process is accompanied on average by $\mu - 1$ processes of little interest, referred to as in-time pile-up, where μ denotes the pile-up parameter.[10] Disentangling interesting physics processes from these pile-up interactions represents a major challenge for the LHC experiments.

The design peak instantaneous luminosity of the LHC is $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$, which was already exceeded during Run 2, when this value reached roughly twice its design value [123]. The integrated luminosity, $\mathcal{L}$, is given by the time integral of Equation (14) and thus quantifies the size of datasets taken by experiments. In total, the LHC delivered an amount of pp collisions corresponding to $\mathcal{L} = 156\,\text{fb}^{-1}$ in Run 2 [123]. This translates to roughly 10^{16} inelastic pp collisions.[11]

After Run 3 of the LHC program, which is expected to provide an integrated luminosity of $\approx 300\,\text{fb}^{-1}$, the LHC will be upgraded to the High-Luminosity Large Hadron Collider (HL-LHC) [125]. The HL-LHC is expected to provide a peak instantaneous luminosity ten times larger than the design luminosity of the LHC, with the goal of offering $3000\,\text{fb}^{-1}$–$4000\,\text{fb}^{-1}$ of pp-collision data at the end of its two runs planned for 2030–2032 and 2035–2038. As this dataset will be several times larger than the combined dataset of the first three runs, the HL-LHC will open the door to studies of much rarer physics processes of the SM that cannot be probed with the current design.

3.2 Production of top quarks in pp collisions at the LHC

The LHC pp collisions are typically described using a right-handed coordinate system, with its origin at the nominal IP. The z-axis is parallel to the beam axis, the x-axis points to the center of the accelerator, and the y-axis points upward. The x–y plane is also referred to as the transverse plane. The azimuthal angle ϕ describes the angle around the beam axis in the transverse plane, while the polar angle θ is the angle relative to the z-axis. In practice, the angle θ is often replaced by the rapidity y of a particle, defined by Equation (15), using the magnitude of the momentum $|\vec{p}|$ and the longitudinal momentum component p_z. This quantity is particularly useful, as differences in rapidity are invariant under Lorentz transformations along the z-axis.

$$y = \frac{1}{2} \ln\left(\frac{|\vec{p}| - p_z}{|\vec{p}| + p_z}\right) \tag{15}$$

For massless particles, the rapidity y is equal to the pseudorapidity η, defined by $\eta = -\ln(\tan\frac{\theta}{2})$. Throughout this thesis, distances in the η–ϕ plane are expressed as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, unless stated otherwise.

[10] Multiple partons may interact simultaneously in a single pp collision, so that these interactions overlap with each other. This is known as multiple parton interactions (MPI).

[11] A cross section of 78 mb for inelastic pp interactions is assumed [124].

Protons are baryons composed of two up quarks and one down quark as their valence quarks, and carry an electric charge of $Q_p = +1$. When protons are brought to collision at high energies, their constituents interact, allowing the substructure of the proton to be resolved. The value of the strong coupling constant, α_s, is $O(0.1)$ at typical hard-scatter energy scales at which the pp constituents interact in LHC pp collisions. Therefore, these interactions can be modeled reliably using perturbation theory. In contrast, the modeling of the initial state, i.e., the dynamics of the confined partons, cannot be performed using perturbative QCD. However, these two aspects can be treated separately, as described by the factorization theorem [126]. The cross section of a pp collision producing a final state X is given by Equation (16), assuming that partons carry only longitudinal momentum.

$$\sigma_{pp\to X} = \sum_{i,j} \int_0^1 \mathrm{d}x_i \int_0^1 \mathrm{d}x_j f_{i,p}(x_i, \mu_\mathrm{F}^2) f_{j,p}(x_j, \mu_\mathrm{F}^2) \sigma_{ij\to X}(x_i, x_j, s, \mu_\mathrm{R}^2, \mu_\mathrm{F}^2) \tag{16}$$

$$\sigma_{ij\to X} = \int \mathrm{d}\Phi_X \frac{1}{2\hat{s}} |\mathcal{M}_{ij\to X}|^2(\Phi_X, \mu_\mathrm{R}, \mu_\mathrm{F}) \tag{17}$$

The cross section $\sigma_{ij\to X}$ of the hard-scattering process of two partons i and j leading to the final state X, which is referred to as the partonic cross section, is weighted by the probabilities that the partons i and j carry momentum fractions x_i and x_j of their respective protons. For Run 2, these fractions are given by $x_i = \frac{p_i}{\sqrt{s}/2} = \frac{p_i}{6.5\,\mathrm{TeV}}$, where p_i is the momentum of parton i. The probabilities are given by the parton distribution functions (PDFs), $f_{i,p}(x_i, \mu_\mathrm{F}^2)$, of the proton, which are universal and independent of the underlying process, and must be derived for each type of proton constituent. Similar to the renormalization scale μ_R (cf. Equation (5)), the factorization scale μ_F is an energy scale introduced when absorbing infrared divergences occurring in the collinear limit of parton scattering into the PDFs. Thus, both the PDFs and the partonic cross section become μ_F-dependent quantities. The dependence of the PDFs on μ_F can be modeled using perturbation theory and is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [127–129]. Conversely, the dependence on x_i cannot be modeled by perturbation theory and is typically extracted empirically from experimental data that is sensitive to the proton's substructure. The PDFs are extracted mainly for two different schemes: the four-flavor scheme (4FS) and the five-flavor scheme (5FS). The 4FS considers gluons and u, d, s, and c quarks as partons, while the 5FS also includes the b quark as a parton. In the 4FS, b quarks are treated as massive particles in the calculation of the partonic cross section, whereas in the 5FS, they are considered massless.

The partonic cross section is given by Equation (17), where $|\mathcal{M}_{ij\to X}|^2$ is integrated over the phase space Φ_X, which is spanned by the four-momenta of the final-state particles, and scaled by the partonic center-of-mass energy, given by $\hat{s} = x_i\, x_j\, s$.

Top quarks are predominantly produced in pp collisions either in pairs, i.e., the production of a top quark and a top anti-quark, denoted by $t\bar{t}$, via the strong interaction or singly via the EW interaction. The ATLAS and CMS collaborations have reported the simultaneous production of four top quarks [130, 131], which occurs at a rate that is four orders of magnitude smaller than that of the predominant production mechanisms. The pair production is the dominant production mechanism, which primarily proceeds either via the interaction of two gluons, referred to as gluon fusion, or via the interaction of a quark and an anti-quark. In roughly 90% of all cases at the LHC, top-quark pairs are produced via gluon fusion. The predicted total production cross section in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$, assuming a top-quark mass of 172.5 GeV, is given by Equation (18). This cross section was obtained by performing a calculation with NNLO precision in QCD, including next-to-next-to-leading logarithm (NNLL) soft-gluon contributions, using the program Top++ 2.0 [132–138] employing the 5FS. The uncertainties correspond to variations of μ_F and μ_R (scales) and of the PDFs and the α_s value (PDF + α_s).

$$\sigma_{t\bar{t}} = 832^{+20}_{-29}(\mathrm{scales}) \pm 35(\mathrm{PDF} + \alpha_s)\ \mathrm{pb} \tag{18}$$

There are three main mechanisms to produce a single top quark via SM EW interactions in pp collisions: the t channel (tq) characterized by the exchange of a W boson between initial-state quarks; the s

channel ($t\bar{b}$) characterized by the production of an intermediate W boson decaying into a $t\bar{b}$ pair; and the W-boson-associated production (tW).[12] Figure 2 shows representative LO Feynman diagrams for each of these production modes using the 5FS.

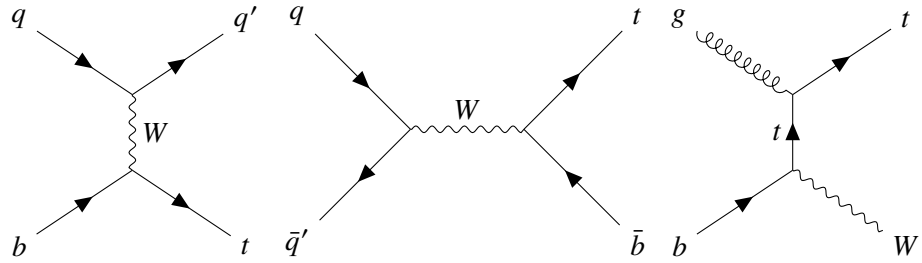


Figure 2: LO Feynman diagrams of single-top-quark production in pp collisions for the different production modes, t channel (left), s channel (middle), and tW channel (right), using the 5FS.

The CDF and DØ collaborations reported the observation of single top quarks produced either via the t channel or s channel in $p\bar{p}$ collisions in Refs. [139, 140]. The observation of single-top-quark production exclusively via the t-channel production was reported by the DØ collaboration [141]. By combining individual measurements, the CDF and DØ collaborations achieved the first observation of the single-top-quark production exclusively via the s channel in Ref. [142]. The ATLAS and CMS collaborations observed single-top-quark production via the t channel and via the tW channel in pp collisions at $\sqrt{s} = 7$ TeV [143, 144] and $\sqrt{s} = 8$ TeV [145, 146], respectively.

The predicted total production cross sections in pp collisions at $\sqrt{s} = 13$ TeV for the different single-top-quark production channels are given by Equations (19), (20), and (21) for the t channel, s channel, and tW channel, respectively. The cross sections for the t channel and s channel are calculated including NLO QCD contributions using the Hathor 2.1 [147, 148] program. The prediction of the cross section for tW production includes NLO QCD contributions and additional NNLL soft-gluon corrections [149, 150]. All cross sections are calculated using the 5FS.

$$\sigma_{tq} = 217^{+7}_{-5}(\text{scales}) \pm 6(\text{PDF} + \alpha_s)\ \text{pb} \tag{19}$$

$$\sigma_{t\bar{b}} = 10.32^{+0.29}_{-0.24}(\text{scales}) \pm 0.27(\text{PDF} + \alpha_s)\ \text{pb} \tag{20}$$

$$\sigma_{tW} = 71.7^{+1.8}_{-1.8}(\text{scales}) \pm 3.4(\text{PDF})\ \text{pb} \tag{21}$$

Therefore, it is expected that at least one top quark was produced ≈ 130 million times via pair production and ≈ 47 million times via single production during Run 2 in pp collisions at the ATLAS experiment. This represents an unprecedentedly large top-quark dataset, so that the LHC can be considered a top-quark factory. This dataset not only allows for entering the precision era of top-quark physics, i.e., top-quark properties such as its mass, width, and production and decay properties can be measured with an excellent statistical precision, it also allows for probing processes that involve top quarks with small production cross sections. While this represents an enormous opportunity, it also represents a challenge, as theory-related uncertainties become dominant limiting factors for the precision of top-quark measurements. Hence, a deeper understanding of the modeling of top-quark physics is required.

It is especially intriguing to use this large dataset to study top-quark production in association with a boson of the EW sector, i.e., with a photon, a Z boson, a $W^{\pm}$ boson, or a Higgs boson. These processes are characterized by cross sections typically suppressed by a few orders of magnitude relative to inclusive top-quark production. They can be used to directly examine the coupling of the top quark

[12] Charge-conjugated processes, e.g., the single production of an anti-top quark, are implicitly included when referring to different production mechanisms throughout this thesis. For quarks, the distinction between a quark (q) and an anti-quark ($\bar{q}$) is maintained where physically relevant. The symbol q typically refers to any quark, with its specific type and charge being clear from the context of the interaction.

to the corresponding boson as well as related properties, such as, in photon-associated production, the top-quark charge, MDM, and EDM, as discussed in Section 2.4. Therefore, these examinations are important tests of the consistency of the SM. Furthermore, as discussed in the previous sections, the top quark is considered a key factor in revealing the mysteries associated with the EWSB. Hence, measurements of these processes not only help to sharpen the understanding of top-quark physics but might also provide an avenue for the discovery of BSM physics. Associated top-quark pair production has been observed with a Z boson ($t\bar{t}Z$) [151, 152], a W boson ($t\bar{t}W$) [151], a photon ($t\bar{t}\gamma$) [153], and a Higgs boson ($t\bar{t}H$) [154, 155]. In contrast, associated single-top-quark production has only been observed with a Z boson in the t-channel topology [156, 157], apart from the well-established tW-channel production.

3.3 Single-top-quark production in association with a photon in pp collisions

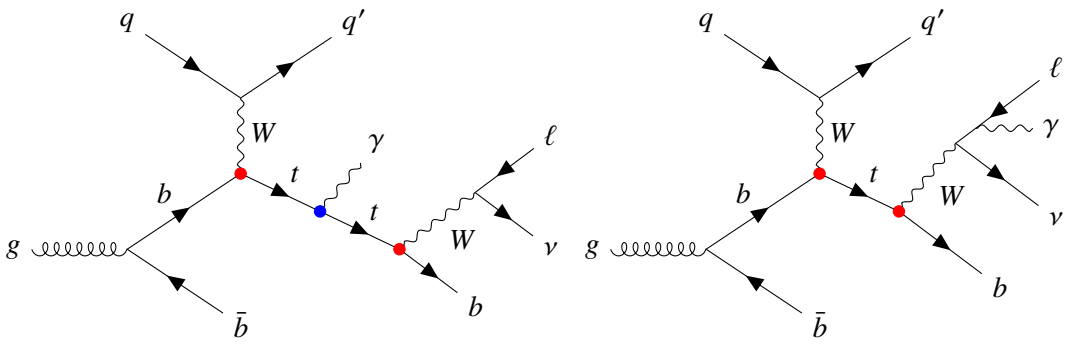


Figure 3: LO Feynman diagrams of t-channel single-top-quark production in association with a photon and the subsequent semileptonic decay of the top quark. The left diagram shows the "production" mode, denoted by $tq\gamma$, where a photon (γ) is either radiated from initial-state particles or during the production of the top quark (t). The right diagram shows the "decay" mode, denoted by $t\,(\to \ell\nu b\gamma)\,q$, where a photon is radiated from any of the electrically charged decay products of the top quark. The red dots mark vertices that are affected by the EFT operators O_{uW}^{33}, while the blue dot marks a vertex that is affected by both the EFT operators O_{uW}^{33} and O_{uB}^{33}. For these diagrams, q denotes an up-type quark and q' denotes a down-type quark. These two modes also arise when the up-type quark is replaced by a down-type anti-quark and the down-type quark is replaced by an up-type anti-quark.

This work investigates single-top-quark production in association with a photon. The observation of this process is one of the primary goals, contributing a missing piece to the observed associated top-quark production processes. The s-channel and tW-channel topologies of single-top-quark production share greater similarity with the $t\bar{t}$ topology than the t-channel topology does, which makes it more challenging to disentangle these processes from the dominant $t\bar{t}$ production. Moreover, the production cross sections for the s channel and tW channel are ≈ 20 and ≈ 3 times smaller than that for the t channel, respectively. Therefore, the t-channel topology offers the best prospects for observing single-top-quark production in association with a photon in pp collisions at the LHC, and for precisely probing the top-quark–photon interaction.

LO Feynman diagrams for the t-channel production of a single top quark in association with a photon are depicted in Figure 3 using the 4FS. An up-type quark (q) from one of the protons interacts with a b quark originating from the splitting of a gluon into a $b\bar{b}$ pair, via the exchange of a W boson. The up-type quark is transformed into a down-type quark (q'), while the b quark is transformed into a top quark. At LO, this process also occurs with an anti-down-type quark being transformed into an anti-up-type quark. As discussed in Section 2.3, the top quark decays almost exclusively into a W boson and a b quark. While the hadronic top-quark decay offers a larger branching ratio ($\approx 66\%$) relative to the semileptonic decay ($\approx 33\%$), it gives rise to a final state consisting only of color-charged particles, apart from the photon. Such final states are produced at high rates in pp collisions, so that the hadronic decay channel offers

little potential for observing the examined process. In fact, none of the single-top-quark production mechanisms have been observed when only hadronic decays of W bosons are involved. Therefore, studying t-channel single-top-quark production using the semileptonic top-quark decay channel is the only reasonable choice.

The photon can be emitted from any of the electrically charged particles participating in the t-channel single-top-quark production, i.e., from all particles except the gluon in the example LO Feynman diagrams of Figure 3. Therefore, the photon can be radiated either from one of the electrically charged top-quark decay products, i.e., the b quark, the W boson, or the lepton, or from any other particle occurring in the initial or final states. The former is referred to as the "production" mode and is denoted by $tq\gamma$, while the latter is referred to as the "decay" mode and is denoted by $t\,(\to \ell\nu b\gamma)\,q$. As both processes give rise to the same final states ($pp \to \ell\nu b\gamma q'\bar{b}$ at LO), they interfere and are, in principle, inseparable. However, the processes are characterized by distinctive kinematic properties, so that interference effects are small, and a separate examination of the processes is reasonable in practice. This will be illuminated in detail in Chapter 6. Furthermore, both processes are sensitive to the EFT operator O_{uW}^{33} due to the occurrence of tW vertices in each of the processes. However, only $tq\gamma$ is sensitive to O_{uB}^{33} because it involves a top-quark–photon vertex, which is also sensitive to O_{uW}^{33} due to the relation shown in Equation (3). Hence, $tq\gamma$ is of special interest for the investigation of the top-quark–photon coupling and related properties, e.g., the MDM and EDM. An investigation of the process in the context of EFT is beyond the scope of this thesis.

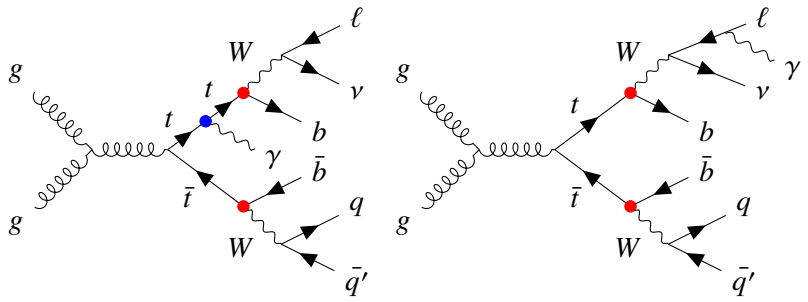


Figure 4: LO Feynman diagrams of double-resonant top-quark pair production in association with a photon with the subsequent decays of the top quarks. While one of the top quarks decays semileptonically, the other top quark decays hadronically. The left diagram shows the "production" mode which is denoted by $t\bar{t}\gamma$ (prod), where a photon is radiated either from initial-state particles or during the production of the top-quark pair. The right diagram shows the "decay" mode denoted by $t\bar{t}\gamma$ (dec), where one of the top quarks undergoes a radiative decay, either $t \to l\nu b\gamma$ or $t \to q\bar{q}'b\gamma$. The red dots mark vertices that are affected by the EFT operator O_{uW}^{33}, while the blue dot marks a vertex that is affected by both the EFT operators O_{uW}^{33} and O_{uB}^{33}.

Photon-associated $t\bar{t}$ production is also sensitive to the top-quark–photon coupling and O_{uW}^{33} and O_{uB}^{33}. Figure 4 shows LO Feynman diagrams for this process. As with photon-associated single-top-quark production, this process can be categorized into the "decay" mode, denoted by $t\bar{t}\gamma$ (dec), where a photon is radiated from one of the electrically charged decay products of the top-quark pair, and the "production" mode, denoted by $t\bar{t}\gamma$ (prod), where the photon is radiated from any other electrically charged particle participating in this process. Both categories give rise to the same final states ($pp \to \ell\nu b\gamma q\bar{q}'\bar{b}$ at LO), and the sum of both categories is referred to as $t\bar{t}\gamma$. Due to its larger cross section in pp collisions, it is possible to investigate $t\bar{t}\gamma$ using smaller pp-collision datasets than those required for photon-associated single-top-quark production. Therefore, it has been examined extensively in experiments and by theorists, and is a well-established process. The most recent cross-section measurement of $t\bar{t}\gamma$ production performed by the ATLAS collaboration reached a precision of 5% [158]. Furthermore, the ATLAS collaboration reported the observation of top-quark pair production in association with two photons for the first time [159].

Refs. [160, 161] investigated the potential contribution of $tq\gamma$ measurements to the bounds on the MDM

and EDM for different data-taking scenarios of the LHC. Both works conclude that measurements of the $tq\gamma$ process can reach a sensitivity similar to those of the $t\bar{t}\gamma$ process. Hence, investigations of $tq\gamma$ are not only complementary to those of $t\bar{t}\gamma$, but also enable a deeper understanding of the top-quark–photon interaction and significantly enhance the sensitivity to potential BSM physics affecting the top-quark–photon interaction.

The CMS collaboration reported evidence of the $tq\gamma$ process with a statistical significance of 4.4σ [20]. The analyzed dataset corresponds to 35.9 fb^{-1} of Run 2 pp collisions, with only $tq\gamma$ final states containing one muon and no other lepton considered. Furthermore, a fiducial cross section was measured in a phase space defined by $p_{\mathrm{T}}^{\gamma} > 25\,\mathrm{GeV}$, $|\eta^{\gamma}| < 1.44$, and $\Delta R(X, \gamma) > 0.5$, where X refers to any final-state object. The result of this measurement and the nominal prediction obtained from the software program MadGraph5_aMC@NLO [162] including NLO QCD contributions are shown in Equations (22) and (23), respectively.[13] Systematic uncertainties are the dominant factors limiting the precision of the result, while the total uncertainty is $\approx 30\%$.

$$\sigma^{\text{meas.}}_{pp\to t(\to\mu\nu b)q\gamma} = 115 \pm 17\,(\text{stat.}) \pm 30\,(\text{syst.})\ \text{fb} \tag{22}$$

$$\sigma^{\text{pred.}}_{pp\to t(\to\mu\nu b)q\gamma} = 81 \pm 4\ \text{fb} \tag{23}$$

The inclusion of final states containing electrons and the use of the full Run 2 pp-collision dataset taken by the ATLAS experiment result in a dataset roughly eight times larger than that used in Ref. [20], so that the observation of the process was considered feasible using this dataset. Ref. [163] studies $tq\gamma$ at NLO QCD as well as the impact of NLO EW contributions, and provides cross-section predictions for the fiducial phase space used in the CMS analysis. Ref. [164] studies $tq\gamma$ at NLO QCD and in addition the impact of approximate NNLO QCD contributions by including soft-gluon resummation, and finds that the resulting SM predictions are in good agreement with the cross section reported by the CMS collaboration. Both works were published after the evidence report of the CMS collaboration and represent, to the authors' best knowledge, the only studies of $tq\gamma$ performed by theorists in the context of the SM beyond LO to date.

Single-top-quark production in association with a photon can also occur via the top-quark–photon interaction through an FCNC. There are mainly two types of such interactions, referred to as the "FCNC decay" mode and the "FCNC production" mode. Example LO Feynman diagrams are depicted in Figure 5. The "FCNC decay" mode refers to $t\bar{t}$ production, where one of the top quarks decays via an FCNC $t \to U\gamma$ transition, with $U = \{u, c\}$, while the "FCNC production" mode refers to the production of a single top quark via an FCNC transition, e.g., $gU \to t\gamma$. The vertices corresponding to FCNC interactions are marked by a red dot. The FCNC interactions can either be described in the context of the SM by the process depicted in Figure 1 or by the EFT operators O_{iB}^{i3} and O_{iW}^{i3} shown in Equations (12) and (13), respectively. As discussed in Section 2.3, the rate of this process within the SM is highly suppressed, so that any observation of such a process with the LHC dataset would represent an observation of unknown physical phenomena.

Previous searches for FCNC top-quark–photon interactions have been conducted in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ by the CMS collaboration [165] and in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ by the ATLAS collaboration using a partial Run 2 dataset [166], focusing on the "FCNC production" mode. The most recent searches performed by the ATLAS [11] and CMS [167] collaborations used the respective full Run 2 datasets. Both works investigate the "FCNC production" and the "FCNC decay" modes. Unlike the CMS collaboration, the ATLAS collaboration investigated left-handed (LH) and right-handed (RH) chiralities separately. In the absence of significant evidence for the presence of these interactions in the analyzed datasets, the main results of these works are formulated as upper bounds on the branching ratios for the FCNC top-quark–photon interactions. The results from the ATLAS collaboration are shown in Equations (24)–(27), and those from the CMS collaboration in Equations (28)–(29). The

[13] Details on the composition of the uncertainty for the predicted value are not given in the reference.

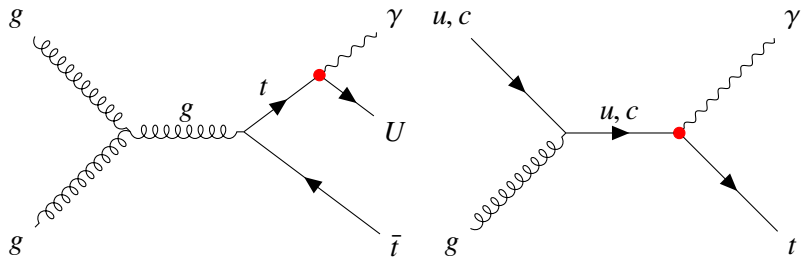


Figure 5: LO Feynman diagrams of single-top-quark production in association with a photon through FCNC interactions. The left diagram shows the "FCNC decay" mode, where a top quark produced in pair production decays via an FCNC. The right diagram shows the "FCNC production" mode, where a single top quark is produced via an FCNC interaction. The red dots mark vertices of FCNC interactions. The FCNC interaction is illustrated in Figure 1 in the context of the SM. In the context of EFT, this vertex is affected by the EFT operators O_{uW}^{l3} and O_{uB}^{l3}, with $l = \{1, 2\}$.

constraints on the branching ratios of $t \to u\gamma$ transitions are of comparable size, whereas the constraint on the branching ratio of $t \to c\gamma$ transitions reported by the CMS collaboration is significantly more stringent. Given the similarity between the topologies of the FCNC top-quark–photon interaction and those examined in this thesis, and the common experimental challenges their investigations face, some methods and results presented in this work were also used in Ref. [11]. These instances are highlighted throughout this thesis.

$$\mathcal{BR}(t \to u\gamma)(\mathrm{LH}) < 0.85 \cdot 10^{-5} \tag{24}$$

$$\mathcal{BR}(t \to u\gamma)(\mathrm{RH}) < 1.22 \cdot 10^{-5} \tag{25}$$

$$\mathcal{BR}(t \to c\gamma)(\mathrm{LH}) < 4.16 \cdot 10^{-5} \tag{26}$$

$$\mathcal{BR}(t \to c\gamma)(\mathrm{RH}) < 4.46 \cdot 10^{-5} \tag{27}$$

$$\mathcal{BR}(t \to u\gamma) < 0.95 \cdot 10^{-5} \tag{28}$$

$$\mathcal{BR}(t \to c\gamma) < 1.51 \cdot 10^{-5} \tag{29}$$

4 Physics with the ATLAS experiment

Four large experiments are installed at the IPs of the LHC to examine the high-energy pp collisions. The experiments are either specialized in a specific area of particle physics or designed as general-purpose experiments that allow for the study of a wide range of particle-physics phenomena. The ATLAS experiment belongs to the latter category and is well-suited for conducting physics studies involving the top quark. In this work, datasets recorded with this experiment are used to observe and study t-channel single-top-quark production in association with a photon. Therefore, this chapter reviews its design, the main sub-detector components, and the data-taking procedure. Accurate simulation of data recorded by the ATLAS experiment is a key element in numerous physics studies. The main steps of this simulation, along with the related software, are briefly discussed. Software algorithms are used to combine the signals from different sub-detector systems into physics objects and associate these with well-known particles, such as photons. This chapter concludes with an overview of the definition and reconstruction of these physics objects.

4.1 Overview of the ATLAS detector

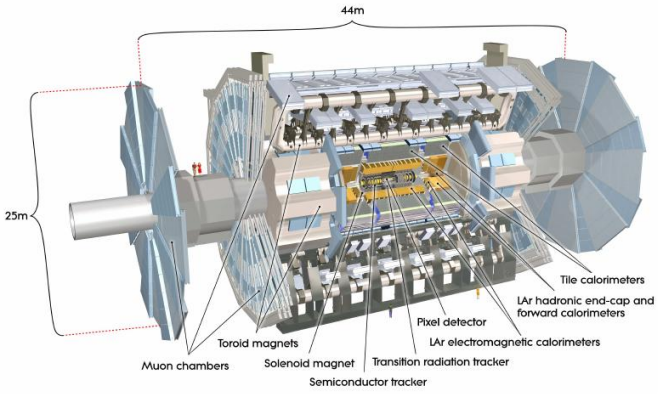


Figure 6: Cut-away view of the ATLAS detector [14]. The design prior to Run 2 is shown, i.e., without the fourth layer of the pixel detector installed as innermost layer ahead of Run 2.

The ATLAS detector is built symmetrically around the beam axis of the LHC and has a cylindrical shape [14]. It is 44 m long, while its diameter is 25 m. Its design offers nearly full solid-angle coverage, allowing for detailed reconstruction of the hard-scattering pp interactions. To describe positions and directions in the detector, a right-handed coordinate system, as defined in Section 3.1, is used, where the origin coincides with the nominal IP in the center of the detector. Figure 6 illustrates the ATLAS detector and its main sub-detector systems. Brief overviews of the sub-detector systems during Run 2, as well as of the data-taking procedure, are provided in the following.

Inner detector Many particles emerge from a single inelastic pp collision, while multiple such interactions occur simultaneously. The inner detector (ID) is the part of the ATLAS detector located closest to the beam pipe. Thus, the ID is exposed to the largest particle flux of all components of the ATLAS detector and its material is therefore required to be characterized by an excellent radiation hardness. The main purpose of the ID is to reconstruct the trajectories of electrically charged particles from its sensor signals. The high particle density requires a finely granulated sensor structure to disentangle the trajectories of different particles. This is an essential requirement for the precise and efficient reconstruction of common origins of particles, referred to as vertices, while ensuring low false-assignment rates. These vertices are either primary vertices (PVs), which correspond to the space points associated with pp interactions, or secondary vertices (SVs), which arise when particles decay before entering the detector volume at a space point significantly displaced from the IP and any PV.

The occurrence of SVs is characteristic of b hadrons, which are produced with a high probability in interactions involving top quarks.

The ID is immersed in a 2 T axial magnetic field that is generated by a thin solenoid magnet, which bends the trajectories of electrically charged particles. This allows measuring the transverse momentum (p_{T}) of such particles and determining the value and sign of their electric charge. Consequently, the particle momenta must be affected as little as possible by the ID material. The ID mainly consists of three parts that will be briefly reviewed in the following.

Pixel detector The innermost part of the ID is the silicon pixel detector, providing the finest granularity and highest spatial resolution of all components of the ATLAS detector. It consists of a barrel part comprising four layers of pixel modules that are arranged concentrically around the beam pipe, and an end-cap part consisting of three discs of pixel modules on each side of the barrel, perpendicular to the beam axis. The size of most (> 90%) pixel sensors in the end-cap discs and in the last three barrel layers is 50 µm × 400 µm, while the remaining pixel sensors have a size of 40 µm × 600 µm. The intrinsic accuracy of pixel modules with the smaller-sized pixel sensors is 10 µm in the r–ϕ plane[14] and 115 µm in the z (in the barrel region) or r (in the end-cap region) directions. The innermost barrel layer, the Insertable B-Layer (IBL) [168], was installed at a distance of ≈ 33 mm from the beam pipe between Run 1 and Run 2 of the LHC program and marks a significant improvement in tracking precision and vertex reconstruction. The pixel sensors in this layer have a size of 50 µm × 250 µm. The intrinsic single-module accuracy for pixel modules consisting of IBL pixel sensors is 10 µm in the r–ϕ plane and 60 µm in the z direction [169].

Semiconductor tracker The semiconductor tracker (SCT) encloses the pixel detector and provides tracking information using silicon microstrip sensors. The structure of the SCT is similar to that of the pixel detector but is characterized by an overall coarser granularity. It consists of four layers in the barrel part and nine discs of silicon microstrip modules on each side of the barrel region, extending the tracking abilities up to $|\eta| = 2.5$. The barrel modules consist of four silicon sensors with a strip pitch of 80 µm and a length of 6.4 cm. Two of these sensors are daisy-chained and form a 12.8 cm-long sensor. One such sensor is installed on both sides of the module. The sensors are rotated by an angle of 40 mrad relative to each other. The end-cap regions use different types of silicon modules with trapezoidal sensors with an average strip pitch of ≈ 80 µm (56.9–90.4 µm), which differ in length between ≈ 52–120 mm (active size). The effective single-module accuracy, obtained by the relative orientation of the back-to-back sensors, is pitch-size dependent and is about 17 µm in the r–ϕ plane and 580 µm in the z and r directions in the barrel and end-cap regions, respectively.

Transition radiation tracker The outermost part of the inner detector is the transition radiation tracker (TRT), which provides tracking information for $|\eta| < 2.0$. It comprises polyamide straw tubes with a diameter of 4 mm filled with a gas mixture of 70% xenon, 27% carbon dioxide, and 3% oxygen, which are operated as proportional drift tubes. The barrel part comprises 73 layers of straw tubes with a length of 144 cm arranged in three rings parallel to the beam pipe. Each of the end-cap parts has 160 planes of straw tubes with a length of 37 cm arranged in wheels with the straw tubes oriented radially to the beam pipe. Electrically charged particles are tracked by determining the drift time of electrons produced by gas ionization with an intrinsic accuracy of 130 µm in the r–ϕ plane. The space between the straw tubes is filled with polypropylene fibers and foils in the barrel and end-cap regions, respectively. Electrically charged particles produce transition radiation (TR) when passing through this material. The energy of the TR photons depends linearly on the Lorentz γ-factor of the traversing particle, which is at least ≈ 200 times larger for electrons than for any other electrically charged particle carrying the same energy. Xenon is highly efficient at capturing these TR photons with typical energies between 6–15 keV when emitted by electrons. The resulting gas ionization produces a signal much larger than the tracking signals, providing a powerful tool for electron identification. After Run 1, gas leaks were discovered in

[14] $r = \sqrt{x^2 + y^2}$

some parts of the TRT. The affected parts were filled with an argon-based gas mixture, which was used during the operation of the detector in Run 2.

Calorimeter system The ATLAS detector uses a calorimeter system, which mainly consists of three parts: the electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL), and forward calorimeter (FCAL). Its purpose is to stop electrons, photons, and hadrons from pp collisions that enter its volume and to measure their energy with high precision. High-energy particles that enter the calorimeters interact with their material and produce cascades of secondary particles, which are referred to as showers. The depth of a shower in the calorimeter and its shape depend on the energy and the type of particle that initiates it. Showers are mainly classified into two categories: EM showers, which are induced by either electrons or photons, and hadronic showers, which are induced by hadrons. Hadronic showers typically start deeper in the calorimeter and are more diffuse relative to EM showers. Therefore, measurements of the longitudinal and lateral shower profile, also referred to as the shower shape, provide powerful tools for particle identification. The individual parts of the calorimeter system will be described in the following.

Presampler The presampler is installed in front of the calorimeter system and provides a measure of the energy loss of particles upstream of the calorimeter. It is thus an important ingredient for measuring their total energy and is realized as a thin, single-layer calorimeter comprising liquid argon (LAr) as the active material. It is divided into a barrel part covering $|\eta| < 1.52$ and an end-cap part covering $1.5 < |\eta| < 1.8$. Both parts have a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$.

Electromagnetic calorimeter The ECAL is the innermost part of the calorimeter system used in the ATLAS detector. Its main task is to fully contain EM showers and to enable precision studies of photons and electrons by providing energy measurements with a fine granularity. The ECAL is divided into a barrel part covering $|\eta| < 1.475$ and an end-cap part covering $1.375 < |\eta| < 3.2$. LAr modules are used as active material, embedded within gaps of lead absorbers that are installed in an accordion geometry. The readout copper electrodes are installed in the middle of these gaps.

The barrel calorimeter is realized as two half-barrels separated by a small gap at $z = 0$. It is segmented into three layers, mainly characterized by different levels of $|\eta|$ granularity. The first layer, also called the strip layer, is characterized by the finest granularity of all calorimeter sections. In its most central section ($|\eta| < 1.40$), it provides a granularity of $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$, which is well suited for precise determination of lateral shower widths. The second layer, which has the greatest depth of the three parts, has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The third layer provides energy measurements up to $|\eta| < 1.35$ and is the coarsest layer with $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$.

The end-cap part is realized on each side of the barrel by two wheels, an inner and an outer wheel, covering $2.5 < |\eta| < 3.2$ and $1.375 < |\eta| < 2.5$, respectively. While the outer wheels consist of three layers, the inner wheels comprise two layers with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in each layer. The granularity of the first layers of the outer wheels varies with $|\eta|$, providing a fine granularity varying from $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$ to $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ in the range $1.5 < |\eta| < 2.5$. The second and third layers are characterized by a finer ϕ granularity of $\Delta\phi = 0.025$, while the η granularity is $\Delta\eta = 0.025$ in the second layers for $1.425 < |\eta| < 2.5$ and $\Delta\eta = 0.05$ elsewhere.

Hadronic calorimeter The HCAL encloses the ECAL and is divided into barrel, extended barrel, and end-cap segments, covering $|\eta| < 1.0$, $0.8 < |\eta| < 1.7$, and $1.5 < |\eta| < 3.2$, respectively. Its main tasks are to stop hadronic showers, as well as EM showers that leak into it, and to measure their energies. The barrel and extended barrel parts are designed as tile sampling calorimeters, where steel tiles are used as absorber material, alternating with plastic scintillator tiles used as active material. These parts are segmented into three layers, with the first two layers providing a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, while the last layer has a coarser granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$. Each of the end-cap section consists of two wheels, a front and a rear wheel, which use LAr as the active and copper plates as the

absorber materials. These two wheels are segmented into two layers each and provide a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $1.5 < |\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $2.5 < |\eta| < 3.2$.

Forward calorimeter The FCAL is the part of the ATLAS calorimeter system that allows for energy measurements of particles that are produced at small polar angles (large $|\eta|$). It covers an η range of $3.1 < |\eta| < 4.9$ and consists of three layers with LAr as the active material. The first layer is designed as an ECAL layer and uses copper as the absorber material to stop EM showers, while the second and third layers are designed as HCAL layers, employing tungsten as the absorber material. A coarse granularity of $\Delta x \times \Delta y = 3.0 \times 2.6\,\mathrm{cm}^2$, $\Delta x \times \Delta y = 3.3 \times 4.2\,\mathrm{cm}^2$, and $\Delta x \times \Delta y = 5.4 \times 4.7\,\mathrm{cm}^2$ is used for the first, second, and third layers in $3.15 < |\eta| < 4.30$, $3.24 < |\eta| < 4.50$, and $3.32 < |\eta| < 4.60$, respectively. The granularity is approximately four times finer in the remaining $|\eta|$ region.

Muon spectrometer The only detectable SM particles that typically escape the calorimeters are muons. The muon spectrometer (MS) is a tracking detector immersed in a magnetic field generated by three superconducting air-core toroid magnets. This design allows for precise momentum and charge determination by using the curvature of the muon tracks.

The barrel part of the MS utilizes monitored drift tubes (MDTs) as precision tracking devices. The barrel toroid magnet is realized as a system of eight coils, placed symmetrically and equidistantly around the beam pipe in racetrack-shaped vessels, providing a bending power of 1.5–5.5 T m in $|\eta| < 1.4$. Three layers of MDT chambers are installed in concentric cylindrical shells, with each layer divided into octants. Each octant contains two sectors of MDT chambers with smaller and larger lateral dimensions.

Each of the two end-cap parts also comprises an air-core toroid magnet with eight racetrack-shaped coils, installed in a common aluminum alloy structure. The coils are smaller than the barrel coils and are rotated by an angle of $\pi/8$ relative to them. These magnets provide a bending power of 1–7.5 T m each for $1.6 < |\eta| < 2.7$. The magnetic fields of the barrel toroid and end-cap toroids overlap in $1.4 < |\eta| < 1.6$, resulting in a weaker bending power in this region. Three wheels of precision tracking chambers are installed in the end-cap parts, one in front of and two behind the end-cap toroid. The inner wheel uses cathode strip chambers (CSCs) as precision tracking devices in $2.0 < |\eta| < 2.7$, while MDT chambers are used elsewhere. The chambers are arranged similarly to those in the barrel toroid, i.e., each wheel is segmented into octants, each containing one smaller and one larger precision tracking chamber.

The MDT chambers and CSCs allow measuring the coordinate of muon tracks in the bending plane with accuracies of 35 µm and 40 µm, respectively. The CSCs also measure the coordinate in the non-bending plane with an accuracy of ≈ 5 mm.

Muon Spectrometer (Trigger Chambers) The MS is instrumented with trigger chambers. These offer a significantly lower resolution of tracks in the bending plane, but they provide excellent timing resolution, allowing for tagging individual bunch crossings. Furthermore, they provide well-defined p_{T} thresholds for triggering and additionally measure the coordinate of tracks in the non-bending plane with an accuracy of 3–10 mm, which can be combined with measurements from the precision tracking chambers.

Two types of trigger chambers are installed in the MS: the resistive plate chambers (RPCs) in the barrel region for $|\eta| < 1.05$ and the thin gap chambers (TGCs) in the end-cap region for $1.05 < |\eta| < 2.4$. Three layers of RPCs are installed in the MS, with two layers sandwiching the middle barrel MDT layer and the third layer installed close to the outermost barrel MDT layer. Four wheels of TGCs are installed in the end-cap part: the first wheel in front of the innermost wheel with precision tracking devices, the second in front of the middle MDT end-cap wheel, and the last two between the second and outermost MDT wheels.

Trigger system There are over 100 million readout channels installed in the ATLAS detector, and the bunch crossing frequency was typically 40 MHz during Run 2. As a consequence, processing every

pp-collision event to storage would produce an enormous data stream, which is not feasible to handle. Therefore, it is necessary to reduce the data rate in a fast and efficient manner. Detector signals that are consistent with the hypothesis of an interesting physics object must be preserved, while events without such signals need to be rejected as fast as possible.

To achieve this, a two-level trigger system is employed [170, 171]. This system consists of a first-level hardware-based trigger (L1) and software-based high-level triggers (HLTs). The L1 trigger receives information from different subsystems, such as the MS and the coarser layers of the calorimeters, and identifies regions of interest (RoIs). For example, it identifies detector regions that have signals consistent with those of a high-energy photon. If the requirements of the L1 trigger are fulfilled, the information about the RoIs is passed to the HLT system. The L1 trigger decision has a latency of $\approx 2.5\,\mu s$ and reduces the data rate to $\approx 100\,kHz$.

The software-based HLT utilizes the full detector granularity in either the RoIs or the entire event. If the event also passes the HLT requirements, the entire detector is read out for that event, and its data is stored permanently on a mass storage device. The rejection rate of the HLTs is $\approx 99\,kHz$, which results in a total data frequency of 1 kHz with a processing time of approximately 200 ms.

Run 2 dataset The LHC provided a total amount of pp collisions that corresponds to an integrated luminosity of $156\,fb^{-1}$ in Run 2. The ATLAS detector recorded $147\,fb^{-1}$ of these pp collisions. The deficit is mainly caused by the "warm start", the time required for the tracking detectors (especially the pixel detector in the ID) to become fully operational. The dataset is further filtered by applying data quality criteria to ensure that the ATLAS detector was fully operational [172]. The remaining dataset is marked as "Good for physics", meaning it is of high quality and suitable for physics studies. Hereafter, this dataset is referred to as the collected data/dataset. A preliminary value of the corresponding integrated luminosity of $139 \pm 2.4\,fb^{-1}$ is used throughout this work [123]. An updated value of $140.1 \pm 1.2\,fb^{-1}$, provided in Ref. [173], agrees well with the preliminary value within the uncertainties. The integrated luminosity as a function of time is shown in Figure 7. Distributions obtained from the collected data are referred to as "observed" distributions.

Figure 7 also shows the profile of the pile-up parameter μ for the different data-taking years. The LHC operated at a higher instantaneous luminosity in the later years. The double-peak structure in 2017 is caused by LHC operations performed at high instantaneous luminosity at the end of that year. A peak is also present for low values of μ, corresponding to operations at low instantaneous luminosity, which are not considered in this work. The year-averaged pile-up parameter is 13.4, 25.1, 37.8, and 36.1 for 2015, 2016, 2017, and 2018, respectively. This results in an overall average pile-up parameter of 33.7 for the full Run 2 dataset, i.e., each hard-scattering event was, on average, accompanied by ≈ 33 inelastic pp interactions.

4.2 Simulation of inelastic high-energy pp collisions at the ATLAS experiment

The data taken with the ATLAS detector are used for a variety of physics studies, mainly for testing the predictions of the SM, deepening the understanding of SM phenomena, and searching for BSM physics phenomena. In many of these analyses, predictions from the SM or BSM models are directly compared to observed distributions. Therefore, the model prediction at the level of elementary particles must be converted into the expected signals of the ATLAS detector. This conversion consists of several sequential steps, which are discussed for the pp collisions under the LHC conditions during Run 2 in this section.

The first step of the simulation chain is the modeling of the hard-scattering process $ij \to X$, based on Equation (16), where i and j represent partons and X represents the final state of interest. For this purpose, Monte Carlo (MC) event generators are used. These generators model hard-scattering processes using perturbation theory according to a specified model. A model typically contain several

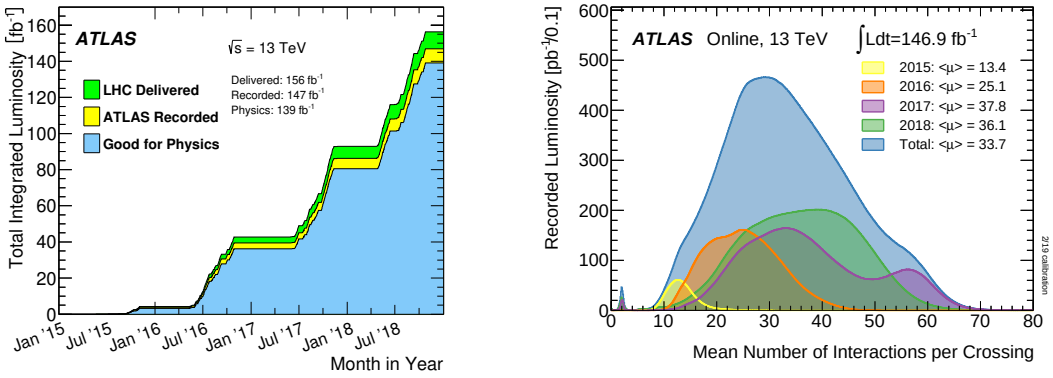


Figure 7: The evolution of the integrated luminosity over time during Run 2 is shown on the left. It is displayed for the *pp* collisions delivered by the LHC, for the dataset recorded by the ATLAS experiment, and for the recorded dataset used for physics studies ("Good for Physics"). The pile-up profiles for the different data-taking years of Run 2 are shown on the right. The double-peak structure observed for 2017 corresponds to high-luminosity LHC operations at the end of that year. The peak at low values corresponds to special LHC operations at low instantaneous luminosity, which are not considered in this work. Both figures are taken from Ref. [172].

free parameters, such as coupling strengths and particle masses, which can be specified by users (e.g., SM parameters can be set to their most precise experimental values). For each initial-state configuration ij that allows the final state X to be realized in the respective model, $|\mathcal{M}_{ij\to X}|^2$ is calculated according to the Feynman rules of the specified model. It is then integrated over the phase space spanned by the four-momenta of the final-state particles, and scaled by the partonic center-of-mass energy. The momenta of the partons i and j are sampled according to the chosen PDF set. As the phase space is typically characterized by high dimensionality, MC integration techniques are employed, which efficiently provide accurate results in such cases. The sampled phase-space points represent the simulated events, which contain all relevant information on the initial- and final-state particles, such as their type and four-momenta. An event weight w_i is assigned to each of these events, so that the sample of generated events reproduces the probability density functions (p.d.f.s) of kinematic and topological properties associated with the hard-scattering process, such as the differential cross section as a function of the transverse momentum of a final-state particle or the angular distance between final-state particles. This type of event is referred to as a parton-level event.

The parton-level events are propagated to a parton shower (PS) algorithm that simulates the dominant additional collinear and soft radiation from initial-state and final-state particles. This is especially important for color-charged particles, as they can radiate gluons. Gluons can also undergo splitting into a quark-antiquark pair. As a result, a cascade of secondary particles is produced. The energy scale of the PS decreases along its evolution. The simulation of additional QED radiation (photon radiation from electrically charged particles) is also performed.

Furthermore, the underlying event (UE) is simulated. This refers to additional interactions that are not associated with the hard-scattering process or with any of the participating particles but arise from the same *pp* collision, e.g., interactions of other partons of the colliding protons. These interactions are predominantly characterized by low-energy transfers and are described by specialized models that are typically tuned to data.

The evolution of the PS stops when the perturbative evolution becomes unreliable and the color-charged particles begin to form hadrons, which is referred to as hadronization. Empirical models are employed to simulate this phenomenon. The most commonly used models are the string model [174, 175] and the cluster model [176, 177]. Hadrons that originate from the same cascade of secondary particles are spatially clustered, forming a cone-like structure known as a jet. The momentum direction of the jet follows that of the original color-charged particle (parton) that initiated the cascade. This is known as the parton-jet duality.

Many particles have short lifetimes and thus typically decay before they can be measured by the ATLAS detector. These decays are also simulated. The events obtained after applying all aforementioned simulation steps consist of leptons, neutrinos, photons, and hadrons, and are referred to as particle-level events.

The final step in the simulation chain is the simulation of the path of the particles through the ATLAS detector and the corresponding interactions in the detector. This response is modeled either with the GEANT4 toolkit, known as full simulation [178], which employs detailed descriptions of the ATLAS detector and interactions with its material, or with a combination of the full simulation for the ID and MS, and a faster simulation for the response of the calorimeter system based on a parametric description, which is provided by the FASTCALOSIM package [179]. The latter type of simulation is referred to as ATLFAST-II (AFII) simulation.

To incorporate effects from additional pp collisions taking place in the same bunch crossing (in-time pile-up) as well as in bunch crossings shortly before and after (out-of-time pile-up), minimum-bias events are overlaid before the digitization [178] of the simulated events. These events represent pp collisions generated without imposing the requirement for a hard-scattering interaction, so that they include non-diffractive and diffractive inelastic collisions. PYTHIA 8 [180], configured with the A3 set of tuned parameters [181] and the NNPDF2.3LO PDF set [182], is used to simulate these events. The "standard pile-up" procedure, as defined in Ref. [183], is applied for the overlaying.

After digitization, the simulated events and events recorded with the ATLAS detector are stored in the same format and processed in the same way. They are propagated through the software algorithms used for event reconstruction and physics analysis. Data and simulation may exhibit different responses to these algorithms. The differences are corrected by dedicated calibration methods [184]. These corrections are typically derived as a function of kinematic or topological variables of the event and adjust the event yield by reweighting using scale factors or the value of event quantities, such as the energy of a reconstructed object. The simulated events contain a record of particle-level information, which allows studying the origin of the detector signatures.

4.3 Definition and reconstruction of physics objects

The design of the ATLAS detector allows for a detailed reconstruction of the hard-scattering process. To achieve this, the final-state objects from the topology under study need to be connected to the corresponding recorded detector signatures. While a large variety of particles are created in LHC pp collisions, only a few types can reach the ATLAS detector. These particles are electrons, muons, photons, neutrinos, mesons, or baryons. Dedicated reconstruction, isolation, and identification algorithms that combine information from the different sub-detectors are applied to form physics objects that most likely correspond to prompt electrons, muons, or photons. The reconstructed objects are referred to as candidates. Electrons, muons, and photons are considered prompt if they are produced in the hard-scattering process or in the decay of W, Z, or Higgs bosons, or the top quark. Photons produced during quark fragmentation are also considered to be prompt photons. As described in the previous section, mesons and baryons (e.g., $\pi^{\pm}$, $K^{\pm}$, p, and n) typically arrive at the detector in the form of jets and are reconstructed as such, i.e., no dedicated algorithms are applied to identify individual hadron types, but detector signatures are identified that are most likely caused by jets. Jets are mainly categorized into three types. Jets compatible with the hypothesis that they contain constituents originating from the decay of a b hadron are referred to as b jets. If this is not the case, but they are compatible with the hypothesis that they contain components originating from the decay of a c hadron, they are referred to as c jets. These two type of jets are commonly referred to as HF jets. All other jets are referred to as LF jets.

The pp collisions provided by the LHC are head-on, i.e., the initial-state particles of the hard-scattering process carry almost zero transverse momentum, and hence, the sum of all transverse momenta of particles that originate from the same hard-scattering process is expected to nearly vanish. As the

ATLAS detector offers nearly full coverage of the solid angle, it is possible to measure the value of this sum for each event. Any significant imbalance is a strong indication of undetected particles. In the SM, neutrinos are the only particles expected to cause such a significant imbalance. Due to their weak interaction with matter, neutrinos escape the ATLAS detector without leaving significant traces in most cases. The negative $\vec{p}_{\mathrm{T}}$ sum is referred to as the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, while its absolute value is referred to as missing transverse energy (MET) denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$.

$$E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}| = \left| -\sum \vec{p}_{\mathrm{T}} \right| \tag{30}$$

The definition of the physics objects and the applied reconstruction, isolation, and identification algorithms are briefly discussed in the following. A more detailed discussion of the identification algorithm applied to photon candidates is provided in Section 5.

4.3.1 Track and primary vertex reconstruction

As discussed in Section 4.1, precise information on particle tracks provided by the ID is an essential ingredient in any physics analysis, as it represents a key element of any particle-identification algorithm, such as those for photons or muons. However, numerous electrically charged particles enter the volume of the ID in a single event, generating numerous individual signals. Correctly combining these signals that originate from the same particle and reconstructing its trajectory are challenging tasks. The track reconstruction algorithm commonly used by the ATLAS collaboration during Run 2 and in this work, is described in detail in Refs. [185, 186].

The reconstructed tracks serve as input for the reconstruction of PVs. Excellent spatial resolution of the PV is crucial for identifying physics objects from different origins, e.g., from hard-scattering or pile-up interactions. The reconstruction of PVs is described in Ref. [187]. A PV is required to be associated with at least two tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$. The PV with the largest sum of squared p_{T} values of the associated tracks is considered the PV of the hard-scattering interaction.

4.3.2 Reconstruction of electrons and photons

As the detector signatures of electrons and photons are similar, both are reconstructed by the same algorithm. Both particle types initiate EM showers in the calorimeters, typically depositing most of their energy in the ECAL. While electrons typically generate one track in the ID, photons typically have either no tracks or are associated with electron tracks when they convert into an electron-positron pair ($\gamma \to ee$) when photons interact with the material of the ID, known as photon conversion.

The reconstruction algorithm of electrons and photons is described in detail in Ref. [3] and is summarized in the following. The algorithm starts by forming topological clusters, referred to as topo-clusters, by grouping calorimeter cells [188].[15] Topo-clusters are rejected if the energy deposited in the ECAL, referred to as EM energy, is below 400 MeV or if the fraction of their total energy from ECAL cells is below 0.5.

After the formation of the topo-clusters, an alternative track-reconstruction step is performed in addition to the standard track reconstruction. RoIs are defined by fixed-size clusters obtained by a sliding-window algorithm [189] that are compatible with the longitudinal and lateral profile of an EM shower. The standard track reconstruction employs the hypothesis that the associated particle is a pion. However, electrons lose a significantly larger fraction of their energy to bremsstrahlung, which renders the pion hypothesis inadequate for electron tracks. The alternative track-reconstruction step accounts for this and provides a list of track candidates in addition to the list provided by the standard track reconstruction. Track candidates loosely associated with an RoI are refitted employing a Gaussian sum filter algorithm [190]. Subsequently, these track candidates are matched to the topo-clusters from the

[15] The cells in the presampler and first ECAL layer are excluded for the formation of topo-clusters.

initial reconstruction step. In the case of multiple track candidates associated with a given topo-cluster, a ranking is performed based on the spatial distance as well as the number and position of hits in the silicon detectors.

Vertices in the ID compatible with a photon conversion are reconstructed using track candidates that are loosely matched to an RoI [1]. Either two-track conversion vertices (CVs) are reconstructed from tracks that are associated with particles having opposite electric charge or single-track vertices are reconstructed from tracks without hits in the innermost ID layer. Photons are reconstructed as one of six different reconstruction types depending on the type of the CV and the location of hits associated with the corresponding tracks. Photon candidates without an associated CV are classified as unconverted photons, and as converted photons otherwise. A summary of the different reconstruction types is given in Table 2. The associated tracks must be highly compatible with being from an electron as determined by the TRT. The probability of being an electron must be higher for tracks that have hits only in the TRT and for single-track CVs. A fit of the CV is performed for two-track CVs. The CVs are then matched to the topo-clusters. In the case of multiple candidates, the two-track CVs with silicon hits are preferred over other two-track types and single-track CVs, while smaller conversion radii are favored for CVs of the same type, i.e., CVs closer to the beam pipe are prioritized. Silicon hits refer to hits in the pixel or SCT detectors.

Table 2: Overview of the reconstruction types of photon candidates. The types differ by the number of tracks assigned to the CV and the location of the hits associated with these tracks. Photon candidates without an associated CV are referred to as unconverted photons and as converted photons otherwise. "Si hits" refer to hits in the pixel or SCT detectors, "TRT hits" refer to hits in the TRT.

Reconstruction type	Number of CV tracks	Location of associated track hits
Unconverted	No track	—
SingleSi	One track	One track with Si hits
SingleTRT	One track	One track only with TRT hits
DoubleSi	Two tracks	Both tracks with Si hits
DoubleTRT	Two tracks	Both tracks only with TRT hits
DoubleSiTRT	Two tracks	One track with Si hits, one track only with TRT hits

Subsequently, superclusters are constructed independently for electron and photon candidates. The first step is to identify topo-clusters that are suitable seed cluster candidates. In the second step, topo-clusters in the vicinity of the seed clusters are identified as satellite clusters likely caused by bremsstrahlung. Topo-clusters are tested as seed candidates in descending order of their EM energy. For electrons, topo-clusters must have a transverse energy (E_{T}) of $E_{\mathrm{T}} > 1$ GeV and a matched track with at least four silicon hits, while for photons, topo-clusters are required to fulfill $E_{\mathrm{T}} > 1.5$ GeV without applying track or CV requirements.[16] If a topo-cluster meets these requirements, clusters within $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ of the barycenter of the seed cluster are added as satellite clusters. In the case of electrons, topo-clusters within a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ are added for which the best-matching track is that of the seed cluster. For photons with an associated CV constructed from tracks containing silicon hits, topo-clusters are added as satellite clusters for which the best-matching track is associated with the CV.

An initial energy calibration and position correction of the superclusters are performed. After this step, tracks and CVs are matched to electron and photon superclusters, respectively, forming the final collection of electron and photon candidates. Since electron and photon candidates are built independently, an electron supercluster can emerge from the same seed cluster as a photon supercluster and vice versa. This ambiguity is resolved by a dedicated algorithm. However, some ambiguities are only resolved later through the overlap removal procedure described in Section 4.3.8.

[16] $E_{\mathrm{T}} = E/\cosh\eta$

The energies of these final electron and photon candidates are calibrated using the algorithm described in Ref. [191]. Variables that describe the shower shape and other discriminating variables used for electron or photon identification are calculated for the final calibrated candidates.

4.3.3 Reconstruction of muons

Muons are minimum ionizing particles at their typical energy values when arising from LHC pp collisions. As such, they lose only a small portion of their energy when traversing the detector material. Muons are expected to leave a track in the ID, cause only low-energy EM showers in the calorimeters, and leave a track in the MS. Depending on which sub-detector information is used and combined, different reconstruction types of muons can be defined [192, 193]. In this work, only muon candidates corresponding to the highest-quality reconstruction type, which combines track information from the ID and the MS, are used. Muons of this reconstruction type are referred to as "combined muons". ID tracks obtained by the standard reconstruction algorithm are combined with tracks that are obtained by independent track reconstruction in the MS to form the full muon trajectory of this reconstruction type. Most of the muons are reconstructed by using an outside-in algorithm, i.e., muon tracks are extrapolated from the MS to the ID and matched to an ID track. An inside-out algorithm is used in a complementary fashion.

The alignment of the ID and MS systems is corrected by dedicated procedures. However, residual misalignments arising from the weak modes, which are correlated geometric distortions that do not affect the χ^2 of a track candidate, introduce biases in the muon-momentum measurements. The sagitta bias refers to geometric distortions in the bending plane that affect the sagitta of positively and negatively charged particles in opposite directions. In-situ corrections are derived as a function of the muon η and ϕ by an iterative procedure using events with two reconstructed muon candidates from Z-boson decays and are applied to each muon candidate with $p_{\mathrm{T}} < 450\,\mathrm{GeV}$ to remove this charge-dependent bias in data. After correcting for the sagitta bias in data, the muon momentum is calibrated in simulation to match data by scale and resolution corrections derived from data enriched in events with two muon candidates from Z-boson and J/ψ-meson decays.

4.3.4 Isolation requirements

Isolation requirements are applied to electron, photon, and muon candidates [3, 192]. The main goal of these requirements is to suppress candidates that arise from hadron decays, which are typically surrounded by more activity than prompt contributions. A combination of track-based and calorimeter-based isolation requirements is used. The track-isolation variables are computed by summing the p_{T} of reconstructed tracks in a cone around the candidate while excluding tracks associated with the candidate. The cone has either a fixed size, corresponding to $\Delta R = X/100$, or a variable size, corresponding to $\Delta R = \min(\frac{10\,\mathrm{GeV}}{p_{\mathrm{T}}}, X/100)$. Track-isolation variables using a fixed cone size and a variable cone size are denoted by $p_{\mathrm{T}}^{\mathrm{cone}X}$ and $p_{\mathrm{T}}^{\mathrm{varcone}X}$, respectively. Reconstructed tracks are considered for track isolation when they fulfill $p_{\mathrm{T}} > 1\,\mathrm{GeV}$, $|\eta| < 2.5$, and a loose primary-vertex-association requirement corresponding to $|\Delta z_0 \sin\theta| < 3\,\mathrm{mm}$, where z_0 is the longitudinal impact parameter relative to the PV. Calorimeter-isolation variables $E_{\mathrm{T}}^{\mathrm{cone}X}$ are computed by summing positive energy contributions from topo-clusters within a fixed-size cone of $\Delta R = X/100$ around the candidate. The position of muon candidates is extrapolated to the calorimeter, while the cluster barycenter is used for electron and photon candidates. The energy associated with the candidates is subtracted, and corrections accounting for pile-up contributions are applied. The criteria applied in this work are given in Table 3.

Photon isolation is calibrated using events with two reconstructed lepton candidates and a reconstructed photon candidate from radiative Z-boson decays as well as single-photon events, i.e., events that contain at least one reconstructed photon candidate and are not required to meet any other criteria. Electron and muon isolation are calibrated in simulation using a tag-and-probe method using events with two reconstructed electron candidates and two reconstructed muon candidates from Z-boson decays,

Table 3: Isolation requirements applied to reconstructed electron, photon, and muon candidates.

Object	Calorimeter isolation	Track isolation
Electrons	$E_{\mathrm{T}}^{\mathrm{cone20}} < 0.06\, p_{\mathrm{T}}$	$p_{\mathrm{T}}^{\mathrm{varcone20}} < 0.06 p_{\mathrm{T}}$
Photons	$E_{\mathrm{T}}^{\mathrm{cone40}} < 0.022\, E_{\mathrm{T}} + 2.45\,\mathrm{GeV}$	$p_{\mathrm{T}}^{\mathrm{cone20}} < 0.05 E_{\mathrm{T}}$
Muons	$E_{\mathrm{T}}^{\mathrm{cone20}} < 0.15\, p_{\mathrm{T}}$	$p_{\mathrm{T}}^{\mathrm{varcone30}} < 0.04 p_{\mathrm{T}}$

respectively. Data-to-simulation scale factors depending on kinematic quantities of the respective candidate are applied to simulated events to correct for differences in the response to the isolation criteria between simulation and collected data.

4.3.5 Identification of electrons and muons

Electron identification Electron-identification criteria are applied to electron candidates and aim to efficiently identify those originating from prompt electrons, while offering excellent rejection of all other candidates, commonly referred to as fake electrons [3].

A likelihood (LH) discriminant is constructed using quantities extracted from the ID and the calorimeters that offer discriminative power against fake electrons. Three types of variables are used: properties of the electron track, variables describing the shower shape, and variables quantifying the spatial compatibility between the primary track and the reconstructed cluster. The p.d.f.s utilized in the LH are obtained from data enriched in events with two reconstructed electron candidates from Z-boson and J/ψ-meson decays for prompt electrons and from data enriched in dijet events for fake electrons. Dijet events are defined as events that contain at least two jets, where the most prominent underlying hard-scattering process is the production of two color-charged particles (quarks or gluons).

Quantities describing the shower development in the calorimeters are known to be mismodeled by the GEANT4 simulation. This mismodeling manifests as either an offset, i.e., the simulated distribution is shifted toward lower or higher values relative to the observed one, or as a discrepancy in the width of the distribution. Some variables are affected in both ways. To compensate for this mismodeling, the values in the simulation are shifted by an offset and/or corrected by a multiplicative factor. The corresponding values of the correction are obtained by χ^2 minimization of the differences between the distributions in simulation and collected data.

To pass the electron identification, candidates are required to fulfill a combination of requirements on the LH discriminant and on variables that are not used as input to the LH. Different working points (WPs) are constructed that differ in the selection efficiency of prompt electrons and the rejection power against fake electrons. The requirements of the WPs are optimized in bins of the cluster $|\eta|$ and E_{T}. Furthermore, electron candidates are required to pass primary-vertex-association requirements. Their transverse impact parameter d_0 relative to the beam spot must satisfy $|d_0/\sigma(d_0) < 5|$, where $\sigma(d_0)$ is the corresponding uncertainty of d_0, while z_0 must fulfill $|z_0\, \sin\theta| < 0.5\,\mathrm{mm}$.

In this work, the tight WP is employed, which offers an average efficiency of 80% and the best rejection power against fake electrons among the calibrated WPs. The WPs and the reconstruction efficiency for prompt electrons are calibrated by computing scale factors corresponding to data-to-simulation efficiency ratios by comparing events with two reconstructed electron candidates from the Z-boson and J/ψ-meson decays using a tag-and-probe method [6]. The scale factors are derived as a function of the $|\eta|$ and E_{T} of the probe electrons.

Muon identification To identify prompt muons originating from the hard-scattering process while rejecting muon candidates from other sources (fake muons), requirements on the number of hits in the different ID layers and MS stations, parameters indicating the quality of the track fit, and the compatibility of the individual measurements in the ID and MS are used [192]. Different WPs are constructed, which differ in the rejection power against fake muons and the selection efficiency of prompt muons.

In this work, the medium WP is utilized. Muon candidates must have at least two precision MS stations, defined as MDT or CSC stations associated with the muon candidate with at least three hits, except for muons with $|\eta| < 0.1$, for which this requirement is loosened to one precision MS station. Only one precision hole station is allowed, which is a precision MS station that is expected to have three hits but has fewer. Furthermore, the value of the ratio of the electric charge and momentum of the muon candidate, measured in the ID must be loosely compatible with that measured in the MS. Additionally, muon candidates must meet the primary-vertex-association requirements $|d_0/\sigma(d_0) < 3|$ and $|z_0 \sin\theta| < 0.5\,\text{mm}$. The muon-reconstruction and -identification efficiencies are calibrated using a tag-and-probe method with events with two muon candidates from Z-boson and J/ψ-meson decays, while the efficiency of passing the primary-vertex-association criteria is calibrated using only the former type of events. Data-to-simulation efficiency ratios are computed as a function of $|\eta|$ and p_{T}, and these ratios are applied to correct the expected contribution of prompt muons in simulation.

4.3.6 Reconstruction of jets

Jets are reconstructed using objects constructed by the particle-flow algorithm [194], summarized in the following. First, topo-clusters are built using the same algorithm as for electron reconstruction. Next, each track candidate fulfilling stringent quality requirements is matched to one of the topo-clusters, excluding tracks that are associated with electron or muon candidates. Since electrically charged particles often deposit their energy in more than one topo-cluster, nearby topo-clusters are added to the track-cluster combination. The expected energy deposited in the calorimeter by the electrically charged particle associated with a track candidate is subtracted from the matched topo-clusters using a cell-by-cell method. After removing potential remnants, topo-clusters are either fully removed or modified if they were matched to a track, or remain unchanged if no track could be matched to them. The particle-flow algorithm hence compiles a list of tracks and topo-clusters, referred to as particle-flow objects.

This list is used as input for jet reconstruction using the anti-k_t algorithm [195] with a radius parameter of $R = 0.4$, implemented in the FastJet software package [196, 197]. The topo-clusters are required to have positive energy, while the tracks must fulfill $|z_0 \sin\theta| < 2.0\,\text{mm}$. The η and ϕ coordinates of the topo-clusters are recomputed prior to executing the jet algorithm, so that their values correspond to the positions of the topo-clusters relative to the PV.

The energy of the jet is calibrated to correspond to the energy of the associated particle-level jet. This calibration is extracted from simulation, while in-situ corrections are applied to data to account for differences between simulation and data. A detailed description of the jet-energy-scale (JES) calibration applied to jets used in this work is given in Ref. [198]. Additionally, the jet energy resolution (JER) is measured in data and simulation. Differences in the JER between data and simulation are accounted for by smearing the energy distribution in simulation in phase spaces where the simulation offers better resolution compared to data.

Reconstructed jets are required to fulfill $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 4.5$ to be considered in the analysis. Furthermore, a tight requirement on the response to the Jet Vertex Tagger (JVT) algorithm is employed [199]. This algorithm, based on the k-nearest neighbor algorithm, identifies jets likely originating from pile-up interactions. A value is assigned to each jet indicating its compatibility with originating from the hard-scattering interaction. This value is required to be larger than 0.5 for jets with $p_{\text{T}} < 60\,\text{GeV}$ and $|\eta| < 2.4$. Differences between collected data and simulation in the JVT efficiency for hard-scatter jets are corrected via dedicated scale factors (SFs).

4.3.7 Identification of b jets

While no identification of individual hadrons is performed in this analysis, a b-tagging algorithm is employed to identify b jets. As b hadrons are characterized by a relatively long mean lifetime, typically $\approx 1.5\,\text{ps}$, they travel an average distance of $\approx 0.5\,\text{mm}$ before decaying. This distance is much larger than

the PV resolution of 30 µm longitudinally and 10–12 µm transversely, allowing the reconstruction of their decay vertex, referred to as SV.

The DL1r algorithm is a high-level b-tagging algorithm and is used in this work [200–202]. It is designed as a fully connected multi-layer feed-forward neural network that receives information from the low-level b-tagging algorithms, IP2D, IP3D, SV1 [203], JetFitter [204], and RNNIP [205] and combines these inputs with the jet p_{T} and $|\eta|$.

The DL1r algorithm is designed as a categorical classifier providing scores for the hypothesis that the jet is a b jet, c jet, and LF jet. These scores are combined into a single discriminant. Four WPs are defined by fixed requirements on the DL1r discriminant, providing inclusive efficiencies (the fraction of correctly identified b jets integrated over p_{T} and $|\eta|$) of 60%, 70%, 77%, and 85%. These four different WPs are calibrated in simulation to data using $t\bar{t}$ events and events from Z-boson production in association with jets (Z+jets). Additionally, the discriminant is divided into five pseudo-continuous bins, defined by the fixed selection cuts of the WPs. Each b jet is assigned to one of these bins. This pseudo-continuous calibration allows the utilization of multiple WPs within a single physics analysis.

4.3.8 Object overlap removal

Most of the aforementioned algorithms for different objects operate independently. Therefore, detector signatures might be assigned to multiple physics objects. This is accounted for by a sequential overlap removal, defined using requirements on the minimum distance between reconstructed physics objects, as described below. The angular distance (ΔR_y) used for the overlap removal is given by $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$.

1. Electron candidates sharing a track with muon candidates are removed.
2. Any jet candidate whose axis is within $\Delta R_y < 0.2$ of an electron candidate is discarded.
3. Any electron within $\Delta R_y < 0.4$ of the axis of any jet is removed.
4. Any jet with fewer than three associated tracks within $\Delta R_y < 0.2$ of a muon candidate is removed.
5. Muon candidates within $\Delta R_y < 0.4$ of the axis of any jet are removed.
6. Photons within $\Delta R_y < 0.4$ of any electron or muon candidate are removed.
7. Jets within $\Delta R_y < 0.4$ of any photon candidate are removed.

4.3.9 Missing transverse momentum

The missing transverse momentum is given by the negative p_{T} sum of all fully reconstructed and calibrated physics objects that pass the criteria introduced in this section, remain after the overlap removal, and are associated with the hard-scattering process. Its reconstruction and the corresponding performance is presented in Ref. [206]. An additional term accounts for low-energy (soft) contributions that arise from the hard-scattering process, but are not associated with any of these objects. This contribution is given by the negative p_{T} sum of tracks that are associated with the PV and is referred to as track soft term (TST). A dedicated overlap-removal procedure is applied to avoid double counting contributions to the missing transverse momentum. The energy scale and resolution are separately calibrated for events that contain no genuine neutrinos and that contain at least one genuine neutrino [206]. Genuine refers to prompt neutrinos that arise from the hard-scattering process or from the decay of a Z or W boson, and not from hadron decays.

5 Measurement of the photon-identification efficiency at high energies

Investigations of physics processes with photons in the final state represent important tests of the SM, for instance, to probe the perturbative and non-perturbative QCD regimes and the EW sector of the SM. Furthermore, such processes are used in the search for BSM phenomena. However, these investigations are particularly challenging using data from pp collisions due to the abundant production of jets. As a consequence, the majority of reconstructed photon candidates are not prompt photons but fake photons. The latter are mainly either genuine but non-prompt photons from the radiative decay of hadrons or diphoton decay of neutral mesons, or hadrons falsely reconstructed as photon candidates. While the relative contributions of underlying physics processes to the total fake-photon contribution are phase-space dependent, the $\pi^0 \to \gamma\gamma$ decay typically contributes the most (over 70%). This is primarily due to the abundant production of π^0 mesons, which are the lightest hadrons and decay almost instantly into a pair of photons in about 99% of all cases. The application of the photon-isolation requirements discussed in Section 4.3.4 rejects a large fraction of fake photons. However, this rejection power is typically not sufficiently high for performing precise studies of physics processes with prompt photons in the final state. Therefore, additional requirements are applied to the properties of reconstructed photon candidates to efficiently identify prompt photons and further suppress fake-photon contributions. These requirements are commonly referred to as photon identification (PID) requirements.

A brief overview of the PID requirements used by the ATLAS collaboration is given in Section 5.1. The efficiency with which prompt photons pass the PID requirements in collected data is measured using three different methods, each providing precise results within specific photon-energy ranges. The PID efficiency may differ for prompt photons in collected data and in simulated data, which can consequently lead to incorrect predictions of the prompt-photon yields and distributions. Therefore, the efficiency measurements are further used to correct the PID efficiency for prompt photons in simulation by applying data-to-simulation SFs, which are derived as functions of the transverse energy (E_{T}^{γ}) and of the absolute pseudorapidity ($|\eta^{\gamma}|$) of the photon candidate. The three methods are briefly discussed in Section 5.2. The measurement of the PID efficiency for high-energy prompt photons is performed using the "inclusive photon method", which is discussed in detail in Section 5.3. The results and improvements to the method presented in that section were developed by the author of this thesis. The results from all three methods using the full Run 2 dataset are published in Ref. [6] by the ATLAS collaboration.[17]

5.1 Photon identification with the ATLAS detector

The lateral and longitudinal developments of EM showers initiated by prompt photons differ from those initiated by fake photons. The fine granularity of the ECAL, especially of the first layer, allows these differences to be exploited to distinguish between prompt and fake photons. Furthermore, the fraction of energy of the photon candidate deposited in the HCAL (hadronic leakage) represents a discriminative feature. The shower development is described by shower shape variables (SSVs), constructed from information of the calorimeter cells of the cluster associated with a photon candidate. Table 4 provides an overview and description of the SSVs. The PID requirements consist of rectangular cuts on the SSV values, designed to reject the majority of fake photons while providing a high selection efficiency for prompt photons. These cuts are optimized in simulation using the TMVA package [207] to maximize the rejection power against fake photons for a given selection efficiency of prompt photons [3]. The SSVs are classified into two main categories: the narrow-strip variables (NSVs) and the relaxed-tight variables (RTVs). The NSVs comprise ΔE, E_{ratio}, f_{side}, and $w_{s\,3}$, which describe properties of the inner core of the cluster of the photon candidate and use only information from the strip layer of the ECAL. The RTVs comprise the remaining variables and are constructed using information from the coarser segments of the ECAL and from the hadronic leakage.

[17] The inclusive photon method is referred to as "matrix method" in Ref. [6]

Three different WPs are designed, which are referred to as the loose, medium, and tight PID WPs. The loose and medium PID WPs are defined using subsets of the SSVs, while the tight PID WP is defined using all SSVs. The loose PID WP makes use of the RTVs R_{had_1} or R_{had}, R_η, and w_{η_2}, while the NSV E_{ratio} is additionally used in the definition of the medium PID WP. The criteria of all PID WPs depend on $|\eta^\gamma|$ to account for differences in the detector geometry and material composition of the detector components that influence the shower development, and are optimized separately for converted and unconverted photon candidates. The criteria for the tight PID WP additionally depend on E_{T}^γ. The loose and medium PID WPs are typically used in photon triggers to reduce the rate of events that only contain fake photons early in the data-taking procedure, while retaining most of the events containing prompt photons. The ATLAS collaboration recommends applying the tight PID WP to reconstructed photon candidates in physics studies of final states with prompt photons.

Table 4: Description of the SSVs used for photon identification. This table is adapted from Ref. [3].

Name	Description
R_{had_1}	Ratio of E_{T}^γ in the first layer of the HCAL to E_{T} of the EM cluster (used over the ranges $\|\eta^\gamma\| < 0.8$ and $\|\eta^\gamma\| > 1.37$)
R_{had}	Ratio of E_{T}^γ in the HCAL to E_{T}^γ of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)
R_η	Ratio of the sum of the energies of the cells contained in a 3×7 $\eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 7×7 rectangle, both centered around the most energetic cell
w_{η_2}	Lateral shower width, $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells
R_ϕ	Ratio of the sum of the energies of the cells contained in a 3×3 $\eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 3×7 rectangle, both centered around the most energetic cell
$w_{s\,\text{tot}}$	Total lateral shower width, $\sqrt{(\Sigma E_i (i - i_{\text{max}})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of $\Delta\eta \approx 0.0625$ and i_{max} is the index of the highest-energy cell
$w_{s\,3}$	Lateral shower width, $\sqrt{(\Sigma E_i (i - i_{\text{max}})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of 3 cells around the highest-energy cell
f_{side}	Energy fraction outside core of three central cells, within seven cells
ΔE_s	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima
E_{ratio}	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies
f_1	Ratio of the energy measured in the first layer of the ECAL to the total energy of the EM cluster

The SSV distributions for prompt photons in simulated and collected data differ due to the aforementioned imperfect GEANT4 simulation of the shower development in the calorimeters. To account for these differences, univariate first-order corrections are applied to the SSVs of photon candidates in simulation. These corrections are referred to as "fudge factors". The value x_i of each SSV is corrected by $x_i + \delta_i$. The mean value of the fudge factor (δ_i) is the value that minimizes the χ^2 test statistic when comparing binned distributions for photon candidates obtained from collected data enriched in prompt photons to those for prompt photons obtained from simulated events [3]. The values of the fudge factors are derived as a function of E_{T}^γ and $|\eta^\gamma|$ separately for converted and unconverted photon candidates [2, 3]. The application of the fudge factors significantly improves the agreement between the SSV distributions of photon candidates in simulation and collected data. However, this procedure is not capable of completely resolving the disagreements, as it does not correct for potential mismodeling of the shape and correlations of the SSVs.

5.2 Overview of the photon-identification efficiency measurement methods

The residual mismodeling of the SSVs of simulated prompt photons limits the agreement between the PID efficiencies for prompt photons ($\varepsilon_{\text{pp}}^{\text{PID}}$) in simulation and collected data. Therefore, evaluating the PID efficiency in collected data is essential not only for validating the PID performance but also for correcting the $\varepsilon_{\text{pp}}^{\text{PID}}$ values in simulation. The ATLAS Collaboration conducts a comprehensive research program including studies of processes with prompt photons in the final state. Consequently, it is critical to ensure that the PID efficiency for prompt photons in simulation is consistent with that in collected data across a broad energy spectrum throughout the entire detector acceptance region for photons. However, a precise calibration of the PID efficiency requires a collected data sample that allows for $\varepsilon_{\text{pp}}^{\text{PID}}$ measurements with small statistical uncertainties. Moreover, the sample must consist predominantly of prompt photons to minimize the impact of systematic uncertainties related to the modeling of the fake-photon contributions, which are typically larger than those related to the modeling of prompt-photon contributions. Defining a single $\varepsilon_{\text{pp}}^{\text{PID}}$-measurement method that provides precise results across the low- and high-E_{T}^{γ} regimes presents a significant challenge as the E_{T}^{γ} spectrum depends on the underlying physics process by which the prompt photons are produced. As a consequence, the ATLAS collaboration employs three distinct measurement methods, each utilizing a different underlying physics process as a prompt-photon source, to achieve the best possible precision of the $\varepsilon_{\text{pp}}^{\text{PID}}$ measurement across various E_{T}^{γ} regimes.

All three methods measure $\varepsilon_{\text{pp}}^{\text{PID}}$ based on Equation (31). Here, N_{γ} and N_{γ}^{PID} denote the numbers of all photon candidates and of photon candidates that pass the PID requirements in collected data, respectively. To measure $\varepsilon_{\text{pp}}^{\text{PID}}$ in collected data, the prompt-photon purity P_{A} of all photon candidates and the prompt-photon purity P_{P} of all reconstructed photon candidates passing the PID requirements are determined by the three different measurement methods. The scale factor (SF^{PID}) is defined by Equation (32). Here, $\varepsilon_{\text{pp}}^{\text{PID, MC}}$ denotes the PID efficiency for prompt photons in simulation. This SF is used to correct the PID efficiency in simulation.

$$\varepsilon_{\text{pp}}^{\text{PID}} = \frac{P_{\text{P}}\, N_{\gamma}^{\text{PID}}}{P_{\text{A}}\, N_{\gamma}} \tag{31}$$

$$\text{SF}^{\text{PID}} = \frac{\varepsilon_{\text{pp}}^{\text{PID}}}{\varepsilon_{\text{pp}}^{\text{PID, MC}}} \tag{32}$$

The PID efficiencies are measured separately for converted and unconverted photon candidates in bins of E_{T}^{γ} and $|\eta^{\gamma}|$. The same binning is used for both reconstruction types. The edges are given by $E_{\text{T}}^{\gamma} = \{10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 80, 100, 125, 150, 175, 250, 350, \infty\}$ GeV and $|\eta^{\gamma}| = \{0, 0.6, 1.37, 1.52, 1.81, 2.37\}$ for the measurement of the efficiencies of the tight PID WP, while the transition region between the barrel and end-cap parts of the ECAL ($1.37 < |\eta^{\gamma}| < 1.52$) is excluded in the measurement. The measurement of the PID efficiencies for the tight PID WP is the main focus in the following. The three methods are briefly discussed below, highlighting the specific underlying physics processes each uses and the E_{T}^{γ} ranges where each provides the most precise measurements.

Radiative Z-boson decays This method uses events with a pair of reconstructed lepton candidates (electron or muon candidates) from Z-boson decays, for which a photon is radiated from one of the leptons. This topology is referred to as radiative Z-boson decay, denoted by $Z \to \ell\ell\gamma$. Consequently, the three-body invariant mass of the photon candidate and the dilepton pair, $m_{\ell\ell\gamma}$, is close to the pole mass of the Z boson (≈ 90 GeV). To reject contributions in which the photon is not radiated in the Z-boson decay, the invariant mass of the dilepton pair must satisfy $40\,\text{GeV} < m_{\ell\ell} < 83\,\text{GeV}$. The corresponding event sample is characterized by a high prompt-photon purity. The values of the prompt-photon purities are extracted by template fits to the observed $m_{\ell\ell\gamma}$ distribution (in the range $80\,\text{GeV} < m_{\ell\ell\gamma} < 100\,\text{GeV}$). This method offers the highest precision in the low-E_{T}^{γ} region and is conducted for $10\,\text{GeV} < E_{\text{T}}^{\gamma} < 100\,\text{GeV}$.

Electron extrapolation To obtain a sample with a high prompt-photon purity providing precise measurements in the medium E_{T}^{γ} region, a sample of events containing two reconstructed electron candidates from Z-boson decays is used. The method exploits the similarities between EM showers that are induced by electrons and photons. Smirnov transformations are applied to the SSV distributions of one of the electron candidates to align them with those of prompt photons. The parameters of this transformation are determined by comparing the SSV distributions of prompt photons and prompt electrons in simulation. After the application of the transformations, the values of the prompt-photon purities are extracted by performing template fits to the observed m_{ee} distributions. This method is conducted for $25\,\mathrm{GeV} < E_{\mathrm{T}}^{\gamma} < 250\,\mathrm{GeV}$.

Inclusive photon method In contrast to the other methods, this method does not rely on selecting events based on a specific physics process. Instead, it utilizes events that are only required to contain at least one photon candidate. A detailed description of the method is given in Section 5.3. This method provides precise results in the high-E_{T}^{γ} region, as a large fraction of photon candidates with low E_{T}^{γ} in the selected events are fake photons and this fraction decreases as E_{T}^{γ} increases. It is conducted for $E_{\mathrm{T}}^{\gamma} > 25\mathrm{GeV}$.

The individual measurements are combined for overlapping E_{T}^{γ} regions using the BLUE method [208, 209]. The results from the full Run 2 pp-collision dataset using the inclusive photon method, which were compiled by the author of this thesis, are discussed in the following section, along with improvements to the method relative to previous iterations of the measurement.

5.3 Photon-identification efficiency measurement with the inclusive photon method

The inclusive photon method was developed by the ATLAS collaboration for measuring the PID efficiency for prompt photons at high energies and has been used in all published PID-efficiency measurements using pp-collision data [1–3]. In the following, the method is described in detail, and improvements to the methodology that were not included in previous measurements are highlighted.

The events that are considered in the inclusive photon method are selected by using the single-photon triggers, which are part of the HLT menu [210]. These triggers filter events with detector signatures consistent with those of prompt photons that pass loose identification requirements. Further details are provided in Appendix A.1. The selected events are required to contain at least one photon candidate that must fulfill loose photon-isolation requirements[18] and the requirements of the loose PID WP. The p_{T}-leading photon candidate is used for the computation of quantities in this method.

The dominant contributions of prompt photons that pass the aforementioned selection criteria arise from the $qg \to q\gamma$ and $qq \to g\gamma$ hard-scattering processes, and from photons that are produced during quark fragmentation, where the most prominent underlying hard-scattering process is the dijet topology. An event sample simulating these contributions is used, in which the hard-scattering process, parton showering, hadronization, and the UE are modeled by using Pythia 8. This event sample is referred to as the γ+jet MC sample in the following.

To measure the prompt-photon purities P_{A} and P_{P} in collected data, the prompt-photon contribution needs to be disentangled from the contribution of fake photons. In the inclusive photon method, the track-isolation efficiency $\hat{\varepsilon}$, i.e., the fraction of photon candidates that pass track-isolation requirements, is used as a discriminating feature. As fake photons mainly arise from hadrons in jets, either through their decay or misreconstruction, they are expected to be accompanied by tracks of electrically charged particles more often than prompt photons. Consequently, the track-isolation efficiency for fake photons is expected to be significantly lower than that for prompt photons. A photon candidate passes the track-isolation requirement if it satisfies $p_{\mathrm{T}}^{\mathrm{cone40}} = 0\,\mathrm{GeV}$. Variables with the subscripts "pp" and "fp" are computed for prompt photons and fake photons, respectively, while variables without a subscript are computed across both photon types. The superscript "PID" denotes that the respective variable is

[18] The loose photon-isolation requirements are $E_{\mathrm{T}}^{\mathrm{cone20}} < 0.065\,E_{\mathrm{T}}$ and $p_{\mathrm{T}}^{\mathrm{cone20}} < 0.05 E_{\mathrm{T}}$.

computed for photon candidates that pass the requirements of the tight PID WP. For instance, $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ denotes the track-isolation efficiency for fake photons passing the requirements of the tight PID WP. The purities P_{A} and P_{P} can be expressed in terms of track-isolation efficiencies, as shown in Equations (33) and (34), respectively.

$$P_{\text{A}} = \frac{\hat{\varepsilon} - \hat{\varepsilon}_{\text{pp}}}{\hat{\varepsilon}_{\text{pp}} - \hat{\varepsilon}_{\text{fp}}} \tag{33}$$

$$P_{\text{P}} = \frac{\hat{\varepsilon}^{\text{PID}} - \hat{\varepsilon}_{\text{pp}}^{\text{PID}}}{\hat{\varepsilon}_{\text{pp}}^{\text{PID}} - \hat{\varepsilon}_{\text{fp}}^{\text{PID}}} \tag{34}$$

The values of $\hat{\varepsilon}$ and $\hat{\varepsilon}^{\text{PID}}$ are computed for photon candidates in the collected data. The values of $\hat{\varepsilon}_{\text{pp}}$ and $\hat{\varepsilon}_{\text{pp}}^{\text{PID}}$ are extracted from the γ+jet MC sample. The determination of $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ is the main challenge of the method. Since the number of fake photons decreases rapidly as E_{T}^{γ} increases, generating an MC sample with a sufficient number of fake photons in the high-E_{T}^{γ} region is computationally prohibitive. Furthermore, the accurate modeling of these contributions is challenging. Therefore, a data-driven approach is used to determine the values of $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, introduced in the following.

Two different selection regions[19] are defined by inverting requirements of the tight PID WP, which are consequently enriched in fake photons. In the first region, denoted by "non-tight region (NTR)", the photon candidate is required to fail at least one of the criteria of the tight PID WP on the SSV values. As only a small fraction of all fake photons satisfies the requirements of the tight PID WP, it is assumed that the value of $\hat{\varepsilon}_{\text{fp}}$ can be approximated by the track-isolation efficiency for fake photons in the NTR, as shown in Equation (35). The second region, denoted by "relaxed-tight region (RTR)", is defined by requiring the photon candidate to pass all criteria on the RTVs of the tight PID WP, while failing at least one of the criteria on the NSVs. As the NSVs describe quantities of the inner core of the photon cluster, they are only weakly correlated with the track-isolation variables. This allows for approximating the value of $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ by determining the track-isolation efficiency for fake photons via Equation (36) in the RTR.

$$\hat{\varepsilon}_{\text{fp}} \approx \hat{\varepsilon}_{\text{fp}}^{\text{NTR}} = \frac{R^{\text{NTR}} \cdot \hat{\varepsilon}^{\text{NTR}} - P_{\text{A}} \cdot f^{\text{NTR}} \cdot \hat{\varepsilon}_{\text{pp}}^{\text{NTR}}}{R^{\text{NTR}} - P_{\text{A}} \cdot f^{\text{NTR}}} \tag{35}$$

$$\hat{\varepsilon}_{\text{fp}}^{\text{PID}} \approx \hat{\varepsilon}_{\text{fp}}^{\text{RTR}} = \frac{R^{\text{RTR}} \cdot \hat{\varepsilon}^{\text{RTR}} - P_{\text{A}} \cdot f^{\text{RTR}} \cdot \hat{\varepsilon}_{\text{pp}}^{\text{RTR}}}{R^{\text{RTR}} - P_{\text{A}} \cdot f^{\text{RTR}}} \tag{36}$$

The quantities introduced in Equations (35) and (36) are defined as follows:

- R^{RTR}, R^{NTR}: fractions of photon candidates that fulfill the RTR and NTR criteria, respectively;
- $\hat{\varepsilon}^{\text{RTR}}$, $\hat{\varepsilon}^{\text{NTR}}$: track-isolation efficiencies for photon candidates in the RTR and in the NTR, respectively;
- f^{RTR}, f^{NTR}: fractions of prompt photons that fulfill the RTR and NTR criteria, respectively;
- $\hat{\varepsilon}_{\text{pp}}^{\text{RTR}}$, $\hat{\varepsilon}_{\text{pp}}^{\text{NTR}}$: track-isolation efficiencies for prompt photons in the RTR and in the NTR, respectively;

Quantities highlighted in red are determined using the γ+jet MC sample, while the others are computed for photon candidates in the collected data sample. Equations (35) and (36) are used to obtain a quadratic equation for $\hat{\varepsilon}_{\text{fp}}$. The ambiguity between the two solutions of the quadratic equation can be resolved by considering the properties of fake photons. First, the track-isolation efficiency is expected to be lower than that of prompt photons. Furthermore, the majority of fake photons arise from hadrons contained in

[19] A selection region refers to a subset of the selected data defined by specific criteria on the event topology and/or properties of the reconstructed physics objects.

jets and the energy of fake photons is positively correlated with the energy of the associated jet. As the particles associated with a jet are more collimated for higher jet energies, the track-isolation efficiency for fake photons decreases as a function of E_{T}^{γ}. The "+" solution for $\hat{\varepsilon}_{\mathrm{fp}}$ is unphysically large across all $|\eta^{\gamma}|$ and E_{T}^{γ} regions considered in the measurement.[20] Consequently, the "-" solution is taken as the value of $\hat{\varepsilon}_{\mathrm{fp}}$ for the subsequent computations. However, in the high-E_{T}^{γ} region it may happen that the resulting values of $\hat{\varepsilon}_{\mathrm{fp}}$ or $\hat{\varepsilon}_{\mathrm{fp}}^{\mathrm{PID}}$ become negative, which violates the self-consistency of the method. A novel handling of these cases is introduced and discussed in Section 5.3.1.

Figure 8 shows the track-isolation efficiencies in collected data, for prompt photons, and for fake photons as determined by Equations (35) and (36) for unconverted photon candidates with $|\eta^{\gamma}| < 0.6$ as a function of E_{T}^{γ}. While $\hat{\varepsilon}_{\mathrm{pp}}$ and $\hat{\varepsilon}_{\mathrm{pp}}^{\mathrm{PID}}$ are almost independent of E_{T}^{γ}, the values of $\hat{\varepsilon}_{\mathrm{fp}}$ and $\hat{\varepsilon}_{\mathrm{fp}}^{\mathrm{PID}}$ decrease monotonically with increasing E_{T}^{γ} values. The values of $\hat{\varepsilon}$ and $\hat{\varepsilon}^{\mathrm{PID}}$ converge toward $\hat{\varepsilon}_{\mathrm{pp}}$ and $\hat{\varepsilon}_{\mathrm{pp}}^{\mathrm{PID}}$, respectively, as the prompt-photon purity increases as a function of E_{T}^{γ}. The values of $\hat{\varepsilon}_{\mathrm{pp}}$ and $\hat{\varepsilon}_{\mathrm{pp}}^{\mathrm{PID}}$ are similar at a given E_{T}^{γ} value, which is also observed for the values of $\hat{\varepsilon}_{\mathrm{fp}}$ and $\hat{\varepsilon}_{\mathrm{fp}}^{\mathrm{PID}}$. The differences between the values of $\hat{\varepsilon}$ and $\hat{\varepsilon}^{\mathrm{PID}}$ are caused by the enhanced prompt-photon purity for candidates that pass the requirements of the tight PID WP, which leads to $\hat{\varepsilon}^{\mathrm{PID}}$ being significantly larger than $\hat{\varepsilon}$. The observation of the small impact of the tight PID WP requirements on the track-isolation efficiencies for prompt and fake photons supports the validity of the approximations in Equations (35) and (36). Appendix A.2 shows these distributions for the other $|\eta^{\gamma}|$ bins and for converted photon candidates. Qualitatively, the above properties are observed for all other distributions as well.

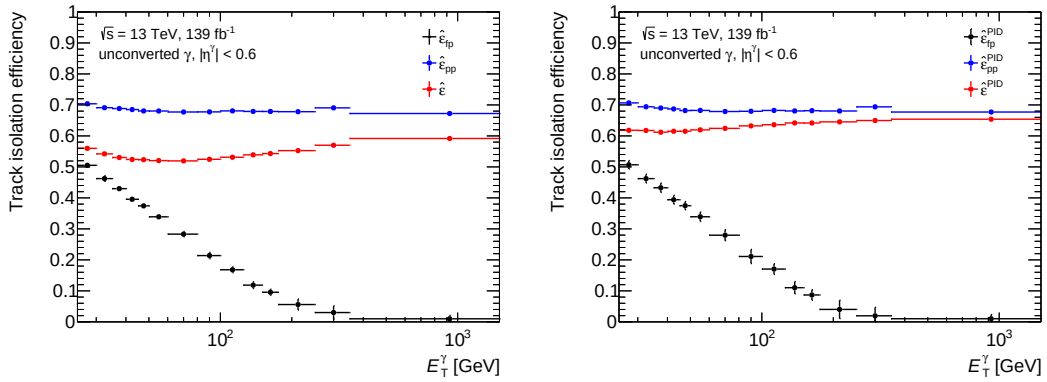


Figure 8: Track-isolation efficiencies for photon candidates in collected data, for prompt photons, and for fake photons determined by Equations (35) and (36) as a function of E_{T}^{γ} for unconverted photon candidates satisfying $|\eta^{\gamma}| < 0.6$.

The aim of the inclusive photon method is to measure the PID efficiency for all reconstructed photon candidates that are prompt photons. However, the selected photon candidates in the collected data sample represent only a subset, as they satisfy requirements of the single-photon triggers and the requirements of the loose PID WP. The PID efficiency for prompt photons in this subset is larger than the PID efficiency for all prompt photons. Therefore, a correction factor c^{presel} is introduced that accounts for this preselection bias, defined by Equation (37). The quantities of this equation represent the number of prompt photons predicted by the γ+jet MC sample that pass the criteria indicated by the superscript. For instance, $N_{\mathrm{pp}}^{\text{loose PID, trigger}}$ denotes the number of prompt photons that pass the requirements of the loose PID WP ("loose PID") and the single-photon triggers ("trigger"). The correction factor and the purities are then used to compute the PID efficiency for prompt photons via Equation (38), where the N^{X} quantities denote the number of photon candidates in the collected data that satisfy the requirements corresponding to the superscript.

[20] The solution of a quadratic equation of the form $x^2 + px + q$ is given by $-p/2 \pm \sqrt{p^2/4 - q}$, where the "+" solution refers to $-p/2 + \sqrt{p^2/4 - q}$ and the "-" solution to $-p/2 - \sqrt{p^2/4 - q}$

$$c^{\text{presel}} = \frac{N_{\text{pp}}^{\text{loose PID, trigger}}}{N_{\text{pp}}} \frac{N_{\text{pp}}^{\text{tight PID}}}{N_{\text{pp}}^{\text{tight PID, trigger}}} \tag{37}$$

$$\varepsilon_{\text{pp}}^{\text{PID}} = \frac{P_{\text{P}}\, N^{\text{tight PID, trigger}}}{P_{\text{A}}\, N^{\text{loose PID, trigger}}} \cdot c^{\text{presel}} \tag{38}$$

The distributions of c^{presel} as a function of E_{T}^{γ} are shown in Appendix A.3 for the four $|\eta^{\gamma}|$ regions and both photon reconstruction types. The distributions start at about 0.94 and 0.98 for $|\eta| < 1.81$ and at about 0.88 and 0.97 for $1.81 < |\eta^{\gamma}| < 2.37$ for converted and unconverted photon candidates, respectively. It is observed that all distributions converge toward 1 in the high-E_{T}^{γ} region.

The uncertainties assigned to the computed $\varepsilon_{\text{pp}}^{\text{PID}}$ values are presented in Section 5.3.2. The results for the P_{A}, P_{P}, $\varepsilon_{\text{pp}}^{\text{PID}}$, and SF^{PID} values are discussed in Section 5.3.3.

5.3.1 Handling of negative track-isolation efficiencies for fake photons

Table 5 lists all bins in which the computed values of $\hat{\varepsilon}_{\text{fp}}$ and/or $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ become negative. This issue only occurs for unconverted photon candidates in the high-E_{T}^{γ} region ($> 250\,\text{GeV}$), i.e., in regions where the prompt-photon purity is large ($> 85\%$) and the track-isolation efficiency for fake photons is small. The investigation of these cases did not uncover a common cause or pattern. A dedicated algorithm is introduced to handle these cases that restores the self-consistency of the method.

Table 5: Regions that are affected by negative values of at least one of the computed track-isolation efficiencies for fake photons. The photon reconstruction type as well as the affected $|\eta^{\gamma}|$ and E_{T}^{γ} ranges are listed.

Photon reconstruction type	$\lvert\eta^{\gamma}\rvert$	E_{T}^{γ} [GeV]
Unconverted	[0, 0.6]	[350, ∞]
Unconverted	[0.6, 1.37]	[350, ∞]
Unconverted	[1.52, 1.81]	[250, 350]
Unconverted	[1.52, 1.81]	[350, ∞]

The concept of the algorithm is to compile distributions of $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ constrained to positive values using pseudo-datasets (PDs). In the first step, 10,000 PDs are created from the collected data, and additional 10,000 PDs are created from the γ+jet MC sample. Each PD is constructed in the following way: a number (N_{PD}^{i}) is randomly sampled from a Poisson distribution whose mean is the nominal number of events in the respective sample. Subsequently, N_{PD}^{i} events are randomly selected from the respective sample using bootstrap sampling. The values of R^{RTR}, R^{NTR}, $\hat{\varepsilon}^{\text{RTR}}$, and $\hat{\varepsilon}^{\text{NTR}}$ are computed for each of the PDs obtained from the collected data, while the values of f^{RTR}, f^{NTR}, $\hat{\varepsilon}_{\text{pp}}^{\text{RTR}}$, and $\hat{\varepsilon}_{\text{pp}}^{\text{NTR}}$ are computed for each of the PDs obtained from the γ+jet MC sample. The corresponding covariance matrices and means of the variables are determined separately for the 10,000 sets of values from the collected data and for the 10,000 sets of values from the γ+jet MC sample.

In the second step, values of the eight variables are sampled from two multivariate normal distributions: one for the variables obtained from the collected data and one for the variables obtained from the γ+jet MC sample, using the means and covariance matrices obtained in the previous step. The values of $\hat{\varepsilon}_{\text{fp}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, and $\varepsilon_{\text{pp}}^{\text{PID}}$ are computed for each sampled set of values. The sampling is performed until 10,000 instances are compiled for which $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ are both positive. Figure 9 shows the resulting number of PDs (N_{PD}) as a function of $\hat{\varepsilon}_{\text{fp}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, and ε^{PID} for unconverted photon candidates satisfying $|\eta^{\gamma}| < 0.6$ and $350\,\text{GeV} < E_{\text{T}}^{\gamma}$. The nominal values of $\hat{\varepsilon}_{\text{fp}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, and $\varepsilon_{\text{pp}}^{\text{PID}}$ are set to the respective means of these distributions.

