Aspects of Rare Production and Decay Processes of the Higgs Boson

Aspekte von seltenen Produktions- und Zerfallsprozessen des Higgs-Bosons

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von

Daniel Rauch

Hamburg
2020

Gutachter/innen der Dissertation:	Prof. Dr. Kerstin Tackmann Prof. Dr. Peter Schleper
Zusammensetzung der Prüfungskommission:	Prof. Dr. Kerstin Tackmann PD Dr. Judith Katzy Prof. Dr. Peter Schleper PD Dr. Markus Diehl Prof. Dr. Dieter Horns
Vorsitzender der Prüfungskommission:	Prof. Dr. Dieter Horns
Datum der Disputation:	28. September 2020
Vorsitzender Fach-Promotionsausschuss PHYSIK:	Prof. Dr. Günter Hans Walter Sigl
Leiter des Fachbereichs PHYSIK:	Prof. Dr. Wolfgang Hansen
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

Abstract

The Higgs boson is the most recent fundamental particle to be experimentally observed. Due to its large mass and its couplings to massive bosons and fermions, it has a wide variety of decay channels. Among them is the still unobserved rare decay $H \to Z(\ell^+\ell^-)\gamma$. In this thesis, a search for the $Z\gamma$ decay mode of the Standard Model Higgs boson as well as potential new higher-mass resonances decaying to the same final state is presented. The analysis is based on a dataset of 36.1 fb^{-1} of proton-proton collisions collected with the ATLAS detector at the CERN Large Hadron Collider at a centre-of-mass energy of 13 TeV. No significant excess of events over the background-only expectation was observed, and consequently, an upper limit at the 95% CL on the cross section times branching ratio of 6.6 times the Standard Model expectation was derived. The expected upper limits were determined to be 5.2 (4.4) times the Standard Model expectation when assuming the presence (absence) of this decay mode.

An alternative Higgs boson decay process that creates the same final state is $H \to \gamma^*(\ell^+\ell^-)\gamma$. In this process, the leptons are predominantly produced with much smaller invariant masses than in the $H \to Z(\ell^+\ell^-)\gamma$ decay. This results in reduced opening angles between the leptons, especially in the final state with electrons. In this thesis, a preliminary, blinded analysis targeting the $H \to \gamma^*(\ell^+\ell^-)\gamma$ decay with the ATLAS detector is presented for the first time. To retain events with close-by electron pairs, a dedicated identification algorithm was developed for such configurations. For an integrated luminosity of 80 fb^{-1} and an LHC centre-of-mass energy of 13 TeV, a significance of 0.9σ and upper limits at the 95% CL on the cross section times branching ratio of 2.9 (2.1) times the Standard Model prediction are expected when assuming the absence (presence) of this decay mode.

At a proton-proton collider, the Higgs boson can be produced via different production processes. About 1% of the Higgs production cross section at the LHC stems from the production in association with a pair of top quarks. This process is very interesting because it gives tree-level access to the top quark Yukawa coupling. The $t\bar{t}H(b\bar{b})$ final state is complicated to analyse because of the presence of multiple jets and b-tags. One of its major backgrounds is $t\bar{t}c\bar{c}$ production, which is theoretically challenging in its own right. A portion of this thesis consists of studies on the Monte Carlo simulation of $t\bar{t}c\bar{c}$ production with massive charm quarks included in the matrix elements.

Zusammenfassung

Das Higgs-Boson ist das zuletzt experimentell entdeckte Elementarteilchen. Aufgrund seiner hohen Masse und seiner Kopplung an massive Bosonen und Fermionen hat es eine Vielzahl von Zerfallskanälen. Einer davon ist der bislang nicht nachgewiesene seltene Zerfall $H \to Z(\ell^+\ell^-)\gamma$. In dieser Dissertation wird eine Suche nach dem $Z\gamma$-Zerfallskanal des Standardmodell-Higgs-Bosons sowie nach möglichen neuen Resonanzen mit größerer Masse, die in den gleichen Endzustand zerfallen, vorgestellt. Die Analyse basiert auf einem Datensatz von 36.1 fb^{-1} Proton-Proton-Kollisionen, die mit dem ATLAS-Detektor am Large Hadron Collider bei einer Schwerpunktsenergie von 13 TeV aufgezeichnet wurden. Es wurde kein signifikanter Überschuss an Ereignissen oberhalb des erwarteten Untergrundes beobachtet und daher wurde eine obere Grenze bei 95% Konfidenzniveau auf das Produkt von Wirkungsquerschnitt und Verzweigungsverhältnis vom 6.6-fachen der Standardmodellerwartung bestimmt. Die erwarteten oberen Grenzen betrugen das 5.2 (4.4)-fache der Standardmodellerwartung bei Annahme des Auftretens (Ausbleibens) dieses Zerfallskanals.

Ein alternativer Zerfallsprozess des Higgs-Boson, der den gleichen Endzustand erzeugt, ist $H \to \gamma^*(\ell^+\ell^-)\gamma$. Bei diesem Prozess werden die Leptonen vorzugsweise mit deutlich geringeren invarianten Massen produziert als beim $H \to Z(\ell^+\ell^-)\gamma$-Zerfall. Dies führt zu reduzierten Öffnungswinkeln zwischen den Leptonen, besonders im Endzustand mit Elektronen. In dieser Dissertation wird zum ersten Mal eine vorläufige, verblindete Analyse des $H \to \gamma^*(\ell^+\ell^-)\gamma$-Zerfalls mit dem ATLAS-Detektor vorgestellt. Um auch Ereignisse mit nahe beieinander liegenden Elektronen nutzen zu können, wurde ein spezieller Identifikationsalgorithmus für solche Konfigurationen entwickelt. Für eine integrierte Luminosität von 80 fb^{-1} und eine LHC-Schwerpunktsenergie von 13 TeV werden eine Signifikanz von 0.9σ und obere Grenzen bei 95% Konfidenzniveau auf das Produkt von Wirkungsquerschnitt und Zerfallsverhältnis vom 2.9 (2.1)-fachen der Standardmodellerwartung unter Annahme des Auftretens (Ausbleibens) dieses Zerfallskanals erwartet.

An einem Proton-Proton-Speicherring kann das Higgs-Boson mittels verschiedener Produktionsprozesse erzeugt werden. Etwa 1% des Higgs-Produktionswirkungsquerschnittes am LHC stammen von der assoziierten Produktion mit einem Top-Quark-Paar. Dieser Prozess ist sehr interessant, weil er einen Zugriff auf die Top-Yukawa-Kopplung auf dem Niveau von Baumgraphen erlaubt. Aufgrund der Vielzahl von Jets und b-Tags ist der $t\bar{t}H(b\bar{b})$ Endzustand kompliziert zu analysieren. Einer seiner Hauptuntergründe ist die Produktion von $t\bar{t}c\bar{c}$, die selbst theoretisch herausfordernd ist. Ebenfalls Bestandteil dieser Dissertation sind Studien zur Monte Carlo Simulation von $t\bar{t}c\bar{c}$ Produktion mit massiven Charm-Quarks in den Matrixelementen.

Acknowledgements

First of all, I would like to thank my advisors Kerstin Tackmann and Judith Katzy, who gave me the opportunity to explore the topics presented in this thesis under their guidance. Kerstin Tackmann supervised my work on the $H \to Z\gamma$ and $H \to \gamma^*\gamma$ analyses and Judith Katzy was my advisor for the work related to Monte Carlo, both during my authorship qualification as well as afterwards, when I was the contact person for the Herwig event generator in ATLAS. Thank you very much for sharing your experience, overviews and drive and for guiding and accompanying me during my thesis. I have always felt supported and that my work and my thoughts were appreciated. I am very grateful for the many fruitful discussions we had and for the chance to learn from your impressive knowledge of so many aspects of LHC physics, while at the same time enjoying large degrees of freedom and trust to work self-dependently. I would also like to sincerely thank Peter Schleper for agreeing to chair my supervision committee and to be a referee for this thesis.

My work on this thesis was performed as part of the ATLAS collaboration. Running an immense collaborative effort of more than 3000 scientists requires an impressive level of organisation, coordination and dedication and I have thoroughly enjoyed working with and learning from so many motivated people from all around the world. It is impossible to track and adequately praise the individual contributions to all aspects of ATLAS, including but not limited to design and construction of detector components, detector operation, data taking, software development and physics data analysis. Therefore, I would like to begin by collectively thanking the entire ATLAS collaboration, the LHC team and CERN for all the effort and dedication that ensure the successful operation of both the accelerator as well as the detector and in consequence made the work presented in this thesis possible.

From the innumerable ATLAS groups organising and carrying out all the work, I would like to particularly mention the $H/X \to Z\gamma$ analysis group and the Physics Modelling Group under whose tutelage this work was performed. While I extend my gratitude to everybody involved, I owe special thanks to Monica D'Onofrio, Yanping Huang, Josh McFayden, Anthony Morley, Marjorie Shapiro and Frank Siegert, who chaired and convened these groups at the time of my involvement. I am also grateful to Dominic Hirschbühl, Ewelina Maria Lobodzinska, Ludovic Scyboz and David Yallup for their collaboration and support when I was the contact person for the Herwig event generator in ATLAS. In the same capacity, I am greatly indebted to Johannes Bellm, Andreas Papaefstathiou, Simon Plätzer and Peter Richardson from the Herwig collaboration for their support and their patience in addressing my endless questions. My personal trajectory brought me from a master's thesis in theoretical particle physics to a PhD project in experimental particle physics and I have enjoyed very much getting to know aspects of both fields as well as their intersections. In this spirit, I would like to express my deep gratitude to Frank Tackmann who was always

available for discussions and readily shared his helpful insights. I am also very grateful to Nicolas Berger for his insights and his support in questions regarding statistics.

It is an invaluable privilege to work in an environment as knowledgeable and diverse as the ATLAS group at DESY. For practically all of my questions there would be a local expert around, able to provide more insights or direct me to further resources, and for this I am very grateful. I would like to direct a special thanks to the Higgs/photon and the top quark / Monte Carlo groups and particularly mention Nedaa Asbah, Jan-Hendrik Arling, Artem Basalaev, Martin Bessner, Daniela Börner, Früd Braren, Kurt Brendlinger, Lydia Brenner, Shuo Han, Judith Katzy, John Keller, Thorsten Kuhl, Jim Lacey, Roger Naranjo and Kerstin Tackmann. Further thanks for helpful discussions and sharing their thoughts on drafts of this thesis go to Nicolas Berger, Kurt Brendlinger, Daniela Börner, Anthony Morley, Shuo Han, Yanping Huang and John Keller.

Besides the impressive individual and collective expertise, being a part of DESY's ATLAS group has been an enormously memorable and enriching experience for me and I immensely appreciate the spirit, the diversity, the kindness and openness that characterise this group. In this spirit, I would like to direct a special thanks to our Hamburg group leader Ingrid-Maria Gregor, who not only ensures the smooth running of day-to-day business but also has a great way of keeping up a very pleasant and fun group atmosphere. And while I am grateful to the entire group for making my time so enjoyable, I would like to particularly mention my officemates and friends Dario Ariza, Nedaa Asbah, Martin Bessner, Früd Braren, Claire David, Nils Flaschel, Roger Naranjo, Kunlin Ran, Alessia Renardi, Annick Renevey, Stefan Richter, Han Shuo, Slavomira Stefkova and Sumera Yamin.

Finally, I would like to thank my parents and family. Without your support in all of my life and my previous education, this thesis would not have been possible.

Contents

Chapter 1

Introduction

For decades, the *Standard Model of particle physics* (SM) has been enormously successful at describing the fundamental particles and their interactions. In *high energy physics* (HEP), particles are brought to collision at ever higher centre-of-mass energies in order to test the SM and search for new particles or phenomena. Since its inception in the 1950s and 1960s, the SM has grown to accommodate all currently known fundamental particles, which comprise three generations of quarks and leptons as well as the gauge bosons mediating the electromagnetic, the weak and the strong interactions. The latest experimentally observed addition to the SM is the Higgs boson, which was discovered by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland in July 2012 [1, 2].

The discovery of the Higgs boson and the subsequent analysis of its properties was one of the main goals of the LHC programme. The LHC is currently the most powerful particle accelerator and has been in operation since the year 2009 [3]. It is located in the same tunnel with 27 km circumference that previously was home to the Large Electron-Positron Collider (LEP). In contrast to its predecessor, proton beams are accelerated and brought to collision at the LHC, at centre-of-mass energies as high as 13 TeV to date.

The most important motivation for the Higgs mechanism is the fact that many fundamental particles, e.g. the W and Z bosons, the electron e and the muon μ, have a non-zero mass, whereas the theoretical framework used to describe the particles and their interactions forbids the explicit introduction of particle masses. Therefore, a way was needed to incorporate the observed masses of the particles without making their description inconsistent. The resulting mechanism is referred to as the *Higgs mechanism*, named after Peter Higgs, who conceived it along with Robert Brout, François Englert, Gerald Guralnik, Carl Hagen and Tom Kibble [4–6]. While the theoretical foundations of the Higgs mechanism were worked out in the 1960s, the particle that it predicted, i.e. the Higgs boson H itself, evaded experimental observation for almost five decades. The reason for this is twofold: Firstly, its mass of 125 GeV requires a very powerful particle accelerator for its production and secondly, the variety and nature of its decays lead to the fact that an enormous

amount of experimental data was necessary to establish the existence of the Higgs boson with sufficient certainty. More specifically, the decay channels with the clearest experimental signatures occur very infrequently, while the most common Higgs decays lead to topologies that are hard to detect.

Despite their rarity, some of the channels with clear experimental signatures, such as the decays to two photons, $H \to \gamma\gamma$, or to four charged leptons, $H \to ZZ^* \to 4\ell$, with $\ell = e, \mu$, contributed the most to the Higgs signal discovered in July 2012. Since then, efforts were dedicated to the exploration of multiple aspects of the newly found particle. While ever larger datasets allow for an increasingly accurate determination of its properties in the most accessible channels, significant work has also been invested in searching for the rarer and the more abundant but complicated Higgs processes. Many of the more challenging decay channels were discovered in the following years. However, to this day, decay modes remain that have not yet been observed experimentally. This thesis is dedicated to aspects of some of the rare Higgs production and decay processes.

Among them is the decay $H \to Z\gamma$, which occurs slightly less frequently than the diphoton decay $H \to \gamma\gamma$, but is complicated by the fact that the Z boson, which is one of the mediators of the weak force, is not a stable particle. In the analysis presented in Part III of this thesis, a search for the $Z\gamma$ decay mode of the Higgs boson is performed using events in which the Z boson decays to pairs of electrons or muons. The same final state is also used to perform a search for a new particle X, which is assumed to be heavier than the Higgs boson and to also decay to the $Z\gamma$ final state.

The $\ell^+\ell^-\gamma$ final state can also be produced in Higgs decays to two photons, where one of the photons undergoes a $\gamma^* \to \ell^+\ell^-$ splitting. While on-shell photons are massless, the intermediate photon γ^* needs a non-vanishing invariant mass in order to decay. Since these invariant masses are much smaller than the mass of the Z boson, the $H \to \gamma^*\gamma$ process typically leads to smaller opening angles between the two charged leptons than the $H \to Z\gamma$ decay. This is especially true for the final state with electrons, as the smallness of the electron mass gives rise to lower invariant γ^* masses than the muon final state. However, the small angular separation leads to a particular signature in the detector, which makes this channel experimentally challenging. In Part IV of this thesis, the groundwork for this analysis is presented, which is attempted for the first time in ATLAS.

Besides its wide range of possible decay modes, the Higgs boson can be created via different production processes. At the LHC, they range from the dominating gluon-gluon fusion process, which accounts for almost 90% of the Higgs production cross section, to the much rarer production in association with a heavy quark. Of the latter, the $t\bar{t}H$ process, i.e. Higgs production in association with a top quark pair, is of particular importance and interest because it provides the most distinctive experimental signature and a direct access to the coupling of the Higgs boson to the top quark t, which is the heaviest of the quarks. At this point, it is worth noting that many analyses rely in one way or another on the simulation of expected signal and background contributions. The simulations are used in various ways, e.g. to estimate event selection efficiencies, to optimise the separation of signal and background contributions, to derive signal and background models or to compare the observed against the expected results. Analyses targeting the $t\bar{t}H$ topology make extensive use of such Monte Carlo simulations, which, due to the complexity of the associated signal and background processes, are quite difficult and subject to substantial uncertainties. Part II of this thesis presents aspects of Monte Carlo simulations in ATLAS as well as a study comparing different Monte Carlo simulations of the $t\bar{t}c\bar{c}$ background

process to the $t\bar{t}H(b\bar{b})$ topology.

The aforementioned chapters are preceded by an introductory part describing the relevant theoretical and experimental foundations. The Standard Model of particle physics and the Higgs mechanism are presented in Chapter 2 and the basics of Monte Carlo simulations are described in Chapter 3. These chapters are followed by descriptions of the LHC and the ATLAS experiment in Chapter 4, the reconstruction of particles in Chapter 5 and an introduction to the statistical methods used throughout this thesis in Chapter 6. After the presentation of the aforementioned analyses, this thesis ends with an overall conclusion in Chapter 17.

Own work and contributions

- Monte Carlo
  - Redesign of the `Herwig7` interface in the ATLAS software framework and contact person for the `Herwig++` and `Herwig7` Monte Carlo event generators in ATLAS
  - Studies on $t\bar{t}c\bar{c}$ production: Generation of samples with different settings, hard process and parton shower cut-off scales
- $H/X \to Z\gamma$ analysis
  - Development of the scheme for the evaluation of the theoretical uncertainties
  - Cross-checks of detector acceptance and event selection efficiencies and signal yields in the different categories
  - Internal documentation and interactions with the review committee
- $H \to \gamma^*\gamma$ analysis
  - Preliminary analysis design and analysis software development
  - Development of a dedicated identification algorithm for close-by electrons
  - Development of a software package for the statistical analysis and evaluation of the expected significance and upper limits on the signal strength as well as projections for larger datasets

Remarks on the notation

In this thesis, scalar quantities are expressed with regular lowercase symbols, while vector quantities are denoted with bold lowercase and matrices and tensors with regular capital symbols. Lists of the mathematical symbols as well as the acronyms used throughout are included after the appendix.

Part I.

Theoretical and Experimental Foundations

Chapter 2

The Standard Model and the Higgs Mechanism

At the beginning of this chapter, a short introduction of the Standard Model of particle physics will be given, followed by a brief account of the physics motivations for the Higgs mechanism and a sketch of the principle of spontaneous electroweak symmetry breaking. The chapter will conclude with an overview of the properties of the Higgs boson and the current status of Higgs measurements and searches for not yet observed decay signatures.

2.1. The Standard Model of Particle Physics

Three of the four currently known fundamental forces of nature are treated in a unified manner in the Standard Model of particle physics [7]. It describes the electromagnetic, the weak, and the strong forces as interactions mediated by their corresponding *gauge bosons*, which have a spin quantum number of one. The matter particles on which those forces act are the *fermions* and have spin one half. Both the fermions as well as the gauge bosons are shown in Figure 2.1 along with their respective masses and electric charges. The fermions are organised in three generations of increasing mass. Each generation contains a pair of leptons, consisting of an electrically charged lepton and a corresponding uncharged neutrino, and a pair of quarks, whose electric charges are multiples of a third of the elementary charge e. Besides their electric charge, quarks also carry one of the three colour charges of the strong interaction. They participate in all three fundamental interactions while the charged leptons only interact via the electromagnetic and weak interactions and the neutrinos are only subject to the weak interaction. The photon is the massless gauge boson that mediates the electromagnetic force. The Z boson and the two charged $W^{\pm}$ bosons are the gauge particles of the weak interaction and have masses of the order of 100 GeV. Finally, the strong interaction is mediated by the eight gluons, which are massless and carry the colour charge of the strong interaction. Due to their colour charge, gluons self-interact. This is also true for the gauge bosons of the weak interaction,

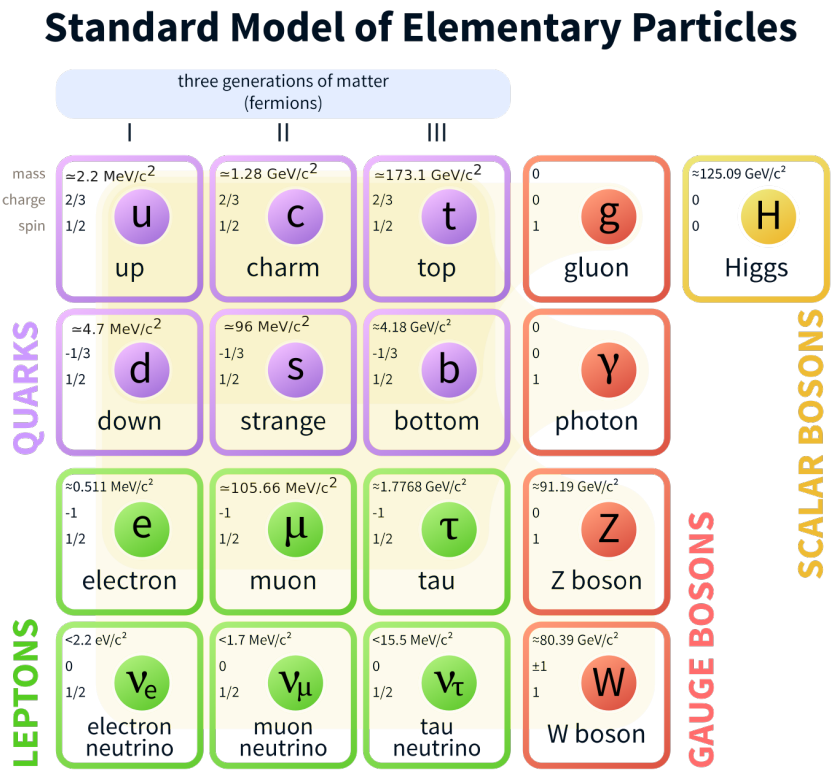


Figure 2.1.: Overview of the particle content of the Standard Model of particle physics. Taken from [9].

where a Z boson couples to a pair of W bosons. The latest addition to the SM is the Higgs boson which is the only fundamental scalar particle. It will be discussed in more detail later in this chapter.

The electromagnetic and strong interactions respect the discrete P, C, and T symmetries individually, i.e. they are invariant under transformations that

- parity P: reverse all spatial directions, i.e. $(t, \boldsymbol{x}) \to (t, -\boldsymbol{x})$,
- charge conjugation C: exchange particles and antiparticles,
- time reversal T: reverse time, i.e. $(t, \boldsymbol{x}) \to (-t, \boldsymbol{x})$.

This, however, is not the case for the weak interaction, which, in fact, maximally violates the P and C symmetries and depends on the chirality of the interacting fermions. Specifically, only particles with left-handed chirality and antiparticles with right-handed chirality will participate in the weak interaction. Chirality is an intrinsic property of a particle and does not depend on the reference frame. In general, it is not exactly the same as the helicity h, which is defined as the projection of the spin $\boldsymbol{s}$ onto the direction of the momentum $\hat{\boldsymbol{p}} = \boldsymbol{p}/|\boldsymbol{p}|$, $h = \boldsymbol{s} \cdot \hat{\boldsymbol{p}}$. Therefore, for massive particles, helicity is frame-dependent, while for massless particles, helicity and chirality are identical.[1]

The SM is based on the mathematical framework of quantum field theory, which combines the theory of special relativity with quantum mechanics. The gauge group of the full SM

[1] A very pedagogical illustration of the concepts of helicity and chirality is given in [8].

is

$$SU(3)_C \times SU(2)_L \times U(1)_Y \ , \tag{2.1}$$

where $SU(3)_C$ is the gauge group of the strong interaction described by the theory of quantumchromodynamics (QCD) and $SU(2)_L \times U(1)_Y$ is the gauge group of the unified electroweak forces. These groups describe the symmetry transformations of the different interactions that leave the SM Lagrange density invariant. The index C refers to the colour charge of QCD, while the index L hints at the fact that only particles with left-handed chirality are subject to the weak interaction. The index Y represents the *weak hypercharge*, which is given by $Y = 2(Q - I_3)$, where Q denotes the electric charge and I_3 the third component of the *weak isospin*. In the $SU(2)$ space of the electroweak interactions, the fermions with left-handed chirality are written in doublets, whose two components have $I_3 = \pm 1/2$.

