technische universität
dortmund

Search for Higgs boson production in association with a single top quark and constraints on the H+b-jets background in the $H \to \gamma\gamma$ decay channel at the ATLAS experiment

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
vorlegt von:

Isabel Nitsche
geboren in Geseke

Lehrstuhl für Experimentelle Physik IV
Fakultät Physik
Technische Universität Dortmund

– 2019 –

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:	Prof. Dr. Dr. Wolfgang Rhode
Vorsitzender der Prüfungskommission:	Prof. Dr. Matthias Schneider
Vertretung der wiss. Mitarbeiter:	Dr. Carsten Nase
Datum des Einreichens der Arbeit:	7. Juni 2019
Datum der mündlichen Prüfung:	29. August 2019

Kurzfassung

In dieser Dissertation wird die Produktion eines einzelnen Top-Quarks in Verbindung mit einem Higgs-Boson (tH) im $H \to \gamma\gamma$-Zerfallskanal bei $\sqrt{s} = 13\,\text{TeV}$ am ATLAS-Experiment untersucht. Neben der Produktion eines Higgs-Bosons in Verbindung mit einem Top-Quark-Paar ($t\bar{t}H$) ermöglicht die tH-Produktion eine direkte Messung der Yukawa-Kopplung des Top-Quarks Y_t. Aufgrund einer destruktiven Interferenz in der tH-Produktion ist dieser Prozess zusätzlich sensitiv auf negative Werte des Kopplungsstärken-Modifikators $\kappa_t = Y_t/Y_t^{SM}$, welcher Abweichungen vom erwarteten Wert im Standardmodell (SM) Y_t^{SM} beschreibt. In der hier vorgestellten Messung der $t\bar{t}H$+tH-Produktion im $H \to \gamma\gamma$-Kanal mit einem Datensatz von $36.1\,\text{fb}^{-1}$ wurden am ATLAS-Experiment erstmals optimierte tH-Kategorien berücksichtigt. Die gemessene Signalstärke für die $t\bar{t}H$+tH-Produktion stimmt mit der SM-Erwartung überein und es wurde kein Hinweis auf ein negatives Vorzeichen von κ_t gefunden.

Die Produktion eines Higgs-Bosons in Verbindung mit zusätzlichen b-jets (H+b-jets) ist ein wichtiger Untergrund für hadronische tH- und $t\bar{t}H$-Endzustände. Dieser Untergrund ist mit einer großen systematischen Unsicherheit verbunden, da nicht bekannt ist wie zutreffend die Vorhersagen aus Monte-Carlo-Simulationen sind. Eine Messung des differentiellen Wirkungsquerschnittes (WQ) der Anzahl von b-jets, $N_{b\text{-jets}}$, für inklusive Higgs-Boson-Produktion im $H \to \gamma\gamma$-Zerfallskanal mit einem Datensatz von $79.8\,\text{fb}^{-1}$ wird präsentiert. Der dominante Beitrag für Ereignisse mit $N_{b\text{-jets}} = 1$ resultiert dabei aus der H+b-jets-Produktion. Der gemessene WQ für $N_{b\text{-jets}} = 1$ stimmt innerhalb der Unsicherheiten mit der Monte-Carlo-Vorhersage überein. Zusätzlich wird ein Ansatz zur Erhöhung der Sensitivität auf H+b-jets präsentiert, der auf einer Trennung von H+b-jets und dem nicht-resonanten Untergrund mittels multivariater Analysemethoden basiert.

Abstract

In this thesis, the production of a single top quark in association with a Higgs boson (tH) is studied at $\sqrt{s} = 13\,\text{TeV}$ in the $H \to \gamma\gamma$ decay channel at the ATLAS experiment. Similar to the production of a Higgs boson in association with a top-antitop quark pair ($t\bar{t}H$), tH production allows for a direct measurement of the Yukawa coupling, Y_t, of the top quark. Due to a destructive interference in tH production, it also provides sensitivity to negative values of the coupling strength modifier $\kappa_t = Y_t/Y_t^{SM}$, which describes deviations from the Standard Model (SM) expectation Y_t^{SM}. For the first time at the ATLAS experiment, dedicated tH categories are included in the measurement of top-quark-associated Higgs boson production in the $H \to \gamma\gamma$ channel, using $36.1\,\text{fb}^{-1}$ of data. The measured signal strength for $t\bar{t}H$+tH production agrees with the SM expectation and no hint for a negative sign of κ_t was found.

An important background for hadronic $t\bar{t}H$ and tH final states in the $H \to \gamma\gamma$ channel is the production of a Higgs boson in association with additional b-jets (H+b-jets). This background is associated with a large systematic uncertainty, as the accuracy of the prediction from Monte Carlo (MC) simulations is not known. A measurement of the differential cross section of the number of b-jets, $N_{b\text{-jets}}$, for inclusive Higgs boson production in the $H \to \gamma\gamma$ decay channel using $79.8\,\text{fb}^{-1}$ of ATLAS data is presented. The measured cross section for $N_{b\text{-jets}} = 1$, with the dominant contribution resulting from H+b-jets production, is in agreement with the MC expectation within the uncertainties. Additionally, an approach to improve the sensitivity to H+b-jets is presented, which is based on multivariate analysis techniques to discriminate between H+b-jets and the non-resonant background.

Contents

Preamble

The measurements presented in this thesis were performed in collaborative work within the ATLAS experiment. Due to the complexity of these measurements, the final results include contributions from several members of the ATLAS collaboration. My specific contributions to the different measurements are listed in the following:

- For the measurement of $t\bar{t}H$ production in the $H \to \gamma\gamma$ channel using $13.3\,\text{fb}^{-1}$ of ATLAS data [1], I worked on the development and validation of a strategy to derive a background model in the $t\bar{t}H$ categories, based on data-driven background templates. Based on this strategy, I derived the background model and the associated systematic uncertainty for the $t\bar{t}H$ categories.
- For the measurement of $t\bar{t}H$+tH production in the $H \to \gamma\gamma$ channel using $36.1\,\text{fb}^{-1}$ of ATLAS data [2], I worked on the development of an analysis strategy for a measurement of tH production and the optimization of the tH categories, which are the first dedicated tH categories that were considered in an ATLAS measurement. In addition, I derived the background model and the associated systematic uncertainty for both, the $t\bar{t}H$ and tH categories.
- For the measurement of $t\bar{t}H(H \to \gamma\gamma)$ production using $79.8\,\text{fb}^{-1}$ of ATLAS data [3], which contributed to the observation of $t\bar{t}H$ production at the ATLAS experiment, I derived the background model and the associated systematic uncertainty for the $t\bar{t}H$ categories.
- For the differential cross section measurement of $N_{b\text{-jets}}$ in the $H \to \gamma\gamma$ decay channel using $79.8\,\text{fb}^{-1}$ of ATLAS data [4], I performed sensitivity studies for a first measurement of Higgs boson production in association with additional b-jets in the $H \to \gamma\gamma$ channel and developed an analysis strategy. I worked on the choice of binning for the $N_{b\text{-jets}}$ distribution and the definition of the event selection on reconstruction and particle level. Furthermore, I performed the signal extraction fit for the different $N_{b\text{-jets}}$ bins and derived the unfolding uncertainty due to the parton shower modeling.
- I performed a measurement of the *tight* photon ID efficiency with the Matrix Method using $79.8\,\text{fb}^{-1}$ of ATLAS data, which is based on the previous measurements performed in Run 1 [5] and Run 2 [6].

List of abbreviations

4FS four-flavor scheme

5FS five-flavor scheme

BDT Boosted Decision Tree

B-only background-only

BSM Beyond the Standard Model

CERN European Centre of Nuclear Research

CKM Cabibbo-Kobayashi-Maskawa

CL confidence level

CSC Cathode-Strip Chamber

DCB double-sided Crystal Ball

EM electromagnetic

EW electroweak

FCal forward calorimeter

FF fudge factor

HL-LHC High-Luminosity LHC

HLT High Level Trigger

IBL Insertable B-Layer

ID identification

InDet Inner Detector

JER jet energy resolution

JES jet energy scale

JVT jet vertex tagger

L1 trigger Level-1 Trigger

LAr Liquid Argon

LHC Large Hadron Collider

LO leading-order

MC Monte Carlo

MDT Monitored Drift Tube

MIP minimum ionizing particles

MS Muon Spectrometer

MVA multivariate analysis

NLO next-to-leading-order

N^{3}LO next-to-next-to-next-to-leading-order

NN neural network

NNLO next-to-next-to-leading-order

NP nuisance parameter

***NTNI* events** events with non-tight and/or non-isolated photons

PDF parton distribution function

p.d.f. probability density function

PER photon energy resolution

PES photon energy scale

PoI parameter of interest

pp proton-proton

PS Proton Synchrotron

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

RoI regions of interest

ROS Read-Out-System

RPC Resistive Plate Chamber

S+B signal+background

SCT Semiconductor Tracker

SF scale factor

SM Standard Model

SPS Super Proton Synchrotron

TDAQ Trigger and Data Acquisition

TGC Thin Gap Chamber

***TI* events** events with tight and isolated photons

TRT Transition Radiation Tracker

WP working point

Chapter 1

Introduction

With the observation of the Higgs boson in 2012 [7, 8], the important, by then, unresolved question of the origin of mass of the elementary particles in the Standard Model (SM) of particle physics was answered. The elementary particles, which are the constituents of matter in our universe, and their interactions are described by the SM. Although the SM has been very successful in predicting the experimental outcome of numerous particle physics experiments, a few questions remain unanswered until now. Among these are the integration of the gravitational force into the SM, the matter-antimatter asymmetry in our universe, the origin of dark matter [9] and dark energy [10, 11] and the hierarchy problem [12]. The Higgs boson was predicted by Peter Higgs, Francois Englert and Robert Brout in 1964 as part of the Higgs mechanism [13–15]. In their theory, the Higgs boson is an excitation of the Higgs field, whose interaction with elementary particles during electroweak symmetry breaking gives mass to these particles. After many decades of searching for the Higgs boson, the observation of a particle with properties consistent with those predicted by Higgs, Englert and Brout and a mass of $m_H = 125.09\,\text{GeV}$ [16] at the Large Hadron Collider (LHC) was one of the greatest achievements of the LHC and the ATLAS and CMS experiments. This was the confirmation of the theory by Higgs, Englert and Brout[1], for which Higgs and Englert were granted the Nobel prize in 2013 [17].

After its observation, the measurement of the properties of the Higgs boson became one of the main goals of the ATLAS and CMS experiments. The production of Higgs bosons at the LHC mainly occurs via the fusion of two gluons (ggH), hence, this production process allows for the most precise measurements of the properties of the Higgs boson until now. As the Higgs boson does not directly couple to massless particles, like gluons or photons, processes like ggH production and the Higgs boson decay into two photons ($H \to \gamma\gamma$) are only possible through loop diagrams involving vector bosons or heavy quarks. For this reason, ggH production and the $H \to \gamma\gamma$ decay channel allow for the indirect measurement of the Higgs boson's coupling to the particles involved in the loops. For the direct measurement of these couplings, more rare processes need to be investigated, which makes these measurements challenging.

The production of the Higgs boson in association with top quarks, which is studied in this thesis, is one example for these rare processes. As the top quark is the most massive fundamental particle in the SM, it exhibits special properties by which it can be easily discriminated from the other quarks. It was observed first at the Tevatron in 1995 [18] and precision measurements of its properties are another important field of research at the LHC. Due to its large mass, the top quark decays before any hadronization processes occur, which offers the possibility to reconstruct the top quark from its decay products. The top quark almost exclusively decays

[1] Brout passed away in 2011, shortly before the discovery of the Higgs boson.

into a b quark and a W boson. In contrast, the lighter quarks underlie hadronization processes and are measured as bundles of particles in the ATLAS detector, so-called jets. Dedicated algorithms are developed in order to tag jets originating from b quarks, whose identification is an important ingredient for measurements involving top quarks. In the SM, the coupling of the Higgs boson to fermions is predicted to be proportional to their mass. As a consequence, the top quark is expected to have a large coupling to the Higgs boson. Therefore, the measurement of the Yukawa coupling between the top quark and the Higgs boson, Y_t, is of major interest, as deviations from the SM expectation might give hints to new physics. Measurements of the cross section of ggH production in the $H \to \gamma\gamma$ decay channel [4] provide no indication for a deviation of Y_t from the SM expectation. However, in order to make a complete statement about Y_t a direct measurement of this quantity is necessary, as yet unknown particles, predicted by Beyond the Standard Model (BSM) theories, might contribute to the loops in ggH production and the $H \to \gamma\gamma$ decay.

There are two processes at the LHC which allow for the direct measurement of Y_t. One of these processes is the production of the Higgs boson in association with a top quark pair via the strong interaction ($t\bar{t}H$). This process was recently observed at the ATLAS and CMS experiments for the first time [3, 19]. The second process which is sensitive to Y_t is the production of a single top quark in association with a Higgs boson (tH). As this process occurs via the electroweak interaction it is more rare than $t\bar{t}H$ production. In tH production, the Higgs boson can not only be radiated off the top quark but also off the W boson from the production. In the SM, there is a destructive interference of these two processes which gives sensitivity to the relative sign between Y_t and the coupling of the W boson to the Higgs boson. Deviations of Y_t from the SM prediction, including a negative sign of Y_t which cannot be excluded by measurements of $t\bar{t}H$ production alone, would lead to a strong increase of the tH production cross section. This is a unique property of tH production, which makes it an interesting process to study despite its small cross section. A measurement of Higgs boson properties in the $H \to \gamma\gamma$ decay channel was performed using a dataset corresponding to $36.1\,\text{fb}^{-1}$, collected by the ATLAS detector in 2015 and 2016 [2]. In this measurement, dedicated tH categories were included for the first time. A combined measurement of top-quark-associated Higgs boson production was performed including a reinterpretation of the results in terms of Y_t. The optimization of the tH categories and the development of an alternative strategy for choosing a background model in categories with a small number of expected events were performed as part of this thesis.

The dominant uncertainty in the measurement of top-quark-associated Higgs boson production in the $H \to \gamma\gamma$ decay channel is the statistical uncertainty, as the cross sections of $t\bar{t}H$ and tH production are comparably small, as well as the branching ratio of the Higgs boson decay into two photons, which amounts to 0.227% [20]. Nevertheless, one of the main systematic uncertainties for final states without leptons, so-called hadronic final states, originates from a background which has not been measured so far. This background is referred to as H+b-jets and is composed of $H \to \gamma\gamma$ events with additional b-jets in the final state. The main contribution to this background arises from ggH production with additional b quarks in the final state, which can e.g. be produced through the splitting of radiated gluons.

The H+b-jets background is usually estimated using Monte Carlo (MC) simulations, but the accuracy of the MC prediction for this kind of background, where the additional jets mostly originate from the parton shower, is not known. Similar measurements of processes with additional b-jets in the final state, such as $t\bar{t} + b(b)$ production [21, 22], observed discrepancies between MC predictions and data. Therefore, a conservative uncertainty of 100% on the H+b-jets background was assumed for several measurements of hadronic final states in the

$H \to \gamma\gamma$ decay channel. For future measurements with larger datasets, the statistical uncertainty will decrease and a more sophisticated approach for estimating the uncertainty on the H+b-jets background will be needed. A first attempt to constrain the H+b-jets background was performed with the 2015-2017 dataset corresponding to $79.8\,\text{fb}^{-1}$, by measuring the differential cross section of the number of b-jets, $N_{b\text{-jets}}$, for hadronic final states in the $H \to \gamma\gamma$ channel [4], and is presented in this thesis. The sensitivity of this measurement is limited by the large contribution of the non-resonant background. Therefore, an alternative approach is presented which aims to increase the sensitivity to H+b-jets by discriminating the Higgs boson signal from the non-resonant background based on multivariate analysis (MVA) techniques.

Although the Higgs boson decay into two photons is one of the most rare decay channels, it benefits from the precise measurement of the photon energy by the ATLAS detector. Additionally, it is characterized by relatively low backgrounds, as the two photons originating from the Higgs boson decay have a high transverse energy due to the large mass of the Higgs boson. Their invariant mass, $m_{\gamma\gamma}$, can be measured as a narrow peak around the mass of the Higgs boson above the falling spectrum of the non-resonant background, which originates from photons which are not produced during a Higgs boson decay. To summarize, it is one of the most promising Higgs boson decay channels and contributed significantly to the observation of the Higgs boson in 2012.

In order to precisely measure photons with the ATLAS detector and to reduce the background from objects which are misidentified as photons, information from the electromagnetic (EM) calorimeter and the Inner Detector (InDet) is exploited. Specific isolation requirements are applied to make sure that the photons are separated from hadronic signatures and hence do not originate from photons inside a jet (e.g. from $\pi^0 \to \gamma\gamma$ decays). In addition, requirements on the shape of the shower induced in the EM calorimeter are applied to further reduce the contribution of hadronic fakes. A set of *tight* photon identification (ID) requirements is defined, which is applied in measurements targeting final states with photons at the ATLAS experiment to improve the signal-to-background ratio. A measurement of the efficiency of this *tight* photon ID requirements with the so-called Matrix Method is presented, which is used to derive correction factors to account for differences between the *tight* ID efficiencies measured in data and predicted by MC simulations.

This thesis is organized as follows: First, Chapter 2 gives a brief overview of the SM and the properties of the top quark and the Higgs boson. Then, the experimental setup is described in Chapter 3, which includes a description of the LHC and the ATLAS experiment. In the following Chapters 4 and 5, the considered dataset and MC simulations are described and the reconstruction and selection of the different objects is explained. In particular, the requirements for the *tight* photon ID requirements are explained, leading to the description of the *tight* photon ID efficiency measurement with the Matrix Method in Chapter 6. The analysis strategy for measurements in the $H \to \gamma\gamma$ decay channel is explained in Chapter 7 and the associated selection of diphoton candidate events is briefly described in Chapter 8. Finally, the different measurements of Higgs boson properties in the $H \to \gamma\gamma$ decay channel are described: In Chapter 9, the combined search for tH and $t\bar{t}H$ production is presented, Chapter 10 describes the differential cross section measurement of the $N_{b\text{-jets}}$ distribution as an approach to constrain the H+b-jets background and in Chapter 11 the alternative MVA-based approach is explained. A summary and discussion of the results obtained in this thesis is presented in Chapter 12.

Chapter 2

The Higgs boson & the top quark in the Standard Model

2.1 A brief overview of the Standard Model

The Standard Model of particle physics (SM) is the theoretical model representing the current state of knowledge which describes the interactions between the elementary particles via three of the four fundamental forces. In the SM, the phenomena observed in particle physics are described based on quantum field theory. It is a very successful theory as it provided accurate predictions for a huge number of particle physics experiments. The SM describes the interactions between different types of elementary particles: the fermions, which represent matter particles, and bosons, which are mediators of the fundamental forces. Furthermore, the SM includes the description of the Higgs boson, which is a remnant of electroweak symmetry breaking, representing the mechanism which generates the masses of the elementary particles. Figure 2.1 gives a schematic overview of the fundamental particles described by the SM and their properties[1].

In the following, first a brief overview of the SM, including the fundamental interactions and the matter particles, is given. The properties of the heaviest particle in the SM, the top quark, are discussed in more detail. In addition, the Higgs mechanism and the properties of the associated Higgs boson, in particular its production, decay and coupling to other particles, are described. Finally, processes which are sensitive to the coupling between the top quark and the Higgs boson are discussed and a framework to parameterize the cross sections of these processes by means of coupling strength modifiers is introduced.

2.1.1 Fundamental interactions and gauge bosons

The SM is a relativistic, renormalizable quantum field theory in which the fundamental interactions are described as gauge fields under the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group and the fermions are introduced as spinor fields. The Lagrangian density of the SM describes the free states of the fundamental particles and their interactions based on the quantum fields of fermions ψ, the different gauge bosons and the Higgs field ϕ. The mathematical description of the SM contains three of the four fundamental interactions observed in nature: the electromagnetic (EM), the strong and the weak interaction. Each of these interactions

[1] Throughout this thesis, natural units with $\hbar = c = 1$ are commonly used, thus energies, momenta and masses are given in units of electronvolts (eV). Similarly, times and distances can be expressed in terms of $\frac{1}{\text{eV}}$.

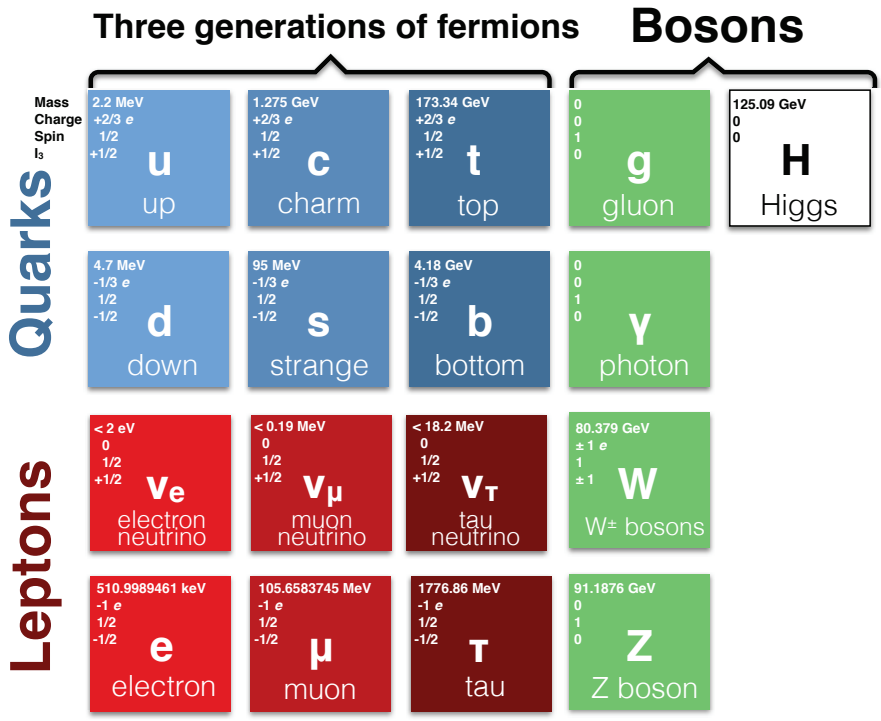


Figure 2.1: A schematic overview of the matter particles and the gauge bosons which are described by the SM. The most recent values for the masses of the particles are taken from Ref. [23], whereas the masses of the top quark and the Higgs boson are taken from Refs. [16, 24]. Additionally, the electric charge, the spin and the third component of the weak isospin I_3 are shown for each particle.

is described through the exchange of gauge bosons, which couple to particles with distinct quantum numbers. Depending on the type of interaction, the gauge bosons may couple to matter particles or also other bosons. The fundamental interactions exhibit characteristic, energy dependent coupling strengths, which are referred to as *running* coupling constants. However, the SM does not include a description of the gravitational force as its formulation within quantum field theory did not succeed until now. The inclusion of gravity into the SM would involve a new hypothetical gauge boson, the graviton, which has not been observed so far. However, the recent observation of gravitational waves allowed to set limits on the mass of the hypothetical graviton [25]. Due to its small coupling to elementary particles, the absence of gravity in the SM is assumed to not significantly affect the validity of its predictions.

The strong interaction is described by Quantum Chromodynamics (QCD) [26–28] based on the $SU(3)_C$ symmetry. It describes interactions between particles carrying a color charge and is mediated by gluons (g), which appear in 8 configurations. The color charge can adopt the three states red, green and blue. Gluons are massless and electrically uncharged bosons with spin 0, which themselves carry a color and an anti-color, enabling gluon self-interactions. The coupling constant of QCD interactions, α_s, is highly energy dependent. Two important effects of QCD, the *confinement* and *asymptotic freedom*, are consequences of the behavior of α_s. The coupling constant increases for large distances and is small for short distances. Similarly, the coupling constant gets smaller for increasing energy scales, while it increases for smaller energy scales. Due to the confinement, color-charged particles (quarks and gluons) are not observed as

free particles in nature and only appear in color-neutral bound states. For increasing distances between color-charged particles, α_s becomes large so that additional color-charged particles are produced through vacuum fluctuations to adopt an energetically lower state. On the other hand, at short distances the coupling constant becomes small and color-charged particles are considered to be asymptotically free.

The EM interaction, which acts on electrically charged particles, is described by Quantum Electrodynamics (QED) based on the $U(1)_{em}$ symmetry [29]. It is mediated by the photon (γ), which couples to the electric charge. The photon itself is a massless and electrically uncharged boson with spin 1. The weak interaction is mediated by the $W^{\pm}$ bosons[2] with charge $\pm\,1\,e$ and the neutral Z boson, which couple to the weak isospin. The W and Z bosons are the only massive gauge bosons in the SM. They interact with fermions either through charged currents, which involve W bosons and represent the only interaction being able to cause transitions between the different quark and lepton flavors, or neutral currents, which are induced by the Z boson. Similar to QCD, self-interactions of the gauge bosons are possible in the weak interaction. The currents of the weak interaction are characterized by their V-A (vector minus axial-vector) structure. As a consequence, only left-handed fermions and right-handed anti-fermions participate in the weak interaction, leading to a violation of the charge conjugation and parity symmetries.

The weak and the EM interaction can be unified into the electroweak (EW) interaction [30–32], based on the combined $SU(2)_L \times U(1)_Y$ symmetry. In the EW theory, the gauge fields of the $SU(2)_L$ component are described by massless vector fields W^a_μ (a = 1,2,3) which couple to the weak isospin, while the massless gauge field of the $U(1)_Y$ component is defined as B_μ and relates to the weak hypercharge. The weak hypercharge $Y = 2(Q - I_3)$ is defined based on the electric charge Q and the third component of the weak isospin I_3. In the EW theory, the W^a_μ bosons acquire their masses through interactions with the Higgs field during EW symmetry breaking, as explained in Section 2.2. The physical W, Z and γ bosons, which are observed in nature, are finally produced through mixing of the W^a_μ and B_μ gauge fields. As a consequence of the mixing, the coupling of the Z boson is also characterized by a component relating to the electric charge.

2.1.2 Matter particles

The matter in our universe is built up of fermions (f), which can be separated into two kinds of particles: quarks (q) and leptons (ℓ). The matter particles have a spin of $\frac{1}{2}$ and appear in three generations, which are characterized by increasing masses. Only the lightest fermions in the SM are stable particles. Among these are the up (u) and down (d) quark, which are the constituents of protons and neutrons and represent the first quark generation, and electrons, which e.g. appear in atomic orbitals. The fermions described in this section are accompanied by their anti-particles[3], which are characterized by the same mass and spin, but opposite sign additive quantum numbers, such as the electric charge, the weak isospin or flavor quantum numbers.

In general, quarks appear in six flavors. In addition to the u and d quarks, the second quark generation includes the charm (c) and strange (s) quark, while the third generation contains

[2] For simplicity reasons, the notation W boson refers to both, the W^+ and W^- bosons.

[3] An antiparticle of a fermion f is usually denoted as $\bar{f}$. Unless specified otherwise, in this thesis the notation f refers to both, the particle and the antiparticle.

the top (t) and the bottom (b) quark. The quark generations are each composed of an up-type and a down-type quark, carrying an electric charge of $2/3\,e$ and $-1/3\,e$, respectively. In the SM, left-handed quarks build up weak isospin doublets $(u, d')_L$, $(c, s')_L$ and $(t, b')_L$ with a third component of the weak isospin I_3 of $+\frac{1}{2}$ and $-\frac{1}{2}$, respectively, whereas the right-handed components are singlets with $I_3 = 0$. Transitions between the different quark flavors can occur through the exchange of a W boson, where the transition probabilities are described by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix [33, 34]. The weak eigenstates (d', s', b') are interpreted as a mixture of the mass eigenstates (d, s, b):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.1}$$

The elements of the CKM matrix, V_{ij}, describe the probability of the transition of a quark with flavor i into a quark with flavor j. Hence, the diagonal elements V_{ud}, V_{cs} and V_{tb} describe the transition between quark states within the same generation. These diagonal elements are close to unity, while the off-diagonal elements, describing the transition between different quark generations, are comparably small.

In contrast to leptons, quarks are color-charged particles which underlie QCD interactions and are not observed as free particles in nature. Due to the confinement, quarks build color-neutral bound states which are referred to as *hadrons*. These hadrons are in general either composed of three quarks with different color charge (qqq or $\bar{q}\bar{q}\bar{q}$) or a quark and an anti-quark ($q\bar{q}$) carrying a color and the respective anti-color. These hadrons are referred to as baryons and mesons, respectively. The proton is a baryon and is composed of the valence quarks uud, which are bound through the exchange of gluons. Additional quarks, so-called *sea* quarks, can be produced by the splitting of gluons inside the proton. Hence, the constituents of the proton include not only the valence quarks, which define its quantum numbers, but also sea quarks and gluons, which are collectively referred to as *partons*. Once a free quark is produced, e.g. through high energy collisions, it underlies hadronization processes and forms color-neutral bound states. Hence, experimentally quarks can only be measured as bundles of particles originating from the hadronization process, so-called *jets*.

The three lepton generations are each composed of a fermion with charge $-1\,e$, which includes electrons (e), muons (μ) and tau leptons (τ), and an associated uncharged neutrino, ν_e, ν_μ and ν_τ. Left-handed leptons form doublets $(\nu_e, e^-)_L$, $(\nu_\mu, \mu^-)_L$ and $(\nu_\tau, \tau^-)_L$ with $I_3 = \pm\frac{1}{2}$, respectively, while the right-handed components are singlets e_R, μ_R, τ_R. There are no right-handed neutrinos in the SM. Leptons carry no color charge and hence, do not interact via the strong interaction. Furthermore, neutrinos only interact via the weak interaction as they do not carry an electric charge. Therefore, neutrino interactions with other particles are rare, which makes their measurement challenging. In the SM, neutrinos are considered to be massless particles, however, the observation of neutrino oscillations [35, 36] has proven that their masses must be different from zero. Similar to the CKM matrix, the mixing of the neutrino states is described by the Pontecorvo-Maki-Nakagawa-Sakata matrix [37, 38]. The masses of the neutrinos have not been measured so far, but experimental boundaries predict these to be very small.

2.2 The Higgs mechanism

Although the W and Z bosons of the weak interaction were experimentally proven to exhibit a mass, the Lagrangian density of the EW interaction does not allow for direct mass terms of the gauge bosons. Hence, the Higgs mechanism [13–15], describing the generation of the W and Z boson masses through EW symmetry breaking induced by the Higgs field, was introduced to resolve this inconsistency in the SM. The observation of the Higgs boson in 2012 [7, 8], which represents an excitation of the Higgs field, provided a proof of the Higgs mechanism. A brief description of the Higgs mechanism is given in the following.

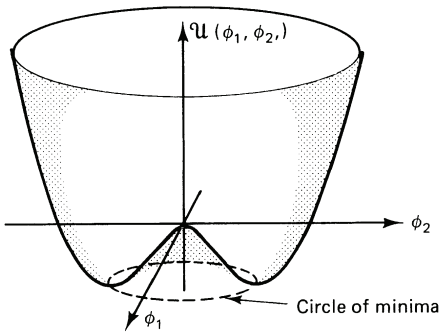


Figure 2.2: Illustration of the potential of the Higgs field which induces EW symmetry breaking [39].

The $SU(2)_L \times U(1)_Y$ symmetry of the EW interaction can be broken by introducing a new complex scalar field of the group $SU(2)_L$, which is referred to as the Higgs field:

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.2}$$

The Lagrangian density of the Higgs field is defined as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \tag{2.3}$$

based on a kinetic term and the potential $V(\phi)$. The kinetic term is defined based on the covariant derivative of the $SU(2)_L \times U(1)_Y$ symmetry group, which introduces couplings between the Higgs field and the gauge fields W_μ^a and B_μ of the EW theory:

$$D_\mu \phi = (\delta_\mu + \frac{i}{2} g \sigma^a W_\mu^a + \frac{i}{2} g' Y B_\mu)\phi, \tag{2.4}$$

with the derivative δ_μ, the Pauli matrices σ^a and the weak hypercharge Y, which is 1 for the Higgs field. Furthermore, the parameters g and g' describe the couplings of the $SU(2)_L$ and the $U(1)_Y$ component, respectively. The potential of the Higgs field is given by

$$V(\phi) = \mu^2(\phi^\dagger \phi) + \lambda(\phi^\dagger \phi)^2 \quad (\lambda > 0), \tag{2.5}$$

where the parameters μ and λ determine the shape of the potential. For $\mu^2 < 0$, the potential has a characteristic shape, which is often referred to as "Mexican hat potential", with an infinite number of minima as illustrated in Figure 2.2. In this scenario, the symmetry is spontaneously broken if one of the ground states of the potential is adopted. The shape of the potential results in an expectation value of the Higgs field which is different from zero:

$$\langle \phi \rangle = \sqrt{-\frac{\mu^2}{2\lambda}} = \frac{\nu}{\sqrt{2}}, \tag{2.6}$$

where ν describes the vacuum expectation value of 246 GeV of the Higgs field. The Higgs field can then be defined by introducing an excitation $H(x)$ around the vacuum expectation value

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{2.7}$$

where H can be identified with the Higgs boson. After EW symmetry breaking, three mass terms proportional to the vacuum expectation of the Higgs field can be identified in the Lagrangian density. Through mixing of the gauge fields W^a_μ and B_μ, the gauge bosons of the weak and EM interaction can be redefined as

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp W^2), \tag{2.8}$$

$$Z = \cos\theta_w W^3 - \sin\theta_w B, \tag{2.9}$$

$$\gamma = \sin\theta_w W^3 + \cos\theta_w B, \tag{2.10}$$

where the parameter θ_w is defined as the weak mixing angle. The masses of the W and Z bosons, which are related through $\cos\theta_w = m_W/m_Z$, are then given by:

$$m_W^2 = \frac{g^2 v^2}{4} \ , \ m_Z^2 = \frac{(g'^2 + g^2)v^2}{4} \tag{2.11}$$

Thus, the overall gauge symmetry of the SM, which is given by $SU(3)_C \times SU(2)_L \times U(1)_Y$, is broken into $SU(3)_C \times U(1)_{em}$. As the $U(1)_{em}$ and the $SU(3)_C$ gauge symmetries are conserved, their gauge bosons, the photons and gluons, remain massless particles. An additional result of the EW symmetry breaking is the existence of a new particle, the Higgs boson, with a mass of $m_H = \sqrt{2\lambda}v$. Finally, the fermions in the SM are also given their masses through a Yukawa coupling term between the fermion fields ψ and the Higgs field, which is defined by the Lagrangian density

$$\mathcal{L}_{\text{Yukawa}} = -\frac{vY_f}{\sqrt{2}}\bar{\psi}\psi - \frac{Y_f}{\sqrt{2}}H\bar{\psi}\psi, \tag{2.12}$$

with $Y_f = \sqrt{2}\frac{m_f}{v}$ describing the Yukawa coupling. While the first term can be identified with the fermion masses of $m_\psi = vY_f/\sqrt{2}$, the second term describes the interactions between the Higgs boson and fermions via the Yukawa coupling. Hence, the origin of the masses of fermions and the massive gauge bosons in the SM can be described through the existence of the Higgs field and the mechanism of EW symmetry breaking.

2.3 Properties of the top quark

The top quark is the elementary particle with the largest mass in the SM and exhibits some special properties in comparison to the lighter quarks. The mass of the top quark has been measured to be $m_t = 173.34 \pm 0.76\,\text{GeV}$ [24], which is more than a factor 40 greater than the mass of the second heaviest quark, the b quark. Top quarks are produced through different processes at the Large Hadron Collider (LHC). The dominant production occurs via QCD interactions based on qq and gg initial states, as illustrated in Figure 2.3a) for a gg initial state, producing pairs of top and anti-top quarks ($t\bar{t}$). In addition, top quarks can be produced singly through the EW interaction, based on the exchange of a W boson. The single top quark production is possible through three processes: the t-channel, the Wt-channel and the s-channel production. The respective Feynman diagrams are shown in Figures 2.3b) - d), respectively. These processes are characterized by smaller cross sections compared to $t\bar{t}$ production, due to the underlying EW interaction. Among the three channels, the single top t-channel production is the dominant process, followed by Wt- and s-channel production, whose cross sections are approximately a factor 3 and 20 smaller than the t-channel cross section [40–42]. The different single top quark production processes were measured at the ATLAS experiment with significances of at least 3σ [43–45] and no deviations from the SM predictions were observed.

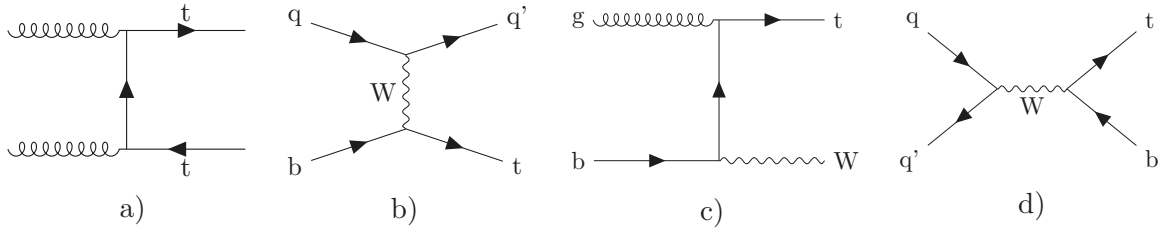


Figure 2.3: Example leading order Feynman diagrams[4] of the different top quark production processes at the LHC: a) top-antitop quark pair production and single top quark production b) in the t-channel, c) in association with a W boson and d) in the s-channel.

In contrast to the other quarks, which can only be measured as jets at the ATLAS experiment, the top quark can be indirectly measured from its decay products. The top quark has a very short lifetime due to its large mass, leading to its decay before hadronization processes may arise. Due to the structure of the CKM matrix, the top quark almost exclusively decays into a b quark and a W boson, as shown in Figure 2.4. The W boson further decays into a pair of quarks ($q\bar{q}'$) or a charged lepton and its neutrino ($\ell\nu_\ell$). The respective branching fractions are $\mathcal{BR}(W \to q\bar{q}') \approx \frac{2}{3}$ and $\mathcal{BR}(W \to \ell\nu_\ell) \approx \frac{1}{3}$.

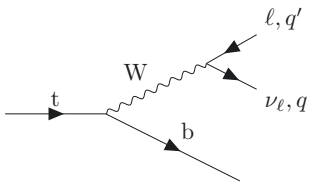


Figure 2.4: Feynman diagram of the most likely decay of the top quark and the subsequent decay of the W boson.

[4]The Feynman diagrams in this thesis were created with the TikZ-Feynman package [46]. The x-axis of the diagrams corresponds to the time axis. Anti-particles are illustrated as particles moving backward in time.

Another consequence of the large mass of the top quark is its large Yukawa coupling. The top quark is expected to be the fermion with the largest Yukawa coupling to the Higgs boson in the SM. Hence, the measurement of the Yukawa coupling of the top quark, Y_t, is an important test of the consistency of the Higgs boson measured at the LHC and the SM predictions. The general properties of the Higgs boson and the top-quark-associated Higgs boson production, which allows for a direct measurement of Y_t, are discussed in the following sections.

2.4 The Higgs boson

As explained in Section 2.2, the existence of the Higgs boson is believed to be a consequence of EW symmetry breaking. A brief overview of the properties of the Higgs boson and its production and decay mechanisms is given in the following.

2.4.1 Properties of the Higgs boson

The Higgs boson (H) is a spin 0 particle which carries neither an electric charge nor a color charge. Theoretically, its mass is given by $m_H = \sqrt{2\lambda}v$, where λ can be identified with the self-coupling parameter of the Higgs boson. No theoretical prediction for the mass of the Higgs boson exists, as λ and v are free parameters in the SM. Experimentally, the mass of the Higgs boson has been measured by the ATLAS and CMS experiments at the LHC based on $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ final states, using the datasets collected at $\sqrt{s} = 7\,\text{TeV}$ and $8\,\text{TeV}$. The measured mass of the Higgs boson yields $m_H = 125.09 \pm 0.21\,(\text{stat.}) \pm 0.11\,(\text{syst.})\,\text{GeV}$ [16].

The Higgs boson couples only to massive particles with a coupling that is proportional to the mass or the squared mass of these particles. Hence, it couples most strongly to heavy particles like the W and the Z boson and the third generation quarks. In addition, it does also couple to itself via trilinear and quartic self-couplings. The interactions between the Higgs boson and fermions f, vector bosons V and itself can be described by the Lagrangian density

$$\mathcal{L} = -g_{Hff}\bar{f}fH + \frac{g_{HHH}}{6}H^3 + \frac{g_{HHHH}}{24}H^4 + \delta_V V_\mu V^\mu \left(g_{HVV}H + \frac{g_{HHVV}}{2}H^2\right), \tag{2.13}$$

with $V = W$ or Z, $\delta_W = 1$ and $\delta_Z = 1/2$. The parameters g_{Hii} describe the coupling of the Higgs boson to particles of type i. The couplings to fermions g_{Hff} is proportional to the mass of the fermion m_f, whereas the couplings to vector bosons, g_{HVV} and g_{HHVV}, or itself, g_{HHH} and g_{HHHH}, are proportional to the squared mass m_V^2 or m_H^2. The different couplings are given by:

$$g_{Hff} = \frac{m_f}{\nu},\ g_{HVV} = \frac{2m_V^2}{\nu},\ g_{HHVV} = \frac{2m_V^2}{\nu^2},\ g_{HHH} = \frac{3m_H^2}{\nu},\ g_{HHHH} = \frac{3m_H^2}{\nu^2} \tag{2.14}$$

The parameter g_{Hff} is related to the Yukawa coupling Y_f through a factor $\sqrt{2}$. As the top quark is the heaviest elementary particle in the SM, it is predicted to have a large Yukawa coupling of $Y_t \approx 1$ in the SM.

2.4.2 Higgs boson production processes at the LHC

The Higgs boson can be produced through different processes at the LHC involving initial states with quarks and gluons. The main production processes are gluon-gluon fusion (ggH), vector boson fusion (VBF) and W- or Z-boson-associated production (WH and ZH). The cross sections of the different processes depend on the center-of-mass energy $\sqrt{s}$ of the proton-proton collisions. An overview of the cross sections for the different Higgs production processes as a function of $\sqrt{s}$ is given in Figure 2.5.

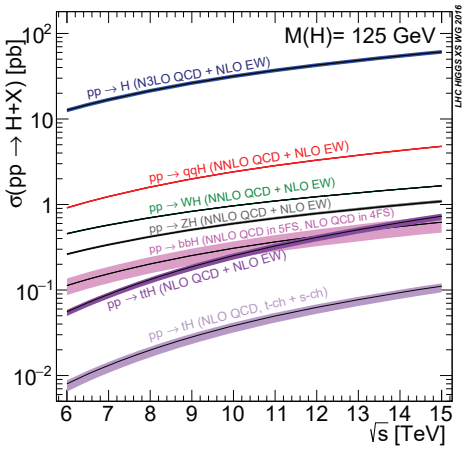


Figure 2.5: SM Higgs boson production cross sections [20] as a function of $\sqrt{s}$, calculated at different orders in QCD with different order EW corrections applied.

The ggH production provides the highest cross section and proceeds via loops, as the Higgs boson does not directly couple to the massless gluons. The respective Feynman diagram is shown in Figure 2.6a). Due to the mass dependency of the Higgs boson's coupling to fermions, the loop process most likely includes top quarks. Hence, the ggH process is sensitive to Y_t and allows for an indirect measurements of the coupling. At lowest order, the Higgs boson is produced without additional final state particles in ggH production. However, additional quarks in the final state may arise e.g. due to initial and final state radiation, as studied in the H+b-jets measurement.

The VBF production provides the second highest cross section, followed by WH and ZH production. In VBF production, two vector bosons are radiated off the initial state quarks and produce a Higgs boson in association with two quarks, as shown in Figure 2.6b). In WH and ZH production, a single vector boson is produced and the Higgs boson is radiated off the vector boson. The dominant contribution to ZH production arises from initial states with quarks, but a minor contribution originates from $gg \to ZH$ production, which is possible through loop and box diagrams. Example Feynman diagrams are shown in Figures 2.6c) and d), respectively.

More rare processes are the Higgs boson production in association with a top quark pair ($t\bar{t}H$) or a pair of b quarks ($b\bar{b}H$), shown in Figure 2.6e), or in association with a single top quark (tH), which is discussed in more detail in Section 2.4.4. The additional contribution of Higgs boson pair production, which is possible either via box diagrams or the Higgs boson self coupling, is neglected due to its very small cross section. The evolution of the cross section as a function of $\sqrt{s}$, illustrated in Figure 2.5, shows that especially the Higgs boson production in association with top quarks benefits from the increase of the center-of-mass energy from $\sqrt{s} = 8$ TeV to

$\sqrt{s} = 13$ TeV at the LHC. The cross section of $t\bar{t}H$ and tH production increases by a factor of approximately 4, in contrast to a factor of approximately 2 for the other Higgs boson production processes.

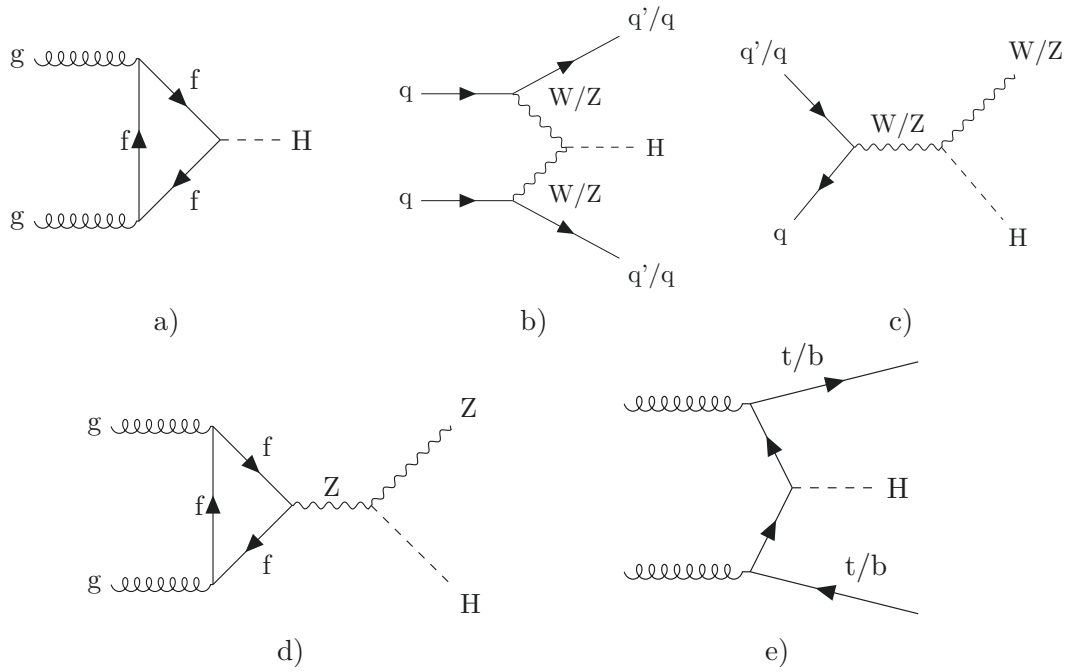


Figure 2.6: Example leading order Feynman diagrams for the different Higgs boson production processes at the LHC: a) gluon-gluon fusion (ggH), b) vector boson fusion (VBF), c) WH/ZH production, d) $gg \to ZH$ production, e) $t\bar{t}H$ and $b\bar{b}H$ production.

2.4.3 Decay channels

The Higgs boson is an unstable particle with a lifetime of approximately 10^{-22} s and provides a large variety of decay channels. The respective branching ratios depend on the strength of the Higgs boson's coupling to the final state particles, therefore final states with heavy particles are favored over final states with light or massless particles. Hence, the dominant decay channels are $H \to b\bar{b}$ and $H \to WW$, as the Higgs boson mass is not high enough to produce a $t\bar{t}$ pair. Other possible final states are gg, ZZ, $\tau\tau$, $c\bar{c}$, $\gamma\gamma$, $Z\gamma$ and $\mu\mu$. An overview of the different decay channels and their branching ratios as a function of the Higgs boson mass is given in Figure 2.7a). The considered $H \to \gamma\gamma$ decay channel is one of the most rare decay channels with a branching ratio of $\mathcal{BR}(H \to \gamma\gamma) = 0.227\%$ for $m_H = 125.09\,\text{GeV}$ [20]. Similar to ggH production, the decay into massless particles is only possible through loops. These loops involve W bosons and heavy quarks, mainly top quarks, as shown in Figure 2.7b).

2.4.4 Top-quark-associated Higgs boson production

The Yukawa coupling of the Higgs boson to the top quark Y_t is an important property of the Higgs boson, as the top quark is the heaviest particle in the SM. Hence, its measurement is essential to test whether the properties of the discovered Higgs boson are consistent with those predicted by the SM. The coupling Y_t can be directly measured through the associated

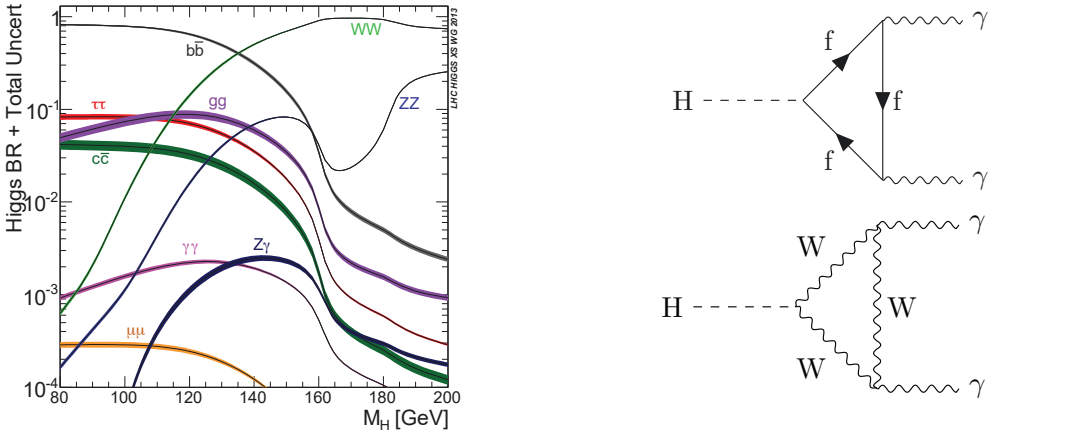


Figure 2.7: a) Branching ratios of the different Higgs boson decay channels as a function of m_H [20], b) example Feynman diagrams of the $H \to \gamma\gamma$ decay involving W bosons or fermions f.

production of the Higgs boson and a top quark. Similar to the SM top quark production mechanisms described in Section 2.3, the Higgs boson can be produced in association with a top-antitop quark pair ($t\bar{t}H$) or in association with a single top quark (tH). There are two main contributions to tH production: the t-channel production ($tHjb$) and the Wt-associated production (WtH). Another small contribution, which is neglected due to its small cross section, arises from the s-channel production. As there is a destructive interference for tH production in the SM and due the fact that it only occurs via EW processes, the cross section of the dominant t-channel contribution is about a factor 7 smaller than the cross section of $t\bar{t}H$ production, which occurs via the strong interaction. In Figure 2.8, the Feynman diagrams for $t\bar{t}H$, $tHjb$ and WtH production are shown.

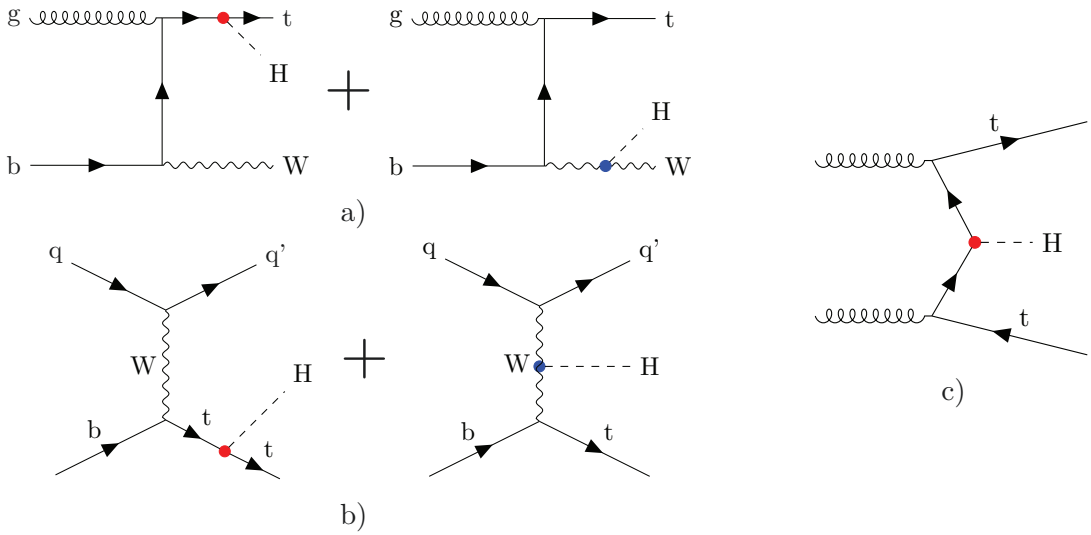


Figure 2.8: Example tree level Feynman diagrams for the interference of a) t-channel tH production ($tHjb$) and b) Wt-associated Higgs boson production (WtH). An example Feynman diagram for $t\bar{t}H$ production is shown in c). The vertices which are sensitive to Y_t and g_{HWW} are highlighted in red and blue, respectively.

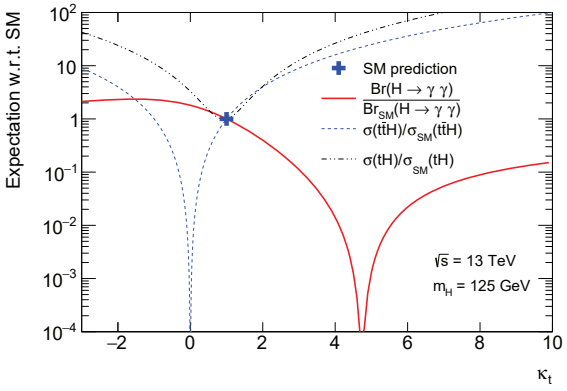


Figure 2.9: Cross sections of $t\bar{t}H$ and tH production and the $H \to \gamma\gamma$ branching ratio as a function of κ_t for $\sqrt{s} = 13\,\text{TeV}$, divided by the respective SM expectations. The SM expectation corresponds to $\kappa_t = 1$, where the tH cross section has a minimum due to the destructive interference in the SM. The κ_t parameterization for the tH cross section is based on Refs. [20, 47], while for $\text{Br}(H \to \gamma\gamma)$ the parameterization is taken from Ref. [48].

Whereas $t\bar{t}H$ production is only sensitive to the magnitude of Y_t, tH production also provides sensitivity to the relative sign of Y_t and the coupling of the Higgs boson to W bosons, g_{HWW}. As shown in Figure 2.8, the Higgs boson in tH production can be radiated off the top quark, giving sensitivity to Y_t, but it can also be radiated off the W boson, which gives sensitivity to g_{HWW}. In the SM, these two contributions interfere destructively leading to the small cross section of tH production. A coupling strength modifier κ_t can be introduced to describe deviations of Y_t from the SM expectation: $Y_t = \kappa_t \times Y_t^{SM}$. These deviations from the SM prediction would result in an increase of the tH production cross section. A special case is a Yukawa coupling with a negative sign with respect to the SM expectation, described by $\kappa_t = -1$, which would highly increase the tH cross section, while not affecting the $t\bar{t}H$ production cross section. In Figure 2.9, the dependence on κ_t of the $t\bar{t}H$ and tH cross sections and the $H \to \gamma\gamma$ branching ratio is shown.

In addition to $t\bar{t}H$ and tH production, also measurements of ggH production and the $H \to \gamma\gamma$ decay provide an indirect proof of the magnitude of Y_t, as they involve loop diagrams with top quarks. However, these indirect measurements have the disadvantage that there could be additional, yet unknown, contributions to the loop diagrams from particles predicted by Beyond the Standard Model (BSM) theories, which would alter the conclusions concerning Y_t. Nevertheless, the dependency of the ggH cross section and the $H \to \gamma\gamma$ branching ratio on κ_t needs to be considered when measuring Y_t. The cross sections of tH and $t\bar{t}H$ production can be parameterized using the minimal parameterization model as suggested in Ref. [48], in order to make a statement about κ_t. Various coupling strength modifiers $\vec{\kappa}$ are defined, which are used to describe deviations from the SM expectation for all relevant processes. The model is based on the parameterization of the cross section of a process i, including the $H \to \gamma\gamma$ branching ratio, in the narrow width approximation:

$$\sigma(i \to H \to \gamma\gamma) = \frac{\sigma_i(\vec{\kappa}) \cdot \Gamma^{\gamma\gamma}(\vec{\kappa})}{\Gamma_H} \tag{2.15}$$

The parameter $\Gamma^{\gamma\gamma}(\vec{\kappa})$ corresponds to the partial width of the Higgs boson decay into two photons and Γ_H to the total width of the Higgs boson. The advantage of this model is that no

parameters are fixed to their SM expectations and no assumptions are made about possible BSM contributions. The loops in ggH production and the $H \to \gamma\gamma$ decay are treated as unresolved loops and effective coupling strength modifiers $\kappa_g^2 = \sigma_{ggH}/\sigma_{ggH,SM}$ and $\kappa_\gamma^2 = \Gamma^{\gamma\gamma}/\Gamma_{SM}^{\gamma\gamma}$ are defined, where the SM expectations are denoted by SM. The total width of the Higgs boson is parameterized using $\kappa_H = \Gamma_H/\Gamma_{H,SM}$, thereby considering possible yet unknown BSM decays of the Higgs boson. Additional coupling strength modifiers are defined for the coupling of the Higgs boson to top quarks, κ_t, and the coupling to vector bosons, κ_V. Based on these coupling strength modifiers, three couplings strength modifier ratios are defined which can be used to fully describe the relevant processes. The parameter $\kappa_{g\gamma}$ describes the effective couplings strengths of ggH production and the $H \to \gamma\gamma$ decay, while λ_{Vg} and λ_{tg} are defined based on κ_V and κ_t and the effective coupling strength modifier κ_g:

$$\kappa_{g\gamma} = \frac{\kappa_g \cdot \kappa_\gamma}{\kappa_H}, \qquad \lambda_{Vg} = \frac{\kappa_V}{\kappa_g} \qquad \text{and} \qquad \lambda_{tg} = \frac{\kappa_t}{\kappa_g}. \tag{2.16}$$

The cross sections for $t\bar{t}H$ and tH production, including the $H \to \gamma\gamma$ branching ratio, can be expressed based on these couplings strength modifiers and their ratios. The $t\bar{t}H$ cross section is proportional to the squares of $\kappa_{g\gamma}$ and λ_{tg}:

$$\sigma(t\bar{t}H) \cdot Br(H \to \gamma\gamma) \quad \propto \quad \frac{\kappa_t^2 \cdot \kappa_\gamma^2}{\kappa_H^2} = \kappa_{g\gamma}^2 \cdot \lambda_{tg}^2 \tag{2.17}$$

In contrast, the tH cross section is parameterized using all three couplings strength modifier ratios, $\kappa_{g\gamma}$, λ_{tg} and λ_{Vg}, and a negative interference term occurs:

$$\sigma(tH) \cdot Br(H \to \gamma\gamma) \quad \propto \quad \frac{(c_1\kappa_t^2 + c_2\kappa_V^2 - c_3\kappa_t\kappa_V) \cdot \kappa_\gamma^2}{\kappa_H^2} \tag{2.18}$$

$$\propto \quad (c_1\lambda_{tg}^2 + c_2\lambda_{Vg}^2 - c_3\lambda_{tg}\lambda_{Vg}) \cdot \kappa_{g\gamma}^2 \tag{2.19}$$

The parameters c_i describe the relative contributions to tH production resulting from the coupling of the Higgs boson to top quarks and W bosons, respectively, and the interference term. These depend on the center-of-mass energy and the considered tH production process.

By using this parameterization, statements about the compatibility between data and the SM predictions of the Higgs boson's coupling to the top quark, vector bosons and the loop contributions in ggH production and the $H \to \gamma\gamma$ decay can be made at the same time. The optimized tH categories presented in this thesis provide sensitivity to negative signs of λ_{tg}, as described in Chapter 9.

Chapter 3

The LHC and the ATLAS experiment

As explained in the previous chapter, most of the elementary particles are unstable or appear only in bound states. In order to study the properties of elementary particles and their interactions, these are produced in particle collisions at high energies. Protons are accelerated at the LHC to almost the speed of light and brought to collision at center-of-mass energies of up to $\sqrt{s} = 13\,\text{TeV}$. The ATLAS detector, which is composed of several subdetectors, is designed to precisely measure the particles produced during these collisions. In the following, the acceleration of protons at the LHC and the detection of the different types of particles at the ATLAS detector are described.

3.1 The Large Hadron Collider

The LHC (Large Hadron Collider) [49] is a circular particle accelerator with a circumference of about 27 km, thus being the largest particle accelerator which has been built so far. It is located at the European Centre of Nuclear Research (CERN) in Geneva. The LHC was built during 1998-2008. One of the main motivations for building the LHC was to reach high enough center-of-mass energies, $\sqrt{s}$, to be able to observe the, by then undiscovered, Higgs boson.

At the LHC, two proton beams are accelerated in opposite directions in a tunnel, which is approximately 100 m underground. The protons are accelerated inside evacuated beam pipes in order to avoid interactions with molecules in the air. The pressure in the beam pipes is about $10^{-13}\,\text{atm}$, which corresponds to an ultrahigh vacuum. In order to keep the protons in a collimated beam and to deflect the beam on the circuit of the LHC, a large number of superconducting magnets is installed over the whole range of the LHC. The magnets are cooled down to 1.9 K using liquid helium. Different types of magnets are used for different purposes. Dipole magnets are used to bend the beams on the circuit, while quadrupole, sextupole, octupole and decapole magnets are used to focus the proton beams and correct the beam position. The protons are accelerated by passing through eight radiofrequency cavities per beam, which are operated at 400 MHz. The cavities are based on superconductivity and they are hence operated at low temperatures of about 4.5 K.

The protons are passed through different pre-accelerators before being injected into the LHC. As a first step, they enter the linear accelerator LINAC2, where the protons reach an energy of about 50 MeV. Then, the PS BOOSTER further increases the energy of the protons to 1.4 GeV. The following steps of the acceleration chain involve the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) accelerators, where energies of 25 GeV and 450 GeV are reached, respectively. The protons with an energy of 450 GeV are finally injected into the LHC. The

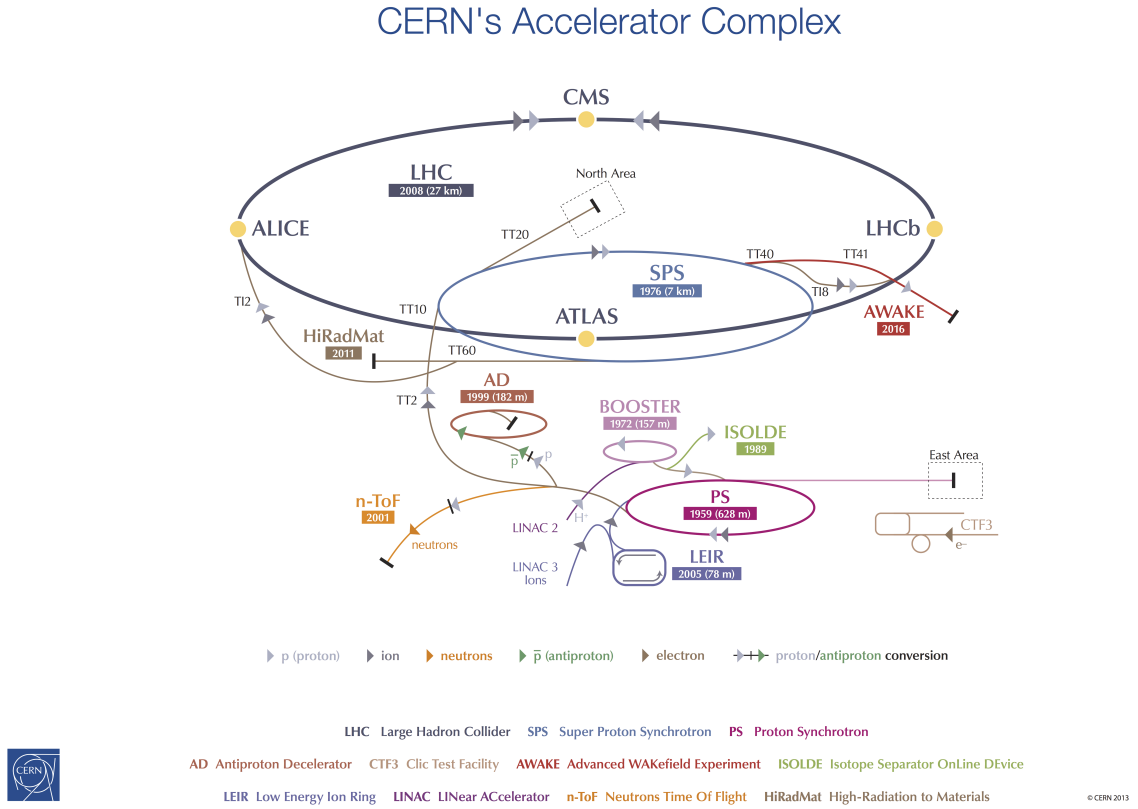


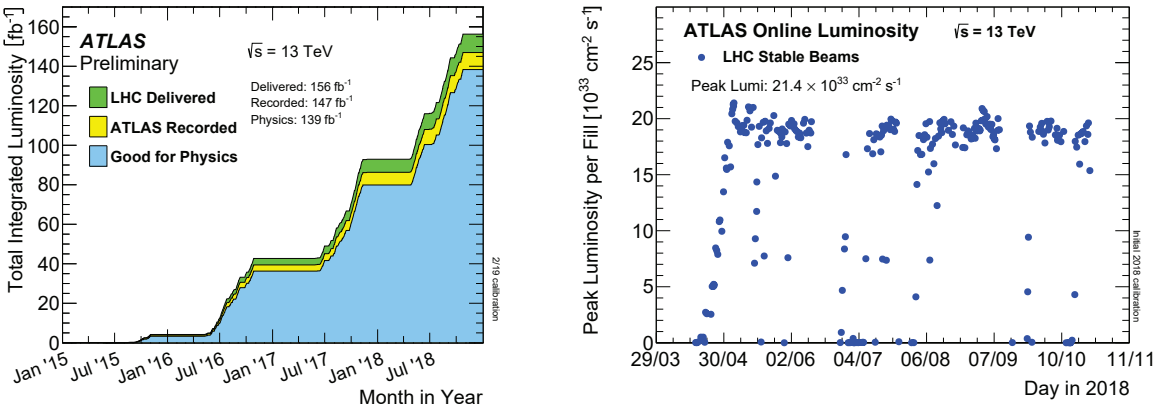
Figure 3.1: An illustration of the CERN accelerator complex [50].

LHC accelerator complex, including the pre-acceleration chain, is illustrated in Figure 3.1. The highest energy at which the LHC was operated so far amounts to 6.5 TeV per proton beam, resulting in a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. The protons are accelerated in bunches of 10^{11} particles with a bunch spacing of 25 ns, which can be directly translated into a frequency of 40 MHz at which the proton bunches are brought to collision. The large number of bunch crossings per second increases the probability to measure a hard scattering process during the proton-proton (pp) interactions.

The proton bunches are brought to collision at four interaction points. At each of them, an experiment is located which is focused on a particular field of research. Among these, there are the two multi-purpose detectors CMS and ATLAS, which are designed to efficiently measure all different kinds of particles in order to cover a broad range of measurements. These include the measurement of the Higgs boson and its properties, precision measurements of several SM parameters, as well as searches for BSM physics. These two experiments exhibit a similar structure and were built with the intention of providing independent measurements in order to confirm each other's results. In addition, there is the LHCb experiment, which focuses on precision measurements of B-hadrons in order to investigate CP violation to explain the matter-antimatter symmetry in the universe. Finally, the ALICE experiment focuses on the production and investigation of the quark-gluon plasma, which is expected to represent the state of the universe shortly after the Big Bang. For this purpose, special heavy ion runs are performed in which primarily lead ions are accelerated and brought to collision.

Different periods of data-taking were performed since the first pp collisions in 2010. The LHC started to operate with $\sqrt{s} = 7\,\text{TeV}$ in 2010, which was increased to $\sqrt{s} = 8\,\text{TeV}$ in 2012. In

2011 and 2012, datasets corresponding to integrated luminosities of $4.5\,\text{fb}^{-1}$ and $20.3\,\text{fb}^{-1}$ were collected, respectively. This data-taking period is referred to as Run 1 and was followed by a technical stop, which lasted until 2015. In 2015, the Run 2 of data-taking started with an increased $\sqrt{s}$ of 13 TeV. The data-taking of Run 2 ended in 2018 with a total amount of data corresponding to $139\,\text{fb}^{-1}$. The operation of the LHC during Run 2 was very successful and the design luminosity of the LHC of $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ was exceeded several times. After another technical stop, the Run 3 data-taking is planned to take place from 2021-2023 before upgrading the LHC to the High-Luminosity LHC (HL-LHC) [51]. In Figure 3.2a), the integrated luminosity provided by the LHC and measured with the ATLAS detector is shown for Run 2. In addition, the amount of data fulfilling certain quality requirements, which is used for physics analyses, is shown. The peak luminosity per fill of the LHC during the 2018 data-taking is shown in Figure 3.2b).



a) b)

Figure 3.2: a) Integrated luminosities provided by the LHC and recorded by the ATLAS detector during the Run 2 data-taking period [52], b) peak luminosity per fill during the 2018 data-taking period [53].

3.2 The ATLAS experiment

The ATLAS detector [54] is a multipurpose detector, which is located at one of the interaction points at the LHC. It has a cylindrical shape with a length of 46 m and a diameter of 25 m. The ATLAS detector is composed of different subdetectors, which are arranged symmetrically around the interaction point. The innermost detectors are dedicated to precisely measuring the tracks of charged particles that are produced in the pp collisions. A calorimeter system is used to measure the energy of electrons, photons and hadrons. The outermost part of the ATLAS detector is the Muon Spectrometer, which measures the tracks of muons as these are the only elementary particles, besides neutrinos, which are not stopped inside the calorimeters. Magnets are used to bend the trajectories of charged particles, thereby allowing to measure the sign of their electric charge and their momentum. Finally, in order to handle the large amount of data which is produced by the ATLAS detector, the Trigger and Data Acquisition system is used to decide which events are saved for data analysis. An overview of the ATLAS detector and its components is given in Figure 3.3. The structure of the different detector components is explained in more detail in the following sections.

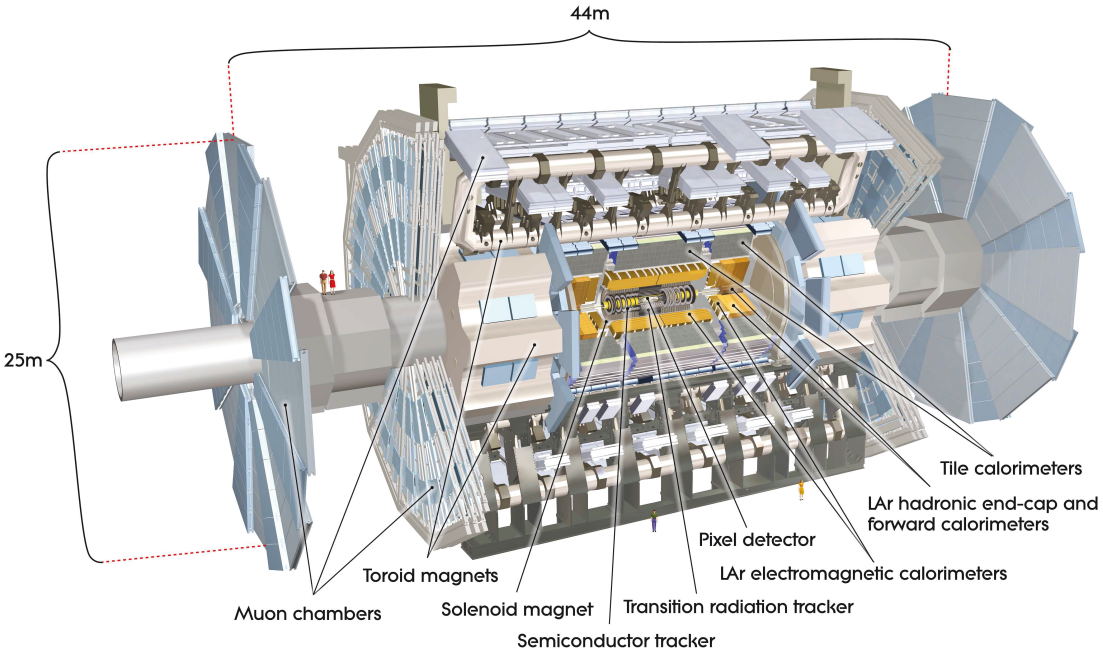


Figure 3.3: An overview of the different components of the ATLAS detector [54].

The coordinate system of the ATLAS detector is arranged so that it has its origin at the interaction point. The z-axis is parallel to the beam pipe, while the x-axis points to the centre of the LHC and the y-axis is directed upwards. The x-y plane describes the plane transverse to the beam axis and is used to define several observables. Due to its geometry, the ATLAS detector is usually described by using cylindrical coordinates with ϕ describing the azimuthal angle and θ describing the polar angle. The pseudorapidity η, which is defined as $\eta = -\ln\tan(\theta/2)$, is commonly used as a measure for the polar angle. While $\eta = 0$ corresponds to the direction transverse to the beam pipe with $\theta = 90°$, $\eta = \infty$ corresponds to the direction of the beam pipe with $\theta = 0°$. Angular distances in the x-y plane are defined by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. At a hadron collider, the absolute momentum of the interacting particles is not known, but its x and y components are known to be zero. Products from the hard scattering process are often characterized by a large momentum perpendicular to the beam axis, p_{T}. In addition, due to momentum conservation, the sum of transverse momenta of the products of a collision are expected to be zero, which allows to define observables like the sum of transverse momenta $\sum p_{\mathrm{T}} = -E_{\mathrm{T}}^{\mathrm{miss}}$, that provides a measure for the presence of undetected particles.