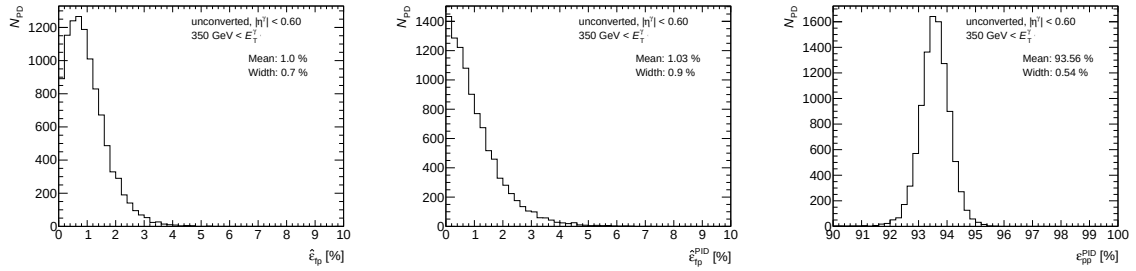


Figure 9: Number of PDs as a function of $\hat{\varepsilon}_{\mathrm{fp}}$, $\hat{\varepsilon}_{\mathrm{fp}}^{\mathrm{PID}}$, and $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ for unconverted photon candidates with $|\eta^{\gamma}| < 0.6$ and $350\,\mathrm{GeV} < E_{\mathrm{T}}^{\gamma}$ obtained by the handling of negative results for the track-isolation efficiencies of fake photons.

5.3.2 Uncertainties

Several types of uncertainties are assigned to the measured $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ values, introduced in the following.

Statistical uncertainties Statistical uncertainties arise from the limited sizes of the collected dataset and the γ+jet MC sample. However, the analytic propagation of these uncertainties is challenging due to the numerous quantities used in Equations (35) and (36), where some of the quantities are at least partially correlated with each other. To circumvent analytic propagation, the statistical uncertainties are estimated by constructing PDs from the collected dataset and the γ+jet MC sample in the same fashion as described in Section 5.3.1. For each of the PDs, $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ is computed and the standard deviation of the resulting distribution is taken as uncertainty. The derivation of the uncertainty is performed separately with the PDs from the γ+jet MC sample and collected data, which results in two independent types of uncertainties denoted by "MC stats" and "stats", respectively. As in previous iterations of the measurements, the number of 200 PDs is found to be sufficient as Gaussian-like shapes are observed for the sampled distributions of $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ across all bins, while the standard deviations do not change significantly using a larger number of PDs. Appendix A.4.1 shows examples of the sampled distributions.

Fudge-factor uncertainties The uncertainties of the individual fudge factors are divided into four groups. Each of the groups affects different subsets of SSVs. The determination of the uncertainties and the grouping is described in Ref. [211]. For each group, the $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ values are recomputed. The differences between the recomputed and the nominal values are considered as uncertainties. The sum in quadrature of these individual uncertainties is assigned as uncertainty ("FF variation"), assuming no correlation between the individual uncertainties.

Track-isolation requirement The computation of the $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ values is repeated using the alternative track-isolation requirement of $p_{\mathrm{T}}^{\mathrm{cone20}} = 0\,\mathrm{GeV}$, which represents a looser requirement compared to the nominal setup. The differences between the recomputed values and the nominal ones are assigned as uncertainties ("track isolation").

Correction factor Two different types of uncertainties are assigned to the values of c^{presel}. The first type is the statistical uncertainty associated with the limited size of the γ+jet MC sample used for computing c^{presel}. Furthermore, the efficiency of the single-photon triggers and the efficiency of the loose PID WP for prompt photons may be mismodeled in simulation, which affects c^{presel}. The impact of the mismodeling of the trigger efficiencies is expected to be small and is therefore neglected, as the trigger efficiency for prompt photons is close to 100% in the considered phase space. Thus, the second type of uncertainty accounts for potential mismodelings of the efficiency of the loose PID WP for prompt photons. Its derivation is discussed in Appendix A.4.2. The total uncertainty of c^{presel} is given by the sum in quadrature of these two types of uncertainties. The corresponding uncertainty ("correction") of the $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ value is given by the difference of the nominal $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ value and the recomputed $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ value using c^{presel} varied by 1σ of its total uncertainty.

Non-closure uncertainty The assumption that the track-isolation efficiencies $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ can be approximated by Equations (35) and (36), respectively, is tested by using a sample of fake photons from simulated dijet events. Like the γ+jet MC sample, this sample is generated using Pythia 8, but with contributions of prompt photons from the fragmentation process removed. However, only a tiny fraction of events contains a reconstructed photon candidate associated with a fake photon satisfying the selection criteria applied in the method. As a consequence, it is prohibitive to produce such a sample that offers a sufficiently large number of events with fake photons, particularly at high E_{T}^{γ}. Therefore, this non-closure test is performed inclusive in E_{T}^{γ}. It is used to assign an uncertainty on the $\varepsilon_{\text{pp}}^{\text{PID}}$ values, referred to as non-closure uncertainty. A detailed description of the computation of the non-closure uncertainty is provided in Appendix A.4.3.

Prompt-photon subtraction A mismodeling of the PID efficiency for prompt photons in simulation is associated with a mismodeling of the rate of prompt photons failing the criteria of the tight PID WP, i.e., a mismodeling of the fractions f^{NTR} and f^{RTR} used in Equations (35) and (36). This potential mismodeling has not been accounted for in previous measurements with the inclusive photon method. An iterative procedure is introduced to estimate the impact of this mismodeling on the measured $\varepsilon_{\text{pp}}^{\text{PID}}$ values and assign a systematic uncertainty to $\varepsilon_{\text{pp}}^{\text{PID}}$ ("Prompt γ subtraction"). This algorithm utilizes the anti-PID SF ($\text{SF}^{\text{anti-PID}}$) defined by Equation (39), which can be used to correct the rate of prompt photons failing the criteria of the tight PID WP in simulation.

$$\text{SF}^{\text{anti-PID}} = \frac{1 - \text{SF}^{\text{PID}}\,\varepsilon_{\text{pp}}^{\text{PID, MC}}}{1 - \varepsilon_{\text{pp}}^{\text{PID, MC}}} \tag{39}$$

A value of $\text{SF}^{\text{anti-PID}}$ different from 1 implies that the values of f^{NTR} and/or f^{RTR} are different for simulated and collected data. It is assumed that both variables are affected by the mismodeling. However, $\text{SF}^{\text{anti-PID}}$ can only be directly applied to f^{NTR}, as f^{RTR} represents only a subfraction of prompt photons failing the criteria of the tight PID WP. The value of f^{NTR} is scaled iteratively by $\text{SF}^{\text{anti-PID}}$, while scaling the value of f^{RTR} accordingly. The following steps are executed for each iteration i. The subscript i indicates the counting index of the iterations. The values of f_0^{NTR} and f_0^{RTR} are given by the uncorrected values of f^{NTR} and f^{RTR} as predicted by the γ+jet MC sample, and $\text{SF}_0^{\text{anti-PID}}$ equals 1.

1. $f_0^{\text{NRTR}} = f_0^{\text{NTR}} - f_0^{\text{RTR}}$
2. $f_i^{\text{NTR}} = \text{SF}_{i-1}^{\text{anti ID}} \cdot f_0^{\text{NTR}}$
3. $f_i^{\text{RTR}} = f_i^{\text{NTR}} - f_0^{\text{NRTR}}$
4. Computation of $\varepsilon_{\text{pp},i}^{\text{PID}}$, SF_i^{PID}, and $\text{SF}_i^{\text{anti-PID}}$ using f_i^{NTR} and f_i^{RTR}

The algorithm is stopped when the value of $\varepsilon_{\text{pp}}^{\text{PID}}$ converges and does not change significantly when performing additional iterations. The difference of the nominal $\varepsilon_{\text{pp}}^{\text{PID}}$ value and its value after the last iteration is assigned as systematic uncertainty to $\varepsilon_{\text{pp}}^{\text{PID}}$. The algorithm is executed for each bin considered in the measurement, except for the bins treated by the handling introduced in Section 5.3.1, as no convergence can be reached for these bins. Fifteen iterations are found to be sufficient, i.e., convergence is observed for all other bins. Figure 10 shows the value of $\varepsilon_{\text{pp}}^{\text{PID}}$ as a function of the number of performed iterations (N_{iter}) for converted photon candidates with $|\eta^{\gamma}| < 0.6$ for two E_{T}^{γ} ranges: $25\,\text{GeV} < E_{\text{T}}^{\gamma} < 30\,\text{GeV}$ and $30\,\text{GeV} < E_{\text{T}}^{\gamma} < 35\,\text{GeV}$. The value of $\varepsilon_{\text{pp}}^{\text{PID}}$ does not change significantly after the sixth iteration for both regions. The difference between the values of the first and last iterations is larger for $25\,\text{GeV} < E_{\text{T}}^{\gamma} < 30\,\text{GeV}$ than for $30\,\text{GeV} < E_{\text{T}}^{\gamma} < 35\,\text{GeV}$. Furthermore, the value of $\varepsilon_{\text{pp}}^{\text{PID}}$ converges in opposite directions for these two energy bins. No pattern is observed for the directions across all bins and the impact of the iterations is larger in the low-E_{T}^{γ} regime. The red lines ("1σ band") represent the uncertainty of $\varepsilon_{\text{pp}}^{\text{PID}}$ due to all aforementioned sources of uncertainties. The value of $\varepsilon_{\text{pp}}^{\text{PID}}$ converges to a value inside this band for all considered bins, i.e., the uncertainty due to the other sources

covers the potential bias due to the mismodeling of the prompt-photon contribution to data where the photon fails the PID requirements.

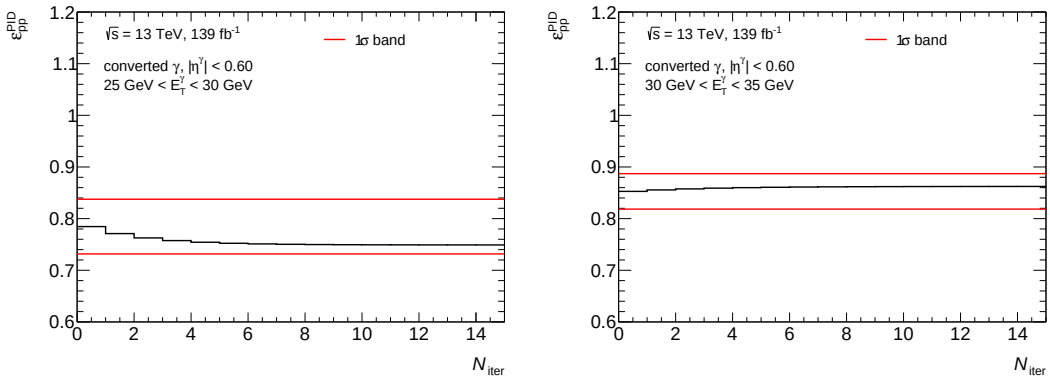


Figure 10: Values of ε_{pp}^{PID} as a function of the number of performed iterations (N_{iter}) in the case of converted photon candidates with $|\eta^{\gamma}| < 0.6$ with 25 GeV $< E_T^{\gamma} <$ 30 GeV (left) and 30 GeV $< E_T^{\gamma} <$ 35 GeV (right). The red lines represent the total uncertainty of ε_{pp}^{PID} due to all other sources of uncertainties.

Total uncertainty The different sources of uncertainty are assumed to be uncorrelated and to affect the values of ε_{pp}^{PID} symmetrically. The total uncertainty of ε_{pp}^{PID} is given by the sum in quadrature of the individual values. Breakdowns of the total uncertainty as a function of E_T^{γ} for all $|\eta^{\gamma}|$ regions for both photon reconstruction types are given in Appendix A.5.

5.3.3 Results

Figure 11 shows the values of P_A, P_P, ε_{pp}^{PID}, and SF^{PID} for converted and unconverted photon candidates with $|\eta^{\gamma}| < 0.6$ as a function of E_T^{γ}. The results for the other $|\eta^{\gamma}|$ bins are provided in Appendix A.5 and are considered in the following discussion. The difference between P_A and P_P is significantly larger for converted photon candidates than for unconverted photon candidates. Furthermore, the values of P_A and P_P are significantly larger for unconverted than for converted photon candidates and increase monotonically with E_T^{γ}. In contrast, the purities for converted photon candidates drop slightly in the high-E_T^{γ} region after an initial monotonic increase. The P_A distributions for converted photon candidates start at values between 22% and 28%, reach maximum values between 42% and 54% for 100 GeV $< E_T^{\gamma} <$ 250 GeV, and drop at high E_T^{γ} to values between 38% and 42%. The P_P distributions for converted photon candidates start at values between 40% and 47%, reach maximum values between 76% and 85% for 100 GeV $< E_T^{\gamma} <$ 250 GeV, and drop at high E_T^{γ} to values between 64% and 81%. The P_A distributions for unconverted photon candidates start at values between 27% and 36% and reach values between 87% and 88%. The P_P distributions for unconverted photon candidates start at values between 46% and 56% and reach values between 91% and 97% at high E_T^{γ}.

The $\varepsilon_{pp}^{PID, MC}$ values extracted from the γ+jet MC sample agree well with the respective ε_{pp}^{PID} values ("Data") within the uncertainty. Consequently, the SF^{PID} values are centered around 1 and agree well with 1 within their respective uncertainty across the whole E_T^{γ} range. The ε_{pp}^{PID} values for converted photon candidates are characterized by a stronger dependence on E_T^{γ}, as the increase of the ε_{pp}^{PID} values toward higher E_T^{γ} values is larger than for unconverted photon candidates. The values of ε_{pp}^{PID} are larger in the low-E_T^{γ} region but smaller in the high-E_T^{γ} region for unconverted photon candidates than for converted photon candidate. For both reconstruction types, the ε_{pp}^{PID} values drop in the highest E_T^{γ} bin. The ε_{pp}^{PID} values for converted photon candidates start at values between 77% and 80%, reach maximum values between 96% and 98% for 150 GeV $< E_T^{\gamma} <$ 250 GeV, and drop to values between 95% and 97% in the highest E_T^{γ} bin. The ε_{pp}^{PID} values for unconverted photon candidates start at values between 83% and 84%, reach maximum values between 92% and 95% in the 150 GeV $< E_T^{\gamma} <$ 250 GeV, and drop to values between 89% and 94% in the highest E_T^{γ} bin. The uncertainty on the SF^{PID} values is

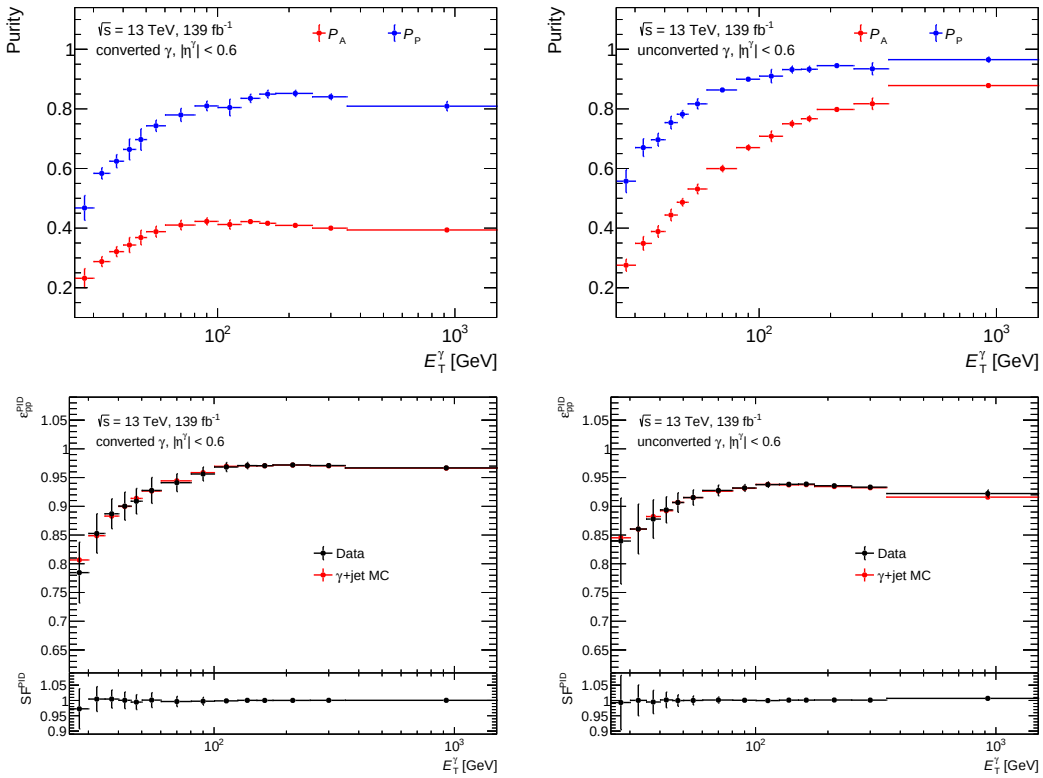


Figure 11: Prompt-photon purities P_{A} and P_{P} (upper row), and $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ and $\mathrm{SF}^{\mathrm{PID}}$ (lower row) as a function of E_{T}^{γ} for converted photon candidates (left) and unconverted photon candidates (right) with $|\eta^{\gamma}| < 0.6$. The uncertainty bars correspond to the total uncertainty.

large in the low-E_{T}^{γ} region (5% to 21%) and small in the high-E_{T}^{γ} region (below 1%). The dominant uncertainty source in the low-E_{T}^{γ} region varies with the $|\eta^{\gamma}|$ regions and the photon reconstruction type. The dominant types are the uncertainties due to the non-closure, track-isolation variation, fudge-factor variation, and prompt-photon subtraction.

The results of the inclusive photon method are combined with the results of the other two methods in the overlapping E_{T}^{γ} regions (25 GeV–250 GeV) using the BLUE method [208, 209]. The combination and the corresponding combined results are documented and discussed in Ref. [6]. While the respective weight of the result of the inclusive photon method is small in the combination for $E_{\mathrm{T}}^{\gamma} < 50\,\mathrm{GeV}$ relative to the other two measurements, the results are assigned the highest weight, i.e., the highest importance, in the $E_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$ region.

6 Simulation and examination of t-channel single-top-quark production in association with a photon

To investigate t-channel single-top-quark production in association with a photon using pp-collision data taken with the ATLAS detector, event samples simulating the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes are produced using a common setup. This setup is introduced in Section 6.1.

The kinematic properties of and differences between these samples are illuminated by using particle-level events. A reconstruction of physics objects is performed at particle level that mimics the reconstruction of physics objects from detector signals discussed in Section 4.3. The definition and reconstruction of these particle-level physics objects are discussed in Section 6.2.

The kinematic and topological properties of the particle-level objects in the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ samples are investigated and compared at LO in Section 6.3. In Section 6.4, the impact of NLO QCD contributions on these properties is examined. MC event generators are currently limited to LO precision for simulating the hard-scattering $t\,(\to \ell\nu b\gamma)\,q$ process, while the hard-scattering $tq\gamma$ process can be simulated incorporating NLO QCD contributions. An alternative approach for simulating the $t\,(\to \ell\nu b\gamma)\,q$ process is investigated that allows for the incorporation of NLO QCD contributions in the simulation. This approach is compared against the simulation of the $t\,(\to \ell\nu b\gamma)\,q$ process at LO.

This chapter concludes in Section 6.5 with a discussion of the main findings and the nominal approach adopted for simulating the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes in this work.

6.1 Common setup for the production of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ samples

A LO "inclusive" sample and a NLO $tq\gamma$ sample[21] are centrally[22] provided by the ATLAS collaboration, with their production following the full simulation chain discussed in Section 4.2. The LO inclusive sample simulates both the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ processes. The author of this thesis was one of the main developers of these samples. LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples are produced by the author without applying detector simulation to the particle-level events. Additionally, a LO sample simulating t-channel single-top-quark production ($pp \to tqb$), where photon radiation in the hard-scattering process is not simulated, is produced.

All the aforementioned samples are produced according to the SM by using a common setup. This means that the values of the SM parameters, the PDF sets, and the kinematic requirements on final-state particles of the hard-scattering process are the same for these samples. Furthermore, the same MC event generator and algorithms for modeling the PS, hadronization, and UE are utilized for the production of these samples. Additionally, a LO sample simulating t-channel single-top-quark production ($pp \to tqb$) is produced with the same setup, where photon-related requirements on the final-state of the hard-scattering process are dropped. This common setup is discussed below.

MadGraph5_aMC@NLO [162] is a commonly used MC event generator for simulating hard-scattering processes in LHC pp collisions. It offers the abilities to generate parton-level events for specific hard-scattering interactions and to interface these events to PS algorithms. Furthermore, it allows for fixed-order calculations, i.e., the automated computation of cross sections for a hard-scattering process at a fixed order of perturbation theory. These computations can be performed inclusively, i.e., without restricting the phase space of the final state, or for specific phase spaces defined by kinematic requirements on the final-state particles of the hard-scattering interaction. The phase spaces are referred

[21] The term "LO sample" refers to a sample that simulates the corresponding hard-scattering process with LO precision. The term "NLO" sample refers to a sample that incorporates NLO QCD contributions in the simulation of the corresponding process.

[22] The production of samples used for reporting results of a physics analysis is coordinated and performed centrally within the ATLAS collaboration.

to as parton-level fiducial phase spaces. The latter type of cross sections are referred to as fiducial cross sections. Kinematic properties of events from these samples are investigated in terms of differential fiducial cross sections.

The pp collisions are simulated under LHC Run 2 conditions, i.e., at a center-of-mass energy of 13 TeV. The NNPDF3.0NLO [212] PDF set, which uses the 4FS, is employed for sample production. The top-quark pole mass (m_t^{pole}) is set to 172.5 GeV and its width to $\Gamma_t = 1.32$ GeV, while the pole mass of the b quark is set to 4.7 GeV. The renormalization and factorization scales are set to $\mu_{\text{R}} = \mu_{\text{F}} = 0.5 \sum_i \sqrt{m_i^2 + p_{\text{T},i}^2}$, where the sum runs over all parton-level final-state particles. The pole mass of the Z boson is set to $m_Z^{\text{pole}} = 91.1876$ GeV. The strong coupling constant is given by $\alpha_{\text{s}}(\mu_0 = m_Z) = 0.1123$ (cf. Equation (5)). The coupling strength of the EM interaction and the Fermi constant are given by $\alpha_{\text{EM}} = 1/132.3489$ and $G_{\text{F}} = 1.16637 \cdot 10^{-5}\,\text{GeV}^{-2}$, respectively. These settings result in a W-boson pole mass of $m_W^{\text{pole}} = 80.399$ GeV. The width of the W boson is set to $\Gamma_W = 2.085$ GeV.[23]

The simulation is performed in a fiducial phase space defined by the following kinematic requirements on the parton-level final-state particles: Photons must fulfill $p_{\text{T}} > 10$ GeV and $|\eta| < 5.0$. The former requirement ensures that infrared divergences are avoided in the calculations of the matrix elements, while the latter is motivated by the $|\eta|$ coverage of the ATLAS detector. Photons are furthermore required to fulfill Frixione isolation using an isolation cone of $\delta_0 = 0.2$ [213].

The parton-level events obtained from MadGraph5_aMC@NLO are propagated to Pythia 8 [180] for the simulation of the PS, hadronization, and UE. Pythia 8 is used with the NNPDF2.3LO [212] PDF set for the PS and the A14 [214] set of tuned parameters is employed for the description of the UE. The decays of b and c hadrons are simulated by employing the EvtGen program [215]. Other decays are modeled internally by Pythia 8.

As discussed in Section 4.2, a weight is assigned to each event, ensuring that the ensemble of events in individual samples models the underlying hard-scattering process according to the chosen parameter settings. To study the impact of varying these settings, alternative weights are assigned to each event. These weights allow for reweighting, so that the ensemble of events reflects the underlying hard-scattering process according to the alternative settings. The alternative weights considered in the following studies correspond to independent variations of μ_{R} and μ_{F} by factors of 2 and 0.5, leading to nine scale variations in total, and to the uncertainties related to the chosen PDF set. These uncertainties are provided as 100 sampled replicas, so that each event is assigned 100 different weights.

6.2 Definition and reconstruction of physics objects at particle level

The particle-level events contain a record of initial- and final-state particles of the hard-scattering process (parton level), intermediate and particle-level final-state particles, particle decays, the PS evolution and hadronization processes. This allows for identifying the origin of the particle-level objects, e.g., determining if a particle-level photon stems from the decay of a hadron and connecting these objects to the parton-level final-state particles.

As described in Section 4.2, numerous particles emerge from a hard-scattering process. Those that are characterized by a mean lifetime larger than 30 ps are considered to be the particle-level final-state particles. Physics objects are formed using these particles as input and are then connected to the parton-level final-state particles. Furthermore, these objects are designed to resemble the physics objects reconstructed from the detector signals described in Section 4.3.

The framework Rivet 3 [216] is used for analyzing the particle-level events, which provides algorithms for the reconstruction of particle-level physics objects described below:

[23] These parameter settings are commonly used by the ATLAS collaboration, aligned with experimental values available at the time they were defined.

- **Photons**: Photons are final-state particles that are identified as prompt photons, i.e., that are not radiated from electrically charged hadrons and do not emerge from hadron decays.
- **Leptons**: Leptons are final-state particles identified as prompt electrons or prompt muons, i.e., they do not originate from a hadron decay or $\gamma \to \ell\ell$ process. Electrons from the decay of prompt tau leptons or prompt muons, and muons originating from the decay of prompt tau leptons are also included. Final-state photons within a distance of $\Delta R = 0.1$ to a lepton are added to the four-momentum of the lepton as the most likely origin of such photons is the radiation from the corresponding lepton. This process is referred to as "lepton dressing".
- **Neutrinos**: Neutrinos are prompt final-state neutrinos that do not originate from hadron or tau-lepton decays.
- **Jets**: Jets are clustered from the final-state particle-level particles using the anti-k_t algorithm with a radius parameter of $\Delta R = 0.4$ [195], as implemented in the FastJet software package [196, 197]. All neutrinos and prompt dressed muons are excluded from the clustering algorithm.
- ***b* jets**: The b hadrons that appeared as intermediate particles are treated as input to the jet clustering algorithm. However, the four-momenta of reconstructed jets are not altered by these b hadrons. A jet is identified as a b jet if at least one of these "ghost" b hadrons with $p_T > 5\,\text{GeV}$ is a constituent of the jet. This procedure is referred to as ghost-matching [217].

Since final-state particles are used as input for the formation of several of the aforementioned physics objects, requirements are applied to the particle-level objects to remove potential overlaps.

1. Jets within $\Delta R = 0.4$ of a lepton are removed.
2. Jets within $\Delta R = 0.4$ of an isolated photon are removed. A photon is considered isolated if the sum of transverse momenta of all charged particles within a $\Delta R = 0.3$ cone around it is less than 10% of its own transverse momentum.
3. Events in which a photon is separated by less than $\Delta R = 0.4$ from either a lepton or a remaining jet are removed.

Furthermore, requirements on the kinematic properties of physics objects and on the event topology are applied, where the latter are motivated by the expected final state shown in the LO Feynman diagrams for $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ in Figure 3.

- Jets are required to fulfill $p_T > 10\,\text{GeV}$ and $|\eta| < 4.9$.
- Each event is required to contain at least one lepton, at least one neutrino, at least one photon with $p_T^\gamma > 15\,\text{GeV}$ and $|\eta^\gamma| < 5.0$, and at least one b jet.

The reconstruction of these physics objects and the aforementioned requirements on these objects define the particle-level fiducial phase space in which the studies discussed in the remaining course of this chapter are performed.

6.3 Investigations of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes at leading order

The hard-scattering $tq\gamma$ process is simulated as a $2 \to 4$ process, i.e., $pp \to tqb\gamma$, where the top quark subsequently decays semileptonically ($t \to Wb \to \ell\nu b$). The hard-scattering $t\,(\to \ell\nu b\gamma)\,q$ process is simulated as a $2 \to 3$ process, i.e., $pp \to tqb$, where the top quark subsequently undergoes a radiative semileptonic decay ($t \to \ell\nu b\gamma$). In this section the particle-level events obtained from the LO $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ samples are examined to study the properties and differences of these two processes.

The $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes lead to the same final state and hence they interfere and are in principle inseparable. While the separate production of $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ samples neglects this

interference, the LO inclusive sample contains contributions from interference terms. In practice, a separate inspection of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ processes is reasonable if the interference effects are found to be small. The sum of kinematic distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples is compared to the respective distribution obtained from the LO inclusive sample to study the impact of the interference. Appendix B.1 shows the comparison of distributions of several kinematic variables. It is found that the distributions agree well within statistical uncertainties. This indicates that the impact of the interference terms is negligible, which justifies neglecting them in practice and allows for a separate inspection of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ processes.

The values for the particle-level fiducial cross section, which is denoted by σ_X^{fid} for process X, are shown in Equations (40) and (41) for $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$, respectively. The cross section corresponding to the $tq\gamma$ process is denoted by the subscript $t\,(\to \ell\nu b)\,q\gamma$ as the value of the cross section takes into account the semileptonic top-quark decay. The values of the cross sections are given by the acceptance fraction (A_X) of events that pass the requirements of the particle-level phase space in the respective sample multiplied by the respective fiducial cross section for the parton-level phase space defined by the requirements on the parton-level final-state particles used in the generation of the event samples ($\sigma_X^{\text{gen-fid}}$).

$$\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma} = A_{t(\to\ell\nu b)q\gamma} \cdot \sigma^{\text{gen-fid}}_{t(\to\ell\nu b)q\gamma} = 363^{+44}_{-38}\ \text{fb} \tag{40}$$

$$\sigma^{\text{fid}}_{t(\to\ell\nu b\gamma)q} = A_{t(\to\ell\nu b\gamma)q} \cdot \sigma^{\text{gen-fid}}_{t(\to\ell\nu b\gamma)q} = 148^{+17}_{-15}\ \text{fb} \tag{41}$$

The particle-level fiducial cross section for $tq\gamma$ is approximately 2.45 times larger than that for $t\,(\to \ell\nu b\gamma)\,q$. The quoted uncertainties arise from the variations of the chosen PDF set and the μ_{R} and μ_{F} values. The values of the acceptance fractions are $A_{t(\to\ell\nu b)q\gamma} = 0.4075^{+0.33\%}_{-0.38\%}$ and $A_{t(\to\ell\nu b\gamma)q} = 0.3577^{+0.31\%}_{-0.03\%}$. The difference in the acceptance fraction is mainly caused by the difference in the rate of passing the p_{T}^{γ} requirement, which is higher for the $tq\gamma$ process. The values of σ_X^{fid} and A_X for the different individual variations are provided in Appendix B.6.

For the following studies, all systematic variations are normalized to the nominal values of σ_X^{fid}, i.e., the sum of the weights of all events in the particle-level fiducial phase space is equal to the value of the nominal σ_X^{fid} for each variation. Therefore, these normalized variations affect only the shape of distributions of kinematic or topological quantities and not the overall cross section in the fiducial phase space. The uncertainty associated with the choice of the μ_{R} and μ_{F} values is given by the envelope of the nine scale variations, while the uncertainty associated with the chosen PDF set is given by the standard deviation of the 100 variations. The total uncertainty is given by the sum in quadrature of these two uncertainties and of the statistical uncertainty due to the limited number of events in the samples. The uncertainties are treated as uncorrelated between different samples. This treatment of uncertainties is also applied for the studies shown in Section 6.4.

The particle-level objects are connected to the parton-level final-state particles by applying a matching procedure. First, the parton-level neutrino, lepton, and b quark emerging from the top-quark decay are extracted from the event record. The particle-level neutrino (ν), lepton (ℓ), and b jet (b) closest in angular distance (ΔR) to the parton-level neutrino, lepton, and b quark emerging from the top-quark decay are selected for the following analysis. In addition, the p_{T}-leading particle-level photon (γ) is selected. Furthermore, the values of the top-quark mass (m_t) and the W-boson mass (m_W) are retrieved from the parton-level record for each event.[24]

If the parton-level final-state particles are correctly matched to their particle-level equivalents, the top-quark and W-boson masses can be reconstructed from the combination of the invariant mass of specific combinations of the particle-level particles. For $tq\gamma$ events, the invariant mass of the lepton and neutrino ($m_{\ell\nu}$) corresponds to m_W and the invariant mass of the lepton, neutrino, and b jet ($m_{\ell\nu b}$)

[24] The values of m_t and m_W are unique for each event and follow a Breit-Wigner distribution that is centered around the corresponding pole mass with the respective decay width.

to m_t. For $t\,(\to \ell\nu b\gamma)\,q$ events, the invariant mass of the lepton, neutrino, b jet, and photon ($m_{\ell\nu b\gamma}$) corresponds to m_t. The invariant mass of the lepton, neutrino, and photon ($m_{\ell\nu\gamma}$) corresponds to m_W when the photon is radiated in the W-boson decay and $m_{\ell\nu}$ corresponds to m_W when this is not the case. The former occurs in approximately 70% of all events in the LO $t\,(\to \ell\nu b\gamma)\,q$ sample.

Due to the effects from the PS and hadronization, the kinematic quantities of the W boson and top quark from the parton-level record, such as their invariant masses, are not expected to be perfectly reproduced, even with a perfect matching of parton-level and particle-level final-state particles. In contrast to the other top-quark decay products, the b quark undergoes parton showering and hadronization. The jet-clustering algorithm aims to cluster all particle-level final-state particles that are associated with the b quark that initiated the jet. However, it might either miss or neglect particle-level final-state particles (e.g., neutrinos arising from hadron decays) that are associated with the b quark, or might assign particles to the jet that are not associated with the b quark. As a result, the properties of the b jet often differ significantly from those of the b quark. In contrast, differences for leptons, neutrinos, and photons are mostly negligible. Furthermore, the b-jet identification via ghost matching is limited in accuracy and there are typically multiple b jets in a particle-level event. These factors can lead to the selection of a b jet that does not originate from the top-quark decay. Therefore, the association of the b quark with the b jet represents the main limitation of the matching procedure and of the reconstruction of kinematic top-quark properties in particle-level events.

The particle-level differential fiducial cross section as a function of a quantity Q of process X is denoted by $\frac{d\sigma_X^{\text{fid}}}{dQ}$. The upper panel of Figure 12 shows the particle-level differential fiducial cross section as a function of the relative difference between the energies of the parton-level b quark (E_b^{parton}) and the matched particle-level b jet (E_b) obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples. The bottom panel shows the ratio of the distribution from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample (dec^{LO}) and the sum of the distributions from the LO $tq\gamma$ (prod^{LO}) and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The distributions are centered around zero but are characterized by significant tails toward larger values, i.e., the particle-level b jet carries less energy than the parton-level b quark in a large fraction of $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ events. This is caused by missing contributions associated with the b quark in the jet clustering or by selecting a b jet that is not from the top-quark decay. Such an incorrectly selected jet typically carries less energy than one originating from a top-quark decay. The tails toward lower values are less pronounced and are mainly caused by the energy contributions of jet constituents wrongly clustered to the b jet. The difference in shape of the distributions from the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples is small with the tail toward lower values being slightly more pronounced for $t\,(\to \ell\nu b\gamma)\,q$.

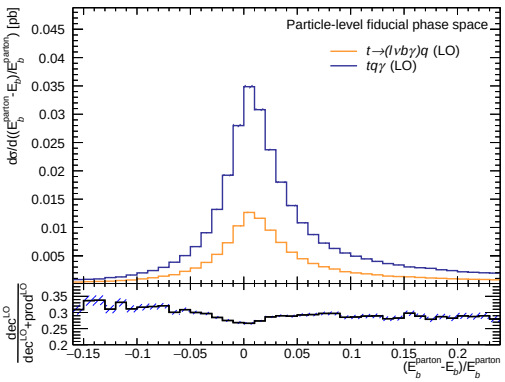


Figure 12: Particle-level differential fiducial cross section obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of the relative energy difference between the parton-level b quark from the top-quark decay and the matched particle-level b jet. The bottom panel shows the ratio of the $t\,(\to \ell\nu b\gamma)\,q$ distribution (dec^{LO}) and the sum of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ (prod^{LO}) distributions. The hashed bands correspond to the uncertainty of the predictions.

Figure 13 shows the particle-level differential fiducial cross section as a function of $m_{\ell\nu}$, $m_{\ell\nu\gamma}$, $m_{\ell\nu b}$,

and $m_{\ell\nu b\gamma}$ obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The red lines indicate the pole W-boson and top-quark masses.

For $tq\gamma$, the $m_{\ell\nu}$ and $m_{\ell\nu b}$ distributions peak at m_W^{pole} and m_t^{pole}, respectively, while peaks are absent for the other two distributions. As the photon does not originate from the top-quark decay, the $m_{\ell\nu\gamma}$ and $m_{\ell\nu b\gamma}$ distributions are shifted toward values significantly larger than m_W^{pole} and m_t^{pole}, respectively. It is also observed that the peak in the $m_{\ell\nu}$ distribution is sharp. In conclusion, the association of particle-level neutrinos and leptons with their parton-level counterparts is successful in the majority of the particle-level events. This is further discussed in Appendix B.2. The $m_{\ell\nu b}$ distribution is characterized by an enhanced tail toward values lower than m_t^{pole}, which is related to the aforementioned limitations in the matching of the b jet with the b quark.

For $t\,(\to \ell\nu b\gamma)\,q$, peaks are observed in the $m_{\ell\nu}$, $m_{\ell\nu\gamma}$, and $m_{\ell\nu b\gamma}$ distributions. Events in which the photon is radiated in the W-boson decay cause the peak at m_W^{pole} in the $m_{\ell\nu\gamma}$ distribution and the tail toward lower values ($m_{\ell\nu} < m_W^{\text{pole}}$) in the $m_{\ell\nu}$ distribution. Events in which the photon is not radiated in the W-boson decay cause the peak at m_W^{pole} in the $m_{\ell\nu}$ distribution and consequently the tail toward larger values ($m_{\ell\nu\gamma} > m_W^{\text{pole}}$) in the $m_{\ell\nu\gamma}$ distribution. As the photon is radiated in the top-quark decay, a peak is absent for the $m_{\ell\nu b}$ distribution. The enhanced tail toward lower values than m_t^{pole} in the $m_{\ell\nu b\gamma}$ distribution is also caused by the limitations in the matching of the b jet with the b quark.

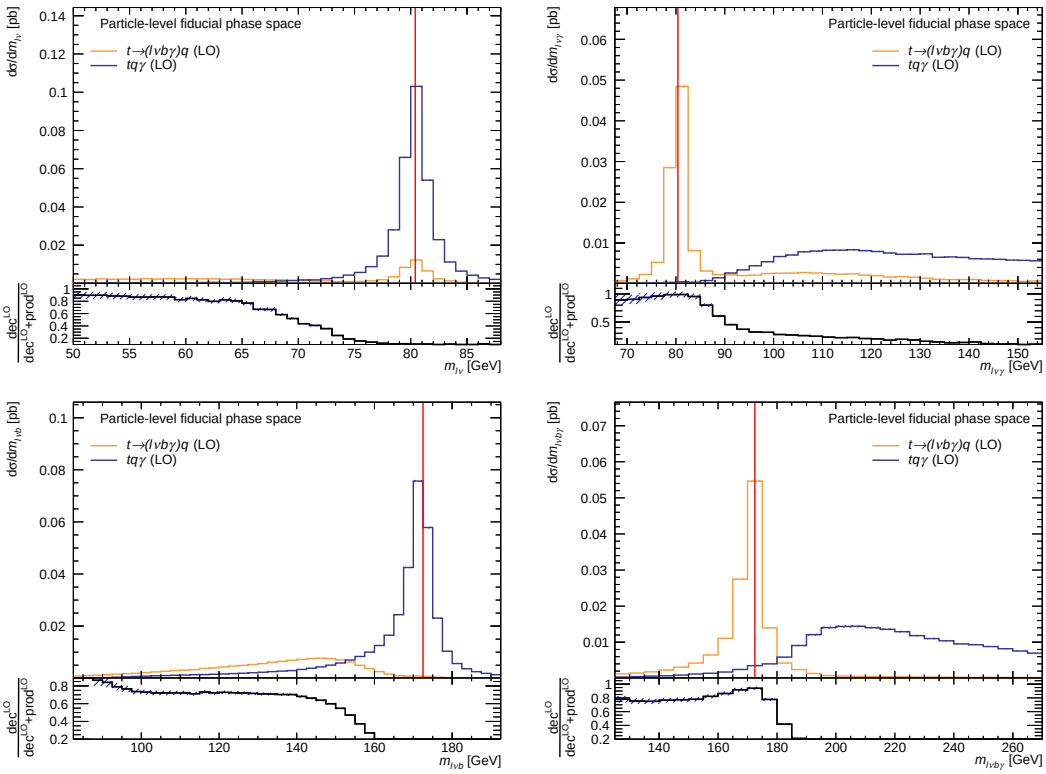


Figure 13: Particle-level differential fiducial cross sections obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of the invariant masses, $m_{\ell\nu}$ (top left), $m_{\ell\nu\gamma}$ (top right), $m_{\ell\nu b}$ (bottom left), and $m_{\ell\nu b\gamma}$ (bottom right). The red vertical lines correspond to the values of m_W^{pole} in the upper row and m_t^{pole} in the lower row. The bottom panel shows the ratio of the $t\,(\to \ell\nu b\gamma)\,q$ distribution (dec^{LO}) and the sum of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ (prod^{LO}) distributions. The hashed bands correspond to the uncertainty of the predictions.

Figure 14 shows the particle-level differential fiducial cross sections as a function of p_{T}^{γ}, p_{T}^{ℓ}, and the angular distances between the photon and the lepton ($\Delta R(\ell,\,\gamma)$), and between the photon and the b jet ($\Delta R(b,\,\gamma)$). Significant shape differences between the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ distributions are observed for all of these variables. The p_{T}^{γ} and p_{T}^{ℓ} distributions from $t\,(\to \ell\nu b\gamma)\,q$ are significantly

softer than the respective $tq\gamma$ distributions, i.e., the typical values of these quantities are much smaller for $t\,(\to \ell\nu b\gamma)\,q$ events. For $t\,(\to \ell\nu b\gamma)\,q$, the energy of the top quark is distributed among four decay products instead of three as in $tq\gamma$, causing the individual decay products to be carry less momentum on average. Furthermore, the angular distance between the lepton and photon is typically much smaller for $t\,(\to \ell\nu b\gamma)\,q$ as the photon is radiated in 70% of all events in the W-boson decay for $t\,(\to \ell\nu b\gamma)\,q$ and hence is in proximity to the lepton. This is also observed for the $\Delta R(b,\,\gamma)$ distributions, however, the difference in the region of small distances is less pronounced relative to that for the $\Delta R(\ell,\,\gamma)$ distributions because the selected photon emerges less often from the b jet than from the lepton in $t\,(\to \ell\nu b\gamma)\,q$ events. The majority of events are contained in the region of medium distances around $\Delta R(b,\,\gamma) = 3$ for $t\,(\to \ell\nu b\gamma)\,q$. As the photon is radiated in the top-quark decay for $t\,(\to \ell\nu b\gamma)\,q$, the rate of larger distances ($\Delta R(b,\,\gamma) >= 4$ and $\Delta R(\ell,\,\gamma) >= 4$) in $t\,(\to \ell\nu b\gamma)\,q$ events is tiny in contrast to $tq\gamma$ events.

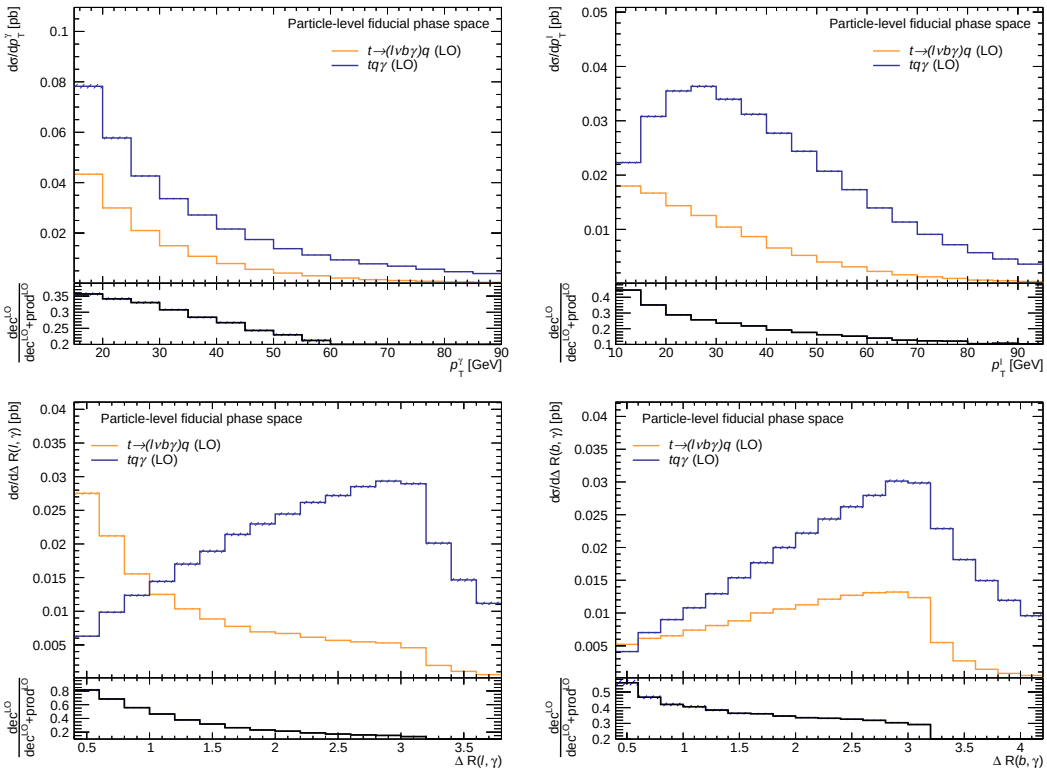


Figure 14: Particle-level differential fiducial cross sections obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of p_{T}^{γ} (upper left), p_{T}^{ℓ} (upper right), $\Delta R(\ell,\,\gamma)$ (lower left), and $\Delta R(b,\,\gamma)$ (lower right). The bottom panel shows the ratio of the $t\,(\to \ell\nu b\gamma)\,q$ distribution ($\mathrm{dec}^{\mathrm{LO}}$) and the sum of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ ($\mathrm{prod}^{\mathrm{LO}}$) distributions. The hashed bands correspond to the uncertainty of the predictions.

Figure 15 shows the particle-level differential fiducial cross section as a function of the p_{T} and η of the b jet, and of the p_{T}-leading jet with $|\eta| > 2.5$, which is referred to as forward jet (fj). The b jet in $t\,(\to \ell\nu b\gamma)\,q$ events is typically less central[25] and carries significantly less p_{T} than in $tq\gamma$ events. The p_{T} distribution of the forward jet is only mildly different for both processes. The forward jet carries slightly more p_{T} on average for $tq\gamma$ events. The difference in shape between the η distributions of the forward jet is only small, while the $|\eta|$ values of the forward jet is slightly larger on average for $t\,(\to \ell\nu b\gamma)\,q$. Figure 16 shows the particle-level differential fiducial cross section as a function of the jet multiplicity (N_{jets}) and b-jet multiplicity ($N_{b\mathrm{jets}}$). The $t\,(\to \ell\nu b\gamma)\,q$ events are characterized by slightly smaller N_{jets} and $N_{b\mathrm{jets}}$ values on average, i.e., the $t\,(\to \ell\nu b\gamma)\,q$ topology leads to the formation of fewer jets that pass the requirements of the fiducial phase space on average.

[25] The $|\eta|$ region close to $|\eta| = 0$ is referred to as central region.

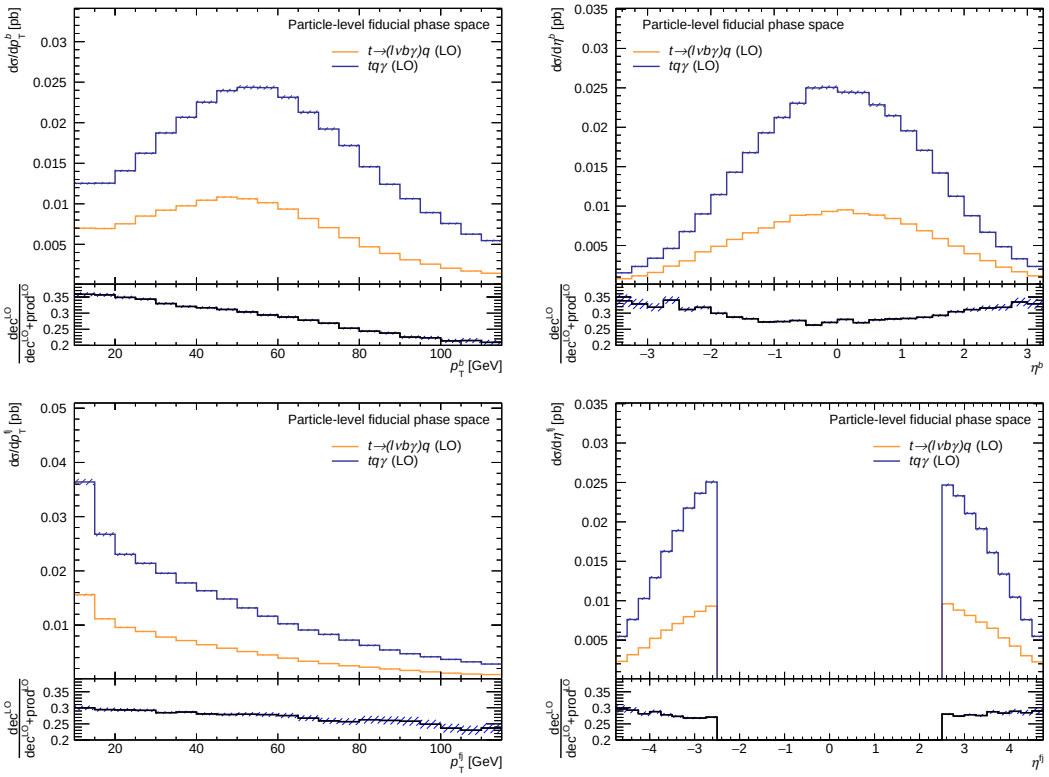


Figure 15: Particle-level differential fiducial cross sections obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of the p_{T} and η of the b jet as well as of the forward jet. The bottom panel shows the ratio of the $t\,(\to \ell\nu b\gamma)\,q$ distribution ($\mathrm{dec}^{\mathrm{LO}}$) and the sum of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ ($\mathrm{prod}^{\mathrm{LO}}$) distributions. The hashed bands correspond to the uncertainty of the predictions.

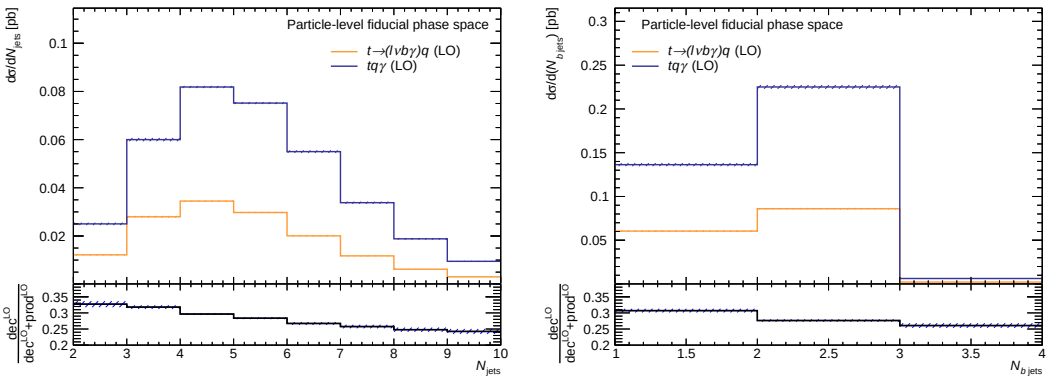


Figure 16: Particle-level differential fiducial cross sections obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of the number of jets (lower left), N_{jets}, and b jets (lower right), $N_{b\,\mathrm{jets}}$. The bottom panel shows the ratio of the $t\,(\to \ell\nu b\gamma)\,q$ distribution ($\mathrm{dec}^{\mathrm{LO}}$) and the sum of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ ($\mathrm{prod}^{\mathrm{LO}}$) distributions. The hashed bands correspond to the uncertainty of the predictions.

6.4 Exploring $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ at NLO QCD

Like for the LO $tq\gamma$ sample, MadGraph5_aMC@NLO is used for the simulation of the $pp \to tqb\gamma$ hard-scattering process for the NLO $tq\gamma$ sample, while NLO QCD contributions are included. The semileptonic top-quark decay is simulated using MadSpin [218, 219] at LO. The branching ratio of the semileptonic top-quark decay is set to $\mathcal{BR}(t \to \ell\nu b) = 32.46\%$.

Currently, MC event generators are not capable to include NLO QCD contributions in the simulation of the $t\,(\to \ell\nu b\gamma)\,q$ hard-scattering process. To study the impact of NLO QCD contributions, an alternative approach is employed. It is based on the use of a sample that simulates the hard-scattering process of t-channel single-top-quark production ($pp \to tqb$) including NLO QCD contributions. This sample, referred to as NLO tq sample, is centrally provided and commonly utilized by the ATLAS collaboration in physics analyses, for instance in Ref. [220]. In this sample, photon radiation is not simulated in the hard-scattering process and is instead modeled exclusively by the PS algorithm. This includes $t\,(\to \ell\nu b\gamma)\,q$ contributions.

The NLO tq sample was produced with the MC event generator Powheg Box v2 [221–224] using the 4FS and the same PDF set as in the production of the LO $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ samples. The nominal values of μ_R and μ_F are set to $4\sqrt{m_b^2 + p_{\mathrm{T},b}^2}$. The semileptonic decay of the top quark is simulated with LO accuracy using MadSpin [218, 219]. No requirements are applied to the kinematic properties of the parton-level final-state particles. Pythia 8 [180] is used for modeling the PS, hadronization, and UE. The decays of b and c hadrons are simulated by employing the EvtGen program [215]. Other decays are modeled internally by Pythia 8. Considering the semileptonic decay of the top quark, the cross section of this process is approximately 70.4 pb, assuming the cross section shown in Equation (19) for the inclusive $pp \to tqb$ process.

The alternative approach has two main limitations. First, the PS algorithms model photon radiation in the collinear limit and lack an accurate description of high-p_T photons radiated at large angles, in contrast to MC event generators simulating photon radiation in the hard-scattering process. Therefore, the kinematic properties of photons and photon-related topological quantities, such as angular distances between the photon and other objects, in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample might differ from those in the LO $t\,(\to \ell\nu b\gamma)\,q$ sample, where photon radiation is modeled as part of the hard-scattering process. Second, the NLO tq sample contains both $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ contributions.

To study the $t\,(\to \ell\nu b\gamma)\,q$ process, a procedure is developed to identify particle-level $t\,(\to \ell\nu b\gamma)\,q$ events in this sample. Particle-level events are tested for their compatibility with the hypothesis of a radiative top-quark decay. For a radiative top-quark decay, $m_{\ell\nu b\gamma}$ is expected to be closer to m_t than $m_{\ell\nu b}$, if the particle-level final-state particles are correctly matched to their parton-level counterparts. Therefore, events are required to fulfill $|m_{\ell\nu b\gamma} - m_t| < |m_{\ell\nu b} - m_t|$. However, due to the aforementioned limitations in the matching of the b jet and b quark, the hypothesis of a radiative W-boson decay is additionally tested. Events are required to fulfill $|m_{\ell\nu\gamma} - m_W| < |m_{\ell\nu} - m_W|$ to be considered compatible with this hypothesis. If an event passes the requirements of at least one of the two hypotheses, which is referred to as the "t or W" category, it is identified as a $t\,(\to \ell\nu b\gamma)\,q$ event. The ensemble of all events identified as $t\,(\to \ell\nu b\gamma)\,q$ events is referred to as NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. The $tq\gamma$ events in the NLO tq sample might also satisfy one of these two hypotheses. This represents an irreducible limitation of this identification procedure, caused mainly by the limitations in the matching of the b jet and b quark. The $t\,(\to \ell\nu b\gamma)\,q$ events can be classified into three orthogonal categories depending on which of the two hypotheses are passed. Events either pass only the hypothesis of a radiative W-boson decay ("only W") or only the hypothesis of a radiative top-quark decay ("only t") or both hypotheses ("t and W"). The "t or W" category represents the sum of these three categories. About 68% of all events from the NLO tq sample passing the requirements of the particle-level fiducial phase space are identified as $t\,(\to \ell\nu b\gamma)\,q$ events.

Table 6: Fraction of events in the particle-level fiducial phase space that are correctly identified as $t\,(\to \ell\nu b\gamma)\,q$ events in the LO $t\,(\to \ell\nu b\gamma)\,q$ sample ($\varepsilon^{t(\to\ell\nu b\gamma)q}$) and that are falsely identified as $t\,(\to \ell\nu b\gamma)\,q$ events in the LO $tq\gamma$ sample ($f^{tq\gamma,\,\mathrm{LO}}$) and in the NLO $tq\gamma$ sample ($f^{tq\gamma,\,\mathrm{NLO}}$). The events either pass both the hypotheses of a radiative top-quark and of a radiative W-boson decay ("t and W"), only the former hypothesis ("only t"), or only the latter hypothesis ("only W"). The last row ("t or W") shows the fraction of events that pass at least one of the hypotheses, i.e., represents the sum of the other categories. The purities P^{LO} and P^{NLO} of $t\,(\to \ell\nu b\gamma)\,q$ events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample obtained with Equation (42) are also listed.

$t\,(\to \ell\nu b\gamma)\,q$ category	$f^{tq\gamma,\,\mathrm{LO}}$	$f^{tq\gamma,\,\mathrm{NLO}}$	$\varepsilon^{t(\to\ell\nu b\gamma)q}$	P^{LO}	P^{NLO}
t and W	0.5%	0.5%	66.9%	98.6%	98.6%
only W	0.7%	0.1%	3.0%	62.8%	92.2%
only t	7.3%	9.8%	26.2%	59.8%	52.8%
t or W	8.5%	10.4%	96.1%	82.4%	79.6%

The accuracy of this identification approach is tested using the LO $t\,(\to \ell\nu b\gamma)\,q$, the LO $tq\gamma$, and the NLO $tq\gamma$ samples. Table 6 shows the fractions of events ($\varepsilon^{t(\to\ell\nu b\gamma)q}$) of all events from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample fulfilling the requirements of the particle-level fiducial phase space in the different categories. Furthermore, the fractions of all $tq\gamma$ events fulfilling the requirements of the particle-level fiducial phase space that are falsely assigned to the different $t\,(\to \ell\nu b\gamma)\,q$ categories obtained from the LO $tq\gamma$ ($f^{tq\gamma,\,\mathrm{LO}}$) and NLO $tq\gamma$ samples ($f^{tq\gamma,\,\mathrm{NLO}}$) are listed as well.

The events in the LO $t\,(\to \ell\nu b\gamma)\,q$ sample are correctly identified as $t\,(\to \ell\nu b\gamma)\,q$ in 96.1% of all cases. Most of the events ($\approx 70\%$) are assigned to the "only W" and "t and W" categories, i.e., the categories containing events that fulfill the hypothesis of a radiative W-boson decay. The overall fraction of misidentified $tq\gamma$ events is larger for the NLO $tq\gamma$ sample (10.4%) than for the LO $tq\gamma$ sample (8.5%). The majority ($> 85\%$) of the misidentified $tq\gamma$ events are contained in the "only t" category due to the shortcomings of the top-quark reconstruction caused by the limitations in the matching of the b jet and b quark.

The values of $\varepsilon^{t(\to\ell\nu b\gamma)q}$ and $f^{tq\gamma,\,(\mathrm{N})\mathrm{LO}}$ can be translated into an estimated purity of $t\,(\to \ell\nu b\gamma)\,q$ events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample ($P^{(\mathrm{N})\mathrm{LO}}$) by Equation (42), where r_σ^{LO} denotes the ratio of the particle-level fiducial cross sections as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples, $r_\sigma^{\mathrm{LO}} = \sigma^{\mathrm{fid}}_{t(\to\ell\nu b\gamma)q} / \sigma^{\mathrm{fid}}_{t(\to\ell\nu b)qy} \approx 0.403$. The computed values for $P^{(\mathrm{N})\mathrm{LO}}$ are also listed in Table 6 for the different $t\,(\to \ell\nu b\gamma)\,q$ categories.

$$P^{(\mathrm{N})\mathrm{LO}} = \frac{\varepsilon^{t(\to\ell\nu b\gamma)q} \cdot r_\sigma^{\mathrm{LO}}}{\varepsilon^{t(\to\ell\nu b\gamma)q} \cdot r_\sigma^{\mathrm{LO}} + f^{tq\gamma,\,(\mathrm{N})\mathrm{LO}}} \tag{42}$$

It should be noted that the values of $P^{(\mathrm{N})\mathrm{LO}}$ are rough estimations. They provide an assessment of the likely purity range. The difference in the modeling of the photon radiation between the LO $t\,(\to \ell\nu b\gamma)\,q$ (photon radiation modeled as part of the hard-scattering process) and the NLO $t\,(\to \ell\nu b\gamma)\,q$ (photon radiation modeled by the PS algorithm) samples and the NLO QCD contributions are likely to affect the $\varepsilon^{t(\to\ell\nu b\gamma)q}$ and r_σ values. However, these effects cannot be quantified.

To test the identification procedure, events from the LO inclusive sample are classified into two categories. Events that fulfill the requirements of the "t or W" category are identified as $t\,(\to \ell\nu b\gamma)\,q$ events and denoted by "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$". Events that fail these requirements are identified as $tq\gamma$ events and denoted by "Inclusive $\to tq\gamma$". Figure 17 shows the particle-level differential fiducial cross section as a function of the reconstructed W-boson mass (m_W^{reco}) and top-quark mass (m_t^{reco}). The distributions of the "Inclusive $\to tq\gamma$" and "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events are compared to those from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples, respectively. The "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events include correctly identified $t\,(\to \ell\nu b\gamma)\,q$ events and misidentified $tq\gamma$ events, and the "Inclusive $\to tq\gamma$" events include correctly identified $tq\gamma$ events and misidentified $t\,(\to \ell\nu b\gamma)\,q$ events. The reconstructed W-boson mass and top-quark mass for the LO $tq\gamma$ and the "Inclusive $\to tq\gamma$" events are given by $m_{\ell\nu}$

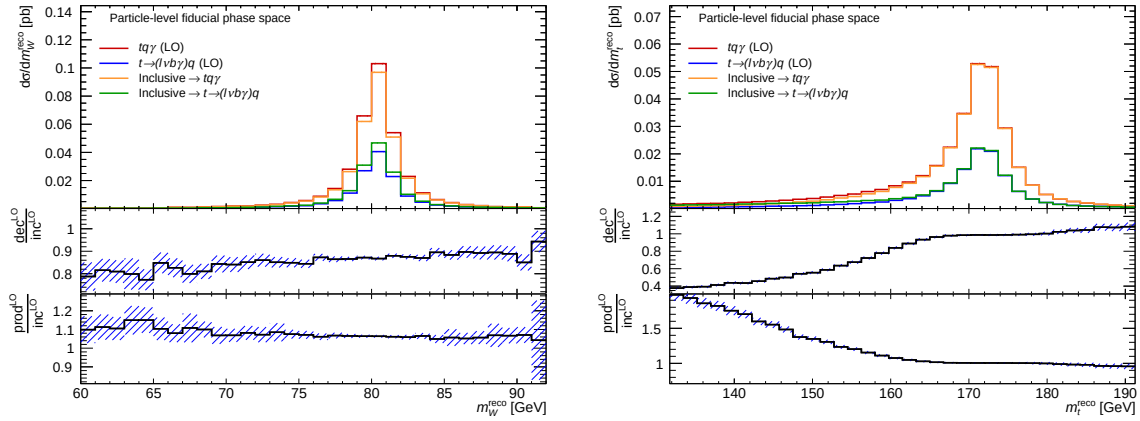


Figure 17: Particle-level differential fiducial cross section as a function of the reconstructed W-boson mass (left) and of the reconstructed top-quark mass (right) as obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples, and obtained from events that are identified as $t\,(\to \ell\nu b\gamma)\,q$ ("Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$") and $tq\gamma$ ("Inclusive $\to tq\gamma$") events in the inclusive sample using the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure. The hashed bands correspond to the uncertainty of the predictions.

and $m_{\ell\nu b}$, respectively. The reconstructed top-quark mass for the LO $t\,(\to \ell\nu b\gamma)\,q$ and the "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events is given by $m_{\ell\nu b\gamma}$, while the reconstructed W-boson mass is given by $m_{\ell\nu\gamma}$ for events in the "t and W" and "only W" categories, and by $m_{\ell\nu}$ for the remaining events. To investigate the impact of the limited accuracy of the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure on these distributions, the bottom panels show the ratio of the distributions obtained from LO $tq\gamma$ and the "Inclusive $\to tq\gamma$" events, and of the distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and the "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events. If the identification procedure had a perfect accuracy, the distributions obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples would agree well within statistical fluctuations with those obtained from the "Inclusive $\to tq\gamma$" and "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events, respectively.

The shapes of the distributions of the reconstructed W-boson mass obtained from the LO $tq\gamma$ sample and "Inclusive $\to tq\gamma$" events agree well within the uncertainties, as well as those obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample and "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events. The offset from 1 in the ratios is caused by the limited accuracy of the identification procedure. The shapes of the top-quark mass distributions agree well near the top-quark pole mass ($165\,\text{GeV} < m_t^{\text{reco}} < 180\,\text{GeV}$), a region that contains the majority of the contributions ($\approx 70\%$). These two observations imply a good overall performance of the identification approach.

Large shape differences are observed in the tails of the distributions of the reconstructed top-quark mass toward lower values. The misidentified $tq\gamma$ events have $m_{\ell\nu b}$ values that are typically lower than the top-quark pole mass. This causes the ratio of the distributions obtained from the LO $tq\gamma$ sample and the "Inclusive $\to tq\gamma$" events to increase toward low reconstructed top-quark masses. The reconstructed top-quark mass distribution ($m_{\ell\nu b\gamma}$) of these misidentified events, which are $\approx 22\%$ of the "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" events, is broader than that obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample, with the tail toward low reconstructed top-quark masses being more pronounced. This causes the ratio of the LO $t\,(\to \ell\nu b\gamma)\,q$ and the "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" distributions to decrease toward low reconstructed top-quark masses. Furthermore, the misidentified $t\,(\to \ell\nu b\gamma)\,q$ events have $m_{\ell\nu b\gamma}$ values that are larger than the top-quark pole mass. This causes the ratio of the LO $t\,(\to \ell\nu b\gamma)\,q$ and the "Inclusive $\to t\,(\to \ell\nu b\gamma)\,q$" distributions to increase toward high reconstructed top-quark masses. The misreconstructed $t\,(\to \ell\nu b\gamma)\,q$ events only contribute a small fraction in the "Inclusive $\to tq\gamma$" events ($\approx 2\%$) and do not impact the shape of the distributions of the "Inclusive $\to tq\gamma$" events significantly.

The values for the particle-level fiducial cross section as predicted by the NLO $tq\gamma$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples are $\sigma^{\text{fid, NLO}}_{t(\to\ell\nu b)q\gamma} = 466^{+29}_{-27}$ fb and $\sigma^{\text{fid, NLO}}_{t(\to\ell\nu b\gamma)q} = 238^{+11}_{-12}$ fb, respectively. The predicted value for

$tq\gamma$ is approximately 1.72 larger than the value for $t\,(\to \ell\nu b\gamma)\,q$ events in this particle-level fiducial phase space. The value of the LO-to-NLO k-factor, defined as the ratio of the NLO fiducial cross section and the respective LO fiducial cross section, is $k_{tq\gamma}^{\text{NLO}} \approx 1.29$ for $tq\gamma$ and $k_{t(\to\ell\nu b\gamma)q}^{\text{NLO}} \approx 1.61$ for $t\,(\to \ell\nu b\gamma)\,q$. The value of the acceptance fraction for NLO $tq\gamma$ events with $A_{t(\to\ell\nu b)q\gamma} = 0.4091^{+0.20\%}_{-0.28\%}$ is close to the value obtained from the LO $tq\gamma$ sample. In contrast, the fraction of events in the NLO tq sample that are contained in the NLO $t\,(\to \ell\nu b\gamma)\,q$ subsample and in the fiducial phase space is tiny with $A_{t(\to\ell\nu b\gamma)q} = (3.375 \cdot 10^{-3})^{+0.20\%}_{-0.03\%}$, as most of the events in the NLO tq sample do not contain any prompt photons satisfying the respective kinematic requirements of the particle-level phase space. The values of σ_X^{fid} and A_X are given in Table 23 of Appendix B.6 for the individual systematic variations.

To compare the shapes of the particle-level differential fiducial cross sections, the ratio $R\left(\frac{i}{j}\right)$ is computed for the samples i and j, which is given by the ratio of the normalized differential fiducial cross sections as a function of the quantity Q, as shown in Equation (43).

$$R\left(\frac{i}{j}\right) = \frac{1}{\sigma_i^{\text{fid}}} \cdot \frac{d\sigma_i^{\text{fid}}}{dQ} \cdot \left(\frac{1}{\sigma_j^{\text{fid}}} \cdot \frac{d\sigma_j^{\text{fid}}}{dQ}\right)^{-1} \tag{43}$$

6.4.1 Examination of NLO QCD contributions for $tq\gamma$

In the following, kinematic properties of the LO $tq\gamma$ and NLO $tq\gamma$ samples are compared to examine the impact of the NLO QCD contributions. Figure 18 shows the particle-level differential fiducial cross section as a function of $m_{\ell\nu}$ and $m_{\ell\nu b}$, i.e., as a function of the reconstructed W-boson and top-quark masses. There are no significant differences in the shapes of the $m_{\ell\nu}$ distributions and the distributions are centered around the pole W-boson mass. The $m_{\ell\nu b}$ distribution for NLO $tq\gamma$ has a significantly more pronounced tail toward values lower than m_t^{pole}. This is related to the kinematic properties of the matched b jet, which, on average, are in worse agreement with the parton-level b quark in the NLO $tq\gamma$ sample. This is illustrated in Figure 19, which shows the differential particle-level fiducial cross section as a function of the relative energy difference between the parton-level b quark and the matched particle-level b jet. The tail toward positive values is significantly larger for NLO $tq\gamma$ events, which corresponds to the b jet being less energetic than the parton-level b quark and causes the larger tail in the $m_{\ell\nu b}$ distribution relative to that in the respective distribution obtained from the LO $tq\gamma$ sample.

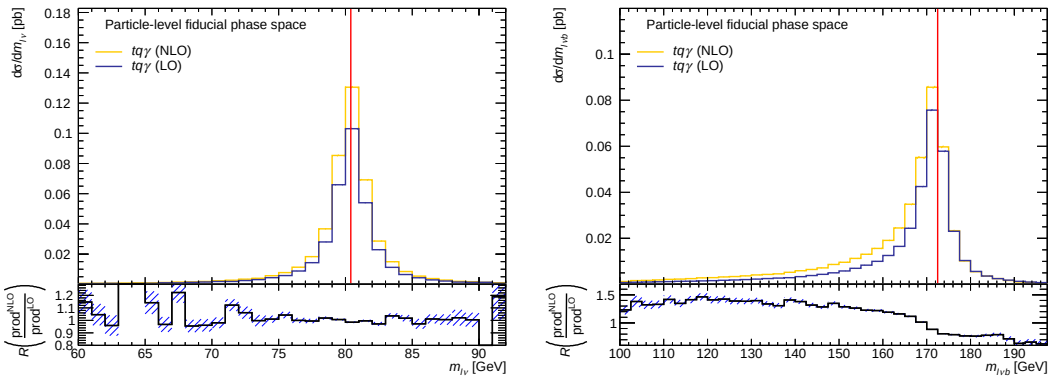


Figure 18: Particle-level differential fiducial cross sections obtained from the NLO $tq\gamma$ (prod^{NLO}) and LO $tq\gamma$ samples and the respective ratio as a function of the invariant masses, $m_{\ell\nu}$ and $m_{\ell\nu b}$. The hashed bands correspond to the uncertainty of the predictions.

Figure 20 shows the particle-level differential fiducial cross section as a function of p_{T}^{γ}, p_{T}^{ℓ}, $\Delta R(\ell,\gamma)$, and $\Delta R(\ell, b)$. The shapes of the p_{T}^{γ} distributions agree excellently with each other. While significant differences in shape are also absent for the p_{T}^{ℓ} distributions, the distributions start to differ for values of $> 70\,\text{GeV}$. However, the uncertainty in the ratio is comparatively large in this region, and hence a

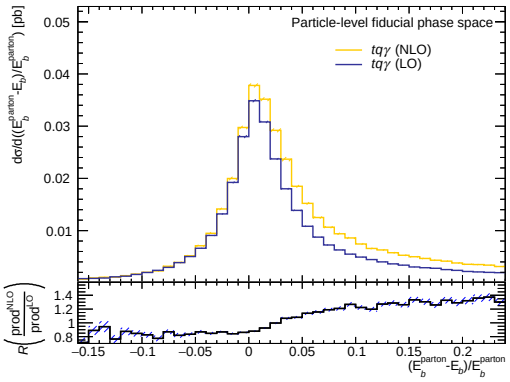


Figure 19: Particle-level differential fiducial cross sections obtained from the NLO $tq\gamma$ and LO $tq\gamma$ samples and the respective ratio as a function of the relative energy difference between the parton-level b quark from the top-quark decay and the matched particle-level b jet. The hashed bands correspond to the uncertainty of the predictions.

difference in shape cannot be concluded definitively. The shapes of the ΔR distributions are also in good agreement. Slight differences are only observed between the $\Delta R(b,\gamma)$ distributions, where the ratio increases with increasing $\Delta R(b,\gamma)$ values, i.e., the NLO $tq\gamma$ is characterized by slightly larger $\Delta R(b,\gamma)$ values on average.

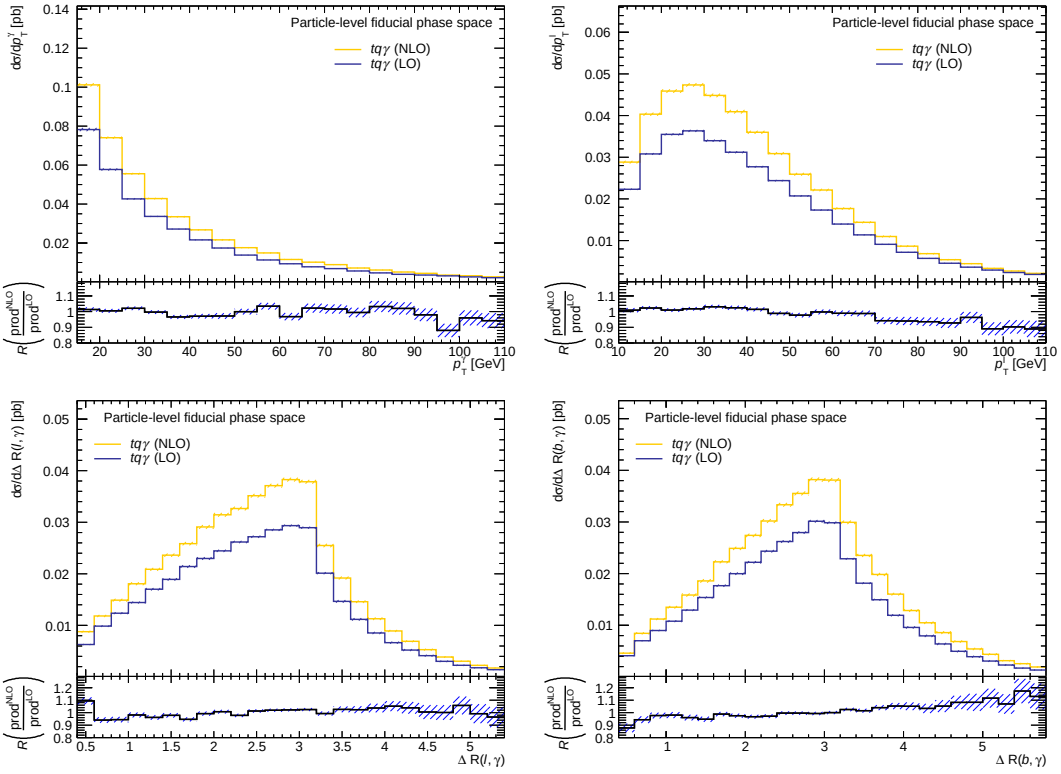


Figure 20: Particle-level differential fiducial cross sections obtained from the NLO $tq\gamma$ and LO $tq\gamma$ samples and the respective ratio as a function of p_{T}^{γ}, p_{T}^{ℓ}, $\Delta R(\ell,\gamma)$, and $\Delta R(\ell,b)$. The hashed bands correspond to the uncertainty of the predictions.

Figure 21 shows the particle-level differential fiducial cross section as a function of jet kinematics, specifically the transverse momentum of the highest-p_{T} jet (lj) that is not identified as a b jet and is not the p_{T}-leading forward jet ($p_{\mathrm{T}}^{\mathrm{lj}}$), of its pseudorapidity (η^{lj}), of $p_{\mathrm{T}}^{\mathrm{fj}}$, and η^{fj}. The $p_{\mathrm{T}}^{\mathrm{lj}}$ and $p_{\mathrm{T}}^{\mathrm{fj}}$ distributions are significantly softer for NLO $tq\gamma$ events, while the difference in shape is more pronounced for $p_{\mathrm{T}}^{\mathrm{lj}}$.

This is caused by the inclusion of additional QCD radiation in the modeling of the hard-scattering in the NLO $tq\gamma$ sample. Furthermore, the forward jet is characterized by larger $|\eta|$ values on average for NLO $tq\gamma$ events, while the shapes of the η^{lj} distributions agree well.

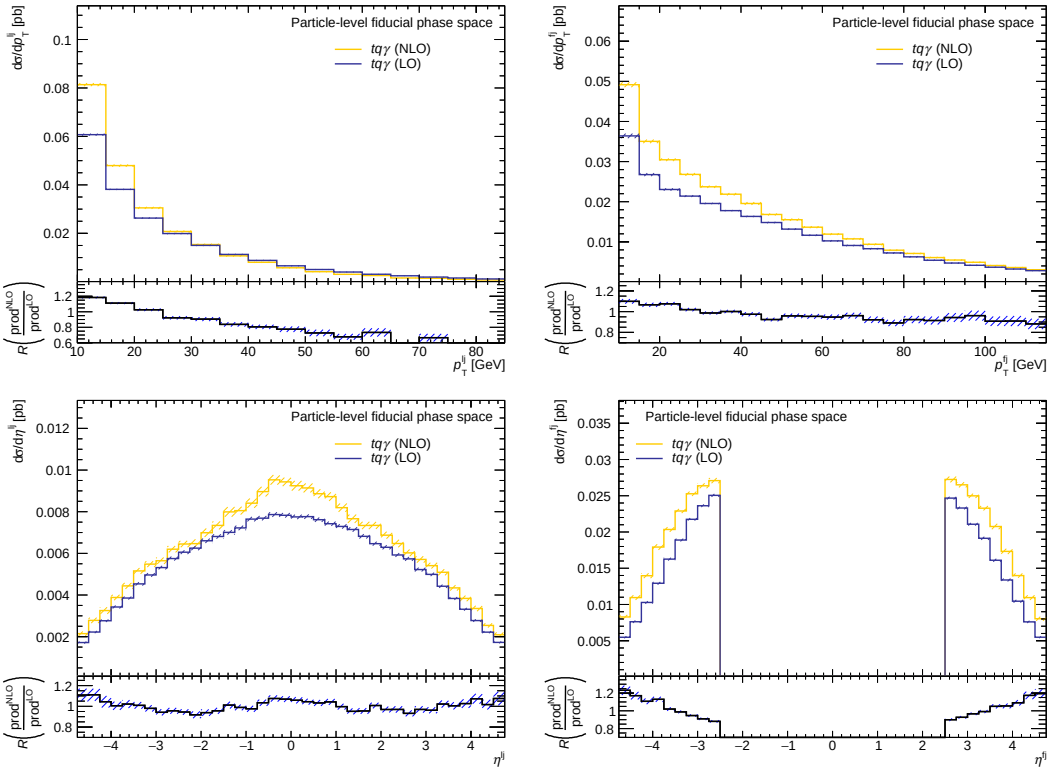


Figure 21: Particle-level differential fiducial cross sections as predicted by the NLO $tq\gamma$ and LO $tq\gamma$ samples and the respective ratio as function of $p_{\mathrm{T}}^{\mathrm{lj}}$, $p_{\mathrm{T}}^{\mathrm{fj}}$, η^{lj}, and η^{fj}. The hashed bands correspond to the uncertainty of the predictions.

Figure 22 shows the particle-level differential fiducial cross section as a function of the multiplicities of jets (N_{jets}) and b jets ($N_{b\ \mathrm{jets}}$). The LO $tq\gamma$ sample predicts a larger number of jets and b jets on average. The differences are most pronounced for $N_{\mathrm{jets}} = 1$ and $N_{b\ \mathrm{jets}} = 1$. The differences in shape observed for these distributions are much larger than for any other quantity, which were found to be mild at most. Therefore, it is concluded that including NLO QCD contributions in the simulation of the $tq\gamma$ process is essential for accurately modeling topological and kinematic quantities that involve jets.

6.4.2 Characteristics of the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample

In the following, kinematic properties of the events from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are studied and compared to those of events from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample. Topological and kinematic differences between those two samples may arise from the difference in the modeling of the photon radiation, from the NLO QCD contributions in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample, and from the contamination of $tq\gamma$ events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. The contamination of $tq\gamma$ events and the modeling of photon radiation solely via the PS algorithm are limitations of the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample, while neglecting NLO QCD contributions is a shortcoming of the LO $t\,(\to \ell\nu b\gamma)\,q$ sample. It is therefore vital to investigate these differences as choosing either the LO $t\,(\to \ell\nu b\gamma)\,q$ sample or the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample for the nominal modeling of the $t\,(\to \ell\nu b\gamma)\,q$ process in the analysis represents a trade-off.

The overall ratio of the fiducial cross sections predicted by the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample and the LO $t\,(\to \ell\nu b\gamma)\,q$ sample is 1.36 for the "t and W" category and 2.49 for the "only t" category, respectively. This large difference is expected to be mainly caused by the different levels of $tq\gamma$ contamination in each category.

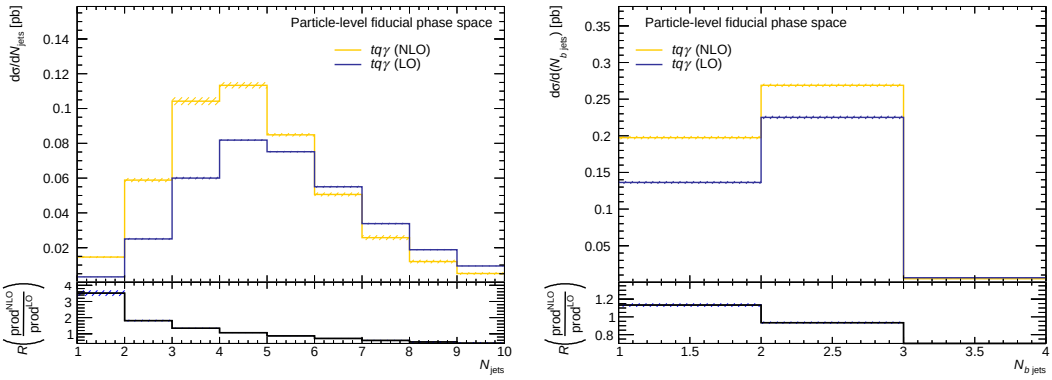


Figure 22: Particle-level differential fiducial cross sections obtained from the NLO $tq\gamma$ and LO $tq\gamma$ samples and the respective ratio as a function of the multiplicities of jets, N_{jets}, and b jets, $N_{b\text{ jets}}$. The hashed bands correspond to the uncertainty of the predictions.

This hypothesis can be tested with a back-of-the-envelope calculation. It is assumed that the NLO k-factor for genuine $t\,(\to \ell\nu b\gamma)\,q$ contributions is equal for both categories and that the $t\,(\to \ell\nu b\gamma)\,q$ purity in the "t and W" category is close to the respective P^{LO} and P^{NLO} values ($\approx 98\%$) listed in Table 6. This results in an estimated purity of $t\,(\to \ell\nu b\gamma)\,q$ events in the "only t" category of $\approx 0.98 \cdot 1.36/2.49 = 54\%$, which is in excellent agreement with the range 52.8%–59.8% estimated in Table 6. This supports the hypothesis that the large difference in the k-factor values is mainly caused by the large $tq\gamma$ contamination in the "only t" category.

In the following, an isolated investigation of the differences in the shape of distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples that may be caused by the difference in the modeling of the photon radiation and by the $tq\gamma$ contamination in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is performed. To isolate these effects, photon-related quantities are examined for the $t\,(\to \ell\nu b\gamma)\,q$ categories that are not expected to be significantly affected by NLO QCD contributions, as discussed in Appendix B.3. The shapes of the distributions obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are compared to those from the LO $t\,(\to \ell\nu b\gamma)\,q$ and the LO inclusive samples. The former comparison allows for assessing how large the combined impact of both limitations in the NLO $t\,(\to \ell\nu b\gamma)\,q$ are. Like the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample, the LO inclusive sample includes $tq\gamma$ contributions. Therefore, differences in shape between the NLO $t\,(\to \ell\nu b\gamma)\,q$ and the LO inclusive samples are expected to arise mainly from the difference in the photon-radiation modeling. It should be noted that the choice for modeling the photon radiation affects both the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ contributions in the LO inclusive and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples. This may contribute to a potential difference in the shape of the distributions from these samples arising from the $tq\gamma$ contamination.

Figure 23 shows the particle-level differential fiducial cross section as a function of p_{T}^{γ}, $\Delta R(\ell\nu\gamma,\,\gamma)$, $\Delta R(\ell,\,\gamma)$, and $\Delta R(b,\,\gamma)$ for the "t and W" category. The impact of the $tq\gamma$ contamination on the shape of the distributions in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is expected to be small for this category as the purity of $t\,(\to \ell\nu b\gamma)\,q$ events is estimated to be close to 100% (cf. Table 6). The difference in shape between the distributions from the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO inclusive samples is negligible. A good agreement between the shapes of the p_{T}^{γ} distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples is observed for $p_{\text{T}}^{\gamma} < 70\,\text{GeV}$. The fraction of the fiducial cross section for $p_{\text{T}}^{\gamma} > 70\,\text{GeV}$ predicted by the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is $\approx 20\%$ lower than that predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ sample. However, the cross section for $p_{\text{T}}^{\gamma} > 70\,\text{GeV}$ is small relative to the cross section for the whole fiducial phase space, so that the overall impact of this discrepancy is insignificant.

The shapes of the $\Delta R(\ell\nu\gamma,\,\gamma)$ distributions agree well with each other. This quantity represents the angular distance between the photon and the reconstructed W boson for this $t\,(\to \ell\nu b\gamma)\,q$ category. Significant shape differences between the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ events are also absent

for $\Delta R(\ell, \gamma)$ and $\Delta R(b, \gamma)$. However, the agreement between the distributions obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples is better than that between the distributions obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO inclusive samples for $\Delta R(\ell\nu\gamma, \gamma) > 3.2$ and $\Delta R(b, \gamma) > 3.4$. These regions are expected to have a larger relative contamination from $tq\gamma$ events and hence the differences are likely caused by a difference between the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO inclusive sample in the modeling of the $tq\gamma$ contamination. However, the cross section is also small for these regions, so that the overall impact of this difference is small as well.

These observations lead to the conclusion that the impact of the difference in the modeling of the photon radiation is small on kinematic and topological quantities of the $t\,(\to \ell\nu b\gamma)\,q$ process for events with a radiative W-boson decay. The shapes of distributions of such quantities predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples agree well in the bulk of the distributions. Mild differences are only observed in the tails of some distributions that have an insignificant overall impact.

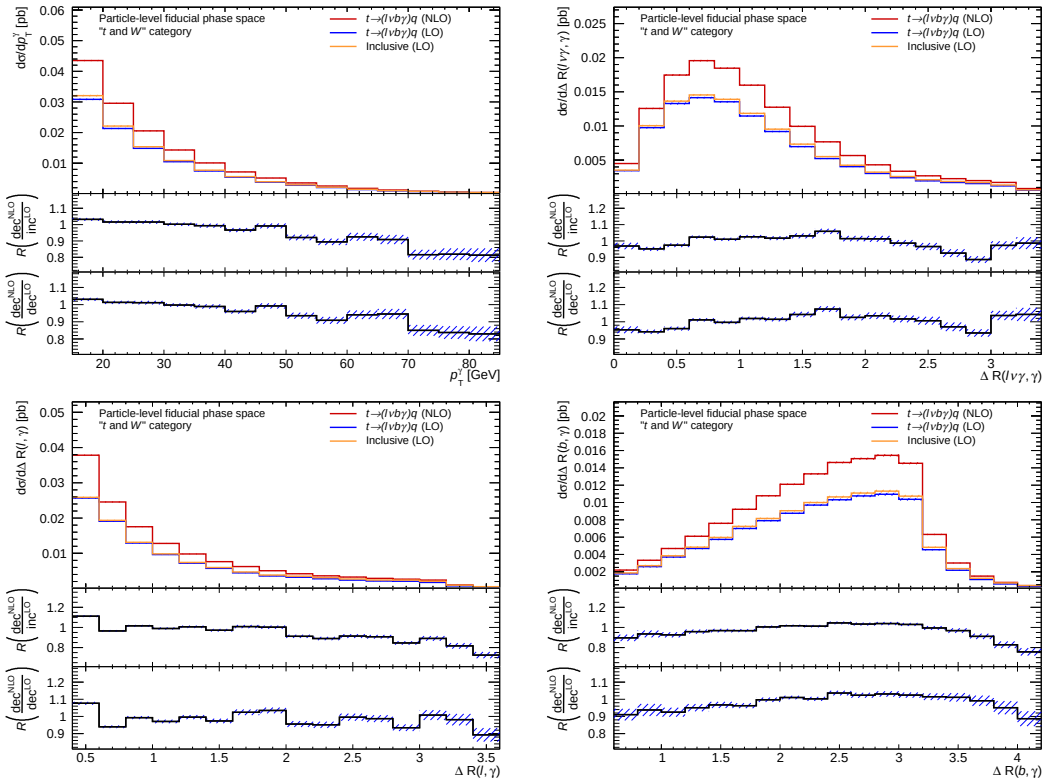


Figure 23: Particle-level differential fiducial cross sections as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the LO inclusive samples and the respective ratios as a function of p_{T}^{γ}, $\Delta R(\ell\nu\gamma, \gamma)$, $\Delta R(\ell, \gamma)$, and $\Delta R(b, \gamma)$ for the "t and W" category. The hashed bands correspond to the uncertainty of the predictions.

In the following, the kinematic properties of the events assigned to the "only t" category from the LO $t\,(\to \ell\nu b\gamma)\,q$, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the inclusive LO samples are compared. This category is expected to have a large relative contribution of $tq\gamma$ events ($> 40\%$) and hence a good agreement between the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ distributions cannot be expected.

Figure 24 shows the particle-level differential fiducial cross sections as a function of p_{T}^{γ}, $\Delta R(\ell\nu, \gamma)$, $\Delta R(\ell, \gamma)$, and $\Delta R(b, \gamma)$. The shapes of the $\Delta R(\ell, \gamma)$ distributions obtained from the LO inclusive sample and the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample agree well with each other. Furthermore, a good agreement between the shapes of the distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample and from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is only observed for this quantity.

The shape of the p_{T}^{γ} distribution obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is significantly softer than the distribution obtained from the LO inclusive sample, while the shapes of the distributions

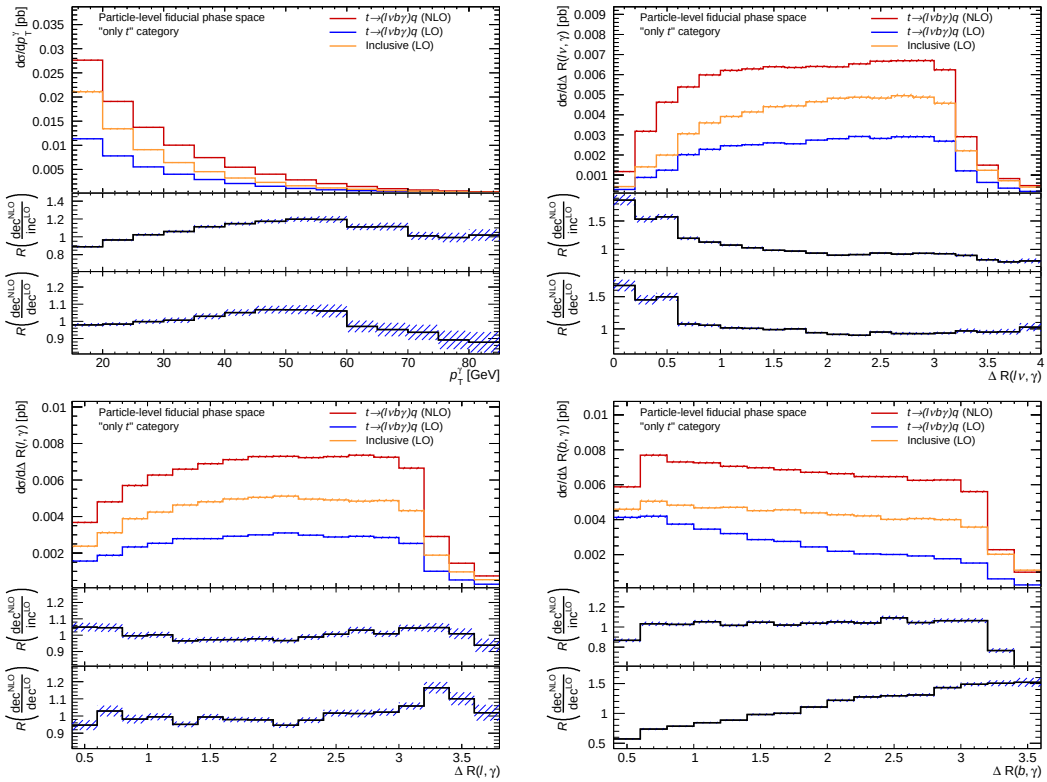


Figure 24: Particle-level differential fiducial cross section as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the LO inclusive samples and the respective ratios as a function of p_{T}^{γ}, $\Delta R(\ell\nu,\,\gamma)$, $\Delta R(\ell,\,\gamma)$, and $\Delta R(b,\,\gamma)$ for the "only t" category. The hashed bands correspond to the uncertainty of the predictions.

obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples agree well up to 60 GeV. The $tq\gamma$ contamination is mainly caused by the limitations in the matching of the b jet and b quark, which can cause $m_{\ell\nu b}$ to be smaller than m_t and $m_{\ell\nu b\gamma}$ to be closer to m_t than $m_{\ell\nu b}$. However, for larger p_{T}^{γ} values a larger fraction of $m_{\ell\nu b\gamma}$ values differs stronger from m_t than $m_{\ell\nu b}$, even when the matching of the b jet and b quark failed. Therefore, the $tq\gamma$ contamination is expected to have a softer p_{T}^{γ} spectrum relative to $t\,(\to \ell\nu b\gamma)\,q$ events and the most likely cause of the softer p_{T}^{γ} spectrum in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample relative to the spectrum in the LO inclusive sample is a difference in the prediction for the $tq\gamma$ contamination.

Both the $\Delta R(\ell\nu,\,\gamma)$ and the $\Delta R(b,\,\gamma)$ distributions exhibit larger deviations between the LO inclusive sample and the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample for values < 0.6. The directions of these deviations are opposite. This indicates that they are primarily caused by the modeling of the photon radiation in $t\,(\to \ell\nu b\gamma)\,q$ events, as the relative $tq\gamma$ contamination is expected to be higher only at larger angular distances. This indicates that the contribution of events where the photon is radiated from the W boson relative to that where the photon is radiated from the b quark is predicted to be higher in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample than in the LO $t\,(\to \ell\nu b\gamma)\,q$ sample.

Across a large value range (0.6–3.2), good agreement between the distributions from the LO inclusive and the NLO $t\,(\to \ell\nu b\gamma)\,q$ samples is observed. However, for larger values (> 3.2), similar deviations as for the respective distributions for the "t and W" category are observed, likely due to differences in the predictions for the $tq\gamma$ contamination. This assumption is supported by the good agreement observed between the shapes of the distributions obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples in this region.

While good agreement is also observed between the $\Delta R(\ell\nu,\,\gamma)$ distributions predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples in the range 0.6–3.2, the shapes of the $\Delta R(b,\,\gamma)$ distribu-

tions differ greatly, with the angular distance between the photon and b jet being significantly smaller on average in events from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample, resulting in a much steeper distribution. This is primarily due to the $tq\gamma$ contamination in the NLO $t\,(\to \ell\nu b\gamma)\,q$ as its relative contribution is expected to increase with increasing $\Delta R(b,\,\gamma)$.

For an isolated examination of the impact of the NLO QCD contributions, kinematic and topological quantities are examined that are expected to be unaffected by the difference in the photon-radiation modeling, as discussed in Appendix B.4. Figure 25 shows the particle-level differential fiducial cross section as a function of $p_{\mathrm{T}}^{\mathrm{lj}}$, η^{lj}, $p_{\mathrm{T}}^{\mathrm{fj}}$, and η^{fj} for events in the "t and W" category. Figure 26 shows the particle-level differential fiducial cross section as a function of N_{jets} and $N_{b\ \mathrm{jets}}$. Appendix B.5 provides these distributions for the "only t" category. No significant deviations are observed between the shapes of the distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and the LO inclusive samples, i.e., the small $tq\gamma$ contamination does not impact the shapes of the distributions significantly. Therefore, shape differences between the distributions obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples are expected to arise solely from the NLO QCD contributions included in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample.

Small shape differences are observed between the distributions of $p_{\mathrm{T}}^{\mathrm{fj}}$ and η^{fj} obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The forward jet is less central and carries slightly less p_{T} for NLO $t\,(\to \ell\nu b\gamma)\,q$ events on average. Significant shape differences are observed for the $p_{\mathrm{T}}^{\mathrm{lj}}$ and η^{lj} distributions. The events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are characterized by significantly lower $p_{\mathrm{T}}^{\mathrm{lj}}$ values on average, while the corresponding jet is produced less centrally on average.

Similarly to the observations in Section 6.4.1, the events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are characterized by significantly smaller multiplicities of jets and b jets on average relative to those from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample. The shape differences for the former are much larger than those observed for the other distributions and are also more pronounced than the differences observed between the respective distributions obtained from the NLO $tq\gamma$ and LO $tq\gamma$ samples.

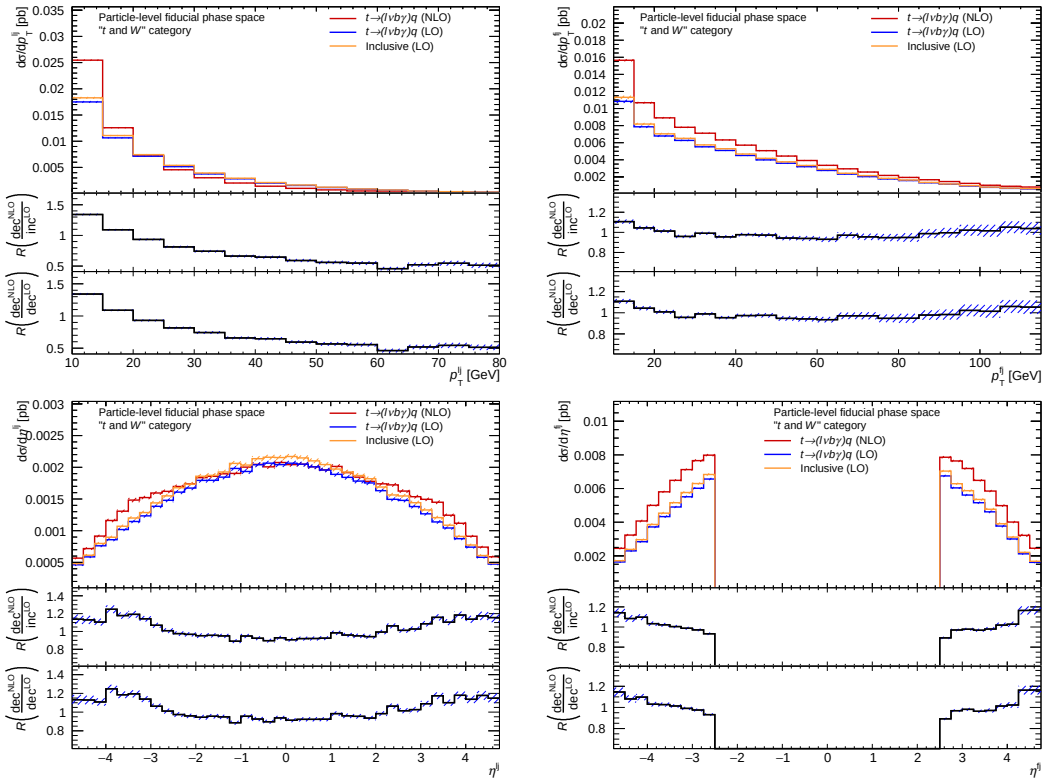


Figure 25: Particle-level differential fiducial cross section as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the LO inclusive samples and the respective ratios as a function of p_{T}^{lj}, $p_{\mathrm{T}}^{\mathrm{fj}}$, η^{lj}, and η^{fj} for the "t and W" category. The hashed bands correspond to the uncertainty of the predictions.

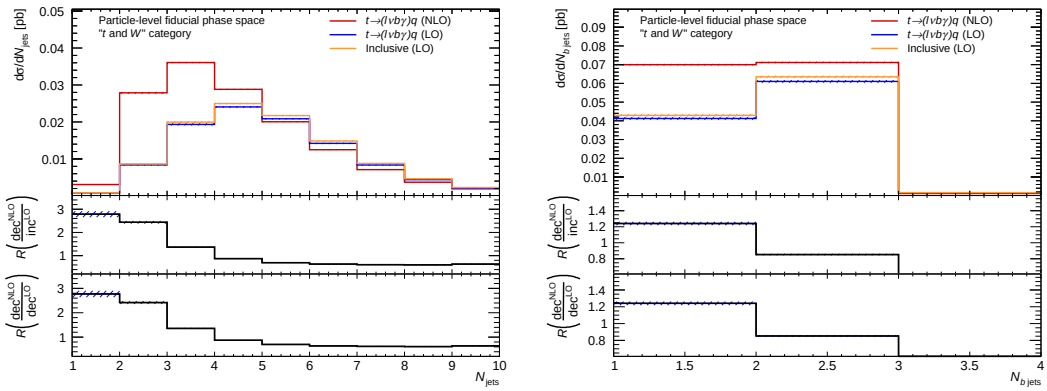


Figure 26: Particle-level differential fiducial cross section as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the LO inclusive samples and the respective ratios as a function the number of jets, N_{jets}, and b jets, $N_{b\ \mathrm{jets}}$, for the "t and W" category. The hashed bands correspond to the uncertainty of the predictions.

6.5 Conclusions about the modeling of $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$

As the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes lead to the same final states, they interfere and are, in theory, inseparable. However, the impact of the interference is found to be small, which allows for a separate simulation and analysis of these two processes in practice.

The MC event generator MadGraph5_aMC@NLO is capable of simulating the hard-scattering $tq\gamma$ process incorporating NLO QCD contributions. While the impact of these contributions on photon-related quantities, such as the p_{T} of the photon and its angular distances to other particles, is small, significant impacts on the shape are observed for kinematic and topological quantities that involve jets, especially for the multiplicity of jets. Therefore, incorporating NLO QCD contributions into the simulation is a necessity for a reliable and precise analysis of the $tq\gamma$ process in collected data.