In quantum field theory, the bosons mediating an interaction and the participating matter particles can be related by the principle of *local gauge symmetry*. In the example case of a fermion field $\psi(x)$ of mass m, the free, non-interacting Lagrange density is given by

$$\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \ , \tag{2.2}$$

where $\bar{\psi} = \psi^\dagger \gamma_0$ and the γ^μ are gamma or Dirac matrices. This Lagrange density is invariant under the *global gauge transformation*

$$\psi(x) \to \psi'(x) = \psi(x) e^{i\theta} \ , \tag{2.3}$$

where θ is a real-valued number, corresponding to a $U(1)$ symmetry group. When promoted to a local gauge transformation

$$\psi(x) \to \psi'(x) = \psi(x) e^{i\theta(x)} \ , \tag{2.4}$$

the derivative in the kinetic term $\bar{\psi} i\gamma^\mu \partial_\mu \psi$ of the Lagrangian will lead to an additional term that spoils the invariance of the free Lagrange density $\mathcal{L}_0$. The principle of local gauge symmetry is enforced by adding another term to the Lagrangian of the theory which is constructed in such a way that it restores the invariance under the local gauge transformation. This is done by introducing a new field $A_\mu(x)$ with a defined transformation behaviour under the gauge transformation so that the full Lagrangian is given by

$$\mathcal{L} = \mathcal{L}_0 - e\bar{\psi}\gamma^\mu A_\mu \psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.5}$$

with

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \frac{1}{e}\partial_\mu \theta(x) \quad \text{and} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{2.6}$$

$F_{\mu\nu}$ is the field strength tensor and e describes the strength of the interaction. In the case of *quantumelectrodynamics* (QED), the coupling strength e is given by the fundamental electric charge. The field A_μ introduces the gauge boson related to the interaction, i.e. the photon in the case of QED or the gluons in the case of QCD. In other words, by enforcing local gauge invariance an interaction between the matter particles is introduced into the gauge theory. This principle is one of the foundations of quantum field theories. $D_\mu = \partial_\mu + ieA_\mu$ is called the *covariant derivative*.

Based on the particle content and the interactions as encoded by the respective Lagrange densities, calculations of observables such as total and differential cross sections or branching ratios are carried out in *perturbation theory* using *Feynman diagrams*. The possible

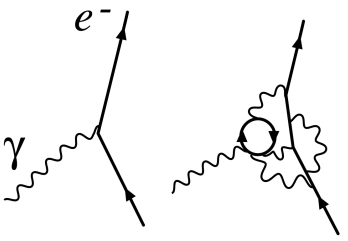


Figure 2.2.: Tree-level diagram (left) and example Feynman diagram with higher-order corrections (right) to the QED coupling α. Taken from [11].

ways in which a certain particle reaction may take place are expanded in powers of the coupling parameter and the resulting series is truncated at a given order. If only the lowest possible order in the coupling parameters is considered, the calculation is said to be carried out at the *leading order* (LO). If terms with one more power of the coupling parameter are added, this is referred to as the *next-to-leading order* (NLO). Each contributing *amplitude* term at each order can be represented graphically by the Feynman diagrams. If a diagram does not contain any closed internal particle loops, it is referred to as a *tree-level diagram*. All allowed patterns as well as their corresponding mathematical expressions can be derived from the Lagrangian of a given quantum field theory. This is described in many standard textbooks, e.g. [10]. Just like with Taylor expansions, adding higher order terms to the calculation is usually expected to result in a better approximation of the exact all-orders result. In general, the smaller the numerical value of the coupling parameter, the more suppressed higher-order terms are and the more rapid the convergence is. While this is especially true for quantumelectrodynamics, the theory of the electromagnetic interaction, the success of this approach is not always guaranteed in QCD as will be described in Section 2.1.1.

A priori, the couplings of the electromagnetic, weak and strong interactions described in the SM are constant parameters of the respective Lagrange densities. However, adding higher-order corrections exposes the effect that the couplings depend on an energy scale, the so-called *renormalisation scale*. This behaviour is commonly referred to as *running coupling*. The procedure of renormalisation essentially consists in redefining the mass and charge parameters in the Lagrange densities so that their initial, bare values are infinite and are subsequently rendered finite and moved to the experimentally observed values by taking into account the aforementioned higher-order corrections. More details as well as the different commonly used renormalisation schemes are detailed in textbooks, e.g. [10]. Let it be noted that the renormalisation scale is an unphysical scale that is introduced in the course of the renormalisation procedure. There is no prescription from first principles that dictates the choice of the functional form of the renormalisation scale for a given calculation. Therefore, usually a choice is made that is related to the kinematics of the process under study. This choice will have an impact on the exact numerical results of the calculations. Consequently, variations of the scale are commonly used to assess the uncertainty due to the truncation of the perturbation series. It is expected that with increasing order of the expansion the dependence on the renormalisation will decrease.

2.1.1. Quantumchromodynamics

QCD is the gauge theory of the strong interaction, which acts on *partons*, i.e. on gluons and quarks. The charge of QCD is called *colour charge* and, inspired by the dimensionality of

the $SU(3)$ gauge group, these (anti-)colours are commonly labelled (anti-)red, (anti-)green and (anti-)blue. As already mentioned, the gluons carry colour charge. In combination with the number of quark flavours ($N_f = 6$ in the SM), this leads to two particular and related properties of the strong interaction. In contrast to the electromagnetic coupling, the strong coupling decreases towards higher scales (*asymptotic freedom*) and increases towards smaller energies, which leads to *confinement*. Because of the latter, quarks and gluons do not exist as freely propagating particles but are confined into colour-neutral hadrons. Another consequence of this behaviour is that fixed-order QCD perturbation theory only is a reasonable approach for sufficiently energetic reactions. Towards lower energies, the increasing QCD coupling leads to the production of many, predominantly collimated particles, the so-called *jets*. In fully exclusive Monte Carlo simulations this effect is approximated by *parton showers* and *hadronisation* models which will be described in more detail in Chapter 3. Eventually, the QCD coupling diverges at the so-called *Landau pole* which is located at approximately 220 MeV [12, 13].

This behaviour of the strong coupling drastically limits the applicability of pure perturbative QCD calculations. In contrast, pure QED processes, such as e.g. Bhabha scattering ($e^+e^- \to e^+e^-$) or the annihilation of electrons into a pair of muons ($e^+e^- \to \mu^+\mu^-$), can be approximated well by fixed-order perturbative calculations involving only the corresponding fundamental particles. If a process involves partons in the initial state, then already the first complication arises: Due to confinement, these initial state particles will not be available as pure beams of fundamental particles, but rather will have to be extracted from some type of hadron. Likewise, no fundamental final-state QCD particles but rather collimated sprays of hadrons will be observed in experimental detectors. To some degree, fixed-order perturbation theory can be used to describe these hadronic jets, because certain properties of the outgoing partons are preserved in the hadronic final state. This *parton-hadron duality* holds for sufficiently inclusive quantities, e.g. the transverse momentum of a sufficiently hard and isolated jet can be approximated by the momentum of the parton it originated from. The duality breaks down, however, when the transverse momentum of the jet becomes too low or when the jets are not sufficiently separated from each other.

The extraction of partons from initial-state hadrons is separated from the perturbative parton interactions using the concept of *factorisation*. A hadronic cross section σ is split into a hadron-level part and the perturbative *hard scattering process* $\hat{\sigma}_{ab}$ that deals with the highly energetic partonic reactions. For the case of collisions of two hadron beams with incoming momenta P_A and P_B this can be written as

$$\sigma = \sum_{a,b} \int_0^1 f_{a/A}(x_a, \mu_F^2) \int_0^1 f_{b/B}(x_b, \mu_F^2) \int \hat{\sigma}_{ab}(x_a P_A, x_b P_B, \phi_f)\ \mathrm{d}x_a \mathrm{d}x_b \mathrm{d}\phi_f\ . \tag{2.7}$$

The sums over a and b run over all possible parton types that may lead to the desired process. The functions $f_{a/A}(x_a, \mu_F^2)$ are the so-called *parton distribution functions* (PDFs). They are universal, i.e. independent of the hard process, and at the LO can be interpreted as the probability density of extracting a parton of type a with a momentum fraction of x_a from a certain type of parent hadron A. They depend on another non-physical scale, the *factorisation scale* μ_F which, like the renormalisation scale μ_R, is usually set to a scale reflecting the kinematics of the process under study. Since partons with very different momentum fractions may contribute, the fractions x_a and x_b are integrated over. The hard scatter process will depend on the momenta of the incoming partons $p_a = x_a P_A$ and $p_b = x_b P_B$ and the phase space ϕ_f of the outgoing final state particles. Due to

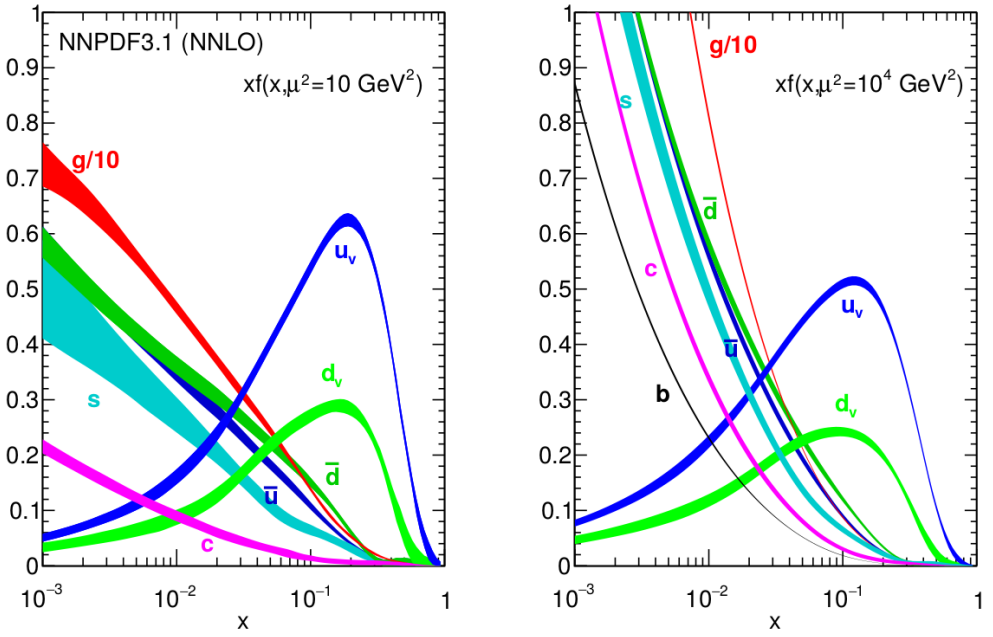


Figure 2.3.: The NNPDF3.1 NNLO PDFs for values of the factorisation scale of $\mu_F^2 = 10\ \text{GeV}^2$ (left) and $\mu_F^2 = 10^4\ \text{GeV}^2$ (right). For display purposes $x \cdot f(x, \mu_F^2)$ is shown instead of $f(x, \mu_F^2)$. Taken from [18].

the large coupling and due to the fact that they describe the dynamics of hadrons, the PDFs cannot be calculated from first principles. But since they are independent of the hard process they can be extracted from experimental data and then be used to calculate predictions for different processes. Also, their evolution from one value of the factorisation scale to another is known and can be calculated perturbatively up to a certain order. A number of collaborations work on the determination of PDFs, some commonly used modern PDFs are the ABMP [14], CT [15], HERAPDF [16], MMHT [17] and NNPDF [18] sets. Some example PDF distributions are shown in Figure 2.3. When evaluating PDF-related uncertainties, variations of the factorisation scale μ_F can be used to assess uncertainties due to the truncation of the perturbation series in the evolution of the PDFs from one factorisation scale to another. Moreover, one given PDF set not only contains the best-fit PDF for each parton type but also a number of auxiliary PDFs that can be used to assess uncertainties such as those due to the modelling and the fitting of the PDFs to the experimental data.

A subtle aspect regarding the interplay between the PDFs and the hard process is the choice of the *flavour scheme* for heavy quarks. There are two options for a given heavy quark flavour: It can either be considered massless and included in the proton and hence the PDFs. In this case, the quark can be part of the hard process initial state. Alternatively, the heavy quark can be treated as massive in the hard process. In this case, it is not included in the proton and the PDFs and can only appear in the final state.[2] The number of light quark flavours is indicated in the name of the flavour scheme, i.e. in the *5-flavour scheme* (5FS), u, d, s, c and b-quarks are considered massless and included in the PDF evolution, while in the *4-flavour scheme* (4FS), the b-quark is treated as massive. In Chapter 8, the *3-flavour scheme* (3FS) will be used, where also the c-quark is considered

[2]Some work to overcome the limitations of these approaches is presented in [19].

massive.

This entire framework for QCD calculations, including the PDFs and the principle of factorisation, is called the *parton model of QCD*. In analogy to the PDFs, the confinement of final state partons into outgoing hadrons may be described by *fragmentation functions* that describe the probability of forming a certain hadron from a parton with a given momentum. As hinted at before, fully exclusive Monte Carlo simulations use parton showers and hadronisation models instead.

2.2. Motivations for the Higgs Mechanism

One milestone in the formulation of the SM was the unification of the electromagnetic and the weak forces by Sheldon Glashow in 1961 [20] based on a $SU(2)_L \times U(1)_Y$ group structure. The associated Z and $W^{\pm}$ vector bosons mediating the weak force were discovered in 1983 by the UA1 [21, 22] and UA2 [23, 24] collaborations at CERN's $Sp\bar{p}S$ accelerator and were found to have masses of $m_Z = 91.2$ GeV and $m_W = 80.4$ GeV [7]. But already in 1961 it was known that the electron and the muon have a non-zero mass. Since explicit mass terms in QFT Lagrangians violate gauge symmetry, this raised the question of how to reconcile the experimental observation of massive fundamental particles and their mathematical description in the framework of quantum field theory. In the same year, Julian Schwinger expressed for the first time the idea that this is indeed possible [25]. Philip Anderson then proposed a non-relativistic formulation of spontaneous symmetry breaking in 1962 [26] that avoided explicit violation of gauge invariance due to non-zero particle masses. Relativistic formulations followed in 1964 in the seminal papers by François Englert and Robert Brout [4], Peter Higgs [5] and Gerald Guralnik, Carl Hagen and Tom Kibble [6] and were incorporated into the electroweak theory by Abdus Salam and Steven Weinberg [27] in 1967.

Besides the non-zero gauge boson masses, there is another motivation for the Higgs mechanism. Without the Higgs boson, the cross section for the scattering of longitudinally polarised W and Z bosons would increase indefinitely with the centre-of-mass energy.

2.3. Electroweak Symmetry Breaking

Consider Yang-Mills-Theory as a minimal example [10, 29], i.e. a complex-valued scalar field $\phi(x)$ and the Lagrange density

$$\mathcal{L} = (\partial^\mu \phi^*)(\partial_\mu \phi) - V(\phi) \quad \text{with} \quad V(\phi) = \mu^2 \phi^* \phi + \lambda^2 (\phi^* \phi)^2 \ ,$$

which is obviously invariant under global unitary $U(1)$ gauge transformations

$$\phi(x) \to \phi'(x) = \phi(x) e^{i\theta} \ .$$

For $\lambda^2 > 0$ and $\mu^2 < 0$ the potential adopts the characteristic *Mexican hat shape* shown in Figure 2.4. The state of lowest energy is then located on a circle with $|\phi| = \sqrt{-\mu^2/2\lambda^2} =: v/\sqrt{2}$, where v is called *vacuum expectation value*. Once a specific ground state ϕ_0 is chosen, this particular configuration stands out, however, which breaks the gauge symmetry. Hence the name *spontaneous symmetry breaking*. The field $\phi(x)$ can be reparameterised as $\phi(x) = (v + H(x)) \cdot e^{i\varphi(x)}$ where $H(x)$ and $\varphi(x)$ are real-valued fields. $H(x)$ then describes a massive

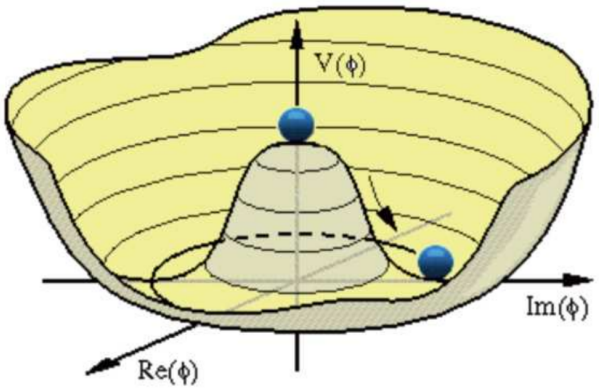


Figure 2.4.: Mexican hat shape of the Higgs potential. Taken from [28].

excitation of the potential and $\varphi(x)$ a massless field, the so-called *Goldstone boson*. The homonymous *Goldstone theorem* [30–32] also applies to higher-dimensional symmetries and predicts the existence of one massless boson for each spontaneously broken symmetry. When enforcing local gauge invariance, the Goldstone boson will combine with a massless spin-1 field to form a massive spin-1 particle. This is precisely the result found by Englert, Brout, Higgs, Guralnik, Hagen and Kibble [4–6].

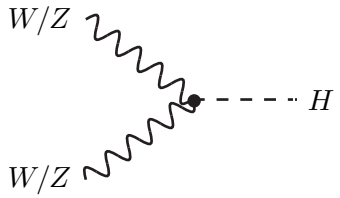
In the SM, the field $\phi(x)$ is a complex doublet, i.e. a field with four degrees of freedom. The Lagrangian containing ϕ has two symmetries, the $SU(2)_L$ symmetry, corresponding to the weak isospin, and the $U(1)_Y$ symmetry, corresponding to the weak hypercharge. The associated gauge bosons are the W_1, W_2, and W_3 bosons and the B field. With this particular group structure, choosing a specific vacuum ground state spontaneously breaks three of the four symmetries of the $SU(2)_L \times U(1)_Y$ group. Only one specific linear combination of the W_3 and B bosons remains unaffected, its gauge group is $U(1)_{\mathrm{em}}$. Thus, there are three would-be Goldstone bosons leading to three massive gauge bosons, while the fourth gauge boson remains massless. Identifying the different mass eigenstates yields the $W^{\pm}$ bosons as a combination of the W_1 and W_2 bosons and the Z boson as a combination of the W_3 and B fields. The other remaining linear combination of W_3 and B has a vanishing mass term and corresponds to the photon [29]. The fourth degree of freedom of the complex doublet ϕ gives rise to a massive scalar field, the Higgs boson.

2.4. Properties of the Higgs Particle

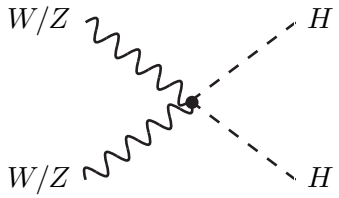
The couplings of the Higgs particle to the $W^{\pm}$ and Z bosons are determined by the kinetic term of the electroweak part of the SM Lagrangian,

$$\mathcal{L}_{\mathrm{EW}} \supset \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) \supset \left(m_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu\right) \cdot \left(1 + \frac{H}{v}\right)^2 , \tag{2.8}$$

where D_μ denotes the covariant derivative which contains terms including the $W^{\pm}$ and Z bosons as well as the photon. Since the photon is massless, there is no tree-level coupling to the Higgs boson. From these terms, the following interaction vertices and couplings arise:

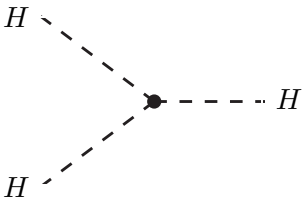


WWH: $\frac{2m_W^2}{v}$ $\quad$ ZZH: $\frac{m_Z^2}{v}$

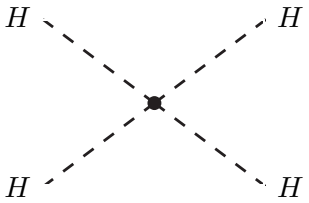


$WWHH$: $\frac{m_W^2}{v^2}$ $\quad$ $ZZHH$: $\frac{m_Z^2}{2v^2}$

From the aforementioned Higgs potential $V(\phi)$, trilinear and quartic self-interactions of the Higgs particle ensue:



HHH: $\frac{m_H^2}{2v}$

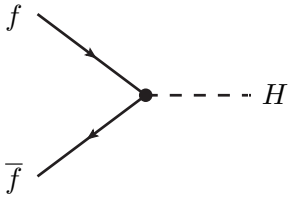


$HHHH$: $\frac{m_H^2}{8v^2}$

While the mechanism outlined so far yields massive bosons, the fermion masses have to be introduced in a different way, by means of the *Yukawa couplings*. They are added via the terms

$$\mathcal{L}_{\text{Yukawa}} = -h_{d_{ij}}\bar{q}_{L_i}\phi d_{R_j} - h_{u_{ij}}\bar{q}_{L_i}i\sigma_2\phi u_{R_j} - h_{\ell_{ij}}\bar{\ell}_{L_i}\phi e_{R_j} \ , \tag{2.9}$$

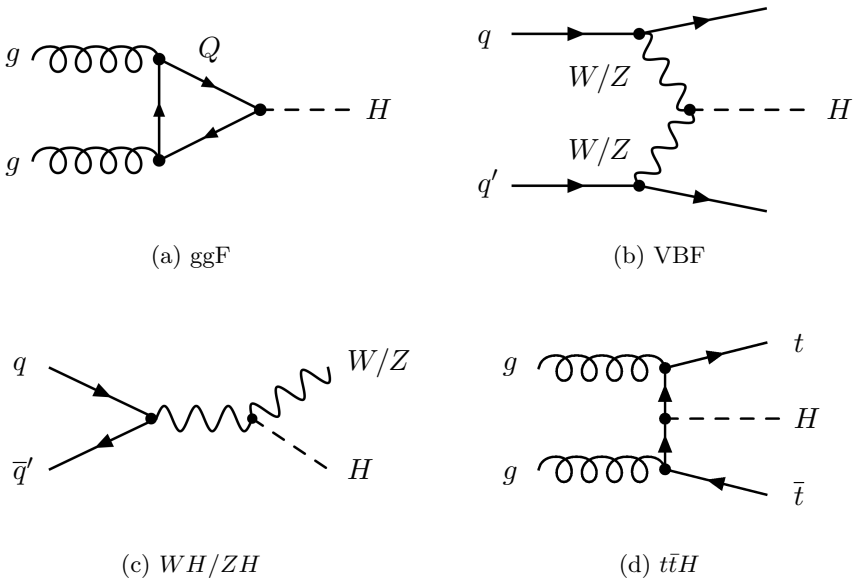
where q_L are the left-handed quarks, u_R and d_R the right-handed quarks, ℓ_L the left-handed leptons, e_R the right-handed charged leptons and σ_2 the second Pauli matrix. The matrices h_d, h_u and h_ℓ encode the Yukawa couplings between the Higgs boson and the massive fermions corresponding to the vertex:



$f\overline{f}H$: $\frac{m_f}{v}$

From all these vertices it can be seen that the coupling strengths to the Higgs boson are determined by the masses of the interacting particles. This has important implications on the relative contributions of the different Higgs production and decay processes. Based on these interactions, the Higgs boson can be produced in four main types of processes at the LHC:

- gluon-gluon fusion (ggF)
- vector boson fusion (VBF)
- production in association with a massive vector boson, i.e. WH and ZH production, also collectively referred to as VH production or *Higgsstrahlung*
- production in association with a heavy quark, where the $t\bar{t}H$ channel gives the most distinctive signature



(a) ggF (b) VBF

(c) WH/ZH (d) $t\bar{t}H$

Figure 2.5.: LO example diagrams for Higgs production.