3.2.1 The Inner Detector

The Inner Detector (InDet) is the part of the detector which is closest to the interaction point and provides the highest granularity. It precisely measures the tracks of charged particles in the region $|\eta| < 2.5$ and allows for the reconstruction of interaction vertices. The central solenoid magnet, which surrounds the InDet, generates a 2 T field inside the detector, which deflects charged particles. Hence, the momentum and the sign of the electric charge of a particle can be reconstructed from the curvature of the measured track. The momentum resolution decreases for particles with high momentum, as their tracks are less curved. The InDet is composed of three subdetectors: The innermost Pixel detector and the surrounding Semiconductor Tracker

(SCT) are precision tracking chambers, while the outermost Transition Radiation Tracker (TRT) has a lower granularity but provides a large number of measurements per particle. Figure 3.4 gives an overview of the composition of the InDet and the distance to the beam pipe of the different subdetectors in the central region.

Pixel Detector

The Pixel detector is the innermost subdetector of the InDet, with a distance of 33.25 mm to 122.5 mm to the beam pipe, and it provides the highest granularity. It is composed of a cylindrical barrel in the most central region and two end-caps, similar to the other components of the InDet. The Pixel detector has four layers in the barrel region, each of them providing one measurement for a traversing particle, while the end-caps are each composed of three disks. It is a semiconductor detector which is composed of silicon detectors with a typical size of $50 \times 400\,\mu m^2$ and has a total number of approximately 80 million read-out channels. The general functioning of these types of detectors is based on semiconductors with a p- and a n-doped region. This structure generates a depletion zone between the two regions and the behavior of the semiconductor is similar to a diode. By applying a voltage in reverse direction, the size of the depletion zone is increased. A traversing ionizing particle produces electron-hole pairs in the depletion zone and the electrons and holes are then attracted to the anode or cathode, respectively, where a current can be measured.

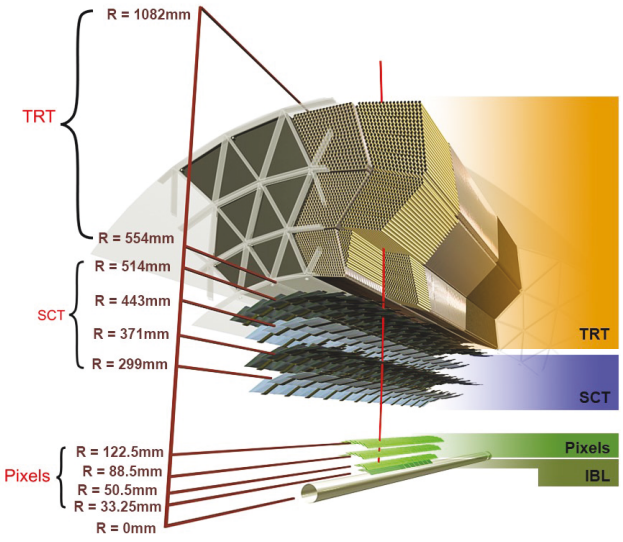


Figure 3.4: An illustration of the distance of the different detector components of the InDet to the beam pipe in the barrel region [55], including the IBL, which was installed during the technical stop between Run 1 and Run 2.

The innermost layer of the Pixel detector is the Insertable B-Layer (IBL) [56]. It is a relatively new component of the detector and was only installed during the shut-down between Run 1 and Run 2. The main motivation for the installation of the IBL were the inefficiencies in the previous innermost layer of the Pixel detector, the B-layer, whose information is in particular important for vertex reconstruction and a precise measurement of impact parameters. Impact parameters describe the closest approach of tracks to the primary vertex, where the hard scattering process took place, and are an important input to algorithms which are optimized to identify jets originating from b quarks.

Semiconductor Tracker

The SCT is an additional semiconductor detector, which is placed around the Pixel detector and is composed of silicon microstrips. Its distance to the beam pipe ranges from 299 mm to 514 mm in the barrel region. The barrel is composed of four layers, while the two end-caps are each composed of 9 disks. The silicon detectors have a size of $6.36 \times 6.4\,\mathrm{cm}^2$ and consist of 768 strips with a pitch of 80 µm. These detectors are used to build 12.8 cm long strips by connecting two pairs of detectors back-to-back at an angle of 40 mrad, allowing for measurements in R-ϕ and z direction. The SCT modules result in a surface of $\approx 61\,\mathrm{m}^2$ of silicon detectors and have about 6.3 million read-out channels. Typically, the SCT provides 8 measurements per track resulting in four measurements of space points.

Transition Radiation Tracker

The TRT is the outermost part of the InDet with a distance of 554 mm to 1082 mm to the beam pipe in the barrel region and is composed of polypropylene straw drift tubes with a diameter of 4 mm. The drift tubes are filled with a xenon-based gas mixture[1]. The anodes inside the drift tubes are tungsten wires, which allow for the measurement of electrons that are produced through ionization processes by traversing particles. In the barrel region, the TRT is composed of 73 straw planes, in which the straw tubes are arranged parallel to the beam axis. In the end-caps, there are 160 straw planes, which are ordered radially around the beam axis, thereby forming a wheel. The TRT exhibits about 351000 read-out channels. The number of measured hits per track is about 36 and therefore large compared to the precision trackers.

In addition, the TRT provides information about the type of particle which traverses the detector. As the name indicates, the TRT measures the transition radiation which is produced by charged relativistic particles passing a boundary layer of two materials with a different absolute permittivity. As electrons are lighter than hadrons, it is possible to distinguish between the two types of particles. The emerging transition radiation photons are absorbed by the gas mixture causing specific signals, which allow the distinction between photons originating from transition radiation and the ionization of a traversing particle.

3.2.2 The calorimeter system

The calorimeter system of the ATLAS detector is composed of two parts: the EM calorimeter and the hadronic calorimeter. The aim of the calorimeters is to measure the energy of particles by inducing particle showers through interactions with a material with a high density. For electrons and photons, these showers are produced through bremsstrahlung and pair production, while several different processes are involved in the development of hadronic showers. Both calorimeters are so-called sampling calorimeters with alternating layers of an active detector material and an absorber material. The incoming particles are supposed to deposit their whole energy inside the calorimeter. While the EM calorimeter is designed to absorb electrons and photons in order to identify these EM interacting particles, hadrons are not stopped inside the EM calorimeter but only inside the hadronic calorimeter. The properties of the showers in the calorimeter, such as the shower width, can be used to define additional criteria to distinguish between the different types of particles.

[1] The xenon gas mixture was replaced with argon in some of the TRT layers during 2016 data-taking.

Electromagnetic calorimeter

The EM calorimeter is a sampling calorimeter, which uses lead absorbers to induce particle showers of photons and electrons and liquid argon (LAr) as detector material to measure the energy of the particles originating from the EM shower. It provides a measurement of the deposited energies in a region $|\eta| < 3.2$. The barrel covers the region $|\eta| < 1.475$, while the end-caps provide measurements for $1.375 < |\eta| < 3.2$. The region between the barrel and the end-caps is usually excluded for the measurement of photons and electrons, as it is characterized by a large amount of material upstream of the calorimeter. The lead absorber plates and the kapton electrodes have an accordion shape, as shown in Figure 3.5, which ensures a full coverage of the EM calorimeter in ϕ. The amount of absorber material inside the calorimeter varies as a function of $|\eta|$ in order to provide a good energy resolution over the whole range covered by the calorimeter.

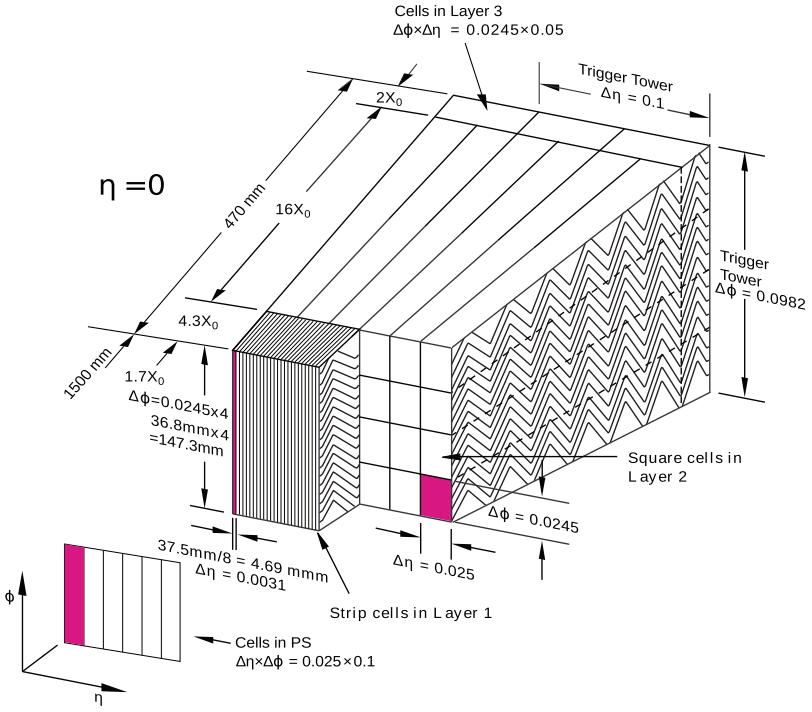


Figure 3.5: An illustration of the three different layers of the EM calorimeter and the presampler in the barrel region [54].

The EM calorimeter exhibits a special detector design in the region $|\eta| < 2.5$, where tracking information from the InDet is available. In this region, precision measurements can be performed for electrons and photons based on a structure of three calorimeter layers with a varying granularity. In Figure 3.5 the structure of the EM calorimeter in the barrel region is presented. The first layer is finely segmented in η and is composed of strips of a size 0.003×0.0245 in $\Delta\eta \times \Delta\phi$ for $0 < |\eta| < 1.4$ and $1.5 < |\eta| < 2.4$. For the remaining η regions, a coarser granularity is chosen. This fine segmentation of the first layer, which is also referred to as the *strip layer*, allows to resolve the overlapping showers of photons originating from neutral hadron decays up to transverse momenta of $\mathcal{O}(100\,\text{GeV})$. This information is in particular important to suppress the contributions of hadronic signatures during the photon reconstruction, as explained in Section 5.1.2. The depth of the first layer corresponds to approximately 4.4 radiation lengths

X_0 at $\eta = 0$. The second layer has a coarser granularity with cell sizes of 0.025×0.0245 in $\Delta\eta \times \Delta\phi$, but is much thicker with $\approx 16\ X_0$ at $\eta = 0$. Electrons and photons deposit most of their energy in this layer. The third layer has an even larger cell size with 0.05×0.0245 in $\Delta\eta \times \Delta\phi$. Its depth amounts to $\approx 2\ X_0$ at $\eta = 0$. The main purpose of this layer is to correct for the leakage of high-energetic EM showers into the hadronic calorimeter. In addition, a presampler is installed in front of the EM calorimeter in the region $|\eta| < 1.8$. The presampler consists of a thin layer of LAr and is used to correct for energy losses of electrons and photons upstream of the EM calorimeter due to interactions with the material of the InDet.

Hadronic calorimeter

The main purpose of the hadronic calorimeter is to measure the energy depositions of hadronically interacting particles, which usually deposit only a small fraction of their energy in the EM calorimeter. It has a coarser granularity than the EM calorimeter of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ in the central region[2], which is sufficient for its main purpose of reconstructing jets and measuring the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$.

The hadronic calorimeter is composed of steel absorbers and scintillating tiles as active detector material in the barrel region $|\eta| < 1.7$. The scintillator turns into an excited state through interactions with a traversing charged particle and photons are emitted through the retransition into the ground state. Wavelength shifting fibres in combination with photomultipliers are used to read out the signal from the scintillators. The barrel is divided into the central barrel for $|\eta| < 1.0$ and two extended barrels for $0.8 < |\eta| < 1.7$, providing space in between for electronics and service pipes. Both consist of three layers in depth, resulting in a total thickness of 9.2 hadronic interaction lengths at $\eta = 0$, including the material of the EM calorimeter. In the end-cap wheels, the hadronic calorimeter is composed of LAr as active material and copper absorbers, as the forward region is characterized by increased radiation. The end-caps each consist of four sampling layers and cover the region $1.5 < |\eta| < 3.2$.

Forward calorimeter

In the region closest to the beam pipe $3.1 < |\eta| < 4.9$, the forward calorimeter (FCal) provides measurements of the energy depositions of photons, electrons and hadronically interacting particles. The FCal needs to be able to cope with the high amount of radiation in the forward region. It is composed of three modules, which use LAr as active material. The first module provides measurements of EM interacting particles and uses copper absorbers, while the remaining modules are built of tungsten and are optimized for measurements of hadronically interacting particles.

3.2.3 The Muon Spectrometer

The outermost part of the ATLAS detector is the Muon Spectrometer (MS), which measures the tracks of charged particles that are not stopped inside the calorimeter system. The muons produced during *pp* collisions at the LHC are minimum-ionizing particles (MIP) and, therefore, rarely interact with the detector material and produce no showers inside the calorimeter. The MS is composed of different parts, which on the one hand measure the tracks of muons and on the other hand provide information to the trigger system. The different parts of the MS are depicted in Figure 3.6. Overall, the MS covers a range of $|\eta| < 2.7$. In the barrel, the muon

[2] The granularity in the third layer of the barrel is coarser with 0.1×0.2, as well as in the region $|\eta| > 1.5$, where it is 0.2×0.2.

chambers are arranged cylindrically around the calorimeter system, while the end-caps consist of large wheels, which are orthogonal to the beam axis. The superconducting toroid magnets generate a magnetic field, which bends the trajectories of the muons. Similar to the InDet, this allows for the measurement of the momentum and the sign of the electric charge of the muons.

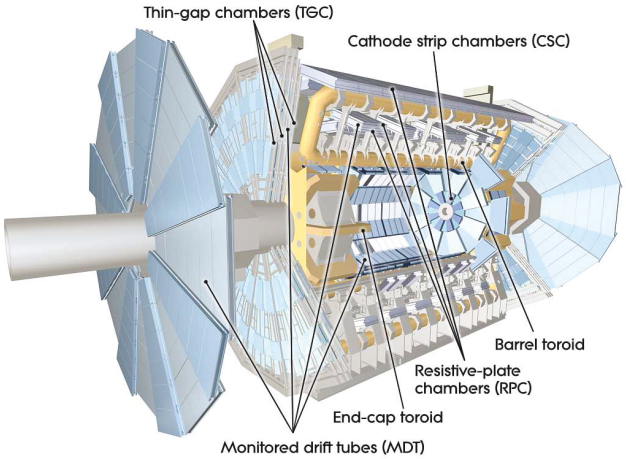


Figure 3.6: An illustration of the different parts of the MS [54]: MDTs ($|\eta| < 2.7$) and CSCs ($2.0 < |\eta| < 2.7$) are used for precision measurements of the muon tracks, while RPCs ($|\eta| < 1.05$) and TGCs ($1.05 < |\eta| < 2.4$) are used to provide information to the trigger system. The superconducting toroid magnets are used to bend the muon trajectories.

The precision muon tracking chambers are composed of Monitored Drift Tubes (MDTs) in the region $|\eta| < 2.7$, except for the innermost layer in the forward region $2.0 < |\eta| < 2.7$, where Cathode-Strip Chambers (CSCs) are used. The MDTs are build of three to eight layers of drift tubes with a diameter of about 30 mm. They are filled with Ar/CO_2 gas at a pressure of 3 bar. When crossing the drift tubes, charged ionizing particles produce free electrons and ions, which are collected at wires with a voltage of about 3000 V. In the forward region, the tracking chambers need to be able to cope with high particle fluxes. Therefore, CSCs, which are characterized by a good time resolution and a higher granularity, are used in this region. The CSCs are multiwire proportional chambers filled with a gas mixture of Ar, CO_2 and CF_4. They contain planes of cathodes with a segmentation into strips and anodes wires which are perpendicular to the cathodes. A traversing muon ionizes the gas, thereby producing avalanches of electrons, which are collected at the anode wires, whereas the positively charged ionized gas atoms induce a charge at the cathode. Due to the orthogonal arrangement of the anodes and cathodes, a measurement of two coordinates is obtained for each traversing particle. An optical alignment system is used to precisely determine the positions of the MDT chambers and the CSC strips with respect to each other, as these need to be known to reconstruct the muon trajectory with a high precision. The information from the alignment system is used to apply corrections during the offline reconstruction.

In addition, the muon system includes specific trigger chambers, which provide fast tracking information. Resistive Plate Chambers (RPCs) are used in the barrel region covering $|\eta| < 1.05$, which are built of two parallel resistive plates with an electric field of 4.5 kV/mm and a gas mixture based on $C_2H_2F_4$ in between. Electrons which are produced through ionization processes are accelerated towards the plate representing the anode, thereby producing electron

avalanches. Thin Gap Chambers (TGCs) are used in the end-cap region $1.05 < |\eta| < 2.4$, whose structure is similar to the multiwire proportional chambers.

3.2.4 Trigger and Data Acquisition system

The amount of data which is collected by the ATLAS detector is too large to be completely read out and stored. Therefore, the Trigger and Data Acquisition (TDAQ) system [57, 58] is used to select events with interesting signatures for permanent storage and reject the rest. Usually, the different types of particles are reconstructed as physics objects using dedicated algorithms. However, the TDAQ system makes use of simplified reconstruction algorithms, allowing for fast decisions. In Run 2 of the LHC, the proton bunches were crossed every 25 ns resulting in an event rate of about 40 MHz, which corresponds to data of the size of $\approx$ 100 Tb per second that would need to be stored. In order to reduce the event rate, a two-level trigger system is used. The different steps, which are performed by various subsystems before an event is stored permanently, are illustrated in Figure 3.7.

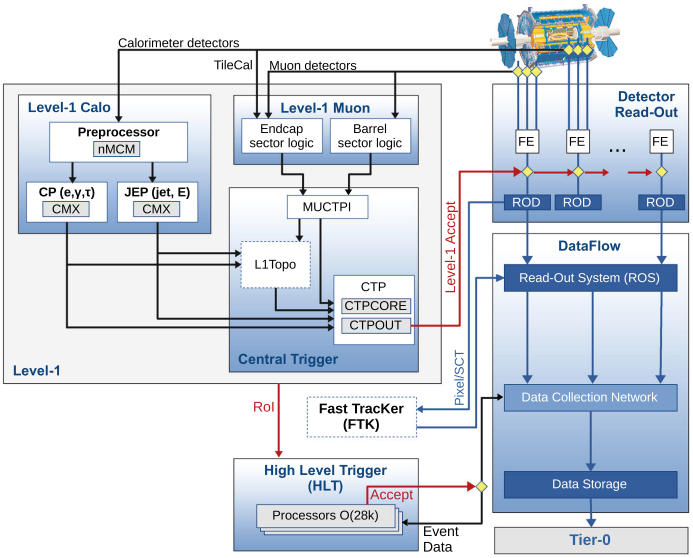


Figure 3.7: An illustration of the stucture of the ATLAS TDAQ system [58], which is used to reduce the event rate provided by the ATLAS detector. Events are processed by the L1 trigger and the HLT to identify interesting events for permanent storage.

First, the hardware-based Level-1 trigger (L1 trigger) rejects events mainly based on information provided by the calorimeter system and the MS. The initial event rate of 40 MHz is reduced to 100 kHz by the L1 trigger. The decisions of the L1 trigger are taken directly after collecting the data and are based on simple detector signatures. Based on these, regions of interest (RoIs) are defined by the L1 trigger. For events which pass the L1 trigger, the full information provided by the different detector components is buffered in the Read-Out-System (ROS). Then, this information is passed to the software-based High Level Trigger (HLT), which further reduces the event rate to 1 kHz. The decision by the HLT is based on simplified reconstruction algorithms for the different types of objects. For that purpose, the RoIs defined by the L1 trigger are passed to the HLT. Finally, if an event is accepted by the HLT, it is saved permanently. For the offline event reconstruction using the full ATLAS software, events are transferred to the Tier-0 computing facility.

Chapter 4

Analyzed dataset & event simulation

In this chapter, an overview of the properties of the analyzed dataset, collected with the ATLAS detector, and the generation of simulated events is given. In order to estimate the expected event yields for the processes of interest, referred to as signal, and contributions from other processes with a similar final state, referred to as background, Monte Carlo (MC) simulations are generated to describe the processes predicted by the SM. These MC simulations describe the kinematics of the different processes and are the basis for the optimization of an analysis. First, the basic principles of MC simulation at hadron colliders are briefly described, followed by a more detailed description of the considered MC samples.

4.1 Analyzed datasets

The considered datasets were collected with the ATLAS detector during pp collisions at a center-of-mass energy of 13 TeV in the years 2015-2016 and 2015-2017 and correspond to integrated luminosities of 36.1 fb^{-1} and 79.8 fb^{-1}, respectively. A diphoton trigger was used to select Higgs boson candidate events, whose requirements are explained in more detail in Chapter 5. Additionally, data quality requirements are applied to ensure that all detector components were fully functioning during the data-taking. As the protons are collided in bunches at the LHC, usually several pp interactions occur in addition to the hard scattering process of interest during one bunch crossing. These additional pp interactions are referred to as pileup. The LHC was operated at slightly different conditions during the different years of data-taking, resulting

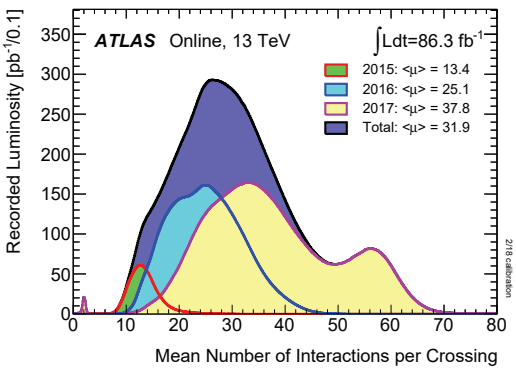


Figure 4.1: The average number of interactions per bunch crossing [59] for data collected with the ATLAS detector during the years 2015, 2016 and 2017.

in different pileup distributions. In Figure 4.1, the average number of interactions per bunch crossing is shown for the datasets collected in 2015, 2016 and 2017.

4.2 Basic principles of event simulation at the LHC

The generation of MC samples for processes induced by pp collisions [60] proceeds in different steps, which are performed by different algorithms. During a pp collision, interactions between the partons inside the protons occur. These partons are quarks or gluons and their properties are described by parton distribution functions (PDFs), $f(x, Q^2)$, which define the probability to find a parton with a certain momentum x inside a proton with momentum Q. Most interactions originate from low-energy partons, which are referred to as soft interactions. For the study of heavy particles, like the top quark or the Higgs boson, hard scattering processes are of interest, which are characterized by momentum transfers which are large compared to the QCD scale. As these hard scattering processes usually have much lower cross sections than soft interactions, MC samples are typically generated for specific processes by defining the initial and final state particles. In Figure 4.2, the general structure of an event produced during pp collisions is illustrated. This example shows a $t\bar{t}H$ event, where the different steps of the MC production are highlighted in different colors.

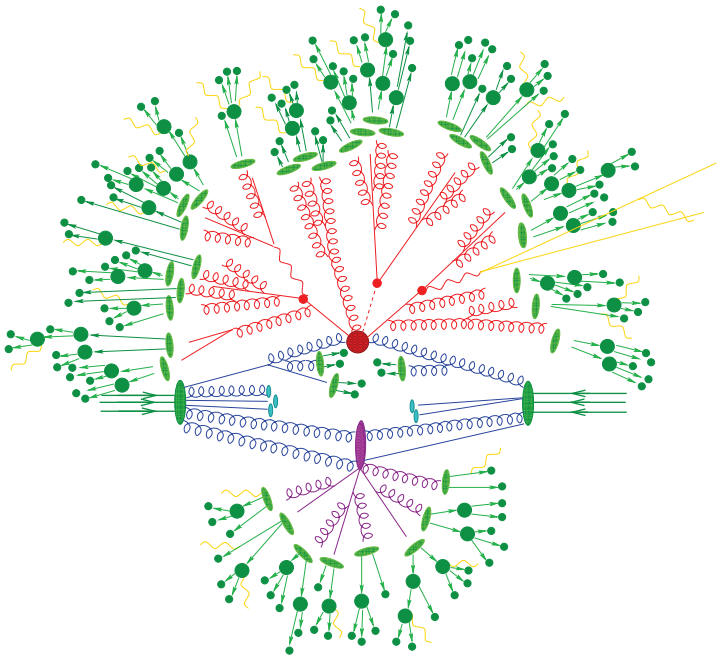


Figure 4.2: Illustration of the structure of an event produced during pp collisions [61]. The dark green ovals show the protons and the outgoing blue lines represent the initial state partons. The large red circle shows the hard scattering process and the red lines represent the generated particles and final state QCD radiation. The small red circles correspond to the decay vertices of the particles produced during the hard scattering process. The purple oval shows additional parton-parton interactions. Finally, the light green ovals represent the hadrons, which are build from all colored objects, and the dark green circles show the decay chains of unstable hadrons. The yellow lines represent the QED radiation, which can occur at any stage.

First of all, the hard scattering process is simulated by using MC event generators, which calculate the relevant transition matrix elements, describing the transition between the initial and final state particles. The squared amplitude of the matrix element describes the respective transition probability. The initial state particles are considered to be partons, which can be valence quarks, sea quarks or gluons, whose properties are described by PDFs. As the interacting partons originating from the proton are assumed to be asymptotically free, their interactions can be described by perturbation theory. In contrast, PDFs cannot be described by perturbative QCD and need to be approximated from data. In order to calculate the cross section of a process, an integration of the probability density function, given by the transition matrix element and the PDFs, over the full phase space needs to be performed, which is approximated by a random sampling of the phase space using the MC method. The resulting sample of events, which describes the process of interest, is referred to as MC sample. The transition matrix element calculation is performed at a fixed order, which is specific for the considered event generator, based on perturbation theory.

Additional QCD radiation is considered by so-called parton shower algorithms, which provide an approximation of higher-order corrections. As the initial state particles carry a color charge, they underlie QCD interactions and gluons can be radiated, which themselves are able to radiate additional gluons or produce quark-antiquark pairs. The same applies to color-charged particles which are produced during the hard scattering process, resulting in cascades of partons. The parton shower algorithms describe the decrease of the momentum transfer scale of the partons, before hadronization processes become relevant. Similar to QCD radiation, QED radiation can occur at any stage of the event generation and is handled by similar shower algorithms.

As the colliding protons are composed of several partons, the structure of an event at the LHC is also characterized by interactions of the remnants of the proton, which do not take part in the hard scattering process and are referred to as underlying event. Most of these interactions are soft and cannot be described by perturbative QCD. Instead, specific models are used, whose parameters are determined from data.

During the evolution of the parton shower, the momentum scale decreases and at a certain point the confinement of the partons becomes relevant. The colored objects originating from the parton shower start to form bound states. These hadronization processes cannot be described by perturbation theory, instead different models can be used to approximate these effects. The most common models are the string model, as implemented in Pythia8 [62], and the cluster hadronization model, which is used in the Herwig++ [63–65] and Sherpa [66] showering programs. The modeling of the parton shower, hadronization and underlying event is usually handled by a single showering program, as described in the following section, with certain parameter sets which are tuned to data.

As a last step, the decays of unstable hadrons are simulated. The hadrons which are formed during the hadronization are mostly not stable on the considered timescales and, therefore, their decays need to be modeled by specific algorithms. These are based on measurements of hadron decays and theoretical calculations of the decay amplitudes. In particular, the simulation of the decays of b- and c-hadrons is important, which is performed by the EvtGen package [67], as these play a major role in physics analyses at the LHC.

The simulated events are then passed through a simulation of the detector, as described in the following section, to transform the generated particles into detector signals. In contrast to data, information about the originally generated particles is available in MC simulation for each event, which is referred to as MC truth information.

4.3 Simulated samples

In the following, the MC samples for the different Higgs boson production processes which are considered in this thesis are described. For all MC samples, the mass of the Higgs boson is set to 125 GeV and the width is assumed to be $\Gamma_H = 4.04\,\text{MeV}$ [48]. An overview of the event generators and showering programs, which are used for the different processes, is given in Table 4.1. Furthermore, the cross sections, which are used for the normalization of the different samples, and their order of calculation are shown. The branching ratio of the $H \to \gamma\gamma$ decay, which is additionally considered for the normalization, amounts to 0.227% [20].

Table 4.1: Summary of the MC samples for the considered Higgs boson production processes in decreasing order of production cross section σ. The cross sections of the different processes are taken from Ref. [20]. For the $t\bar{t}H$ and $b\bar{b}H$ production processes, different MC samples were used for the measurements based on the 2015-2016 and the 2015-2017 dataset.

Process	Generator	Showering	σ [pb]	Order of σ calculation
ggH	Powheg NNLOPS	Pythia8	48.52	N^3LO(QCD)+NLO(EW)
VBF	Powheg-Box	Pythia8	3.78	NNLO(QCD)+NLO(EW)
WH	Powheg-Box	Pythia8	1.37	NNLO(QCD)+NLO(EW)
$q\bar{q}' \to$ ZH	Powheg-Box	Pythia8	0.76	NNLO(QCD)+NLO(EW)
$gg \to$ ZH	Powheg-Box	Pythia8	0.12	NLO+NLL(QCD)
$t\bar{t}H$	MG5_aMC@NLO	Pythia8	0.51	NLO(QCD)+NLO(EW)
$b\bar{b}H$	MG5_aMC@NLO	Pythia8	0.49	5FS(NNLO)+4FS(NLO)
$tHjb$	MG5_aMC@NLO	Pythia8	0.07	5FS(NLO)
WtH	MG5_aMC@NLO	Herwig++	0.02	5FS(NLO)
Improved MC samples for 2015-2017 analyses				
$t\bar{t}H$	Powheg-Box	Pythia8	0.51	NLO(QCD)+NLO(EW)
$b\bar{b}H$	Powheg-Box	Pythia8	0.49	5FS(NNLO)+4FS(NLO)

The t-channel production of a single top quark in association with a Higgs boson, $tHjb$, is simulated using MG5_aMC@NLO [68] with leading-order (LO) accuracy in QCD. The sample is produced using the four-flavor scheme (4FS), where the incoming b quark is assumed to originate from a gluon splitting. In contrast, the five-flavor scheme (5FS) assumes that the incoming b quark originates directly from the proton. It was shown that the use of the 4FS results in a more accurate description of the kinematics of the process [69]. The CT10 PDF set [70] and the Pythia8 showering algorithm [62], which provides a description of the parton shower, hadronization and underlying event, are used in combination with the A14 tune parameter set [71]. The cross section for $tHjb$ production is calculated at next-to-leading-order (NLO) in QCD [72].

For the simulation of Wt-associated production of the Higgs boson, the MG5_aMC@NLO event generator and the CT10 PDF set are used to generate events at NLO. In contrast to the $tHjb$ sample, the 5FS is used. The generator is interfaced to Herwig++ [63–65] for showering and hadronization. For the underlying event, the Herwig++ UEEE5 parameter set [73] is used. The sample is normalized to the cross section calculated at NLO in QCD. As the final state of WtH and $t\bar{t}H$ production is very similar, an overlap removal between the two samples is performed using a diagram regularization technique [74].

Different MC samples are used to describe $t\bar{t}H$ production in the measurements performed with the 2015-2016 and the 2015-2017 dataset, respectively. For the 2015-2016 dataset, $t\bar{t}H$ events

are generated at NLO in QCD using MG5_aMC@NLO and the NNPDF3.0 PDF set [75]. The event generator is interfaced to Pythia8 for showering, hadronization and underlying event and the A14 parameter set is used. The cross section is calculated at NLO in QCD and EW corrections are applied at NLO [76–79]. For the 2015-2017 dataset, the Powheg event generator [80–82] and the PDF4LHC15 PDF set [83] are used instead. In addition, the Pythia8 showering program is used with the AZNLO parameter set [84].

Also, the MC sample for $b\bar{b}H$ production is different for the two analyzed datasets. For data taken in 2015-2016, $b\bar{b}H$ production is simulated using MG5_aMC@NLO and the CT10 PDF set. The event generator is interfaced to Pythia8 for showering, hadronization and underlying event. The normalization is based on the calculation of the cross section using the 5FS at next-to-next-to-leading-order (NNLO) in QCD and the 4FS at NLO in QCD, which are matched using the *Santander* scheme [85–87]. The interference with the ggH MC sample is accounted for. Similar to $t\bar{t}H$ production, the sample used for the 2015-2017 dataset is generated using Powheg and the PDF4LHC15 PDF set.

For ggH production, the Powheg NNLOPS event generator [88] and the PDF4LHC15 PDF set are used to generate events at NNLO in QCD. Pythia8 and the AZNLO parameter set are used for showering, hadronization and the underlying event. The cross section for ggH production is calculated at next-to-next-to-next-to-leading order (N^3LO) in QCD with EW corrections applied at NLO [89–92].

VBF and ZH/WH production through qq or qg initial states and the additional contribution from $gg \to ZH$ are generated at NLO in QCD using Powheg-Box [93, 94]. For all samples, the PDF4LHC15 PDF set and the AZNLO tune are used. The $gg \to ZH$ sample only describes processes without additional partons in the final state, which include box and loop diagrams. For the simulation of the parton shower Pythia8 is used, as well as for hadronization and the underlying event. Cross sections calculated at NNLO in QCD with NLO EW corrections are used for the normalization of the ZH, WH and VBF samples [95–99]. For the $gg \to ZH$ sample, the cross section is calculated at NLO and next-to-leading-logarithm accuracy in QCD [96, 100].

Furthermore, the pileup distribution observed in data is considered for the simulation of MC samples in order to obtain a similar distribution of the mean number of interactions per bunch crossing. The additional interactions are accounted for by simulating inelastic pp interactions using Pythia8 in combination with the MSTW2008lo PDF set [101] and the A2 parameter set [102].

For all Higgs boson production MC samples, the simulated events are passed through a full simulation of the ATLAS detector [103] based on Geant4 [104], providing a detailed description of all detector components. After this last simulation step, events from MC samples are treated similar to data and the same object reconstruction and selection requirements are applied.

Additional MC samples are produced for the dominant $\gamma\gamma$ component of the continuum background. These diphoton MC samples are generated using Sherpa at LO, allowing for a maximum number of two or three real emissions of additional partons. The default underlying event parameter set and the CT10 PDF set are used. A merging of the generated events and the Sherpa parton shower [66] is performed using the Meps@lo scheme [105]. Due to the much higher cross section of non-resonant SM diphoton production, an MC sample with large statistics is needed. Therefore, in contrast to the Higgs boson production samples, a fast parametric simulation of the ATLAS detector [103] is used instead of a full Geant4 simulation of the detector response, as this step of the MC production requires the most computing resources.

Chapter 5

Object reconstruction and selection

Dedicated algorithms are used for the reconstruction of the different types of physics objects in the data collected by the ATLAS detector and equally in MC simulations. Different types of particles exhibit specific signatures in the ATLAS detector, which allow for their precise identification. These objects are used to select events for an analysis depending on the final state of interest. The reconstruction of photons is particularly important for measurements in the $H \to \gamma\gamma$ decay channel and proceeds in different steps, as explained in the following. In addition, the reconstruction of jets, electrons and muons is explained, as these objects are relevant for the final states under study. Jets originating from b quarks are of special interest, both for the measurement of the tH and $t\bar{t}H$ processes and H+b-jets, and can be identified by using dedicated flavor tagging algorithms. Furthermore, the reconstruction of the missing transverse momentum is explained, as this quantity provides a measure for the presence of undetected particles.

5.1 Photons

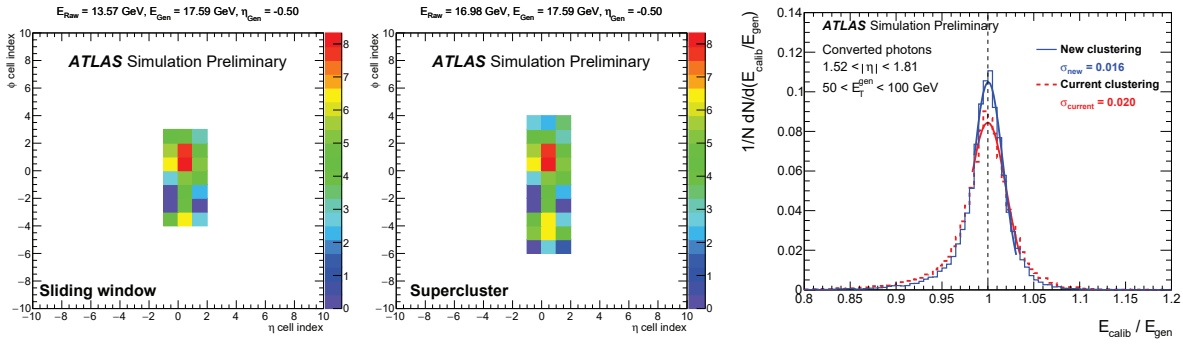
Different approaches are used for the photon reconstruction at the ATLAS experiment, which are briefly described in the following. Additionally, photon identification, isolation and trigger requirements are defined in order to efficiently identify photons, while reducing the fraction of objects misidentified as photons.

5.1.1 Reconstruction algorithms

As the signatures of electrons and photons in the ATLAS detector are very similar, the same algorithm is used for their reconstruction. After the basic reconstruction, a distinction is made between electrons, converted and unconverted photons. Unconverted photons are measured only in the EM calorimeter and have no associated tracks. Converted photons, on the other hand, interact with the material of the InDet, thereby producing an e^+e^- pair. Hence, converted photons are reconstructed from energy depositions in the EM calorimeter and tracks from the conversion. Therefore, their signature is similar to those of electrons, which are also characterized by a track in the InDet which is associated with an energy deposition in the EM calorimeter.

In general, the reconstruction of photons is based on energy clusters in the EM calorimeter. These clusters are built from a group of calorimeter cells by using dedicated clustering algorithms. Different reconstruction algorithms were used for the 2015-2016 and 2015-2017 datasets. For the

first two years of data-taking in Run 2, the so-called sliding-window [106] clustering algorithm was used. This algorithm constructs seed clusters in the EM calorimeter based on a rectangular window with a fixed size, which is located such that it contains a local maximum of the transverse energy, E_{T}. The fixed cluster size is defined in order to make a compromise between including most of the energy deposited by the EM shower and reducing the noise, which increases for larger cluster sizes. The cluster size is chosen to be $\Delta\eta \times \Delta\phi = 0.075 \times 0.123$ and an energy threshold of 2.5 GeV is applied to the seed cluster. The final EM cluster is built around the seed cluster with a size that varies for converted and unconverted photons and depends on the position of the seed cluster in the detector.



a) b)

Figure 5.1: a) Comparison of the EM clusters of an electron built from calorimeter cells using the sliding-window and the supercluster algorithm. The energy of the uncalibrated EM cluster, E_{Raw}, is compared to the true energy of the particle generated in MC, E_{Gen}. The z-axis is defined as the natural logarithm of E_{Raw} in units of MeV. In b) a comparison of the energy resolution of converted photons reconstructed with the sliding-window (red curve) and the new supercluster algorithm (blue curve) is shown [107].

The sliding-window algorithm was replaced by a dynamic, topological cell clustering approach [108] for data taken in 2015-2017. Topoclusters [109], which are based on patterns of calorimeter cell significances, are also used for the reconstruction of jets, as explained in more detail in Section 5.4. Small modifications were implemented to the algorithm for building topoclusters based on information from the EM calorimeter only. First, seed clusters are selected by iterating over the topological clusters, starting with the highest-p_{T}-cluster, and imposing requirements on the minimum energy of the cluster and the presence of InDet tracks in the case of electrons. For each seed cluster, a search for satellite clusters with a lower p_{T} is performed. For photons, the satellite clusters are required to be within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the barycenter of the seed cluster, whereas for electrons a slightly larger window is considered. The satellite clusters need to fulfill additional requirements based on InDet information, like the presence of matched tracks or conversion vertices, to distinguish between electrons and photons when building the clusters. Finally, superclusters are constructed based on the selected seed clusters and the associated satellite clusters. A limitation of the cluster size in $|\eta|$ direction is applied, restricting the cluster size to 3 cells in η direction in the barrel and 5 cells in the end-caps. Using dynamic clusters with a variable size improves the energy resolution for electrons and converted photons, as e.g. losses due to bremsstrahlung in the InDet are considered. Moreover, combined superclusters can be defined by merging the clusters of e^+e^- pairs produced by a photon conversion. A comparison of EM clusters that are built using the sliding-window and the supercluster algorithm is shown in Figure 5.1a) for electrons. Additional

cells in $|\phi|$ direction are considered by the supercluster algorithm for the determination of the energy, leading to an improved energy resolution compared to the sliding-window algorithm, which is illustrated in Figure 5.1b) for converted photons.

The reconstruction of unconverted photons is straightforward, as these are defined as EM clusters without associated tracks. The reconstruction of converted photons is more challenging as their detector signature is similar to those of electrons. In order to distinguish between converted photons and electrons, conversion vertices are reconstructed based on electron tracks in the InDet. These conversion vertices can have either two associated tracks, as expected for e^+e^- pairs, or only one associated track, if the electrons cannot be separated or the reconstruction of one electron fails. If the tracks associated with conversion vertices can be matched to an EM cluster, the object is classified as a converted photon. However, most of the photons which are reconstructed as converted photons are also reconstructed as electrons. In addition, there is a non-significant fraction of unconverted photons which are reconstructed as electrons. Therefore, additional criteria are defined to solve this ambiguity [110] and to discriminate between electrons, converted and unconverted photons. These criteria are based on the presence of a conversion vertex, the p_{T} of the object, the ratio of the cluster energy and the track momentum E/p, the number of hits in the B-layer of the Pixel detector and the presence of hits only in the TRT. The probability of a conversion increases with an increasing amount of material which has been passed by the photon. Therefore, if the object is e.g. matched to a TRT-only track it is classified as a photon, whereas the presence of hits in the B-layer, which is closer to the interaction point, indicates the presence of an electron.

The calibration of photons and electrons [111, 112] proceeds in various steps. A multivariate algorithm is used to perform an e/γ calibration based on the photon and electron energies in MC simulations. The training of the multivariate algorithm is performed separately for converted and unconverted photons and electrons. Additionally, an intercalibration of the different layers of the EM calorimeter is performed to account for their different scales. Uniformity corrections are applied to account for variations in the detector response in certain detector regions. Finally, the absolute energy scale is determined based on $Z \to ee$ events by adjusting the data to agree with the MC expectation. Observed differences between the resolution of the Z boson peak in data and MC simulation are accounted for by deriving correction factors, which are applied to MC simulation. The same scale factors (SFs), derived from $Z \to ee$ events, are applied to photons and electrons and the approach is validated using $Z \to \ell\ell\gamma$ and $J/\psi \to ee$ events, respectively.

The considered photons are required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and to be reconstructed within $|\eta| < 2.37$, excluding the region $1.37 < |\eta| < 1.52$ which corresponds to the transition region between the barrel and the end-caps of the EM calorimeter.

5.1.2 Identification based on shower shape variables

Requirements are imposed on the reconstructed photons to ensure that these originate directly from the hard scattering process. These photons are referred to as prompt photons and are of interest for most physics analyses. Photons can also be produced inside a jet, e.g. from $\pi^0 \to \gamma\gamma$ decays, and jets which deposit a significant fraction of their energy in the EM calorimeter may be wrongly identified as photons. These reconstructed photon objects are referred to as *fake* photons in the following. It is possible to distinguish between prompt and *fake* photons based on information provided by the calorimeter system. Nine different variables describing the

longitudinal and lateral shower development are chosen to define photon identification (ID) requirements [110]. Their definition is based on information of the first and second layer of the EM calorimeter and the leakage into the hadronic calorimeter. The discriminating variables are briefly described in the following.

Prompt photons are expected to deposit most of their energy in the EM calorimeter, whereas *fake* photons usually come along with large energy depositions in the hadronic calorimeter. Therefore, two variables describing the leakage into the hadronic calorimeter are used to discriminate between prompt and *fake* photons:

- R_{had1} is defined by the ratio of E_{T} measured in the first layer of the hadronic calorimeter and the E_{T} measured in the EM calorimeter.
- R_{had} is defined by the ratio of E_{T} measured in the whole hadronic calorimeter and the E_{T} measured in the EM calorimeter.

The variable R_{had} was found to provide a better discrimination in the region $0.8 < |\eta| < 1.37$, while for the remaining $|\eta|$ ranges a requirement on R_{had1} is defined. In addition, several variables can be defined based on information provided by the finely segmented first layer of the EM calorimeter. Prompt photons are characterized by a narrower calorimeter shower compared to *fake* photons. In addition, the high resolution of the first layer in particular allows for the definition of variables based on a second energy maximum in the EM cluster, which is a typical signature for $\pi^0 \to \gamma\gamma$ decays. The EM strip layer variables are defined as follows:

- $\omega_{s\,3}$ and $\omega_{s\,\text{tot}}$ describe the lateral width of the shower based on $\eta \times \phi = 3 \times 2$ strips or 20×2 strips around the strip containing the largest energy deposit, respectively. The shower width tends to be smaller for prompt photons.
- f_{side} is defined as the relative difference between the energy deposited in the seven strips around the maximum and the energy deposited in the three central strips around the maximum. As the showers of prompt photons tend to be narrower, f_{side} tends to be smaller for prompt photons.
- $\Delta E = E^{S1}_{max,2}$ - E^{S1}_{min} describes the difference between the energy of the second-highest maximum and the energy measured in the strip which is located in the minimum between the two maxima. In the case of prompt photons only one maximum is present and ΔE becomes small, while it tends to have higher values for *fake* photons.
- $E_{\text{ratio}} = \frac{E^{S1}_{max,1} - E^{S1}_{max,2}}{E^{S1}_{max,1} + E^{S1}_{max,2}}$ describes the ratio between the difference of the first and second energy maximum and their sum. For prompt photons, this value is close to one due to the absence of a second significant energy maximum. For *fake* photons, a wider distribution is observed.

Finally, additional variables are defined using information from the second layer of the EM calorimeter. This layer has a coarser granularity than the strip layer, but it contains the largest fraction of energy deposited in the EM calorimeter by a photon, thereby providing additional discrimination against *fake* photons. The following three additional variables are used to define the photon ID criteria:

- R_η describes the ratio of the energy deposited in $\eta \times \phi = 3 \times 7$ cells and the energy deposited in a region of 7×7 cells. As the showers of prompt photons are narrower, the fraction of energy deposited in 3×7 cells is larger and, therefore, R_η tends to have larger values for prompt photons than for *fake* photons.

- R_ϕ describes the ratio between the energy deposited in 3×3 cells and the energy deposited in a region of 3×7 cells. Similar to R_η, prompt photons tend to have higher values of R_ϕ than *fake* photons.
- ω_2 describes the lateral width of the shower in η in a window of 3×5 cells. The showers of *fake* photons tend to be broader than those of prompt photons.

Differences between the distributions in data and MC simulations are observed for most of these shower shape variables. Therefore, a correction of the MC distributions is performed by applying a shift to these distributions [6]. These shifts are applied separately for converted and unconverted photons in bins of $|\eta|$ and E_T and are referred to as *fudge factors* (FFs).

Two different working points (WPs) are defined based on requirements on the shower shape variables: the *loose* and the *tight* ID WP. The requirements on the shower shape variables were optimized in seven different $|\eta|$ regions, which are defined by considering the amount of material upstream of the EM calorimeter and the geometry of the detector. The *loose* ID WP is defined using only a subset of the variables described above. The criteria on the shower shape variables are designed such that they provide a high signal efficiency, as the *loose* ID criterion is mainly used for trigger purposes. At this stage of the photon reconstruction, it is not distinguished between converted and unconverted photons. Therefore, the *loose* ID is based only on those shower shape variables which show a similar behavior for converted and unconverted photons. The variables which are used to define the *loose* ID requirement are the hadronic leakage, R_η and ω_2.

The *tight* ID requirement is based on the full set of shower shape variables. In contrast to the *loose* ID, the *tight* ID requirement is optimized to provide a high rejection of *fake* photons, as it is applied in physics analyses. For the *tight* ID criterion, the requirements on the shower shape variables are optimized for converted and unconverted photons separately. As the shower shape variables vary as function of E_T, considering their E_T dependence further improves the performance of the *tight* ID requirement. An E_T-dependent *tight* ID menu was optimized at the ATLAS experiment for the first time and the respective *tight* ID efficiencies, which were evaluated as part of this thesis, are presented in Chapter 6. For the measurements of Higgs boson properties presented in this thesis, the previous E_T-independent *tight* ID requirement is considered.

5.1.3 Isolation requirements

Additional criteria to discriminate between prompt photons and *fake* photons, originating from hadronic signatures, are defined based on isolation requirements [110], which are applied to the reconstructed photons. Prompt photons are usually isolated from hadronic activity, whereas *fake* photons inside jets are accompanied by several other particles that are produced during the hadronization process. Two different criteria can be used to define the photon isolation, which are based on the calorimeter isolation and the track isolation. Three different photon isolation WPs are defined [6]. For all measurements presented in this thesis, the loosest WP is considered, which is based on the requirements described below.

The calorimeter isolation, $E_\mathrm{T}^\mathrm{iso}$, is defined as the sum of E_T of clusters in a certain cone with radius ΔR around the photon: $E_\mathrm{T}^\mathrm{iso} = \sum_{\Delta R} E_T$. The contributions to $E_\mathrm{T}^\mathrm{iso}$ from the considered photon itself, the underlying event and pileup are accounted for and subtracted for

each photon. In the measurements presented in this thesis, a calorimeter isolation criterion of $E_{\mathrm{T}}^{\mathrm{iso}} < 0.065 \cdot E_{\mathrm{T}}$ is applied. The cone size is chosen to be $\Delta R = 0.2$.

The track isolation, $p_{\mathrm{T}}^{\mathrm{iso}}$, is defined as the sum of p_{T} of tracks in a certain cone around the photon direction. These tracks are required to have $p_{\mathrm{T}} > 1$ GeV and fulfill certain track quality requirements. Tracks associated with the conversion vertex of a photon are excluded from the computation of $p_{\mathrm{T}}^{\mathrm{iso}}$. The considered track isolation criterion is defined by $p_{\mathrm{T}}^{\mathrm{iso}} < 0.05 \cdot p_{\mathrm{T}}$. Similar to the calorimeter isolation criterion, the cone size is set to be $\Delta R = 0.2$.

5.1.4 Photon trigger

Specialized triggers [58] are used to efficiently identify events with potential photon signatures in data. First, the L1 calorimeter trigger is used to define a RoI in the EM calorimeter by defining trigger tower clusters with $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The RoI is defined by 4 neighboring towers which exceed a certain energy threshold. Isolation requirements are defined based on the energy of the surrounding towers in the EM calorimeter and towers in the hadronic calorimeter.

For the HLT, a fast reconstruction of the photons is performed first, followed by a more precise reconstruction. For the fast reconstruction, clusters are built from calorimeter cells based on the RoI using a fast calorimeter algorithm. These clusters need to fulfill requirements on a subset of the shower shape variables and the E_{T} of the cluster.

Finally, the more precise reconstruction makes use of simplified versions of the nominal photon reconstruction algorithms, as described above. The cluster reconstruction is performed based on the sliding-window algorithm and a calibration based on the multivariate method is performed, without distinguishing between converted and unconverted photons. Additionally, requirements on the photon E_{T} and the photon ID, based on a subset of the shower shape variables, need to be fulfilled. Hence, the photon triggers do not use any tracking information provided by the InDet.

5.2 Electrons

The reconstruction of electrons is based on the presence of an EM cluster which is matched to a track in the InDet and proceeds similar to the reconstruction of photons, as explained in Section 5.1.1. The reconstructed electrons need to fulfill $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.47$. Similar to photons, the region $1.37 < |\eta| < 1.52$ is excluded.

For the electron identification, a likelihood-based discriminant [113] is used in order to suppress the fraction of converted photons and jets which are misidentified as electrons. Several input variables are considered, which include shower shape variables, information about the number of hits in the different components of the InDet, a likelihood probability which is based on transition radiation information provided by the TRT and variables describing the compatibility between the direction of the reconstructed track and the EM cluster. The medium WP of the likelihood-based discriminant is considered. In addition, loose track and calorimeter isolation requirements are applied, based on the values $E_{\mathrm{T}}^{\mathrm{iso}}$ and $p_{\mathrm{T}}^{\mathrm{iso}}$ calculated in a cone $\Delta R = 0.2$ around the electron, which provide an E_{T} independent efficiency of 99%. The efficiencies of the electron identification and isolation criteria are corrected in MC simulation to match the values measured in data using $Z \to ee$ and $J/\psi \to ee$ events [114].

Additionally, requirements on the so-called impact parameters are defined to ensure that the electron is consistent with originating from the primary vertex. The impact parameters are defined as the distance of closest approach of the considered track to the primary vertex in different projections: z_0 and d_0 describe the longitudinal and the transverse impact parameter, respectively. The requirements $|z_0 \sin\theta| < 0.5\,\text{mm}$ and $|d_0|/\sigma_{d_0} < 5$ need to be fulfilled, with σ_{d_0} describing the uncertainty on d_0. Electrons are discarded if they are reconstructed within $\Delta R = 0.4$ of a jet, after applying the overlap removal criteria for jets described in Section 5.4, or an isolated photon.

5.3 Muons

The reconstruction of muons [115] is based on tracking information provided by the InDet and the MS. After separately reconstructing tracks in both detector components, muons are either reconstructed by performing a combined fit to the hits measured in the MS and the InDet or, in the absence of an InDet track, by an MS track which can be matched to the interaction point. Muons without a reconstructed track in the MS, which can still be reconstructed by using information from the EM calorimeter and track segments in the MS, are not considered. Muon candidates are required to have $p_\text{T} > 10\,\text{GeV}$ and to be reconstructed in the detector region $|\eta| < 2.7$.

Muon identification requirements are applied to suppress muons originating from the decay of light hadrons. The medium muon identification WP is used, which is based on requirements on the number of hits measured in the different components of the InDet and the MS and the compatibility of the momenta of the candidate muon measured in the two detector components. Additionally, loose muon isolation criteria are applied, including a track isolation criterion on p_T^iso, calculated in a cone of $\Delta R = 0.3$ around the muon, and a calorimeter isolation criterion on E_T^iso calculated in a cone of $\Delta R = 0.2$ around the muon. These isolation criteria provide an efficiency $\geq 95\%$ (99%) for muons with $p_\text{T} > 25\,\text{GeV}$ (60 GeV). In simulation, a correction of the muon momentum scale and resolution and the isolation and identification efficiencies is performed, based on measurements performed using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events [115], to obtain a better agreement between data and MC simulation.

Similar to electrons, muons are required to be compatible with originating from the selected primary vertex by requiring $|z_0 \sin\theta| < 0.5\,\text{mm}$ and $|d_0|/\sigma_{d_0} < 3$. Muons which are reconstructed within $\Delta R = 0.4$ of an isolated photon or a jet are rejected.

5.4 Jets

Jets are reconstructed using topological clusters [109] at the EM energy scale, which are built from energy depositions in the EM calorimeter and the hadronic calorimeter. These topological clusters are formed by assigning a significance to each calorimeter cell, which is based on the magnitude of the cell signal and the average noise. Cells with a high significance are defined as seed cells and topoclusters are formed by adding neighboring cells which pass a certain threshold on their cell significance. In addition, the spatial structure of the energy depositions is considered. Then, the anti-k_t algorithm [116] with a radius parameter $\Delta R = 0.4$ is used to build jets from the topological clusters by using the FastJet package [117], which clusters jets based on the distance of the topoclusters starting with those that have the highest p_T. A correction

is applied to the jets in each event to account for energy depositions from pileup [118] and its direction is corrected to match the position of the selected primary vertex. Finally, the jets are calibrated [119, 120] by applying various corrections, which are based on MC simulations as well as data. Simulation-based corrections are applied to the jet energy scale (JES) and the jet η based on comparisons of the distributions of truth jets and reconstructed isolated jets in data and dijet MC simulation. Additional corrections based on MC simulations are applied to various longitudinal and transverse properties of the jet, as they have an influence on the JES. Finally, in-situ corrections are applied, which are derived from the differences of the jet p_{T} distribution observed in 2012 data and predicted by MC simulations for γ/Z+jet and multijet events.

In order to suppress jets originating from pileup, the so-called jet vertex tagger (JVT) algorithm [121] is used. It is applied to jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$, as it relies on information from the InDet. The JVT provides a multivariate discriminant based on two variables: the corrected jet-vertex fraction, which is calculated as the sum of p_{T} of tracks associated with the jet and the primary vertex and the sum of p_{T} of all associated tracks with corrections applied to the p_{T} spectrum of pileup tracks, and the variable R_{pT}, which is defined as the ratio of the sum of p_{T} of tracks associated with the primary vertex and the total jet p_{T} after the calibration, during which a pileup subtraction is applied. The value of the JVT discriminant is required to be larger than 0.59 for jets fulfilling the kinematic requirements. Corrections are applied to the JVT efficiency in simulation to match the values measured in data.

In addition, kinematic requirements are applied. Jets are in general required to fulfill $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 4.5$. The jets which are measured in the region $|\eta| < 2.5$, where InDet information is available, are referred to as central jets, while jets with $|\eta| > 2.5$ are referred to as forward jets. Jets which overlap with reconstructed photons or electrons are rejected, by requiring a distance ΔR larger than 0.4 (0.2) to isolated photon (electron) candidates.

5.5 Missing transverse energy

Although neutrinos cannot be directly measured by the ATLAS detector, the magnitude of the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, provides a measure for the presence of undetected particles. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated [122, 123] as the negative sum of transverse momenta of all reconstructed objects described in this chapter and additional soft contributions, which are estimated from tracks that are consistent with originating from the primary vertex. The overlap between the different objects is accounted for during the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

5.6 Identification of *b*-jets

The use of flavor tagging algorithms allows for the identification of jets which have properties that are consistent with the decay of a b-hadron (b-tagging). Throughout this thesis, the MV2 algorithm [124, 125] is used which provides a multivariate discriminant by combining the information from different b-tagging algorithms. These algorithms are optimized to distinguish between jets originating from b quarks (b-jets), c quarks (c-jets) and u, d and s quarks and gluons (light flavor jets). Hadrons originating from a b-quark are characterized by a relatively long lifetime of about 1.5 ps. Specific properties of the associated tracks and the presence of a

secondary vertex enable the identification of b-jets. Three types of b-tagging algorithms based on different quantities are considered:

- **Impact Parameter:** The IP2D and IP3D algorithms [126] are based on the longitudinal and transverse impact parameters d_0 and z_0, as introduced in Section 5.2. Due to the long lifetime of the b-hadron, its decay vertex can be resolved and is displaced from the primary vertex, leading to large impact parameters of the tracks originating from the decay vertex.
- **Secondary vertex:** The secondary vertex finding algorithm [126] aims for reconstructing the displaced vertex within the jet. For all pairs of tracks inside the jet, a two-track vertex hypothesis is tested in order to reconstruct the secondary vertex which corresponds to the decay vertex of the b-hadron.
- **Decay chain multi-vertex algorithm:** The JetFitter algorithm [127] attempts to reconstruct the full decay chain of the b-hadron. The flight path of the b-hadron is thereby determined by connecting the primary vertices with the decay vertices by using a Kalman filter.

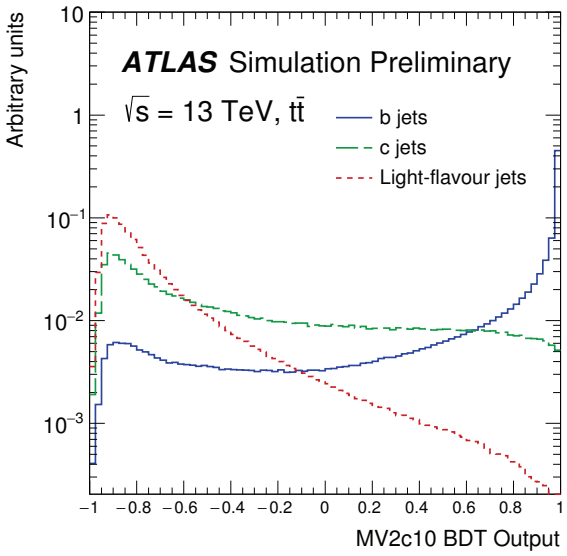


Figure 5.2: The MV2c10 discriminant for b-jets, c-jets and ℓ-jets obtained from $t\bar{t}$ MC simulation [125].

The different outputs, which are provided by these algorithms, and additional information about the kinematic properties of the jet are combined by using a multivariate algorithm. A Boosted Decision Tree (BDT) is trained considering b-jets as signal and c-jets and light flavor jets (ℓ-jets) as background. The proportion of c-jets and ℓ-jets in the background sample can be varied to define discriminants which do either provide a better rejection of c-jets or of ℓ-jets, respectively. For the MV2c10 algorithm, which is considered in this thesis, the background sample is composed of 93% ℓ-jets and 7% c-jets. The output discriminant of the MV2c10 algorithm is shown in Figure 5.2 for b-jets, c-jets and ℓ-jets. In MC simulation, it is possible to make a statement about the flavor of the considered jet, based on MC truth information, by looking for a b-hadron or a c-hadron in a certain cone with opening angle ΔR around the jet.

Based on the MV2c10 discriminant, four different WPs are defined, which offer different b-tagging efficiencies and rejection rates for c-jets and ℓ-jets. The rejection rates are defined as

Table 5.1: WPs of the MV2c10 algorithm and the respective efficiencies for b-jets and the rejection rates for c-jets and ℓ-jets evaluated in $t\bar{t}$ MC simulations [125].

BDT cut value	b-jet efficiency [%]	c-jet rejection	ℓ-jet rejection
0.9349	60	34	1538
0.8244	70	12	381
0.6459	77	6	134
0.1758	85	3.1	33

the inverse efficiencies. The cut values on the MV2c10 discriminant and the efficiencies and rejection rates of the different WPs are summarized in Table 5.1. Different methods [128–133] are used to measure the b-tagging efficiency for b-, c- and ℓ-jets in data. These measured efficiencies are compared to the values predicted by MC simulation and SFs are derived to correct the MC simulation.

Chapter 6

Photon identification efficiency measurements

As explained in the previous chapter, photon ID criteria are defined by requirements on shower shape variables in order to distinguish between prompt photons and hadronic signatures, which are reconstructed as photons (*fake* photons). The *tight* ID requirement is applied in most physics analyses, investigating final states with photons, to improve the signal to background ratio. Differences between the *tight* ID efficiency measured in data and predicted by MC simulation may occur e.g. due to mismodeling of the shower shape variables. There are well-known differences between the distributions of the shower shape variables in data and MC simulation, which are accounted for by correcting the distributions in MC simulation using the FFs, which shift the shower shape distributions in MC simulation to match the data distributions. Additional differences are taken into account by applying SFs to MC simulation, which are defined as the ratio of the *tight* photon ID efficiency measured in data and predicted by MC simulation: $\varepsilon_{data}^{\text{tight-ID}}/\varepsilon_{MC}^{\text{tight-ID}}$. These corrections are particularly important to precisely estimate signal and background contributions from MC samples. The uncertainty on the *tight* ID SFs is accounted for as a systematic uncertainty. The measurement of the *tight* ID efficiency in data is based on three different methods, which are briefly described in this chapter, followed by a more detailed description of one of these methods, the Matrix Method, which was used to measure the *tight* ID efficiency in the 2015-2017 dataset as part of this thesis.

6.1 Overview of the different methods

In order to measure the *tight* ID efficiency over a large range in photon E_{T}, three different measurements are combined which exploit the properties of different processes. In general, the measurement of the photon ID efficiency requires a sample which is sufficiently pure in prompt photons and provides a sufficient number of events for the efficiency measurement. However, the photon ID efficiency can also be measured from a sample of electrons, as explained in the following. The *tight* ID efficiency is measured separately for converted and unconverted photons and as a function of E_{T} in four different $|\eta|$ regions. The precision of the different measurements depends on the photon E_{T} and also the conversion type of the photons. Therefore, a combination of the three measurements is performed in order to achieve the best possible precision over a large E_{T} range.

A very pure sample of prompt photons is obtained from leptonic Z boson decays, where a photon is radiated off from one of the two leptons $Z \to \ell\ell\gamma$ ($\ell = e, \mu$). These events are used by the *Radiative Z method* [5, 6] to measure the *tight* ID efficiency for photons with a relatively low transverse energy, $10\,\text{GeV} \leq E_{\text{T}} \leq 100\,\text{GeV}$. This efficiency measurement benefits from the high

purity of the selected photon sample, allowing for a precise measurement even in the low E_{T} region, which is usually contaminated by a large fraction of *fake* photons. In contrast, it suffers from large statistical uncertainties for photons with a high E_{T}. Candidate events are selected by requiring two electrons or two muons with opposite electric charge and a photon fulfilling isolation requirements. After applying these requirements, the selected sample is dominated by $Z(\to \ell\ell) + jet$ events. However, these $Z(\to \ell\ell) + jet$ events can be easily discriminated from $Z \to \ell\ell\gamma$ events by considering the invariant masses of the system formed by the two selected charged leptons, $m_{\ell\ell}$, and the system formed by the two charged leptons and the selected photon candidate, $m_{\ell\ell\gamma}$. Whereas $Z \to \ell\ell\gamma$ events are characterized by an invariant mass $m_{\ell\ell\gamma}$ which is similar to the Z boson mass, m_Z, and $m_{\ell\ell} < m_Z$, $Z(\to \ell\ell) + jet$ events are characterized by $m_{\ell\ell} \approx m_Z$ and $m_{\ell\ell\gamma} > m_Z$. After imposing requirements on $m_{\ell\ell\gamma}$ and $m_{\ell\ell}$, the signal and background contributions are estimated by performing a maximum-likelihood fit to the $m_{\ell\ell\gamma}$ distribution, assuming certain templates for the shape of the signal and the background. The efficiency is then calculated as the ratio of the number of signal photons fulfilling the *tight* ID requirement and the total number of signal photons.

For the *Electron Extrapolation method* [5, 6], a sample of $Z \to ee$ events is used to measure the *tight* ID efficiency by exploiting similarities in the shower shape variables of reconstructed electrons and photons. This efficiency measurement is performed in a range $25\,\mathrm{GeV} \leq E_{\mathrm{T}} \leq 150\,\mathrm{GeV}$. Events containing a pair of oppositely charged electrons with an invariant mass of $70 < m_{ee} < 110\,\mathrm{GeV}$ are selected and a tag-and-probe approach is used. In order to account for differences of the shower shape variables of electrons and photons, Smirnov transformations are derived based on the shower shape variables observed for the *probe* electron in data and the shower shape variables for prompt photons predicted by MC simulation. Furthermore, small background contributions are considered by performing a signal+background fit to the invariant mass spectrum m_{ee}. The *tight* ID efficiency is then derived as the ratio of transformed electrons fulfilling the *tight* photon ID requirement and the total number of electron candidates. The measured efficiencies are characterized by relatively low uncertainties over the whole considered E_{T} range.

The *tight* ID efficiency measurement with the *Matrix Method* is based on a sample of photons which are produced either during the hard scattering process, via $qg \to \gamma q$ or $q\bar{q} \to \gamma g$, or during the fragmentation of a parton. For this efficiency measurement, the track isolation of the photons is used to discriminate between prompt and *fake* photons. The purity of the sample, which can be expressed by the track isolation efficiencies, is extracted before and after applying the *tight* ID criteria. This method has the advantage that it covers a comparably large E_{T} range of $25\,\mathrm{GeV} \leq E_{\mathrm{T}} \leq 1500\,\mathrm{GeV}$. However, due to large systematic uncertainties for low values of E_{T}, this measurement becomes important for the photon ID efficiency combination only for high values of E_{T}. The measurement presented in this thesis is based on the work of previous analyzers, who performed the measurement during Run 1 [5] and with the Run 2 dataset collected in 2015-2016 [6]. For this thesis, the measurement is performed for the first time using photons reconstructed with a new algorithm, based on superclusters [108], and data taken during 2017 is studied for the first time. In addition, the estimation of systematic uncertainties is reinvestigated. A systematic uncertainty which was previously evaluated by varying the material of the InDet during the detector simulation is replaced by variations of the shower shape variables in the prompt photon MC sample based on the uncertainties of the FFs. In addition, a new systematic uncertainty is added to take into account differences between data and MC simulation of the *loose* ID efficiency, when correcting the measured efficiency for the preselection requirements. The details of the measurement are described in the following.

6.2 Efficiency measurement with the Matrix Method

6.2.1 Analyzed dataset & MC samples

The data which are used to measure the *tight* ID efficiency were taken with the ATLAS detector during 2015-2017 and correspond to an integrated luminosity of $79.8\,\text{fb}^{-1}$. Only events containing prompt photon candidates are selected by using single photon triggers, which are based on the *loose* photon ID criteria and have different E_T thresholds ranging from 10 GeV to 140 GeV with 5 GeV to 10 GeV intervals. Most of these triggers are prescaled, which reduces the trigger rate for photons with low E_T. Therefore, the data events are weighted by considering the trigger prescales, in order to obtain a continuous E_T distribution. A trigger matching is applied, requiring that the selected leading E_T photon in the event is close to the object that fired the trigger. Finally, data quality requirements are applied to make sure that all detector components were working properly during the data-taking.

In addition, MC samples for prompt and *fake* photons are used for the *tight* ID efficiency measurement. The prompt photon MC sample consists of $\gamma + jet$ events, which are generated with Pythia8 [62]. This sample is composed of several subsamples, which are produced with different requirements on the photon E_T, to ensure a sufficient number of simulated events over the whole E_T range. For the *fake* photon MC sample, a Pythia8 dijet sample composed of four subsamples is used, which covers a smaller E_T range up to ≈ 100 GeV. For both samples, the NNPDF2.3LO PDF set [134] and the A14 tune parameter set [71] are used. Additionally, both samples are processed by the full ATLAS simulation using GEANT4 [103, 104] and the same trigger requirements are applied to data and MC samples.

6.2.2 Single photon selection

Certain preselection requirements are applied to the photon candidates in order to reduce the fraction of *fake* photons in the analyzed dataset. In each event, the E_T leading photon is considered for the efficiency measurement. Only photons fulfilling *loose* ID requirements are considered, thereby rejecting a large fraction of *fake* photons. Kinematic requirements on the transverse energy, $E_\text{T} \geq 25$ GeV, and the pseudorapidity, $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, need to be fulfilled. In addition, track and calorimeter isolation criteria are applied to the photon candidate, in order to reduce the fraction of jets which are misidentified as photons. The calorimeter isolation is based on the sum of E_T of clusters in the EM calorimeter in a cone with $\Delta R = 0.2$ around the photon candidate, which is required to be smaller than $0.065 \cdot E_\text{T}$. For the track isolation, the sum of p_T of tracks with $p_\text{T} > 1$ GeV in a cone with $\Delta R = 0.2$ around the photon candidate is required to be smaller than $0.05 \cdot E_\text{T}$.

The efficiency is measured in four bins of $|\eta|$, taking into account the material distribution of the InDet: [0.00, 0.60), [0.60, 1.37), (1.52, 1.81), [1.81, 2.37). The amount of material in the InDet increases for certain $|\eta|$ regions, e.g. due to service panels such as cables and cooling pipes, and influences the fraction of converted photons and their shower shapes. In addition, the *tight* ID efficiency is measured as a function of E_T. The E_T distribution is split into 13 bins with an increasing width, to take into account the decreasing number of photon candidates at high E_T. In MC simulation, a truth matching is applied to the photons to make sure that only real or *fake* photons are selected in the respective MC samples.

6.2.3 Efficiency calculation with the Matrix Method

The efficiency calculation is performed using the sample of preselected photons. In general, the *tight* ID efficiency in the considered sample is defined as the fraction of prompt photons that fulfill the *tight* ID requirement, N^s_{ID}, among all prompt photons, N^s:

$$\varepsilon^{tight-ID} = \frac{N^s_{ID}}{N^s} \tag{6.1}$$

However, the considered sample does not only include prompt photons but also *fake* photons. Therefore, the total number of photons in the sample is given by the sum of prompt photons, N^s, and *fake* photons, N^b:

$$N^{total} = N^s + N^b. \tag{6.2}$$

The number of photons fulfilling the *tight* ID requirement, N^{total}_{ID}, can be defined analogously based on N^s_{ID} and N^b_{ID}. In order to consider the composition of the data sample, the *tight* ID efficiency can be expressed through the signal purity of the sample before applying the *tight* ID criteria, A, and after applying the *tight* ID criteria, P:

$$\varepsilon^{\text{tight-ID}} = \frac{P \cdot N^{total}_{ID}}{A \cdot N^{total}} \quad \text{with} \quad P = \frac{N^s_{ID}}{N^{total}_{ID}},\ A = \frac{N^s}{N^{total}} \tag{6.3}$$

To calculate the purities A and P, the fractions of prompt and *fake* photons in the considered sample need to be estimated. In order to distinguish between signal and background photons, a track isolation criterion is introduced which requires no tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ to be present in a cone of $\Delta R = 0.4$ around the photon, excluding tracks that are associated with the conversion of a photon. The *tight* ID efficiency can then be expressed through the track isolation efficiencies $\hat{\varepsilon}^i$, which are defined based on the fractions of photons fulfilling the additional track isolation criterion. Similar to Equation (6.2), the number of track isolated photons in the sample, $\hat{N}^{total}$, can be expressed through the number of prompt and *fake* photons and the respective track isolation efficiencies $\hat{\varepsilon}^s$ and $\hat{\varepsilon}^b$:

$$\hat{N}^{total} = \hat{\varepsilon}^s \cdot N^s + \hat{\varepsilon}^b \cdot N^b \tag{6.4}$$

The number of photons passing the *tight* ID and the track isolation criteria, $\hat{N}^{total}_{ID}$, is defined analogously, based on N^s_{ID}, N^b_{ID} and the track isolation efficiencies for photons fulfilling the *tight* ID requirement, $\hat{\varepsilon}^s_{ID}$ and $\hat{\varepsilon}^b_{ID}$. When solving these equations for N^s or N^s_{ID} and entering the new terms into the formulas of the purities A and P, these can be fully expressed through the track isolation efficiencies for prompt photons, *fake* photons and for the inclusive sample, $\hat{\varepsilon}$ and $\hat{\varepsilon}_{ID}$, as shown in Equations (6.5) and (6.6).

$$P = \frac{\hat{\varepsilon}_{ID} - \hat{\varepsilon}^b_{ID}}{\hat{\varepsilon}^s_{ID} - \hat{\varepsilon}^b_{ID}} \tag{6.5}$$

$$A = \frac{\hat{\varepsilon} - \hat{\varepsilon}^b}{\hat{\varepsilon}^s - \hat{\varepsilon}^b} \tag{6.6}$$

Finally, the *tight* ID efficiency can be expressed through the total number of photons in the sample before and after applying the *tight* ID requirements, N^{total} and N_{ID}^{total}, and the respective track isolation efficiencies, $\hat{\varepsilon}^{i}$ and $\hat{\varepsilon}_{ID}^{i}$, for prompt photon s, *fake* photons b, and for the inclusive sample, $\hat{\varepsilon} = \hat{N}/N^{total}$ and $\hat{\varepsilon}_{ID} = \hat{N}_{ID}/N_{ID}^{total}$, as shown in Equation (6.7).

$$\varepsilon^{tight-ID} = \frac{\frac{\hat{\varepsilon}_{ID}-\hat{\varepsilon}_{ID}^{b}}{\hat{\varepsilon}_{ID}^{s}-\hat{\varepsilon}_{ID}^{b}} \cdot N_{ID}^{total}}{\frac{\hat{\varepsilon}-\hat{\varepsilon}^{b}}{\hat{\varepsilon}^{s}-\hat{\varepsilon}^{b}} \cdot N^{total}} \tag{6.7}$$

This formula can be used to calculate the *tight* ID efficiency, as the different parameters can be retrieved either from data or from prompt photon MC samples. The number of photons before and after passing the *tight* ID criteria, as well as the track isolation efficiencies of the inclusive sample, $\hat{\varepsilon}$ and $\hat{\varepsilon}_{ID}$, are taken from data. The prompt photon track isolation efficiencies, $\hat{\varepsilon}^{s}$ and $\hat{\varepsilon}_{ID}^{s}$, are retrieved from the prompt photon MC samples.