In contrast, to date no MC event generator is capable of simulating the hard-scattering $t\,(\to \ell\nu b\gamma)\,q$ process with NLO QCD contributions. However, the impact of NLO QCD contributions on jet-related quantities is expected to be at a similar level for the $t\,(\to \ell\nu b\gamma)\,q$ process as for the $tq\gamma$ process, so that neglecting these contributions significantly limits the reliability of the predictions obtained from the simulation. An alternative approach is investigated that simulates the $t\,(\to \ell\nu b\gamma)\,q$ process with a sample that simulates the hard-scattering t-channel single-top-quark production incorporating NLO QCD contributions, while the photon radiation in the top-quark decay is solely simulated using the PS algorithm. The sample simulates $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ contributions. A procedure is defined to identify the $t\,(\to \ell\nu b\gamma)\,q$ events in this sample by requiring events to be compatible with the hypothesis of a radiative top-quark decay and/or a radiative W-boson decay. The ensemble of events being compatible forms the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. The performance of this identification approach is tested using the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples. While a good overall performance is observed, an irreducible contamination of $\approx 20\%$ from $tq\gamma$ events remains in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. This is primarily caused by the limited quality in the reconstruction of the top quark at particle level, due to the limited accuracy of the matching of the parton-level b quark from the top-quark decay with a particle-level b jet, and the limited agreement of their kinematic properties.

The kinematic and topological properties in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are compared to those of the LO $t\,(\to \ell\nu b\gamma)\,q$ sample. Differences may arise from the difference in the modeling of the photon radiation (via the PS algorithm versus in the simulation of the hard-scattering process), from the $tq\gamma$ contamination in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample, and from the NLO QCD contributions considered in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. The $tq\gamma$ contamination significantly impacts only those events in which the photon is not radiated in the W-boson decay, which are about 40% of all events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. For events with a radiative W-boson decay, the $tq\gamma$ contamination in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is small and the differences in the modeling of the photon radiation are found to only have small overall impacts on the shape of distributions. For the other events, the impact of the $tq\gamma$ contamination on the shape of distributions is found to be large for some kinematic and topological quantities, for instance, for the angular distance between the photon and the b jet emerging from the top-quark decay. Furthermore, mild impacts on the shapes due to the difference in the modeling of the photon radiation are observed. However, the impact of the NLO QCD contributions on jet-related quantities affects all types of $t\,(\to \ell\nu b\gamma)\,q$ events and the overall impact is found to be significantly larger than the other effects.

Therefore, it is concluded that the alternative approach offers a better nominal description of the kinematic and topological properties of the $t\,(\to \ell\nu b\gamma)\,q$ process relative to the simulation of the process via the LO $t\,(\to \ell\nu b\gamma)\,q$ sample. To account for the main limitation of this approach, the contamination from $tq\gamma$ events, conservative uncertainties are assigned to the predictions from this sample and are considered in the statistical analysis. These are discussed in Section 7.6.2. This approach for simulating the $t\,(\to \ell\nu b\gamma)\,q$ process is also used in Ref. [11].

7 Strategy for the analysis of t-channel single-top-quark production in association with a photon

This section outlines the strategy for examining single-top-quark production in association with a photon using pp-collision data taken with the ATLAS detector. The main goals are the precise measurement of fiducial cross sections of this process and its observation. A brief overview of this strategy is given below, while detailed descriptions of individual parts are given in dedicated sections.

First, a subset of the collected data is selected that is enriched in $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events. This event selection is shaped by the distinctive features of the expected detector signature of these signal topologies depicted in Figure 3. These distinctive features are a b jet, a photon, and a lepton, all characterized by high p_{T}, and high $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the neutrino from the W-boson decay. Furthermore, a significant fraction of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events are expected to contain a jet that is produced in the forward direction ($|\eta| > 2.5$). This is characteristic of t-channel production with one valence quark in the initial state, which occurs frequently for the signal topologies, as valence quarks typically carry large longitudinal momenta. The specific criteria for the event selection are introduced and defined in Section 7.1.

The detector signatures of the signal topologies are mimicked by several other hard-scattering processes. These processes are commonly referred to as "background" topologies. The modeling of those that are expected to significantly contribute to the selected data is discussed in Section 7.2. Furthermore, a classification scheme of simulated events based on the origin of the photon candidate is introduced. Dedicated data-driven techniques are used to correct the prediction for background contributions in which the selected photon candidate is not a prompt photon. Additionally, a data-driven estimation of background contributions in which the lepton candidate is not a prompt lepton is performed. These data-driven estimations are presented in Section 8.

The predicted contributions of the signal and background processes to the selected data are studied in Section 7.3. Furthermore, the selected data is further categorized into different selection regions characterized by a large relative contribution from either the signal or the dominant background processes. The former type of region is referred to as signal region (SR), while the latter is referred to as control region (CR). Both types are commonly referred to as analysis region (AR).

Measurements of inclusive fiducial cross sections of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes are performed by unfolding the respective measured number of events in the ARs to parton-level and particle-level fiducial phase spaces. The strategy for these measurements, the definition of the fiducial phase spaces, and the predicted values of the respective fiducial cross sections used for comparison with the measured values are discussed in Section 7.4.

The majority of events in the SRs are background events. To achieve precise measurements of the fiducial cross sections and conclusively establish the discovery of $tq\gamma$, NN models are used to efficiently discriminate between $tq\gamma$ and all other contributions to the selected data. The design, fitting, and optimization of the NN models are described in Section 9. The NN models are used to assign a value ($\mathrm{NN_{out}}$) to each simulated event and each event in the collected data, indicating its $tq\gamma$-likeliness.

The numbers of $tq\gamma$, $t\,(\to \ell\nu b\gamma)\,q$, and background events in the ARs, along with their respective uncertainties, are measured by performing a statistical analysis using binned profile likelihood (BPL) fits of the predicted $\mathrm{NN_{out}}$ distributions to the respective observed distributions. These measured values are used for the determination of the fiducial cross sections. The BPL fits are further used to test for the discovery of the $tq\gamma$ process. The technical setup of the statistical analysis and the treatment of uncertainties are discussed in Section 7.5. Section 7.6 reports on the systematic uncertainties considered in the statistical analysis.

7.1 Event selection

The aim of the event selection is to define a subset of the collected dataset that is enriched in events from the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes. As outlined in Section 4.1, multiple triggers, specialized for different detector signatures, decide whether an event is recorded by the ATLAS detector. The first step of the event selection is therefore the definition of trigger requirements that events must satisfy to be considered in this analysis.

It is expected that exactly one high-p_{T} lepton is produced in the majority of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events. Single-lepton triggers are designed to efficiently accept events containing at least one prompt lepton while suppressing those that do not. Accurately measuring tau leptons is more challenging than measuring electrons and muons, as these decay before reaching the detector. The event selection is tailored to final states with either a prompt electron or a prompt muon. However, small contributions from final states with tau leptons may arise through their leptonic decays, which produce an electron or a muon in $\approx 35\%$ of the cases [36]. Such electrons and muons carry significantly lower p_{T} on average relative to those produced in the decay of a massive EW boson, as is the case for the signal topologies.

A combination of different single-electron [210] and single-muon triggers [225] is used. Events are required to pass the requirements of at least one of these triggers. An overview of the single-lepton triggers utilized in this analysis for the different data-taking periods is given in Table 7 using the naming convention employed by the ATLAS collaboration. Differences between collected and simulated data in the trigger efficiencies are corrected by reweighting simulated events with scale factors derived from data-to-simulation efficiency ratios.

The muon HLTs employed for data taken in 2015 are seeded by the L1 muon trigger using a p_{T} threshold of 15 GeV (L1MU15). A low-p_{T} HLT (HLT_mu20_iloose) with a threshold of 20 GeV applying loose isolation ($p_{\mathrm{T}}^{\mathrm{cone20}} < 0.12\,p_{\mathrm{T}}$) requirements is used in combination with a high-p_{T} HLT (HLT_mu50) that uses a p_{T} threshold of 50 GeV and no isolation requirements. For data taken during the years 2016–2018, the p_{T} threshold for the L1 trigger is increased to 20 GeV. The threshold of the low-p_{T} HLT (HLT_mu26_ivarmedium) is increased to 26 GeV and medium isolation requirements are employed ($p_{\mathrm{T}}^{\mathrm{varcone30}} < 0.07\,p_{\mathrm{T}}$), while the same requirements are used for the high-p_{T} HLT as in 2015.

The electron HLTs employed for data taken in 2015 are seeded by the L1 electron trigger (L1EM20) using a nominal electron-p_{T} threshold of 20 GeV. This threshold slightly varies with η. Furthermore, the L1 trigger includes a requirement to reject hadronic activity. Three HLTs are used with p_{T} thresholds of 24 GeV, 60 GeV, and 120 GeV. The former two apply the requirements of the medium LH identification WP (lhmedium) to candidate signatures, while the latter applies those of the loose LH identification WP (lhloose). A threshold of 22 GeV is used for the L1 trigger for data taken during the years 2016–2018 that applies additional isolation requirements to electron candidates with $p_{\mathrm{T}} < 50$ GeV. Three HLTs are also used during this period with p_{T} thresholds of 26 GeV (HLT_e26_lhtight_nod0_ivarloose), 60 GeV (HLT_e60_lhmedium_nod0), and 140 GeV (HLT_e140_lhloose_nod0) and apply the requirements of the tight (lhtight), medium, and loose LH identification WPs, respectively. All three triggers exclude d_0 information (nod0) as input to the LH. The low-p_{T} HLT additionally applies loose isolation requirements (ivarloose) to the candidate signatures ($p_{\mathrm{T}}^{\mathrm{varcone20}} < 0.1\,p_{\mathrm{T}}$).

Events that pass the trigger selection are further filtered based on the object definitions given in Section 4.3, guided by the expected detector signature of signal events. The following requirements form the pre-selection region (PSR) and are used to select a subset of the collected data for further analysis: Events are required to contain exactly one electron candidate with $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.5$, excluding electron candidates with $1.37 < |\eta| < 1.52$, or exactly one muon candidate with $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.5$.[26] The selected lepton candidate needs to match a detector signature that fired one of the single-lepton triggers. Furthermore, events are required to contain at least one photon candidate

[26] The p_{T} requirements are motivated by the p_{T} threshold of the low-p_{T} triggers and are slightly higher to avoid "turn-on" effects.

Table 7: List of single-lepton triggers used in this analysis [210, 225].

Data-taking period	Electron trigger	Muon trigger
2015	HLT_e24_lhmedium_L1EM20VH	HLT_mu20_iloose_L1MU15
	HLT_e60_lhmedium	HLT_mu50
	HLT_e120_lhloose	
2016–2018	HLT_e26_lhtight_nod0_ivarloose	HLT_mu26_ivarmedium
	HLT_e60_lhmedium_nod0	HLT_mu50
	HLT_e140_lhloose_nod0	

with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.37$, excluding candidates with $1.37 < |\eta| < 1.52$, and to satisfy $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$. At least one jet with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$ passing the 85% WP of the DL1r b-tagging algorithm needs to be present in an event.

7.2 Background modeling and event categorization

While it is possible to filter a signal-enriched subset from the recorded dataset with the PSR requirements, it remains inevitable that residual contributions from other underlying hard-scattering processes persist within this selected dataset. To perform precise examinations of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ with the selected data, it is hence essential to reliably estimate this contamination. Since there is no inherent way to estimate the relative contributions from different underlying processes, simulation-based techniques, data-driven techniques or a combination of both are used for the modeling of the individual contributions.

Most background processes are estimated in a simulation-based fashion. The respective underlying hard-scattering processes and the characteristics of the MC samples are discussed below. The MC samples are mainly categorized into two classes: the $X\gamma$ samples, in which photon radiation is modeled in the simulation of the hard-scattering process, and the X samples, in which photon radiation is modeled exclusively by the PS algorithm. Here, X denotes the hard-scattering process simulated by the sample. For all background MC samples, the 5FS and the corresponding NNPDF3.0NLO PDF set [212] are employed in the modeling of the hard-scattering processes. For simulations with Pythia 8, the NNPDF2.3LO PDF set [212], the A14 tune for parameters describing the UE, and EvtGen for the modeling of the decay of b and c hadrons are used. All background MC samples are centrally provided by the ATLAS collaboration for physics analyses to its researchers.

Small background contributions arise from events with reconstructed lepton candidates that are not prompt leptons. These contributions are referred to as "fake leptons" and are estimated in a data-driven fashion. This estimation is discussed in Section 8.1.

7.2.1 $X\gamma$ samples

Similar to the photon-associated t-channel single-top-quark production, the $t\bar{t}\gamma$ (dec) and $t\bar{t}\gamma$ (prod) processes are modeled using two dedicated samples. Both processes are simulated using MadGraph5_aMC@NLO for modeling the hard scatter while Pythia 8 is utilized for modeling the PS, hadronization, and UE. Furthermore, at least one of the top quarks is required to decay semileptonically. The hard-scattering $t\bar{t}\gamma$ (dec) process and the subsequent decay of the top-quark pair are simulated at LO. The hard-scattering $t\bar{t}\gamma$ (prod) process is modeled incorporating NLO QCD contributions. The decay of the top-quark pair is simulated with MadSpin at LO. The nominal prediction for the cross section of the $t\bar{t}\gamma$ (prod) process is given by the prediction obtained from the MC simulation of the $t\bar{t}\gamma$ (prod) sample. The prediction of the respective cross section of the $t\bar{t}\gamma$ (dec) process obtained from a LO calculation with MadGraph5_aMC@NLO is adjusted by a k-factor, so that the sum of the cross

sections of the $t\bar{t}\gamma$ (prod) and $t\bar{t}\gamma$ (dec) processes equals the prediction for the fiducial cross section reported in Ref. [226] for the inclusive process, i.e., for the combined $t\bar{t}\gamma$ (prod) and $t\bar{t}\gamma$ (dec) process. Further details about these samples are given in Ref. [227].

Samples simulating the production of W and Z bosons in association with jets and a photon are produced using the MC event generators Sherpa 2.2.2 and Sherpa 2.2.4 [228], respectively. The samples simulate the hard-scattering processes with up to three additional final-state partons beyond those present at LO, while the corresponding matrix elements are calculated incorporating NLO QCD contributions for up to one additional parton using the OpenLoops 2 [229–231] library and with LO accuracy for the remaining contributions using the Comix [232] library. The hard-scattering events are merged with the internal PS algorithm based on the Catani-Seymour dipole factorization [232, 233] using the MEPS@NLO prescription [234–237]. Hadronization, the UE, and the decay of hadrons are also modeled internally by the MC event generators. The parameters of the PS algorithm and the models for hadronization and the UE are adjusted with a tune developed by the authors. Only leptonic decays of the vector bosons are considered. The samples are referred to as $V\gamma$+jets ($V = \{W, Z\}$) samples in the following.

7.2.2 X samples

Samples simulating inclusive top-quark pair production as well as inclusive single-top-quark production via the s and tW channels are produced with Powheg Box v2 for modeling the hard scatter and Pythia 8 for modeling the PS, hadronization, and UE. These samples are referred to as $t\bar{t}$, tW, and s-channel samples, respectively. The hard-scattering processes are simulated including NLO QCD contributions for all of these samples. The h_{damp} parameter, which regulates the high-p_{T} emission of gluons in Powheg Box v2, is set to $1.5\,m_t$ [238] for simulating $t\bar{t}$ production. At least one top quark is required to decay semileptonically in the $t\bar{t}$ sample. The decay of top quarks in the s-channel sample is also restricted to the semileptonic decay channel, while no decay channel restrictions are applied for the tW-channel sample. The tW sample contains contributions that interfere with $t\bar{t}$ production and hence are present in both samples. This overlap is removed in the tW sample by applying the diagram removal (DR) scheme [239]. The events in the NLO sample simulating t-channel single-top-quark production that are not contained in the NLO $t\,(\rightarrow \ell\nu b\gamma)\,q$ sample (cf. Section 6.4) are also used in the analysis. The ensemble of these events is referred to as the t-channel sample.

Samples simulating the production of W bosons and Z bosons in association with jets are simulated with the Sherpa 2.2.1 [228] generator. The samples simulate the respective hard-scattering processes with up to four additional final-state partons beyond those present at LO, while the corresponding matrix elements are calculated including NLO QCD contributions for up to two additional partons using the OpenLoops 2 library and with LO accuracy for the remaining contributions using the Comix library. Only leptonic decays of the vector boson are considered. The samples are referred to as V+jets ($V = \{W, Z\}$) samples in the following.

Hard-scattering processes in which two massive vector bosons are produced, i.e., WZ, WW, and ZZ production, are simulated with Sherpa 2.2.2 when both vector bosons decay leptonically, while Sherpa 2.2.1 is used in all other cases. The samples simulate these processes with up to three additional final-state partons beyond those present at LO, while the corresponding matrix elements are calculated incorporating NLO QCD contributions for up to one additional parton using the OpenLoops 2 library and using LO accuracy for the remaining contributions using the Comix library. Samples simulating the loop-induced processes $gg \rightarrow VV$ are generated using LO-accurate matrix elements for the emission of up to one additional parton.

7.2.3 Removal of overlap between X and $X\gamma$ samples

As PS algorithms simulate QED radiation from initial- and final-state particles, an X sample might simulate contributions that are also present in the respective $X\gamma$ sample. To remove this overlap, a dedicated procedure was developed by the ATLAS collaboration that is applied to all events in the X

and $X\gamma$ samples containing a prompt photon. This overlap removal procedure favors contributions modeled by the $X\gamma$ samples for two main reasons: First, PS algorithms model the QED radiation in the collinear limit and hence do not offer an accurate description of high-p_{T} photon emission at large angles. Furthermore, only a small fraction of events in the X samples contain a high-p_{T} prompt photon and the $X\gamma$ samples typically model prompt-photon contributions with a higher statistical precision. However, kinematic requirements on the photon are typically employed in the production of the $X\gamma$ samples, for instance a minimum p_{T} requirement to avoid infrared divergences in the computation of the matrix elements of the respective hard-scattering process. Therefore, PS-modeled photon radiation represents the only source of prompt photons in the simulation that fail these kinematic requirements.

As a first step of the overlap removal procedure, a list of particle-level final-state prompt photons and prompt leptons is compiled for each simulated event. An overlap region is defined by the requirements $\Delta R(\ell, \gamma) > 0.2$ and $p_{\mathrm{T}}^{\gamma} > 15\,\mathrm{GeV}$. Events that satisfy these requirements for any of the photon–lepton combinations are considered to be part of the overlap region. Events in the overlap region are removed for the X samples and kept for the $X\gamma$ samples. In contrast, events that do not fulfill these requirements are removed for the $X\gamma$ samples and kept for the X samples.

7.2.4 Event categorization

The majority of fake photons are rejected by the reconstruction, isolation, and identification algorithms discussed in Sections 4.3 and 5. However, irreducible and significant fake-photon contributions remain as part of the selected data. These contributions can be mainly classified into two categories, which are described in the following.

Electrons may be falsely reconstructed and identified as photon candidates due to the similarity of the detector signatures of these two particles. As discussed in Ref. [2], this is mainly caused by the limited efficiencies of the reconstruction of electron tracks in the ID and of the matching of tracks to the cluster in the ECAL. For instance, an electron might be reconstructed as an unconverted photon if the associated track is not reconstructed or not matched to the cluster, or as a converted photon when at least one of the tracks forming a CV is matched to an electron cluster. Furthermore, electrons radiate bremsstrahlung when interacting with material. The emitted bremsstrahlung photons might also be misreconstructed as photon candidates and subsequently pass the isolation and identification criteria. These contributions are commonly referred to as "electron-to-photon" ($e \to \gamma$) fake contributions.

The contributions of events with fake photons arising from hadronic activity (cf. Chapter 5) are commonly referred to as "hadron-to-photon" ($h \to \gamma$) fake contributions. Although the photon-isolation and photon-identification requirements offer excellent rejection power against these contributions, $h \to \gamma$ events typically represent a significant background contribution to analyses studying final states with prompt photons.

Typically, $X\gamma$ and X samples simulate $e \to \gamma$ and $h \to \gamma$ contributions. The particle from the underlying particle level associated with the p_{T}-leading reconstructed photon candidate is used to identify these contributions. This association is performed by internal software of the ATLAS collaboration and stored in the event record. The events are classified as $e \to \gamma$ or $h \to \gamma$ events by the following scheme, where at least one of the conditions within each category must be fulfilled:

- $e \to \gamma$ event:
    - the matched particle-level object is an electron
    - the matched particle-level object is a photon while the maximum angular distance to the closest particle-level electron is smaller than $\Delta R = 0.05$
- $h \to \gamma$ event:
    - the matched particle-level object is a hadron

- the matched particle-level object is a non-prompt photon originating from a hadron

The remaining events are referred to as prompt-photon (prompt-γ) events. The majority of all p_{T}-leading photon candidates in the $X\gamma$ samples are prompt photons and hence the simulated $e \to \gamma$ and $h \to \gamma$ contributions mainly stem from the X samples.

There are no calibrations for the simulated $e \to \gamma$ and $h \to \gamma$ contributions that are centrally provided by the ATLAS collaboration. Dedicated data-driven estimations are performed, and data-to-simulation SFs are computed parameterized as a function of kinematic properties of the photon candidate and of its reconstruction type (cf. Table 2). These are used to correct for potential differences between collected data and simulation. The estimations are discussed in Sections 8.2 and 8.3, respectively.

7.3 Definition and examination of signal and control regions

Figure 27 shows the predicted relative signal and background contributions to the PSR. The predicted $tq\gamma$ contribution of $\approx 3\%$ is about three times larger than the predicted $t\,(\to \ell\nu b\gamma)\,q$ contribution of $\approx 1\%$. The dominant background contributions arise from the $t\bar{t}\gamma$ (27%), $t\bar{t}$ (21%), and $W\gamma$+jets (21%) topologies. In contrast to the signal processes, the $t\bar{t}\gamma$ (prod) contribution of 11% is smaller than the $t\bar{t}\gamma$ (dec) contribution of 16%. Most of the selected $t\bar{t}$ events (75%) are classified as $e \to \gamma$ events. In most of these cases, both top quarks decay semileptonically, where at least one electron is produced that is misreconstructed as photon candidate. Approximately 20% of the selected $t\bar{t}$ events are single-lepton events classified as $h \to \gamma$ events. These are events for which one of the top quarks decays leptonically, while the other decays hadronically. The remaining $\approx 5\%$ of the total $t\bar{t}$ yield arises from events with prompt photons that survive the $X/X\gamma$ overlap removal. Sub-dominant contributions to the PSR arise from the Z+jets (10%) and $Z\gamma$+jets (6%) processes, while the other processes only provide small individual contributions.

Figure 27 also shows the predicted relative contributions of the different event categories for all, converted, and unconverted photon candidates and the respective predicted absolute number of events. The overall contribution of events with an unconverted photon candidate is 40% larger than the contribution of events with a converted photon. Approximately 67% of all simulated events are classified as prompt-γ event, while the absolute contribution of events containing an unconverted prompt photon is ≈ 2.3 times larger than that of events containing a converted prompt photon. More than a quarter of all events arise from $e \to \gamma$ contributions. Therefore, the precise and reliable estimation of this background is a crucial part of this analysis. The absolute $e \to \gamma$ contribution with converted photons is approximately twice as large as the absolute $e \to \gamma$ contribution with unconverted photons. A smaller, but significant overall contribution of 6% arises from $h \to \gamma$ events. Hence, the reliable estimation of this background is an important part of the analysis as well. The absolute $h \to \gamma$ contribution with unconverted photons is about 40% larger than the absolute $h \to \gamma$ contribution with converted photons.

The requirements of the PSR are further optimized and extended to enhance the sensitivity to contributions of the signal processes. The $N_S/\sqrt{N_B}$ ratio is chosen as measure for the sensitivity, where N_S represents the predicted total signal yield and N_B the predicted total background yield. The predicted yields for the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes are denoted by $N_{tq\gamma}$ and $N_{t(\to\ell\nu b\gamma)q}$, respectively. The value of this ratio corresponds to the statistical significance in counting experiments when systematic uncertainties in the predictions are neglected. Therefore, a higher value of $N_S/\sqrt{N_B}$ indicates an enhanced sensitivity. The value of the ratio in the PSR is 14.2 for $N_S = N_{tq\gamma}$ when considering $t\,(\to \ell\nu b\gamma)\,q$ as additional background contribution, 4.2 for $N_S = N_{t(\to\ell\nu b\gamma)q}$ when considering $tq\gamma$ as additional background contribution, and 18.3 for $N_S = N_{tq\gamma} + N_{t(\to\ell\nu b\gamma)q}$. The additional requirements of the extended pre-selection region (ePSR) are introduced in the following.

The PSR requires at least one jet passing the 85% DL1r WP to be present in an event. This requirement is denoted by $N_b^{85} \geq 1$. Table 8 shows the values for $N_S/\sqrt{N_B}$ when the 85% WP is replaced with a

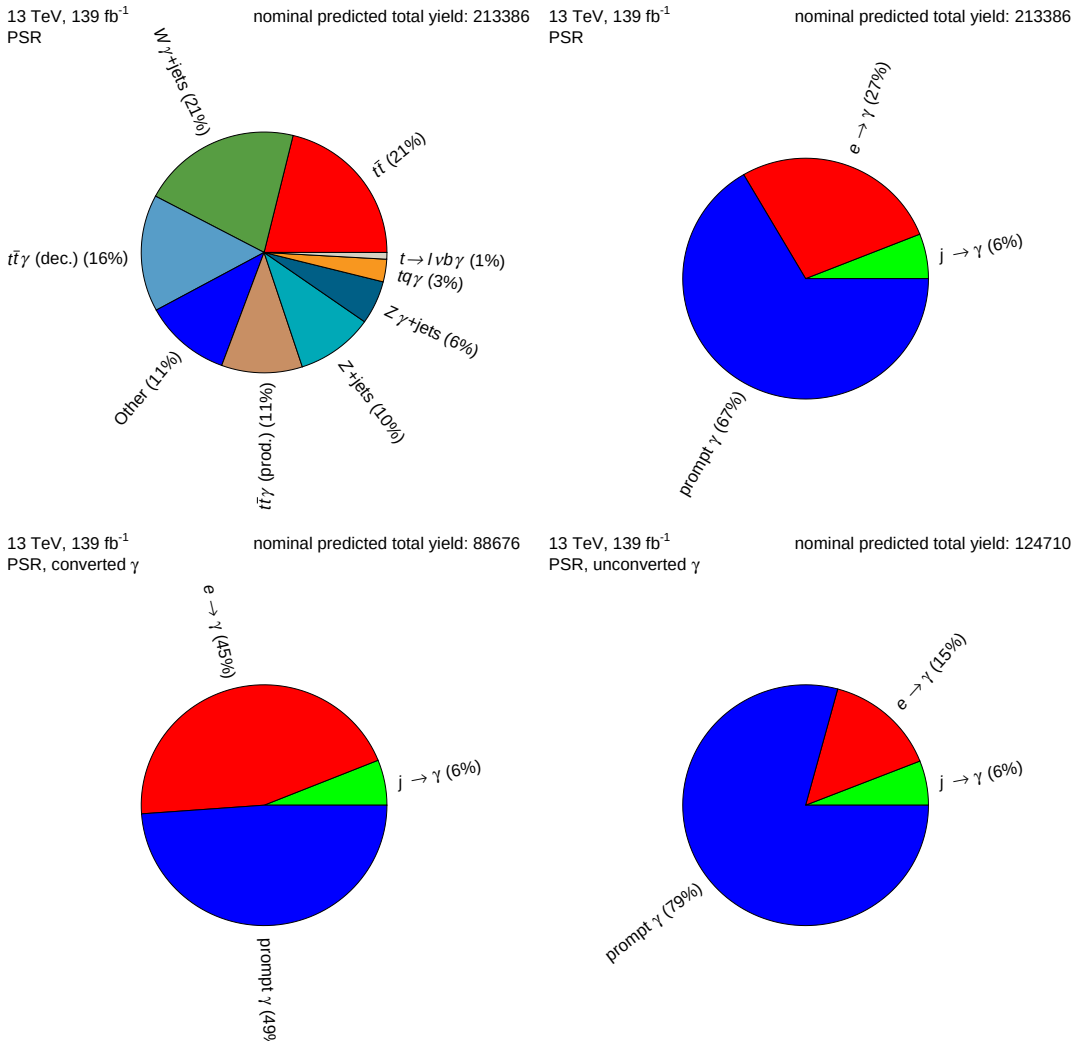


Figure 27: Predicted relative contributions in the PSR of the signal and background processes (left) as well as of events classified as prompt-γ, $e \to \gamma$ or $h \to \gamma$ events (right). Relative background contributions below 5% are summarized as "Other". The sum of all expected signal and background yields is given in the top right corner.

Table 8: Dependence of the $N_S/\sqrt{N_B}$ ratio on the chosen WP (XX) of the DL1r b-tagging algorithm for the $N_b^{XX} \geq 1$ requirement of the event selection.

	DL1r WP			
	$N_b^{60} \geq 1$	$N_b^{70} \geq 1$	$N_b^{77} \geq 1$	$N_b^{85} \geq 1$
$N_{tq\gamma}/\sqrt{N_B + N_{t(\to\ell\nu b\gamma)q}}$	14.8	15.3	15.2	14.2
$N_{t(\to\ell\nu b\gamma)q}/\sqrt{N_B + N_{tq\gamma}}$	4.2	4.4	4.4	4.2
$(N_{tq\gamma} + N_{t(\to\ell\nu b\gamma)q})/\sqrt{N_B}$	19.2	19.9	19.7	18.3

tighter WP, i.e., a WP offering a better rejection of non-b-jets at the cost of a smaller efficiency for b jets. The events that pass this alternative requirement are hence a subset of the PSR events. While the differences in the $N_S/\sqrt{N_B}$ values are small for the different options, choosing the 70% WP provides the best sensitivity and therefore the requirement $N_b^{70} \geq 1$ is used for the ePSR.

The improvement of the sensitivity is predominantly caused by the reduction of the number of $V(\gamma)$+jets background events. Most such events that satisfy the $N_b^{85} \geq 1$ requirement contain a jet that is falsely tagged as a b jet and is either a c jet or a LF jet. This is exemplarily shown for the $W\gamma$+jets events in Figure 28. The events are classified using particle-level information into events that contain at least one b jet ("bottom"), at least one c jet, but no b jet ("charm"), and only LF jets ("light"). The latter two

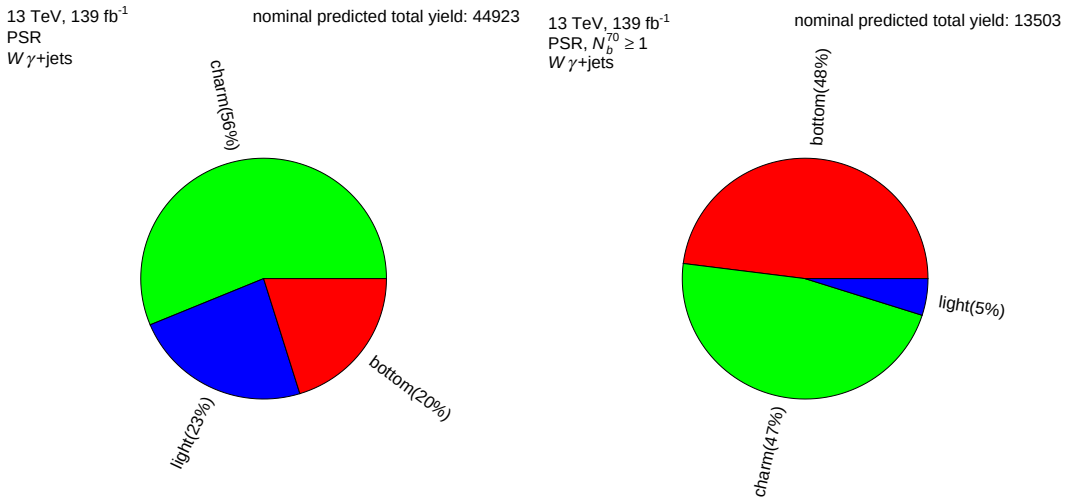


Figure 28: Predicted relative contributions to the total predicted $W\gamma$+jets contribution that contain at least one b jet ("bottom"), at least one c jet, but no b jet ("charm"), and only LF jets ("light") at particle level. The left figure shows these for events that pass the PSR requirements and the right figure for events that pass the PSR and the $N_b^{70} \geq 1$ requirements.

categories correspond to events for which a jet is falsely tagged as b jet. The fraction of events for which a jet is falsely tagged as b jet is reduced from 79% to 52% when altering the requirement from $N_b^{85} \geq 1$ to $N_b^{70} \geq 1$. This reduces the overall predicted relative contribution of $W\gamma$+jets events from 21% to 10% to the respective dataset.

Most of the Z+jets events are $e \to \gamma$ events ($\approx 96\%$) for which one electron of the Z-boson decay is misreconstructed as a photon candidate. This contribution is denoted by $Z \to e(e \to)\gamma$. Here, $(e \to)\gamma$ denotes the misreconstruction and misidentification of one of the final-state electrons. Figure 29 shows the composition of the total $e \to \gamma$ contribution. Only $t\bar{t}$ and Z+jets contribute significantly to the total $e \to \gamma$ yield with relative contributions of about 72% and 21%, respectively. A sub-dominant contribution arises from tW events (6%). Figure 29 also shows the distribution of the invariant mass of the lepton and photon candidate ($m_{\ell\gamma}$). A peak is observed in the $e \to \gamma$ distribution close to the Z-boson pole mass ($\approx 90\,\text{GeV}$) which is caused by the $Z \to e(e \to)\gamma$ events. A veto on events with $80\,\text{GeV} < m_{e\gamma} < 100\,\text{GeV}$ is added to the ePSR criteria, which reduces the $Z \to e(e \to)\gamma$ contribution to the total $e \to \gamma$ contribution to $\approx 3\%$. The requirement is referred to as "$Z \to e(e \to)\gamma$ veto". The total contribution of Z+jets events is reduced to 1%.

Adding the $Z \to e(e \to)\gamma$ veto results in sensitivity values of $N_{tq\gamma}/\sqrt{N_B} = 15.2$, $N_{t(\to\ell\nu b\gamma)q}/\sqrt{N_B} = 4.5$, and $(N_{tq\gamma} + N_{t(\to\ell\nu b\gamma)q})/\sqrt{N_B} = 19.9$. The sensitivity to $tq\gamma$ slightly decreases, while the sensitivity to $t\,(\to \ell\nu b\gamma)\,q$ is slightly enhanced by the $Z \to e(e \to)\gamma$ veto. This is caused by the difference in the shape of the predicted $m_{\ell\gamma}$ distributions of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events. As illustrated in the ratio panel of the right figure in Figure 29, the fraction of events in the $t\,(\to \ell\nu b\gamma)\,q$ sample that are contained in the $80\,\text{GeV} < m_{e\gamma} < 100\,\text{GeV}$ region is smaller than in the $tq\gamma$ sample. Therefore, the relative fraction of events surviving the $Z \to e(e \to)\gamma$ veto is larger for $t\,(\to \ell\nu b\gamma)\,q$ than for $tq\gamma$. As the overall sensitivity to the signal processes remains the same, the introduction of the veto is considered positive overall, as it contributes to the strong reduction of a less well-modeled background.

Based on the ePSR criteria, SRs and CRs are defined, which are used in the statistical analysis. The former are characterized by having the highest relative contributions from the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes compared to other regions. The SRs offer the best prospects for the discovery of the $tq\gamma$ process and for precise measurements of fiducial cross sections of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes. The CRs are used to estimate the main background contributions to the SRs more precisely, thereby increasing the accuracy of the investigation of the signal processes. The main background contributions

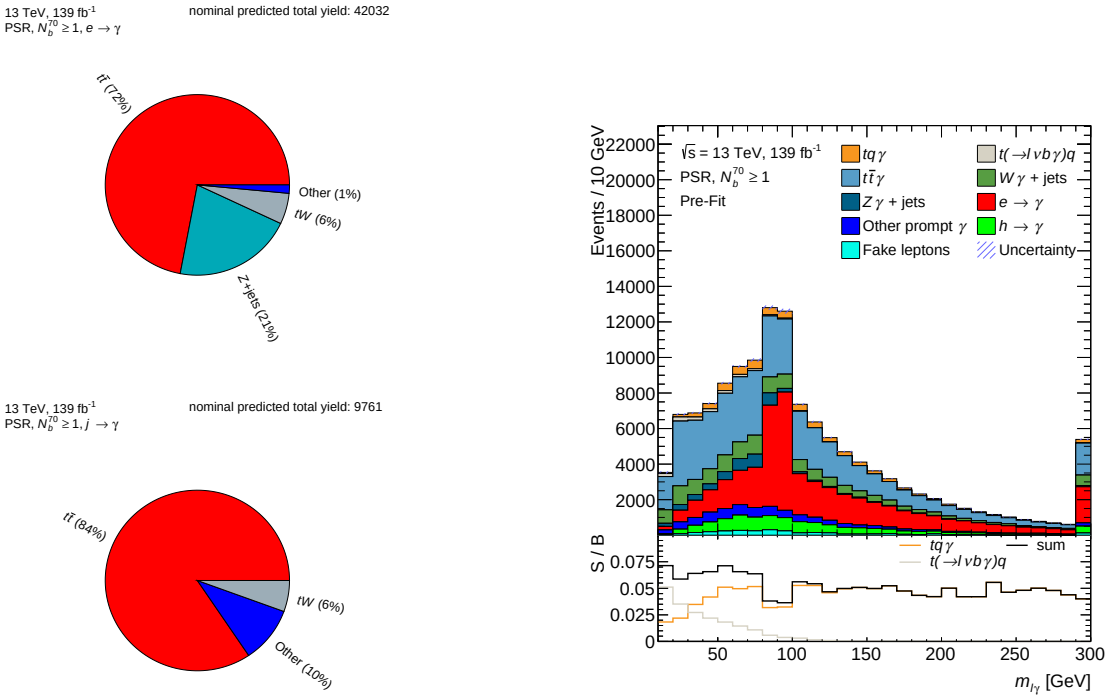


Figure 29: Predicted relative contributions of individual processes to the overall $e \to \gamma$ contribution (upper left) as well as to the $h \to \gamma$ events (lower left) and the predicted $m_{\ell\gamma}$ distribution (right). Events are required to pass the requirements of the PSR and $N_b^{70} \geq 1$. The first and last bins of the $m_{\ell\gamma}$ distribution contain the under- and overflow contributions, respectively. The uncertainty band corresponds to the uncertainty due to the limited number of events. The ratio panel illustrates the predicted $tq\gamma$, $t\,(\to \ell\nu b\gamma)\,q$, and the sum of these two contributions relative to the background contributions.

arise from the $t\bar{t}$, $t\bar{t}\gamma$, and $W\gamma$+jets topologies. The $t\bar{t}$ contribution mostly consist of $e \to \gamma$ and $h \to \gamma$ events, which are estimated by dedicated data-driven estimation techniques. Therefore, CRs are introduced for the $t\bar{t}\gamma$ and $W\gamma$+jets contributions.

One of the main distinguishing features between these background topologies and the signal topologies is the number and origin of b-tagged jets. The $W\gamma$+jets events rarely contain high-p_{T} b jets and, as shown in Figure 28, approximately half of the b-tagged jets are mistagged c jets in the ePSR. Therefore, a CR enriched in $W\gamma$+jets can be designed either by inverting or loosening the b-tagging requirements. Conversely, two high-p_{T} b jets are expected to emerge from the decay of the top-quark pair in $t\bar{t}\gamma$ events. For the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes, only one high-p_{T} b jet is expected from the top-quark decay, while a second b jet that emerges from gluon splitting ($g \to b\bar{b}$) in the initial state is typically softer and produced with $|\eta| > 2.5$ in a significant fraction of events. The b-tagging efficiency is lower for b jets with low p_{T} than for those with high p_{T} and the b-tagging algorithm is only applied to jets with $|\eta| < 2.5$. A second b-tagged jet is hence expected to occur less often for the signal topologies relative to $t\bar{t}\gamma$. Therefore, regions requiring a higher number of b-tagged jets are suited for controlling the $t\bar{t}\gamma$ background.

Figure 30 shows the expected contributions of all signal and background processes in bins defined by the number of "loose" (N_b^{l}) and "tight" b-tagged jets (N_b^{t}) when applying the PSR requirements and the $Z \to e(e \to)\gamma$ veto. A tight b-tagged jet is considered to be a jet that passes the 70% WP, while a loose b-tagged jet is a jet that passes the criteria of the 85% WP while failing the criteria of the 70% WP. The $N_b^{\mathrm{t}} = 0\,\&\,N_b^{\mathrm{l}} \geq 1$ bin is highly enriched in $W\gamma$+jets events with a relative overall contribution of $\approx 50\%$. Therefore, the CR for $W\gamma$+jets ($W\gamma$ CR) is defined by using this requirement in addition to the

Table 9: Values of the $N_S/\sqrt{N_B}$ ratio for the 0fj SR and $\geq$ 1fj SR for the different N_S values.

	0fj SR	$\geq$ 1fj SR
$N_S = N_{tq\gamma}$	8.70	12.05
$N_S = N_{t(\to\ell\nu b\gamma)q}$	2.23	2.89
$N_S = N_{tq\gamma+t(\to\ell\nu b\gamma)q}$	11.02	15.19

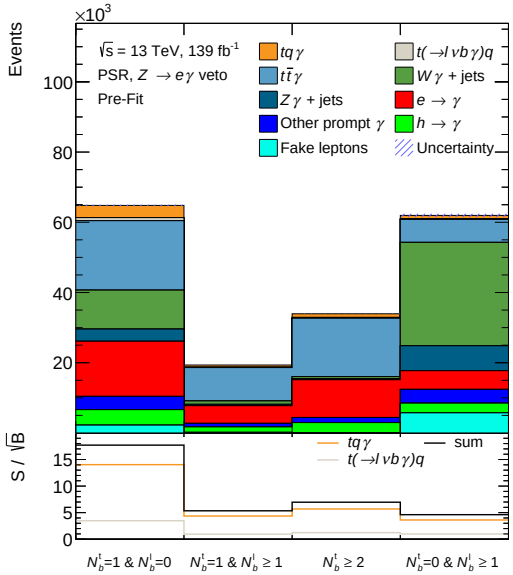


Figure 30: Predicted contributions of the signal and background processes for different requirements on the number of loose and tight b-tagged jets, which are shown on the x-axis. Events are required to pass the PSR requirements and the $Z \to e(e \to)\gamma$ veto. The hashed band corresponds to the uncertainty due to the limited sample sizes of the different processes.

PSR requirements and the $Z \to e(e \to)\gamma$ veto. The $N_b^{\mathrm{l}} = 0 \,\&\, N_b^{\mathrm{t}} = 1$ requirement offers the highest sensitivity to both the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes and is used to define the SR.

The respective b-tagging requirements of the other two bins are both well-suited for defining a CR for $t\bar{t}\gamma$ as about 50% of all events are predicted to arise from this process in each of the regions. Overall, a larger number of events pass the $N_b^{\mathrm{t}} \geq 2$ requirement. However, the $N_b^{\mathrm{t}} = 1 \,\&\, N_b^{\mathrm{l}} \geq 1$ requirement is more similar to the b-tagging requirements of the SR as events satisfy $N_b^{\mathrm{t}} = 1$ for both regions. Therefore, the latter b-tagging requirements are preferred for defining the $t\bar{t}\gamma$ CR as the respective extrapolation of kinematic features from this region to the SR is expected to be smaller. This is supported by Figure 31 which shows the relative contribution of $e \to \gamma$, $h \to \gamma$, and prompt-γ events to the regions displayed in Figure 30. The compositions are almost identical for the SR and the $t\bar{t}\gamma$ CR, while the composition is significantly different from the SR for the $W\gamma$ CR and the $N_b^{\mathrm{t}} \geq 2$ region.

To further enhance the sensitivity to the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes, the events of the SR are sorted into two orthogonal SRs, which are denoted by $\geq$ 1fj SR and 0fj SR. Events containing at least one forward jet are sorted into the former region, while the remaining events are sorted into the latter. As the presence of a forward jet is a distinctive feature of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ topologies, the $\geq$ 1fj SR is characterized by high $N_S/\sqrt{N_B}$ values, while the 0fj SR is characterized by a significantly lower value. The respective values are given in Table 9. An overview of the requirements related to b tagging and the number of forward jets of the different regions is given in Table 10, highlighting those that are used in the statistical analysis. Hereafter, "SRs" refers to the 0fj SR and $\geq$ 1fj SR, while "SR" refers to the combined 0fj SR and $\geq$ 1fj SR.

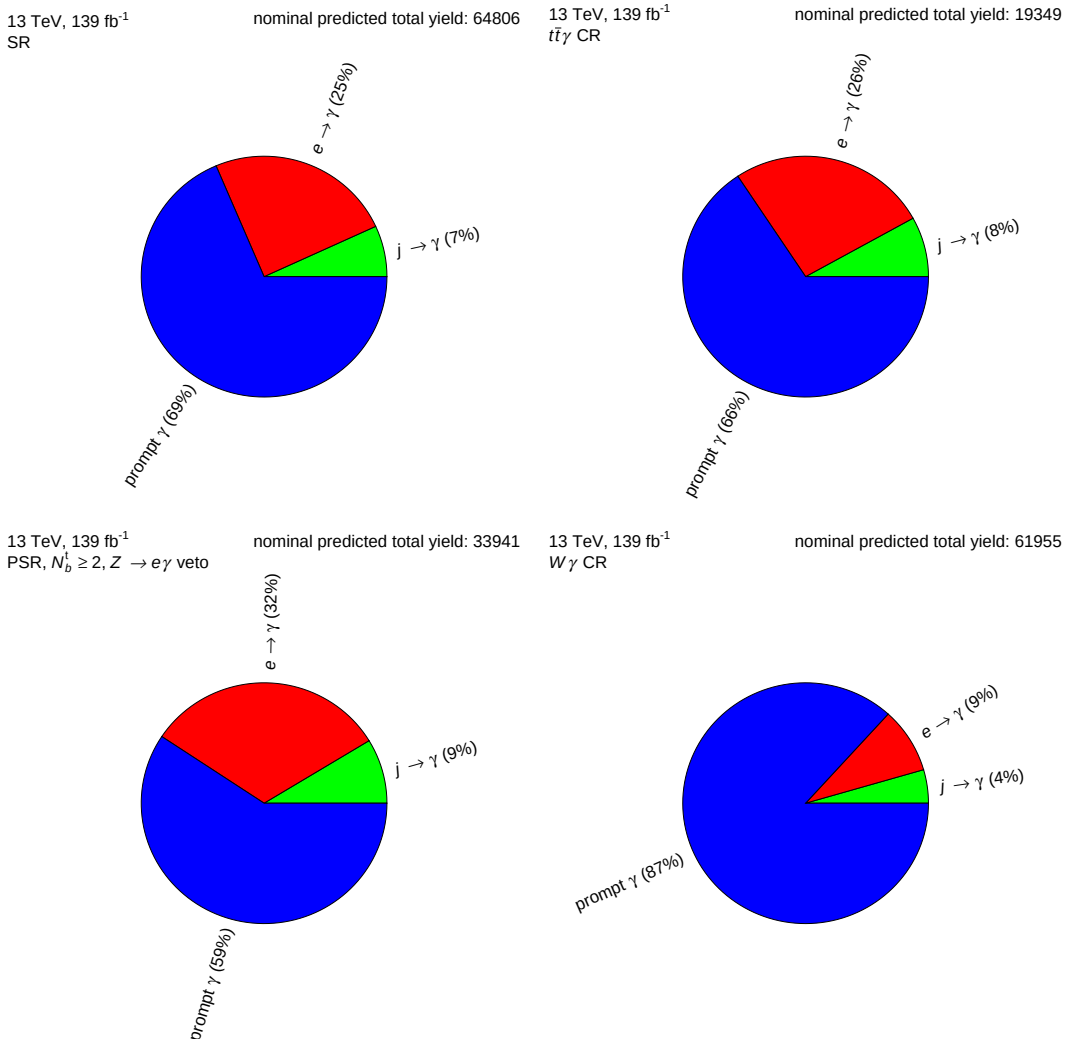


Figure 31: Predicted relative contributions of events classified as prompt-γ, $e \to \gamma$, or $h \to \gamma$ events for the SR, the $W\gamma$ CR, the $t\bar{t}\gamma$ CR, and the $N_b^{\mathrm{t}} \geq 2$ region. The sum of all expected signal and background yields is given in the top right corner.

Table 10: Overview of the b-tagging and forward-jet requirements used in the definition of the PSR, ePSR, and the different ARs, which are highlighted. The number of loose (N_b^{l}) and tight (N_b^{t}) b-tagged jets as well as the number of forward jets (N_{fj}) are used to define the requirements.

Region	b-tagging requirement	Forward-jet requirement
PSR	$N_b^{\mathrm{l}} + N_b^{\mathrm{t}} \geq 1$	$N_{\mathrm{fj}} \geq 0$
ePSR	$N_b^{\mathrm{t}} \geq 1$	$N_{\mathrm{fj}} \geq 0$
SR	$N_b^{\mathrm{t}} = 1 \,\&\, N_b^{\mathrm{l}} = 0$	$N_{\mathrm{fj}} \geq 0$
0fj SR	$N_b^{\mathrm{t}} = 1 \,\&\, N_b^{\mathrm{l}} = 0$	$N_{\mathrm{fj}} = 0$
≥ 1fj SR	$N_b^{\mathrm{t}} = 1 \,\&\, N_b^{\mathrm{l}} = 0$	$N_{\mathrm{fj}} \geq 1$
$t\bar{t}\gamma$ CR	$N_b^{\mathrm{t}} = 1 \,\&\, N_b^{\mathrm{l}} \geq 1$	$N_{\mathrm{fj}} \geq 0$
$W\gamma$ CR	$N_b^{\mathrm{t}} = 0 \,\&\, N_b^{\mathrm{l}} \geq 1$	$N_{\mathrm{fj}} \geq 0$

7.4 Fiducial cross-section measurements

A fundamental property of a particle physics process is its cross section. Measurements of this property are either performed for the entire final-state phase space or for a fiducial phase space defined by kinematic or topological restrictions of the final state. These restrictions, used to define the fiducial phase space, are typically tailored to the experimental acceptance of the corresponding objects, which reduces the size of the extrapolation of measured quantities to the fiducial phase space. Therefore, the main advantage of measurements for fiducial phase spaces is that the uncertainties associated with the theoretical modeling of contributions in the ARs (e.g., the modeling of the PS) are typically smaller relative to those in cross section measurements of the unrestricted phase space. This usually results in the measurements of fiducial cross sections being more precise. The measured values are compared to theoretical predictions to guide their development, constrain theory parameters of the SM or BSM models, and discover potential BSM physics. The gain in knowledge is limited by the respective accuracies of the measurements and predictions. In this work, fiducial cross section measurements are performed for two different phase spaces.

To measure the fiducial cross section, the number of signal events in the collected data is measured in the ARs. In Section 7.5 it is described how this measurement is performed. The measured values are unfolded into a fiducial cross section via Equation (44), where L denotes the integrated luminosity and C denotes the correction factor explained in the following. The signal events in the ARs are composed of those that fulfill the requirements of the fiducial phase space and those that do not. The respective numbers of signal events are denoted by $N(\text{sel}|\text{fid})$ and $N(\text{sel}|\text{non-fid})$, respectively. The total number of signal events in the ARs is denoted by $N(\text{sel})$. Furthermore, signal contributions that fulfill the requirements of the fiducial phase space, denoted by $N(\text{fid})$, may fail to meet the selection criteria of the ARs, for instance due to the limited detector acceptance, resolution effects, or reconstruction inefficiencies. The correction factor, defined by Equation (45), accounts for the limited selection efficiency (ε_{sel}) of events that satisfy the criteria of the fiducial phase space and for the fraction of selected events (f_{mig}) that do not, which need to be subtracted from the measured number of signal events.

$$\sigma^{\text{fid}} = \frac{N_{\text{S}}}{L\,C} \tag{44}$$

$$C = \frac{\varepsilon_{\text{sel}}}{1 - f_{\text{mig}}}, \qquad f_{\text{mig}} = \frac{N(\text{sel}|\text{non-fid})}{N(\text{sel})}, \qquad \varepsilon_{\text{sel}} = \frac{N(\text{sel}|\text{fid})}{N(\text{fid})} \tag{45}$$

Fiducial cross sections are measured for phase spaces that are typically defined either at particle level or at parton level. In contrast to parton-level phase spaces, the particle-level phase spaces include the modeling of the PS and hadronization. Particle-level objects correspond to observable final states, while parton-level objects are theoretical, and essentially non-physical, constructs. Measurements performed for particle-level phase spaces are typically more precise than measurements performed for parton-level phase spaces. This is primarily due to the theoretical signal-related uncertainties considered in the measurements being usually smaller for particle-level measurements, as the extrapolation from particle-level phase spaces to ARs is smaller. However, due to the necessary modeling of the PS and hadronization, it is more challenging to compile precise predictions for particle-level measurements than for parton-level measurements. The main advantage of parton-level measurements is that they can be compared to fixed-order theory calculations, which are typically more precise and less complex to compute. Therefore, parton-level measurements are more straightforward to reinterpret, e.g., in the context of EFTs.

One measurement is performed for a fiducial phase space that is defined at particle level and is referred to as "particle-level measurement". The respective fiducial cross section is measured for the combined $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes. The other one is performed for a parton-level fiducial phase space and is referred to as "parton-level measurement". The nominal $t\,(\to \ell\nu b\gamma)\,q$ sample (cf. Section 6) does not

simulate the $t\,(\to \ell\nu b\gamma)\,q$ process at parton level. As the hard-scattering $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes lead to the same final state, the processes interfere, and are, in principle, inseparable. A measurement of the cross section of the combined $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes represents the correct way from a physics point of view as it includes these potential effects. However, as discussed in Chapter 6 the potential interference effects are small and do not notably affect the kinematic properties, allowing for a separate treatment and investigation of these processes. Therefore, the parton-level fiducial cross section is measured for $tq\gamma$ only. This type of measurement is especially intriguing for the $tq\gamma$ process as the comparison with theory predictions allows to improve the knowledge about the top-quark–photon interaction and to constrain parameters of the SM and/or BSM theories, e.g., the EFT parameters O_{uW}^{33} and O_{uB}^{33} (cf. Section 3.3). In the following, the definition of the two fiducial phase spaces and the predictions for the respective fiducial cross sections are presented.

7.4.1 Particle-level measurement

A measurement of the cross section of the combined $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes is performed for a particle-level fiducial phase space. This phase space is defined by requirements on particle-level objects that resemble the PSR requirements on reconstructed physics objects (cf. Section 7.3). The same definition and overlap removal of particle-level objects as presented in Section 6.1 are used. The definition of the particle-level fiducial phase space is given by the following requirements:

- Exactly one lepton with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 25\,\mathrm{GeV}$
- At least one photon with $|\eta| < 2.37$ and $p_{\mathrm{T}} > 20\,\mathrm{GeV}$
- At least one b jet with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$
- At least one neutrino

Nominal predictions of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ quantities in the fiducial phase space are computed by using the NLO $tq\gamma$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples introduced in Chapter 6. Uncertainties in these predictions are retrieved by computing the differences between the nominal and alternative predictions, which are compiled by varying settings and parameters in the production of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ samples. These variations are independent of each other and are described below. If not stated otherwise, a type of variation is derived for both $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$.

The uncertainty associated with the chosen PDF set is given by the standard deviation of the alternative predictions obtained from the 100 sampled replicas of the PDF set. The values of μ_{R} and μ_{F} are each varied by factors of 0.5 and 2 and the differences of the predictions for these four alternative settings are considered as uncertainties. Samples produced using Herwig 7 instead of Pythia 8 are used for compiling alternative predictions. Herwig 7 is a commonly used alternative to Pythia 8 in the simulation of pp collisions for modeling the PS, hadronization, and UE. The cluster hadronization model is used [176, 177]. For simplicity, this type of variation is referred to as the "PS model variation". Furthermore, the Var3c variations of the parameters of the A14 tune of Pythia 8 are considered [214]. The corresponding $\pm 1\sigma$ variations are assessed via reweighting for $t\,(\to \ell\nu b\gamma)\,q$, and by using two samples that are produced with settings corresponding to the $+1\sigma$ and -1σ variations for $tq\gamma$. Moreover, alternative predictions for the $t\,(\to \ell\nu b\gamma)\,q$ process are compiled by using a sample in which the hard-scattering process is modeled with MadGraph5_aMC@NLO instead of Powheg Box v2 ("matrix element (ME) variation"). This sample includes NLO QCD contributions in the simulation of the hard-scattering process.

An additional uncertainty is assigned to the prediction of the fiducial cross section of the $t\,(\to \ell\nu b\gamma)\,q$ process accounting for the limited accuracy of the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure. As discussed in Section 6.4.2, the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample obtained by the application of this procedure is contaminated by $tq\gamma$ contributions. Furthermore, this procedure is also limited in efficiency, i.e., not all $t\,(\to \ell\nu b\gamma)\,q$ events are correctly identified as such. Therefore, the predicted fiducial cross

section might be either overestimated or underestimated due to these effects. To estimate the respective impact on the fiducial cross section, the purity of $t\,(\to \ell\nu b\gamma)\,q$ events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is computed with Equation (42) for this fiducial phase space. The obtained values of $r_\sigma^{\mathrm{LO}} = 0.22$, $f^{tq\gamma,\,\mathrm{NLO}} = 0.05$, and $\varepsilon^{t(\to\ell\nu b\gamma)q} = 0.89$ lead to an estimated $t\,(\to \ell\nu b\gamma)\,q$ purity of $P^{\mathrm{NLO}} = 0.797$. The values are obtained in the same way as described in Section 6.4.2. The value of r_σ^{LO} for this phase space is approximately half as large as the value for the phase space defined in Section 6.4.2. This is primarily caused by the stricter photon-p_{T} requirement of this phase space, as the photon-p_{T} values in $t\,(\to \ell\nu b\gamma)\,q$ events are significantly lower on average relative to those in $tq\gamma$ events. The estimated purity corresponds to an overestimation of the fiducial cross section due to the $tq\gamma$ contamination by a factor of $\varepsilon^{t(\to\ell\nu b\gamma)q}/P^{\mathrm{NLO}} \approx 1.11$. However, as discussed in Section 6.4.2, the value of P^{NLO} is a rough estimation, and therefore, a conservative uncertainty of -20% is assigned to the prediction of the fiducial cross section.

The total $+1\sigma$ and -1σ uncertainties of a quantity are given by the sum in quadrature of the individual uncertainties from variations that enhance and decrease the predicted value, respectively. The different sources of uncertainties are treated as uncorrelated across the processes except for those corresponding to the PS model variations that are treated as fully correlated.

The values of the acceptance fractions A (cf. Section 6.3) are $A_{t(\to\ell\nu b)q\gamma} = 0.1467^{+2.3\%}_{-0.5\%}$ for $tq\gamma$ events and $A_{t(\to\ell\nu b\gamma)q} = (7.16^{+48.4\%}_{-0.2\%}) \cdot 10^{-4}$ for $t\,(\to \ell\nu b\gamma)\,q$ events. These values are roughly three and five times smaller than the corresponding values for the phase space analyzed in Chapter 6, which is defined by significantly looser kinematic requirements. The dominant contribution to the $+1\sigma$ uncertainty arises from the PS model variation for $tq\gamma$ and from the ME variation for $t\,(\to \ell\nu b\gamma)\,q$.

The predicted values for the particle-level fiducial cross section are $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma} = 167^{+11}_{-9}$ fb and $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b\gamma)q} = 50^{+24}_{-10}$ fb. The $\pm1\sigma$ values of the total uncertainty of the $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma}$ prediction are dominated by the μ_{R} variations. A subdominant contribution to the $+1\sigma$ value arises from the PS model variation, which is roughly half the size of the μ_{R} variations, while the other variations are negligibly small. For $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b\gamma)q}$, the $+1\sigma$ value is dominated by the ME variation and the -1σ variation by the uncertainty accounting for the limited accuracy of the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure. A subdominant contribution to the $+1\sigma$ value arises from the PS model variation, which is about half the size of the ME variation. Other contributions to the $\pm1\sigma$ values are negligible.

The values of $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b\gamma)q}$ and $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma}$ are combined into the prediction for the particle-level fiducial cross section of the combined process, which is 217^{+27}_{-15} fb. The relative $t\,(\to \ell\nu b\gamma)\,q$ contribution is hence predicted to be $\approx 23\%$.

The values of $\varepsilon_{\mathrm{sel}}$ and f_{mig} obtained from the NLO $tq\gamma$ sample are $0.194^{+2.5\%}_{-2.5\%}$ and $0.090^{+4\%}_{-6\%}$, respectively. These values translate into a value of the correction factor of $C = 0.2135^{+2\%}_{-2\%}$. The PS model variation and the -1σ Var3c variation are the most dominant sources of uncertainties for all these quantities.

The values of $\varepsilon_{\mathrm{sel}}$ and f_{mig} obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are $0.1466^{+0.3\%}_{-16.5\%}$ and $0.187^{+20\%}_{-20\%}$, respectively. These values translate into a value of the correction factor of $C = 0.1802^{+0.2\%}_{-15.5\%}$, which agrees within the uncertainties with the value of C obtained for the NLO $tq\gamma$ sample, while the -1σ value of the uncertainty for $t\,(\to \ell\nu b\gamma)\,q$ is significantly larger. The largest sources of uncertainty are the PS model and ME variations, which are of similar absolute size for all these quantities, while other sources of uncertainties are negligible in comparison.

Appendix C provides tables that list the size of the uncertainty for all individual variations for each of the aforementioned quantities.

7.4.2 Parton-level measurement

A measurement of the fiducial cross section of the $tq\gamma$ process is performed for a parton-level phase space while treating $t\,(\to \ell\nu b\gamma)\,q$ as a background process. The parton-level fiducial phase space is defined by kinematic requirements on the parton-level final-state photon. The photon is required to fulfill $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 2.37$, and Frixione-isolation using $\Delta R = 0.2$. The two kinematic requirements are the same as for the reconstructed photon candidates in the ARs (cf. Section 7.1). Furthermore, the top quark is required to decay semileptonically.

A prediction for the parton-level fiducial cross section is computed by using the NLO $tq\gamma$ sample introduced in Chapter 6. The acceptance fraction is $A_{t(\to\ell\nu b)q\gamma} = 0.420^{+0.5\%}_{-0.4\%}$, where the uncertainties are obtained from combining the PDF, μ_{R} and μ_{F} variations. This value is about three times larger than the corresponding value for the particle-level fiducial phase space. The predicted value of the parton-level fiducial cross section is $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma} = 479^{+28}_{-37}\,\mathrm{fb}$, where the value of the uncertainty is dominated by the μ_{R} and μ_{F} variations, while the PDF uncertainty is negligible in comparison. This value is used as nominal prediction of the parton-level fiducial cross section in the statistical analysis.

The values of $\varepsilon_{\mathrm{sel}}$ and f_{mig} obtained from the NLO $tq\gamma$ sample are $0.0714^{+1.7\%}_{-4.25\%}$ and $0.0423^{+22\%}_{-1\%}$, respectively. The values of these quantities are smaller than their values for the particle-level phase space. This is because the extrapolation from the parton-level phase space to the ARs is larger. The parton-level phase space only uses criteria related to the photon and is therefore less closely defined to the ARs relative to the particle-level fiducial phase space. These values are translated into a value of the correction factor of $C = 0.075^{+2.6\%}_{-3.8\%}$. The most dominant contribution to the total uncertainties arises from the PS model and Var3c variations. The predicted C and $\varepsilon_{\mathrm{sel}}$ values are larger for the PS model variation and smaller for the Var3c variations relative to the respective values for the nominal predictions. While the corresponding absolute differences are of similar size for C, the PS model variation is half the size of the Var3c variations for $\varepsilon_{\mathrm{sel}}$. The predicted f_{mig} values are larger for the PS model variation and Var3c variations relative to those for the nominal prediction, while the uncertainty associated with the PS model variation is roughly three times larger than the Var3c uncertainties. The PDF, μ_{R}, and μ_{F} variations are small for each of these values. Appendix C provides tables that list the size of the uncertainty for all individual variations for each of the aforementioned quantities.

The measured value of $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma}$ is compared to a prediction obtained from a fixed-order calculation using MadGraph5_aMC@NLO that includes NLO QCD contributions. This fixed-order calculation is performed for the $pp \to tq\gamma$ process, i.e., without including the top-quark decay. A setup for this type of calculation is provided in Ref. [163]. With the help of the authors, the corresponding technical setup was reproduced and used for the calculation of the parton-level fiducial cross section in this work. Apart from the kinematic requirements on the photon, the same settings for SM parameters, PDF sets, and μ_{R} and μ_{F} are used for the computation. This setup is summarized in the following.

- Flavor scheme choice: The 5FS is used for the computation of the nominal result. The result using the 4FS is used for deriving an uncertainty on the prediction which is discussed below.
- Scale values: μ_{R} and μ_{R} are defined by

$$\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = \frac{1}{6}\sum_{i=t,b,\gamma} m^{i}_{\mathrm{T}}$$

  where m^{i}_{T} denotes the transverse mass of final-state particle i. One of the main features of this definition is a smaller difference between predictions computed in the 4FS and 5FS relative to those for other common choices.
- PDF choice: NNLO PDF sets from the NNPDF3.1 set are used, while the set including a photon density function based on the LUXqed parametrization is used in the 5FS calculations [240, 241].

- Uncertainties: Apart from the PDF uncertainty (denoted by "PDF") that is calculated as standard deviation obtained from the 100 replicas that are provided by the NNPDF sets, a combined uncertainty due to the choice of the scale values and the flavor scheme is assigned to the prediction (denoted by "scale+FS"). This uncertainty is given by the envelope of the nine-point variations of for both flavor schemes, i.e., the envelope is calculated from 18 scale variations in total. The nine-point variations are obtained by varying the values of μ_{R} and μ_{F} independently by factors 0.5 and 2 for each flavor scheme.

The values of the SM parameters are listed in Section 3.1 of Ref. [163]. The predictions for the fiducial cross section of the $pp \to tq\gamma$ process using the requirements on the final-state photon of the parton-level fiducial phase space computed using the 5FS ($\sigma_{tq\gamma}^{\mathrm{5FS}}$) and 4FS ($\sigma_{tq\gamma}^{\mathrm{4FS}}$) are listed below. The $\sigma_{tq\gamma}^{\mathrm{pred.}}$ value is the same as $\sigma_{tq\gamma}^{\mathrm{5FS}}$ except that the scale+FS uncertainty is listed instead of the scale uncertainty only.

$$\sigma_{tq\gamma}^{\mathrm{5FS}} = 1.588^{+7.0\%}_{-2.1\%}\,(\mathrm{scale})^{+0.4\%}_{-0.4\%}\,(\mathrm{PDF})\,\mathrm{pb} \tag{46}$$

$$\sigma_{tq\gamma}^{\mathrm{4FS}} = 1.539^{+4.1\%}_{-5.2\%}\,(\mathrm{scale})\,\mathrm{pb} \tag{47}$$

$$\sigma_{tq\gamma}^{\mathrm{pred.}} = 1.588^{+7.0\%}_{-8.1\%}\,(\mathrm{scale+FS})^{+0.4\%}_{-0.4\%}\,(\mathrm{PDF})\,\mathrm{pb} \tag{48}$$

The prediction for $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma}$ is obtained by scaling $\sigma_{tq\gamma}^{\mathrm{pred.}}$ with $\mathcal{BR}(t \to \ell\nu b) = 32.46\%$. This is the same value as used in the simulation of the top-quark decay via MadSpin for the nominal NLO $tq\gamma$ sample. Therefore, the final prediction is $\sigma^{\mathrm{fid}}_{t(\to\ell\nu b)q\gamma} = 515^{+36}_{-42}$ fb. This value and the value predicted by the NLO $tq\gamma$ sample agree well within the uncertainties, while the former is roughly 8% larger.

7.5 Strategy and technical setup for the statistical analysis

A statistical analysis is performed in the ARs to measure the number of background and signal events in the collected data, as well as to determine their associated uncertainties and to quantify the agreement between the predictions and the collected data, a process known as "hypothesis testing". In this work, BPL fits are used for these tasks. The predictions of the signal and background contributions are fitted simultaneously to the collected data in the SRs and CRs, considering systematic uncertainties as nuisance parameters (NPs) in the fit model. The numbers of signal events obtained from the BPL fits are translated into the measured fiducial cross sections. Furthermore, the results of the fit from the parton-level measurement are used to test for the discovery of the $tq\gamma$ process. The fit model, along with the technical setups for the BPL fit and the discovery test, are discussed in the following. The framework TRExFitter is used, which is provided and maintained by the ATLAS collaboration [242]. It employs the HistFactory software package [243], based on the RooStats [244] and RooFit [245] libraries of the ROOT software framework [246].

7.5.1 Binned-profile-likelihood fit

This section describes the specific model and the technical details of the BPL fits. The fit model is a likelihood function L, which is built from binned distributions obtained from the signal and background predictions. It is fitted to the corresponding binned distributions of the collected data by performing a maximum-likelihood estimate. Equation (49) displays the function, where bold symbols represent grouped sets of parameters or of event numbers in the collected data. The individual components of L are discussed in the following.

$$\begin{aligned} L(\boldsymbol{n}^{\mathrm{data}}|\boldsymbol{\mu},\boldsymbol{\theta},\boldsymbol{\gamma}) = &\prod_{r}^{N^{r}} \prod_{b}^{N^{r,b}} \mathcal{P}\left(n^{r,b}|m^{r,b}(\boldsymbol{\mu},\boldsymbol{\theta},\boldsymbol{\gamma^{r,b}})\right) \\ &\times \prod_{y}^{N_y} \mathcal{P}\left(k_y^{r,b}|\gamma_y^{r,b}\tau_y^{r,b}\right) \prod_{j}^{N_j} \mathcal{G}\left(\theta_j\right) \end{aligned} \tag{49}$$

$$L^{\text{stat.}}(\boldsymbol{n}^{\text{data}}|\boldsymbol{\mu}) = \prod_{r}^{N^r} \prod_{b}^{N^{r,b}} \mathcal{P}\left(n^{r,b}|m^{r,b}(\boldsymbol{\mu})\right) \tag{50}$$

The probability to observe $n^{r,b}$ events in collected data, when $m^{r,b}(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma^{r,b}})$ events are predicted, is modeled for each bin b of each AR r by a Poisson p.d.f. ($\mathcal{P}$). The number of regions considered in the fit model is denoted by N^r, while $N^{r,b}$ denotes the number of bins considered in a specific region r. The predicted number of events $m^{r,b}$ is the sum of the predicted signal and background yields and depends on two types of parameters, the unconstrained parameters ($\boldsymbol{\mu}$) and the constrained parameters ($\boldsymbol{\theta}$ and $\boldsymbol{\gamma^{r,b}}$). Both types of parameters are adjusted by the fit so that L is maximized. A penalty term is introduced into L for each constrained parameter, penalizing it for deviating significantly from a priori defined boundaries. Conversely, the unconstrained parameters are adjusted without any penalty by the fit and hence these parameters are also referred to as free parameters of the fit. All parameters are considered to be uncorrelated with each other. Equation (50) shows the version of the likelihood function ($L^{\text{stat.}}$) in which only statistical uncertainties are considered. In this case, the predicted number of events $m^{r,b}$ only depends on the unconstrained $\boldsymbol{\mu}$-parameters.

The $\boldsymbol{\mu}$-parameters adjust the predicted contributions of individual processes in all bins by the same relative amount. In this work, three such parameters are considered in L. Two are assigned to the most dominant prompt-photon background contributions, the $t\bar{t}\gamma$ and $W\gamma$+jets processes, and are denoted by $\mu_{t\bar{t}\gamma}$ and $\mu_{W\gamma}$, respectively. The other one is assigned to the predicted signal contribution. In the parton-level measurement it is assigned to the $tq\gamma$ prediction and is denoted by $\mu_{tq\gamma}$, while in the particle-level measurement it is assigned to the sum of the $tq\gamma$ and $t\,(\rightarrow \ell\nu b\gamma)\,q$ predictions and is denoted by μ_S. The parameters $\mu_{tq\gamma}$ and μ_S are also referred to as signal strength parameters.

The constrained $\boldsymbol{\theta}$-parameters (NPs), account for systematic variations of the signal and background predictions. The systematic variations are provided as binned distributions obtained from the respective alternative predictions, which correspond to the $\pm1\sigma$ variations. Gaussian p.d.f.s ($\mathcal{G}\left(\theta_j\right)$) are used to constrain each NP (θ_j). The mean is given by $\theta_j = 0$ and corresponds to the nominal prediction. Furthermore, the p.d.f.s have unit variance, so that $\theta_j = \pm1$ correspond to the respective $\pm1\sigma$ variations. The Gaussian p.d.f.s are hence given by $\mathcal{G}(\theta_j) = 1/\sqrt{2\pi}\exp(-\theta_j^2/2)$.

The constrained $\boldsymbol{\gamma}$-parameters account for the limited number of events in the samples used to construct the predicted signal and background distributions, and are constructed following the prescription provided by Barlow and Beeston in Ref. [247]. A multiplicative scale factor is assigned to the predicted number of events for each bin considered in the fit, referred to as γ-factor. However, introducing individual γ-factors for all signal and background samples would result in an impractically large number of parameters. Individual γ-factors are only assigned to the $tq\gamma$ and $t\,(\rightarrow \ell\nu b\gamma)\,q$ processes, denoted by $\gamma^{r,b}_{tq\gamma}$ and $\gamma^{r,b}_{t(\rightarrow\ell\nu b\gamma)q}$, respectively. A third set of γ-factors is assigned to the predicted sum of all other contributions ($\gamma_B^{r,b}$). The value of each γ-factor is regulated by a Poisson p.d.f., $\mathcal{P}\left(k_y^{r,b}|\gamma_y^{r,b}\tau_y^{r,b}\right)$, where $k_y^{r,b}$ is given by $(\nu_y^{r,b}/\delta_y^{r,b})^2$. Here, $\nu_y^{r,b}$ is the number of events in the sample and $\delta_y^{r,b}$ is the corresponding uncertainty. The subscript y is a placeholder for the process(es) to which the $\gamma_y^{r,b}$ parameters belong and N_y denotes the number of processes for which individual γ-factors are assigned. The value $\tau_y^{r,b}$ is also given by $(\nu_y^{r,b}/\delta_y^{r,b})^2$. However, this value is fixed, i.e., if pseudo-experiments were constructed, $k_y^{r,b}$ would fluctuate around $\tau_y^{r,b}$, while the latter remains fixed.

Equation (51) shows the parameterized predicted total number of events in bin b of region r. For the particle-level measurement, the equation is used with the red parentheses and μ_S. For the parton-level measurement, the equation is used without the red parentheses and with $\mu_{tq\gamma}$. The final term is a sum over the N_B remaining background processes (B_i).

$$m^{r,b}(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma^{r,b}}) = \mu_{tq\gamma\,(S)}(\boldsymbol{\theta})\left(\gamma^{r,b}_{tq\gamma}\eta^r_{tq\gamma}(\boldsymbol{\theta})\sigma_{tq\gamma}(\boldsymbol{\theta}) + \gamma^{r,b}_{t(\rightarrow\ell\nu b\gamma)q}\eta^r_{t(\rightarrow\ell\nu b\gamma)q}(\boldsymbol{\theta})\sigma_{t(\rightarrow\ell\nu b\gamma)q}(\boldsymbol{\theta})\right)$$

$$+ \gamma_B^{r,b} \left(\mu_{t\bar{t}\gamma} \eta_{t\bar{t}\gamma}^r(\boldsymbol{\theta}) \sigma_{t\bar{t}\gamma}^{r,b}(\boldsymbol{\theta}) + \mu_{W\gamma} \eta_{W\gamma}^r(\boldsymbol{\theta}) \sigma_{W\gamma}^{r,b}(\boldsymbol{\theta}) + \sum_{B_i}^{N_B} \eta_{B_i}^r(\boldsymbol{\theta}) \sigma_{B_i}^{r,b}(\boldsymbol{\theta}) \right) \tag{51}$$

The effect of the NPs on the predicted contribution of an individual process p is divided into a normalization component $\eta_p^r(\boldsymbol{\theta})$ and a shape component $\sigma_p^{r,b}(\boldsymbol{\theta})$. Both only depend on $\boldsymbol{\theta}$ and are consequently fully correlated. The normalization component alters the prediction of the overall contributions of the affected processes to the respective region, while the shape component alters the relative contributions in each bin of the respective region and not the prediction of the overall contribution. Both the normalization and shape components are composed of the normalization and shape components of the individual NPs. As the systematic variations are provided as binned distributions corresponding to the $\pm 1\sigma$ variations, an interpolation procedure is applied which defines the impact of the normalization and shape components inside and outside the $\pm 1\sigma$ boundaries. This interpolation procedure is applied to each NP. For the shape components a piecewise linear interpolation is chosen, while for the normalization components a polynomial interpolation and exponential extrapolation are chosen. The respective parameterizations are given in the documentation of the HistFactory software package [243].

The parameter space of L is typically characterized by a high dimensionality and hence the maximization procedure is a complex task. In practice, the negative log likelihood $-\ln L$ is minimized to find the set of parameters that best describes the distributions in collected data. The minimization of the constrained parameters is based on the MINUIT framework [248] employing the MIGRAD minimization technique, which uses the Davidon-Fletcher-Powell approach [249–251]. This technique leads to symmetric posterior (post-fit) uncertainties of the parameters of L. For the unconstrained parameters the MINOS technique is employed that may lead to asymmetric post-fit uncertainties [248].

The fitted values of $\mu_{tq\gamma}$ and μ_S are translated into the measured fiducial cross sections by Equation (52), where $N_X^{\text{meas.}}$ and $N_X^{\text{pred.}}$ are the fitted number of events and the predicted number of events, respectively.

$$\sigma_X = \frac{N_X^{\text{meas.}}}{L \cdot C} = \frac{\mu_X \cdot N_X^{\text{pred.}}}{L \cdot C} = \mu_X \sigma_X^{\text{pred.}} \tag{52}$$

The uncertainties in L and C are considered in the measurement as described in Section 7.6. Hence, the measured fiducial cross section is the product of the measured value of μ_X and the predicted fiducial cross section ($\sigma_X^{\text{pred.}}$), where the relative uncertainty of the measured fiducial cross section is given by the relative uncertainty of μ_X.

7.5.2 Treatment of systematic uncertainties

The NPs account for systematic variations of the signal and background predictions. These systematic variations are provided either as one-sided variations, i.e., one alternative distribution is provided, treated as $+1\sigma$ variation, or as two-sided variations, i.e., two alternative distributions are provided, which correspond to the $\pm 1\sigma$ variations. For one-sided variations, the $+1\sigma$ variation is mirrored with respect to the nominal prediction to construct the -1σ variation. The -1σ variation, $\delta_{i,r,b}^-$, for the systematic variation i in bin b of region r is therefore given by $\delta_{i,r,b}^- = 2\,\delta_{i,r,b}^0 - \delta_{i,r,b}^+$, where $\delta_{i,r,b}^0$ and $\delta_{i,r,b}^+$ are the nominal prediction and the prediction of the $+1\sigma$ variation, respectively. For two-sided variations, the absolute differences of the $\pm 1\sigma$ variations from the nominal prediction may be different, so that the systematic uncertainty impacts the prediction in an asymmetric fashion. A two-sided symmetrization technique is applied to these variations to center the impact of the systematic uncertainty around the nominal value. Therefore, the provided $\delta_{i,r,b}^{\pm}$ values are adjusted to symmetrized values $\hat{\delta}_{i,r,b}^{\pm}$ according to Equation (53). Unless stated otherwise, these symmetrization techniques are applied to all considered systematic variations.

$$\hat{\delta}_{i,r,b}^{\pm} = \pm \frac{\delta_{i,r,b}^+ - \delta_{i,r,b}^-}{2} + \delta_{i,r,b}^0 \tag{53}$$

The majority of the systematically varied signal and background predictions are constructed by using MC samples and are hence constructed from limited number of simulated events. Therefore, the respective distributions are subject to statistical fluctuations. These can lead to the under- or overestimation of systematic impacts in individual bins. To mitigate this effect, a smoothing algorithm is applied to the distributions of the systematic variations. The first step of this algorithm is to compute a smoothing threshold δM_b for each bin b. This threshold is either computed by Equation (54), assuming that the statistical uncertainties of the nominal prediction and the systematic variations are independent, or by Equation (55), assuming that the statistical uncertainties are fully correlated. The variables δN_b and δS_b denote the statistical uncertainty of the nominal prediction and the systematic variation, respectively.

$$\delta M_b = \sqrt{\delta S_b^2 + \delta N_b^2} \tag{54}$$

$$\delta M_b = \max(\delta S_b, \delta N_b) \tag{55}$$

In a second step, the absolute difference $x_{b-1,b}$ between the relative systematic variation of bin b and the previous bin $b-1$ is calculated by Equation (56).

$$x_{b-1,b} = \left| \frac{S_b - N_b}{N_b} - \frac{S_{b-1} - N_{b-1}}{N_{b-1}} \right| \tag{56}$$

A relative uncertainty $\delta x_{b-1,b}$ is assigned to each of these values by Equation (57).

$$\delta x_{b-1,b} = \sqrt{\left(\frac{\delta M_b}{N_b}\right)^2 + \left(\frac{\delta M_{b-1}}{N_{b-1}}\right)^2} \tag{57}$$

The algorithm then proceeds to search for bins that satisfy $x_{b-1,b} < \delta x_{b-1,b}$. If at least one such bin is present, the bin (b) with the largest $\delta x_{b-1,b}/x_{b-1,b}$ is merged with the respective previous bin $(b-1)$. The algorithm is repeated until no such bins are present in the distribution. The last step is to apply a smoothing procedure, referred to as "353QH twice" algorithm, which produces a smooth distribution for the merged bins [252]. This algorithm is applied four times when the statistical uncertainties are treated as uncorrelated and once when the statistical uncertainties are treated as correlated. The smoothing algorithm is applied independently to the $\pm 1\sigma$ variations. If the systematic uncertainty is provided as two-sided variation, symmetrization is applied after the smoothing algorithm. For one-sided variations, the order is reversed.

Each NP that is considered in the fit enhances the dimensionality of the minimization procedure. However, individual NPs might impact the predictions in an insignificant fashion and can therefore be neglected in the fitting procedure. This reduction of the dimensionality of the fit, which is referred to as pruning, saves computational time and improves the stability of the fit. If the impact of the normalization component of an individual NP on η_p^r is smaller than 0.1%, this component is neglected for the corresponding region and process. If the impact of its shape component on $\sigma_p^{r,b}$ is less than 0.1% for all bins b in region r for process p, this component is neglected for this region for process p. Therefore, the pruning procedure might either remove the shape component, the normalization component, or both components of an individual NP for a process p in region r.

7.5.3 Discovery test

To quantify the compatibility of a hypothesized model, which is given by a fixed set of specific parameters of L, with the collected data, hypothesis testing is performed. Two hypotheses are typically of special interest: the background-only (b-only) hypothesis, constructed by fixing the signal strength parameters to zero, and the signal-plus-background (s+b) hypothesis, constructed by fixing at least one signal strength parameter to a value > 0. To test for the discovery of $tq\gamma$, the compatibility of the b-only hypothesis is probed using the parton-level setup, i.e., $t\,(\to \ell\nu b\gamma)\,q$ is treated as background process.

The test statistic for the discovery test is constructed from the profile likelihood ratio λ given by Equation (58). The numerator is the conditional likelihood, which is maximized for a fixed value $\mu_{tq\gamma} = \mu^h_{tq\gamma}$, while the denominator is the unconditional likelihood, which is maximized without fixing any of the parameters. The parameters in both likelihoods are the same, but they are marked with a hat in the conditional likelihood to indicate that their fitted values differ from those in the unconditional likelihood. Note that $\mu_{tq\gamma}$ is not included in $\hat{\boldsymbol{\mu}}$.

$$\lambda(\mu^h_{tq\gamma}) = \frac{L\left(\mu^h_{tq\gamma}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}}\right)}{L\left(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma}\right)} \tag{58}$$

The test statistic q_0 used for probing the $tq\gamma$ discovery is given by Equation (59).

$$q_0 = \begin{cases} -2\ln(\lambda(\mu^h_{tq\gamma} = 0)), & \mu_{tq\gamma} \geq 0 \\ +2\ln(\lambda(\mu^h_{tq\gamma} = 0)), & \mu_{tq\gamma} < 0. \end{cases} \tag{59}$$

To quantify the level of agreement of the b-only hypothesis and the collected data, the p-value is calculated by Equation (60). The function f represents the p.d.f. of the test statistic q_0 and q_0^{obs} denotes the value of q_0 obtained by fitting the conditional and unconditional likelihoods to the collected data and evaluating Equations (58) and (59).

$$p = \int_{q_0^{\text{obs}}}^{\infty} f\left(q_0|0, \hat{\theta}(\mu_{tq\gamma} = 0)\right) dq_0 \tag{60}$$

The p-value is translated into the significance Z, which is expressed in units of the standard deviation (σ)

$$Z = \Theta^{-1}(1 - p), \tag{61}$$

where Θ^{-1} is the inverse of the cumulative distribution for a unit Gaussian.

In the large sample size limit, as is the case in this work, the asymptotic approximation can be applied [253–255] and Z is expressed by Equation (62).

$$Z = \begin{cases} \sqrt{q_0}, & \mu_{tq\gamma} \geq 0 \\ -\sqrt{-q_0}, & \mu_{tq\gamma} < 0. \end{cases} \tag{62}$$

By convention, $Z \geq 5\sigma$ is a mandatory criterion for rejecting the b-only hypothesis and establishing the discovery of a physics process.

7.6 Systematic uncertainties

To quantify the agreement of the prediction with collected data in the ARs, various types of uncertainties are considered as NPs in the BPL fits. Apart from statistical uncertainties, which arise from the limited number of events in a dataset, there are various sources of systematic uncertainties. These are mainly classified into two types, the experimental uncertainties and the modeling uncertainties. The former includes any uncertainty that is related to the modeling of the response of the ATLAS detector, such as uncertainties in particle detection and reconstruction efficiency, energy scale and resolution calibrations, and particle-identification and -isolation efficiencies, as well as the limited knowledge of pile-up conditions and luminosity. The latter comprises uncertainties that are related to the modeling of physics processes, such as those of the values of SM parameters, parameter settings in the generation of MC samples, or the choice of MC event generators. Due to the diversity of systematic uncertainties, a plethora of different sources can be considered. However, it is impractical to consider all possible

individual uncertainties. Therefore, it is necessary to identify those that have a significant impact on the measurement, while ensuring that no uncertainty that could significantly affect the assessment of the agreement between the prediction and the collected data is omitted.

The considered systematic uncertainties and the corresponding NPs are discussed below. A reference tag is assigned to the NPs that appear in the following sections. Unless stated otherwise, uncertainties are provided as two-sided variations to the setup of the BPL fit.

7.6.1 Experimental uncertainties

Each type of experimental uncertainty is treated as fully correlated across the different ARs and across the different processes that are modeled by MC simulations. None of the experimental uncertainties are considered for the prediction of the fake-lepton background, as its estimation is based on reweighted events from collected data. All other processes are modeled via MC simulation.

Luminosity The considered uncertainty of the integrated luminosity is 1.7% [123], affecting the predicted overall contribution to all ARs of all processes by this amount. The uncertainty of the latest measurement of the integrated luminosity is 0.83% [173]. This improvement in the uncertainty has no significant impact on the precision of the measurement results. The corresponding NP is denoted by "Luminosity".

Pile-up reweighting Each event in the simulation is overlaid with pile-up interactions. Therefore, a pile-up weight is assigned to each event in simulation, which ensures that the pile-up distribution in simulated data agrees with the distribution in collected data (cf. Figure 7). Two alternative sets of weights are provided that are used to construct the distributions corresponding to the $\pm 1\sigma$ uncertainties via reweighting. These alternative sets account for the uncertainty in the ratio of the predicted and measured cross sections of inelastic pp interactions [124]. The corresponding NP is denoted by "Pile-up SF".

Leptons and Photons The uncertainties of the simulation-to-data SFs associated with the isolation and identification algorithms for electrons, muons, and photons, the reconstruction algorithms for muons and electrons, the single-lepton triggers, and the primary-vertex-association requirements applied to muon tracks are considered (cf. Section 4.3). The individual sources of uncertainties associated with electrons and photons are considered in the statistical analysis via one NP for each type of uncertainty corresponding to the total uncertainty of the SF values. In contrast, those associated with muons are considered via two NPs, which correspond to the statistical and systematic uncertainties of the corresponding SF values. The NPs corresponding to the variations of the photon-identification and photon-isolation SFs are denoted by "Photon identification SF" and "Photon efficiency isolation SF", respectively.

To account for variations in the sagitta-bias correction, the momentum of muons in simulation is varied according to the uncertainties of this correction [193]. Two variations are considered, which are associated with the residual difference between simulated and collected data, and with the uncertainty in the procedure for the correction itself. Furthermore, the muon-momentum calibration in simulation is varied by taking into account variations of the ID-track resolution, the MS-track resolution, and a variation of the momentum-scale correction factor for combined muons.

As outlined in Section 4.3, the algorithms used to correct the energy scale and the energy resolution are the same for electrons and photons [191]. Therefore, uncertainties of these calibrations are treated as fully correlated between electrons and photons, i.e., the energy scale and energy resolution are each varied simultaneously for electrons and photons. The uncertainties of the energy-resolution and energy-scale calibrations are considered via one NP each.

Jets and missing transverse momentum Uncertainties associated with jets arise from the efficiency of jet identification based on the JVT variable, as well as from the JES and JER. Uncertainties of the

associated SF values include the statistical uncertainty and a 30% uncertainty on the estimation of the residual contamination from pile-up jets after pile-up suppression [199]. These two uncertainties are combined into a single uncertainty that is considered as $\pm 1\sigma$ variation of the respective NP.

The JES and its uncertainties were derived by combining information from test-beam data, LHC collision data, and simulation [198]. The uncertainties are grouped based on the *Category reduction* scheme, which results in a set of 32 NPs. However, since only samples produced using the full GEANT4 detector simulation are considered for the nominal signal and background predictions, two variations corresponding to calibrations using AFII simulations are neglected. Furthermore, the uncertainty due to the flavor composition is decorrelated into four NPs. These affect the predictions for the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$, $t\bar{t}\gamma$, $W\gamma$+jets, and all other background processes, respectively. The four NPs are referred to as "JES flavor comp. X", where X denotes the processes that are affected by the specific NP. Therefore, 33 NPs for JES are considered in total.

About 120 NPs are considered in the derivation of the JER corrections. A simplified uncertainty scheme is used in this analysis to account for uncertainties in the calibration of the JER. The NPs are reduced to seven uncorrelated NPs via eigenvector decomposition, which are denoted by "JER effective NPY", where Y is an integer number from 1 to 7. An additional uncertainty is derived that accounts for the difference in resolution between collected and simulated data in phase spaces, where the resolution is better in data than in simulation. The corresponding NP is denoted by "JER effective NP rest". The JER uncertainties are provided as one-sided variations.

Missing transverse energy As $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from the selected physics objects, it is affected by uncertainties associated with those objects that impact their reconstructed four-momenta, such as uncertainties related to the energy resolution of electrons and photons. Therefore, $E_{\mathrm{T}}^{\mathrm{miss}}$ is recalculated for each of these uncertainties. Additional uncertainties are assigned to the TST that are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ values [206]. Three NPs are considered, one accounts for the uncertainty in the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale, while the other two account for resolution uncertainties.

***b*-tagging algorithm** As the SRs and CRs are defined by requirements on multiple WPs of the DL1r b-tagging algorithm, the pseudo-continuous calibration is used (cf. Section 4.3). The responses of b jets, c jets, and LF jets to the b-tagging algorithm are separately calibrated by applying data-to-simulation SFs as a function of the jet p_{T}. Systematic variations of the SFs are obtained via eigenvector decomposition. These are fully correlated across the p_{T} spectrum and treated as uncorrelated with one another in the statistical analysis. The uncertainties of the SFs are considered by 45 NPs for b jets, and 20 NPs each for c jets and LF jets [200, 256, 257].

7.6.2 Modeling uncertainties

Several types of modeling uncertainties affect the predictions for processes modeled via MC simulation. However, considering all such uncertainties for every process is impractical as it leads to a highly complex fit setup. Therefore, only uncertainties that are expected to have a significant impact on the measurements are considered. As a consequence, the number of different uncertainty types that are considered in the statistical analysis is larger for processes with a large predicted contribution to the ARs, like the $t\bar{t}\gamma$ process, than for those with small predicted contributions, like the Z+jets process.

In the following, the considered modeling uncertainties are discussed. While some types of uncertainties are common and derived in the same fashion for several processes like those associated with the choices of the μ_{R} and μ_{F} values and PDF sets, other uncertainties are specific for each process. Unless stated otherwise, each uncertainty is considered in the statistical analysis by one NP, i.e., the respective uncertainties are treated as fully correlated across all different bins and ARs.

Scale and PDF variations The uncertainties due to the chosen μ_{R} and μ_{F} values and PDF set are derived in the same way as described in Section 7.4.1. All of these uncertainty types are considered for the $t\,(\to \ell\nu b\gamma)\,q$, $tq\gamma$, $t\bar{t}\gamma$ (prod), $t\bar{t}\gamma$ (dec), $t\bar{t}$, $W\gamma$+jets, and $Z\gamma$+jets processes, while only the

uncertainties associated with μ_R and μ_F are considered for the tW process. One individual NP is assigned to each process for each considered type of uncertainty. The NPs that correspond to the μ_R and μ_F variations are denoted by "X μ_R" and "X μ_F", respectively, where X denotes the process that a specific NP affects.

tqγ The alternative prediction obtained from using Herwig 7 instead of Pythia 8 ("$tq\gamma$ PS model") and the Var3c variations ("$tq\gamma$ ISR/FSR") are considered as uncertainties in the statistical analysis. The former is provided as one-sided variation. Further details on these variations are given in Section 7.4.

t (→ ℓνbγ) q The alternative prediction obtained from using Herwig 7 instead of Pythia 8, the variation of the MC event generator ("$t\,(\to \ell\nu b\gamma)\,q$ ME generator"), and the Var3c variations are considered as uncertainties in the statistical analysis. The former two are provided as one-sided variations. Further details on these variations are given in Section 7.4. An additional uncertainty is assigned to the predicted $t\,(\to \ell\nu b\gamma)\,q$ distributions by comparing the nominal $t\,(\to \ell\nu b\gamma)\,q$ prediction with the $t\,(\to \ell\nu b\gamma)\,q$ prediction obtained from the LO inclusive sample using the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure (cf. Chapter 6). This uncertainty is provided as one-sided variation and denoted by "$t\,(\to \ell\nu b\gamma)\,q$ LO vs. NLO". The normalization component of this uncertainty is not considered in the BPL fits. A conservative uncertainty of 30% is assigned to the overall predicted $t\,(\to \ell\nu b\gamma)\,q$ contribution to account for the limitations in the simulation of the $t\,(\to \ell\nu b\gamma)\,q$ contribution due to the $tq\gamma$ contamination in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample (cf. Section 6.4). For this uncertainty, one individual NP is assigned for each of the ARs. The corresponding NPs are denoted by "$t\,(\to \ell\nu b\gamma)\,q$ fraction X", where X indicates the specific AR that is affected.

tqγ+t (→ ℓνbγ) q The $tq\gamma$ variations are normalized to the nominal predicted fiducial parton-level cross section in the parton-level measurement. The $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ variations are normalized to the respective nominal predicted fiducial particle-level cross section in the particle-level measurement. This normalization ensures that uncertainties in the predicted fiducial cross sections are not included and only those related to the correction factor C are considered in the statistical analysis.

tt̄γ Differences between the nominal predictions and alternative predictions obtained by using Herwig 7 instead of Pythia 8, as well as those from the Var3c variations, are considered as uncertainties. The former differences are provided as one-sided variations and are decorrelated across the ARs by assigning an individual NP for each AR, denoted as "$t\bar{t}\gamma$ PS model X", where X denotes the respective region. These two types of uncertainties are treated as fully correlated across the $t\bar{t}\gamma$ (prod) and $t\bar{t}\gamma$ (dec) predictions. As discussed in Section 7.2, the prediction of the overall $t\bar{t}\gamma$ (dec) contribution is adjusted using a k-factor based on the fiducial cross section reported in Ref. [226]. An alternative k-factor based on the fiducial cross section reported in Ref. [258] deviates by approximately 18% from the nominal. This deviation is included in the statistical analysis as an uncertainty on the prediction of the overall $t\bar{t}\gamma$ (dec) contribution. As the k-factor may vary across ARs, one individual NP is assigned for each AR. These NPs are denoted by "$t\bar{t}\gamma$ (rad. dec.) k-factor X", where X indicates the specific AR that is affected.

tt̄ The difference between the nominal prediction and an alternative prediction obtained by setting the h_{damp} parameter in the production of the $t\bar{t}$ sample to $3\,m_t$ instead of $1.5\,m_t$ is considered as uncertainty ("$t\bar{t}$ h_{damp}"). Furthermore, an uncertainty due to using Powheg Box v2 for the production of the sample is derived by compiling an alternative prediction using MadGraph5_aMC@NLO as the MC event generator instead ("$t\bar{t}$ ME generator"). Additionally, the alternative prediction obtained from using Herwig 7 instead of Pythia 8 and the Var3c variations are considered as uncertainties. Apart from the Var3c variations, the aforementioned $t\bar{t}$ variations are provided as one-sided variations. Independent NPs associated with the PS model variation are assigned to $e \to \gamma$, $h \to \gamma$, and prompt-γ contributions. Moreover, the impact of the PS model variation on the $h \to \gamma$ contributions is included in the statistical analysis via an individual NP for each of the ARs. These NPs are denoted by "$t\bar{t}$ PS model $h \to \gamma$ X", where X indicates the specific AR that is affected. Another NP ("$t\bar{t}$ σ") is introduced, which accounts for the uncertainty of the $t\bar{t}$ production cross section of $\approx 6\%$ [138, 182, 259–261].

***Wγ*+jets** The production of W bosons in association with a photon and HF jets is less well understood. Therefore, an uncertainty of 30% is assigned to the fraction of $W\gamma$+jets events with a genuine b jet. Three individual NPs are assigned for this uncertainty, where one is associated with both SRs, one with the $t\bar{t}\gamma$ CR, and one with the $W\gamma$ CR. The NPs are denoted by "$W\gamma$+jets b-Hadron norm. X", where X indicates the specific AR that is affected. This is motivated by the difference in the b-tagging requirements used for the different ARs, which impacts the fraction of $W\gamma$+jets events with a genuine b jet.

tW An uncertainty is assigned to the choice of the procedure to remove $t\bar{t}$ contributions in the nominal tW sample that overlap with contributions simulated in the $t\bar{t}$ sample. It is derived by compiling an alternative prediction by using the diagram subtraction (DS) scheme in the production of the tW sample and by comparing it to the nominal prediction ("tW DS scheme"). Another alternative prediction is compiled by using MadGraph5_aMC@NLO instead of Powheg Box v2 as MC event generator in the production of the tW sample ("tW ME generator"). Furthermore, a variation is obtained by using Herwig 7 instead of Pythia 8 in the production of the sample. This variation is decorrelated into three NPs, which correspond to the $e \to \gamma$ ("tW PS model $e \to \gamma$"), $h \to \gamma$ ("tW PS model $h \to \gamma$"), and prompt-γ ("tW PS model prompt γ") contributions, respectively. All these types of tW uncertainties are provided as one-sided variations. An additional NP is introduced which accounts for the uncertainty of the cross section of 5.3% [149].

Other For other processes, only uncertainties in the predicted overall contribution are considered. Conservative ad-hoc uncertainties of 30% and 50% are assigned to the predicted number of $Z\gamma$+jets ("$Z\gamma$+jets σ") and VV events, respectively. This choice is motivated by the less well understood production of $Z\gamma$ and VV events in association with HF jets. An uncertainty of 5.3% on the predicted cross section is assigned to both the s-channel and the t-channel single-top-quark production processes [262, 263]. The latter is only assigned to the predictions from the t-channel sample and not those from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. Furthermore, a 5% uncertainty is assigned to each of the V+jets processes [264].

Data-driven background estimations Data-driven estimations are performed for the $e \to \gamma$, $h \to \gamma$, and fake-lepton backgrounds. These are discussed in Section 8 along with the treatment of the respective uncertainties in the statistical analysis.

8 Data-driven background estimations

The $e \to \gamma$, $h \to \gamma$, and fake-lepton contributions are significant background contributions to the ARs. While there are several established methods to estimate these contributions, it is challenging to provide such estimations centrally, as their underlying composition is difficult to model via simulation and depends strongly on the physics-object definitions and event-selection criteria. Therefore, common methods are individually adapted to this analysis. The data-driven estimation techniques employed for the fake-lepton and the fake-photon ($e \to \gamma$ and $h \to \gamma$) backgrounds in this work and the respective results are discussed in this chapter.

8.1 Estimation of background contributions with fake leptons

Physics studies that examine final states with prompt leptons using pp-collision data typically face large background contributions, where detector signatures are incorrectly identified as those of prompt leptons. These contributions can be mainly classified into two categories. One category consists of events with non-prompt leptons, which are genuine leptons that do not directly originate from the hard-scattering process or from the decay of a massive boson, but, for instance, from the decay of hadrons. The other category consists of events with misreconstructed lepton candidates that are not genuine leptons but are, for instance, hadrons. For simplicity, both types of contributions are commonly referred to as "fake leptons" hereafter. A detailed overview of the different sources of fake leptons, which typically have significant contributions in physics studies using pp-collision data taken with the ATLAS detector, is provided in Ref. [265]. A brief summary of these different sources is given in the following.

Common sources of fake leptons are non-prompt leptons arising from EW decays of b hadrons, c hadrons, LF hadrons, and tau leptons. Another fake-muon contribution arises from electrically charged hadrons that are not stopped in the calorimeters and enter the MS. However, this contribution is typically negligibly small. In contrast, electrically charged hadrons that mimic the signature of prompt electrons are typically a significant source of fake electrons. Furthermore, fake-electron contributions arise from photons that convert into electron-positron pairs. These photons are either prompt photons ("prompt γ conversions") or non-prompt photons that are radiated from a hadron or produced in a hadron decay, where the most dominant contribution arises from $\pi^0 \to \gamma\gamma$ decays. Dalitz decays of pions lead to a signature that is similar to $\pi^0 \to \gamma\gamma$ decays, where one of the photons converts into an electron-positron pair, and therefore, represent another source of fake electrons. Contributions from Dalitz decays of pions and photon conversions from $\pi^0 \to \gamma\gamma$ decays are commonly referred to as "π^0 decays" in the following.

As discussed in Ref. [265], it is challenging to accurately simulate the fake-lepton contributions. Furthermore, the fraction of events containing a reconstructed lepton candidate that passes the requirements of the ARs and is a fake lepton is tiny. As a consequence, it is impractical to simulate a sufficiently large number of such events. These two aspects prohibit a reliable estimation of the fake-lepton contributions to these regions based solely on simulation. Therefore, a data-driven approach is chosen. Detailed descriptions of the methods for estimations of fake-lepton contributions commonly used in measurements performed by the ATLAS collaboration are provided in Ref. [265]. In this work, the "asymptotic matrix method" is employed and is described in the following.

8.1.1 Estimation of fake-lepton contributions using the asymptotic matrix method

Two sets of requirements for lepton candidates are used in this method, the "loose" requirements, which represent a set of quality requirements that any lepton candidate must satisfy to be considered in the estimation, and the "tight" requirements, which represent an extended and stricter set of quality requirements with respect to the loose ones. The tight requirements are given by the lepton-quality requirements described in Section 4.3. The loose requirements are given by the tight requirements except

that the isolation requirements for electron and muon candidates are dropped and electron candidates are required to fulfill the medium LH identification WP instead of the tight LH identification WP. Lepton candidates fulfilling the loose criteria are referred to as "loose leptons", and those fulfilling the tight criteria are referred to as "tight leptons".