Example LO diagrams for all production modes are shown in Figure 2.5. The production cross sections and the relative contributions to the total Higgs production cross section are given in Table 2.1 and shown as a function of the centre-of-mass energy in Figure 2.6. The most frequent production mode is gluon-gluon fusion with a contribution of almost 90% to the total Higgs production cross section. This process contains a loop already at the leading order. This leads to significant theoretical uncertainties from QCD scale variations of still about 4% at the N^{3}LO as listed in Table 12.2. Non-SM particles in the loop may lead to modifications of the cross section or deviations from the expected kinematic distributions. Besides its abundance, this possible BSM sensitivity makes the ggF channel interesting. The next most abundant process is VBF, which has a very characteristic event topology with two forward back-to-back jets and a gap between them, where little jet activity is expected. These jets as well as the jets or leptons coming from the vector bosons in the VH production channels can be used as part of the event selection strategies in experimental analyses. Finally, the $t\bar{t}H$ production mode provides a distinctive topology for directly accessing the tree-level coupling of the Higgs boson to fermions.

The aforementioned couplings are also responsible for the possible decay channels of the Higgs boson. The final decay products have to be on their mass shells, whereas decays with off-shell intermediate particles are suppressed. Therefore, due to its relatively low mass of approximately 125 GeV, the Higgs boson cannot decay to a top quark pair. This makes $H \to b\bar{b}$ the most frequent decay channel with a fraction of almost 60%. Despite its abundance, this decay channel is extremely challenging experimentally because of the very large SM jet background at the LHC. The decay through two W bosons benefits from a cleaner signature with two charged leptons and two neutrinos but does not allow for a full reconstruction of all Higgs decay products and therefore has a very limited mass resolution. The decay to two gluons proceeds via a top loop, but still has a sizeable fraction of about 9%. Due to the SM jet background, it is experimentally inaccessible at the LHC, though. Two very interesting channels for experimental analyses are the decays to two photons and to four charged leptons via two intermediate Z bosons. Despite being relatively rare, these channels have clear signatures, manageable ($\gamma\gamma$) or even very low ($ZZ \to 4\ell$) backgrounds

$\sqrt{s}$	7 TeV		8 TeV		13 TeV		14 TeV	
Production mode	σ [pb]	f [%]	σ [pb]	f [%]	σ [pb]	f [%]	σ [pb]	f [%]
ggF	16.83	88.2	21.39	88.2	48.51	88.1	54.60	88.1
VBF	1.240	6.5	1.600	6.6	3.779	6.9	4.275	6.9
WH	0.576	3.0	0.701	2.9	1.370	2.5	1.510	2.4
ZH	0.338	1.8	0.420	1.7	0.882	1.6	0.984	1.6
$t\bar{t}H$	0.089	0.5	0.133	0.6	0.507	0.9	0.613	1.0
Total	19.1		24.2		55.0		62.0	

Table 2.1.: Inclusive Higgs cross sections σ and relative contributions f to the total inclusive Higgs cross section for the different production modes at different LHC centre-of-mass energies $\sqrt{s}$. The numbers are given for a Higgs mass of 125.09 GeV and are taken from [33].

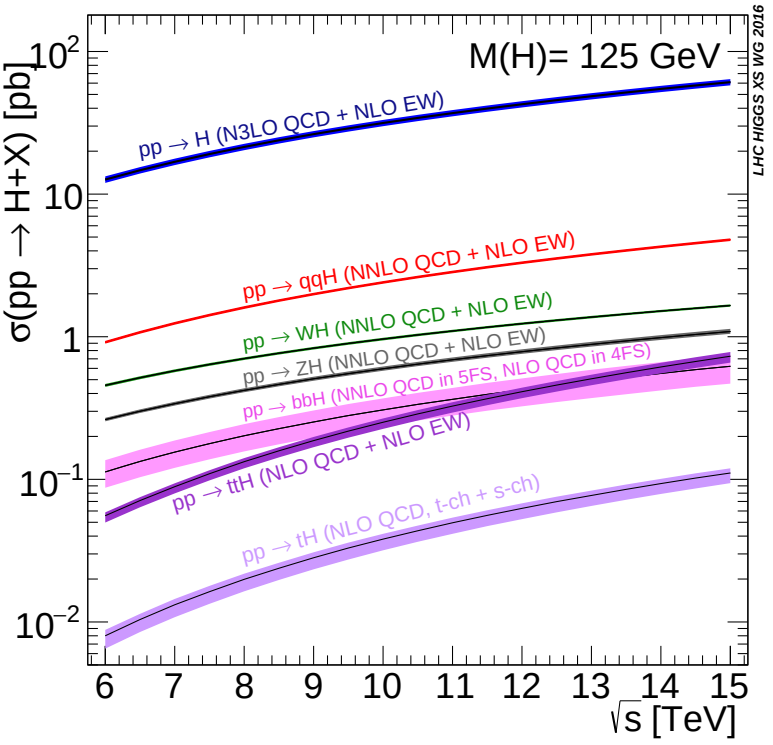


Figure 2.6.: Standard Model Higgs boson production cross sections as a function of the centre-of-mass energy. Taken from [34].

Decay channel	Branching fraction
$H \to b\bar{b}$	58.1 %
$H \to W^+W^{-*}$	21.5 %
$H \to gg$	8.18 %
$H \to \tau^+\tau^-$	6.26 %
$H \to c\bar{c}$	2.88 %
$H \to ZZ^*$	2.64 %
$H \to \gamma\gamma$	0.227 %
$H \to Z\gamma$	0.154 %
$H \to \mu^+\mu^-$	0.0217%
Total decay width	4.07 MeV

Table 2.2.: Higgs boson branching fractions for a mass of $m_H = 125$ GeV [33].

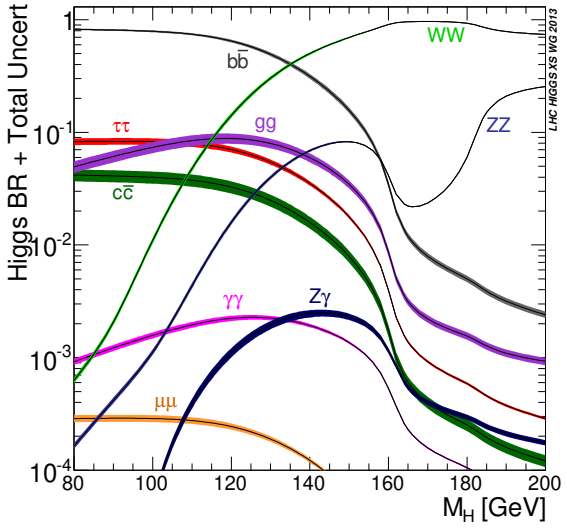


Figure 2.7.: Standard Model Higgs boson branching fractions as a function of the mass. Taken from [34].

and allow for a full reconstruction of the Higgs decay products. These features make those channels most suited for the determination of the Higgs mass. The branching fractions for all decay modes are given in Table 2.2 and shown as a function of the Higgs mass in Figure 2.7. Overall, the Higgs boson has a very narrow decay width of approximately 4 MeV.

2.5. Current Status of Experimental Measurements and Searches

The discovery of a new boson in the search for the Higgs boson was announced by ATLAS [1] and CMS [2] on July 4th, 2012. The corresponding analyses are based on a combination of the results from the channels $H \to \gamma\gamma$, $H \to ZZ^*$, $H \to W^+W^{-*}$, $H \to \tau^+\tau^-$ and $H \to b\bar{b}$. Ever since then, an enormous effort has been put into the

scrutiny of the properties of this particle and into checking whether it is compatible with the SM prediction for the Higgs boson. So far all analyses indicate that this is indeed the case within the uncertainties.

One aspect under investigation is the mass of the new boson. Based on the full LHC Run 1 dataset at centre-of-mass energies at 7 and 8 GeV and the decay modes $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$, ATLAS and CMS determined the mass to be $m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})$ GeV [35]. Ongoing LHC Run 2 measurements have not yet reached a higher level of precision [36,37]. Other important aspects are the spin and parity quantum numbers. In [38] and [39], various spin and parity combinations were tested and all configurations except the SM one, $J^P = 0^+$, were excluded.[3,4]

The couplings of the Higgs boson to other particles were treated in the so-called *κ-framework* in Run 1 of the LHC. The coupling modifiers are defined as

$$\kappa_i^2 = \frac{\sigma_i}{\sigma_i^{\text{SM}}} , \qquad \kappa_f^2 = \frac{\Gamma_f}{\Gamma_f^{\text{SM}}} \qquad \text{and} \qquad \kappa_H^2 = \frac{\Gamma}{\Gamma^{\text{SM}}} \tag{2.10}$$

for a production process i, a decay process f and - if only SM decays of the Higgs boson are allowed - the total decay width. Both tree-level couplings as well as the effective couplings relevant for $gg \to H$ production and the $H \to \gamma\gamma$ decay can be treated in this manner. Based on Run 1 data and considering all Higgs production modes as well as the decays to ZZ, WW, $\gamma\gamma$, $\tau\tau$, bb, and $\mu\mu$, all couplings are compatible with the SM Higgs boson hypothesis [41].

Going beyond coupling modifiers, the *simplified template cross section* (STXS) and the differential cross section measurements provide a more detailed look into the kinematic properties of the Higgs particle. Measurements at 13 TeV were performed in the $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ channels by ATLAS [42–46] and CMS [37,47]. In the STXS analyses, the events are divided into a set of mutually exclusive categories targeted at different Higgs production modes as well as, if possible, kinematic properties within a given mode. For each of the different categories, the cross section is extracted and compared against the SM expectation. Differential cross section measurements are performed in a defined fiducial volume and do not divide the events into different categories. Typical observables include the transverse momentum and the rapidity of the Higgs boson, the number of jets produced as well as the transverse momentum of the leading jet. All these results again show compatibility with the SM Higgs boson.

While the more accessible final states and more abundant production modes are being exploited to gain a better understanding of the Higgs boson properties, searches for the rarer production modes and decay channels continue. One of the most important milestones in these efforts is the observation of the $t\bar{t}H$ production mode. Combining the $b\bar{b}$, WW, $\tau\tau$, $\gamma\gamma$, and ZZ decay channels, ATLAS reports an observation with a significance of 5.8 standard deviations [48]. The latest corresponding CMS publication reports an observed significance of 5.2 standard deviations [49].

A comprehensive overview of the results of experimental analyses by ATLAS and CMS for the different production modes and decay channels of the Higgs boson is given in Tables 2.3 and 2.4.

[3] If C-parity is assumed to be conserved, the decay of the Higgs boson to two photons indicates that the Higgs boson has positive $C = +1$.

[4] However, CP-even ($J^{CP} = 0^{++}$) and CP-odd ($J^{CP} = 0^{+-}$) admixtures to the SM Higgs boson are not ruled out by data in the $H \to WW$ decay channel [40].

Production mode	Status in ATLAS	CMS
ggF	observed arXiv:1507.04548 [50]	observed, 6.6σ (7.4σ) arXiv:1412.8662 [51]
VBF	observed, 5.4σ (4.6σ) arXiv:1606.02266 [41]	
VH	observed, 5.3σ (4.8σ) 1808.08238 [52]	
$t\bar{t}H$	observed, 5.8σ (4.9σ) arXiv:1806.00425 [48]	observed, 5.2σ (4.2σ) arXiv:1804.02610 [49]

Table 2.3.: Status of experimental analyses of Higgs production modes. Expected significances are given in parentheses.

Decay channel	Status in ATLAS	Status in CMS
$H \to b\bar{b}$	observation, 5.4σ (5.5σ) 1808.08238 [52]	observation, 5.6σ (5.5σ) arXiv:1808.08242 [53]
$H \to WW^*$	observation, 6.1σ (5.8σ) arXiv:1412:2641 [54]	observation, 9.1σ (7.1σ) arXiv:1806.05246 [55]
$H \to \tau\tau$	observation, 6.4σ (5.4σ) arXiv:1811.08856 [56]	observation, 4.9σ (4.7σ) arXiv:1708.00373 [57]
$H \to c\bar{c}$	upper limit, $\mu_{95\%} \leq 110(150)$ arXiv:1802.04329 [58]	
$H \to ZZ$	observation, 6.6σ (4.4σ) arXiv:1307.1427 [59]	observation, 6.5σ (6.3σ) arXiv:1412.8662 [51]
$H \to \gamma\gamma$	observation, 7.4σ (4.3σ) arXiv:1307.1427 [59]	observation, 5.6σ (5.3σ) arXiv:1412.8662 [51]
$H \to Z\gamma/\gamma^*\gamma$	upper limit, $\mu_{95\%} \leq 6.6(5.2_1)$ arXiv:1708.00212 [60]	upper limit, $\mu_{95\%} \leq 3.9(2.9_1)$ arXiv:1806.05996 [61]
$H \to \mu\mu$	upper limit, $\mu_{95\%} \leq 2.1(2.0_0)$ ATLAS-CONF-2018-026 [62]	upper limit, $\mu_{95\%} \leq 2.95(2.45_0)$ arXiv:1807.06325 [63]
$H \to$ invisible	upper limit, $\mathcal{BR}_{95\%} \leq 0.26(0.17)$ ATLAS-CONF-2018-054 [64]	upper limit, $\mathcal{BR}_{95\%} \leq 0.24(0.23)$ arXiv:1610.09218 [65]
$H \to J/\Psi\gamma, \Upsilon\gamma$	upper limits arXiv:1807.00802 [66]	upper limits arXiv:1507.03031 [67]
$H \to \phi\gamma, \rho\gamma$	upper limits arXiv:1712.02758 [68]	

Table 2.4.: Status of experimental analyses of Higgs decay modes. Expected limits and significances are given in parentheses. For the expected limits an index indicates the hypothesised SM Higgs signal strength, i.e. a subscript of 0 indicates that the corresponding expected limit is based on the assumed absence of the Higgs boson while a subscript of 1 denotes that presence of the SM Higgs boson was assumed.

Chapter 3

Monte Carlo Event Generation

In its broadest sense, the term *Monte Carlo method* refers to a numerical calculation technique which makes use of random or pseudo-random numbers.[1] One application of the Monte Carlo (MC) method, which is of particular interest in high energy particle physics, is the calculation of integrals, which e.g. arise in the evaluation of total or differential cross sections. A *Monte Carlo event generator* is a software program based on the Monte Carlo method which allows for the simulation of particle collisions and provides fully detailed final states at the level of observable particles. In doing so, it essentially calculates a multi-dimensional cross section integral.

There are both physical and mathematical reasons for the use of Monte Carlo event generators in HEP. From a physics viewpoint, the fully detailed final states produced by a MC event generator allow to subsequently run a detector simulation which contains a full model of the experimental apparatus and enables the calculation of the detector response to the incident and traversing particles, down to the level of electrical signals in individual detector components. This procedure allows experimental data and simulated events to be treated on an equal footing and used interchangeably. This is very important e.g. when setting up an experimental analysis, as it can be designed and optimised using Monte Carlo events and later be applied to experimental data. Furthermore, Monte Carlo events are also used to study analysis properties such as the detector acceptance and event selection efficiencies and comparing the experimental results to theoretical predictions. The mathematical motivation for using the Monte Carlo method is related to the efficiency of such calculations. The number of integrals in the calculation of cross sections grows linearly with the number of involved final state particles. This behaviour renders other numerical integration algorithms impractical as their complexity grows too quickly with the number of integrations. Therefore, Monte Carlo techniques often are the only practical tools in the context of such calculations.

[1] Pseudo-random numbers (PRNs) are generated by means of algorithms and are deterministic, rather than truly random. The sequence of pseudo-random numbers generated successively by an algorithm depends on a starting condition, the so-called *seed*. For a given algorithm and seed, the sequence of pseudo-random numbers is completely determined, which allows for reproducible calculations based on a PRN generator.

In order to arrive at a description at the level of observable particles, the event generation is split up into different stages that are sketched in Figure 3.1. The following steps are considered in Monte Carlo event generators when making predictions of physics processes at the LHC:

- The hard scattering process is calculated using fixed-order perturbation theory based on Feynman diagrams, corresponding to a certain process of interest, such as Higgs or top quark pair production. It is represented by the red vertex in the sketch.
- The parton shower mostly generates QCD radiation, in which a number of partons are emitted as the event is evolved from the high energy scale of the hard interaction down to a much lower *parton shower cut-off scale* of the order of 1 or 2 GeV. This stage is shown in green and blue. Parton shower emissions off outgoing particles is referred to as *final state radiation* (FSR), while emissions off incoming particles is termed *initial state radiation* (ISR).
- The remainders of the protons after the extraction of the partons participating in the ISR are called *proton* or *beam remnants*. The remnants are depicted as brown ellipses in the sketch. Further partons are extracted from the remnants and are subjected to ISR, a scattering, typically based on LO QCD $2 \to 2$ processes, and FSR. This is referred to as *underlying event* (UE) or *multiple parton interactions* (MPI).
- As perturbation theory ceases to be meaningful around the parton shower cut-off scale, hadronisation models are used to form hadrons out of the available partons. High-mass hadrons undergo decays to lighter and longer-lived particles. This stage is sketched with the black ellipses and yellow decay cascades.

In Monte Carlo event generators, particles with mean lifetimes of more than 30 ps are typically considered to be stable. The behaviour of the longer lived particles, their energy loss and further reactions taking place when entering and traversing the detector is modelled by a dedicated subsequent detector simulation, based on `Geant4` [69], which will not be described here. The fully detailed simulation of the ATLAS detector is referred to as *FullSim*. In order to reduce the computing demands, an alternative, less detailed and therefore faster detector simulation, referred to as *FastSim*, was created, in which low-energy electromagnetic particles are removed from the calorimeter and replaced with pre-simulated cascades [70]. The FastSim was designed in such a way, that the standard reconstruction software can still be run and therefore, both types of simulations can be used side by side in an analysis. A detailed description of the ATLAS detector simulation infrastructure is available here [71]. Descriptions of the ATLAS detector and the reconstruction of physics objects will be described in Chapters 4 and 5.

In this chapter, some basic properties of the Monte Carlo method will be reviewed, followed by a discussion of the different parts of a modern HEP Monte Carlo event generator. Parts of this chapter follow the discussion in [73]. Further details can be found in [74, 75].

3.1. Monte Carlo Techniques

The idea of using the Monte Carlo technique for solving integrals is based on the Riemann formulation of integration

$$\int_a^b f(x)dx = \lim_{N\to\infty} \frac{b-a}{N} \sum_{i=1}^{N} f(x_i) \ , \tag{3.1}$$

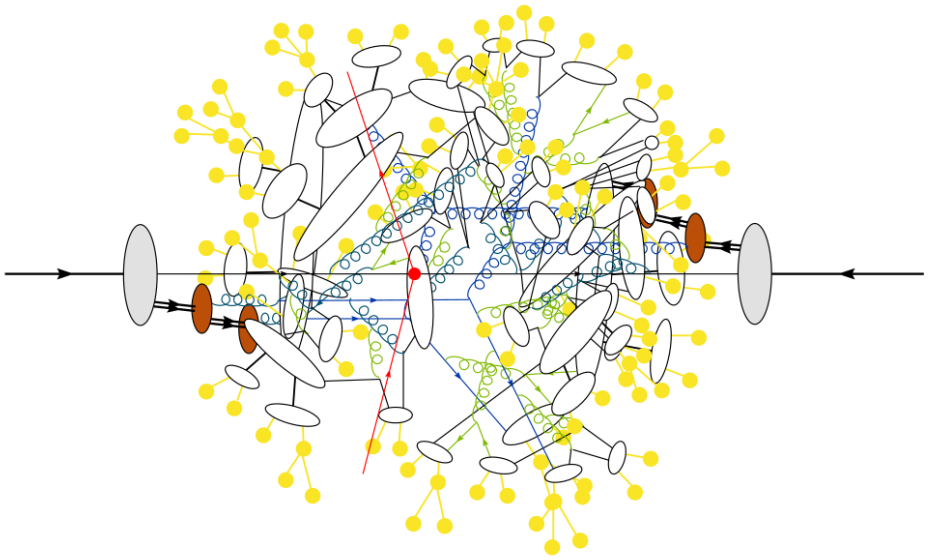


Figure 3.1.: Sketch of the different stages of a modern Monte Carlo event generator used for the simulation of scattering processes at the LHC. The different parts of the simulation are described in the main text. The overlap of the different elements of this sketch is due to Taken from [72].

where the $\{x_i\}$ may e.g. be located at the left or right edges of the N intervals into which the integration range $[a, b]$ is divided. In Monte Carlo integration, the sequence $\{x_i\}_{i=1}^N$ is replaced by a sequence of N (pseudo-)random points $\{r_i\}_{i=1}^N$. It can be shown that the statistical uncertainty δ of the Monte Carlo estimate of an integral is given by

$$\delta = \frac{V}{\sqrt{N(N-1)}} \sqrt{\sum_{i=1}^{N} f_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} f_i\right)^2}\,, \tag{3.2}$$

where $V = b - a$ is the size of the integration range [73]. The exact same result also holds for the case of multiple integrations on a domain Ω, where V denotes the volume of Ω and the integrand is evaluated at the points $\{\boldsymbol{r}_i\}_{i=1}^N$. The points $\{r_i\}$ and $\{\boldsymbol{r}_i\}$ are often referred to as *phase space points*. A derivation of these results can be found in [73]. Note that the expression for the statistical uncertainty δ does not depend on the dimensionality of the integration. Instead, the uncertainty decreases as $1/\sqrt{N}$ independently of the number of integrals. This is precisely the property which makes the Monte Carlo technique the appropriate tool for complex calculations involving possibly many integrals. The same expression also provides further hints on how to reduce the uncertainty, besides only increasing the number of sampling points N. The square root term in Equation (3.2) describes the variance of the integrand f. Thus, modifying the integration in a way that reduces the variance of the integrand will result in a reduced uncertainty. A few techniques for doing this will be introduced in the following.

3.1.1. Importance Sampling

In importance sampling, the integrand $f(x)$ is approximated by a function $g(x)$ chosen in such a way that its integral is known analytically and invertible. Rewriting the integration

and substituting $\mathrm{d}y := g(x)\mathrm{d}x$ and $y = G(x)$ yields

$$\int_{x_1}^{x_2} f(x)\mathrm{d}x = \int_{x_1}^{x_2} \frac{f(x)}{g(x)} g(x)\mathrm{d}x = \int_{y_1=G(x_1)}^{y_2=G(x_2)} \frac{f(G^{-1}(y))}{g(G^{-1}(y))}\mathrm{d}y \ , \tag{3.3}$$

where the integrand f/g will have a smaller variance than f, if g is chosen properly. More intuitively, the purpose of importance sampling is to sample more densely in those regions where the original integrand f takes on its largest values.

3.1.2. Multi-Channelling

Often times, it will not be possible to find a single function g that approximates well the behaviour of f and is simple enough to have an analytically known and invertible integral G. This may e.g. be the case when the integrand has multiple peaks in different regions of the integration domain. The multi-channel method consists in using a combination of m different variable transformations $y_i = g_i(x)$ that collectively are able to approximate the behaviour of the integrand. In principle, this is similar to the method of importance sampling. The different transformations g_i are called channels and are combined by means of the weights α_i as

$$g(x) = \sum_{i=1}^{m} \alpha_i g_i(x) \quad \text{with} \quad 1 = \sum_{i=1}^{m} \alpha_i \ . \tag{3.4}$$

For a given sampling point, one then selects the channel i with probability α_i, uses the mapping $y = g_i(x)$ to transform the integration variable and then evaluates the modified integrand $f(y)/g(y)$. Note that not the estimate of the integral value itself, but only the estimate of the uncertainty depends on the values of the $\{\alpha_i\}$. Therefore, they can be modified during the calculation so as to reduce the variance. [76]

3.1.3. Stratified Sampling

The idea behind stratified sampling is to divide the integration region into multiple smaller regions and calculate the integral separately in each of them. In one dimension and for a division into n smaller regions this may be written as

$$\int_{x_0}^{x_n} f(x)\mathrm{d}x = \int_{x_0}^{x_1} f(x)\mathrm{d}x + \int_{x_1}^{x_2} f(x)\mathrm{d}x + \ldots + \int_{x_{n-1}}^{x_n} f(x)\mathrm{d}x \ , \tag{3.5}$$

where $\{x_i\}_{i=1}^{n-1}$ denote the boundaries of the smaller integration regions with $x_0 < x_1 < \ldots < x_n$. According to [76] the total uncertainty δ is then given by

$$\delta^2 = \sum_{i=1}^{n} \delta_i^2 \ , \tag{3.6}$$

where δ_i denotes the uncertainty of the calculation in each of the smaller regions.