The *fake* photon track isolation efficiencies are estimated from *fake*-photon-enriched regions in data, in order to avoid a bias due to the modeling of *fake* photons in MC simulation. Furthermore, no MC samples exist that provide information about *fake* photons over the whole considered E_{T} range. The *fake*-photon-enriched regions are defined by splitting the shower shape variables into two categories. The so-called *narrow-strip* variables, f_{side}, $\omega_{s\,3}$, ΔE and E_{ratio}, are constructed from information collected in the finely segmented first layer of the EM calorimeter. The remaining shower shape variables are the so-called *relaxed-tight* variables, R_{had}, R_{η}, R_{ϕ}, ω_2 and $\omega_{s\,\mathrm{tot}}$, which are mostly defined from information in the middle layer of the EM calorimeter and the leakage into the hadronic calorimeter. In the following, the fact that the *narrow-strip* variables are only weakly correlated to the track isolation criterion is exploited. *Fake*-photon-enriched regions in data are constructed by requiring the photons to either

- pass the *tight* ID (region 1),
- pass only the *narrow-strip* requirements (region 2),
- pass only the *relaxed-tight* requirements (region 3) or
- fail both the *relaxed-tight* and *narrow-strip* requirements (region 4).

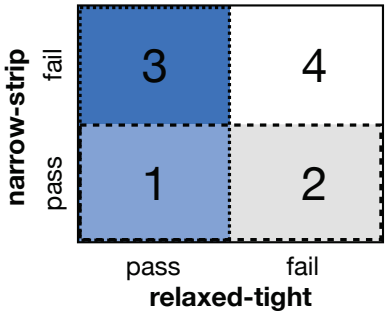


Figure 6.1: Illustration of the four regions defined in data, among which the *fake*-photon-enriched regions 2, 3 and 4 are used for the estimation of the *fake* photon track isolation efficiencies, $\hat{\varepsilon}^{b}$ and $\hat{\varepsilon}_{ID}^{b}$. The definition of the regions is based on the shower shape variables, which are split into *narrow-strip* variables (f_{side}, $\omega_{s\,3}$, ΔE, E_{ratio}) and *relaxed-tight* variables (R_{had}, R_{η}, R_{ϕ}, ω_2, $\omega_{s\,\mathrm{tot}}$).

The four different regions are illustrated in Figure 6.1. Due to the small correlation of the *narrow-strip* variables and the track isolation, the track isolation efficiency for *fake* photons is assumed to be similar for photons passing the *tight* ID requirement and photons passing the subset of *relaxed-tight* criteria. Hence, the track isolation efficiency for *fake* photons fulfilling the *tight* ID requirement can be estimated from the *fake*-photon-enriched region 3:

$$\hat{\varepsilon}^b_{ID} = \frac{\hat{N}^b_1}{N^b_1} \approx \frac{\hat{N}^b_3}{N^b_3} \tag{6.8}$$

Similarly, the track isolation efficiency for all *fake* photons can be estimated from the *fake*-photon-enriched regions 2, 3 and 4, since the number of *fake* photons in region 1 is negligible compared to those in regions 2, 3 and 4:

$$\hat{\varepsilon}^b = \frac{\hat{N}^b_{1+2+3+4}}{N^b_{1+2+3+4}} \approx \frac{\hat{N}^b_{2+3+4}}{N^b_{2+3+4}} \tag{6.9}$$

The signal leakage into these *fake*-photon-enriched regions is estimated from prompt photon MC simulation and subtracted. The assumptions made in Equations (6.8) and (6.9) are validated using *fake* photon MC samples and a systematic uncertainty is defined, based on the differences of the *fake* photon track isolation efficiencies in the different regions observed in *fake* photon MC simulation.

6.2.4 Correction of the measured *tight* ID efficiencies

The *tight* ID efficiency, which is calculated based on Equation (6.7), needs to be corrected for the *loose* ID and trigger requirements, which are imposed during the preselection. While these requirements are unavoidable in data in order to reduce the large fraction of *fake* photons, prompt photon MC simulation can be used to estimate the efficiency of these preselection requirements. In order to obtain the *tight* ID efficiency for the full sample of reconstructed photons, the measured *tight* ID efficiencies, ε_{data}, are multiplied with a correction factor C as a function of E_{T} in each $|\eta|$ region for converted and unconverted photons, respectively, as shown in Equation (6.10).

$$\varepsilon^{\text{tight-ID}} = C \cdot \varepsilon_{data} = C \cdot \frac{N_\gamma^{\text{tight-ID,trigger}}}{N_\gamma^{\text{loose-ID,trigger}}} \tag{6.10}$$

$$C = \frac{\frac{N^{\text{tight-ID}}_{\gamma,MC}}{N_{\gamma,MC}}}{\frac{N^{\text{tight-ID,trigger}}_{\gamma,MC}}{N^{\text{loose-ID,trigger}}_{\gamma,MC}}} = \frac{\frac{N^{\text{loose-ID,trigger}}_{\gamma,MC}}{N_{\gamma,MC}}}{\frac{N^{\text{tight-ID,trigger}}_{\gamma,MC}}{N^{\text{tight-ID}}_{\gamma,MC}}} \tag{6.11}$$

The correction factor is defined as shown in Equation (6.11) and is derived using prompt photon MC simulation. The numerator in Equation (6.11) is defined based on the fraction of prompt photons fulfilling the *tight* ID requirement, $N^{\text{tight-ID}}_{\gamma,MC}/N_{\gamma,MC}$, which corresponds to the value of interest. The denominator is defined as the fraction of photons fulfilling *tight*

ID and trigger requirements among the photons fulfilling *loose* ID and trigger requirements, $N_{\gamma,MC}^{\text{tight-ID,trigger}}/N_{\gamma,MC}^{\text{loose-ID,trigger}}$, which is similar to ε_{data}. The correction factor can be rewritten as the ratio of the *loose* ID and trigger efficiency for all photons and the trigger efficiency for photons fulfilling the *tight* ID requirement.

The correction factors depend on the photon E_{T} and get very close to 1 with increasing E_{T}, as shown in Figure 6.2. As the correction factors are determined using prompt photon MC samples, differences between the *loose* ID and trigger efficiencies in data and MC simulation may occur, which are accounted for by introducing a systematic uncertainty on the correction factor. The single photon triggers, which are used in this measurement, are only applied in regions of E_{T} where their efficiency is very close to 100% for both data and MC simulation [135]. Therefore, the differences between the trigger efficiencies in data and MC simulation are considered to be negligible.

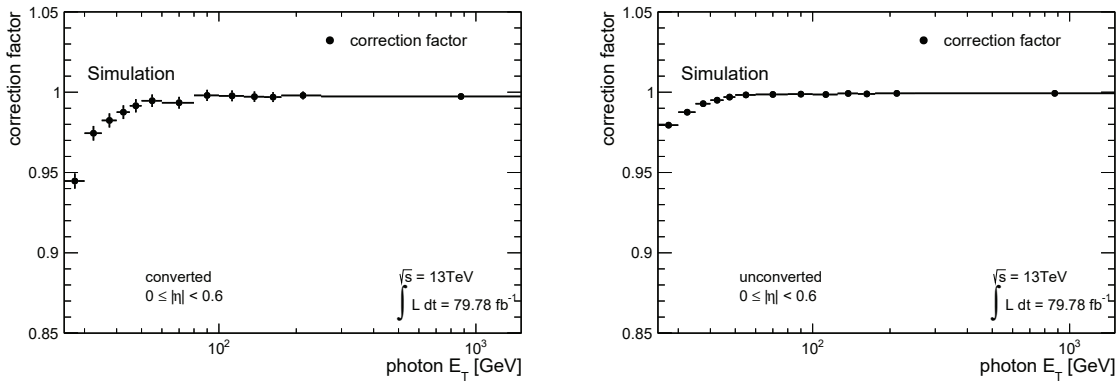


Figure 6.2: Correction factors in the most central $|\eta|$ region for converted and unconverted photons, which are used to correct the measured *tight* ID efficiencies for the trigger and *loose* ID preselection requirements. The uncertainties include the statistical uncertainty and the systematic uncertainty due to differences of the *loose* ID efficiency in data and MC simulation.

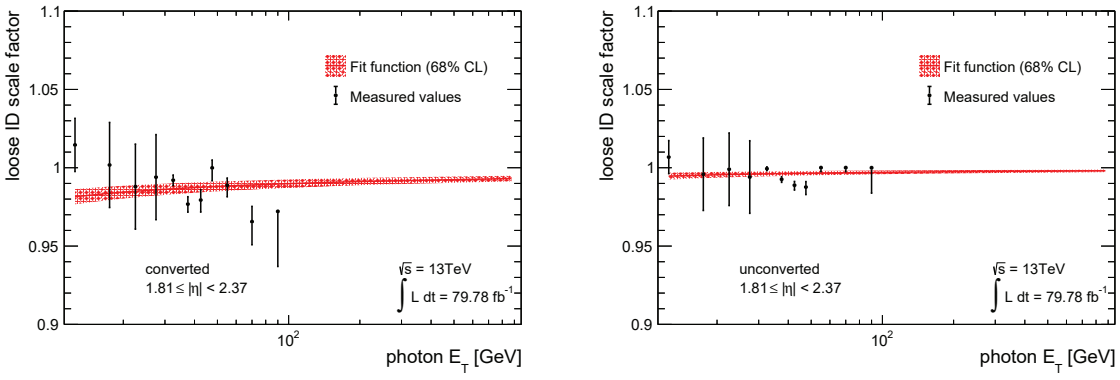


Figure 6.3: Fit to the *loose* ID SFs, which were evaluated by the Radiative Z method, for the most forward $|\eta|$ region for converted and unconverted photons, respectively. The fit function is defined as $\alpha/log(E_T) + 1$, with α being the only free parameter. The largest difference of the fit function, including the 68% confidence level interval, from 1 is considered as systematic uncertainty in each bin of E_{T} and $|\eta|$.

Hence, the systematic uncertainty on the correction factor only takes into account differences between data and MC simulation of the *loose* ID requirement. Deriving SFs for the *loose* ID requirement is challenging, due to the large fraction of *fake* photons in the sample of photons fulfilling the *loose* ID. The Radiative Z method allows for a measurement of the *loose* ID SFs in a region between $10\,\text{GeV} \leq E_\text{T} \leq 100\,\text{GeV}$. However, these SFs cover a much smaller E_T range than the range considered by the Matrix Method and show relatively large fluctuations. Therefore, they are not directly applied to the prompt photon MC sample. Instead, a fit is performed to the *loose* ID SFs, using a function of the form $\alpha/log(E_T) + 1$ with only one free parameter, to compensate the fluctuations of the *loose* ID SFs. The functional form is chosen so that it represents the expected behavior of the *loose* ID SFs. At high E_T, the shower shape variables get narrower and the *loose* ID efficiency is very close to 100%, therefore differences between data and MC simulation are expected to decrease and the *loose* ID SFs get close to 1. The chosen fit function slowly approaches to the expected value of 1, as shown in Figure 6.3. The largest difference between the fit function, including its 68% confidence level interval, and the assumed value of 1 is considered as systematic uncertainty in each bin of E_T and $|\eta|$. The distributions of the correction factors and the fits to the *loose* ID SFs for the remaining $|\eta|$ regions are shown in Appendix A.1.

6.2.5 Sources of uncertainty

Different sources of uncertainty are considered in this measurement, which are explained in the following. The statistical uncertainty of the considered dataset is evaluated by performing pseudoexperiments. A pseudo dataset is generated by taking a random number N_{events} from the considered dataset, where N_{events} is assumed to follow a Poisson distribution with the mean being the number of events observed in data. For each of these pseudoexperiments, the photon ID efficiency measurement is repeated. The *tight* ID efficiencies for the different pseudoexperiments result in a Gaussian distribution for each $|\eta|$ region and each E_T bin. The width of the Gaussian distribution is assigned as statistical uncertainty to the *tight* ID efficiency in the respective bin. In total, 200 pseudoexperiments are performed to estimate this uncertainty.

In addition to the statistical uncertainty of the analyzed dataset, various sources of systematic uncertainties are considered:

- **MC statistics:** A systematic uncertainty is assigned due to the limited size of the prompt photon MC sample, which is used to evaluate the *tight* ID efficiency. A number of 200 pseudoexperiments is generated from the prompt photon MC sample and the *tight* ID efficiency is evaluated for each pseudoexperiment. The uncertainty is then defined similar to the statistical uncertainty of the dataset.
- **Track isolation uncertainty:** In order to take into account the influence of the choice of the track isolation criterion on the efficiency measurement, the measurement is repeated with a track isolation criterion with a different cone size. The nominal track isolation requirement corresponds to the most stringent setup, therefore the cone size is reduced from $\Delta R = 0.4$ to $\Delta R = 0.2$, which corresponds to a loosened track isolation criterion. The uncertainty is given by the difference of the *tight* ID efficiency for the nominal case and the *tight* ID efficiency evaluated with the varied setup in each bin of E_T and $|\eta|$.
- **Closure uncertainty:** An additional uncertainty is assigned due to the assumptions which are made to estimate the track isolation efficiencies of *fake* photons from

fake-photon-enriched regions in data. The relative differences of the track isolation efficiencies in region 1 and 3 and in regions 1+2+3+4 and 2+3+4, as defined in Equation (6.12), are evaluated using *fake* photon MC samples.

$$\Delta\hat{\varepsilon}^b_{ID} = \frac{|\hat{\varepsilon}^b_1 - \hat{\varepsilon}^b_3|}{\hat{\varepsilon}^b_1}, \;\; \Delta\hat{\varepsilon}^b = \frac{|\hat{\varepsilon}^b_{1+2+3+4} - \hat{\varepsilon}^b_{2+3+4}|}{\hat{\varepsilon}^b_{1+2+3+4}} \tag{6.12}$$

The relative differences are evaluated over the whole E_{T} range for each $|\eta|$ region, due to the large statistical fluctuations in the *fake* photon MC sample. The *fake* photon track isolation efficiencies are varied up and down by these differences during the calculation of the *tight* ID efficiency. The down variation of $\hat{\varepsilon}^b$ results in a lower *tight* ID efficiency, while the up variation increases the efficiency. For the track isolation efficiency of *fake* photons fulfilling the *tight* ID requirement, $\hat{\varepsilon}^b_{ID}$, the up and down variations result in the opposite behavior. Finally, the difference of each of the four variations and the nominal *tight* ID efficiency is evaluated, and the variations which result in an increase or a decrease of the efficiency are added in quadrature, respectively. The largest sum of variations is defined as the closure uncertainty in each bin of E_{T} and $|\eta|$. In Figure 6.4, the *fake* photon track isolation efficiencies estimated from the different regions in MC simulation and from data are shown, as an example, for the most central $|\eta|$ region for converted photons. Similar distributions for the remaining regions are shown in Appendix A.2. The efficiencies agree with each other within statistical uncertainties, therefore, the method to estimate $\hat{\varepsilon}^b$ and $\hat{\varepsilon}^b_{ID}$ is considered to be validated.

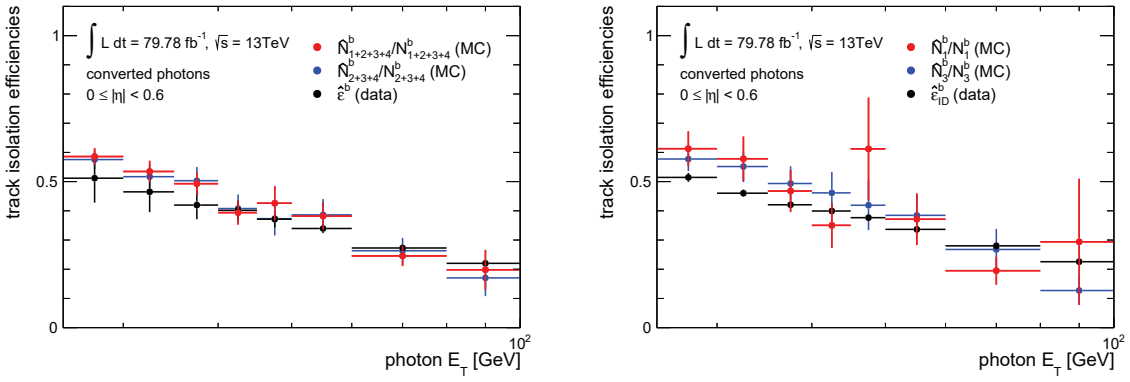


Figure 6.4: Track isolation efficiencies for *fake* photons measured in data and estimated from *fake* photon MC samples for the regions defined in data.

- **Fudge factor uncertainty:** This systematic uncertainty takes into account the uncertainties on the FFs, which are used to correct the shower shape distributions for the differences observed between data and MC simulation, and is applied in the Matrix Method for the first time. Four variations of the FFs are applied to the prompt photon MC sample. For each of these variations, a different set of shower shape variables is varied in order to take into account their correlations: $\{R_{\mathrm{had}}\}$, $\{R_\phi\}$, $\{R_\eta,\, \omega_2\}$ and $\{\omega_{s\,3},\, \omega_{s\,\mathrm{tot}},\, f_{\mathrm{side}}\}$. The FF uncertainties are added to the shower shape variables before applying the nominal *tight* ID requirements to the varied shower shape variables. Then, the *tight* ID efficiencies are calculated using the prompt photon MC sample with the

varied shower shape variables and the difference between the varied setup and the nominal *tight* ID efficiency is considered as systematic uncertainty. The absolute values of the uncertainties of the four different FF variations are finally added in quadrature. An example for the breakdown of the total FF uncertainty into the different FF variations is shown in Figure 6.5, while the distributions for the remaining $|\eta|$ regions are shown in Appendix A.3.

- **Correction factor uncertainty:** The uncertainty on the correction factor is directly translated into an uncertainty on the *tight* ID efficiency.

Finally, the total uncertainty is evaluated by adding the individual sources of uncertainty in quadrature. In Figure 6.6, the different contributions to the total uncertainty of the *tight* ID efficiency are shown for the most central $|\eta|$ region for converted and unconverted photons, respectively. The distributions for the remaining $|\eta|$ regions are shown in Appendix A.4. The dominant uncertainties at low E_{T} are the closure uncertainty and the track isolation uncertainty. The total uncertainty strongly decreases with increasing E_{T}, as the closure uncertainty becomes less important due to the smaller *fake* photon contribution at higher E_{T}. At high E_{T}, the MC statistics and correction factor uncertainties become more important.

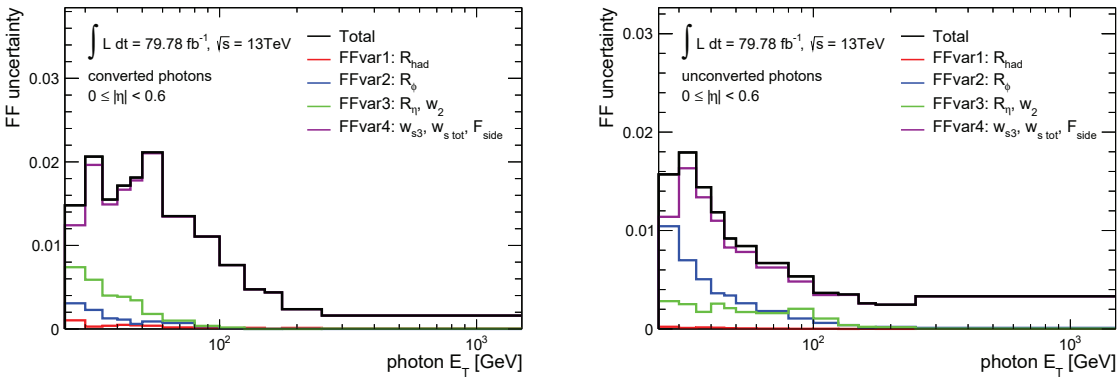


Figure 6.5: Uncertainty resulting from the four different FF variations for the most central $|\eta|$ region for converted and unconverted photons. The total uncertainty is derived by adding the absolute values of the four uncertainties in quadrature.

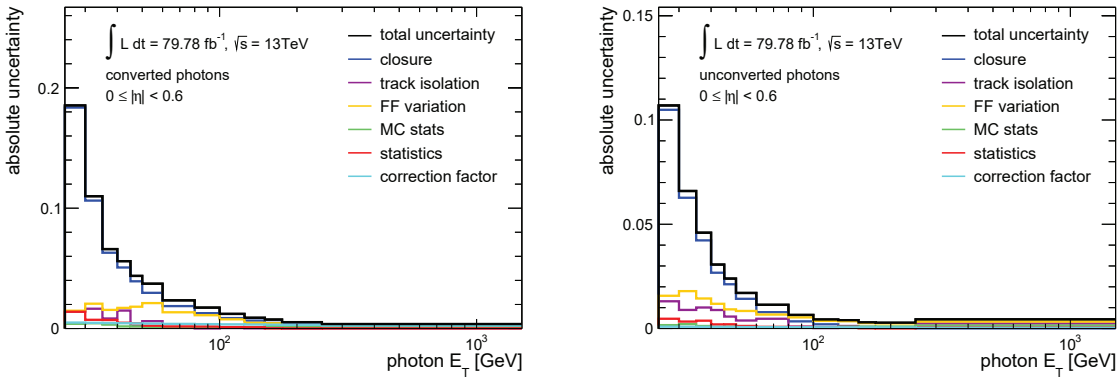
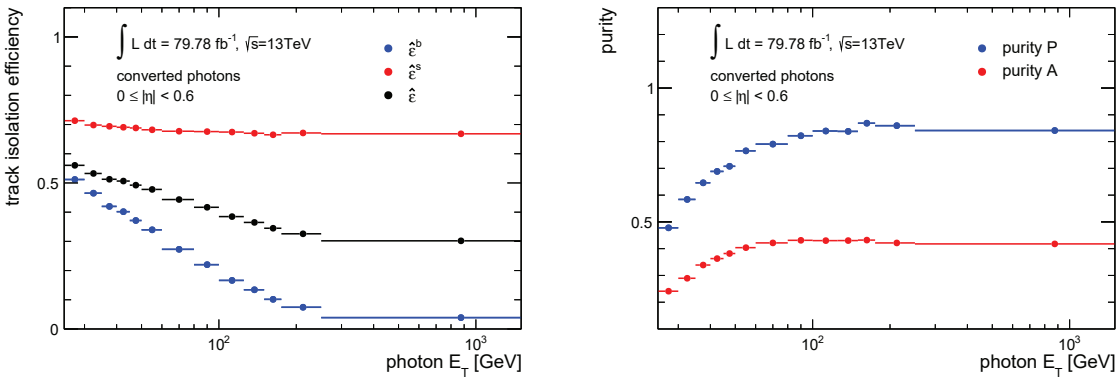


Figure 6.6: Uncertainty breakdown for the *tight* ID efficiency measurement for the most central $|\eta|$ region for converted and unconverted photons.

6.2.6 Measured *tight* ID efficiencies

In order to calculate the *tight* ID efficiencies, the track isolation efficiencies for prompt photons, *fake* photons and the inclusive sample are estimated from prompt photon MC simulation, *fake*-photon-enriched regions in data and the whole dataset, respectively. An example for the distributions of the track isolation efficiencies and the corresponding purities, A and P, is shown in Figure 6.7. While the track isolation efficiency for prompt photons is almost independent from the photon E_{T}, for *fake* photons it decreases with increasing E_{T}. As the individual objects inside a jet become more collimated for higher E_{T}, the probability for a photon inside a jet to be track isolated decreases. As the considered dataset is composed of prompt and *fake* photons, the track isolation efficiency for the whole dataset is composed of both components. The purities A and P increase with the photon E_{T}, due to the decreasing fraction of *fake* photons.



a) b)

Figure 6.7: a) Track isolation efficiencies for the inclusive sample and *fake* photons, retrieved from data, and for prompt photons, estimated from MC simulation, b) purities of the sample before and after applying the *tight* ID requirements, A and P. These plots are examples showing the most central $|\eta|$ region for converted photons.

From these purities, the *tight* ID efficiencies can be calculated, which are shown in Figure 6.8 for the most central $|\eta|$ region for converted and unconverted photons, respectively. In addition to the *tight* ID efficiencies measured in data, the *tight* ID efficiencies determined using prompt photon MC samples are shown. The blue markers show the *tight* ID efficiency in prompt photon MC simulation with the FF corrections applied, while the red markers show the corresponding efficiencies without applying the FFs. It can be nicely seen that the agreement of the data and MC efficiencies improves when the FF-based MC corrections are applied, especially at low E_{T}. The uncertainties of the *tight* ID efficiencies for converted photons, which are mostly dominated by the closure uncertainty, are larger than for unconverted photons. The *tight* ID efficiencies for both, converted and unconverted photons, increase for higher E_{T} until approximately 100 GeV. For higher values of E_{T}, the *tight* ID efficiency is almost constant. The *tight* ID efficiencies for unconverted photons range from $(84.3 \pm 2.9)\%$ to $(96.7 \pm 2.2)\%$ for the first to the last E_{T} bin, while for converted photons it ranges from $(80.3 \pm 13.9)\%$ to $(97.8 \pm 0.3)\%$. Due to the good agreement of data and the corrected prompt photon MC simulation, the final *tight* ID SFs are very close to 1, as shown in Figure 6.9. The *tight* ID efficiencies and SFs for the remaining $|\eta|$ regions are shown in Appendix A.5.

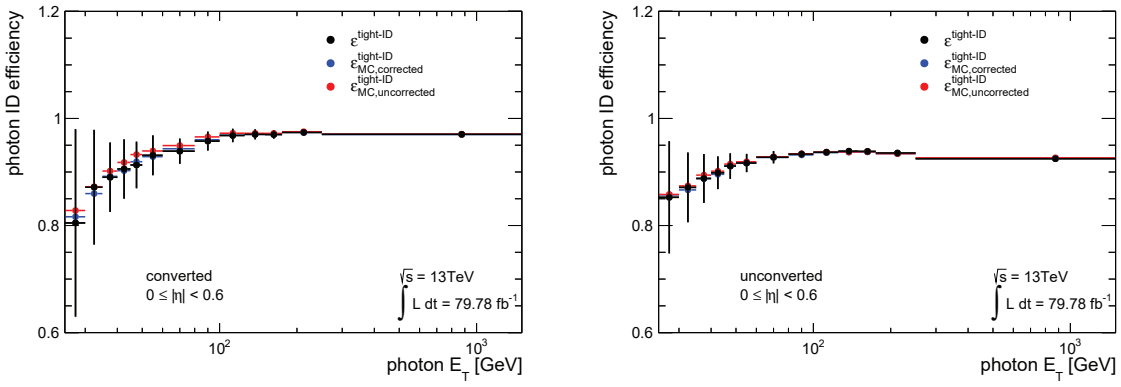


Figure 6.8: *Tight* ID efficiencies measured in 2015-2017 data with the Matrix Method for converted and unconverted photons in the most central $|\eta|$ region. The error bars include all considered uncertainties as explained in Section 6.2.5. The efficiencies measured in data are compared to the efficiencies of the prompt photon MC simulation with and without corrections applied to the shower shape variables.

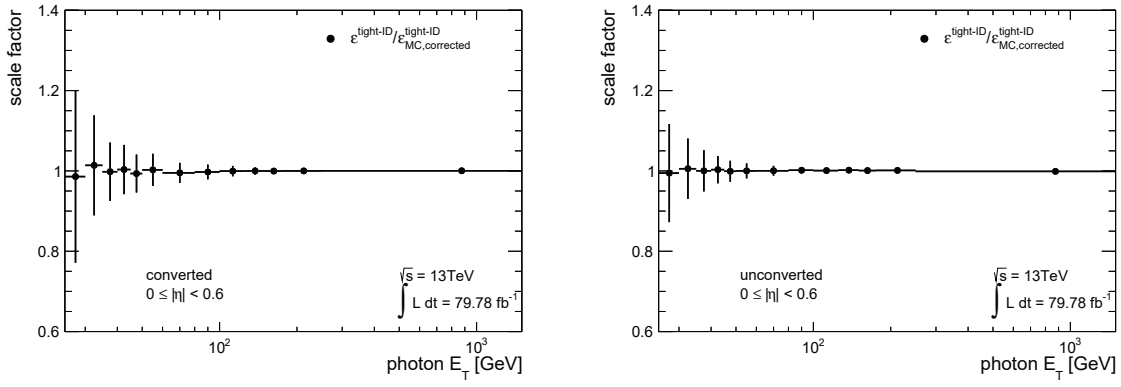


Figure 6.9: Scale factors measured in 2015-2017 data with the Matrix Method for converted and unconverted photons in the most central $|\eta|$ region.

Chapter 7

Analysis strategy in the $H \to \gamma\gamma$ channel

Measurements in the $H \to \gamma\gamma$ decay channel benefit from the good photon energy resolution of the ATLAS detector and allow to measure the Higgs boson signal as a narrow peak on top of a falling background spectrum. Measurements of different quantities in the $H \to \gamma\gamma$ decay channel are presented in this thesis, which share a similar analysis strategy. In the following, the common analysis strategy based on an unbinned likelihood fit to the invariant mass of the diphoton system is presented, including the approaches for defining a signal and background model, and the statistical model which is used to quantify the sensitivity of an analysis is explained.

7.1 Unbinned likelihood fit to $m_{\gamma\gamma}$

The invariant mass of the two photons originating from the Higgs boson decay, $m_{\gamma\gamma}$, is used as discriminant between the Higgs boson signal and background processes for most measurements in the $H \to \gamma\gamma$ decay channel. There are only a few exceptions of analyses where other variables offer a better discrimination, among these are primarily BSM searches for heavy new particles. The diphoton invariant mass is calculated based on the four-momenta of the two p_{T}-leading photons in each event, which fulfill *tight* ID and isolation requirements. The $m_{\gamma\gamma}$ distribution is characterized by a narrow Higgs boson signal peak, which is located around the Higgs boson mass of 125.09 GeV and has a width of approximately 3 GeV, depending on the topology of the selected events. The Higgs boson signal peak is located above the so-called continuum background, which is characterized by a monotonically decreasing spectrum. The dominant contribution to the continuum background are events with two high-p_{T} photons which do not originate from Higgs boson decays. Additional small contributions arise from events with jets that are misidentified as photons.

A large number of processes provide contributions to the continuum background, so that it is challenging to obtain a sufficient prediction based on MC samples. Instead, its contribution is determined from a fit to data. An unbinned likelihood fit to $m_{\gamma\gamma}$ with a signal and a background component is performed to simultaneously estimate the Higgs boson signal and the continuum background. Prior to that, models for the signal and the continuum background need to be defined. While the signal model is determined based on Higgs boson MC samples, the procedure for choosing a background model is more complicated and based on background templates which are derived from diphoton MC samples and data control regions. The choice of background model needs to be well motivated, as it might introduce a significant bias on the fitted Higgs boson signal. Hence, a general strategy for the background modeling was developed

by the ATLAS collaboration. In the final signal+background (S+B) fit, the parameters of the signal model are fixed, except for its normalization, while the parameters of the selected background model are free. An example for the $m_{\gamma\gamma}$ distribution for inclusive Higgs boson production in the $H \to \gamma\gamma$ decay channel is shown in Figure 7.1, illustrating the Higgs boson signal and continuum background contributions which are estimated from an S+B fit to the $m_{\gamma\gamma}$ distribution observed in data. The $m_{\gamma\gamma}$ region between 120 and 130 GeV, which contains the main proportion of the Higgs boson signal, is usually blinded during the optimization of an analysis. However, the $m_{\gamma\gamma}$ regions outside of this range, referred to as the data sidebands, can be used for studies of the continuum background.

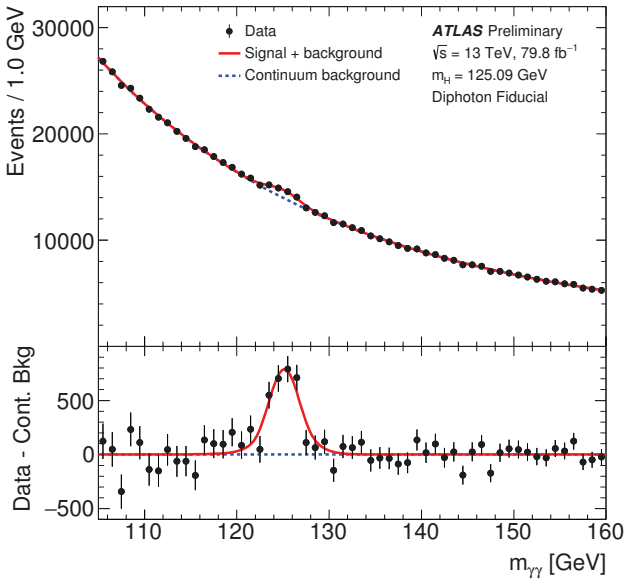


Figure 7.1: The diphoton invariant mass distribution for inclusive Higgs boson production in the $H \to \gamma\gamma$ decay channel using the 2015-2017 dataset [4]. The fitted continuum background is shown as a dashed blue line, while the result from the S+B fit is indicated by a continuous red line, illustrating the narrow Higgs boson signal peak on top of the continuum background.

7.2 Signal modeling

The shape of the $H \to \gamma\gamma$ resonance, as expected from theory, is a Breit-Wigner distribution with a very narrow width of about 4 MeV. However, the width and the shape of the observed $m_{\gamma\gamma}$ distributions differ from this theory expectation due to experimental limitations, such as the photon energy resolution of the detector, resulting in a much broader asymmetric signal distribution. Hence, the Higgs boson signal is modeled with a double-sided Crystal Ball (DCB) function. The DCB function is characterized by a Gaussian core, whereas the tails of the distribution follow a power law distribution. This functional form accounts for converted photons where the electrons from the conversion loose a significant amount of their energy through bremsstrahlung in the InDet, which causes the tails for the $m_{\gamma\gamma}$ distribution to be non-Gaussian. The DCB function is defined as shown in Equation (7.1).

$$f_{DCB}(m_{\gamma\gamma}) = N \times \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{low} \leq t \leq \alpha_{high}, \\ e^{-\frac{1}{2}\alpha_{low}^2} \left[\frac{1}{R_{low}}(R_{low} - \alpha_{low} - t)\right]^{-n_{low}} & \text{if } t < -\alpha_{low}, \\ e^{-\frac{1}{2}\alpha_{high}^2} \left[\frac{1}{R_{high}}(R_{high} - \alpha_{high} + t)\right]^{-n_{high}} & \text{if } t > \alpha_{high}. \end{cases} \tag{7.1}$$

The parameter t is defined as $t = (m_{\gamma\gamma} - \mu_{CB})/\sigma_{CB}$, where μ_{CB} and σ_{CB} describe the mean and the width of the Gaussian distribution, respectively. The normalization of the function is accounted for by the parameter N, while the remaining parameters are properties of the functions describing the non-Gaussian tails of the distribution: α_{low} and α_{high} are defined by the positions where the Gaussian distribution turns into the power law tails and n_{low} and n_{high} are the exponents of the power law functions describing the lower and upper tail, respectively. Based on these parameters, the ratios $R_{low} = \frac{\alpha_{low}}{n_{low}}$ and $R_{high} = \frac{\alpha_{high}}{n_{high}}$ are defined.

The parameters of the signal model are determined by fitting the DCB function to the $m_{\gamma\gamma}$ distribution predicted by MC simulations of all contributing Higgs boson production processes, scaled to their cross sections. The signal model needs to be evaluated for each event selection separately, as its parameters depend on the topology of the selected events. As the Higgs boson signal MC is produced for a Higgs boson mass of exactly 125 GeV, the fitted signal model is shifted by 90 MeV to match the measured value of the ATLAS and CMS combination.

An example for the signal modeling is shown in Figure 7.2 for a simple event selection requiring two photons and either at least one lepton or no leptons and exactly zero jets. It can be seen that the tails of this distribution are asymmetric and, hence, cannot be described by a simple Gaussian function. In addition, the different parameters of the DCB function are indicated, as well as the dependencies of the distribution in the central part and the tails, in order to illustrate the properties of the DCB function.

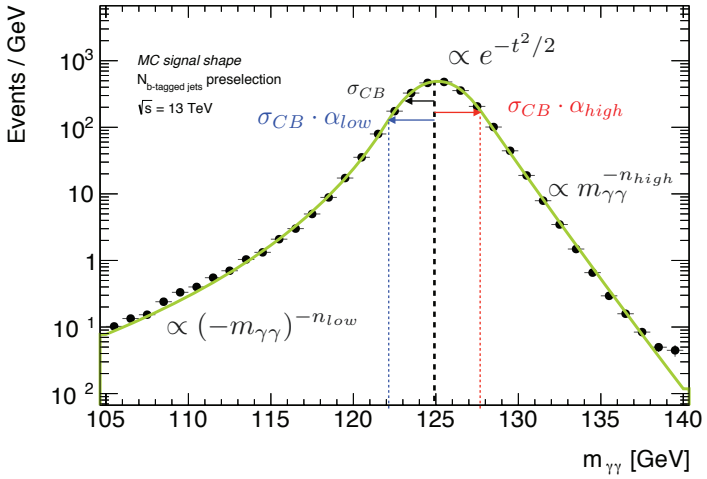


Figure 7.2: An example signal model determined from fitting a DCB function to MC simulation for a simple event selection, requiring two photons and either at least one lepton or no leptons and exactly zero jets. The core of the distribution follows a Gaussian function with the mean μ_{CB} and the width σ_{CB}, while the tails can be described by power law functions with the exponents n_{low} and n_{high}. The parameters α_{low} and α_{high} describe the positions where the transitions of the Gaussian function to the power law tails occur.

7.3 Background modeling

In the following, the strategy for choosing a background model is explained, which is composed of three different steps: First, the method is described which is used to estimate the fraction of real photons in the continuum background and the fraction of jets which are misidentified as photons. Then, these fractions are used to construct a background template. The shapes of the different contributions are extracted from diphoton MC samples and control regions in data. Finally, different functional forms are fitted to this background template to estimate the bias, the so-called spurious signal, which results from the choice of a specific function. Among these functional forms, the background model is chosen by minimizing this bias and the number of degrees of freedom of the considered functional forms.

7.3.1 Background decomposition

In order to construct a background-only (B-only) template, the relative contributions of the different components of the continuum background need to be estimated. These include the non-resonant production of a pair of photons ($\gamma\gamma$), the production of a photon and a *fake* photon ($\gamma j/j\gamma$) and the production of two jets which are misidentified as photons (jj). The relative fractions of the $\gamma\gamma$, $\gamma j/j\gamma$ and jj contributions are estimated with the data-driven $2 \times 2D$ sideband method [136, 137], by defining different orthogonal regions based on the *tight* ID and isolation requirements of the photons.

The two p_{T}-leading photons in each event are classified into one of four regions, depending on whether they pass or fail the *tight* ID and isolation requirements. The region A defines the signal region, where a photon passes the *tight* ID and isolation requirements, whereas region B denotes the isolation sideband where the photon fails the isolation requirement. Photons which fail the *tight* ID requirement are classified into the identification sideband region C, while the remaining photons are classified into region D. The definition of the four regions is illustrated in Figure 7.3.

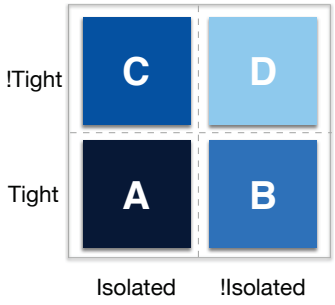


Figure 7.3: Illustration of the four regions which are defined by the $2 \times 2D$ sideband method for the p_{T}-leading and -subleading photon, respectively. The region A corresponds to the signal region where the photon fulfills *tight* ID and isolation requirements, while for the control regions either one of the requirements or both of them fail. Region B is referred to as the isolation sideband, C is the identification sideband and D is the combined sideband.

In total, there are 16 possibilities to assign the p_{T}-leading and the p_{T}-subleading photon to one of the regions, resulting in 16 equations describing the number of events in each region N_{ij}. The subscripts i and j indicate the region into which the leading and subleading photons

are classified, respectively. Only events where both photons pass a subset of the *tight* ID requirements are considered, which are referred to as *loose*' (*loose-prime*) events[1]. Hence, photons failing the *tight* ID are still required to fulfill the *loose*' requirement.

The event yield in each region can be expressed by introducing efficiencies for the leading and subleading photon to pass the *tight* ID and isolation requirements, fake rates for the leading and subleading jet in $\gamma j/j\gamma$ and jj events and various correlation coefficients. A simplified example for the event yield N_{AA}, where both photons are classified into region A, is presented in Equation (7.2). In this simplified example the correlation coefficients are not considered.

$$\begin{aligned} N_{AA} =& N_{\gamma\gamma}^{TI} + N_{\gamma j}^{TI} + N_{j\gamma}^{TI} + N_{jj}^{TI} \\ =& N_{\gamma\gamma}^{loose'} \varepsilon_{iso,\gamma_1}\varepsilon_{tight,\gamma_1}\varepsilon_{iso,\gamma_2}\varepsilon_{tight,\gamma_2} + N_{\gamma j}^{loose'} \varepsilon_{iso,\gamma_1}\varepsilon_{tight,\gamma_1} f_{iso,\gamma_2} f_{tight,\gamma_2} + \\ & N_{j\gamma}^{loose'} \varepsilon_{iso,\gamma_2}\varepsilon_{tight,\gamma_2} f_{iso,\gamma_1} f_{tight,\gamma_1} + N_{jj}^{loose'} f'_{iso,\gamma_1} f'_{tight,\gamma_1} f'_{iso,\gamma_2} f'_{tight,\gamma_2} \end{aligned} \tag{7.2}$$

The total number of events in data N_{AA} is given by the sum of the $\gamma\gamma$, γj, $j\gamma$ and jj contributions of events with photons fulfilling the *tight* ID and isolation requirements. These can be expressed through the number of events with photons fulfilling the *loose*' requirement for each contribution. The efficiencies for passing the *tight* ID or isolation requirement are given by $\varepsilon_{tight,\gamma_i}$ and $\varepsilon_{iso,\gamma_i}$ and are estimated from diphoton MC samples. The fake rates for the *tight* ID and isolation requirements are described by the parameters f_{tight,γ_i} and f_{iso,γ_i} for $\gamma j/j\gamma$ events and by f'_{tight,γ_i} and f'_{iso,γ_i} for jj events.

In total, 32 parameters are used to describe the event yields N_{ij} observed in data, from which some are estimated from diphoton MC and some parameters are fixed to a constant value, such as some of the correlation coefficients which are fixed to 1. The remaining free parameters in the set of 16 equations, which include the fake rates in $\gamma j/j\gamma$ and jj events, are determined by minimizing a χ^2, which is defined based on the number of observed and predicted events in each region N_{ij}.

Different sources of systematic uncertainty are assigned to the $\gamma\gamma$, $\gamma j/j\gamma$ and jj fractions, which are estimated through a variation of the *loose*' definition, a variation of the shower shape variables by using the uncorrected values, and a variation of the event generator by using Pythia8 instead of the nominal generator Sherpa. However, as the $\gamma\gamma$, $\gamma j/j\gamma$ and jj fractions are only used to construct the background templates for the selection of a background model, these uncertainties on the relative fractions are not considered in the following. For the inclusive $H \to \gamma\gamma$ sample, the $\gamma\gamma$, $\gamma j/j\gamma$ and jj continuum background components roughly result in fractions of 78%, 19% and 3%, respectively.

7.3.2 Background templates

Once the contributions of the different components of the continuum background are estimated, these can be used to construct the B-only template. To derive the $m_{\gamma\gamma}$ shape of the $\gamma\gamma$ contribution, the diphoton MC sample can be used, which in general offers a reasonable description of the shape. An exception occurs for final states with a large number of jets, where the diphoton MC sample does not provide enough events to sufficiently describe the continuum background. This is for example the case for tH and $t\bar{t}H$ final states, where an alternative

[1] The requirements on f_{side}, ΔE and $\omega_{s\,3}$ are dropped.

strategy is used to derive a background template. This alternative approach is discussed in Chapter 9.

For the $\gamma j/j\gamma$ (collectively referred to as γj in the following) and jj contributions, no MC samples exist that provide a sufficient number of events to extract the respective $m_{\gamma\gamma}$ shapes. Therefore, data control regions enriched with *fake* photons are defined. This is done by reversing the *tight* ID requirement and replacing it with the *loose*' requirement, which was introduced in the previous section. The *loose*' and isolation requirements are kept, to ensure that the selected *fake* photons have a signature similar to real photons. The reversed requirement is then applied to either the p_{T}-leading photon, the p_{T}-subleading photon or both photons to extract the γj and jj shapes, respectively. As the $m_{\gamma\gamma}$ distributions of these data control regions are in general still characterized by large statistical fluctuations, they are not directly transferred to the B-only template. Instead, a reweighting is applied to the $m_{\gamma\gamma}$ distribution of the $\gamma\gamma$ component, to match the shape of the γj and jj distributions. This is done by dividing the $m_{\gamma\gamma}$ distribution of the $\gamma\gamma$ template by the γj or the jj template derived from data, respectively, and fitting their ratio with a linear function. This linear function, which describes the slope difference between the different templates, is then used to derive weights which are applied to the $\gamma\gamma$ template.

Finally, the templates for the $\gamma\gamma$, γj and jj contributions are added according to the relative fractions which are derived as explained in the previous section. The overall normalization of this B-only template is estimated from the number of observed events in the $m_{\gamma\gamma}$ data sidebands.

7.3.3 Spurious signal test

For the final selection of a background model, an S+B fit is performed to the B-only template. The amount of fitted signal, which is referred to as the spurious signal, N_{spur}, is a measure for the bias which is introduced by the chosen model. It is considered as a systematic uncertainty in the final analysis.

Several different functions with varying complexity are considered for the background model. The simplest functions, such as an exponential function with only one degree of freedom, are expected to be an adequate choice for signal regions with a low number of expected events. The small number of degrees of freedom prevents the fit function from adapting to statistical fluctuations of the $m_{\gamma\gamma}$ distribution in data. In contrast, for regions with a large number of expected events, these simple functions do not sufficiently describe the continuum background and more complex functions are considered to be a reasonable choice. In general, the functional form with the smallest number of degrees of freedom which provides a reasonable description of the background is chosen. The following functional forms, which are arranged in increasing order of complexity, with the free parameters c_i are considered as background model:

- A simple exponential function: $f_{Exponential}(m_{\gamma\gamma}) = e^{(c_1 \cdot m_{\gamma\gamma})}$
- Power law function: $f_{PowerLaw}(m_{\gamma\gamma}) = m_{\gamma\gamma}^{c_1}$
- Dijet function: $f_{Dijet}(m_{\gamma\gamma}) = x^{c_1} \cdot (1-x)^{c_2}$ with $x = \left(\frac{m_{\gamma\gamma}}{13 \cdot 10^3\,\mathrm{GeV}}\right)$
- An exponential function of a polynomial of second order:

  $f_{ExpPoly2}(m_{\gamma\gamma}) = e^{(c_1 \cdot x^2 + c_2 \cdot x)}$ with $x = \left(\frac{m_{\gamma\gamma} - 100\,\mathrm{GeV}}{100\,\mathrm{GeV}}\right)$

- Bernstein polynomials of 3rd to 5th order:
  $B_N(m_{\gamma\gamma}) = \sum_{i=0}^{N} c_i \cdot b_{i,N}$ with $b_{i,N} = \binom{N}{i} m_{\gamma\gamma}^{i}(1 - m_{\gamma\gamma})^{N-i}$

To make sure that the chosen background model does not introduce a large bias on the fitted number of signal events, certain requirements need to be fulfilled in order to accept a functional form as a background model candidate. An S+B fit to the B-only template is performed as a function of $m_{\gamma\gamma}$ in the range between 121 and 129 GeV. For this S+B fit, the signal model which is determined from MC simulation, as described in this chapter, is assumed. The maximum amount of fitted spurious signal in this region, $|N_{\text{spur}}^{\text{max}}|$, is used as a systematic uncertainty in the final fit to $m_{\gamma\gamma}$. Two criteria on $|N_{\text{spur}}^{\text{max}}|$ are defined: $|N_{\text{spur}}^{\text{max}}|$ needs to be small compared to either the expected number of signal events, S_{ref}, as shown in Equation (7.3), or the expected statistical uncertainty on the signal yield, δS, as shown in Equation (7.4).

$$|N_{\text{spur}}^{\text{max}}|/S_{\text{ref}} < 10\% \tag{7.3}$$

$$|N_{\text{spur}}^{\text{max}}|/\delta S < 20\% \tag{7.4}$$

The function with the lowest number of degrees of freedom which fulfills one of these requirements is chosen as background model. If one of the requirements is fulfilled for two or more functions with the same number of degrees of freedom, the one which is causing the smaller bias is chosen.

In some cases, it is challenging to choose a background model with this procedure if the B-only template is characterized by large statistical fluctuations. In this case, the fitted spurious signal is dominated by statistical fluctuations in the template, resulting in large values for $|N_{\text{spur}}^{\text{max}}|$. As a consequence, it may happen that none of the functional forms passes the requirements in Equations (7.3) and (7.4) or only functions with a high number of degrees of freedom, which are not considered a reasonable choice.

For these cases, where the spurious signal test fails due to statistical fluctuations in the B-only template, relaxed spurious signal criteria were developed which take into account the statistical uncertainty of the background template. A new variable ζ_{spur} is defined, which replaces $N_{\text{spur}}^{\text{max}}$ in Equations (7.3) and (7.4), and compensates local $2\,\sigma$ fluctuations in the background template:

$$\zeta_{\text{spur}} = \begin{cases} N_{\text{spur}} + 2\Delta_{MC} & \text{if } N_{\text{spur}} + 2\Delta_{MC} < 0, \\ N_{\text{spur}} - 2\Delta_{MC} & \text{if } N_{\text{spur}} - 2\Delta_{MC} > 0, \\ 0 & \text{other cases.} \end{cases} \tag{7.5}$$

In this new definition, Δ_{MC} is the statistical uncertainty of the B-only template for a certain value of $m_{\gamma\gamma}$. An additional criterion is introduced for the relaxed spurious signal criteria to ensure that a function offers a sufficient description of the background template in the whole considered $m_{\gamma\gamma}$ region, requiring the χ^2 probability to be larger than 1%.

7.4 Statistical model

In this section, the statistical model which is used to describe and interpret the expected or observed data is explained. First, the structure of the likelihood function is described, which

is used to determine the parameter of interest (PoI) and its uncertainty based on a profile likelihood ratio. Then, the test statistic, which is based on this profile likelihood ratio, and its asymptotic approximation are presented. Finally, the procedure for estimating the expected limit and the expected significance based on a representative dataset, the so-called Asimov dataset, and the CL_S method are described.

7.4.1 Definition of the likelihood function

As explained in Section 7.1, the $m_{\gamma\gamma}$ distribution is described by a likelihood function with signal and background components. The likelihood function is defined based on the poisson probabilities in each category i or, in case of a differential cross section measurement, in each bin i of a distribution [2, 138]. For the optimization studies presented in this thesis, the likelihood function shown in Equation (7.6) is used, which only considers statistical uncertainties. This is considered to be a valid approach, as the statistical uncertainty is expected to be dominant for the considered processes in the $H \to \gamma\gamma$ decay channel.

$$L_i(\mu) = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \cdot \prod_{j=1}^{n_i} f_i(m_{\gamma\gamma}^j) \tag{7.6}$$

The likelihood function depends on the signal strength μ, which is the PoI in this example. It is defined as the ratio of the observed or hypothesized cross section for a specific process, σ_{obs}, and the cross section predicted by the SM: $\mu = \frac{\sigma_{obs}}{\sigma_{SM}}$. The expected signal yield s_i is scaled by μ in order to measure deviations from the SM expectation. The expected number of background events is given by b_i, while n_i describes the observed number of events. The function $f_i(m_{\gamma\gamma}^j)$, defined in Equation (7.7), is the probability density function ($p.d.f.$) of $m_{\gamma\gamma}$ for each event j, which is composed of the signal component, f_i^s, and the background component, f_i^b.

$$f_i(m_{\gamma\gamma}^j) = \frac{\mu s_i \cdot f_i^s(m_{\gamma\gamma}^j) + b_i \cdot f_i^b(m_{\gamma\gamma}^j)}{\mu s_i + b_i} \tag{7.7}$$

While for the optimization studies it is sufficient to consider only statistical uncertainties, for the final analysis several systematic uncertainties need to be taken into account. These are defined as a set of nuisance parameters (NPs) $\boldsymbol{\theta}$ in the likelihood function. The parameters which are affected by a systematic uncertainty of the size σ are multiplied with a response factor $F_G(\sigma, \theta) = (1 + \sigma \cdot \theta)$, in the case of a Gaussian constraint, or $F_{LN}(\sigma, \theta) = e^{\sqrt{\ln(1+\sigma^2)}\theta}$ in the case of a log-normal constraint. In addition, the NPs enter the constraints term $G(\boldsymbol{\theta})$, which is defined as the product of unit Gaussian constraints that are introduced for each NP. The complete likelihood function, which considers the systematic uncertainties, is shown in Equation (7.8).

$$L_i(\mu, \boldsymbol{\theta}) = \frac{(\mu s_i + b_i + n_i^{\text{spur}} \cdot \theta_i^{\text{spur}})^{n_i}}{n_i!} e^{-(\mu s_i + b_i + n_i^{\text{spur}} \cdot \theta_i^{\text{spur}})} \cdot \prod_{j=1}^{n_i} f_i(m_{\gamma\gamma}, \theta_i^{\text{spur}}, \boldsymbol{\theta_i^s}, \boldsymbol{\theta_i^b}) \cdot G(\boldsymbol{\theta}) \tag{7.8}$$

The expected event yield in the Poisson probability term is modified by adding the spurious signal uncertainty, given by $n_i^{\text{spur}} \cdot \theta_i^{\text{spur}}$, as an additional component. Here, the absolute value of the maximum amount of fitted spurious signal is denoted by n_i^{spur}. Additional NPs are considered by the *p.d.f.* describing the $m_{\gamma\gamma}$ shape, $f_i(m_{\gamma\gamma}^j, \theta_i^{\text{spur}}, \boldsymbol{\theta_i^s}, \boldsymbol{\theta_i^b})$, as shown in Equation (7.9). These include NPs affecting the signal shape, $\boldsymbol{\theta_i^s}$, which describe uncertainties affecting the width and the position of the DCB function, and NPs affecting the background shape, $\boldsymbol{\theta_i^b}$, which include the free parameters of the background model. The nuisance parameters related to the background model do not enter the constraints term. For a combined fit of various categories or the different bins of a differential distribution, the product of the individual likelihood functions L_i is considered.

$$f_i(m_{\gamma\gamma}^j, \theta_i^{\text{spur}}, \boldsymbol{\theta_i^s}, \boldsymbol{\theta_i^b}) = \frac{(\mu s_i + n_i^{\text{spur}} \cdot \theta_i^{\text{spur}}) \cdot f_i^s(m_{\gamma\gamma}^j, \boldsymbol{\theta_i^s}) + b_i \cdot f_i^b(m_{\gamma\gamma}^j, \boldsymbol{\theta^b})}{\mu s_i + b_i + n_i^{\text{spur}} \cdot \theta_i^{\text{spur}}} \tag{7.9}$$

7.4.2 Profile likelihood based test statistic

A test statistic, based on a profile likelihood ratio, is defined which allows to estimate the uncertainty on the PoI and to test different values of μ for the calculation of expected limits or the expected significance. The test statistic, which is used for the measurements presented in this thesis, has the form:

$$t_\mu = -2 \ln \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{7.10}$$

The denominator of the profile likelihood ratio describes the unconditional likelihood function which is maximized for the values $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. The numerator $L(\mu, \hat{\hat{\boldsymbol{\theta}}})$ describes the conditional likelihood function, which is maximized by $\hat{\hat{\boldsymbol{\theta}}}$ for a hypothesized value of μ. The profile likelihood ratio $L(\mu, \hat{\hat{\boldsymbol{\theta}}})/L(\hat{\mu}, \hat{\boldsymbol{\theta}})$ adopts values between 0 and 1, where a value of 1 corresponds to a good agreement between the estimator $\hat{\mu}$ and the tested value μ. For the entire test statistic, increasing values of t_μ correspond to a decreasing compatibility.

The test statistic for a hypothesized value of μ, for a dataset which is characterized by a signal strength μ', can be expressed through the asymptotic approximation [138], assuming that $\hat{\mu}$ follows a Gaussian distribution with a mean μ' and a width σ_μ:

$$t_\mu = \frac{(\mu - \hat{\mu})^2}{\sigma_\mu^2} + \mathcal{O}(1/\sqrt{N}), \tag{7.11}$$

where $\mathcal{O}(1/\sqrt{N})$ can be neglected for large sample sizes N. The test statistic t_μ can then be interpreted as a non-central χ^2-distribution. This allows to easily construct confidence intervals of the profile likelihood test statistic, in order to estimate the uncertainty on the PoI. Furthermore, a so-called Asimov dataset can be defined based on the assumption $\hat{\mu} = \mu'$, such that the estimators are equal to the expected values for this dataset. Using the asymptotic approximation and the Asimov dataset allows to easily estimate the *p.d.f.* of the test statistic $f(t_\mu|\mu')$ for a hypothesized value μ. This is in particular helpful for the calculation of expected limits, where several different hypotheses of μ are tested, as explained in the following section.

7.4.3 Calculation of expected significances and limits

Statistical tests can be performed in order to quantify the significance of a signal or, in the absence of a signal, to derive limits on specific model parameters. Therefore, different hypotheses are defined and their compatibility with data is tested. The B-only hypothesis describes the background processes expected by the SM, while the S+B hypothesis describes the background processes and yet unobserved signal contributions. For the discovery of a new process, the B-only hypothesis is required to be rejected. In order to set 95% confidence level (CL) limits, the compatibility of the S+B hypothesis and data is quantified. In this thesis, limits on the signal strength μ are derived by scaling the expected signal with μ in order to find the largest value of μ which is still compatible with the data within the 95% CL. An Asimov dataset is used to calculate the expected values for the limits and the significance during the optimization studies.

In order to quantify the level of agreement between the hypothesis of interest and either the observed data or the Asimov dataset, p-values are calculated, which describe the probability for a certain hypothesis to find data which are equally or even more incompatible with the prediction. The p-value can be expressed as a significance Z, by reinterpreting it as the area under the upper tail of a Gaussian function with center at 0 and unit width in a range $[Z, \infty]$. In this definition, Z describes the distance from the mean of the Gaussian distribution in terms of standard deviations. A new signal is discovered if the B-only hypothesis, defined by $\mu = 0$, is rejected with a significance of at least $5\,\sigma$ or equally a p-value of at most $2.87 \cdot 10^{-7}$. For limit setting, the S+B hypothesis is considered to be compatible with data for a p-value ≥ 0.05, which corresponds to $1.64\,\sigma$.

For the calculation of the expected limit and significance, the profile likelihood ratio based test statistic, as shown in Equation (7.10), is used. While the general definition of t_μ is unchanged, slightly modified test statistics, which consider some special cases, are used for the different hypothesis tests as explained in the following. For the estimation of the expected limit on μ, an alternative test statistic $\tilde{t}_\mu$ is defined, which considers two special cases:

$$\tilde{t}_\mu = \begin{cases} 0 & \text{for } \mu < \hat{\mu}, \\ -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(0,\hat{\boldsymbol{\theta}}(0))} & \text{for } \hat{\mu} < 0, \\ t_\mu & \text{otherwise.} \end{cases} \tag{7.12}$$

Hence, values of μ are not considered if they are smaller than the estimator $\hat{\mu}$. For these values, $\tilde{t}_\mu$ is set to 0. In addition, if the estimator $\hat{\mu}$ is smaller than 0, the test statistic is evaluated for $\hat{\mu} = 0$, which corresponds to the B-only hypothesis.

For the calculation of an expected significance, the test statistic is evaluated for $\mu = 0$, assuming the B-only hypothesis. A test statistic t_0 is defined, which considers upward and downward fluctuations of the data with respect to the background prediction:

$$t_0 = \begin{cases} +t_\mu & \text{for } \hat{\mu} > 0 \\ -t_\mu & \text{for } \hat{\mu} \leq 0 \end{cases} \tag{7.13}$$

Finally, the p-value is calculated to quantify the agreement of the tested hypothesis with data. The p-value for an observed value $t_{\mu,obs}$ is calculated as the integral of the $p.d.f.$ of the test statistic $f(t_\mu|\mu)$ for a hypothesized value μ:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu)\mathrm{d}t_\mu \tag{7.14}$$

Hence, the p-value corresponds to the area under the $f(t_\mu|\mu)$ distribution for $t_\mu \geq t_{\mu,obs}$. For estimating the expected limits, a special procedure is used for analyses with a small sensitivity, e.g. due to the small cross section of the signal process. The CL_S method [139] is adopted for experiments where the signal expectation s is small compared to the background expectation b, which is the case for the rare processes investigated in this thesis. For these experiments, the S+B hypothesis is by construction rejected in 5% of the cases, although there is no sensitivity. In this case, there is a large overlap of the $p.d.f$.s for the B-only and the S+B hypothesis. The p-value for the S+B hypothesis $p_{s+b} = P(t_\mu \geq t_{\mu,obs}|s+b)$ becomes very small, while the associated p-value of the B-only hypothesis $p_b = P(t_\mu \leq t_{\mu,obs}|b)$ gets large if there is a downward fluctuation of the data.

The CL_S method, which is defined based on the p-values p_{s+b} and p_b, introduces a penalty term to prevent the exclusion of the S+B hypothesis in these cases:

$$CL_S = \frac{p_{s+b}}{1-p_b} \tag{7.15}$$

The requirement for the exclusion of an S+B hypothesis $p_{s+b} < 0.05$ is replaced by the requirement $CL_S < 0.05$. For experiments with a low sensitivity, the denominator $1-p_b$ gets small for downward fluctuations of the data, which leads to larger values of CL_S and, hence, prevents a hypothesis from being falsely rejected. For other cases, where a separation between the S+B and B-only hypothesis is possible, p_b gets small and the value of CL_S is similar to the nominal value p_{s+b}. In Figure 7.4, the calculation of the p-values for the S+B and the B-only hypothesis is illustrated.

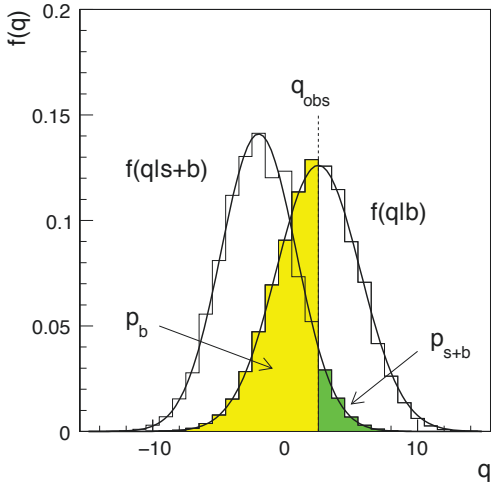


Figure 7.4: Illustration of the $p.d.f$.s of a test statistic q (which represents t_μ) for the S+B and the B-only hypothesis and the corresponding p-values, p_{s+b} and p_b, for an observed value q_{obs} [138].

Chapter 8

Selection of diphoton events

In order to reduce the amount of data which needs to be analyzed and to improve the ratio of $H \to \gamma\gamma$ signal events to the continuum background, only events containing two photons which are likely to originate from $H \to \gamma\gamma$ decays are selected. The ATLAS analyses in the $H \to \gamma\gamma$ decay channel share a common diphoton selection, which is optimized to efficiently reduce the large non-resonant background and to offer a sufficient signal efficiency at the same time. The photon selection proceeds in various steps, which are explained in the following. A distinction is made between the basic diphoton preselection, based on loosely identified photons, and the final selection of diphoton events which are used for measurements in the $H \to \gamma\gamma$ decay channel. In addition, the reconstruction of the diphoton primary vertex is shortly explained, which is used to determine the direction of the photon candidates more precisely.

As an example, the efficiencies of the different steps of the diphoton selection are shown in Figure 8.1 for the $tHjb$ MC sample. The overall efficiency for the diphoton selection amounts to about 33% for this process. Due to differences in the kinematics of the photons originating from the Higgs boson decay, the efficiency of the diphoton selection depends on the Higgs boson production mechanism and therefore slightly differs for the other processes.

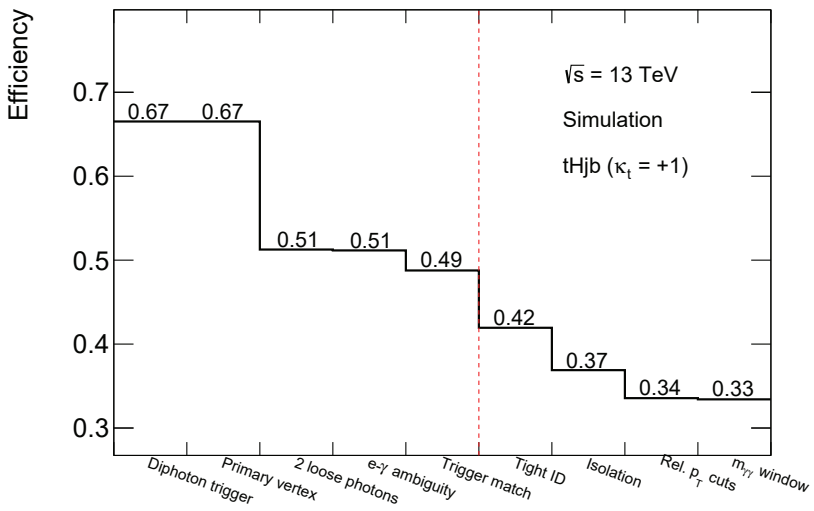


Figure 8.1: The efficiencies of the individual selection steps of the diphoton selection as an example for the $tHjb$ MC sample. The first five requirements correspond to the diphoton preselection, as indicated by the dashed red line.

8.1 Diphoton preselection

A basic preselection is performed in order to select events with two photon candidates within the detector acceptance. These basic selection requirements need to be fulfilled for all events which are considered in the following chapters. The events which are obtained after the preselection are characterized by loosely identified photons and are used for studies of the continuum background in regions with a very small number of expected events, which in particular applies to the tH and $t\bar{t}H$ event selections. The following requirements define the diphoton preselection:

- **Diphoton trigger:** As a first step, a trigger requirement is applied to MC simulation and data. A diphoton trigger is used which requires two clusters in the EM calorimeter to be present with E_{T} thresholds of 35 GeV and 25 GeV for the E_{T}-leading and the E_{T}-subleading cluster, respectively. These clusters are used to define trigger level photon objects which are required to fulfill the *loose* photon ID criteria for 2015-2016 data, while for 2017 data the *medium* photon ID criteria[1] were considered. This first selection step is particularly important for the selection of Higgs boson candidate events in data, while in signal MC simulation this requirement is fulfilled for a large fraction of events.
- **Primary vertex:** As the Higgs boson is produced during a hard scattering process, selected events are required to have a reconstructed primary vertex.
- **Two photons fulfilling *loose* ID:** The events are required to contain at least two reconstructed photons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.37$, excluding $1.37 < |\eta| < 1.52$, which fulfill the *loose* ID requirements. The two p_{T}-leading photons in each event are considered as Higgs boson candidate.
- ***e*-γ ambiguity:** As the same reconstruction algorithms are used for electrons and photons, an object can be reconstructed as both, an electron and a photon. Based on the procedure explained in Section 5.1.1, this ambiguity is resolved by using additional criteria mainly based on tracking and vertex information.
- **Trigger matching:** The two selected photon candidates are required to match the objects which fired the diphoton trigger.

8.2 Diphoton primary vertex reconstruction

The two photon candidates are used in order to reconstruct the diphoton primary vertex [2]. Commonly, the hardest vertex is chosen as a candidate for the hard-scattering vertex, which is defined as the vertex with the highest sum of squares of the p_{T} of associated tracks ($\sum p_{\mathrm{T}}^2$). As photons do not leave tracks inside the InDet, this choice often results in a misidentification of the $H \to \gamma\gamma$ decay vertex. Therefore, a neural network (NN) based approach is used in order to identify the vertex of the $H \to \gamma\gamma$ decay among all reconstructed vertices. The inputs which are provided to the NN are information about the direction of the two photon candidates, tracking and primary vertex information. The efficiency for selecting the correct vertex[2] using the NN-based diphoton primary vertex selection is about 79% for simulated ggH events. The efficiency increases if additional jets or charged leptons are produced and lies within 84% and

[1] The photon ID criteria were tightened for the trigger used during 2017 data-taking, to take into account the on average higher pileup. An additional requirement on E_{ratio} is added for the *medium* ID criteria.

[2] The selected vertex is required to be within 0.3 mm of the truth vertex to be classified as the correct vertex.

97% for the other Higgs boson production processes. The validation of the NN is performed using MC simulations and data containing $Z \to e^+e^-$ events. Finally, a correction is applied to the direction of the photon candidates based on the reconstructed diphoton primary vertex, leading to an improvement of the $m_{\gamma\gamma}$ resolution.

8.3 Final diphoton selection

Among the candidate events containing two photons which fulfill *loose* ID requirements, only events fulfilling additional quality criteria are used for the final analyses. These requirements include tighter requirements on the two photon candidates for a further reduction of the continuum background and in particular the fraction of jets which are misidentified as photons:

- **Tight ID:** Both photons are required to fulfill the *tight* ID requirements, which are defined based on the shape of the calorimeter showers, in order to reduce the contribution from hadronic fakes.
- **Photon isolation:** Both photons need to fulfill the isolation criteria which are described in Section 5.1.1, including track and calorimeter isolation requirements, in order to make sure that the photons are isolated from hadronic activity.
- **Relative p_{T} cuts** : The photons originating from $H \to \gamma\gamma$ decays are characterized by a relatively high p_{T} in contrast to the photons in the continuum background, due to the large mass of the Higgs boson. Relative p_{T} cuts are applied to the p_{T}-leading (p_{T}-subleading) photon, requiring the p_{T} to be larger than 0.35 (0.25) $\times\, m_{\gamma\gamma}$.
- **Diphoton invariant mass window:** Events with a diphoton invariant mass in the range $m_{\gamma\gamma} \in [105, 160]\,\mathrm{GeV}$ are selected. This window is asymmetric around the Higgs boson mass of approximately 125 GeV. As the continuum background is estimated from a fit to $m_{\gamma\gamma}$, a sufficient number of data events needs to be available below and above the Higgs boson mass peak. As the spectrum is continuously falling, the number of events with $m_{\gamma\gamma} > 130\,\mathrm{GeV}$ is significantly smaller than the number of events with $m_{\gamma\gamma} > 120\,\mathrm{GeV}$. Therefore, the $m_{\gamma\gamma}$ range above the Higgs boson mass is chosen to be larger to ensure a sufficient number of events in the tail of the $m_{\gamma\gamma}$ distribution for a reliable estimation of the continuum background. Furthermore, when considering lower values of $m_{\gamma\gamma}$, the continuum background distribution is not continuously falling anymore, but instead shows a turn-on due to the E_{T} requirements applied by the diphoton trigger.

In a nutshell, the Higgs boson candidate events are characterized by two high-p_{T} photons fulfilling *tight* ID and isolation requirements. These events are referred to as TI events (**t**ight and **i**solated), whereas for optimization and background modeling studies also events where at least one of the photons fails the *tight* ID or isolation criteria are used, which are referred to as $NTNI$ events (**n**on-**t**ight, **n**on-**i**solated). The region with $105\,\mathrm{GeV} \leq m_{\gamma\gamma} \leq 160\,\mathrm{GeV}$ around the Higgs boson mass is considered to allow for a fit of the continuum background spectrum using the selected background model. In the following, only events which pass the full diphoton selection or the diphoton preselection are considered. Unless otherwise stated, the full diphoton selection is applied.