The contribution of fake leptons is estimated by reweighting events in the collected dataset that contain a loose lepton. The weight applied to the i-th event is given by Equation (63), where $\varepsilon_{\text{fake}}$ and $\varepsilon_{\text{prompt}}$ denote the efficiencies for fake and prompt loose leptons to satisfy the tight requirements, respectively. The former is also referred to as the fake-lepton efficiency and the latter as the prompt-lepton efficiency. The value of δ_i is 1 if the lepton candidate satisfies the tight requirements and 0 otherwise.

$$w_{\text{fake},\,i} = \frac{\varepsilon_{\text{fake}}}{\varepsilon_{\text{prompt}} - \varepsilon_{\text{fake}}} \cdot \left(\varepsilon_{\text{prompt}} - \delta_i\right) \tag{63}$$

The $\varepsilon_{\text{prompt}}$ values are extracted from simulated data, while the $\varepsilon_{\text{fake}}$ values are extracted from collected data and are computed with Equation (64). In this context, $N^{\text{loose}}_{\text{prompt}}$ and $N^{\text{tight}}_{\text{prompt}}$ represent the predicted number of events with prompt loose leptons and prompt tight leptons, respectively, based on simulation. Similarly, N^{loose} and N^{tight} correspond to the number of events with loose lepton candidates and tight lepton candidates in the collected data, respectively.

$$\varepsilon_{\text{fake}} = \frac{N^{\text{tight}}_{\text{fake}}}{N^{\text{loose}}_{\text{fake}}} = \frac{N^{\text{tight}} - N^{\text{tight}}_{\text{prompt}}}{N^{\text{loose}} - N^{\text{loose}}_{\text{prompt}}} \tag{64}$$

8.1.2 Estimation of prompt-lepton and fake-lepton efficiencies

The rejection factors of the lepton-isolation and lepton-identification requirements may differ significantly for the different fake-lepton sources. As a consequence, $\varepsilon_{\text{fake}}$ strongly depends on the relative contributions from each source. To ensure similar relative contributions, the criteria for selecting data used in the estimation should closely align with those of the ARs, while remaining orthogonal to them. Furthermore, the estimated value of $\varepsilon_{\text{fake}}$ relies on the modeling of the prompt-lepton contribution. This dependence should be as minimal as possible in the estimation. Meeting these conditions is challenging, especially when requiring the presence of a photon candidate in the selected data. However, it is assumed that the requirement for the presence of a photon candidate does not significantly alter the composition of fake leptons or the efficiencies for individual fake-lepton sources. Therefore, the prompt- and fake-lepton efficiencies are determined in events that satisfy all PSR requirements except for the photon-related ones. A test of these assumptions using simulated fake-lepton contributions is discussed below.

Events containing exactly one lepton candidate that satisfies the loose criteria are extracted from an MC sample that simulates top-quark-pair production, where both top quarks decay hadronically.[27] This $t\bar{t}$ topology yields a substantial number of events with fake leptons due to the high number of jets in the final state, including b jets, and the expected absence of a prompt lepton in the majority of events. These events allow testing the assumptions in a kinematic region that is close to the ARs. Specifically, the extracted events must include at least one jet with $p_{\text{T}} > 25\,\text{GeV}$ that passes the 85% WP of the DL1r algorithm. The remaining events are split into two categories: those including at least one photon candidate satisfying the same quality requirements as for the ARs (denoted by "$\geq 1\gamma$") and those without such a candidate (denoted by "no γ"). To enhance the number of events with photon candidates, the p_{T} requirement for the photon candidate is lowered from 20 GeV to 7 GeV. The relative contributions and the fake-lepton efficiency for the different fake-lepton sources are computed for both event categories.

Figure 32 illustrates the relative contributions of the different fake-lepton sources as a function of the transverse momentum of the lepton (p^{ℓ}_{T}) for both event categories. Additionally, it displays the

[27] The technical setup of this sample is the same as for the $t\bar{t}$ sample described in Section 7.2, so that the only difference between the setup of the samples are the considered decay channels of the top quark.

total fake-lepton efficiency and the fake-lepton efficiency for the two sources with the largest relative contributions as a function of p_{T}^{ℓ}. The upper row shows these distributions for fake electrons, the lower row for fake muons. Only sources contributing more than 1% to the respective total fake-lepton contribution are displayed, as statistical precision is insufficient for other sources, preventing conclusive insights from their distributions.

For both fake electrons and fake muons, the dominant contributions come from non-prompt leptons originating from EW decays of b hadrons (denoted by "bottom"), accounting for approximately 55% and 75% of the total fake-electron and fake-muon contributions, respectively. The second-largest contribution of fake electrons arises from misreconstructed hadrons ($\approx$ 24%), while the second-largest contribution of fake muons are non-prompt muons produced in EW decays of c hadrons ($\approx$ 22%). All remaining individual contributions are below 10% across the entire p_{T}^{ℓ} range. The compositions of fake electrons and fake muons in no-γ events agree well with the respective compositions in $\geq 1\gamma$ events. Furthermore, the fake-lepton efficiencies for the two dominant sources, as well as the total fake-lepton efficiency, agree well between the two event categories within the statistical uncertainties due to the limited sample sizes. These observations strongly suggest that the composition of fake-lepton events and the fake-lepton efficiencies are independent of the presence of a photon candidate in the selected data. This validates the approach of estimating the values of $\varepsilon_{\text{prompt}}$ and $\varepsilon_{\text{fake}}$ using events that are not required to contain a photon candidate and using these for the estimation of the fake-lepton contributions in the ARs.

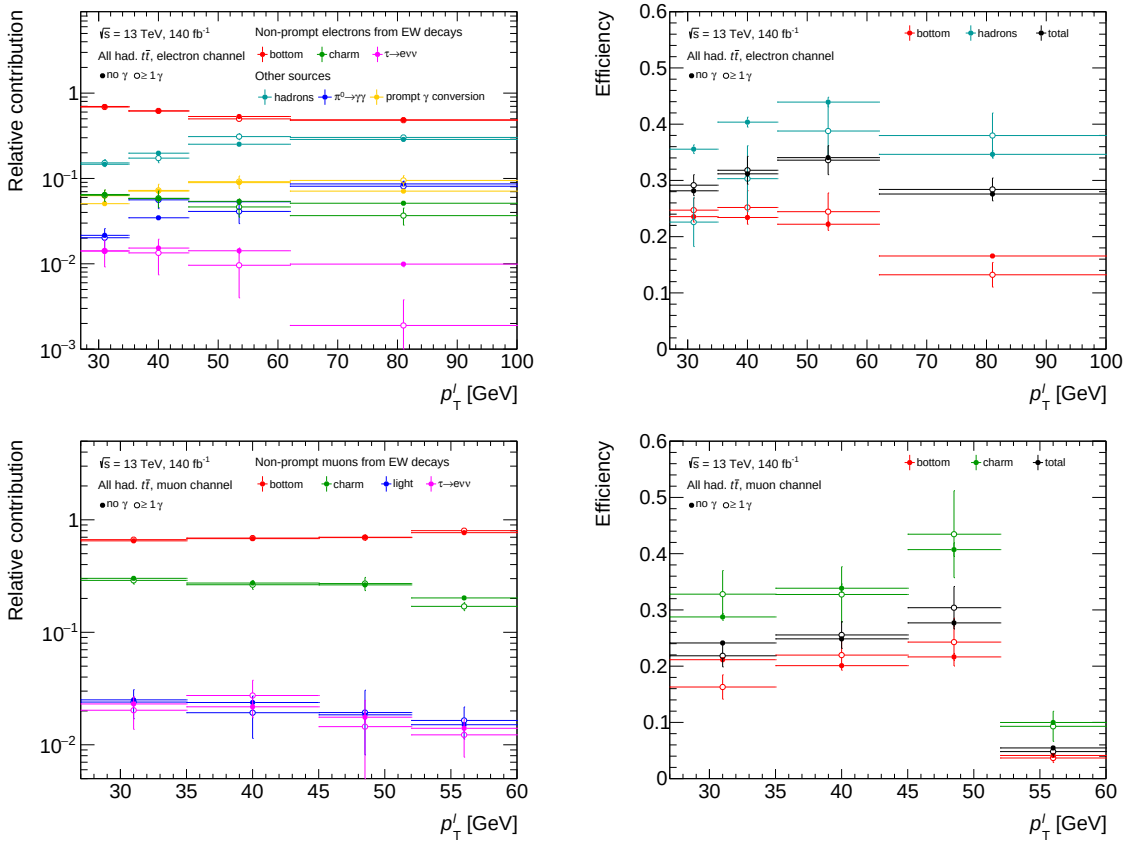


Figure 32: Relative contributions of the different fake-lepton sources (left column) with a relative contribution of at least 1% and the fake-lepton efficiencies for the two dominant sources and the total fake-lepton contribution (right column) for electrons (upper row) and muons (lower row). The values are extracted from a $t\bar{t}$ MC sample, in which both top quarks are required to decay hadronically. The values displayed by filled markers correspond to events for which the presence of a photon candidate is not required ("no γ"), while the others correspond to events for which the presence of at least 1 photon candidate ($\geq 1\gamma$) is required.

Based on this validation, the values of $\varepsilon_{\text{fake}}$ and $\varepsilon_{\text{prompt}}$ are extracted from two regions separately. The "1b70" and the "0b70,1b85" regions are defined by the same criteria as the SR and $W\gamma$ CR, respectively, except that events are not required to contain a photon candidate and must fulfill $30\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 50\,\text{GeV}$ instead of $E_{\text{T}}^{\text{miss}} > 30\,\text{GeV}$. The modification of the $E_{\text{T}}^{\text{miss}}$ requirement decreases the relative prompt-lepton contribution to both regions, so that the impact of the modeling of the prompt-lepton contribution on the estimation of the fake-lepton efficiencies is minimized. A discussion about this modification is provided in Appendix D.1. The prompt-lepton contributions in these regions are simulated by the X samples introduced in Section 7.2. The relative prompt-lepton contribution from $t\bar{t}$ is 33% in the 1b70 region and 5% in the 0b70,1b85 region. The relative prompt-lepton contribution from W+jets is 47% in the 1b70 region, while the relative prompt-electron and prompt-muon contributions are 74% and 82% in the 0b70,1b85 region, respectively. The single-top-quark production channels (t-channel, s-channel, and tW-channel production) contribute 10% to the predicted total prompt-lepton contribution in the 1b70 region and 2% in the 0b70,1b85 region. In the 1b70 region, Z+jets contributes 9% to the total prompt-lepton contribution, while in the 0b70,1b85 region, it contributes 19% to the total prompt-electron contribution and 10% to the total the prompt-muon contribution. The relative contribution from diboson production is small (0.5%–1%) in both regions across both lepton types.

The values of $\varepsilon_{\text{fake}}$ and $\varepsilon_{\text{prompt}}$ depend on kinematic and topological properties of the event and are therefore derived as a function of kinematic variables typically related to quantities of the lepton candidate. However, the choice of these variables and the corresponding parameterization of $\varepsilon_{\text{fake}}$ and $\varepsilon_{\text{prompt}}$ is individual for each physics analysis. The p_{T}^{ℓ} and η of the lepton candidate are two of the most commonly chosen variables for the parameterization. In this work, a parameterization depending on p_{T}^{ℓ} and the transverse mass, computed from the four-momentum of the lepton candidate and the missing transverse momentum, is found to provide better results than a parameterization depending on p_{T}^{ℓ} and η. For the dominant prompt-lepton contributions, arising from the W+jets and $t\bar{t}$ processes, the missing transverse momentum mainly arises from the final-state neutrino (ν). The missing transverse momentum is therefore attributed to this particle and the transverse mass is denoted by $m_{\text{T}}(\ell,\,\nu)$.

In the following, predicted kinematic distributions of events with prompt leptons are compared to the respective observed distributions. As the prediction does not include fake-lepton contributions, it is not expected that the predicted distributions align with the observed ones. The p_{T}^{ℓ} and $m_{\text{T}}(\ell,\,\nu)$ distributions in the 1b70 region for loose and tight leptons are depicted along with the ratios of the observed ("Data") and the predicted distributions ("Pred.") in Figure 33. The corresponding distributions for the 0b70,1b85 region can be found in Appendix D.1. The observations discussed below for the 1b70 region are qualitatively the same for the 0b70,1b85 region. The observed p_{T}^{ℓ} distributions and the respective predicted distributions disagree significantly at low p_{T}^{ℓ}, while showing good agreement at high p_{T}^{ℓ}. These observations are caused by the fact that fake leptons are characterized by a significantly softer p_{T}^{ℓ} spectrum relative to that of prompt leptons. The observed p_{T}^{ℓ} distributions are monotonically decreasing for events with tight lepton candidates. For events with loose muon and loose electron candidates, the observed p_{T}^{ℓ} distributions initially decrease monotonically until they suddenly rise at 50 GeV and 60 GeV, respectively, before resuming their monotonic decrease. These values correspond to the p_{T}^{ℓ} thresholds of the high-p_{T} muon and medium-p_{T} electron triggers (cf. Section 7.1), respectively. These triggers apply no isolation criteria to the candidate signature, unlike the respective low-p_{T} triggers, causing the sudden rise in collected data. This sudden rise is absent in the predicted prompt-lepton distributions, indicating that the efficiency of prompt leptons passing the isolation requirements of the low-p_{T} triggers is close to 100%.

The $m_{\text{T}}(\ell,\,\nu)$ distributions of events with prompt leptons originating from W-boson decays (e.g., W+jets or $t\bar{t}$ events) peak just below the W-boson pole mass ($\approx 80\,\text{GeV}$). This feature is absent for fake-lepton contributions and other prompt-lepton sources. As the $E_{\text{T}}^{\text{miss}}$ and p_{T}^{ℓ} values of events with fake leptons are typically smaller relative to those of events with prompt leptons, $m_{\text{T}}(\ell,\,\nu)$ is smaller for fake-lepton events on average. This is also reflected in the distributions as the ratio of the observed and predicted distributions is particularly large for $m_{\text{T}}(\ell,\,\nu) < 40\,\text{GeV}$ and decreases monotonically as a function

of $m_{\mathrm{T}}(\ell,\,\nu)$. This ratio is larger for loose muons than for loose electrons, while its decrease as a function of $m_{\mathrm{T}}(\ell,\,\nu)$ is steeper for loose muon events. Furthermore, the difference in the ratio between the distributions of loose muon and tight muon events is larger than that between the distributions of loose electron and tight electron events. In conclusion, fake-muon events are characterized by a significantly softer $m_{\mathrm{T}}(\ell,\,\nu)$ distribution than fake-electron events, and the fake-muon efficiency is smaller than the fake-electron efficiency. The ratio reaches similar values for loose and tight muon events for $m_{\mathrm{T}}(\ell,\,\nu) > 70\,\mathrm{GeV}$ and converges to 1 in both cases, which indicates that the fake-muon contribution is small in this region. Conversely, the value of the ratio is significantly higher for loose electron events relative to that for tight electron events across the entire $m_{\mathrm{T}}(\ell,\,\nu)$ range. The ratio remains at a constant level for $m_{\mathrm{T}}(\ell,\,\nu) > 80\,\mathrm{GeV}$ for loose and tight electron events. This level is about 1.3 for tight electron events indicating that the fake-electron contribution in the high-$m_{\mathrm{T}}(\ell,\,\nu)$ region is significant.

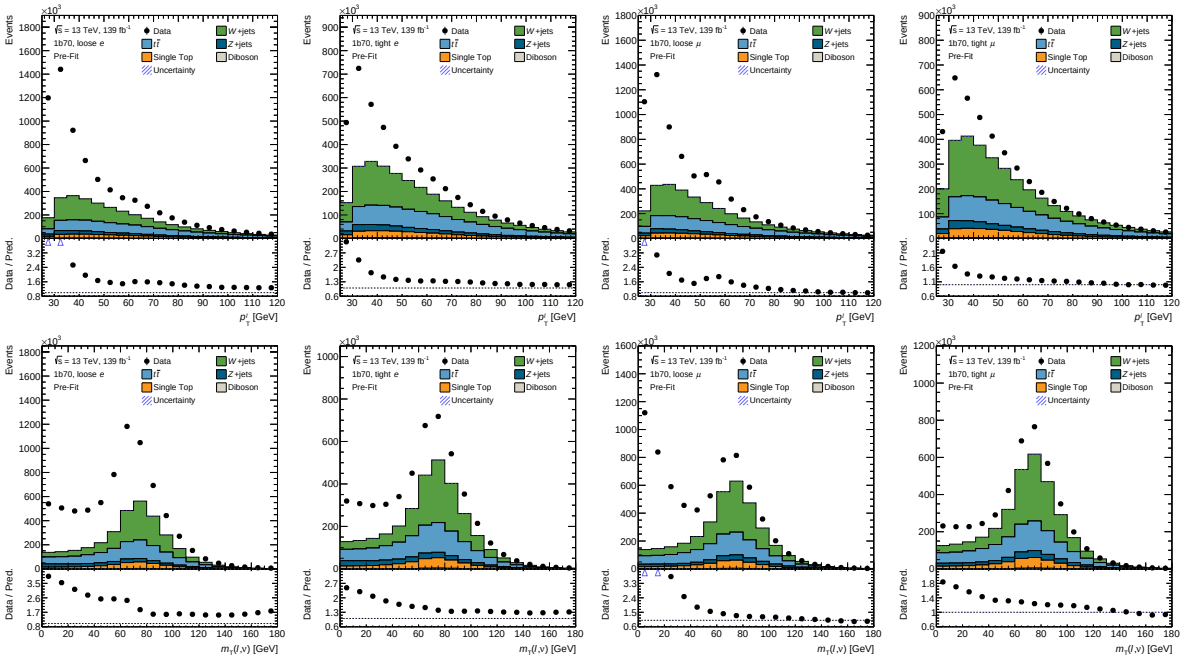


Figure 33: Distribution of events with loose and tight leptons in the 1b70 region as a function of the lepton p_{T}^{ℓ} (upper row) and $m_{\mathrm{T}}(\ell,\,\nu)$ (lower row). The first two distributions in each row correspond to events with an electron candidate, while the last two correspond to events with a muon candidate. The bottom panels illustrate the ratio of the observed ("Data") and predicted ("Pred.") distributions. The uncertainty bands correspond to the uncertainty due to the limited number of events in the MC samples.

Figure 34 shows the prompt-lepton efficiencies in the 1b70 region and the fake-lepton efficiencies in the 1b70 and 0b70,1b85 regions, parameterized as a two-dimensional function of p_{T}^{ℓ} and $m_{\mathrm{T}}(\ell,\,\nu)$. The prompt-lepton efficiencies in the 0b70,1b85 region agree well with those from the 1b70 region and are provided in Appendix D.1. Both the prompt-lepton and fake-lepton efficiencies are derived using the same parameterization. However, the fake-muon efficiencies are only derived for $m_{\mathrm{T}}(\ell,\,\nu) < 25\,\mathrm{GeV}$. The chosen parameterization is the result of testing different parameterizations and was found to provide the best performance among those tested. A parameterization is considered to perform well when good agreement is observed between the observed and predicted distributions, including the estimation of fake-lepton contributions, in data enriched in fake-lepton events (e.g., low-$E_{\mathrm{T}}^{\mathrm{miss}}$ and/or low-p_{T}^{ℓ} regions) in the SRs and $W\gamma$ CR. This is discussed in Section 8.1.3 for the chosen parameterization.

The uncertainty of the prompt-lepton efficiencies displayed in Figure 34 correspond to the uncertainty due to the limited number of simulated events. A systematic uncertainty is assigned to the fake-lepton efficiencies accounting for the modeling of the prompt-lepton contribution. The size of the overall contribution of the W+jets process is less well known in regions enriched in HF jets, such as the 1b70 and 0b70,1b85 regions, and hence a conservative uncertainty is assigned to it. Due to the large W+jets

contribution in the 1b70 and 0b85,1b70 regions, this uncertainty is expected to be the largest uncertainty in the modeling of the prompt-lepton contributions and, as a consequence, to have the largest impact on the fake-lepton efficiencies. The fake-lepton efficiencies are recomputed while altering the overall W+jets contribution by 30% and the difference to the nominal values is considered as uncertainty on the fake-lepton efficiencies. The total uncertainty on the fake-lepton efficiencies is given by the sum in quadrature of this uncertainty and the respective statistical uncertainty due to the limited sample sizes. The systematic uncertainty is significantly larger for most of the fake-lepton efficiencies.

The fake-muon efficiencies are significantly smaller than the fake-electron efficiencies. This is because the loose-electron definition is stricter in terms of the quality criteria than the loose-muon definition. Furthermore, it is observed that the fake-lepton efficiencies are larger in the 0b70,1b85 region relative to those in the 1b70 region. This observation is related to differences in the jet composition. The 0b70,1b85 region is expected to contain fewer events with genuine b jets, as shown in Figure 28 for the $W\gamma$+jets contribution to the $W\gamma$ CR, which uses the same jet and b-jet requirements as the 0b70,1b85 region. The efficiency of fake leptons from b-hadron decays is typically smaller relative to other significant sources, as illustrated in Figure 32. This causes the fake-lepton efficiencies to be larger in the 0b70,1b85 region.

The uncertainties of the fake-electron efficiencies are larger than those of the fake-muon efficiencies. The relative prompt-electron contributions are larger than the respective prompt-muon contributions in both regions. Therefore, the impact of the variation of the overall W+jets contribution on the prediction of the prompt-electron contribution is larger than the impact on the prediction of the prompt-muon contribution. This results in the larger uncertainties of the fake-electron efficiencies.

The fake-lepton efficiencies increase as a function of p_{T}^{ℓ} up to the aforementioned trigger thresholds, at which they drop. The lower edge of the last p_{T}^{ℓ} bin in the parameterization is chosen to be slightly larger than the trigger thresholds to avoid turn-on effects. The dependence on $m_{\mathrm{T}}(\ell,\, \nu)$ is smaller for the fake-electron efficiencies than for the fake-muon efficiencies. For $p_{\mathrm{T}}^{\ell} > 35\,\mathrm{GeV}$, the fake-electron efficiencies agree within the uncertainties within a constant p_{T}^{ℓ} region, while a clear trend for the efficiencies as a function of $m_{\mathrm{T}}(\ell,\, \nu)$ for $p_{\mathrm{T}}^{\ell} < 35\,\mathrm{GeV}$ is not observed. On the contrary, the fake-muon efficiencies increase as a function of $m_{\mathrm{T}}(\ell,\, \nu)$.

8.1.3 Estimation of the fake-lepton contribution in the SRs and CRs

The fake-lepton contributions in the SRs and CRs are estimated by assigning a weight to each event in collected data fulfilling the respective selection criteria, except that the lepton candidate is only required to satisfy the loose criteria. This weight is computed by Equation 63. The fake- and prompt-lepton efficiencies are given by $\varepsilon_X = \varepsilon_X(p_{\mathrm{T}}^{\ell},\, m_{\mathrm{T}}(\ell,\, \nu))$ using the parameterization shown in Figure 34, where $X = \{\text{fake, prompt}\}$. The efficiencies determined in the 0b85,1b70 region are used for computing the weights in the $W\gamma$ CR, while the efficiencies determined in the 1b70 region are used for the SRs and the $t\bar{t}\gamma$ CR. When a value of p_{T}^{ℓ} or $m_{\mathrm{T}}(\ell,\, \nu)$ exceeds the respective upper limit used in the parameterization, that upper limit is used for the extraction of the efficiencies. For instance, if the p_{T}^{ℓ} value of the loose lepton is larger than 100 GeV, the efficiencies are given by $\varepsilon_X(p_{\mathrm{T}}^{\ell} = 100\,\mathrm{GeV},\, m_{\mathrm{T}}(\ell,\, \nu))$.

Two uncertainties are assigned to the prediction for the fake-lepton contributions, which are considered in the statistical analysis and the investigations conducted in this section. An alternative prediction is compiled by using the fake-lepton efficiencies obtained by varying the W+jets contribution and the difference to the nominal prediction is used as uncertainty. The impact on the predicted overall fake-lepton contribution is neglected for this uncertainty, so that only the shape component of the respective NP is considered in the statistical analysis. The respective NP is denoted by "Fake leptons prompt subtr.". Additionally, a 50% uncertainty is assigned to the predicted overall fake-lepton contribution, denoted by "Fake leptons norm.". This value is found to be of the size of the maximum relative difference between the nominal and alternative fake-lepton predictions across the distributions in all ARs and both lepton types. These two uncertainties are treated as uncorrelated.

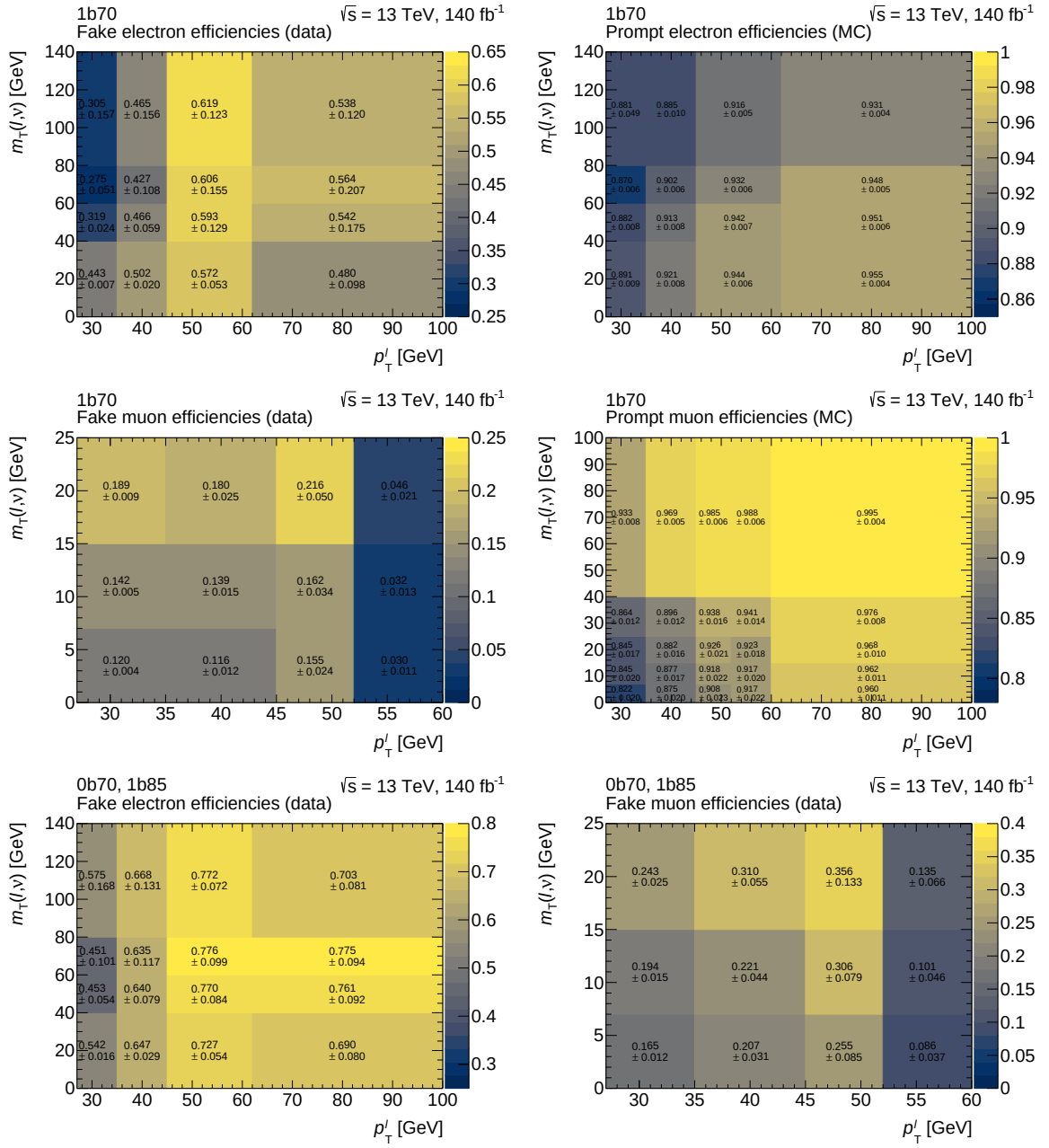


Figure 34: Fake-lepton efficiencies and prompt-lepton efficiencies as a two-dimensional function of p_{T}^{ℓ} and $m_{\mathrm{T}}(\ell,\,\nu)$. The upper row shows the fake-electron and prompt-electron efficiencies in the 1b70 region, the middle row the fake-muon and prompt-muon efficiencies in the 1b70 region, and the lower row the fake-electron and fake-muon efficiencies in the 0b70,1b85 region. The uncertainties on the fake-lepton efficiencies correspond to the sum in quadrature of the statistical uncertainties and the uncertainty associated with the variation of the prompt-lepton contribution from W+jets. The uncertainty on the prompt-lepton efficiencies correspond to the statistical uncertainties.

The estimation of the fake-lepton contribution is tested in the SR and $W\gamma$ CR using events satisfying $m_{\mathrm{T}}(\ell,\, \nu) < 25\,\mathrm{GeV}$. This additional requirement preserves approximately 29% and 25% of the predicted fake-lepton contribution in the SR and $W\gamma$ CR, while rejecting 89% and 88% of the predicted remaining contributions, respectively. Figure 35 shows, as examples, the observed and predicted p_{T}^{ℓ} (upper row) and $\Delta R(\ell, b)$ (lower row) distributions in both regions separately for events with an electron candidate ($1e$) and events with a muon candidate (1μ). The uncertainty corresponds to the sum in quadrature of the uncertainty due to the limited sample sizes and the two uncertainties assigned to the estimation of the fake-lepton contributions.

All predicted distributions agree well within the uncertainty with the corresponding observed distributions in the SR. Therefore, it is concluded that the fake-lepton contribution is well-modeled in the SR for both lepton types. The observed and predicted distributions agree well in the $W\gamma$ CR for events with one electron candidate. While the observed and predicted distributions agree well within the uncertainties for the p_{T}^{ℓ} distribution in the case of muons, the predicted distribution constantly overestimates the observed distribution for $p_{\mathrm{T}}^{\ell} < 70\,\mathrm{GeV}$. This indicates that the alternative set of fake-muon efficiencies might be better suited in this region as it provides lower predicted numbers of fake-muon events. However, as this discrepancy is covered by the systematic uncertainties of the estimation of the fake-muon contribution and the observed and predicted $\Delta R(\ell, b)$ distributions agree well within the uncertainty, it is concluded that the fake-electron and fake-muon contributions are well-modeled in the $W\gamma$ CR as well. The distributions for other quantities are shown in Appendix D.1.

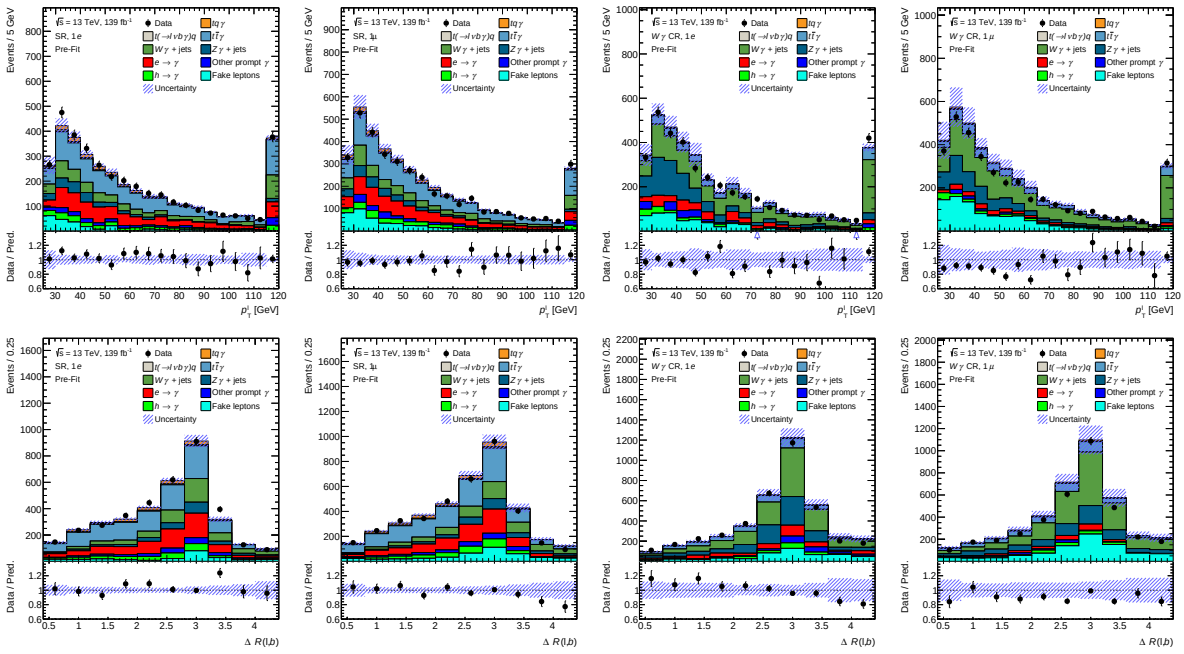


Figure 35: Predicted and observed p_{T}^{ℓ} and $\Delta R(\ell, b)$ distributions in the fake-lepton-enriched subsets of events in the SR and the $W\gamma$ CR. The distributions are shown separately for events with an electron candidate ($1e$) and events with a muon candidate (1μ). The uncertainty band corresponds to the sum in quadrature of the uncertainties due to limited sample sizes, the uncertainty in the shape of the fake-lepton contributions ("Fake leptons prompt subtr."), and the uncertainty in the prediction of the overall fake-lepton contribution ("Fake leptons norm."). The last bins include the overflow content.

8.2 Estimation of $e \to \gamma$ fake background

Electrons and photons are characterized by similar detector signatures and are reconstructed by a common algorithm. Therefore, electrons may be reconstructed as photon candidates that pass the

photon-isolation and photon-identification criteria. Furthermore, electrons may undergo bremsstrahlung, in which the radiated photon may also be reconstructed as a photon candidate and classified as a prompt photon. These contributions are commonly referred to as $e \to \gamma$ events, with the former type accounting for about 97% and 94% of the total $e \to \gamma$ contribution for converted and unconverted photon candidates, respectively. Ref. [2] studied the ratio ($f_{e\to\gamma}$) of prompt electrons that are falsely reconstructed as photon candidates and identified as prompt photons over prompt electrons that are correctly reconstructed as electron candidates and identified as prompt electrons.[28] Its value is about 1.5% for unconverted and 3% for converted photon candidates.[29]

This fake ratio may be mismodeled in simulation and needs to be calibrated to data for a precise prediction of the $e \to \gamma$ contribution. The calibration performed for this analysis is documented in Ref. [7] and was performed by the author of that reference. An overview of the calibration method, the uncertainties that are considered in the calibration, and the results is given in the following.

The aim of the calibration procedure is to obtain data-to-simulation scale factors ($\mathrm{SF}^{e\to\gamma}$) that are used to correct the predicted $e \to \gamma$ contribution in simulation via reweighting. The values of $\mathrm{SF}^{e\to\gamma}$ are computed by Equation (65) as the ratio of the values of the fake ratio determined using collected data ($f^{\mathrm{data}}_{e\to\gamma}$) and using simulated data ($f^{\mathrm{MC}}_{e\to\gamma}$).

$$\mathrm{SF}^{e\to\gamma} = \frac{f^{\mathrm{data}}_{e\to\gamma}}{f^{\mathrm{MC}}_{e\to\gamma}} \tag{65}$$

The values of $f_{e\to\gamma}$ are determined using events enriched in $Z \to ee$ decays, where the ee pair is either reconstructed as ee pair or as $e(e\to)\gamma$ pair. Events are selected that pass the requirements of the single-electron triggers and contain at least one reconstructed electron candidate associated with the signature that fired the trigger. To obtain samples characterized by a high purity of $Z \to ee$ events, events are further required to fulfill $E_{\mathrm{T}}^{\mathrm{miss}} < 30\,\mathrm{GeV}$ and not to contain a jet that passes the b-tagging requirements of the 70% DL1r WP. Events that meet these requirements are sorted into two regions: the $Z \to ee$ CR, where events contain exactly two electron candidates with opposite measured electric charges and no photon candidates fulfilling the criteria of the ARs, and the $Z \to e\gamma$ CR, where events are required to contain exactly one electron candidate and exactly one photon candidate satisfying these criteria. The invariant mass of the electron pair (m_{ee}) must satisfy $70\,\mathrm{GeV} < m_{ee} < 110\,\mathrm{GeV}$ in the $Z \to ee$ CR and the invariant mass of the electron-photon pair ($m_{e\gamma}$) must satisfy $70\,\mathrm{GeV} < m_{e\gamma} < 110\,\mathrm{GeV}$ in the $Z \to e\gamma$ CR. The kinematic requirements as well as the isolation and identification requirements for electron and photon candidates are the same as for the ARs. The majority of all events are predicted to arise from $Z \to ee$ decays that are reconstructed as an ee pair in the $Z \to ee$ CR (> 99%) and as an $e\gamma$ pair in the $Z \to e\gamma$ CR (> 86%).

The fake ratio is computed by Equation (66) and the $\mathrm{SF}^{e\to\gamma}$ by Equation (67). Here, $N(Z \to ee)$ and $N(Z \to e(e\to)\gamma)$ denote the number of $Z \to ee$ and $Z \to e(e\to)\gamma$ events in the $Z \to ee$ CR and $Z \to e\gamma$ CR, respectively. The superscripts in Equation (67) whether the respective value is predicted by simulation ("MC") or obtained from collected data ("data"). The factor of 2 in the denominator in Equation (66) is a combinatorial factor accounting for the fact that either of the two final-state electrons may be misreconstructed and misidentified. The values of $\mu_{e\gamma}$ and μ_{ee} are the ratios of the observed and predicted numbers of $Z \to ee$ events and $Z \to e(e\to)\gamma$ events, respectively.

$$f_{e\to\gamma} = \frac{N(Z \to e(e\to)\gamma)}{2 \cdot N(Z \to ee)} \tag{66}$$

$$\mathrm{SF}^{e\to\gamma} = \frac{N^{\mathrm{data}}(Z \to e(e\to)\gamma)/N^{\mathrm{MC}}(Z \to ee)}{N^{\mathrm{data}}(Z \to ee)/N^{\mathrm{MC}}(Z \to e(e\to)\gamma)} = \frac{\mu_{e\gamma}}{\mu_{ee}} \tag{67}$$

[28] Note that $f_{e\to\gamma}$ is denoted as "fake rate" in Ref. [2], although it is not defined as such.

[29] Note that in the text of the reference, the values for converted and unconverted photon candidates are incorrectly swapped.

The values of μ_{ee} and $\mu_{e\gamma}$ are determined by performing BPL fits of the predicted m_{ee} and $m_{e\gamma}$ distributions in the $Z \to ee$ CR and $Z \to e\gamma$ CR to the corresponding observed distributions, respectively. The m_{ee} and $m_{e\gamma}$ distributions are provided with a constant bin width of 1 GeV. The predicted distributions are obtained from the Z+jets, $Z\gamma$+jets, and $W\gamma$+jets samples. Furthermore, residual contributions to the $Z \to e\gamma$ CR are accounted for by using the third-order Bernstein polynomial in the respective fit. Uncertainties due to the limited size of the MC samples are considered by assigning one γ-factor to the total prediction per bin b. The parameterizations of the likelihood function used for these fits are adapted versions of the likelihood function discussed in Section 7.5 and are given in Appendix D.2.

The predicted and observed m_{ee} distributions are displayed in the upper row of Figure 36, where the left plot shows the nominal predicted distributions and the right plot shows the predicted distributions after performing the BPL fit. A value of $\mu_{ee} = 1.0500 \pm 0.0002$ is obtained and the predicted and observed distributions agree well after the BPL fit.

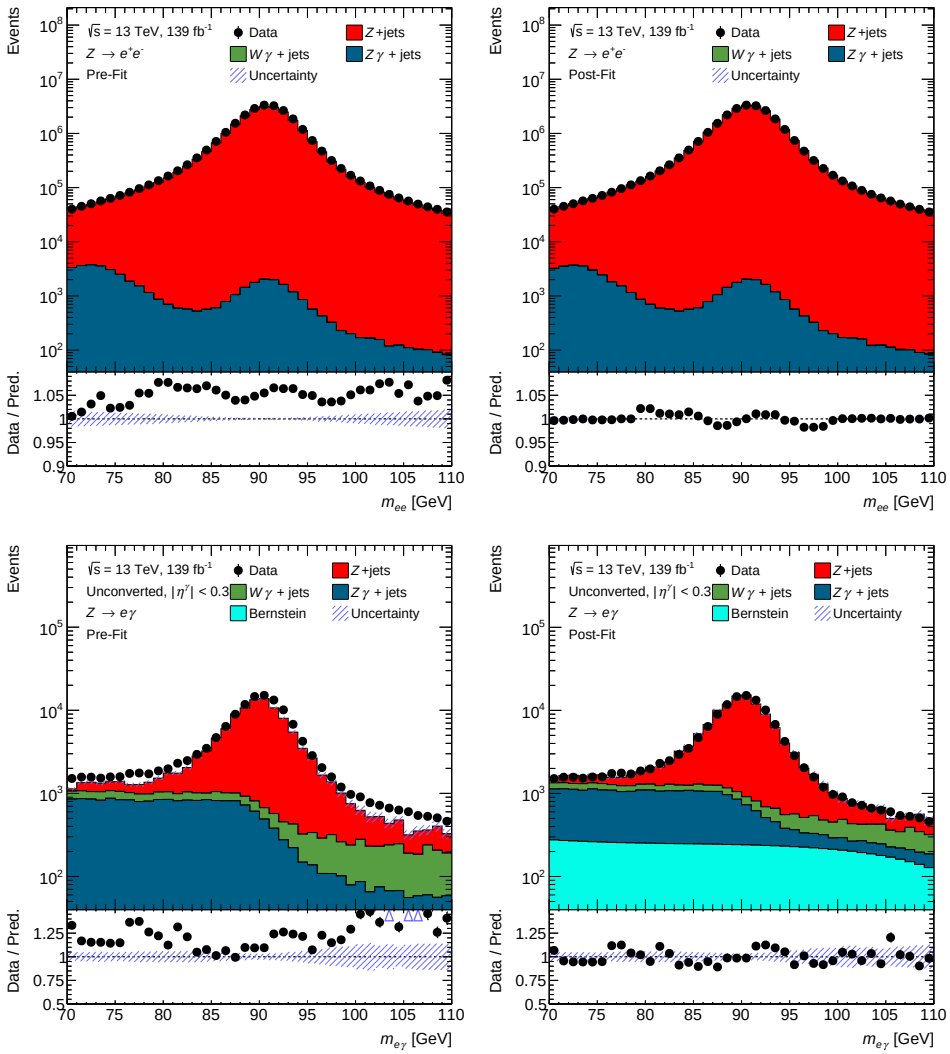


Figure 36: The upper row shows the predicted m_{ee} distribution and the respective observed distribution in the $Z \to ee$ CR. The lower row shows the $m_{e\gamma}$ distributions in the $Z \to e\gamma$ CR for events where the photon candidate is reconstructed as unconverted photon candidate with $|\eta^{\gamma}| < 0.3$. The bottom panels display the ratios of the observed ("Data") and predicted ("Pred.") distributions. The left figures show the predicted distributions before performing the BPL fits, the right figures after performing the BPL fits. The hashed band shows the total uncertainty in the prediction.

The values of $\mu_{e\gamma}$ are computed for six $|\eta^{\gamma}|$ ranges and for the six photon reconstruction types (cf. Table 2),

i.e., in total 36 $\mu_{e\gamma}$ values are computed. The $|\eta^{\gamma}|$ ranges are given by: $|\eta^{\gamma}| < 0.3$, $0.3 < |\eta^{\gamma}| < 0.6$, $0.6 < |\eta^{\gamma}| < 1.0$, $1.0 < |\eta^{\gamma}| < 1.37$, $1.52 < |\eta^{\gamma}| < 1.81$, and $1.81 < |\eta^{\gamma}| < 2.37$. These ranges are closely aligned with the ranges used in the measurement of the PID efficiency (cf. Section 5), optimized based on the comparison of the predicted and observed $|\eta^{\gamma}|$ distributions in the $Z \to e\gamma$ CR. The lower row of Figure 36 shows the $m_{e\gamma}$ distributions for events where the photon candidate is reconstructed as unconverted photon candidate with $|\eta^{\gamma}| < 0.3$, where the left plot shows the nominal predicted distributions and the right plot shows the predicted distributions after performing the BPL fit. The predicted and observed distributions show good agreement after the BPL fit.

8.2.1 Systematic uncertainties

Systematic uncertainties are assigned to the values of $\mathrm{SF}^{e\to\gamma}$ by recomputing their values using several alternative setups for the BPL fits. The differences between these alternative values and the nominal values are considered as systematic uncertainties. These alternative setups are described in the following:

- The values of $\mu_{e\gamma}$ are determined without using the third-order Bernstein polynomials.
- The values of μ_{ee} and $\mu_{e\gamma}$ are determined neglecting the predicted contributions from the $W\gamma$+jets and $Z\gamma$+jets processes.
- The energy of the photon candidate is calibrated under the assumption that it is a prompt photon, while in reality it is an electron for $Z \to e(e \to)\gamma$ events. This bias is accounted for by varying the energy of the photon candidate by 1%.[30]
- The fit ranges are reduced to $80\,\mathrm{GeV} < m_{ee} < 100\,\mathrm{GeV}$ and $80\,\mathrm{GeV} < m_{e\gamma} < 100\,\mathrm{GeV}$.
- The nominal Z+jets sample generated with Sherpa 2.2.4 is replaced with a sample that is generated using Powheg Box v2 and Pythia 8.

The total uncertainty of the $\mathrm{SF}^{e\to\gamma}$ values is given by the sum in quadrature of the individual uncertainties, while the different sources of uncertainty are treated as uncorrelated. Two NPs are considered in the BPL fits for the parton-level and particle-level measurements accounting for the $\mathrm{SF}^{e\to\gamma}$ uncertainties: one regulating the value of $\mathrm{SF}^{e\to\gamma}$ for events with converted photon candidates, and one regulating the value of $\mathrm{SF}^{e\to\gamma}$ for events with unconverted photon candidates within the constraints given by the respective uncertainties. The uncertainties of the $\mathrm{SF}^{e\to\gamma}$ values for the different $|\eta^{\gamma}|$ bins and for all subcategories of converted photon candidates are treated as fully correlated.

8.2.2 Results

Table 11 shows the determined values of $\mathrm{SF}^{e\to\gamma}$ for the different photon-reconstruction categories and $|\eta^{\gamma}|$ ranges. The values of $\mathrm{SF}^{e\to\gamma}$ tend to be closer to 1 in the end-cap region of the calorimeters ($|\eta^{\gamma}| > 1.52$). The most significant deviations from 1 are observed for the $\mathrm{SF}^{e\to\gamma}$ values in the central region ($|\eta^{\gamma}| < 1.37$) for the SingleSi and SingleTRT reconstruction types. The design of the detector offers the best performance for photon and electron measurements and the reconstruction of particle tracks in the central region. Therefore, the SingleSi and SingleTRT categories are the most atypical photon-reconstruction categories in this central region while the respective detector signatures (single ID track matched to an EM cluster) are the closest to the signature of an electron. Hence, a more pronounced mismodeling of the $e \to \gamma$ contributions to these categories is not unexpected. The nominal $\mathrm{SF}^{e\to\gamma}$ values for the DoubleSiTRT reconstruction type also deviate more strongly from 1 for $|\eta^{\gamma}| < 0.6$. However, the uncertainties of the corresponding $\mathrm{SF}^{e\to\gamma}$ values are large as the relative contribution of this reconstruction type is the smallest, so that the deviations are compatible with 1 at the level of $< 1.5\sigma$ of the $\mathrm{SF}^{e\to\gamma}$ uncertainties.

[30] This value is based on a dedicated study that investigated this bias performed by other researchers of the ATLAS collaboration.

Table 11: Measured values of the $SF^{e\to\gamma}$, with their uncertainties, in bins of $|\eta^{\gamma}|$ and the reconstruction type of the photon candidates ("Reco. type").

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

The values of $SF^{e\to\gamma}$ are found not to depend significantly on the photon p_{T}. The values of $SF^{e\to\gamma}$ are validated using events that fulfill the criteria of the $Z \to e\gamma$ CR except that events are required to contain at least one jet that passes the 70% WP of the DL1r b-tagging algorithm instead of none of such jets. A good agreement between the predicted and observed contributions, where the $e \to \gamma$ contribution is corrected by the $SF^{e\to\gamma}$, is observed. As the criteria for this region are similar to the criteria for the ARs, this observation validates that applying the computed $SF^{e\to\gamma}$ provides a reliable estimation of the $e \to \gamma$ contributions to the ARs. The $SF^{e\to\gamma}$ values are also applied in Ref. [11] to correct the predicted $e \to \gamma$ contribution.

8.3 Estimation of $h \to \gamma$ fake background

As discussed in Chapter 5, physics studies that require events to contain a reconstructed photon candidate typically face a significant background contribution arising from events where the photon candidate is a fake photon originating from hadronic activity. This contribution is referred to as the $h \to \gamma$ contribution. The composition of the $h \to \gamma$ contribution is phase-space dependent and, hence, specific to each physics study. The precise modeling of this composition via simulation is challenging as it requires a deep understanding of the composition of jets and calibrations of the responses of $h \to \gamma$ contributions to the reconstruction, isolation, and identification algorithms for photons, which, to date, are not centrally provided by the ATLAS collaboration to its researchers.

Therefore, the $h \to \gamma$ contribution is estimated in a data-driven fashion by using the ABCD method employing four orthogonal selection regions, which are denoted by "HFR A", "HFR B", "HFR C" and "HFR D". Regions HFR A, HFR B, and HFR C are enriched in $h \to \gamma$ events and are used to extrapolate the $h \to \gamma$ contribution to the target region (HFR D), which, in this work, is the PSR. The extrapolation exploits the weak correlation between the NSVs used for PID and the photon-isolation variables ($E_{\mathrm{T}}^{\mathrm{cone40}}$ and $p_{\mathrm{T}}^{\mathrm{cone20}}$), as discussed in Section 5.3. Therefore, HFR A, HFR B, and HFR C are defined by inverted photon-isolation criteria and/or partially inverted requirements on the NSVs of the tight PID WP. All other requirements on selected physics objects are the same as in the PSR. The predicted $e \to \gamma$ contribution to these regions is corrected by applying the $SF^{e\to\gamma}$ and the prediction for the fake-lepton contribution is included. The PID and photon-isolation requirements for the HFR A, HFR B, and HFR C are listed below.

- HFR A: the photon candidate is required to pass the requirements on the RTVs, to fail at least one of the criteria on the NSVs, and to pass the track-isolation and calorimeter-isolation requirements.
- HFR B: the photon candidate is required to pass the requirements on the RTVs, to fail at least one of the criteria on the NSVs, and to fail both the track-isolation and calorimeter-isolation requirements.

- HFR C: the photon candidate is required to pass the requirements of the tight PID WP, and to fail both the track-isolation and calorimeter-isolation requirements.

Figure 37 shows the yields in collected data and the predicted yields for the $h \to \gamma$, $e \to \gamma$, and prompt-γ contributions in the HFR A, HFR B, and HFR C separately for events with a converted and unconverted photon candidate. For simplicity, the small fake-lepton contributions are included in the prompt-γ contributions. For converted photon candidates, the predicted yields agree excellently with the observed yields. For unconverted photon candidates, the predicted and observed yields agree well, where the difference only just exceeds 1σ of the uncertainty of the predicted yields (HFR B and HFR C).

The $e \to \gamma$ contribution is significantly smaller in the case of unconverted photon candidates relative to the respective contribution in the case of converted photon candidates. The relative prompt-γ contributions for unconverted photon candidates are approximately two times larger for all regions than those for converted photon candidates. The relative $h \to \gamma$ contributions are of similar size for converted and unconverted photon candidates for all regions and are large, with at least 43% across all three regions. The relative $h \to \gamma$ contribution is at least five times larger relative to that in the PSR (cf. Figure 27). The HFR B is characterized by the largest relative $h \to \gamma$ contribution of about 93% and 95% for events with converted and unconverted photon candidates, respectively. This is expected, as photon candidates are required to fail both the photon-isolation and the PID requirements.

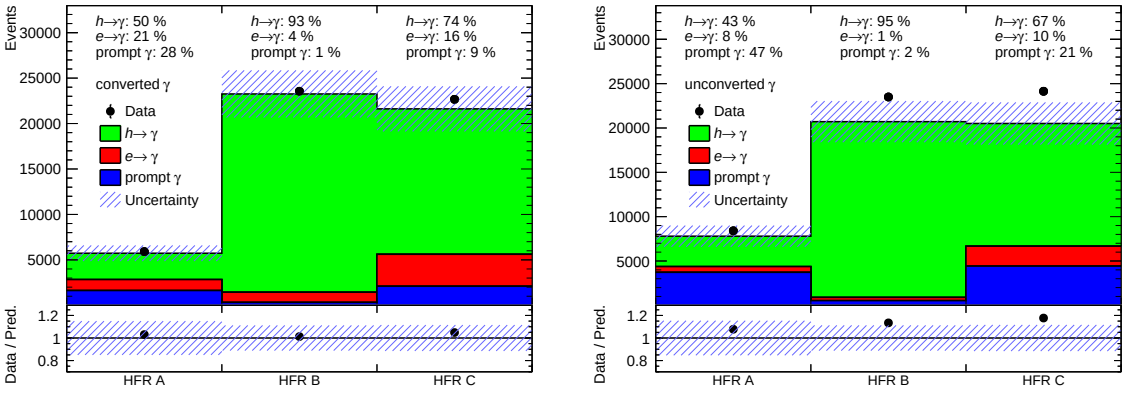


Figure 37: The number of events in collected data with converted photon (left) and unconverted photon candidates in the HFR A, HFR B, and HFR C (right). The respective predicted number of $e \to \gamma$, $h \to \gamma$, and prompt-γ events are also shown. The numbers listed at the top are the predicted relative contributions of $e \to \gamma$, $h \to \gamma$, and prompt-γ events. The ratio panel illustrates the ratio of the observed ("Data") and predicted yields ("Pred."). The hashed bands correspond to the total uncertainties due to all uncertainties described in Section 8.3.1, except for the variation of θ_{MC}.

Figure 38 shows the composition of the $h \to \gamma$ contributions in terms of the underlying physics process in all ABCD regions separately for events with a converted and an unconverted photon candidate. The four largest contributions are displayed, which arise from non-prompt photons from $\pi^0 \to \gamma\gamma$ decays ("π^0"), decays of other LF mesons ("LF meson"), decays of b mesons ("b meson"), and hadrons that are misreconstructed as photon candidates ("Hadrons"). The contribution from $\pi^0 \to \gamma\gamma$ decays is the largest for all regions and varies between 85% and 90% for converted photon candidates, and 80% and 87% for unconverted photon candidates. The contribution from b-meson decays is small and is smaller for converted photon candidates. The contributions from LF-meson decays and hadrons are of similar size for unconverted photon candidates. While the size of the relative contribution of LF-meson decays is similar for converted and unconverted photon candidates, the relative $h \to \gamma$ contribution from hadrons is smaller for converted photon candidates relative to that for unconverted photon candidates. The largest contribution of misreconstructed hadrons arises from electrically neutral Kaons, K_S and K_L, which are less likely to be associated with a CV, as they do not generate a particle track and are consequently less likely to be reconstructed as a converted photon candidate.

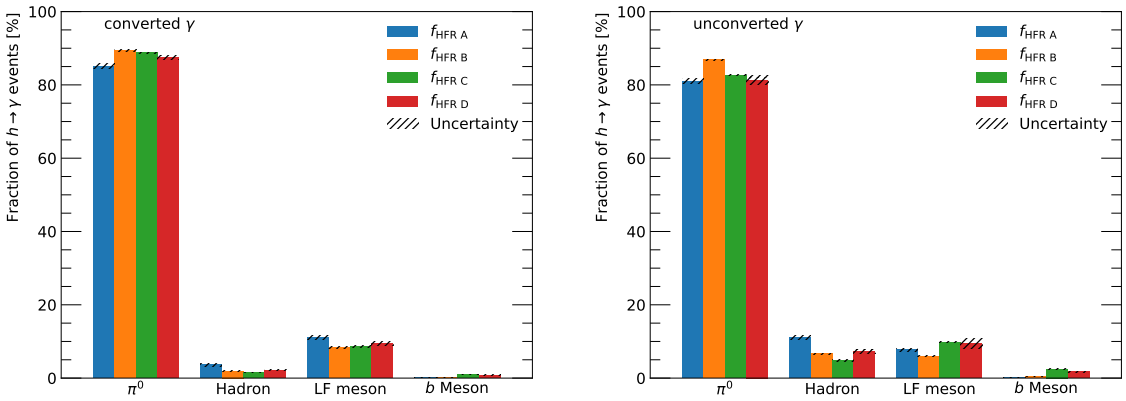


Figure 38: Relative contributions of different $h \to \gamma$ origins to the total $h \to \gamma$ contribution in all regions used for the ABCD method. The hashed uncertainty bands correspond to the statistical uncertainty due to the limited number of simulated events in the different categories.

The absolute $h \to \gamma$ contributions ($N_X^{h\to\gamma,\text{data}}$) to the HFR A, HFR B, and HFR C to the collected data are estimated by using Equation (68), with X indicating the region for which the contribution is estimated. The predicted numbers of events with prompt photons ($N_X^{\text{prompt }\gamma,\text{ MC}}$), with $e \to \gamma$ fakes ($N_X^{e\to\gamma,\text{ MC}}$) corrected by the $\text{SF}^{e\to\gamma}$ values, and with a fake lepton ($N_X^{\text{fake leptons}}$) are subtracted from the number of events in the collected data (N_X^{data}).

$$N_X^{h\to\gamma,\text{data}} = N_X^{\text{data}} - N_X^{\text{prompt }\gamma,\text{ MC}} - \text{SF}^{e\to\gamma} N_X^{e\to\gamma,\text{ MC}} - N_X^{\text{fake leptons}} \tag{68}$$

The $h \to \gamma$ contribution to the HFR D (PSR) is computed using these estimated contributions by the approximation given in Equation (69). This approximation assumes no correlation between the photon-isolation variables and the NSVs. This implies that the ratio of the efficiencies (ε) for photon candidates satisfying the photon-isolation requirements (I) and for photon candidates that do not (NI) to pass the requirements on the RTVs (RT) equals the ratio of the respective efficiencies to pass all requirements of the tight PID WP.

$$\frac{N_{\text{HFR A}}^{h\to\gamma,\text{data}}}{N_{\text{HFR B}}^{h\to\gamma,\text{data}}} = \frac{\varepsilon_{\text{RT}}^{\text{I}}}{\varepsilon_{\text{RT}}^{\text{NI}}} \approx \frac{\varepsilon_{\text{PID}}^{\text{I}}}{\varepsilon_{\text{PID}}^{\text{NI}}} = \frac{N_{\text{HFR D}}^{h\to\gamma,\text{data}}}{N_{\text{HFR C}}^{h\to\gamma,\text{data}}} \tag{69}$$

Figure 39 displays the values of the predicted efficiencies for all $h \to \gamma$ events, and individually for the most dominant $h \to \gamma$ sources separately for events with converted and unconverted photon candidates. The uncertainties of the efficiencies correspond to the uncertainties due to the limited sample sizes. The values of the efficiencies for events with converted photon candidates are similar to the respective values for events with unconverted photon candidates for all $h \to \gamma$ sources. The values of $\varepsilon_{\text{PID}}^{\text{NI}}$ and $\varepsilon_{\text{RT}}^{\text{NI}}$ are slightly smaller for events with unconverted photon candidates and consequently the values of $\varepsilon_{\text{PID}}^{\text{I}}$ and $\varepsilon_{\text{RT}}^{\text{I}}$ are slightly larger than the respective values for events with converted photon candidates. The predicted relations between $\varepsilon_{\text{PID}}^{\text{I}}$ and $\varepsilon_{\text{RT}}^{\text{I}}$, and $\varepsilon_{\text{PID}}^{\text{NI}}$ and $\varepsilon_{\text{RT}}^{\text{NI}}$ are the same for events with converted and unconverted photon candidates.

The values of $\varepsilon_{\text{RT}}^{\text{I}}$ and $\varepsilon_{\text{PID}}^{\text{I}}$ agree well within the uncertainties for $h \to \gamma$ contributions from misreconstructed hadrons and from LF-meson decays. The same is observed for the values of $\varepsilon_{\text{RT}}^{\text{NI}}$ and $\varepsilon_{\text{PID}}^{\text{NI}}$. The uncertainties of the efficiencies for $h \to \gamma$ events from $\pi^0 \to \gamma\gamma$ decays are the smallest, as this source contributes the majority of the $h \to \gamma$ events. For $h \to \gamma$ events from $\pi^0 \to \gamma\gamma$ decays, the values of $\varepsilon_{\text{PID}}^{\text{NI}}$ are slightly smaller than $\varepsilon_{\text{RT}}^{\text{NI}}$ while the values of $\varepsilon_{\text{PID}}^{\text{I}}$ are consequently slightly larger than the values of $\varepsilon_{\text{RT}}^{\text{I}}$. These slight discrepancies are due to the weak correlation of the NSVs and photon-isolation variables. The values of $\varepsilon_{\text{PID}}^{\text{NI}}$ and $\varepsilon_{\text{RT}}^{\text{I}}$ are significantly larger than $\varepsilon_{\text{RT}}^{\text{NI}}$ and $\varepsilon_{\text{PID}}^{\text{I}}$ for

$h \to \gamma$ events from b-meson decays, respectively. However, the overall contribution from this source is negligible (cf. Figure 38) and hence does not significantly impact the approximation. Therefore, the investigations of the predicted values of $\varepsilon^{\mathrm{NI}}_{\mathrm{PID}}$, $\varepsilon^{\mathrm{NI}}_{\mathrm{RT}}$, $\varepsilon^{\mathrm{I}}_{\mathrm{PID}}$, and $\varepsilon^{\mathrm{I}}_{\mathrm{RT}}$ and their predicted relations overall support the validity of the approximation given by Equation (69).

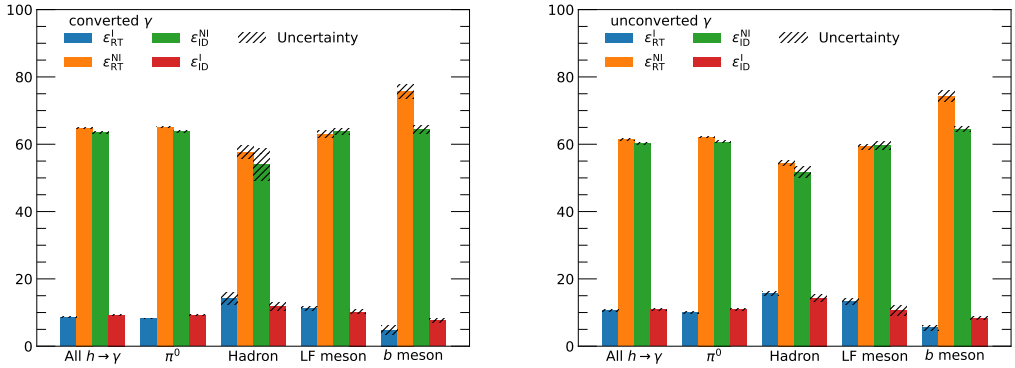


Figure 39: Predicted values of $\varepsilon^{\mathrm{NI}}_{\mathrm{RT}}$, $\varepsilon^{\mathrm{NI}}_{\mathrm{PID}}$, $\varepsilon^{\mathrm{I}}_{\mathrm{RT}}$, and $\varepsilon^{\mathrm{I}}_{\mathrm{PID}}$ for all $h \to \gamma$ events, and for individual $h \to \gamma$ origins. The values are shown separately for events with a converted fake photon (left) and an unconverted fake photon (right).

The residual small correlation of the NSVs and the photon-isolation variables are accounted for by θ^{MC} defined in Equation (70), also referred to as non-closure.

$$\theta^{\mathrm{MC}} = \frac{N^{h\to\gamma,\mathrm{MC}}_{\mathrm{HFR\ B}}}{N^{h\to\gamma,\mathrm{MC}}_{\mathrm{HFR\ A}}} \cdot \frac{N^{h\to\gamma,\mathrm{MC}}_{\mathrm{HFR\ D}}}{N^{h\to\gamma,\mathrm{MC}}_{\mathrm{HFR\ C}}} \tag{70}$$

The $h \to \gamma$ contribution to the collected data in HFR D is then estimated by Equation (71).

$$N^{h\to\gamma,\ \mathrm{data}}_{\mathrm{HFR\ D}} = \frac{N^{h\to\gamma,\ \mathrm{data}}_{\mathrm{HFR\ A}}}{N^{h\to\gamma,\ \mathrm{data}}_{\mathrm{HFR\ B}}} \cdot N^{h\to\gamma,\ \mathrm{data}}_{\mathrm{HFR\ C}} \cdot \theta^{\mathrm{MC}} \tag{71}$$

The predictions for the $h \to \gamma$ contributions, based on simulation, are corrected via data-to-simulation SFs ($\mathrm{SF}^{h\to\gamma}$) that are given by Equation (72).

$$\mathrm{SF}^{h\to\gamma} = \frac{N^{h\to\gamma,\mathrm{data}}_{\mathrm{HFR\ D}}}{N^{h\to\gamma,\mathrm{MC}}_{\mathrm{HFR\ D}}} \tag{72}$$

These SFs are determined as a function of $|\eta^{\gamma}|$ using the same parameterization as for the $e \to \gamma$ SFs, separately for converted and unconverted photon candidates. In contrast to the derivation of the $e \to \gamma$ SFs, no separate $h \to \gamma$ SFs are derived for the subcategories of converted photons, due to the lack of sufficient statistics in collected data for the individual reconstruction types. Furthermore, the $h \to \gamma$ SFs are derived separately for $p^{\gamma}_{\mathrm{T}} < 40\,\mathrm{GeV}$ and $p^{\gamma}_{\mathrm{T}} \geq 40\,\mathrm{GeV}$. The motivation for this choice is discussed in the following.

Figure 40 shows the computed values of $\mathrm{SF}^{h\to\gamma}$ for events where the photon candidate fulfills $p^{\gamma}_{\mathrm{T}} < X$ and $p^{\gamma}_{\mathrm{T}} \geq X$ as a function of the value of X, separately for events with converted and unconverted photon candidates. The computation is performed inclusively in η^{γ}. The value of X is varied between 30 GeV and 48 GeV in steps of 2 GeV. The error bars correspond to the total uncertainty of the $\mathrm{SF}^{h\to\gamma}$ values, considering all uncertainties discussed in Section 8.3.1. For events fulfilling $p_{\mathrm{T}} \geq X$, the computed $\mathrm{SF}^{h\to\gamma}$ values have no strong dependence on the value of X. In contrast, the values of $\mathrm{SF}^{h\to\gamma}$

decrease with increasing values of X for events fulfilling $p_{\mathrm{T}}^{\gamma} < X$, while this trend stops at about 40 GeV. Therefore, $X = 40\,\mathrm{GeV}$ is chosen and the values of $\mathrm{SF}^{h\to\gamma}$ used in the measurement are computed for events with $p_{\mathrm{T}}^{\gamma} < 40\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\gamma} \geq 40\,\mathrm{GeV}$. Due to the lack of sufficient statistics in the collected dataset, it is not considered reasonable to perform the $\mathrm{SF}^{h\to\gamma}$ computation for additional p_{T}^{γ} ranges.

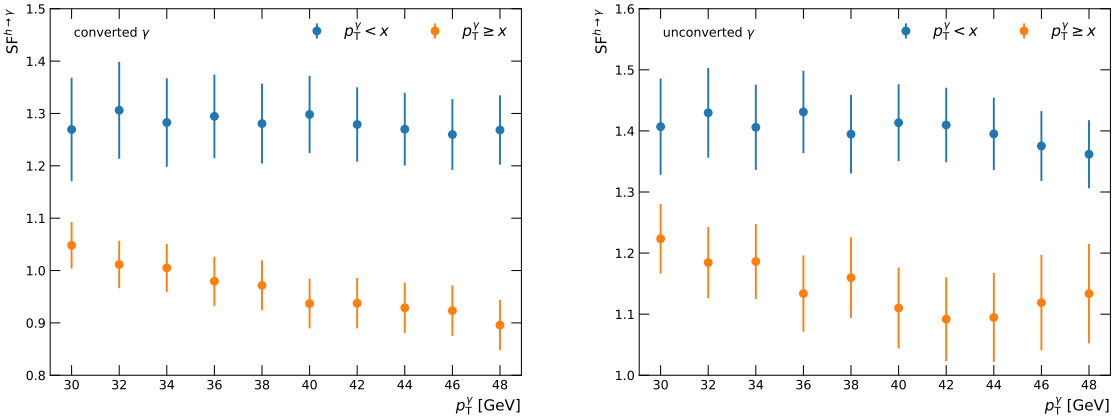


Figure 40: Computed $\mathrm{SF}^{h\to\gamma}$ values for photon candidates with $p_{\mathrm{T}}^{\gamma} < X$ and $p_{\mathrm{T}}^{\gamma} \geq X$ as a function of X, for events with converted and unconverted photon candidates. The error bars correspond to the total uncertainty of the $\mathrm{SF}^{h\to\gamma}$.

8.3.1 Systematic uncertainties

Seven different types of systematic uncertainties are assigned to the $\mathrm{SF}^{h\to\gamma}$ values, described below.

Limited data and MC statistics Similar to the inclusive photon method discussed in Section 5.3, several quantities and ratios are used in the computation of the $\mathrm{SF}^{h\to\gamma}$ values, which makes the analytic computation of the statistical uncertainties challenging. Therefore, the uncertainty due to the limited sizes of the collected dataset and simulated datasets is estimated by using 1,000 PDs for each of both dataset types. The PDs are constructed by bootstrap sampling: Each PD i is constructed by randomly drawing $N^{i}_{\mathrm{PD,\,data}}$ and $N^{i}_{\mathrm{PD,\,MC}}$ events from the respective sample. The values of $N^{i}_{\mathrm{PD,\,X}}$ follow Poisson distributions, where the mean values are given by the nominal number of events in the respective sample. The computation of the $\mathrm{SF}^{h\to\gamma}$ values is repeated for each of these PDs. The standard deviation of the resulting $\mathrm{SF}^{h\to\gamma}$ distribution is used as uncertainty. The number of 1,000 PDs is found to be sufficient as the resulting distributions of the $\mathrm{SF}^{h\to\gamma}$ values are Gaussian-like for all bins and the respective standard deviations are not altered significantly by enhancing the number of PDs. Appendix D.3 shows examples of these distributions.

Variation of $\mathrm{SF}^{e\to\gamma}$ The values of the $\mathrm{SF}^{e\to\gamma}$ are varied by $\pm 1\sigma$ of their respective uncertainties and the values of $\mathrm{SF}^{h\to\gamma}$ are recomputed. The differences between the alternative and nominal values are considered as uncertainty.

Variation of θ_{MC} As θ_{MC} is highly dependent on the modeling of the correlation of the photon-isolation variables and the NSVs, which are known to be mismodeled in simulation (cf. Section 5.1), conservative uncertainties are estimated for these values. These are obtained by recomputing the θ_{MC} values using the following alternative requirements on the SSVs for the HFR A and HFR B, while the other requirements remain the same:

1. The photon candidate additionally must satisfy the requirement on the NSV E_{ratio} of the tight PID WP.

2. The photon candidate additionally must satisfy the requirements on the NSVs ΔE and E_{ratio} of the tight PID WP.

3. The photon candidate is not required to satisfy the requirement for the RTV $\omega_{s,\,\mathrm{tot}}$.

The calculation of the $\mathrm{SF}^{h\to\gamma}$ values is repeated while setting the θ_{MC} value to the maximum and minimum values across all bins and variations (including the nominal values). The maximum absolute difference of the resulting two $\mathrm{SF}^{h\to\gamma}$ variations and the respective nominal value is considered as uncertainty. This is denoted as "non-closure" uncertainty in the following.

Prompt-photon subtraction in anti-tight regions Dedicated corrections of the prompt-γ contributions to regions where the photon candidate is required to fail either the photon-isolation or PID requirements are not available, while these contributions may be mismodeled. Furthermore, the modeling of the underlying processes impacts the $\mathrm{SF}^{h\to\gamma}$ values. To account for these potential impacts, the predicted contributions of prompt photons in the HFR A, HFR B, and HFR C are varied. The approach used in Ref. [266] is followed. The overall dominant prompt-γ contribution from $t\bar{t}\gamma$ is varied by ±15%, and prompt-γ contribution from other EW processes, i.e., $V\gamma$+jets, are varied simultaneously by ±30%. The maximum absolute difference of the resulting two $\mathrm{SF}^{h\to\gamma}$ variations and the respective nominal $\mathrm{SF}^{h\to\gamma}$ value is considered as uncertainty.

Variation of $\sigma_{t\bar{t}}$ Most of the $h \to \gamma$ events originate from $t\bar{t}$ production. The overall contribution of the process is varied by ±6% accounting for the uncertainty in the cross section of $t\bar{t}$ production [138, 182, 259–261]. The computation of the $\mathrm{SF}^{h\to\gamma}$ values is repeated for each of the two corresponding variations. The maximum difference of these variations and the nominal $\mathrm{SF}^{h\to\gamma}$ value is considered as uncertainty.

Scale uncertainties of $t\bar{t}$ and $t\bar{t}\gamma$ The impact due to the choice of the μ_{F} and μ_{R} values in the production of the $t\bar{t}$ and $t\bar{t}\gamma$ samples on the $\mathrm{SF}^{h\to\gamma}$ values is evaluated by repeating the calculation for each of the seven variations $(\mu_{\mathrm{F}}/\mu_{\mathrm{F}}^{\mathrm{nom}}, \mu_{\mathrm{R}}/\mu_{\mathrm{R}}^{\mathrm{nom}})$: (0.5, 0.5), (1.0, 0.5), (0.5, 1.0), (1.0, 2.0), (2.0, 1.0), and (2.0, 2.0), excluding the nominal case (1.0, 1.0). The envelopes of the resulting variations are used as uncertainties.

Total uncertainty The total uncertainties of the $\mathrm{SF}^{h\to\gamma}$ values are given by the sum in quadrature of the individual components.

Treatment of $\mathrm{SF}^{h\to\gamma}$ uncertainties in statistical analysis The total uncertainties of the value of $\mathrm{SF}^{h\to\gamma}$ are derived individually for each bin. In the statistical analysis, the uncertainties of the $\mathrm{SF}^{h\to\gamma}$ values for the different $|\eta^{\gamma}|$ bins are considered fully correlated. Additionally, the $\mathrm{SF}^{h\to\gamma}$ values for the same reconstruction type and for the same p_{T}^{γ} bin are treated as fully correlated. Therefore, four independent NPs are considered for the BPL fits: two regulate the values of $\mathrm{SF}^{h\to\gamma}$ for events with converted photon candidates fulfilling $p_{\mathrm{T}}^{\gamma} < 40\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\gamma} \geq 40\,\mathrm{GeV}$, denoted as "$\mathrm{SF}^{h\to\gamma}$ conv. low p_{T}" and "$\mathrm{SF}^{h\to\gamma}$ conv. high p_{T}", respectively, while the other two regulate the respective values for events with unconverted photon candidates, denoted as "$\mathrm{SF}^{h\to\gamma}$ unconv. low p_{T}" and "$\mathrm{SF}^{h\to\gamma}$ unconv. high p_{T}", respectively.

8.3.2 Results

Figure 41 shows the $\mathrm{SF}^{h\to\gamma}$ values for events with converted and unconverted photon candidates. The figure also displays the total uncertainties (tot. unc.), the statistical uncertainties from limited sample sizes (stat. unc.), and the combined systematic uncertainties (syst. unc.). Furthermore, it shows the breakdown of the total uncertainty into the individual contributions to the respective sum in quadrature $((\Delta\mathrm{SF}^{h\to\gamma})^2)$.

The values of $\mathrm{SF}^{h\to\gamma}$ agree with 1 within the respective total uncertainties across all bins. For converted photon candidates, the nominal $\mathrm{SF}^{h\to\gamma}$ values are close to 1 across all bins except for the most central region $|\eta| < 1$ and $p_{\mathrm{T}}^{\gamma} < 40\,\mathrm{GeV}$, where the values are between 1.5 and 1.7. A clear trend of the $\mathrm{SF}^{h\to\gamma}$ values is absent for unconverted photon candidates. The values tend to be closer to 1 for $p_{\mathrm{T}}^{\gamma} \geq 40\,\mathrm{GeV}$ in

the central region, while the value in the end-cap region is characterized by the largest deviation from 1. For $p_{\mathrm{T}}^{\gamma} < 40\,\mathrm{GeV}$, $\mathrm{SF}^{h\to\gamma}$ values fluctuate around 1.4.

The uncertainty due to the limited sample sizes is significantly smaller than the uncertainty due to the other sources of systematic uncertainties across all bins. The non-closure uncertainty is the dominant uncertainty and is significantly larger than the other uncertainties across almost all bins. The prompt-photon subtraction uncertainty is the only other uncertainty that is of similar size for $p_{\mathrm{T}}^{\gamma} < 40\,\mathrm{GeV}$ and for unconverted photon candidates with $p_{\mathrm{T}}^{\gamma} \geq 40\,\mathrm{GeV}$ in the end-cap region, and is the subdominant uncertainty in most of the bins. Other uncertainties are small relative to these two types of uncertainties and contribute little to the total uncertainty.

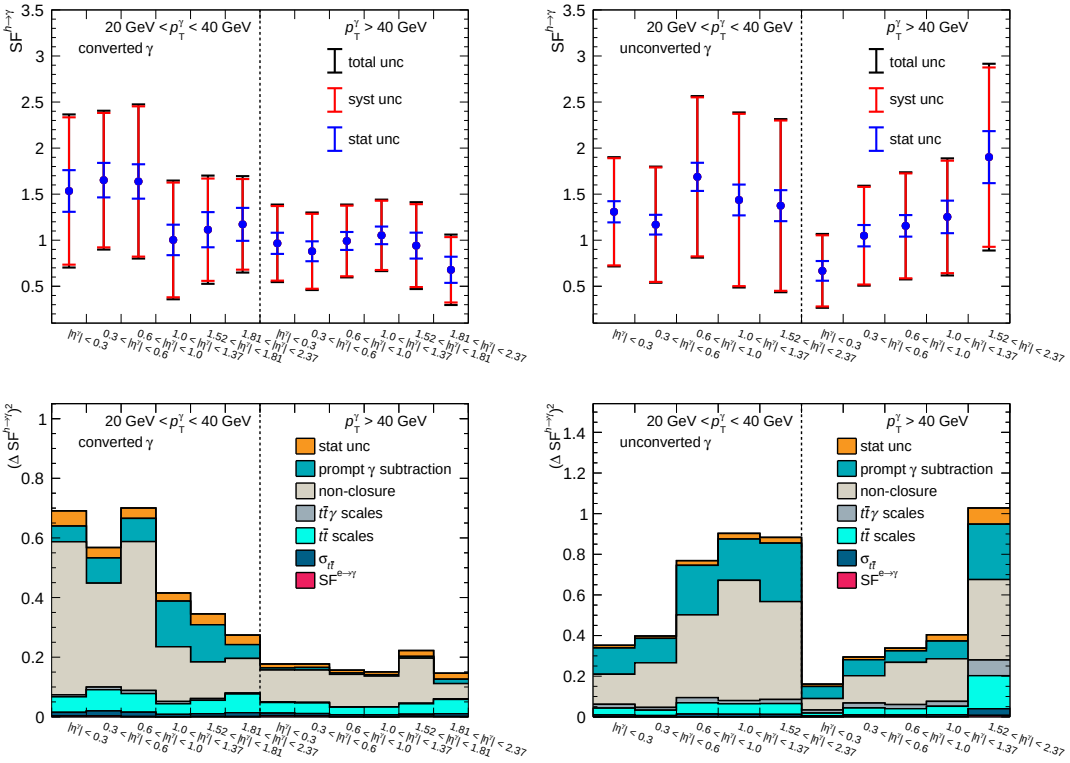


Figure 41: The upper row shows the computed $\mathrm{SF}^{h\to\gamma}$ values for events with converted and unconverted photon candidates. The error bars represent the uncertainties due to the limited sizes of the collected and simulated datasets (stat. unc.), due to all systematic uncertainties (syst. unc.), and due to all uncertainties (tot. unc.). The lower row shows the breakdown of the total uncertainty into the individual contributions.

The upper panels of Figure 42 shows the predicted and observed yields in the HFR D (PSR) for all p_{T}^{γ} and $|\eta^{\gamma}|$ bins that are considered for the computation of the $\mathrm{SF}^{h\to\gamma}$ values. The $e \to \gamma$ and $h \to \gamma$ contributions are corrected by the respective $\mathrm{SF}^{e\to\gamma}$ and $\mathrm{SF}^{h\to\gamma}$ values. The uncertainty bands correspond to the uncertainties due to the limited samples sizes and the $\mathrm{SF}^{e\to\gamma}$ and $\mathrm{SF}^{h\to\gamma}$ uncertainties. The predicted yields agree well with the observed yields within the uncertainty across all bins for converted photon candidates. The agreement is worse for unconverted photon candidates, where the prediction underestimates the observed yield across all bins. This discrepancy is especially pronounced in the end-cap region. However, the majority of all events with unconverted photon candidates are prompt-γ events. A potential mismodeling of this contribution causing this discrepancy is not expected to arise from the modeling of the prompt-γ contribution itself as it is corrected by the dedicated calibrations discussed in Sections 4.3 and 5. But it may arise from the mismodeling of the underlying physics processes (e.g., $t\bar{t}\gamma$ or $W\gamma$+jets production), which is not accounted for in these predicted distributions.

The lower panel shows the ratio of observed and predicted yields for the cases when the $\mathrm{SF}^{e\to\gamma}$ and

$SF^{h\to\gamma}$ are not applied ("no fake γ SFs"), only the $SF^{e\to\gamma}$ is applied ("$e \to \gamma$"), and both the $SF^{e\to\gamma}$ and $SF^{h\to\gamma}$ are applied ("$e \to \gamma$ & $h \to \gamma$"). This provides insights into how each of the different types of SFs contributes to the improvement of the predictions. The application of the $SF^{e\to\gamma}$ significantly improves the agreement for events with converted photon candidates. The additional application of the $SF^{h\to\gamma}$ further improves the agreement for $|\eta^\gamma| < 1.0$ and $p_T^\gamma < 40\,\text{GeV}$, while no further improvement is observed in the other bins. As the $e \to \gamma$ contribution is significantly smaller for events with unconverted photon candidates, the improvement in the agreement by the application of $SF^{e\to\gamma}$ is less pronounced. Especially for $p_T^\gamma < 40\,\text{GeV}$, applying the $SF^{h\to\gamma}$ leads to a larger improvement of the agreement than applying the $SF^{e\to\gamma}$. For $p_T^\gamma \geq 40\,\text{GeV}$, no improvement in the agreement is observed in the central region, while a significant improvement in the agreement is obtained by applying the $SF^{h\to\gamma}$ in the end-cap region.

In conclusion, the application of the $SF^{e\to\gamma}$ and $SF^{h\to\gamma}$ provides significant improvements in the modeling of the $e \to \gamma$ and $h \to \gamma$ contributions in the PSR. The presented method for deriving the $SF^{h\to\gamma}$ was adapted for Ref. [11] and the corresponding results were compiled by the author of this work.

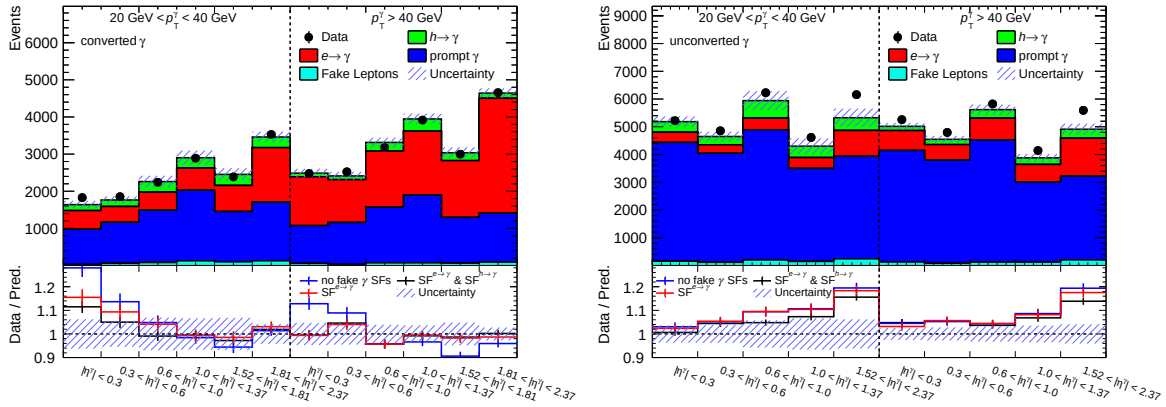


Figure 42: The upper panels show the yields in collected data and predicted yields in all p_T^γ and $|\eta^\gamma|$ bins used for the computation of the $SF^{h\to\gamma}$ values in the HFR D (PSR). The uncertainty band represents the uncertainties of the prediction due to the limited sample sizes and the $SF^{e\to\gamma}$ and $SF^{h\to\gamma}$ uncertainties. The bottom panels show the ratio of event yields in collected data and the predicted yields for different cases: no fake-γ SFs, only $SF^{e\to\gamma}$, and both $SF^{e\to\gamma}$ and $SF^{h\to\gamma}$ applied to the prediction.

9 Discrimination of $tq\gamma$ and background contributions

The primary goals of observing the $tq\gamma$ process and precisely measuring the fiducial cross section of single-top-quark production in association with a photon are challenging to achieve, as the expected signal-to-background ratio is small even in the SRs. Therefore, the signal contributions need to be efficiently separated from the background contributions. The $tq\gamma$ topology is characterized by features that are distinctive from those of the background topologies, e.g., the presence of a high-p_{T} forward jet. However, identifying an individual variable, i.e., a kinematic or topological quantity, or a combination of such variables that offers sufficient discriminative power is a difficult task. Thus, this challenge is circumvented by constructing a discriminator for each SR, which combines multiple individual features into a single output quantity by employing supervised machine learning (ML) techniques. These output quantities provide greater discriminative power than any of the individual features and are used as the final discriminants in the statistical analysis. The strategy for designing these discriminators as well as their optimization and performance will be discussed in this chapter.

9.1 Strategy for the construction of machine learning discriminators

The chosen ML models for efficiently discriminating $tq\gamma$ from background events are feed-forward deep neural networks (NNs) consisting of fully connected dense layers. An individual NN model is constructed for each SR. Each NN model performs a binary classification of the input data using logistic regression. The input data consists of either the simulated events or events from the collected data. The parameters of the NN models are obtained by fitting them to the simulated events. The $tq\gamma$ events are assigned a target value of 1, while background events are assigned a target value of 0. The fitted NN models predict whether an event is a $tq\gamma$ event by assigning a value to it, referred to as NN output ($\mathrm{NN_{out}}$), based on the values of a set of individual variables that are referred to as "input features". Consequently, $\mathrm{NN_{out}}$ values close to 1 correspond to events that are the most "$tq\gamma$-like". Table 12 presents the total number of simulated $tq\gamma$, $t\,(\rightarrow \ell\nu b\gamma)\,q$, and background events available for the 0fj SR and $\geq$ 1fj SR.

Table 12: Total number of simulated $tq\gamma$, $t\,(\rightarrow \ell\nu b\gamma)\,q$, and background events available for the 0fj SR and $\geq$ 1fj SR. The number of background events includes the number of $t\,(\rightarrow \ell\nu b\gamma)\,q$ events.

Region	$tq\gamma$	$t\,(\rightarrow \ell\nu b\gamma)\,q$	background
0fj SR	54415	9792	533416
$\geq$ 1fj SR	48552	8813	288961

The kinematic properties of $t\,(\rightarrow \ell\nu b\gamma)\,q$ and $tq\gamma$ events are significantly different from each other, as discussed in Chapter 6. Therefore, treating both processes simultaneously as signal may compromise the discriminative power of the NN models compared to treating only $tq\gamma$ as signal. As the number of events in the $t\,(\rightarrow \ell\nu b\gamma)\,q$ sample and the expected $t\,(\rightarrow \ell\nu b\gamma)\,q$ contributions to the SRs are small, designing an ML model that discriminates between $t\,(\rightarrow \ell\nu b\gamma)\,q$ events and any other contribution is neither promising nor beneficial for achieving the goals of this work. As a consequence, the $t\,(\rightarrow \ell\nu b\gamma)\,q$ events are also treated as part of the background when fitting the NN models.

In the statistical analysis, the predicted $\mathrm{NN_{out}}$ distributions are compared to those obtained from collected data. Using any of these collected data events in the fitting of the NN models introduces a bias to the statistical analysis. Therefore, the fake-lepton contribution is excluded from the background in the fitting of the NN models, as it is estimated using collected data that is also used in the statistical analysis.

The design and selection of input features is the first step in constructing the NN models. Comprehensive lists of candidate variables are compiled that represent features of the individual selected physics objects, combinations of these objects, or the event topology. While it is possible to use all candidate

variables as input features, disregarding variables reduces the complexity of the NN models. This decreases the computation time needed to fit these models to the input data and mitigates the risk of significantly "overfitting" the models, i.e., the models adapt to statistical fluctuations in the input data. However, choosing which variables to disregard as input features is a challenging task, as variables that provide greater discriminative power should be kept. This decision is guided by employing an ML algorithm, specifically by using gradient boosting decision trees (GBDTs) implemented in the open-source XGBoost library [267]. This procedure is described in Section 9.2.

The reduced lists of candidate variables are used as initial input features for benchmark NN models, which are discussed in Section 9.3. Starting with the benchmark NN models, the list of input features is further reduced in an iterative fashion, which is described in Section 9.4. Subsequently, the structures of the NN models, i.e., the number and design of the dense layers, are optimized. Finally, the NN models using the final sets of input features and the optimized structures are fitted to a subset of the input data. The fitting and performance of these final NN models are presented in Section 9.6. The $\mathrm{NN_{out}}$ distributions of these final models obtained from the signal and background samples and the dataset are used in the statistical analysis.

9.2 Selection of initial input features

This section describes the procedure for compiling initial lists of variables that are used as input features for the NN models. As a first step, comprehensive lists of candidate variables are assembled for both SRs, from which subsets are selected to form these initial lists. These variables represent either "low-level" features, which refer to individual kinematic quantities of the selected physics objects, or "high-level" features, which are combinations of at least two low-level features. Ideally, the low-level information is sufficient to build powerful discriminators as NN models are able to adapt to the correlations between their input features and therefore use high-level information intrinsically. However, as the correlations between features may have intricate shapes, this typically requires NN models to be more complex, i.e., to contain more free parameters, compared to NN models that directly use high-level information as input. Since the available number of simulated $tq\gamma$ events is relatively small, as listed in Table 12, which limits the maximum level of model complexity, adding high-level information as support is found to be mandatory for the construction of powerful and reliable NN discriminators.

The low-level features that are considered in the initial lists of input features comprise the p_{T}, η, and ϕ of the lepton candidate, photon candidate, b-tagged jet, and forward jet[31] as well as $E_{\mathrm{T}}^{\mathrm{miss}}$ and ϕ of the missing transverse momentum (ϕ^{miss}), the type of the charged lepton (PID^{ℓ}), the information about which of the b-tagging WPs are passed by the b-tagged jet (b-tagging bin)[32], and whether the p_{T}-leading jet (lj) that fails the tight b-tagging requirements of the SRs is a forward jet (lj = fj?).

While the number of low-level features that can be extracted from an event is usually limited, there is no limit to the number of high-level features that can be constructed from the low-level features. Therefore, comprehensive lists of candidate high-level features are compiled and then reduced to lists containing only those that provide strong discriminative power and added to the initial lists of input features.

The $tq\gamma$ topology involves the semileptonic decay of a top quark. The top quark and the W boson from the top-quark decay are reconstructed by identifying selected physics objects as their decay products. Since several background contributions do not involve a semileptonic top-quark decay, properties of the reconstructed top quark and W boson are discriminative high-level features. The selected lepton candidate and b jet are identified with the lepton from the W-boson decay and the b quark from the top-quark decay, respectively. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is identified as the $\vec{p}_{\mathrm{T}}$ of the neutrino from the W-boson decay.

[31] Properties related to the forward jet are only considered for the NN model fitted to simulated events in the ≥ 1fj SR due to the absence of forward jets in the 0fj SR.

[32] As the b-tagged jet is required to pass the 70% WP in the SRs and hence also passes the 77% and 85% WPs, this information represents a flag indicating whether the b-tagged jet passes the 60% WP.

The missing z-component of its momentum (p_z^ν) is given by the value for which the invariant mass of the lepton and neutrino ($m_{\ell\nu}$) matches the W-boson pole mass ($m_W = 80.379\,\text{GeV}$), under the assumption that the neutrino is massless. The resulting four-momentum of the neutrino and the four-momentum of the lepton candidate are then combined into the reconstructed four-momentum of the W boson. Furthermore, the four-momenta of the neutrino, the lepton candidate, and the b-tagged jet are combined into the reconstructed four-momentum of the top quark.

The following high-level features are considered as candidate variables for the initial lists of input features:

- the p_T, η, ϕ, mass (m), and energy (E) of all pairwise combinations of the four-momenta of the lepton candidate, photon candidate, b-tagged jet, and forward jet,
- all pairwise angular distances ΔR, $\Delta\eta$, and $\Delta\phi$ of the lepton candidate, photon candidate, b-tagged jet, and forward jet,
- the cosine of the angle α between the momentum vectors ($\cos\alpha$) of each pair of the lepton candidate, photon candidate, b-tagged jet, and forward jet,
- the scalar sum (H_T) of the p_T of all jets,
- the rapidity (y) of the pairwise combinations of the four-momenta of the lepton candidate, photon candidate, and b-tagged jet,
- the transverse mass of the lepton candidate and neutrino ($m_\text{T}(\ell, \nu)$), of the lepton candidate, neutrino, and photon candidate ($m_\text{T}(\ell, \nu, \gamma)$), and of the lepton candidate, neutrino, and b-tagged jet ($m_\text{T}(\ell, \nu, b)$),
- the ΔR, $\Delta\phi$, and $\cos\alpha$ associated with the lepton-neutrino pair and with the pair of reconstructed W boson and b-tagged jet,
- the p_T, η, ϕ, y, and E of the reconstructed W boson,
- the p_T, η, ϕ, y, m, and E of the reconstructed top quark and of the combinations of the four-momenta of the reconstructed W boson and the photon candidate, of the reconstructed top quark and the photon candidate, and of the reconstructed top quark and the forward jet,
- the ΔR, $\Delta\phi$, $\Delta\eta$, and $\cos\alpha$ associated with the pair of the reconstructed W boson and photon candidate, the pair of the reconstructed top quark and photon candidate, and the pair of the reconstructed top quark and forward jet.

While all low-level features are used for the initial list of input features for the NN models, only the subset of the high-level features characterized by large discriminative powers is selected. The total discriminative power of a feature is given by its individual discriminative power, i.e., by differences between signal and background events in the distribution of its values, and by differences between signal and background events in the correlations with other features. While a measure for the former can typically be defined in a straightforward fashion, the latter is challenging to assess, as the correlations may have complex forms. Therefore, ML models are used for the selection of high-level features as these are capable of learning complex correlations between input features, while not assuming their functional form, and also exploit the individual discriminative power. The "feature importance" is used as a measure of the contribution of individual features to the discriminative power of an ML model. High-level features with large contributions are considered to be the most important ones for the model and are selected for the initial input feature lists.

Multiple methods exist for evaluating the feature importance, and the choice of the method depends on the type of the ML model. For NN models, it is a challenging task to evaluate the feature importance in a direct fashion, as the inner workings, e.g., how and when features are used and interact within complex NN models, are difficult to understand. Instead of using NN models, the feature importance is

evaluated with GBDT models that are fitted to the simulated signal and background events. Like the NN models, these GBDT models perform binary classification using logistic regression. There are mainly two advantages that motivate the choice of GBDT models for this task. First, it is straightforward to comprehend whether and how features are used in GBDT models, which enables a direct computation of the feature importance. Second, the number of input features is not a model parameter, i.e., in contrast to NN models, using longer lists of input features does not inherently increase the complexity of the model.

The selection of candidate high-level features for the initial lists of input features is based on the feature importance rankings of GBDT models fitted to a preliminary set of simulated events. This selection procedure is documented in Ref. [7] and was performed by the author of that work. The initial input feature lists of low-level and high-level features are provided in Table 13 for each SR. The validation of this selection of high-level features, using GBDT models fitted to the finalized set of simulated events, is discussed in Appendix E.1.

Table 13: Lists of low-level and high-level features used as initial input features for the NN models in the 0fj SR and ≥ 1fj SR. The "X" indicates whether the input feature is used in the respective SR.

	0fj SR	≥1fj SR
Low-level features		
p_{T}, η, ϕ of ℓ, γ, and b	X	X
$E_{\mathrm{T}}^{\mathrm{miss}}$	X	X
ϕ^{miss}	X	X
PID^{ℓ}	X	X
b-tagging bin	X	X
lj=fj?		X
$p_{\mathrm{T}}^{\mathrm{fj}}$, η^{fj}, ϕ^{fj}		X
High-level features		
$\Delta\eta(b,\gamma)$	X	X
$m_{\mathrm{T}}(\ell,\nu)$, $m_{\mathrm{T}}(\ell,\nu,\gamma)$, $m_{\mathrm{T}}(\ell,\nu,b)$	X	X
m_t	X	X
$p_{\mathrm{T}}^{b+\gamma}$	X	X
$m_{\ell b}$	X	X
$m_{b\gamma}$	X	X
$p_{\mathrm{T}}^{t+\gamma}$	X	X
$\Delta\eta(b,\ell)$	X	X
$\Delta R(b,\ell)$, $\Delta R(\gamma,\ell)$	X	X
$\Delta R(\mathrm{fj},\ \ell)$, $\Delta R(\gamma,\mathrm{fj})$		X
$m_{b+\mathrm{fj}}$, $m_{\gamma+\mathrm{fj}}$		X
$m_{\ell+\gamma}$	X	X
$\cos\alpha(\gamma,\mathrm{fj})$		X
$\cos\alpha(\gamma,\mathrm{t})$	X	X
H_{T}	X	X
$\Delta\eta(\mathrm{fj},\gamma)$		X
$E^{\gamma,\mathrm{fj}}$		X
E^{W}	X	X
Total number of input features	29	40

9.3 Benchmark neural networks

The initial lists of input features depicted in Table 13 are used to train benchmark NN models in the 0fj SR and ≥ 1fj SR. All NN models introduced in the following share a common setup, differing only in their hidden-layer architecture, which refers to the number of hidden layers and the number of neurons in the hidden layers. An overview of this setup is given in the following.

The NN models are implemented in Keras [268] with the TensorFlow [269] back end. They consist of n layers where the k-th layer is denoted by the row vector α^k, whose entries are referred to as neurons. The first layer α^1 consists of m neurons. Each of the m input features is assigned to exactly one of these neurons. For each event in the input data, the i-th neuron takes the scaled value x_i of the i-th input feature. The value of x_i is obtained by Equation (73). Here, f_i is the raw value of the input feature and $f_i^{\max}$ and $f_i^{\min}$ are the maximum and minimum values of the feature across all simulated signal and background events that the model is fitted to. This scaling ensures that all feature values are in between 0 and 1, while conserving relative differences between values of the same feature.

$$x_i = \frac{f_i - f_i^{\min}}{f_i^{\max} - f_i^{\min}} \tag{73}$$

All layers except the first and last layers are referred to as hidden layers, i.e., the NN models consist of $n - 2$ hidden layers. The k-th hidden layer consists of $\mathcal{N}_k$ neurons, where $\mathcal{N}_k$ is a user-defined quantity. The value of the j-th neuron is given by applying an activation function g to a linear combination of the values of all neurons in the $(k-1)$-th layer, as shown in Equation (75). Here, ω_{ij}^k and b_j^k are referred to as NN weight and NN bias, respectively.

$$\alpha^1 = \begin{pmatrix} x_0 & x_1 & \cdots & x_m \end{pmatrix} \tag{74}$$

$$\alpha_j^k = g\left(\sum_i \alpha_i^{k-1}\,\omega_{ij}^k + b_j^k\right) \tag{75}$$

$$\alpha_0^n = \mathrm{NN_{out}} = h\left(\sum_i \alpha_i^{n-1}\,\omega_{i0}^n + b_0^n\right) \tag{76}$$

$$h(x) = \frac{1}{1 + e^{-x}} \tag{77}$$

The Leaky Rectified Linear Unit (LeakyReLU), which is depicted in Equation (78), is chosen as activation function. The last layer α^n consists of a single neuron α_0^n. Its value is determined by using Equation (76), i.e., in the same fashion as for neurons in the hidden layers except that the LeakyReLU function is replaced by the sigmoid function h given by Equation (91). Therefore, the output values ($\mathrm{NN_{out}}$) of the NN models represent the confidence that events belong to the $tq\gamma$ process.

$$g(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0.05\,x, & \text{if } x < 0 \end{cases} \tag{78}$$

The NN weights and NN biases are the free parameters of the NN models and the number of free parameters is given by $\sum_{k=2}^{n}(\mathcal{N}_{k-1} + 1) \cdot \mathcal{N}_k$, where $\mathcal{N}_k$ is the number of neurons in the k-th layer. The prediction error of the NN models is evaluated by the binary cross entropy as the loss function shown in Equation (90), where w_i and y_i are the weight and the target value of the i-th of N events in the input data.

$$\mathcal{B} = -\sum_i^N w_i \cdot \Big[y_i \log(\mathrm{NN_{out}}) + (1 - y_i) \log\left(1 - \mathrm{NN_{out}}\right) \Big] \tag{79}$$

The NN models are fitted in sequential steps, which are referred to as epochs. At the start of each epoch, the input data is randomly shuffled and split into equally sized batches, where the batch size is constant

throughout the fitting. In each epoch, the NN models are sequentially evaluated on each of the batches. The values of the free parameters, i.e., the NN weights and NN biases, are updated after each of these evaluations using the Adam optimizer [270]. This is a stochastic gradient descent method with the objective of minimizing the loss function by employing the gradient of the loss function with respect to the free parameters at their current values ("backpropagation"). The number of updates of the free parameters, referred to as fitting steps, of the NN models is hence given by the number of batches times the number of epochs.

Before performing the fitting, the NN weights of each layer are initialized by using the Glorot uniform initializer [271], while all NN biases are set to zero. Furthermore, the input dataset is split into two statistically independent datasets using stratified random sampling, a training dataset and a test dataset. The former contains 80% of the input data and is used in the fitting of the NN models, and a test dataset, which is solely used for performance evaluation in the optimization of the NN models discussed in Sections 9.4 and 9.5. The $tq\gamma$ event weights are scaled, so that the sum of $tq\gamma$ event weights equals the sum of background event weights.