3.1.4. Unweighting

So far, only the integrand was evaluated and its values, the so-called *weights*, for different phase space points were summed up to obtain the value of the integral. As mentioned before, the integral may correspond to a total or a differential cross section. In principle, the same procedure could be used repeatedly to separately and independently calculate

various differential cross sections. However, it is usually much more convenient to store the *events*, i.e. the combinations of the phase space points and the corresponding weights. This has the great advantage that the integration only has to be performed once and the stored events can be used to obtain the distributions of all conceivable variables later on. As long as they are non-negative, the event weights have an intuitive probabilistic interpretation[2]: Events that are more likely to happen have a larger weight than rare events. In particular, a sample of *weighted events* may consist of many events with low weights and few events with large weights. One may, however, wish to generate a sample of *unweighted events*, i.e. where all events have equal weight and occur just as frequently as appropriate for their corresponding phase space points. The procedure to obtain unweighted events from a set of weighted events is called *accept-or-reject*, *hit-or-miss* or *von Neumann algorithm*. It consists in choosing a reference weight w_{ref} that is at least as large as any of the individual event weights and then accepting an event with weight w_i with the probability w_i/w_{ref} and rejecting it with the probability $1 - w_i/w_{\text{ref}}$. The closer w_{ref} is to the maximum weight and the smaller the variance of the $\{w_i\}$, the more efficient the unweighting is. The unweighting can either be performed after the event generation, in which case the largest event weight is known, or during the event generation. The latter approach may be preferable in order to ensure that a given number of unweighted events is indeed generated. In this case, however, the largest event weight is not known and the reference weight has to be estimated. Therefore, it may happen that during the event generation, the reference weight $w_{\text{ref}}^{(\text{old})}$ is exceeded by an event weight w_j. This can be accounted for by increasing the reference weight, $w_{\text{ref}}^{(\text{new})} = w_j$, and subjecting the previous events to the hit-or-miss algorithm again, now using the probability $w_{\text{ref}}^{(\text{old})}/w_{\text{ref}}^{(\text{new})}$ for accepting an event.

3.2. Fixed-Order Calculations

The all-orders perturbative expansion of a total cross section σ has the structure

$$\sigma = \sum_{k=0}^{\infty} \int_{m+k} \mathrm{d}\Phi_{m+k} \left| \sum_{l=0}^{\infty} \mathcal{M}_{m+k}^{(l)}(\Phi_{m+k}) \right|^2 . \tag{3.7}$$

In this notation, the sums and integrals over the different contributing PDFs, that appear if hadronic beams are present and that were written explicitly in Equation (2.7), were omitted for the sake of clarity. m describes the minimal number of particles in the final state for the given process, k the number of additional particle emissions, l the number of loops in the Feynman diagrams, Φ_{m+k} the $m+k$-particle phase space, and $\mathcal{M}_{m+k}^{(l)}$ the scattering amplitude obtained from summing all contributing Feynman diagrams with $m+k$ final state particles and l loops. The real-valued product of two scattering amplitudes is called *matrix element* (ME). Differential cross sections are calculated in a very similar manner, the only difference being that the integrations over the differential variables are not carried out. For an exemplary observable $\mathcal{O}$ this reads:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{O}} = \sum_{k=0}^{\infty} \int_{m+k} \mathrm{d}\Phi_{m+k} \left| \sum_{l=0}^{\infty} \mathcal{M}_{m+k}^{(l)}(\Phi_{m+k}) \right|^2 \delta(\mathcal{O} - \mathcal{O}(\Phi_{m+k})) . \tag{3.8}$$

[2] This is only strictly true at the LO. Due to the contributions with different final state multiplicities at NLO as well as the modifications needed for treating the occurring divergences, events at NLO and beyond do not have a straightforward interpretation.

In practice, the perturbation series is truncated at a certain order, as mentioned in Section 2.1. The leading order approximation of the cross section typically reads

$$\sigma_{\mathrm{LO}} = \int_m \mathrm{d}\Phi_m \mathcal{B}(\Phi_m) \quad \text{with} \quad \mathcal{B}(\Phi_m) = \left|\mathcal{M}_m^{(0)}(\Phi_m)\right|^2 \tag{3.9}$$

and only contains the *Born* contribution $\mathcal{B}$.[3] The next-to-leading order approximation includes one more power in the coupling constant and contains the terms

$$\sigma_{\mathrm{NLO}} = \int_m \mathrm{d}\Phi_m \left[\mathcal{B}(\Phi_m) + \mathcal{V}(\Phi_m)\right] + \int_{m+1} \mathrm{d}\Phi_{m+1} \mathcal{R}(\Phi_{m+1}) \tag{3.10}$$

with

$$\mathcal{V}(\Phi_m) = 2\mathrm{Re}\left(\mathcal{M}_m^{(0)\dagger}\mathcal{M}_m^{(1)}\right) \quad \text{and} \quad \mathcal{R}(\Phi_{m+1}) = \left|\mathcal{M}_{m+1}^{(0)}\right|^2 . \tag{3.11}$$

$\mathcal{V}$ and $\mathcal{R}$ are called the virtual and real emission corrections, respectively. Note that amplitudes with an additional loop contain two additional vertices which is why their product with a Born-level amplitude is of the same order as the square of an amplitude with one additional emission.

It is important to keep in mind the different types of singularities that either possibly or definitely arise at the different orders of this perturbative expansion. In principle, the Born cross section is finite. Depending on the process under consideration it may be divergent, however, unless kinematic cuts are applied to remove those regions of phase space that give rise to the singularities. Such configurations include the presence of soft, i.e. low energetic, or collinear massless partons in the Born-level final state. Applying *jet cuts*, that enforce a minimum transverse momentum and a minimum separation of the corresponding partons, avoids such configurations and provides a well-defined process with a finite cross section.

The next class of divergences appear in the virtual corrections at NLO. Each loop comes with an integration over the four-momentum carried around in the loop and integrals over unbounded loop momenta may lead to divergences. For this reason, they are commonly referred to as *ultraviolet* (UV) divergences. These divergences are treated by means of renormalisation, which was introduced in Section 2.1.

The third and last type of singularities are the so-called *infrared* (IR) divergences that appear between the virtual and real emission contributions to the NLO cross section. In the real emission corrections, they occur if one of the associated final state particles is soft or both associated outgoing particles are collinear. This is, in principle, exactly the same behaviour mentioned previously that makes jet cuts necessary for processes containing massless partons at the Born level. It can be shown, however, that the IR divergences from the virtual and real emission corrections cancel each other at each order in perturbation theory. This is known as the KLN theorem [77, 78]. For numerical perturbative calculations, these divergences are problematic, however, since the virtual and real emission corrections are connected to phase spaces with different particle multiplicities. Therefore, performing these calculations in a naive manner would result in subtracting two possibly large opposite-sign numbers which is numerically unstable. Different methods, going by the names of *phase space slicing* and *subtraction* methods, were developed to overcome this challenge. An overview of these methods is given in [74, 75] and the subtraction method by Catani and Seymour is summarised in [73].

[3] Certain processes, such as e.g. gluon-gluon fusion Higgs production, contain loops already at the leading order. For the sake of simplicity, tree-level Born processes, i.e. Born processes without loops, are assumed in the discussion presented here.

3.3. Fundamentals of Parton Showers

As previously mentioned, the strong coupling increases towards lower energy scales. This particular feature gives rise to a striking behaviour, which is the appearance of jets from QCD interactions. When comparing the scattering amplitudes for QCD reactions such as $q \to qg$ or $g \to q\bar{q}$, it can be seen that they are very similar in structure to their QED counterparts $e \to e\gamma$ and $\gamma \to e^+e^-$. But it is the increase of the strong coupling towards lower energy scales that leads to an increased number of emissions in QCD final states compared to QED final states and this is precisely the origin of the jets.

This means that fixed-order calculations for the hard process are not sufficient to describe observables that are sensitive to the presence and structure of jets. But while it is in principle possible to generate and evaluate matrix elements for $2 \to n$ processes with $n = \mathcal{O}(10)$, this would be very time consuming and computationally involved and therefore impossible in practice. For this reason, the parton shower approach was developed, which provides a link and transition between the fixed-order perturbative calculation at the high energy scales of the hard scattering process and the heuristic hadronisation models which are applied at energy scales of only a few GeV. Parton showers are recursive algorithms that apply the same generic matrix elements for $1 \to 2$ splittings over and over again to subsequently build up a multi-parton final state. They start off from the highly energetic configuration generated in the hard scattering process and evolve it down to ever lower scales by performing increasingly soft and collinear emission reactions. This evolution is tracked with the so-called *evolution* or *ordering variable* q. Different choices are possible for this variable, such as e.g. the transverse momentum, the opening angle between two partons or the virtuality [72]. The recursion stops once a configurable lower cut-off scale μ_{IR} is reached, which is typically of the order of 1-2 GeV. At that stage, the event is passed to the hadronisation model which does not rely on perturbative physics. Let it be emphasised at this point that parton showers are perturbative algorithms. They are, however, not a fixed-order approach as their stopping criterion is given by a scale of the evolution variable rather than a specific number of emissions corresponding to a particular order in perturbation theory.

Technically, parton showers are based on the observation that cross sections of different particle multiplicities factorise in the soft and collinear limits. This can be written as

$$\mathrm{d}\sigma_{n+1} \approx \mathrm{d}\sigma_n \times \sum_{\substack{\text{emitters and} \\ \text{splittings}}} \frac{\alpha_s}{2\pi} \cdot \text{phase space} \cdot \text{splitting function} \,. \tag{3.12}$$

The splitting functions can be obtained from calculating the $2 \to 3$ processes that contain the splitting in question and extracting the expressions describing the splitting from the factorised cross sections in the collinear or soft limit. Due to this factorisation, the same splitting functions can then be used to calculate parton emissions off a different hard process with a different final state. Different formulations of the QCD splitting functions can be found in e.g. [74, 75, 79].

While splitting functions describe parton emissions, the probability for having no splitting and emission when evolving the event from an upper scale q_1 to a lower scale q_2 is given by the *Sudakov form factor*

$$\Delta(q_1^2, q_2^2, \Phi_n) = \exp\left[-\frac{\alpha_s}{2\pi}\int_{q_2^2}^{q_1^2}\frac{\mathrm{d}q^2}{q^2}\sum_{i,j}\int P_{ij}(z, q^2)\mathrm{d}z\right] \,. \tag{3.13}$$

q is the evolution variable, the sum over i and j encodes the different emitting particles and splittings that may occur, P_{ij} are the corresponding splitting functions and z is the variable that together with q defines the splitting phase space.[4] With this, all ingredients are at hand to write down the parton shower operator formulation

$$\begin{aligned} \mathrm{PS}[Q^2, \mu_{\mathrm{IR}}^2, \Phi_n, \mathcal{O}] &= \Delta(Q^2, \mu_{\mathrm{IR}}^2, \Phi_n)\ \mathcal{O}_n(\Phi_n) \\ &+ \int_{\mu_{\mathrm{IR}}^2}^{Q^2} \mathrm{d}\mathcal{P}(q^2)\ \Delta(Q^2, q^2, \Phi_n)\ \mathrm{PS}[q^2, \mu_{\mathrm{IR}}^2, \Phi_{n+1}, \mathcal{O}] \end{aligned} \tag{3.14}$$

with

$$\mathrm{d}\mathcal{P}(q^2) = \frac{\alpha_s}{2\pi}\frac{\mathrm{d}q^2}{q^2}\sum_{i,j}\int P_{ij}(z, q^2)\mathrm{d}z\ . \tag{3.15}$$

Q denotes the scale corresponding to the configuration obtained from the hard scattering and $\mathcal{O}$ is the observable that is to be calculated. This operator formulation has a straightforward interpretation. At each level of the recursion either of two things may happen: Either there is no splitting between the upper scale and the cut-off scale μ_{IR} as encoded by the Sudakov form factor in the first line of Equation (3.14) or there is no splitting down to an intermediate scale q where an emission occurs. Since the intermediate scale may be anywhere between the upper and the cut-off scale, it is integrated over. The parton shower operator is then recursively applied on the resulting new $n+1$ particle final state.

It is also worth pointing out that it is not possible to enforce exact four-momentum conservation in a $1 \to 2$ splitting while at the same time keeping all three involved particles on their respective mass shells. Therefore, one of those conditions has to be relaxed. Traditional parton showers that work in the $1 \to 2$ scheme typically allow particles to depart from their mass shells and then perform a global momentum reshuffling once the cut-off scale is reached in order to restore the nominal masses of the final state particles. A different approach is the one of dipole showers which work in the $2 \to 3$ splitting picture where a momentum recoil is absorbed by a spectator particle so that all nominal masses are retained [74].

3.4. Matching and Merging

As explained in the previous section, the parton shower generates additional emissions on top of the configuration obtained from the hard scattering. It is an approximation based on the factorisation of the cross sections in the soft and collinear limits. Due to this approximation, parton showers do not have the same formal accuracy as full matrix element calculations for the same topology. Therefore, while parton showers work well in the realm of their approximation, they fail to describe hard and well-separated emissions, which are better described by the corresponding matrix elements. This leads to the idea of combining matrix elements of subsequently higher orders and multiplicities with parton showers so as to profit from the strengths of both methods. Care has to be taken, however, when a hard process cross section calculated beyond the LO is to be combined with a parton shower. If done naively, parton shower emissions on top of Born configurations will mimic and double-count real emission NLO corrections and parton shower non-emissions will double-count virtual corrections. The procedures developed to combine NLO matrix elements

[4] The integration over the third variable, the azimuthal angle ϕ, is not explicitly written as the terms do not depend on it.

with parton shower calculations in a consistent manner go by the names of *matching* and *merging*. While there are some historical exceptions to the nomenclature, matching usually refers to the combination of an NLO cross section for a single particle multiplicity with a parton shower, while merging is used to refer to the combination of NLO and LO matrix elements for various particle multiplicities with a parton shower. An example for the latter would be the combination of matrix elements for $pp \to Z + 0/1/2j$ at NLO in one hard process calculation to which a parton shower is applied.

Expanding the parton shower operator and truncating after the first additional order in the strong coupling yields

$$\mathrm{PS}[Q^2, \mu_{\mathrm{IR}}^2, \Phi_n, \mathcal{O}] \approx \left[1 - \int_{\mu_{\mathrm{IR}}^2}^{Q^2} \mathrm{d}\mathcal{P}(q^2)\right] \mathcal{O}_n(\Phi_n) + \int_{\mu_{\mathrm{IR}}^2}^{Q^2} \mathrm{d}\mathcal{P}(q^2) \cdot 1 \cdot \mathcal{O}_{n+1}(\Phi_{n+1}) \ . \quad (3.16)$$

Naively combining this with the NLO cross section and truncating all terms beyond NLO yields

$$\begin{aligned} \mathrm{PS}[Q^2, \mu_{\mathrm{IR}}^2, \Phi_n, \mathcal{O}]\ \mathcal{O}_{\mathrm{NLO}} &= \int \mathrm{d}\Phi_n \left(\mathcal{B} + \mathcal{V}\right) \mathcal{O}_n + \int \mathrm{d}\Phi_{n+1} \mathcal{R} \mathcal{O}_{n+1} \quad (3.17) \\ &- \int \mathrm{d}\Phi_n \mathcal{B} \int \mathrm{d}\mathcal{P}(q^2) \mathcal{O}_n + \int \mathrm{d}\Phi_n \mathcal{B} \int \mathrm{d}\mathcal{P}(q^2) \mathcal{O}_{n+1} \ . \end{aligned}$$

Note that throughout this chapter terms related to the subtraction of the fixed-order IR divergences are not written for the sake of clarity. In this formulation, the double counting of terms at NLO is manifest: The first term in the second line of Equation (3.17) mimics the virtual correction and the second term in the second line approximates the real emission corrections.

There are two main algorithms aimed at combining NLO matrix elements of one particle multiplicity with parton showers. They go by the names of MC@NLO [80–82] and Powheg [83–85].

The MC@NLO method essentially works by modifying the hard process cross section and subtracting from it the doubly counted parton shower terms:

$$\mathrm{d}\sigma_{\mathrm{MC@NLO}}^{\text{hard process}} = \mathrm{d}\Phi_n \left[\left(\mathcal{B} + \mathcal{V}\right) + \mathcal{B}\, \mathrm{d}\mathcal{P}(q^2)\right] \mathcal{O}_n + \left(\mathrm{d}\Phi_{n+1}\, \mathcal{R} - \mathrm{d}\Phi_n\, \mathcal{B}\, \mathrm{d}\mathcal{P}(q^2)\right) \mathcal{O}_{n+1} \ . \quad (3.18)$$

Therefore, in the MC@NLO method, the hard process event generation needs to be specific for the exact parton shower which is supposed to be applied afterwards and hard process events generated e.g. for the `Pythia8` [86] parton shower cannot be reused and showered with `Herwig7` [87].

The Powheg-matched cross section is given by the expression [74]

$$\begin{aligned} \mathrm{d}\sigma_{\mathrm{Powheg}} &= \mathrm{d}\Phi_n \bar{\mathcal{B}}(\phi_n) \quad (3.19) \\ &\times \left[\bar{\Delta}(Q^2, \mu_{\mathrm{IR}}^2) \mathcal{O}_n + \int \mathrm{d}\Phi_{\mathrm{rad}}(q^2)\ \theta(q^2 > \mu_{\mathrm{IR}}^2)\ \frac{\mathcal{R}(\Phi_{n+1})}{\mathcal{B}(\Phi_n)} \bar{\Delta}(Q^2, q^2) \mathcal{O}_{n+1}\right] \end{aligned}$$

with

$$\bar{\mathcal{B}}(\Phi_0) = \mathcal{B}(\Phi_n) + \mathcal{V}(\Phi_n) + \int \mathrm{d}\Phi_{\mathrm{rad}} \mathcal{R}(\Phi_{n+1}) \quad (3.20)$$

and

$$\bar{\Delta}(Q_1^2, Q_2^2) = \exp\left[-\int \mathrm{d}\Phi_{\mathrm{rad}}(q^2)\ \theta(q^2 > Q_2^2)\ \frac{\mathcal{R}(\Phi_{n+1})}{\mathcal{B}(\Phi_n)}\right] \ . \quad (3.21)$$

Φ_{rad} denotes the phase space associated to an emission. Technically, for each event an n-particle Born-type configuration is generated with the kinematic configuration Φ_n and is assigned the (NLO-) weight $\bar{\mathcal{B}}(\Phi_n)$, i.e. the full integration over the real emission phase space Φ_{rad} is performed for every single phase space point [85].

More double counting occurs when matrix elements of different particle multiplicities are combined, since a tree-level matrix element of a given parton multiplicity contributes both as the Born contribution for the process with the same multiplicity but also as the real emission correction to the next lower multiplicity. Different methods and algorithms were developed for dealing with this. The FxFx merging [88] combines MC@NLO-matched NLO matrix elements of different parton multiplicities. A merging scale μ_Q is used to classify different event configurations into different jet multiplicities. The main idea behind the merging can be sketched as

$$\mathrm{d}\bar{\sigma}_i = \mathrm{d}\sigma_i \, D(d_{i+1}) \, (1 - D(d_i)) \, \theta(d_{i-1} - \mu_Q) \tag{3.22}$$

where $\mathrm{d}\sigma_i$ is the MC@NLO-matched i-parton cross section, $D(\mu) = \theta(\mu_Q - \mu)$ and d_i marks the scale at which a given event transitions from being interpreted as an i-jet to an $i-1$-jet configuration. This implies that i jets must be harder than the merging scale μ_Q and one jet should be softer than μ_Q, thus ensuring that any jets harder than the merging scale μ_Q are described by NLO matrix elements. This condition, encoded by the factor $D(d_{i+1})$ has to be relaxed for the highest multiplicity matrix element. The reason is that if NLO matrix elements for up to n partons are considered, then the real emission correction for the n-th jet contains the Born, i.e. LO description of the $n+1$-st parton and there is no higher-order description available. Thus, the $n+1$-st jet can only be described by a LO matrix element. In practice one may wish to choose a different functional form for $D(\mu)$. According to [88] any smooth function

$$D(\mu) = \begin{cases} 1 & \mu \leq \mu_1 \\ \text{monotonic} & \mu_1 < \mu \leq \mu_2 \\ 0 & \mu > \mu_2 \end{cases} \tag{3.23}$$

with $\mu_1 \leq \mu_2$ can be considered.

3.5. The Underlying Event

After the partons leading to the hard interaction have been extracted from the incoming protons, the so-called *beam remnants* are left and partons extracted from them may undergo additional interactions. These are collectively called the *Underlying Event* (UE) or *Multiple Parton Interactions* (MPI). Technically, the UE / MPI simulation is performed before hadronisation, but after the initial-state and final-state parton showers were applied to the hard process. Partons participating in the UE undergo ISR, typically a LO QCD $2 \to 2$ scattering process and FSR. Heuristic models are added in to provide realistic descriptions of relevant observables such as the distribution of the number of charged particles. More information on the different components and aspects of UE simulations can be found in [74].

3.6. Hadronisation

Hadronisation refers to the transition from partonic final states after the parton shower to hadronic final states containing baryons and mesons. This part of the event simulation is performed after the parton shower evolution variable reaches the lower cut-off scale μ_{IR}. At this scale, which as mentioned before is typically of the order of 1-2 GeV, perturbation theory ceases to be valid because the strong coupling constant becomes too large. Therefore, heuristic models with free parameters are used to describe the transition from partons to hadrons. The procedure of adjusting the model parameters in order to arrive at predictions consistent with experimental results is called *tuning*. A consistent set of event generator parameters obtained from this procedure is often referred to as a *tune*. There are two main hadronisation models in use in current Monte Carlo event generators: The *Lund string* model and the *cluster hadronisation* model [74].

The Lund string model [86, 89–91] directly creates hadrons from combinations of partons. It is based on the linear part of the QCD potential, $V(r) = \kappa r$, where r is the distance between two colour charges and $\kappa \approx 1$ GeV/fm the string constant. In this model, the field between colour charges is often illustrated as a string or flux tube spanning between quarks and antiquarks. Gluons are represented as kinks in this string. When two quarks q and $\bar{q}$ are separated, the potential energy increases until a new quark-antiquark pair $q'\bar{q}'$ is formed between them, which can be combined with the original quarks to form colour-neutral singlets $q\bar{q}'$ and $\bar{q}q'$. This is illustrated as the breaking of a string. Since these singlets carry no net colour charge, there is no colour field between them and they can separate. As long as the invariant mass of the partons forming a string is large enough, the string is allowed to expand and break. Otherwise, the string is turned into a meson, which may subsequently decay. Baryons may be produced via string breaks to diquark-antidiquark pairs or from combining successive string breaks to quark-antiquark pairs. A downside of the string model is the fact that it requires many adjustable parameters, which on the other hand can be used to accurately tune the model to experimental data [74]. The Lund string model is implemented in the `Pythia` [86, 91] event generator family.