Chapter 9

Search for Higgs boson production in association with a single top quark

As explained in Chapter 2, the coupling of the Higgs boson to the top quark Y_t is expected to be large in the SM, as the Higgs boson's coupling to fermions is proportional to their mass and the top quark is the heaviest elementary particle in the SM. Therefore, the coupling Y_t is an important property of the Higgs boson and its measurement is of paramount importance to test whether the properties of the discovered Higgs boson are consistent with those predicted by the SM. The $t\bar{t}H$ and tH processes both provide the possibility to directly measure Y_t. However, tH production additionally provides sensitivity to the relative sign between Y_t and the coupling of the Higgs boson to the W boson, g_{HWW}, due to a destructive interference in the SM. This property of tH production makes it an interesting process to study, despite its small cross section, providing sensitivity to BSM scenarios for Y_t. In this chapter, a combined search for $t\bar{t}H$ and tH production with a dataset of $36.1\,\text{fb}^{-1}$ is presented. At the ATLAS experiment, this is the first time of defining a dedicated event selection targeting the tH topology.

Several measurements were performed at the LHC during Run 1 and Run 2 by the ATLAS and CMS experiments in order to find an evidence for top-quark-associated Higgs boson production. Due to the small cross section of tH production, most tH measurements in Run 1 focused on the BSM scenario of $\kappa_t = -1$ which comes along with an increased tH cross section, while for $t\bar{t}H$ production the SM scenario was considered. An overview of the different measurements, performed by the ATLAS and CMS experiments, is given in the following section.

9.1 Measurements of top-quark-associated Higgs boson production at the LHC

At the ATLAS experiment, in Run 1 a search for Higgs boson production in association with top quarks, including $t\bar{t}H$ and tH production, was performed in the $H \to \gamma\gamma$ decay channel using data corresponding to $4.5\,\text{fb}^{-1}$ and $20.3\,\text{fb}^{-1}$ [140] collected at $\sqrt{s} = 7\,\text{TeV}$ and $8\,\text{TeV}$, respectively. Two categories were defined, which were optimized for events with leptonic and hadronic final states. While the leptonic event selection is characterized by relatively loose requirements, thereby selecting $t\bar{t}H$ as well as tH events, the hadronic event selection uses tight requirements on the number of jets in order to reduce the ggH background, which also reject tH events. No evidence was found for $t\bar{t}H$ and tH production and, hence, a 95% CL limit of $6.7\,\mu$ was set on the signal strength. In addition, 95% CL limits were set on the coupling strength modifier κ_t, excluding values $\kappa_t < -1.3$ and $\kappa_t > +8.0$.

In Run 2, two $t\bar{t}H$ categories were included in the $H \to \gamma\gamma$ measurement of production cross sections at the ATLAS experiment using data corresponding to $13.3\,\text{fb}^{-1}$ [1] collected during 2015 and the early data-taking period in 2016. The two categories were targeting leptonic and hadronic final states, respectively. A signal strength of $\mu_{t\bar{t}H} = -0.25\,^{+1.26}_{-0.99}$ was measured and no evidence for $t\bar{t}H$ production in the $H \to \gamma\gamma$ decay channel was found. Another measurement was performed at the ATLAS experiment using the $36.1\,\text{fb}^{-1}$ dataset collected in 2015 and 2016 and included $t\bar{t}H$ as well as tH categories [2]. The analysis strategy, the optimization of the tH and $t\bar{t}H$ categories and the results are presented and discussed in this chapter.

Finally, a measurement of $t\bar{t}H(H \to \gamma\gamma)$ production was performed with the 2015-2017 dataset [3, 4], corresponding to $79.8\,\text{fb}^{-1}$. A significance larger than $3\,\sigma$ was expected for this measurement, thus it provided the possibility to obtain evidence for $t\bar{t}H$ production in the $H \to \gamma\gamma$ decay channel for the first time. For this reason, tH production was only considered as background in this analysis. The measurement included seven $t\bar{t}H$ categories targeting leptonic and hadronic final states. The categories were defined based on BDT discriminants, which were used to reduce the ggH and continuum background contributions. The observed (expected) significance for $t\bar{t}H$ production in the $H \to \gamma\gamma$ decay channel was $4.1\,\sigma$ $(3.7\,\sigma)$. A combination with other Run 2 measurements of $t\bar{t}H$ production in the channels targeting Higgs boson decays into $b\bar{b}$, WW, ZZ and $\tau\tau$ was performed, resulting in an observed significance of $5.8\,\sigma$ and thereby providing the first observation of $t\bar{t}H$ production at the ATLAS experiment. The $t\bar{t}H(H \to \gamma\gamma)$ measurement was updated with the full Run 2 dataset of $139\,\text{fb}^{-1}$ [141], resulting in an observed (expected) significance of $4.9\,\sigma$ $(4.2\,\sigma)$.

At the CMS experiment, $t\bar{t}H$ production was also observed for the first time by combining analyses considering the $b\bar{b}$, WW, ZZ, $\tau\tau$ and $\gamma\gamma$ final states [19], which were performed using datasets collected at center-of-mass energies of 7, 8 and 13 TeV, and resulted in an observed significance of $5.2\,\sigma$. Similar to the ATLAS measurement, multivariate analysis (MVA) techniques were used in the $H \to \gamma\gamma$ decay channel at $\sqrt{s} = 13\,\text{TeV}$ [142] for hadronic final states, to separate $t\bar{t}H$ events from multijet backgrounds. This measurement of $t\bar{t}H$ in the $H \to \gamma\gamma$ channel was performed with a dataset corresponding to $35.9\,\text{fb}^{-1}$ and a signal strength of $2.2\,^{+0.9}_{-0.8}$ was measured.

In addition, searches for tH production with $\kappa_t = -1$ were performed at the CMS experiment in various Higgs boson decay channels using the dataset collected during Run 1 [143]. The considered Higgs boson decay channels include $\gamma\gamma$, $b\bar{b}$, multilepton and $\tau\tau$ final states. An MVA discriminant was used in the $H \to \gamma\gamma$ channel to distinguish between tH and $t\bar{t}H$ final states. No evidence was found for tH production with $\kappa_t = -1$, therefore limits were set on the signal strength. The observed (expected) upper 95% CL limit on $\mu_{\kappa_t=-1}$, after combining the different decay channels, amounts to 2.8 (2.0). The $H \to \gamma\gamma$ decay channel provided the highest sensitivity with an observed and expected upper limit on $\mu_{\kappa_t=-1}$ of 4.1.

A separate search for SM-like tH production was performed at the CMS experiment using the 2016 dataset corresponding to $35.9\,\text{fb}^{-1}$ [144]. The Higgs boson decays into WW, $\tau\tau$, ZZ and $b\bar{b}$ were considered in separate measurements, which were finally combined to constrain κ_t. In addition, the $\gamma\gamma$ channel was considered by using the results from the $t\bar{t}H$ measurement at $35.9\,\text{fb}^{-1}$, as the $t\bar{t}H$ categories also include significant fractions of $tHjb$ and WtH events. For an SM-like coupling of the Higgs boson to vector bosons, positive values of κ_t were favored by the data, while excluding values outside the regions [-0.9, -0.5] and [1.0, 2.1] at 95% CL.

9.2 Strategy for the measurement of tH production

The measurement of tH production is challenging due to its small cross section. First studies on the prospects for the measurement of tH production in the $H \to \gamma\gamma$ channel were presented in Ref. [145]. It was shown, that $t\bar{t}H$ production is the main background for tH production due to the similar final states. Especially requirements on the number of jets help to reduce the $t\bar{t}H$ contamination, but due to its much larger cross section it remains the dominant process for the event selection presented in Ref. [145]. Therefore, a combined measurement of tH and $t\bar{t}H$ production is considered to be the most promising approach to draw conclusions on Y_t with the available 2015+2016 dataset, due to the large overlap of the tH and $t\bar{t}H$ event selections. In order to improve the sensitivity to tH and $t\bar{t}H$ production, several categories with varying signal-to-background ratio S/B are defined. For the optimization of the tH and $t\bar{t}H$ categories, the following aspects are considered:

- **Categorization:** Categories with varying S/B are defined based on different event topologies. In order to ensure orthogonality of the different categories, these are filled in a predefined order so that every event can only enter exactly one category.
- **Ordering of categories:** Categories aiming for the selection of rare processes are defined first, to improve the S/B in these categories. Therefore, the categories targeting tH final states are filled before the $t\bar{t}H$ categories. Similarly, categories targeting final states with leptons are filled before categories targeting purely hadronic final states.
- **Considered signal processes:** Only $tHjb$ and WtH production are considered as signal processes, while the s-channel tH production is neglected due to its small cross section. In the optimization studies, $t\bar{t}H$ production is treated as background. As the WtH final state is very similar to $t\bar{t}H$ production and can only be differentiated by the presence of one additional jet in the $t\bar{t}H$ topology, the optimization studies focus on the $tHjb$ final state. This is further motivated by the fact that the $tHjb$ cross section is a factor ≈ 3 higher than the WtH cross section.
- **Maximize expected limit for tH categories:** Previous optimization studies showed that it is unlikely to measure SM-like tH production in the $H \to \gamma\gamma$ decay channel even with the full Run 2 dataset. Therefore, the tH categories are optimized by maximizing the expected limit on μ.
- **Maximize expected significance for $t\bar{t}H$ categories:** Unlike tH production, achieving an evidence for $t\bar{t}H$ production in the $H \to \gamma\gamma$ channel is within reach with the Run 2 dataset. Therefore, the $t\bar{t}H$ categories are optimized in order to maximize the expected significance.

The search for tH and $t\bar{t}H$ production is part of an analysis which targets various Higgs boson production modes in order to measure their cross sections, their signal strengths and the couplings of the Higgs boson. Therefore, additional categories targeting VH, VBF and ggH production are part of the analysis. In total, 31 categories are defined in a well-defined order, which is determined by the cross sections of the processes. The tH and $t\bar{t}H$ categories are defined first due to their small cross sections. Finally, a simultaneous fit to the $m_{\gamma\gamma}$ distributions of all categories is performed, which helps to constrain the Higgs boson background processes in the $t\bar{t}H$ and tH categories.

9.3 Optimization of tH categories

First prospective studies for a measurement of tH production were performed in Ref. [145] with $\sqrt{s} = 8\,\text{TeV}$ MC samples, which were PDF-reweighted to $\sqrt{s} = 13\,\text{TeV}$ by scaling the considered PDFs to higher proton energies. The results of these prospective studies are used as a basis for the studies presented in this chapter. They are validated using $\sqrt{s} = 13\,\text{TeV}$ MC samples and a reoptimization of the event selection is performed, which is presented in the following. In these studies it was shown, that an MVA approach for the separation of $t\bar{t}H$ and $tHjb$ production works very well, but due to the small cross section of tH production it does not significantly improve the expected limits. Hence, a cut-based approach is presented in the following. The combination of the proposed cut-based leptonic and hadronic categories in Ref. [145] results in an expected limit of $9.8\,\mu$ for a dataset corresponding to $100\,\text{fb}^{-1}$.

The event selection is targeting the $tHjb$ final state, where the top quark in almost 100% of the cases decays into a b quark and a W boson, which further decays either leptonically into a charged lepton and the corresponding neutrino or hadronically into two quarks. Therefore, two event selections are developed which target leptonic and hadronic final states separately. In Figure 9.1, the Feynman diagram for $tHjb$ production in the 4FS, including all subsequent decays, is shown. The $tHjb$ final state is characterized by the presence of

- a light flavor jet from the production,
- a b-jet from the top quark decay,
- a possible second b-jet produced during the gluon splitting,
- a charged lepton and $E_{\text{T}}^{\text{miss}}$ **or** two additional jets.

For the calculation of expected limits and significances during the optimization studies, it is assumed that the shape of events originating from Higgs boson production processes are described by a Gaussian $m_{\gamma\gamma}$ distribution with a mean of 125 GeV and a width of 1.5 GeV, while the shape of the continuum background is described by a simple exponential function. Only statistical uncertainties are considered during the optimization studies, as these are expected to be dominant. Furthermore, the optimization is performed assuming the SM coupling with $\kappa_t = +1$ for tH production. The kinematics of the tH final state particles are very similar for $\kappa_t = +1$ and the BSM scenario with $\kappa_t = -1$, as shown in Appendix B.1. Hence, the optimized event selection also provides sensitivity to tH production with $\kappa_t = -1$.

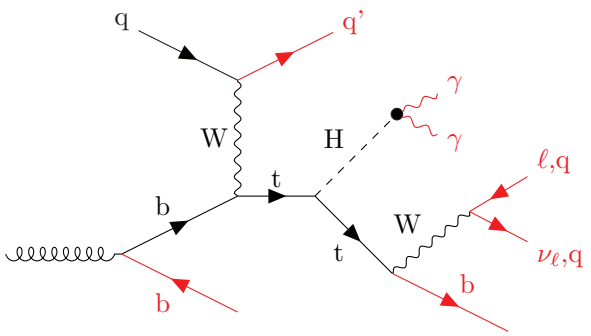


Figure 9.1: Example Feynman diagram for $tHjb$ production in the 4FS and the subsequent decays. The expected final state particles are highlighted in red.

In order to perform a combined measurement of tH+$t\bar{t}H$ production, the optimized tH categories are finally combined with categories optimized for $t\bar{t}H$ production. During this harmonization of the tH and $t\bar{t}H$ categories, the definition of the optimized tH categories is slightly changed. In the following, first the data-driven strategy to estimate the continuum background contribution is explained. In contrast to the Higgs boson backgrounds, the continuum background cannot be determined from MC samples. Then, the optimization of the leptonic and hadronic tH categories is presented, as well as the final categorization, which is used for the combined measurement of tH and $t\bar{t}H$ production.

9.3.1 Continuum background estimation

The composition of the continuum background for tH and $t\bar{t}H$ final states is barely known, due to the small number of events in the data $m_{\gamma\gamma}$ sidebands and the large number of processes that might contribute to this background, such as $t\bar{t}$ and single top quark production, W/Z+jets production or QCD multijet production with additional prompt or *fake* photons in the final state. Hence, no MC samples exist which sufficiently describe the continuum background for the investigated final states. Likewise, the data $m_{\gamma\gamma}$ sidebands with $m_{\gamma\gamma} \notin [120, 130]$ GeV are characterized by large statistical fluctuations, so that they are not suitable to draw conclusions on the continuum background during the optimization studies. Therefore, in order to avoid optimizing on statistical fluctuations in the $m_{\gamma\gamma}$ sidebands, the continuum background contribution is estimated using the $ABCD$ method based on control regions in data that are characterized by a larger number of events. These control regions consist of events failing *tight* ID and/or isolation requirements, so-called $NTNI$ events. The ratio of the number of events in the $m_{\gamma\gamma}$ sidebands (TI events) and the number of $NTNI$ events, N^{TI}/N^{NTNI}, is considered. This ratio is assumed to be similar for two different event selections $sel1$ and $sel2$:

$$\frac{N^{TI}_{sel1}}{N^{NTNI}_{sel1}} \approx \frac{N^{TI}_{sel2}}{N^{NTNI}_{sel2}} \tag{9.1}$$

Hence, if a second event selection $sel2$ is defined with $N^{TI}_{sel2} > N^{TI}_{sel1}$, the continuum background expectation N^{TI}_{sel1} can be evaluated for the tested event selection $sel1$ by rearranging Equation (9.1). The continuum background is estimated using 36.1 fb^{-1} of data collected at $\sqrt{s} = 13$ TeV. For the optimization studies, a hadronic $t\bar{t}H$ event selection[1] is used as reference event selection $sel2$. However, this reference event selection was found to underestimate the number of continuum background events in the leptonic categories. For the final evaluation of the expected sensitivity of the full tH+$t\bar{t}H$ categorization, a leptonic reference event selection[2] is thus used to estimate the continuum background in these categories. As the proportion of the estimated number of continuum background events in the different leptonic categories remains unchanged, the choice of reference event selection does not affect the validity of the optimization studies.

[1] The hadronic reference event selection requires no charged leptons, at least 5 central jets with $p_{\mathrm{T}} > 30$ GeV and at least 1 b-tagged jet (77% WP).

[2] The leptonic reference event selection requires the presence of at least one charged lepton, at least two central jets with $p_{\mathrm{T}} > 25$ GeV and at least one loosely b-tagged jet (85% WP).

9.3.2 Leptonic categories

Event selections targeting leptonic signatures are usually characterized by a high signal purity, as the requirement of a charged lepton provides an efficient suppression of the large QCD backgrounds. In the prospective studies for tH production, it was shown that the main backgrounds for the leptonic final state result from $t\bar{t}H$ production and the continuum background. The leptonic tH event selection presented in Ref. [145] consists of the following requirements: exactly one electron or muon in the final state, $E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$, at least one b-tagged jet with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ (85% WP) and a maximum number of 3 central jets, $N_{\mathrm{jets}}^{\mathrm{central}}$, with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. This event selection is referred to as the baseline event selection in the following. The requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$, the number of b-tagged jets ($N_{b\text{-tag}}$) and the number of charged leptons (N_{lep}) efficiently reduce the Higgs boson backgrounds without top quarks in the final state, while the requirement on the maximum number of central jets reduces the $t\bar{t}H$ background. The b-tagging WP with the highest efficiency was found to provide the highest sensitivity. In addition, a second b-tagged jet is allowed in the event selection to account for b-jets which

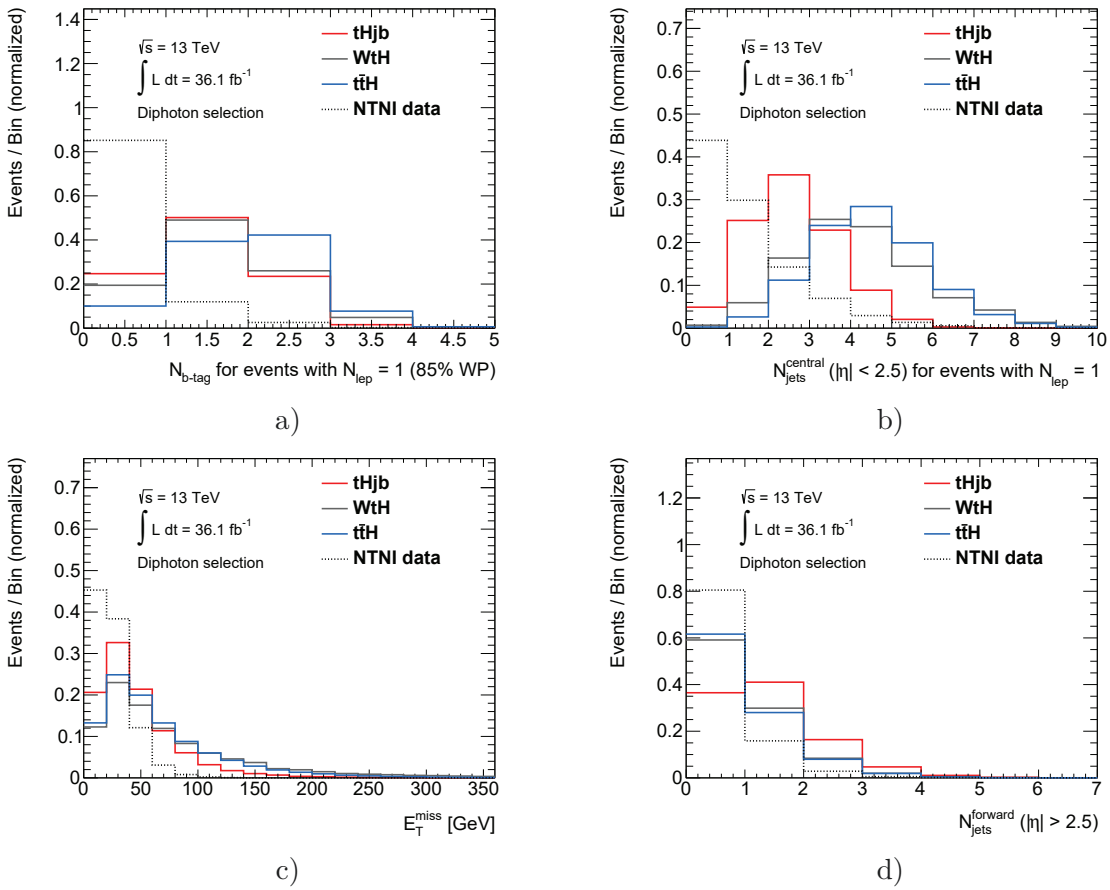


Figure 9.2: Distributions of the quantities which are used to define the leptonic tH event selection for the signal processes $tHjb$ and WtH, the main Higgs boson background $t\bar{t}H$ and $NTNI$ data representing the continuum background: a) the number of b-tagged jets (85% WP) for events with exactly one charged lepton in the final state, b) the number of central jets for events with exactly one charged lepton in the final state, c) the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$, d) the number of forward jets, which are characterized by $|\eta| > 2.5$.

are produced by the gluon splitting in $tHjb$ events. The distributions of the quantities which are used to define the leptonic event selection are shown in Figure 9.2a)-c) for the two signal processes, $tHjb$ and WtH, and the two main backgrounds, $t\bar{t}H$ and the continuum background, which is represented by $NTNI$ data events.

Table 9.1: Validation of the leptonic event selection presented in Ref. [145], which was optimized using PDF-reweighted $\sqrt{s} = 8\,\text{TeV}$ MC samples, for an integrated luminosity of $100\,\text{fb}^{-1}$. The expected number of events for the signal processes S, the Higgs boson background B_{Higgs}, the data-driven continuum background estimation B_{ContBG} and the expected significance and the limit on μ are compared to numbers obtained using $\sqrt{s} = 13\,\text{TeV}$ MC samples.

MC samples	S	B_{Higgs}	S/B_{Higgs}	B_{ContBG}	Exp. sig. $[\sigma]$	Exp. limit $[\mu]$
8 TeV (PDF rew.)	0.70	3.9	0.18	56	0.21	11.2
13 TeV	0.79	4.7	0.17	87	0.20	11.5

A validation of the results obtained from the optimization studies with the PDF-reweighted 8 TeV MC samples is performed by comparing the expected event yields and the expected sensitivity of the leptonic baseline event selection to numbers obtained with the 13 TeV MC samples. This comparison is shown in Table 9.1, where S refers to the number of expected events for the signal processes $tHjb$ and WtH, B_{Higgs} to the Higgs boson backgrounds[3] and B_{ContBG} to the continuum background expectation in the full 105 to 160 GeV mass window. The number of expected events obtained from $\sqrt{s} = 13\,\text{TeV}$ MC samples is larger for signal as well as for background, while the ratio S/B_{Higgs} is almost unchanged. In particular, the approximation for B_{ContBG} from the 8 TeV data samples seems to underestimate the continuum background at $\sqrt{s} = 13\,\text{TeV}$. In MC simulation, the increased center-of-mass energy between Run 1 and Run 2 is considered by applying the PDF-reweighting. But additional changes were made e.g. to the detector setup, the reconstruction algorithms and the considered MC generators, so that differences in the events yields are expected for both, the Higgs boson processes, which are estimated from MC samples, and the continuum background estimated from data. Nevertheless, the expected limit and the expected significance are very similar, as both the signal and the background expectation increased for the $\sqrt{s} = 13$ TeV data and MC samples.

In general, the numbers show a reasonable agreement, so that the prospective tH studies from Ref. [145] are considered to be valid. In Table 9.2, the efficiencies of the baseline event selection for the different Higgs boson production processes and the number of expected events for an integrated luminosity of $100\,\text{fb}^{-1}$ are shown. The efficiencies for Higgs boson processes without top quarks in the final state are smaller than or equal to 0.64%, while the efficiencies for $t\bar{t}H$ and tH production range from approximately 8.4% to 11%.

During the reoptimization of the leptonic event selection, the presence of a forward jet in $tHjb$ production is exploited. The light flavor jet of the production (q' in Figure 9.1) tends to be produced in forward direction, while $t\bar{t}H$ production is characterized by mostly central jets. Therefore, the requirement of a forward jet helps to reduce the $t\bar{t}H$ contamination in the tH categories. Forward jets are defined in the region $|\eta| > 2.5$, which is not covered by the InDet. The distribution of the number of forward jets for the signal processes and the two main

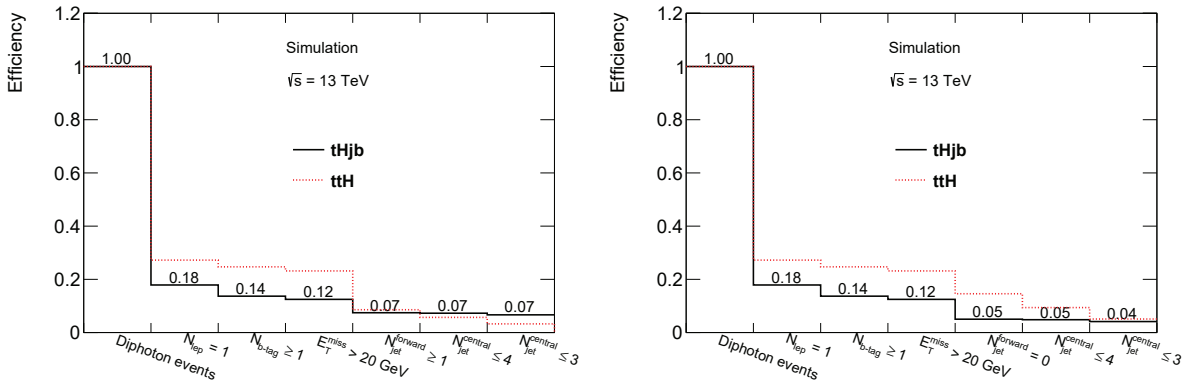
[3] Higgs boson production via $b\bar{b}H$ is neglected during the optimization studies, as no MC samples were available at the time when the studies were performed. Due to the small contribution of $b\bar{b}H$ events in the final event selection, the absence of the $b\bar{b}H$ background is not expected to affect the validity of the optimization studies.

Table 9.2: Efficiencies of the leptonic baseline event selection for the different Higgs boson production processes and the number of expected events for a dataset corresponding to $100\,\text{fb}^{-1}$, evaluated using $\sqrt{s} = 13\,\text{TeV}$ MC samples.

Process	Efficiency	N_{exp} for $100\,\text{fb}^{-1}$
ggH	0.0035%	0.15
VBF	0.0039%	0.013
WH	0.64%	0.66
ZH	0.26%	0.18
$t\bar{t}H$	8.4%	3.7
$tHjb$	11%	0.65
WtH	9.6%	0.14

backgrounds is shown in Figure 9.2d). More than 60% of the $tHjb$ events are characterized by the presence of at least one forward jet in the final state, while more than 60% of the $t\bar{t}H$ events contain only central jets. In the following, the expected event yields are normalized to $35\,\text{fb}^{-1}$, which approximately corresponds to the dataset collected during 2015 and 2016.

Two categories are defined, requiring either $N_{\text{jets}}^{\text{forward}} \geq 1$ or $N_{\text{jets}}^{\text{forward}} = 0$. The requirements on $N_{b\text{-tag}}$ and $E_{\text{T}}^{\text{miss}}$ remain unchanged, whereas the requirement on the maximum number of jets is reinvestigated. Only a small fraction of $tHjb$ events is characterized by a number of central jets larger than 4, therefore, the requirements on the maximum number of jets $N_{\text{jets}}^{\text{central}} \leq 3$ and $N_{\text{jets}}^{\text{central}} \leq 4$ are considered. The efficiencies of the different requirements for the two categories are shown in Figure 9.3a) and b), respectively. As illustrated in Figure 9.3a), the requirement of $N_{\text{jets}}^{\text{forward}} \geq 1$ significantly reduces the $t\bar{t}H$ efficiency, thereby allowing for a more loose requirement on the maximum number of central jets of 4. For the $N_{\text{jets}}^{\text{forward}} = 0$ category, a tighter requirement, allowing for a maximum number of 3 central jets, is chosen to reduce the larger $t\bar{t}H$ fraction.



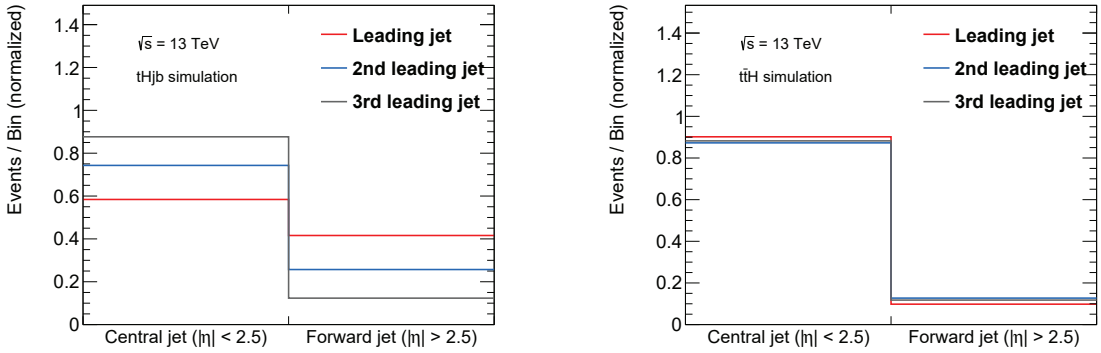
a) b)

Figure 9.3: Efficiencies of the different requirements, which are used to define the leptonic event selection, for $tHjb$ and $t\bar{t}H$ production and a) the category requiring at least one forward jet and b) the category with a veto on forward jets.

Table 9.3: Summary of the definitions of the different leptonic tH categories, which are investigated during the optimization studies.

Categories	Additional requirements for leptonic categories ($N_{\text{lep}} = 1$)
Baseline leptonic	$N_{b\text{-tag}} \geq 1$, $E_{\text{T}}^{\text{miss}} > 20\,\text{GeV}$, $N_{\text{jets}}^{\text{central}} \leq 3$
Cat($N_{\text{jet}}^{\text{forward}} = 0$)	$N_{b\text{-tag}} \geq 1$, $E_{\text{T}}^{\text{miss}} > 20\,\text{GeV}$, $N_{\text{jets}}^{\text{forward}} = 0$, $N_{\text{jets}}^{\text{central}} \leq 3$
Cat($N_{\text{jet}}^{\text{forward}} \geq 1$)	$N_{b\text{-tag}} \geq 1$, $E_{\text{T}}^{\text{miss}} > 20\,\text{GeV}$, $N_{\text{jets}}^{\text{forward}} \geq 1$, $N_{\text{jets}}^{\text{central}} \leq 4$
Cat($\lvert\eta(j_{1,untagged})\rvert > 2.5$)	$N_{b\text{-tag}} \geq 1$, $E_{\text{T}}^{\text{miss}} > 20\,\text{GeV}$, $\lvert\eta(j_{1,untagged})\rvert > 2.5$, $N_{\text{jets}}^{\text{central}} \leq 4$

Furthermore, it is investigated if the forward jet in $tHjb$ production can be identified among the selected jets, in order to define additional requirements for a further reduction of the $t\bar{t}H$ contribution in the $N_{\text{jets}}^{\text{forward}} \geq 1$ category. As the jet is expected to be a light flavor jet, only jets which are not b-tagged are considered. In Figure 9.4, the classification of the three p_{T}-leading untagged jets into central and forward jets is shown for $tHjb$ and $t\bar{t}H$ production. For $tHjb$ production, the p_{T}-leading untagged jet can be identified as a forward jet in about 40% of the cases. In contrast, for $t\bar{t}H$ events there is no tendency which jet is more likely produced in forward direction. Therefore, an additional category is defined, where the $N_{\text{jets}}^{\text{forward}} \geq 1$ requirement is replaced by the requirement $|\eta(j_{1,untagged})| > 2.5$, with $j_{1,untagged}$ referring to the p_{T}-leading untagged jet.



a) b)

Figure 9.4: Probability for the three p_{T}-leading untagged jets to be classified as a forward jet ($|\eta| > 2.5$) or a central jet ($|\eta| < 2.5$) for a) $tHjb$ production and b) $t\bar{t}H$ production.

The requirements of the leptonic baseline event selection and the three leptonic tH categories, which are defined based on the presence of a forward jet, are summarized in Table 9.3. For each of these categories, the expected number of events for signal and background is estimated and the expected limit on the signal strength is calculated for $35\,\text{fb}^{-1}$, as shown in Table 9.4. Both, the category based on $N_{\text{jets}}^{\text{forward}} \geq 1$ and $|\eta(j_{1,untagged})| > 2.5$, show an increased ratio S/B_{Higgs} compared to the baseline event selection. The expected limit of the $N_{\text{jets}}^{\text{forward}} \geq 1$ category is similar to the baseline event selection, while the expected limit of the $|\eta(j_{1,untagged})| > 2.5$ category is slightly worse, although this category provides the highest S/B_{Higgs}. The $N_{\text{jets}}^{\text{forward}} \geq 1$ category and the $|\eta(j_{1,untagged})| > 2.5$ category are both orthogonal to the $N_{\text{jets}}^{\text{forward}} = 0$ category and, hence, can be combined to further increase the expected sensitivity. Finally, combining the $N_{\text{jets}}^{\text{forward}} \geq 1$ and the $N_{\text{jets}}^{\text{forward}} = 0$ category provides the best expected limit with $18.9\,\mu$,

compared to an expected limit of 20.6 μ for the other combination. This corresponds to a 13% improvement with respect to the leptonic baseline event selection. Hence, the two leptonic categories requiring at least one or exactly zero forward jets are considered for the final event selection. With the improved continuum background estimation based on a leptonic reference event selection, as described in the previous section, the expected limit of the two combined leptonic categories yields 23.2 μ. The expected limits for the individual leptonic categories obtained with the improved continuum background estimation are shown in Appendix B.2.

Table 9.4: Comparison of the expected event yields for the signal S, the Higgs boson background B_{Higgs}, the continuum background B_{ContBG} and the expected limit for the different leptonic tH categories for an integrated luminosity of 35 fb^{-1}. The efficiencies and expected event yields for the different Higgs boson production processes are shown in Appendix B.2

Event selection	S	B_{Higgs}	S/B_{Higgs}	B_{ContBG}	Exp. limit [μ]
Baseline leptonic	0.28	1.6	0.17	31	21.4
Cat($N_{\text{jets}}^{\text{forward}} = 0$)	0.12	1.1	0.11	22	43.6
Cat($N_{\text{jets}}^{\text{forward}} \geq 1$)	0.19	0.97	0.19	9.5	22.9
Cat($\lvert\eta(j_{1,untagged})\rvert > 2.5$)	0.13	0.38	0.34	5.5	27.0

9.3.3 Hadronic categories

The hadronic tH final state is characterized by the presence of several jets in the final state, among which one or two may be b-tagged. The main backgrounds in this channel are, besides the continuum background, $t\bar{t}H$ and ggH production. Although the $t\bar{t}H$ final state is characterized by a higher number of jets, it becomes similar to the tH final state if some of the jets are not reconstructed, such as jets which are not within the detector acceptance. In contrast, the ggH final state is very different to tH production, as jets in the final state are only expected from additional radiation. These jets are mostly light flavor jets, which are characterized by a falling p_{T} spectrum. Nevertheless, ggH is an important background for hadronic tH final states, as its cross section is significantly higher than the $tHjb$ cross section. For processes with additional heavy flavor jets in the final state, such as ggH and VBF and WH, it is unknown how well the additional heavy flavor jets are modeled by MC simulation. Therefore, a conservative 100% uncertainty is associated to this kind of background. An attempt is hence made to minimize the ggH contribution in the hadronic categories during the optimization. As a benchmark, the ggH contribution is required to be smaller than approximately 10%.

The optimization of the hadronic categories proceeds similarly to the approach presented in Ref. [145]. For the hadronic categories, a veto on events with charged leptons is applied. Different requirements on the number of central jets, $N_{\text{jets}}^{\text{central}}$, and b-tagged jets, $N_{b\text{-tag}}$, are considered. As an example, Figure 9.5 shows the $N_{\text{jets}}^{\text{central}}$ and the $N_{b\text{-tag}}$ distributions for the 77% b-tagging WP for $tHjb$, $t\bar{t}H$, ggH and $NTNI$ data. Most of the $tHjb$ events contain between 2 and 5 central jets with $p_{\text{T}} > 25$ GeV, whereas $t\bar{t}H$ events are characterized by on average higher jet multiplicities. In contrast, ggH production and the $NTNI$ data are characterized by lower jet multiplicities, with more than 50% of the events containing no central jets with $p_{\text{T}} > 25$ GeV in the final state and approximately 30% with exactly one such jet. The $N_{b\text{-tag}}$ distribution shows that the requirement of at least one b-tagged jet is highly efficient for a reduction of ggH and the continuum background. The $tHjb$ events mostly have exactly one b-tagged jet in the final state, while for $t\bar{t}H$ events the probability of finding events with one or

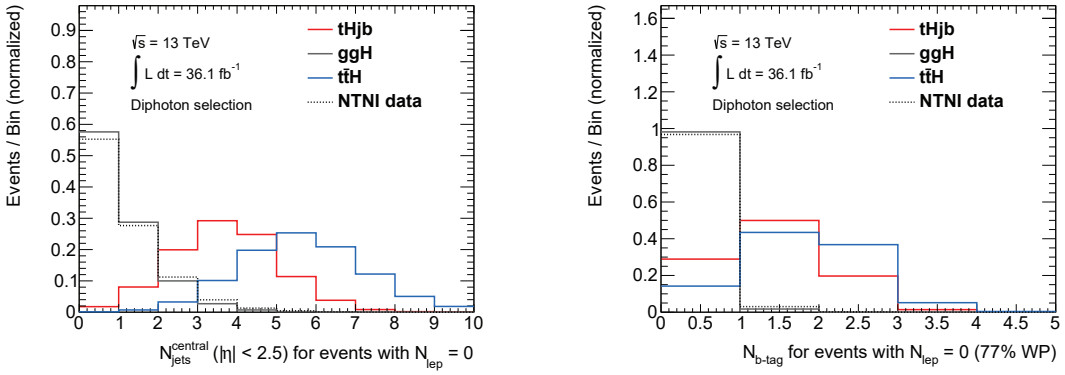
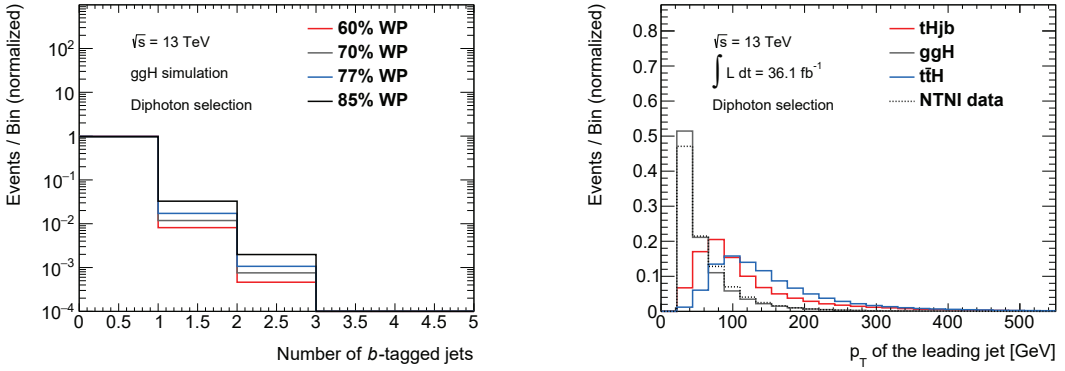


Figure 9.5: Distributions of the number of central jets and b-tagged jets for events without leptons in the final state for the $tHjb$ signal, the ggH and $t\bar{t}H$ background and $NTNI$ data, representing the continuum background.



a) b)

Figure 9.6: a) Distribution of the number of b-tagged jets for the different b-tagging WPs for ggH production, b) p_{T} distribution of the leading jet for $tHjb$ signal events, ggH and $t\bar{t}H$ background events and $NTNI$ data, representing the continuum background.

two b-tagged jets is similar. In the following, for the labeling of the hadronic categories the number of central jets is denoted by j and the number of b-tagged jets is denoted by b, so that e.g. $4j1b$ corresponds to a category with 4 central jets and 1 b-tagged jet. For $2b$ final states, the requirement of ≥ 2 b-tagged jets is used. Categories with 3 to 5 central jets in the final state and either 1 or ≥ 2 b-tagged jets are tested. For lower and higher jet multiplicities, the ggH and $t\bar{t}H$ fractions increase, respectively. In order to reduce the ggH contribution in each category, the following requirements are varied:

- ***b*-tagging WP:** The probability to find b-tagged jets in ggH production is smaller than in tH production, therefore choosing a tighter b-tagging WP helps to reduce the ggH fraction. The distribution of the number of b-tagged jets in ggH production for the different b-tagging WPs is shown in Figure 9.6a).
- **Jet p_{T} cut:** Jets in ggH production on average have a smaller p_{T} than in tH production, as shown in Figure 9.6b), therefore, increasing the jet p_{T} requirement helps to reduce the ggH fraction. Jet p_{T} requirements of 25 GeV, 30 GeV and 35 GeV are considered.

After testing the various combinations, only the $4j$ and $5j$ categories remain candidates for

the hadronic event selection, as for lower jet multiplicities the ggH fraction is always larger than 10%. As an example, the $3j1b$ category in association with the tightest jet p_{T} requirement and the tightest b-tagging WP shows a ggH fraction of about 31%. In general, the b-tagging WP and the jet p_{T} cut which provide the highest amount of signal events and a ggH fraction smaller than or close to 10% are chosen for each category. However, the jet p_{T} requirements and b-tagging WPs for categories with the same jet multiplicity, j, and the same number of b-tagged jets, b, are harmonized, respectively, in order to avoid large overlaps between the different categories.

Table 9.5: Summary of the chosen jet p_{T} and b-tagging requirements for the hadronic tH categories. The categories are filled in the order which is shown here.

Category	Jet p_{T} requirement	b-tagging WP
$4j1b$	35 GeV	60%
$4j2b$	35 GeV	77%
$5j1b$	25 GeV	60%
$5j2b$	25 GeV	77%

A summary of the selected jet p_{T} requirements and the b-tagging WPs for the $4j1b$, $4j2b$, $5j1b$ and $5j2b$ categories is shown in Table 9.5. It is noticeable that for the $1b$ categories it is necessary to use a tighter b-tagging WP than for the $2b$ categories. In addition, a higher jet p_{T} requirement is chosen for $4j$ categories than for the $5j$ categories. In Table 9.6, the fractions of the different Higgs boson production processes are shown for the four reoptimized hadronic tH categories. In addition, the number of signal, Higgs boson background and continuum background events is shown and the expected limit for each category. The $4j$ categories are filled prior to the $5j$ categories, so that each event can enter only one category. The expected limits of the individual hadronic categories are worse compared to the leptonic categories, as they are characterized by a lower signal-to-background ratio. The combined limit for the hadronic categorization amounts to $32.3\,\mu$. However, during the harmonization of the tH and $t\bar{t}H$ categories the definition of the hadronic tH categories is modified and a larger ggH fraction is allowed eventually, as described in the following section.

Table 9.6: Composition of the different hadronic categories and the expected number of signal, Higgs boson background and continuum background events for $35\,\mathrm{fb}^{-1}$. In addition, the expected limit is shown for each category.

	$4j1b$	$4j2b$	$5j1b$	$5j2b$
ggH	11%	7.9%	11%	10%
VBF	1.1%	0.84%	1.5%	0.76%
WH	1.2%	1.1%	1.6%	1.3%
ZH	5.0%	5.6%	3.6%	4.0%
$t\bar{t}H$	73%	76%	74%	77%
$tHjb$	6.7%	6.8%	5.7%	5.0%
WtH	2.1%	1.5%	2.4%	1.5%
S	0.17	0.067	0.087	0.035
B_{Higgs}	1.8	0.75	0.98	0.51
B_{ContBG}	98	24	63	17
Exp. limit [μ]	50.5	77.6	81.7	132.6

9.4 Harmonization of tH and $t\bar{t}H$ categories

As explained in Section 9.2, a combined search for tH and $t\bar{t}H$ production is performed, due to the similarity of the $t\bar{t}H$ and tH final states and the comparably small cross section of tH production, which makes a measurement of $tH(H \to \gamma\gamma)$ alone with the available dataset challenging. In the optimization studies presented in the previous chapter, only the tH final state is considered. Thus, the optimized tH categories need to be merged with the categories optimized for $t\bar{t}H$ production, which comes along with small changes in the definition of the tH categories. In order to use different b-tagging WPs at the same time, a special calibration of the b-tagging algorithm is needed. This calibration was not available on the timescale of this measurement and, hence, a single b-tagging WP needs to be used for all categories. Therefore, the 70% b-tagging WP is used throughout the analysis, as it provides the best compromise. In addition, the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement in the leptonic categories is dropped, as it mainly reduces backgrounds which are already well suppressed by the requirement of a charged lepton, such as ggH+jets and VBF+jets.

For the leptonic final state, two tH categories and one $t\bar{t}H$ category are defined, as shown in Table 9.7. In addition to the tH categories with the $N_{\mathrm{jets}}^{\mathrm{forward}} = 0$ and $N_{\mathrm{jets}}^{\mathrm{forward}} \geq 1$ requirements, which are referred to as tH lep 0fwd and tH lep 1 fwd, a $t\bar{t}H$ category is defined, which is filled after these two categories. The requirements of this category, which is referred to as $t\bar{t}H$ lep, include a minimum number of two central jets, at least one b-tagged jet and at least one charged lepton. A veto is applied to events with two same-flavor leptons with an invariant mass between 86 and 96 GeV to reduce the contribution from events with Z bosons decaying into a pair of charged leptons ($Z \to \ell\ell$).

Table 9.7: An overview of the tH and $t\bar{t}H$ categories targeting leptonic and hadronic final states. The hadronic $t\bar{t}H$ categories include requirements on the BDT discriminant ($\mathrm{BDT}_{t\bar{t}H}$), the number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 4.5$ (N_{jets30}) and the number of b-tagged jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ ($N_{b\text{-tag30}}$). The categories are filled in the order which is shown here.

Leptonic categories ($N_{\mathrm{lep}} \geq 1$)	Additional requirements
tH lep 0fwd	$N_{\mathrm{lep}} = 1$, $N_{\mathrm{jets}}^{\mathrm{central}} \leq 3$, $N_{b\text{-tag}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{forward}} = 0$
tH lep 1fwd	$N_{\mathrm{lep}} = 1$, $N_{\mathrm{jets}}^{\mathrm{central}} \leq 4$, $N_{b\text{-tag}} \geq 1$, $N_{\mathrm{jets}}^{\mathrm{forward}} \geq 1$
$t\bar{t}H$ lep	$N_{\mathrm{jets}}^{\mathrm{central}} \geq 2$, $N_{b\text{-tag}} \geq 1$, $Z \to \ell\ell$ veto
Hadronic categories ($N_{\mathrm{lep}} = 0$)	Additional requirements
$t\bar{t}H$ had BDT1	$N_{\mathrm{jets30}} \geq 3$, $N_{b\text{-tag30}} \geq 1$, $\mathrm{BDT}_{t\bar{t}H} > 0.92$
$t\bar{t}H$ had BDT2	$N_{\mathrm{jets30}} \geq 3$, $N_{b\text{-tag30}} \geq 1$, $0.83 < \mathrm{BDT}_{t\bar{t}H} < 0.92$
$t\bar{t}H$ had BDT3	$N_{\mathrm{jets30}} \geq 3$, $N_{b\text{-tag30}} \geq 1$, $0.79 < \mathrm{BDT}_{t\bar{t}H} < 0.83$
$t\bar{t}H$ had BDT4	$N_{\mathrm{jets30}} \geq 3$, $N_{b\text{-tag30}} \geq 1$, $0.52 < \mathrm{BDT}_{t\bar{t}H} < 0.79$
tH had 4j1b	$N_{\mathrm{jets}}^{\mathrm{central}} = 4$, $N_{b\text{-tag}} = 1$
tH had 4j2b	$N_{\mathrm{jets}}^{\mathrm{central}} = 4$, $N_{b\text{-tag}} \geq 2$

For the hadronic $t\bar{t}H$ final state, a BDT was trained in order to separate $t\bar{t}H$ events from ggH and continuum background events. A preselection is applied to events entering the BDT, requiring no leptons to be present, at least three jets with $|\eta| < 4.5$ and $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, among which one jet is required to be b-tagged. The input variables of the BDT include the sum of transverse momenta of all jets, the invariant mass of all jets, the total number of jets and the number of central jets and b-tagged jets. Based on the BDT discriminant, four $t\bar{t}H$ categories

with different S/B are defined by cutting on the distribution. In addition, two hadronic cut-based $4j1b$ and $4j2b$ categories, targeting tH final states, are defined. The BDT categories are filled prior to the hadronic tH categories, as otherwise their significance is significantly reduced. Likewise, defining the BDT categories prior to the tH categories degrades the expected limit on tH, as several events from the cut-based categories are selected by the BDT categories. Therefore, the requirements of the $4j1b$ and $4j2b$ categories are loosened by using a jet p_{T} threshold of 25 GeV in addition to the common 70% b-tagging WP, thereby allowing for a higher ggH fraction, to preserve the sensitivity to tH. The additional hadronic $5j1b$ and $5j2b$ categories are discarded, as most of the events which were previously selected by these categories enter the BDT categories. A summary of the definitions of the hadronic categories is also shown in Table 9.7.

Table 9.8: Expected event yields for 35 fb^{-1} for $t\bar{t}H$, $tHjb$ and WtH, the remaining Higgs boson processes, $B_{\mathrm{Higgs,rest}}$, and the continuum background, B_{ContBG}.

Category	$t\bar{t}H$	$tHjb$	WtH	$B_{\mathrm{Higgs,rest}}$	B_{ContBG}	Exp. limit tH [μ]
tH lep 0fwd	0.76	0.082	0.025	0.13	37	54.1
tH lep 1fwd	0.85	0.14	0.028	0.047	16	28.0
$t\bar{t}H$ lep	2.2	0.022	0.059	0.0098	14	67.4
Leptonic						22.2
$t\bar{t}H$ had BDT1	1.3	0.0093	0.029	0.029	9.5	115.9
$t\bar{t}H$ had BDT2	1.5	0.031	0.041	0.11	27	83.1
$t\bar{t}H$ had BDT3	0.50	0.018	0.013	0.047	11	>200
$t\bar{t}H$ had BDT4	1.9	0.10	0.056	0.32	86	52.2
tH had1	0.89	0.21	0.031	1.3	320	57.6
tH had2	0.23	0.063	0.0035	0.25	40	84.3
Hadronic						26.8
Combination						15.9

Furthermore, Table 9.8 shows the expected event yields for $tHjb$, WtH, $t\bar{t}H$ and the remaining Higgs boson backgrounds, $B_{\mathrm{Higgs,rest}}$, and the expected limit on tH for the merged tH+$t\bar{t}H$ categorization. For the modified leptonic categories, the sensitivity to tH production is almost unchanged, as the requirements are on the one hand tightened by using a tighter b-tagging requirement, but on the other hand loosened by discarding the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. In contrast, the expected limit of the hadronic categorization improves due the loosened jet p_{T} requirement in the $4j$ categories. However, it should be noted that these expected limits are calculated considering only statistical uncertainties and, hence, the large systematic uncertainty on ggH+b is not considered here. The total expected limit on tH production, provided by the nine tH+$t\bar{t}H$ categories, yields approximately 16 μ, with the dominant contribution originating from the leptonic categorization. Finally, Figure 9.7 shows the fraction of the different Higgs boson production processes in the tH and $t\bar{t}H$ categories, as well as the remaining categories targeting WH/ZH, VBF and ggH Higgs boson production modes. The definition of the categories targeting final states without top quarks is shown in Appendix B.3.

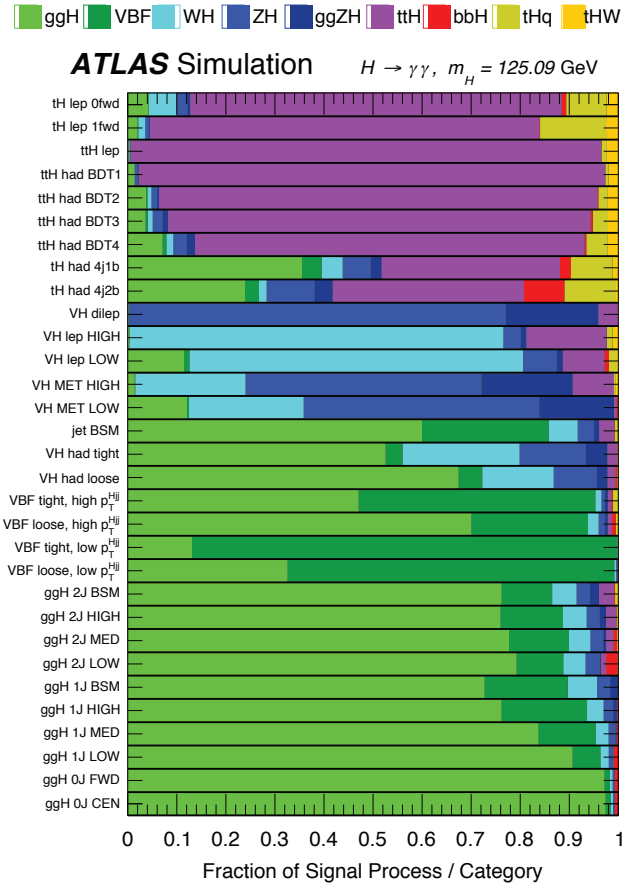


Figure 9.7: Relative contributions of the different Higgs boson production processes to the categories of the Higgs boson couplings measurement [2]. These include the tH and $t\bar{t}H$ categories and categories targeting WH/ZH, VBF and ggH final states. The categories are filled in the order which is shown here.

9.5 Continuum background modeling

A key element for measurements in the $H \to \gamma\gamma$ decay channel is the choice of background model, which is used to estimate the continuum background contribution in the S+B fit to the $m_{\gamma\gamma}$ distribution of each category. In contrast to the methodology described in Section 7.3, data-driven background templates are used to choose a background model for the tH and $t\bar{t}H$ categories, due to the absence of MC samples that sufficiently describe the continuum background. This strategy is similar to the approach that was used for the Run 1 measurement of top-quark-associated production in the $H \to \gamma\gamma$ channel [140]. For this thesis, the approach has been reinvestigated and validated for the datasets collected at $\sqrt{s} = 13\,\text{TeV}$. Additionally, the background model and the associated systematic uncertainty have been evaluated, based on this approach, for the measurements of $t\bar{t}H$ using $13.3\,\text{fb}^{-1}$ of data [1], the $t\bar{t}H$+tH measurement with $36.1\,\text{fb}^{-1}$ of data [2] and the measurement of $t\bar{t}H$ production using $79.8\,\text{fb}^{-1}$ of data [4], which contributed to the observation of $t\bar{t}H$ production at the ATLAS experiment. In the following, the strategy for choosing a background model in the tH and $t\bar{t}H$ categories is described and the results for the $t\bar{t}H$+tH measurement using $36.1\,\text{fb}^{-1}$ are presented.

9.5.1 Data-driven background templates

As described in Section 7.3.3, the continuum background model is chosen by minimizing the bias on the extracted signal yield due to the chosen fit function. This is done by performing an S+B fit to a high statistics B-only template and imposing certain requirements on the maximum amount of fitted signal, $N_{\text{spur}}^{\text{max}}$. If the background template is characterized by large statistical fluctuations, the fit function adapts to the fluctuations in the $m_{\gamma\gamma}$ distribution and the S+B fit typically results in a large amount of fitted signal. In this case, the value of $N_{\text{spur}}^{\text{max}}$ does not represent the real bias due to the chosen fit function and the criteria on the maximum amount of fitted signal fail. Due to the absence of diphoton MC samples that provide a sufficient number of events for final states with leptons and a high jet multiplicity, a data-driven approach is chosen. The background templates are defined using data events with $NTNI$ photons, which are characterized by loosened photon ID and isolation criteria. Hence, the sample of $NTNI$ photons is characterized by a larger number of events than the sample of TI photons and shows a falling $m_{\gamma\gamma}$ spectrum over the full range from 105 to 160 GeV, due to the negligible fraction of $H \to \gamma\gamma$ events. Furthermore, events with $NTNI$ photons are assumed to have an $m_{\gamma\gamma}$ shape similar to the TI continuum background events, as both are dominated by non-Higgs boson production processes with additional prompt or misidentified photons in the final state. However, as there are differences in the kinematics of prompt and *fake* photons and the fraction of *fake* photons differs between data events with TI and $NTNI$ photons, a validation of this approach is performed by comparing the shape of the respective $m_{\gamma\gamma}$ distributions.

For the available dataset, the background templates, which are defined based on $NTNI$ events and the other nominal selection requirements of the tH and $t\bar{t}H$ categories, are usually characterized by large statistical fluctuations and are hence not suitable for choosing a background template. This is indicated by failed spurious signal criteria for most of the tested functional forms. Hence, the nominal event selection requirements are successively loosened to obtain background templates with smaller statistical fluctuations. Reversing or dropping certain selection requirements might influence the shape of the $m_{\gamma\gamma}$ distribution of the continuum background. Therefore, the requirements are loosened step-wise, in order to test whether the obtained background template is suitable for choosing a background model after each step. The shape differences between the nominal TI sideband distribution and the respective background template are investigated for each category to validate this approach. Three different types of templates, denoted by QCD::i, are defined based on $NTNI$ photons by changing the following requirements:

- **QCD::1:** Only the requirements on the photons are loosened, by selecting events with $NTNI$ instead of TI photons. The remaining selection requirements remain unchanged.
- **QCD::2:** Events with $NTNI$ photons are selected and the b-tagging requirements are dropped.
- **QCD::3:** Events with $NTNI$ photons are selected, the b-tagging requirements are dropped and lepton requirements are replaced by jet requirements in the leptonic categories, while in the hadronic categories the jet p_{T} cuts are loosened.

It is assumed that the systematic uncertainty on the chosen background model, which is defined based on $N_{\text{spur}}^{\text{max}}$, is conservatively large for regions with a small number of expected events, as the result of the S+B fit is dominated by statistical fluctuations. This assumption is supported by the results of the spurious signal fits, which are discussed in the following sections. In

addition, due to the low number of expected events in the data $m_{\gamma\gamma}$ sidebands of the tH and $t\bar{t}H$ categories, only functions with a low number of degrees of freedom are considered to be a reasonable choice.

Table 9.9: Definition of the QCD::2 background templates for the tH and $t\bar{t}H$ categories, which are constructed based on $NTNI$ events.

Category	Requirements for QCD::2 background templates
tH lep 0fwd	$N_{\text{lep}} = 1$, $N_{\text{jets}}^{\text{central}} \leq 3$, $N_{\text{jets}}^{\text{forward}} = 0$
tH lep 1fwd	$N_{\text{lep}} = 1$, $N_{\text{jets}}^{\text{central}} \leq 4$, $N_{\text{jets}}^{\text{forward}} \geq 1$
$t\bar{t}H$ lep	$N_{\text{lep}} \geq 1$, $N_{\text{jets}}^{\text{central}} \geq 2$
$t\bar{t}H$ had BDT	$N_{\text{lep}} = 0$, $N_{\text{jets30}} \geq 3$, $N_{b\text{-tag30}}$ increased by 1, $\text{BDT}_{t\bar{t}H} > 0.52$
tH had $4j1b$, $4j2b$	Veto on events from BDT categories, $N_{\text{lep}} = 0$, $N_{\text{jets}}^{\text{central}} = 4$

While for the $t\bar{t}H$ measurement with $13.3\,\text{fb}^{-1}$ only the QCD::3 templates were suitable for choosing a background model, for the increased datasets of $36.1\,\text{fb}^{-1}$ and $79.8\,\text{fb}^{-1}$ the QCD::2 templates already provide an $m_{\gamma\gamma}$ distribution with a sufficient number of events. For the definition of the QCD::2 templates of the leptonic categories and the cut-based hadronic $4j1b$ and $4j2b$ categories, the b-tagging requirements are dropped. Thus, the tH had $4j1b$ and $4j2b$ categories share a common background template. For the hadronic BDT-based categories, loosening the b-tagging requirement is more complicated, as the number of b-tagged jets is one of the input variables of the BDT. First, the b-tagging requirement in the preselection is dropped. Then, the number of b-tagged jets in each event is increased by one before applying the BDT weights, as the BDT is trained to classify events with at least one b-tagged jet in the final state as signal. As for some of the BDT categories the spurious signal criteria still fail when using the QCD::2 templates, the four templates of the different BDT categories are merged into one inclusive template, which is used to derive the background model for each of the categories. As there is only a small correlation between the BDT discriminant and the $m_{\gamma\gamma}$ distribution, this is considered to be a valid approach. In Table 9.9, a summary of the definitions of the QCD::2 background templates for the different tH and $t\bar{t}H$ categories is shown.

9.5.2 Validation of the data-driven strategy

The data-driven strategy for the background modeling, described in the previous section, is validated by investigating the shape differences between the $m_{\gamma\gamma}$ distribution of the nominal TI continuum background sidebands and the background templates. In Figure 9.8, the TI sidebands and the QCD::1 and QCD::2 templates are shown as an example for the tH lep 1fwd category. While the nominal sidebands only include 17 events, this number is increased by a factor of approximately 10 and 100 for the QCD::1 and QCD::2 templates, respectively. The $m_{\gamma\gamma}$ distribution of the QCD::1 template shows large statistical fluctuations and several bins with either 0 or 1 entries. In contrast, the QCD::2 template shows a smoothly falling $m_{\gamma\gamma}$ distribution. The $m_{\gamma\gamma}$ distributions of the TI sidebands and the QCD::1 and QCD::2 background templates for the remaining categories are presented in Appendix C.1.

The TI sidebands and the QCD::1 and QCD::2 templates are fitted with an Exponential and a PowerLaw function, in order to compare the shapes of the fit functions and to investigate if there are significant differences in the slopes of the $m_{\gamma\gamma}$ distributions. Due to the small number of events in the $m_{\gamma\gamma}$ sidebands, these two functional forms are expected to be the

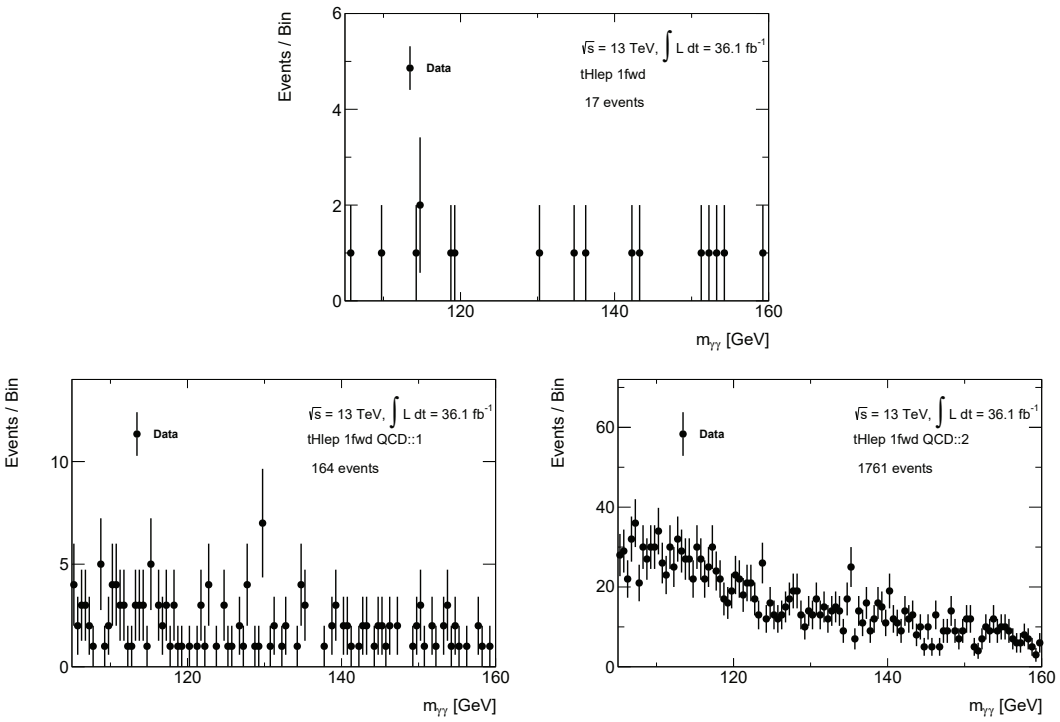


Figure 9.8: Example $m_{\gamma\gamma}$ distributions of the TI sidebands and the QCD::1 and QCD::2 background templates for the tH lep 1fwd category. Before performing the S+B fit, the background templates are normalized to the number of TI sideband events in the $m_{\gamma\gamma}$ distribution of the respective category.

only reasonable choices for the background model, as they have the fewest number of free parameters among the considered functional forms. Furthermore, the leptonic categories and the hadronic BDT categories are merged for this comparison, respectively, due to the small number of events in the $m_{\gamma\gamma}$ sidebands. In Figure 9.9, a comparison of the Exponential function fitted to the TI sidebands and the QCD::1 and QCD::2 background templates is shown. For all types of categories, small differences between the slopes of the Exponential fit function are observed. However, the differences are relatively small and in order to quantify if these differences are statistically significant, the slope parameters of the Exponential function, including their statistical uncertainties, are investigated.

Figure 9.10 shows a comparison of the slope parameters for the Exponential function fitted to the TI sidebands and the QCD::1 and QCD::2 background templates. A tendency towards a more steeply falling function is observed for the QCD::1 and QCD::2 background templates for the leptonic categories and the hadronic BDT $t\bar{t}H$ categories. However, the slope parameters for the different background templates are still in agreement within the large statistical uncertainties. Therefore, the QCD::2 background templates are considered to be a reasonable choice for selecting a background model. In Appendix C.2, equivalent plots are presented assuming a PowerLaw background model, which show a similar trend for the slope of the background templates.

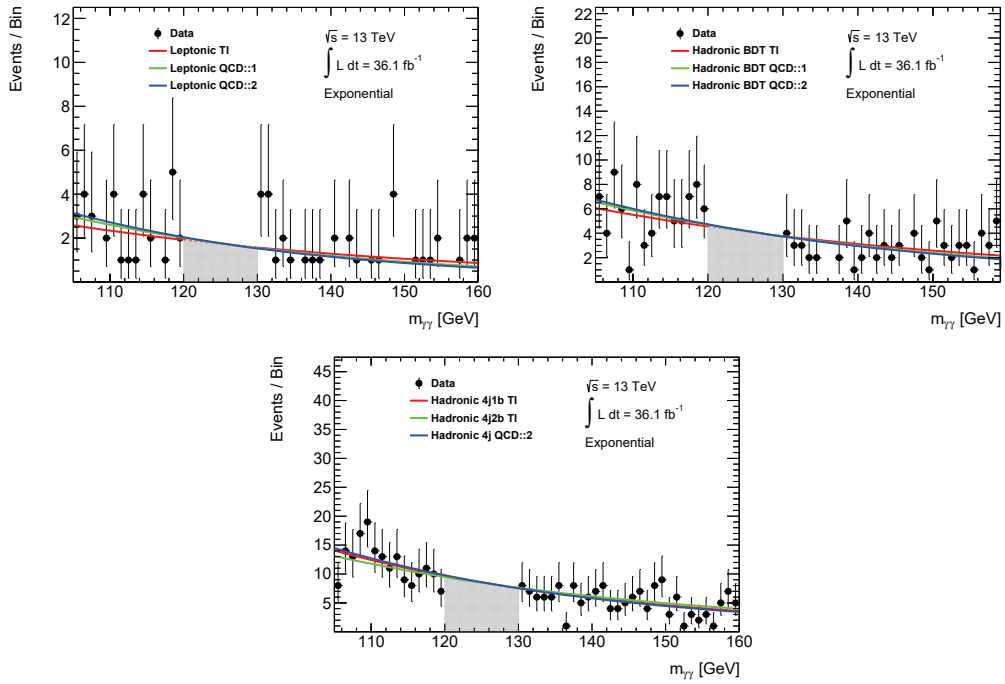


Figure 9.9: Comparison of the Exponential functions fitted to the TI sideband events and the QCD::1 and QCD::2 background templates for the combined leptonic tH+$t\bar{t}H$ categories, the combined hadronic $t\bar{t}H$ BDT categories and the cut-based hadronic tH categories.

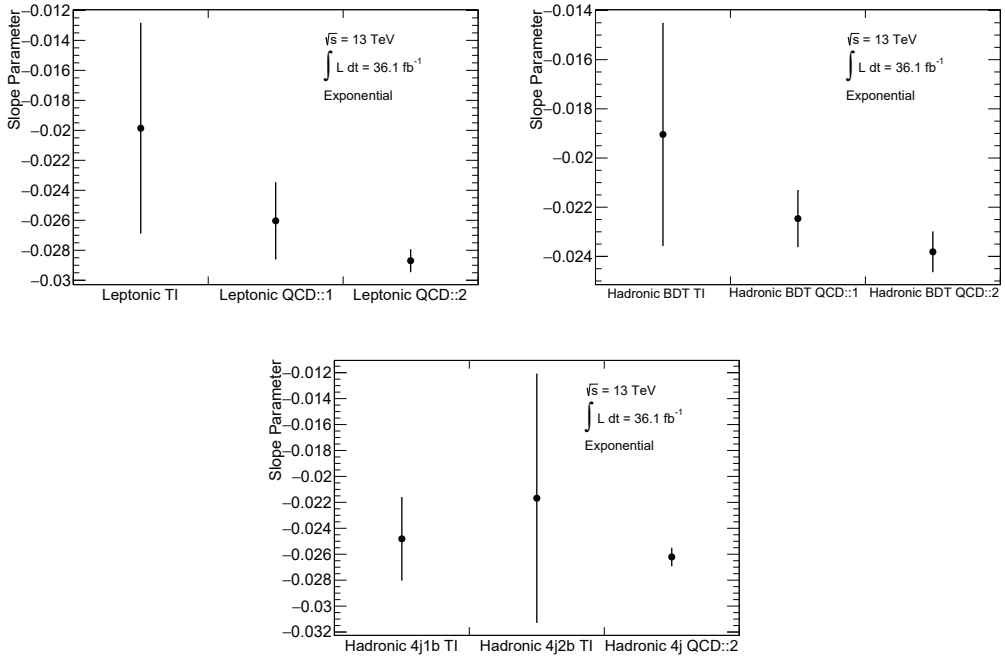


Figure 9.10: Comparison of the slope parameters, including statistical uncertainties, for the Exponential function fitted to the TI sideband events and the QCD::1 and QCD::2 background templates for the combined leptonic tH+$t\bar{t}H$ categories, the combined hadronic $t\bar{t}H$ BDT categories and the cut-based hadronic tH categories.

9.5.3 Spurious signal test

Finally, an S+B fit is performed to the QCD::2 background templates, assuming the different background models described in Section 7.3.3. The signal model is determined, as described in Section 7.2, by fitting a DCB function to the $m_{\gamma\gamma}$ distribution obtained from the MC samples of all contributing Higgs boson production processes for each category. First, the background templates are normalized to the observed number of TI sideband events for each category. Then, an S+B fit is performed as a function of $m_{\gamma\gamma}$ for each considered background model and criteria are imposed on the maximum amount of fitted signal, $N_{\text{spur}}^{\text{max}}$, as defined in Equations (7.3) and (7.4). Among the functions which pass at least one of the criteria, the functional form with the fewest number of degrees of freedom is chosen as background model. Finally, $N_{\text{spur}}^{\text{max}}$ is assigned as a systematic uncertainty on the background model. In Figure 9.11, the fitted signal is shown as a function of $m_{\gamma\gamma}$ for the QCD::2 background templates of the two leptonic tH categories. It is noticeable that the fitted spurious signal has a similar shape for all tested functional forms, although their number of free parameters ranges from one to five. This can be explained by the comparably small number of events in the B-only templates, as the shape of the fitted signal is dominated by the statistical fluctuations in the background template. When fitting a B-only template that is characterized by even larger statistical fluctuations, the fitted spurious signal strongly increases and the two criteria on $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ and $N_{\text{spur}}^{\text{max}}/\delta S$ are usually not fulfilled anymore. Hence, the QCD::2 background templates, which are characterized by relatively large statistical fluctuations, are expected to provide a conservative background modeling uncertainty. Further loosening the requirements for the background templates would decrease the statistical fluctuations, but it would also potentially create a bias in the shape of the background template compared to the initial $m_{\gamma\gamma}$ distribution. Hence, a compromise was made. The considered approach is assumed to be valid, as the shape of the background model for TI events and the different QCD::i background templates agree with each other within the statistical uncertainties, as shown in the previous section.

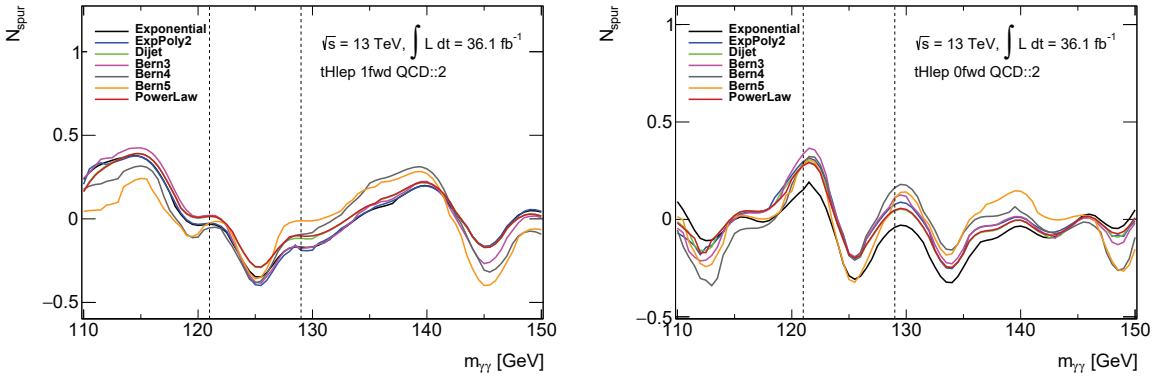


Figure 9.11: Fitted signal as a function of $m_{\gamma\gamma}$ for the tH lep 1fwd and the tH lep 0fwd QCD::2 background templates. Various functional forms with a different number of free parameters are considered. $N_{\text{spur}}^{\text{max}}$ is defined as the maximum fitted signal in the region between 121 and 129 GeV, which is indicated by the dashed lines.

The same procedure is repeated for each category, in order to select a functional form based on the amount of fitted spurious signal. A summary of the chosen functional forms for each category is shown in Table 9.10, including $N_{\text{spur}}^{\text{max}}$ and the two criteria on $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$. For all categories, functions with only one free parameter are chosen, i.e. the Exponential or the

PowerLaw function. The systematic uncertainty due to the choice of background model, which is assigned to the number of fitted Higgs boson signal events in each category, ranges between 0.037 and 0.29 events.