The two benchmark models fitted separately in the 0fj SR and $\geq$ 1fj SR share a common hidden-layer architecture. Three hidden layers are used, consisting of 48, 24, and 15 neurons, respectively.[33] The fitting of both NN models is performed in 1,000 epochs using a batch size of 10,000 events. A validation dataset, containing 25% of the training dataset, is obtained via stratified random sampling and used for performance evaluation and monitoring of the fitting progress. The benchmark NN models are fitted to the remaining 75% of the training dataset.

Figure 43 shows the values of the binary cross entropy divided by the sum of event weights ($\mathcal{B}_{\text{norm}}$) as a function of the number of completed epochs for the training and validation datasets in both SRs. The bands correspond to the uncertainty due to the limited size of the corresponding dataset. This uncertainty is the standard deviation of the values of $\mathcal{B}_{\text{norm}}$ obtained from 100 PDs. These PDs are constructed by bootstrap sampling in the same way as described in Section 5.3.1. The $\mathcal{B}_{\text{norm}}$ values decrease monotonically for both datasets in both SRs, and the values for the training and validation datasets agree within their uncertainties. Therefore, the performance of the NN models, i.e., the power of separating $tq\gamma$ from background events, increases steadily during the fitting process. The small differences between the performance for the training and validation datasets indicate that the performance increase is mainly based on general patterns that are different for $tq\gamma$ and background events, which are learned by the NN models. However, the difference between the $\mathcal{B}_{\text{norm}}$ values for the different datasets starts to increase after about 700 epochs in the 0fj SR and about 500 epochs in the $\geq$ 1fj SR, while the values for the validation datasets stop decreasing significantly thereafter. This indicates that the NN models are starting to adapt to patterns in the training datasets that are not present in the validation datasets, which may be caused by statistical fluctuations. Therefore, the number of epochs is considered to be sufficiently large as no further significant improvement on the validation dataset is expected.

Furthermore, Figure 43 shows the fraction of $tq\gamma$ events and background events as a function of NN_{out} separately for the training and validation datasets for both SRs. The ratio panels show the differences in the event fractions between the training and validation datasets (val − train) relative to the uncertainty due to the limited number of simulated events ($\sigma_{\text{MC stat.}}$). The NN models are found to discriminate well between $tq\gamma$ and background events, as the fraction of $tq\gamma$ events is significantly larger than that of background events for $\text{NN}_{\text{out}} > 0.5$. The difference between the fractions of $tq\gamma$ and background events increases monotonically with increasing NN_{out} values in this region. The opposite behavior is observed for $\text{NN}_{\text{out}} < 0.5$. The values shown in the ratio panels fluctuate around zero, with only one value larger than two across all distributions, indicating a good agreement between equivalent distributions of the training and validation datasets. Here, equivalent distributions refer to distributions of the same quantity of the same process type (e.g., the NN_{out} distributions for $tq\gamma$ events) obtained from statistically independent datasets (e.g., the training and validation datasets).

[33] This hidden-layer architecture represents an initial guess, which is found to provide a good performance for both NN models.

The agreement of equivalent distributions from the training and validation datasets is quantified by performing two-sample Kolmogorov-Smirnov (2-KS) tests of the hypothesis that the equivalent distributions from the different datasets can be drawn from the same p.d.f.. Small p-values of these tests indicate that the shapes of the distributions disagree significantly.[34] This would likely indicate that the corresponding NN model suffers from significant overfitting, i.e., the model has adapted to patterns in the training dataset that are not general features of the input data. All p-values of the 2-KS tests are larger than 0.23 and, therefore, no significant disagreement between equivalent distributions from the different datasets and no hint for significant overfitting are found.[35]

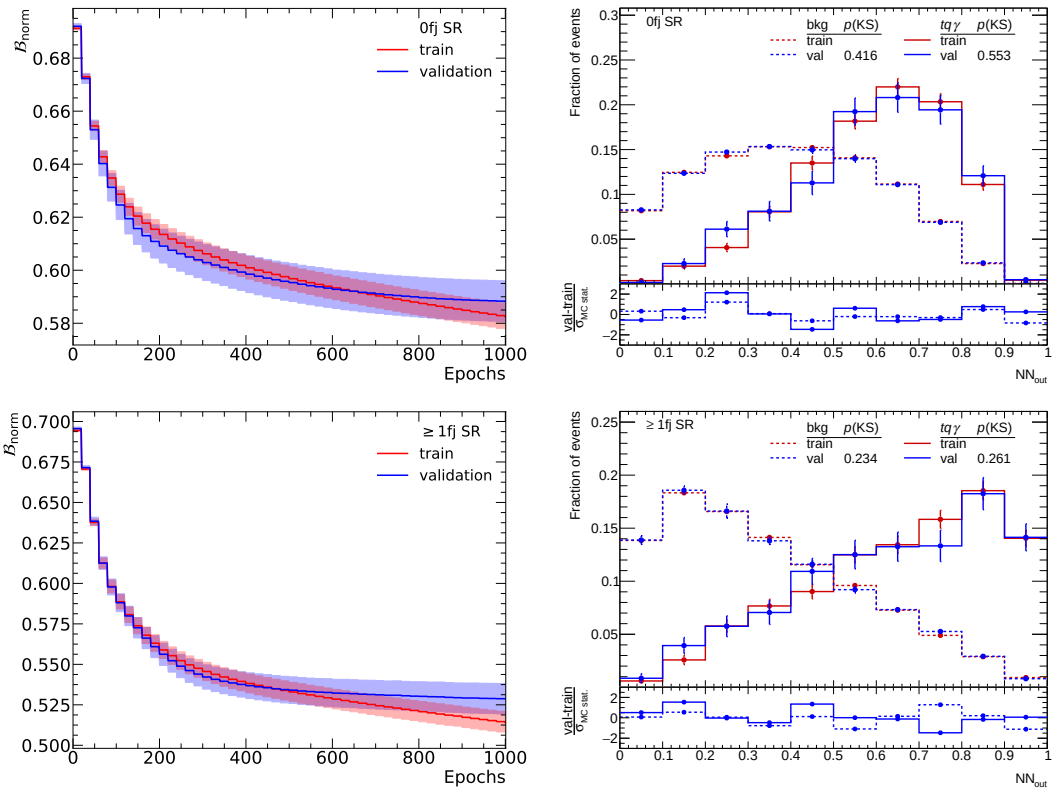


Figure 43: The figures on the left show $\mathcal{B}_{\text{norm}}$ as a function of the number of completed epochs ("Epochs") for the training and validation datasets. The uncertainty bands correspond to the statistical uncertainty evaluated using bootstrap sampling. The figures in the right column display the fraction of events as a function of NN_{out} separately for $tq\gamma$ ("sig") and background ("bkg") events from the training ("train") and validation datasets ("val"). The uncertainty bars correspond to the uncertainty due to the limited number of simulated events. Furthermore, the p-values of 2-KS tests (p(KS)) between equivalent distributions from the training and validation datasets are displayed in the legend. The ratio panels show the differences in the event fractions between equivalent distributions of the training and validation datasets (val−train) relative to the total uncertainty ($\sigma_{\text{MC stat.}}$) due to the limited number of simulated events of both datasets. The upper and lower rows show the distributions for the 0fj SR and ≥ 1fj SR, respectively.

9.4 Removal of input features

The computation time of the benchmark NN models is small and the performance investigations of the fitted benchmark NN models concluded that these models offer a reliable and powerful discrimination

[34] Typically, values below 0.05 are considered to be small, which correspond to a statistical significance of $> 1.6\sigma$ for one-tailed statistical hypothesis tests.

[35] It should be noted that overfitting can, in general, not be fully avoided as statistical models will adapt to statistical fluctuations to some degree in the fitting procedure. However, for statistical models to provide reliable predictions, it is tested whether this discriminative power is mainly based on general patterns and whether the contribution of the statistical fluctuations is small by comparing the performance of the fitted models on the input data and on statistically independent data.

of $tq\gamma$ and background events. Consequently, the initial lists of input features are considered to be well-suited. However, there are two main advantages of further reducing the lists to minimal sets of input features, while conserving most of the discriminative power. First, it is crucial to ensure that all input features are well-modeled in the simulated events used for fitting the models, i.e., that the simulation accurately reproduces the corresponding distributions in the collected data. Otherwise, the NN models will be biased toward the simulation and adapt to patterns that are not present in the collected data, which prevents precise and reliable measurements of the $tq\gamma$ process using the $\mathrm{NN_{out}}$ distributions. Furthermore, the NN models may be highly susceptible to systematic uncertainties, i.e., the difference in the $\mathrm{NN_{out}}$ distributions between the nominal prediction and a systematic uncertainty is unexpectedly large relative to the corresponding difference for individual features. If this is the case, it is mandatory to identify the origin of the difference. Both of these tasks are less complex with a smaller number of features, making it feasible to perform thorough investigations of each one.

An iterative procedure is used to disregard input features that do not significantly contribute to the discriminative power of the NN models, where the feature with the lowest feature importance among the remaining ones is removed from the input list after each iteration. In contrast to the GBDT models, the feature importance is calculated in an indirect fashion using a leave-one-feature-out (LOFO) algorithm, referred to as LOFO importance. The choice of the metric to evaluate the LOFO importance represents a key element of this iterative procedure and is discussed in the following.

As one of the main goals of this work is the observation of $tq\gamma$, the statistical significance Z defined in Equation (62) of the discovery test obtained from binned BPL fits to the $\mathrm{NN_{out}}$ distributions is chosen as the metric. The $\mathrm{NN_{out}}$ distributions are divided into five equally-sized bins in the range between 0 and 1. This default binning configuration is an ad-hoc choice and is used throughout this chapter. The BPL is given by Equation (50), so that only statistical uncertainties are considered in the BPL fits. The number of events $m^{b,r}(\mu_{tq\gamma})$ is given by the predicted sum of $tq\gamma$ and background contributions for each bin and only depends on $\mu_{tq\gamma}$, i.e., the predictions for all background contributions are fixed. The performance of intermediate NN models is not evaluated on the collected data as this would bias the design of the NN models. Therefore, the observed number of events $n^{b,r}$ is extracted from an Asimov dataset, which is defined by $n^{b,r} = m^{b,r}(\mu_{tq\gamma} = 1)$, i.e., the observed number of events is given by the predicted number of events, where the $tq\gamma$ contribution is given by its nominal prediction. As an example, Figure 44 shows the predicted $\mathrm{NN_{out}}$ distributions corresponding to the benchmark NN model in the $\geq$ 1fj SR for $tq\gamma$ and background events from the validation dataset and the corresponding Asimov dataset. The predicted distributions are shown before ("Pre-Fit") and after ("Post-Fit") performing a BPL fit to the Asimov dataset. The uncertainty band corresponds to the uncertainty associated with $\mu_{tq\gamma}$, while the uncertainties on the Asimov dataset are the statistical uncertainties (Poisson uncertainties) of the yields. The value of $\mu_{tq\gamma}$ after performing the BPL fit is 1.0 ± 0.06.

A detailed description of the iterative feature-removal procedure is given in the following. At the start of each iteration, the training and validation datasets are combined and split into four equally-sized subsets using stratified random sampling. Next, the following steps are executed for each of the m input features.

1. The input feature is temporarily excluded from the list of input features.
2. The NN model is fitted using four-fold cross-validation, where each of the four subsets is used as a validation dataset once.
3. The statistical significance is evaluated for each of the four fitted NN models on the corresponding training (Z_{train}), validation (Z_{val}), and test (Z_{test}) datasets.
4. The mean value is computed for each set of Z_{train}, Z_{val}, and Z_{test} values.
5. The LOFO importance is given by $1/Z_{\text{test}}^{\text{mean}}$. This ensures that features whose exclusion leads to a significant decrease in the NN model performance have a high LOFO importance.

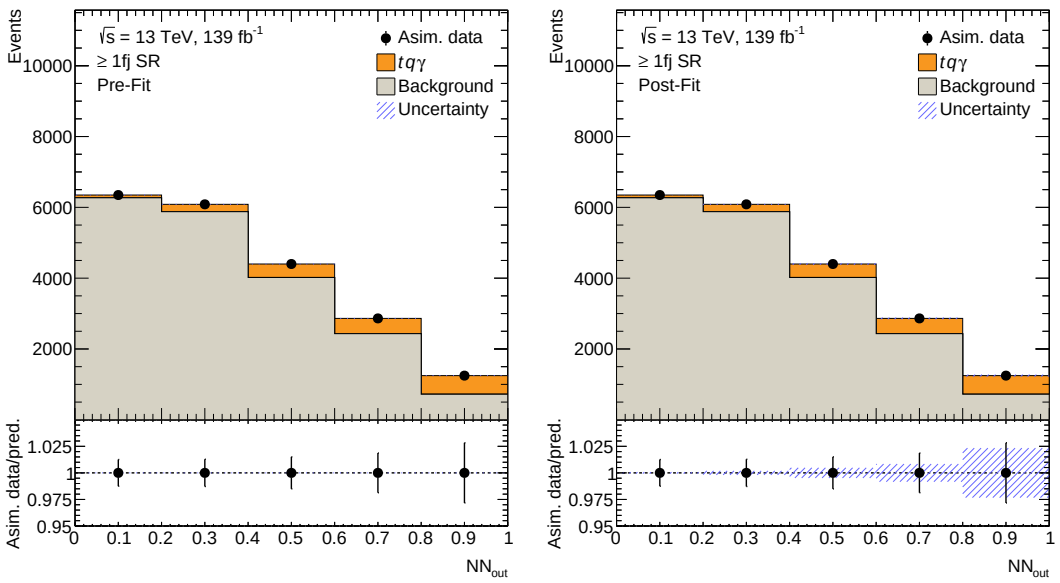


Figure 44: NN_{out} distributions of $tq\gamma$ and background events from the validation dataset in the $\geq$ 1fj SR corresponding to the benchmark NN model before (Pre-Fit) and after (Post-Fit) performing a BPL fit to the Asimov dataset.

At the end of each iteration, the feature with the lowest LOFO importance is permanently removed from the list of input features, and the next iteration is executed considering only the $n-1$ remaining input features.

Figure 45 represents an illustration of the progress of the iterative feature-removal procedure for both SRs. The i-th tick on the x-axis corresponds to the i-th iteration. Its label represents the input feature that is removed in the i-th iteration. The values of $Z^{\text{mean}}_{\text{train}}$, $Z^{\text{mean}}_{\text{val}}$, and $Z^{\text{mean}}_{\text{test}}$ for the i-th iteration, displayed on the y-axis, correspond to the values obtained from the NN model when that feature is removed. The differences between the values of $Z^{\text{mean}}_{\text{train}}$, $Z^{\text{mean}}_{\text{val}}$, and $Z^{\text{mean}}_{\text{test}}$ decrease as the iterative feature removal-procedure progresses, especially in the $\geq$ 1fj SR. This indicates that the reduction of the complexity of the NN models by removing less important features improves the generality of the performance and limits the overfitting of the NN models.

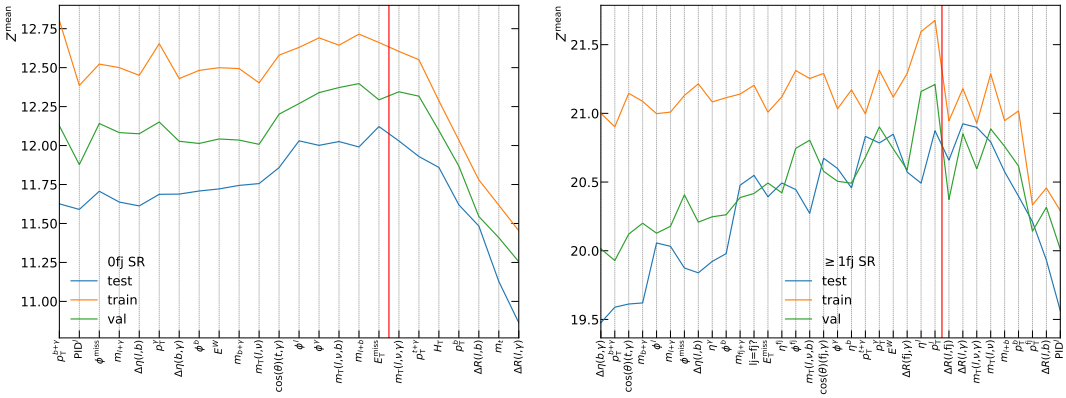


Figure 45: Illustration of the iterative feature-removal procedure in the 0fj SR (left) and the $\geq$ 1fj SR (right). The i-th tick on the x-axis corresponds to the i-th iteration. Its label represents the input feature that is removed in the i-th iteration. The values of $Z^{\text{mean}}_{\text{train}}$, $Z^{\text{mean}}_{\text{val}}$, and $Z^{\text{mean}}_{\text{test}}$ of the i-th iteration displayed on the y-axis correspond to this removed feature. The vertical red line displays the chosen stopping points, i.e., variables listed on the left of the line are ultimately removed from the list of input features. The lines between Z values of adjacent iterations represent linear interpolations between the corresponding points and are shown for illustrative purpose only.

In the 0fj SR, the values of $Z^{\text{mean}}_{\text{train}}$ change by less than 3% up to the 17th iteration ($E^{\text{miss}}_{\text{T}}$) indicating that the removal of features does not lead to a significant drop of the NN model performance. The corresponding values of $Z^{\text{mean}}_{\text{val}}$ change by less than 1.4%, while $Z^{\text{mean}}_{\text{test}}$ slightly improves by 4.3%. The values of $Z^{\text{mean}}_{\text{test}}$ decrease steadily in the subsequent iterations and, hence, the 17th iteration is chosen as the stopping point of the iterative procedure, i.e., the feature list as given after the 17th iteration is used as the final list of input features.

In the ≥ 1fj SR, the values of $Z^{\text{mean}}_{\text{train}}$ change by less than 1.5% up to the 23rd iteration ($\Delta R(\text{fj}, \gamma)$) and hence no significant drop in the NN model's performance is observed either. The corresponding values of $Z^{\text{mean}}_{\text{test}}$ and $Z^{\text{mean}}_{\text{val}}$ improve by 6% and 3%, respectively. The value of $Z^{\text{mean}}_{\text{test}}$ reaches its maximum value in the 27th iteration ($\Delta R(l, \gamma)$), which corresponds to an improvement of 7.4%, and drops steadily in the subsequent iterations. However, the value of $Z^{\text{mean}}_{\text{val}}$ reaches its maximum in the 25th iteration (p^{l}_{T}), which corresponds to an improvement of 6%, and is 2% smaller in the 27th iteration. As the value of $Z^{\text{mean}}_{\text{test}}$ is only insignificantly smaller ($\approx 0.2\%$) in the 25th iteration relative to its maximum, the 25th iteration is chosen as the stopping point of the iterative procedure. Consequently, the final input feature lists for the NN models in the 0fj SR and the ≥ 1fj SR consist of 12 and 15 input features, respectively. Table 14 lists all variables that are removed in both SRs. Notably, 14 of the 17 variables removed in the 0fj SR are also removed in the ≥ 1fj SR. This large overlap implies that the corresponding features consistently contribute little to the performance of the NN model across both SRs.

Table 14: Lists of variables that are excluded as input features depicted in Table 13 for the 0fj SR and ≥ 1fj SR. Variables are highlighted that are removed for both SRs. The numbers represent the number of iterations after which the corresponding feature is excluded.

0fj SR			≥ 1fj SR			
1. $\boldsymbol{p^{b+\gamma}_{\text{T}}}$	8. $\boldsymbol{\phi^{b}}$	15. $\boldsymbol{m_{\text{T}}(\ell, \nu, b)}$	1. $\boldsymbol{\Delta\eta(b,\gamma)}$	8. $\boldsymbol{\Delta\eta(\ell,b)}$	15. ϕ^{fj}	22. $\boldsymbol{E^{W}}$
2. PID^{ℓ}	9. $\boldsymbol{E^{W}}$	16. $m_{\ell+b}$	2. $\boldsymbol{p^{b+\gamma}_{\text{T}}}$	9. η^{γ}	16. $\boldsymbol{m_{\text{T}}(l,\nu,b)}$	23. $\Delta R(\text{fj}, \gamma)$
3. $\boldsymbol{\phi^{\text{miss}}}$	10. $\boldsymbol{m_{b+\gamma}}$	17. $\boldsymbol{E^{\text{miss}}_{\text{T}}}$	3. $\boldsymbol{\cos(\theta)(t,\gamma)}$	10. $\boldsymbol{\phi^{b}}$	17. $\cos(\theta)(\text{fj}, \gamma)$	24. η^{ℓ}
4. $\boldsymbol{m_{\ell+\gamma}}$	11. $m_{\text{T}}(\ell, \nu)$	-	4. $\boldsymbol{m_{b+\gamma}}$	11. $m_{\text{fj}+\gamma}$	18. $\boldsymbol{\phi^{\gamma}}$	25. p^{l}_{T}
5. $\boldsymbol{\Delta\eta(\ell,b)}$	12. $\boldsymbol{\cos(\theta)(t,\gamma)}$	-	5. $\boldsymbol{\phi^{\ell}}$	12. lj=fj?	19. η^{b}	-
6. $\boldsymbol{p^{\gamma}_{\text{T}}}$	13. $\boldsymbol{\phi^{\ell}}$	-	6. $\boldsymbol{m_{\ell+\gamma}}$	13. $\boldsymbol{E^{\text{miss}}_{\text{T}}}$	20. $p^{t+\gamma}_{\text{T}}$	-
7. $\boldsymbol{\Delta\eta(b,\gamma)}$	14. $\boldsymbol{\phi^{\gamma}}$	-	7. $\boldsymbol{\phi^{\text{miss}}}$	14. η^{fj}	21. $\boldsymbol{p^{\gamma}_{\text{T}}}$	-
17			25			
14 in common						

Figure 46 shows the predicted and observed distributions of different final NN input features as examples in the SRs and CRs. The predicted distributions include the prediction for the fake-lepton contributions. The distributions of m_t are shown for all SRs and CRs, while the distributions of η^{b} and η^{γ} are shown for the 0fj SR, and those of $m_{b+\text{fj}}$ and H_T are shown for the ≥ 1fj SR. The predicted distributions agree well with the observed ones within the uncertainties, as the majority of observed values are contained within the uncertainty in the prediction. Furthermore, no trends are observed in the ratio of observed ("Data") and predicted distributions ("Pred."). The predicted total contributions differ by 1.7%, 5.8%, 4.6%, and 2.6% from the observed yields for the ≥ 1fj SR, 0fj SR, $t\bar{t}\gamma$ CR, and $W\gamma$ CR, respectively. These differences are smaller than 0.73 times the total uncertainty in the prediction and are therefore insignificant. Additionally, the overall contributions of the signal, $t\bar{t}\gamma$, and $W\gamma$+jets processes are treated as free parameters in the statistical analysis. In the distributions shown in Figure 46, the contributions of the three processes correspond to the nominal predictions and the uncertainty in the predictions does not include uncertainties in the prediction of their overall contributions. Therefore, slight differences between the predicted and observed overall contributions are not unexpected. The distributions of all NN input features are displayed in Appendix E.2 for the SRs and CRs. None of the distributions show significant differences in the shape between the predicted distributions and the observed distributions. Therefore, all NN input features are considered to be well-modeled.

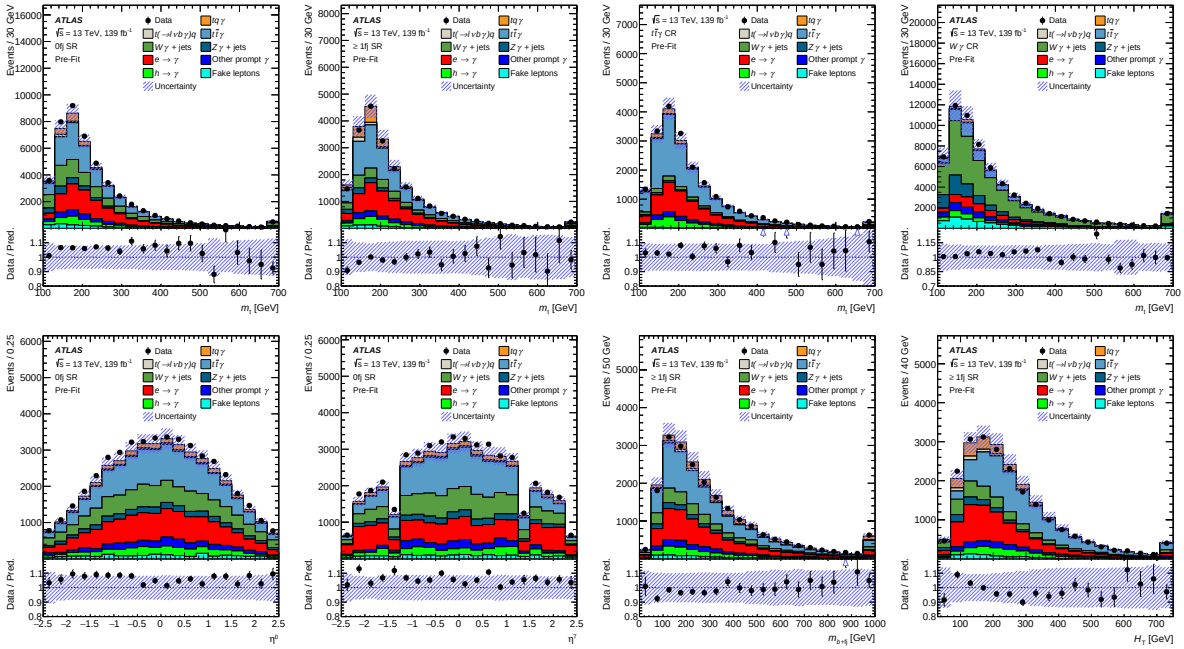


Figure 46: Predicted and observed distributions of different final NN input features. The distributions of m_t are shown for all SRs and CRs. The distributions of η^b and η^γ are shown for the 0fj SR, and those of $m_{b+\mathrm{fj}}$ and H_T are shown for the ≥ 1fj SR. The last bins of the m_t, $m_{b+\mathrm{fj}}$, and H_T distributions include the overflow, while the first bins of the m_t and $m_{b+\mathrm{fj}}$ include the underflow. The ratio panels show the ratio of the observed ("Data") and predicted distributions ("Pred."). The hashed bands correspond to the total uncertainties in the prediction, considering all systematic uncertainties introduced in Section 7.6. Except for the m_t distributions in the $t\bar{t}\gamma$ CR, all distributions are taken from Ref. [10].

9.5 Optimization of hidden-layer architecture

After the finalization of the input feature lists of the NN models, the hidden-layer architectures of the NN models are optimized. However, there is an infinite number of choices for the architecture of an NN model, while only architectures with a number of NN parameters significantly larger than the amount of available input data can be immediately excluded. Practically, only a limited number of these possibilities can be tested. For the optimization of these NN models, 774 architectures are tested. The number of hidden layers is varied between two and four, while the number of neurons in the hidden layers is varied as follows:

- First hidden layer: 48, 64, 128
- Other hidden layers: 8, 16, 32, 64, 128, 256.

The NN models are fitted using each of these configurations in both SRs. Their performance is evaluated by stratified four-fold cross-validation as in the previous section. The different configurations are ranked by $Z^{\mathrm{mean}}_{\mathrm{test}}$. The three pairwise differences between $Z^{\mathrm{mean}}_{\mathrm{train}}$, $Z^{\mathrm{mean}}_{\mathrm{val}}$, and $Z^{\mathrm{mean}}_{\mathrm{test}}$ must satisfy Equation (80). Configurations that do not fulfill this criterion are disregarded to ensure that only configurations offering consistent performance across the three statistically independent datasets are considered.

$$|Z^{\mathrm{mean}}_{\mathrm{i}} - Z^{\mathrm{mean}}_{\mathrm{j}}| < 1\sigma \quad \text{with i}\neq\text{j and i,j=\{train,val,test\}} \tag{80}$$

This criterion is passed by 438 and 122 architectures in the 0fj SR and ≥ 1fj SR, respectively. Table 15 lists the 15 hidden-layer architectures providing the largest values of $Z^{\mathrm{mean}}_{\mathrm{test}}$ for each SR that fulfill the aforementioned criterion in ascending order of $Z^{\mathrm{mean}}_{\mathrm{test}}$.

Table 15: Lists of the 15 hidden-layer architectures of the NN models that provide the largest value of $Z_{\text{test}}^{\text{mean}}$ for each SR and of the corresponding $Z_{\text{test}}^{\text{mean}}$ values. The rows that are highlighted by red margins correspond to the chosen architectures.

≥ 1fj SR		0fj SR	
architecture	$Z_{\text{test}}^{\text{mean}}$	architecture	$Z_{\text{test}}^{\text{mean}}$
64,8,32,8	20.76 ± 0.17	64,32,8	12.11 ± 0.06
128,16,8,8	20.76 ± 0.11	64,32,8,64	12.11 ± 0.06
64,32,64,8	20.77 ± 0.18	128,32,8	12.11 ± 0.10
128,16,64,16	20.77 ± 0.17	128,256,16	12.12 ± 0.05
48,8,32,8	20.77 ± 0.25	128,16,256	12.12 ± 0.09
128,8,8,32	20.78 ± 0.27	128,16,16,128	12.12 ± 0.11
48,32,16	20.79 ± 0.13	64,16,8,8	12.12 ± 0.08
64,8,16	20.80 ± 0.21	128,32,64	12.12 ± 0.07
128,16,8,64	20.84 ± 0.20	128,16,16	12.13 ± 0.06
128,32,16,16	20.85 ± 0.23	128,64,8	12.13 ± 0.09
128,8,16	20.85 ± 0.24	64,32,8,8	12.13 ± 0.10
128,8,16,32	20.85 ± 0.30	128,16,64,8	12.13 ± 0.09
128,8,32,8	20.87 ± 0.17	128,256,16,8	12.13 ± 0.07
48,64,8,8	20.89 ± 0.27	64,32,64	12.13 ± 0.14
128,8,32	20.89 ± 0.15	128,64,32,32	12.14 ± 0.01

The number of parameters of the NN models associated with these hidden-layer architectures is typically larger for the 0fj SR than for the ≥ 1fj SR. This is expected as the 0fj SR offers larger numbers of simulated events as shown in Table 12, which allows for a higher level of NN model complexity. In both regions, no models with only two hidden layers are among those with the highest performance. There is no clear performance difference between configurations with three or four hidden layers. In both regions, configurations with the largest number of neurons (128) in the first layer tend to provide a better performance, while no trend is observed for the number of neurons in the other hidden layers. Furthermore, the performance difference between the listed configurations is small for each region and covered by the uncertainty on $Z_{\text{test}}^{\text{mean}}$. Models with three hidden layers are preferred over those with four, as they are considered less complex. If multiple three-layer architectures yield the same value of $Z_{\text{test}}^{\text{mean}}$, the configuration with the fewest parameters is chosen. The chosen configurations are highlighted in Table 15 and correspond to 3,401 and 4,017 free parameters of the NN models in the ≥ 1fj and 0fj SR, respectively.

Table 16 shows the value of $Z_{\text{test}}^{\text{mean}}$ evaluated with stratified four-fold cross-validation for the NN models at the different optimization stages, i.e., for the benchmark NN models ("Benchmark"), for the NN models using the final input feature lists ("Feature removal") and the architecture of the benchmark models, and for the NN models using the final input feature lists and the optimized hidden-layer architectures ("Architecture optimization"). The performance of the NN models improves after each of the optimization steps. While the improvements of $Z_{\text{test}}^{\text{mean}}$ from the architecture optimization in both SRs and from the feature removal in the 0fj SR are small (≈ 1%), the improvement of $Z_{\text{test}}^{\text{mean}}$ from the feature removal in the ≥ 1fj SR is large (≈ 10%). This improvement is mainly caused by the improved generalization ability of the NN models.

9.6 Performance of the optimized NN models

The NN models are fitted to a subset of the input data using the variables listed in Table 17 as input features and the optimized hidden-layer architectures that are highlighted in Table 15. Distributions of these final fitted NN models are used in the statistical analysis. To avoid any overfitting bias, any

Table 16: Comparison of the values of $Z_{\text{test}}^{\text{mean}}$ for the NN models at the different optimization stages in both SRs, i.e., for the benchmark NN models ("Benchmark"), for the NN models using the final input feature lists ("Feature removal"), and for the NN models using the final input feature lists and the optimized hidden-layer architectures ("Architecture optimization").

Optimization step	0fj SR	$\geq$ 1fj SR
Benchmark	11.85 ± 0.04	18.67 ± 0.13
Feature removal	11.99 ± 0.07	20.70 ± 0.16
Architecture optimization	12.13 ± 0.06	20.89 ± 0.15

simulated data that the NN models are fitted to is not used in the statistical analysis. However, the statistical precision in the prediction is decreased as a consequence. The final NN models are fitted to training datasets containing only 20% of the input data to minimize this decrease, while ensuring that the performance of the NN models does not significantly drop. Their performances are evaluated on validation datasets, which consist of the remaining 80% of the input events and are used for obtaining the predicted $tq\gamma$ and background NN_{out} distributions. The batch size in the $\geq$ 1fj SR is adjusted to 5,000, as the previous size of 10,000 events is found to provide an insufficient number of fitting steps. Increasing the number of epochs, which is an alternative to decreasing the batch size for increasing the number of fitting steps, leads to worse performance. A detailed analysis of the NN models is provided in the following, including an analysis of the input features.

The separation strength $\mathcal{S}$ is used as a measure for the individual discriminative power of NN_{out} and the input features and is computed by Equation (81). This metric can take values between 0 and 1, where larger values correspond to stronger discriminative power. The range of each feature is divided into 100 bins of variable width, so that each bin contains approximately 1% of the total $tq\gamma$ contribution. The fractions of $tq\gamma$ and background events contained in the i-th bin are denoted by s_i and b_i, respectively.

$$\mathcal{S} = \sum_{\text{bins}} \frac{(s_i - b_i)^2}{2\,(s_i + b_i)} \tag{81}$$

Table 17 lists the values of $\mathcal{S}$ in both SRs for each of the respective input features. In both SRs, m_t has the largest $\mathcal{S}$ value, with 0.104 in the 0fj SR and 0.09 in the $\geq$ 1fj SR. This is expected as there are several background contributions (e.g., $W\gamma$+jets, $Z\gamma$+jets) that do not involve a top-quark decay and hence the deviation of the reconstructed top-quark mass from the top-quark pole mass is significantly larger on average for these background contributions than for $tq\gamma$. Additionally, m_t discriminates $tq\gamma$ from contributions involving radiative top-quark decays, such as $t\,(\rightarrow \ell\nu b\gamma)\,q$ or $t\bar{t}\gamma$ (dec), as the radiated photon must be included for an accurate top-quark reconstruction for these contributions, but it is neglected in the computation of m_t. Furthermore, for background contributions involving multiple top-quark or W-boson decays, such as $t\bar{t}$, $t\bar{t}\gamma$, or tW, it is less likely that the b-tagged jet, lepton candidate, and reconstructed neutrino are associated with the same top-quark decay than it is for $tq\gamma$. In the 0fj SR, the second-highest $\mathcal{S}$ value of 0.041 is associated with p_{T}^{ℓ} and is significantly smaller than the value of m_t. In the $\geq$ 1fj SR, the $\mathcal{S}$ values of $m_{b+\text{fj}}$, m_{l+b}, and $E^{\text{fj},\gamma}$ are of similar size (0.063–0.064) and are the next highest values.

Figure 47 shows, as examples, the fraction of $tq\gamma$ and of background events in the validation dataset as a function of m_t in the 0fj SR and as a function of $E^{\text{fj},\gamma}$ in the $\geq$ 1fj SR. The ratio panels illustrate the differences between the $tq\gamma$ and background distributions ("sig − bkg") normalized to the statistical uncertainty ($\sigma_{\text{MC stat.}}$) due to the limited number of simulated events. The m_t distributions of $tq\gamma$ and background events peak close to the top-quark pole mass, which is set to 172.5 GeV in the simulation, while the peak of the $tq\gamma$ distribution is significantly sharper and the tail toward larger m_t values is significantly smaller than for background contributions. The values of $E^{\text{fj},\gamma}$ are significantly larger for $tq\gamma$ events on average than for background events. The presence of a highly energetic forward jet is a distinctive feature of $tq\gamma$. Furthermore, photons radiated in the production of a massive particle,

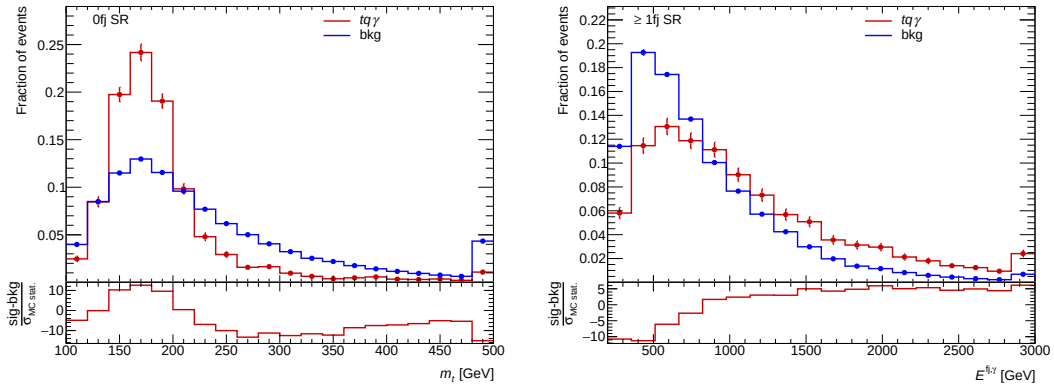


Figure 47: Fraction of $tq\gamma$ ("sig") and background ("bkg") events from the validation dataset as a function of m_t in the 0fj SR (left) and of $E^{\mathrm{fj},\gamma}$ in the $\geq$ 1fj SR (right). The uncertainty bars correspond to the uncertainty due to the limited number of simulated events. The ratio panels show the differences between the $tq\gamma$ and background distributions ("sig − bkg") relative to the uncertainty due to the limited number of events ($\sigma_{\text{MC stat.}}$).

as in $tq\gamma$ production, are expected to be more energetic than those from radiative decays, such as in $t\,(\rightarrow \ell\nu b\gamma)\,q$ production. These two effects lead to the observed differences in the $E^{\mathrm{fj},\gamma}$ distributions.

The distribution of the background rejection $(1 - \varepsilon^{\mathrm{bkg}})$ as a function of the signal efficiency $(\varepsilon^{\mathrm{sig}})$ is a receiver operating characteristic (ROC), which is commonly used as a performance measure of a statistical model in data science.[36] The background efficiency $(\varepsilon^{\mathrm{bkg}})$ and signal efficiency denote the fractions of background and signal events that fulfill $\mathrm{NN}_{\mathrm{out}} > X$, respectively, where X is a threshold value between 0 and 1. A statistical model operates as a perfect binary classifier if its background rejection is 1 for all $\varepsilon^{\mathrm{sig}}$ values. Conversely, a statistical model for binary classification characterized by $\varepsilon^{\mathrm{bkg}} \approx \varepsilon^{\mathrm{sig}}$ performs at the level of random guessing. The area under the ROC curve, $\mathcal{A}$, therefore represents a general measure of the discriminative power of the fitted NN models as binary classifiers, with $\mathcal{A} = 1$ and $\mathcal{A} = 0.5$ corresponding to a perfect binary classifier and to random guessing, respectively.

Figure 48 illustrates the performance of the final NN models in the 0fj SR and $\geq$ 1fj SR. Separately for the training and validation datasets, it shows the values of $\mathcal{B}_{\mathrm{norm}}$ as a function of completed epochs, the ROC curve, and the fraction of signal and background events as a function of $\mathrm{NN}_{\mathrm{out}}$. The performance of both NN models improves after each epoch, as $\mathcal{B}_{\mathrm{norm}}$ monotonically decreases for the training and validation datasets.

In the $\geq$ 1fj SR, the difference in the $\mathcal{B}_{\mathrm{norm}}$ values from the training and validation datasets starts to steadily become larger after $\approx$ 400 epochs. However, the values agree well within the uncertainties throughout the fitting process. A similarly good agreement between the ROC curves is observed. Furthermore, the ROC curves indicate that the NN model provides large discriminative power, as it significantly outperforms random guessing, represented by the black line. For example, a requirement of $\mathrm{NN}_{\mathrm{out}} > X$, which classifies 50% of all $tq\gamma$ events of the validation dataset correctly ($\varepsilon^{\mathrm{sig}} = 50\%$), rejects $\approx$ 90% of the respective background events, i.e., improves the $tq\gamma$-background ratio by a factor of five. Moreover, the equivalent distributions of the fraction of events as a function of $\mathrm{NN}_{\mathrm{out}}$ agree excellently. The differences are at most $1.5\sigma_{\text{MC stat.}}$ across all bins, with no trend observed in the differences. The p-values of the 2-KS tests for the distributions of background and $tq\gamma$ events are 0.432 and 0.332, respectively. This indicates that there is no significant disagreement between the equivalent distributions. In conclusion, the performance of the NN model is consistently high across the statistically independent datasets.

[36] The background rejection corresponds to the "true negative rate", i.e., the fraction of instances of the "negative" class being correctly classified. The signal efficiency corresponds to the "true positive rate", i.e., the fraction of instances of the "positive" class being correctly classified. In this context, the background events correspond to the negative class and the $tq\gamma$ events correspond to the positive class.

In the 0fj SR, the difference of the $\mathcal{B}_{\text{norm}}$ values from the training and validation datasets starts to steadily become larger after $\approx$ 300 epochs. The values of $\mathcal{B}_{\text{norm}}$ agree well within the uncertainties only up to approximately the 500th epoch and significantly disagree at the end of the fitting process. This is also reflected in the ROC curves, which disagree significantly in the range $70\% < \varepsilon^{\text{sig}} < 95\%$. However, the ROC curves agree well outside this range and indicate that the NN model provides large discriminative power. It rejects $\approx$ 82% of all background events in the validation dataset for $\varepsilon^{\text{sig}} = 50\%$, which corresponds to an improvement of the $tq\gamma$-background ratio by a factor of 2.8.

The distributions of the fraction of background events as a function of NN_{out} from the different datasets agree well within the uncertainties for the 0fj SR. There is no trend in the differences, where only three of the ten differences are larger than $1.5\sigma_{\text{MC stat.}}$, with one outlier at $2.4\sigma_{\text{MC stat.}}$. The p-value of the corresponding 2-KS test is 0.469. Therefore, the NN_{out} distributions for background events are consistent across the training and validation datasets. The distributions of the fraction of signal events from the different datasets agree well within the uncertainties for $\text{NN}_{\text{out}} > 0.4$. No trend in the differences is observed, which remain below $2\sigma_{\text{MC stat.}}$ across all bins in this region. In contrast, the agreement of the distributions is significantly worse for $\text{NN}_{\text{out}} \leq 0.4$, as the differences are between $1.2\sigma_{\text{MC stat.}}$ and $2\sigma_{\text{MC stat.}}$ across all bins in this region. This region contains $\approx$ 20% of the $tq\gamma$ events and thus only affects the part of the ROC curve corresponding to $\varepsilon^{\text{sig}} > 80\%$. It causes the discrepancies between the $\mathcal{B}_{\text{norm}}$ values and the ROC curves from the different datasets, and the low p-value of 0.064 of the corresponding 2-KS test. This p-value corresponds to a statistical significance of 1.55σ, representing a sign of mild overfitting for $tq\gamma$ events. The $\text{NN}_{\text{out}} < 0.4$ region is insensitive to the $tq\gamma$ process, as the corresponding $tq\gamma$-background ratio is below 2%. Furthermore, it is shown in Appendix E.3 that the $\mathcal{B}_{\text{norm}}$ values from the different datasets agree well across all epochs when events with $\text{NN}_{\text{out}} < 0.4$ are excluded in their computation. The p-values of the 2-KS test performed on the $tq\gamma$ and background distributions for $\text{NN}_{\text{out}} > 0.4$ are 0.87 and 0.174, respectively. Therefore, there is no significant disagreement between the NN_{out} distributions for $tq\gamma$ events from the different datasets in the region sensitive to the $tq\gamma$ process. This indicates that this potential mild overfitting has a negligible impact on the sensitivity of the NN model to the $tq\gamma$ process. It is concluded that the performance of the NN model is also consistently high across the statistically independent datasets for the 0fj SR.

The $\mathcal{A}$ values of 0.737 ± 0.004 for the 0fj SR and 0.793 ± 0.005 for the $\geq$ 1fj SR evaluated on the validation datasets indicate that both NN models perform well as binary classifiers. Furthermore, the respective $\mathcal{S}$ values of 0.176 and 0.264 for NN_{out} are significantly larger than the corresponding values for m_{t}. Hence, the NN models separate $tq\gamma$ from background events significantly better than any individual feature. The statistical significance of the fitted NN models, evaluated by BPL fits of the predicted NN_{out} distributions obtained from the validation datasets to the distributions of the corresponding Asimov datasets, is 11.99 ± 0.24 for the 0fj SR and 19.70 ± 0.48 for the $\geq$ 1fj SR. These values correspond to small performance drops of $\approx$ 1% and $\approx$ 5% caused by the decreased size of the training dataset relative to the values for the optimized NN models listed under "Architecture optimization" in Table 16. The sizes of these drops are insignificant when all systematic uncertainties are considered in the BPL fits.

Table 17 displays the lists of input features in both regions ordered by their LOFO importance (f_i^{LOFO}). The table lists the difference (Δf_i^{LOFO}) between f_i^{LOFO} and the largest LOFO importance value ($f_{\text{max}}^{\text{LOFO}}$) across all input features relative to f_i^{LOFO}. In both SRs, it is observed that a large $\mathcal{S}$ value does not imply a large LOFO importance, which indicates that the NN models exploit the differences in the correlations between input features. In the 0fj SR, η^b and η^γ are the highest ranked features. Their f_i^{LOFO} values are at least 5.5% and 4.1% larger, respectively, than that of any other feature, while the differences between the remaining features in the f_i^{LOFO} values are at most 3.7%. In the $\geq$ 1fj SR, $m_{b+\text{fj}}$, H_T, and $\Delta\eta(\text{fj}, \gamma)$ are the highest ranked features. Their f_i^{LOFO} values are at least 7.7%, 6.1%, and 4.8% larger, respectively, than that of any other feature, while the differences in the f_i^{LOFO} values between the remaining features are at most 6.7%.

As the correlations between input features may have complex shapes and the interpretation of the

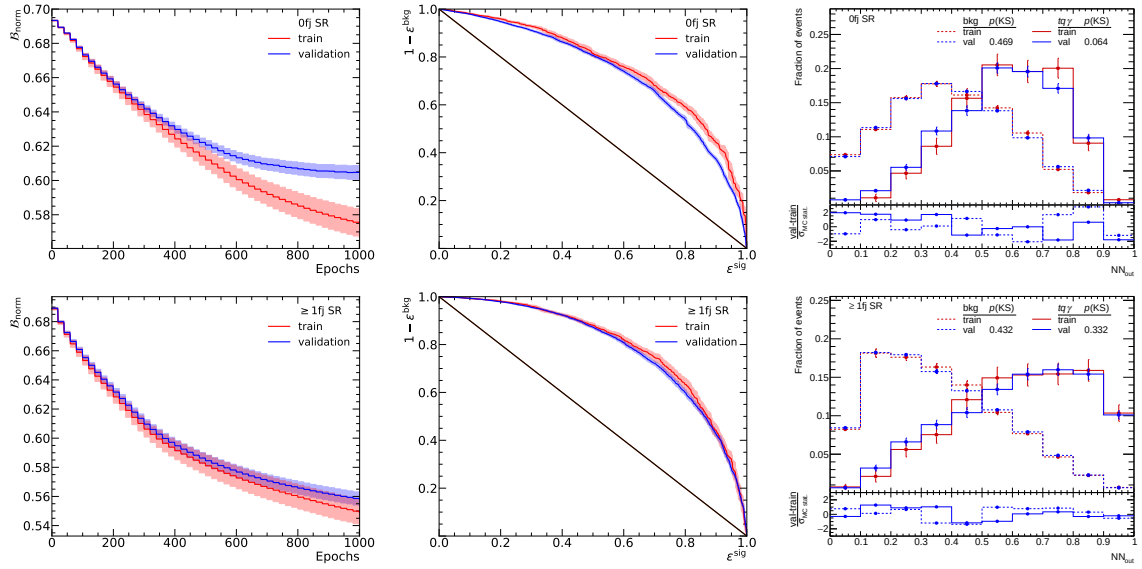


Figure 48: The figures in the left column show $\mathcal{B}_{\text{norm}}$ as a function of the number of completed epochs ("Epochs") and the figures in the middle column show the background rejection ($1-\varepsilon^{\text{bkg}}$) as a function of the signal efficiency (ε^{sig}) of the NN models for the training ("train") and validation datasets. The uncertainty bands correspond to the statistical uncertainty evaluated using bootstrap sampling. The figures in the right column display the fraction of events as a function of NN_{out} separately for $tq\gamma$ ("sig") and background ("bkg") events from the training ("train") and validation datasets ("val"). The uncertainty bars correspond to the uncertainty due to the limited number of simulated events. Furthermore, the p-values of 2-KS tests (p(KS)) between equivalent distributions from the training and validation datasets are displayed in the legend. The ratio panels show the differences in the event fractions between the training and validation datasets (val−train) relative to the uncertainty due to the limited number of simulated events ($\sigma_{\text{MC stat.}}$). The upper and lower rows show the distributions for the 0fj SR and ≥ 1fj SR, respectively.

Table 17: Lists of input features used for the NNs in the SRs ordered by the LOFO feature importance (f_i^{LOFO}). The corresponding $\Delta f_i^{\text{LOFO}}/f_i^{\text{LOFO}}$ values are the differences between f_i^{LOFO} and the largest LOFO feature importance value ($f_{\text{max}}^{\text{LOFO}}$) relative to f_i^{LOFO}. The corresponding separation strength $\mathcal{S}$ is also given for each input feature. Both metrics are evaluated using the validation dataset.

0fj SR			≥ 1fj SR		
feature	$\Delta f_i^{\text{LOFO}}/f_i^{\text{LOFO}}$ [%]	$\mathcal{S}$	feature	$\Delta f_i^{\text{LOFO}}/f_i^{\text{LOFO}}$ [%]	$\mathcal{S}$
η^b	0	0.001	$m_{b+\text{fj}}$	0	0.064
η^γ	−1.4	0.002	H_{T}	−1.6	0.043
b-tagging bin	−5.5	0.013	$\Delta\eta(\text{fj},\gamma)$	−3.9	0.018
m_t	−5.8	0.104	m_t	−7.7	0.090
p_{T}^b	−5.9	0.019	$\Delta R(\ell,b)$	−8.4	0.017
$\Delta R(\ell,\gamma)$	−5.9	0.022	p_{T}^b	−9.7	0.008
H_{T}	−6.9	0.029	$E^{\text{fj},\gamma}$	−9.8	0.063
$\Delta R(\ell,b)$	−7.0	0.030	$\Delta R(\ell,\text{fj})$	−9.9	0.020
η^ℓ	−7.1	0.008	$\Delta R(\ell,\gamma)$	−11.4	0.023
p_{T}^ℓ	−7.9	0.041	b-tagging bin	−11.5	0.009
$p_{\text{T}}^{t+\gamma}$	−8.3	0.011	p_{T}^{fj}	−11.7	0.047
$m_{\text{T}}(\ell,\nu,\gamma)$	−9.2	0.036	$m_{\text{T}}(\ell,\nu,\gamma)$	−12.2	0.021
-	-	-	$m_{\text{T}}(\ell,\nu)$	−12.6	0.060
-	-	-	$m_{\ell+b}$	−13.1	0.064
-	-	-	PID^ℓ	−14.4	0.020

inner workings of NN models is difficult, understanding the reason for the order in the LOFO feature importance is challenging. There are several commonly used options to quantify the correlation between input features. The computation of the Pearson correlation coefficients ρ represents one such option. Figure 49 shows their values for each pair of the NN input features separately for $tq\gamma$ (signal) and background events for the 0fj SR and $\geq$ 1fj SR, respectively. However, as these correlation coefficients quantify the linear dependence of two features, their values in combination with the $\mathcal{S}$ values represent only rough indicators for this order.

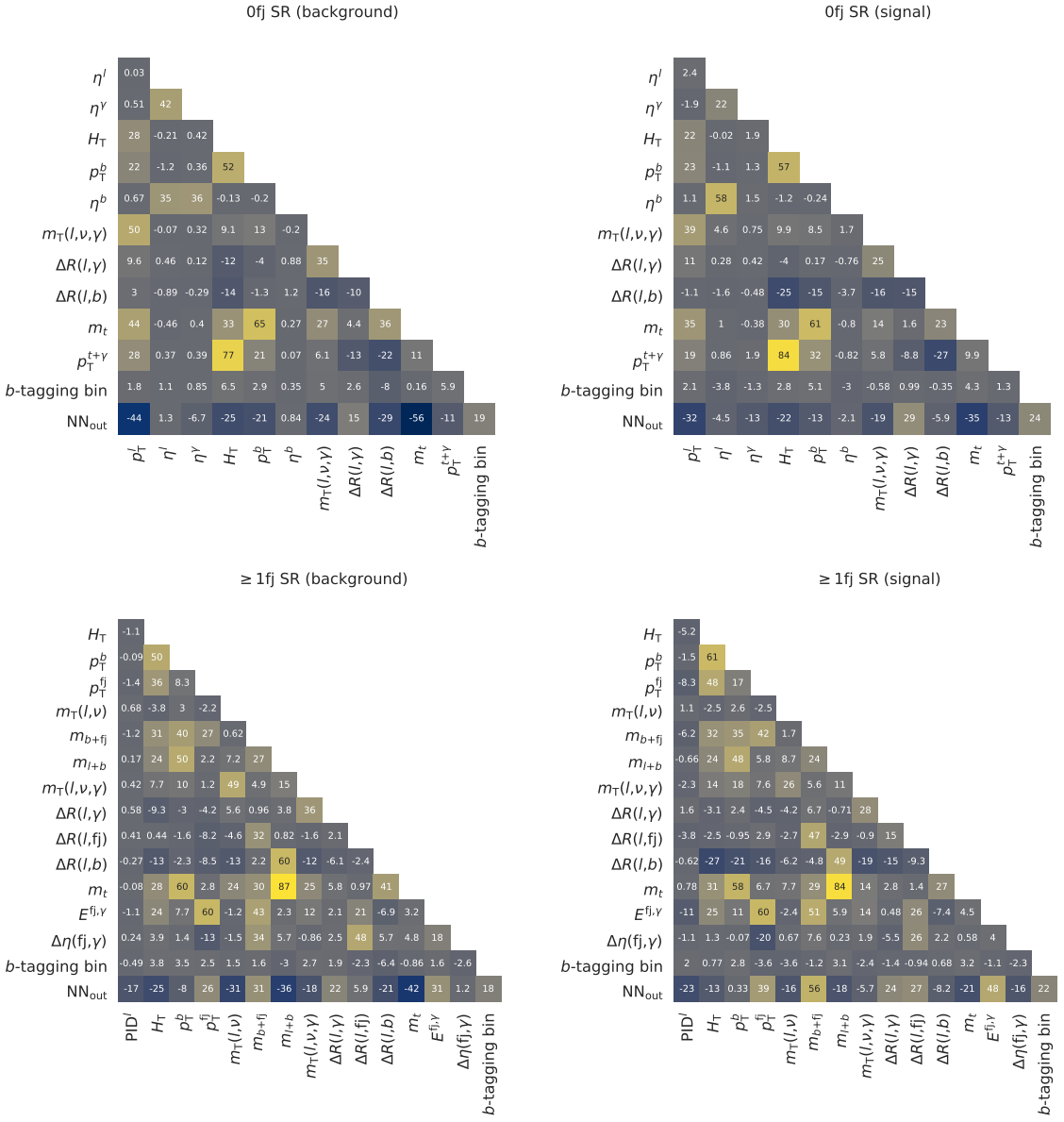


Figure 49: Pearson correlation coefficients between each of the input features and $\mathrm{NN_{out}}$ for $tq\gamma$ (signal) events (left) and background events (right) in the 0fj SR (upper row) and in the $\geq$ 1fj SR (lower row). The values are multiplied by 100 for illustrative purposes.

In the 0fj SR, most of the correlations between input features are either small ($|\rho| < 0.1$), weak ($0.1 < |\rho| < 0.3$), or moderate ($0.3 < |\rho| < 0.5$). A strong positive correlation ($0.5 < \rho < 0.7$) is observed between p_{T}^{b} and H_{T}, and p_{T}^{b} and m_t for both $tq\gamma$ and background events. This is due to p_{T}^{b} being used in the calculation of H_{T} and the four-momentum of the b-tagged jet being used in the calculation of m_t. A strong positive correlation is also observed between η^{b} and η^{ℓ} for $tq\gamma$ events, as the b-tagged jet and lepton candidate originate from the same top-quark decay in the majority of $tq\gamma$ events. The largest

correlation ($\rho > 0.7$) is observed between $p_{\mathrm{T}}^{t+\gamma}$ and H_{T} for both $tq\gamma$ and background events. The p_{T} of the reconstructed top-quark–photon system is compensated by the p_{T} of the remaining objects in an event due to momentum conservation. For $tq\gamma$ and the dominating semileptonic $t\bar{t}\gamma$ background contribution, these objects are expected to be mainly jets, and consequently, H_{T} becomes larger for larger $p_{\mathrm{T}}^{t+\gamma}$ values.

The absolute differences in the correlation coefficients between $tq\gamma$ and background contributions are mostly small (< 0.14) in the 0fj SR. The largest differences in the correlation coefficients of -0.34, 0.23, and -0.20 are observed for the coefficients of η^b and η^γ, η^b and η^ℓ, and η^ℓ and η^γ, respectively. While the correlations between η^b and η^γ, and η^ℓ and η^γ are moderate for background events, they are small and weak for $tq\gamma$ events, respectively. This is caused by the significant contributions of events involving radiative decays in the background that are absent for $tq\gamma$ events. These differences indicate the large LOFO importance observed for η^b and η^γ.

In the $\geq$ 1fj SR, most of the correlations between input features are also at most at a moderate level. Strong positive correlations are observed for p_{T}^b and H_{T}, p_{T}^b and m_t, and $p_{\mathrm{T}}^{\mathrm{fj}}$ and $E^{\mathrm{fj},\gamma}$ for both $tq\gamma$ and background events. The correlations between the former two feature pairs are also observed in the 0fj SR. The z-component of the momentum of the forward jet is always larger than its transverse component. Hence, larger $p_{\mathrm{T}}^{\mathrm{fj}}$ values typically imply larger p_z^{fj} values, so that the overall momentum is larger. This implies larger values of E^{fj} as well, causing the strong correlation between $p_{\mathrm{T}}^{\mathrm{fj}}$ and $E^{\mathrm{fj},\gamma}$. The largest correlation is observed for $m_{\ell+b}$ and m_t. This is due to the four-momenta of the b-tagged jet and the lepton candidate being used in the computation of both features. Furthermore, a strong correlation between $\Delta R(\ell, b)$ and $m_{\ell+b}$ is observed for background events, while the correlation is moderate for $tq\gamma$ events. The angular distance between particles originating from the same particle decay tends to be larger when the decaying particle has smaller momentum or larger invariant mass. As a consequence, the invariant mass of two particles that do not originate from the same particle decay is larger for larger angular distances. As the lepton candidate and b-tagged jet arise from the same top-quark decay more often for $tq\gamma$ events than for background events, their kinematic dependence, e.g., between η^b and η^ℓ, is larger for $tq\gamma$ events. This causes the correlation between $\Delta R(\ell, b)$ and $m_{\ell+b}$ to be smaller for $tq\gamma$ events than for background events.

The largest difference in the correlation coefficients between $tq\gamma$ and background events is observed for $m_{b+\mathrm{fj}}$ and $\Delta\eta(\mathrm{fj}, \gamma)$ pair with -0.27 in the $\geq$ 1fj SR. The angular distance between the forward jet and b-tagged jet is correlated with $m_{b+\mathrm{fj}}$. A moderate correlation between η^b and η^γ is observed for background events while it is at a negligible level for signal events. These two aspects cause the moderate correlation between $m_{b+\mathrm{fj}}$ and $\Delta\eta(\mathrm{fj}, \gamma)$ for background events, absent for $tq\gamma$ events. The second-largest difference of -0.23 is observed for the correlation coefficients of $m_{\mathrm{T}}(\ell,\ \nu)$ and $m_{\mathrm{T}}(\ell,\ \nu,\ \gamma)$. The correlation is moderate for background events and weak for signal events. The correlation of these two quantities is larger for events involving a radiative W-boson decay. The fraction of events where the photon candidate originates from such a decay is significantly larger for background contributions than for $tq\gamma$ events. For the $tq\gamma$ process, this decay is expected to be absent in the majority of events. This causes the difference in the correlation coefficients. The third-largest difference of -0.21 is observed for the correlation coefficients of $\Delta\eta(\mathrm{fj}, \gamma)$ and $\Delta R(\ell, \mathrm{fj})$, which is mainly caused by the previously discussed differences in the correlation of η^γ and η^ℓ. These observations represent indications for the large LOFO importances of $m_{b+\mathrm{fj}}$ and $\Delta\eta(\mathrm{fj}, \gamma)$ for the $\geq$ 1fj SR.

Table 18 lists the separation strengths of $\mathrm{NN}_{\mathrm{out}}$ when they are computed considering only individual background contributions, while neglecting all others. The predicted prompt-γ, $e \to \gamma$, and $h \to \gamma$ contributions from the different underlying processes modeled by the X samples are grouped and commonly referred to as "Other prompt γ", $e \to \gamma$, and $h \to \gamma$, respectively, in the following. The weakest separation power is observed when only $t\,(\to \ell\nu b\gamma)\,q$ is considered. This results from the strong topological similarity between $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$, as well as the relatively small overall $t\,(\to \ell\nu b\gamma)\,q$ contribution to the total background contribution. As a result, the NN models tend to prioritize the

separation of $tq\gamma$ events from other, more dominant background processes. Therefore, $t\,(\rightarrow \ell\nu b\gamma)\,q$ represents one of the most challenging but least impactful background processes for the NN models to separate from $tq\gamma$, resulting in the observed weak separation power. The largest separation strengths are observed for $W\gamma$+jets, as this background contribution does not contain any top quarks, while it is the third-largest background contribution in the 0fj SR and $\geq$ 1fj SR, respectively. The $t\bar{t}\gamma$ contribution is the largest background contribution and hence the most important for the NN models to separate from the signal to minimize the loss function. Furthermore, its topology is significantly different from the $tq\gamma$ topology, while it is slightly more similar to it than the $W\gamma$+jets topology due to the presence of top quarks. Therefore, the separation strengths associated with $t\bar{t}\gamma$ are large but slightly smaller than those associated with $W\gamma$+jets. The "Other prompt γ" contribution consists of events with a prompt photon mainly from the $t\bar{t}$ and tW samples and is hence similar to the $t\bar{t}\gamma$ contribution in terms of the event topology. Therefore, the corresponding separation strengths for this background contribution are of similar size compared to those for $t\bar{t}\gamma$. The $e \rightarrow \gamma$ and $h \rightarrow \gamma$ contributions mainly consist of $t\bar{t}$ events with a fake photon and are associated with similar separation strengths. The $Z\gamma$+jets background is associated with the smallest separation strength in the 0fj SR, while the separation strength is at the level of that for the $e \rightarrow \gamma$ contribution in the $\geq$ 1fj SR. While the $Z\gamma$+jets topology is fairly distinctive from the $tq\gamma$ topology, its contribution is the second-smallest in both SRs and hence the events only contribute a small amount to the loss function.

Table 18: Separation strengths $\mathcal{S}$ of $\mathrm{NN_{out}}$ when computed considering only one of the listed background contributions at a time for the 0fj SR and $\geq$ 1fj SR.

Background process	0fj SR	$\geq$ 1fj SR
$t\,(\rightarrow \ell\nu b\gamma)\,q$	0.03	0.06
$t\bar{t}\gamma$	0.19	0.30
$W\gamma$+jets	0.24	0.34
$Z\gamma$+jets	0.11	0.23
$e \rightarrow \gamma$	0.16	0.23
$h \rightarrow \gamma$	0.14	0.26
Other prompt γ	0.17	0.31

In conclusion, both fitted NN models excellently separate $tq\gamma$ events from all other events in the SRs. The corresponding output variable $\mathrm{NN_{out}}$ offers a significantly larger separation power than any other tested low-level or high-level feature. The performance of the NN models is found to be consistent across statistically independent datasets. This indicates that the NNs models adapted to patterns that are generally different between $tq\gamma$ events and all other events and not significantly to statistical fluctuations during the fitting. Furthermore, the order of the importance of the individual input features for the NN and the level of NN performance in terms of discriminative power against individual background contributions are understood. Hence, the NN models represent powerful and reliable discriminators in the SRs, making them well-suited for performing precise measurements of the fiducial cross sections of single-top-quark production in association with a photon and for testing for the discovery of $tq\gamma$.

10 Results of the fiducial cross-section measurements and the $tq\gamma$ discovery

To measure the numbers of signal and background events in the collected data for the parton-level and the particle-level measurements, BPL fits of the predicted distributions to the corresponding observed distributions are performed. The number of signal events is translated into the corresponding fiducial cross section via Equation (52). The predicted $\mathrm{NN_{out}}$ distributions in the SRs and the $t\bar{t}\gamma$ CR, and the predicted overall contribution ("Inclusive yield") to the $W\gamma$ CR are simultaneously fitted to the corresponding observed distributions for each measurement. The $\mathrm{NN_{out}}$ distribution in the $t\bar{t}\gamma$ CR is given by the predictions of the NN model fitted to simulated data in the 0fj SR for events without a forward jet and by the predictions of the NN model fitted to simulated data in the $\geq$ 1fj SR for events with at least one forward jet (cf. Section 9). The $\mathrm{NN_{out}}$ distributions in the SRs and the $t\bar{t}\gamma$ CR are divided into multiple bins that are included in the fit. The optimization of the binning scheme is discussed in Section 10.1. The results of the BPL fits using this optimized binning scheme, the measured values of the fiducial cross sections, and the test for the discovery of the $tq\gamma$ process are presented in Section 10.2.

10.1 Optimization of the binning scheme

The choice of the binning scheme is one of the key elements of the configuration of the BPL fits. The two most relevant candidate metrics for guiding the optimization in this work are the predicted statistical significance obtained in the test for the discovery of the $tq\gamma$ process and the precision of the measured fiducial cross sections obtained by the BPL fits. The predicted statistical significance is obtained from a fit of the predicted distributions to the respective distributions from an Asimov dataset using the default binning (cf. Section 9.4). Considering all systematic uncertainties, this value is 6.2σ. Since this value is well above the 5σ observation threshold, the precision of the fiducial cross sections is the more appropriate choice for the optimization. The predicted $t\,(\rightarrow \ell\nu b\gamma)\,q$ contribution is about 4 times smaller than the predicted $tq\gamma$ contribution. Furthermore, the uncertainty of the predicted overall $t\,(\rightarrow \ell\nu b\gamma)\,q$ contribution of 45%–54% is much larger than the uncertainty of the predicted overall $tq\gamma$ contribution of 6%–12%, depending on the AR. Therefore, the optimization of the binning scheme is focused on the $tq\gamma$ process. The setup of the parton-level measurement is used in the optimization, treating $t\,(\rightarrow \ell\nu b\gamma)\,q$ as a background process. The optimization of the binning scheme is guided by the precision of the fitted $\mu_{tq\gamma}$ value, as its precision determines the precision of the measured parton-level fiducial cross section for $tq\gamma$ (cf. Equation (52)). The binning scheme obtained from the optimization is also used for the particle-level measurement.

The "Transformation D" algorithm [272] is designed to automatically generate binning schemes of continuous distributions, based on the values of three user-defined parameters. Each set of parameter values leads to a specific binning scheme. Further details about this algorithm are provided in Appendix F.1. The binning configurations of the $\mathrm{NN_{out}}$ distributions in the SRs and the $t\bar{t}\gamma$ CR are optimized in three sequential steps using this algorithm: In the first step, the binning configuration of the $\mathrm{NN_{out}}$ distribution in the $\geq$ 1fj SR is optimized, as it is the most sensitive to $tq\gamma$. Subsequently, the binning configuration is optimized in the 0fj SR and then in the $t\bar{t}\gamma$ CR. In each step of the optimization, 195 configurations are tested. For each configuration, a BPL fit of the predicted distributions to the respective distributions from an Asimov dataset is performed, considering the full setup for systematic uncertainties introduced in Sections 7.5 and 7.6. The default binning configuration is used for the $\mathrm{NN_{out}}$ distributions in regions not yet optimized, while the optimized binning configuration is used for regions that have been optimized. The configuration that provides the smallest uncertainty of the fitted $\mu_{tq\gamma}$ value is used for that region in the statistical analysis. The $\mathrm{NN_{out}}$ distributions are displayed using the optimized binning scheme in the following.

10.2 Results of the binned-profile-likelihood fits and discovery test

In total 153 γ-factors, 230 NPs, and three free parameters are considered in the BPL given by Equation (49). The application of the pruning procedure, discussed in Section 7.5.2, reduces the number of NPs to 176 that are ultimately considered in the fitting. Figure 50 shows the predicted and observed $\mathrm{NN_{out}}$ distributions for the 0fj SR, $\geq$ 1fj SR, and $t\bar{t}\gamma$ CR, and the "Inclusive yield" in the $W\gamma$ CR. The predicted signal and background distributions and their respective uncertainties are shown before (pre-fit) and after (post-fit) their adjustment through the BPL fit. Furthermore, the ratio of the observed ("Data") and predicted ("Pred.") distributions is illustrated in the bottom panels. The displayed predicted distributions correspond to the parton-level measurement. The respective distributions for the particle-level measurement are presented in Appendix F.2. There are no visible differences between the predicted distributions for the two measurements.

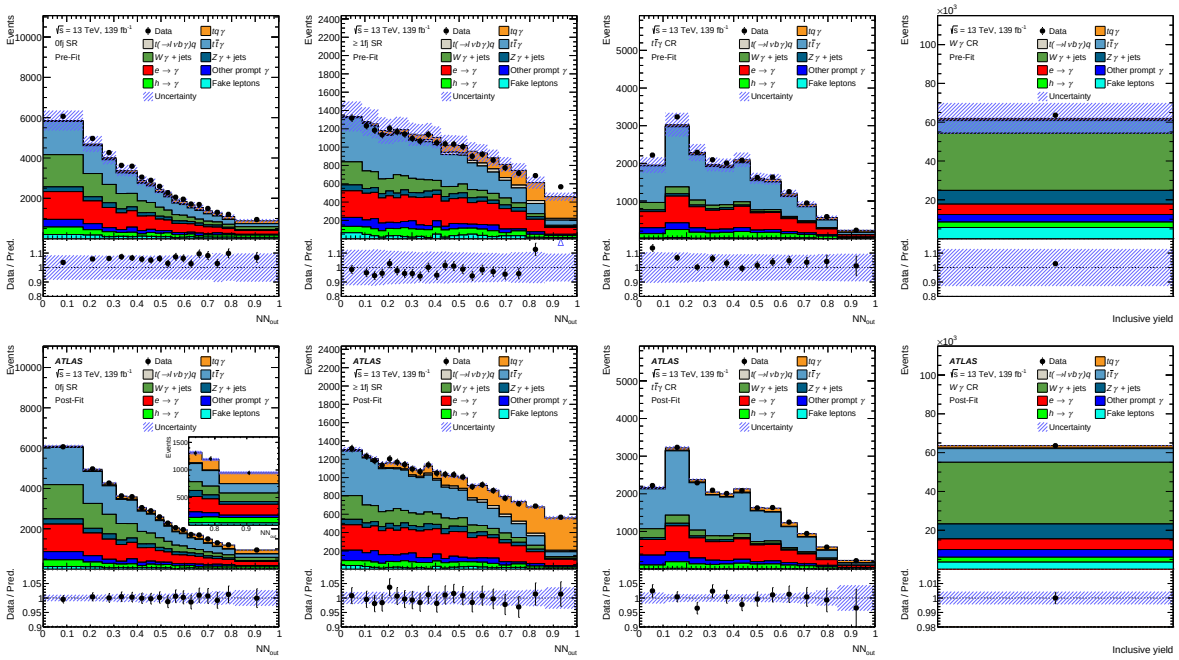


Figure 50: Predicted and observed ("Data") distributions in the ARs that are considered for the BPL fit for the parton-level measurement. The upper row shows the predicted distributions before performing the BPL fit and the lower row shows the predicted distributions adjusted by the BPL fit to the observed distributions. The $\mathrm{NN_{out}}$ distributions for the 0fj SR, $\geq$ 1fj SR, and $t\bar{t}\gamma$ CR, and the overall contributions ("Inclusive yield") for the $W\gamma$ CR are displayed. The hashed band corresponds to the total uncertainty of the sum of the predicted signal and background distributions. The lower panel displays the ratio of the observed and predicted distributions ("Pred."). The inset in the lower left plot presents a close-up of the last three bins of the $\mathrm{NN_{out}}$ distribution. The plots from the lower row are taken from Ref. [10].

The pre-fit distributions are discussed in the following. The predicted and observed inclusive yield in the $W\gamma$ CR agree excellently. The difference between these two values is much smaller than the uncertainty of the prediction. The predicted and observed $\mathrm{NN_{out}}$ distributions agree well in the 0fj SR, where the differences fluctuate around a constant offset and are covered by the total uncertainty of the prediction. In the $t\bar{t}\gamma$ CR, the ratio slightly exceeds the uncertainty band in the first bin. However, for the remaining bins, the differences are within the uncertainty band, and there are no systematic trends in the distribution of the ratio. The predicted and observed distributions also agree excellently in the $\geq$ 1fj SR for $\mathrm{NN_{out}} < 0.785$, which corresponds to a region dominated by the background contributions. The differences between predicted and observed distributions are much smaller than the uncertainty of the prediction, while systematic trends are absent in the distribution of the respective ratio. All the aforementioned findings indicate that the overall size of the background contributions to the ARs and the shapes of the respective $\mathrm{NN_{out}}$ distributions of background events are well-modeled. However, a

large disagreement between the predicted and observed $\mathrm{NN_{out}}$ distributions is observed in the $\geq$ 1fj SR for $\mathrm{NN_{out}} \geq 0.785$, especially for the last bin ($\mathrm{NN_{out}} > 0.86$). This region is characterized by the largest S/B ratio (≈ 1.07) across all ARs and is hence highly sensitive to the $tq\gamma$ process. Therefore, this disagreement indicates that the actual cross section of $tq\gamma$ production is larger than its predicted nominal value.

The fitted predicted distributions agree excellently with the observed distributions across all ARs, where most differences are contained within the uncertainty of the prediction. The "Saturated model" is used to evaluate the goodness-of-fit [273]. The corresponding p-value is 0.97, which validates the absence of any significant discrepancies between predicted and observed distributions that are not covered by the uncertainties considered in the fit model.

Values of $\mu_{W\gamma} = 1.08^{+0.22}_{-0.18}$ and $\mu_{t\bar{t}\gamma} = 1.08^{+0.15}_{-0.14}$ are obtained in the parton-level measurement, and of $\mu_{W\gamma} = 1.08^{+0.22}_{-0.18}$ and $\mu_{t\bar{t}\gamma} = 1.06^{+0.15}_{-0.14}$ in the particle-level measurement. Hence, the values of these parameters are compatible across the measurements and agree well with 1 within the uncertainties. The fitted values of the signal strength parameters are $\mu_{tq\gamma} = 1.438^{+0.163}_{-0.156} = 1.438 \pm 0.047\,(\mathrm{stat.})^{+0.156}_{-0.148}\,(\mathrm{syst.})$ in the parton-level measurement and $\mu_S = 1.395^{+0.159}_{-0.151} = 1.395 \pm 0.040\,(\mathrm{stat.})^{+0.154}_{-0.146}\,(\mathrm{syst.})$ in the particle-level measurement. These measured values translate into the fiducial cross sections shown below.

$$\sigma^{\text{fid., meas.}}_{t(\to\ell\nu b)q\gamma} = 688 \pm 23\,(\mathrm{stat.})\,^{+75}_{-71}\,(\mathrm{syst.})\ \mathrm{fb}$$
$$\sigma^{\text{fid., meas.}}_{t(\to\ell\nu b)q\gamma} + \sigma_{t(\to\ell\nu b\gamma)q} = 303 \pm 9\,(\mathrm{stat.})\,^{+33}_{-32}(\mathrm{syst.})\ \mathrm{fb}$$

The precision of the values of 11% is mainly limited by systematic uncertainties, while the statistical precision is about 3%. The values are compatible with their predictions, presented in Section 7.4, within 2.1 and 2.0 standard deviations, respectively. The observed significance of the discovery test corresponds to 9.3σ, exceeding the predicted significance of 6.8σ due to the significantly larger measured overall $tq\gamma$ contribution relative to its nominal prediction.

Table 19 lists the predicted yields of each contribution as obtained from the BPL fit for the parton-level measurement for all ARs and the corresponding observed yields. The respective predicted yields for the particle-level measurement are only listed for the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ processes, as the differences between the predicted yields from the particle-level and parton-level measurements are much smaller than the corresponding uncertainties for all contributions except for the $t\,(\to \ell\nu b\gamma)\,q$ process. The parton-level measurement is barely sensitive to the $t\,(\to \ell\nu b\gamma)\,q$ contribution, so that the adjustment of its prediction by the BPL fit is small. In the particle-level measurement, $t\,(\to \ell\nu b\gamma)\,q$ is treated as part of the signal and, hence, its overall predicted contribution is adjusted by the same signal strength parameter as $tq\gamma$. These two aspects primarily cause the large discrepancy in the nominal measured $t\,(\to \ell\nu b\gamma)\,q$ contribution obtained from the two measurements. However, the measured overall $t\,(\to \ell\nu b\gamma)\,q$ contributions are consistent within the uncertainty across the measurements.

To ensure that the measurement results obtained from the fit are consistent, it is crucial to analyze the behavior of the NPs in the fitting to identify any unexpected or inconsistent behavior in the fit parameters. Such behavior could arise from errors in the fit setup, poorly modeled contributions, or systematic uncertainties that are improperly neglected in the fit model. This analysis is discussed in the following.

The impact of each NP and each γ-factor on the fitted values of $\mu_{tq\gamma}$ and μ_S is determined to identify which of these parameters most significantly limits the precision of the measurements. These insights can be used to deduce measures for improving the precision of future measurements. The BPL fit is performed four times for each NP and twice for each γ-factor, with the respective parameter set to a constant value in each fit. The four constant values used for the NPs are the fitted nominal value ($\hat{\theta}$) shifted by ± 1 of the pre-fit ($\Delta\theta$) and post-fit uncertainties ($\Delta\hat{\theta}$). The values of the γ-factors are set to their $\pm 1\sigma$ post-fit variations. For each fit, the absolute change ($\Delta\mu$) of the measured value of the respective

Table 19: Predicted yields and their total uncertainties in the SRs and CRs as obtained by the BPL fits for the two measurements [10]. For the particle-level measurement, only the yields for the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ are listed, as the differences in the yields obtained from the two measurements are < 3% for all other contributions across all regions.

Parton-level measurement				
	$\geq$ 1fj SR	0fj SR	$t\bar{t}\gamma$ CR	$W\gamma$ CR
$tq\gamma$	2410 ± 250	2500 ± 320	890 ± 120	1290 ± 150
$t\,(\to \ell\nu b\gamma)\,q$	370 ± 160	460 ± 230	130 ± 50	230 ± 110
$t\bar{t}\gamma$ (prod)	3200 ± 500	4800 ± 700	4300 ± 600	2700 ± 400
$t\bar{t}\gamma$ (dec)	3900 ± 600	9400 ± 1400	5700 ± 600	4300 ± 900
$W\gamma$+jets	2500 ± 400	9200 ± 1400	1160 ± 320	31 800 ± 3000
$Z\gamma$+jets	980 ± 310	2700 ± 800	430 ± 150	7700 ± 2400
$e \to \gamma$	5100 ± 500	10 300 ± 800	4800 ± 400	5400 ± 500
$h \to \gamma$	1100 ± 400	2600 ± 900	1300 ± 500	2500 ± 800
Other prompt γ	1330 ± 350	2600 ± 900	1300 ± 400	4000 ± 600
Fake leptons	390 ± 190	1000 ± 500	110 ± 50	3500 ± 1700
Total	21 250 ± 150	45 720 ± 250	20 180 ± 150	63 590 ± 280
Data	21 227	45 723	20 194	63 592
Particle-level measurement				
	$\geq$ 1fj SR	0fj SR	$t\bar{t}\gamma$ CR	$W\gamma$ CR
$tq\gamma$	2360 ± 250	2450 ± 310	880 ± 120	1260 ± 140
$t\,(\to \ell\nu b\gamma)\,q$	500 ± 170	660 ± 210	180 ± 60	330 ± 120
Total	21 250 ± 150	45 720 ± 240	20 180 ± 150	63 590 ± 280
Data	21 227	45 723	20 194	63 592

signal strength is calculated. The parameters are sorted in descending order according to the maximum of the two respective $\Delta\mu$ values obtained when setting the parameters to the post-fit variations.

Figure 51 illustrates the resulting rankings for the parton-level and particle-level measurements. The "$t\bar{t}\gamma$ PS model $\geq$ 1fj SR" NP has the largest effect on the values of $\mu_{tq\gamma}$ and μ_S. Its impact is significantly larger than that of any other parameter (more than twice as large). The imapct of this variation on the total prediction, i.e., the sum of all signal and background predictions, in the $\geq$ 1fj SR is illustrated in the lower left plot in Figure 52. The dashed distribution corresponds to the $\pm1\sigma$ variations before applying smoothing and symmetrization, and the solid distributions after applying these two algorithms. The bottom panel shows the relative difference between the $\pm1\sigma$ variations and the nominal prediction. The absolute relative differences between the respective $\pm1\sigma$ variations and the nominal prediction increase monotonically as a function of $\mathrm{NN_{out}}$ and is about 5% in the last two bins, which are the most sensitive ones to the $tq\gamma$ process. These properties cause the large impact on the signal strength parameters. The remaining parameters listed in the rankings decrease steadily in their impact.

Most of the illustrated NPs are listed in both rankings. The largest differences in the order and impact are observed for the NPs related to the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ predictions. The $t\,(\to \ell\nu b\gamma)\,q$ process is being considered as signal in the particle-level measurement. This causes larger impacts of the $t\,(\to \ell\nu b\gamma)\,q$-related NPs on μ_S relative to those on $\mu_{tq\gamma}$ in the parton-level measurement, in which it is treated as a background process. As discussed in Section 7.6.2, the signal-related variations are normalized differently for the two measurements. This results in different overall impacts of these variations on the predicted $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ contributions in the two measurements. This may also

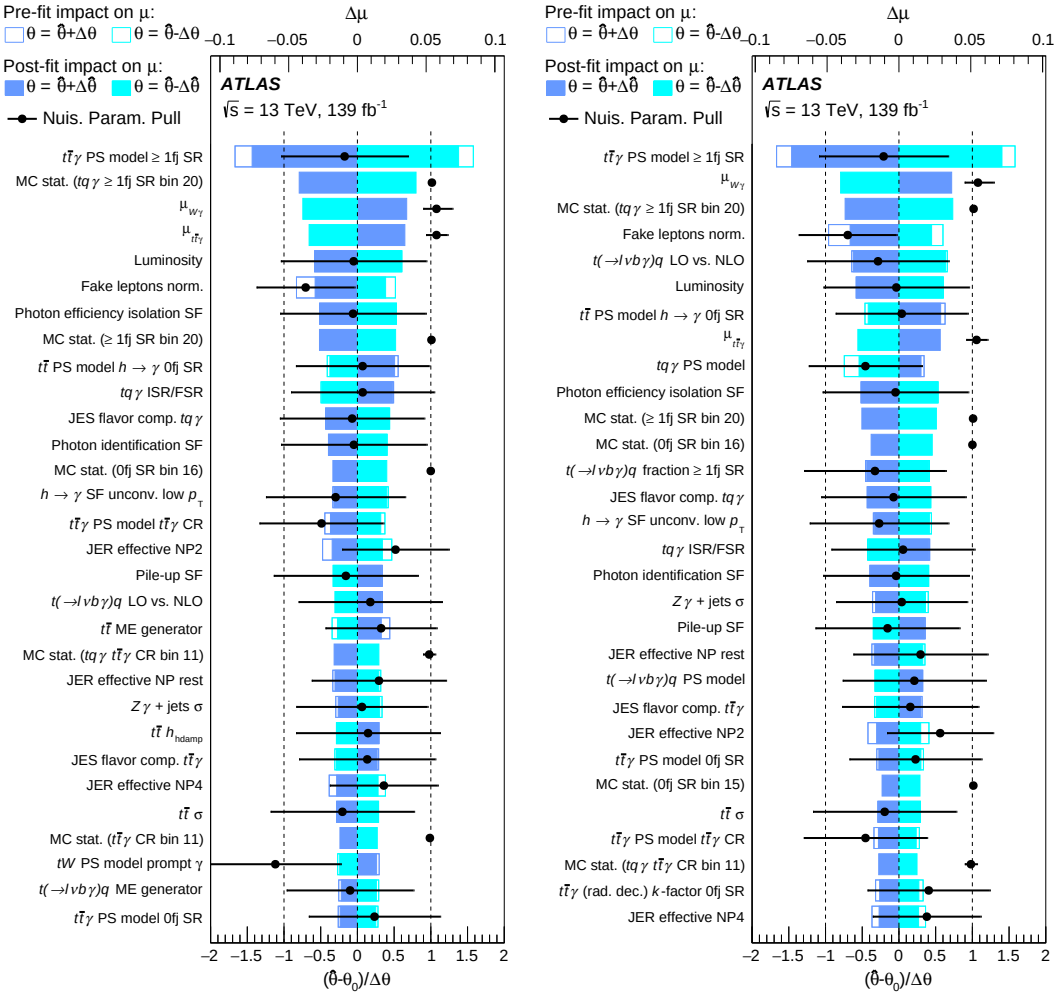


Figure 51: Rankings of the 30 parameters (NPs and γ-factors) that impact the measured values of $\mu_{tq\gamma}$ (left) and μ_S (right) in the parton-level and particle-level (right) measurements, respectively, the most. Both signal strength parameters are denoted by μ in the rankings. The γ-factors are denoted by "MC stat.". The impact of each parameter, $\Delta\mu$, is computed by comparing the nominal post-fit value of μ with the result of the fit when fixing the considered parameter to a specific value. For NPs, the parameter is fixed to its post-fit value ($\hat{\theta}$) shifted by its pre-fit ($\pm\Delta\theta$) and post-fit uncertainties ($\pm\Delta\hat{\theta}$). For γ-factors, the parameter is fixed to its post-fit value shifted by its post-fit uncertainty. The empty boxes show the pre-fit impact while the filled boxes show the post-fit impact of each parameter on the result. For NPs, the black dots represent the post-fit value (pull) of each NP where the pre-fit value is subtracted (θ_0), while the black line represents the post-fit uncertainties normalized to the respective pre-fit uncertainties. For γ-factors, the black dots and black lines represent the post-fit values and uncertainties, respectively. The figures are taken from Ref. [10].

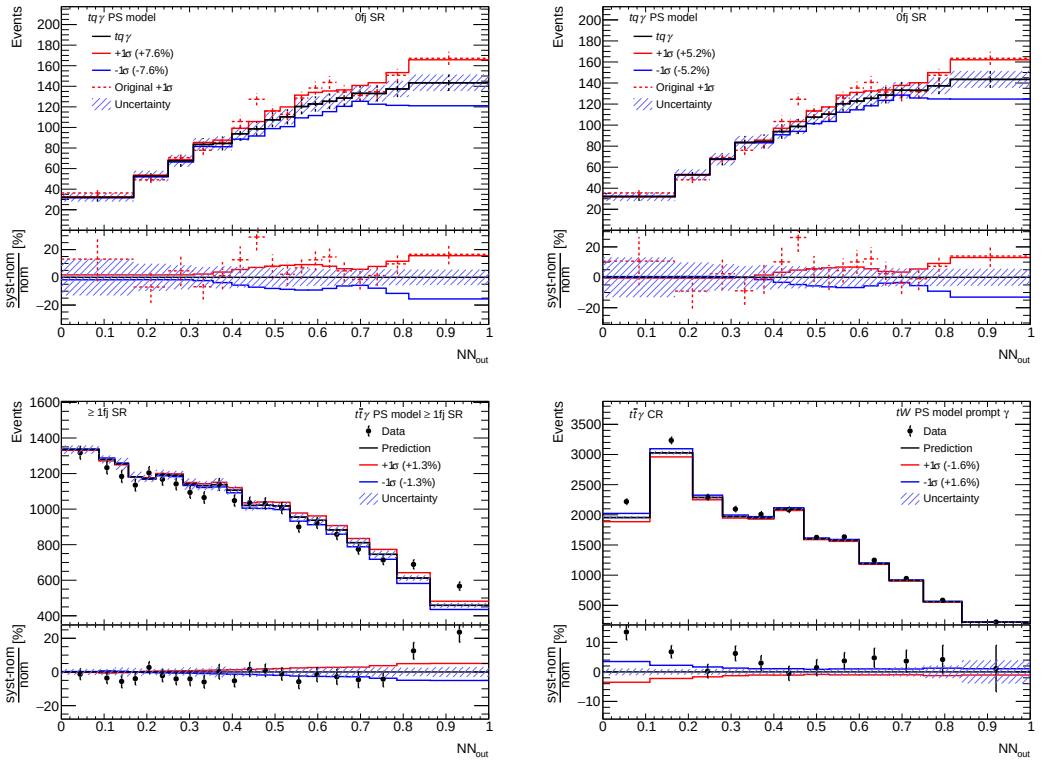


Figure 52: The upper row shows the impact of the "$tq\gamma$ PS model" variation on the prediction of the $\mathrm{NN_{out}}$ distributions of $tq\gamma$ in the 0fj SR for the parton-level measurement (left) and the particle-level measurement (right). The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing. The error bars represent the uncertainty of the predictions due to the limited sample sizes. The ratio panel compare the $\pm 1\sigma$ variations ("syst") to the nominal prediction ("nom") in terms of relative differences.
The lower row shows the impacts of the "$t\bar{t}\gamma$ PS model $\geq$ 1fj SR" (left) and the "tW PS model prompt γ" (right) variations on the overall prediction given by the sum of all signal and background predictions for the $\mathrm{NN_{out}}$ distributions, and the respective observed distributions. The former is shown for the $\geq$ 1fj SR and the latter for the $t\bar{t}\gamma$ CR. The error bars on the observed distributions correspond to the uncertainty due to the limited number of events in the collected data, while the hashed blue bands illustrate the uncertainty due to the limited sample sizes. The ratio panel compares the distributions obtained from the variations and observed distributions to the nominal prediction in terms of the relative difference to the nominal prediction.

impact the shape of the variation obtained after the application of the smoothing procedure. The upper row of Figure 52 shows the impact of the "$tq\gamma$ PS model" variation on the predicted $\mathrm{NN_{out}}$ distribution of $tq\gamma$ events in the 0fj SR for the parton-level (left) and particle-level (right) measurements. The uncertainty of the predicted overall $tq\gamma$ contribution is smaller in the particle-level measurement, while the shape of the uncertainty distributions agree well for the different measurements. Appendices F.3 and F.4 show the impact of all signal-related modeling uncertainties on the predicted $\mathrm{NN_{out}}$ distributions of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events for both measurements, respectively.

To assess the impact of individual types of systematic uncertainties on the measurements, the NPs are grouped into different categories, and the impact of each of these groups is evaluated in the following way:

1. The BPL fit is performed again while the values of the NPs associated with the respective group are fixed to their best-fit values.

2. The uncertainty of the signal strength parameter corresponding to this alternative fit is subtracted in quadrature from the nominal uncertainty. The result is considered as the uncertainty caused by the respective group. The larger this value the larger the impact of the respective group.

The groups with the largest impacts are shown in descending order of their impact in Table 20, along with their impacts on the uncertainty of the signal strength parameters. A description of the displayed groups is given in the following. The "$t\bar{t}\gamma$ modeling" and "$t\,(\to \ell\nu b\gamma)\,q$ modeling" groups correspond to the impact of all NPs associated with the modeling of $t\bar{t}\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$, introduced in Section 7.6.2, respectively. The former is the group with the largest impact in both measurements, as $t\bar{t}\gamma$ is the largest background in the SRs. The "Background MC statistics" and "$tq\gamma$ MC statistics" groups correspond to the impacts of the γ-factors associated with the background and $tq\gamma$ predictions, respectively. The "Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$" group include NPs that are associated with reconstructed jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, as introduced in Section 7.6.1. The impact on μ_S is similar to the respective impact on $\mu_{tq\gamma}$ for all groups except for the "$t\,(\to \ell\nu b\gamma)\,q$ modeling", where the impact on μ_S is significantly larger. The remaining groups are introduced and discussed in Appendix F.2.

Table 20: List of groups of NPs with the largest impacts on $\mu_{tq\gamma}$ in the parton-level measurement and on μ_S in the particle-level measurement. The impact of each group is quantified in the following way: The BPL fit is performed while the values of the NPs of the respective group are fixed to their best-fit values. The uncertainty of the fitted $\mu_{tq\gamma}$ (μ_S) values are subtracted in quadrature from the respective nominal uncertainty. The square root of this value is denoted by $\Delta\sigma_{\mu_{tq\gamma(S)}}$ and the impact is given by $\Delta\sigma_{\mu_{tq\gamma}}/\mu_{tq\gamma}$ ($\Delta\sigma_{\mu_S}/\mu_S$). The values of the impacts are taken from Ref. [10].

Uncertainty	$\Delta\sigma_{\mu_{tq\gamma}}/\mu_{tq\gamma}$	$\Delta\sigma_{\mu_S}/\mu_S$
$t\bar{t}\gamma$ modeling	±5.5%	±5.5%
Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$	±3.6%	±3.5%
Background MC statistics	±3.5%	±3.6%
$tq\gamma$ MC statistics	±3.3%	±3.0%
$t\,(\to \ell\nu b\gamma)\,q$ modeling	±1.9%	±3.3%

Figure 51 also shows the pulls and constraints of each of the listed parameters. A pull is defined by the difference of the fitted value ($\hat{\theta}$) of the NP from its nominal value (θ_0) in units of its pre-fit uncertainty ($\Delta\theta$). A constraint refers to the shrinkage of the uncertainty after the fit ($\Delta\hat{\theta}$) with respect to its nominal size, and is hence only observed for NPs and not γ-factors. Large pulls or large constraints might indicate an inconsistent behavior of the corresponding NP and are therefore analyzed.

Figure 53 provides a detailed overview of the pulls and constraints for all NPs that are listed in either of the rankings or with an absolute pull value ($|\hat{\theta} - \theta_0|/\Delta\theta$) larger than 0.5. The majority of the pulls

are at most mild (< 0.5), while only a few larger pulls are observed. The absolute value of the largest pull slightly exceeds 1 and corresponds to the "tW PS model prompt γ" NP. The second-largest pull, which has an absolute value slightly smaller than 1, is also related to tW production and corresponds to the "tW DS scheme" NP. The investigation of these pulls concludes that they are related to events with small $\mathrm{NN_{out}}$ values in the $t\bar{t}\gamma$ CR. The agreement between the predicted and observed distributions is worse in this region than elsewhere in the $t\bar{t}\gamma$ CR, and these uncertainties are found to impact the overall prediction the most of all uncertainties. These pulls have no significant impact on the measured values of $\mu_{tq\gamma}$ and μ_S, which is indicated by their absence in the ranking lists. The lower right figure in Figure 52 displays the impact of the "tW PS model prompt γ" variation on the overall prediction in the $t\bar{t}\gamma$ CR. It causes an uncertainty of 4% in the first bin of the distribution, while the shape of the predicted distribution corresponding to the -1σ variation follows the shape of the observed distribution. These two aspects cause the relatively large pull of this NP.

Furthermore, the pulls and constraints are similar for the parton-level and particle-level measurements. The largest difference between values of the pulls for the different measurements is observed for the "$t\,(\rightarrow \ell\nu b\gamma)\,q$ LO vs. NLO" NP. The pulls are in opposite directions for the two measurements. The particle-level measurement is more sensitive to this uncertainty, which causes the slightly larger absolute value of the pull.

An NP is considered to be constrained when the value of the ratio $\Delta\hat{\theta}/\Delta\theta$ is significantly smaller than 1, i.e., smaller values of the ratio correspond to larger constraints. Most of the NPs are at most slightly constrained ($\Delta\hat{\theta}/\Delta\theta > 0.7$). The largest constraints are observed for the "tW ME generator" ($\Delta\hat{\theta}/\Delta\theta = 0.63$) and the "Fake leptons norm." ($\Delta\hat{\theta}/\Delta\theta = 0.67$) NPs. The "$tW$ ME generator" NP is mainly constrained due to its large impact of 2.5% on the overall prediction in the 0fj SR. It further impacts the overall prediction by 1.5% in the $\geq$ 1fj SR and by 1.6% in the $t\bar{t}\gamma$ CR. The $W\gamma$ CR is especially sensitive to the fake-leptons background. The pre-fit impact of the "Fake leptons norm." NP on the overall prediction in this region is 4.6%. This is the largest impact of an individual NP on the overall prediction across all ARs, which leads to the significant post-fit constraint of the NP. Appendix F.2 discusses further constraints of NPs with $0.7 < \Delta\hat{\theta}/\Delta\theta < 0.8$.

Figure 53 shows the post-fit correlation coefficients between the pairs of selected NPs and $\mu_{tq\gamma}$, $\mu_{t\bar{t}\gamma}$, and $\mu_{W\gamma}$ for the parton-level measurement.[37] It is not possible to display the correlation coefficients for all NP pairs as $\approx$ 16,000 of such pairs exist in the fit model. The NPs must fulfill at least one of the following two conditions to be selected for display:

1. The absolute correlation coefficient is at least 0.2 with at least one of the three free parameters.
2. The absolute correlation coefficient is at least 0.25 with at least one other NP.

Therefore, only NPs that have at least a mild correlation with any other NP are displayed. The absolute value of only 18 of the selected correlation coefficients from the $\approx$ 16,000 are $\geq$ 0.2 and only 8 are $\geq$ 0.3. The largest correlation coefficient of 0.64 is observed for the pair of $\mu_{W\gamma}$ and the $W\gamma$+jets μ_{F} NP. The μ_{F} and μ_{R} variations are found to have a large impact on the predicted overall contribution of the $W\gamma$+jets process. The μ_{F} variation has an especially large impact on the predicted contribution in the $W\gamma$ CR, which causes this large post-fit correlation with $\mu_{W\gamma}$. As the respective $+1\sigma$ variation is lowering the prediction of the $W\gamma$+jets contribution, the correlation coefficient is positive. In contrast to the μ_{F} variation, the μ_{R} variation has a small impact on the overall $W\gamma$+jets contribution in the $W\gamma$+jets CR, but large impacts in the SR and $t\bar{t}\gamma$ CR. This results in a significantly smaller correlation coefficient, which nevertheless still exceeds the 0.2 threshold. The second-largest absolute value of the correlation coefficient is observed for the pair of $\mu_{W\gamma}$ and the "$W\gamma$+jets b-Hadron norm. SR" NP. The $+1\sigma$ variation enhances the predicted $W\gamma$+jets contribution in the SRs and is hence negatively correlated with $\mu_{W\gamma}$. Consequently, the correlation coefficient of the pair of this NP and the $W\gamma$+jets μ_{F} NP variation is also one of the larger ones with -0.34.

[37] The NPs are treated as uncorrelated before the fit and all pre-fit correlation coefficients are zero.

The "$t\bar{t}\gamma$ (rad. dec.) k-factor X" NPs impact the predictions of the $t\bar{t}\gamma$ contributions and are negatively correlated with $\mu_{t\bar{t}\gamma}$ as the $+1\sigma$ variation enhances the predicted overall $t\bar{t}\gamma$ contribution to an AR. The absolute values of the correlation coefficients decrease with decreasing sensitivity to $t\bar{t}\gamma$ of the corresponding region. The correlation coefficient is the largest for the $t\bar{t}\gamma$ CR (-0.65), second-largest for the 0fj SR (-0.36), and third-largest for the $\geq$ 1fj SR (-0.31). The only large correlation with $\mu_{tq\gamma}$ is observed for the "$t\bar{t}\gamma$ PS model $\geq$ 1fj SR" NP with a correlation coefficient of -0.48, as the $\geq$ 1fj SR is the most sensitive to $tq\gamma$. This NP has the largest overall impact on the $\mu_{tq\gamma}$ value, as shown in Figure 51. The $+1\sigma$ variation of the NP enhances the $t\bar{t}\gamma$ (dec) prediction, resulting in the two parameters to be negatively correlated.

The corresponding matrix of correlation coefficients for the particle-level measurement is shown in Appendix F.2. The correlation coefficients do not deviate significantly from those for the parton-level measurement. However, due to the enhanced sensitivity to the $t\,(\to \ell\nu b\gamma)\,q$-related NPs in the fit model of the particle-level measurement relative to that of the parton-level measurement, "$t\,(\to \ell\nu b\gamma)\,q$ LO vs. NLO" NP, which is not selected for display for the parton-level measurement, has a larger correlation coefficient with μ_S (-0.21).

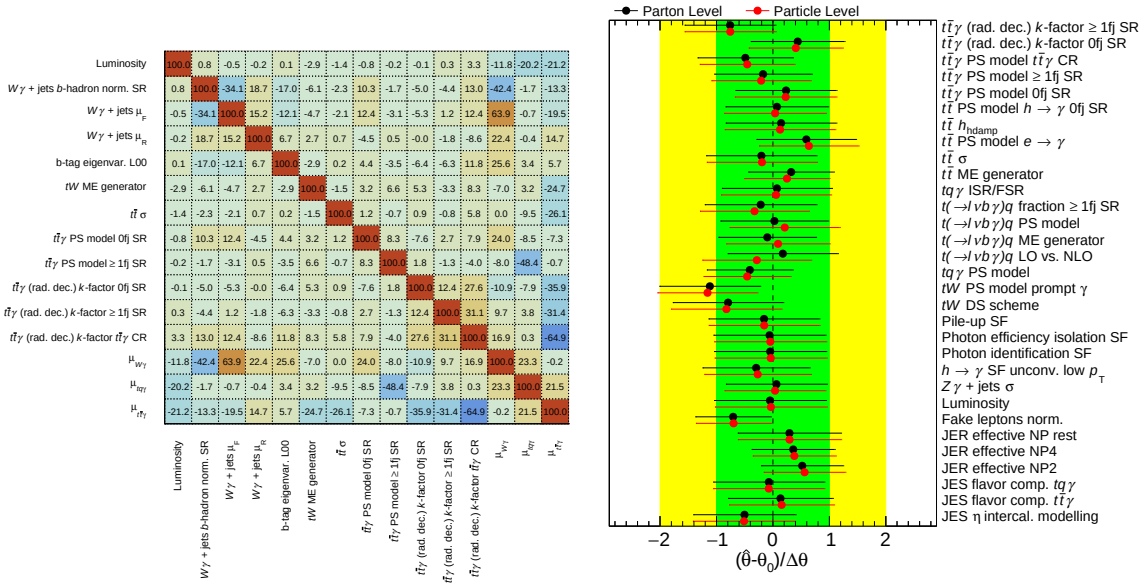


Figure 53: The left figure shows the matrix of the correlation coefficients of the displayed NPs and free-floating parameters. The correlation coefficients are scaled by a factor 100 for better readability. The right figure shows the results of the fitted NP values and their uncertainties. The black dots represent the post-fit mean value of the NP, the lines represent the post-fit NP uncertainties. The green and the yellow bands represent 1σ and 2σ of the respective pre-fit uncertainties, respectively.