The cluster hadronisation model [87, 92–97] is based on the observation that after parton showering, colour singlet quark-antiquark combinations can be formed in such a way that their invariant mass distribution asymptotically tends towards a universal distribution, that is independent of the hard process and the details at large values of the parton shower evolution variable [74,98]. This behaviour is referred to as the *preconfinement property of parton showers*. At the end of the parton shower evolution, cluster hadronisation begins by forcing gluons to split into quark-antiquark pairs. From all quarks and antiquarks, colour singlet pairs are formed. These clusters mostly undergo two-body decays to meson or baryon pairs. Another way of producing baryons is via gluon splittings to light diquark-antidiquark pairs. In general, kinematic effects from the quark and cluster masses are sufficient to adequately dampen and suppress the production of heavy flavours and strangeness, which can nevertheless be produced, albeit at a reduced rate, in cluster decays. One advantage of the cluster hadronisation model is the lower number of tuneable parameters, which however in some cases may result in tensions between the model predictions and experimental data, which in turn may trigger extensions and improvements of the hadronisation model [74]. Cluster hadronisation models are implemented in the `Herwig7` [87, 95, 96] and `Sherpa` [97] event generators.

Chapter 4

The LHC and the ATLAS Experiment

4.1. The Large Hadron Collider

The LHC [99–101] is the world's largest and most energetic particle accelerator. It is located at CERN, in the same circular tunnel with 26.7 km circumference that had previously housed the Large Electron-Positron Collider. After passing a series of pre-accelerators, the proton beams are injected into the LHC with an energy of 450 GeV, accelerated in opposite directions and brought to collision at four interaction points where the ALICE, ATLAS, CMS and LHCb experiments are located. While the design centre-of-mass energy of the LHC is 14 TeV, the highest energy reached so far is 13 TeV, which has been used for the Run 2 collisions starting from 2015. At this energy, the protons complete more than 11000 revolutions per second [102]. The proton beams are held on their trajectory and focused by more than 1200 dipole and almost 400 quadrupole magnets. In order to reach the aforementioned centre-of-mass energies with the given tunnel, the main dipoles need to create magnetic fields of more than 8 T, which is why superconducting magnets are used. The proton beams consist of different bunches with a separation of 25 ns, each containing approximately 100 billion particles. In the Run 2 phase, fills with up to about 2800 bunches are being used, which is sufficient to exceed the LHC design instantaneous luminosity of 10^{34} cm^{-2}s^{-1} in the ATLAS detector. The *instantaneous luminosity* $\mathscr{L}$ is a measure for the beam intensity. It is related to the cross section σ and the event rate $\dot{N}$ via $\dot{N} = \sigma \cdot \mathscr{L}$. In terms of the collider parameters it is given by [103]

$$\mathscr{L} = \frac{N_p^2 N_b f_{\text{rev}}}{2\pi\sigma_x\sigma_y} \,, \tag{4.1}$$

where N_p is the number of protons per bunch, N_b the number of colliding bunches, f_{rev} the revolution frequency of the accelerator and σ_x and σ_y the horizontal and vertical widths of the beam spot profiles. For Gaussian beam profiles, σ_x and σ_y represent the standard deviations of the corresponding distributions. The *integrated luminosity* is given by $\mathcal{L} = \int \mathscr{L} \mathrm{d}t$. Apart from protons, also lead and xenon atoms can be brought to collision in the LHC.

LHC Run	Year	CM energy $\sqrt{s}$	Bunch spacing	LHC delivered luminosity	ATLAS recorded luminosity	Average pile-up $\langle\mu\rangle$
1	2010	7 TeV	150 ns	47 pb^{-1}	45 pb^{-1}	≈ 2
	2011	7 TeV	50 ns	5.5 fb^{-1}	5.23 fb^{-1}	≈ 9
	2012	8 TeV	50 ns	22.7 fb^{-1}	21.3 fb^{-1}	20.7
2	2015	13 TeV	50 ns	4.2 fb^{-1}	173 pb^{-1}	19.6
	2015	13 TeV	25 ns		3.7 fb^{-1}	13.5
	2016	13 TeV	25 ns	38.5 fb^{-1}	35.6 fb^{-1}	25.1
	2017	13 TeV	25 ns	50.4 fb^{-1}	47.1 fb^{-1}	37.8
	2018	13 TeV	25 ns	65.0 fb^{-1}	62.2 fb^{-1}	37.0

Table 4.1.: Overview of LHC proton-proton physics runs [103–106].

4.2. The ATLAS Experiment

The ATLAS experiment [107] is a cylinder-symmetric, general-purpose detector covering almost the entire solid angle around the interaction point, where both particle beams are crossed and brought to collision. It has a diameter of about 25 m, a length of about 44 m and weighs approximately 7000 tons. It consists of several subdetectors, sketched in Figure 4.1, that are arranged in barrel and disk or endcap layers and serve to detect and measure different particle types and properties. Going outwards from the interaction point, the three main detector subsystems are:

- the inner detector (ID), where the tracks of charged particles are measured,
- the calorimeters, where electrons, photons and hadrons are contained, and
- the muon spectrometer.

The hermeticity of the detector is not only important in order to measure particles in as large a region of phase space as possible, but also in order to detect neutrinos. These cannot be measured directly by any of the detector subsystems and therefore have to be reconstructed from transverse momentum imbalances. This procedure can only work if as many of the detectable particles as possible are indeed measured.

Due to the structure of the bunches, multiple pairs of protons may collide during a single bunch crossing. While usually the most energetic reaction is of interest in physics analyses, the other interactions in a given bunch crossing are referred to as *in-time pile-up*. Reactions from previous bunch crossings that contribute to a given event are termed *out-of-time pile-up*. The latter effect is due to the fact that many subdetectors are sensitive to time spans larger than the 25 ns between two subsequent bunch crossings [108]. The average number of interactions per bunch crossing $\langle\mu\rangle$ is given in Table 4.1 along with the integrated luminosities for different data taking periods.

ATLAS makes use of a right-handed coordinate system centred at the nominal interaction point in the middle of the detector. The positive x-axis points towards the centre of the accelerator ring, the positive y-axis upwards and the z-axis along the beampipe. Besides the corresponding Cartesian coordinates, it is often useful to use cylindrical coordinates, given by the radial distance r in the xy-plane, the azimuthal angle ϕ and the polar angle θ. The latter is related to the pseudorapidity $\eta = -\ln\tan(\theta/2)$.

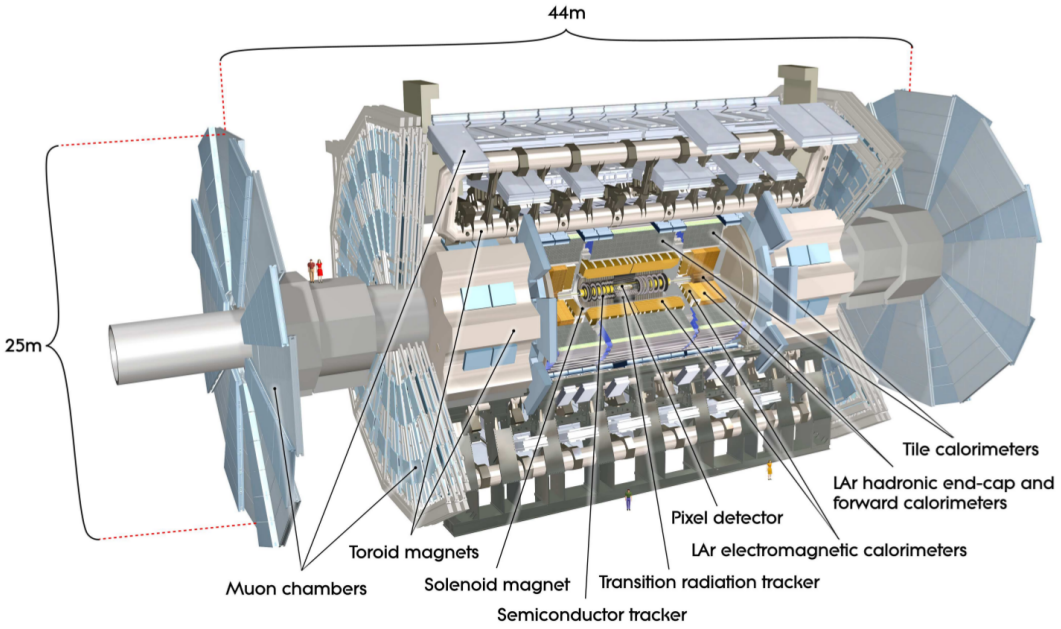


Figure 4.1.: Cut-away overview of the ATLAS detector. Picture taken from [107].

4.2.1. The Inner Detector

The ID consists of different types of tracking detectors, that are used to measure electrically charged particles. It covers the region of $|\eta| < 2.5$ and spans from a radius of approximately 3 cm to about 1 m. It is located inside a superconducting solenoid magnet, which creates a field of up to 2 T parallel to the z-axis. Therefore, electrically charged particles are deflected in the ϕ-direction but move in straight lines in the η-direction. The ID is made up of three components, the *pixel detector* (PIX), the *semiconductor tracker* (SCT), also commonly referred to as the *silicon strip detector*, and the *transition radiation tracker* (TRT), which all consist of barrel sections covering the central detector region and endcap disks covering the more forward regions. The barrel sections of the different subsystems and their locations as a function of the radius are sketched in Figure 4.2.

The PIX [110] is the ID subsystem that is located the closest to the beam pipe and the interaction point. It is made of silicon semiconductor pixel sensors and its barrel section consists of four layers of detectors, while the endcap sections on either side have three layers of disks each. Apart from the innermost barrel layer, the pixels have a size of 50×400 μm^2 and provide a spatial resolution on an individual track hit of $8 - 10$ μm along the ϕ-direction and $40 - 115$ μm along the z-direction in the barrel section and the r-direction in the endcaps. The innermost barrel layer is called *insertable B-layer* (IBL) [111, 112] and was inserted into the ID prior to the start of LHC Run 2 in 2015. Its pixels have a size of 50×250 μm^2. Due to it being located in close proximity to the beampipe, the IBL helps to improve the identification of jets containing b-flavoured hadrons.

The cylinder-shaped SCT [113] encases the PIX and is made of silicon microstrip sensors. It features four cylindrical layers in its barrel section and nine layers of disks in each endcap. Each layer has one sensor on either side and both sensors are rotated against each other at an angle of 40 mrad in order to improve the resolution along the strip direction, which is along the z-axis in the barrel section and along the r-direction in the endcaps.

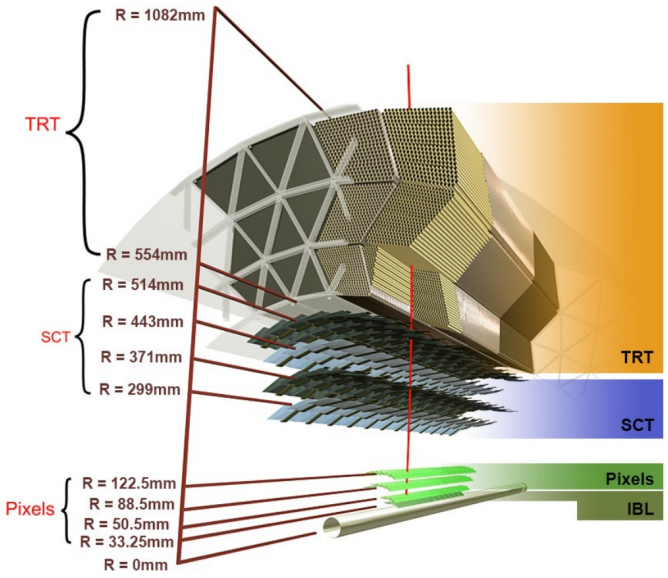


Figure 4.2.: The barrel sections of the different subdetector systems of the ID and their locations as a function of the radius. Taken from [109].

There are 8448 rectangular sensors in the barrel section and 6944 wedge-shaped sensors in the endcaps. Each sensor has a length between 5 and 13 cm and contains 768 readout strips, whose widths range between 16 and 22 μm. In the barrel section, the pitch, i.e. the distance between the centres of neighbouring readout strips, is 80 μm, while in the endcaps, the pitch ranges between 57 and 94 μm due to the sensors being wedge-shaped. The resolution is 17 μm along the ϕ-direction and 580 μm along z and r.

The outermost subdetector of the ID is the TRT [114]. It covers the region of $|\eta| < 2.0$ and consists of straw-tubes, which have a diameter of 4 mm and are filled with a gas mixture of 70% Xe, 27% CO_2 and 3% O_2. Elements suffering from gas leakage due to radiation damage have been refilled and operated with a modified gas mixture of Ar, CO_2 and O_2 since the beginning of Run 2. The barrel region is made of 72 layers of tubes with a length of 144 cm, which are parallel to the z-axis, and the endcaps consist of 160 layers of radially oriented tubes with a length of 36 cm. There is a total number of about 350 000 straw tubes. Compared to the silicon-based subdetectors, the number of layers is larger in the TRT, which results in an average of more than 30 detectable hits per track in this part of the ID. The different dielectric indices of the gas mixture inside the tubes and the scintillating fibres (barrel) and foils (endcaps) in between the tubes lead to x-ray transition radiation when electrically charged particles traverse the TRT. This radiation can ionise the Xe and Ar atoms, thus creating a detectable electric signal above an upper threshold, which can be distinguished from the low-threshold signal stemming from ionisation along the particle tracks. The spatial hit resolution in the plane perpendicular to the tubes is about 130 μm, but there is no resolution along the length of the straws. Since the intensity of the transition radiation is related to the Lorentz factor, the TRT provides separation power for distinguishing electrons from heavier charged particles, e.g. pions.

4.2.2. The Calorimeters

The next detector component outside of the ID is the calorimeter system, which is sketched in Figure 4.3 and consists of the *electromagnetic* (ECal) and the surrounding *hadronic*

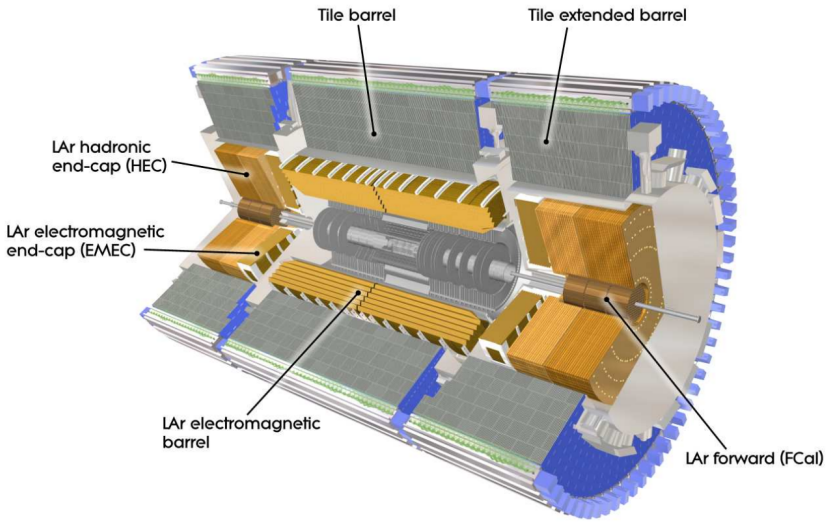


Figure 4.3.: Overview of the calorimeter system. Picture taken from [107].

calorimeter (HCal). Just like the rest of the detector, the calorimeter system is cylindrical and consists of barrel and endcap sections. It measures the energy of all particles except for muons and neutrinos. In particular, it also measures the energy of electrically neutral particles, which are not detected in the ID. But also in the case of electrons, the energy information from the calorimeter system is used to recover energy losses from bremsstrahlung. The ATLAS calorimeters are sampling calorimeters, i.e. they consist of alternating active and passive layers. Due to their density, the passive absorber layers create cascades of secondary particles, while the active layers measure the deposited energy. The energy of electrons and photons is measured in the ECal, whereas hadron showers usually begin to form in the ECal, but then proceed to the HCal, where they deposit most of their energy.

4.2.2.1. The Electromagnetic Calorimeter

The active material in the ECal is liquid argon and the absorbers are made of lead and steel. The ECal structure is depicted in Figure 4.4. It consists of three layers, which in the region $0 < |\eta| < 1.8$ are preceded by an additional *pre-sampling* layer, whose purpose is to estimate the energy losses due to material in front of it, i.e. the ID, the solenoid and the cryostat. The barrel section extends to $|\eta| < 1.475$ and the endcaps span the region of $1.375 < |\eta| < 3.2$. Between them, there is the so-called *crack region* at $1.37 < |\eta| < 1.52$, which is not fully sensitive. The three main layers are organised in an accordion-type structure in order to avoid insensitive cracks. The first layer, the so-called *strip layer*, is finely segmented in η, with each central cell spanning $\Delta\eta = 0.0031$, which is exploited in the identification of photons. The second layer has a much lower granularity of 0.025 in η and a larger depth of approximately 16 radiation lengths, as opposed to 4.3 radiation lengths in the first layer. The third ECal layer is thinner and has a depth of two radiation lengths.

The radiation length X_0 is a material property, which characterises the longitudinal evolution of an electromagnetic shower. It denotes the distance over which the energy of an electron is reduced to a factor of $1/e$ due to bremsstrahlung[1] and 7/9 of the mean free

[1] Here, e represents Euler's number, not the elementary charge.

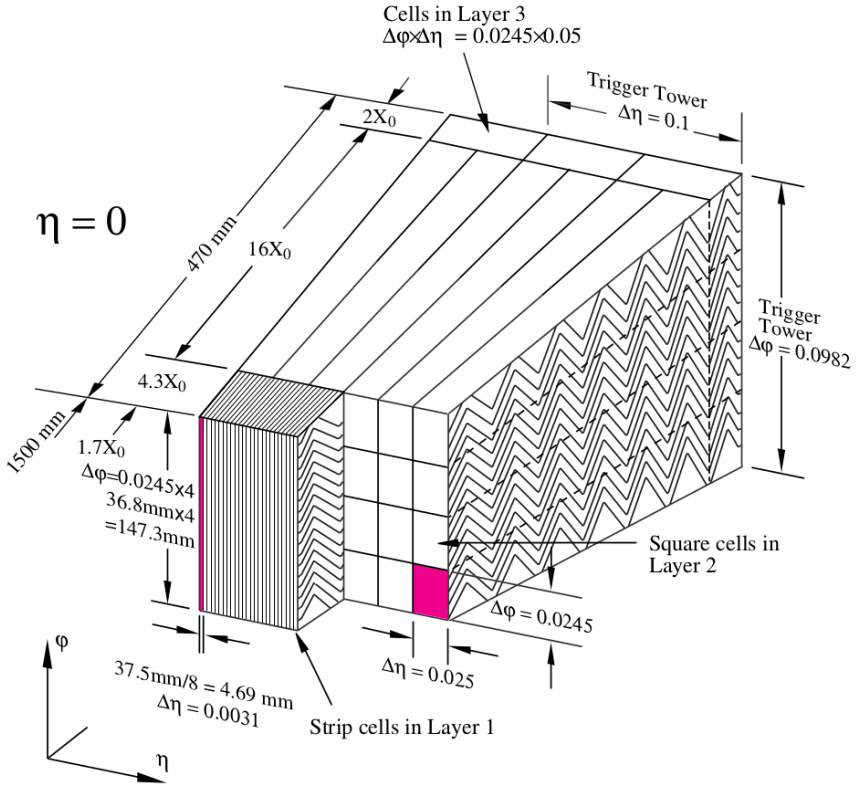


Figure 4.4.: Schematic of the barrel section of the electromagnetic calorimeter. The first layer is finely grained in the η-direction. The second layer captures the bulk of the energy depositions and has similar granularity in the η and ϕ-directions. Taken from [107].

path for pair production from photons. The ECal was designed to be at least 22 radiation lengths everywhere, so as to reliably capture most of the energy of electrons and photons. The relative energy resolution of the ECal is quantified as [115–117]

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/\mathrm{GeV}}} \oplus 1..2\% \ , \tag{4.2}$$

where the direct sum symbol implies adding the different terms in quadrature. The first contribution is called the *sampling term* and stems from the stochastic nature of the shower cascades, in particular fluctuations in the showers and fluctuations due to the sampling nature of the calorimeter. The second contribution, the *constant term*, is mainly due to non-uniformities in the detector [118].

4.2.2.2. The Hadronic Calorimeter

The purpose of the hadronic calorimeter is to contain and measure the energy of jets originating from quarks and gluons. It consists of three different parts:

- the *tile calorimeter* (TileCal) in the barrel section
- the *hadronic endcaps* (HEC) and
- the *forward calorimeters* (FCal)

The TileCal is a sampling calorimeter with steel absorbers and scintillators as the active material. It is segmented into cells of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ and three radial layers. It has a depth of about 2m, corresponding to 7.4 hadronic interaction lengths. The TileCal is split into the barrel and two extended barrel regions, covering $|\eta| < 1.0$ and $0.8 < |\eta| < 1.7$, respectively. The hadronic endcaps cover the region of $1.5 < |\eta| < 3.2$ and have copper absorbers and liquid argon as the active material. The latter is also true for the FCal, which covers the forward region of $3.1 < |\eta| < 4.9$. Its first layer uses copper absorbers and extends the ECal, whereas the second and third layers have absorbers made from tungsten.

4.2.3. The Muon Spectrometer

The *muon spectrometer* (MS) [120, 121], sketched in Figure 4.5, is the outermost and by volume the largest part of the ATLAS detector. It serves to detect muons with transverse momenta of at least 3 GeV. Its barrel section ($|\eta| < 1.05$) has three layers and the endcaps ($1.05 < |\eta| < 2.7$) consist of a total of four wheels. The MS is equipped with air core toroid magnets, which create a field of up to 2T. Due to the field configuration, muons get deflected in the η-direction but propagate along straight lines in the ϕ-direction. The barrel magnets extend up to $|\eta| = 1.6$, while the endcap magnets cover the region of $1.4 < |\eta| < 2.7$. The MS features four different types of detectors: *Monitored Drift Tubes* (MDT) [122] and *Cathode Strip Chambers* (CSC) [123] are used for tracking with a spatial resolution of about 50 μm. For triggering, faster *Resistive Plate Chambers* (RPC) [124] and *Thin Gap Chambers* (TGC) are used. Typical momentum resolutions of the MS are 4 % for muons with 100 GeV of transverse momentum and 10 % for 1 TeV [119].

4.2.4. The Trigger System

As mentioned before, the LHC provides proton bunches with a 25 ns bunch spacing, which corresponds to a collision frequency of 40 MHz. With a single event requiring a few MB of

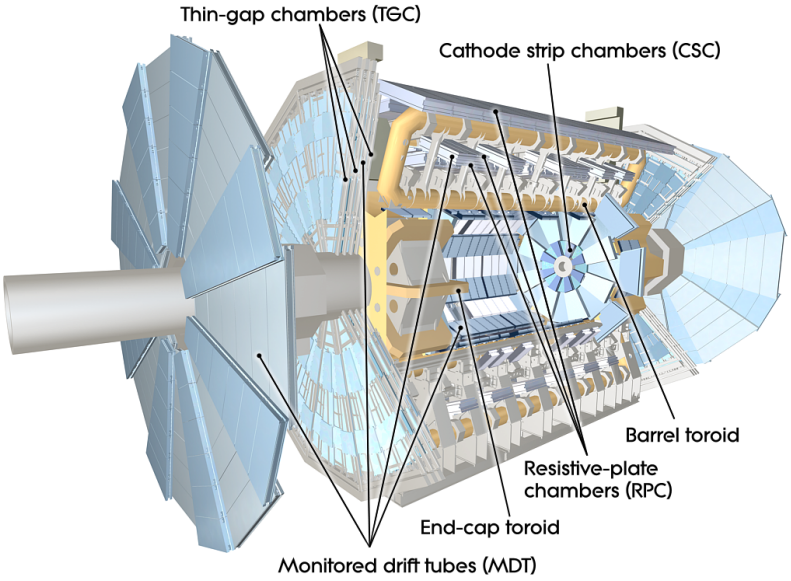


Figure 4.5.: Schematic of the muon spectrometer. Picture taken from [119].

storage space, neither the data bandwidth nor the storage space nor the computing power are available to record, permanently store and reconstruct all collision events for analysis. In order to reduce the data stream to manageable levels, a two-stage trigger system is used, consisting of the *level-1 trigger* (L1) and the *high-level trigger* (HLT) [125, 126].