The criterion on $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ fails for several categories, while the criterion on $N_{\text{spur}}^{\text{max}}/\delta S$ is generally fulfilled. In Figures 9.12 and 9.13, the two criteria are illustrated for the different QCD::2 background templates and the Exponential and PowerLaw background models. In order to fulfill one of the criteria, the distribution needs to be fully contained in the black square indicating $121 \leq m_{\gamma\gamma} \leq 129\,\text{GeV}$ and $|N_{\text{spur}}/\delta S| < 20\%$ or $|N_{\text{spur}}/S_{\text{ref}}| < 10\%$, respectively. In Appendix C.3, the distributions of the fitted spurious signal as a function of $m_{\gamma\gamma}$ are shown for all QCD::2 background templates. In addition, the detailed results for each tH and $t\bar{t}H$ category and each tested functional form are shown in Appendix C.4.

Table 9.10: Chosen continuum background models for the $t\bar{t}H$ and tH categories, derived from an S+B fit to the QCD::2 templates, and the associated systematic uncertainty on the background model, $N_{\text{spur}}^{\text{max}}$. Additionally, the values of the two criteria on $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}/S_{\text{ref}}$ are shown.

Category	Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$
tH lep 0fwd	PowerLaw	+12.9	+28.5	+0.29
tH lep 1fwd	PowerLaw	−19.6	–26.2	–0.29
$t\bar{t}H$ lep	PowerLaw	−16.2	–10.7	–0.25
$t\bar{t}H$ had BDT1	Exponential	+3.57	+3.75	+0.053
$t\bar{t}H$ had BDT2	Exponential	+4.55	+4.77	+0.082
$t\bar{t}H$ had BDT3	Exponential	+3.72	+9.72	+0.057
$t\bar{t}H$ had BDT4	Exponential	+10.1	+15.4	+0.38
tH had 4j1b	PowerLaw	–3.47	–9.73	–0.25
tH had 4j2b	PowerLaw	–1.38	–6.52	–0.037

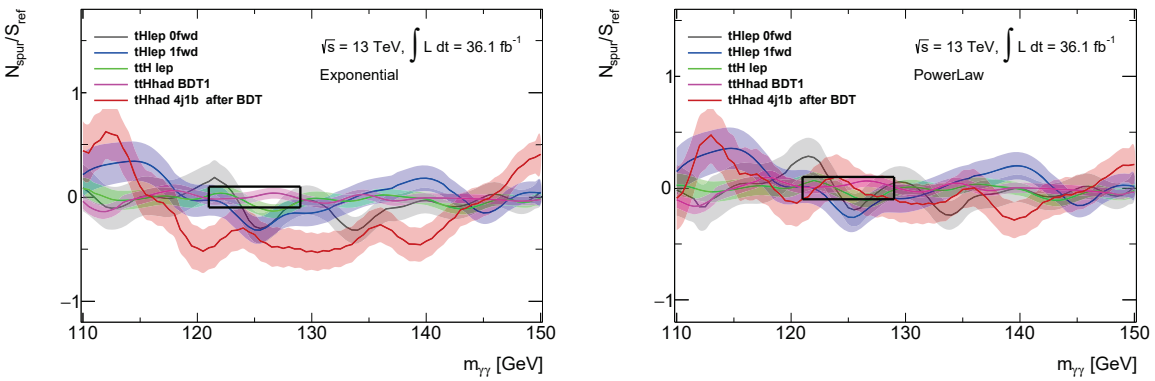


Figure 9.12: Illustration of the $N_{\text{spur}}/S_{\text{ref}}$ criterion for the Exponential and the PowerLaw function for the different QCD::2 background templates. The ratio is required to be smaller than 10% in the region $121 \leq m_{\gamma\gamma} \leq 129\,\text{GeV}$, which is indicated by the black square.

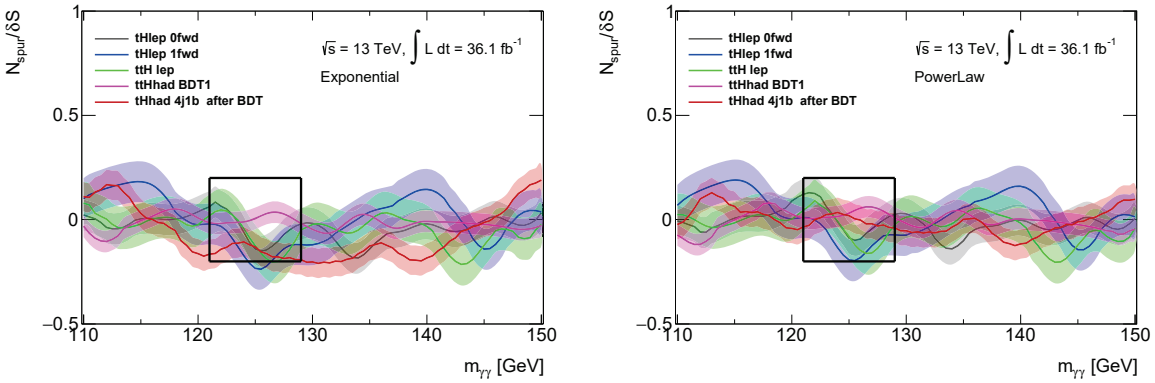


Figure 9.13: Illustration of the $N_{\text{spur}}/\delta S$ criterion for the Exponential and the PowerLaw function for the different QCD::2 background templates. The ratio is required to be smaller than 20% in the region $121 \leq m_{\gamma\gamma} \leq 129\,\text{GeV}$, which is indicated by the black square.

9.6 Results

A simultaneous unbinned likelihood fit is performed to the $m_{\gamma\gamma}$ distributions of the 31 categories, including the 9 categories targeting tH and $t\bar{t}H$ final states. Several sources of systematic uncertainty are considered in addition to the background modeling uncertainty, described in the previous section, which are treated as NPs in the fit. These systematic uncertainties include uncertainties related to signal and background modeling, detector related uncertainties and theory uncertainties, which are discussed in detail in Ref. [2]. However, due to the small number of expected and observed events in the tH and $t\bar{t}H$ categories, the statistical uncertainty is dominant in these categories.

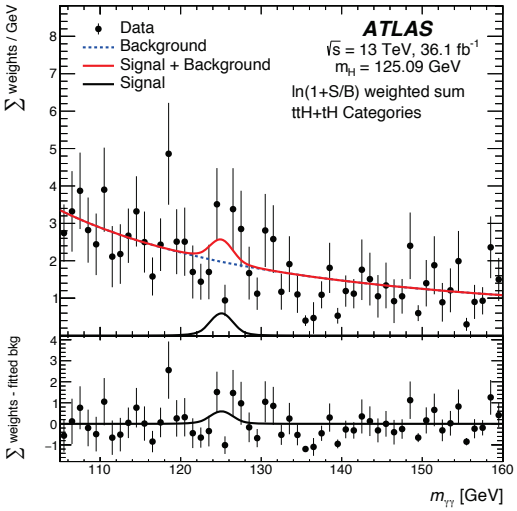


Figure 9.14: Diphoton invariant mass distribution of the $t\bar{t}H$+tH categories [2]. The events are weighted by $\ln(1 + S/B)$, with S/B representing the expected signal to background ratio in each category. The red line indicates the result of the S+B fit.

In Figure 9.14, the $m_{\gamma\gamma}$ distribution of the $t\bar{t}H$+tH categories is shown, including the signal and background components of the S+B fit. In order to illustrate the effect of the categorization, the

events in the $m_{\gamma\gamma}$ distribution are weighted by $\ln(1+S/B)$, with S/B representing the expected signal to background ratio in each category. Based on the models which are determined for the Higgs boson signal and the continuum background, the Higgs boson signal is extracted for the combined $t\bar{t}H$ and tH categories.

9.6.1 Signal strength and cross section measurement

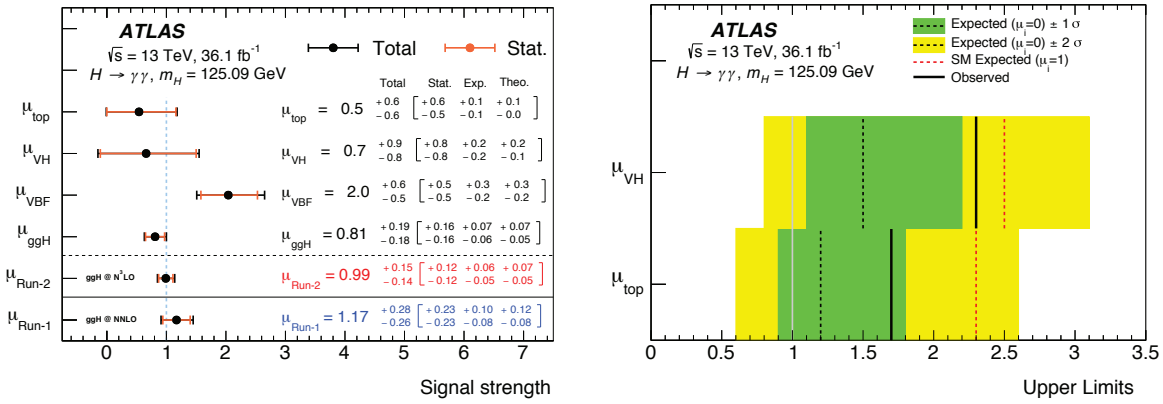
The production cross section of top-quark-associated Higgs boson production, σ_{top}, and the associated signal strength, μ_{top}, are extracted from the Higgs boson signal, which is measured from a simultaneous fit to all categories. The production cross section is measured for a Higgs boson rapidity $|y_H| < 2.5$, for which the detector acceptance is close to 100%, and amounts to

$$\sigma_{\text{top}} \times \mathcal{BR}(H \to \gamma\gamma) = 0.7\,^{+0.9}_{-0.7}\,\text{fb}.$$

A comparison to the corresponding SM expectation, $\sigma^{SM}_{\text{top}} \times \mathcal{BR}(H \to \gamma\gamma) = 1.3\,^{+0.9}_{-0.8}\,\text{fb}$, shows that the measured cross section σ_{top} is almost a factor 2 smaller than the cross section expected by the SM, but the two values agree with each other within the uncertainties. The agreement between the measured value and the SM expectation can be quantified in terms of a signal strength:

$$\mu_{\text{top}} = 0.5\,^{+0.6}_{-0.5}\,(\text{stat.})\,^{+0.1}_{-0.1}\,(\text{exp.})\,^{+0.1}_{-0.0}\,(\text{theo.}).$$

The measured value of μ_{top} is associated with a large statistical uncertainty, while the experimental and theoretical uncertainties provide minor contributions to the total uncertainty. As no significant deviation from the SM expectation of $\mu_{\text{top}} = 1$ is found, the measured value of μ_{top} is interpreted as a downward fluctuation of the data. In Figure 9.15a), a comparison of the measured signal strengths for the different Higgs boson production processes and the inclusive



a) b)

Figure 9.15: a) Signal strengths for the different Higgs boson production processes, which are measured from the 31 categories, and a comparison of μ for inclusive Higgs boson production with $H \to \gamma\gamma$ in Run 1 and Run 2 [2], b) asymptotic limits on the signal strength at 95% CL for top-quark-associated Higgs boson production, μ_{top}, and the production of a Higgs boson in association with a vector boson, μ_{VH} [2].

Run 1 and Run 2 values are shown. Similar to top-quark-associated Higgs boson production, the measured signal strength for VH production indicates a downward fluctuation of the data, while for ggH the signal strength is in good agreement with the SM expectation and for VBF the measured cross section is slightly larger than the cross section predicted by the SM. However, the signal strength for inclusive Higgs boson production in the $H \to \gamma\gamma$ channel, $\mu_{\text{Run-2}}$, is in very good agreement with the SM expectation and with the value measured during Run 1, $\mu_{\text{Run-1}}$. The statistical uncertainty of the inclusive Run 2 value is almost halved compared to the Run 1 measurement, but it remains the dominant uncertainty for the analyzed dataset. The results of the combined measurement of $t\bar{t}H$ and tH production correspond to an observed (expected) significance of $1.0\,\sigma$ ($1.8\,\sigma$). Hence, no evidence for top-quark-associated Higgs boson production is found and an upper limit is set on the signal strength μ_{top}, as shown in Figure 9.15b). The observed limit amounts to $1.7\,\mu$ and can be compared to an expected limit of $2.3\,\mu$, when assuming the S+B hypothesis with $\mu = 1$, or an expected limit of $1.2\,\mu$, when the B-only hypothesis with $\mu = 0$ is assumed.

9.6.2 Couplings interpretation

The measured value of μ_{top} provides no indication for a Yukawa coupling of the top quark with a negative sign, $\kappa_t = -1$, which would be accompanied by a large increase of the cross section for tH production and an excess in the combined $t\bar{t}H$+tH measurement. In order to draw quantitative conclusions on Y_t, an interpretation of the data in terms of the couplings strength modifier κ_t is performed based on the minimal parameterization model described in Chapter 2.

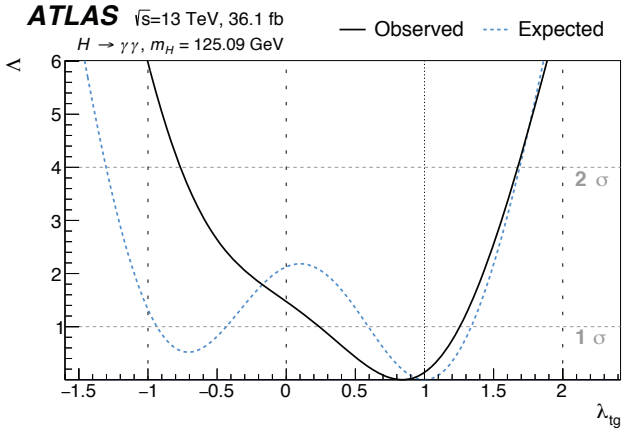


Figure 9.16: Expected and observed negative profile log-likelihood function Λ of the parameter λ_{tg}, which is given by the ratio of the coupling strength modifiers κ_t and κ_g [2].

A log-likelihood fit is performed in which the three coupling strength modifier ratios are profiled. The parameter λ_{tg} is allowed to be negative to exploit the sensitivity to negative values of κ_t through tH production. In addition to tH production, $gg \to ZH$ production also provides sensitivity to negative values of λ_{tg}, due to an interference of processes where the Higgs boson is radiated off a Z boson and processes involving box diagrams with top quarks. The best fit value of the coupling strength modifier ratio is measured to be

$$\lambda_{tg} = 0.8\,^{+0.4}_{-0.6}\,(\text{stat.})\,\pm 0.1\,(\text{exp.})\,^{+0.1}_{-0.0}\,(\text{theo.})$$

and is close to the SM expectation of 1. The observed and expected profile log-likelihood functions for the parameter λ_{tg} are shown in Figure 9.16. Values of $\lambda_{tg} < 0$ are not excluded, but positive values are favored by the analyzed data.

9.7 Discussion

The measurement of tH and $t\bar{t}H$ production in the $H \to \gamma\gamma$ channel, presented in this chapter, makes use of several new concepts. Dedicated tH categories, which were optimized as part of this thesis, were included in an ATLAS measurement for the first time. In contrast to $t\bar{t}H$ production, these provide sensitivity to the BSM scenario of $\kappa_t = -1$. With the considered dataset of $36.1\,\text{fb}^{-1}$, negative values of κ_t could not be excluded, but, similar to the Run 2 CMS results, positive values are favored by the data. In contrast to the κ_t interpretation of the CMS analysis targeting tH production [144], no assumptions were made on the coupling of the Higgs boson to other particles.

The background models for both, the tH and $t\bar{t}H$ categories, were derived based on data-driven background templates by estimating the bias due to the choice of fit function. The associated background modeling uncertainty is dominated by the statistical fluctuations in the background templates, but it has a minor impact on the final results for which the statistical uncertainty is dominant. The hadronic $t\bar{t}H$ categories were defined based on a BDT discriminant. While increasing the sensitivity to $t\bar{t}H$ in the hadronic channel significantly, the usage of this BDT is suboptimal for a combined measurement of $t\bar{t}H$+tH as it degrades the sensitivity to tH in the cut-based tH categories. An additional discrimination between $t\bar{t}H$ and tH events would need to be achieved by either filling the hadronic tH categories prior to the BDT-based $t\bar{t}H$ categories or using additional MVA techniques to discriminate between $t\bar{t}H$ and tH. However, most of the sensitivity to tH production is provided by the leptonic categories, which target $tHjb$ final states either with or without a forward jet.

Due to the low expected significance of tH production, the results are quoted for inclusive top-quark-associated Higgs boson production, for which no significant deviations from the SM prediction are observed. The observation of tH production in the $H \to \gamma\gamma$ decay channel alone is unlikely with the dataset collected during Run 2 and will probably be still without reach in Run 3. In order to be able to measure this rare process, a larger dataset and/or the study of additional decay channels will be necessary.

In contrast to tH production, $t\bar{t}H$ has been measured in the $H \to \gamma\gamma$ decay channel[4] with a significance of $4.1\,\sigma$ using a dataset corresponding to $79.8\,\text{fb}^{-1}$. Despite the larger amount of data compared to the result obtained with $36.1\,\text{fb}^{-1}$, the statistical uncertainty is still dominant in this measurement. However, this is going to change in the near future with the much larger datasets collected during Run 3 and, at a later stage, by the HL-LHC. An important systematic uncertainty for hadronic final states results from Higgs boson production, in particular ggH, in association with additional b-jets in the final state. As explained in Chapter 9.3, a conservative 100% uncertainty is imposed on this kind of background, as the accuracy of the MC predictions is not known. In the following two chapters, a first attempt to measure this background and studies on possible improvements of this measurement, based on MVA techniques, are presented.

[4] This measurement was updated with a dataset corresponding to $139\,\text{fb}^{-1}$, where $t\bar{t}H(H \to \gamma\gamma)$ production was measured with a significance of $4.9\,\sigma$ [141].

Chapter 10

Differential cross section measurement of $N_{b\text{-jets}}$

The production of a Higgs boson in association with additional b-jets (H+b-jets), which do not originate from decays of particles produced during the hard scattering process, is an important background in several $H \to \gamma\gamma$ measurements which study hadronic final states with b-jets. These include measurements of $t\bar{t}H$ and tH production [1–4, 140, 141], Higgs boson pair production in the $\gamma\gamma bb$ final state [146, 147] or the search for vector-like B quarks with $B \to bH(H \to \gamma\gamma)$ [148]. Similar to the $t\bar{t}H$+tH measurement presented in the previous chapter, these measurements assume a 100% systematic uncertainty on processes where a Higgs boson is produced in association with additional b-jets, such as ggH, VBF and WH. At the moment, the dominant uncertainty in these measurements is the statistical uncertainty, but the uncertainty on H+b-jets is among the dominant systematic uncertainties and will become more important for larger datasets. The 100% uncertainty on H+b-jets is motivated by the discrepancies between data and MC predictions which were observed in the measurements of other processes with additional b-jets in the final state, such as $t\bar{t} + b(b)$ [21]. In MC simulation, the additional b quarks are mostly not simulated during the Matrix Element calculation, but only during the parton showering, which provides an approximation for higher-order corrections in QCD. Hence, it is not known how well this kind of background is modeled in MC simulation. The dominant contribution to the H+b-jets background is ggH+b, thus a measurement of this process is of particular interest. An example Feynman diagram for ggH+b production, where the additional b quarks are produced by a radiated gluon, is shown in Figure 10.1.

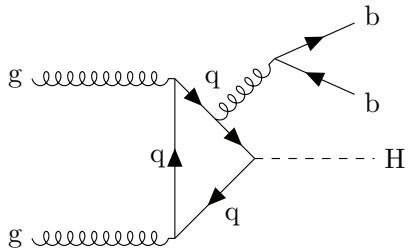


Figure 10.1: Example Feynman diagram for Higgs boson production via ggH with additional b quarks in the final state.

In this chapter, a first measurement of the differential cross section of the number of b-jets, $N_{b\text{-jets}}$, produced in association with a Higgs boson in the $H \to \gamma\gamma$ decay channel using $79.8\,\text{fb}^{-1}$ of ATLAS data is presented [4]. This measurement provides information about H+b-jets in the $N_{b\text{-jets}} = 1$ bin, as explained in the following. First, the strategy for the measurement is explained, followed by a description of the different steps of the measurement. These include the

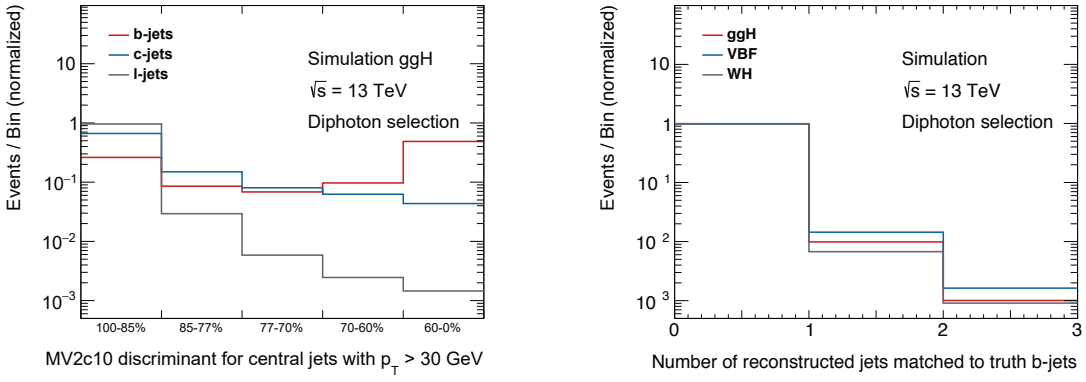
definition of the event selection, the modeling of signal and background, the signal extraction fit, the unfolding of the measured cross section to particle level and the estimation of systematic uncertainties. Finally, the results of the measurement are discussed.

10.1 Strategy to constrain H+b-jets

The production of ggH in association with b-jets in the final state is a comparably rare process with relatively large backgrounds and has not been measured so far. The production of VBF+b-jets and WH+b-jets is even more rare, but these processes are also less important backgrounds in $H \to \gamma\gamma$ measurements. Different approaches were studied for a measurement of H+b-jets production in the $H \to \gamma\gamma$ decay channel. The most distinct features of this final state are the two high-p_{T} photons, which are specific to the $H \to \gamma\gamma$ channel, and the b-jets, which can be identified using the MV2c10 b-tagging algorithm. While the $m_{\gamma\gamma}$ distribution provides a good separation between H+b-jets and the continuum background, the b-tagging information allow to discriminate between events with jets originating from b, c and light flavor quarks (b-, c-, and ℓ-jets) in the final state. For this reason, two approaches were investigated:

1. An unbinned likelihood fit to the $m_{\gamma\gamma}$ distribution of events with b-tagged jets
2. A binned likelihood fit to the MV2c10 b-tagging discriminant

The second approach is motivated by measurements of other processes, such as $t\bar{t}$, with additional b-jets in the final state [22]. An example distribution of the pseudo-continuous MV2c10 discriminant for ggH production, which illustrates the discrimination between b-jets, c-jets and ℓ-jets, is shown in Figure 10.2a). The pseudo-continuous MV2c10 discriminant consists of five bins which correspond to the different calibrated b-tagging WPs.



a) b)

Figure 10.2: a) Pseudo-continuous MV2c10 discriminant, where the five bins correspond to the calibrated b-tagging WPs, for reconstructed jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ ΔR-matched to truth b-jets, c-jets and ℓ-jets in ggH events, b) the number of reconstructed jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ ΔR-matched to truth b-jets for ggH, VBF and WH production.

For a fit to the MV2c10 discriminant, a selection of events in a narrow region around the Higgs boson mass is performed and templates, describing the shape of the discriminant, are derived for the H+b-jets, H+c-jets and H+ℓ-jets contributions and the continuum background. The normalization of the H+b-jets, H+c-jets and H+ℓ-jets contributions is then adjusted in a fit

to data. Hence, this approach allows to directly measure the H+b-jets contribution. But as the templates for the Higgs boson processes are derived from simulation, this approach strongly relies on the modeling of b-jets, c-jets and ℓ-jets in MC simulation. The other approach for a measurement of H+b-jets is a fit to the $m_{\gamma\gamma}$ distribution of b-tagged jets. As the invariant mass of the diphoton system provides a good separation between the Higgs boson signal and the continuum background, a fit to $m_{\gamma\gamma}$ is the usual approach in $H \to \gamma\gamma$ measurements. In this case, the fitted Higgs boson signal includes contributions from events with b-jets, c-jets and ℓ-jets, whose fractions depend on the chosen b-tagging WP. However, in a similar way the 100% uncertainty on the H+b-jets background is assigned to events with b-tagged jets in the final state, which also include a certain fraction of mis-tagged c-jets and ℓ-jets. The contribution of c-jets and ℓ-jets can be accounted for by unfolding the measured cross section to the particle level cross section, as explained in the following. For a first attempt to measure H+b-jets, the unbinned likelihood fit to the $m_{\gamma\gamma}$ distribution is chosen as it is a well-studied approach in the $H \to \gamma\gamma$ channel and it less strongly relies on assumptions from MC simulation. Furthermore, for a fit to the MV2c10 discriminant, a special pseudo-continuous calibration of the different b-tagging WPs is required, which was not available on the timescale of this measurement.

The measurement is performed as a differential cross section measurement of the $N_{b\text{-jets}}$ distribution, based on jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ selected by the 70% b-tagging WP. This approach is similar to a previous measurement of the differential $N_{b\text{-jets}}$ cross section in the $H \to ZZ^* \to 4\ell$ channel with a dataset corresponding to $36.1\,\mathrm{fb}^{-1}$ [149]. In this measurement, the $N_{b\text{-jets}}$ distribution was measured for inclusive Higgs boson production with $H \to ZZ^*$. The distribution was split into two bins with $N_{b\text{-jets}} = 0$ and ≥ 1, respectively. In both bins, the measured cross section was slightly larger than the prediction from MC simulations, although still being in agreement with the prediction within the uncertainties. For the measurement presented in this chapter, the relatively tight 70% b-tagging WP is chosen in order to reduce the fraction of misidentified c-jets and ℓ-jets, which increases for b-tagging WPs with higher efficiencies. Due to the limited sensitivity to ggH+b, as discussed in more detail in Section 10.2.3, the $N_{b\text{-jets}}$ distribution is measured for inclusive Higgs boson production in the $H \to \gamma\gamma$ channel. Preselection requirements are defined for the events which enter the $N_{b\text{-jets}}$ distribution, in order to reduce the fraction of events where the b-jets originate from decays of particles produced during the hard scattering process, in particular $t\bar{t}H$. The preselection includes a veto on events with charged leptons ($N_{\mathrm{lep}} = 0$) and requires at least one central jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ to be present in each event ($N_{\mathrm{jets30}}^{\mathrm{central}} \geq 1$), as explained in more detail in the following section. The $N_{b\text{-jets}}$ distribution is then split into three bins, which are dominated by different processes:

- $N_{b\text{-jets}} = 0$: Dominated by ggH+ℓ-jets
- $N_{b\text{-jets}} = 1$: Enriched in ggH+b-jets
- $N_{b\text{-jets}} \geq 2$: Dominated by $t\bar{t}H$

The $N_{b\text{-jets}} = 1$ bin is enriched in ggH+b events and, hence, the cross section measured in this bin provides information about H+b-jets production. In Figure 10.2b), the number of reconstructed jets which can be ΔR-matched to b-hadrons based on MC truth information is shown for ggH, VBF and WH events. Most of theses events are characterized by only one reconstructed jet in the final state which originates from a b quark. This may happen e.g. when two b quarks from a gluon splitting are reconstructed as a single jet, due to their small angular separation. Hence, most ggH+b events enter the $N_{b\text{-jets}} = 1$ bin, while $t\bar{t}H$ events are often characterized by two reconstructed b-jets. For this reason, the last bin of the $N_{b\text{-jets}}$

distribution is defined based on the requirement $N_{b\text{-jets}} \geq 2$, resulting in an $N_{b\text{-jets}} = 1$ bin which is characterized by a smaller $t\bar{t}H$ fraction compared to a bin with $N_{b\text{-jets}} \geq 1$. In Figure 10.3, the contribution of each Higgs boson production process to the different bins of the $N_{b\text{-jets}}$ distribution is illustrated.

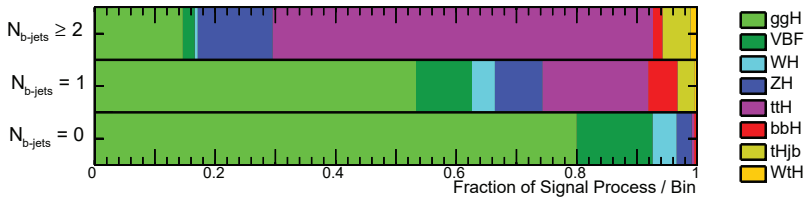


Figure 10.3: Relative contributions of the different Higgs boson production processes to the three bins of the $N_{b\text{-jets}}$ distribution, with additional requirements on N_{lep} and $N_{\text{jets30}}^{\text{central}}$ applied. The $N_{b\text{-jets}} = 0$ bin is dominated by ggH+light flavor jets, the $N_{b\text{-jets}} = 1$ bin is enriched in ggH+b, while the $N_{b\text{-jets}} \geq 2$ bin is dominated by $t\bar{t}H$ production.

It is not differentiated between the different Higgs boson production processes in the differential cross section measurement of $N_{b\text{-jets}}$, therefore the total number of Higgs boson signal events is extracted in each bin by performing an unbinned likelihood fit to the $m_{\gamma\gamma}$ distribution. The measured $N_{b\text{-jets}}$ distribution is then unfolded to the particle level cross section by deriving correction factors based on MC truth information, in order to correct for detector effects. As the correction factors are derived from events containing true b-jets, this also corrects the measured cross section for the amount of mis-tagged c-jets and ℓ-jets. The details of the measurement are presented in the following.

10.2 Event selection

In the following, the preselection requirements for the $N_{b\text{-jets}}$ measurement on reconstruction level and the definition of the fiducial volume on particle level are described, which were defined as part of this thesis. In addition, the $N_{b\text{-jets}}$ variable is specified based on the definitions of b-jets on reconstruction and particle level. While the reconstruction level requirements are applied to both, data and MC simulation, the particle level definition is used to correct the measured $N_{b\text{-jets}}$ cross section, based on MC truth information, to match the definition of the fiducial volume. In Table 10.1, the requirements, which are imposed to events on reconstruction and particle level, are summarized.

10.2.1 Reconstruction level requirements

At reconstruction level, the nominal diphoton selection requirements are applied. A veto on events containing charged leptons is applied, as no leptons are expected in the H+b-jets final state. In addition, this veto slightly reduces the $t\bar{t}H$ contamination in the $N_{b\text{-jets}} = 1$ bin from approximately 24% to 18%. Although the $N_{b\text{-jets}}$ differential cross section is measured for the inclusive Higgs boson signal, it is preferable to reduce the contribution of processes that do not match the actual definition of H+b-jets.

Additionally, at least one central jet with $p_{\text{T}} > 30\,\text{GeV}$ is required to be present in each event, in order to measure the $N_{b\text{-jets}}$ distribution only for events that contain jets for which b-tagging

Table 10.1: Summary of the preselection requirements on recontruction level and the definition of the fiducial volume on particle level, as well as the respective definitions of b-jets.

	Reconstruction level	Particle level
Photons	Nominal diphoton selection	Two truth photons fulfilling truth isolation requirements & relative p_{T} cuts
Leptons	Veto on electrons & muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$	Veto on truth electrons & muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$
Jets	≥ 1 central jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$	≥ 1 central truth jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$
b-jets	b-tagged jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and the 70% b-tagging WP	Truth jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $\lvert y\rvert < 2.5$ ΔR-matched to a b-hadron with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ within $\Delta R = 0.4$

information is available. This requirement only influences the first bin of the $N_{b\text{-jets}}$ distribution and results in a less steeply falling distribution. The jet p_{T} requirement of 30 GeV is chosen for consistency with the other differential cross section measurements presented in Ref. [4], where this requirement was chosen due to the smaller contribution of pileup jets with respect to a jet p_{T} criterion of 25 GeV.

Finally, b-jets are defined as jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ that pass the 70% b-tagging WP. The considered b-tagging WP, which is characterized by a relatively low fraction of misidentified c-jets and ℓ-jets, provides a compromise between maximizing the expected significance to inclusive Higgs boson production and the sensitivity to $ggH{+}b$ in the $N_{b\text{-jets}} = 1$ bin. While a tighter b-tagging WP that provides an even higher rejection of c-jets and ℓ-jets slightly increases the sensitivity to $ggH{+}b$, it decreases the sensitivity to inclusive Higgs boson production. A comparison of the sensitivity in the $N_{b\text{-jets}} = 1$ bin, provided by the different b-tagging WPs, is shown in Appendix D.1.

10.2.2 Particle level requirements

The fiducial volume for the $N_{b\text{-jets}}$ measurement is defined on particle level, based on MC truth information. In general, the selection on reconstruction and particle level is required to be similar, in order to avoid an extrapolation between different phase spaces during the unfolding. Only stable particles with a lifetime of $\tau > 10\,\mathrm{ps}$ are considered for the definition of objects on particle level. The fiducial volume is defined based on the particle level diphoton selection and the additional requirements of at least one central truth jet and a veto on truth leptons. The same kinematic requirements are applied to objects defined on reconstruction and particle level. A more detailed description of the particle level definitions of the different objects is given in the following. These are the common definitions in fiducial and differential cross section measurements in the $H \to \gamma\gamma$ channel [4].

Similar to the reconstruction level event selection, events with two truth photons are selected. These photons are required to not originate from hadronization processes, which is ensured by checking the type of the parent particles in MC truth information. In addition, a truth isolation requirement needs to be fulfilled by the two photons, which is optimized to closely match the reconstruction level requirement. Therefore, the $p_{\mathrm{T}}^{\mathrm{iso,truth}}$ is calculated as the sum of p_{T} of all charged particles with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ in a cone with $\Delta R = 0.2$ around the photon direction, which is then required to be smaller than 5% of the photon p_{T}.

A veto is applied to events with charged leptons, based on truth electrons and muons which do not originate from hadronization processes. The lepton four-momenta are corrected by adding the four-momenta of photons within a cone of $\Delta R < 0.1$. Similar to reconstruction level, an overlap removal is performed by rejecting electrons which are in a cone with $\Delta R < 0.4$ around a photon.

Truth jets are defined by clustering all stable particles, except for muons and neutrinos, using the anti-k_t clustering algorithm with a radius parameter of 0.4. Muons and neutrinos are excluded from the possible constituents of jets on particle level, as they do not produce significant energy depositions in the EM and hadronic calorimeters. Only central jets with a rapidity of $|y| < 2.5$ are considered. Among these, truth b-jets are defined as truth jets which are ΔR-matched to a b-hadron with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ within a cone of $\Delta R = 0.4$. Truth jets are rejected if they are within a cone with $\Delta R < 0.4$ or 0.2 around a truth photon or a truth electron, respectively.

10.2.3 Expected event yields

The expected number of events in the different bins of the $N_{b\text{-jets}}$ distribution for a dataset corresponding to $79.8\,\mathrm{fb}^{-1}$ is shown in Table 10.2. The contributions of the different Higgs boson production processes are shown, as well as the expected number of continuum background events in the full 105 to 160 GeV $m_{\gamma\gamma}$ window, referred to as B_{ContBG}, which is extrapolated from a fit to the $m_{\gamma\gamma}$ sideband distribution of the respective bin. The total number of events from Higgs boson production is denoted by S_{Higgs}. The expected significance for inclusive Higgs boson production, considering only statistical uncertainties, is estimated using the approach detailed in Chapter 7. There is a large continuum background expectation in each bin, which in particular limits the sensitivity to final states with either one or at least two b-jets. This is the case because the expected number of signal events has a similar size to background fluctuations in these bins. The $N_{b\text{-jets}} = 0$ bin provides an expected significance of approximately $7.6\,\sigma$, while the $N_{b\text{-jets}} = 1$ and $N_{b\text{-jets}} \geq 2$ bins yield expected significances of approximately $1.8\,\sigma$ and $1.3\,\sigma$. The expected significance to ggH production alone amounts to approximately $1.0\,\sigma$ in

Table 10.2: Expected number of events on reconstruction level for the different Higgs boson production processes and the continuum background for $79.8\,\mathrm{fb}^{-1}$. The continuum background contribution in the full 105 to 160 GeV $m_{\gamma\gamma}$ window is extrapolated from a fit to the $m_{\gamma\gamma}$ sidebands. In addition, the expected sensitivity to inclusive Higgs boson production for each bin of the $N_{b\text{-jets}}$ distribution is shown.

Process	$N_{b\text{-jets}} = 0$	$N_{b\text{-jets}} = 1$	$N_{b\text{-jets}} \geq 2$
ggH	1000	34	1.6
VBF	160	5.9	0.22
WH	51	2.4	0.056
ZH	32	5.0	1.4
$t\bar{t}H$	5.2	11	7.0
$tHjb$	1.4	1.8	0.51
WtH	0.27	0.37	0.13
$b\bar{b}H$	6.0	3.1	0.17
S_{Higgs}	1270	63.4	11
B_{ContBG}	234000	9950	583
Exp. sig. $[\sigma]$	7.6	1.8	1.3

the $N_{b\text{-jets}} = 1$ bin. Hence, there is no sensitivity to $ggH{+}b$ alone with the considered dataset. Nevertheless, the dominant contribution in the $N_{b\text{-jets}} = 1$ bin results from ggH production, so that the compatibility of the measured and the predicted cross section in this bin provides a measure for the accuracy of the $H{+}b$-jets prediction from MC simulation.

As explained in Section 5.6, the WPs of the b-tagging algorithm are characterized by a certain efficiency for identifying b-jets and certain mis-tag rates for c-jets and ℓ-jets. Hence, the number of selected events also includes a certain fraction of events without real b-jets in the final state. Based on MC truth information it can be investigated which fraction of the selected events contains truth b-, c- and ℓ-jets under the considered b-tagging WP. Similar to b-jets, c-jets can be identified in MC simulation by performing a ΔR-matching[1] to c-hadrons based on MC truth information. Jets which are not identified as b-jets or c-jets are assumed to be ℓ-jets. Based on these definitions, the selected events are classified into

- $H{+}b$ events if $N_{b\text{-jets}} > 0$,
- $H{+}c$ events if $N_{b\text{-jets}} = 0$ and $N_{c\text{-jets}} > 0$,
- $H{+}\ell$ events if $N_{b\text{-jets}} = 0$ and $N_{c\text{-jets}} = 0$.

In Table 10.3 the fraction of events with b-, c- and ℓ-jets is shown for the three bins of the $N_{b\text{-jets}}$ distribution and the ggH, VBF and WH processes. In Appendix D.2, the equivalent table is shown for all Higgs boson production processes and for the different b-tagging WPs in the $N_{b\text{-jets}} = 1$ bin. In the $N_{b\text{-jets}} = 1$ bin, approximately 57% of the selected ggH events are classified as $H + b$, so that ggH events with b-jets provide the dominant contribution in this bin. For VBF and WH production, the fraction of $H + c$ events is dominant as in both processes c-jets may be produced by the decays of vector bosons. The fraction of misidentified b-jets is corrected for during the unfolding to particle level. As the estimated fraction of b-jets, c-jets and ℓ-jets might depend on the considered MC sample, a systematic uncertainty is assigned to the correction factor based on a variation of the parton shower algorithm.

Table 10.3: Classification of the selected ggH, VBF and WH events into $H{+}b$, $H{+}c$ and $H{+}\ell$ in the different bins of the $N_{b\text{-jets}}$ distribution.

	$N_{b\text{-jets}} = 0$			$N_{b\text{-jets}} = 1$			$N_{b\text{-jets}} \geq 2$		
Process	$H{+}\ell$	$H{+}c$	$H{+}b$	$H{+}\ell$	$H{+}c$	$H{+}b$	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH	92%	6.8%	1.4 %	18%	26%	57%	4.5%	4.0%	91%
VBF	83%	16%	1.0 %	15%	48%	37%	1.7%	8.5%	90%
WH	70%	29%	0.39 %	12%	76%	12%	1.8%	32%	66%

10.3 Signal & background modeling

The models for signal and background, which are used to define the likelihood function for the S+B fit to $m_{\gamma\gamma}$, are determined as described in Chapter 7. The parameters of the DCB function are determined from a fit to the $m_{\gamma\gamma}$ distribution of all Higgs boson production MC samples.

[1] A slightly different ΔR-matching criterion of $\Delta R < 0.3$ is used here, as this information is provided along with the b-tagging information.

In Table 10.4, the mean and the width of the Gaussian component of the DCB functions for the different $N_{b\text{-jets}}$ bins are shown. While the mean of the DCB function is similar for the three bins, the $m_{\gamma\gamma}$ distributions become narrower for the $N_{b\text{-jets}} = 1$ and ≥ 2 bins. The signal model for the $N_{b\text{-jets}} = 1$ and ≥ 2 bins, including the full set of DCB parameters, is shown in Figure 10.4. The DCB function provides a good description of the Higgs boson processes in the $m_{\gamma\gamma}$ region between 120 and 130 GeV, which contains the largest fraction of the Higgs boson signal. A disagreement between the MC prediction and the DCB function is observed in the tails of the DCB function, which is expected to have no significant impact on the measurement, due to the tiny fraction of signal events in these regions. The signal model for the remaining bins is shown in Appendix D.3.

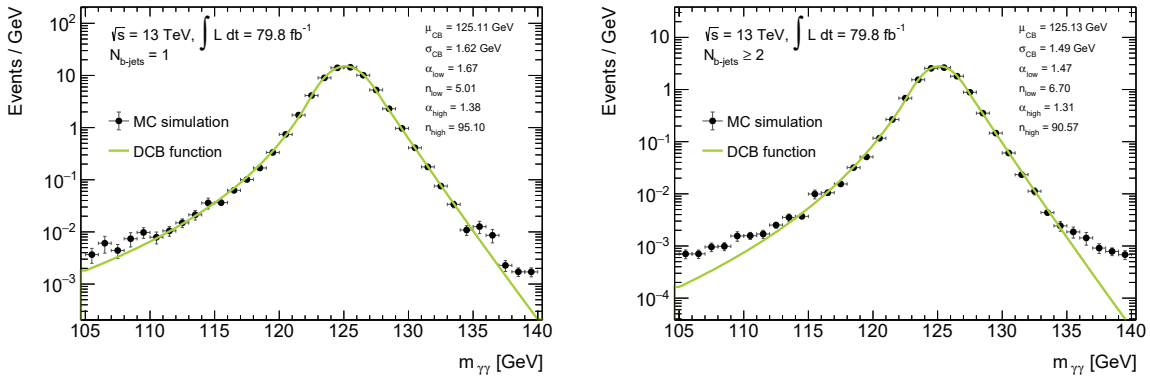


Figure 10.4: The DCB signal model for the $N_{b\text{-jets}} = 1$ and ≥ 2 bins as determined from a fit to the $m_{\gamma\gamma}$ distribution in MC simulations. In addition, the different fit parameters of the DCB function are shown.

The background model in each $N_{b\text{-jets}}$ bin is determined based on an S+B fit to a B-only template, that is derived from diphoton MC simulation and data control regions. As the background templates are characterized by relatively large statistical fluctuations, especially in the $N_{b\text{-jets}} = 1$ and $N_{b\text{-jets}} \geq 2$ bins, the relaxed spurious signal criteria are used to chose a background model. Functions with either one or two free parameters are chosen for the background model, as summarized in Table 10.4. The associated background modeling uncertainties, which are defined based on the fitted spurious signal, are comparably large, as shown in the following section.

Table 10.4: Summary of the signal and background models for the different bins of the $N_{b\text{-jets}}$ distribution. The mean and the width of the Gaussian part of the DCB function are shown, as well as the functional forms which are used to model the continuum background, including the number of free parameters N_{par}.

	Signal		**Background**	
Bin	μ_{CB} [GeV]	σ_{CB} [GeV]	Functional form	N_{par}
$N_{b\text{-jets}} = 0$	125.11	1.68	ExpPoly2	2
$N_{b\text{-jets}} = 1$	125.11	1.62	PowerLaw	1
$N_{b\text{-jets}} \geq 2$	125.13	1.49	PowerLaw	1

10.4 Signal extraction

The Higgs boson signal is extracted in each $N_{b\text{-jets}}$ bin by simultaneously performing an S+B fit to the respective $m_{\gamma\gamma}$ distributions, based on the signal and background models described in the previous section. This signal extraction fit for the $N_{b\text{-jets}}$ distribution was performed as part of this thesis. The statistical model is based on a likelihood function $\mathcal{L}$, describing the Poisson probabilities in each bin, which is used to define a profile likelihood ratio Λ as described in Section 7.4. The signal and background contributions of the $m_{\gamma\gamma}$ distribution observed in data and their uncertainties are then determined with the profile likelihood test statistic $-2\ln\Lambda$. Different systematic uncertainties are considered in the signal extraction fit, which are either treated as NPs in the fit or as an uncertainty on the fitted signal yield. The considered sources of systematic uncertainty are

- the photon energy scale (PES) and photon energy resolution (PER) uncertainties,
- the Higgs boson mass uncertainty and
- the background modeling uncertainty.

The systematic uncertainties related to PES and PER affect the shape of the signal model and, hence, NPs are defined for each of these uncertainties. While the PES variations generate a shift in $m_{\gamma\gamma}$ of the DCB function, the PER variations influence the width of the DCB function. For both, PES and PER, the uncertainties are split into several different components in order to avoid overconstraining these NPs due to an oversimplified treatment. In total, there are 64 NPs related to PES and 9 NPs related to PER. The respective uncertainties, which correspond to the $\pm\,1\,\sigma$ variations of each of these NPs, are determined by deriving the signal model for each variation. For the PES uncertainties, the difference between the mean of the DCB functions for the nominal case and the variation defines the uncertainty, while for PER uncertainties the difference between the widths is considered. The PES uncertainties are defined as Gaussian constraints in the likelihood function, while log-normal constraints are used for the PER uncertainties.

An additional NP is defined based on the assumed Higgs boson mass value, which is given by the combination of the measured values at the ATLAS and CMS experiments and amounts to $(125.09 \pm 0.24)\,\text{GeV}$. This measured value is used to shift the $m_{\gamma\gamma}$ distribution in MC simulations, which are nominally produced with $m_H = 125\,\text{GeV}$. Hence, the corresponding uncertainty of $\pm\,0.24\,\text{GeV}$ is considered in the fit.

Additional free parameters in the fit are those related to the background model. While the functional form of the background model is chosen based on diphoton MC samples, the final parameters of the background model are determined from the S+B fit to the $m_{\gamma\gamma}$ distribution. Depending on the $N_{b\text{-jets}}$ bin, there are one or two additional free parameters in the fit to account for the background model. Additionally, a systematic uncertainty related to the background modeling, which is defined based on the maximum amount of fitted spurious signal, is considered as an uncertainty on the fitted signal yield.

In Figure 10.5, the $m_{\gamma\gamma}$ distribution for each $N_{b\text{-jets}}$ bin and the fitted signal and background components, which are determined as described above, are shown. The lower panel in these plots shows the difference between the number of events observed in data and the fitted number of continuum background events. In the $N_{b\text{-jets}} = 0$ bin, the Higgs boson signal peak is well visible above the falling continuum background spectrum. In contrast, the $m_{\gamma\gamma}$ distributions in

the $N_{b\text{-jets}} = 1$ and $N_{b\text{-jets}} \geq 2$ bins are characterized by larger statistical fluctuations, leading to a less visible Higgs boson signal peak. In addition to the three $N_{b\text{-jets}}$ bins, the underflow bin, containing the remaining $H \to \gamma\gamma$ events that failed the preselection requirements, is fitted along with the other bins in order to help constraining the PES/PER NPs.

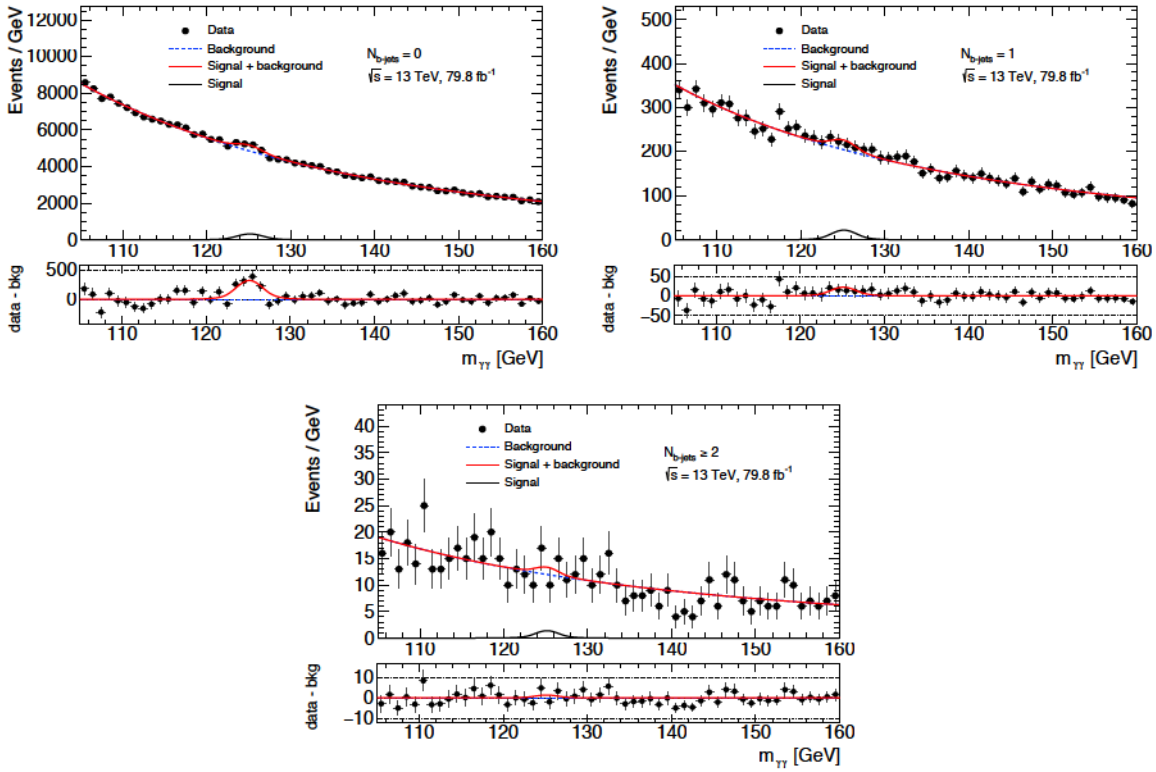
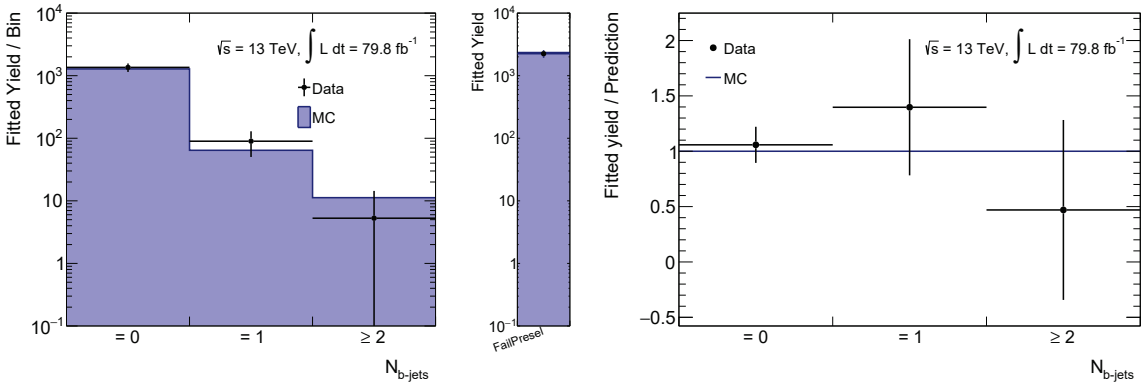


Figure 10.5: S+B fit to the $m_{\gamma\gamma}$ distribution of the different bins of the $N_{b\text{-jets}}$ distribution, which is performed to extract the Higgs boson signal yield in each bin. The $m_{\gamma\gamma}$ distributions observed in data, as well as the fitted signal and background contributions are shown. The lower panel shows the difference between the number of events observed in data and the fitted continuum background yield.

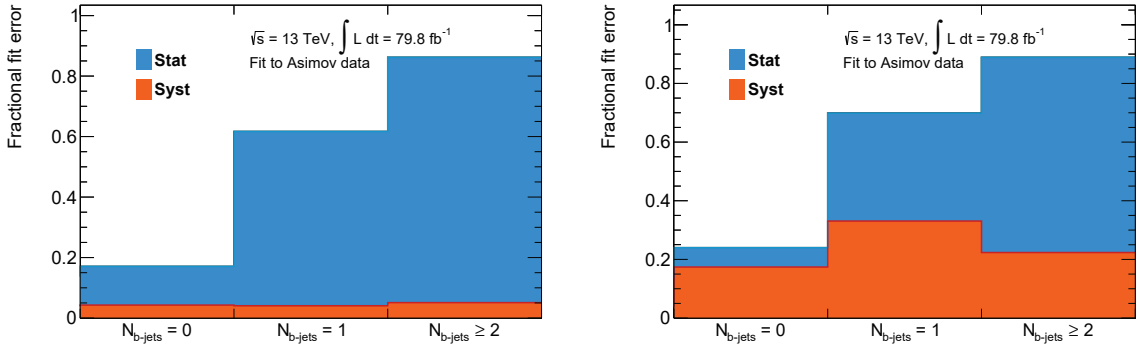
The number of fitted signal events in the different bins of the $N_{b\text{-jets}}$ distribution, including the fit uncertainties, is shown in Figure 10.6a), while the ratio of the fitted Higgs boson signal yields and the MC expectation is shown in Figure 10.6b). In the $N_{b\text{-jets}} = 1$ bin, the number of observed signal events is about 40% larger than the prediction, whereas a downward fluctuation of about 50% is observed in the $N_{b\text{-jets}} \geq 2$ bin. However, in each bin of the $N_{b\text{-jets}}$ distribution the number of fitted Higgs boson signal events is in good agreement with the prediction within the fit uncertainties, which include the statistical, the PES/PER and Higgs boson mass uncertainties.

A scan of the -2 ln Λ function is performed as a function of the number of signal events N^{sig}, which corresponds to the parameter of interest (PoI). While the minimum of the scan curve corresponds to the best-fit value, the width of the scan curve for -2 ln Λ = 1 corresponds to the 68% CL, which is used to define the uncertainty on N^{sig}. The expected contributions of the individual fit uncertainties to the total uncertainty on the extracted signal yield are studied in each $N_{b\text{-jets}}$ bin, based on a fit to an Asimov dataset. The total uncertainty, σ_{tot}, is determined based on the width of the scan curve, when leaving all NPs free floating in the fit. In contrast,



a) b)

Figure 10.6: a) Fitted number of Higgs boson signal events in data and the number of signal events predicted by MC simulations in the different bins of the $N_{b\text{-jets}}$ distribution and the underflow bin, b) ratio of the fitted number of Higgs boson signal events and the MC expectation in the three $N_{b\text{-jets}}$ bins. The error bars include the fit-related uncertainties, i.e. the statistical, the PES/PER and the Higgs boson mass uncertainties.



a) b)

Figure 10.7: Breakdown of the expected fractional uncertainties on the extracted signal yield, evaluated using an Asimov dataset. The systematic uncertainty is shown in orange, while the statistical uncertainty is shown in blue. In a) the systematic uncertainties only include the PES/PER and Higgs boson mass uncertainties, while in b) the background modeling uncertainty is included as well.

the statistical uncertainty, σ_{stat}, is defined based on the width of the scan curve when fixing the NPs related to PES/PER and the Higgs boson mass uncertainty to their best-fit values. The systematic component of the total uncertainty is then estimated based on the equation $\sigma_{\text{tot}} = \sqrt{\sigma^2_{\text{stat}} + \sigma^2_{\text{syst}}}$.

In Figure 10.7, the expected fractional fit uncertainties in each bin of the $N_{b\text{-jets}}$ distribution are shown. While Figure 10.7a) shows the statistical uncertainty and the systematic uncertainty based on the NPs in the S+B fit, in Figure 10.7b) the systematic uncertainty does also include the background modeling uncertainty. In each bin, the statistical uncertainty is large compared to the PES/PER uncertainties. The background modeling uncertainty provides the dominant contribution to the systematic uncertainty on the fitted signal. While the statistical uncertainty

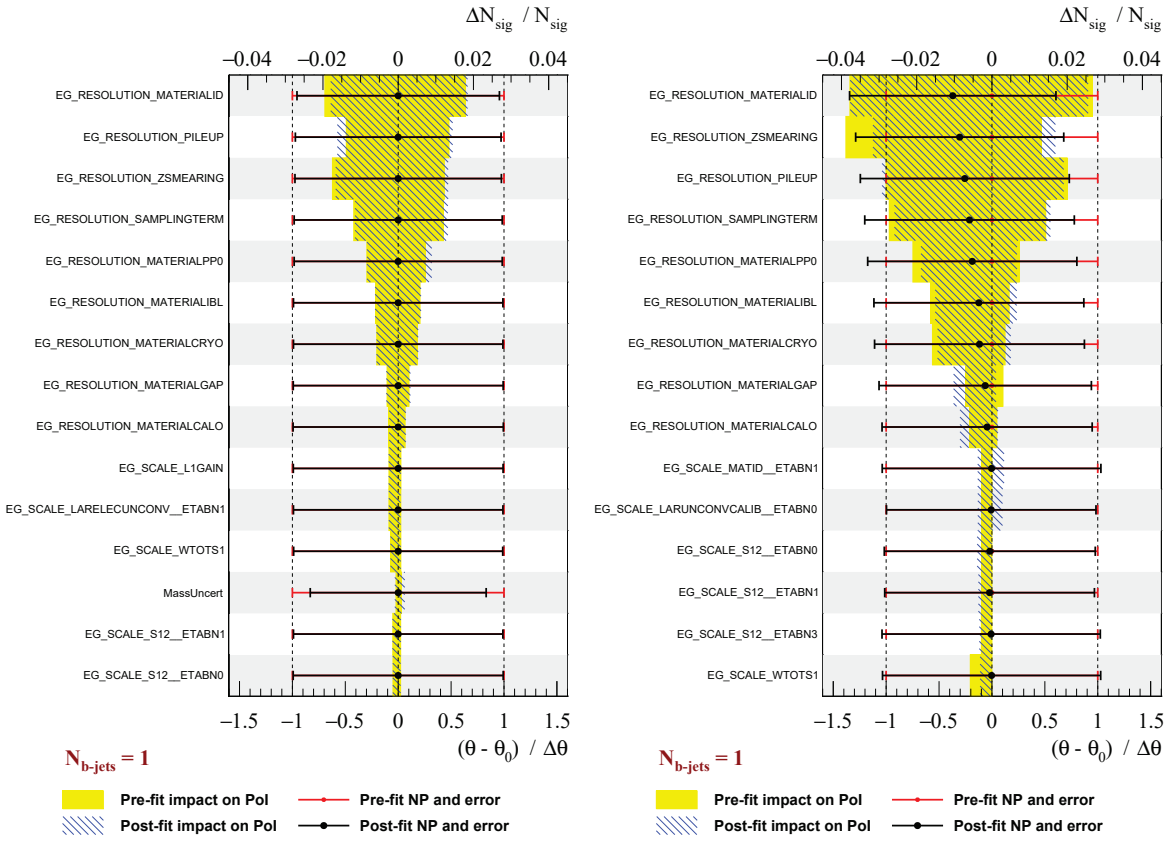


Figure 10.8: Ranking plots obtained from a) a fit to an Asimov dataset or b) the fit to data, showing the 15 most important NPs for the $N_{b\text{-jets}} = 1$ bin. The number of signal events, N^{sig}, corresponds to the PoI. The pre-fit and post-fit values of the nuisance parameters θ and the associated uncertainties are shown as red and black points and lines, respectively. The yellow and blue hatched bands show the pre-fit and post-fit uncertainties on N^{sig}, respectively.

is expected to be dominant in the $N_{b\text{-jets}} = 1$ and $N_{b\text{-jets}} \geq 2$ bins, in the $N_{b\text{-jets}} = 0$ bin the systematic uncertainty becomes dominant when including the background modeling uncertainty. Similar plots of the signal extraction uncertainty are shown in Appendix D.4 for the actual fit to data. The magnitude of the individual uncertainties obtained from the fit to data is very similar to the expectation, except for the $N_{b\text{-jets}} \geq 2$ bin. As the observed number of events in this bin is smaller than expected, a larger statistical uncertainty is observed.

In Figure 10.8 a ranking of the NPs for the $N_{b\text{-jets}} = 1$ bin is shown, which is obtained either from a fit to an Asimov dataset or the fit to data. The PES and PER uncertainties are denoted by EG_RESOLUTION and EG_SCALE, respectively. Due to the large number of NPs, these plots only show the 15 NPs with the largest impact on the PoI. The red and black points show the expected and fitted values of the NPs, which are both 0 for a fit to an Asimov dataset, and the associated uncertainties, respectively. In data, the fitted NPs related to the PER uncertainties are slightly pulled toward negative values, resulting in a signal peak which is narrower than predicted by MC simulation. The yellow and blue hatched bands show the pre-fit and post-fit uncertainties of the NPs, indicating their expected and observed impact. The pre-fit impact of a given NP is determined by fixing the respective NP to its $\pm 1\,\sigma$ values, while leaving

the other parameters free in the fit, whereas the post-fit impact is determined by fixing the other NPs to their best-fit values. The most important NPs are related to the PER uncertainties, such as EG_RESOLUTION_MATERIALID, which is related to the modeling of the material in the InDet. However, as discussed in the previous paragraph, the PES/PER uncertainties only provide a comparably small contribution to the total signal extraction uncertainty of the different $N_{b\text{-jets}}$ bins, due to the much larger background modeling and statistical uncertainties. In general, the same trends are observed in the ranking plots obtained from the fit to an Asimov dataset and the fit to data. In Appendix D.4, similar ranking plots are shown for the remaining $N_{b\text{-jets}}$ bins.

10.5 Unfolding

The distribution of the number of b-jets, for which the amount of Higgs boson signal is extracted as described in the previous section, is then unfolded to the particle level cross section. The unfolding corrects the measured cross section for effects like the detector response and detector inefficiencies, which cause migrations between the different bins of the distribution. This is done by using the bin-by-bin unfolding method, where the extracted signal yield in each bin is divided by a correction factor C. The correction factor in bin i is determined from MC simulation based on

$$C_i = \frac{n_i^{reco}}{n_i^{ptcl}}, \tag{10.1}$$

where n_i^{reco} describes the number of events in MC simulation that fulfill the reconstruction level requirements and n_i^{ptcl} describes the respective number of events fulfilling the particle level requirements in bin i. Then, the particle level cross section in each bin is calculated as

$$\frac{d\sigma_i}{dN_{b\text{-jets}}} = \frac{N_i^{\text{sig}}}{C_i \cdot \int \mathcal{L}\,\mathrm{dt}}, \tag{10.2}$$

where $\int \mathcal{L}\,\mathrm{dt}$ describes the integrated luminosity of the considered dataset and N_i^{sig} the number of extracted signal events in bin i. The unfolding, based on the bin-by-bin method, introduces a bias as it strongly depends on the modeling of the particle level distribution in MC simulation. Therefore, different bias scenarios are investigated, which are considered as theoretical uncertainties on the correction factors. However, the bin-by-bin method is assumed to be a valid approach in measurements which are statistically limited, as in those cases the bias is small compared to the statistical uncertainty of the measurement. In the measurement of the differential $N_{b\text{-jets}}$ cross section, migrations between the different bins are primarily induced by the use of b-tagging. On the one hand, the limited b-tagging efficiency leads to b-jets which are not selected on reconstruction level, leading to correction factors < 1. On the other hand, c-jets and ℓ-jets jets may falsely be identified as b-jets on reconstruction level by the considered b-tagging algorithm. In Figure 10.9, the migration matrix of the $N_{b\text{-jets}}$ distribution is shown, where the number of b-jets on reconstruction level is plotted versus the number of b-jets on particle level. The migration matrix describes the probability for an event which is classified in bin i on reconstruction level to be classified in bin j on particle level. The matrix is asymmetric due to the two effects of unidentified b-jets and mis-tagged c-jets and ℓ-jets.

The correction factor also accounts for Dalitz events, $H \to \gamma\gamma^* \to \gamma f \bar{f}$, which are generated by Pythia8 along with the $H \to \gamma\gamma$ decay. These events are rejected at particle level due to the absence of a stable diphoton final state. However, at reconstruction level approximately 0.4% of the selected events correspond to Dalitz events. Hence, this fraction is also accounted for during the unfolding.

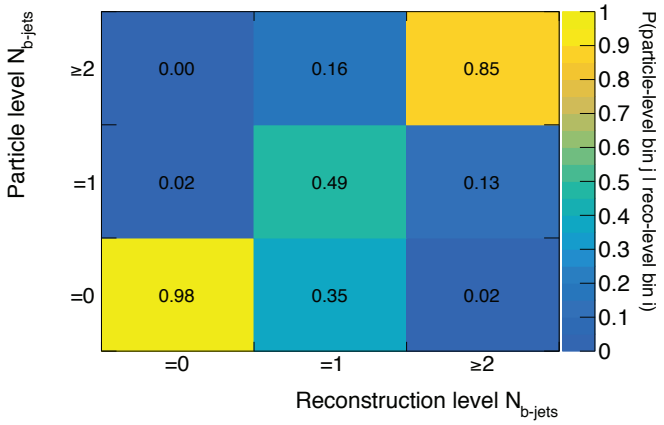


Figure 10.9: Unfolding matrix of the $N_{b\text{-jets}}$ distribution, where the number of b-jets selected on reconstruction level is plotted versus the number of b-jets on particle level. The off-diagonal elements are non-zero due to the limited efficiency of the considered b-tagging algorithm and misidentified c-jets and ℓ-jets.

10.6 Systematic uncertainties

Different sources of uncertainty are considered for the measurement of the differential $N_{b\text{-jets}}$ cross section. In general, these can be summarized by the following components:

1. Statistical uncertainty on the extracted signal yield
2. Signal extraction: PES/PER, background model, Higgs boson mass uncertainties
3. Luminosity uncertainty
4. Correction factor: Physics modeling and experimental uncertainties

The uncertainties related to the signal extraction fit were described in detail in Section 10.4. These include the statistical uncertainty on the extracted signal yield and systematic uncertainties related to the PES/PER, the background model and the assumed Higgs boson mass. In addition, a constant uncertainty of 2.0% is imposed on the integrated luminosity of the considered 2015-2017 dataset [150]. Finally, there are systematic uncertainties related to the unfolding, which include theoretical as well as experimental uncertainties, as described in the following. The estimation of the parton shower related systematic uncertainty was performed as part of this thesis.

10.6.1 Physics modeling uncertainties

Signal composition

A model dependence of the differential cross section measurement is introduced if the correction factors depend on the Higgs boson production processes. A systematic uncertainty is assigned to the correction factor to account for this potential bias. This uncertainty is derived by varying the composition of the Higgs boson signal in each bin, based on the experimental boundaries of the different Higgs boson production processes as measured by the ATLAS and CMS combination of Higgs boson production cross sections [151].

Parton shower modeling

The additional b-jets in ggH production are mostly simulated during the parton showering. Hence, their modeling depends on the choice of the parton shower algorithm. The nominal correction factors are determined with the default parton shower algorithm Pythia8. A systematic uncertainty is assigned to the modeling of the parton shower, by rederiving the correction factors with a ggH MC sample produced with the alternative parton shower generator Herwig7. The relative difference between the correction factors obtained with the nominal and the alternative MC sample is considered as systematic uncertainty on the correction factor in each $N_{b\text{-jets}}$ bin.

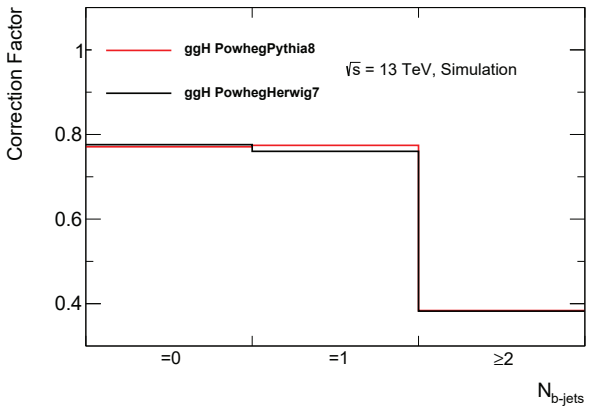


Figure 10.10: Correction factors for the different $N_{b\text{-jets}}$ bins determined using ggH MC samples generated with the Pythia8 and Herwig7 parton shower algorithms.

In Figure 10.10, the inclusive correction factors for the nominal and the alternative case are shown for the three $N_{b\text{-jets}}$ bins. The uncertainties in the $N_{b\text{-jets}} = 0$, 1 and ≥ 2 bins amount to 0.65%, 1.83% and 0.37%, respectively. The uncertainty increases from the $N_{b\text{-jets}} = 0$ to the $N_{b\text{-jets}} = 1$ bin due to the stricter requirement on the number of jets in the final state. On the other hand, it decreases again for $N_{b\text{-jets}} \geq 2$, as $t\bar{t}H$ production is dominant in this bin. As the b-jets in $t\bar{t}H$ production originate from the top quark decay and not from the parton shower, this behavior is expected. As an additional cross check, the b-jet, c-jet and ℓ-jet contributions to the different bins of the $N_{b\text{-jets}}$ distribution are compared for the ggH MC samples produced with the Pythia8 and the Herwig7 parton shower algorithms. The respective numbers are shown in Table 10.5. The difference between the $H + b$ contribution in the $N_{b\text{-jets}} = 1$ bin amounts to 1.8%, which supports the accuracy of the estimated uncertainty on the correction factor.

Table 10.5: Comparison of the $H{+}b$, $H{+}c$ and $H{+}\ell$ contributions to ggH events generated with POWHEG in combination with the PYTHIA8 and the HERWIG7 parton shower algorithm.

	$N_{b\text{-jets}} = 0$			$N_{b\text{-jets}} = 1$			$N_{b\text{-jets}} \geq 2$		
Parton shower	$H{+}\ell$	$H{+}c$	$H{+}b$	$H{+}\ell$	$H{+}c$	$H{+}b$	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH PYTHIA8	91.9%	6.7%	1.4%	18.2%	25.8%	56.1%	4.9%	4.4%	90.7%
ggH HERWIG7	93.2%	5.5%	1.4%	19.4%	22.6%	57.9%	4.0%	4.4%	91.6%

Higgs boson p_{T} and rapidity mismodelling

An additional uncertainty is defined to account for a possible mismodeling of the Higgs boson p_{T} and rapidity distributions in MC simulations. To estimate this uncertainty, the distributions in MC simulations are reweighted to match the distributions observed in data and the correction factors are rederived based on these MC samples. The difference to the nominal correction factor is then defined as a systematic uncertainty. As the $N_{b\text{-jets}}$ distribution does not strongly depend on the modeling of the Higgs boson p_{T} and rapidity, this uncertainty only has a minor impact.

Dalitz decays

An uncertainty is assigned to the fraction of Dalitz events in the considered MC samples. As the branching ratio of Dalitz decays is barely known, a conservative 100% uncertainty is assigned.

10.6.2 Experimental uncertainties

Object related uncertainties

As described in Chapter 5, the efficiencies of several object definition and selection criteria in MC simulation are corrected to match the values measured in data. The corresponding data-driven SFs are applied to the MC simulations and their associated uncertainties are considered in the $N_{b\text{-jets}}$ measurement. In addition, uncertainties related to the calibration of photons and jets are considered. Overall, systematic uncertainties related to photon ID and isolation, PES/PER, the jet energy scale and resolution (JES/JER), the JVT requirement and flavor tagging are accounted for. These systematic uncertainties are estimated by rederiving the correction factors with MC samples containing the up and down variations of the SFs and calibration related quantities.

Pileup reweighting

An uncertainty is assigned to the pileup reweighting which is applied to MC simulation. This uncertainty is estimated by varying the average number of pileup events in MC simulation based on the pileup distribution observed in data.

Modeling of pileup jets in simulation

An additional uncertainty is assigned to the modeling of pileup jets in MC simulation. The fraction of pileup jets, which do not originate from the hard scattering process, is estimated using jets that cannot be ΔR-matched to a truth jet with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ within a cone with $\Delta R < 0.4$. The influence of the pileup modeling on the correction factor is estimated by randomly removing 20% of the pileup jets in MC simulation.

Diphoton vertex selection

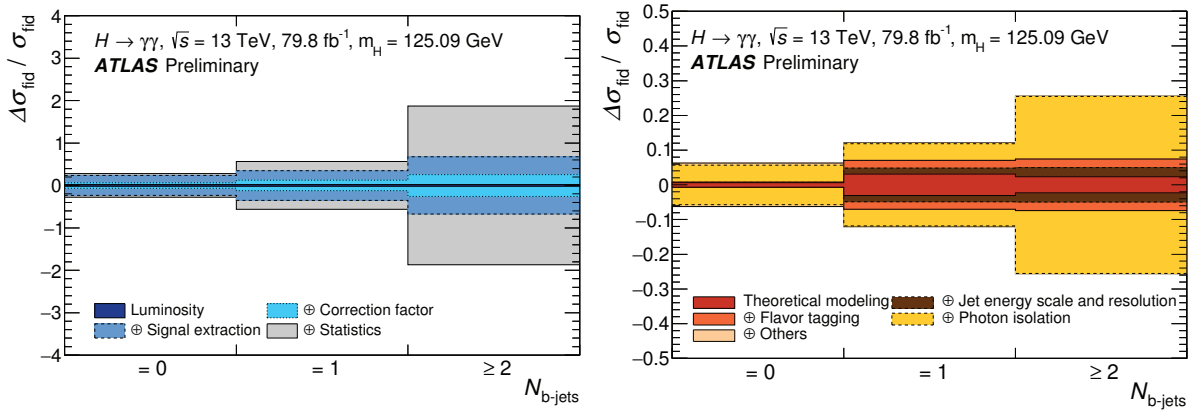
Furthermore, an uncertainty is assigned to the diphoton vertex selection. For each MC event, the z value of the truth vertex is compared to the respective value of the selected vertex at reconstruction level. A matching is applied based on the requirement $\Delta z < 3\,\text{mm}$ and the weights of events, which do not fulfill this criterion, is increased by approximately 10%. These increased weights account for the difference of the vertex selection efficiency in data and MC simulation. The uncertainty is then derived by reevaluating the correction factors with the varied event weights.