In conclusion, the analysis of the pulls, constraints, and correlations of the NPs does not reveal any indications of issues that challenge the stability and consistency of the fit model. The behavior of the parameters is understood, thus validating the fit model.

11 Conclusions

The data of proton-proton (pp) collisions at a center-of-mass energy of 13 TeV, collected during Run 2 of the LHC run program by the ATLAS experiment, provides an extensive amount of hard-scattering events in which at least one top quark is produced. This dataset not only allows for precision studies of top-quark properties, but also for the exploration and observation of rare top-quark processes. The t-channel single-top-quark production in association with a photon belongs to this category and is examined in this thesis. This process is sensitive to the coupling of the top quark to the photon. Primarily, this key element of electroweak physics has been studied in top-quark pair production in association with a photon ($t\bar{t}\gamma$). Refs. [160, 161] concluded that measurements of the two complementary processes using data from pp collisions provided by the LHC are expected to reach similar sensitivity to the top-quark–photon interaction in the future, so that the combination of measurements of both processes is expected to reach a significantly higher sensitivity than individual measurements.

Prior to this work, single-top-quark production in association with a photon had not been observed and was only measured by the CMS collaboration using a partial Run-2 dataset of pp collisions corresponding to 35.9 fb^{-1} [20]. The CMS collaboration reported evidence for the process corresponding to a statistical significance of 4.4σ, and measured a fiducial cross section of this process with a precision of 30%. The measured value is 1.42 ± 0.43 times the respective Standard Model (SM) prediction and is in agreement with it. The prediction includes next-to-leading order (NLO) contributions in quantum chromodynamics (QCD).

The pp collisions provided by the LHC represent a challenging environment for studying physics processes with photons in the final state, as the majority of reconstructed photon candidates are fake photons. Photon-isolation and photon-identification requirements are applied to the reconstructed photon candidates to select signal, prompt photons with a high efficiency while rejecting the majority of fake photons. The measurement of the efficiency of the photon identification for prompt photons at high energies using the inclusive photon method is presented. This method was developed and established by the ATLAS collaboration and uses a sample of events that are only required to contain at least one reconstructed photon candidate. To measure the photon-identification efficiency for prompt photons, they must be disentangled from fake-photon contributions. The efficiency of passing track-based isolation requirements is significantly higher for prompt photons than for fake photons. Therefore, it is used as a discriminating variable. A data-driven estimation of the efficiency for fake photons is performed by using data in which reconstructed photon candidates are required to fail specific criteria of the photon identification. The prompt-photon contribution to this data (non-PID data) is modeled using simulation. With respect to previous measurements, two improvements of the method are introduced. A novel approach is implemented that restores the self-consistency of the method in cases of unphysical results. Furthermore, the impact of a potential mismodeling of the simulated prompt-photon contribution to the non-PID data on the measured efficiencies has not been accounted for in previous measurements. A procedure is implemented that introduces data-to-simulation scale factors (SFs) correcting the simulated prompt-photon contribution to this data, and assigns an uncertainty to the measured photon-identification efficiencies for prompt photons accounting for this potential mismodeling. These novel scale factors can be provided to the researchers of the ATLAS collaboration to correct the modeling of the prompt-photon contribution in analyses using non-PID data. The measured photon-identification efficiencies are further used to compute data-to-simulation SFs for correcting this modeling in physics analyses that require photon candidates to pass the photon-identification requirements. The presented values are the most precise for photon energies > 100 GeV with uncertainties below 1% relative to those from previous measurements. The presented results are published in Ref. [6].

The t-channel single-top-quark production in association with a photon is investigated considering only semileptonic top-quark decays, as using hadronic top-quark decays offers little potential to separate contributions of this process from background contributions in pp-collision data. In the SM, the top quark interacts with the photon either via a flavor-conserving neutral current or a flavor-changing neutral

current, leading to different final-state topologies. The former is predicted to occur at a frequency more than ten orders of magnitude higher, so that any observation of the latter in this dataset represents an indication for beyond the SM physics. This work focuses on the former interaction type, while the latter is examined in Ref. [11].

The process is divided into two sub-processes based on whether the photon is radiated in the semileptonic top-quark decay. The sub-process in which this is the case is denoted by $t\,(\to \ell\nu b\gamma)\,q$ and the other by $tq\gamma$. While these two sub-processes are in principle inseparable, it is shown that noticeable interference effects are absent justifying a separate treatment in physics analyses. The hard-scattering $tq\gamma$ process is modeled including NLO QCD contributions, as in Refs. [20, 163, 164]. A novel strategy is designed for the modeling of the $t\,(\to \ell\nu b\gamma)\,q$ process. Common frameworks for automated computations of cross sections and event generation are not capable of modeling the hard-scattering $t\,(\to \ell\nu b\gamma)\,q$ process beyond leading order (LO) accuracy to date. However, studies of the impact of NLO QCD contributions on the $tq\gamma$ process revealed that the incorporation of such contributions is essential for a reliable and precise prediction of kinematic properties of this process. The $t\,(\to \ell\nu b\gamma)\,q$ process is therefore modeled by identifying $t\,(\to \ell\nu b\gamma)\,q$ contributions in a sample simulating the inclusive t-channel single-top-quark production including NLO QCD contributions in which photon radiation is solely handled via a parton shower algorithm. This sample includes both $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ contributions, which are identified with limited accuracy based on whether events in this sample fulfill the hypothesis of a radiative top-quark decay. The events identified as $t\,(\to \ell\nu b\gamma)\,q$ in this sample are used for the prediction of the $t\,(\to \ell\nu b\gamma)\,q$ contributions. The recall of the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure is about 89% at a precision of about 80% for the relevant phase space. Despite the limited precision, it is concluded that this approach for modeling the $t\,(\to \ell\nu b\gamma)\,q$ process is more reliable than the modeling via a dedicated sample simulating the $t\,(\to \ell\nu b\gamma)\,q$ process with LO accuracy. Conservative uncertainties are assigned to the $t\,(\to \ell\nu b\gamma)\,q$ predictions to account for the limited accuracy of the identification procedure. This approach for the modeling of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes is also used in Ref. [11], where these processes are background contributions.

A subset from the collected dataset is selected based on the expected detector signature of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events. Events are selected that contain a photon candidate, exactly one electron or muon candidate, a particle jet compatible with the hypothesis of being initiated by a b quark, and high missing transverse momentum. The background processes with a prompt photon in the final state contributing the most to this selected subset are $t\bar{t}\gamma$ and W-boson production in association with a photon and jets ($W\gamma$+jets). The selected subset is divided into two signal regions (SRs) enriched in $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ events, and into two control regions (CRs) that are enriched in $t\bar{t}\gamma$ ($t\bar{t}\gamma$ CR) and $W\gamma$+jets ($W\gamma$ CR) events, respectively. One SR ($\geq$ 1fj SR) contains events with a reconstructed jet characterized by a large pseudorapidity ($|\eta| > 2.5$), while events in the other SR (0fj SR) do not contain any such jets. Their presence is typical for the signal processes and provides large discriminative power against background contributions. The $tq\gamma$ contribution is about 11% in the $\geq$ 1fj SR and 5% in the 0fj SR. It is roughly four times larger than the $t\,(\to \ell\nu b\gamma)\,q$ contribution.

Additional significant background contributions arise from events with fake photons and fake leptons. The fake photons are classified into electrons misidentified as photons ($e \to \gamma$) and photons originating from hadron decays or hadrons misidentified as photons ($h \to \gamma$). The majority of the events with fake photons originate from top-quark pair production ($t\bar{t}$). All three contributions are estimated by data-driven techniques. The $e \to \gamma$ and $h \to \gamma$ contributions are estimated by correcting the simulated contributions by data-to-simulation SFs derived as a function of the kinematics and reconstruction type of the photon candidate. The $e \to \gamma$ SFs are determined by measuring the rate of electrons being misreconstructed and misidentified as prompt photons in data enriched in events with a Z boson decaying to an electron-positron pair, and dividing it by the respective predicted rate in simulation [7]. The $h \to \gamma$ SFs are determined by estimating the $h \to \gamma$ contribution based on data in which the photon candidate is required to fail specific photon-identification and/or photon-isolation criteria, and by dividing this estimation by the corresponding simulation-based prediction. The fake-lepton contribution

is modeled by reweighting events in collected data that must satisfy the same selection criteria as the aforementioned subset except that the lepton candidate must only satisfy looser lepton-identification and/or lepton-isolation requirements. The corresponding weights are computed with the asymptotic matrix method [265] using events in which the requirement for the presence of a photon candidate is dropped. The predicted relative $e \to \gamma$ and $h \to \gamma$ contributions to the SRs are 24% and 7%, respectively, and the predicted relative fake-lepton contribution is 3.4%. The $e \to \gamma$ SFs, as well as the methodology for deriving the $h \to \gamma$ SFs, are also used in Ref. [11].

Deep neural networks (NNs) consisting of fully-connected neuron layers are fitted to the predicted contributions of the 0fj SR and the $\geq$ 1fj SR. The NNs are designed to separate the events from $tq\gamma$ from any other contribution and assign each event a value ($\mathrm{NN_{out}}$) continuously distributed between 0 (background-like) and 1 ($tq\gamma$-like). Initial sets of input features for the NNs are selected based on their feature importance in gradient boosted decision trees fitted to the predicted contributions, presented in Ref. [7]. This work presents the iterative removal of input features that do not contribute significantly to the discriminative power of the NNs and the optimization of their neuron-layer structures. The separation strengths of the optimized NNs are significantly larger than the individual ones of the respective input features.

If the NNs are used to select the 70% of $tq\gamma$ events with the highest $\mathrm{NN_{out}}$, the $tq\gamma$-to-background ratio increases by a factor of 2 in the 0fj SR and by a factor of 2.4 in the $\geq$ 1fj SR, relative to the respective overall predicted ratio. This demonstrates the excellent discriminative power of the NNs. The larger increase for the $\geq$ 1fj SR is attributed to the large discriminative power of the characteristic features of the forward jet that are used as input features, which is absent in the 0fj SR.

Two cross-section measurements are conducted for fiducial phase spaces. For each measurement, a binned profile likelihood fit of the predicted $\mathrm{NN_{out}}$ distributions in the 0fj SR, $\geq$ 1fj SR, and $t\bar{t}\gamma$ CR, and the predicted overall yields in the $W\gamma$ CR to the respective distributions in the collected data is performed. Systematic uncertainties are considered as nuisance parameters and the overall contributions of the $t\bar{t}\gamma$ and $W\gamma$+jets processes to the collected data are considered as free-floating parameters. In the first measurement, the combined predicted overall $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ contribution is an additional free parameter of the fit. Its measured value is unfolded to a fiducial phase space defined at particle level. In the second measurement, only the overall $tq\gamma$ contribution is a free parameter of the fit, while $t\,(\to \ell\nu b\gamma)\,q$ is considered a background process. Its measured value is unfolded to an additional fiducial phase space defined at parton level, i.e., to a phase space for which parton showering and hadronization are not considered. The fiducial phase space at particle level is defined by criteria that resemble the selection criteria for the collected dataset as closely as possible. The fiducial phase space at parton level is defined solely by requirements on the final-state photon. Both fiducial phase spaces require the presence of a photon with a transverse momentum greater than 20 GeV.

The measured fiducial cross section at particle level is $303 \pm 9\,(\mathrm{stat.})\,^{+33}_{-32}(\mathrm{syst.})$ fb, which agrees with the prediction of 217^{+27}_{-15} fb including NLO QCD contributions at the level of 2.0σ. This value is to be compared to fixed-order theory calculations. The measured fiducial cross section at parton level is $688 \pm 23\,(\mathrm{stat.})\,^{+75}_{-71}\,(\mathrm{syst.})$ fb, which agrees with the prediction of 515^{+36}_{-42} fb including NLO QCD contributions at the level of 2.1σ.

The measured values of the fiducial cross sections exceed the prediction by about 40%. This is consistent with the excess of 1.42 ± 0.43 reported by the CMS collaboration in Ref. [20]. It needs to be studied whether this discrepancy will be resolved in future measurements and/or by improved theory predictions or if it will become more significant. The findings of Ref. [164] suggest that including higher-order corrections in the $tq\gamma$ prediction may decrease this discrepancy.

A test for the discovery of $tq\gamma$ is performed in the context of the parton-level measurement. The observed significance corresponds to 9.3σ. This marks the first observation of single-top-quark t-channel production in association with a photon. The results can further be interpreted within the

effective-field-theory (EFT) framework, providing a model-independent probe of operators related to top-quark–photon interactions. In Ref. [274], the result of the measurement for the parton-level phase space is included in a global fit of the corresponding Wilson coefficients to experimental results sensitive to these operators. It is shown that this result has a small sensitivity to the coefficients relative to the other included measurements, especially those of the $t\bar{t}\gamma$ process.

The precision of 11% of the measured fiducial cross sections is mainly limited by systematic uncertainties. The most recent measurement of $t\bar{t}\gamma$ process reached a precision of 5% [158], where statistical uncertainties are at the level of 1%. Assuming the same level of statistical precision, the systematic uncertainties of the results presented in this work need to be decreased by about 50% to reach a similar overall precision. The most dominant contribution to the overall size of systematic uncertainties arises from the uncertainties related to the modeling of the background contributions, especially to the modeling of the $t\bar{t}\gamma$ contribution, and the limited number of simulated events in the $tq\gamma$ sample. The former can be reduced by improving the understanding of the modeling, so that the related uncertainties become smaller, and by developing more powerful and/or more robust machine learning (ML) models for the signal and background separation, so that the measurements become less sensitive to the uncertainties in the background modeling. The latter can be decreased to a negligible level by enhancing the size of the $tq\gamma$ sample. This is also beneficial for the development of more powerful ML models, as larger sample sizes allow for more complex models to be fitted to them. Assuming a reduction of the impact of the $t\bar{t}\gamma$-related modeling uncertainties by 50% and neglecting the uncertainty due to the limited size of the $tq\gamma$ sample leads to a decrease of the overall size of the systematic uncertainties by 14%. Therefore, it is concluded that an overall precision close to 5% is achievable for the presented fiducial cross-section measurements with the larger datasets from LHC pp collisions that will become available in the future.

The observation of single-top-quark production in association with a photon is a milestone of particle physics and adds a tool for deepening the understanding of electroweak physics. The precision of the presented results and the predicted $tq\gamma$ yields suggest that differential cross-section measurements using this pp-collision dataset have a high potential. Such measurements will further improve the understanding of the process and of the top-quark–photon interaction in general. With the prospect of reaching a similar experimental precision as measurements of the $t\bar{t}\gamma$ process, it will be intriguing to study this process experimentally and theoretically in detail in the future.

References

[1] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, Eur. Phys. J. C **76** (2016) 666, arXiv: 1606.01813 [hep-ex].

[2] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, Eur. Phys. J. C **79** (2019) 205, arXiv: 1810.05087 [hep-ex].

[3] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex].

[4] B. Wendland, *Photon identification with the ATLAS detector*, 2021, URL: https://cds.cern.ch/record/2777847.

[5] B. Wendland, *Photon identification with the ATLAS detector*, PoS **EPS-HEP2021** (2022) 332.

[6] ATLAS Collaboration, *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, JHEP **05** (2024) 162, arXiv: 2308.13362 [hep-ex].

[7] H. Potti, *Seeing a single top quark: Search for $pp \to tq\gamma$ production and The combination of ATLAS searches for flavor changing neutral currents in $t \to Hq$ decays*, 2020, URL: https://cds.cern.ch/record/2748120.

[8] ATLAS Collaboration, *Observation of single-top-quark production in association with a photon at the ATLAS detector*, ATLAS-CONF-2022-013, 2022, URL: https://cds.cern.ch/record/2805217.

[9] B. Wendland, *Single top production ATLAS+CMS*, 2022, URL: https://cds.cern.ch/record/2810238.

[10] ATLAS Collaboration, *Observation of Single-Top-Quark Production in Association with a Photon Using the ATLAS Detector*, Phys. Rev. Lett. **131** (2023) 181901, arXiv: 2302.01283 [hep-ex].

[11] ATLAS Collaboration, *Search for flavour-changing neutral-current couplings between the top quark and the photon with the ATLAS detector at $\sqrt{s}$ = 13 TeV*, Phys. Lett. B **842** (2023) 137379, [Erratum: Phys. Lett. B **847** (2024) 138286], arXiv: 2205.02537 [hep-ex].

[12] J. Dalton, *A New System of Chemical Philosophy*, 1808, DOI: 10.1017/CBO9780511736391.

[13] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

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

[15] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[16] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex].

[17] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex].

[18] CDF Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626, arXiv: hep-ex/9503002.

[19] D0 Collaboration, *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) 2632, arXiv: hep-ex/9503003.

[20] CMS Collaboration, *Evidence for the Associated Production of a Single Top Quark and a Photon in Proton–Proton Collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett. **121** (2018) 221802, arXiv: 1808.02913 [hep-ex].

[21] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B **44** (1972) 189.

[22] S. Weinberg, *New Approach to the Renormalization Group*, Phys. Rev. D **8** (1973) 3497.

[23] S. Weinberg, *Non-Abelian Gauge Theories of the Strong Interactions*, Phys. Rev. Lett. **31** (1973) 494.

[24] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346.

[25] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343.

[26] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the color octet gluon picture*, Phys. Lett. B **47** (1973) 365.

[27] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[28] A. Salam, *Weak and electromagnetic interactions*, Conf. Proc. C **680519** (1968) 367.

[29] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[30] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[31] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[32] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[33] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[34] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156.

[35] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. **155** (1967) 1554.

[36] Particle Data Group, *Review of particle physics*, Phys. Rev. D **110** (2024) 030001.

[37] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[38] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[39] R. P. Feynman, *Quantum theory of gravitation*, Acta Phys. Polon. **24** (1963) 697.

[40] B. S. DeWitt, *Quantum Theory of Gravity. 1. The Canonical Theory*, Phys. Rev. **160** (1967) 1113.

[41] B. S. DeWitt, *Quantum Theory of Gravity. 2. The Manifestly Covariant Theory*, Phys. Rev. **162** (1967) 1195.

[42] G. 't Hooft and M. J. G. Veltman, *One loop divergencies in the theory of gravitation*, Ann. Inst. H. Poincare A Phys. Theor. **20** (1974) 69.

[43] M. H. Goroff and A. Sagnotti, *Quantum gravity at two loops*, Phys. Lett. B **160** (1985) 81.

[44] A. E. M. van de Ven, *Two loop quantum gravity*, Nucl. Phys. B **378** (1992) 309.

[45] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: hep-ex/9807003.

[46] SNO Collaboration, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv: nucl-ex/0204008.

[47] KATRIN Collaboration, *Direct neutrino-mass measurement based on 259 days of KATRIN data*, Science **388** (2025) adq9592, arXiv: 2406.13516 [nucl-ex].

[48] V. C. Rubin, W. K. Ford Jr. and N. Thonnard, *Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa through Sc*, Astrophys. J. Lett. **225** (1978) L107.

[49] J. P. Ostriker and P. J. Steinhardt, *Cosmic concordance*, 1995, arXiv: astro-ph/9505066.

[50] N. A. Bahcall, J. P. Ostriker, S. Perlmutter and P. J. Steinhardt, *The Cosmic Triangle: Revealing the State of the Universe*, Science **284** (1999) 1481.

[51] Planck Collaboration, *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, [Erratum: Astron. Astrophys. **652** (2021) C4], arXiv: 1807.06209 [astro-ph.CO].

[52] G. Jungman, M. Kamionkowski and K. Griest, *Supersymmetric dark matter*, Phys. Rept. **267** (1996) 195, arXiv: hep-ph/9506380.

[53] J. E. Kim, *Light Pseudoscalars, Particle Physics and Cosmology*, Phys. Rept. **150** (1987) 1.

[54] A. G. Riess et al., *A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SH0ES Team*, Astrophys. J. Lett. **934** (2022) L7, arXiv: 2112.04510 [astro-ph.CO].

[55] H0LiCOW Collaboration, *H0LiCOW – XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3σ tension between early- and late-Universe probes*, Mon. Not. Roy. Astron. Soc. **498** (2020) 1420, arXiv: 1907.04869 [astro-ph.CO].

[56] J. D. Barrow and M. S. Turner, *Baryosynthesis and the Origin of Galaxies*, Nature **291** (1981) 469.

[57] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[58] G. 't Hooft, *Symmetry Breaking Through Bell-Jackiw Anomalies*, Phys. Rev. Lett. **37** (1976) 8.

[59] V. A. Kuzmin, V. A. Rubakov and M. E. Shaposhnikov, *On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe*, Phys. Lett. B **155** (1985) 36.

[60] C. D. Froggatt and H. B. Nielsen, *Hierarchy of Quark Masses, Cabibbo Angles and CP Violation*, Nucl. Phys. B **147** (1979) 277.

[61] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, Phys. Rev. Lett. **38** (1977) 1440.

[62] LHCb Collaboration, *Measurement of CP-Averaged Observables in the $B^0 \to K^{*0}\mu^+\mu^-$ Decay*, Phys. Rev. Lett. **125** (2020) 011802, arXiv: 2003.04831 [hep-ex].

[63] LHCb Collaboration, *Comprehensive analysis of local and nonlocal amplitudes in the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, JHEP **09** (2024) 026, arXiv: 2405.17347 [hep-ex].

[64] LHCb Collaboration, *Determination of short- and long-distance contributions in $B^0 \to K^{*0}\mu^+\mu^-$ decays*, Phys. Rev. D **109** (2024) 052009, arXiv: 2312.09102 [hep-ex].

[65] LHCb Collaboration, *Amplitude Analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ Decay*, Phys. Rev. Lett. **132** (2024) 131801, arXiv: 2312.09115 [hep-ex].

[66] CMS Collaboration,
Angular analysis of the $\mathrm{B}^0 \to \mathrm{K}^*(892)^0\mu^+\mu^-$ *decay in proton-proton collisions at* $\sqrt{s}$ *= 13 TeV*,
Phys. Lett. B **864** (2025) 139406, arXiv: 2411.11820 [hep-ex].

[67] ATLAS Collaboration, *Angular analysis of* $B_d^0 \to K^*\mu^+\mu^-$ *decays in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **10** (2018) 047, arXiv: 1805.04000 [hep-ex].

[68] Belle Collaboration, *Test of Lepton-Flavor Universality in* $B \to K^*\ell^+\ell^-$ *Decays at Belle*,
Phys. Rev. Lett. **126** (2021) 161801, arXiv: 1904.02440 [hep-ex].

[69] Belle Collaboration, *Lepton-Flavor-Dependent Angular Analysis of* $B \to K^*\ell^+\ell^-$,
Phys. Rev. Lett. **118** (2017) 111801, arXiv: 1612.05014 [hep-ex].

[70] LHCb Collaboration,
Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ *decays*,
JHEP **06** (2014) 133, arXiv: 1403.8044 [hep-ex].

[71] CMS Collaboration, *Test of lepton flavor universality in* $B^\pm \to K^\pm\mu^+\mu^-$ *and* $B^\pm \to K^\pm e^+e^-$ *decays in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Rept. Prog. Phys. **87** (2024) 077802,
arXiv: 2401.07090 [hep-ex].

[72] Belle Collaboration,
Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ *decays*,
JHEP **03** (2021) 105, arXiv: 1908.01848 [hep-ex].

[73] LHCb Collaboration,
Branching Fraction Measurements of the Rare $B_s^0 \to \phi\mu^+\mu^-$ *and* $B_s^0 \to f_2'(1525)\mu^+\mu^-$*- Decays*,
Phys. Rev. Lett. **127** (2021) 151801, arXiv: 2105.14007 [hep-ex].

[74] LHCb Collaboration, *Test of lepton universality in beauty-quark decays*,
Nature Phys. **18** (2022) 277, arXiv: 2103.11769 [hep-ex].

[75] LHCb Collaboration, *Test of lepton universality in* $b \to s\ell^+\ell^-$ *decays*,
Phys. Rev. Lett. **131** (2023) 051803, arXiv: 2212.09152 [hep-ex].

[76] CMS Collaboration,
High-precision measurement of the W boson mass with the CMS experiment at the LHC,
(2024), Submitted to Nature, arXiv: 2412.13872 [hep-ex].

[77] CDF Collaboration, *High-precision measurement of the W boson mass with the CDF II detector*,
Science **376** (2022) 170.

[78] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay,
Evidence for the 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (1964) 138.

[79] S. W. Herb et al.,
Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions,
Phys. Rev. Lett. **39** (1977) 252.

[80] CDF Collaboration, *Exclusion of exotic top-like quarks with* $-4/3$ *electric charge using jet-charge tagging in single-lepton* $t\bar{t}$ *events at CDF*, Phys. Rev. D **88** (2013) 032003.

[81] D0 Collaboration, *Measurement of the Electric Charge of the Top Quark in* $\boldsymbol{t\bar{t}}$ *Events*,
Phys. Rev. D **90** (2014) 051101, [Erratum: Phys.Rev. D **90** (2014) 079904],
arXiv: 1407.4837 [hep-ex].

[82] ATLAS Collaboration,
Measurement of the top quark charge in pp collisions at $\sqrt{s} = 7$ *TeV with the ATLAS detector*,
JHEP **11** (2013) 031, arXiv: 1307.4568 [hep-ex].

[83] CMS Collaboration, *Constraints on the Top-Quark Charge from Top-Pair Events*,
CMS-PAS-TOP-11-031, 2012, URL: https://cds.cern.ch/record/1429970.

[84] ATLAS, CDF, CMS and D0 Collaborations,
First combination of Tevatron and LHC measurements of the top-quark mass,
ATLAS-CONF-2014-008, 2014, URL: https://cds.cern.ch/record/1669819.

[85] CMS Collaboration, *Measurement of the top quark mass using a profile likelihood approach with the lepton + jets final states in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*,
Eur. Phys. J. C **83** (2023) 963, arXiv: 2302.01967 [hep-ex].

[86] CMS Collaboration,
Measurement of the top quark mass using proton–proton data at $\sqrt{s} = 7$ *and* 8 *TeV*,
Phys. Rev. D **93** (2016) 072004, arXiv: 1509.04044 [hep-ex].

[87] M. Jeżabek and J. Kühn, *QCD corrections to semileptonic decays of heavy quarks*,
Nucl. Phys. B **314** (1989) 1.

[88] A. Czarnecki and K. Melnikov, *Two-loop QCD corrections to top quark width*,
Nucl. Phys. B **544** (1999) 520.

[89] K. G. Chetyrkin, R. Harlander, T. Seidensticker and M. Steinhauser,
Second order QCD corrections to $\Gamma(t \to Wb)$, Phys. Rev. D **60** (1999) 114015.

[90] J. Gao, C. S. Li and H. X. Zhu, *Top Quark Decay at Next-to-Next-to Leading Order in QCD*,
Phys. Rev. Lett. **110** (2013) 042001, arXiv: 1210.2808 [hep-ph].

[91] CMS Collaboration,
Measurement of the ratio $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq)$ *in pp collisions at* $\sqrt{s} = 8$ *TeV*,
Phys. Lett. B **736** (2014) 33, arXiv: 1404.2292 [hep-ex].

[92] ATLAS Collaboration, *Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, ATLAS-CONF-2019-038, 2019,
URL: https://cds.cern.ch/record/2684952.

[93] C. Herwig, T. Ježo and B. Nachman,
Extracting the Top-Quark Width from Nonresonant Production,
Phys. Rev. Lett. **122** (2019) 231803, arXiv: 1903.10519 [hep-ex].

[94] J. Haller et al.,
Update of the global electroweak fit and constraints on two-Higgs-doublet models,
Eur. Phys. J. C **78** (2018) 675, arXiv: 1803.01853 [hep-ph].

[95] J. A. Aguilar-Saavedra, *A minimal set of top anomalous couplings*,
Nucl. Phys. B **812** (2009) 181.

[96] W. Bernreuther et al., *QCD corrections to static heavy quark form-factors*,
Phys. Rev. Lett. **95** (2005) 261802, arXiv: hep-ph/0509341.

[97] A. Soni and R. M. Xu, *Electric dipole moment of the top quark in Higgs-boson-exchange models of CP nonconservation*, Phys. Rev. Lett. **69** (1992) 33.

[98] F. Hoogeveen, *The standard model prediction for the electric dipole moment of the electron*,
Nucl. Phys. B **341** (1990) 322.

[99] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*,
Phys. Rev. D **2** (1970) 1285.

[100] J. A. Aguilar-Saavedra,
Top flavor-changing neutral interactions: Theoretical expectations and experimental detection,
Acta Phys. Polon. B **35** (2004) 2695, arXiv: hep-ph/0409342.

[101] J. F. Kamenik, M. Papucci and A. Weiler, *Constraining the dipole moments of the top quark*,
Phys. Rev. D **85** (2012) 071501.

[102] A. O. Bouzas and F. Larios, *Electromagnetic dipole moments of the top quark*,
Phys. Rev. D **87** (2013) 074015.

[103] C. Abel et al., *Measurement of the Permanent Electric Dipole Moment of the Neutron*, Phys. Rev. Lett. **124** (2020) 081803, arXiv: 2001.11966 [hep-ex].

[104] B. Graner, Y. Chen, E. G. Lindahl and B. R. Heckel, *Reduced Limit on the Permanent Electric Dipole Moment of* ^{199}Hg, Phys. Rev. Lett. **116** (2016) 161601.

[105] B. K. Sahoo and B. P. Das, *Relativistic Normal Coupled-Cluster Theory for Accurate Determination of Electric Dipole Moments of Atoms: First Application to the* ^{199}Hg *Atom*, Phys. Rev. Lett. **120** (2018) 203001.

[106] K. Yanase, *Screening of nucleon electric dipole moments in atomic systems*, Phys. Rev. C **103** (2021) 035501, arXiv: 2008.03678 [nucl-th].

[107] V. Cirigliano, W. Dekens, J. de Vries and E. Mereghetti, *Is there room for CP violation in the top-Higgs sector?*, Phys. Rev. D **94** (2016) 016002.

[108] T. Ibrahim and P. Nath, *Top quark electric dipole moment in a minimal supersymmetric standard model extension with vectorlike multiplets*, Phys. Rev. D **82** (2010) 055001.

[109] W. Hollik, J. Illana, S. Rigolin, C. Schappacher and D. Stöckinger, *Top dipole form factors and loop-induced CP violation in supersymmetry*, Nucl. Phys. B **551** (1999) 3.

[110] K. Agashe, G. Perez and A. Soni, *Collider signals of top quark flavor violation from a warped extra dimension*, Phys. Rev. D **75** (2007).

[111] T. Ibrahim and P. Nath, *Chromoelectric dipole moment of the top quark in models with vectorlike multiplets*, Phys. Rev. D **84** (2011).

[112] J. Montaño, H. Novales-Sánchez and J. J. Toscano, *Effects of universal extra dimensions on top-quark electromagnetic interactions*, J. Phys. G **47** (2019) 015002, arXiv: 1908.06226 [hep-ph].

[113] J. M. Yang, B.-L. Young and X. Zhang, *Flavor-changing top quark decays in R-parity-violating supersymmetric models*, Phys. Rev. D **58** (1998) 055001, arXiv: hep-ph/9705341.

[114] D. Delépine and S. Khalil, *Top flavour violating decays in general supersymmetric models*, Phys. Lett. B **599** (2004) 62, arXiv: hep-ph/0406264.

[115] J. J. Cao et al., *Supersymmetry-induced flavor-changing neutral-current top-quark processes at the CERN Large Hadron Collider*, Phys. Rev. D **75** (2007) 075021, arXiv: hep-ph/0702264.

[116] J.-L. Yang, T.-F. Feng, H.-B. Zhang, G.-Z. Ning and X.-Y. Yang, *Top quark decays with flavor violation in the B-LSSM*, Eur. Phys. J. C **78** (2018) 438, arXiv: 1806.01476 [hep-ph].

[117] D. Atwood, L. Reina and A. Soni, *Phenomenology of two Higgs doublet models with flavor-changing neutral currents*, Phys. Rev. D **55** (1997) 3156, arXiv: hep-ph/9609279.

[118] S. Béjar, J. Guasch and J. Solà, *Loop induced flavor changing neutral decays of the top quark in a general two-Higgs-doublet model*, Nucl. Phys. B **600** (2001) 21, arXiv: hep-ph/0011091.

[119] T. Appelquist and J. Carazzone, *Infrared singularities and massive fields*, Phys. Rev. D **11** (1975) 2856.

[120] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, JHEP **10** (2010) 085, arXiv: 1008.4884 [hep-ph].

[121] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[122] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[123] ATLAS Collaboration,
Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC,
ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054.

[124] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s}$ = 13 TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002,
arXiv: 1606.02625 [hep-ex].

[125] O. Aberle et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*,
2020, URL: https://cds.cern.ch/record/2749422.

[126] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*,
Adv. Ser. Direct. High Energy Phys. **5** (1989) 1, arXiv: hep-ph/0409313.

[127] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*,
Sov. Phys. JETP **46** (1977) 641.

[128] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*,
Sov. J. Nucl. Phys. **15** (1972) 438.

[129] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*,
Nucl. Phys. B **126** (1977) 298.

[130] ATLAS Collaboration,
Observation of four-top-quark production in the multilepton final state with the ATLAS detector,
Eur. Phys. J. C **83** (2023) 496, [Erratum: Eur. Phys. J. C **84** (2024) 156],
arXiv: 2303.15061 [hep-ex].

[131] CMS Collaboration,
Observation of four top quark production in proton-proton collisions at s=13TeV,
Phys. Lett. B **847** (2023) 138290, arXiv: 2305.13439 [hep-ex].

[132] M. Beneke, P. Falgari, S. Klein and C. Schwinn,
Hadronic top-quark pair production with NNLL threshold resummation,
Nucl. Phys. B **855** (2012) 695, arXiv: 1109.1536 [hep-ph].

[133] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*,
Phys. Lett. B **710** (2012) 612, arXiv: 1111.5869 [hep-ph].

[134] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$*,
Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph].

[135] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 054, arXiv: 1207.0236 [hep-ph].

[136] M. Czakon and A. Mitov,
NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction,
JHEP **01** (2013) 080, arXiv: 1210.6832 [hep-ph].

[137] M. Czakon, P. Fiedler and A. Mitov,
Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$,
Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[138] M. Czakon and A. Mitov,
Top++: A program for the calculation of the top-pair cross-section at hadron colliders,
Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[139] CDF Collaboration, *First Observation of Electroweak Single Top Quark Production*, Phys. Rev. Lett. **103** (2009) 092002, arXiv: 0903.0885 [hep-ex].

[140] D0 Collaboration, *Observation of Single Top Quark Production*, Phys. Rev. Lett. **103** (2009) 092001, arXiv: 0903.0850 [hep-ex].

[141] D0 Collaboration, *Model-independent measurement of t-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Lett. B **705** (2011) 313, arXiv: 1105.2788 [hep-ex].

[142] CDF and D0 Collaborations, *Observation of s-channel production of single top quarks at the Tevatron*, Phys. Rev. Lett. **112** (2014) 231803, arXiv: 1402.5126 [hep-ex].

[143] CMS Collaboration, *Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s}$ = 7 TeV*, JHEP **12** (2012) 035, arXiv: 1209.4533 [hep-ex].

[144] ATLAS Collaboration, *Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Phys. Lett. B **717** (2012) 330, arXiv: 1205.3130 [hep-ex].

[145] CMS Collaboration, *Observation of the Associated Production of a Single Top Quark and a W Boson in pp Collisions at $\sqrt{s}$ = 8 TeV*, Phys. Rev. Lett. **112** (2014) 231802, arXiv: 1401.2942 [hep-ex].

[146] ATLAS Collaboration, *Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment*, JHEP **01** (2016) 064, arXiv: 1510.03752 [hep-ex].

[147] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034, arXiv: 1007.1327 [hep-ph].

[148] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191** (2015) 74, arXiv: 1406.4403 [hep-ph].

[149] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph].

[150] N. Kidonakis, *Top Quark Production*, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons, 2014 139, arXiv: 1311.0283 [hep-ph].

[151] CMS Collaboration, *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **08** (2018) 011, arXiv: 1711.02547 [hep-ex].

[152] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: 1901.03584 [hep-ex].

[153] 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. D **91** (2015) 072007, arXiv: 1502.00586 [hep-ex].

[154] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv: 1806.00425 [hep-ex].

[155] CMS Collaboration, *Observation of $t\bar{t}H$ Production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv: 1804.02610 [hep-ex].

[156] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **07** (2020) 124, arXiv: 2002.07546 [hep-ex].

[157] CMS Collaboration, *Observation of Single Top Quark Production in Association with a Z Boson in Proton–Proton Collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. Lett. **122** (2019) 132003, arXiv: 1812.05900 [hep-ex].

[158] ATLAS Collaboration, *Measurements of inclusive and differential cross-sections of $t\bar{t}\gamma$ production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **10** (2024) 191, arXiv: 2403.09452 [hep-ex].

[159] ATLAS Collaboration, *Observation of $t\bar{t}\gamma\gamma$ production at $\sqrt{s}$ =13 TeV with the ATLAS detector*, (2025), Submitted to Phys. Lett. B, arXiv: 2506.05018 [hep-ex].

[160] M. Fael and T. Gehrmann, *Probing top quark electromagnetic dipole moments in single-top-plus-photon production*, Phys. Rev. D **88** (2013) 033003, arXiv: 1307.1349 [hep-ph].

[161] S. M. Etesami, S. Khatibi and M. Mohammadi Najafabadi, *Measuring anomalous WW γ and $t\bar{t}\gamma$ couplings using top+ γ production at the LHC*, Eur. Phys. J. C **76** (2016) 533, arXiv: 1606.02178 [hep-ph].

[162] 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].

[163] D. Pagani, H.-S. Shao, I. Tsinikos and M. Zaro, *Automated EW corrections with isolated photons: $t\bar{t}\gamma$, $t\bar{t}\gamma\gamma$ and $t\gamma j$ as case studies*, JHEP **09** (2021) 155, arXiv: 2106.02059 [hep-ph].

[164] N. Kidonakis and N. Yamanaka, *QCD corrections in $tq\gamma$ production at hadron colliders*, Eur. Phys. J. C **82** (2022) 670, arXiv: 2201.12877 [hep-ph].

[165] CMS Collaboration, *Search for anomalous single top quark production in association with a photon in pp collisions at $\sqrt{s}$ = 8 TeV*, JHEP **04** (2016) 035, arXiv: 1511.03951 [hep-ex].

[166] ATLAS Collaboration, *Search for flavour-changing neutral currents in processes with one top quark and a photon using 81 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS experiment*, Phys. Lett. B **800** (2020) 135082, arXiv: 1908.08461 [hep-ex].

[167] CMS Collaboration, *Search for flavor changing neutral current interactions of the top quark in final states with a photon and additional jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. D **109** (2024) 072004, arXiv: 2312.08229 [hep-ex].

[168] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19; CERN-LHCC-2010-013, 2010, URL: https://cds.cern.ch/record/1291633, Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: https://cds.cern.ch/record/1451888.

[169] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run-2*, Eur. Phys. J. C **80** (2020) 1194, arXiv: 2007.07624 [hep-ex].

[170] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[171] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [hep-ex].

[172] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[173] ATLAS Collaboration,
Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC,
Eur. Phys. J. C **83** (2023) 982, arXiv: 2212.09379 [hep-ex].

[174] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand,
Parton Fragmentation and String Dynamics, Phys. Rept. **97** (1983) 31.

[175] B. Andersson, *The Lund Model*, Nucl. Phys. A **461** (1987) 513C.

[176] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*,
Nucl. Phys. B **238** (1984) 492.

[177] J.-C. Winter, F. Krauss and G. Soff, *A Modified cluster hadronization model*,
Eur. Phys. J. C **36** (2004) 381, arXiv: hep-ph/0311085.

[178] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
arXiv: 1005.4568 [physics.ins-det].

[179] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, 2010,
URL: https://cds.cern.ch/record/1300517.

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

[181] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965.

[182] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244,
arXiv: 1207.1303 [hep-ph].

[183] ATLAS Collaboration, *Emulating the impact of additional proton–proton interactions in the ATLAS simulation by presampling sets of inelastic Monte Carlo events*,
Comput. Softw. Big Sci. **6** (2022) 3, arXiv: 2102.09495 [hep-ex].

[184] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*,
ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187.

[185] T. Cornelissen et al.,
Concepts, Design and Implementation of the ATLAS New Tracking (NEWT),
ATL-SOFT-PUB-2007-007; ATL-COM-SOFT-2007-002, 2007,
URL: https://cds.cern.ch/record/1020106.

[186] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex].

[187] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332,
arXiv: 1611.10235 [hep-ex].

[188] ATLAS Collaboration,
Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1,
Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[189] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*,
ATL-LARG-PUB-2008-002, 2008, URL: https://cds.cern.ch/record/1099735.

[190] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012,
URL: https://cds.cern.ch/record/1449796.

[191] ATLAS Collaboration,
Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data,
JINST **19** (2024) P02009, arXiv: 2309.05471 [hep-ex].

[192] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex].

[193] ATLAS Collaboration, *Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **83** (2023) 686, arXiv: 2212.07338 [hep-ex].

[194] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex].

[195] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[196] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph].

[197] M. Cacciari, *FastJet: A Code for fast k_t clustering, and more*, Proceedings, Deep inelastic scattering, 14th International Workshop, DIS 2006, 2006 487, arXiv: hep-ph/0607071 [hep-ph].

[198] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv: 2007.02645 [hep-ex].

[199] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s}$ = 8 *TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex].

[200] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with* $t\bar{t}$ *events in pp collisions at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex].

[201] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281.

[202] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, Eur. Phys. J. C **83** (2023) 681, arXiv: 2211.16345 [physics.data-an].

[203] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, ATL-PHYS-PUB-2017-011, 2017, URL: https://cds.cern.ch/record/2270366.

[204] ATLAS Collaboration, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, ATL-PHYS-PUB-2018-025, 2018, URL: https://cds.cern.ch/record/2645405.

[205] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, ATL-PHYS-PUB-2017-003, 2017, URL: https://cds.cern.ch/record/2255226.

[206] ATLAS Collaboration, *The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using 140 fb^{-1} of* $\sqrt{s}$ = 13 *TeV pp collisions*, Eur. Phys. J. C **85** (2025) 606, arXiv: 2402.05858 [hep-ex].

[207] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007, arXiv: physics/0703039 [physics.data-an].

[208] L. Lyons, D. Gibaut and P. Clifford,
How to combine correlated estimates of a single physical quantity,
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **270** (1988) 110.

[209] A. Valassi and R. Chierici, *Information and treatment of unknown correlations in the combination of measurements using the BLUE method*, Eur. Phys. J. C **74** (2014) 2717, arXiv: 1307.4003 [physics.data-an].

[210] ATLAS Collaboration,
Performance of electron and photon triggers in ATLAS during LHC Run 2,
Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

[211] F. Braren, *Measurement and Interpretation of Higgs Boson Differential Cross Sections in the Diphoton Decay Channel and Measurement of the Photon Identification Efficiency in the ATLAS Experiment*, 2020, URL: https://ediss.sub.uni-hamburg.de/handle/ediss/8622.2.

[212] NNPDF Collaboration, *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph].

[213] S. Frixione, *Isolated photons in perturbative QCD*, Phys. Lett. B **429** (1998) 369, arXiv: hep-ph/9801442.

[214] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[215] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. A **462** (2001) 152.

[216] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, SciPost Phys. **8** (2020) 026, arXiv: 1912.05451 [hep-ph].

[217] M. Cacciari, G. P. Salam and G. Soyez, *The catchment area of jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph].

[218] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: hep-ph/0702198.

[219] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk,
Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations,
JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph].

[220] ATLAS Collaboration, *Measurement of t-channel production of single top quarks and antiquarks in pp collisions at 13 TeV using the full ATLAS Run 2 data sample*,
JHEP **05** (2024) 305, arXiv: 2403.02126 [hep-ex].

[221] R. Frederix, E. Re and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP **09** (2012) 130, arXiv: 1207.5391 [hep-ph].

[222] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[223] 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].

[224] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

[225] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*,
JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

[226] K. Melnikov, M. Schulze and A. Scharf,
QCD corrections to top quark pair production in association with a photon at hadron colliders,
Phys. Rev. D **83** (2011) 074013, arXiv: 1102.1967 [hep-ph].

[227] ATLAS Collaboration, *Measurement of the charge asymmetry in top-quark pair production in association with a photon with the ATLAS experiment*, Phys. Lett. B **843** (2023) 137848,
arXiv: 2212.10552 [hep-ex].

[228] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034,
arXiv: 1905.09127 [hep-ph].

[229] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: 1907.13071 [hep-ph].

[230] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*,
Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph].

[231] A. Denner, S. Dittmaier and L. Hofer,
Collier: A fortran-based complex one-loop library in extended regularizations,
Comput. Phys. Commun. **212** (2017) 220, arXiv: 1604.06792 [hep-ph].

[232] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039,
arXiv: 0808.3674 [hep-ph].

[233] S. Schumann and F. Krauss,
A parton shower algorithm based on Catani–Seymour dipole factorisation, JHEP **03** (2008) 038,
arXiv: 0709.1027 [hep-ph].

[234] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
A critical appraisal of NLO+PS matching methods, JHEP **09** (2012) 049,
arXiv: 1111.1220 [hep-ph].

[235] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
QCD matrix elements + parton showers. The NLO case, JHEP **04** (2013) 027,
arXiv: 1207.5030 [hep-ph].

[236] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD Matrix Elements + Parton Showers*,
JHEP **11** (2001) 063, arXiv: hep-ph/0109231.

[237] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*,
JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph].

[238] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*,
ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168.

[239] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber,
Single-top hadroproduction in association with a W boson, JHEP **07** (2008) 029,
arXiv: 0805.3067 [hep-ph].

[240] NNPDF Collaboration, *Parton distributions from high-precision collider data*,
Eur. Phys. J. C **77** (2017) 663, arXiv: 1706.00428 [hep-ph].

[241] NNPDF Collaboration,
Illuminating the photon content of the proton within a global PDF analysis,
SciPost Phys. **5** (2018) 008, arXiv: 1712.07053 [hep-ph].

[242] ATLAS Collaboration, *TRExFitter*, 2016,
URL: https://gitlab.cern.ch/TRExStats/TRExFitter.

[243] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats, (2012).

[244] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057,
arXiv: 1009.1003 [physics.data-an].

[245] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: physics/0306116 [physics.data-an].

[246] R. Brun and F. Rademakers, *ROOT – An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81.

[247] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219.

[248] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.

[249] W. C. Davidon, *Variable Metric Method for Minimization*, SIAM Journal on Optimization **1** (1991) 1.

[250] R. Fletcher, *A new approach to variable metric algorithms*, The Computer Journal **13** (1970) 317.

[251] M. J. D. Powell, *Variable Metric Methods for Constrained Optimization*, Mathematical Programming The State of the Art (1983) 288, URL: https://doi.org/10.1007/978-3-642-68874-4_12.

[252] J. H. Friedman, *Data Analysis Techniques for High-Energy Particle Physics*, 3rd CERN School of Computing (1974) 271.

[253] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, [Erratum: Eur. Phys. J. C **73** (2013) 2501], arXiv: 1007.1727 [physics.data-an].

[254] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60.

[255] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Trans. Amer. Math. Soc. **54** (1943) 426.

[256] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2022) 95, arXiv: 2109.10627 [hep-ex].

[257] ATLAS Collaboration, *Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139 fb^{-1} of ATLAS proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **83** (2023) 728, arXiv: 2301.06319 [hep-ex].

[258] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s}$ = 13 TeV in ATLAS*, Eur. Phys. J. C **79** (2019) 382, arXiv: 1812.01697 [hep-ex].

[259] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph].

[260] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections*, Eur. Phys. J. C **64** (2009) 653, arXiv: 0905.3531 [hep-ph].

[261] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. D **89** (2014) 033009, arXiv: 1302.6246 [hep-ph].

[262] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph].

[263] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv: 1103.2792 [hep-ph].

[264] ATLAS Collaboration, *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*, 2016, URL: https://cds.cern.ch/record/2120133.

[265] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, JINST **18** (2023) T11004, arXiv: 2211.16178 [hep-ex].

[266] ATLAS Collaboration, *Measurement of the $Z(\to \ell^+\ell^-)\gamma$ production cross-section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **03** (2020) 054, arXiv: 1911.04813 [hep-ex].

[267] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM **KDD '16** (2016) 785, URL: http://doi.acm.org/10.1145/2939672.2939785.

[268] F. Chollet et al., *Keras*, 2015, URL: https://keras.io.

[269] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, 2015, URL: https://www.tensorflow.org.

[270] D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, CoRR **abs/1412.6980** (2014), arXiv: 1412.6980.

[271] X. Glorot and Y. Bengio, *Understanding the difficulty of training deep feedforward neural networks*, Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, PMLR **9** (2010) 249, URL: https://proceedings.mlr.press/v9/glorot10a.html.

[272] D. Buscher, *Search for Higgs bosons with b-jets in the final state in proton-proton collisions with the ATLAS experiment*, 2016, URL: https://freidok.uni-freiburg.de/data/11321.

[273] R. D. Cousins, *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms*, 2013, URL: https://api.semanticscholar.org/CorpusID:5936965.

[274] M. Miralles López, *Probing the top quark couplings within the ATLAS detector and EFT global fits*, 2023, URL: https://hdl.handle.net/10550/88737.

Acknowledgements

My physics studies and this thesis would not have been possible in this form without the support of exceptional people who accompanied me along the way. At this point, I would like to express my gratitude to them.

First and foremost, I would like to thank Prof. Dr. Kevin Kröninger, who gave me the opportunity to complete all my academic theses at his department and served as the primary referee of this thesis. I am also grateful for the trust and support he extended to me during my time there. I also wish to thank Dr. Dominik Mitzel, who kindly agreed to review this thesis as the second referee.

A very special thanks goes to Prof. Dr. Johannes Erdmann, who supervised all my theses. Thank you for your support, the many insightful discussions, and your helpful feedback. I look back on a very enjoyable and inspiring collaboration.

In particular, I would also like to thank the "$tq\gamma$" analysis team, which included Johannes Erdmann, Peter Onyisi, and Harish Potti. Working together was always smooth, productive, and constructive. Overall, I greatly enjoyed our collaboration, and I am convinced that, just like in sports, success can only be achieved when the team is in harmony. At this point, I also want to thank Maria Jose Costa, Knut Zoch, and Leonardo Carminati for their meticulous and constructive review of the analysis.

I am deeply grateful to the members of the ATLAS collaboration who create a lively and inspiring research environment. I especially want to thank the Top Group and the Photon Identification Group for providing continuous feedback and constructive discussions on the progress of the work presented in this thesis.

I would also like to thank Davide Pagani and Marco Zaro, who are co-authors of Ref. [163] and helped me reproduce the technical setup for their work used for the theoretical predictions in Section 7.4.2.

The path to my physics studies at the TU Dortmund University was laid already during my final years of school. My interest in studying physics and my enthusiasm for particle physics were evoked particularly through the Masterclass program of Netzwerk Teilchenwelt, which gave me the opportunity early in my career to visit CERN and thus gain insights into modern particle physics research at one of the largest research facilities. I warmly thank the members and organizers of Netzwerk Teilchenwelt for their commitment and enthusiasm in introducing modern research to aspiring scientists. I also want to thank my former physics teacher Christian Wolf, who made my participation possible and supervised me through it.

Throughout my studies, I was accompanied by many companions, each of whom helped shape me and made sure I look back on this time with many good memories. First, I want to thank my fellow students who have been with me since day one of my journey at the university: Kevin Foryt, Fabian Hundehege, Felix Brauers, and Ben Ponczeck. A very special thanks goes to Rigo Bause. What would all the lab courses have been without the best lab partner? For sure much less fun! Our work together on countless reports taught me so much, and thanks to you, even the rather tedious lab courses are remembered positively. Thank you also for your feedback on chapters 2 and 3 of my thesis. And then there is someone who has been a guiding presence in my life for 20 years, through school, university, and beyond: Marcel Golz. Thank you for the many shared experiences and for everything I was able to learn from you. I also appreciate your feedback on chapters 2 and 3.

There are no superlatives that can adequately describe the time I spent with Kevin Sedlaczek during my PhD phase. Together we braved the jungle of the IT infrastructure of our department, battled through the Machine Learning School, ATLAS workshops, and the everyday madness. The joyful memories are countless, and this time remains unforgettable thanks to you. I cannot thank you enough for reading a large portion of my thesis and providing feedback on the abstract, chapters 1, 2, 3, 4, 6, 7.4, 7.6, and 9–11.

I would also like to thank Jan Lukas Späh, whom I had the pleasure of supervising during both his bachelor's and master's theses. I sometimes wonder who learned more from whom, as I have not learned more about particle physics from any other fellow student than I did from him. Thank you for this inspiring and enjoyable collaboration. I am also grateful for the very helpful and detailed feedback on chapters 2, 3, and 4.2.

Furthermore, I would like to thank my additional proofreaders for their helpful and meticulous reading of the thesis: Christopher Krause for feedback on the chapter 4.1, Benedikt Gocke for feedback on chapters 7.5, 8.1, and 8.2, and Nils Julius Abicht for feedback on the chapters 5, 7–7.3, and 8.3.

I also wish to thank the entire AG Kröninger department (formerly Experimental Physics IV) for providing a consistently positive and calm environment throughout my time there. Thank you for the support and the many wonderful experiences that I will always carry with me.

Finally, I would like to thank my friends and family who supported me during my studies and doctoral work. In particular, I wish to thank my partner, Elvira Zentner, who stood by me during all the good and challenging phases of my PhD, offering support, encouragement, and strength.

Appendices

A Measurement of the photon-identification efficiency with the inclusive photon method

This appendix provides additional material for the measurement of the PID efficiency for prompt photons conducted with the inclusive photon method presented in Section 5.3.

A.1 Overview of the single-photon triggers

The events used in the inclusive photon method are required to pass the criteria of one of the high-level single-photon triggers [210]. The triggers differ by the applied E_{T}^{γ} threshold, which is denoted by "gX", meaning that the detector signatures compatible with those expected from prompt photons fulfill $E_{\mathrm{T}}^{\gamma} > X\,\mathrm{GeV}$. These triggers apply the requirement of the loose PID WP, denoted by "loose" in the naming. Some triggers are seeded by the L1 trigger, denoted by "L1EMY", where Y denotes the E_{T}^{γ} threshold used by the L1 trigger. Most of the triggers are prescaled, i.e., events passing the respective trigger requirements are rejected with a constant frequency during data taking. A prescale weight is assigned to events in the collected data to correct for the loss of events by the prescaling.

- `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`.

A.2 Track-isolation efficiencies

Figures 54, 55, 56, and 57 show the track-isolation efficiencies for photon candidates in collected data, for prompt photons, and for fake photons for all $|\eta^{\gamma}|$ bins as a function of E_{T}^{γ}. The efficiencies for fake photons are determined by Equations (35) and (36). The former two figures display the efficiencies for converted photon candidates, the latter two for unconverted photon candidates.

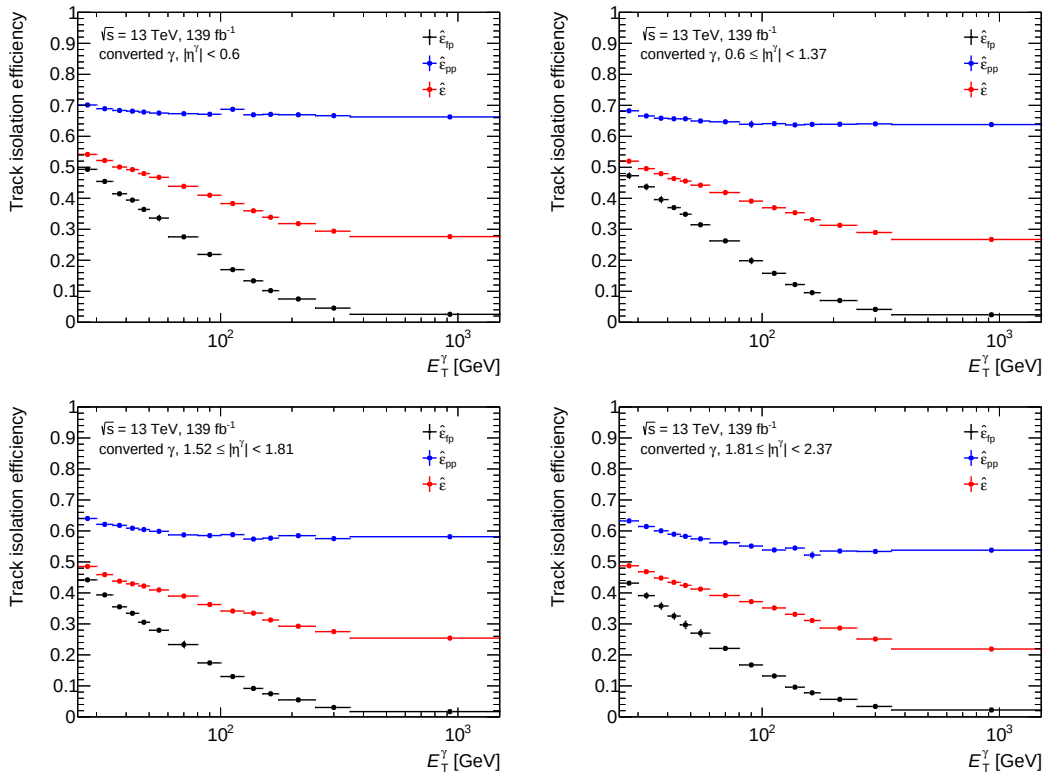


Figure 54: Track-isolation efficiencies for photon candidates in collected data, for prompt photons, and for fake photons as a function of E_{T}^{γ} for the different $|\eta^{\gamma}|$ bins. The efficiencies are shown for converted photon candidates and those for fake photons are determined by Equation (35).

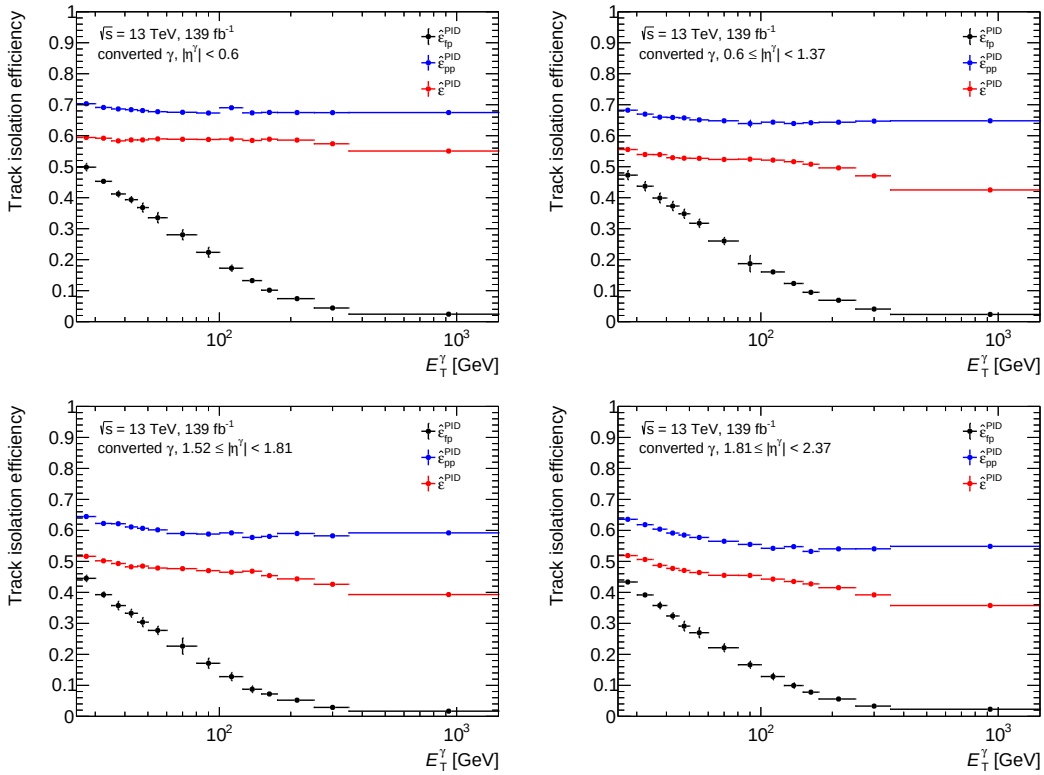


Figure 55: Track-isolation efficiencies for photon candidates in collected data, for prompt photons, and for fake photons determined by Equations (36) that the requirements of the tight PID WP as a function of E_{T}^{γ} for the different $|\eta^{\gamma}|$ bins. The efficiencies are shown for converted photon candidates and those for fake photons are determined by Equation (36).

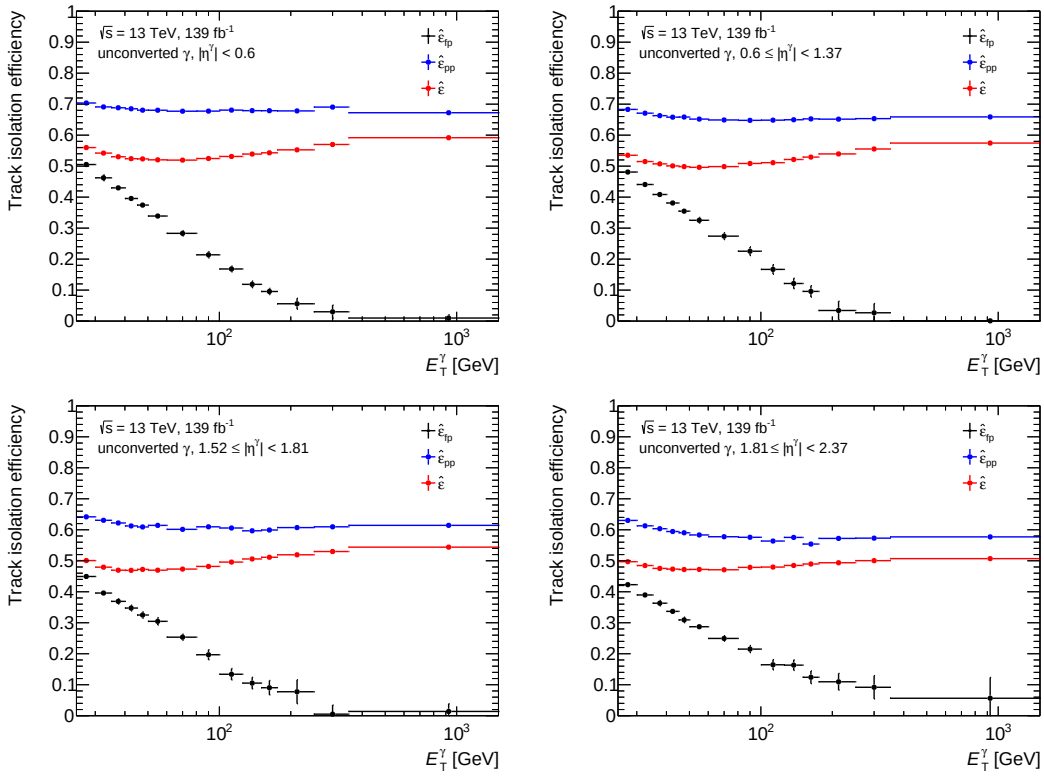


Figure 56: Track-isolation efficiencies for photon candidates in collected data, for prompt photons, and for fake photons as a function of E_{T}^{γ} for the different $|\eta^{\gamma}|$ bins. The efficiencies are shown for unconverted photon candidates and those for fake photons are determined by Equation (35).

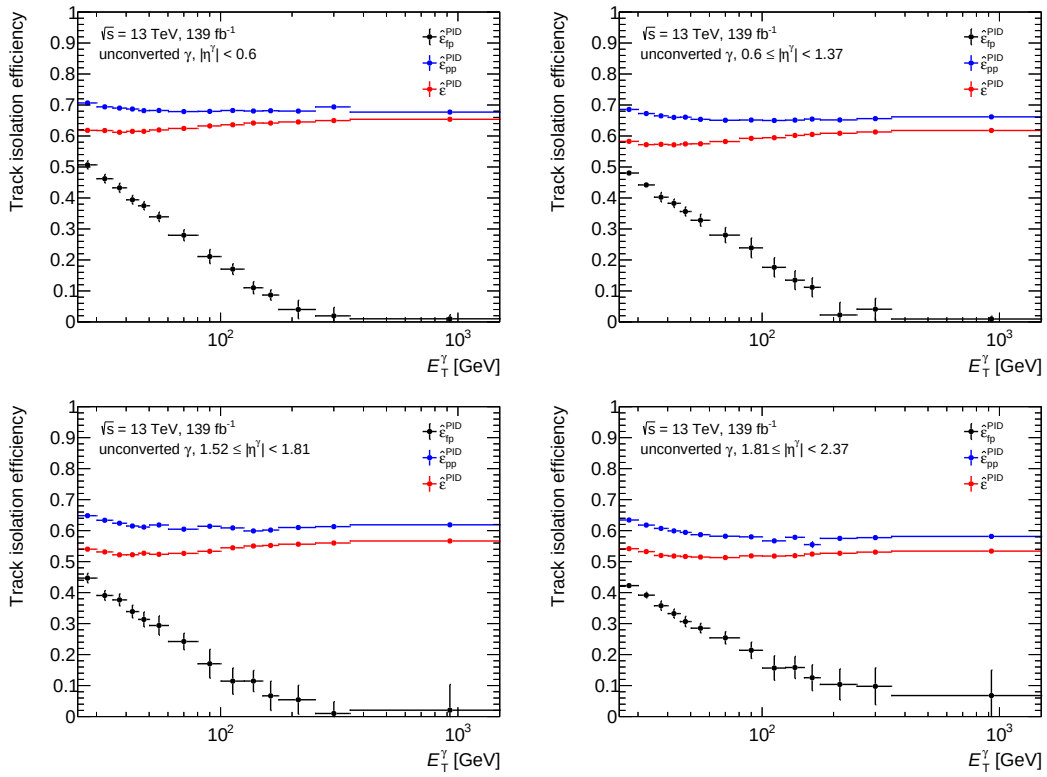


Figure 57: Track-isolation efficiencies for photon candidates in collected data, for prompt photons, and for fake photons determined by Equations (36) that the requirements of the tight PID WP as a function of E_{T}^{γ} for the different $|\eta^{\gamma}|$ bins. The efficiencies are shown for unconverted photon candidates and those for fake photons are determined by Equation (36).

A.3 Correction factor

Figures 58 and 59 show the correction factor c_{presel} computed by Equation (37) using events with prompt photons from the γ+jet MC sample as a function of E_{T}^{γ} for all $|\eta^{\gamma}|$ bins for converted and unconverted photon, respectively.

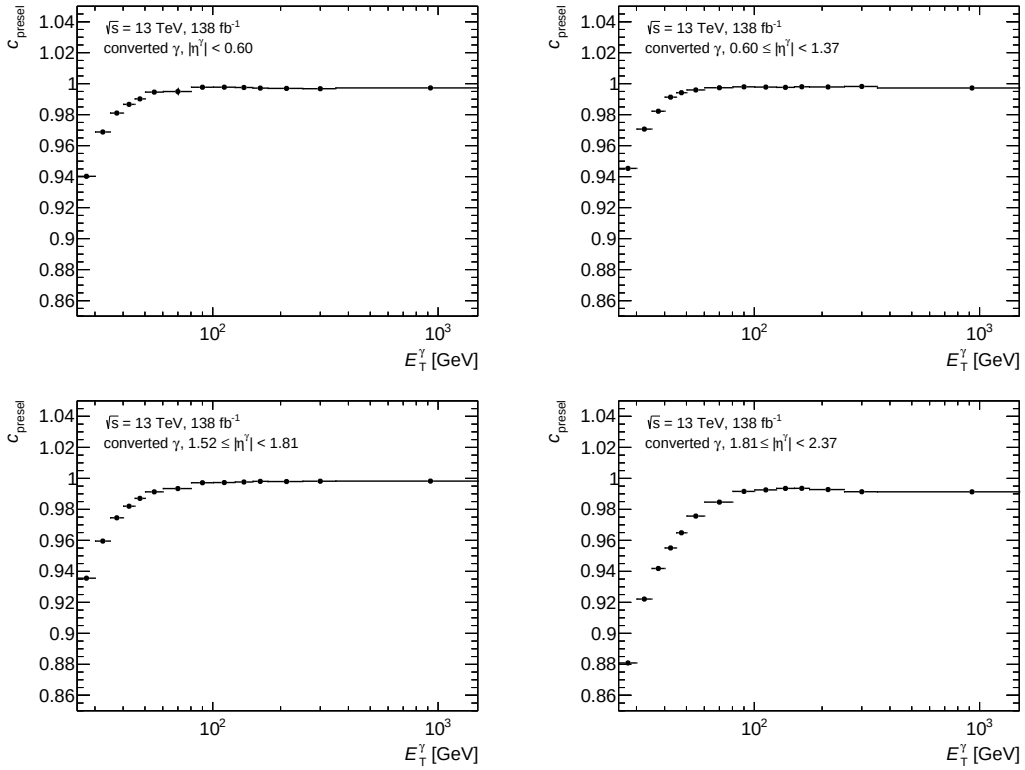


Figure 58: Correction factor c_{presel} as a function of E_{T}^{γ} for converted photons for the different $|\eta^{\gamma}|$ bins considered in the measurement. The statistical uncertainties due to the limited size of the γ+jet MC sample are displayed. However, these are tiny, so that they are covered by the size of the data point and not visible.

A.4 Additional information about the systematic uncertainties

In the following, additional information about the computation of individual types of systematic uncertainties considered in the $\varepsilon_{\text{pp}}^{\text{PID}}$ measurement are given.

A.4.1 Statistical uncertainties

Figure 60 shows example distributions of the number of PDs (N_{PD}) as a function of $\varepsilon_{\text{pp}}^{\text{PID}}$ obtained by the bootstrap sampling as described in Section 5.3.1. The standard deviation of these distributions are assigned as statistical uncertainties to the $\varepsilon_{\text{pp}}^{\text{PID}}$ values as described in Section 5.3.2.

A.4.2 Correction-factor uncertainty

The correction factor depends on the efficiency for prompt photons passing the loose PID WP and is computed using events with prompt photons from the γ+jet MC sample. Hence, a mismodeling of this efficiency in simulation affects the correction factor. An uncertainty is assigned to the correction factor that accounts for the impact of this mismodeling and is derived based on the SFs for the loose PID WP ($\text{SF}^{\text{loose-PID}}$) that are used to correct the efficiency for prompt photons of passing the loose PID WP in simulation. These SFs are computed with the method using radiative Z-boson decays (cf. Section 5.2)

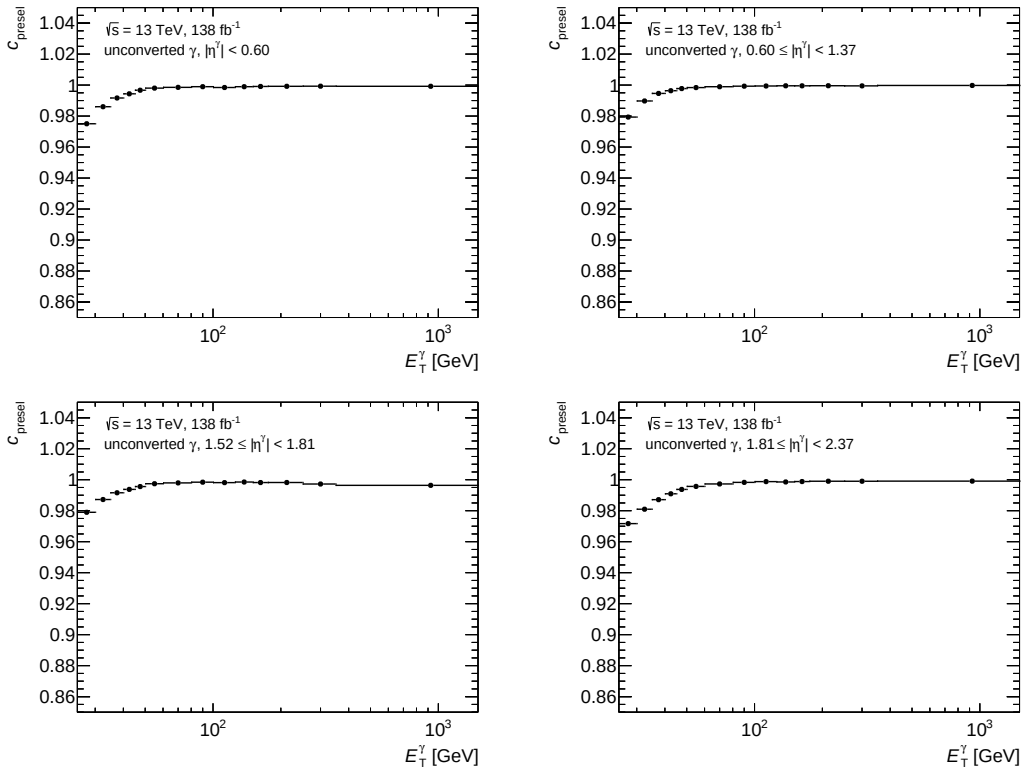


Figure 59: Correction factor c_{presel} as a function of E_{T}^{γ} for unconverted photons for the different $|\eta^{\gamma}|$ bins considered in the measurement. The statistical uncertainties due to the limited size of the γ+jet MC sample are displayed. However, these are tiny, so that they are covered by the size of the data point and not visible.

and are provided by the ATLAS collaboration. They are only available for $10\,\text{GeV} < E_{\text{T}}^{\gamma} < 100\,\text{GeV}$ and are therefore not available for the entire high-E_{T}^{γ} range ($E_{\text{T}}^{\gamma} > 100\,\text{GeV}$) used in the inclusive photon method. Furthermore, they are characterized by large statistical uncertainties. To perform an extrapolation of the SFs to the high-E_{T}^{γ} region that is robust against statistical fluctuations, a fit to $\text{SF}^{\text{loose-PID}}$ as a function of E_{T}^{γ} is performed. The fit function is required to meet the following criteria:

- It is expected that the differences in the efficiencies of the loose PID WP between simulated and collected data decrease as a function of E_{T}^{γ}, as the EM showers of photons become narrower for larger E_{T}^{γ} and the correction factor converges to 1 at high E_{T}^{γ} (cf. Section A.3). Therefore, the fit function is required to converge to 1 in the high-E_{T}^{γ} regime.
- The extrapolation of the SF values is performed over a large E_{T}^{γ} range. Therefore, the fit function is required to have a logarithmic dependence on E_{T}^{γ}, as such functions are characterized by slowly changing slopes.
- Due to the large statistical uncertainties of the $\text{SF}^{\text{loose-PID}}$ values, the fit model should be of low variance, i.e., the fit function should only have a few free parameters.

The chosen parameterization of the fit function that meets all the aforementioned is given by Equation (82), where p represents the only free parameter.

$$\text{SF}^{\text{loose-PID}}(E_{\text{T}}^{\gamma}) = p/\log(E_{\text{T}}^{\gamma}) + 1 \tag{82}$$

The fit is performed for each of the $|\eta^{\gamma}|$ bins and for converted and unconverted photons candidates, separately. For each E_{T}^{γ} bin, the uncertainty of the correction factor is given by the maximum deviation

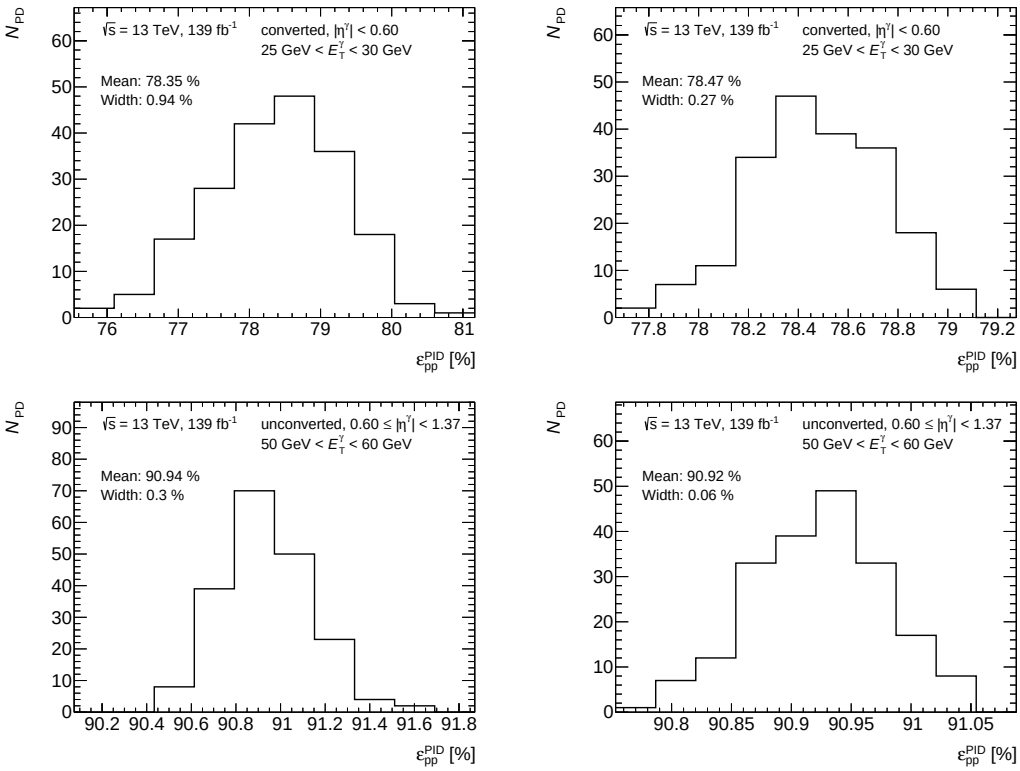


Figure 60: Distributions of the number of PDs (N_{PD}) as a function of $\varepsilon_{\text{pp}}^{\text{PID}}$ for converted photon candidates with $|\eta^{\gamma}| < 0.6$ and $25\,\text{GeV} < E_{\text{T}}^{\gamma} < 30\,\text{GeV}$ (upper row), and unconverted photon candidates with $0.6 \leq |\eta^{\gamma}| < 1.37$ and $50\,\text{GeV} < E_{\text{T}}^{\gamma} < 60\,\text{GeV}$ (lower row). The left column shows the distributions for the PDs obtained from the collected dataset and the right column shows the distributions for the PDs obtained from the γ+jet MC sample. The means and the standard deviations of the distributions are given in the legend.

($\Delta\kappa$) of all SF values at the central E_{T}^{γ} values of that bin within the 68% confidence level interval of the fitted function from 1. The absolute uncertainty of the correction factor is given by $\Delta c_{\text{presel}} = c_{\text{presel}} \cdot \Delta\kappa$. Figures 61 and 62 show the value of $\text{SF}^{\text{loose-PID}}$ and the corresponding fitted function of Equation (82) as a function of E_{T}^{γ} for all $|\eta^{\gamma}|$ bins for converted and unconverted photon candidates.

A.4.3 Non-closure uncertainty

The assumption that the track-isolation efficiencies $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ can be approximated by Equations (35) and (36), respectively, is challenged by performing a non-closure test, described in the following. The track-isolation efficiencies $\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{NTR, MC}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{PID, MC}}$, and $\hat{\varepsilon}_{\text{fp}}^{\text{RTR, MC}}$ are computed for fake photons from the dijet MC sample (cf. Section 5.3.2). The first superscript indicates which of the criteria of the tight PID WP are required to be fulfilled by the fake photons, i.e., either none, all, or only those of the RTR or NTR. If the approximations are valid, the value of $\hat{\varepsilon}_{\text{fp}}^{\text{NTR, MC}}$ approximates $\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$, and $\hat{\varepsilon}_{\text{fp}}^{\text{RTR, MC}}$ approximates $\hat{\varepsilon}_{\text{fp}}^{\text{ID, MC}}$. The relative differences $\Delta\hat{\varepsilon}_{\text{fp}}$ and $\Delta\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ are computed by Equations (83) and (84).

$$\Delta\hat{\varepsilon}_{\text{fp}} = \frac{|\hat{\varepsilon}_{\text{fp}}^{\text{MC}} - \hat{\varepsilon}_{\text{fp}}^{\text{NTR, MC}}|}{\hat{\varepsilon}_{\text{fp}}^{\text{MC}}} \tag{83}$$

$$\Delta\hat{\varepsilon}_{\text{fp}}^{\text{PID}} = \frac{|\hat{\varepsilon}_{\text{fp}}^{\text{PID, MC}} - \hat{\varepsilon}_{\text{fp}}^{\text{RTR, MC}}|}{\hat{\varepsilon}_{\text{fp}}^{\text{PID, MC}}} \tag{84}$$

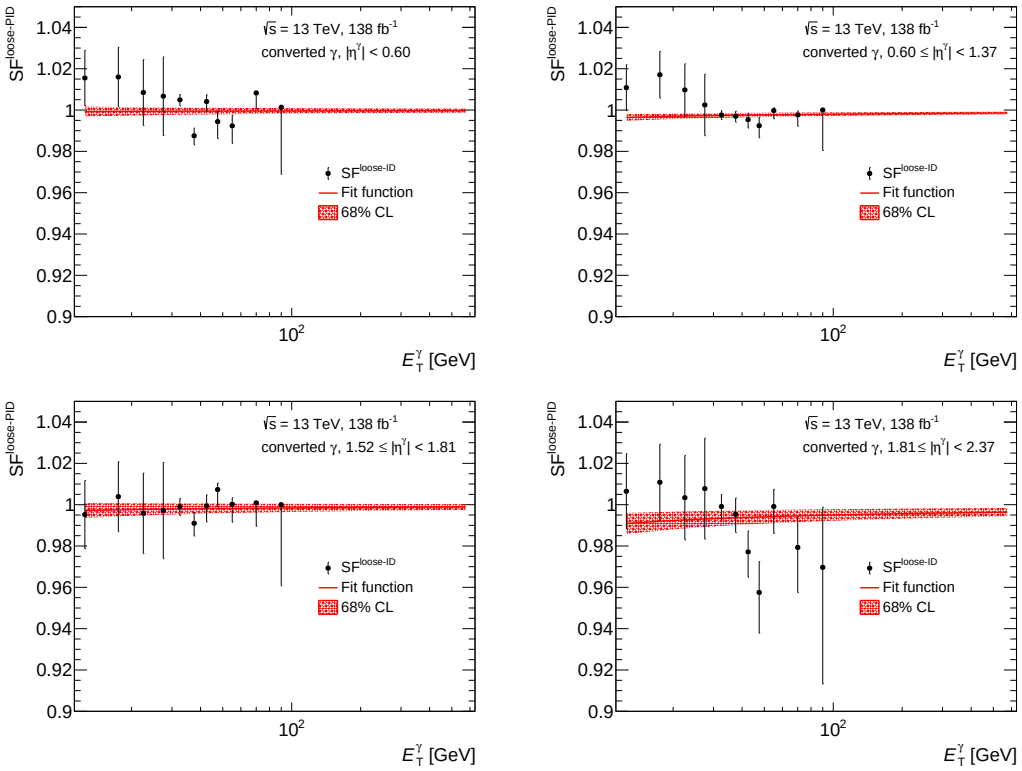


Figure 61: The SFs of the loose PID WP computed by the method using radiative Z-boson decays as a function of E_{T}^{γ} for converted photon candidates. The uncertainties of the SFs correspond to the statistical uncertainty. The fitted function (82) and the respective 68% confidence level interval shown by the hashed red band.

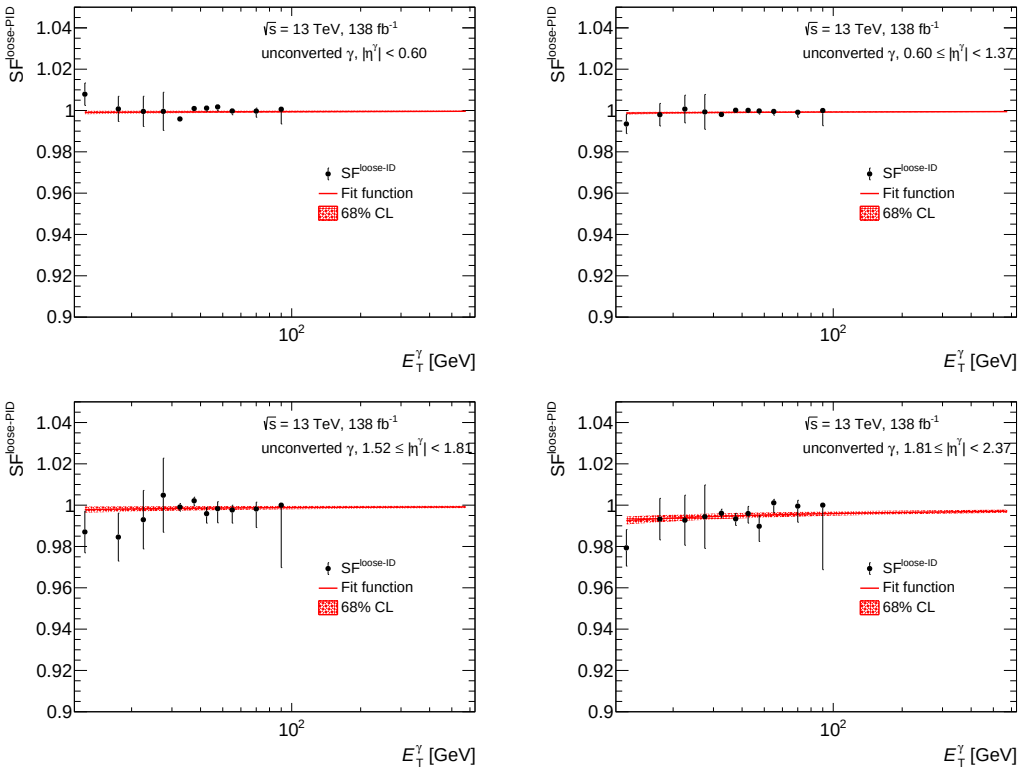


Figure 62: The SFs of the loose PID WP computed by the method using radiative Z-boson decays as a function of E_{T}^{γ} for unconverted photon candidates. The uncertainties of the SFs correspond to the statistical uncertainty. The fitted function (82) and the respective 68% confidence level interval shown by the hashed red band.

Table 21: Values of $\Delta\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$ and $\Delta\hat{\varepsilon}_{\text{fp}}^{\text{ID, MC}}$ for the different $|\eta^{\gamma}|$ bins considered in the measurement for converted and unconverted photons. The uncertainty corresponds to the uncertainty due to the limited size of the simulated fake-photon sample.

$\lvert\eta^{\gamma}\rvert$	unconverted		converted	
	$\Delta\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$ [%]	$\Delta\hat{\varepsilon}_{\text{fp}}^{\text{ID, MC}}$ [%]	$\Delta\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$ [%]	$\Delta\hat{\varepsilon}_{\text{fp}}^{\text{ID, MC}}$ [%]
$[0, 0.6)$	0.89 ± 1.48	2.54 ± 2.44	0.53 ± 2.58	0.04 ± 4.23
$[0.6, 1.37)$	0.42 ± 1.46	0.62 ± 2.38	1.71 ± 1.93	2.92 ± 2.83
$[1.52, 1.81)$	1.10 ± 3.14	3.18 ± 5.33	0.02 ± 2.54	1.54 ± 3.87
$[1.81, 2.37)$	0.05 ± 2.56	1.08 ± 3.73	0.86 ± 2.56	0.60 ± 3.40

Table (21) lists the values determined with Equations (83) and (84) for the different $|\eta^{\gamma}|$ bins inclusive in E_{T}^{γ}. All values are close to zero and compatible with zero within two times of their statistical uncertainty. This supports the validity of the approximations used in the inclusive photon method for computing the $\varepsilon_{\text{pp}}^{\text{PID}}$ values.

The nominal values of $\hat{\varepsilon}_{\text{fp}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ are independently varied up and down by $\Delta\varepsilon_{\text{fp}}$ and $\Delta\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, respectively. The values of $\varepsilon_{\text{pp}}^{\text{PID}}$ are recomputed for each of these four variations, and the corresponding differences to the nominal $\varepsilon_{\text{pp}}^{\text{PID}}$ values are quantified. The differences of the variations that enhance the $\varepsilon_{\text{pp}}^{\text{PID}}$ values, i.e., the negative variation of $\hat{\varepsilon}_{\text{fp}}$ and the positive variation of $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, are summed in quadrature, as are the other two. The larger of the two resulting differences is assigned as the non-closure uncertainty.

Figures 63, 64, 65, and 66 compare the values of $\hat{\varepsilon}_{\text{fp}}$ to the values of $\hat{\varepsilon}_{\text{fp}}^{\text{NTR, MC}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$ and the values of $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$ to the values of $\hat{\varepsilon}_{\text{fp}}^{\text{RTR, MC}}$ and $\hat{\varepsilon}_{\text{fp}}^{\text{PID, MC}}$ as a function of E_{T}^{γ} for the different $|\eta^{\gamma}|$ bins and for converted and unconverted photon candidates.

While the values of the track-isolation efficiencies computed in simulation agree within the large statistical uncertainties, which supports the validity of the approximations, the values determined by the inclusive photon method are smaller, especially in the low-E_{T}^{γ} region ($< 60\,\text{GeV}$) across all $|\eta^{\gamma}|$ bins and for both reconstruction types. In the high-E_{T}^{γ} regime the statistical uncertainties become very large, so that it is challenging to deduce conclusive insights. However, the same trend is still observed.

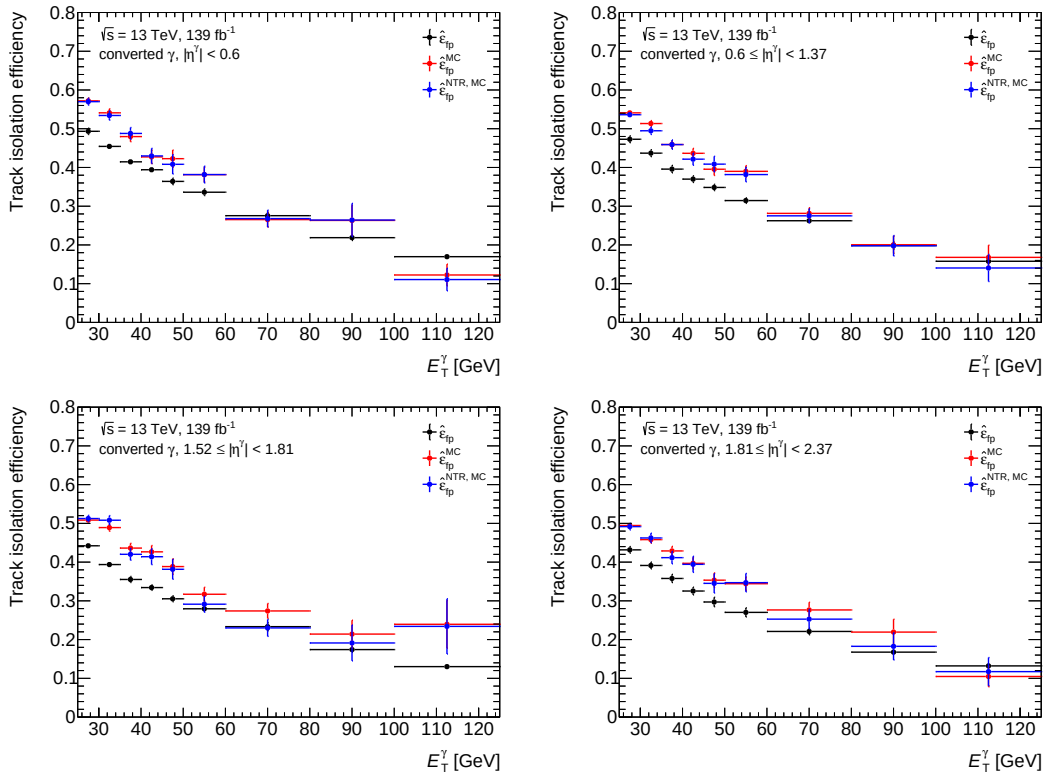


Figure 63: The track-isolation efficiencies $\hat{\varepsilon}_{\text{fp}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{NTR, MC}}$, and $\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$ as a function of E_{T}^{γ} for converted photon candidates for the different $|\eta^{\gamma}|$ bins considered in the measurement.

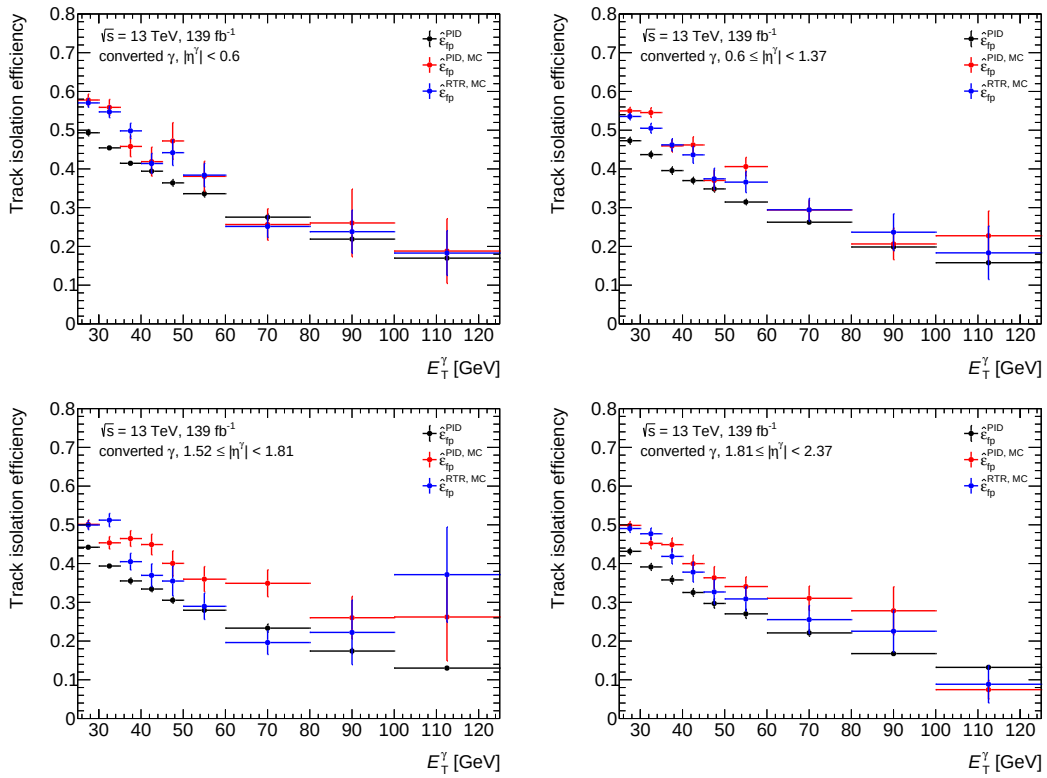


Figure 64: The-track isolation efficiencies $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{RTR, MC}}$, and $\hat{\varepsilon}_{\text{fp}}^{\text{PID, MC}}$ as a function of E_{T}^{γ} for converted photon candidates for the different $|\eta^{\gamma}|$ bins considered in the measurement.

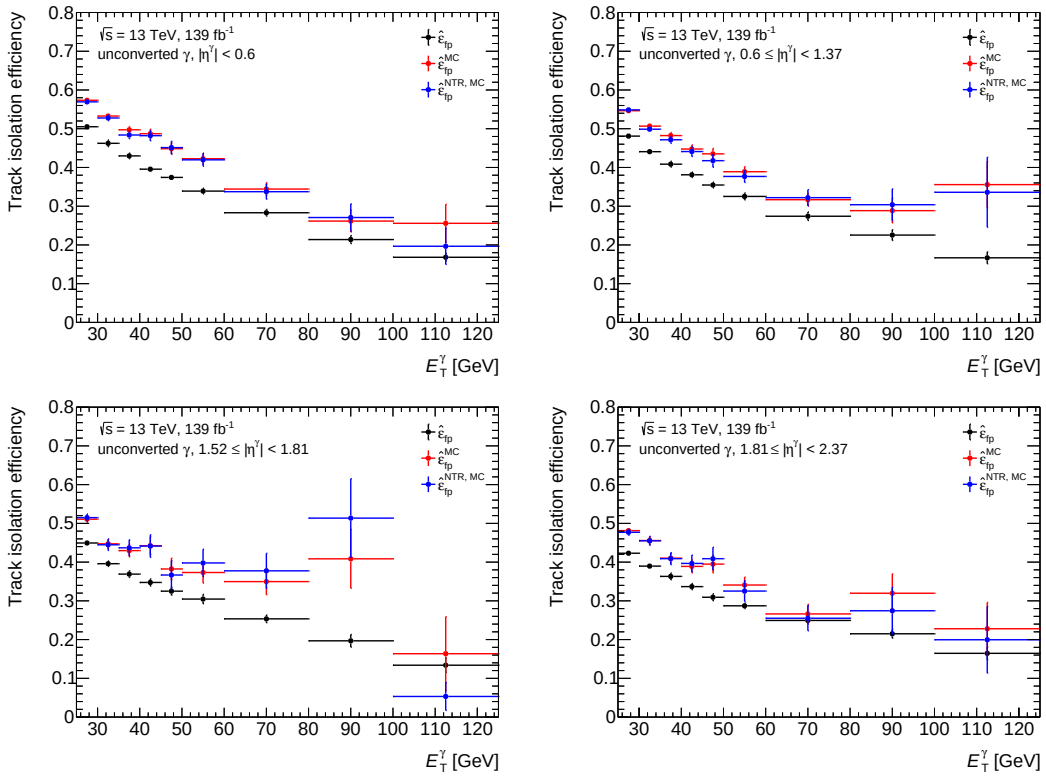


Figure 65: The track-isolation efficiencies $\hat{\varepsilon}_{\text{fp}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{NTR, MC}}$, and $\hat{\varepsilon}_{\text{fp}}^{\text{MC}}$ as a function of E_{T}^{γ} for unconverted photon candidates for the different $|\eta^{\gamma}|$ bins considered in the measurement.

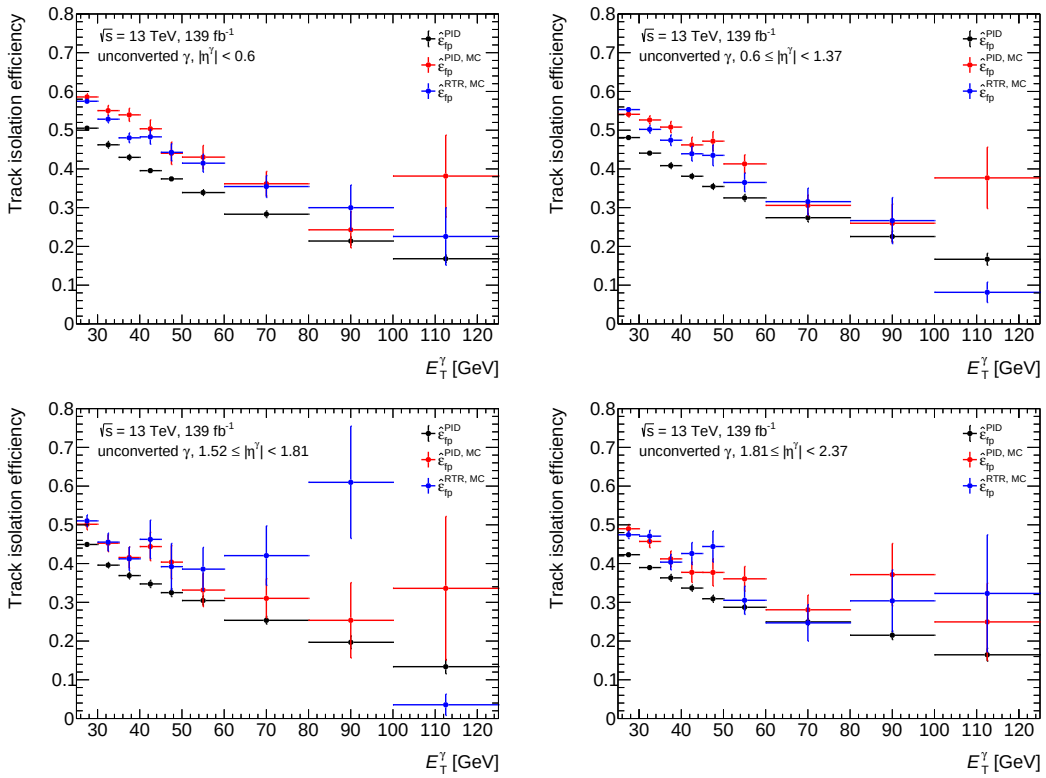


Figure 66: The track-isolation efficiencies $\hat{\varepsilon}_{\text{fp}}^{\text{PID}}$, $\hat{\varepsilon}_{\text{fp}}^{\text{RTR, MC}}$, and $\hat{\varepsilon}_{\text{fp}}^{\text{PID, MC}}$ as a function of E_{T}^{γ} for unconverted photon candidates for the different $|\eta^{\gamma}|$ bins considered in the measurement.

A.5 Results

Figures 67 and 68 show the value of the prompt-photon purities P_{A} and P_{P} as a function of E_{T}^{γ} for all $|\eta^{\gamma}|$ bins for converted photon and unconverted photon candidates, respectively.

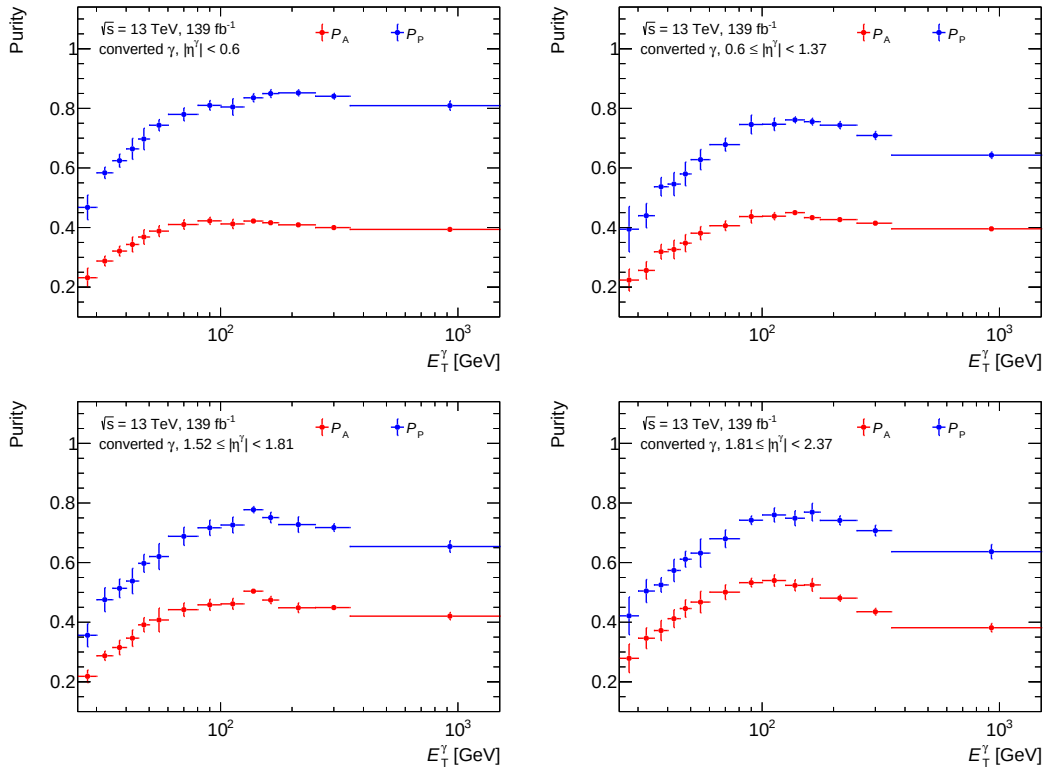


Figure 67: Prompt-photon purities P_{A} and P_{P} as a function of E_{T}^{γ} for converted photon candidates for the different $|\eta^{\gamma}|$ considered in the measurement. The uncertainty bars correspond to the total uncertainty.