The L1 trigger system is implemented in custom hardware and reduces the event rate to about 70-100 kHz. Until the L1 trigger decision is taken, all detector information has to be kept in the corresponding buffers, which creates strict time requirements. Therefore, the L1 trigger stage only uses coarse and partial detector information to decide whether to keep or discard an event. The following objects are considered, if the corresponding energy or momentum thresholds are exceeded:

- muons, as reconstructed from the RPCs and TGCs of the MS,
- energy deposits in multiple, adjacent cells of the calorimeter system and
- missing transverse momentum, calculated from the sum of the transverse components of all energy deposits in the calorimeters. The transverse component of an energy deposit E in a calorimeter cell at the polar angle θ is given by $E_{\mathrm{T}} = E \sin\theta$.

These criteria and combinations thereof are implemented with different thresholds in various *trigger lines*. If a collision event passes one or more trigger lines, the full detector information is read out and passed to the HLT along with so-called *regions of interest* (RoI), which encode detector regions that are to be analysed in more detail by the software-based HLT. This location information is used by time-critical HLT trigger algorithms, while faster algorithms may use the full detector information. Internally, the HLT stage is split up into a fast first-pass reconstruction and a slower precision reconstruction, which is only applied to events selected in the faster first-pass reconstruction. In the HLT, information from the ID is used for the first time and tracks are reconstructed in RoIs identified by the L1 trigger stage. HLT-level calorimeter reconstruction uses the *sliding window algorithm*, which uses rectangular grids of calorimeter cells, for electron and photon candidate reconstruction. Tau lepton and jet candidates as well as the missing transverse momentum are reconstructed by means of the *topo-clustering algorithm* [127], which starts from a seed cell and

iteratively adds neighbouring cells based on the significance of their energy depositions, thereby creating clusters of variable, irregular shapes. For HLT muon reconstruction, the precise MDT and CSC tracking modules are used in RoIs and ID tracking information is taken into account. Similar to ID track reconstruction, the muon track reconstruction is split into a fast and a more precise stage.

In order to keep the overall trigger rates at an acceptable level, the instantaneous luminosity can be capped on purpose. This technique, called *luminosity levelling*, is especially used during the beginning of a physics fill. In typical Run 2 conditions with a levelling at an instantaneous luminosity of $1.5 \cdot 10^{34}$ cm^{-2}s^{-1}, the peak number of interactions per bunch crossing is approximately 60 and the peak L1 trigger rate around 80 kHz. In such conditions ATLAS can collect about 5 fb^{-1} per week. Without luminosity levelling and a peak instantaneous luminosity of up to $2 \cdot 10^{34}$ cm^{-2}s^{-1} up to 80 interactions per bunch crossing are possible. These numbers as well as the precise values of the p_{T} thresholds for the different trigger lines depend on the running conditions of the LHC, e.g. the filling scheme, and therefore have to be adjusted accordingly.

Chapter 5

Reconstruction of Physics Objects

As described in Section 4.2.4, the events are partially reconstructed with different levels of granularity already during data taking in order to evaluate trigger decisions. This is referred to as *online reconstruction*. The full-detail reconstruction is later run on the events which were stored for analysis. This is called *offline reconstruction* and the corresponding software is part of ATLAS' `Athena` framework [128]. In this chapter, aspects of the offline reconstruction will be described for the physics objects relevant to the analyses presented later in this thesis.

5.1. Tracking

The reconstruction of tracks [129, 130] is based on hits from charged particles in the inner detector, which was described in Section 4.2.1. Combinations of three hits in the PIX and SCT subdetectors are formed and used as track seeds, if they satisfy a number of quality criteria. These criteria include requirements on the transverse momentum, impact parameter and compatibility with one additional hit [130]. A Kalman filter [131] is used to extrapolate the track seeds outwards to the TRT and add matching hits along the trajectory. At first, this is attempted assuming that the track candidate stems from a pion, i.e. only a minimal amount of energy is lost in interactions with the detector material. Only if this is not possible, a second attempt is performed with a different model, which allows for energy loss. This Kalman extrapolation and fitting stage is performed in a deliberately loose manner in order to identify all reasonably conceivable track candidates. The resulting track candidates are scored, ranked and subjected to a disambiguation algorithm. Passing track candidates are refitted using the Global χ^2 Track Fitter [132]. While the Kalman filter proceeds from one hit to another along the track during the forward filtering and backward smoothing stages, the Global χ^2 Track Fitter minimises a χ^2 function describing the entire track, including residuals between expected and measured hit positions as well as scattering angles and energy losses at each traversed material surface, in a simultaneous fit. It is about a factor of two slower than the Kalman filter track fitting but yields

estimates of the scattering angles and does not need initial estimates of the errors of the track parameters. The refitted tracks are accepted if they still pass the disambiguation. Due to the orientation of the procedure, this is referred to as the *inside-out* algorithm. In order to capture tracks from converted photons and long-lived particles, a similar track fitting stage is performed in the opposite direction, i.e. *outside-in*, starting from TRT track seeds built from hits that were not associated to any track during the inside-out stage. Afterwards, vertices are reconstructed from the tracks and fitted in order to identify the vertex positions and the corresponding uncertainties. Vertices whose location is compatible with the beam spot are labelled as *primary vertices* (PV) [133]. Typically, the PV with the largest sum of squared track p_{T} associated is taken to be the main vertex of the collision, the remaining PVs are assumed to be from pile-up. Vertices not compatible with the beam spot region are considered to be secondary vertices.

5.2. Jets

As described in Section 2.1.1, quarks and gluons cannot exist as freely propagating particles, but give rise to jets, i.e. sprays of collimated hadrons. The energy of these jets is measured mainly in the hadronic calorimeter. The reconstruction of jets proceeds in two steps. First, calorimeter cells, which detect significant energy depositions E_{cell} above the noise threshold σ_{noise}, are grouped together in clusters. For jets, this is done by means of the topo-clustering algorithm [127, 134], which starts from seed cells with a high signal significance $\zeta = E_{\mathrm{cell}}/\sigma_{\mathrm{noise}} > 4$, to which all laterally and radially neighbouring or overlapping cells are added. If these added cells pass a reduced significance threshold of $\zeta > 2$, their respective neighbours are added. This procedure is repeated iteratively, until no new cells can be added or no clusters can be merged. Since hadronic jets may begin to form showers already in the ECal, cells from the entire calorimeter system, i.e. from both the ECal and the HCal, are considered. Due to the nature of this procedure, topo-clusters do not have a predetermined shape. To each topo-cluster, a massless four-vector is associated, pointing from the geometric centre of the detector towards the energy-weighted barycentre of the cluster. In a second step, jet algorithms are applied to the resulting four-vectors in order to cluster them into well-defined jets. A common choice is the *anti-k_t* algorithm [135] with a radius parameter of $R = 0.4$.

The determination of the jet energy from the signals of the cells associated to a given jet relies on a calibration procedure. Since the calorimeters are sampling calorimeters, not all of the jet energy is detected in the active material layers. Therefore, the jet energy is approximated as a weighted sum of the energies in the relevant calorimeter cells. The weights are given in bins of p_{T} and η and are determined from simulations in such a way that the reconstructed jet energy deviates as little as possible from the truth-level jet energy. Furthermore, a series of calibration corrections is applied, accounting for the displacement of the primary vertex with respect to the geometric centre of the detector and the contributions from pile-up. Finally, data-driven corrections, so-called *in-situ* corrections, are applied, which serve to validate the simulation-based calibration and account for mismodelling in the Monte Carlo simulations with respect to data. The in-situ corrections are derived from different procedures comparing reconstructed jets to well-measured reference objects. They are based on the *η-intercalibration*, which is derived from reference events with two jets in different detector regions, the *vector boson balancing*, which exploits the momentum balance between a jet and a photon or two electrons or muons from a Z boson decay, and the multi-jet calibration, in which a number of low and

medium p_{T} jets is balanced against few high p_{T} jets in order to extend the calibration to the hard regime.

Further details are presented in [136].

5.3. Electrons and Photons

5.3.1. Reconstruction

In Run 1 and Run 2 analyses using `Athena` releases up to and including version 20, the reconstruction of electrons[1] and photons is based on rectangular clusters formed using the *sliding window algorithm*. Later analyses based on `Athena` releases beginning with version 21 make use of the *topo-cluster* approach, which has already been in use since Run 1 in the reconstruction of hadronic jets. Both algorithms are introduced in [134]. Since both procedures are used in the analyses presented later in this thesis, they are both described briefly in this section.

The reconstruction of electrons and photons begins with the identification of *seed clusters*. In the sliding-window algorithm, grids of 3×5 cells of the second ECal layer, corresponding to a size of $\Delta\eta\times\Delta\phi = 0.075\times0.1225$, are used [137]. The formation of electron and photon topo-clusters proceeds as described in Section 5.2. In order to preselect topo-clusters for the electron and photon reconstruction, the electromagnetic energy fraction, f_{EM}, is considered, which describes the fraction of the total cluster energy that was deposited in the ECal. A preselection cut is used to reject clusters from pile-up interactions and hadronic jets. These preselected clusters are then restricted to only contain ECal cells.

After cluster formation, ID tracks are loosely matched to the seed clusters, using geometric $\Delta\eta$ and $\Delta\phi$ matching. The presence of tracks matching the seed clusters is used to identify electron and converted photon candidates. In order to compensate for energy losses due to bremsstrahlung, the track momenta are rescaled to the cluster energies. Loosely matching tracks with hits in the silicon detectors are refitted using the *Gaussian Sum Filter* (GSF) algorithm [138], which allows for non-Gaussian energy losses from bremsstrahlung due to material interactions and improves the track parameter resolution in comparison to the previous track fits [139].

Double-track conversion vertex candidates are built from track pairs with opposite charges that are likely to be electrons. Track pairs passing a set of quality criteria are subjected to a constrained conversion vertex fit in order to determine the vertex location. For photons converting in the outermost ID layers, it is more likely that one of the two tracks is not reconstructed. In order to recover such conversions, *single-track conversion candidates* can be formed based on individual tracks without hits in the innermost ID layers and the location of the first measurement of the reconstructed track is taken to be the vertex location. Finally, an arbitration is performed to decide whether an object is classified as a converted or unconverted photon or an electron. Essentially, electrons are defined from a cluster and at least one matching track, converted photons from a cluster and a matching conversion vertex and unconverted photons from a cluster not matched to a track or a conversion vertex. More details are given in [139].

In the topo-cluster reconstruction, an additional step is performed in which superclusters are now built from one or more topo-clusters. Proceeding along a p_{T}-ordered list of

[1] Often, the term *electron* is used to refer collectively to both the negatively charged electron and the positively charged positron.

topo-clusters, supercluster seed candidates are picked and topo-clusters with lower p_{T} are considered as satellite clusters. For both photons and electrons, topo-clusters are added to the seed cluster to form a supercluster if they are located within a 3×5 cell window corresponding to $\Delta\eta \times \Delta\phi = 0.075 \times 0.1225$ around the seed cluster. For electrons, topo-clusters in a wider window are also added if they are associated to the same track as the seed cluster. For converted photons, also topo-clusters associated to the same conversion vertex as the seed cluster are added. The lateral size of all topo-clusters in ϕ is then restricted to three cells in the barrel section and to five cells in the endcap region. This aims at reducing the systematic uncertainties of the energy calibration.

As a last step of both reconstruction methods, the energy is measured from deposits in the relevant ECal cells. In the sliding-window approach, a 3×7 cell grid corresponding to $\Delta\eta \times \Delta\phi = 0.075 \times 0.172$ is used in the barrel. In the endcaps, a 5×5 cell grid corresponding to $\Delta\eta \times \Delta\phi = 0.125 \times 0.125$ is used for all photon candidates. For electron candidates, grids of 3×7 and 5×5 cells are used in the barrel and the endcaps, corresponding to $\Delta\eta \times \Delta\phi = 0.075 \times 0.172$ and 0.125×0.125, respectively [107]. Similar to jets, also the energy of photons and electrons is calibrated in order to compensate for energy losses due to interactions with upstream material, lateral and longitudinal leakage and to also account for the sampling nature of the calorimeter. The calibration is based on simulated events as well as data-driven correction factors and is performed separately for electrons, converted and unconverted photons [116, 140].

A major part of the calibration for both electrons and photons is based on a multivariate approach using boosted decision trees [140]. In addition to the raw, uncalibrated energy, other variables denoting the energy fractions deposited in the different ECal layers as well as the cluster positions are considered. Among the other steps of the calibration procedure are the intercalibration of the different longitudinal layers and a compensation for detector inhomogeneities. More details are presented in [116].

5.3.2. Isolation

Most often, analyses are interested in *prompt* photons and electrons, which are produced in the hard scattering process or originate from the decay of heavy resonances, such as the W or the Z boson, as opposed to a background originating from hadron decays inside jets or radiating quarks. Rejection power against this non-prompt background is provided by so-called *isolation* (ISO) criteria, which are based on two quantities, the *calorimeter isolation* and the *track isolation*. For the calorimeter isolation, the transverse energies of all topo-clusters, whose barycentres lie within a cone of size ΔR[2] around the direction of the photon or electron candidate, are summed. From this, the candidate's energy itself is subtracted along with the contributions from the underlying event and pile-up, which are calculated individually for each event. This defines the *isolation transverse energy*, $E_{\mathrm{T}}^{\mathrm{iso}}$. For the track isolation, $p_{\mathrm{T}}^{\mathrm{iso}}$, all tracks with $p_{\mathrm{T}} > 1$ GeV within a cone of size ΔR around the photon or electron candidate are summed and contributions from any tracks from conversion electrons and from the track of the electron candidate itself are subtracted. In order to reject pile-up contributions, only tracks compatible with the relevant primary vertex are considered. Further details and references as well as lists with the most common photon and electron isolation working points are given in [137] and [141].

[2] $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where $\Delta\phi$ and $\Delta\eta$ are the differences in azimuthal angle and pseudorapidity between the centre and border of the cone.

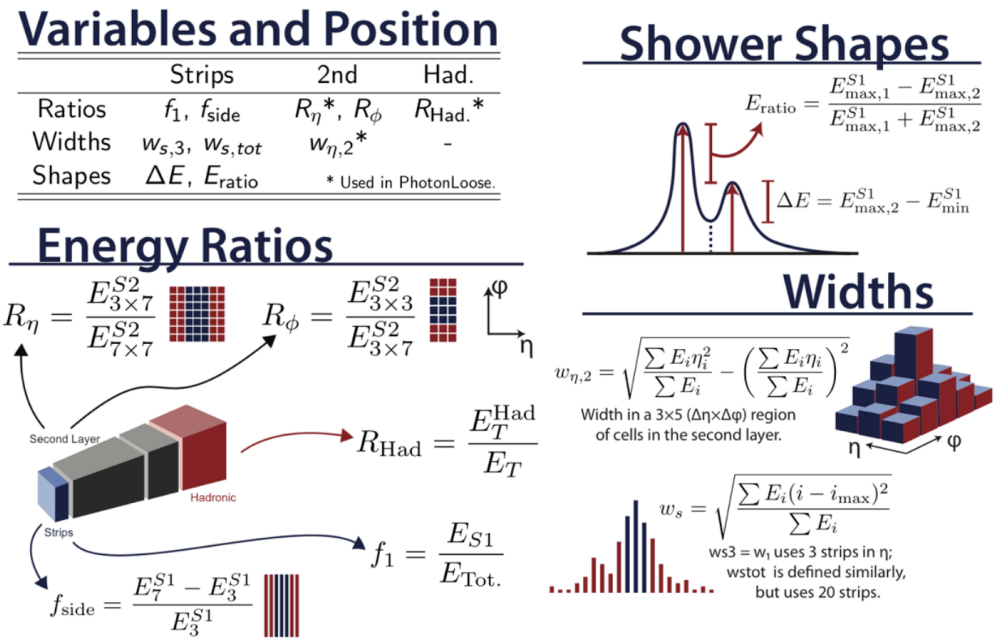


Figure 5.1.: Illustration of the shower shape variables that describe the structure of the showers from photons and electrons in the calorimeters. Picture taken from [142].

5.3.3. Identification

The purpose of electron and photon identification (ID) is to establish whether or not a particular reconstructed object is a prompt signal photon or electron. The identification procedure is based on so-called *shower shape* variables, sketched in Figure 5.1 and described in Table 5.1, which characterise the structure of the associated showers in the calorimeters. They are mostly calculated from cells in the first and second layers of the ECal and comprise ratios of energies in different layers and in grids of different sizes as well as variables describing the width and the peak structure of the energy distributions across different cells. Besides the shower shape variables, electron identification also makes use of track-related variables.

5.3.3.1. Photon Identification

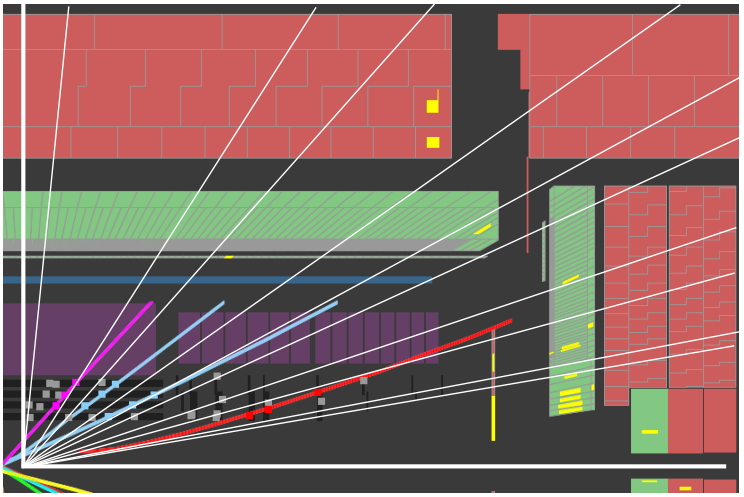
The identification of photons is based on a series of subsequent selection criteria on the shower shape variables. Its aim is to distinguish signal photons from the background due to photons from hadron decays inside jets or hadronic jets mimicking photons. The ID is optimised separately for converted and unconverted photons in bins of $|\eta|$ to take into consideration variations in the geometry of the detector. The $|\eta|$-bins are sketched in Figure 5.2. In particular, the transition region between the ECal barrel and endcap sections at $1.37 < |\eta| < 1.52$ is excluded. A table specifying which shower shape variable is used for which photon ID working point can be found in [137].

5.3.3.2. Electron Identification

The electron ID [141] strives to distinguish between prompt electrons and background electrons from photon conversions, jets mimicking prompt electrons and non-prompt electrons from the decays of hadrons containing heavy-flavour quarks. In contrast to the photon

Shower shape	Description	γ ID *Loose*	γ ID *Tight*	e ID
$R_{\text{had},1}$	longitudinal ratio of the cluster E_{T} in the first HCal layer over the cluster E_{T} in the ECal, used for $\lvert\eta\rvert < 0.8$ and $\lvert\eta\rvert > 1.52$.	✓	✓	✓
R_{had}	longitudinal ratio of the cluster E_{T} in the entire HCal over the cluster E_{T} in the ECal, used for $0.8 < \lvert\eta\rvert < 1.37$	✓	✓	✓
f_3	longitudinal ratio of the cluster energy in the third ECal layer over the total cluster energy in the ECal			✓
R_η	lateral ratio along the η-direction of the energy in 3×7 over 7×7 cells in the second ECal layer	✓	✓	✓
R_ϕ	lateral ratio along the ϕ-direction of the energy in 3×3 over 3×7 cells in the second ECal layer		✓	✓
$w_{\eta,2}$	lateral shower width along the η-direction calculated in 3×5 cells in the second ECal layer	✓	✓	✓
f_1	longitudinal ratio of the energy in the first ECal layer over the total cluster energy in the ECal		✓	✓
f_{side}	lateral ratio along the η-direction of the energy in the two sideband cells at both sides over the central three cells in the first ECal layer		✓	
ΔE	difference between the energy of the second maximum and the minimum between the first and second maxima in the first ECal layer		✓	
E_{ratio}	ratio of the difference over the sum of the energies of the first and second maxima in the first ECal layer		✓	✓
$w_{s,3}$	Lateral shower width in a grid of 3×2 strip cells in the first ECal layer		✓	
$w_{s,\text{tot}}$	Lateral shower width in a grid of 20×2 strip cells in the first ECal layer		✓	✓

Table 5.1.: Description of the shower shape variables and their use in photon and electron identification [137, 141]. Cell window sizes are given in units of $\Delta\eta \times \Delta\phi$ of the size of the calorimeter cells in the relevant layer around the cluster barycentre. Besides the indicated shower shape variables, electron identification also makes use of variables related to the track quality and track-cluster matching.



Eta	Angle
0	90°
0.1	84.3
0.6	57.5
0.8	48.4
1.15	35.1
1.37	28.5
1.52	24.7
1.81	18.6
2.01	15.3
2.37	10.7
2.47	9.7

Figure 5.2.: Schematic of the ATLAS inner detector and the electromagnetic and hadronic calorimeters. Superimposed are the bins in pseudorapidity η and the corresponding polar angles θ used in electron and photon identification. Taken from [144].

ID, it is based on a multi-variate likelihood procedure. Besides the shower shape variables indicated in Table 5.1, information from the silicon detectors and the TRT as well as variables relating the tracks to the calorimeter clusters are used. From the likelihoods L_S and L_B, that an electron candidate is from signal or from background, the discriminant $d_L = L_S/(L_S + L_B)$ is formed. After applying an inverse sigmoid function [143], the transformed discriminant is used to define the ID working points in bins of $|\eta|$ and E_{T}. The $|\eta|$-bins are sketched in Figure 5.2. Further details can be found in [141].

5.4. Muons

Muon tracks are first reconstructed separately and independently in both the inner detector and the muon spectrometer. For the ID track candidates, the collection of tracks built using the procedure outlined in Section 5.1 is considered. For MS track reconstruction, pattern recognition methods are used to form track segments in all MS layers and chambers individually. Matching track segments are then assembled into track candidates, starting from seed segments in the middle layer. In the next step, requirements on the track quality are imposed and a disambiguation of segments shared between different tracks is carried out. Finally, an iterated χ^2-fit is performed while adding and removing individual hits in order to obtain the best-fitting tracks and their parameters. Where possible, muon tracks from the ID and the MS are then matched, combined and refitted to form *combined* (CB) muons. There are other muon types, for which MS tracks are combined with other detector information, in order to at least partially recover muons passing through non-instrumented detector regions. In particular, muons that are formed entirely from MS information and that have tracks compatible with the beam spot region are referred to as *extrapolated* (ME) muons. Due to the coverage of the ID, combined muons are limited to $|\eta| < 2.5$, whereas extrapolated muons extend up to $|\eta| < 2.7$.

Muon isolation is based on the same variables as electron and photon isolation, which was discussed in Section 5.3.2.

The purpose of muon identification is both to ensure an accurate momentum measurement and to retain prompt muons, while rejecting muons from hadron decays, which predominantly originate from decaying pions and kaons. The identification procedure is based on the fact that tracks from muons originating from charged hadron decays exhibit characteristic tracks with a kink structure, which typically leads to a bad track fit quality and differences between the track momentum measurements from the ID and the MS. Different variables sensitive to this are exploited in muon identification. Additionally, requirements on the number of ID and MS hits and holes are imposed, where a hole is defined as an active sensor without a hit along the track.

More details on muon reconstruction, isolation and identification can be found in [121].

Chapter 6

Statistical Methods

Given experimental data, such as a collection of n observed events $\{\boldsymbol{y}_i\}_{i=1}^n$, and an expected result derived from an assumed theory, statistical methods can be used to analyse the data, estimate or constrain parameters of the theory, quantify the agreement between the experimental observation and the theory and, eventually, to possibly reject the theory as incompatible with the data. In this chapter, a few of the methods commonly used in high energy physics and in the analyses described in the following parts of this thesis will be introduced.

The discussion in this chapter is partly based on Chapters 38 and 39 of [7] where further details can be found. A few common random distributions that will also appear in this chapter are listed in Appendix A.1.

6.1. Parameter Estimation

The goal of parameter estimation is to extract the best-fitting or most probable values $\boldsymbol{\theta}$ of a model or theory from a set of experimental observations. The variables $\boldsymbol{\theta}$ are also called *parameters of interest*.