Trigger efficiency

An uncertainty is assigned to the efficiency of the diphoton trigger, based on the uncertainty of the efficiency measured in data using a bootstrap method [58], and amounts to approximately 0.7%.

10.6.3 Summary

An overview of the impact of the different kinds of uncertainties on the measured $N_{b\text{-jets}}$ cross section is given in Figure 10.11a). The statistical uncertainty is the largest uncertainty in each bin of the $N_{b\text{-jets}}$ distribution, except for the $N_{b\text{-jets}} = 0$ bin, where the systematic uncertainties have a similar size. The uncertainties related to the signal extraction are the second largest contribution in each bin. This uncertainty is dominated by the background modeling uncertainty, while the PES/PER uncertainties provide a relatively small contribution. The uncertainty on the correction factor, whose composition is shown in Figure 10.11b), is comparably small in each bin of the $N_{b\text{-jets}}$ distribution. Among the different contributions, the experimental uncertainties due to photon isolation, flavor tagging and JES/JER are dominant, compared to the minor contributions of the theoretical uncertainties.



a) b)

Figure 10.11: Breakdown of the different kinds of uncertainties on the measured differential cross section of the $N_{b\text{-jets}}$ distribution [4]. In a) the total uncertainty breakdown is shown, including the signal extraction, correction factor, luminosity and statistical uncertainties, while b) illustrates the composition of the correction factor uncertainty.

10.7 Results and Discussion

The unfolded $N_{b\text{-jets}}$ distribution, including the total systematic and statistical uncertainties, is shown in Figure 10.12. The nominal MC expectations for ggH, $t\bar{t}H$ and the remaining Higgs boson production processes, which are denoted by XH, are compared to the cross section measured in data. In addition, the ratio of the cross section measured in data and predicted by MC simulation is shown in the bottom panel. The measured cross section in the $N_{b\text{-jets}} = 1$ bin is about 40% higher than the prediction. However, within the uncertainties the measured values are generally in good agreement with the SM prediction in each $N_{b\text{-jets}}$ bin. The agreement is quantified in terms of a χ^2 probability, which amounts to 84% [4].

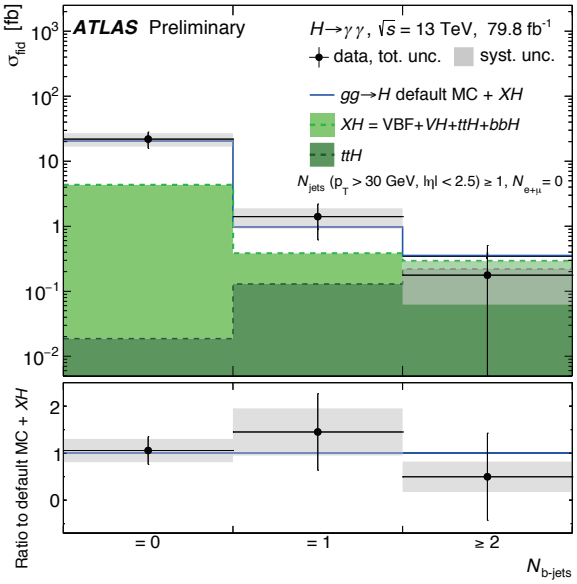


Figure 10.12: Measured differential cross section of the $N_{b\text{-jets}}$ distribution using 79.8 fb^{-1} of ATLAS data [4]. Additional requirements are applied on N_{lep} and $N^{\text{central}}_{\text{jets30}}$. The MC predictions, estimated from the nominal MC samples, are shown for ggH, $t\bar{t}H$ and the remaining Higgs boson production processes, XH.

Especially in the $N_{b\text{-jets}} = 1$ and $N_{b\text{-jets}} \geq 2$ bin, the statistical uncertainty is the dominant uncertainty in this measurement. The total uncertainty in the $N_{b\text{-jets}} = 1$ bin, which is designed to provide sensitivity to H+b-jets, amounts to approximately 56%. The 100% uncertainty, which is currently assigned to H+b-jets, approximately corresponds to the largest deviation between the prediction and the measured cross section for $N_{b\text{-jets}} = 1$, including the total uncertainty. Compared to the measurement of the $N_{b\text{-jets}}$ cross section in the $H \to ZZ^* \to 4\ell$ channel, a better agreement between the measured and predicted cross section is observed and the uncertainties are decreased due to the larger considered dataset. Additionally, the requirement on N_{lep} and the choice of binning slightly reduce the $t\bar{t}H$ contribution in the $N_{b\text{-jets}} = 1$ bin.

Due to the small cross section of ggH+b and the comparably large non-resonant background, there is no sensitivity to this single process, yet. Therefore, the measurement of the differential $N_{b\text{-jets}}$ cross section is considered to be a first step towards a measurement of ggH+b. In the next chapter, an attempt to separate H+b-jets production from the continuum background, based on MVA techniques, is presented.

Chapter 11

Alternative strategy for a measurement of H+b-jets using MVA techniques

In this chapter, a strategy to obtain a more precise measurement of H+b-jets, with respect to the $N_{b\text{-jets}}$ measurement presented in the previous chapter, is presented. In order to define a signal region with a reduced $t\bar{t}H$ contribution, an additional event selection criterion on the maximum number of jets is investigated. Additionally, as the sensitivity of the $N_{b\text{-jets}}$ measurement is limited by the large contribution of continuum background events, an attempt to separate H+b-jets from the continuum background using MVA techniques is presented. In the following, first the strategy for improving the sensitivity of the H+b-jets measurement is discussed. Then, the alternative event selection and the application of MVA techniques in the H+b-jets measurement are described in detail. This is followed by a discussion of the impact of this approach on the background modeling and the systematic uncertainties.

11.1 Strategy for improvements

In contrast to the differential cross section measurement of $N_{b\text{-jets}}$, in this chapter the measurement of a fiducial cross section in a region with $N_{b\text{-jets}} \geq 1$ is studied, as this corresponds to the final state of interest for H+b-jets. Although the cross section measured in the $N_{b\text{-jets}} = 1$ bin is characterized by a significant contribution from ggH+b production, it contains a non-negligible fraction of $t\bar{t}H$ events of approximately 18%. The ggH final state can be distinguished from $t\bar{t}H$ events, which are characterized by considerably higher jet multiplicities, by introducing a requirement on the maximum number of jets. Therefore, the event selection is modified accordingly, in order to obtain a signal region which is more pure in ggH.

The sensitivity of the $N_{b\text{-jets}}$ measurement is limited by the amount of continuum background, which is large compared to the Higgs boson signal for the considered final states. An enhanced separation between Higgs boson signal and continuum background events would help to increase the sensitivity to H+b-jets and, simultaneously, decrease the statistical uncertainty of the measurement. After the preselection, the Higgs boson and continuum background events have very similar properties and, hence, no significant improvement of the sensitivity to H+b-jets can be obtained through additional requirements on single kinematic variables. Variables describing the kinematics of the individual final state particles and more complex variables, which e.g. characterize the structure of the whole event, are studied. Several of these variables provide at least a small discrimination between H+b-jets and the continuum background. The application of MVA techniques allows to combine the information from these variables into one discriminant which provides an improved separation between signal and background by

exploiting the correlations between the discriminating variables. In the following, a BDT is trained to separate H+b-jets events from the continuum background. Events which pass the preselection on reconstruction level enter the BDT and a weight between -1 and 1 is assigned to each of them. A BDT weight close to -1 indicates a more background-like final state, while a BDT weight close to 1 is characteristic for signal-like events. Based on the BDT discriminant, the events are divided into three categories with a different Higgs boson signal to continuum background ratio S/B_{ContBG}. Then, an S+B fit to the $m_{\gamma\gamma}$ distributions of the respective categories is performed, in order to extract the Higgs boson signal.

In general, the fiducial volume should be defined based on simple kinematic requirements, which can be defined similarly on reconstruction and particle level. For this reason, it is not possible to reject continuum background events by imposing requirements on the BDT discriminant, as the measured cross section would correspond to a phase space which cannot be easily defined on particle level. In order to keep the definition of the fiducial volume independent of the BDT categorization, the number of fitted signal events and the associated uncertainties of the three BDT categories need to be combined. The measured signal yield in the combined BDT categories, however, can then be unfolded to particle level and the unfolding related uncertainties remain independent of the BDT categorization. An overview of the different steps of the alternative analysis strategy is given in Figure 11.1.

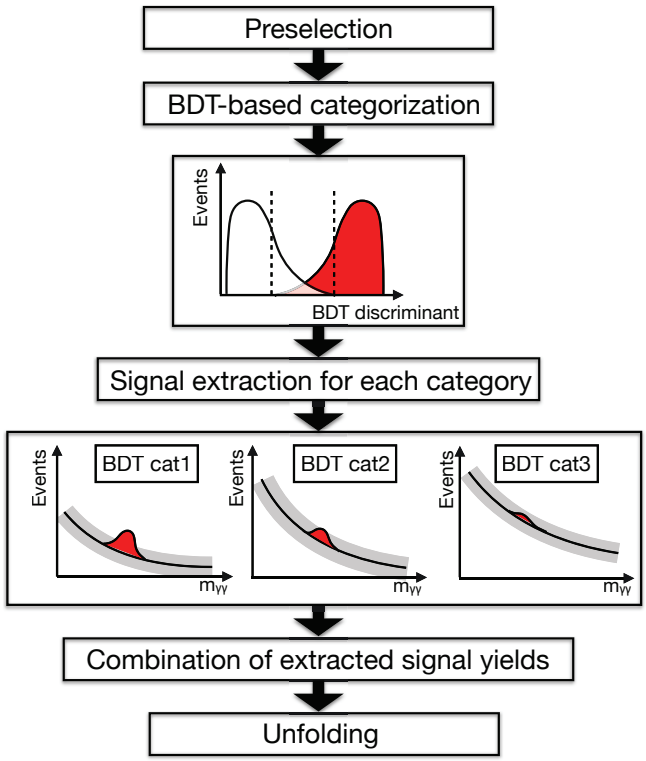


Figure 11.1: A sketch illustrating the different steps of the alternative strategy for a measurement of H+b-jets. After the preselection, events enter a BDT which discriminates between continuum background and H+b-jets events. Based on the BDT discriminant, three categories with a varying S/B_{ContBG} are defined. The signal is first extracted in each of these categories through a fit to $m_{\gamma\gamma}$ and is then combined into the total number of signal events in the fiducial volume. Finally, the measured cross section is unfolded to particle level.

11.2 Event selection

The event selection criteria on reconstruction and particle level remain similar to the requirements in the $N_{b\text{-jets}}$ measurement, described in the previous chapter. On reconstruction level, a veto on charged leptons is applied and at least one b-tagged jet, considering the 70% b-tagging WP, is required. In each event, the jet with the highest MV2c10 weight (the "MV2-leading jet"), which is the jet with the highest probability of originating from a b quark, is required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. The properties of this jet are used in the following to define variables which are used for the training of the BDT. In general, only jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.5$ are considered. An additional requirement on the maximum number of jets is introduced, which is dedicated to reducing the fraction of $t\bar{t}H$ events in the signal region. In Table 11.1, the efficiencies of different requirements on $N_{\mathrm{jets30}}^{\mathrm{central}}$ are shown for the different Higgs boson production processes. A compromise has to be made, in order to keep the selection efficiency of ggH high, while reducing the $t\bar{t}H$ contribution. A requirement of $N_{\mathrm{jets30}}^{\mathrm{central}} \leq 3$ is chosen, as the selection efficiency of ggH remains close to 100%, while the efficiency for selecting $t\bar{t}H$ events is reduced to 25%. This results in a $t\bar{t}H$ contribution of approximately 6% after the preselection.

The fiducial volume is defined similarly to the reconstruction level event selection, requiring a veto on truth leptons, a maximum number of three central truth jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and at least one central truth jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, which is ΔR-matched to a b-hadron.

Table 11.1: Efficiencies of different requirements on the maximum number of central jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $N_{\mathrm{jets30}}^{\mathrm{central}}$, on reconstruction level for the different Higgs boson production processes and events without leptons and at least one central jet in the final state.

	Efficiency for events with $N_{\mathrm{lep}} = 0$, $N_{\mathrm{jets30}}^{\mathrm{central}} \geq 1$ and			
Process	$N_{\mathrm{jets30}}^{\mathrm{central}} = 1$	$N_{\mathrm{jets30}}^{\mathrm{central}} \leq 2$	$N_{\mathrm{jets30}}^{\mathrm{central}} \leq 3$	$N_{\mathrm{jets30}}^{\mathrm{central}} \leq 4$
ggH	0.73	0.94	0.99	1.00
VBF	0.69	0.94	0.99	1.00
WH	0.44	0.81	0.95	0.99
ZH	0.39	0.77	0.93	0.98
$t\bar{t}H$	0.02	0.08	0.25	0.50
$tHjb$	0.13	0.40	0.71	0.91
WtH	0.03	0.13	0.32	0.57
$b\bar{b}H$	0.73	0.93	0.98	0.99

As the modeling of jets in ggH MC simulation strongly relies on the choice of the parton shower algorithm, it is investigated if the requirement on the maximum number of jets introduces a model dependence during the unfolding. Therefore, the distributions of the number of central truth jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ for the ggH MC samples showered with Pythia8 and Herwig7 are compared for events without truth leptons, as shown in Figure 11.2. Relatively large differences are observed in the individual bins of the distribution for small truth jet multiplicities. However, the probability of finding ggH events with more than three central jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ is very small, resulting in an efficiency of the $N_{\mathrm{jets30}}^{\mathrm{central}} \leq 3$ requirement close to 100% for both MC samples. The relative difference between the number of selected events after applying the $N_{\mathrm{jets30}}^{\mathrm{central}} \leq 3$ requirement amounts to approximately 0.8%. Hence, only a minor model dependence is introduced by this additional requirement, which is accounted for by the parton shower modeling uncertainty on the correction factor.

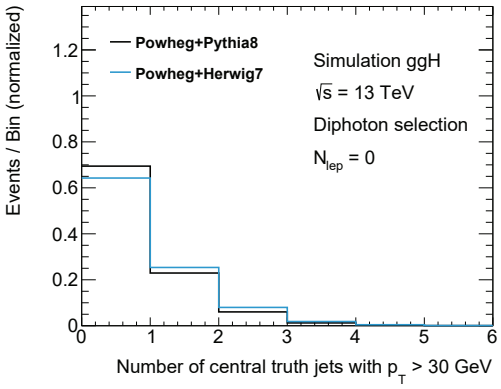


Figure 11.2: Comparison of the number of central truth jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ for events without truth leptons, obtained with the ggH Powheg+Pythia8 and Powheg+Herwig7 MC samples.

11.3 Separation between H+b-jets and the continuum background

The application of MVA techniques has become a common approach in high energy physics. These techniques allow to efficiently separate rare signal processes from background processes, thereby increasing the sensitivity of an analysis. Popular choices for MVAs are artificial neural networks, which can be built of several layers of connected *neurons*, and Boosted Decision Trees (BDTs), which usually consist of a large set of binary decision trees. The MVA algorithms are optimized for a specific application in a process which is referred to as training. The training is based on discriminating variables, which are provided to the MVA algorithm via signal and background MC simulations. This has the advantage that the true classification into signal and background of the events is known from the training sample. The MVA assigns a weight between 1 and -1 to each event, indicating if the event is signal-like or background-like, and the distribution of weights is referred to as the MVA discriminant. In the following, a BDT-based approach to separate the H+b-jets signal from the continuum background is presented. First, the basic principles of BDTs are briefly described, followed by a more detailed description of the BDT training, the discriminating variables and the final discriminant, which is used to classify the preselected events into different categories.

11.3.1 Boosted Decision Trees

A decision tree is a classifier which discriminates between signal and background based on binary decisions. It has a tree structure, as illustrated in Figure 11.3, with *nodes* that represent different criteria on the provided input variables. At each node, a *true* or *false* decision is taken, depending on whether the criterion is fulfilled or not for the considered event. The criteria are optimized during the training of the decision tree. Based on the decisions taken at each node, an event passes through the tree until it reaches the final node, where is it classified either as signal or background. Single decision trees are inefficient and unstable with respect to statistical fluctuations in the training sample, whereas BDTs take advantage of a large number of decision trees that are iteratively trained on the same sample and whose output is combined. For this thesis, the BDT implementation provided by the TMVA package [152] is used.

For the training, samples for signal and background containing several discriminating variables are provided to the BDT. The decision trees are then trained in an iterative procedure by successively optimizing the criteria that split the sample into signal-like and background-like events for each tree. The training starts with the optimization of the criterion at the first node of a tree, where it chooses the variable that provides the best discrimination between signal and background. The criterion on this discriminating variable is optimized to provide the best possible discrimination. For the subsequent nodes, this step is repeated on the remaining subsamples. The training of the decision tree stops if the number of remaining training events is too small or the predefined maximum depth of the tree has been reached. An event entering the trained BDT is finally classified as signal or background, depending on whether the majority of training events in the final node was classified as signal or background. The separation power of the different variables is assessed by the *Gini Index*, which is defined as $G = p \cdot (1 - p)$ with the purity p. The Gini Index gets small for a good separation between signal and background, indicated by $p = 0$ or $p = 1$, and has a maximum if signal and background events are randomly distributed, which corresponds to a purity of $p = 0.5$. The Gini index is symmetric with respect to the classification into signal and background events, so that criteria with a high signal efficiency are treated similar to criteria with a high background efficiency.

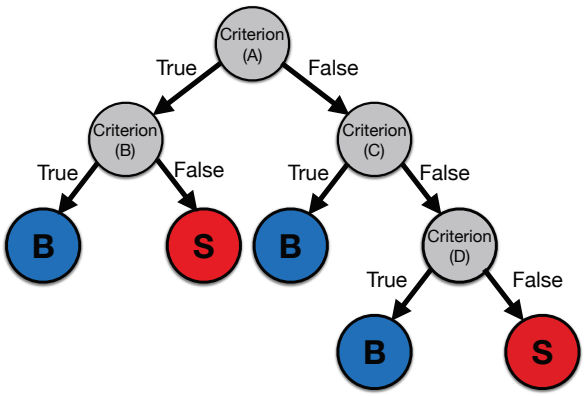


Figure 11.3: Sketch of the structure of a decision tree, which consists of different nodes that represent certain criteria on the discriminating variables. Based on these criteria, binary decisions are taken at each node in order to classify events reaching the final nodes as signal (S) or background (B).

The performance and stability of decision trees can be enhanced by the use of *boosting* and *bagging* techniques. Bagging refers to a technique where subsets of the original training sample are created randomly. The training of the classifier is then repeated on each of these subsamples and a combined classifier is built. The use of bagging during the training helps to stabilize the classification algorithm, as it corresponds to a smearing of statistical fluctuations in the training sample.

Boosting techniques are used to increase the performance of weak classifiers, by combining a large number of weak classifiers into one strong classifier. Two common choices of boosting algorithms are the adaptive boost (AdaBoost) [153] and the Gradient boosting [154]. The AdaBoost algorithm improves the performance of decision trees through modifications of the event weights during the training. If an event is misclassified by a single classifier, it is given a higher event weight. The following tree is then trained using the training sample with the modified event weights, thereby improving the classification of previously misclassified events.

The different trees are then combined into a single classifier through the weighted sum of the individual classifiers h_i:

$$y_{AdaBoost}(\boldsymbol{x}) = \frac{1}{N} \cdot \sum_{i}^{N} \ln(\alpha_i) \cdot h_i(\boldsymbol{x}), \tag{11.1}$$

where the weights α_i are calculated as $\alpha_i = (1 - r_i)/r_i$, with r_i describing the misclassification rate. The number of individual classifiers is given by N and $\boldsymbol{x}$ describes the set of input parameters.

Similarly, the Gradient boosting combines several weak classifiers through the weighted sum of the individual decision trees $f(\boldsymbol{x}; a_m)$:

$$F(\boldsymbol{x}; P) = \sum_{m=0}^{M} \beta_m f(\boldsymbol{x}; a_m); \qquad P \in \{\beta_m; a_m\}_0^M, \tag{11.2}$$

The boosting is then performed by optimizing the parameter set $P \in \{\beta_m; a_m\}_0^M$, such that the difference between the response of the classifier $F(\boldsymbol{x})$ and the true value y is minimized. The parameters a_m describe the properties of the individual classifiers, while the parameters β_m describe the impact of the individual classifiers in the weighted sum. The number of individual classifiers is given by M. The difference between $F(\boldsymbol{x})$ and y is quantified by the loss-function $L(F, y)$. For Gradient boosting, a binomial log-likelihood loss function is used in the TMVA package, which is defined as follows:

$$L(F, y) = \ln\left(1 + e^{-2F(\boldsymbol{x})y}\right) \tag{11.3}$$

The loss function is minimized during the training with the help of a numerically approximation, based on the steepest-descent approach. This step is repeated for a large number of decision trees, which are added successively, in order to obtain a combined classifier providing an improved discrimination between signal and background. Hence, the classification is improved by the use of boosting algorithms, as these assess the weaknesses of the classifier in each step of the training either through the misclassification rate, in the case of AdaBoost, or the gradient of the loss function, in the case of Gradient boosting. In the different steps of the training, additional classifiers are added successively to account for these weaknesses and improve the classification algorithm. Instead of classifying an event as signal or background, as in the case of a single decision tree, a BDT assigns a weight between -1 and 1 to each event.

A common issue of MVA algorithms is the so-called overtraining, which describes the effect that the classifier learns features in the training sample that result from statistical fluctuations. This might happen e.g. if too many training iterations are performed, which is equivalent to a too large number of nodes in the case of decision trees. The overtraining of decision trees can be regulated by the use of *pruning* methods, which remove statistically insignificant nodes after the training. Furthermore, in order to prove the presence of overtraining, the MC sample is usually split into a training and a testing sample. The testing sample is statistically independent of the training sample and allows to separately evaluate the performance of the classifier. If the

performance of the classifier on the testing sample is significantly different compared to the training sample, this provides an indication for overtraining. However, as BDTs are built using weak classifiers, meaning simple trees with a small tree depth, overtraining is a minor problem of BDTs and a pruning of the single decision trees is not performed.

11.3.2 Training of the Boosted Decision Tree

The BDT is trained using events from signal and background MC samples, aiming for a discrimination between the H+b-jets signal and the continuum background. A preselection is applied to both, signal and background events, before performing the training. This preselection ensures that only events are used for the training which are loosely consistent with the H+b-jets final state. Otherwise the training becomes inefficient due to the large variety of possible final states. The BDT preselection requirements are summarized in Table 11.2. Similar to the nominal preselection, a veto on leptons is applied and a maximum number of three central jets is allowed. In addition, at least one jet is required to be b-tagged with the 85% WP. This loose WP is chosen to preserve a sufficient number of signal events for the training.

Finally, only events with a diphoton invariant mass between 122.5 GeV and 127.5 GeV are used for the training. While this requirement is fulfilled for most Higgs boson signal events, it significantly reduces the number of continuum background events. It is applied to ensure that the BDT discriminant remains uncorrelated to the $m_{\gamma\gamma}$ distribution. A correlation between the BDT discriminant and $m_{\gamma\gamma}$ might deform the shape of the $m_{\gamma\gamma}$ distribution of the continuum background and, thus, make it difficult to find an accurate background model for the BDT-based categories. As the BDT input variables contain information about the two p_{T}-leading photons, information about the shape of the $m_{\gamma\gamma}$ distribution is indirectly provided to the BDT. Additionally, the photon kinematics of the continuum background differ over the nominal $m_{\gamma\gamma}$ range of 105 GeV to 160 GeV. A discrimination between the continuum background and H+b-jets is, however, of particular interest in the region close to the Higgs boson signal peak, where the photon kinematics are very similar for signal and background.

Table 11.2: Summary of the preselection requirements, which are applied to the signal and background MC samples before the training of the BDT.

BDT preselection requirements
Veto on electrons & muons
$1 \leq N_{\mathrm{jets}}^{\mathrm{central}} \leq 3$
≥ 1 b-tagged jet (85% WP)
$122.5\,\mathrm{GeV} \leq m_{\gamma\gamma} \leq 127.5\,\mathrm{GeV}$

Although the continuum background is usually estimated from data in $H \to \gamma\gamma$ analyses, the training is performed using diphoton MC samples. This allows using information about the photons in the Higgs boson mass window, which corresponds to the region of interest. The distributions of the discriminating variables were found to be in good agreement for diphoton MC simulation and the data $m_{\gamma\gamma}$ sidebands, as shown in Figure 11.4 in the following section. For the signal, events from the ggH and VBF samples are used, which are characterized by truth-matched b-jets or c-jets. Events with c-jets in the final state are included in the signal training sample, as the kinematics of events with b-jets and c-jets is similar and otherwise the number of available training events is not sufficient. A summary of the training samples is

presented in Table 11.3. The total number of signal events after the BDT training preselection amounts to 16248, among which 12000 events are used for the training and the remaining events are used for testing. For the background sample, a relatively large number of 50000 training events is available and 21587 events are used for testing.

Table 11.3: Considered signal & background MC samples for the BDT training and additional requirements for the signal training samples.

Classification	MC sample	Additional requirements
Signal	ggH	Truth-matched b-jets or c-jets
Signal	VBF	Truth-matched b-jets or c-jets
Background	Diphoton MC	-

Finally, the training is performed using eight input variables, which are described in the following section. Prior to the training, a decorrelation of the input variables is performed, which accounts for linear correlations between approximately Gaussian distributed variables. This preprocessing of the input variables is a common approach, as the performance of the BDT is worse in the presence of strong correlations between the input variables. The BDT is trained using Gradient boosting in combination with bagging techniques, in order to obtain a stable and well performing classification algorithm, despite the low number of events in the signal training sample. Furthermore, due to the low number of events in the signal training sample, decision trees with a small maximum depth of 2 are trained in order to avoid overtraining. A number of 200 trees is found to be sufficient, as for a larger number of trees the classification performance of the BDT converges. This setup was found to provide the best discrimination, among different tested boosting algorithms and BDT parameters. Additionally, a comparison to an NN-based approach, using the same set of discriminating variables, was performed and a very similar performance was obtained.

11.3.3 Discriminating variables

In order to find quantities that provide a discrimination between H+b-jets and the continuum background, several variables were defined and tested as input variables. Among these were relatively simple kinematic variables, such as the p_{T} and η of photons and jets, the ratios of the p_{T} of two particles and ΔR, $\Delta\eta$ and $\Delta\phi$ distributions for pairs of particles. Additionally, so-called event shape variables were considered, which do not only provide information about a single particle or a pair of particles, but instead parameterize the structure of the whole event. These event shape variables included e.g. the sphericity and aplanarity of an event, describing how equally the final state particles are distributed in the detector space. In total, about 80 variables were tested among which the 8 most discriminating variables were chosen for the training. These input variables are described in more detail in the following. The final set of input variables includes

- the ΔR and $\Delta\eta$ of the $\gamma\gamma$ system,
- the ΔR between the p_{T}-leading and -subleading photons (γ_1 and γ_2) and the MV2-leading and -subleading jets ($j_{1,\mathrm{MV2}}$ and $j_{2,\mathrm{MV2}}$),
- the first order FoxWolfram moment of the system built of $\gamma\gamma$ and the central jets and
- the effective mass of the system built of $\gamma\gamma$ and the central jets.

Variables based on the two p_{T}-leading photons and the two MV2-leading jets, which are the jets with the highest probability of originating from a b quark in each event, were found to provide the best discrimination between H+b-jets and the continuum background. The angular distance between the two photons generally provides discrimination between $H \to \gamma\gamma$ events and the continuum background. It is related to $m_{\gamma\gamma}$ by the equation $m_{\gamma\gamma} = \sqrt{2E_{\gamma 1}E_{\gamma 2}(1-\cos\alpha)}$, with α representing the opening angle between the two photons. In SM diphoton production, the two photons are on average characterized by larger angular separations ΔR and $\Delta\eta$ than in $H \to \gamma\gamma$ events. Due to the different kinematics of SM diphoton production and $H \to \gamma\gamma$ decays, the angular distances ΔR between the photons and the jets provide additional discrimination. These ΔR distributions show larger tails toward lower values of ΔR for SM diphoton production.

Among the considered event shape variables, the FoxWolfram moment of first order ($FoxW_1$) and the effective mass (m_{eff}) of the particle system consisting of the two photons and the central jets in each event were found to provide the best discrimination. The effective mass is defined based on the four-vectors p_i of the considered particles:

$$m_{\mathrm{eff}} = \sqrt{\left(\sum_i p_i\right)^2} \tag{11.4}$$

This variable mainly provides separation between VBF production and the continuum background. For VBF production, m_{eff} adopts higher values compared to ggH and the continuum background, due to the on average higher number of jets in the final state.

In contrast, the FoxWolfram moments $FoxW_i$ [155, 156], which were introduced as observables describing the geometrical structure of multijet final states, are based on the angular correlations between pairs of particles in an event. These are parameterized in terms of spherical harmonics, so that the FoxWolfram moments of order ℓ are defined as follows:

$$FoxW_\ell = \sum_{i,j=1}^{N} \frac{|\vec{p_i}||\vec{p_j}|}{E_{tot}} \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{+\ell} Y_\ell^m(\Omega_i) Y_\ell^{m*}(\Omega_j) \tag{11.5}$$

$$= \sum_{i,j=1}^{N} \frac{|\vec{p_i}||\vec{p_j}|}{E_{tot}} P_\ell(\cos\Omega_{ij}) \tag{11.6}$$

$$\tag{11.7}$$

The product of the spherical harmonics Y_ℓ^m can be rewritten in terms of Legendre polynomials P_ℓ, based on the addition theorem of spherical harmonics. The parameters Ω_i describe the angle between particle i and a reference axis, while Ω_{ij} describes the angle between a pair of particles i and j. The momentum vector $\vec{p_i}$ is defined based on the spatial component of the four-vector of a particle i and E_{tot} describes the sum of the energies of all considered particles in the event. Hence, the FoxWolfram moments combine information about the number of particles in an event, their momenta and angular correlations. FoxWolfram moments of first to fourth order were considered, which can adopt values between 0 and 1. The $FoxW_1$ values of the $\gamma\gamma$+central jets particle system tend to be closer to 0 for the signal training samples.

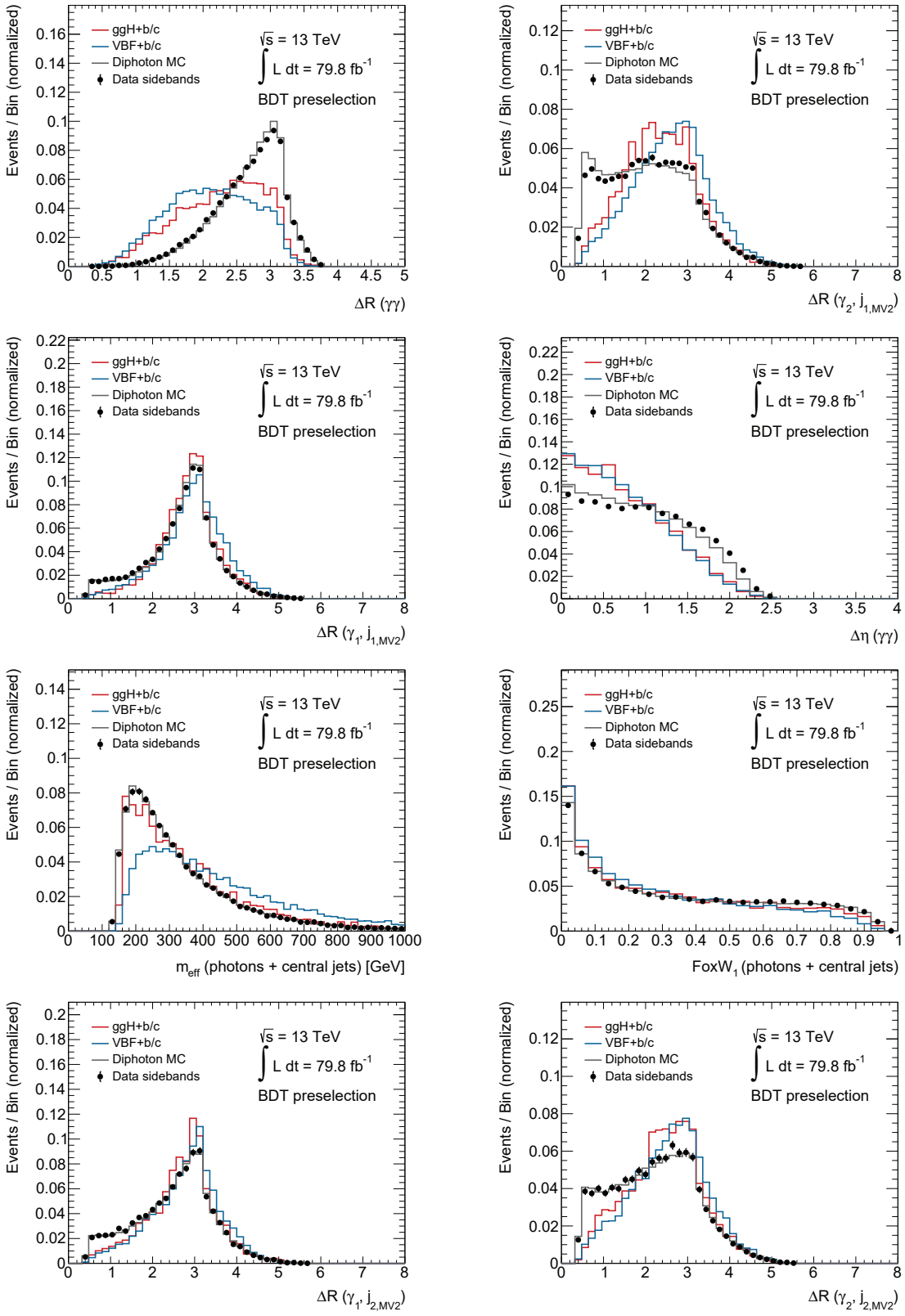


Figure 11.4: Distributions of the BDT input variables, after applying the BDT preselection, for signal and background training samples and a comparison to the distributions of the data $m_{\gamma\gamma}$ sidebands. The order of the distributions represents the importance of the individual variables for the BDT-based classification.

The distributions of the BDT input variables are shown in Figure 11.4 for the signal and background training samples. In addition, the distributions are shown for the data $m_{\gamma\gamma}$ sidebands, in order to check if these are in agreement with the distributions in the diphoton MC samples. As the $m_{\gamma\gamma}$ region between 120 to 130 GeV is blinded in data, the $m_{\gamma\gamma}$ sidebands are used for this comparison. In general, a good agreement is observed between the background distributions in the data sidebands and the diphoton MC simulation. However, some distributions show small discrepancies, which can most probably be explained by the absence of the γj and jj background components in diphoton MC simulation. In the following section, a comparison of the BDT discriminant obtained for the diphoton MC samples and data sidebands is shown, in order to check if these small discrepancies cause larger deviations in the BDT discriminant.

The correlations between the BDT input variables are mostly small. An exception occurs for $\Delta R(\gamma_1, j_{2,\mathrm{MV2}})$ and $\Delta R(\gamma_2, j_{2,\mathrm{MV2}})$, as these variables are both set to a default value of -999 when only one central jet is present in the event, resulting in a large correlation. The correlation matrices for the signal and background training samples are shown in Appendix E.1.

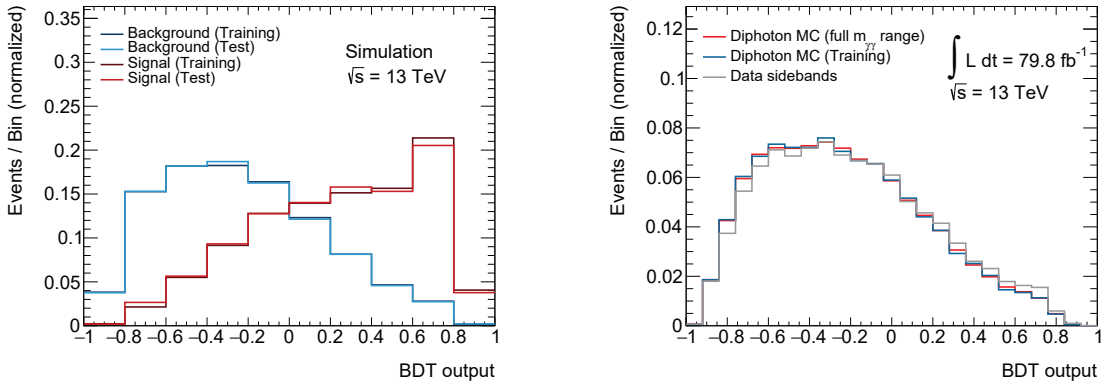
Finally, the discriminating variables are ranked by the TMVA package according to their importance for the BDT-based classification, as shown in Table 11.4. This importance is calculated based on how frequent a certain variable is used to define a criterion at one of the nodes. Additionally, this number is weighted by the separation which is gained by the respective criterion and the statistical significance of the nodes. The three most important discriminating variables are the ΔR between the two p_{T}-leading photons and the ΔR between the MV2-leading jet and the p_{T}-leading and -subleading photon, respectively.

Table 11.4: Ranking of the BDT input variables according to their importance, which is calculated based on how frequent a certain variable is used to define a criterion at one of the nodes, the separation which is gained by these criteria and the statistical significance of the nodes.

	Input variable	Importance
1.	$\Delta R(\gamma\gamma)$	0.168
2.	$\Delta R(\gamma_2, j_{1,\mathrm{MV2}})$	0.129
3.	$\Delta R(\gamma_1, j_{1,\mathrm{MV2}})$	0.120
4.	$\Delta\eta(\gamma\gamma)$	0.120
5.	$m_{\mathrm{eff}}(\gamma\gamma$+central jets)	0.119
6.	$FoxW_1(\gamma\gamma$+central jets)	0.118
7.	$\Delta R(\gamma_1, j_{2,\mathrm{MV2}})$	0.114
8.	$\Delta R(\gamma_2, j_{2,\mathrm{MV2}})$	0.112

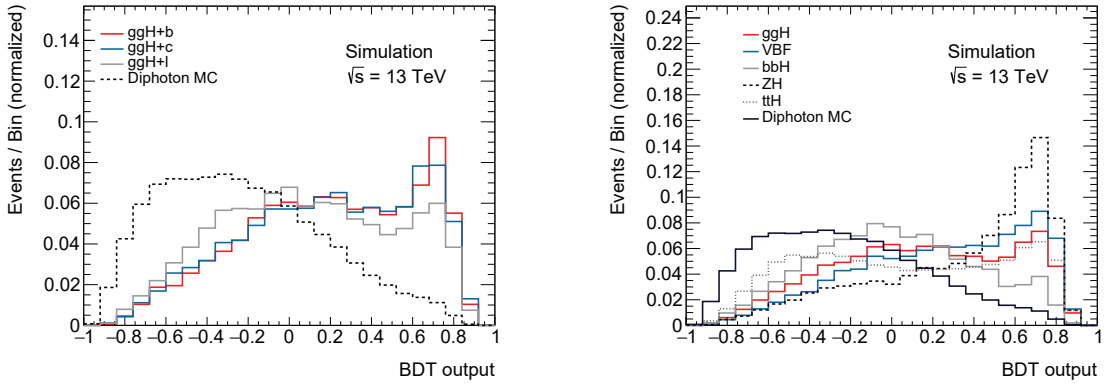
11.3.4 Final discriminant & categorization

After the training, the BDT classification results in a single discriminant, which separates H+b-jets from the continuum background. In Figure 11.5a), the BDT discriminant is shown for the training and testing samples of signal and background, respectively. The BDT output for the signal sample is shifted to the right and has its maximum at approximately 0.7, while the background distribution is shifted to the left and has its maximum at approximately -0.3. Hence, for higher values of the BDT discriminant the signal contribution is significantly enhanced, while the background is reduced in this region. Additionally, in Figure 11.5a) a very



a) b)

Figure 11.5: a) Illustration of the BDT discriminant of the training and testing samples for signal and background, respectively, b) comparison of the BDT discriminant for the background training sample, the diphoton MC sample in the full 105-160 GeV diphoton invariant mass range and the data $m_{\gamma\gamma}$ sidebands.



a) b)

Figure 11.6: a) Comparison of the BDT discriminant for diphoton MC simulations and the ggH MC sample, where the MV2-leading jet is classified as a b-jet, c-jet or ℓ-jet, b) comparison of the BDT discriminant for various Higgs boson production processes and the diphoton MC sample.

good agreement between the BDT discriminants obtained with the training and testing samples is shown, so there is no indication of overtraining.

Furthermore, the shape of the BDT discriminant is investigated for the background in more detail, as the diphoton MC sample is used for the BDT training whereas the final estimation of the continuum background is performed using data. A comparison of the BDT discriminant for the background training sample, the diphoton MC sample in the full $m_{\gamma\gamma}$ window from 105 GeV to 160 GeV and the data $m_{\gamma\gamma}$ sidebands is presented in Figure 11.5b). A good agreement between the BDT discriminant obtained with the diphoton MC sample and the data sidebands is observed, despite the small discrepancies in the distributions of the input variables.

Additionally, the performance of the BDT for final states with b-jets, c-jets and ℓ-jets and different Higgs boson production processes is investigated. In Figure 11.6a), the BDT discriminant

is compared for ggH events in which the MV2-leading jet is either identified as a truth b-jet, c-jet or ℓ-jet. It is noticeable that the performance of the BDT is very similar for all ggH events, regardless of the truth flavor of the jet. Some of the most important BDT input variables only rely on photon information and, hence, provide a general discrimination between Higgs boson production and continuum background processes. But nevertheless, the separation is slightly better for events with b-jets and c-jets in the final state. The similarity of the BDT discriminants for events with b-jets and c-jets verifies the approach of also using events with c-jets during the training. In Figure 11.6b), a comparison of the BDT discriminant for various Higgs boson production processes is shown. While the BDT discriminants for ggH and VBF production are very similar, the BDT performs slightly worse for processes like $b\bar{b}H$ and $t\bar{t}H$ probably due to the different origin of the b-jets in these processes. Surprisingly, the best discrimination is obtained for ZH production, although this process was not considered during the training.

Table 11.5: Expected number of events in the three BDT categories for Higgs boson production and the continuum background for $79.8\,\text{fb}^{-1}$. The continuum background contribution is estimated from the number of events in the data sidebands and the number of events in the blinded 120-130 GeV region, which is estimated by calculating the integral of an ExpPoly2 fit function in this region. In addition, the expected significance, considering only statistical uncertainties, to inclusive Higgs boson production is shown for each category.

	BDTcat1	BDTcat2	BDTcat3
ggH	9.69	16.7	6.31
VBF	2.18	2.78	0.79
WH	0.696	0.86	0.442
ZH	2.12	2.26	1.05
$t\bar{t}H$	0.806	1.21	0.919
$tHjb$	0.43	0.608	0.251
WtH	0.0416	0.0349	0.0222
$b\bar{b}H$	0.46	1.84	0.768
S	16.4	26.3	10.6
B_{ContBG}	818	4140	4230
Exp. sig. $[\sigma]$	1.6	1.2	0.5

As explained in the previous sections, events in the background-dominated region of the BDT discriminant are not rejected, as this would result in a signal region which is not well-defined on particle level. Instead, the BDT discriminant is used to split the phase space into three regions with varying S/B_{ContBG}. The requirements on the BDT discriminant are optimized in order to maximize the expected significance to events in the BDT signal training sample, which is quantified using the approach described in Chapter 7. First, the BDT discriminant is scanned from the right to the left, in order to find the cut which provides the highest sensitivity. This region is used to define the first BDT category, referred to as BDTcat1. Then, starting from the first cut value, the distribution is scanned again from the right to the left in order to find the second cut-value, which maximizes the expected significance when being combined with the first category. The region between the first and the second requirement on the BDT output is used to define the category BDTcat2. The remaining events define the third BDT category, BDTcat3, which is dominated by continuum background events. The cut-values, which are found to be the optimal choices, are 0.46 for BDTcat1 and -0.26 for BDTcat2.

In Table 11.5, the expected event yields for the three BDT categories and the corresponding expected sensitivities to inclusive Higgs boson production are shown. These event yields are evaluated for the full reconstruction level event selection, as described in Section 11.2. The corresponding BDT discriminant, which is obtained after applying the full event selection, is shown in Appendix E.2. The continuum background contribution is estimated from the number of events in the data $m_{\gamma\gamma}$ sidebands and the number of events in the blinded 120-130 GeV region, which is estimated by calculating the integral of an ExpPoly2 fit function in this region. The BDT categorization increases the expected significance, considering only statistical uncertainties, from $1.6\,\sigma$ to $2.1\,\sigma$. The influence of the BDT categorization on the background modeling and the signal extraction uncertainties is discussed in the following sections.

11.4 Influence on the background modeling

The BDT-based categorization, as described in the previous section, is applied for the signal extraction fit in order to increase the sensitivity to the Higgs boson signal. As the BDT uses information about the two p_{T}-leading photons, which are also used to define the $m_{\gamma\gamma}$ discriminant, it is investigated if the BDT categorization influences the shape of the $m_{\gamma\gamma}$ distribution in the individual categories. This might e.g. cause problems when choosing a background model for the final S+B fit to $m_{\gamma\gamma}$. Hence, the shape of the $m_{\gamma\gamma}$ distribution is studied in diphoton MC simulation for the three BDT categories, as shown in Figure 11.7. For each BDT category, a monotonically falling $m_{\gamma\gamma}$ distribution is observed in the full $m_{\gamma\gamma}$ range from 105 to 160 GeV.

For choosing a background model, an S+B fit is performed to the B-only templates of each category to estimate the bias which is introduced by the choice of fit function. The B-only templates are derived from diphoton MC samples and the shape of the γj and jj components is estimated from control regions in data, as described in Chapter 7.3. The signal is modeled with a DCB function, which provides a good description of the Higgs boson signal for the three BDT categories, as shown in Appendix E.3.

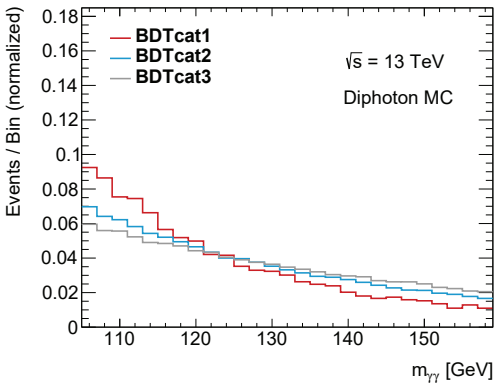
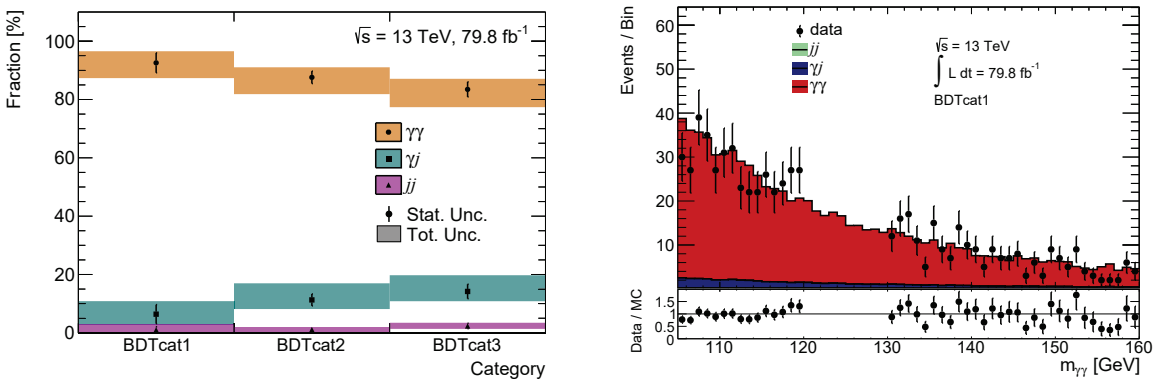


Figure 11.7: Comparison of the shape of the $m_{\gamma\gamma}$ distributions in diphoton MC simulation for the three BDT categories.

In order to construct the background templates, the background decomposition of the three BDT categories into $\gamma\gamma$, γj and jj components is derived using the $2 \times 2D$ sideband method, which was described in Chapter 7.3. The resulting $\gamma\gamma$, γj and jj fractions of the different

categories are shown in Figure 11.8a). It is noticeable that BDTcat1 is characterized by a comparably high $\gamma\gamma$ purity of about 92.6%, which decreases to 87.6% and 83.5% for BDTcat2 and BDTcat3. Hence, the BDT output does not only distinguish between Higgs boson processes and the continuum background, but also between events with prompt photons and *fake* photons in the final state.



a) b)

Figure 11.8: a) Background decomposition into $\gamma\gamma$, γj and jj components for the three BDT categories, b) example background template for BDTcat1, where the $\gamma\gamma$ contribution is taken from diphoton MC simulation and the shape of the γj and jj components is derived from control regions in data, as described in Chapter 7.3.

The shape of the $\gamma\gamma$ component of the background template is then taken from diphoton MC simulation. The γj and jj components are derived by reweighting the shape of the $\gamma\gamma$ template to the shape of the γj and jj control regions in data, as described in Chapter 7.3. Due to the small number of events in the jj control region, it is challenging to properly derive the shape of this component. However, due to its small contribution, the jj component has a minor influence on the overall background template. An example for the resulting background template is shown in Figure 11.8b) for BDTcat1. Within the statistical uncertainties, the shape of the background template is in good agreement with the data sidebands. The background templates for the remaining BDT categories and supplementary plots for the reweighting of the γj and jj components are shown in Appendix E.4.

Table 11.6: Chosen functional forms for the three BDT categories, the maximum amount of fitted signal, $N_{\text{spur}}^{\text{max}}$, and the two ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$. As the nominal requirements on $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ are not fulfilled, the relaxed spurious signal criteria are considered.

Category	Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$
BDTcat1	PowerLaw	-31.6	-18.5	-3.0
BDTcat2	PowerLaw	49.9	47.5	12.3
BDTcat3	Exponential	-23.2	-59.1	-6.3

Finally, a background model is chosen in each category based on the maximum amount of fitted spurious signal, $N_{\text{spur}}^{\text{max}}$. Due to the relatively large statistical fluctuations in the background templates, the relaxed spurious signal criteria are applied, which take into account the statistical uncertainty of the background template. For each category, either the PowerLaw or the Exponential function passes the relaxed spurious signal criteria and is chosen as background

model, as summarized in Table 11.6. The two criteria on the amount of fitted signal are illustrated in Figure 11.9 for the chosen functional forms.

In general, the $m_{\gamma\gamma}$ distribution in the different categories is not strongly affected by the BDT categorization and can still be described by the common functional forms. However, the spurious signal uncertainties in the three categories are relatively large, ranging from approximately 19% to 59% with respect to the expected signal yield, due to statistical fluctuations in the background templates.

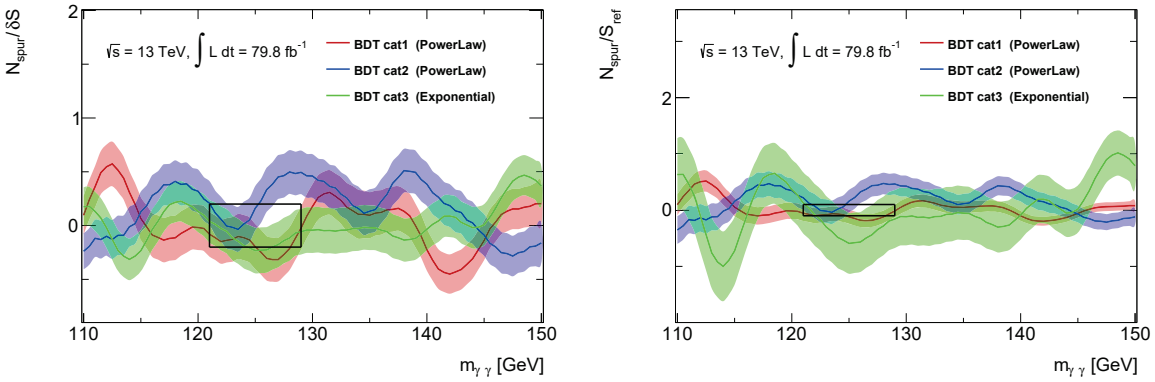
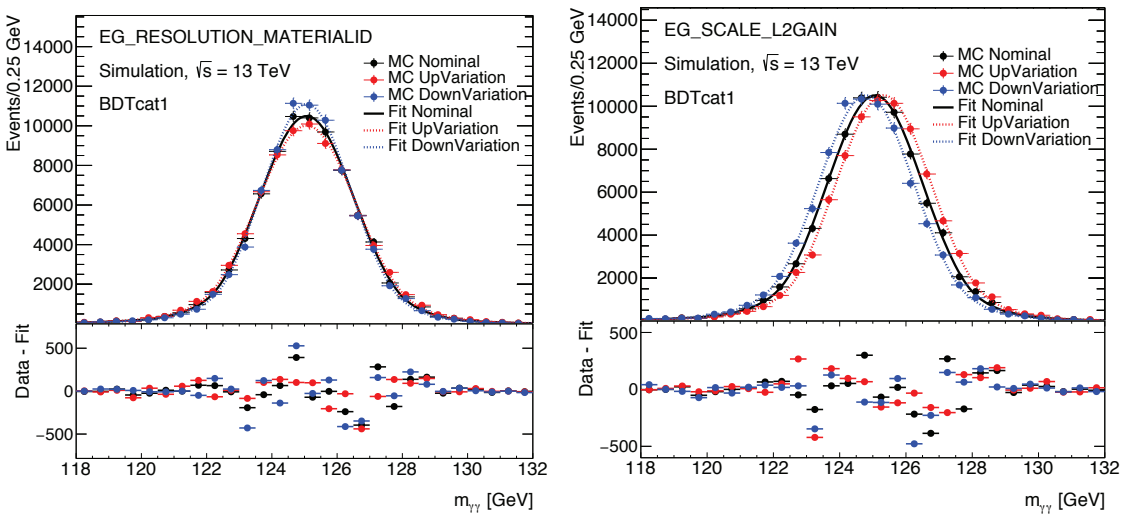


Figure 11.9: Illustration of the criteria on $N_{\text{spur}}/\delta S$ and $N_{\text{spur}}/S_{\text{ref}}$ (including the 1σ uncertainty bands), which are indicated by the black squares, for the chosen functional forms of the three BDT categories. The relaxed spurious signal criteria are applied, which take into account the $2\,\sigma$ uncertainty bands of the fitted spurious signal.

11.5 Expected signal extraction uncertainties

As presented in Chapter 10.6, the uncertainties resulting from the signal extraction fit are dominant in the $N_{b\text{-jets}}$ measurement. The largest component is the statistical uncertainty on the extracted signal yields, followed by the background modeling uncertainty and the PES/PER uncertainties. An attempt to decrease the statistical uncertainty by categorizing the events, based on the BDT discriminant, is presented in this chapter. However, while expecting a decrease of the statistical uncertainty, the BDT categorization may also have an impact on the size of the systematic uncertainties. For this reason, the impact of the BDT categorization on both, the statistical and systematic signal extraction uncertainties, is investigated. In contrast, the BDT categorization has no impact on the unfolding uncertainties, as it is only applied for the signal extraction. The unfolding uncertainties are expected to have a similar size to those in the $N_{b\text{-jets}} = 1$ bin, presented in the previous chapter, as the definitions of the preselection and the fiducial volume are very similar. Hence, the unfolding uncertainties, which play a minor role in statistically limited cross section measurements, are not discussed in this chapter. The method for estimating the signal extraction uncertainties is similar to the procedure in the $N_{b\text{-jets}}$ measurement. The expected statistical uncertainty is estimated based on a profile likelihood test statistic by performing a fit to an Asimov dataset where the NPs are fixed to their best-fit values. The Asimov dataset is generated based on the MC expectation for the Higgs boson processes and the data $m_{\gamma\gamma}$ sidebands

Additionally, the PES/PER uncertainties affecting the signal model are derived for the three BDT categories. Therefore, a DCB function is fitted to the $m_{\gamma\gamma}$ distribution of the respective



a) b)

Figure 11.10: Example for the nominal $m_{\gamma\gamma}$ distribution of BDTcat1 in MC simulations and the up- and down-variations of a) one of the PER uncertainties, which changes the width of the distribution, and b) one of the PES uncertainties, causing a shift of the $m_{\gamma\gamma}$ distribution.

category for each PES/PER variation. In Figure 11.10, an example for the influence of the PES/PER variations on the $m_{\gamma\gamma}$ distribution is shown for BDTcat1. The up and down PER variations result in a change of the width of the DCB function, while the up and down PES variations generate a shift in the mean of the DCB function. The uncertainties are defined as the relative difference between the mean or the width of the DCB functions, respectively, for each PES/PER variation. The uncertainties for all PES/PER variations, which are evaluated with this approach, are shown in Appendix E.5. In addition, an uncertainty on the assumed Higgs boson mass of 125.09 GeV is considered in the signal extraction fit.

The expected statistical uncertainty in the inclusive fiducial volume, without applying the BDT categorization, amounts to approximately 70%. This uncertainty is larger than the statistical uncertainty in the $N_{b\text{-jets}} = 1$ bin, due to the additional criterion on $N_{\text{jets30}}^{\text{central}}$. Based on a fit to an Asimov dataset, the expected statistical and systematic uncertainties on the extracted signal yield are estimated for the three BDT categories. The fractional statistical and systematic uncertainties for the different categories are shown in Figure 11.11a), when excluding the background modeling uncertainty, and in Figure 11.11b) for the full set of systematic uncertainties. In addition, the three contributions to the total uncertainty are summarized in Table 11.7.

The expected statistical uncertainty (σ_{stat}) for BDTcat1 alone is already smaller than the expected statistical uncertainty for the case that no BDT categorization is applied. In general, the statistical uncertainty is predicted to be by far the largest uncertainty in each category, followed by the background modeling uncertainty (σ_{bkg}). The PES/PER uncertainties and the Higgs boson mass uncertainty, summarized as σ_{syst}, are expected to have a size of approximately 5% in the first two BDT categories, while these uncertainties are expected to be slightly increased for the last BDT category. The extracted signal yields for the three BDT categories need to be combined before performing the unfolding, thus the uncertainties on the extracted signal yields need to be combined as well. As each data event can only enter one of the categories,

Table 11.7: Expected signal extraction uncertainties, including the statistical uncertainty (σ_{stat}), the systematic uncertainties due to the PES/PER variations and the Higgs boson mass uncertainty (σ_{syst}) and the background modeling uncertainty (σ_{bkg}) for the three BDT categories. These numbers are combined following the procedure described in the text.

Category	σ_{stat}	σ_{syst}	σ_{bkg}	Total
BDT cat1	65.7%	5.5%	18.5%	68.5%
BDT cat2	97.8%	5.0%	47.5%	108.8%
BDT cat3	253.8%	12.1%	59.1%	260.9%
Combination	53.3%	12.1%	26.6%	60.8%

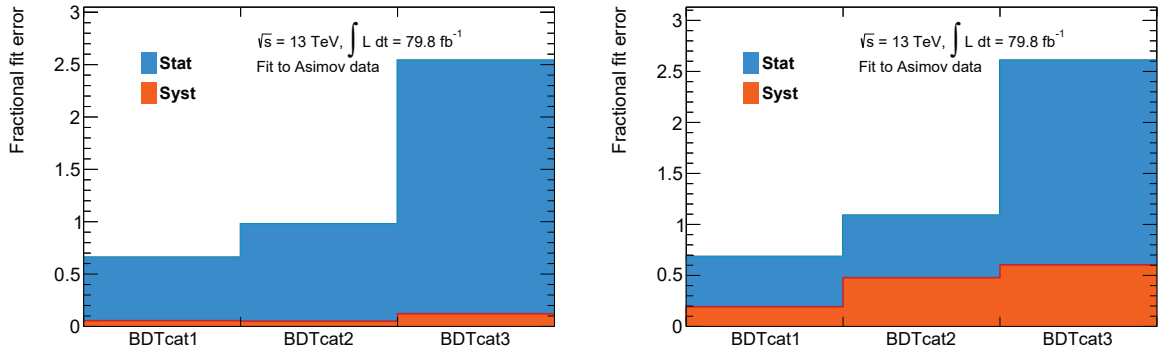


Figure 11.11: Fractional fit uncertainties of the three BDT-based categories. The statistical uncertainty is shown in blue, while the systematic uncertainty is shown in orange. In a) the systematic uncertainty only includes the PES/PER and the Higgs boson mass uncertainties, while in b) the background modeling uncertainty is included as well.

the statistical uncertainties are not correlated among the different categories. Hence, the total statistical uncertainty can be evaluated using the formula:

$$\sigma_{\text{stat,total}} = \sqrt{\frac{1}{\sum_i 1/\sigma^2_{\text{stat},i}}} \tag{11.8}$$

For the combination of the systematic uncertainties, in general the correlations between the different categories need to be taken into account. The background modeling uncertainties are estimated from independent background templates and are, hence, treated as uncorrelated between the different categories. In contrast, the Higgs boson mass uncertainty and the PES/PER uncertainties are correlated between the different categories and cannot be easily combined. However, as these uncertainties only have a minor impact on the total uncertainty, the largest value of σ_{syst} is considered for a rough estimation of the total signal extraction uncertainty. When applying the BDT categorization, the statistical uncertainty decreases to about 53%, thereby providing a significant improvement of the dominant contribution to the total signal extraction uncertainty. The combination of the statistical uncertainty and the systematic uncertainties yields approximately 61%.

11.6 Discussion

In this chapter, a strategy for an improved measurement of H+b-jets was presented, based on the application of a BDT to discriminate between H+b-jets and the continuum background. The use of MVA techniques has not been studied before in the context of fiducial cross section measurements in the $H \to \gamma\gamma$ channel at the ATLAS experiment, as it might introduce a model dependence during the unfolding. However, as the strategy presented in this chapter only involves the application of a BDT to categorize events for the signal extraction fit, it has no influence on the unfolding. This approach decreases the expected statistical uncertainty on the measured cross section from approximately 70% to 53%. The reduction of the statistical uncertainty is equivalent to an increase of the amount of data of approximately 70%.

However, the BDT categorization is accompanied by relatively large background modeling uncertainties, due to statistical fluctuations in the background templates. For larger datasets, these large background modeling uncertainties will limit the improvement obtained by the BDT categorization. In addition, for a combination of the uncertainties on the extracted signal yields, the correlations between the PES/PER uncertainties should be taken into account properly. For the general studies presented in this thesis, an estimation of these correlations is not important, as the impact of the PES/PER and the Higgs boson mass uncertainties on the total uncertainty is small. Based on the approximate calculation presented in the previous section, an improvement of the total signal extraction uncertainty is expected when applying the BDT categorization.

Additionally, a requirement on the maximum number of jets was investigated to reduce the $t\bar{t}H$ contribution. A requirement of $N_{\text{jets30}}^{\text{central}} \leq 3$ was found to efficiently reduce $t\bar{t}H$ events, while keeping the efficiency for ggH production close to 100%. This additional requirement is not expected to introduce a significant model dependency during the unfolding, as the fraction of ggH events with more than three central jets with $p_{\text{T}} > 30\,\text{GeV}$ is found to be very small in MC samples produced with different parton shower generators.

Further improvements could be obtained by decreasing the background modeling uncertainty, whose size is mainly determined by the statistical fluctuations in the background templates. The use of larger MC samples or the application of smoothing techniques could reduce the statistical fluctuations in the background templates.

Chapter 12

Conclusion & Outlook

In this thesis, the Higgs boson production in association with a single top quark (tH) and the H+b-jets background were investigated in the $H \to \gamma\gamma$ decay channel at the ATLAS experiment. Datasets collected at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ during Run 2 of the Large Hadron Collider (LHC) were considered, which correspond to integrated luminosities of $36.1\,\text{fb}^{-1}$ and $79.8\,\text{fb}^{-1}$. As the Higgs boson does not directly couple to massless particles like photons, the decay of the Higgs boson into photons is rare. Nevertheless, the $H \to \gamma\gamma$ channel was one of the discovery channels of the Higgs boson [7, 8] and provides a promising final state to study the properties of the Higgs boson. A fit to the diphoton invariant mass, $m_{\gamma\gamma}$, allows to measure the Higgs boson signal as a narrow resonance above the decreasing continuum background.

In the $H \to \gamma\gamma$ decay channel, events with two isolated, high-p_{T} photons that fulfill certain requirements on the shape of the calorimeter shower, which are referred to as *tight* photon identification (ID) criteria, are considered. The photon ID criteria are used to distinguish between prompt photons and so-called *fake* photons, such as photons inside a jet or jets which are misidentified as photons. A measurement of the *tight* photon ID efficiency in data was performed as part of this thesis. This measurement makes use of a strict track isolation requirement to distinguish between prompt and *fake* photons. The *tight* photon ID efficiency was measured with the 2015-2017 dataset and new sources of systematic uncertainty were studied and included in the measurement. The measured *tight* photon ID efficiencies range from approximately $(80.3 \pm 13.9)\%$ to $(98.2 \pm 0.5)\%$ and a good agreement between the efficiencies extracted from the data and predicted by MC simulation is observed.

The coupling of the Higgs boson to the top quark, Y_t, is one of the fundamental properties of the Higgs boson. As the top quark is the heaviest particle in the Standard Model (SM), Y_t is expected to be large in the SM. Hence, the measurement of Y_t is important to prove the compatibility of the Higgs boson observed at the LHC and the properties predicted by the SM. Due to a destructive interference in tH production, this process provides sensitivity to negative values of the coupling strength modifier $\kappa_t = Y_t/Y_t^{SM}$. A measurement of SM-like tH production is challenging due to its small cross section. However, deviations from Y_t^{SM} would result in an enhanced tH cross section. The CMS experiment performed searches for tH production during Run 1 and Run 2 [143, 144], assuming either the anomalous coupling $\kappa_t = -1$ or the SM coupling. At the ATLAS experiment, dedicated tH categories were considered for the first time in the combined measurement of top-quark-associated production in the $H \to \gamma\gamma$ channel, using $36.1\,\text{fb}^{-1}$ [2]. Four tH categories targeting final states with leptons and fully hadronic final states were optimized as part of this thesis. For the definition of the leptonic categories, the presence of a forward jet in the t-channel production is exploited to improve the sensitivity to tH production. The main backgrounds are the continuum background and the

production of a top-antitop quark pair in association with a Higgs boson ($t\bar{t}H$). In addition, Higgs boson production via gluon-gluon fusion (ggH) with additional b-jets in the final state is an important background for hadronic final states. Due to the similarity of the tH and $t\bar{t}H$ final states and the small sensitivity to tH production alone, a combined measurement of tH+$t\bar{t}H$ production was performed. Furthermore, a strategy for choosing a background model in the tH and $t\bar{t}H$ categories based on data-driven background templates was developed as part of this thesis. Choosing a proper background model is an important ingredient for analyses in the $H \to \gamma\gamma$ channel, as a bias on the fitted number of Higgs boson events is introduced by the choice of fit function. The measured signal strength, $\mu = \sigma_{obs}/\sigma_{SM}$, for $t\bar{t}H$+tH production is in agreement with the SM expectation of $\mu = 1$. The best-fit value of the coupling strength modifier ratio $\lambda_{tg} = \kappa_t/\kappa_g$, with $\kappa_g^2 = \sigma_{ggH}/\sigma_{ggH,SM}$, was measured to be $0.8\,^{+0.4}_{-0.6}$, providing no hint for a negative sign of Y_t. In this measurement, no assumptions were made on the coupling of the Higgs boson to other particles. This result shows a similar tendency for the sign of Y_t as the result of the latest search for tH production at the CMS experiment, where positive values of κ_t were found to be favored by the data as well [144]. In the CMS measurement, the coupling of the Higgs boson to vector bosons is assumed to be SM-like. Based on this assumption, values of κ_t outside the regions $[-0.9, -0.5]$ and $[1.0, 2.1]$ were excluded.

The production of H+b-jets is an important background in the tH and $t\bar{t}H$ categories targeting final states without leptons. This background includes processes where the additional b-jets do not originate from decays of particles produced during the hard scattering process, in particular ggH. It is associated with a large systematic uncertainty, as the accuracy of the modeling of the additional b-jets in MC simulation is not known. A measurement of the differential cross section of the number of b-jets, $N_{b\text{-jets}}$, was performed as a first attempt to constrain the H+b-jets background in the $H \to \gamma\gamma$ channel [4]. The differential cross section is measured for inclusive Higgs boson production, as the sensitivity of the measurement is limited by the large continuum background contribution. In the $N_{b\text{-jets}} = 1$ bin, which is enriched in H+b-jets, the measured cross section is approximately 40% higher than the cross section predicted by MC simulations. However, the measured and predicted values are in agreement within the uncertainties. The signal extraction uncertainty, in particular its statistical component, provides the dominant contribution to the total uncertainty and yields approximately 50%. An approach to decrease the large statistical uncertainty was investigated, which is based on the use of multivariate analysis techniques to discriminate between H+b-jets and the continuum background.

To conclude, based on the studies for tH production presented in thesis, dedicated tH categories were included in an ATLAS measurement for the first time. No indication for a negative sign of the coupling Y_t was found. However, the measurement of SM-like tH production in the $H \to \gamma\gamma$ channel alone is not yet possible, due to the small cross section of this process. A larger dataset and the study of additional Higgs boson decay channels will be necessary to be able to measure tH production. An extrapolation of the expected sensitivity to $tH(H \to \gamma\gamma)$ production with the analysis strategy presented in this thesis and $\sqrt{s} = 13\,\text{TeV}$ results in an expected limit of approximately $7\,\mu$ for the full Run 2 dataset and $4\,\mu$ for an assumed dataset of 450 fb^{-1}, collected during Run 2 and the future Run 3 of the LHC. Furthermore, first studies of the H+b-jets background in the $H \to \gamma\gamma$ channel were presented. This background is relevant for hadronic tH and $t\bar{t}H$ final states and is associated with a large systematic uncertainty. The measurement of the differential $N_{b\text{-jets}}$ cross section in the $H \to \gamma\gamma$ channel represents a first step towards the definition of a better-motivated uncertainty on the H+b-jets background. In general, the new concepts which are presented in this thesis represent a basis for future measurements of the rare tH and H+b-jets processes.

Appendix A

Photon ID efficiency measurement

A.1 Correction factor

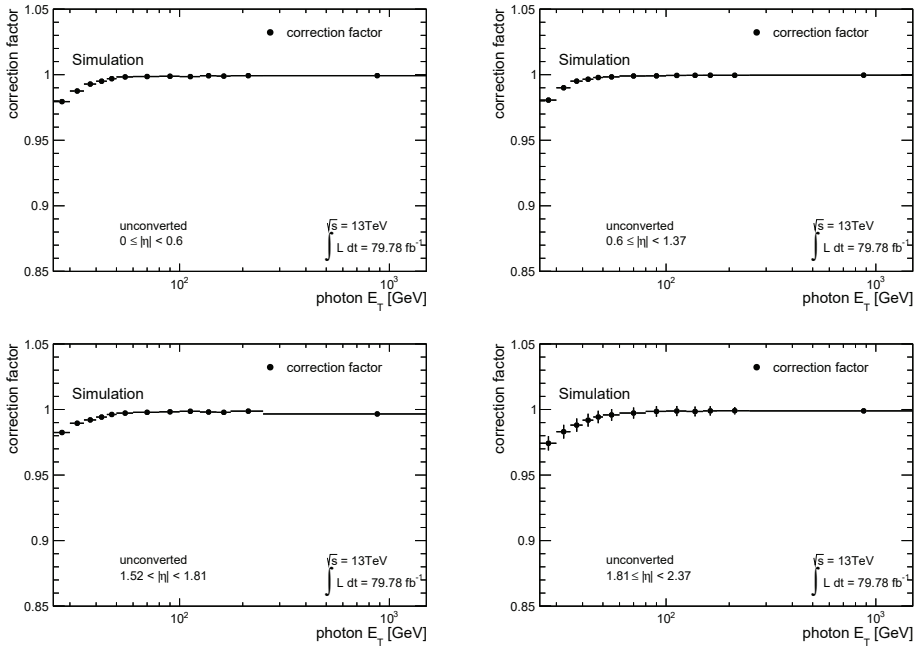


Figure A.1: Correction factors for unconverted photons, which are used to correct the measured *tight* ID efficiencies for the trigger and *loose* ID preselection requirements. The uncertainties include the statistical uncertainty and the systematic uncertainty due to differences of the *loose* ID efficiency in data and MC simulation.