Figures 69 and 70 show the measured $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ values and the $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ value predicted by the γ+jet MC sample as a function of E_{T}^{γ} for all $|\eta^{\gamma}|$ bins for converted and unconverted photon candidates, respectively. The bottom panels show the respective distributions of the $\mathrm{SF}^{\mathrm{PID}}$ values.

Figures 71 and 72 show the absolute values of the total uncertainty and of the individual uncertainty as a function of E_{T}^{γ} for all $|\eta^{\gamma}|$ bins for converted and unconverted photon candidates, respectively.

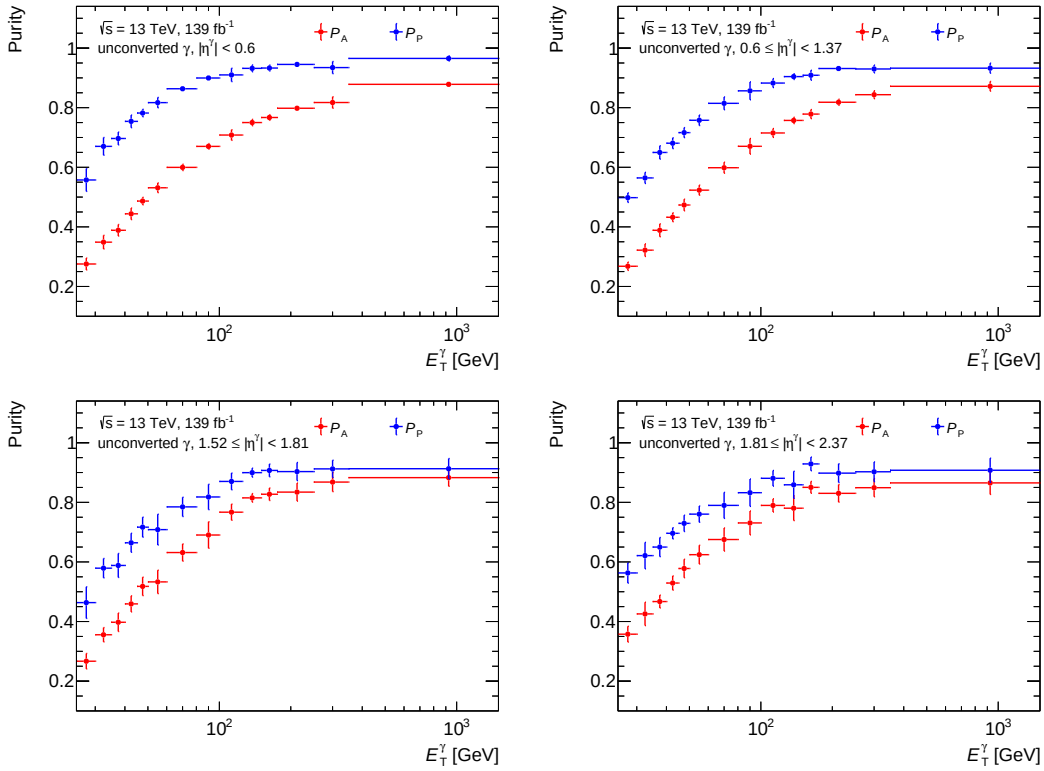


Figure 68: Prompt-photon purities P_{A} and P_{P} as a function of E_{T}^{γ} for unconverted photon candidates for the different $|\eta^{\gamma}|$ considered in the measurement. The uncertainty bars correspond to the total uncertainty.

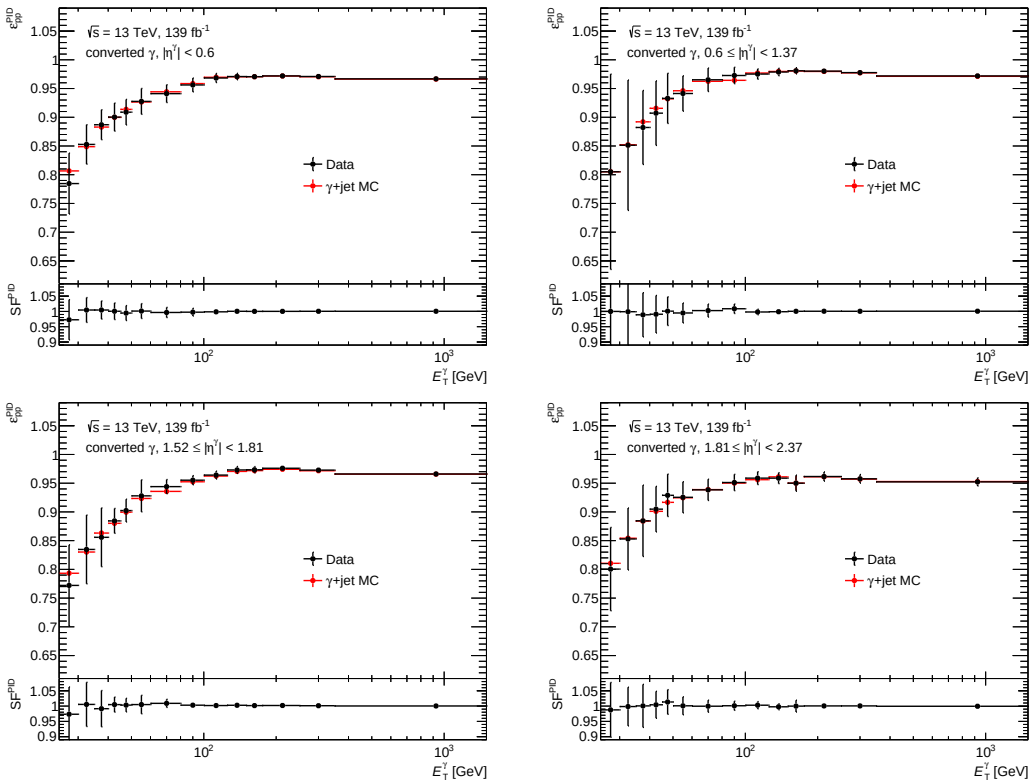


Figure 69: The values of $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ (upper panel) and $\mathrm{SF}^{\mathrm{PID}}$ (lower panel) as a function of E_{T}^{γ} for converted photons for the different $|\eta^{\gamma}|$ considered in the measurement. The uncertainty bars correspond to the total uncertainty.

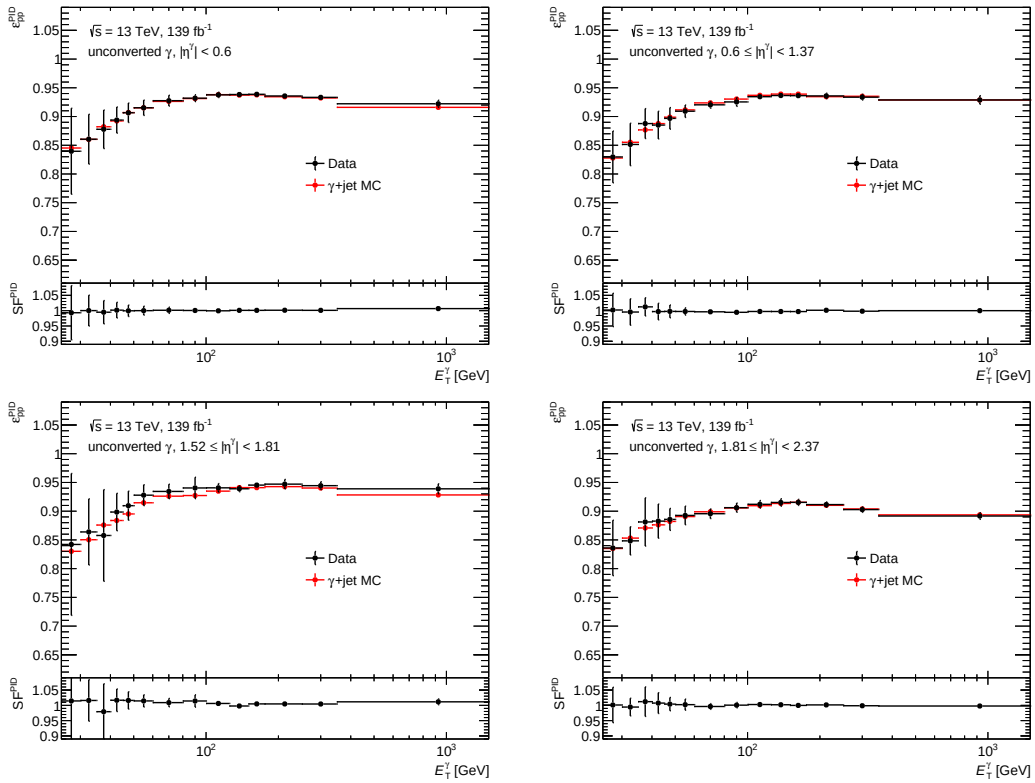


Figure 70: The values of $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ (upper panel) and $\mathrm{SF}^{\mathrm{PID}}$ (lower panel) as a function of E_{T}^{γ} for unconverted photon candidates for the different $|\eta^{\gamma}|$ considered in the measurement. The uncertainty bars correspond to the total uncertainty.

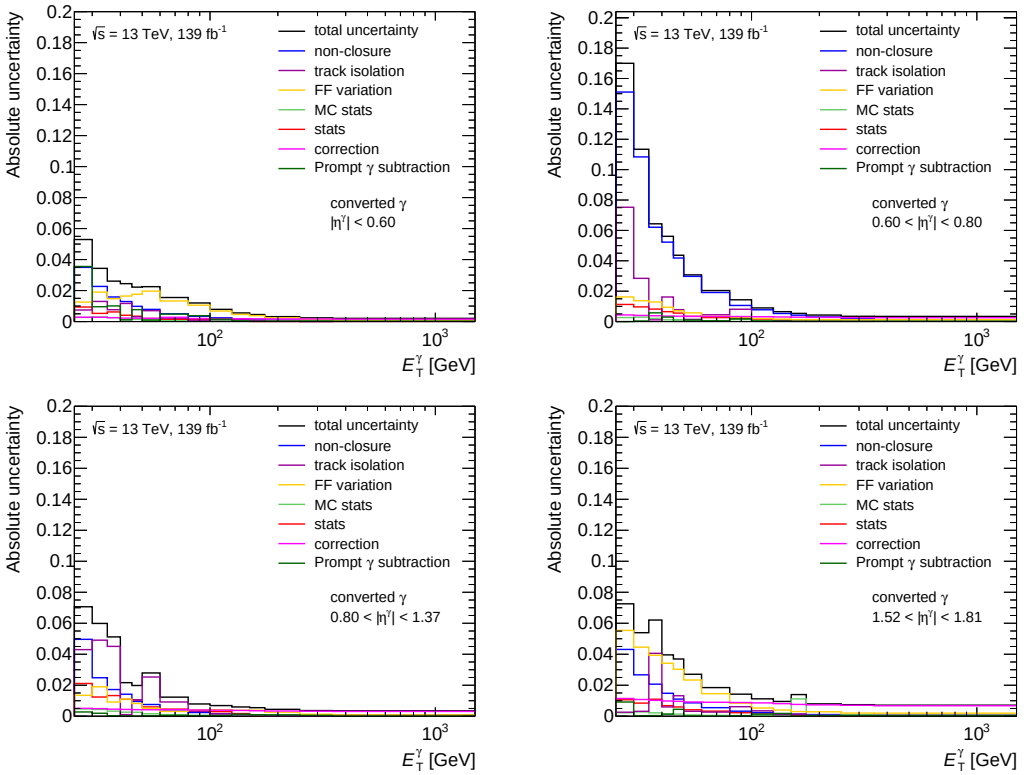


Figure 71: Absolute differences in the values of $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ as obtained from the individual systematic uncertainties as a function of E_{T}^{γ} for converted photon candidates for the different $|\eta^{\gamma}|$ bins. Furthermore, the total uncertainty obtained from the sum in quadrature of the individual uncertainties is shown.

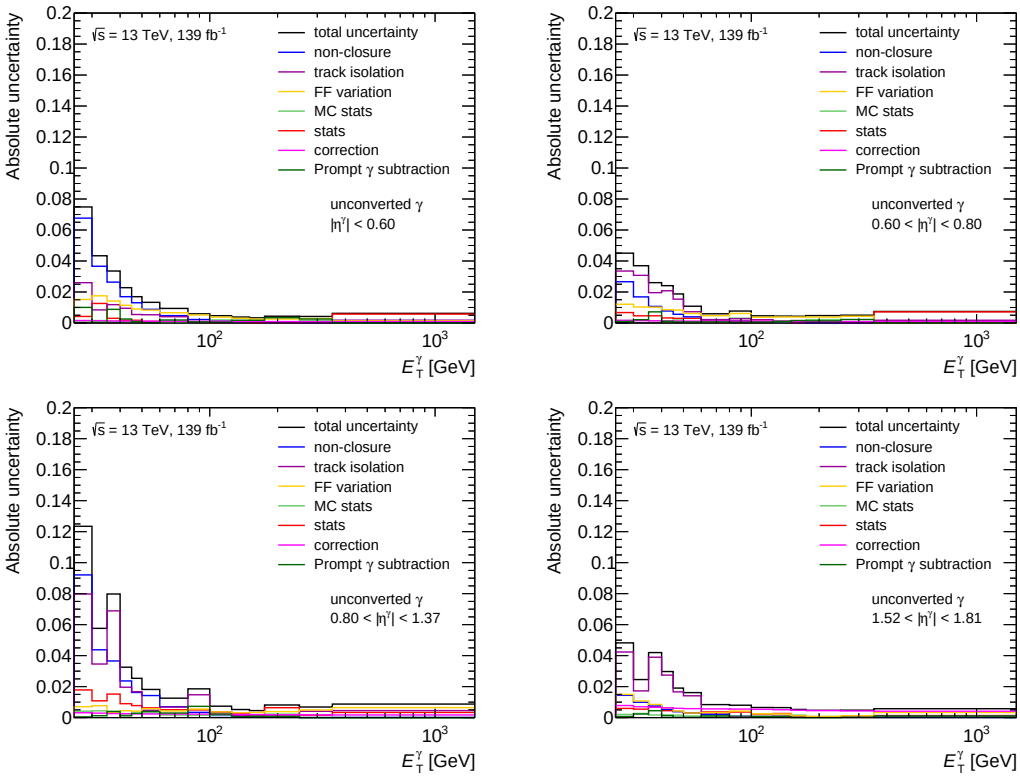


Figure 72: Absolute differences in the values of $\varepsilon_{\mathrm{pp}}^{\mathrm{PID}}$ as obtained from the individual systematic uncertainties as a function of E_{T}^{γ} for unconverted photon candidates for the different $|\eta^{\gamma}|$ bins. Furthermore, the total uncertainty obtained from the sum in quadrature of the individual uncertainties is shown.

B Simulation of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes

This appendix provides details about additional studies of the simulation of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ processes presented in Chapter 6.

B.1 Examination of interference effects

In the following, the overall size of interference effects between the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ sample are assessed. The sum of the kinematic distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and LO $tq\gamma$ samples are compared to those obtained from the LO inclusive sample. This tests for the impact of potential interference effects of the $t\,(\to \ell\nu b\gamma)\,q$ and $tq\gamma$ topologies that are simulated in the LO inclusive sample and neglected in the separate production of these samples. Figures 73 and 74 show the comparison of the predictions for the particle-level differential fiducial cross sections obtained from these samples as a function of photon-related kinematic and topological quantities, and for jet-related kinematic and topological quantities, respectively. The distributions agree well within the statistical uncertainty due to the limited sample sizes. This result indicates that the interference impacts are negligible, which justifies the treatment of $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ as separate processes in practice.

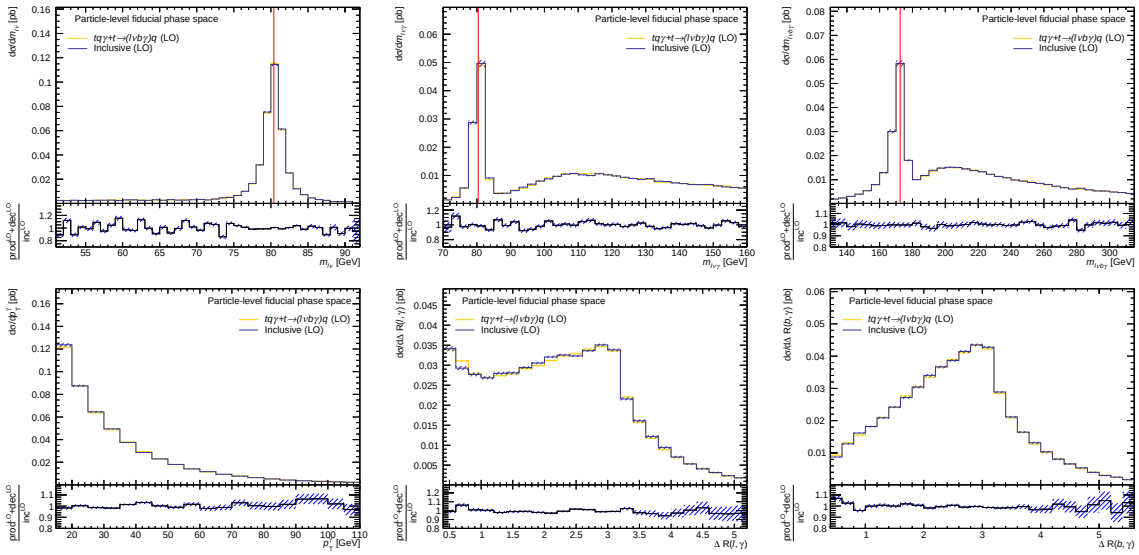


Figure 73: Particle-level differential fiducial cross sections as predicted by the LO $tq\gamma$, $t\,(\to \ell\nu b\gamma)\,q$, and inclusive samples as a function of the $m_{\ell\nu}$, $m_{\ell\nu\gamma}$, $m_{\ell\nu b\gamma}$, p_{T}^{γ}, $\Delta R(\ell,\,\gamma)$, and $\Delta R(b,\,\gamma)$. The bottom panels show the ratio of the sum of the distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ ($\mathrm{dec}^{\mathrm{LO}}$) and LO $tq\gamma$ ($\mathrm{prod}^{\mathrm{LO}}$) samples to the respective distribution from the LO inclusive sample ($\mathrm{inc}^{\mathrm{LO}}$). The hashed bands correspond to the statistical uncertainty due to the limited sample sizes.

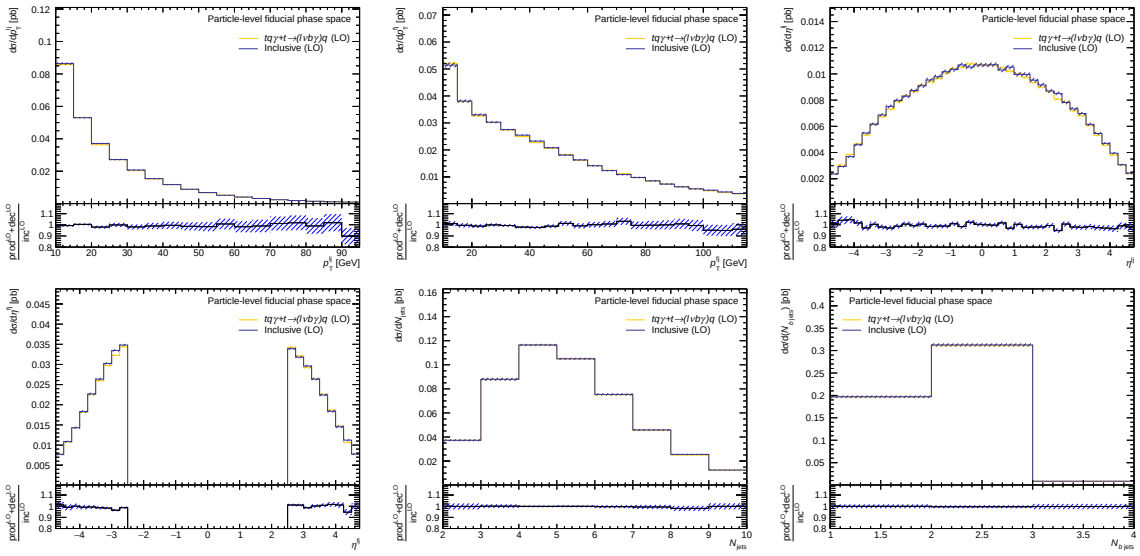


Figure 74: Particle-level differential fiducial cross sections as predicted by the LO $tq\gamma$, $t\,(\to \ell\nu b\gamma)\,q$ and inclusive samples as a function of p_{T}^{lj}, $p_{\mathrm{T}}^{\mathrm{fj}}$, η^{lj} and η^{fj}, of the number of jets (N_{jets}), and b jets ($N_{b\ \mathrm{jets}}$). The bottom panels show the ratio of the sum of the distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ ($\mathrm{dec}^{\mathrm{LO}}$) and LO $tq\gamma$ ($\mathrm{prod}^{\mathrm{LO}}$) samples to the respective distribution from the LO inclusive samples ($\mathrm{inc}^{\mathrm{LO}}$). The hashed bands correspond to the statistical uncertainty due to the limited sample sizes.

B.2 Reconstruction of the W-boson and top-quark masses at leading order

Figure 75 shows the particle-level differential fiducial cross section as a function of the difference between the reconstructed W-boson mass and the truth W-boson mass, and of the difference between the reconstructed top-quark mass and truth top-quark mass as obtained from the LO $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ samples. The truth W-boson and top-quark masses are the values as obtained from the parton-level event record and are specific to each simulated event. The distributions of the difference in the W-boson mass are centered around zero and the values of the difference are tiny for the majority of the events in both samples. Conclusively, the W boson is correctly reconstructed in almost all cases.

The distributions of the difference in the top-quark mass peak at 0 and the majority of the distribution lies close to 0. However, the distributions are characterized by large tails toward values < 0 and small tails toward values > 0. The tails are caused by the limitations of the matching of the b jet with the b quark, as discussed in Section 6.3. The tail toward lower values is less pronounced for $t\,(\to \ell\nu b\gamma)\,q$ events due to the energy of the b jet being lower on average for $t\,(\to \ell\nu b\gamma)\,q$ events than for $tq\gamma$ events.

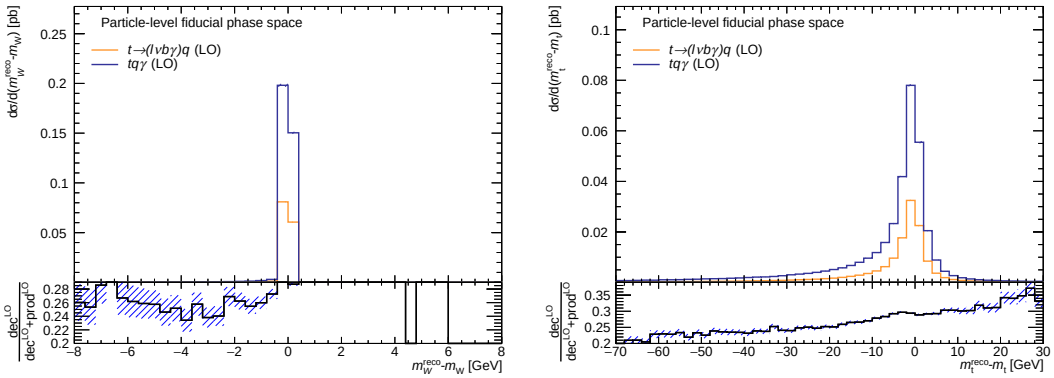


Figure 75: Particle-level differential fiducial cross sections as a function of the difference between the reconstructed W-boson mass and truth W-boson mass, and of the difference between the reconstructed top-quark mass and truth top-quark mass as obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The bottom panel shows the ratio of the distribution from the LO $t\,(\to \ell\nu b\gamma)\,q$ sample (dec^{LO}) and the sum of the distributions from the LO $tq\gamma$ (prod^{LO}) and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The hashed bands correspond to the uncertainty of the predictions.

B.3 Investigation of NLO QCD contributions for $t\,(\to \ell\nu b\gamma)\,q$

To study the impact of NLO QCD contributions on the differential particle-level cross sections as a function of photon-related quantities, the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is compared to a LO $t\,(\to \ell\nu b\gamma)\,q$ sample, which is obtained by the $t\,(\to \ell\nu b\gamma)\,q$ identification procedure (cf. Section 6.4) from a sample that simulates the hard-scattering t-channel single-top-quark production ($pp \to tqb$). As in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample, photon radiation is solely modeled via the PS algorithm. This sample is referred to as LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) sample, where "P8" indicates that the photon radiation is simulated by Pythia 8. The only difference between the NLO $t\,(\to \ell\nu b\gamma)\,q$ and LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) samples is the inclusion of NLO QCD contributions in the hard-scattering process in the former sample.

The particle-level differential cross sections as a function of $m_{\ell\nu}$, $m_{\ell\nu\gamma}$, $m_{\ell\nu b\gamma}$, p_{T}^{γ}, $\Delta R(\gamma,\,\ell)$, and $\Delta R(b,\,\gamma)$ obtained from these samples are compared for the "t and W" and "only t" categories in Figures 76 and 77 respectively. These figures also show the respective ratio of the differential cross sections. Across all distributions, the shapes agree within the uncertainty of the predictions. Hence, it is concluded that the NLO QCD contribution have only an insignificant influence on the shape of these photon-related quantities. The particle-level fiducial cross section predicted by the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample is 1.28 times larger than that predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) sample. This value is only slightly smaller than the k-factor for $t\,(\to \ell\nu b\gamma)\,q$ of 1.36 (cf. Section 6.4.2) that is computed based on the estimated $t\,(\to \ell\nu b\gamma)\,q$ purities in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample. Furthermore, the k-factor is consistent for the "t and W" and "only t" categories. These observations qualitatively support the estimated values of the purity of $t\,(\to \ell\nu b\gamma)\,q$ events in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample discussed in Section 6.4.2.

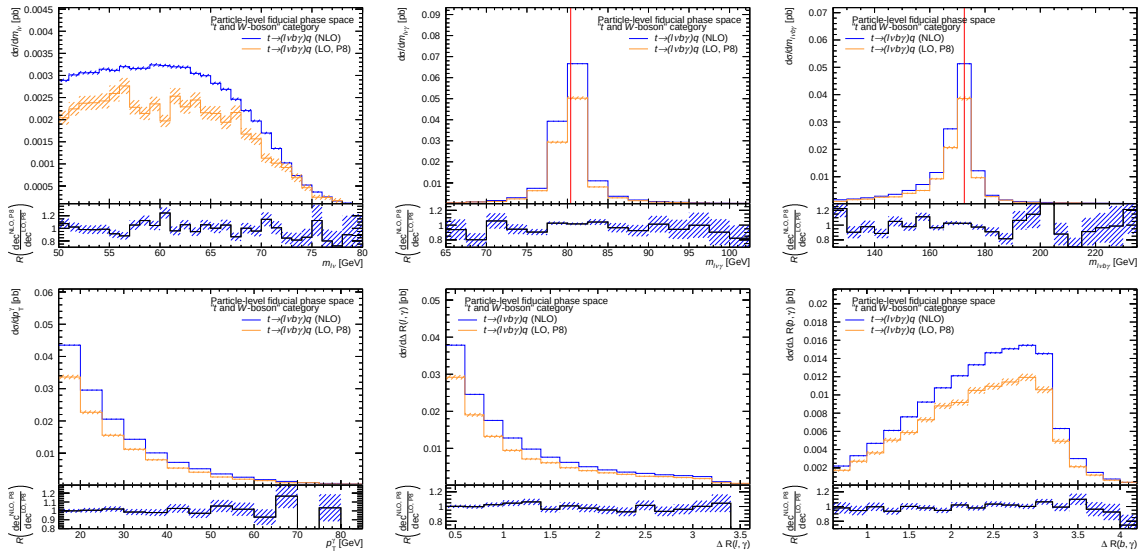


Figure 76: Particle-level differential fiducial cross sections as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of the $m_{\ell\nu}$, $m_{\ell\nu\gamma}$, $m_{\ell\nu b\gamma}$, p_{T}^{γ}, $\Delta R(\ell,\,\gamma)$, and $\Delta R(b,\,\gamma)$ for the "t and W" category. The bottom panels show the ratio of the two distributions denoted by "$\mathrm{dec}^{\mathrm{LO,\,P8}}$" and "$\mathrm{dec}^{\mathrm{NLO,\,P8}}$", respectively. The hashed bands correspond to the uncertainty of the predictions.

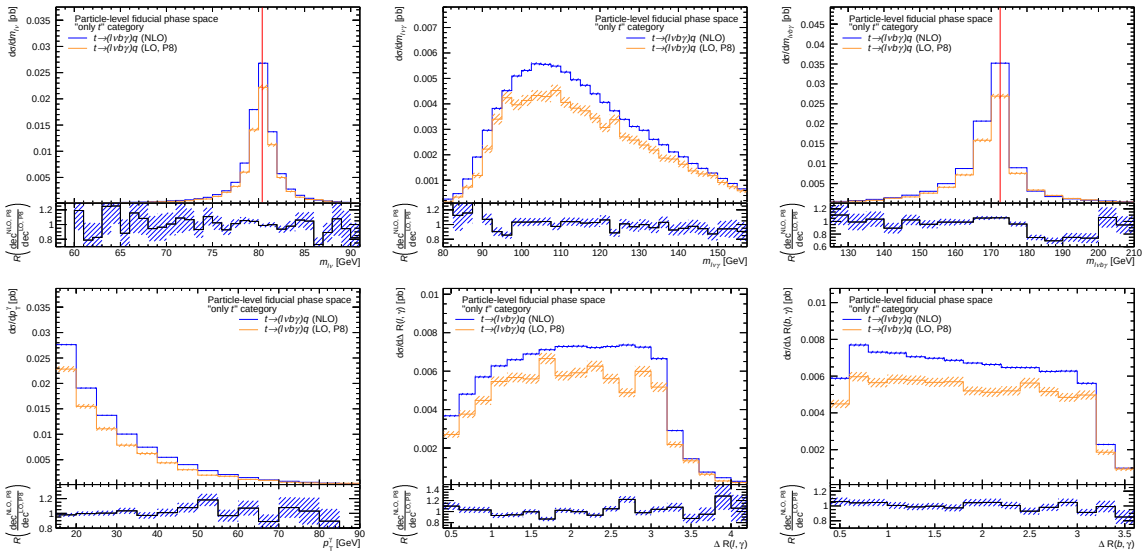


Figure 77: Particle-level differential fiducial cross sections as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples as a function of $m_{\ell\nu}$, $m_{\ell\nu\gamma}$, $m_{\ell\nu b\gamma}$, p_{T}^{γ}, $\Delta R(\ell,\,\gamma)$, and $\Delta R(b,\,\gamma)$ for the "only t" category. The bottom panels show the ratio of the two distributions denoted by "$\mathrm{dec}^{\mathrm{LO,\,P8}}$" and "$\mathrm{dec}^{\mathrm{NLO,\,P8}}$", respectively. The hashed bands correspond to the uncertainty of the predictions.

B.4 Impact of the photon-radiation modeling on the $t\,(\to \ell\nu b\gamma)\,q$ process

To study whether the difference in the modeling of the photon radiation (via PS algorithm versus in the simulation of the hard-scattering process) affects the shapes of distributions of jet-related quantities, the differential particle-level fiducial cross sections as a function of these quantities obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) sample are compared to those obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ and inclusive samples. Figure 78 shows these distributions for the "t and W" category. The shapes of the distributions and the overall rate obtained from the three samples agree well within the uncertainties. Figure 79 shows these distributions in the "only t" category. While the shapes of the distributions obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) sample and the LO $t\,(\to \ell\nu b\gamma)\,q$ sample show significant differences in the N_{jet} distribution, a good agreement is observed for all other distributions. As the distributions for N_{jet} obtained from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO inclusive samples agree well, it is concluded that this significant difference is caused by the $tq\gamma$ contamination in the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) sample.

In conclusion, the jet-related quantities are found to be unaffected by the difference in the modeling of the photon radiation and are suitable tools to probe the impact of NLO QCD contributions on the $t\,(\to \ell\nu b\gamma)\,q$ process in isolation.

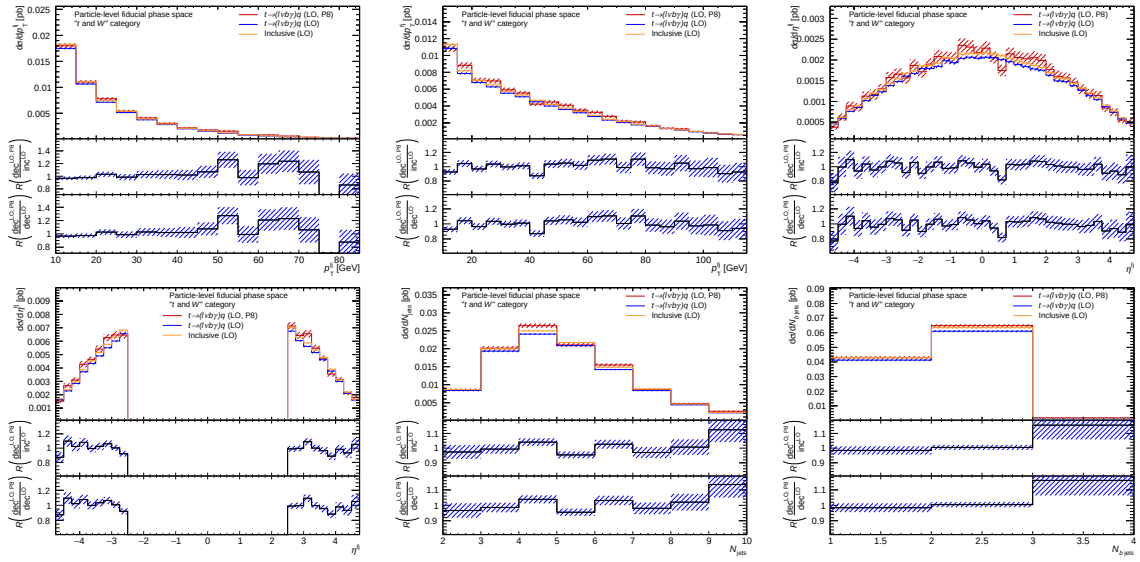


Figure 78: Particle-level differential fiducial cross sections as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO inclusive samples as a function of p_{T}^{lj}, p_{T}^{fj}, η^{lj} and η^{fj}, of the number of jets (N_{jets}) and b jets ($N_{b\text{ jets}}$) for the "t and W" category. The bottom panels show the ratio of the distributions from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO inclusive samples, and the ratio of the distributions from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The hashed bands correspond to the uncertainty of the predictions.

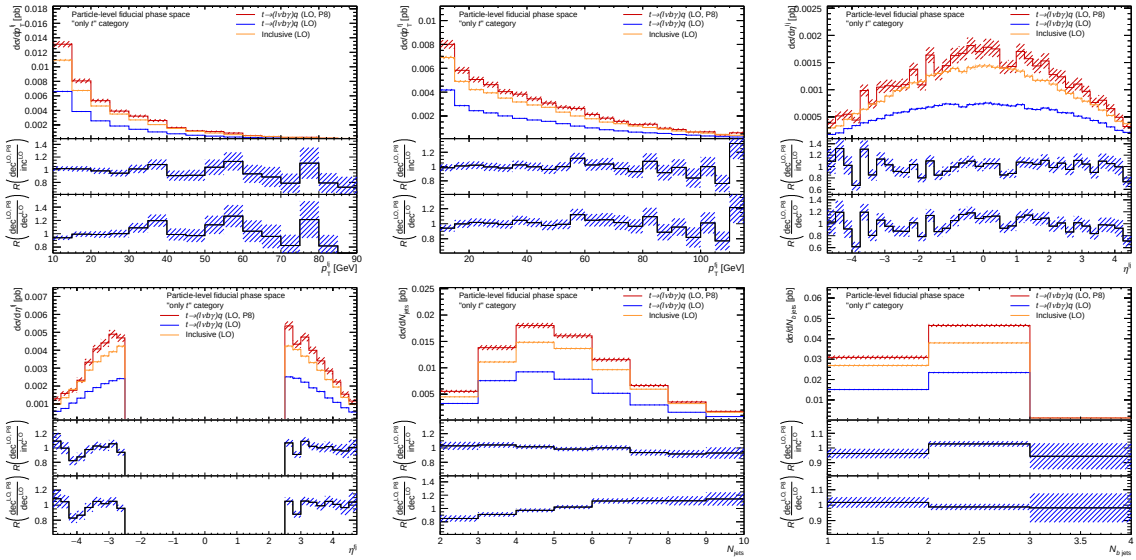


Figure 79: Particle-level differential fiducial cross sections as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO inclusive samples as a function of p_{T}^{lj}, $p_{\mathrm{T}}^{\mathrm{fj}}$, η^{lj} and η^{fj}, of the number of jets (N_{jets}), and b jets ($N_{b\ \mathrm{jets}}$) for the "only t" category. The bottom panels show the ratio of the distributions from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO inclusive samples, and the ratio of the distributions from the LO $t\,(\to \ell\nu b\gamma)\,q$ (P8) and LO $t\,(\to \ell\nu b\gamma)\,q$ samples. The hashed bands correspond to the uncertainty of the predictions.

B.5 Investigation of jet-related properties in the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample

Figure 80 and 81 show the particle-level differential fiducial cross section as predicted by the samples as a function of jet-related quantities for the "only t" category. The behavior of the differences between the differential cross sections from the different samples are qualitatively consistent with those described in Section 6.4.2.

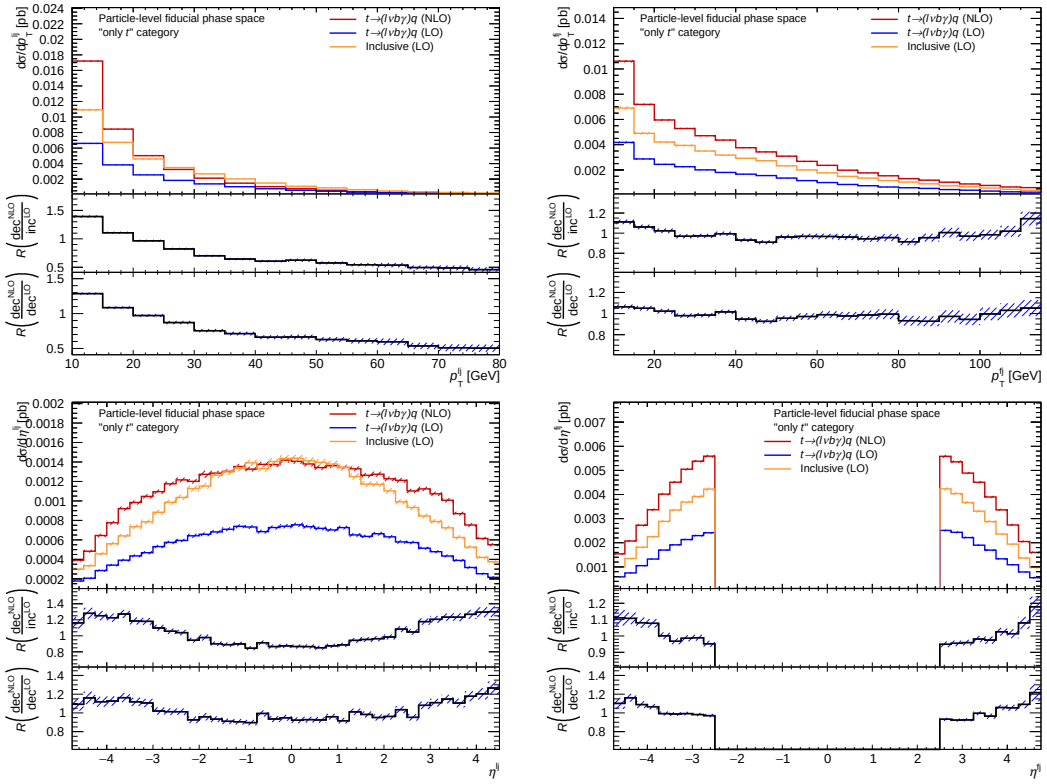


Figure 80: Particle-level differential fiducial cross section as predicted by the LO $t\,(\to \ell\nu b\gamma)\,q$, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the LO inclusive samples and the respective ratios as a function of p_{T}^{lj}, $p_{\mathrm{T}}^{\mathrm{fj}}$, η^{lj}, and η^{fj} for the "only t" category. The hashed bands correspond to the uncertainty of the predictions.

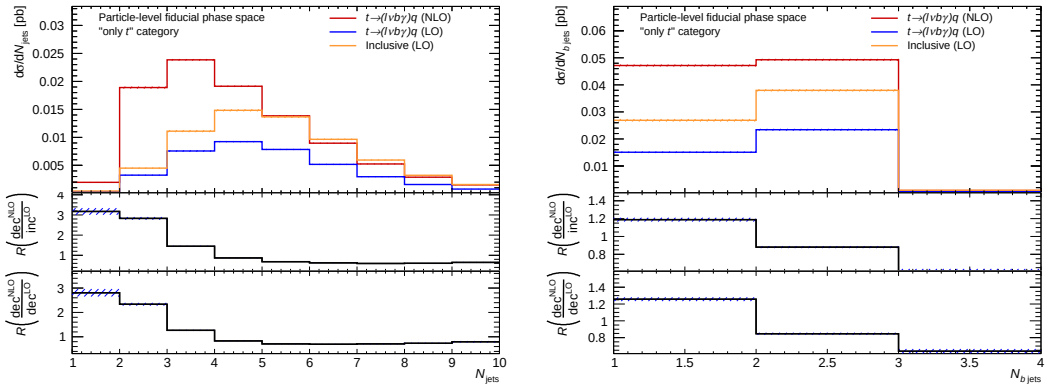


Figure 81: Particle-level differential fiducial cross section as predicted by the LO, NLO $t\,(\to \ell\nu b\gamma)\,q$, and the LO inclusive samples and the respective ratios as a function the number of jets (N_{jets}), and b jets ($N_{b\ \mathrm{jets}}$) for the "only t" category. The hashed bands correspond to the uncertainty of the predictions.

B.6 Acceptance fractions and fiducial cross sections for the particle-level phase space

Table 22 lists the values of the acceptance fractions (A) and the predicted fiducial cross sections ($\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$ and $\sigma^{\text{fid}}_{tq\gamma}$) obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples for different choices of the scale values (μ_{R} and μ_{F}). Furthermore, the relative differences between the values from the variations and the nominal values, and those corresponding to the standard deviation of the PDF variations are also listed. Additionally, the total $+1\sigma$ ("Combined up") and -1σ variations ("Combined down") obtained from the combination of the individual variations that increase and decrease the values, respectively, are given as well. Table 23 lists the corresponding values for the NLO $tq\gamma$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples.

Table 22: Values of the acceptance fraction (A) and the particle-level fiducial cross section (σ^{fid}) obtained from the LO $tq\gamma$ and LO $t\,(\to \ell\nu b\gamma)\,q$ samples for different scale choices as well as the relative uncertainty with respect to the nominal prediction. The relative uncertainties corresponding to the standard deviation of the PDF variations are also listed.

variation	$A_{t(\to\ell\nu b)q\gamma}$		$A_{t(\to\ell\nu b\gamma)q}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b\gamma)q}$	
	value	rel. uncertainty	value	rel. uncertainty	value/fb	rel. uncertainty	value/fb	rel. uncertainty
nominal	0.4075	-	0.3577	-	363	-	148	-
$\mu_{\text{F}} \times 2.0$	0.4061	−0.35%	0.3582	+0.16%	347	−4.36%	145	−1.85%
$\mu_{\text{F}} \times 0.5$	0.4089	+0.33%	0.3584	+0.20%	378	+4.40%	150	+1.70%
$\mu_{\text{R}} \times 2.0$	0.4074	−0.04%	0.3579	+0.07%	329	−9.29%	134	−9.38%
$\mu_{\text{R}} \times 0.5$	0.4070	−0.12%	0.3582	+0.16%	403	+11.29%	165	+11.90%
PDF	-	±0.07%	-	±0.03%	-	±0.87%	-	±0.88%
Combined up	0.4089	+0.33%	0.3588	+0.31%	407	+12.15%	165	+12.05%
Combined down	0.4060	−0.38%	0.3576	−0.03%	325	−10.30%	133	−9.60%

Table 23: Values for the acceptance fraction (A) and the predicted values for the particle-level fiducial cross section (σ^{fid}) obtained from the NLO $tq\gamma$ and NLO $t\,(\to \ell\nu b\gamma)\,q$ samples for the different scale choices as well as the relative uncertainty with respect to the nominal prediction. The relative uncertainties corresponding to the standard deviation of the PDF variations are also listed.

variation	$A_{t(\to\ell\nu b)q\gamma}$		$A_{t(\to\ell\nu b\gamma)q}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b\gamma)q}$	
	value	rel. uncertainty	value/10^{-2}	rel. uncertainty	value/fb	rel. uncertainty	value/fb	rel. uncertainty
nominal	0.4091	-	0.3375	-	466	-	238	-
$\mu_{\text{F}} \times 2.0$	0.4081	−0.23%	0.3378	+0.08%	463	−0.52%	239	+0.56%
$\mu_{\text{F}} \times 0.5$	0.4097	+0.15%	0.3379	+0.11%	470	+0.78%	238	+0.13%
$\mu_{\text{R}} \times 2.0$	0.4085	−0.14%	0.3378	+0.08%	439	−5.74%	226	−4.91%
$\mu_{\text{R}} \times 0.5$	0.4096	+0.12%	0.3379	+0.12%	494	+6.11%	249	+4.63%
PDF	-	±0.06%	-	±0.03%	-	±0.86%	-	±0.93%
Combined up	0.4099	+0.20%	0.3382	+0.20%	495	+6.22%	249	+4.76%
Combined down	0.4080	−0.28%	0.3374	−0.03%	439	−5.83%	226	−4.99%

C Details about systematic variations for the fiducial cross-section measurements

Table 24 lists the acceptance fraction (A) for $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ and the predicted particle-level fiducial cross sections ($\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$ and $\sigma^{\text{fid}}_{tq\gamma}$) for different variations. For both processes, values obtained from variations of the scales (μ_{R} and μ_{F}), from the Var3c variations, and from using Herwig 7 instead of Pythia 8 ("PS variation") in the production of the samples are listed. Additionally, the values obtained from using MadGraph5_aMC@NLO instead of Powheg Box v2 ("ME variation") in the production of the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample are given. Furthermore, the relative differences between the values from the variations and the nominal values, and those corresponding to the standard deviation of the PDF variations are also listed. Furthermore, the total $+1\sigma$ ("Combined up") and -1σ variations ("Combined down") obtained from the combination of the individual variations that increase and decrease the values, respectively, are given as well.

Table 24: Values of the acceptance fraction (A) and the particle-level fiducial cross section (σ^{fid}) for different scale choices as well as the relative uncertainty with respect to the nominal prediction. The relative uncertainties corresponding to the standard deviation of the PDF variations are also listed.

variation	$A_{t(\to\ell\nu b)q\gamma}$		$A_{t(\to\ell\nu b\gamma)q}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b\gamma)q}$	
	value	rel. uncertainty	value/10^{-3}	rel. uncertainty	value/fb	rel. uncertainty	value/fb	rel. uncertainty
nominal	0.1467	-	0.7115	-	167.1	-	50.1	-
$\mu_{\text{R}} \times 2.0$	0.1470	+0.16%	0.7119	+0.06%	157.9	−5.47%	47.6	−4.93%
$\mu_{\text{R}} \times 0.5$	0.1465	−0.18%	0.7115	−0.01%	176.7	+5.79%	52.4	+4.51%
$\mu_{\text{F}} \times 2.0$	0.1465	−0.17%	0.7108	−0.09%	166.3	−0.46%	50.3	+0.39%
$\mu_{\text{F}} \times 0.5$	0.1470	+0.18%	0.7120	+0.07%	168.4	+0.81%	50.2	+0.10%
PS variation	0.1501	+2.27%	0.8819	+23.96%	170.8	+2.27%	62.1	+23.96%
ME variation	-	-	1.0104	+42.02%	-	-	71.2	+42.02%
Var3c up	0.1462	−0.35%	0.7114	−0.01%	166.5	−0.35%	50.1	−0.01%
Var3c down	0.1467	+0.02%	0.7118	+0.05%	167.1	−0.02%	50.1	+0.05%
PDF variation	0.1467	±0.22%	0.7115	±0.21%	167.1	±0.92%	50.1	±1.04%
Combined up	0.1501	+2.29%	1.0556	+48.37%	177.6	+6.34%	74.5	+48.59%
Combined down	0.1460	−0.48%	0.7098	−0.23%	157.7	−5.57%	47.6	−5.04%

Tables 25 and 26 list the values of the efficiency ε_{sel}, the migration fraction f_{mig}, and the correction factor C (cf. Equations (45)) for all considered variations for the NLO $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ samples for the particle-level measurement, respectively.

Table 25: Values of the efficiency ε_{sel}, the migration fraction f_{mig}, and the correction factor C for different variations obtained from the NLO $tq\gamma$ sample for the particle-level measurement. The relative uncertainties with respect to the nominal prediction are also listed.

variation	ε_{sel}		f_{mig}		C	
	value	rel. uncertainty	value	rel. uncertainty	value	rel. uncertainty
nominal	0.1942	-	0.0902	-	0.2135	-
$\mu_{\text{R}} \times 2.0$	0.1934	−0.45%	0.0907	+0.59%	0.2127	−0.39%
$\mu_{\text{R}} \times 0.5$	0.1954	+0.60%	0.0895	−0.80%	0.2146	+0.53%
$\mu_{\text{F}} \times 2.0$	0.1940	−0.13%	0.0901	−0.12%	0.2132	−0.13%
$\mu_{\text{F}} \times 0.5$	0.1944	−0.10%	0.0904	+0.17%	0.2137	+0.11%
PDF variation	0.1942	±0.12%	0.0902	±0.45%	0.2135	±0.13%
PS variation	0.1949	+2.44%	0.0889	−5.72%	0.2139	+1.83%
Var3c up	0.1875	−0.79%	0.0953	+0.95%	0.2086	−0.69%
Var3c down	0.1858	−2.37%	0.0977	+3.56%	0.2059	−2.00%
Combined up	0.1990	+2.47%	0.0937	+3.93%	0.2175	+1.89%
Combined down	0.1894	−2.49%	0.0847	−6.06%	0.2090	−2.13%

Table 27 lists the acceptance fraction and the predicted parton-level fiducial cross sections ($\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$) obtained from the NLO $tq\gamma$ sample for different variations. The values from variations of the scales (μ_{R} and μ_{F}) in the production of the sample are listed. Furthermore, the relative differences between the values from the variations and the nominal values, and those corresponding to the standard deviation of the PDF variations are also listed. Additionally, the total $+1\sigma$ ("Combined up") and -1σ variations ("Combined down") obtained from the combination of the individual variations that increase and

Table 26: Values of the efficiency ε_{sel}, the migration fraction f_{mig}, and the correction factor C for different variations obtained from the NLO $t\,(\to \ell\nu b\gamma)\,q$ sample for the particle-level measurement. The relative uncertainty with respect to the nominal prediction are also listed.

variation	ε_{sel}		f_{mig}		C	
	value	rel. uncertainty	value	rel. uncertainty	value	rel. uncertainty
nominal	0.14658	-	0.1866	-	0.1802	-
$\mu_{\text{R}} \times 2.0$	0.14662	+0.03%	0.1866	+0.02%	0.1803	+0.04%
$\mu_{\text{R}} \times 0.5$	0.14644	−0.09%	0.1869	+0.15%	0.1801	−0.06%
$\mu_{\text{F}} \times 2.0$	0.14649	−0.06%	0.1863	−0.13%	0.1800	−0.09%
$\mu_{\text{F}} \times 0.5$	0.14672	+0.10%	0.1867	−0.07%	0.1804	+0.12%
PS variation	0.13539	−7.63%	0.1498	−19.72%	0.1593	−11.62%
ME variation	0.12768	−12.89%	0.2248	+20.49%	0.1647	−8.60%
Var3c up	0.14690	+0.22%	0.1861	−0.26%	0.1805	+0.16%
Var3c down	0.14625	−0.22%	0.1872	+0.31%	0.1799	−0.15%
PDF variation	0.14658	±0.10%	0.1866	±0.15%	0.1802	±0.08%
Combined up	0.14697	+0.27%	0.2248	+20.49%	0.1806	+0.22%
Combined down	0.12243	−16.47%	0.1484	−20.44%	0.1523	−15.49%

decrease the values, respectively, are given as well.

Table 27: Values of the acceptance fraction $A_{t(\to\ell\nu b)q\gamma}$ and the parton-level fiducial cross section ($\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$) for different scale choices as well as the relative uncertainty with respect to the nominal prediction. The relative uncertainties corresponding to the standard deviation of the PDF variations are also listed.

variation	$A_{t(\to\ell\nu b)q\gamma}$		$\sigma^{\text{fid}}_{t(\to\ell\nu b)q\gamma}$	
	value	rel. uncertainty	value/fb	rel. uncertainty
nominal	0.4204	-	478.62	-
$\mu_{\text{R}} \times 2.0$	0.4195	−0.21%	506.19	+5.76%
$\mu_{\text{R}} \times 0.5$	0.4197	−0.15%	476.57	−0.43%
$\mu_{\text{F}} \times 2.0$	0.4204	+0.01%	448.37	−6.32%
$\mu_{\text{F}} \times 0.5$	0.4219	+0.37%	458.2	−4.27%
PDF variation	0.4204	±0.27%	478.62	±0.87%
Combined up	0.4223	+0.45%	506.50	+5.83%
Combined down	0.4188	−0.38%	441.83	−7.69%

Table 28 lists the values of the efficiency ε_{sel}, the migration fraction f_{mig}, and the correction factor C (cf. Equations (45)) for all variations obtained from the NLO $tq\gamma$ sample for the parton-level measurement. The values obtained from the Var3c variations, and from using Herwig 7 instead of Pythia 8 ("PS variation") in the production of the $tq\gamma$ sample are also listed.

Table 28: Values of the efficiency ε_{sel}, the migration fraction f_{mig}, and the correction factor C for different variations for $tq\gamma$ for the parton-level measurement. The relative uncertainties with respect to the nominal prediction are also listed.

variation	ε_{sel}		f_{mig}		C	
	value	rel. uncertainty	value	rel. uncertainty	value	rel. uncertainty
nominal	0.07137	-	0.04231	-	0.07452	-
$\mu_{\text{R}} \times 2.0$	0.07138	+0.01%	0.04199	−0.76%	0.07451	−0.02%
$\mu_{\text{R}} \times 0.5$	0.07169	+0.45%	0.04272	+0.96%	0.07489	+0.49%
$\mu_{\text{F}} \times 2.0$	0.07115	−0.31%	0.04221	−0.24%	0.07429	−0.32%
$\mu_{\text{F}} \times 0.5$	0.07131	−0.08%	0.04240	+0.22%	0.07447	−0.08%
PDF variation	0.07137	±0.21%	0.04231	±0.33%	0.07452	±0.21%
PS variation	0.07252	+1.61%	0.05078	+20.01%	0.07640	+2.52%
Var3c up	0.06960	−2.48%	0.04521	+6.86%	0.07290	−2.18%
Var3c down	0.06892	−3.43%	0.04537	+7.23%	0.07220	−3.12%
Combined up	0.0726	+1.69%	0.05178	+22.39%	0.07644	+2.58%
Combined down	0.0683	−4.25%	0.04195	−0.86%	0.07168	−3.82%

D Data-driven background estimations

D.1 Estimation of background contributions with fake leptons

Motivation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement in the 1b70 and 0b70,1b85 regions

In the following, the modification of the requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ from $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$, which is used in the ARs, to $30\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$, which is used in the 1b70 and 0b70,1b85 regions, is discussed. For this, both regions are analyzed in the following without applying any $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement.

In the 1b70 region, $t\bar{t}$ and W+jets production contribute about 40% of events with prompt leptons each, and single-top-quark production (t-channel, s-channel, and tW-channel production) and Z+jets production contribute about 10% each. In the 0b70,1b85 region, $t\bar{t}$ contributes about 6% and single-top-quark production about 2% of all events with prompt leptons. W+jets production contributes 70% of all events with prompt electrons and 79% of those with prompt muons, while Z+jets production contributes 21% of the events with prompt electrons and 11% of those with prompt muons. The relative contribution from diboson production is small (0.5%–1%) in both regions across both lepton types.

Figure 82 shows the unconstrained $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in collected data and for the sum of the predicted prompt-lepton contributions for both regions for events with exactly one loose lepton candidate and exactly one tight lepton candidate, along with the ratios of the observed ("Data") and the predicted distributions ("Pred."). The $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are expected to be significantly different for events with fake leptons and prompt leptons. The prompt leptons originating from a leptonic decay of a W boson, which represent the majority of the prompt-lepton contribution, are accompanied by a genuine neutrino. This typically causes larger $E_{\mathrm{T}}^{\mathrm{miss}}$ values compared to other $E_{\mathrm{T}}^{\mathrm{miss}}$ sources (e.g., non-selected objects due to the limited reconstruction efficiencies). Hence, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of prompt-lepton events is expected to be harder, i.e., characterized by larger values on average, than the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of events with fake leptons. This is observed for both the 1b70 and the 0b70,1b85 regions. The observed distributions exceed the predicted prompt-lepton distributions by factors between 2 and 5 for loose leptons and by factors between 1.8 and 2.8 for tight leptons for $E_{\mathrm{T}}^{\mathrm{miss}} < 10\,\mathrm{GeV}$.[38] The disagreement decreases as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$, which indicates that the relative contribution of prompt-lepton events rises in the collected data, saturating at factors close to 100% at about 70 GeV for tight leptons and loose electrons. The observed distribution also exceeds the predicted distribution at high $E_{\mathrm{T}}^{\mathrm{miss}}$ in the case of loose muons, indicating that the contribution of fake loose muons is significant in this region. Furthermore, the disagreement between the collected data and the prediction for loose leptons is larger for the 1b70 than for the 0b70,1b85 region, while the disagreement is similar for tight leptons. This observation indicates that the fake-lepton efficiency is larger for the 0b70,1b85 region than for the 1b70 region.

Following the aforementioned observations, the $E_{\mathrm{T}}^{\mathrm{miss}}$ value of events for the 1b70 and 0b70,1b85 regions is restricted to values below 50 GeV to minimize the prompt-lepton contribution. The ARs definitions require events to satisfy $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$. To ensure that the compositions of fake leptons in the 1b70 and 0b70,1b85 regions resemble those in the ARs, this requirement is also used in the 1b70 and 0b70,1b85 regions. Therefore, events in these regions must fulfill $30\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$.

Additional kinematic distributions and prompt-lepton efficiencies in the 0b70,1b85 region

Figure 83 shows the p_{T}^{ℓ} and $m_{\mathrm{T}}(\ell,\,\nu)$ distributions in the 0b70,1b85 region for loose and tight leptons.

Figure 84 shows the prompt-lepton efficiencies in the 0b70,1b85 region.

Additional kinematic distributions in the SR and $W\gamma$ CR

[38] The uncertainty of the prediction is given only by the statistical uncertainty due to the limited sizes of the simulated samples.

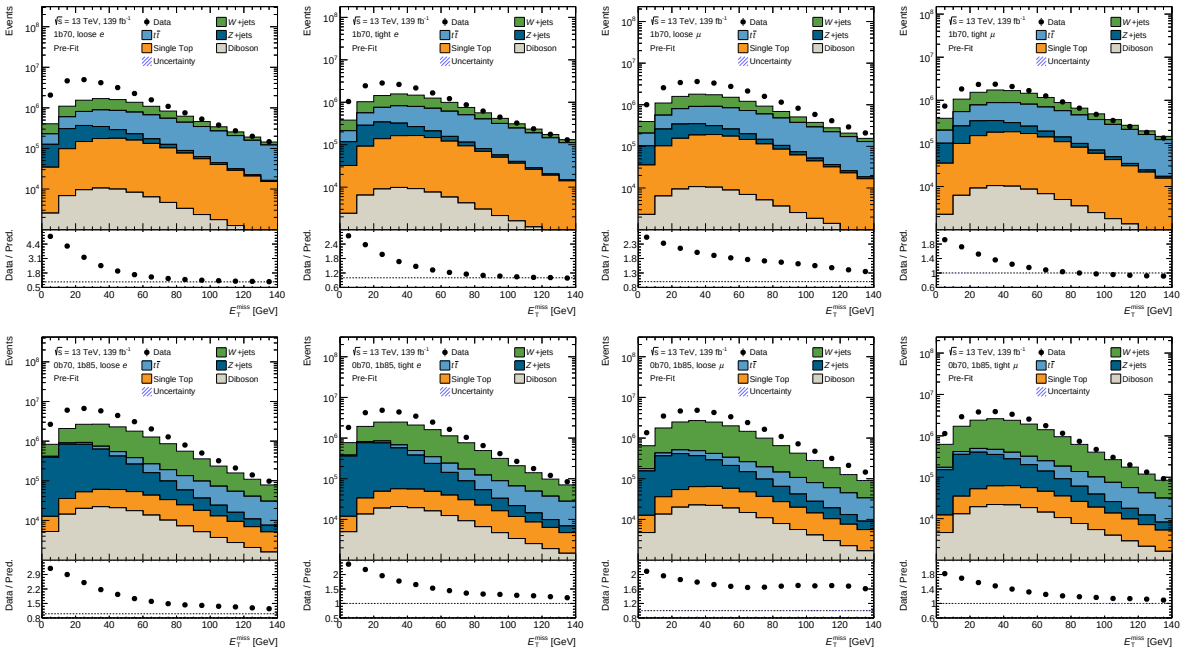


Figure 82: Distribution of events with loose and tight leptons in the 1b70 (upper row) and 0b70,1b85 (lower row) regions as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$. The first two distributions in each row correspond to events with an electron candidate, while the last two correspond to events with a muon candidate. The bottom panels illustrate the ratio of the observed ("Data") and predicted ("Pred.") distributions. The uncertainty bands correspond to the uncertainty due to the limited number of events in the MC samples.

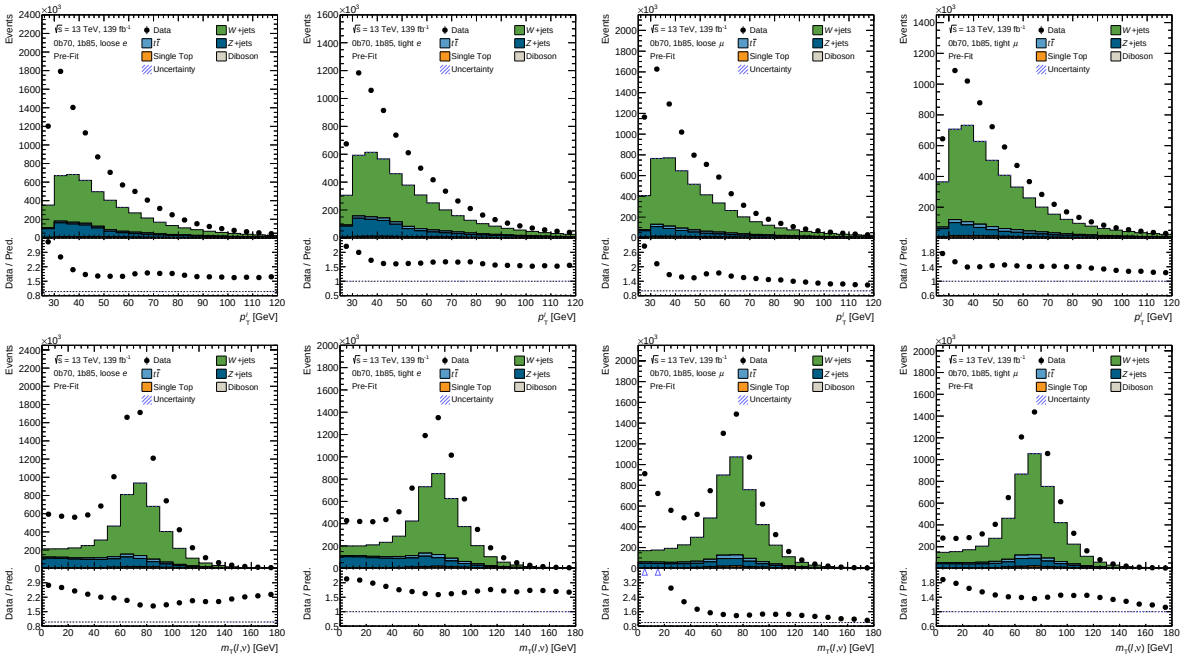


Figure 83: Distributions of events with loose and tight leptons in the 0b70,1b85 region as a function of p_{T}^{ℓ} (upper row) and $m_{\mathrm{T}}(\ell,\, \nu)$ (lower row). The first two distributions in each row correspond to events with an electron candidate, while the last two correspond to events with a muon candidate. The uncertainty bands correspond to the uncertainty due to the limited number of events in the MC samples.

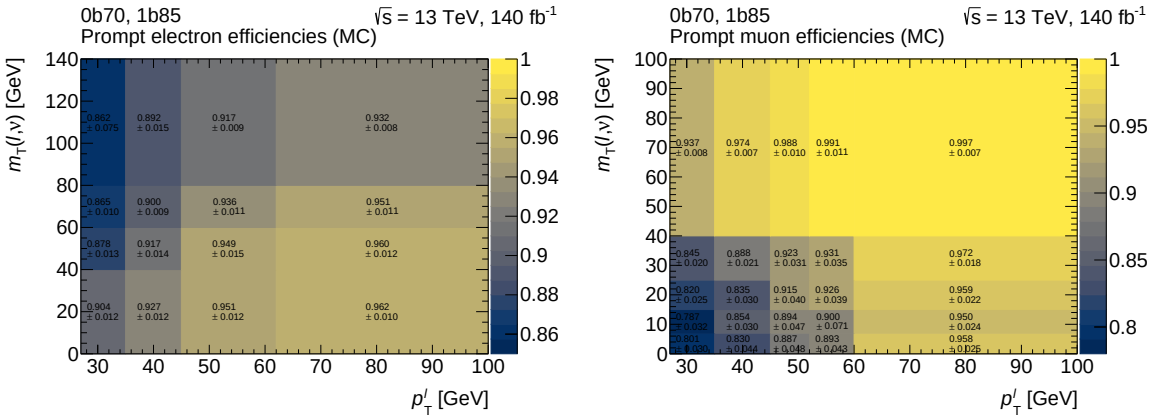


Figure 84: Prompt-lepton efficiencies as a two-dimensional function of p_{T}^{ℓ} and $m_{\mathrm{T}}(\ell,\ \nu)$ in the 0b70,1b85 region. The uncertainty on the prompt-lepton efficiencies correspond to the statistical uncertainties.

Figures 85, 86, and 87 show the observed and predicted distributions in the SR and $W\gamma$ CR for events with an electron candidate ($1e$) and events with a muon candidate (1μ) satisfying $m_{\mathrm{T}}(\ell,\ \nu) < 25\,\mathrm{GeV}$. The distributions are shown for the p_{T} of the photon candidate and the p_{T}-leading b-tagged jet, the number of jets, the reconstructed top-quark mass (m_t), the reconstruction type of the photon candidate ("photon conversion type"), and the p_{T} obtained from the combined four-momenta of the reconstructed top quark and the photon candidate. The reconstruction of the W boson and the top quark is discussed in Section 9.2. The resulting four-momentum of the top quark is used for the computation of m_t and $p_{\mathrm{T}}^{t+\gamma}$.

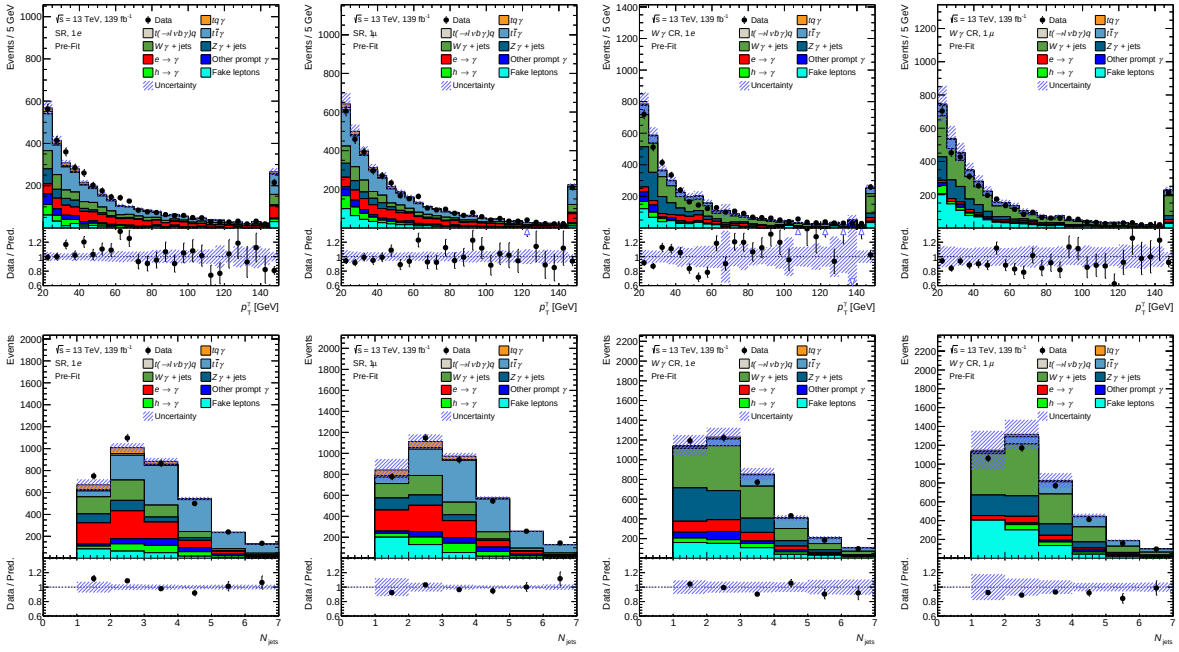


Figure 85: Predicted and observed p_{T}^{γ} and N_{jets} distributions in the fake-lepton enriched subsets of events in the SR and the $W\gamma$ CR. The distributions are shown separately for events with an electron candidate ($1e$) and events with a muon candidate (1μ). The uncertainty band corresponds to the sum in quadrature of the uncertainties due to limited sample sizes, the uncertainty in the shape of the fake-lepton contributions, and the uncertainty in the prediction of the overall fake-lepton contribution. The last bins include the overflow content.

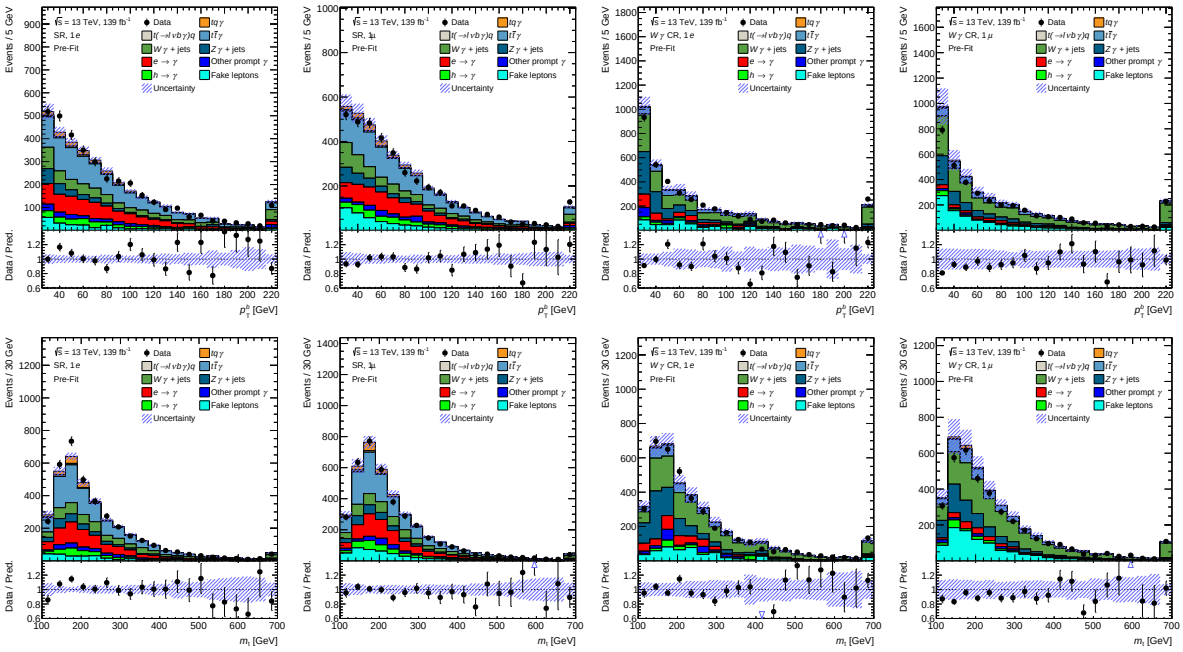


Figure 86: Predicted and observed p_{T}^{b} and m_t distributions in the fake-lepton enriched subsets of events in the SR and the $W\gamma$ CR. The distributions are shown separately for events with an electron candidate ($1e$) and events with a muon candidate (1μ). The uncertainty band corresponds to the sum in quadrature of the uncertainties due to limited sample sizes, the uncertainty in the shape of the fake-lepton contributions, and the uncertainty in the prediction of the overall fake-lepton contribution. The last bins include the overflow content. The first bin of the m_t distributions include the underflow content.

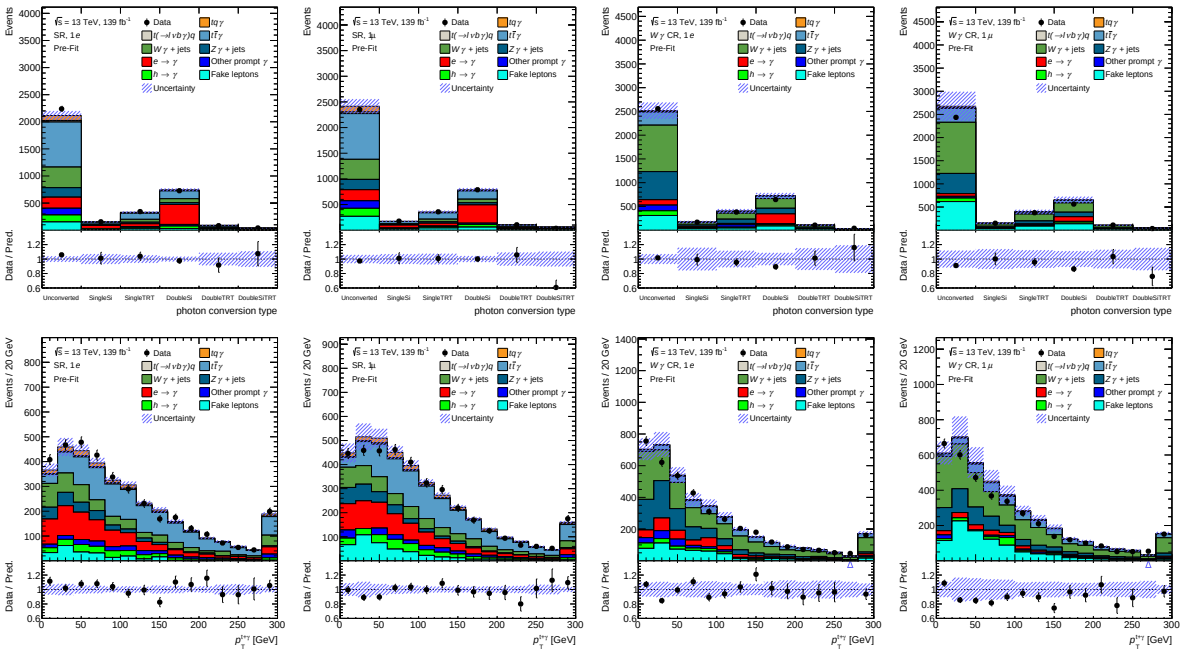


Figure 87: Predicted and observed "photon conversion type" and $p_{\mathrm{T}}^{t+\gamma}$ distributions in the fake-lepton enriched subsets of events in the SR and the $W\gamma$ CR. The distributions are shown separately for events with an electron candidate ($1e$) and events with a muon candidate (1μ). The uncertainty band corresponds to the sum in quadrature of the uncertainties due to limited sample sizes, the uncertainty in the shape of the fake-lepton contributions, and the uncertainty in the prediction of the overall fake-lepton contribution. The last bins of the $p_{\mathrm{T}}^{t+\gamma}$ distribution include the overflow content.

D.2 Estimation of $e \rightarrow \gamma$ contributions

The form of the likelihoods used in the BPL fits is adapted from Equation (49) and is given by Equation (85).

$$L(\boldsymbol{n}^{\text{data}}|\mu,\boldsymbol{\gamma}) = \prod_{b}^{N_b} \mathcal{P}\left(n^b|m^b(\mu,\gamma^b)\right) \times \mathcal{P}\left(k^b|\gamma^b\tau^b\right) \tag{85}$$

The number of predicted events in bin b in the $Z \rightarrow ee$ CR is given by Equation (86), where μ_{ee} is a free parameter of the fit regulating the overall prediction for the contribution $Z \rightarrow ee$ events. The $\sigma_X^{0,b}$ values are given by the nominal prediction in bin b for the respective processes X.

$$m^b(\mu_{ee},\gamma^b) = \gamma^b\left(\mu_{ee}\sigma^{0,b}_{Z\rightarrow ee} + \sigma^{0,b}_{B}\right) \tag{86}$$

The number of predicted events in bin b in the $Z \rightarrow e\gamma$ CR is given by Equation (87).

$$\begin{aligned} m^b(\mu_{e\gamma},\gamma^b,\boldsymbol{\beta},N_{\text{BP}}) &= \gamma^b\left(\mu_{e\gamma}\sigma^{0,b}_{Z\rightarrow e\gamma} + \sigma^{0,b}_{B}\right) \\ &\quad + N_{\text{BP}}\cdot B_3(\boldsymbol{\beta},\varsigma^b) \\ B_3(\varsigma^b) &= \beta_0\cdot b_{0,3}(\varsigma^b) + \beta_1\cdot b_{1,3}(\varsigma^b) + \beta_2\cdot b_{2,3}(\varsigma^b) + b_{3,3}(\varsigma^b) \end{aligned} \tag{87}$$

The variables $\mu_{e\gamma}$, N_{BP}, β_0, β_1, and β_2 are free parameters of the fit. The $\mu_{e\gamma}$ parameter adjusts the overall prediction for the contribution of $Z \rightarrow e\gamma$ events. The N_{BP} parameter regulates the overall predicted background contribution modeled by the third-order Bernstein polynomial $B_3(\varsigma_b)$. The β_i parameters regulate the relative contribution of the individual basis polynomials $b_{i,3}$. The basis polynomials are given by:

$$\begin{aligned} b_{0,3}(x) &= (1-x)^3 \\ b_{1,3}(x) &= 3x(1-x)^2 \\ b_{2,3}(x) &= 3x^2(1-x) \\ b_{3,3}(x) &= x^3. \end{aligned}$$

The argument of the polynomials is given by Equation 88, where x^b represents the value of the center of bin b, x^{min} and x^{max} represent the minimum and maximum x values of the distributions considered in the fit, respectively.

$$\varsigma^b = (x^b - x^{\text{min}})/(x^{\text{max}} - x^{\text{min}}) \tag{88}$$

D.3 Estimation of $h \rightarrow \gamma$ contributions

Figure 88 shows example distributions of the number of PDs (N_{PD}) as a function of $\text{SF}^{h\rightarrow\gamma}$, which are used to compute the statistical uncertainties of the $\text{SF}^{h\rightarrow\gamma}$ values.

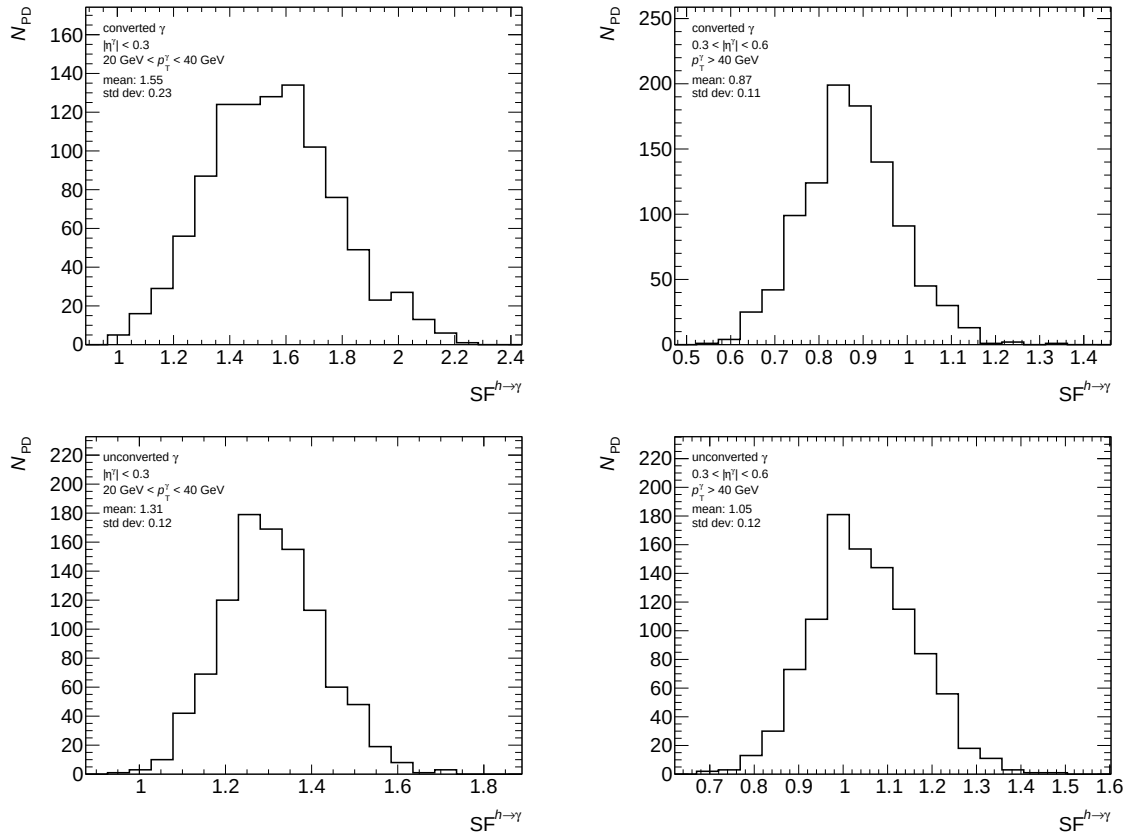


Figure 88: Distributions of the number of PDs (N_{PD}) as a function of $\mathrm{SF}^{h\to\gamma}$ for converted photon candidates (upper row) and unconverted photon candidates (lower row). The left column shows the distribution for photon candidates satisfying $E_{\mathrm{T}}^{\gamma} < 40\,\mathrm{GeV}$ and $|\eta^{\gamma}| < 0.3$ and the right column shows the distributions for photon candidates satisfying $E_{\mathrm{T}}^{\gamma} > 40\,\mathrm{GeV}$ and $0.3 < |\eta^{\gamma}| < 0.6$. The mean and the standard deviation of the distributions are given in the legend.

E Discrimination of $tq\gamma$ and background contributions

This appendix provides additional material associated with the studies presented in Chapter 9.

E.1 Validation of the selection of input features

The initial selection of input features of the NN models is based on the feature-importance rankings of GBDT models fitted to preliminary input data by the author of Ref. [7]. The validation of the selection of high-level features with GBDT models fitted to the finalized input data is discussed below.

A brief general overview of the GBDT models as implemented in the XGBoost framework is given in the following. GBDT is an ensemble ML algorithm, where decision trees are fitted to the input data in t sequential steps and are combined into a single model. In each step, a single decision tree is constructed which is required to minimize an objective function $\mathcal{O}(\theta)$ depicted in Equation (89), which is the sum of a loss function $\mathcal{B}(\theta)$ and a regularization term $\Omega(\theta)$ that depend on the free parameters θ of the model. The loss function is a measure of the prediction error and is given by the logistic loss function shown in Equation (90) for the GBDT models used in this work. Here, w_i and y_i are the weight and the true value of the i-th of N events in the input data and p_i represents the corresponding $tq\gamma$ probability that is obtained from the GBDT model prediction $\hat{y}_i$ using the sigmoid function h shown in Equation (91). This probability is referred to as the GBDT output (GBDT_{out}). The logistic loss function is also referred to as binary cross entropy in the following. The regularization term limits the complexity of the model by penalizing the objective function for higher levels of model complexity.

$$\mathcal{O}(\theta) = \mathcal{B}(\theta) + \Omega(\theta) \tag{89}$$

$$\mathcal{B}(\theta) = -\sum_{i}^{N} w_i \cdot \Big[y_i \log p_i(\theta) + (1 - y_i) \log \left(1 - p_i(\theta)\right) \Big] \tag{90}$$

$$p_i(\theta) = h(\hat{y}_i(\theta)) = \frac{1}{1 + e^{-\hat{y}_i(\theta)}} \tag{91}$$

In each step, the growth of the decision tree starts from a single root node. Decision nodes are added to the decision tree iteratively through a greedy algorithm. In each iteration, a decision node is added and the associated decision rule is defined, so that splitting the input data into two parts based on the decision rule maximizes the "gain" score, which is a measure of the improvement in the objective function. The decision rule is a requirement on the value of one of the input features and assigns each tested event unambiguously to one of two child nodes. Therefore, each decision node is unambiguously associated with one of the input features. When the growth of the tree is stopped because, e.g., no decision node is found to improve the objective function, all nodes, i.e., child nodes or the root node, that were not transformed into decision nodes become leaf nodes. A leaf weight is assigned to each leaf node. The value of the leaf weights and the "gain" score depend on the first (g_i) and second derivatives (h_i) of the loss function with respect to the prediction value $\hat{y}_i^{t-1}$ of the $(t-1)$-th step at $\hat{y}_i^{t-1}$ ("gradient boosting") for the t-th tree. The prediction value $\hat{y}_i^t$ for the i-th event after the t-th fitting step is given by the sum of the leaf weights of the leaf nodes of all t trees that the i-th event is assigned to.

The XGBoost package provides several parameters, referred to as hyperparameters, which can be adjusted by the user and are employed for the user-controlled configuration of the fitting of the GBDT models. Only those that are not set to default values are discussed in the following. The γ parameter is set to 0.9, which is the required size of the gain score for each additional decision node. Furthermore, a decision node is only added when the sum of the second derivatives h_i of all events assigned to a child node is larger than 2 for both child nodes ("min_child_weight"). The α and λ parameters correspond to the L1- and L2-regularization of the leaf weights, respectively, and are both set to 200. The values of all newly added leaf weights are scaled by the learning rate η, which is set to 0.1, at the end of each step.

Before performing the t-th fitting step, a subset is randomly sampled from the input data consisting of 80% of all events. Furthermore, 80% of the input features are randomly selected from the list of input features. Only this subset is considered for the growth of the t-th decision tree. Moreover, each decision tree is limited to a depth of four, i.e., only up to four consecutive decision nodes are allowed for each tree.

In each SR, the GBDT model is built by using $t = 500$ sequential steps and by considering all candidate features as input features. The input data is split through stratified random sampling into two statistically independent datasets, a training dataset and a validation dataset. While the former is used as input data for fitting the GBDT model and contains 70% of the input data, the latter is used for monitoring the fitting progress and for evaluating the model performance. The ratio of the number of $tq\gamma$ events and the number of background events is conserved when the input data is split. As the event weights are considered in the loss function, the $tq\gamma$ event weights are scaled, so that the sum of $tq\gamma$ event weights and the sum of background event weights is the same. This ensures that the reduction of the prediction error for $tq\gamma$ and for background events are of equal importance in the fitting.

Figure 89 shows the fraction of $tq\gamma$ and of background events as function of GBDT_{out} of the fitted GBDT models separately for the training and validation datasets for both SRs. The GBDT models are found to discriminate well between $tq\gamma$ and background events as the values of the $tq\gamma$ distributions are significantly larger than the values of the background distributions for $\text{GBDT}_{\text{out}} > 0.7$, i.e., high $tq\gamma$ probabilities, while the corresponding difference increases monotonically with increasing GBDT_{out} values in this region. An opposite behavior is observed for $\text{GBDT}_{\text{out}} < 0.6$, i.e. the values of the background distributions are significantly larger for low $tq\gamma$ probabilities. The separation in the $\geq$ 1fj SR is more powerful, as the GBDT model exploits the unique attributes of the high-p_{T} forward jet in this region.

The ratio panels show the differences in the event fractions between the training and validation datasets (val−train) relative to the uncertainty due to the limited number of simulated events ($\sigma_{\text{MC stat.}}$). A good agreement between equivalent distributions from different datasets is observed, as no value of the ratio is larger than two while no trend is visible for the distribution of its values. The agreement is quantified by performing two-sample Kolmogorov-Smirnov (2-KS) tests of the hypothesis that the equivalent distributions from the different datasets can be drawn from the same p.d.f.. Small p-values of these tests indicate that the shape of the distributions disagree significantly, which is likely caused by the corresponding GBDT model suffering from significant overfitting, i.e., the model has adapted to patterns in the training dataset that are not general features of the input data. All corresponding p-values are larger than 0.3 and, therefore, no significant disagreement between the tested distributions and no hint for significant overfitting are found. Conclusively, the fitted GBDT models are suitable for the validation of the selection of high-level input features.

The "total gain" (F_i) is chosen as the measure of the feature importance of an input feature i, which corresponds to the sum of the gain scores of all decision nodes in a GBDT model associated with feature i. Figure 90 lists the 30 largest F_i values normalized to the largest value F_{max} across all input features, and the corresponding variables. In both SRs, m_t is the input feature with the maximum total gain F_{max}. The orange bars mark the variables that are selected for the initial lists of input features shown in Table 13. In the 0fj SR, the list of input features contains the nine highest-ranked features and 17 of the top 20 ranked features. In the 1fj SR, the list of input features contains the twelve highest-ranked features and 18 of the top 20 ranked features. In both regions, m_t, $m_{\ell+b}$, $m_{\text{T}}(\ell, \nu)$, and $\Delta\eta(b, \gamma)$ are among the seven most important features. The two highest-ranked features are characterized by similar F_i values while these values are significantly larger than the values for the lower ranked features. Furthermore, the majority of all highly ranked variables are high-level features, while lj=fj? is the only low-level feature with a large feature importance in one of the SRs.

The contribution to the total gain of the GBDT models from features that are highly ranked (within the 20 most important features) but not selected (blue bars) is small relative to the contribution from

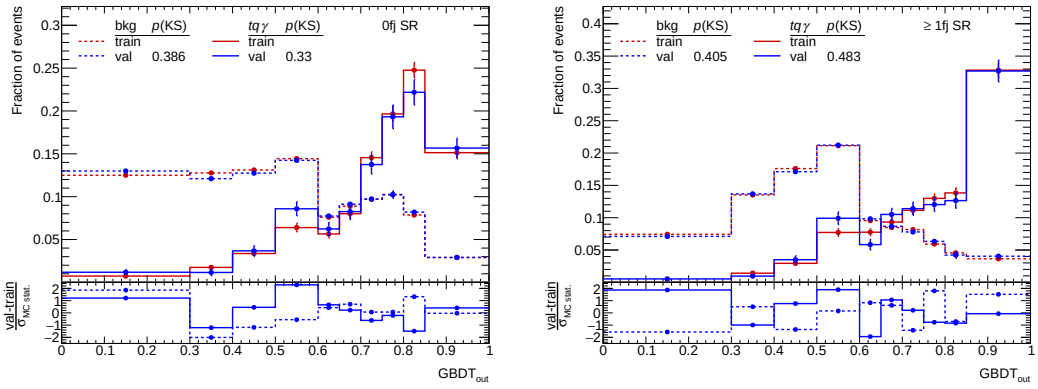


Figure 89: Fraction of $tq\gamma$ events and of background events ("bkg") as a function of the output value of the GBDT models (GBDT_{out}) for the training ("train") and validation ("val") datasets for the 0fj SR (left) and ≥ 1fj SR. The uncertainty bars correspond to the uncertainty due to the limited number of simulated events. Furthermore, the p-values of two-sample Kolmogorov-Smirnov (p(KS)) tests performed on all equivalent distributions from the different datasets are displayed in the legend. The ratio panels show the differences in the event fractions between equivalent distributions of the training and validation datasets relative to the total uncertainty ($\sigma_{\text{MC stat.}}$) due to the limited number of simulated events of both datasets.

the highest-ranked selected features. Moreover, fitting the GBDT models only using the features of the initial input lists (c.f. Table 13) leads to tiny, insignificant increases of the loss functions by 0.1% and 0.3% in the 0fj SR and ≥ 1fj SR, respectively. Therefore, the disregarded features do not have a significant contribution to the separation power of the GBDT models. In conclusion, this validates that no candidate high-level feature offering significant discriminative power was disregarded from the list of initial input features.

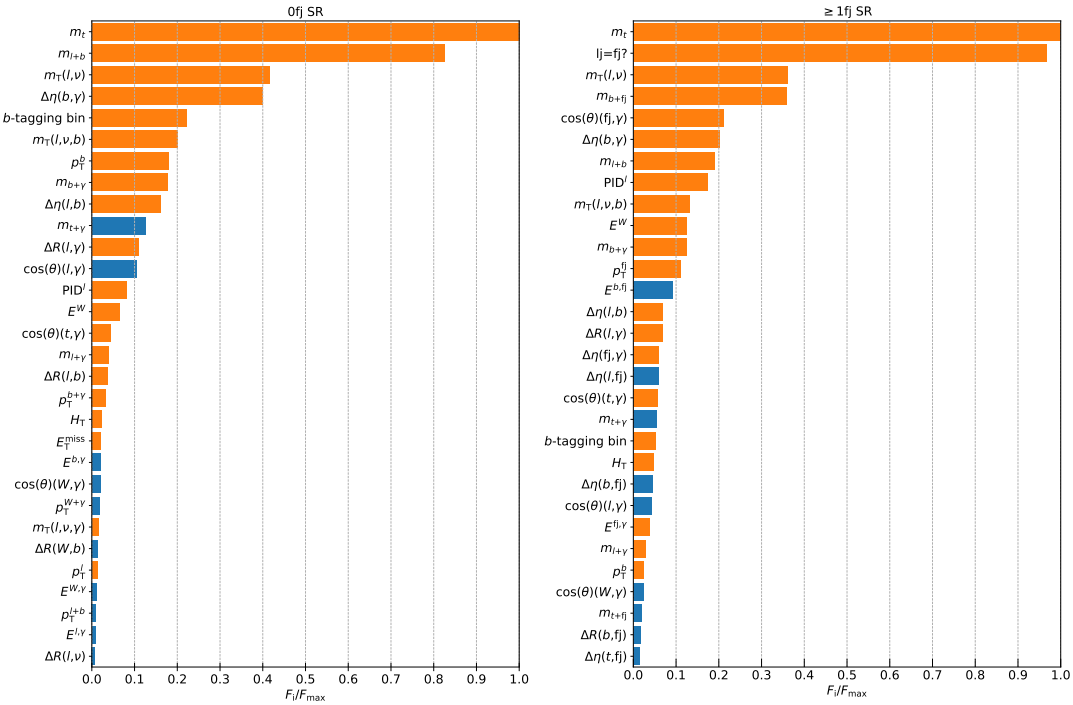
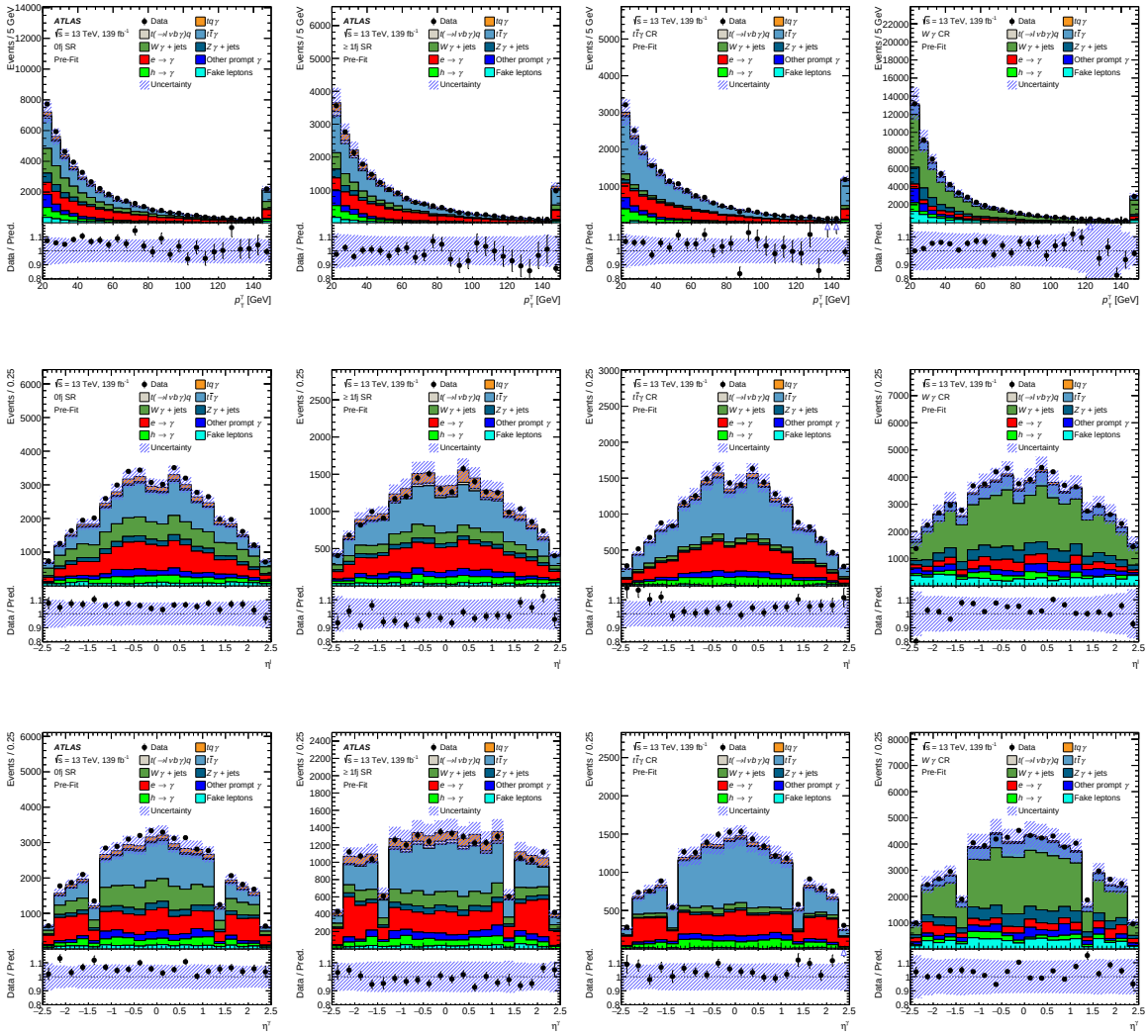
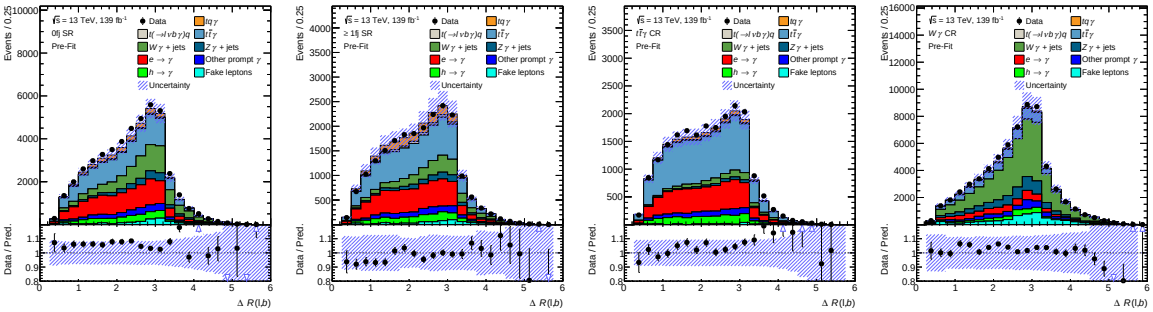


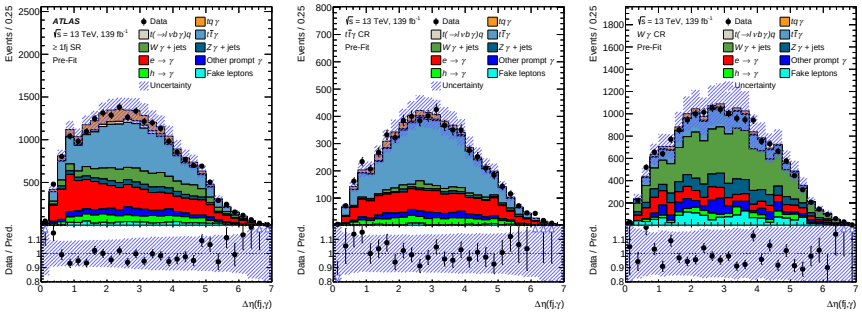
Figure 90: Lists of the 30 variables with the highest feature importance score F_i for the 0fj SR (left) and ≥ 1fj SR (right). The bars represent the values of F_i normalized to the highest value F_{max} across all features (m_t in both SRs). The orange bars correspond to variables that are selected for the initial lists of input features while blue bars correspond to disregarded candidate features.

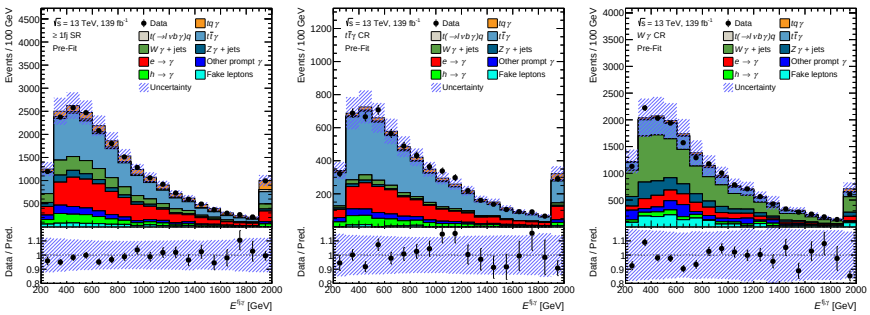
E.2 Distributions of input features

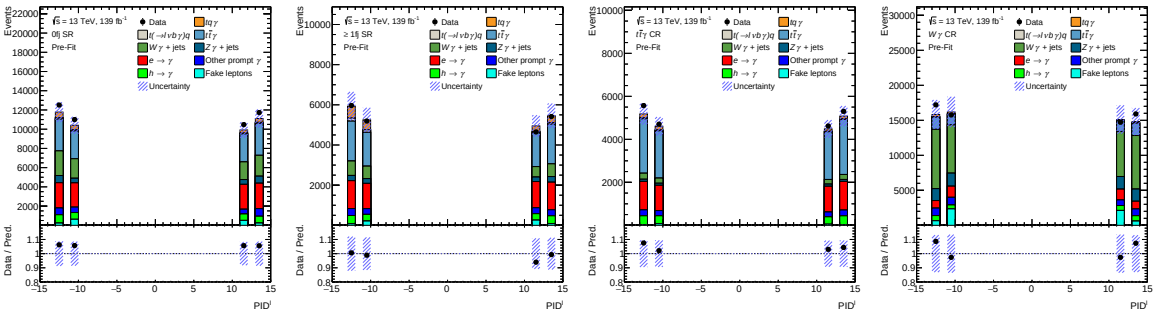
Distributions of the final NN input features for the collected data and the sum of the predicted signal and background contributions. The distributions are shown for all SRs and CRs, except for kinematic features involving the forward jet that are not shown for the 0fj SR, as events in that region do not contain a forward jet. The type of the charged lepton (PID^l) is provided as integer values where -11 and -13 refer to electrons and muons, and 11 and 13 to the corresponding antiparticles, respectively. The first and last bins of the distributions include any overflow and underflow contributions, respectively. The ratio panels show the ratios ("Data/Pred.") of the observed and predicted distributions. The hashed bands correspond to the total uncertainties on the prediction. The figures labelled with "ATLAS" are taken from [10].