6.1.1. The Method of Least Squares

The method of least squares is also known as the χ^2-method for reasons that will become clear later in this section. A common application is the comparison of experimental data binned into a histogram $\{n_i\}_{i=1}^b$ with b bins to the expected or predicted values $\{N_i\}_{i=1}^b$ in each of the bins. The expectation may depend on a number of parameters $\boldsymbol{\theta}$ of the underlying theory. The test statistic is called χ^2 and takes the form

$$\chi^2(\boldsymbol{\theta}) = \sum_{i=1}^{b} \left(\frac{n_i - N_i(\boldsymbol{\theta})}{\delta_i} \right)^2 \quad , \tag{6.1}$$

where δ_i denotes the uncertainty for each bin i. The smaller the value of χ^2, the better the agreement between the data and the expectation. Consequently, χ^2 can be used as the figure of merit in a fit, where it is minimised by varying the parameters of the assumed theory so as to identify those parameter values that match best the experimental data. The lowest, best-fit value of χ^2 can then be used to quantify the level of agreement between the theory and the data.

χ^2 is also the name of a random distribution which is related to the aforementioned χ^2-method. If a set of random variables $\{x_i\}_{i=1}^k$ are normally distributed with a mean of 0 and a variance of 1 and independent of each other, then their quadratic sum $\sum_{i=1}^k x_i^2$ will follow a χ^2 distribution with k degrees of freedom. In the method of least squares, the number of degrees of freedom is in principle given by the number of bins b. However, these are not all independent but rather correlated in some way that is encoded in the theory that underlies the expectation. Therefore, the number of degrees of freedom is given by $k = b - m$ where m is the number of fitted theory parameters $\boldsymbol{\theta}$. Using this known behaviour, one can compare the obtained best-fit χ^2-value and the number of degrees of freedom against the expected χ^2-distribution to calculate how probable a χ^2-value at least as large as the observed one is, assuming the expected model is correct. This concept will be elaborated in more detail in Section 6.2.

Besides estimating the true values of the theory parameters $\boldsymbol{\theta}$, it is also important to define the uncertainties of these estimates. Procedures to do this, also in case the method of least squares is used, will be introduced in Section 6.1.2.

6.1.2. The Likelihood Method

A different method for estimating model parameters is the likelihood method which owes its name to the likelihood function L describing the probability of making a certain observation when assuming a specific theory or model:

$$L = L(\boldsymbol{\theta}) = P(\text{data}|\text{theory}) \ . \tag{6.2}$$

The likelihood method comes in a binned and an unbinned variant. As the name suggests and just like the method of least squares, the former is based on binned data, whereas the latter uses individual events, which prevents the loss of information that is inevitable when binning data. Therefore, the unbinned likelihood method is better suited for experiments or analyses with a smaller number of events.

For practical reasons, partly related to numerical precision, instead of maximising the likelihood L, usually the *negative log-likelihood* $-2\ln L$ is minimised.

6.1.2.1. Binned Likelihood Fits

The basic formulation of the binned likelihood method reads as follows:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{b} P(n_i|N_i(\boldsymbol{\theta})) \tag{6.3}$$

The product runs over the histogram bins and P is a probability distribution function that relates the expected and observed bin contents. P usually is taken to be a Poisson distribution

$$\mathcal{P}(n;N) = \frac{N^n e^{-N}}{n!} \ . \tag{6.4}$$

The reason for this is given in Appendix A.2.

6.1.2.2. Unbinned Likelihood Fits

As mentioned before, the unbinned likelihood uses individual events and in its simplest form reads

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} P(\boldsymbol{y}_i|\boldsymbol{\theta}) \ , \tag{6.5}$$

i.e. one or more observables $\boldsymbol{y}_i$ are compared to their expected distribution for each event individually as opposed to using the number of times that $\boldsymbol{y}_i$ takes on a value in certain ranges which is done in the binned case.

If the total number of expected events N depends on the model parameters, i.e. $N = N(\boldsymbol{\theta})$, the number of observed events can also be used to constrain the model parameters in the fit. This can be achieved by using the *extended likelihood method*

$$L(\boldsymbol{\theta}) = \frac{N(\boldsymbol{\theta})^n e^{-N(\boldsymbol{\theta})}}{n!} \prod_{i=1}^{n} P(\boldsymbol{y}_i|\boldsymbol{\theta}) \ . \tag{6.6}$$

In high energy physics, one is often interested in identifying a signal process against one or more background processes. The unbinned likelihood function in this case reads

$$L(\boldsymbol{\theta}) = \frac{N(\boldsymbol{\theta})^n e^{-N(\boldsymbol{\theta})}}{n!} \prod_{i=1}^{n} \left(\frac{N_{\text{sig}}(\boldsymbol{\theta})}{N(\boldsymbol{\theta})} P_{\text{sig}}(\boldsymbol{y}_i|\boldsymbol{\theta}) + \frac{N_{\text{bkg}}(\boldsymbol{\theta})}{N(\boldsymbol{\theta})} P_{\text{bkg}}(\boldsymbol{y}_i|\boldsymbol{\theta}) \right) \ , \tag{6.7}$$

where P_{sig} and P_{bkg} are the distributions of the observable $\boldsymbol{y}$ for the signal and background processes and $N = N_{\text{sig}} + N_{\text{bkg}}$.

6.1.2.3. Uncertainty Estimation

For independent and identically distributed random variables $\{\boldsymbol{y}_i\}_{i=1}^{n}$, the likelihood function factorises:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} P(\boldsymbol{y}_i|\boldsymbol{\theta}) \ . \tag{6.8}$$

By virtue of the central limit theorem, in the large sample limit ($n \to \infty$) the likelihood function approaches a multi-dimensional Gaussian distribution with a mean given by the best-fit parameter values $\hat{\boldsymbol{\theta}}$ and the covariance matrix V, which describes the uncertainties and correlations of the parameters. The negative log-likelihood is therefore given by

$$-2\ln L(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T V^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) + c \ , \tag{6.9}$$

where the constant c describes the value of the minimum. As can readily be seen in the one-dimensional case, departing from the best-fit parameter value $\hat{\theta}$ by $\pm Z\sigma$ leads to an increase of the negative log-likelihood of

$$-2\ln L(\hat{\theta} \pm Z\sigma) + 2\ln L(\hat{\theta}) = Z^2 \ , \tag{6.10}$$

where σ denotes the uncertainty of the parameter θ. Therefore, in the one-dimensional case the boundaries of the $Z\sigma$ confidence intervals can be determined by identifying those parameter values of θ for which the negative log-likelihood is Z^2 units larger than at its minimum. The same procedure can also be applied for a higher number of dimensions which under the aforementioned assumptions results in confidence regions described by (hyper-)ellipsoids. Note that for an increasing number of dimensions, the volume fraction

Number of dimensions	Z				
	1	2	3	4	5
1	0.6827	0.9545	0.9973	$1 - 6.33 \cdot 10^{-5}$	$1 - 5.73 \cdot 10^{-7}$
2	0.3935	0.8647	0.9889	$1 - 3.35 \cdot 10^{-4}$	$1 - 3.73 \cdot 10^{-6}$
3	0.1987	0.7385	0.9707	$1 - 1.13 \cdot 10^{-3}$	$1 - 1.54 \cdot 10^{-5}$

Table 6.1.: Integrals of the Gaussian distribution for different deviations $Z\sigma$ from the mean value and numbers of dimensions. For the one-dimensional case, these correspond to the well-known 68% and 95% confidence levels. For a higher number of dimensions, larger values of Z have to be chosen in order to identify the regions corresponding to the same confidence levels.

enclosed by the region of constant likelihood for a given value of Z decreases. While e.g. in one dimension the regions bounded by $Z = 1$ and $Z = 2$ contain the well-known fractions of approximately 68% and 95% of the Gaussian distribution, these values are reduced to about 39% and 86% in two dimensions as listed in Table 6.1.

The relation between the uncertainty on a parameter and the shape of the likelihood function close to the minimum is also plausible intuitively: The better constrained one or more parameters are, i.e. the smaller their uncertainties are, the narrower the likelihood function will be along the dimensions corresponding to those parameters. Since this is directly related to the curvature of the log-likelihood, it is possible to determine the covariance matrix from the second derivatives:

$$V_{ij}^{-1} = -\frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j} = \frac{1}{2}\frac{\partial^2}{\partial\theta_i \partial\theta_j}(-2 \ln L) \ . \tag{6.11}$$

For sufficiently well-behaved likelihood functions, this strategy necessarily yields symmetric or (hyper-)ellipsoidal uncertainties and confidence regions.

Finally, it is worth noting that when using an extended likelihood function, the negative log-likelihood function will in general not be quadratic or even symmetric around the minimum. In contrast to when deriving the covariance matrix from the second derivatives, the method based on Equation (6.10) can be used to determine asymmetric or non-ellipsoidal confidence regions.

6.1.2.4. Relation Between the Negative Log-Likelihood Function and the Method of Least Squares

For Gaussian measurements the (binned) negative log-likelihood and the χ^2 test statistic are identical up to a constant $l_{\min}$

$$-2 \ln L(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T V^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) + l_{\min} = \chi^2 + l_{\min} \ . \tag{6.12}$$

Therefore, when using the method of least squares, the uncertainty regions can be determined using the same prescriptions as for the likelihood method.

6.2. Hypothesis Testing and Confidence Intervals

Besides estimating the values and uncertainties of parameters of a given model, a common question is whether the observed data is adequately described by said model or whether the data can be used to refute a theory. Such efforts go by the name of hypothesis testing.

After an introduction into the basics of hypothesis testing, the procedures used for the analysis of LHC data are sketched in the following sections.

6.2.1. Introduction

The hypothesis that is put to the test is commonly referred to as the *null hypothesis* H_0. It may depend on one or more parameters of interest $\boldsymbol{\theta}$ that are part of the theory which underlies the hypothesis. One may also have an *alternative hypothesis* H_1 that is compared to H_0 and adopted in case the null hypothesis is rejected.

Hypotheses can be accepted or rejected at a given *confidence level* (CL), which has to be decided upon before conducting the test. This is related to the choice of a criterion whose evaluation then leads to one decision or the other. Since the input for that decision is based on experimental data, hypotheses can never be accepted or rejected with absolute certainty. Specifically, there are two kinds of possible errors:

- The *type-1 error* indicates the probability of falsely rejecting the null hypothesis H_0 even though it is in reality true. This error is also referred to as α-error and α is also used to denote the size of this error.
- The *type-2 error* gives the probability of falsely accepting the null hypothesis while in reality it is in fact wrong. It is also called β-error with β at the same time being used to indicate the size of this error.

The quantity $1 - \alpha$ is the aforementioned confidence level and $1 - \beta$ is called the *power* of a specific hypothesis test. It is worth pointing out that both types of errors are not independent and minimising one of them will often come at the price of increasing the other one. This is also sketched in Figure 6.1, where the expected distributions of an observable $\mathcal{O}$ as predicted by the theories corresponding to the two different hypotheses are shown. The decision criterion is encoded in a boundary value $\mathcal{O}_{\text{crit}}$ that is defined by the value of α chosen by the analyser – an observable value $\mathcal{O}_1$ larger than $\mathcal{O}_{\text{crit}}$ will lead to rejection of H_0 even if H_0 is in fact true (type-1 error) and an observable value $\mathcal{O}_2$ smaller than $\mathcal{O}_{\text{crit}}$ will lead to acceptance of H_0 even if it is false (type-2 error). Changing the decision criterion may reduce α but will at the same time increase β and reduce the test's power – and vice versa.

The null hypothesis may be a single, fixed statement H_0, but it may also be a collection of possibly continuous statements $H_0(\boldsymbol{\theta})$ in which case it will depend on one or more parameters of interest $\boldsymbol{\theta}$. In the latter case, the goal of the hypothesis testing is to determine one or more parameter intervals or regions for which the hypothesis can be accepted or rejected at a previously decided confidence level $1 - \alpha$. In the simplest cases, such a region may be defined based on the best-fit values of the parameters and the associated uncertainties, but in general the situation will be more complicated. One such example would be the presence of additional restrictions on the allowed parameter values, e.g. a parameter that may be required to be non-negative by the underlying theory may have a positive best-fit value but a larger uncertainty. In that case, such a naive lower limit would extend into the disallowed region. Therefore, more sophisticated methods are required in general. Typical examples for the different kinds of hypotheses in high energy physics are summarised in Table 6.2.

One other important concept in the context of limit setting on parameters is *coverage*, which describes the fraction of times in which – for a large number of hypothetical or actual repetitions of the entire experiment – the unknown true parameter value lies within

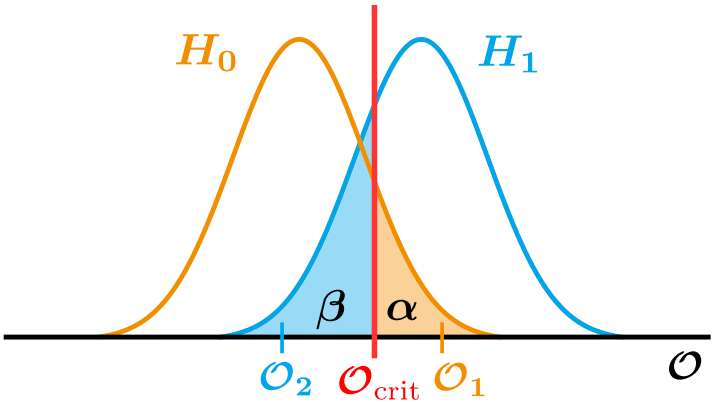


Figure 6.1.: Sketch of the type-1 and type-2 errors of hypothesis testing. The type-1 error describes the probability of falsely rejecting the true null hypothesis H_0, while the type-2 error gives the probability of falsely accepting the false null hypothesis. In the frequentist approach, the decision to accept or reject the null hypothesis is made by comparing the observed value to a critical value $\mathcal{O}_{\text{crit}}$ which is derived from the desired level of confidence $1 - \alpha$.

Type	Application / example
Fixed hypothesis	testing for a new discovery by checking the compatibility of an observation with the expectation from the Standard Model
1-sided confidence intervals	upper / lower bounds on the signal strength $\mu = \sigma_{\text{obs}}/\sigma_{\text{exp}}$ for a yet unobserved process
2-sided confidence intervals	parameter determination, e.g. mass or coupling strength of the Higgs boson
Exclusion or confidence regions	probing of the parameter space of a new theory with multiple parameters

Table 6.2.: Examples of hypothesis tests in high energy physics

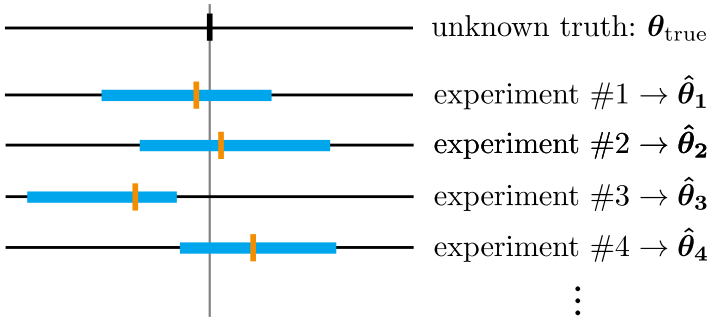


Figure 6.2.: Coverage describes the fraction of times in which the unknown true parameter value lies within the confidence intervals of regions of the individual experiments. Over-coverage describes the situation in which the coverage is larger than the confidence level, while for under-coverage it is smaller.

the confidence intervals or regions of the individual experiments. This is sketched in Figure 6.2. Note that the confidence interval obtained from experiment #3 does not include the true value. This by itself is not harmful but rather expected since confidence regions are obtained from experiments which are random processes. The question rather is whether or not for a large number of repetitions of an experiment the true parameter value is contained within the confidence intervals exactly as often as promised – with the promised value being given by the confidence level $1-\alpha$. It must be stressed that adequate coverage is a property of the chosen statistical procedure, not of the individual experiments and their outcomes. Over-coverage describes the situation in which the coverage is larger than the confidence level. Thus, the confidence regions derived in each of the individual experiments are larger than promised and necessary, which is overly conservative. Likewise, under-coverage means that the true value is contained in the individual measurements' confidence regions less often than implied by the given confidence level. Especially under-coverage is not only wrong but also dangerous as quoted results will systematically report smaller than adequate uncertainties for each measurement that was analysed with that specific statistical method.

Beside the parameters of interest $\boldsymbol{\theta}$, one or more *nuisance parameters* $\boldsymbol{\nu}$ may be introduced into the theory in order to improve its agreement with the data and reflect the limited knowledge about certain inputs to the model. These additional parameters are not of primary interest and therefore no specific values of $\boldsymbol{\nu}$ are tested and no constrained parameter regions for $\boldsymbol{\nu}$ are quoted as a result of the measurement[1]. Typical nuisance parameters in high energy physics measurements at the LHC are e.g. associated to uncertainties on energy scales and resolutions, the luminosity, the size and composition of the background or also to theoretical uncertainties.

6.2.1.1. Bayesian Approach

Due to Bayes' theorem, the *posterior probability density* $P(\boldsymbol{\theta}|\{\boldsymbol{y}_i\})$, which is the quantity that provides the basis for statements about the theory after its confrontation with the

[1] Looking at the pulls of the nuisance parameters and how they are constrained by the data may be a useful cross-check of the measurement, though.

data, can be obtained from the likelihood as

$$P(\boldsymbol{\theta}|\{\boldsymbol{y}_i\}) = \frac{P(\{\boldsymbol{y}_i\}|\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\theta})}{P(\{\boldsymbol{y}_i\})} = \frac{P(\{\boldsymbol{y}_i\}|\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\theta})}{\int P(\boldsymbol{y}_i|\boldsymbol{\theta}')\pi(\boldsymbol{\theta}')d\boldsymbol{\theta}'} \ , \tag{6.13}$$

where $\pi(\boldsymbol{\theta})$, the so-called *prior probability distribution*, quantifies the previous belief in the theory and has to be decided on before conducting the experiment. $P(\{\boldsymbol{y}_i\})$ effectively ensures the correct normalisation and using the *law of total probability*[2] can be expressed as an integral over all allowed values of the theory parameters. A common point of criticism from opponents to the Bayesian approach is the subjectivity of the prior density for whose choice there is no generally valid prescription. It is true, however, that the more overwhelming the experimental evidence is, the less important the choice of the prior density becomes.

Confidence intervals or regions are commonly called *credible intervals* or *regions* in the Bayesian approach. They can be derived directly from the posterior density by identifying those regions Θ in the parameter space for which $\int_\Theta P(\boldsymbol{\theta}|\{\boldsymbol{y}_i\})\mathrm{d}\boldsymbol{\theta} = 1 - \alpha$ for a desired level of significance $1 - \alpha$. In general, such regions are not unique and may e.g. be chosen to be symmetric around the best-fit value $\hat{\boldsymbol{\theta}}$, as small as possible or so as to contain the largest values of $P(\boldsymbol{\theta}|\{\boldsymbol{y}_i\})$.

6.2.1.2. Frequentist Approach

In the frequentist picture of statistics, probability is defined as the limit of the relative frequency of a certain outcome for an increasingly large number of repetitions of the experiment. Because of this and since the Bayesian posterior density is calculated from a single realisation of a certain experiment, the quantity $P(\boldsymbol{\theta}|\{\boldsymbol{y}_i\})$ cannot be interpreted as a probability distribution in the frequentist sense.

Therefore, the frequentist approach adopts a different method: The outcome of the measurement is condensed into a single random variable q, the so-called *test statistic*, which is usually defined in such a way that larger values of q correspond to an increasingly bad agreement of the theory with the data. Before conducting the experiment, a *critical* or *rejection interval* or *region* for q is defined which depends on the desired confidence level $1 - \alpha$. If the observed value of the test statistic, q_{obs}, which is obtained from the experimental data, happens to land in the critical region, the hypothesis under test is rejected. In order to establish if the observed value of the test statistic is inside the critical region, the *p-value*

$$p = \int_{q_{\mathrm{obs}}}^{\infty} f(q|\boldsymbol{\theta})\mathrm{d}q \tag{6.14}$$

is used. It depends on the expected distribution $f(q|\boldsymbol{\theta})$ of the test statistic as predicted by the theory corresponding to the null hypothesis and the values of the parameters of interest under test, i.e. f is a function of the random variable q for a fixed value of $\boldsymbol{\theta}$. Therefore, p denotes the probability of finding a disagreement between theory and data which is as least as bad as the one observed, under the assumption that the null hypothesis is in fact true. If $p < \alpha$, i.e. if it is sufficiently unlikely to make the given observation, the null hypothesis is rejected. In general, the form of f can only be obtained from Monte Carlo simulations, so-called *pseudo-experiments*. Under certain conditions it may, however, approach analytically known distributions which brings an enormous saving in terms of the

[2] The law of total probability states that $\sum_i P(A|B_i) \cdot P(B_i) = P(A)$ if the outcomes B_i are mutually exclusive and cover the full associated samples space.

required computing times and resources. In case of a set of null hypotheses $H_0(\boldsymbol{\theta})$, these are tested separately for each possible parameter point $\boldsymbol{\theta}$ so that eventually the desired confidence intervals or regions are obtained.

There are different options for the treatment of the nuisance parameters $\boldsymbol{\nu}$. In the strict frequentist approach a hypothesis can only be rejected if $p < \alpha$ for all allowed values of the nuisance parameters. Depending on the exact functional dependencies, this may be undesired, however. Many times there is either some previous knowledge about allowed or preferred ranges of the nuisance parameters or such knowledge is obtained from auxiliary measurements. Such additional requirements on the nuisance parameters may in fact be incorporated into the likelihood function via so-called *penalty terms*. These penalise large deviations of the nuisance parameters from their nominal or expected values by adding increasingly negative contributions to the likelihood function. A commonly used approach is to use the *profile likelihood* in which those nuisance parameter values $\hat{\boldsymbol{\nu}}(\boldsymbol{\theta})$ are used that maximise the likelihood function for a given value of $\boldsymbol{\theta}$ – this procedure is also referred to as *profiling of the nuisance parameters*. These best-fit values $\hat{\boldsymbol{\nu}}$ are then also used for the generation of pseudo-experiments. Alternatively, the nuisance parameters may also be treated in a Bayesian manner in which their distribution is modelled with a prior $\pi(\boldsymbol{\nu})$.

As described above, evaluating the observed value of the test statistic from the data involves fitting the likelihood function. This will yield the best-fit values $\hat{\boldsymbol{\nu}}$ of the nuisance parameters as well as their covariance matrix. The list of the best-fit nuisance parameter values, compared and normalised to their pre-fit estimates with means of zero and uncertainties of one, is called the *pull distribution*. It may be used to check if there is tension between the model encoded in the likelihood function and one or more of the nuisance parameters, which would be indicated by a large pull on the affected parameters, or to check if certain parameters are constrained by the fit, in which case their post-fit uncertainties would be reduced compared to the pre-fit expectations. The *pre-fit impacts* of the different nuisance parameters can be derived by fixing one single nuisance parameter at a time at one standard deviation above or below its expected nominal value and repeating the fit for all the remaining parameters. The resulting best-fit parameter of interest will then be different compared to the case where none of the nuisance parameters are fixed. The differences between the global best-fit parameter of interest values and those values for fixed individual nuisance parameter variations give the list of impacts from which it can be established which uncertainties are the dominant ones. Similarly, fixing one nuisance parameter at a time at one post-fit standard deviation above or below its best-fit value yields the distribution of the *post-fit impacts*.

6.2.2. Some Theoretical Foundations of the Frequentist Approach

The Neyman-Pearson lemma [145] asserts that for a desired significance level $1 - \alpha$ the likelihood ratio

$$\frac{P(\{\boldsymbol{y}_i\}|H_1)}{P(\{\boldsymbol{y}_i\}|H_0)} \tag{6.15}$$

is the most powerful test statistic for discriminating between two hypotheses H_0 and H_1, i.e. the test that minimises the type-2 error β. Since the logarithm is a strictly monotonic function, the Neyman-Pearson lemma also holds for log-likelihood ratios.