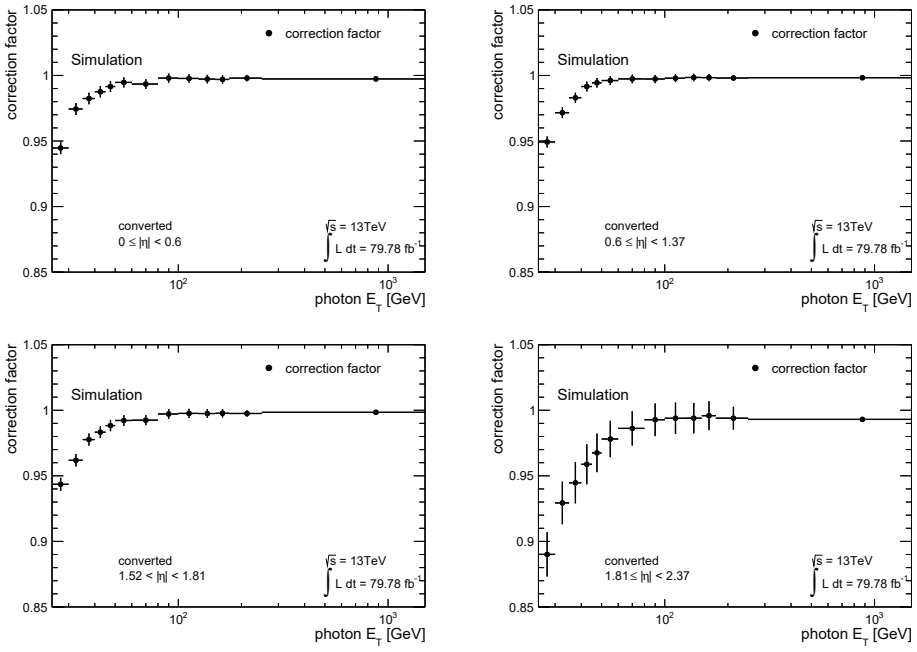


Figure A.2: Correction factors for converted photons, which are used to correct the measured *tight* ID efficiencies for the trigger and *loose* ID preselection requirements. The uncertainties include the statistical uncertainty and the systematic uncertainty due to differences of the *loose* ID efficiency in data and MC simulation.

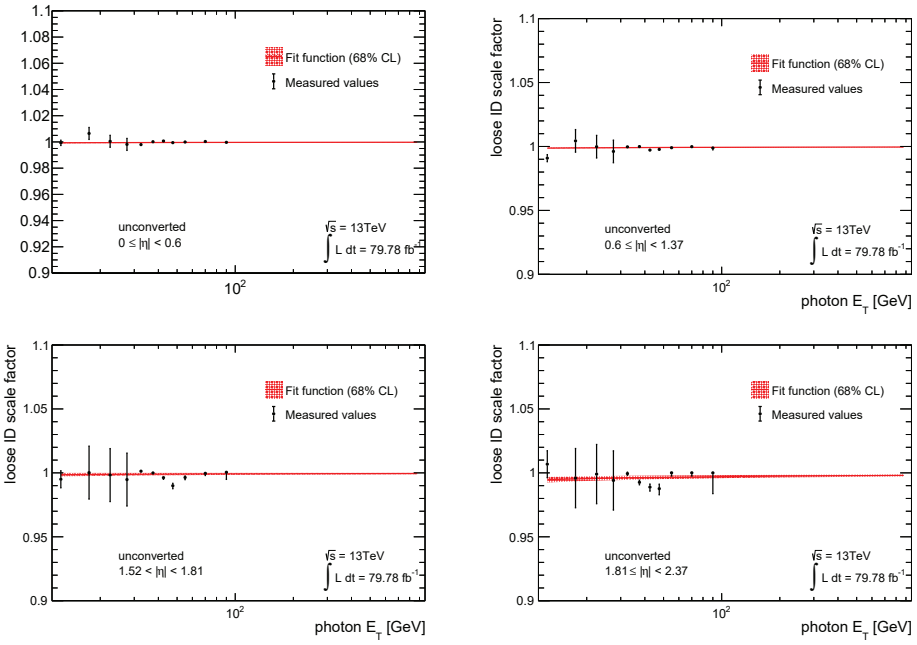


Figure A.3: Fit to loose ID SFs, which are evaluated with the Radiative Z method, for unconverted photons. The fit function is defined as $\alpha/log(E_T) + 1$, with α being the only free parameter. The largest difference of the fit function, including the 68% confidence level interval, from 1 is considered as systematic uncertainty.

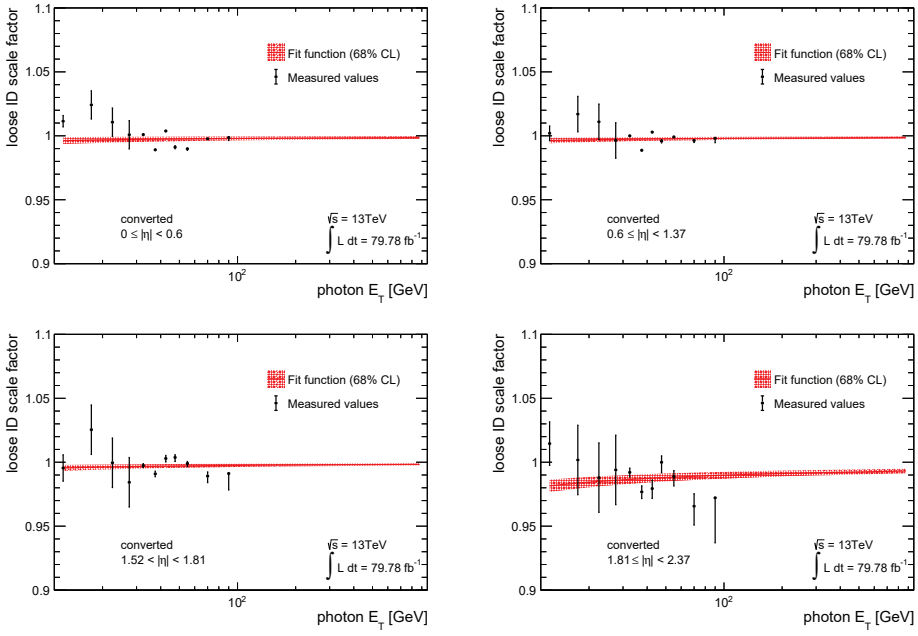


Figure A.4: Fit to loose ID SFs, which are evaluated with the Radiative Z method, for converted photons. The fit function is defined as $\alpha/log(E_T) + 1$, with α being the only free parameter. The largest difference of the fit function, including the 68% confidence level interval, from 1 is considered as systematic uncertainty.

A.2 Closure uncertainty

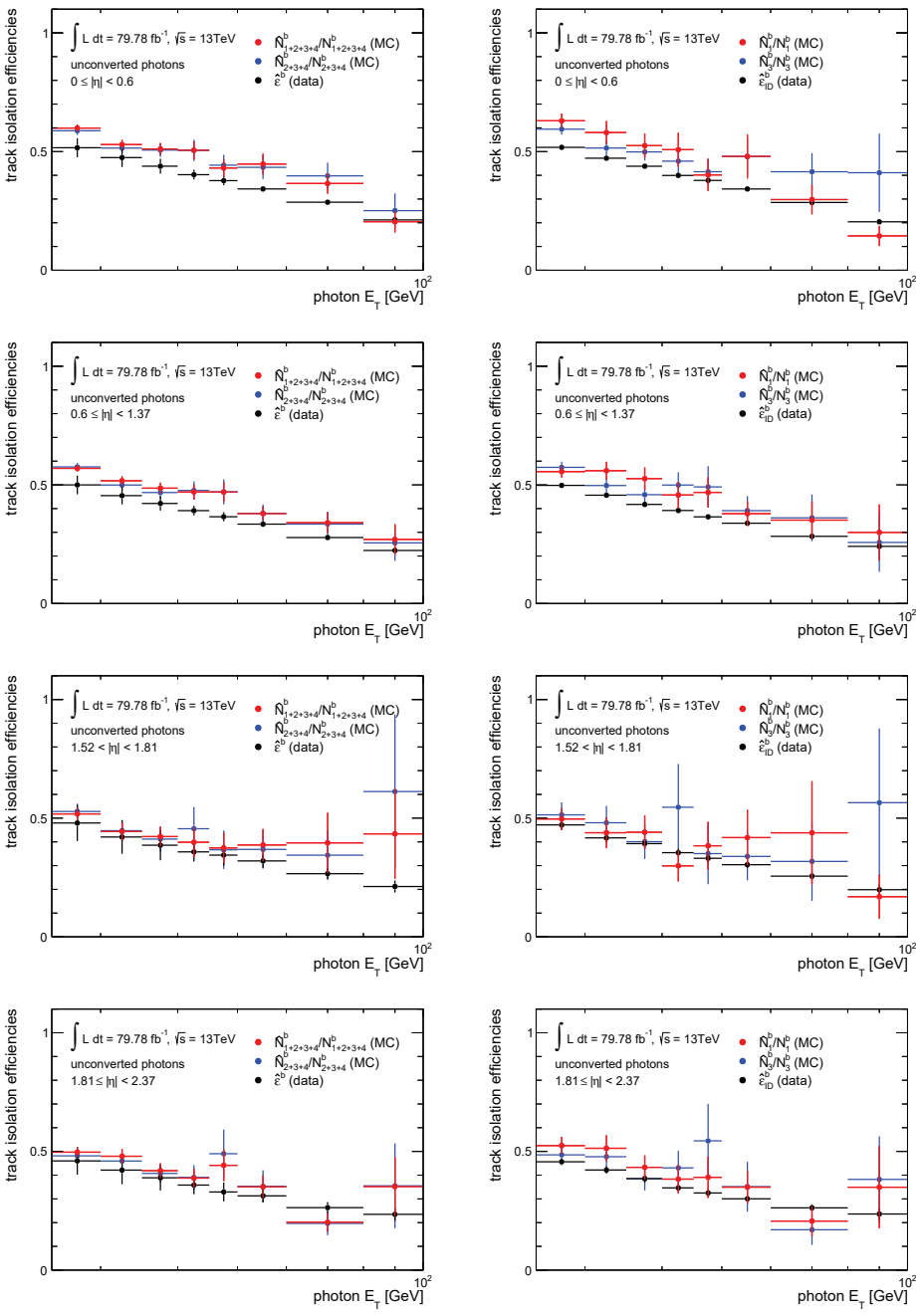


Figure A.5: Track isolation efficiencies for unconverted *fake* photons measured in data and estimated from MC simulation for the regions defined in data.

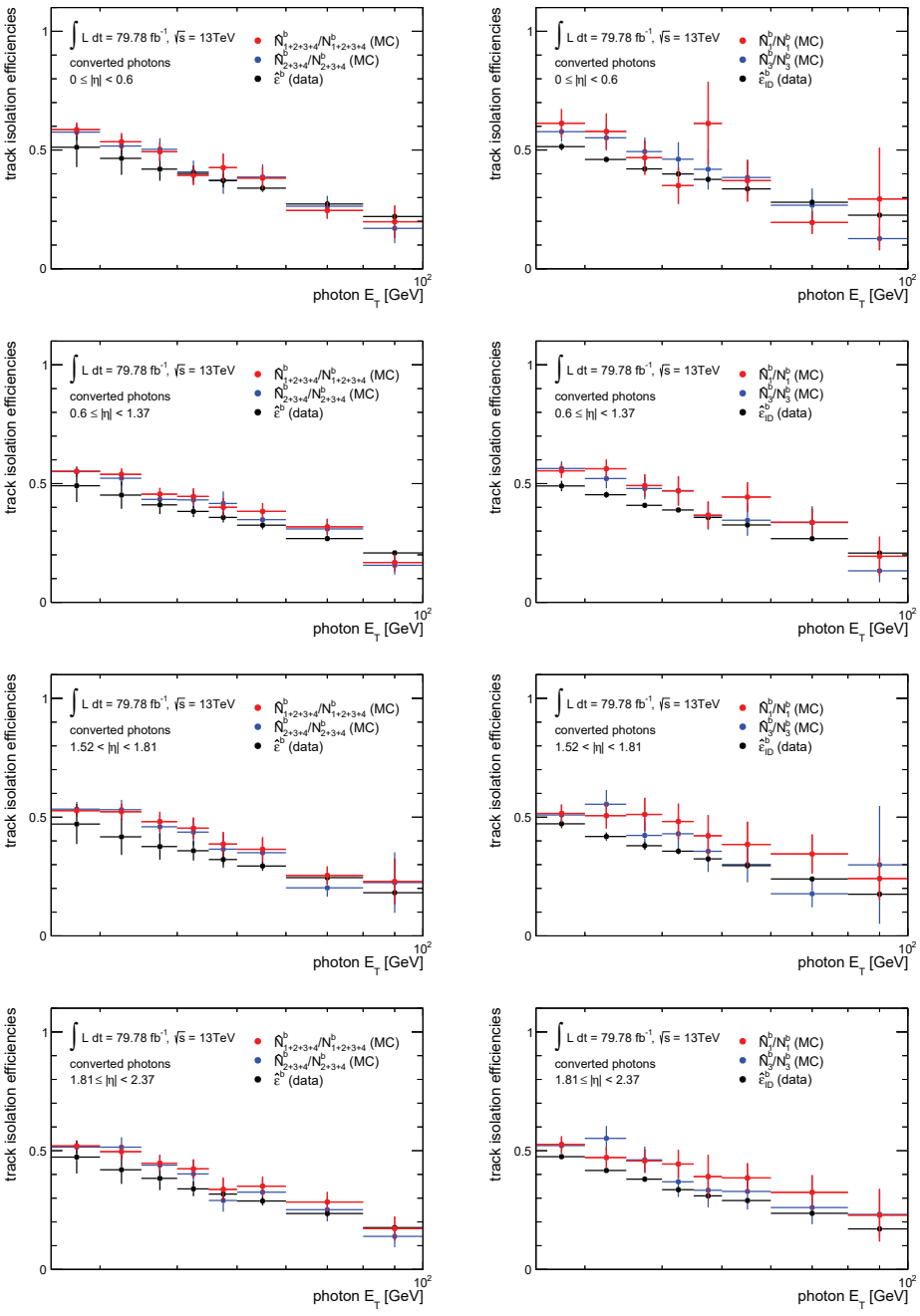


Figure A.6: Track isolation efficiencies for converted *fake* photons measured in data and estimated from MC simulation for the regions defined in data.

A.3 Fudge factor uncertainties

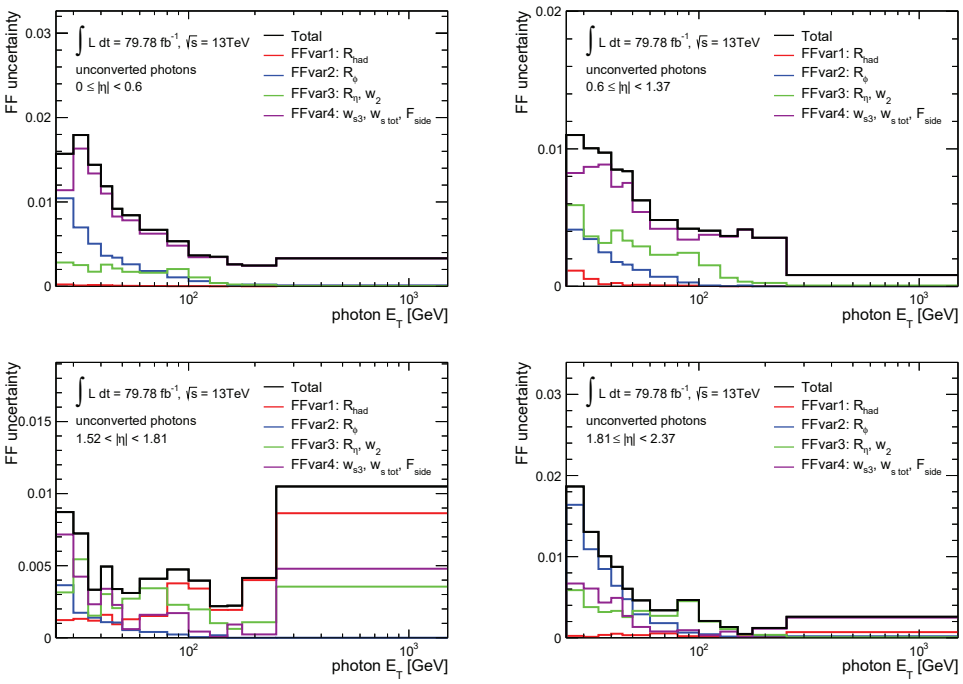


Figure A.7: Uncertainty resulting from the four different FF variations for unconverted photons. The total uncertainty is derived by adding the four uncertainties in quadrature.

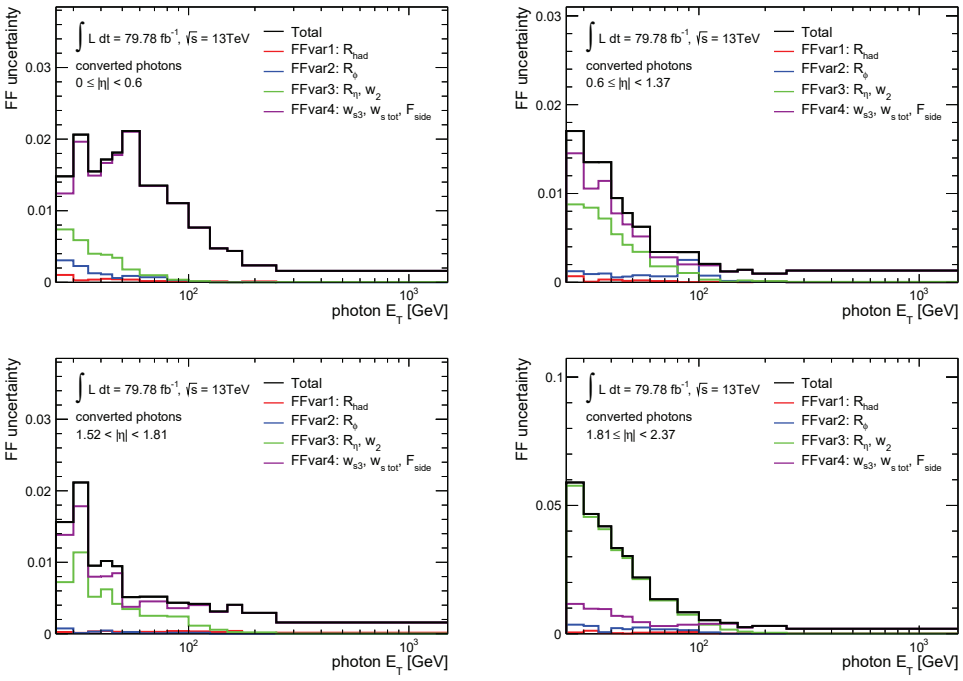


Figure A.8: Uncertainty resulting from the four different FF variations for converted photons. The total uncertainty is derived by adding the four uncertainties in quadrature.

A.4 Total uncertainty

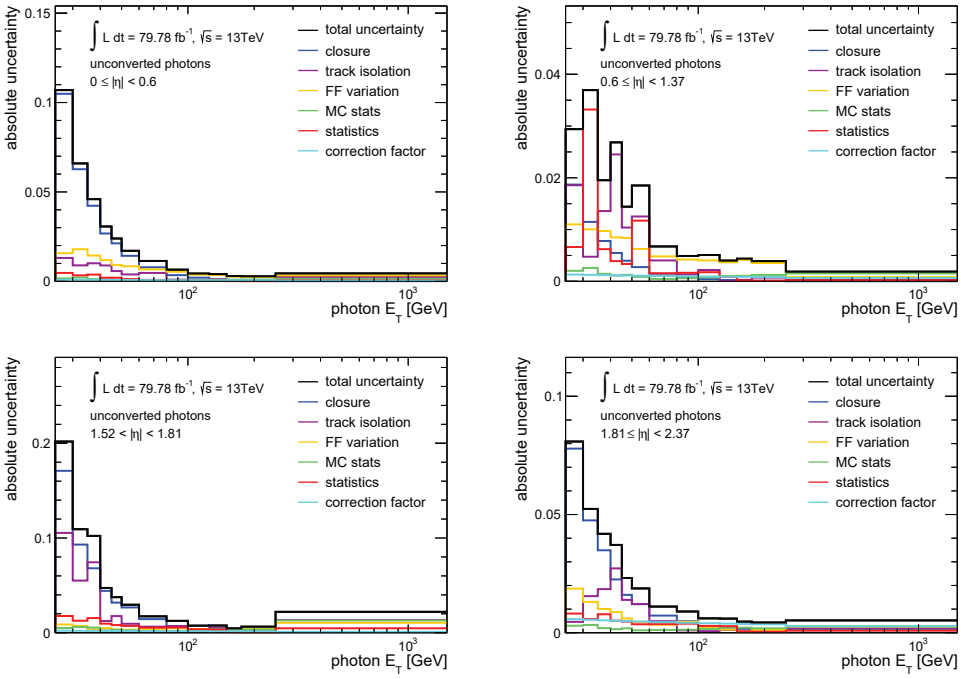


Figure A.9: Uncertainty breakdown for unconverted photons in the four different $|\eta|$ regions of the photon ID efficiency measurement.

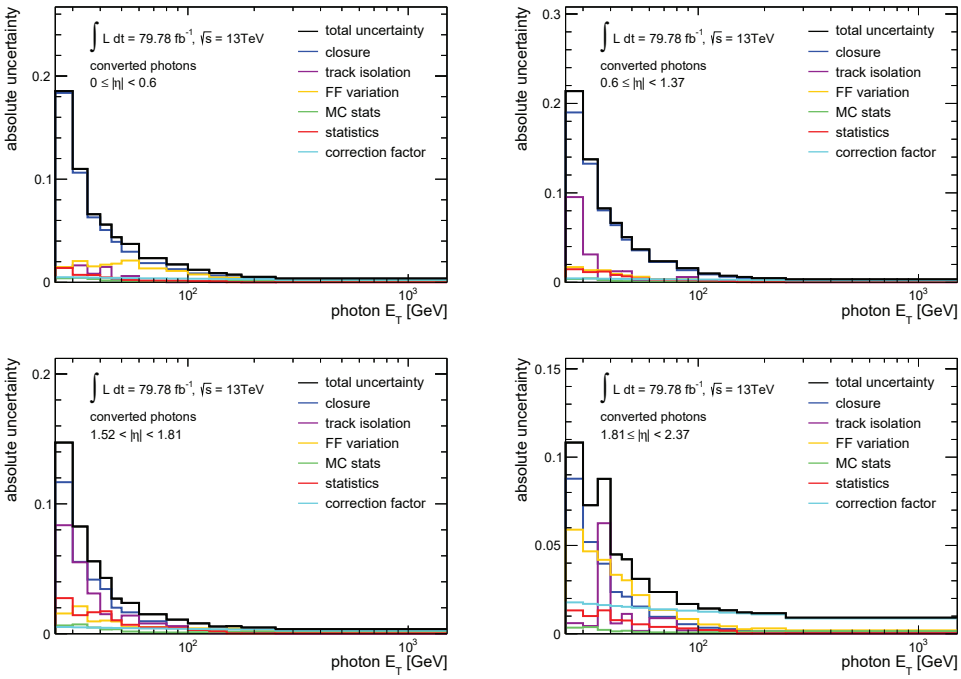


Figure A.10: Uncertainty breakdown for converted photons in the four different $|\eta|$ regions of the photon ID efficiency measurement.

A.5 *Tight* ID efficiencies and scale factors

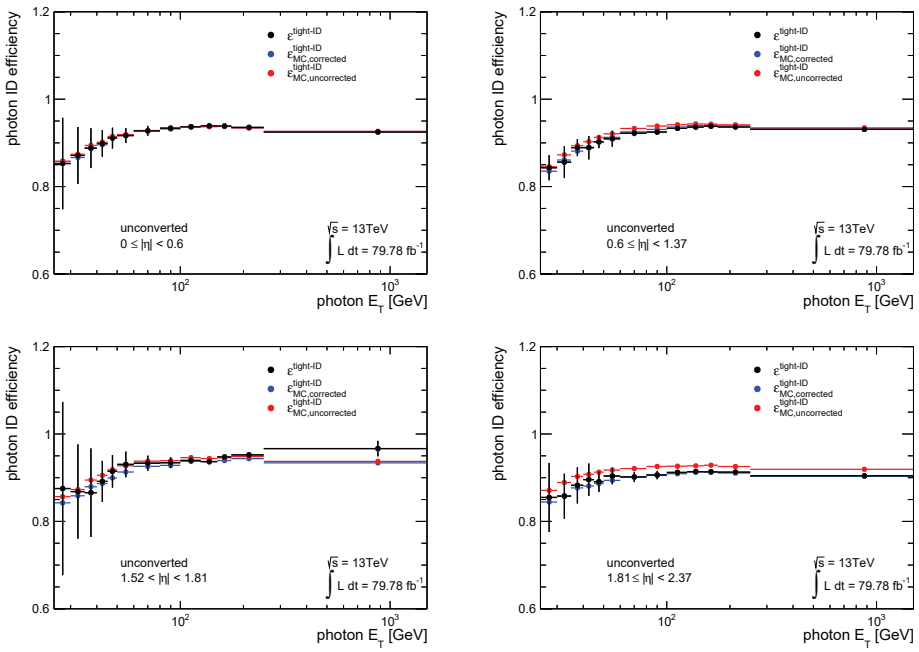


Figure A.11: Photon ID efficiencies measured in 2015-2017 data with the Matrix Method for unconverted photons in four different $|\eta|$ regions. The error bars include all considered uncertainties. The efficiencies measured in data are compared to the efficiencies of the prompt photon MC sample with and without corrections applied to the shower shape variables.

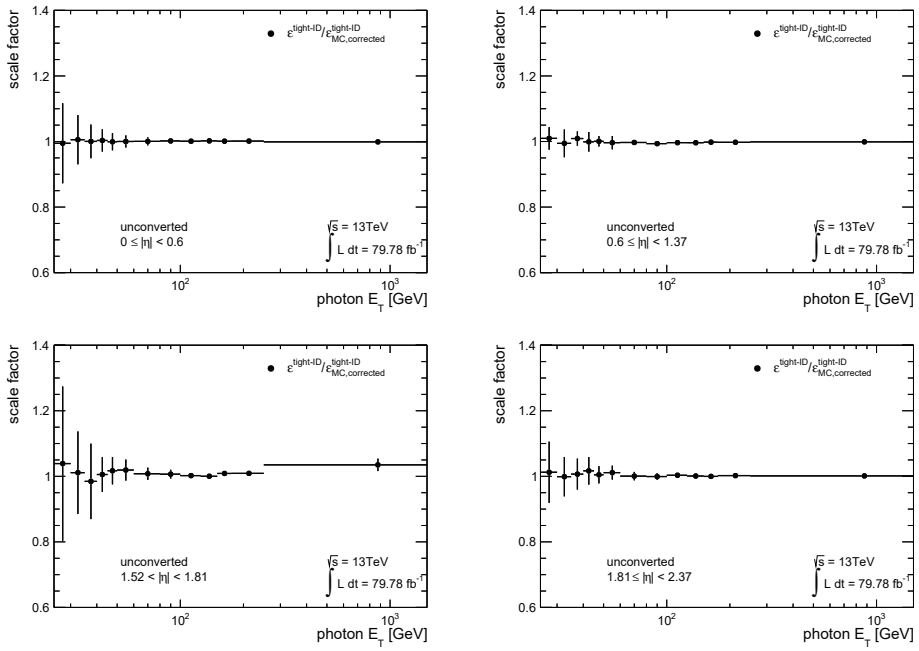


Figure A.12: Scale factors measured in 2015-2017 data with the Matrix Method for unconverted photons in four different $|\eta|$ regions.

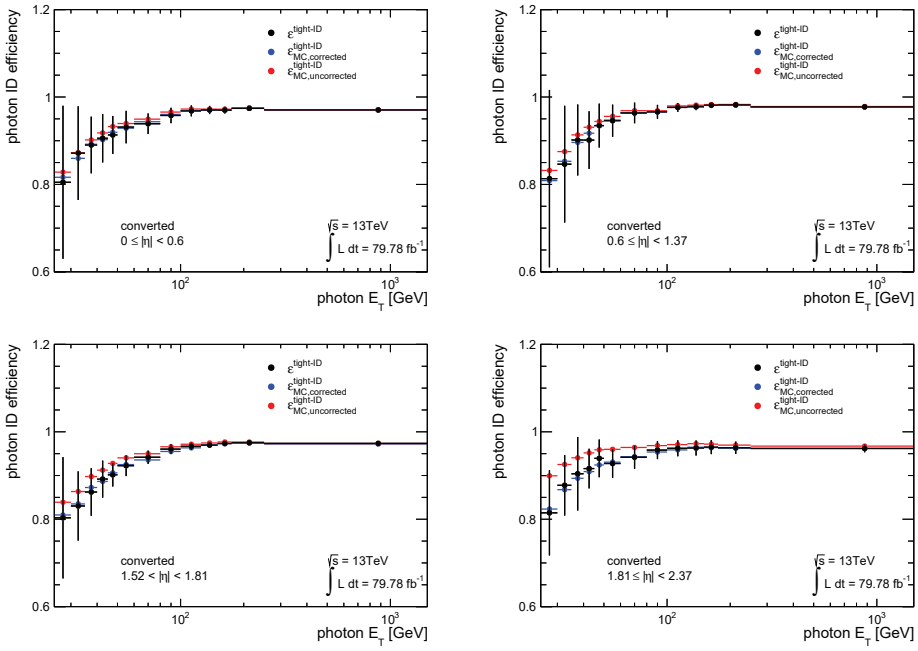


Figure A.13: Photon ID efficiencies measured in 2015-2017 data with the Matrix Method for converted photons in four different $|\eta|$ regions. The error bars include all considered uncertainties. The efficiencies measured in data are compared to the efficiencies of the prompt photon MC sample with and without corrections applied to the shower shape variables.

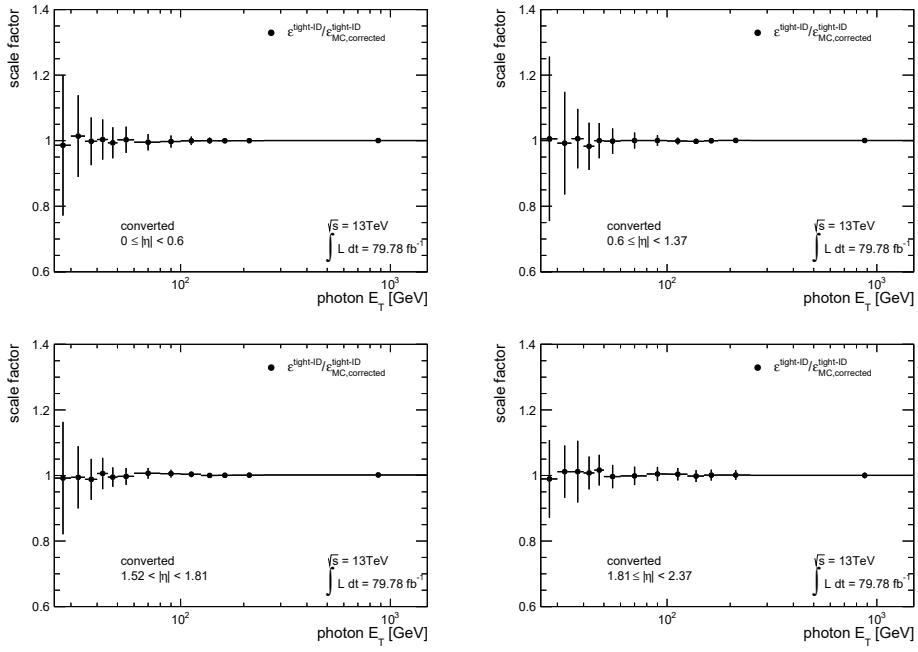


Figure A.14: Scale factors measured in 2015-2017 data with the Matrix Method for converted photons in four different $|\eta|$ regions.

Appendix B

Optimization of tH categories

B.1 Comparison of kinematics for $\kappa_t = \pm 1$

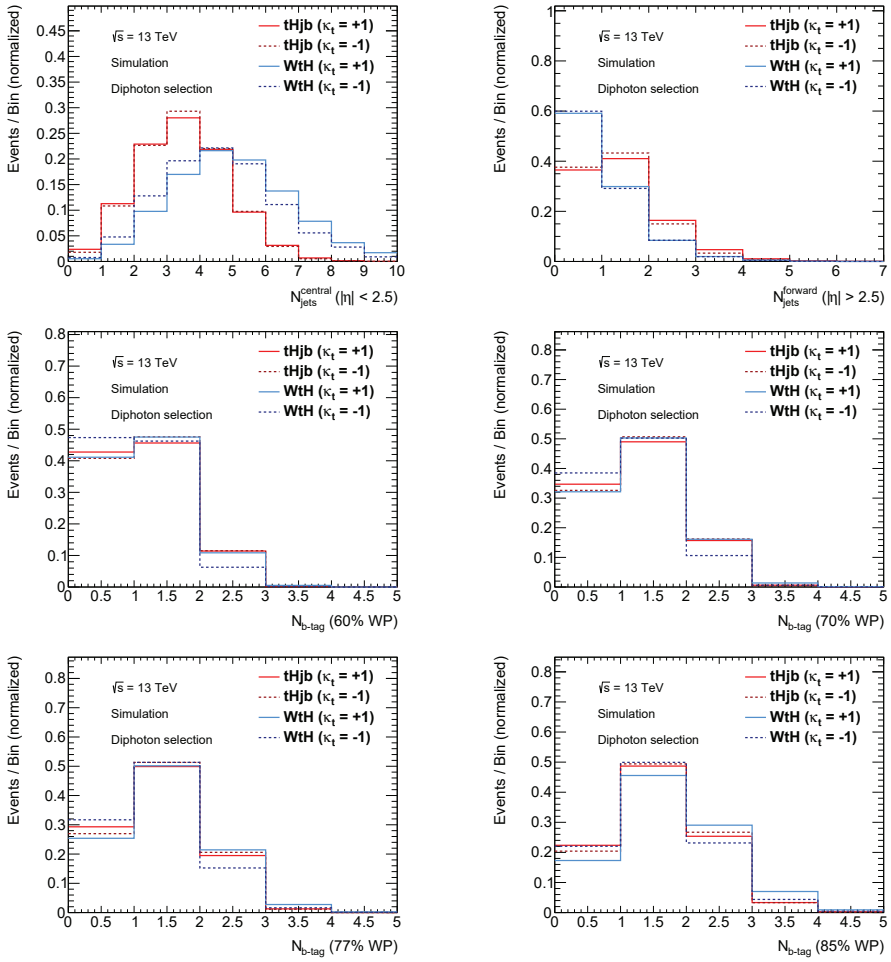


Figure B.1: Comparison of the distributions of the number of central jets, forward jets and b-tagged jets for coupling strength modifiers $\kappa_t = \pm 1$ and $tHjb$ and WtH production.

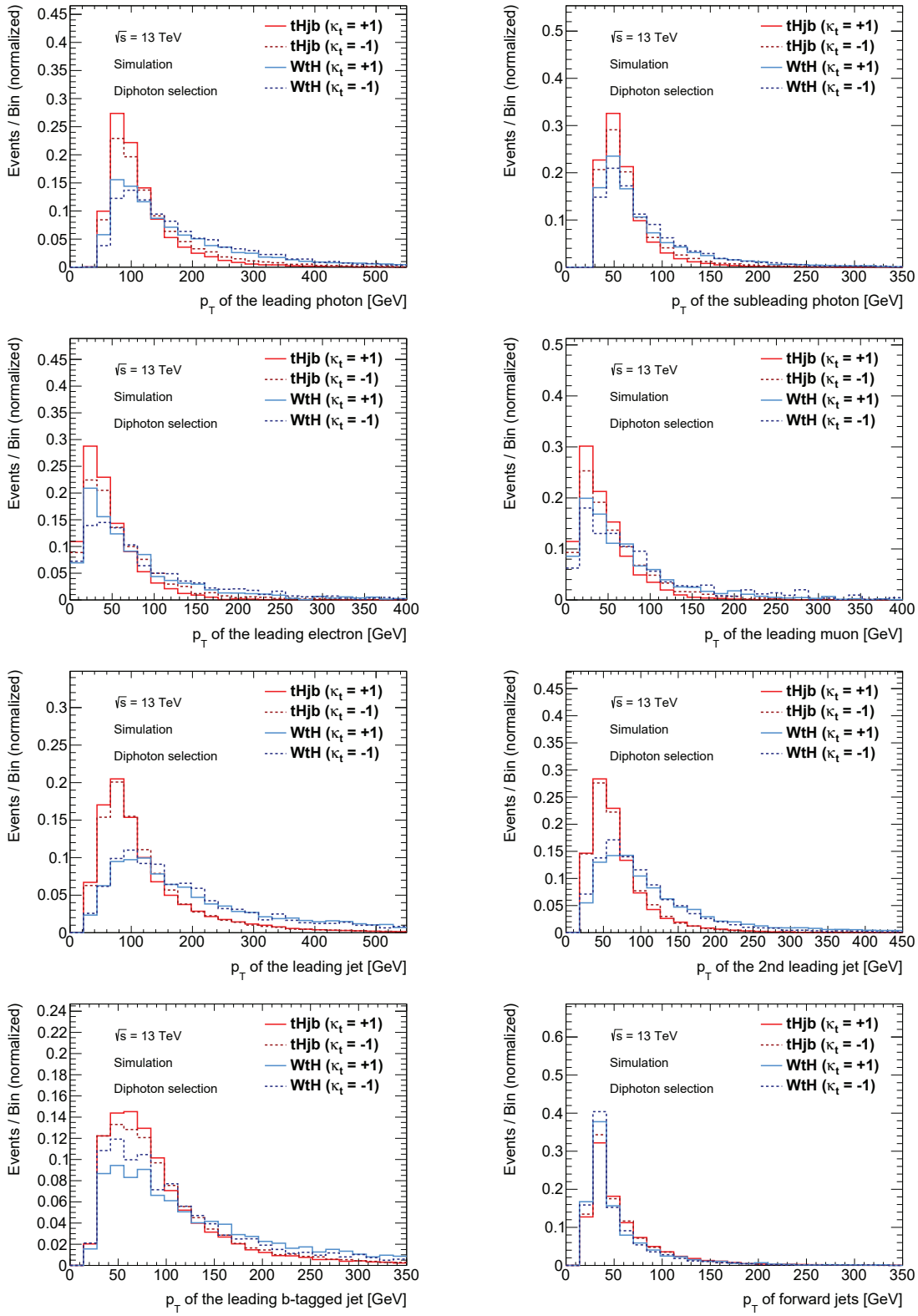


Figure B.2: Comparison of the p_{T} distributions of the two p_{T}-leading photons, the p_{T}-leading electrons and muons, the two p_{T}-leading jets, the p_{T}-leading b-tagged jet and forward jets for coupling strength modifiers $\kappa_t = \pm 1$ and $tHjb$ and WtH production.

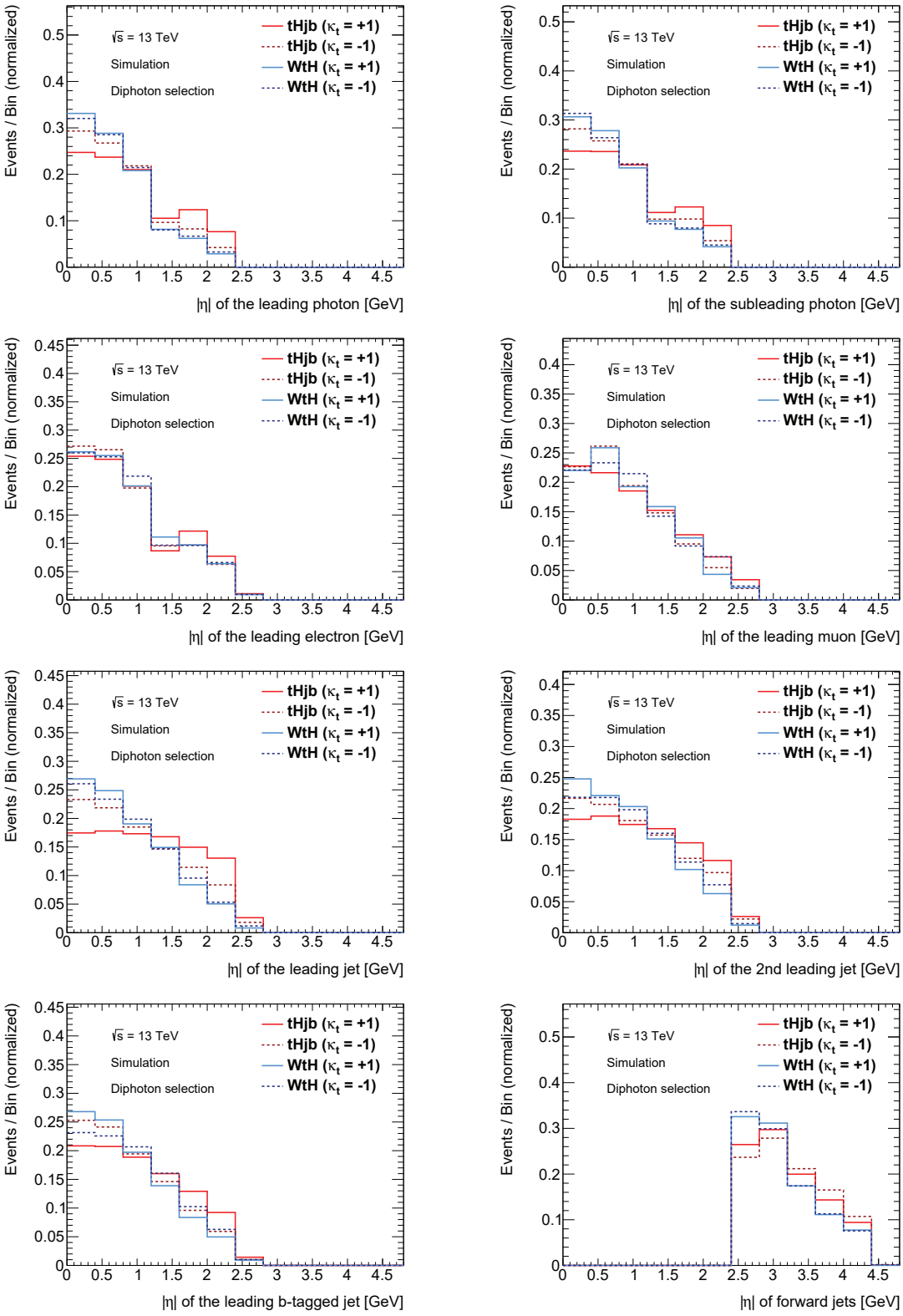


Figure B.3: Comparison of the $|\eta|$ distributions of the two p_{T}-leading photons, the p_{T}-leading electrons and muons, the two p_{T}-leading jets, the p_{T}-leading b-tagged jet and forward jets for coupling strength modifiers $\kappa_t = \pm 1$ and $tHjb$ and WtH production.

B.2 Leptonic categories

Table B.1: Comparison of the expected event yields for the signal S, the Higgs boson background B_{Higgs}, the continuum background B_{ContBG} and the expected limit for the different leptonic tH categories for an integrated luminosity of $35\,\text{fb}^{-1}$. The continuum background is estimated using a leptonic reference event selection, resulting in a more accurate estimate with respect to the numbers presented in Table 9.4.

Event selection	S	B_{Higgs}	S/B_{Higgs}	B_{ContBG}	Exp. limit [μ]
Baseline leptonic	0.28	1.6	0.17	71	27.4
Cat($N_{\text{jets}}^{\text{forward}} = 0$)	0.12	1.1	0.11	51	55.4
Cat($N_{\text{jets}}^{\text{forward}} \geq 1$)	0.19	0.97	0.19	22	27.4
Cat($\lvert\eta(j_{1,untagged})\rvert > 2.5$)	0.13	0.38	0.34	13	32.3

Table B.2: Efficiencies and number of expected events for the different Higgs boson production processes for the leptonic baseline event selection.

Process	Efficiency	N_{exp} for $35\,\text{fb}^{-1}$
ggH	0.0035%	0.052
VBF	0.0039%	0.0045
WH	0.64%	0.23
ZH	0.26%	0.063
$t\bar{t}H$	8.4%	1.3
$tHjb$	11%	0.23
WtH	9.6%	0.051

Table B.3: Efficiencies and number of expected events for the different Higgs boson production processes for the leptonic $N_{\text{jets}}^{\text{forward}} = 0$ category.

Process	Efficiency	N_{exp} for $35\,\text{fb}^{-1}$
ggH	0.0023%	0.035
VBF	0.0018%	0.0021
WH	0.5%	0.18
ZH	0.19%	0.047
$t\bar{t}H$	5.1%	0.79
$tHjb$	4.1%	0.087
WtH	5.6%	0.03

Table B.4: Efficiencies and number of expected events for the different Higgs boson production processes for the leptonic $N_{\text{jets}}^{\text{forward}} \geq 1$ category.

Process	Efficiency	N_{exp} for $35\,\text{fb}^{-1}$
ggH	0.0012%	0.017
VBF	0.0028%	0.0032
WH	0.14%	0.051
ZH	0.075%	0.018
$t\bar{t}H$	5.7%	0.88
$tHjb$	7.3%	0.15
WtH	6.1%	0.032

Table B.5: Efficiencies and number of expected events for the different Higgs boson production processes for the leptonic $|\eta(j_{1,untagged})| > 2.5$ category.

Process	Efficiency	N_{exp} for $35\,\text{fb}^{-1}$
ggH	0.0014%	0.02
VBF	0.0026%	0.003
WH	0.078%	0.028
ZH	0.052%	0.013
$t\bar{t}H$	2.1%	0.32
$tHjb$	5.7%	0.12
WtH	1.9%	0.01

B.3 Definition of all categories

Category	Selection
tH lep 0fwd	$N_{\text{lep}} = 1$, $N_{\text{jets}}^{\text{cen}} \leq 3$, $N_{b-\text{tag}} \geq 1$, $N_{\text{jets}}^{\text{fwd}} = 0$ ($p_{\text{T}}^{\text{jet}} > 25\,\text{GeV}$)
tH lep 1fwd	$N_{\text{lep}} = 1$, $N_{\text{jets}}^{\text{cen}} \leq 4$, $N_{b-\text{tag}} \geq 1$, $N_{\text{jets}}^{\text{fwd}} \geq 1$ ($p_{\text{T}}^{\text{jet}} > 25\,\text{GeV}$)
ttH lep	$N_{\text{lep}} \geq 1$, $N_{\text{jets}}^{\text{cen}} \geq 2$, $N_{b-\text{tag}} \geq 1$, $Z_{\ell\ell}$ veto ($p_{\text{T}}^{\text{jet}} > 25\,\text{GeV}$)
ttH had BDT1	$N_{\text{lep}} = 0$, $N_{\text{jets}} \geq 3$, $N_{b-\text{tag}} \geq 1$, $\text{BDT}_{\text{ttH}} > 0.92$
ttH had BDT2	$N_{\text{lep}} = 0$, $N_{\text{jets}} \geq 3$, $N_{b-\text{tag}} \geq 1$, $0.83 < \text{BDT}_{\text{ttH}} < 0.92$
ttH had BDT3	$N_{\text{lep}} = 0$, $N_{\text{jets}} \geq 3$, $N_{b-\text{tag}} \geq 1$, $0.79 < \text{BDT}_{\text{ttH}} < 0.83$
ttH had BDT4	$N_{\text{lep}} = 0$, $N_{\text{jets}} \geq 3$, $N_{b-\text{tag}} \geq 1$, $0.52 < \text{BDT}_{\text{ttH}} < 0.79$
tH had 4j1b	$N_{\text{lep}} = 0$, $N_{\text{jets}}^{\text{cen}} = 4$, $N_{b-\text{tag}} = 1$ ($p_{\text{T}}^{\text{jet}} > 25\,\text{GeV}$)
tH had 4j2b	$N_{\text{lep}} = 0$, $N_{\text{jets}}^{\text{cen}} = 4$, $N_{b-\text{tag}} \geq 2$ ($p_{\text{T}}^{\text{jet}} > 25\,\text{GeV}$)
VH dilep	$N_{\text{lep}} \geq 2$, $70\,\text{GeV} \leq m_{\ell\ell} \leq 110\,\text{GeV}$
VH lep High	$N_{\text{lep}} = 1$, $\lvert m_{e\gamma} - 89\,\text{GeV}\rvert > 5\,\text{GeV}$, $p_{\text{T}}^{\ell+E_{\text{T}}^{\text{miss}}} > 150\,\text{GeV}$
VH lep Low	$N_{\text{lep}} = 1$, $\lvert m_{e\gamma} - 89\,\text{GeV}\rvert > 5\,\text{GeV}$, $p_{\text{T}}^{\ell+E_{\text{T}}^{\text{miss}}} < 150\,\text{GeV}$, $E_{\text{T}}^{\text{miss}}$ significance > 1
VH MET High	$150\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 250\,\text{GeV}$, $E_{\text{T}}^{\text{miss}}$ significance > 9 or $E_{\text{T}}^{\text{miss}} > 250\,\text{GeV}$
VH MET Low	$80\,\text{GeV} < E_{\text{T}}^{\text{miss}} < 150\,\text{GeV}$, $E_{\text{T}}^{\text{miss}}$ significance > 8
jet BSM	$p_{\text{T,j1}} > 200\,\text{GeV}$
VH had tight	$60\,\text{GeV} < m_{\text{jj}} < 120\,\text{GeV}$, $\text{BDT}_{\text{VH}} > 0.78$
VH had loose	$60\,\text{GeV} < m_{\text{jj}} < 120\,\text{GeV}$, $0.35 < \text{BDT}_{\text{VH}} < 0.78$
VBF tight, high p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\text{j1}} + \eta_{\text{j2}})\rvert < 5$, $p_{\text{T}}^{Hjj} > 25\,\text{GeV}$, $\text{BDT}_{\text{VBF}} > 0.47$
VBF loose, high p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\text{j1}} + \eta_{\text{j2}})\rvert < 5$, $p_{\text{T}}^{Hjj} > 25\,\text{GeV}$, $-0.32 < \text{BDT}_{\text{VBF}} < 0.47$
VBF tight, low p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\text{j1}} + \eta_{\text{j2}})\rvert < 5$, $p_{\text{T}}^{Hjj} < 25\,\text{GeV}$, $\text{BDT}_{\text{VBF}} > 0.87$
VBF loose, low p_{T}^{Hjj}	$\lvert\Delta\eta_{jj}\rvert > 2$, $\lvert\eta_{\gamma\gamma} - 0.5(\eta_{\text{j1}} + \eta_{\text{j2}})\rvert < 5$, $p_{\text{T}}^{Hjj} < 25\,\text{GeV}$, $0.26 < \text{BDT}_{\text{VBF}} < 0.87$
ggH 2J BSM	≥ 2 jets, $p_{\text{T}}^{\gamma\gamma} \geq 200\,\text{GeV}$
ggH 2J High	≥ 2 jets, $p_{\text{T}}^{\gamma\gamma} \in [120, 200]$ GeV
ggH 2J Med	≥ 2 jets, $p_{\text{T}}^{\gamma\gamma} \in [60, 120]$ GeV
ggH 2J Low	≥ 2 jets, $p_{\text{T}}^{\gamma\gamma} \in [0, 60]$ GeV
ggH 1J BSM	$= 1$ jet, $p_{\text{T}}^{\gamma\gamma} \geq 200\,\text{GeV}$
ggH 1J High	$= 1$ jet, $p_{\text{T}}^{\gamma\gamma} \in [120, 200]$ GeV
ggH 1J Med	$= 1$ jet, $p_{\text{T}}^{\gamma\gamma} \in [60, 120]$ GeV
ggH 1J Low	$= 1$ jet, $p_{\text{T}}^{\gamma\gamma} \in [0, 60]$ GeV
ggH 0J Fwd	$= 0$ jets, one photon with $\lvert\eta\rvert > 0.95$
ggH 0J Cen	$= 0$ jets, two photons with $\lvert\eta\rvert \leq 0.95$

Figure B.4: The definition of the 31 categories in the measurement of Higgs boson couplings in the $H \to \gamma\gamma$ channel using $36.1\,\text{fb}^{-1}$ [2], targeting tH, $t\bar{t}H$, WH/ZH, VBF and ggH final states. The categories are filled in the order shown here, so that they are orthogonal by construction.

Appendix C

Background modeling for tH and $t\bar{t}H$ categories

C.1 $m_{\gamma\gamma}$ distributions of the TI sidebands and the background templates

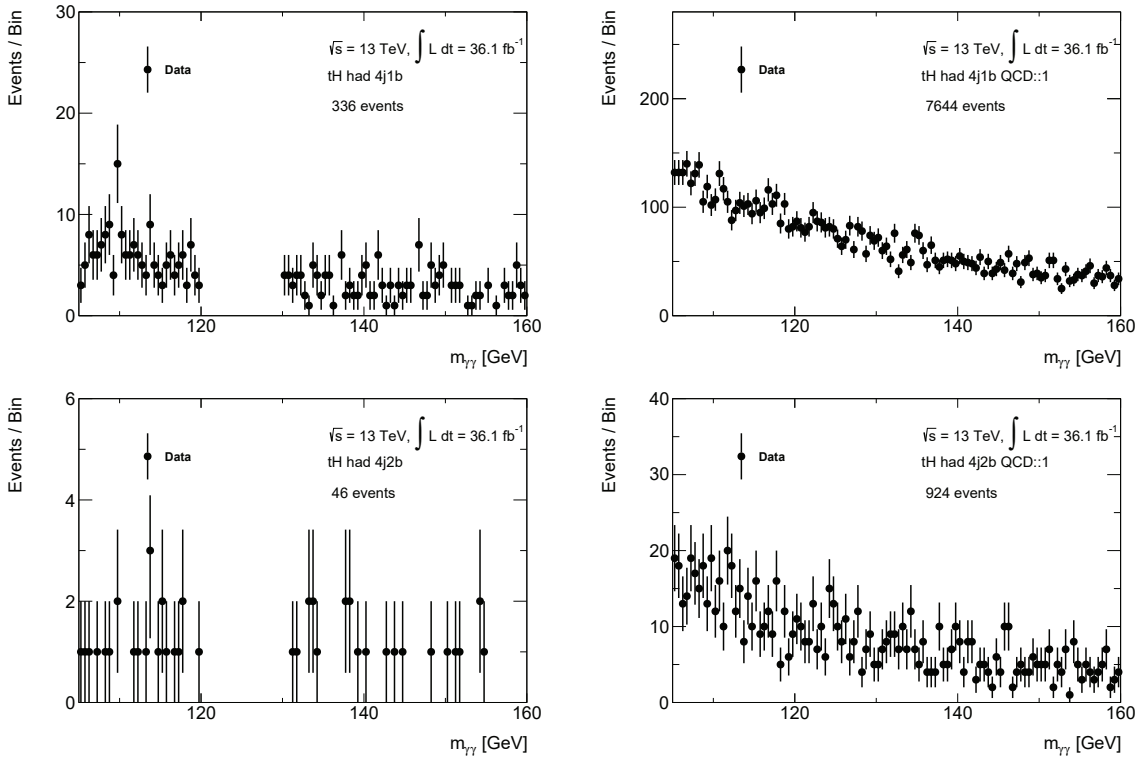


Figure C.1: Comparison of the $m_{\gamma\gamma}$ distributions of a) the TI sidebands and b) the QCD::1 background templates for the hadronic cut-based tH categories. Before performing the S+B fit, the background templates are normalized to the number of TI sideband events of the $m_{\gamma\gamma}$ distribution of the respective category.

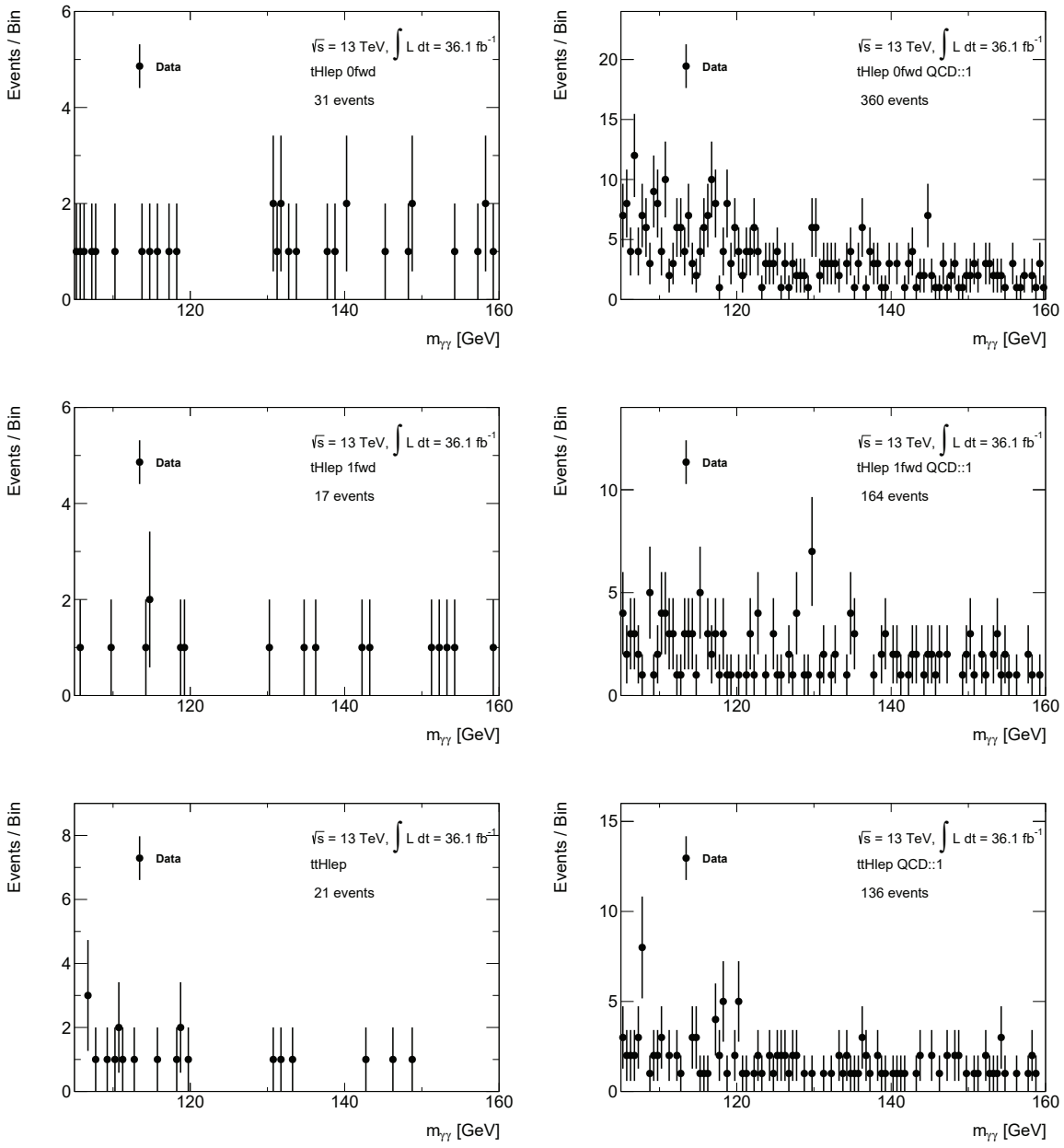


Figure C.2: Comparison of the $m_{\gamma\gamma}$ distributions of a) the TI sidebands and b) the QCD::1 background templates for the leptonic tH and $t\bar{t}H$ categories. Before performing the S+B fit, the background templates are normalized to the number of TI sideband events of the $m_{\gamma\gamma}$ distribution of the respective category.

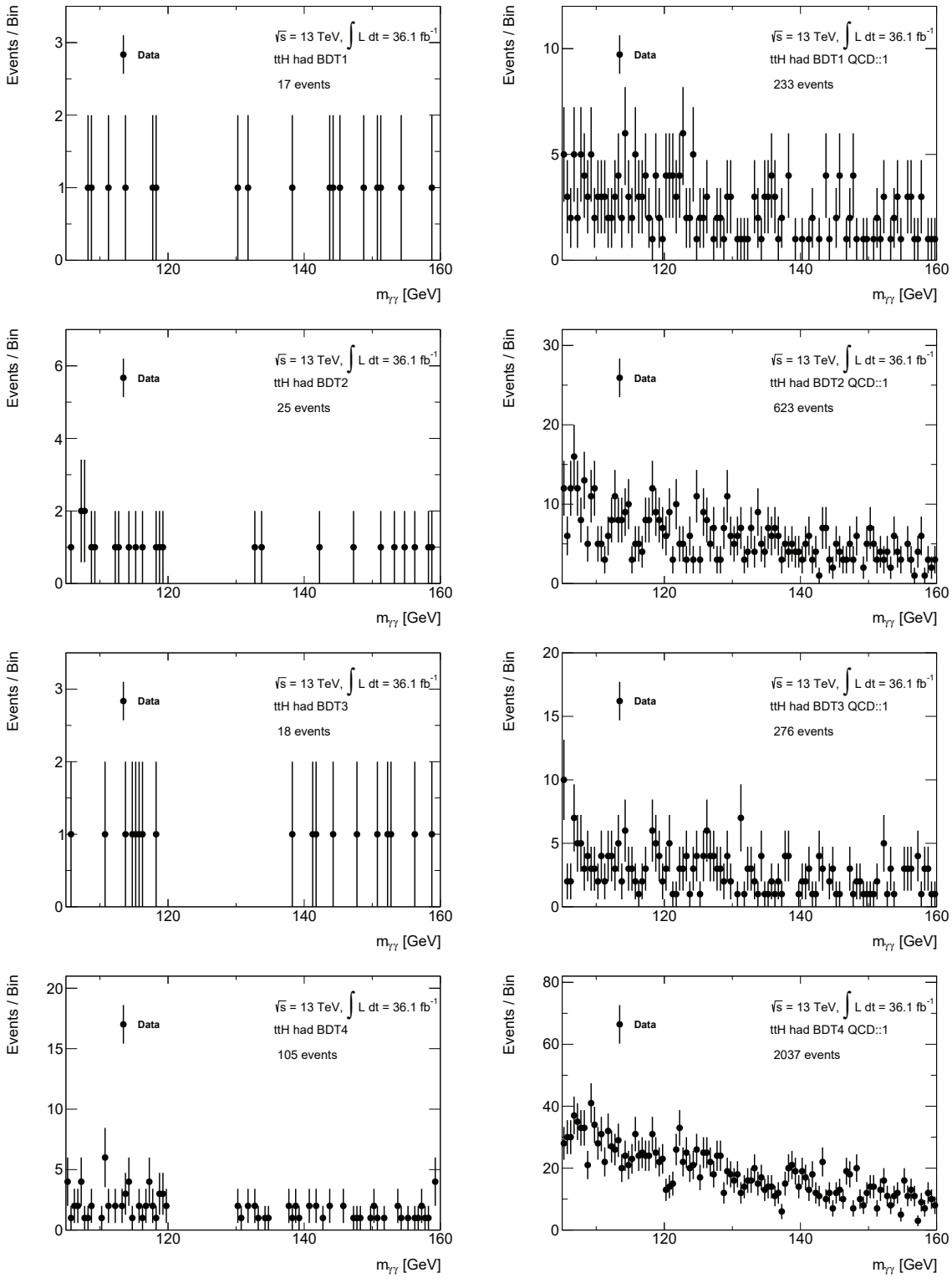


Figure C.3: Comparison of the $m_{\gamma\gamma}$ distributions of a) the TI sidebands and b) the QCD::1 background templates for the hadronic $t\bar{t}H$ BDT categories. Before performing the S+B fit, the background templates are normalized to the number of TI sideband events of the $m_{\gamma\gamma}$ distribution of the respective category.

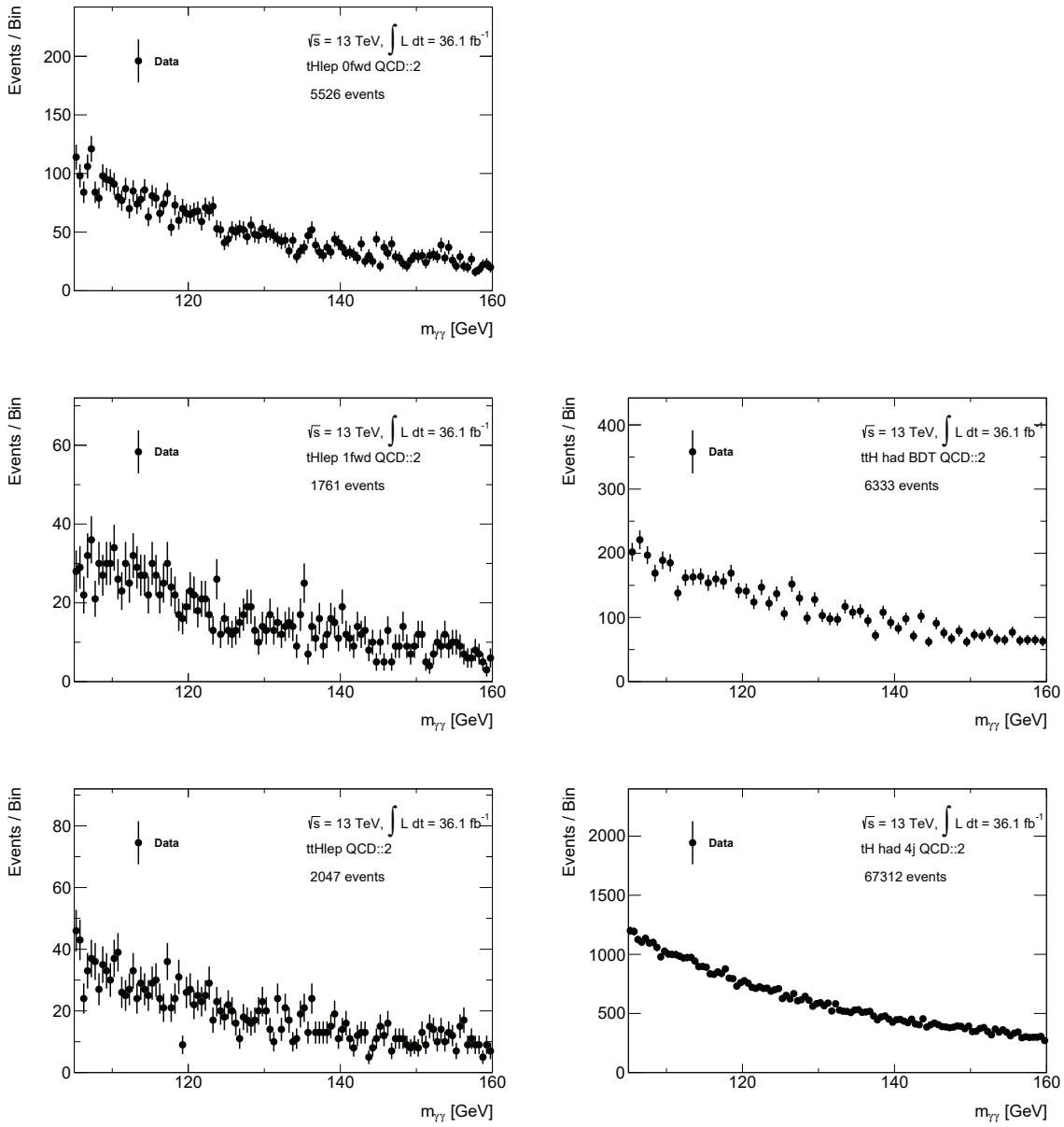


Figure C.4: QCD::2 background templates, as defined in Section 9.5, for the leptonic $t\bar{t}H$ and tH categories, the combined hadronic $t\bar{t}H$ BDT categories and the combined hadronic cut-based tH categories. These background templates are used to derive the background models for the categories described in Section 9.4. Before performing the S+B fit, the background templates are normalized to the number of TI sideband events of the $m_{\gamma\gamma}$ distribution of the respective category.

C.2 Validation of the data-driven background modeling strategy

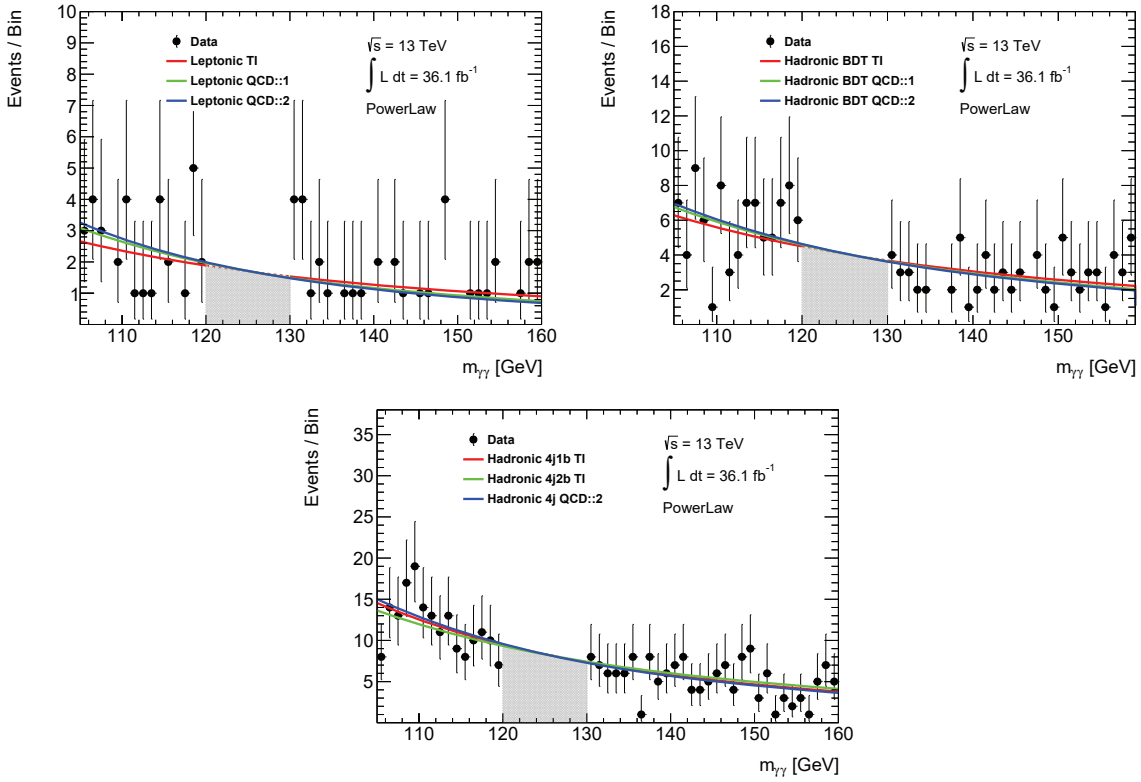


Figure C.5: Comparison of the PowerLaw functions fitted to the TI sideband events and the QCD::1 and QCD::2 background templates for the combined leptonic tH+$t\bar{t}H$ categories, the combined hadronic $t\bar{t}H$ BDT categories and the cut-based hadronic tH categories.

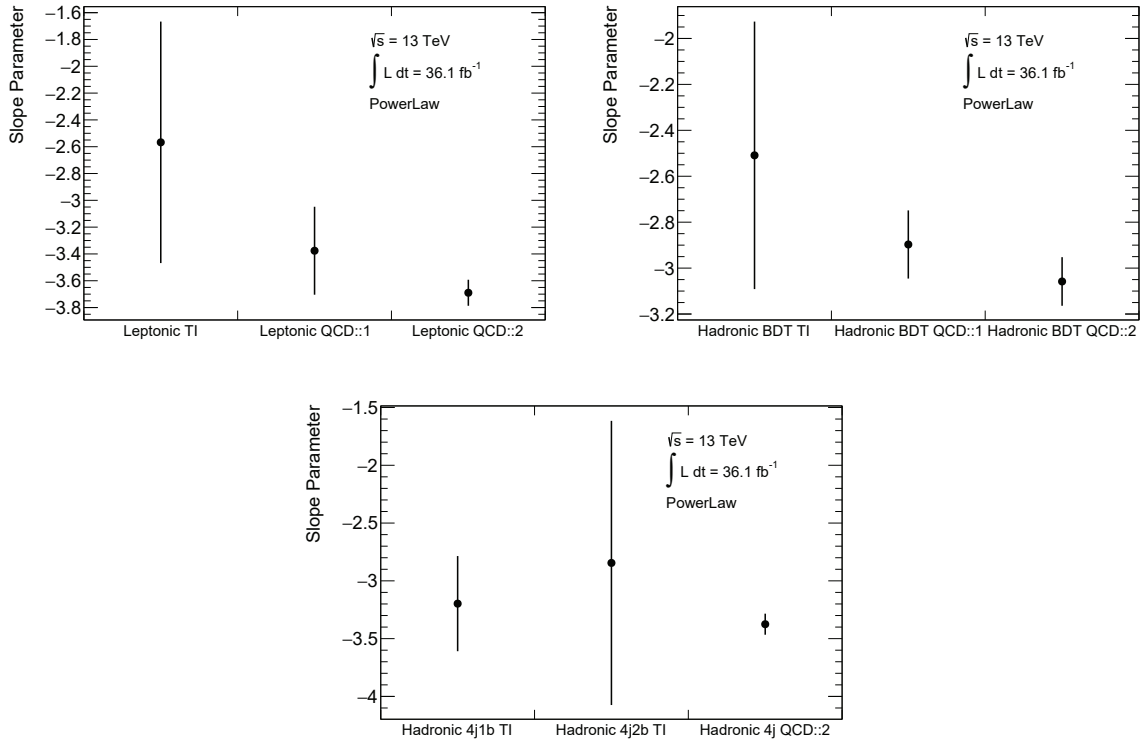


Figure C.6: Comparison of the slope parameters, including statistical uncertainties, of the PowerLaw functions fitted to the TI sideband events and the QCD::1 and QCD::2 background templates for the combined leptonic tH+$t\bar{t}H$ categories, the combined hadronic $t\bar{t}H$ BDT categories and the cut-based hadronic tH categories.