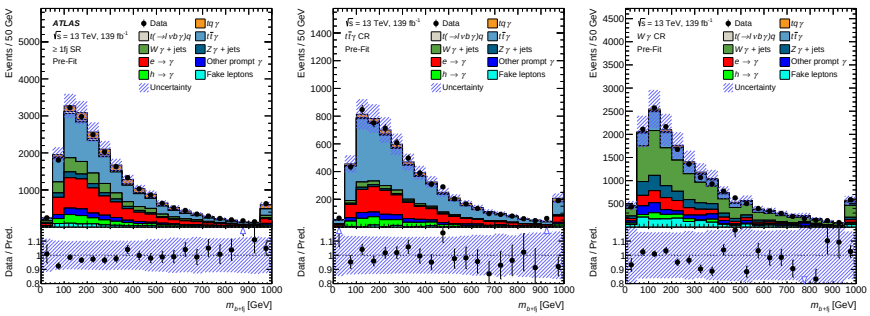


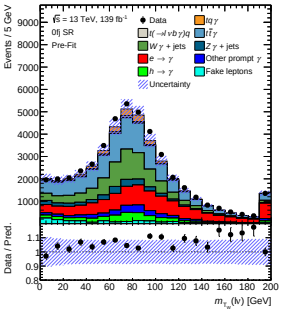
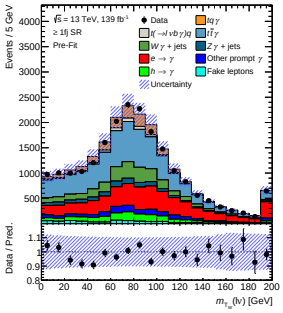
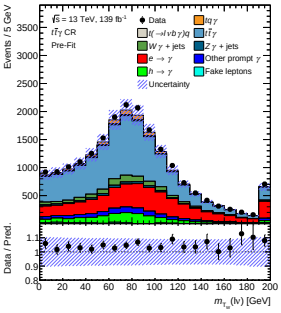
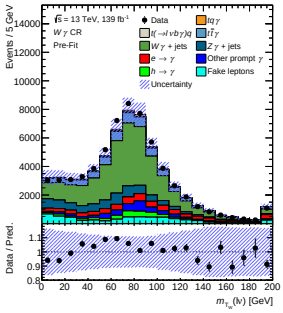
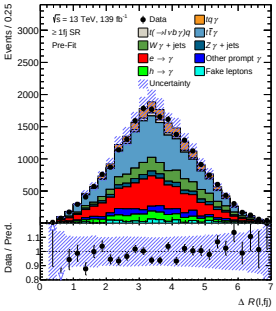
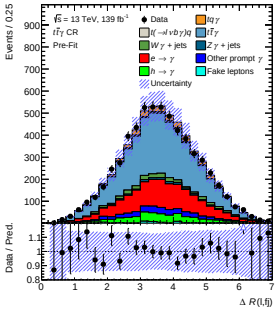
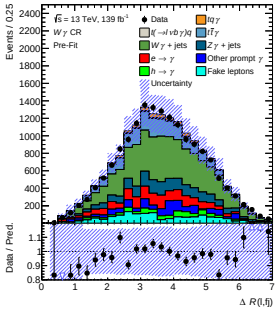
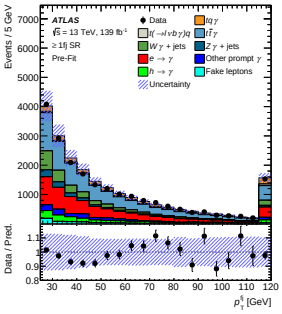
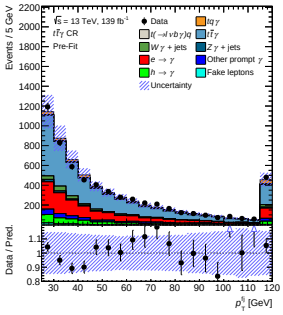
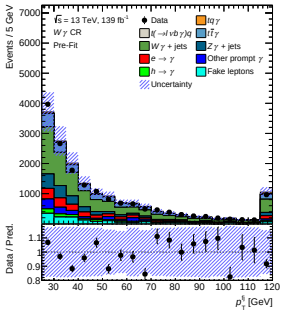
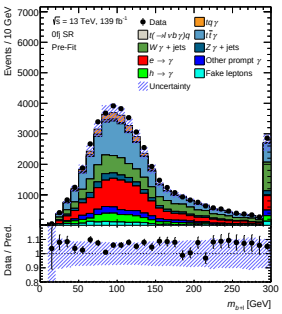
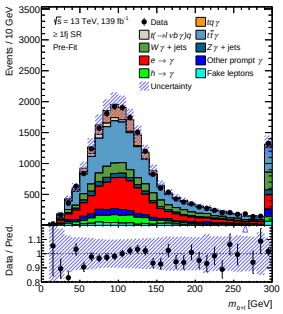
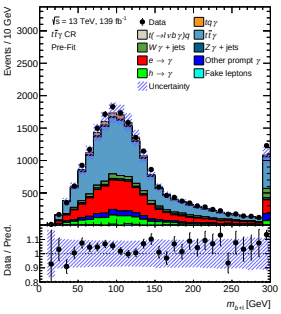
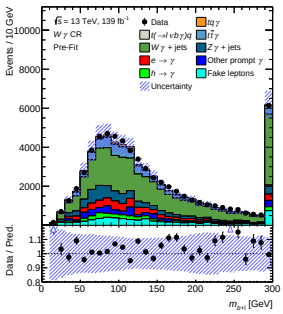
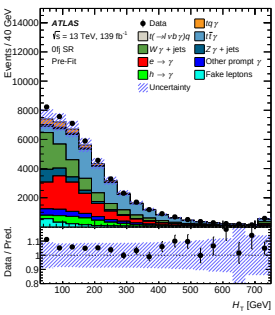
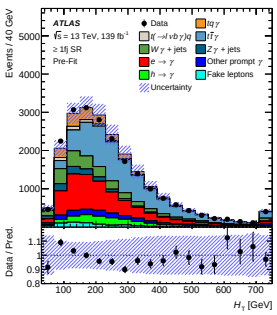
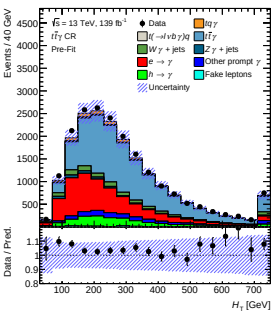
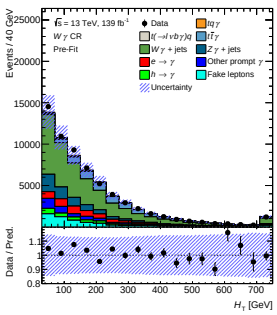

Events / 0.25
Data / Pred.
Δ R(l,b)
0j SR
≥ 1j̄ SR
tt̄γ CR
Wγ CR
Pre-Fit
Data
tqγ
l(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty

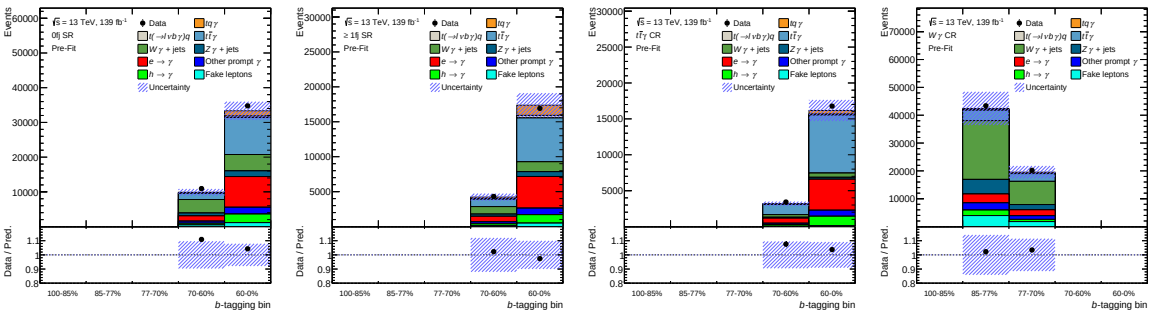

ATLAS
Events / 0.25
Data / Pred.
Δη(l̄j,γ)
≥ 1j̄ SR
tt̄γ CR
Wγ CR
Pre-Fit
Data
tqγ
l(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty

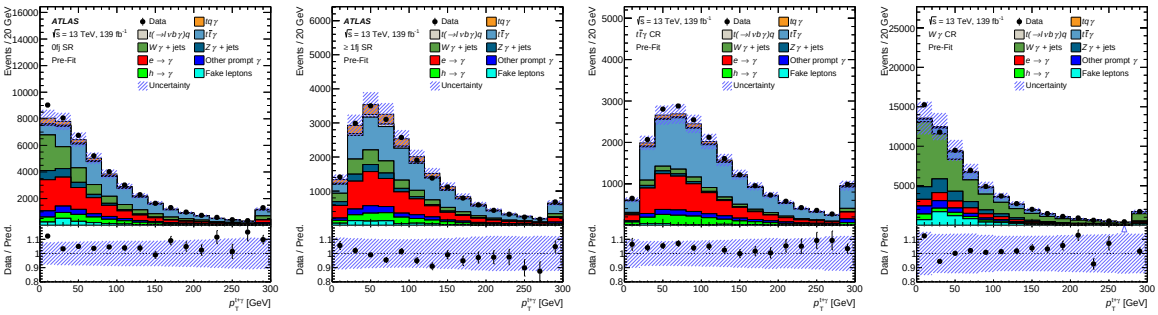

ATLAS
Events / 100 GeV
Data / Pred.
E^l̄j [GeV]
≥ 1j̄ SR
tt̄γ CR
Wγ CR
Pre-Fit
Data
tqγ
l(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty

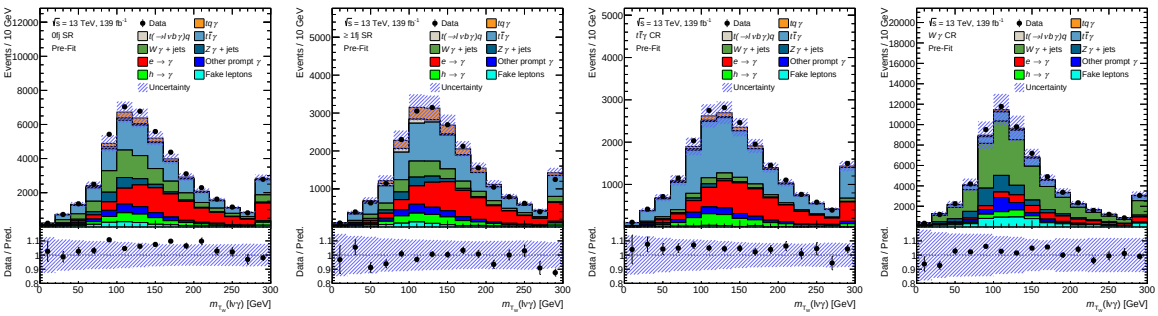

Events
Data / Pred.
PID^l
0j SR
≥ 1j̄ SR
tt̄γ CR
Wγ CR
Pre-Fit
Data
tqγ
l(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty

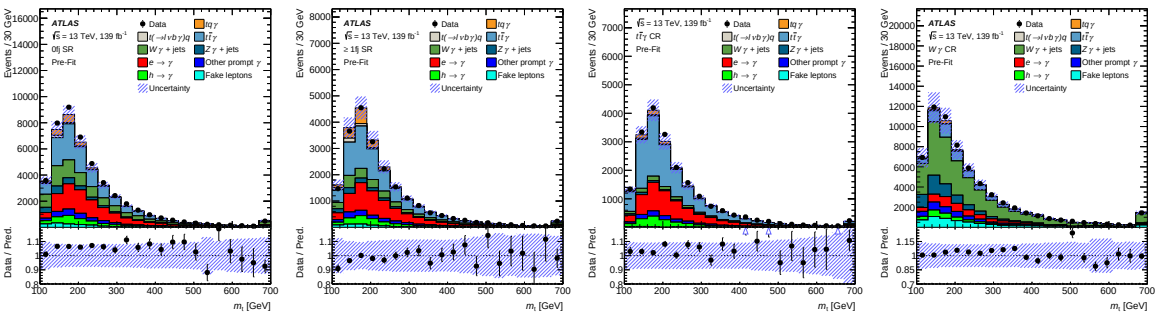

ATLAS
Events / 50 GeV
Data / Pred.
m_bγj [GeV]
≥ 1j̄ SR
tt̄γ CR
Wγ CR
Pre-Fit
Data
tqγ
l(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty

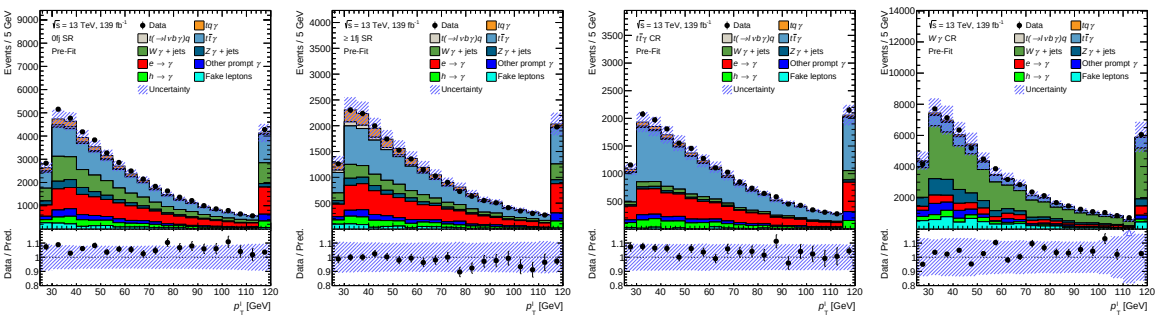

√s = 13 TeV, 139 fb⁻¹
0ℓ SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
m_Tω(lν) [GeV]

√s = 13 TeV, 139 fb⁻¹
≥ 1ℓj SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
m_Tω(lν) [GeV]

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
m_Tω(lν) [GeV]

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
m_Tω(lν) [GeV]

√s = 13 TeV, 139 fb⁻¹
≥ 1ℓj SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,fj)

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,fj)

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,fj)

ATLAS
√s = 13 TeV, 139 fb⁻¹
≥ 1ℓj SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^ℓ [GeV]

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^ℓ [GeV]

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^ℓ [GeV]

√s = 13 TeV, 139 fb⁻¹
0ℓ SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 10 GeV
Data / Pred.
m_bH [GeV]

√s = 13 TeV, 139 fb⁻¹
≥ 1ℓj SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 10 GeV
Data / Pred.
m_bH [GeV]

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 10 GeV
Data / Pred.
m_bH [GeV]

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 10 GeV
Data / Pred.
m_bH [GeV]

ATLAS
√s = 13 TeV, 139 fb⁻¹
0ℓ SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 40 GeV
Data / Pred.
H_T [GeV]

ATLAS
√s = 13 TeV, 139 fb⁻¹
≥ 1ℓj SR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 40 GeV
Data / Pred.
H_T [GeV]

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 40 GeV
Data / Pred.
H_T [GeV]

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→ℓνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 40 GeV
Data / Pred.
H_T [GeV]

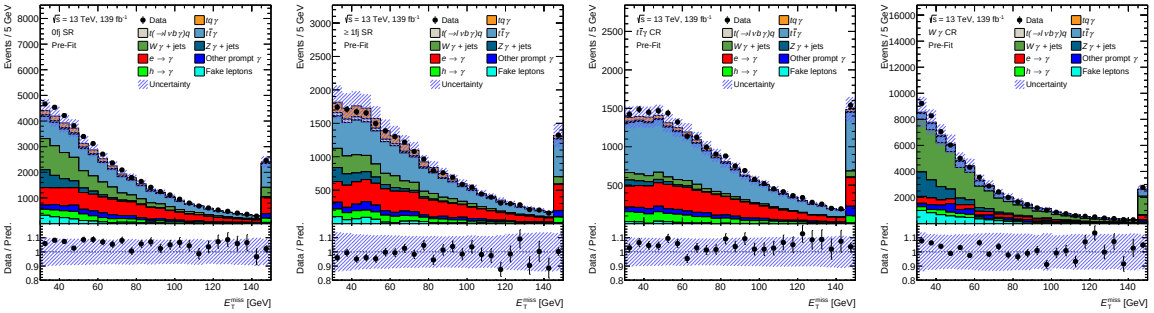

Events
Data / Pred.
b-tagging bin
100-85%
85-77%
77-70%
70-60%
60-0%


ATLAS
Events / 20 GeV
Data / Pred.


Events / 10 GeV
Data / Pred.


ATLAS
Events / 30 GeV
Data / Pred.

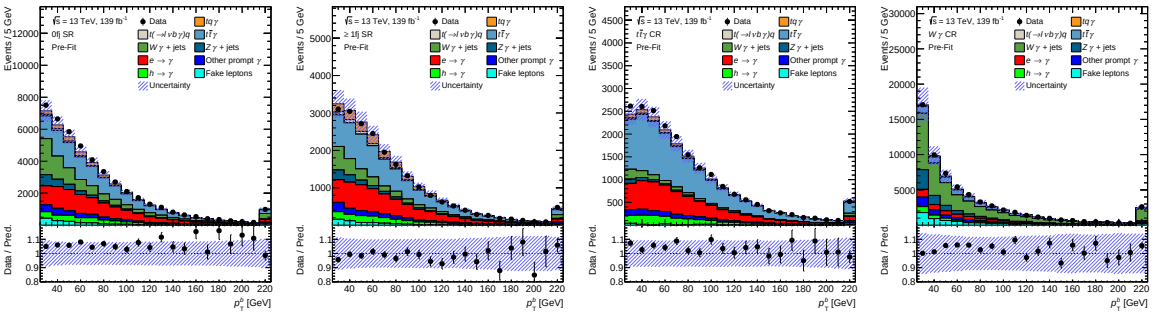

Events / 5 GeV
Data / Pred.


√s = 13 TeV, 139 fb⁻¹
0fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
E_T^miss [GeV]

√s = 13 TeV, 139 fb⁻¹
≥ 1fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
E_T^miss [GeV]

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
E_T^miss [GeV]

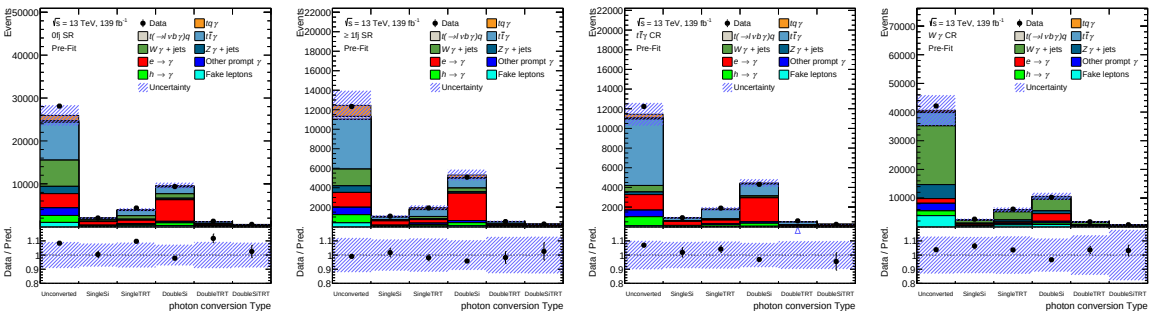
√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
E_T^miss [GeV]


√s = 13 TeV, 139 fb⁻¹
0fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^b [GeV]

√s = 13 TeV, 139 fb⁻¹
≥ 1fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^b [GeV]

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^b [GeV]

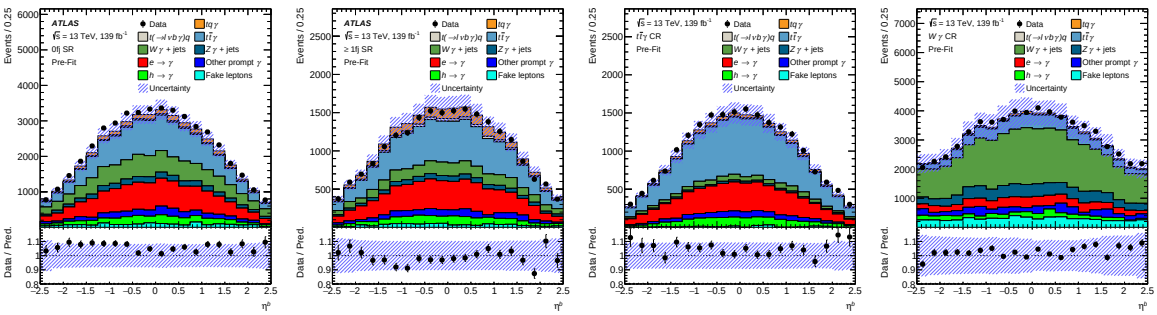
√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 5 GeV
Data / Pred.
p_T^b [GeV]


√s = 13 TeV, 139 fb⁻¹
0fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events
Data / Pred.
Unconverted
SingleSi
SingleTRT
DoubleSi
DoubleTRT
DoubleSiTRT
photon conversion Type

√s = 13 TeV, 139 fb⁻¹
≥ 1fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events
Data / Pred.
Unconverted
SingleSi
SingleTRT
DoubleSi
DoubleTRT
DoubleSiTRT
photon conversion Type

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events
Data / Pred.
Unconverted
SingleSi
SingleTRT
DoubleSi
DoubleTRT
DoubleSiTRT
photon conversion Type

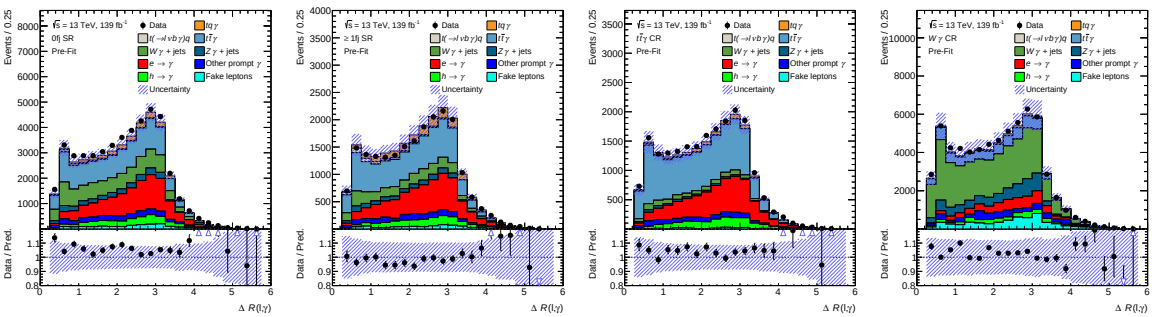
√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events
Data / Pred.
Unconverted
SingleSi
SingleTRT
DoubleSi
DoubleTRT
DoubleSiTRT
photon conversion Type


ATLAS
√s = 13 TeV, 139 fb⁻¹
0fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
η^b

ATLAS
√s = 13 TeV, 139 fb⁻¹
≥ 1fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
η^b

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
η^b

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
η^b


√s = 13 TeV, 139 fb⁻¹
0fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,γ)

√s = 13 TeV, 139 fb⁻¹
≥ 1fj SR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,γ)

√s = 13 TeV, 139 fb⁻¹
tt̄γ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,γ)

√s = 13 TeV, 139 fb⁻¹
Wγ CR
Pre-Fit
Data
tqγ
ℓ(→lνbγ)q
tt̄γ
Wγ + jets
Zγ + jets
e → γ
Other prompt γ
h → γ
Fake leptons
Uncertainty
Events / 0.25
Data / Pred.
Δ R(l,γ)

E.3 Investigation of the NN fitting in the 0fj SR

For the NN model in the 0fj SR, the difference between the values of the binary cross entropy obtained from the training and validation datasets increases as the fitting of the NN progresses (cf. Section 9.6). To investigate whether this behavior is mainly caused by the performance of the NN in the background-dominated region that is insensitive to the $tq\gamma$ process ($\mathrm{NN_{out}} < 0.4$), Figure 69 shows the value of the binary cross entropy as a function of the number of completed epochs obtained from the two datasets considering only events with $\mathrm{NN_{out}} > 0.4$. Both distributions agree well within their uncertainties and only differ insignificantly at the end of the fitting process. In conclusion, the observed differences between the values from the training and validation datasets when all events are considered are mainly caused by events satisfying $\mathrm{NN_{out}} < 0.4$.

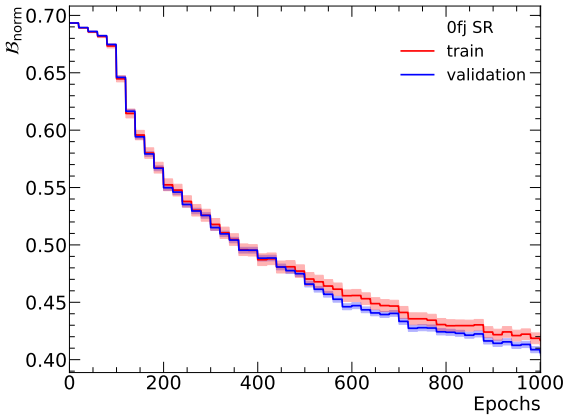


Figure 69: Binary cross entropy as a function of the number of completed epochs considering only events satisfying $\mathrm{NN_{out}} > 0.4$ for the 0fj SR. The uncertainty bands correspond to the statistical uncertainty evaluated with bootstrap sampling.

F Details of the statistical analysis

This appendix is associated with Chapter 10 and provides additional material about the investigation of the BPL fits and the behavior of the NPs used in the BPL fits. Furthermore, Section F.1 discusses the details of the algorithm used for the binning optimization of the $\mathrm{NN_{out}}$ distributions used in the SR and CRs (cf. Section 10.1).

F.1 Binning optimization

The "Transformation D" algorithm [272] is designed to automatically generate binnings of continuous distributions, based on the values of three user-defined parameters. It is used in this work to optimize the binnings of the $\mathrm{NN_{out}}$ distributions used in the BPL fits. An overview of this algorithm is provided in the following.

The values of the lower (k_i) and upper (l_i) boundaries of the i-th bin are guided by the value of ζ_i given by Equation (92).

$$\zeta_i = z_{\mathrm{sig}} \frac{n_{\mathrm{sig}}(k_i, l_i)}{N_{\mathrm{sig}}} + z_{\mathrm{bkg}} \frac{n_{\mathrm{bkg}}(k_i, l_i)}{N_{\mathrm{bkg}}} \tag{92}$$

The ζ_i value represents a weighted average of the predicted fractions of $tq\gamma$ and of background events contained in the i-th bin. The quantities $n_{\mathrm{X}}(k_i, l_i)$ and N_{X} denote the corresponding predicted number of events in the i-th bin and the predicted total number of events, respectively, where X denotes whether the numbers are computed for signal ("sig") or background ("bkg") contributions. The weights z_{sig} and z_{bkg} represent the parameters of the algorithm. The values of the lower and upper boundaries of the bins are sequentially defined, where the upper boundary of the i-th bin is given by the lower boundary of the $(i-1)$-th bin, i.e., $l_i = k_{i-1}$. The lower boundary of the i-th bin is given by $k_i = l_i - n_i\,\delta$, where δ denotes a variable step size and n_i is the smallest integer number for which ζ_i is larger than one. Two boundary conditions are applied: the upper boundary of the first bin is defined by the maximum value of the $\mathrm{NN_{out}}$ distribution, i.e., $l_1 = 1$, and the lower boundary of the last bin j is defined by the minimum value of the $\mathrm{NN_{out}}$ distributions, i.e., $k_j = 0$. However, if ζ_j is smaller than one, this bin is removed and the lower boundary of the new last bin, the $(j-1)$-th bin, is set to zero, i.e., $k_{j-1} = 0$. These conditions ensure that the definition of the remaining bins is unambiguous for any given parameter configuration. If z_{sig} and z_{bkg} are integer numbers, the number of bins defined by the algorithm equals the sum of the weights.

F.2 Results of the particle-level measurement and additional investigations of the systematic uncertainties

Figure 70 shows the predicted and observed $\mathrm{NN_{out}}$ distributions for the 0fj SR, $\geq$ 1fj SR, and $t\bar{t}\gamma$ CR, and the "Inclusive yield" in the $W\gamma$ CR, where the predicted signal and background distributions and their respective uncertainties are shown before (pre-fit) and after (post-fit) their adjustment through the BPL fit. The displayed predicted distributions correspond to the particle-level measurement.

Figure 71 shows the matrix of correlation coefficients for the particle-level measurement, analogously to that shown in Figure 53 for the parton-level measurement.

Table 29 provides a list of NPs that are constrained to be below 80% of their pre-fit uncertainty. The constraints of the "tW ME generator" and "Fake leptons norm." NPs are discussed in Section 10.2. The JER NPs affect the prediction of all processes except for fake leptons. The "JER effective NP2" affects the overall prediction by 1.6% in the $\geq$ 1fj SR and by 2.0% in the $W\gamma$ CR. The "JER effective NP4" affects the overall prediction by 1.8% in the $W\gamma$ CR. The "'$t\bar{t}$ ME generator" NP affects the overall

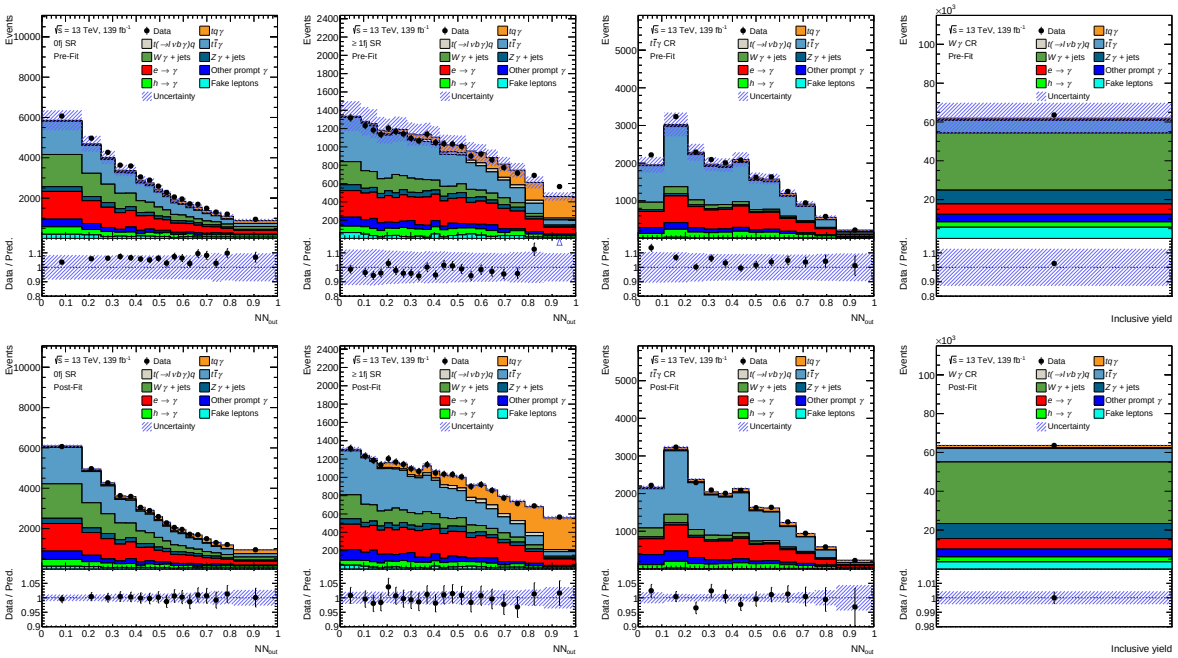


Figure 70: Predicted and observed ("Data") distributions in the ARs that are considered for the BPL fit for the particle-level measurement. The upper row shows the predicted distributions before performing the BPL fit and the lower row shows the predicted distributions adjusted by the BPL fit to the observed distributions. The NN_{out} distributions for the 0fj SR, $\geq$ 1fj SR (upper right), and $t\bar{t}\gamma$ CR (lower left), and the overall contributions ("Inclusive yield") for the $W\gamma$ CR are displayed. The hashed band corresponds to the total uncertainty of the sum of the predicted signal and background distributions. The lower panel displays the ratio of the observed and predicted distributions. The plots from the lower row are taken from Ref. [10].

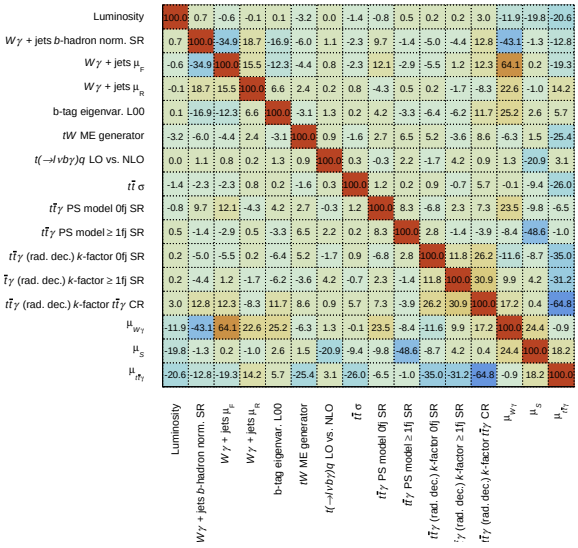


Figure 71: The matrix of the correlation coefficients of the displayed NPs and free-floating parameters for the particle-level measurement. The correlation coefficients are scaled by a factor 100 for better readability.

Table 29: List of nuisance parameters (NPs) constrained to ≤ 0.8 of their pre-fit uncertainty by the BPL fits. The values shown are the constraints, defined as the ratio of the post-fit to pre-fit uncertainty.

NP	Parton-level measurement	Particle-level measurement
tW ME generator	0.63	0.63
Fake leptons norm.	0.67	0.67
JER effective NP2	0.73	0.72
JER effective NP4	0.74	0.74
$t\bar{t}$ ME generator	0.76	0.75
$tq\gamma$ PS model	0.76	0.77

prediction by 1.4% in the $t\bar{t}\gamma$ CR. The "$tq\gamma$ PS model" NP especially affects the overall prediction in the last bin of the 0fj SR and the corresponding uncertainty on the overall prediction reaches up to 2%. Since these pre-fit impacts are relatively large, the NPs become significantly constrained by the BPL fits.

Table 30 provides a full list of NP groups. A description of all displayed groups is given in the following. The "$t\bar{t}\gamma$ modeling", "$t\,(\to \ell\nu b\gamma)\,q$ modeling", "$tq\gamma$ modeling", and "$t\bar{t}$ modeling" groups correspond to the impact of all NPs assigned to the predictions for the $t\bar{t}\gamma$, $t\,(\to \ell\nu b\gamma)\,q$, $tq\gamma$, and $t\bar{t}$ introduced in Section 7.6.2, respectively. The "Additional background uncertainties" group includes the NPs that are associated with the predictions of all background processes except the aforementioned ones.

Table 30: List of the impact of the groups of systematic uncertainties on $\mu_{tq\gamma}$ in the parton-level measurement and on μ_S in the particle-level measurement. The impact of each group is quantified in the following way: The BPL fit is performed while the values of the NPs of the respective group are fixed to their best-fit values. The uncertainty of the fitted $\mu_{tq\gamma(S)}$ is subtracted in quadrature from the nominal uncertainty of $\mu_{tq\gamma(S)}$. The square root of this value is denoted by $\Delta\sigma_{\mu_{tq\gamma(S)}}$ and the impact is given by $\Delta\sigma_{\mu_{tq\gamma(S)}}/\mu_{tq\gamma(S)}$. The values of the impacts are taken from Ref. [10].

Uncertainty	$\Delta\sigma_{\mu_{tq\gamma}}/\mu_{tq\gamma}$	$\Delta\sigma_{\mu_S}/\mu_S$
$t\bar{t}\gamma$ modeling	±5.5%	±5.5%
Background MC statistics	±3.5%	±3.6%
$tq\gamma$ MC statistics	±3.3%	±3.3%
$t\bar{t}$ modeling	±2.4%	±2.3%
$tq\gamma$ modeling	±2.0%	±2.3%
$t\,(\to \ell\nu b\gamma)\,q$ modeling	±1.9%	±3.3%
Additional background uncertainties	±1.9%	±2.0%
$t\,(\to \ell\nu b\gamma)\,q$ MC statistics	±0.3%	±0.3%
$h \to \gamma$ photon fakes	±2.0%	±2.2%
Lepton fakes	±1.9%	±2.1%
$e \to \gamma$ photon fakes	±0.6%	±0.6%
Luminosity	±2.2%	±2.2%
Pileup	±1.2%	±1.3%
Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$	±3.6%	±3.5%
Photons	±2.5%	±2.5%
Leptons	±0.9%	±0.9%
b-tagging	±0.9%	±0.7%
Total systematic uncertainty	±10.6%	±10.7%

The "Background MC statistics", "$tq\gamma$ MC statistics", and "$t\,(\to \ell\nu b\gamma)\,q$ MC statistics" groups

correspond to the impacts of the γ-factors associated with the background, $tq\gamma$ predictions and $t\,(\to \ell\nu b\gamma)\,q$ predictions, respectively.

The "Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$" group includes NPs that are associated with reconstructed jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. The "Photons", "Leptons", and "b-tagging" groups correspond to the NPs accounting for the uncertainties in the calibration of photons, leptons, and the b-tagging algorithm, respectively. The "Luminosity" and "Pileup" groups correspond to the NPs accounting for the uncertainties in the measured luminosity and in the pileup reweighting, respectively. The uncertainties that are related to these groups are discussed in Section 7.6.1.

The "$e \to \gamma$ fakes", "$h \to \gamma$ fakes", and "Lepton fakes" groups include the NPs accounting for the uncertainties of the $e \to \gamma$ SFs (cf. Section 8.2), $h \to \gamma$ SFs (cf. Section 8.3), and the fake-lepton estimation (cf. Section 8.1).

F.3 Modeling uncertainties of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ predictions for the parton-level measurement

In the following, the impacts of the NPs affecting the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ predictions are displayed for the parton-level measurement.

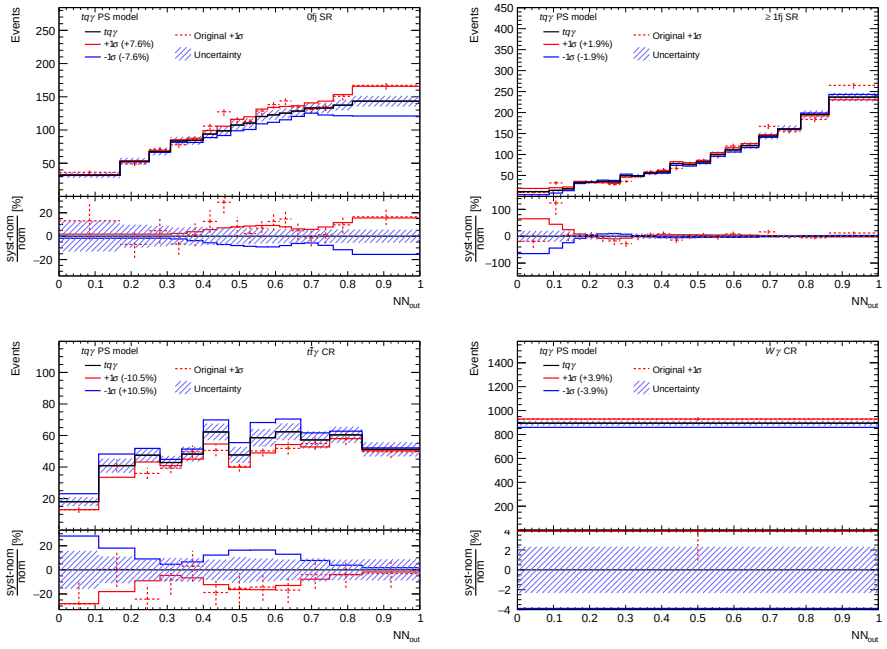


Figure 72: Impact of the variation of the PS model variation on the prediction of the $\mathrm{NN_{out}}$ distributions of $tq\gamma$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panels compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

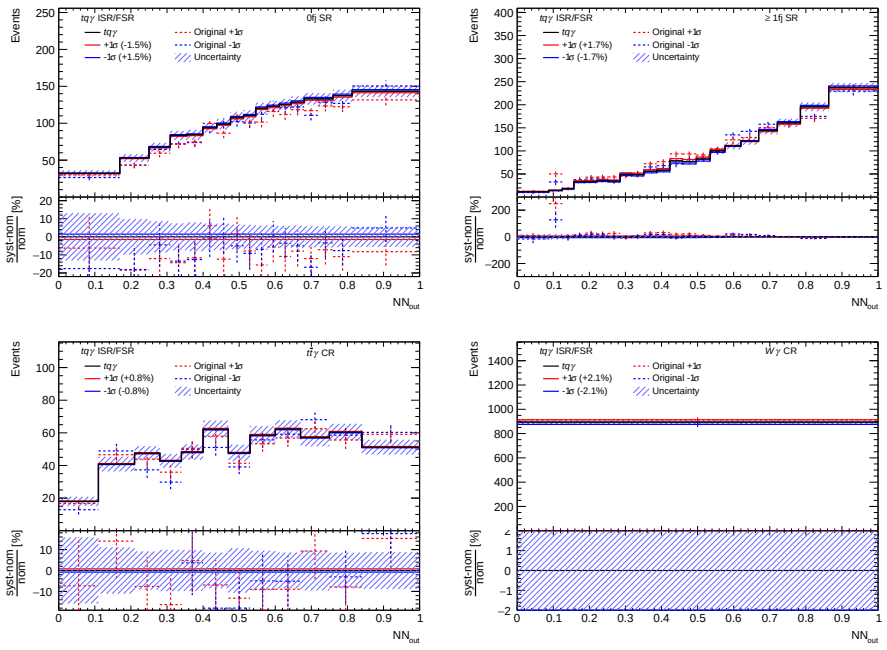


Figure 73: Impact of the Var3c variation (ISR/FSR) on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

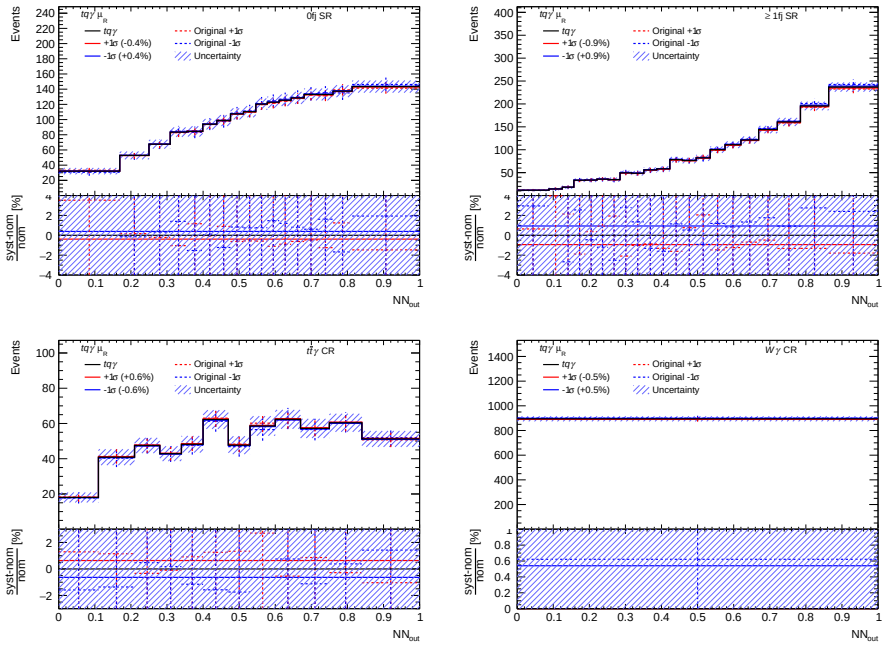


Figure 74: Impact of the μ_R variation on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

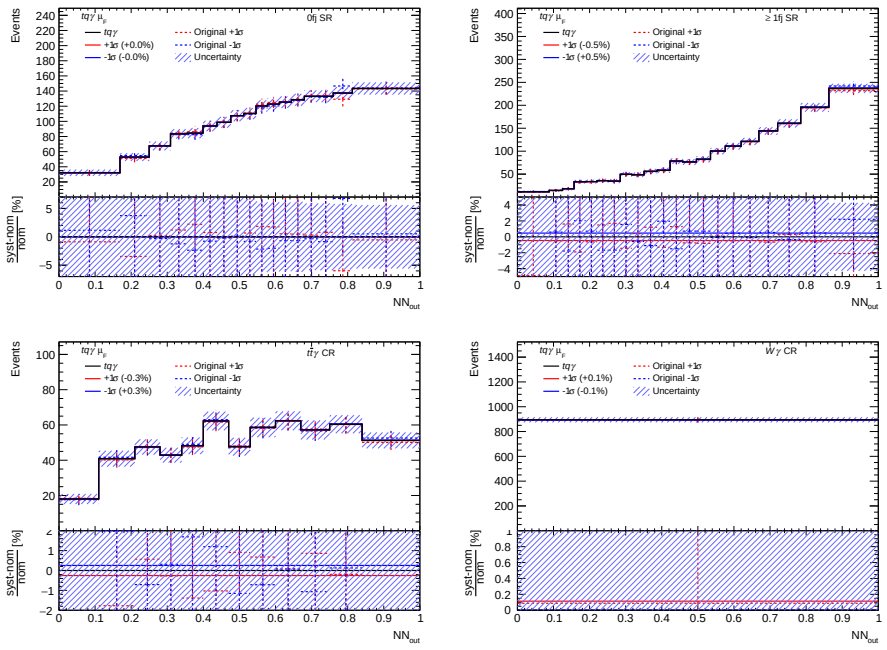


Figure 75: Impact of the μ_F variation on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

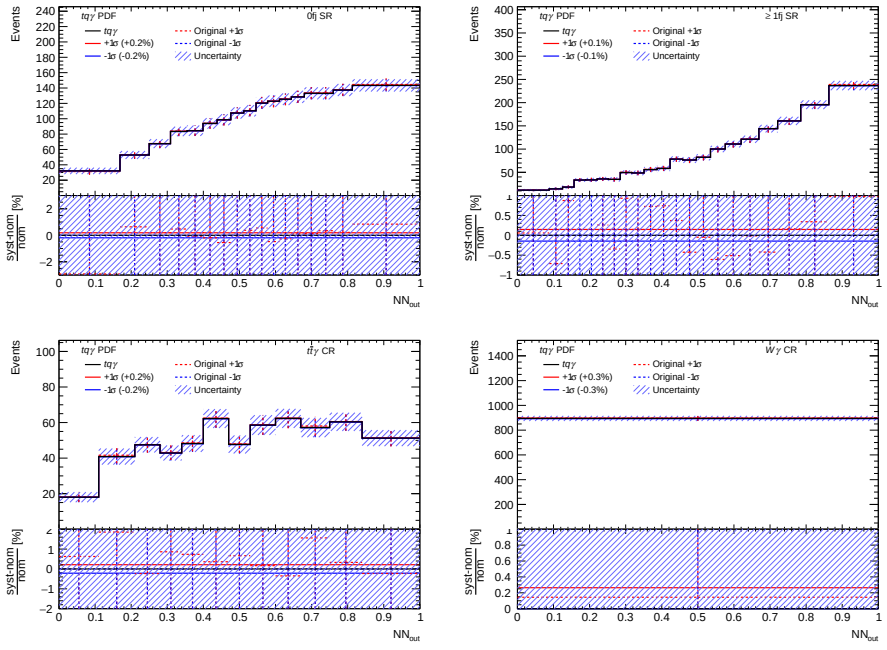


Figure 76: Impact of the PDF uncertainty on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

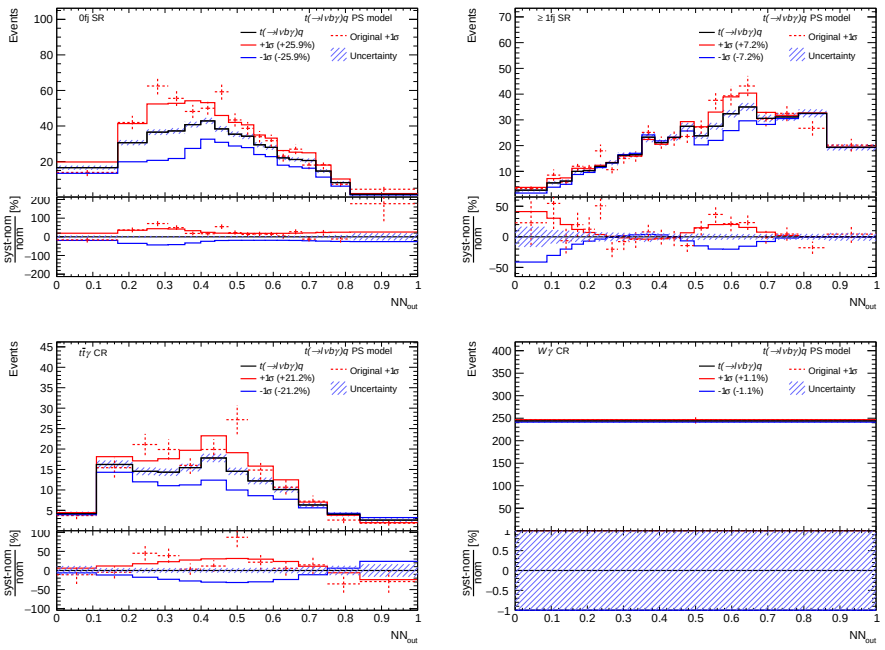


Figure 77: Impact of the variation of the PS model variation on the prediction of the NN_{out} distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

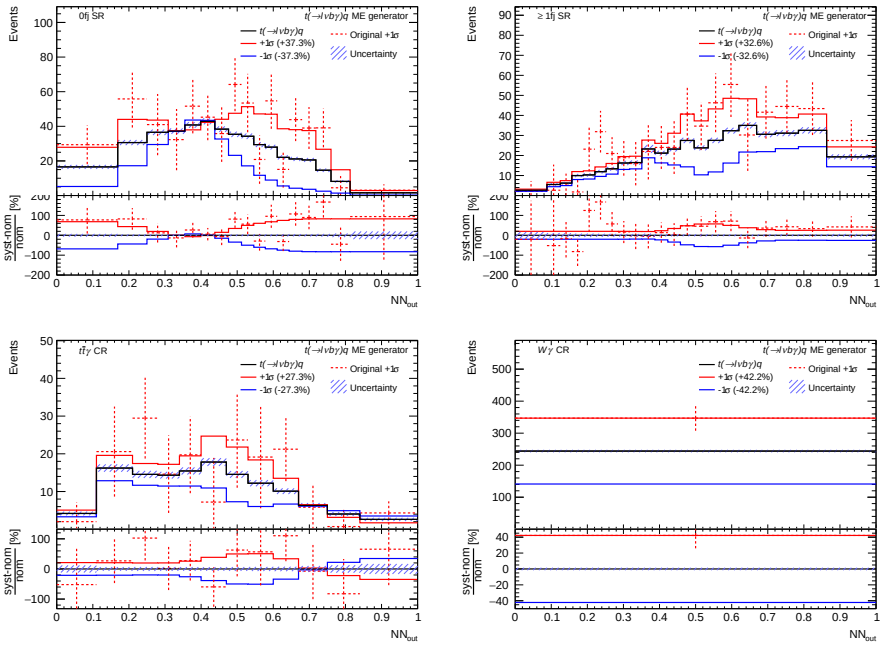


Figure 78: Impact of the variation of the ME generator variation on the prediction of the NN_{out} distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

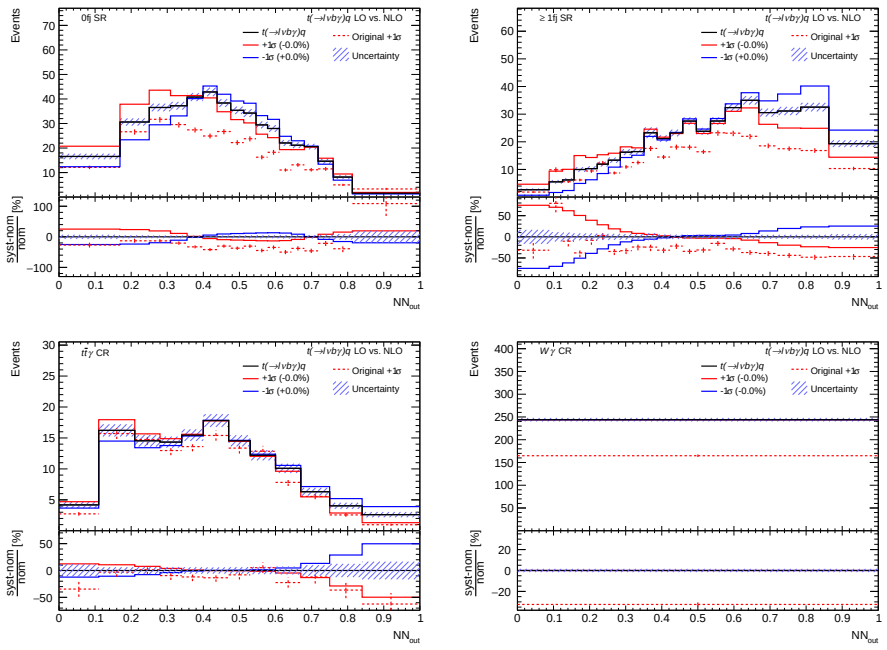


Figure 79: Impact of the "$t\,(\to \ell\nu b\gamma)\,q$ LO vs. NLO" NP on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

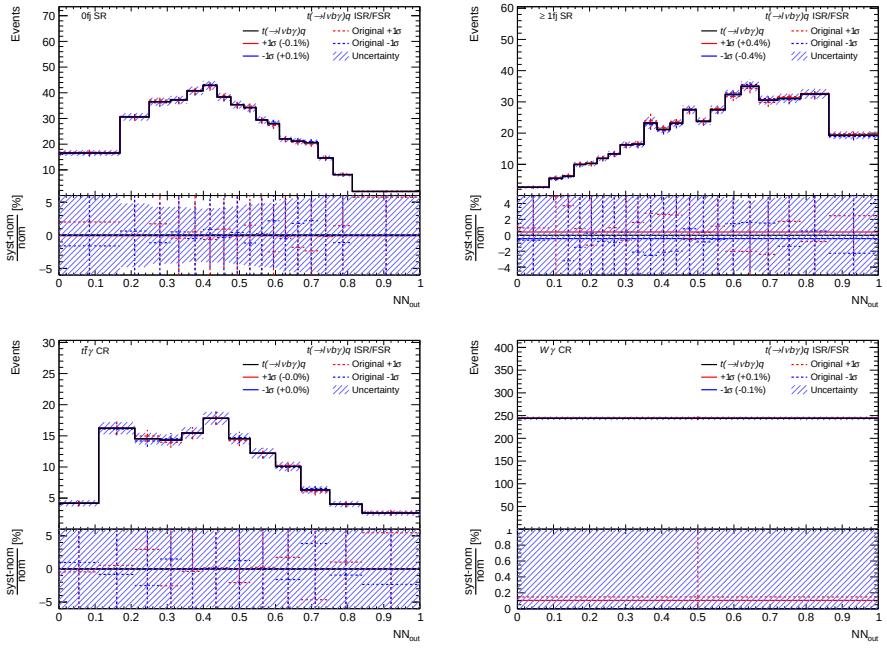


Figure 80: Impact of the Var3c variation (ISR/FSR) on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

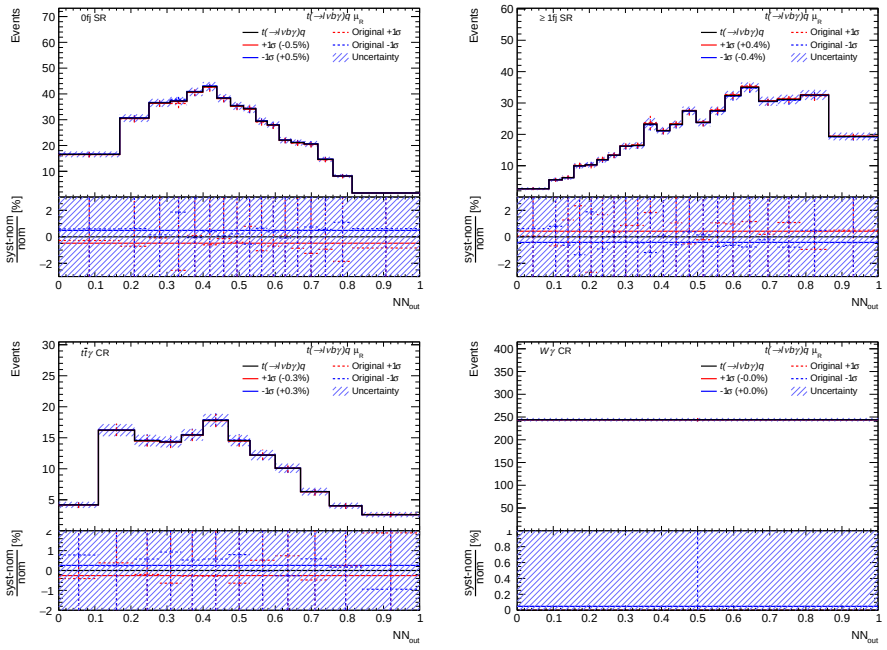


Figure 81: Impact of the μ_{R} variation on the prediction of the $\mathrm{NN}_{\mathrm{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

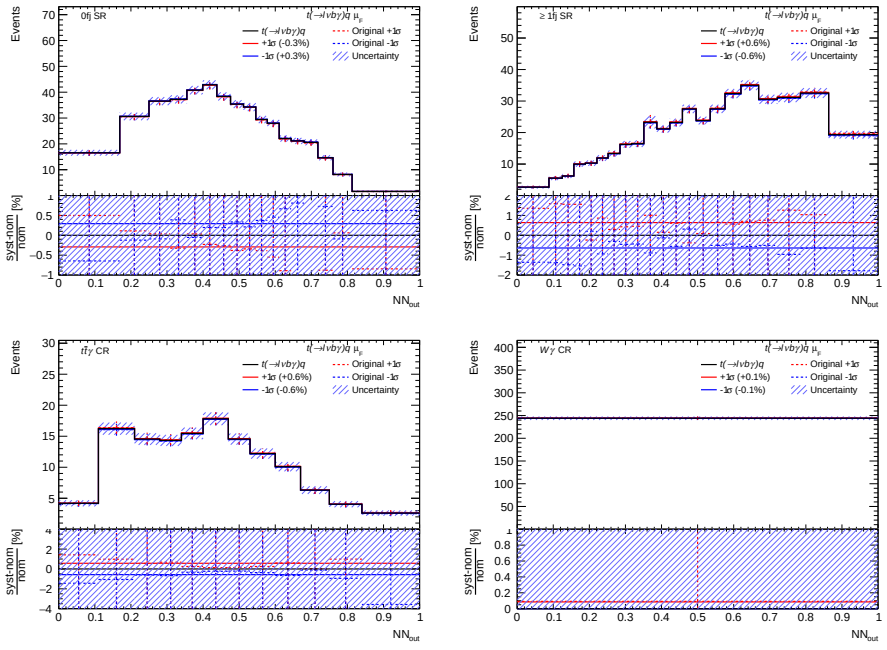


Figure 82: Impact of the μ_{F} variation on the prediction of the $\mathrm{NN}_{\mathrm{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

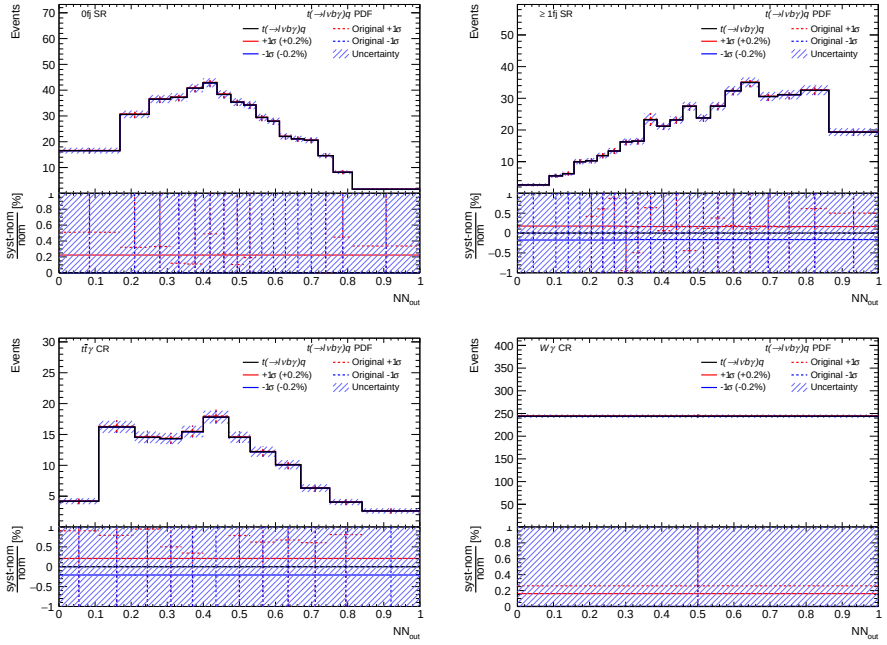


Figure 83: Impact of the PDF uncertainty on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the parton-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

F.4 Modeling uncertainties of the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ predictions for the particle-level measurement

In the following, the impacts of the NPs affecting the $tq\gamma$ and $t\,(\to \ell\nu b\gamma)\,q$ predictions are displayed for the particle-level measurement.

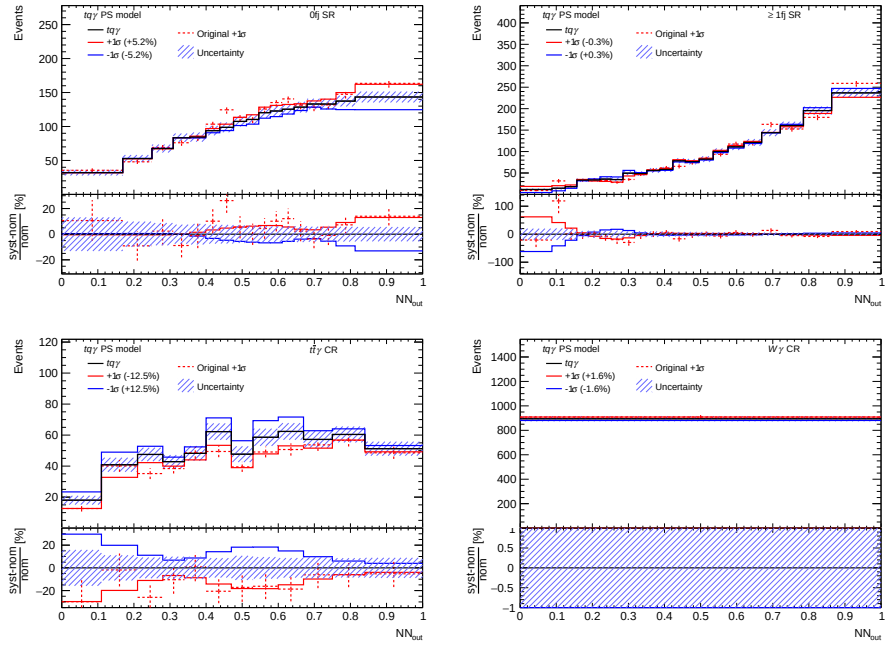


Figure 84: Impact of the variation of the PS model variation on the prediction of the $\mathrm{NN_{out}}$ distributions of $tq\gamma$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

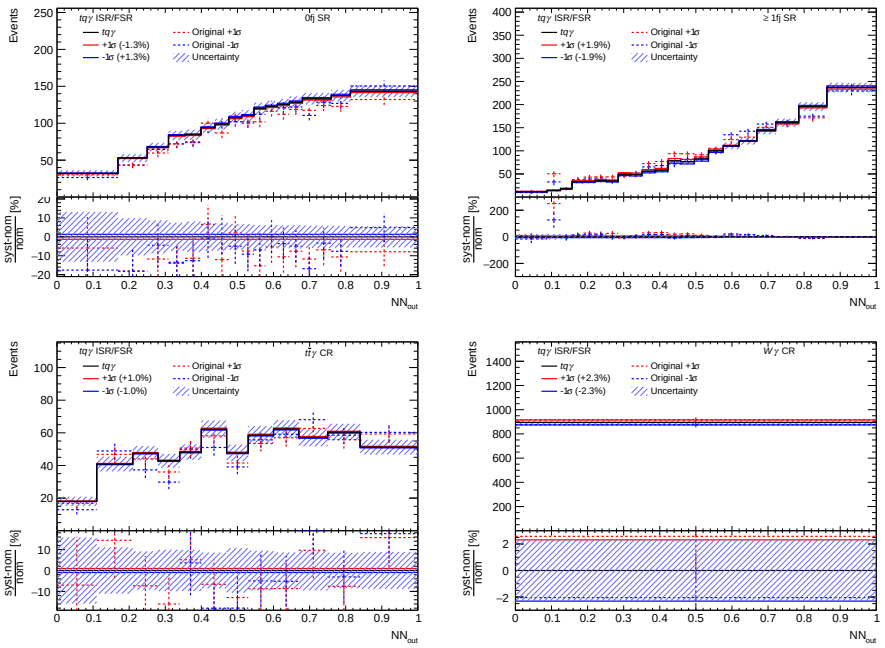


Figure 85: Impact of the Var3c variation (ISR/FSR) on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

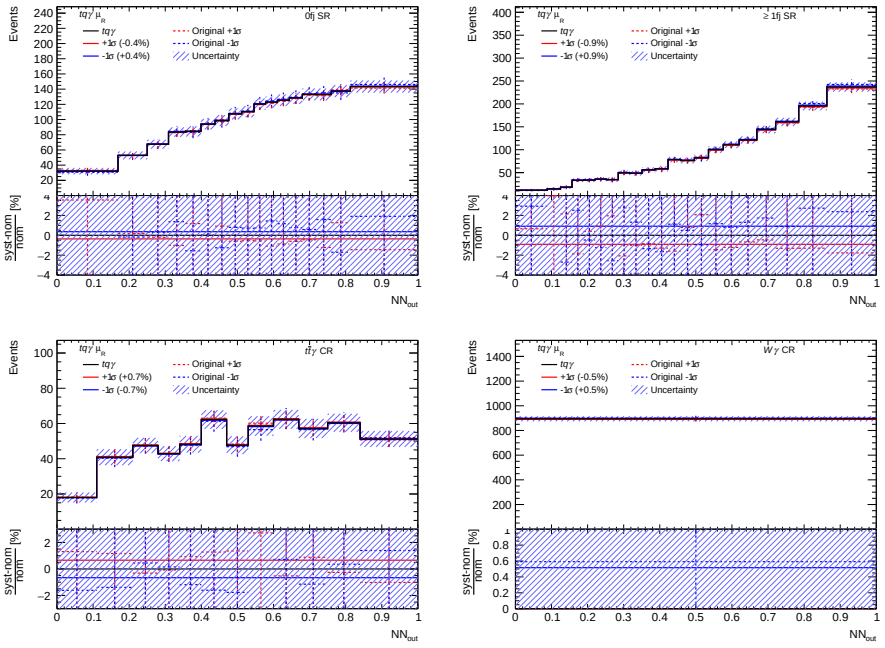


Figure 86: Impact of the μ_R variation on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

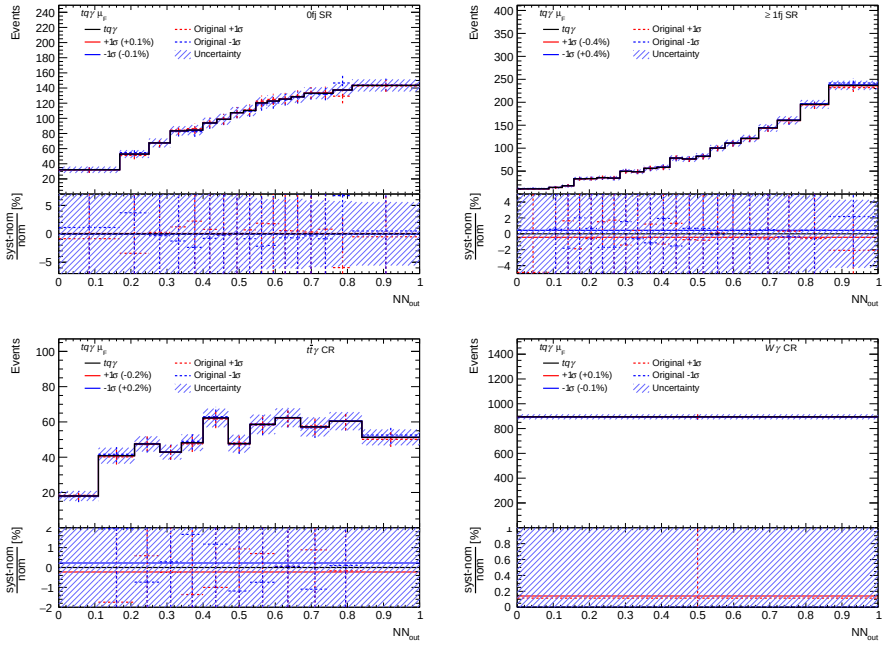


Figure 87: Impact of the μ_F variation on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

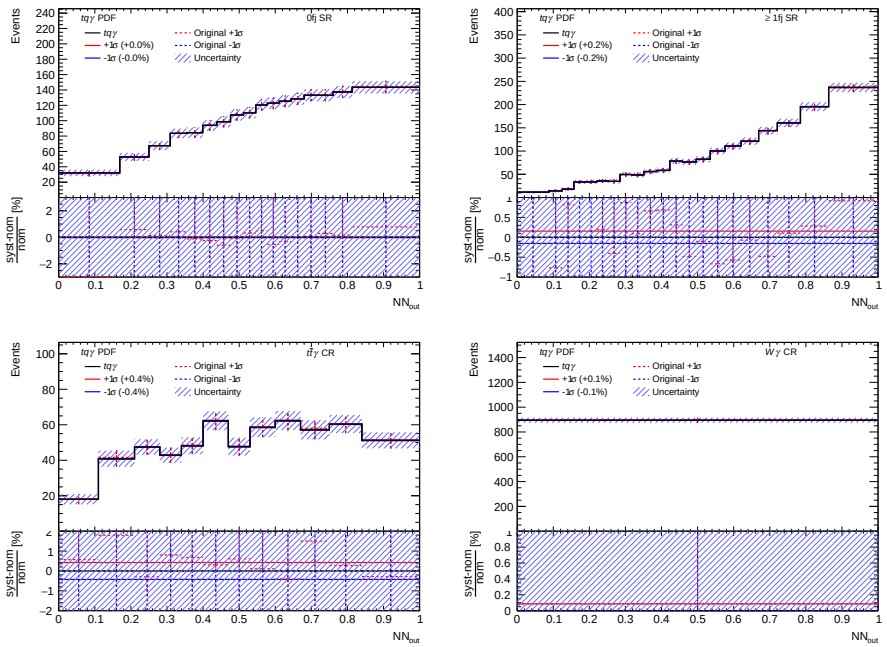


Figure 88: Impact of the PDF uncertainty on the prediction of the NN_{out} distributions of $tq\gamma$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

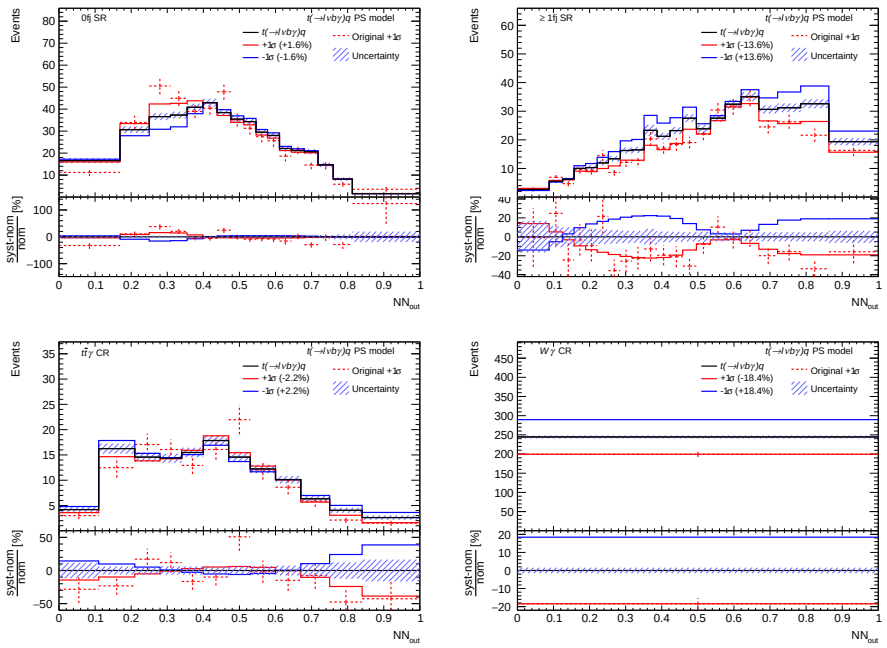


Figure 89: Impact of the variation of the PS model variation on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

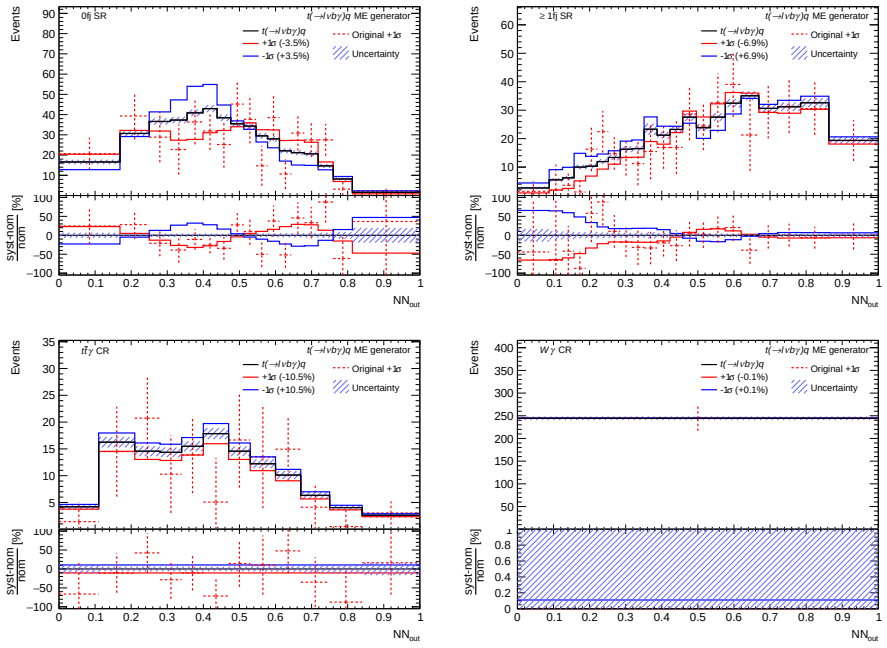


Figure 90: Impact of the variation of the ME generator variation on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

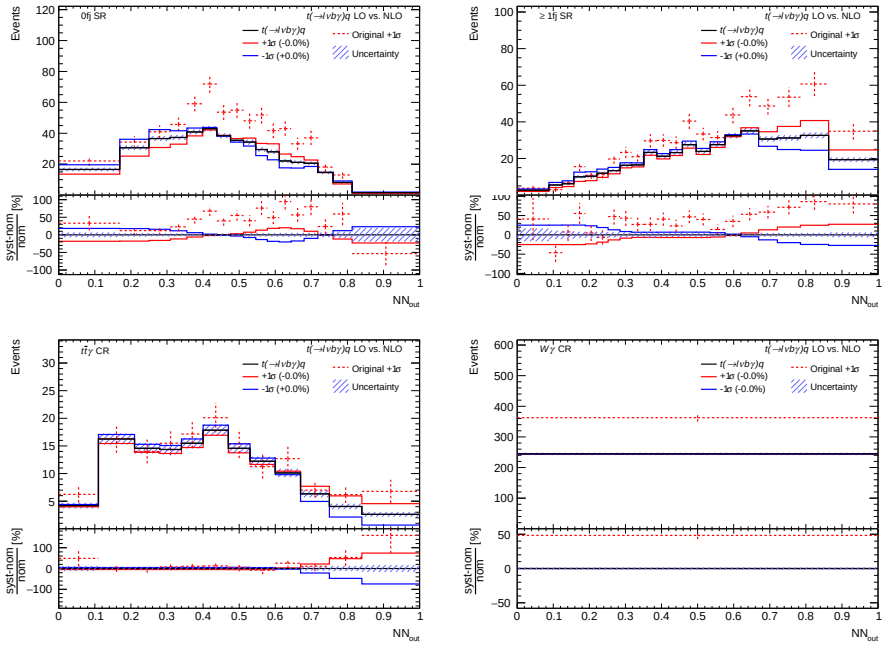


Figure 91: Impact of the "$t\,(\to \ell\nu b\gamma)\,q$ LO vs. NLO" NP on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the prediction as obtained from the alternative sample. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

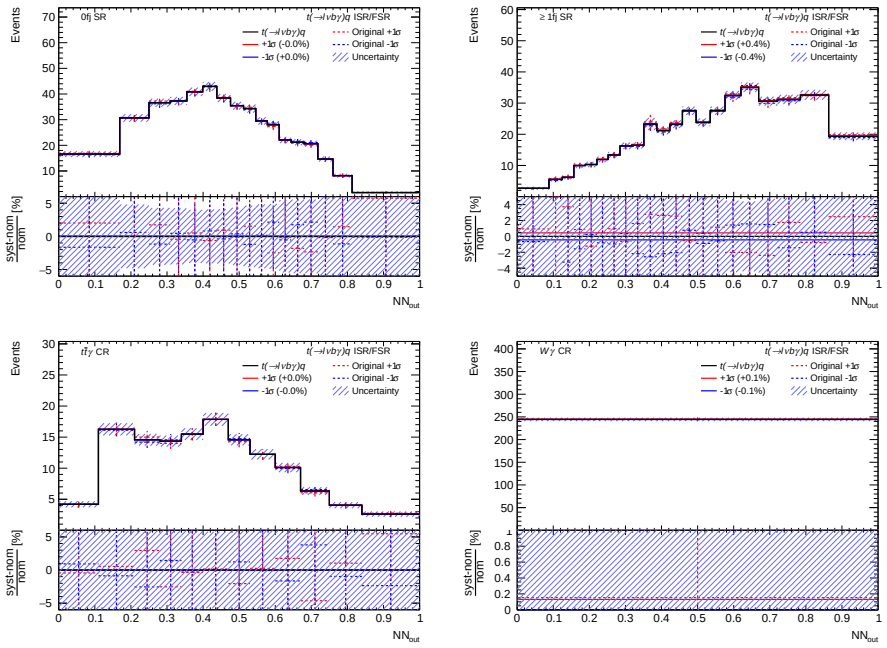


Figure 92: Impact of the Var3c variation (ISR/FSR) on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm 1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm 1\sigma$ variations to the nominal prediction in terms of relative differences.

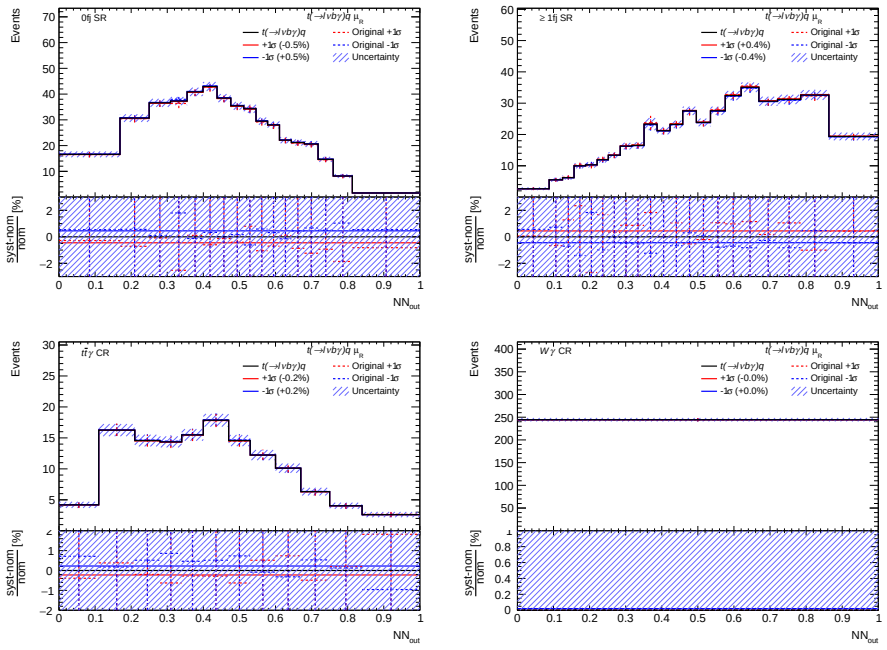


Figure 93: Impact of the μ_{R} variation on the prediction of the $\mathrm{NN}_{\mathrm{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

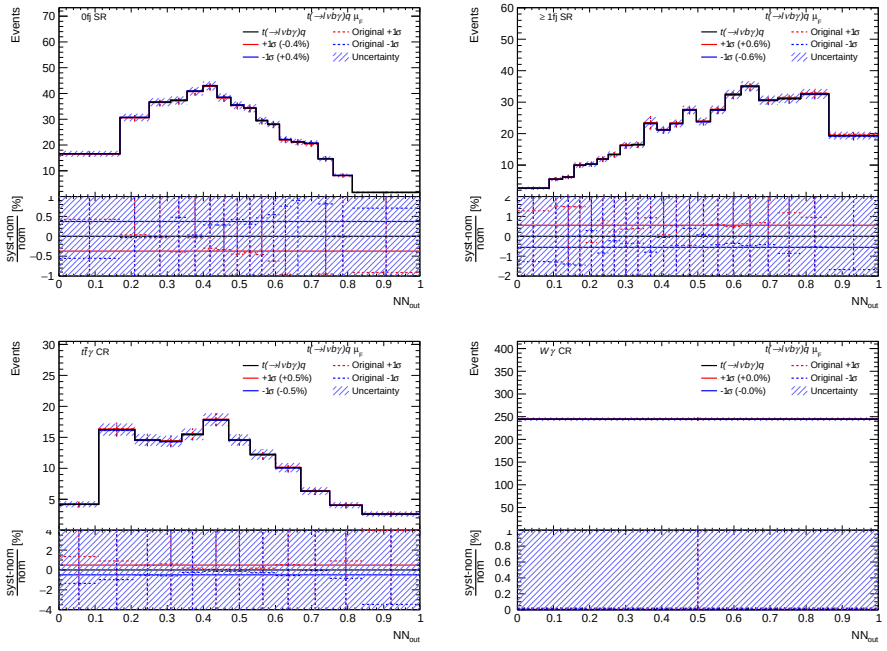


Figure 94: Impact of the μ_{F} variation on the prediction of the $\mathrm{NN}_{\mathrm{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

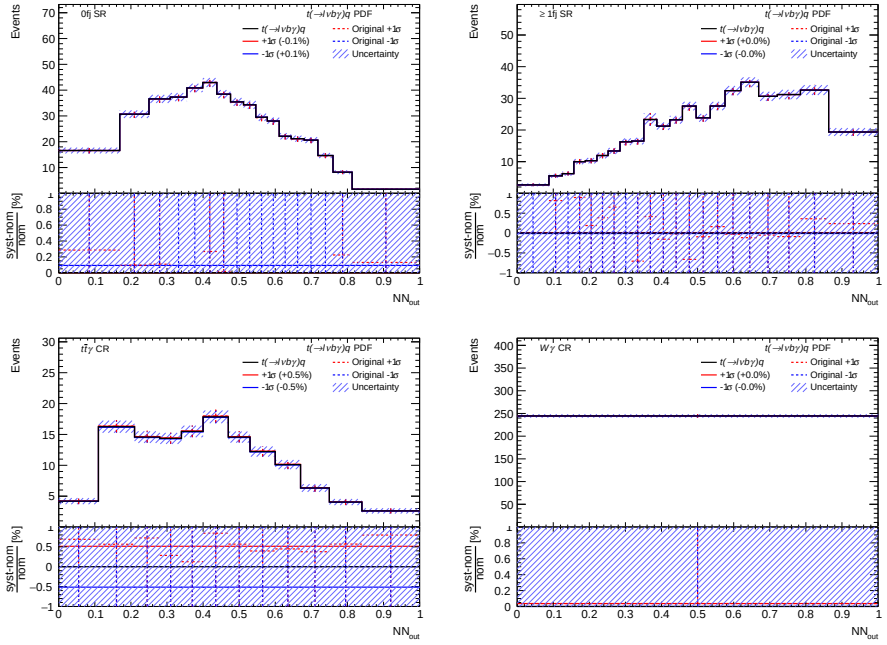


Figure 95: Impact of the PDF uncertainty on the prediction of the $\mathrm{NN_{out}}$ distributions of $t\,(\to \ell\nu b\gamma)\,q$ in all ARs for the particle-level measurement. The solid black line in the upper panel represents the nominal prediction, while the dashed red line corresponds to the original $+1\sigma$ variation and the blue line corresponds to the original -1σ variation. The solid red and blue lines correspond to the $\pm1\sigma$ variations after applying symmetrization and smoothing to the varied distributions. The error bars represent the uncertainty in the predictions due to the limited sample size. The ratio panel compare the $\pm1\sigma$ variations to the nominal prediction in terms of relative differences.

F.5 Impact of the uncertainties of the data-driven background estimations

In the following, the impacts of the NPs corresponding to the uncertainties of the data-driven background estimations on the overall prediction are displayed. The impacts are the same for the parton-level and particle-level measurements.

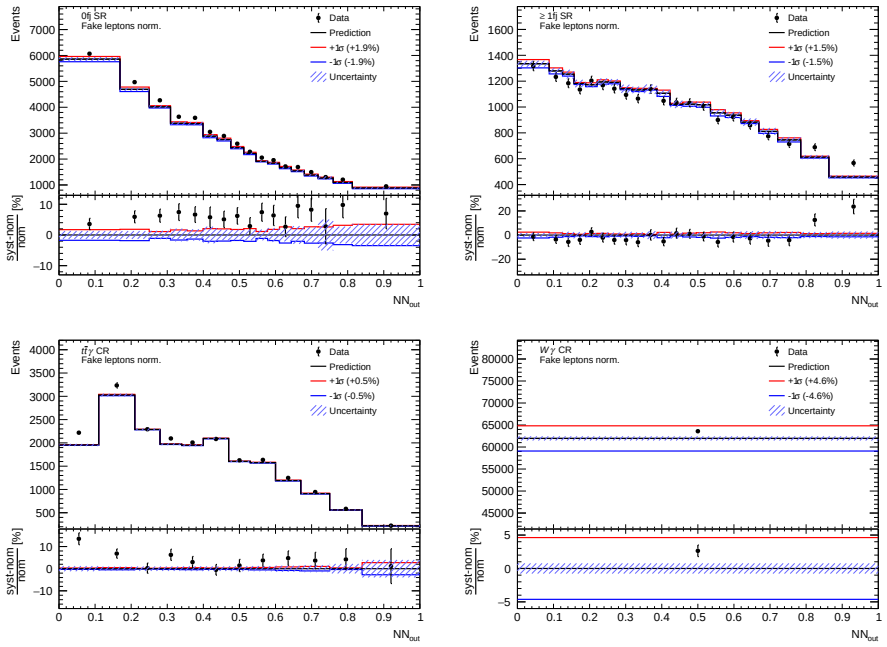


Figure 96: Impact of the "Fake leptons norm." NP (cf. Section 8.1) on the overall prediction of the NN_{out} distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

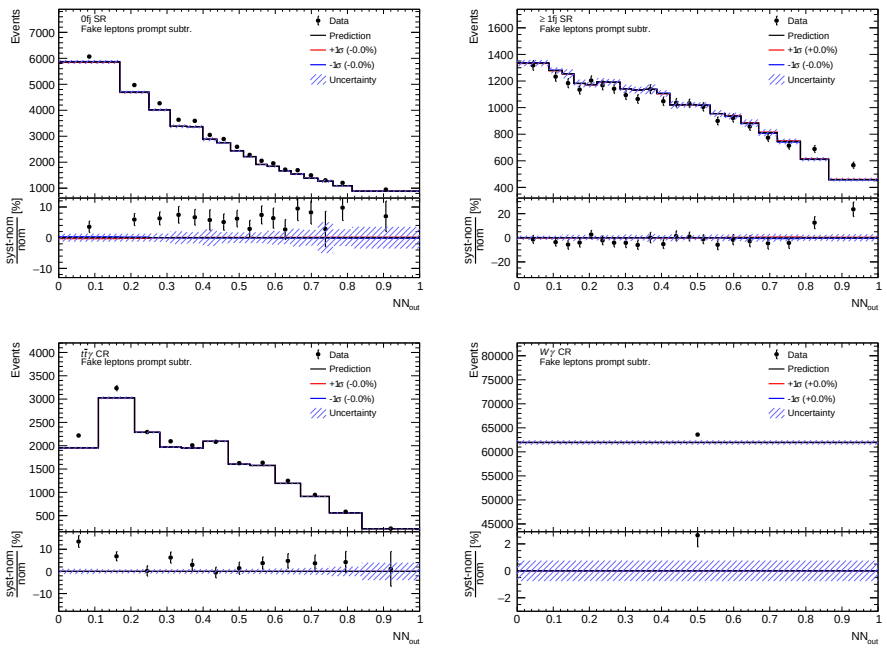


Figure 97: Impact of the "Fake leptons prompt subtr." NP (cf. Section 8.1) on the overall prediction of the NN_{out} distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

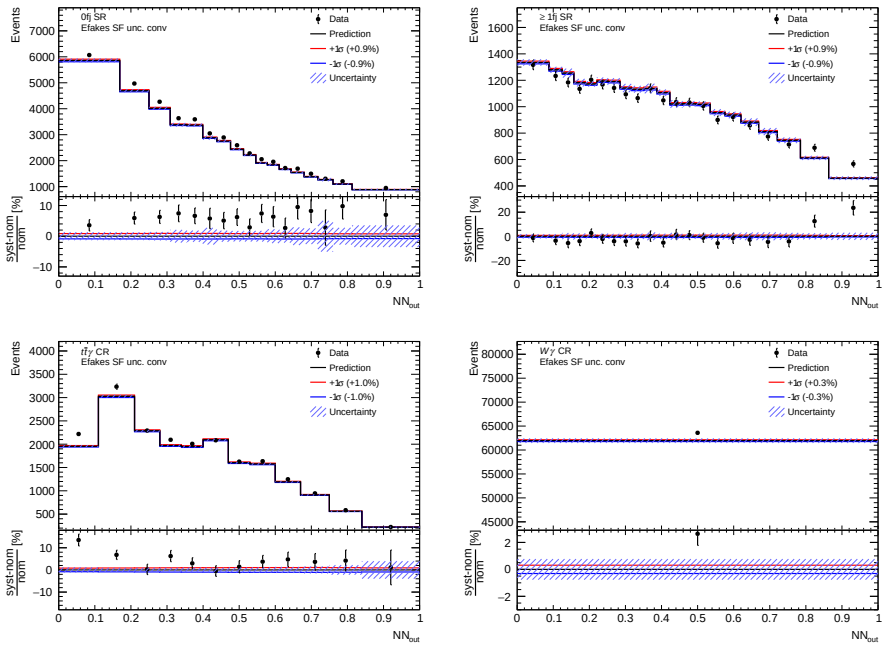


Figure 98: Impact of the uncertainty of the $e \to \gamma$ SF for converted reconstructed $e \to \gamma$ photons on the overall prediction of the NN_{out} distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

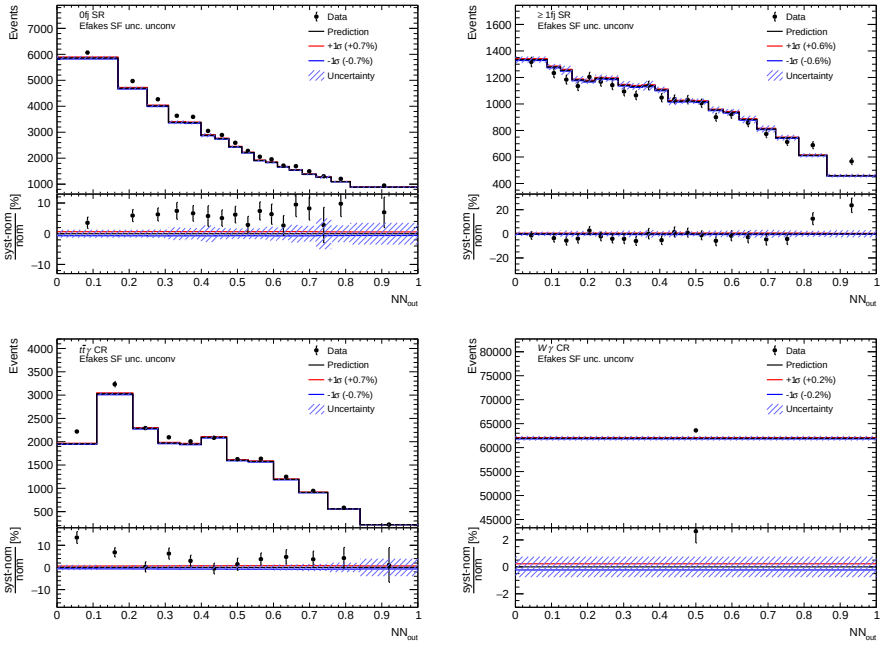


Figure 99: Impact of the uncertainty of the $e \to \gamma$ SF for unconverted reconstructed $e \to \gamma$ photons on the overall prediction of the $\mathrm{NN_{out}}$ distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

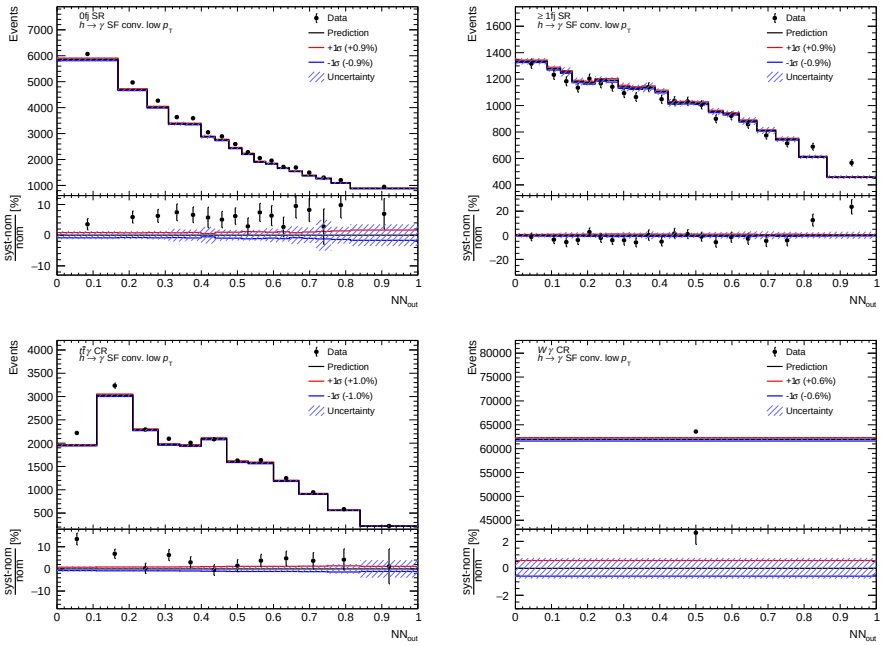


Figure 100: Impact of the uncertainty of the $h \to \gamma$ SF for converted reconstructed $h \to \gamma$ photons with $20\,\mathrm{GeV} < p_\mathrm{T} < 40\,\mathrm{GeV}$ on the overall prediction of the $\mathrm{NN_{out}}$ distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

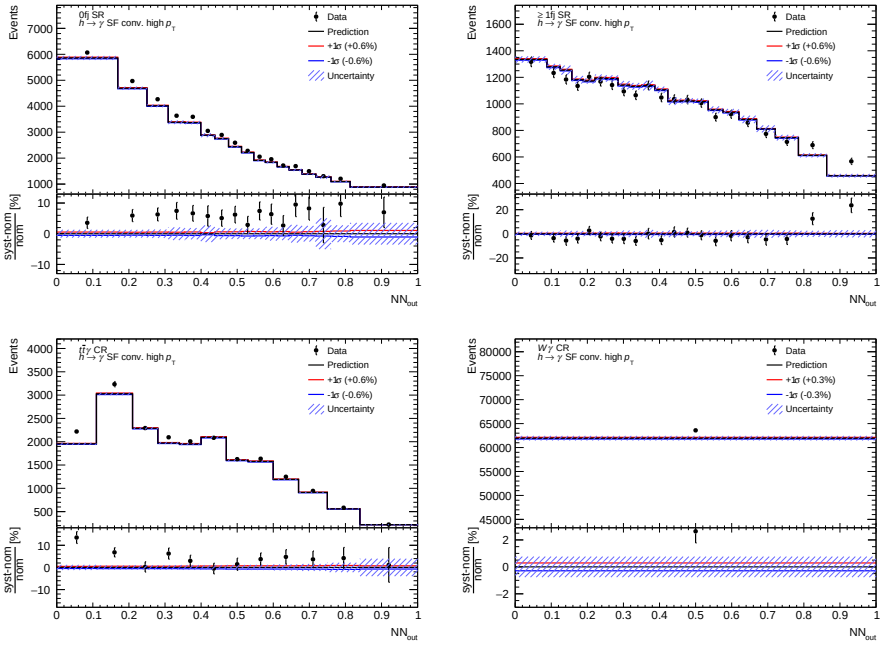


Figure 101: Impact of the uncertainty of the $h \to \gamma$ SF for converted reconstructed $h \to \gamma$ photons with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ on the overall prediction of the $\mathrm{NN}_{\mathrm{out}}$ distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

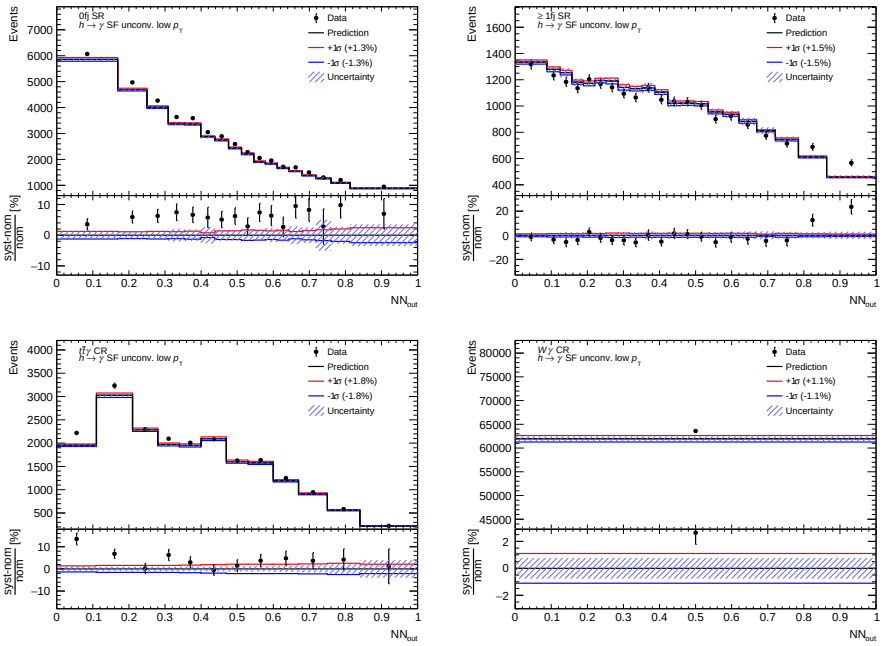


Figure 102: Impact of the uncertainty of the $h \to \gamma$ SF for unconverted reconstructed $h \to \gamma$ photons with $20\,\mathrm{GeV} < p_{\mathrm{T}} < 40\,\mathrm{GeV}$ on the overall prediction of the $\mathrm{NN}_{\mathrm{out}}$ distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.

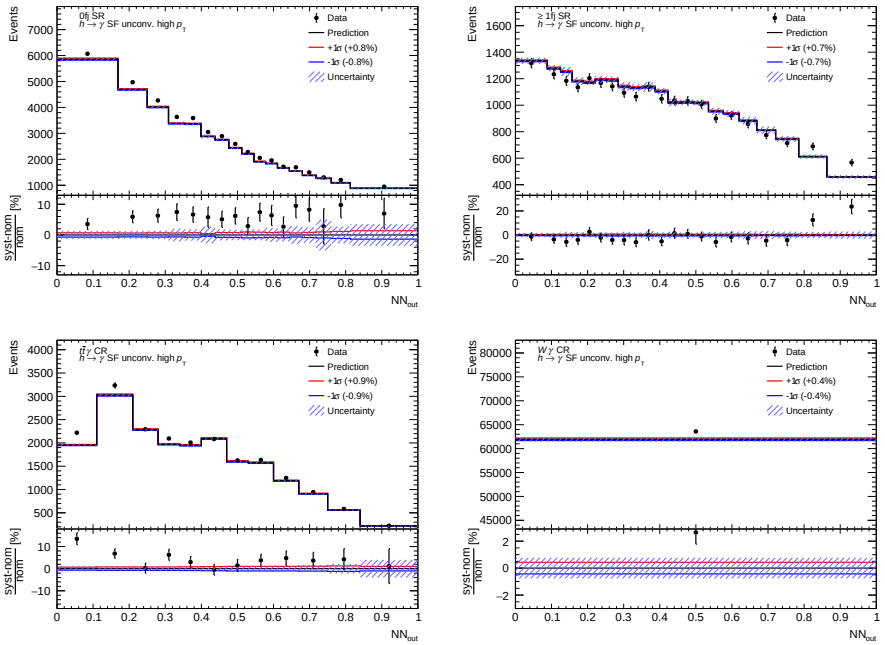


Figure 103: Impact of the uncertainty of the $h \to \gamma$ SF for unconverted reconstructed $h \to \gamma$ photons with $p_T > 40\,\text{GeV}$ on the overall prediction of the NN_{out} distributions in all ARs for the parton-level measurement. The black line ("Prediction") in the upper panel represents the nominal prediction, while the red and blue liness correspond to the $\pm 1\sigma$ variations. Furthermore, the observed distributions are shown ("Data"). The ratio panels compare the $\pm 1\sigma$ variations to the nominal predictions, and the relative difference of the nominal predictions to the observed distributions in terms of relative differences. The blue hashed bands represent the uncertainty in the predictions due to the limited sample size.