C.3 Fitted spurious signal

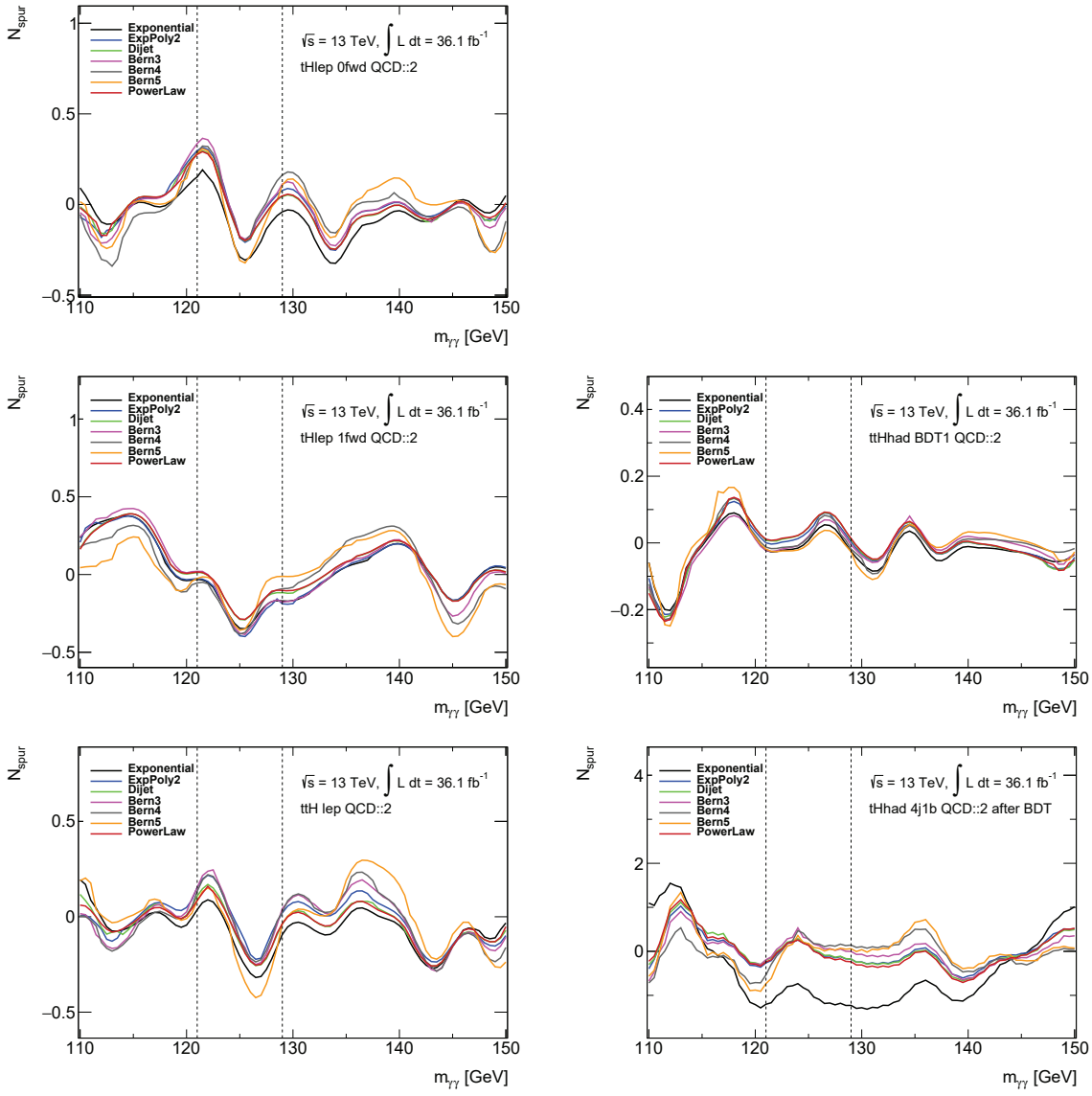


Figure C.7: Fitted spurious signal as a function of $m_{\gamma\gamma}$ for a) the QCD::2 background templates of the leptonic tH and $t\bar{t}H$ categories and b) the QCD::2 background templates of the hadronic $t\bar{t}H$ had BDT1 and the tH had 4j1b category. As the same background templates are used for the other hadronic categories, the shape of the fitted signal looks similar for these but the y-axis has a different scale. Hence, only one example plot is shown for each of these background templates.

C.4 Detailed results of the spurious signal tests

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
PowerLaw	12.9	28.5	0.29	1	PASS
Exponential	-15.0	-30.1	-0.30	1	PASS
Dijet	12.7	29.0	0.29	2	PASS
ExpPoly2	13.4	30.8	0.32	2	PASS
Bern3	14.8	35.7	0.37	3	PASS
Bern4	12.6	31.5	0.33	4	PASS
Bern5	-12.7	-31.6	-0.32	5	PASS

Table C.1: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the tH lep 0fwd category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
PowerLaw	-19.6	-26.2	-0.29	1	PASS
Dijet	-19.8	-26.5	-0.29	2	PASS
Bern5	-19.4	-32.6	-0.36	5	PASS
Exponential	-23.7	-31.7	-0.34	1	FAIL
ExpPoly2	-25.2	-36.3	-0.40	2	FAIL
Bern3	-23.0	-35.1	-0.38	3	FAIL
Bern4	-22.6	-34.9	-0.38	4	FAIL

Table C.2: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the tH lep 1fwd category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
PowerLaw	-16.2	-10.7	-0.25	1	PASS
ExpPoly2	-13.2	-9.4	-0.22	2	PASS
Dijet	-16.4	-10.8	-0.26	2	PASS
Bern3	-13.8	-10.6	0.25	3	PASS
Bern4	-13.3	-10.0	-0.24	4	PASS
Exponential	-20.3	-13.4	-0.32	1	FAIL
Bern5	-23.3	-17.9	-0.43	5	FAIL

Table C.3: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the $t\bar{t}H$ lep category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
Exponential	3.6	3.8	0.05	1	PASS
PowerLaw	6.2	6.4	0.09	1	PASS
Dijet	6.2	6.3	0.09	2	PASS
ExpPoly2	5.8	6.2	0.09	2	PASS
Bern3	4.1	4.8	0.07	3	PASS
Bern4	5.2	5.7	0.09	4	PASS
Bern5	1.6	2.6	0.03	5	PASS

Table C.4: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the $t\bar{t}H$ had BDT1 category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
Exponential	4.5	4.8	0.08	1	PASS
PowerLaw	7.3	8.0	0.13	1	PASS
ExpPoly2	6.2	7.3	0.12	2	PASS
Dijet	8.3	8.7	0.15	2	PASS
Bern3	7.2	8.1	0.14	3	PASS
Bern4	6.1	7.1	0.12	4	PASS
Bern5	-4.0	-4.5	-0.08	5	PASS

Table C.5: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the $t\bar{t}H$ had BDT2 category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
Exponential	3.7	9.7	0.06	1	PASS
PowerLaw	6.4	16.5	0.10	1	PASS
Dijet	6.4	16.3	0.10	2	PASS
ExpPoly2	6.0	16.1	0.10	2	PASS
Bern3	4.5	12.7	0.08	3	PASS
Bern4	5.6	16.0	0.10	4	PASS
Bern5	2.1	7.3	0.04	5	PASS

Table C.6: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the $t\bar{t}H$ had BDT3 category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
Exponential	10.1	15.4	0.38	1	PASS
PowerLaw	17.4	26.4	0.65	1	PASS
Dijet	16.7	25.9	0.64	2	PASS
ExpPoly2	16.8	27.5	0.67	2	PASS
Bern3	14.2	24.0	0.59	3	PASS
Bern4	13.1	22.3	0.55	4	PASS
Bern5	9.0	16.8	0.41	5	PASS

Table C.7: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the $t\bar{t}H$ had BDT4 category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
PowerLaw	-3.5	-9.7	-0.25	1	PASS
Exponential	-19.2	-49.8	-1.22	1	PASS
ExpPoly2	-3.7	-10.6	-0.28	2	PASS
Dijet	4.1	11.6	0.28	2	PASS
Bern3	6.9	21.5	0.54	3	PASS
Bern4	-6.4	-21.7	-0.54	4	PASS
Bern5	-8.6	-29.7	-0.73	5	PASS

Table C.8: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the tH had 4j1b category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Model	$N_{\text{spur}}^{\text{max}}/\delta S$ [%]	$N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$ [%]	$N_{\text{spur}}^{\text{max}}$	N_{par}	Result
PowerLaw	-1.4	-6.5	-0.04	1	PASS
Exponential	-6.7	-27.0	-0.16	1	PASS
ExpPoly2	-1.4	-6.8	-0.04	2	PASS
Dijet	-1.7	-8.2	-0.05	2	PASS
Bern3	1.7	8.7	0.05	3	PASS
Bern4	-1.6	-9.6	-0.05	4	PASS
Bern5	-3.6	-19.9	-0.12	5	PASS

Table C.9: Maximum fitted spurious signal between 121 and 129 GeV, $N_{\text{spur}}^{\text{max}}$, and the ratios $N_{\text{spur}}^{\text{max}}/\delta S$ and $N_{\text{spur}}^{\text{max}}/S_{\text{ref}}$, which are required to be smaller than 20% or 10%, respectively, for the different tested functional forms for the tH had 4j2b category. In addition, the number of free parameters N_{par} of the tested functional form, excluding its normalization, and the results of the spurious signal test are shown.

Appendix D

Measurement of the differential $N_{b\text{-jets}}$ cross section

D.1 Comparison of *b*-tagging WPs

Table D.1: The expected number of events for $79.8\,\text{fb}^{-1}$ is shown for the inclusive Higgs boson signal, S_{Higgs}, the contribution of ggH with true b-jets in the final state, S_{ggH+b}, and B_{ContBG}, estimated from a $m_{\gamma\gamma}$ sideband fit as shown below. The requirements $N_{\text{lep}} = 0$ and $N_{b\text{-tag}} = 1$ are applied. In addition, the expected significance for S_{Higgs} and S_{ggH+b} is shown.

b-tagging WP	S_{Higgs}	S_{ggH+b}	B_{ContBG}	Exp. sig. to S_{Higgs} $[\sigma]$	Exp. sig. to S_{ggH+b} $[\sigma]$
60%	46.6	16.0	5960	1.75	0.59
70%	63.4	18.9	9950	1.85	0.55
77%	86.2	20.8	15400	2.02	0.48
85%	157	22.9	29400	2.67	0.38

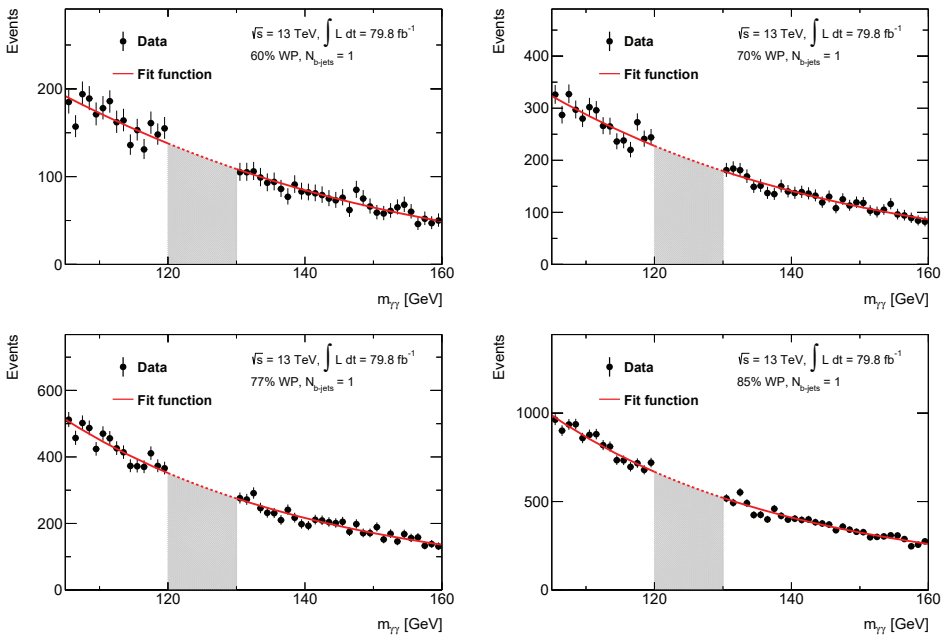


Figure D.1: Fit of an ExpPoly2 function to the $m_{\gamma\gamma}$ distribution of events with $N_{\text{lep}} = 0$ and $N_{b\text{-tag}} = 1$ for the different b-tagging WPs. The number of sideband events and the integral over the fit function in the shaded area are used to estimate B_{ContBG} in the full $m_{\gamma\gamma}$ range.

D.2 Flavor composition

Table D.2: Classification of the selected events from the different Higgs boson production processes into $H{+}b$, $H{+}c$ and $H{+}\ell$ in the different bins of the $N_{b\text{-jets}}$ distribution.

	$N_{b\text{-jets}} = 0$			$N_{b\text{-jets}} = 1$			$N_{b\text{-jets}} \geq 2$		
Process	$H{+}\ell$	$H{+}c$	$H{+}b$	$H{+}\ell$	$H{+}c$	$H{+}b$	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH	92%	6.8%	1.4%	18%	26%	57%	4.5%	4.0%	91%
VBF	83%	16%	1.0%	15%	48%	37%	1.7%	8.5%	90%
WH	70%	29%	0.39%	12%	76%	12%	1.8%	32%	66%
ZH	76%	18%	6.4%	4.0%	18%	78%	0.1%	2.1%	98%
$t\bar{t}H$	15%	15%	71%	0.11%	1.1%	99%	0.002%	0.047%	100%
$tHjb$	33%	15%	52%	0.29%	1.5%	98%	0%	0.098%	100%
WtH	23%	18%	59%	0.4%	2%	98%	0%	0.11%	100%
$b\bar{b}H$	52%	2.2%	46%	0.64%	0.57%	99%	0.46%	0%	100%

Table D.3: Classification into $H{+}b$, $H{+}c$ and $H{+}\ell$ of events with $N_{\text{lep}} = 0$ and $N_{b\text{-tag}} = 1$ for the different b-tagging WPs.

60% WP

Process	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH	11%	18%	71%
VBF	7.8%	36%	56%
WH	7.6%	70%	22%
ZH	1.3%	9.4%	89%
$t\bar{t}H$	0.031%	0.48%	99%
$tHjb$	0.046%	0.68%	99%
WtH	0.16%	0.79%	99%
$b\bar{b}H$	0.45%	0.27%	99%

70% WP

Process	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH	18%	26%	56%
VBF	15%	48%	37%
WH	12%	77%	12%
ZH	3.9%	18%	78%
$t\bar{t}H$	0.12%	1.1%	99%
$tHjb$	0.27%	1.6%	98%
WtH	0.4%	1.9%	98%
$b\bar{b}H$	0.92%	0.58%	98%

77% WP

Process	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH	28%	30%	42%
VBF	23%	53%	24%
WH	16%	77%	6.9%
ZH	8.5%	26%	66%
$t\bar{t}H$	0.31%	2.1%	98%
$tHjb$	0.65%	2.7%	97%
WtH	0.74%	3.5%	96%
$b\bar{b}H$	1.9%	0.64%	97%

85% WP

Process	$H{+}\ell$	$H{+}c$	$H{+}b$
ggH	52%	26%	22%
VBF	41%	47%	13%
WH	29%	68%	3.3%
ZH	23%	31%	47%
$t\bar{t}H$	1.3%	4.4%	94%
$tHjb$	2.4%	5.1%	93%
WtH	2.8%	6.7%	90%
$b\bar{b}H$	4.6%	1.1%	94%

D.3 Signal model

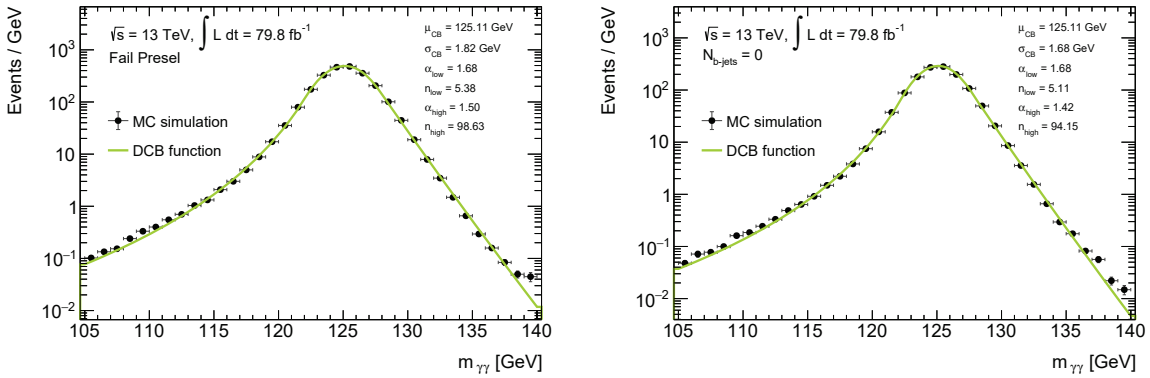


Figure D.2: The DCB signal model derived from Higgs boson MC simulation for the $N_{b\text{-jets}} = 0$ bin and the underflow bin.

D.4 Signal extraction uncertainties

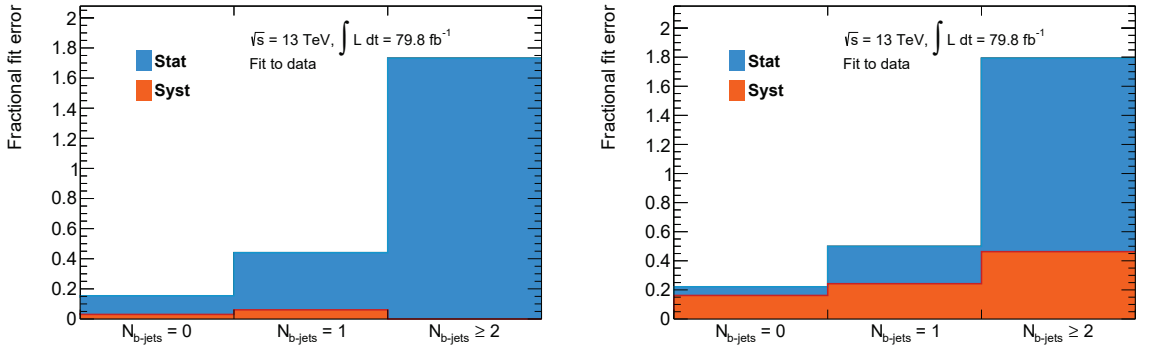
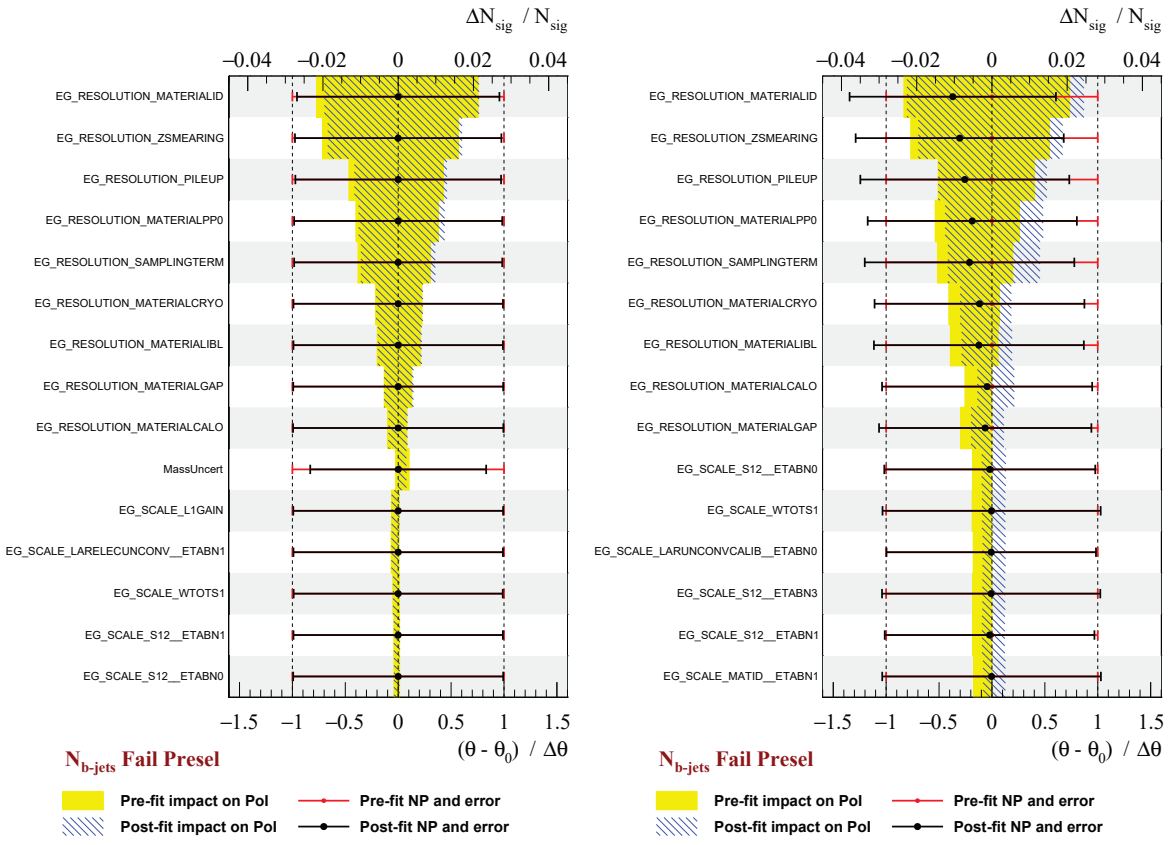


Figure D.3: Breakdown of the expected fractional uncertainties on the extracted signal yield, evaluated from the fit to data. The systematic uncertainty is shown in orange, while the statistical uncertainty is shown in blue. In a) the systematic uncertainties only include the PES/PER and Higgs boson mass uncertainties, while in b) the background modeling uncertainty is included as well.



a) b)

Figure D.4: Ranking plots, obtained from a) a fit to an Asimov dataset and b) the fit to data, showing the 15 most important NPs for the underflow bin. The number of signal events, N^{sig}, corresponds to the PoI. The pre-fit and post-fit values of the NPs and the associated uncertainties are shown as red and black points and lines, respectively. The yellow and blue hatched bands show the pre-fit and post-fit uncertainties on N^{sig}, respectively.

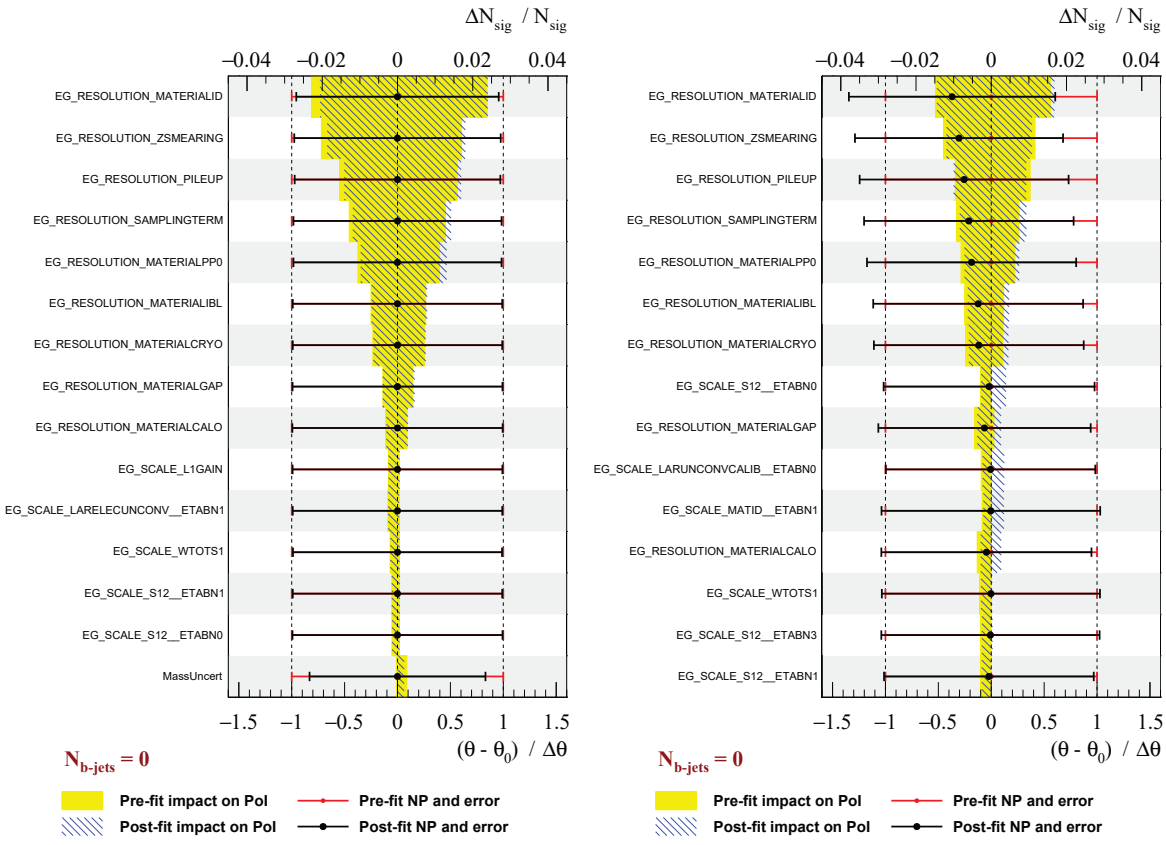


Figure D.5: Ranking plots, obtained from a) a fit to an Asimov dataset and b) the fit to data, showing the 15 most important NPs for the $N_{b\text{-jets}} = 0$ bin. The number of signal events, N^{sig}, corresponds to the PoI. The pre-fit and post-fit values of the NPs and the associated uncertainties are shown as red and black points and lines, respectively. The yellow and blue hatched bands show the pre-fit and post-fit uncertainties on N^{sig}, respectively.

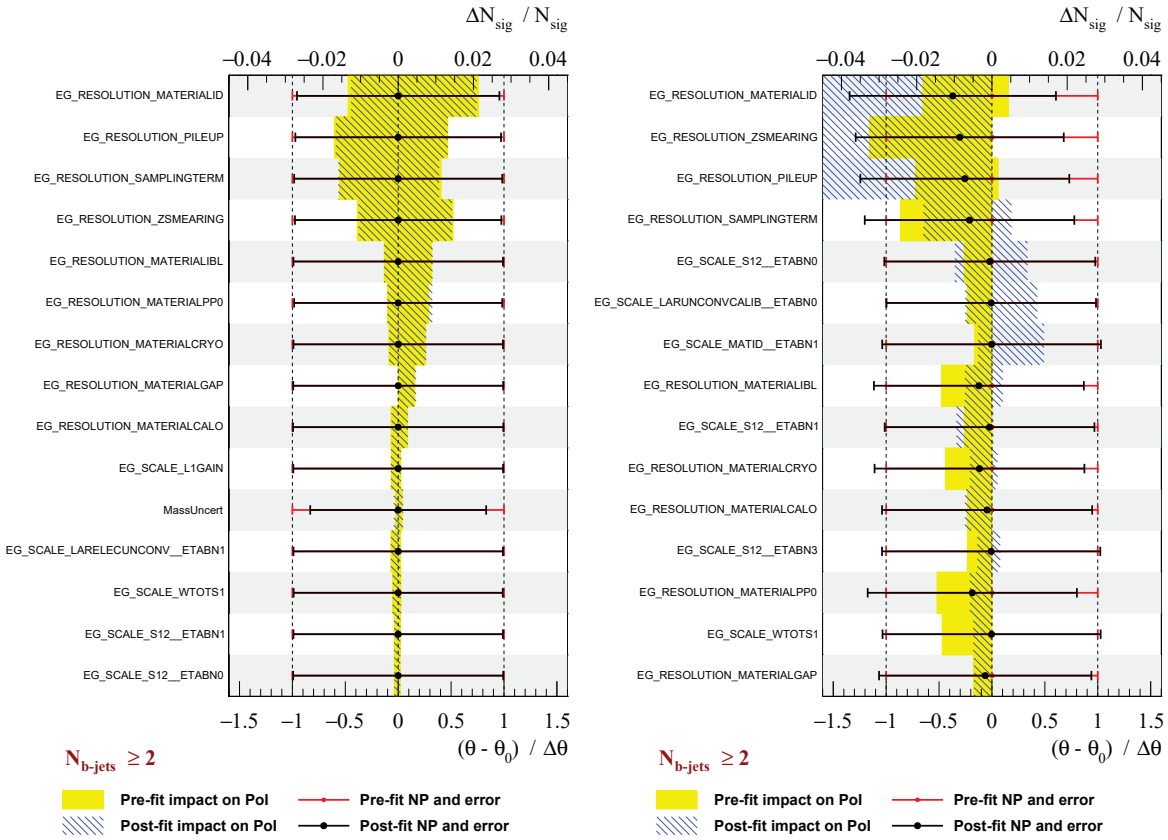


Figure D.6: Ranking plots, obtained from a) a fit to an Asimov dataset and b) the fit to data, showing the 15 most important NPs for the $N_{b\text{-jets}} \geq 2$ bin. The number of signal events, N^{sig}, corresponds to the PoI. The pre-fit and post-fit values of the NPs and the associated uncertainties are shown as red and black points and lines, respectively. The yellow and blue hatched bands show the pre-fit and post-fit uncertainties on N^{sig}, respectively. Due to the small number of observed events in the $N_{b\text{-jets}} \geq 2$ bin and the large statistical uncertainty, the estimation of the post-fit impact of the individual systematic uncertainties in this bin gets unstable for the fit to data.

Appendix E

MVA-based analysis strategy for H+b-jets

E.1 Correlations between the BDT input variables

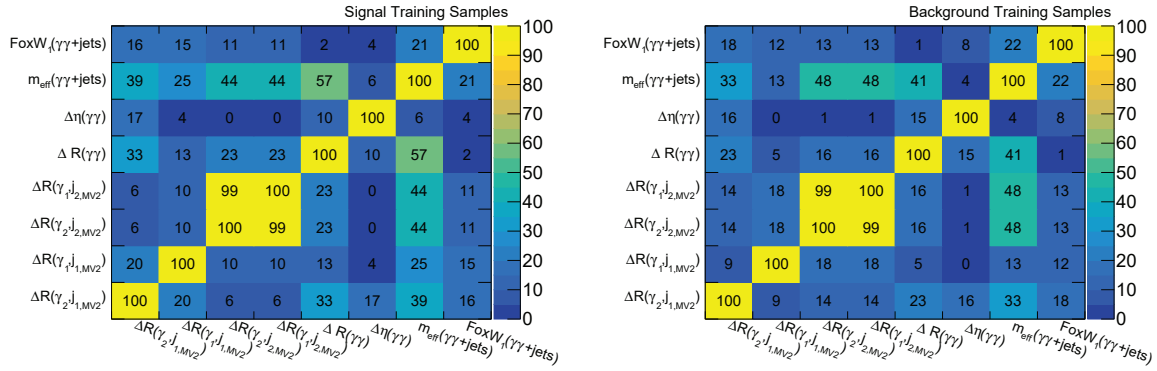


Figure E.1: Correlation matrices of the signal and background training samples for the different BDT input variables, showing the absolute values of the linear correlation coefficients.

E.2 BDT output after applying the full event selection

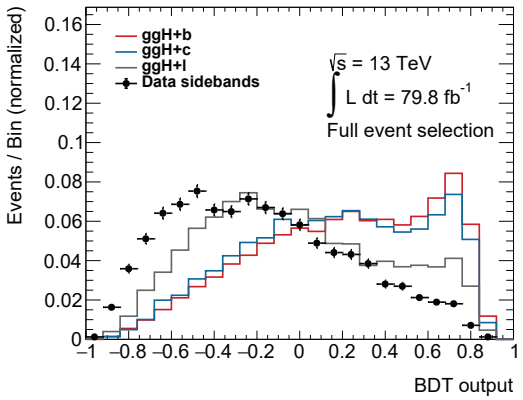


Figure E.2: BDT output for the full event selection, described in Section 11.2, for ggH MC simulation, where the MV2-leading jet is classified as a b-, c- or ℓ-jet, and the data sidebands.

E.3 Signal modeling

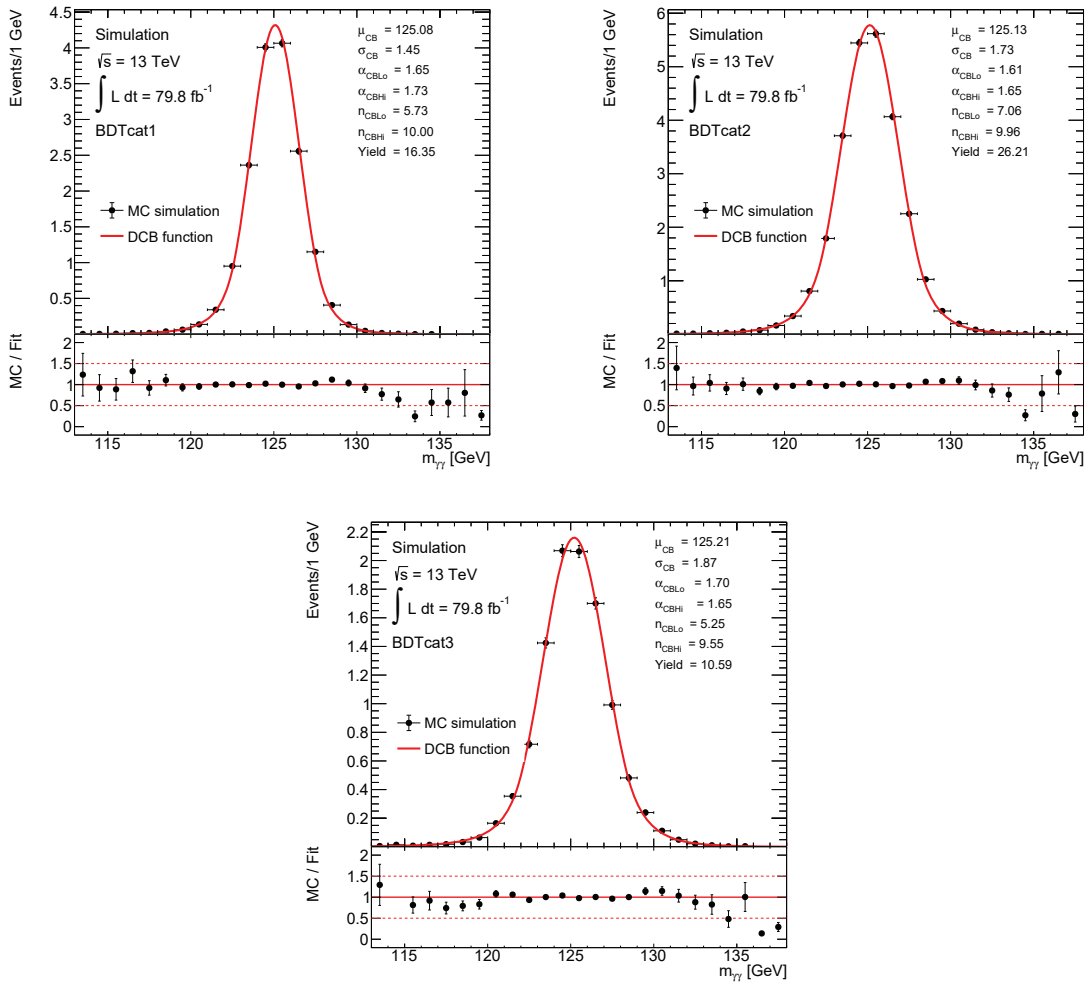


Figure E.3: Signal model for the three BDT categories, which is determined by fitting a DCB function to the $m_{\gamma\gamma}$ distribution predicted by the MC simulation of all contributing Higgs boson production processes.

E.4 Background modeling

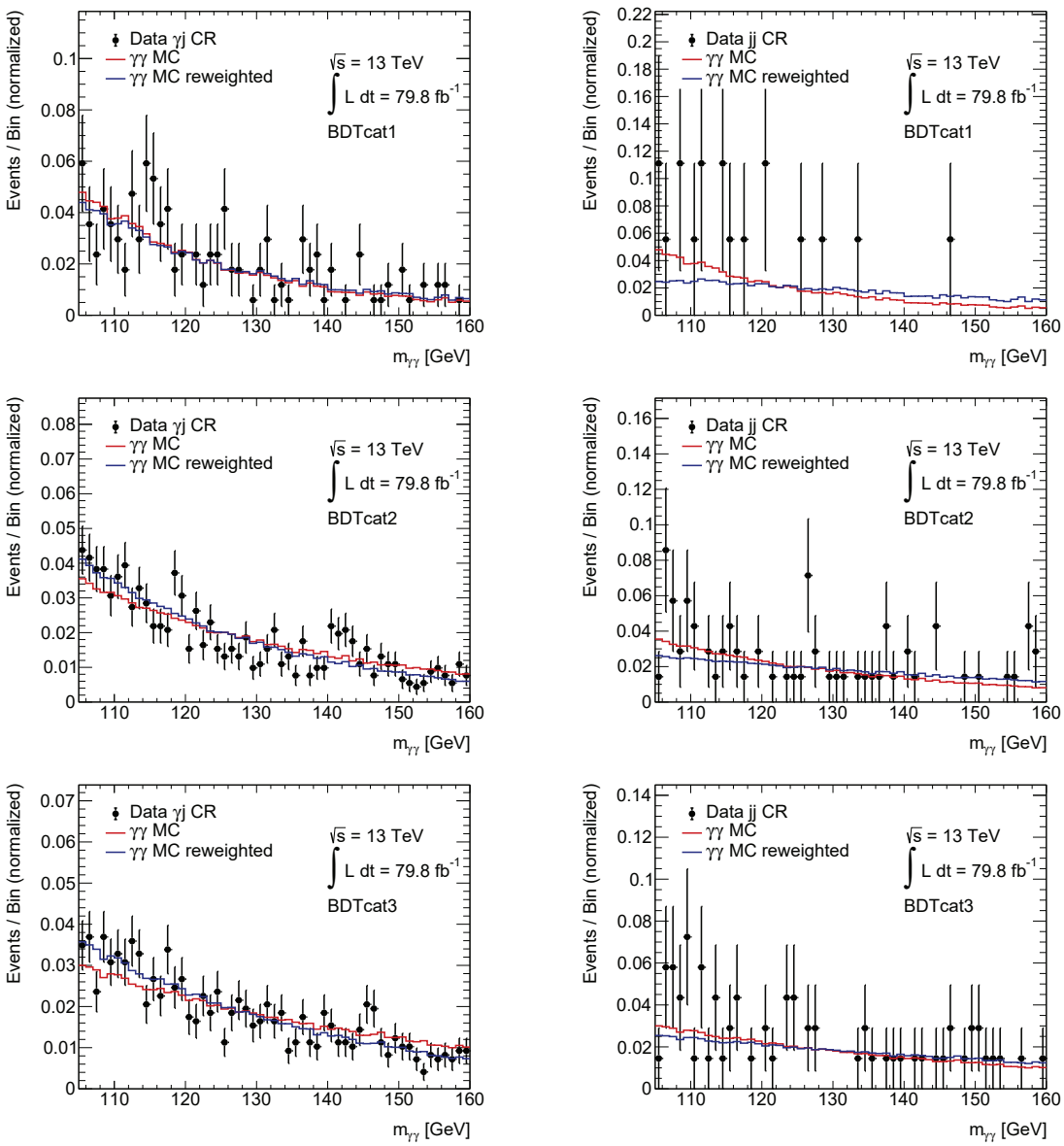


Figure E.4: Reweighting of the $m_{\gamma\gamma}$ distributions in diphoton MC simulation to the shape observed in the γj and jj control regions for the three BDT categories. For illustration purposes, the distributions are normalized to unit area. Due to the small number of events in the jj control region, it is challenging to properly derive the shape of this component. However, due to its small contribution, the jj component has a minor influence on the overall background template.

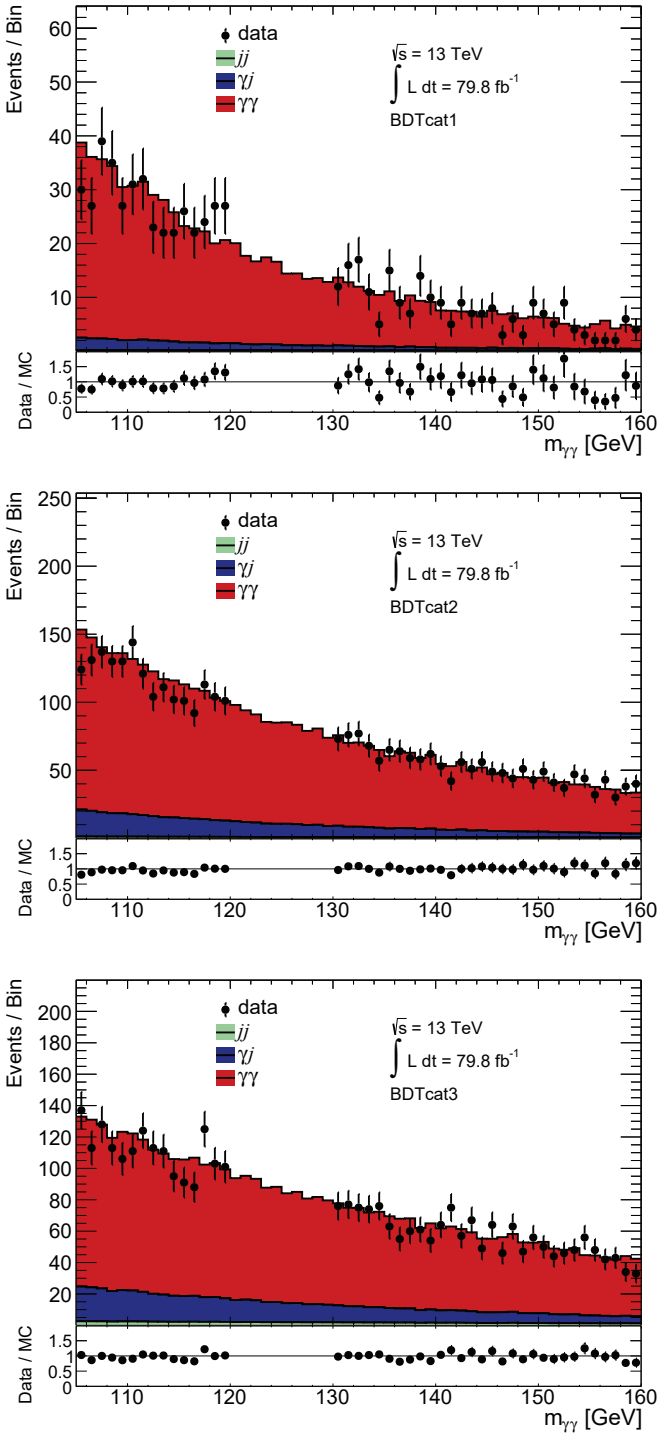


Figure E.5: Background templates for the three BDT categories, for which the $\gamma\gamma$ contribution is taken from diphoton MC simulation and the shape of the γj and jj components is derived from control regions in data, as described in Chapter 7.3.

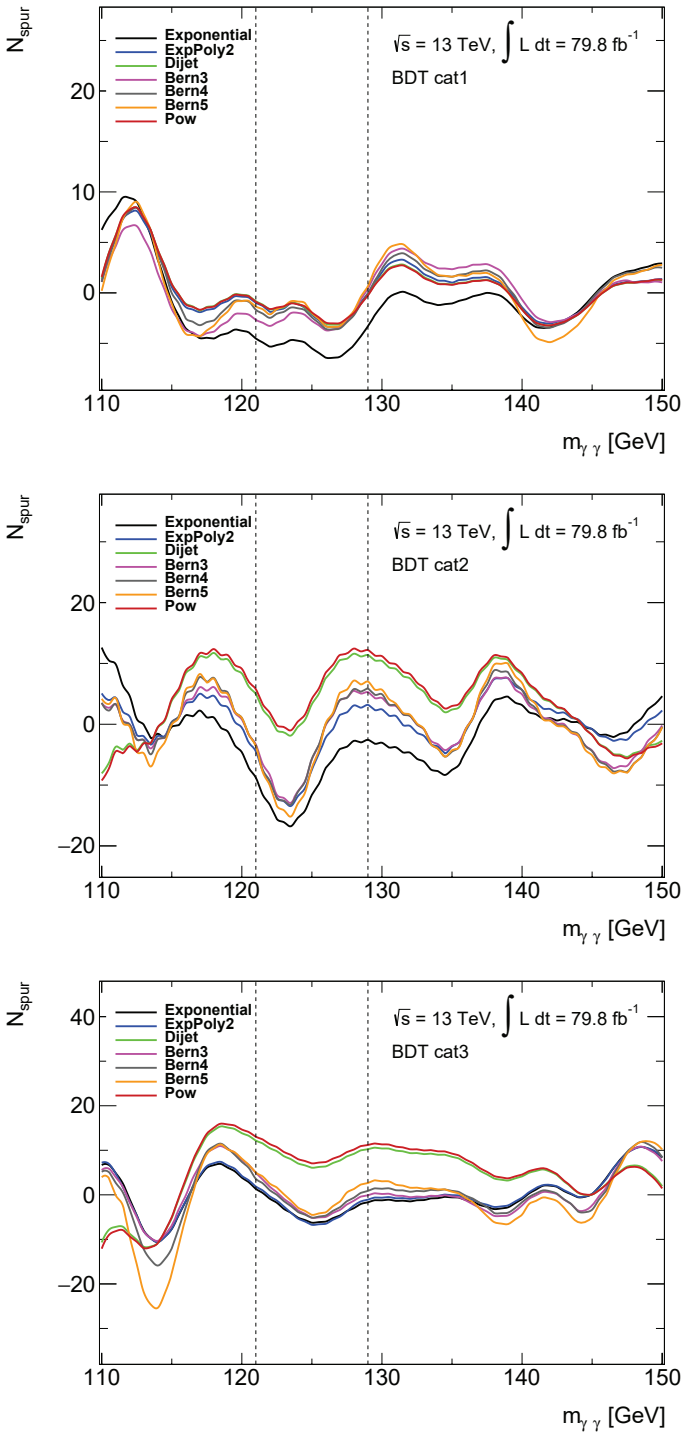


Figure E.6: Fitted spurious signal as a function of $m_{\gamma\gamma}$ for the three BDT categories. The region between 121 and 129 GeV, which is used to define $N_{\text{spur}}^{\text{max}}$, is indicated by the dashed lines.

E.5 PES/PER uncertainties

Table E.1: Uncertainties for the three BDT categories resulting from the PES/PER variations. The PES variations are denoted by EG_SCALE, while the PER variations are denoted by EG_RESOLUTION. The uncertainties are derived based on the relative change of the mean or width of the DCB function for each PES or PER variation, respectively.

Variation	Uncertainty in %		
	BDTcat1	BDTcat2	BDTcat3
EG_RESOLUTION_MATERIALCALO_up	0.76	0.44	0.42
EG_RESOLUTION_MATERIALCALO_down	0.86	0.46	0.38
EG_RESOLUTION_MATERIALCRYO_up	1.61	1.30	1.13
EG_RESOLUTION_MATERIALCRYO_down	1.68	1.28	1.06
EG_RESOLUTION_MATERIALGAP_up	0.79	0.66	0.76
EG_RESOLUTION_MATERIALGAP_down	0.92	0.71	0.78
EG_RESOLUTION_MATERIALIBL_up	1.88	1.14	0.99
EG_RESOLUTION_MATERIALIBL_down	2.10	1.28	0.95
EG_RESOLUTION_MATERIALID_up	4.83	3.87	3.36
EG_RESOLUTION_MATERIALID_down	5.46	4.05	3.41
EG_RESOLUTION_MATERIALPP0_up	2.08	1.97	2.03
EG_RESOLUTION_MATERIALPP0_down	2.36	1.97	2.02
EG_RESOLUTION_PILEUP_up	5.03	2.35	1.57
EG_RESOLUTION_PILEUP_down	4.99	2.57	1.83
EG_RESOLUTION_SAMPLINGTERM_up	4.68	2.03	1.33
EG_RESOLUTION_SAMPLINGTERM_down	4.34	1.85	1.48
EG_RESOLUTION_ZSMEARING_up	4.11	3.54	3.46
EG_RESOLUTION_ZSMEARING_down	3.79	3.04	2.71
EG_SCALE_PEDESTAL_up	0.01	0.01	<0.00
EG_SCALE_PEDESTAL_down	0.01	0.01	0.01
EG_SCALE_PS_BARREL_B12_up	0.07	0.06	0.05
EG_SCALE_PS_BARREL_B12_down	0.07	0.06	0.06
EG_SCALE_PS___ETABN0_up	0.01	0.01	0.01
EG_SCALE_PS___ETABN0_down	0.01	0.01	0.01
EG_SCALE_PS___ETABN1_up	0.01	0.01	<0.00
EG_SCALE_PS___ETABN1_down	0.01	0.01	0.01
EG_SCALE_PS___ETABN2_up	0.01	0.01	0.01
EG_SCALE_PS___ETABN2_down	0.01	0.01	0.01
EG_SCALE_PS___ETABN3_up	0.01	0.01	0.01
EG_SCALE_PS___ETABN3_down	0.01	0.01	0.01
EG_SCALE_PS___ETABN4_up	0.01	<0.00	<0.00
EG_SCALE_PS___ETABN4_down	0.01	<0.00	<0.00
EG_SCALE_PS___ETABN5_up	0.01	0.01	0.01
EG_SCALE_PS___ETABN5_down	0.01	0.01	0.01
EG_SCALE_PS___ETABN6_up	0.07	0.06	0.06
EG_SCALE_PS___ETABN6_down	0.07	0.06	0.06
EG_SCALE_PS___ETABN7_up	<0.00	<0.00	<0.00
EG_SCALE_PS___ETABN7_down	<0.00	<0.00	<0.00
EG_SCALE_PS___ETABN8_up	0.01	0.01	0.01
EG_SCALE_PS___ETABN8_down	0.01	0.01	0.01
EG_SCALE_AF2_up	<0.00	<0.00	<0.00
EG_SCALE_AF2_down	<0.00	<0.00	<0.00
EG_SCALE_E4SCINTILLATOR___ETABN0_up	<0.00	<0.00	<0.00
EG_SCALE_E4SCINTILLATOR___ETABN0_down	<0.00	<0.00	<0.00
EG_SCALE_E4SCINTILLATOR___ETABN1_up	<0.00	<0.00	<0.00
EG_SCALE_E4SCINTILLATOR___ETABN1_down	<0.00	<0.00	<0.00
EG_SCALE_E4SCINTILLATOR___ETABN2_up	0.01	<0.00	<0.00
EG_SCALE_E4SCINTILLATOR___ETABN2_down	0.01	<0.00	<0.00
EG_SCALE_G4_up	0.01	0.01	0.01
EG_SCALE_G4_down	0.01	0.01	0.01
EG_SCALE_L1GAIN_up	0.03	0.04	0.06
EG_SCALE_L1GAIN_down	0.03	0.04	0.05
EG_SCALE_L2GAIN_up	0.21	0.11	0.08
EG_SCALE_L2GAIN_down	0.21	0.12	0.08

Table E.2: Uncertainties for the three BDT categories resulting from the PES/PER variations. The PES variations are denoted by EG_SCALE, while the PER variations are denoted by EG_RESOLUTION. The uncertainties are derived based on the relative change of the mean or width of the DCB function for each PES or PER variation, respectively.

Variation	Uncertainty in %		
	BDTcat1	BDTcat2	BDTcat3
EG_SCALE_LARCALIB___ETABN0__up	0.10	0.08	0.06
EG_SCALE_LARCALIB___ETABN0__down	0.10	0.07	0.07
EG_SCALE_LARCALIB___ETABN1__up	0.02	0.02	0.03
EG_SCALE_LARCALIB___ETABN1__down	0.02	0.02	0.03
EG_SCALE_LARELECCALIB__up	<0.00	<0.00	0.01
EG_SCALE_LARELECCALIB__down	<0.00	<0.00	<0.00
EG_SCALE_LARELECUNCONV___ETABN0__up	0.09	0.08	0.07
EG_SCALE_LARELECUNCONV___ETABN0__down	0.09	0.08	0.08
EG_SCALE_LARELECUNCONV___ETABN1__up	0.04	0.04	0.05
EG_SCALE_LARELECUNCONV___ETABN1__down	0.04	0.05	0.05
EG_SCALE_LARUNCONVCALIB___ETABN0__up	0.03	0.03	0.03
EG_SCALE_LARUNCONVCALIB___ETABN0__down	0.03	0.03	0.03
EG_SCALE_LARUNCONVCALIB___ETABN1__up	<0.00	<0.00	<0.00
EG_SCALE_LARUNCONVCALIB___ETABN1__down	<0.00	<0.00	<0.00
EG_SCALE_S12___ETABN0__up	0.04	0.03	0.02
EG_SCALE_S12___ETABN0__down	0.04	0.03	0.03
EG_SCALE_S12___ETABN1__up	0.04	0.03	0.03
EG_SCALE_S12___ETABN1__down	0.04	0.03	0.03
EG_SCALE_S12___ETABN2__up	<0.00	<0.00	<0.00
EG_SCALE_S12___ETABN2__down	<0.00	<0.00	<0.00
EG_SCALE_S12___ETABN3__up	0.03	0.03	0.04
EG_SCALE_S12___ETABN3__down	0.03	0.03	0.04
EG_SCALE_S12___ETABN4__up	<0.00	<0.00	<0.00
EG_SCALE_S12___ETABN4__down	<0.00	<0.00	<0.00
EG_SCALE_TOPOCLUSTER_THRES__up	0.01	0.01	0.01
EG_SCALE_TOPOCLUSTER_THRES__down	<0.00	<0.00	<0.00
EG_SCALE_WTOTS1__up	0.03	0.04	0.04
EG_SCALE_WTOTS1__down	0.03	0.04	0.04
EG_SCALE_MATCALO___ETABN0__up	0.01	0.01	0.01
EG_SCALE_MATCALO___ETABN0__down	0.01	0.01	0.01
EG_SCALE_MATCALO___ETABN1__up	0.01	0.01	0.01
EG_SCALE_MATCALO___ETABN1__down	0.01	0.01	<0.00
EG_SCALE_MATCALO___ETABN10__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN10__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN11__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN11__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN2__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN2__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN3__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN3__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN4__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN4__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN5__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN5__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN6__up	0.01	0.01	0.01
EG_SCALE_MATCALO___ETABN6__down	0.01	0.01	0.01
EG_SCALE_MATCALO___ETABN7__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN7__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN8__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN8__down	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN9__up	<0.00	<0.00	<0.00
EG_SCALE_MATCALO___ETABN9__down	<0.00	<0.00	<0.00
EG_SCALE_MATID___ETABN0__up	0.06	0.05	0.06
EG_SCALE_MATID___ETABN0__down	0.06	0.06	0.05
EG_SCALE_MATID___ETABN1__up	0.02	0.02	0.02
EG_SCALE_MATID___ETABN1__down	0.02	0.02	0.02
EG_SCALE_MATID___ETABN2__up	0.01	0.02	0.02
EG_SCALE_MATID___ETABN2__down	0.01	0.02	0.02
EG_SCALE_MATID___ETABN3__up	<0.00	<0.00	0.01
EG_SCALE_MATID___ETABN3__down	<0.00	<0.00	0.01

Table E.3: Uncertainties for the three BDT categories resulting from the PES/PER variations. The PES variations are denoted by EG_SCALE, while the PER variations are denoted by EG_RESOLUTION. The uncertainties are derived based on the relative change of the mean or width of the DCB function for each PES or PER variation, respectively.

Variation	Uncertainty in %		
	BDTcat1	BDTcat2	BDTcat3
EG_SCALE_MATPP0___ETABN0__up	0.02	0.02	0.02
EG_SCALE_MATPP0___ETABN0__down	0.02	0.02	0.02
EG_SCALE_MATPP0___ETABN1__up	0.05	0.07	0.08
EG_SCALE_MATPP0___ETABN1__down	0.05	0.07	0.08
EG_SCALE_ZEESTAT__up	<0.00	<0.00	<0.00
EG_SCALE_ZEESTAT__down	<0.00	<0.00	<0.00
EG_SCALE_ZEESYST__up	0.06	0.07	0.07
EG_SCALE_ZEESYST__down	0.06	0.07	0.07
EG_SCALE_MATCRYO___ETABN0__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN0__down	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN1__up	0.01	0.01	0.01
EG_SCALE_MATCRYO___ETABN1__down	0.01	0.01	0.01
EG_SCALE_MATCRYO___ETABN10__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN10__down	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN11__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN11__down	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN2__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN2__down	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN3__up	0.01	0.01	0.01
EG_SCALE_MATCRYO___ETABN3__down	0.01	0.01	0.01
EG_SCALE_MATCRYO___ETABN4__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN4__down	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN5__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN5__down	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN6__up	0.01	<0.00	0.01
EG_SCALE_MATCRYO___ETABN6__down	0.01	0.01	<0.00
EG_SCALE_MATCRYO___ETABN7__up	0.02	0.02	0.02
EG_SCALE_MATCRYO___ETABN7__down	0.02	0.02	0.02
EG_SCALE_MATCRYO___ETABN8__up	0.02	0.02	0.02
EG_SCALE_MATCRYO___ETABN8__down	0.02	0.02	0.02
EG_SCALE_MATCRYO___ETABN9__up	<0.00	<0.00	<0.00
EG_SCALE_MATCRYO___ETABN9__down	<0.00	<0.00	<0.00

Bibliography

[1] ATLAS collaboration, *Measurement of fiducial, differential and production cross sections in the $H \to \gamma\gamma$ decay channel with 13.3 fb^{-1} of 13 TeV proton-proton collision data with the ATLAS detector*, ATLAS-CONF-2016-067 (2016), http://cds.cern.ch/record/2206210.

[2] ATLAS collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **98**, 052005 (2018).

[3] 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**, 173 (2018).

[4] ATLAS collaboration, *Measurements of Higgs boson properties in the diphoton decay channel using 80 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2018-028 (2018), http://cds.cern.ch/record/2628771.

[5] ATLAS collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, Eur. Phys. J. C **76**, 666 (2016).

[6] 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**, 205 (2019).

[7] 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**, 1 (2012).

[8] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716**, 30 (2012).

[9] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, *A Direct Empirical Proof of the Existence of Dark Matter*, Astrophys. J. **648**, L109 (2006).

[10] S. Perlmutter et al., *Measurements of Ω and Λ from 42 High-Redshift Supernovae*, Astrophys. J. **517**, 565 (1999).

[11] A. G. Riess et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, Astron. J. **116**, 1009 (1998).

[12] K. G. Wilson, *The Renormalization Group and Strong Interactions*, Phys. Rev. D **3**, 1818 (1971).

[13] F. Englert, and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13**, 321 (1964).

[14] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964).

[15]G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13**, 585 (1964).

[16] ATLAS and CMS collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114**, 191803 (2015).

[17] Nobelprize.org, Nobel Media AB 2019, *The Nobel Prize in Physics 2013*, https://www.nobelprize.org/prizes/physics/2013/summary/ (visited on 01/22/2019).

[18] D0 collaboration, *Observation of the Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995).

[19] CMS collaboration, *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. **120**, 231801 (2018).

[20] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN-2017-002-M (2016), https://cds.cern.ch/record/2227475.

[21] ATLAS collaboration, *Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in pp collisions at $\sqrt{s}$ =8 TeV using the ATLAS detector*, Eur. Phys. J. C **76**, 11 (2016).

[22] ATLAS collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **04**, 046 (2019).

[23] Particle Data Group, *Review of Particle Physics*, Phys. Rev. D. **98**, 030001 (2018).

[24] ATLAS, CDF, CMS and D0 collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass* (2014), arXiv:1403.4427.

[25] LIGO scientific collaboration and Virgo collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, 061102 (2016).

[26]H. D. Politzer, *Asymptotic freedom: An approach to strong Interactions*, Phys. Rept. **14**, 129 (1974).

[27]H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30**, 1346 (1973).

[28]D. J. Gross, and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30**, 1343 (1973).

[29]R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, Phys. Rev. **80**, 440 (1950).

[30]S. L. Glashow, *Partial-symmetries of weak Interactions*, Nucl. Phys. **22**, 579 (1961).

[31]S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967).

[32]A. Salam, and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13**, 168 (1964).

[33]N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963).

[34]M. Kobayashi, and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[35] Super-Kamiokande collaboration, *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998).

[36] SNO collaboration, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89**, 011301 (2002).

[37]Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Prog. Theor. Phys. **28**, 870 (1962).

[38]B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Sov. Phys. JETP **26**, 984 (1968).

[39]D. Griffiths, *Introduction to elementary particles*, Copyright by WILEY-VCH Verlag GmbH & Co. KGaA (Wiley-VCH, Weinheim, 2014).

[40] LHC Top Physics Working Group, *NLO single-top channel cross sections*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec (visited on 04/30/2019).

[41]P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer, *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191**, 74 (2015).

[42]N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82**, 054018 (2010).

[43] ATLAS collaboration, *Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **04**, 086 (2017).

[44] ATLAS collaboration, *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s} = 13$ TeV with ATLAS*, Eur. Phys. J. C **78**, 186 (2018).

[45] ATLAS collaboration, *Search for s-channel single top-quark production in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Lett. B **740**, 118 (2015).

[46]J. P. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, Comput. Phys. Commun. **210**, 103 (2017).

[47] LHC Higgs Cross Section Working Group, *κ coupling modifiers*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG2KAPPA (visited on 04/30/2019).

[48] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN-2013-004 (2013), https://cds.cern.ch/record/1559921.

[49]O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2004), https://cds.cern.ch/record/782076.

[50]J. Haffner, *The CERN accelerator complex.* OPEN-PHO-ACCEL-2013-056 (Oct. 2013), https://cds.cern.ch/record/1621894.

[51]G. Apollinari, I. Béjar Alonso, O. Brüning, P. Fessia, M. Lamont, L. Rossi, and L. Tavian, *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report Version 0.1*, CERN Yellow Reports: Monographs (CERN, Geneva, 2017), https://cds.cern.ch/record/2284929.

[52] ATLAS collaboration, *Total integrated luminosity and data quality in 2015-2018*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 2018/01/27).

[53] ATLAS collaboration, *Peak luminosity per fill in 2018*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 2018/01/27).

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

[55] ATLAS collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-018 (2015), https://cds.cern.ch/record/2037683.

[56] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-019 (2010), https://cds.cern.ch/record/1291633, Addendum: ATLAS-TDR-19-ADD-1 (2012), https://cds.cern.ch/record/1451888.

[57] ATLAS TDAQ collaboration, *The ATLAS Data Acquisition and High Level Trigger system*, JINST **11**, P06008 (2016).

[58] ATLAS collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77**, 317 (2017).

[59] ATLAS collaboration, *2017 pp collisions: Number of Interactions per Crossing*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 01/28/2019).

[60] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504**, 145 (2011).

[61] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02**, 007 (2009).

[62] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178**, 852 (2008).

[63] S. Gieseke, A. Ribon, M. H. Seymour, P. Stephens, and B. Webber, *Herwig++ 1.0: an Event generator for* e^+e^- *annihilation*, JHEP **02**, 005 (2004).

[64] J. Bellm et al., *Herwig++ 2.7 Release Note* (2013), arXiv:1310.6877.

[65] M. Bähr et al., *Herwig++ physics and Manual*, Eur. Phys. J. C **58**, 639 (2008).

[66] S. Schumann, and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03**, 038 (2008).

[67] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462**, 152 (2001).

[68] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07**, 079 (2014).

[69] R. Frederix, E. Re, and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP **09**, 130 (2012).

[70] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, *New parton distributions for collider physics*, Phys. Rev. D **82**, 074024 (2010).

[71] ATLAS collaboration, *ATLAS Run 1 Pythia8 tunes*, ATL-PHYS-PUB-2014-021 (2014), https://cds.cern.ch/record/1966419.

[72] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, *Higgs production in association with a single top quark at the LHC*, Eur. Phys. J. C **75**, 267 (2015).

[73] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. C **72**, 2225 (2012).

[74]F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, *tWH associated production at the LHC*, Eur. Phys. J. C **77**, 34 (2017).

[75] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04**, 040 (2015).

[76]W. Beenakker, S. Dittmaier, M. Krämer, B. Plümper, M. Spira, and P. M. Zerwas, *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*, Nucl. Phys. B **653**, 151 (2003).

[77]S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackeroth, *Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections*, Phys. Rev. D **68**, 034022 (2003).

[78]Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen, and L. Guo, *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, Phys. Lett. B **738**, 1 (2014).

[79] S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP **06**, 184 (2015).

[80]P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11**, 040 (2004).

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

[82]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**, 043 (2010).

[83] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43**, 023001 (2016).

[84] ATLAS collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09**, 145 (2014).

[85]S. Dawson, C. B. Jackson, L. Reina, and D. Wackeroth, *Exclusive Higgs boson production with bottom quarks at hadron colliders*, Phys. Rev. D **69**, 074027 (2004).

[86]S. Dittmaier, M. Krämer, and M. Spira, *Higgs radiation off bottom quarks at the Tevatron and the CERN LHC*, Phys. Rev. D **70**, 074010 (2004).

[87]R. Harlander, M. Krämer, and M. Schumacher, *Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach*, CERN-PH-TH-2011-134 (2011), https://cds.cern.ch/record/1407669.

[88] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, JHEP **10**, 222 (2013).

[89]C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, Phys. Rev. Lett. **114**, 212001 (2015).

[90]C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, and B. Mistlberger, *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, JHEP **05**, 058 (2016).

[91]S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, Phys. Lett. B **670**, 12 (2008).

[92]C. Anastasiou, R. Boughezal, and F. Petriello, *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*, JHEP **04**, 003 (2009).

[93]P. Nason, and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02**, 037 (2010).

[94]K. Mimasu, V. Sanz, and C. Williams, *Higher order QCD predictions for associated Higgs production with anomalous couplings to gauge bosons*, JHEP **08**, 039 (2016).

[95]A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK*, JHEP **03**, 075 (2012).

[96]O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. B **579**, 149 (2004).

[97]M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and Electroweak Corrections to the Production of Higgs Boson + 2 Jets via Weak Interactions at the Large Hadron Collider*, Phys. Rev. Lett. **99**, 161803 (2007).

[98]M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, Phys. Rev. D **77**, 013002 (2008).

[99]P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. **105**, 011801 (2010).

[100]L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke, *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, JHEP **02**, 078 (2013).

[101]A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63**, 189 (2009).

[102] ATLAS collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003 (2012), https://cds.cern.ch/record/1474107.

[103] ATLAS collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70**, 823 (2010).

[104] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506**, 250 (2003).

[105]S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05**, 053 (2009).

[106]W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, ATL-COM-LARG-2008-003 (2008), https://cds.cern.ch/record/1099735.

[107] ATLAS collaboration, *Expected electron and photon energy calibration performance with the ATLAS detector using Super-Clusters*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2017-001/ (visited on 01/24/2019).

[108] ATLAS collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022 (2017), http://cds.cern.ch/record/2298955.

[109] ATLAS collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77**, 490 (2017).

[110] ATLAS collaboration, *Expected photon performance in the ATLAS experiment*, ATL-PHYS-PUB-2011-007 (2011), https://cds.cern.ch/record/1345329.

[111] ATLAS collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74**, 3071 (2014).

[112] ATLAS collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at* $\sqrt{s} = 13$ TeV, ATL-PHYS-PUB-2016-015 (2016), https://cds.cern.ch/record/2203514.

[113] ATLAS collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024 (2016), https://cds.cern.ch/record/2157687.

[114] ATLAS collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s}$ *= 13 TeV*, CERN-EP-2018-273, Accepted by Eur. Phys. J. (2019), http://cds.cern.ch/record/2657964.

[115] ATLAS collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. C **76**, 292 (2016).

[116] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-*k_t *jet clustering algorithm*, JHEP **04**, 063 (2008).

[117] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C **72**, 1896 (2012).

[118] 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**, 581 (2016).

[119] ATLAS collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-015 (2015), https://cds.cern.ch/record/2037613.

[120] ATLAS collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. C **75**, 17 (2015).

[121] ATLAS collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018 (2014), https://cds.cern.ch/record/1700870.

[122] ATLAS collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-023 (2015), https://cds.cern.ch/record/2037700.

[123] ATLAS collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at* $\sqrt{s}$*= 13 TeV*, ATL-PHYS-PUB-2015-027 (2015), https://cds.cern.ch/record/2037904.

[124] ATLAS collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022 (2015), https://cds.cern.ch/record/2037697.

[125] ATLAS collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012 (2016), https://cds.cern.ch/record/2160731.

[126] ATLAS collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102 (2011), https://cds.cern.ch/record/1369219.

[127] ATLAS collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, CERN-OPEN-2008-020 (2009), https://cds.cern.ch/record/1125884.

[128] ATLAS collaboration, *Performance of b-jet identification in the ATLAS Experiment*, JINST **11**, P04008 (2016).

[129] ATLAS collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using* $t\bar{t}$ *events at* $\sqrt{s} = 13$ *TeV*, JHEP **08**, 089 (2018).

[130] ATLAS collaboration, *Calibration of the b-tagging efficiency on charm jets using a sample of* $W+c$ *events with* $\sqrt{s}$ *= 13 TeV ATLAS data*, ATLAS-CONF-2018-055 (2018), https://cds.cern.ch/record/2652195.

[131] ATLAS collaboration, *Calibration of the ATLAS b-tagging algorithm in $t\bar{t}$ semi-leptonic events*, ATLAS-CONF-2018-045 (2018), https://cds.cern.ch/record/2638455.

[132] ATLAS collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006 (2018), https://cds.cern.ch/record/2314418.

[133] ATLAS collaboration, *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*, ATLAS-CONF-2018-001 (2018), https://cds.cern.ch/record/2306649.

[134] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867**, 244 (2013).

[135] ATLAS collaboration, *Public Egamma Trigger Plots for Collision Data: Electron and photon trigger efficiencies using the full 2016 ATLAS data [ATL-COM-DAQ-2017-015]*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults (visited on 04/29/2019).

[136] ATLAS collaboration, *Measurement of the isolated diphoton cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **85**, 012003 (2012).

[137] ATLAS collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **83**, 052005 (2011).

[138] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71**, 1554 (2011), Erratum: Eur. Phys. J. C **73**, 2501 (2013).

[139] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. G **28**, 2693 (2002).

[140] ATLAS collaboration, *Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, Phys. Lett. B **740**, 222 (2015).

[141] ATLAS collaboration, *Measurement of Higgs boson production in association with a $t\bar{t}$ pair in the diphoton decay channel using 139 fb^{-1} of LHC data collected at $\sqrt{s} = 13$ TeV by the ATLAS experiment*, ATLAS-CONF-2019-004 (2019), http://cds.cern.ch/record/2668103.

[142] CMS collaboration, *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **11**, 185 (2018).

[143] CMS collaboration, *Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JHEP **06**, 177 (2016).

[144] CMS collaboration, *Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at $\sqrt{s} = 13$TeV*, Phys. Rev. D **99**, 092005 (2019).

[145] I. Nitsche, *Studies on the prospects for the search for tH-production with $H \to \gamma\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Master's thesis, TU Dortmund (September 2015).

[146] ATLAS collaboration, *Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector*, Phys. Rev. Lett. **114**, 081802 (2015).

[147] ATLAS collaboration, *Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment*, JHEP **11**, 040 (2018).

[148] ATLAS collaboration, *Search for single production of a vector-like B quark decaying into a bottom quark and a Higgs boson which decays into a pair of photons*, ATLAS-CONF-2018-024 (2018), https://cds.cern.ch/record/2628759.

[149] ATLAS collaboration, *Measurement of inclusive and differential cross sections in the $H \to ZZ^* \to 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **10**, 132 (2017), ($N_{b\text{-jets}}$ measurement is part of the auxiliary material, which is available at `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2016-25/`).

[150] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76**, 653 (2016).

[151] ATLAS and CMS collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08**, 045 (2016).

[152] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT**, 040 (2007), arXiv:`physics/0703039`.

[153] Y. Freund, and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55**, 119 (1997).

[154] J. H. Friedman, *Greedy function approximation: a gradient boosting machine*, The Annals of Statistics **29**, 1189 (2001).

[155] G. C. Fox, and S. Wolfram, *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes*, Phys. Rev. Lett. **41**, 1581 (1978).

[156] C. Bernaciak, M. S. A. Buschmann, A. Butter, and T. Plehn, *Fox-Wolfram Moments in Higgs Physics*, Phys. Rev. D **87**, 073014 (2013).

Acknowledgments

The completion of this thesis would not have been possible without the support that I received from various people during the last years. I am very grateful for this support and would like to use this last page of my thesis to say thank you. First of all, I would like to thank Prof. Dr. Kevin Kröninger for providing me a PhD position in his group at the TU Dortmund and giving me the opportunity to work in this interesting field of research. In addition, I would like to thank Prof. Dr. Dr. Wolfgang Rhode for agreeing to read and review my thesis. Special thanks go to Dr. Johannes Erdmann for the great supervision during the last years, starting from the beginning of my Master's thesis. I would like to express my gratitude for introducing me to the interesting field of Higgs boson physics, always being open for discussions and having an open ear for questions. Last but not least, thanks for proof-reading my thesis.

In addition, I would like to thank everyone from the E4 group for the excellent working atmosphere during the last years. Thanks go to Andrea and Markus for their kind support in solving organizational and technical problems. Furthermore, I would like to thank all the members of the analysis group for the interesting discussions relating to physics but also for the non-physics related conversations during our coffee or tea breaks. Special thanks go to my colleagues Elena Freundlich, Gregor Geßner, Tobias Kupfer and Björn Wendland for helping me proof-reading my thesis. Furthermore, I would like to thank Tobias and Gregor for providing me a room during my stay at CERN, I really enjoyed all the barbecues and cooking sessions in Chevry. Finally, thanks go to my present (and previous) office mates Elena, Björn, Tobias and Frederic for the nice working atmosphere in our office, the helpful discussions about work-related problems and the numerous reasons to laugh even in stressful periods.

Furthermore, I would like to thank my colleagues from the ATLAS collaboration who made the work in this interesting field of research pleasant and exciting despite all the urgent deadlines. I owe thanks to the $t\bar{t}H/tH$ analysis team, the $H \to \gamma\gamma$ differential cross section group and the photon ID group for the excellent cooperation. Special thanks go to Jahred Adelman, Elizabeth Brost, Jared Vasquez and Diane Cinca for the great teamwork and the interesting discussions. I would particularly like to thank Diane for her supervision during my time at CERN. Besides our analysis-related discussions, I enjoyed the French lessons and the Sushi Friday.

Last but not least, I would like to thank my family, in particular my parents, for their support and encouragement during the last years. Your kind words of advice helped me a lot whenever I had doubts. I would also like to thank my friends who kept on trying to understand what I was actually doing during the last years. Furthermore, I thank Hermine, my clumsy fellow, for giving me a reason to smile every day and reminding me that there is something else than physics in life. Lastly, I owe thanks to Richard for being at my side during the last five years, enduring my moods, encouraging me to continue, providing me with chocolate when necessary and so much more.