As will be clear from Section 6.2.5, the distribution $f(q_\mu|\mu')$ of the test statistic q_μ may be needed for different combinations of the parameter value μ that is tested for and the

assumed parameter value μ'. Since obtaining f from pseudo-experiments may be computationally expensive, it is very beneficial to have approximate analytic formulations. For $\mu = \mu'$, Wilks [146] showed in 1938 that the negative log-likelihood ratio

$$-2\ln\lambda(\mu) = -2\ln\frac{L(\mu, \hat{\boldsymbol{\nu}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} \tag{6.16}$$

approximates a chi-square distribution with one degree of freedom. A similar result also holds for the case of multiple parameters of interest and

$$-2\ln\lambda(\boldsymbol{\theta}) = -2\ln\frac{L(\boldsymbol{\theta}, \hat{\boldsymbol{\nu}}(\boldsymbol{\theta}))}{L(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\nu}}(\hat{\boldsymbol{\theta}}))} \tag{6.17}$$

follows a chi-square distribution with $n_{\boldsymbol{\theta}}$ degrees of freedom, where $n_{\boldsymbol{\theta}}$ describes the number of parameters of interest $\boldsymbol{\theta}$. These results allow for the calculation of p-values without pseudo-experiments. The approximation holds in the large sample limit, i.e. for large numbers of events. According to [147], based on findings from Wald [148] it can be shown that for $\mu \neq \mu'$ the negative log-likelihood ratio approximates a non-central chi-squared distribution with $n_{\boldsymbol{\theta}}$ degrees of freedom.[3] This result then allows for the analytic approximation of f when calculating expected sensitivities.

6.2.3. LHC-Era Test Statistics for Signal Discovery and Limit Setting

The test statistics described in the following section are based on the discussion in [147] and briefly summarised in Table 6.3 where they are classified in terms of their purpose, the associated rejection region and the sign of the signal that is tested for. Note that in contrast to the notation adopted in [147], here the symbol $\boldsymbol{\nu}$ is used for the nuisance parameters. In accordance with the presentation in [147], the symbol μ is used to denote the single parameter of interest. Inspired by the application in the search for the Higgs boson in the beginning of the LHC running or searches for so far unobserved decay modes of the Higgs boson to this day, in this section the parameter μ is interpreted as the signal strength $\mu = \sigma_{\text{obs}}/\sigma_{\text{exp}}$, which describes the ratio of the observed and the expected cross sections. The results do, however, in principle also hold for multiple parameters of interest $\boldsymbol{\mu}$ in which case the exact form of the distributions f of the test statistics will be different. Furthermore, although the aforementioned Higgs searches are not commonly referred to as *searches for new physics*, as far as the statistics methodology is concerned, there is no conceptual difference between them and searches for *Beyond the Standard Model* (BSM) physics. Therefore, in this chapter, the term *new physics* is used in a wider sense than the commonly used jargon for BSM physics.

Their common advantage compared to the test statistics used at LEP and the Tevatron is that the LHC test statistics include systematic uncertainties by using profiling of the nuisance parameters in the likelihood ratios in such a way that asymptotic closed-form formulae can be used to describe the distributions of the test statistics in the large sample limit. Specifically, this is achieved by combining profiling with the use of the best-fit parameters of interest $\hat{\boldsymbol{\mu}}$ and the corresponding best-fit values of the nuisance parameters $\hat{\boldsymbol{\nu}}(\hat{\boldsymbol{\mu}})$ in the denominator of the log-likelihood ratios, as shown previously in Equations (6.16) and (6.17).

[3] For one parameter of interest μ with $\mu \neq \mu'$ and a sample of size N, Wald [148] established that $-2\ln\lambda(\mu) = (\mu - \hat{\mu})^2/\delta^2 + \mathcal{O}(1/\sqrt{N})$, where $\hat{\mu}$ is distributed according to a Gaussian with mean μ' and variance δ^2.

Symbol	Purpose	Rejection region	Signal	Remarks
t_0	discovery	two-sided	pos. / neg.	
q_0	discovery	one-sided	pos.	
t_μ	limit setting	two-sided (confidence intervals)	pos. / neg.	
$\tilde{t}_\mu$	limit setting	two-sided (confidence intervals)	pos.	analogous to Feldman & Cousins
q_μ	limit setting	one-sided (upper limits)	pos. / neg.	
$\tilde{q}_\mu$	limit setting	one-sided (upper limits)	pos.	analogous to Feldman & Cousins

Table 6.3.: Summary of LHC-era test statistics and classification in terms of their properties as described in more detail in Section 6.2.3.

6.2.3.1. Discovery of a (Positive) Signal

In the context of searches for new physics, one is often interested in testing for the existence of a positive signal above a background. This is done by assuming the non-existence of the signal as the null hypothesis, corresponding to a value of $\mu = 0$. If the data rise sufficiently above the background to be significantly discrepant from this background-only hypothesis, the latter is rejected and the existence of the signal is established. This situation is sketched in the main plot of Figure 6.3(a) where the best-fit signal strength significantly exceeds the expectation from background-only fluctuations. The fact that only upward fluctuations are interpreted as contradicting the background-only hypothesis is suggested by the asymmetric filled blue region. Inspired by typical applications in Higgs analyses, the sketches in Figure 6.3 show mass spectra – however, the same test statistics may also be used for plain counting experiments or distributions of other observables.

This desired behaviour can be implemented with the following *capped* test statistic:

$$q_0 = \begin{cases} -2 \ln \frac{L(0, \hat{\boldsymbol{\nu}}(0))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} & \text{if } \hat{\mu} \geq 0 \\ 0 & \text{if } \hat{\mu} < 0 \ . \end{cases} \tag{6.18}$$

It is built from a likelihood ratio where in the numerator the hypothesis under test is considered by fixing μ to the value that is to be tested, i.e. $\mu = 0$, whereas the same parameter is fitted in the denominator as indicated by the notation $\hat{\mu}$. In both cases the resulting best-fit values of the nuisance parameters $\hat{\boldsymbol{\nu}}$ will be different in general. The assumption of a positive signal is specifically encoded by setting the test statistic q_0 equal to zero whenever the best-fit signal strength $\hat{\mu}$ is negative. This way, a downward fluctuation of the data (or a negative signal) will not contribute to the rejection region.

The inlay in Figure 6.3(a) shows the distribution $f(q_0|0)$ of the test statistic q_0 from statistical fluctuations in the background-only case ($\mu = 0$). As discussed before, q_0 is set to zero for downward fluctuations of the data, thus leading to the stylised spike. The region of large values of q_0, which is the rejection region, is indicated in orange and has the integral α which is related to the confidence level CL through $\alpha = 1 - \mathrm{CL}$. The p-value is defined as

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0'|0) \mathrm{d}q_0' \tag{6.19}$$

where $q_{0,\text{obs}}$ is the value of the test statistic obtained from data. For $p_0 < \alpha$, the background-only hypothesis is rejected, while it is accepted for $p_0 > \alpha$.

In order to avoid the spike at $q_0 = 0$, an *uncapped* test statistic can be defined as [149]

$$q_0^{(\text{uncapped})} = \begin{cases} -2 \ln \frac{L(0, \hat{\boldsymbol{\nu}}(0))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} & \text{if } \hat{\mu} \geq 0 \\ +2 \ln \frac{L(0, \hat{\boldsymbol{\nu}}(0))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} & \text{if } \hat{\mu} < 0 \ . \end{cases} \tag{6.20}$$

For upward fluctuations, this test statistic will yield the same results as the capped test statistic q_0. Downward fluctuations, i.e. cases with $\hat{\mu} < 0$, however, will result in negative values of the test statistic $q_0^{(\text{uncapped})}$, which do not lead to a rejection of the background-only hypothesis. Therefore, the uncapped test statistic still has a one-sided rejection region, like the capped test statistic q_0, and only upward fluctuations are considered as opposing the background-only hypothesis. However, the uncapped test statistic allows to calculate meaningful p-values also for downward fluctuations.

If not only positive signals are to be considered, the two-sided test statistic

$$t_0 = -2 \ln \frac{L(0, \hat{\boldsymbol{\nu}}(0))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} \ , \tag{6.21}$$

can be used, which is sketched in Figure 6.3(b). Note that in this case both upward and downward fluctuations of the data contribute to the rejection region since both reduce the likelihood ratio and thus yield a large value for t_0. This is hinted by the symmetric blue filled tail regions. The test statistic t_0 therefore also allows for the discovery of a negative signal, e.g. in case of a destructive interference between the SM and contributions from a new physics model. If one is only interested in positive signals, the data has to be compared to the background fit to establish the direction of a given fluctuation.

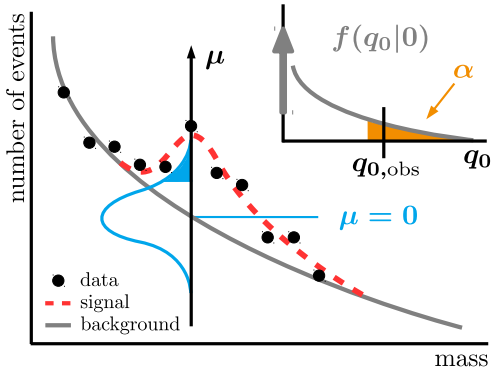
Independently of whether a one-sided or two-sided test statistic is used, p_0 is often translated into the significance Z, which gives the number of standard deviations for a standard normal distribution that a given value of p_0 corresponds to. The relation between p_0 and the significance Z is sketched in Appendix A.3. By convention, *evidence* for a new signal may only be claimed when the background-only hypothesis is rejected at the 3σ confidence level. For claiming a *discovery*, the threshold is at the 5σ confidence level.

6.2.3.2. Two-Sided Confidence Intervals

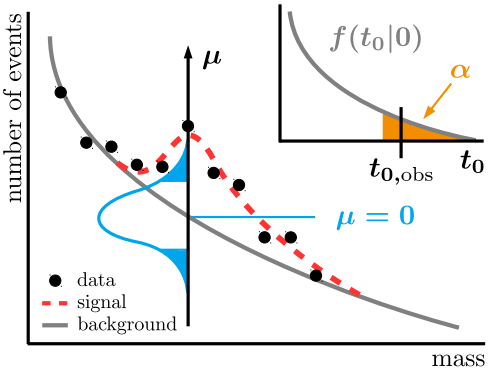
For testing a null hypothesis $H_0(\mu)$ described by some parameter μ and deriving a two-sided confidence interval for it, the following test statistic can be used:

$$t_\mu = -2 \ln \frac{L(\mu, \hat{\boldsymbol{\nu}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} \ . \tag{6.22}$$

The likelihood ratio is similar in structure to the aforementioned case of testing for the presence of a signal, but now also non-zero values of the signal strength μ can be evaluated. In case the value of μ that is tested is close to the best-fit value $\hat{\mu}$, the likelihood ratio will be close to unity, thus leading to a small value for t_μ. Should the tested μ be in disagreement with the $\hat{\mu}$ preferred by the data, the likelihood ratio will be smaller than one and t_μ will take on larger values. It is also obvious from Equation (6.22) that testing values of μ both sufficiently much larger or smaller than the value suggested by the data will result in a rejection of the null hypothesis which is why the resulting limits are two-sided, i.e. a confidence interval is determined for the parameter μ.



(a) Capped one-sided test statistic q_0.



(b) Two-sided test statistic t_0.

Figure 6.3.: Test statistics for establishing a signal. In the main plots, the expected background and signal contributions are shown with grey and red lines while the observed number of events are indicated by black points. The expected distribution of the number of events corresponding to the background-only hypothesis is indicated by the blue curve. The critical regions corresponding to the one-sided and two-sided formulations are indicated by the filled blue regions. The smaller inlays show the distributions of the respective test statistics. In the case of the capped one-sided test statistic, the setting of $q_0 = 0$ for $\hat{\mu} < 0$ leads to a δ-function as indicated by the stylised spike.

In case the parameter of interest can only take on non-negative values, the test statistic

$$\tilde{t}_\mu = -2 \ln \tilde{\lambda}(\mu) = \begin{cases} -2 \ln \frac{L(\mu, \hat{\boldsymbol{\nu}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} & \text{if } \hat{\mu} \geq 0 \\ -2 \ln \frac{L(\mu, \hat{\boldsymbol{\nu}}(\mu))}{L(0, \hat{\boldsymbol{\nu}}(0))} & \text{if } \hat{\mu} < 0 \ . \end{cases} \tag{6.23}$$

can be used. In contrast to t_μ, for downward fluctuations of the background ($\hat{\mu} < 0$) a signal strength of zero is used when evaluating the likelihood function.

6.2.3.3. Upper Limits

Upper limits can be derived with a test statistic that is a one-sided variant of the one defined in Equation (6.22):

$$q_\mu = \begin{cases} -2 \ln \frac{L(\mu, \hat{\boldsymbol{\nu}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\nu}}(\hat{\mu}))} & \text{if } \mu \geq \hat{\mu} \\ 0 & \text{if } \mu < \hat{\mu} \ . \end{cases} \tag{6.24}$$

This is typically of interest in a scenario where no significant positive signal rising above the background can be established. In such a situation, one will want to exclude large positive signal strengths μ that significantly exceed the best-fit $\hat{\mu}$ by the data as sketched with the black curve in Figure 6.4. Smaller positive hypothesised values of μ that are closer to but still larger than the signal strength preferred by data can not be excluded (blue curve) whereas signal strengths μ smaller than $\hat{\mu}$ are accepted by construction. This becomes plausible when recalling that setting upper limits corresponds to negating the question *Given the data, can the signal strength be equal to μ or larger?*. Obviously, this question has to be affirmed for $\mu < \hat{\mu}$. In summary, the upper limit on the parameter μ indicates the largest value that cannot be rejected.

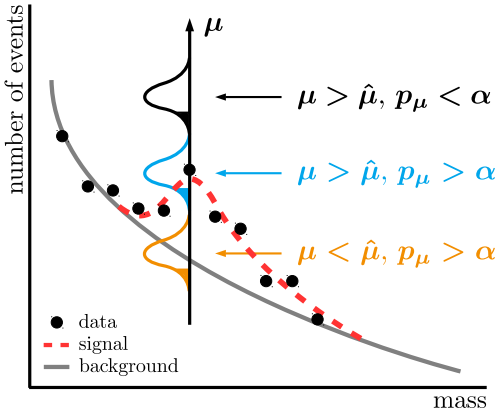


Figure 6.4.: Test statistic for the determination of upper limits. The expected background and signal contributions are shown with grey and red lines while the observed number of events are indicated by black points. The expected distributions of the number of events for different hypothesised signal strengths μ are indicated by the orange, blue and black curves. The critical regions are sketched with filled areas of the same colours. By construction, for one-sided upper limits only too large hypothesised signal strength μ are rejected (black curve), while hypothesised lower than observed signal strengths are not (orange curve).

In case the parameter of interest μ is assumed to be non-negative, the test statistic

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda} & \text{if } \mu \geq \hat{\mu} \\ 0 & \text{if } \mu < \hat{\mu} \end{cases} = \begin{cases} \begin{cases} -2\ln\frac{L(\mu,\hat{\boldsymbol{\nu}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\nu}}(\hat{\mu}))} & \text{if } \hat{\mu} \geq 0 \\ -2\ln\frac{L(\mu,\hat{\boldsymbol{\nu}}(\mu))}{L(0,\hat{\boldsymbol{\nu}}(0))} & \text{if } \hat{\mu} < 0 \end{cases} & \text{if } \mu \geq \hat{\mu} \\ 0 & \text{if } \mu < \hat{\mu} \end{cases} \tag{6.25}$$

may be used.

6.2.4. Modifications of the Classical Frequentist Method

When setting upper limits on a parameter μ using a classical likelihood ratio similar to the test statistic q_μ, the following situation may happen: If the number of expected signal events is small, then just from statistical fluctuations occasionally the observed best-fit value $\hat{\mu}$ will be negative. As a consequence, even a hypothesised signal strength of $\mu = 0$ may be excluded.[4] In the context of searches for which the parameter μ can only take on non-negative values this behaviour is undesired and therefore, extensions of the classical limit setting procedures were developed which will be briefly introduced in the following subsection.[5]

6.2.4.1. The Method of Feldman and Cousins

The key idea of the method suggested by Feldman and Cousins [150] is to restrict the fitted signal strength to be non-negative, i.e. $\hat{\mu} \geq 0$, and to not evaluate the likelihood

[4] Note that the test statistics $\tilde{t}_\mu$ and $\tilde{q}_\mu$ are defined in such a way that they do not suffer from this shortcoming. This becomes clear from Equations 6.23 and 6.25 where it can be seen that both test statistics yield zero for $\mu = 0$ and $\hat{\mu} < 0$.

[5] Note that the description of the various test statistics and hypothesis testing procedures in this chapter does not proceed chronologically. The development of the extensions described in this section historically occurred before the unified treatment as outlined in Section 6.2.3 was formulated.

ratio for negative signal strengths. This prevents the likelihood ratio from increasing and thus protects against a false rejection of a small hypothesised signal strength in the case of a downward fluctuation of the background. This approach is in fact used by the test statistics $\tilde{t}_\mu$ and $\tilde{q}_\mu$ described in Section 6.2.3.

6.2.4.2. The CL_s Method

As described before, in the classical frequentist procedures a hypothesis $H_0(\mu)$ is rejected if the associated observed p_μ-value is smaller than α for a desired confidence level $1-\alpha$. In contrast to this, in the CL_s method, which was introduced in [151], a hypothesis $H_0(\mu)$ is rejected if

$$\mathrm{CL_s} = \frac{\mathrm{CL_{s+b}}}{\mathrm{CL_b}} < \alpha \tag{6.26}$$

where

$$\mathrm{CL_{s+b}} = p_\mu = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q'_\mu|\mu)\mathrm{d}q'_\mu \qquad \text{and} \qquad \mathrm{CL_b} = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q'_\mu|0)\mathrm{d}q'_\mu \ . \tag{6.27}$$

In particular, when testing $H_0(\mu = 0)$, $\mathrm{CL_{s+b}} = \mathrm{CL_b}$ and therefore $\mathrm{CL_s} = 1$. Thus, a signal strength of $\mu = 0$ will never be excluded, not even in the presence of a downward fluctuation of the background. From this, it is obvious that the CL_s method is only suited for deriving upper limits on a non-negative parameter.

On the other hand, dividing the p_μ-value by $\mathrm{CL_b}$ and therefore making rejection of the null hypothesis less likely, introduces an over-coverage which means that e.g. 95% CL exclusion limits do not correspond to parameter regions that are excluded with 95% confidence. Rather, the exclusion limits will in general be conservative, i.e. they may correspond to regions obtained at a higher confidence level with the classical methods. In particular, when testing for $\mu = 0$ while the observed signal strength is $\hat{\mu} = 0$, then $\mathrm{CL_{s+b}} = p_{\mu=0} = 0.5$ corresponding to an over-coverage by a factor of two.

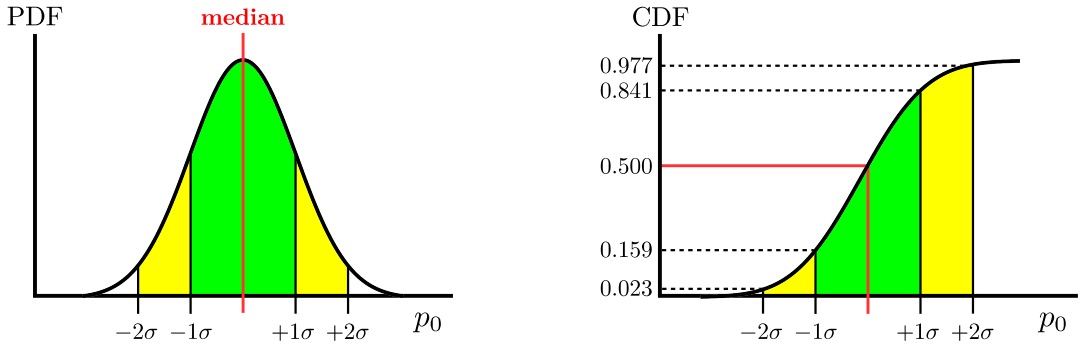
6.2.5. Quantifying the Sensitivity of an Experiment

Before actually looking at experimental data, it is interesting and instructive to assess the expected sensitivity of the planned measurement. The following will describe the procedure for testing a new physics model $H_1(\mu)$ characterised by the signal strength μ against the background-only assumption H_0 with $\mu = 0$. Example applications are the search for the Higgs boson before its discovery as well as searches for yet unobserved decay modes of the Higgs boson. The following description is based on the presentation in [152] and therefore corresponds to procedures adopted in the context of Higgs searches in ATLAS.

6.2.5.1. Expected Sensitivity for Discovery

For some non-zero benchmark signal strength μ, which in the context of the Higgs boson is $\mu = 1$, the distribution $f(q_0|\mu)$ of the appropriate test statistic q_0 (or t_0) can be obtained either from asymptotic formulae or from a set of k pseudo-experiments.[6] For each pseudo-experiment out of an additional set of m pseudo-datasets, the 'observed' value of the test statistic q_0^{obs} and the corresponding value of p_0 are calculated. From the resulting distribution of p_0, the median value as well as the surrounding 68% and 95% intervals can be obtained as sketched in Figure 6.5. From these, the expected median significance and its statistical uncertainty can be quantified.

[6] Even in the former case, it is usually a good idea to verify the known analytic relation with pseudo-experiments to check that the large sample limit is in fact applicable.



(a) PDF of the expected p_0-value.

(b) CDF of the expected p_0-value.

Figure 6.5.: Probability and cumulative distribution functions of the expected p_0-value for testing the discovery significance for a specific signal strength μ. From these distributions, the median expected p_0-value and the corresponding significance as well as the 68% and 95% uncertainty intervals can be determined.

6.2.5.2. Expected Sensitivity for Upper Limits

For the background-only assumption, a number of $k + m$ pseudo-datasets is generated. Depending on the chosen method and test statistic q, the distribution of the test statistic $f(q|0)$ may either be obtained from analytic formulae or from the first k pseudo-datasets. For each of the next m toy datasets, the 'observed' value of the test statistic q_{obs} is calculated. Subsequently, depending on the chosen method, the value of p_μ or the $\text{CL}_\text{s}(\mu)$ are calculated for a range of signal strengths μ in order to find that value $\mu_{95\%}$ which marks the transition between the acceptance and the rejection regions at the 95% confidence level. These values are entered into a histogram to obtain the expected distribution of the limit on the signal strength and from it the median expected limit and the 68% and 95% percentiles that characterise the statistical uncertainty on the expected limit.

6.2.5.3. The Asimov Dataset

Since the generation of pseudo-datasets can be computationally demanding, it is often beneficial to use approximate asymptotic expressions for the distribution of the test statistics, as described and given in [147]. As mentioned in Section 6.2.3, these approximate closed-form formulae hold in the large sample limit. In general, for their construction, the variance σ^2 of the best-fit signal strength $\hat{\mu}$ is needed, which can be determined from the *Asimov* dataset. The Asimov dataset was introduced in [147]. According to its abstract definition, the parameter values obtained from fitting a model to the corresponding Asimov dataset yield the true values of the parameters. A practical realisation is a dataset in the form of a binned histogram without statistical fluctuations, i.e. where the event numbers in all bins are exactly equal to the expected event numbers for a given signal strength and the given nuisance parameters. As described in [147], $\sigma^2 = (\mu - \mu')^2/q_A$, where μ describes the signal strength of the hypothesis that is tested, μ' the signal strength assumed in the data and the Asimov dataset and q_A denotes the relevant test statistic evaluated using the Asimov dataset. Thus, when estimating the expected significance of a search, $\mu = 0$, as required by the test statistic q_0 or t_0, and the benchmark signal strength μ' is commonly taken to be 1. When evaluating expected upper limits, different values of μ will be tested, whereas the benchmark signal strength μ' is kept fixed at 0 (1) if the upper limit is meant

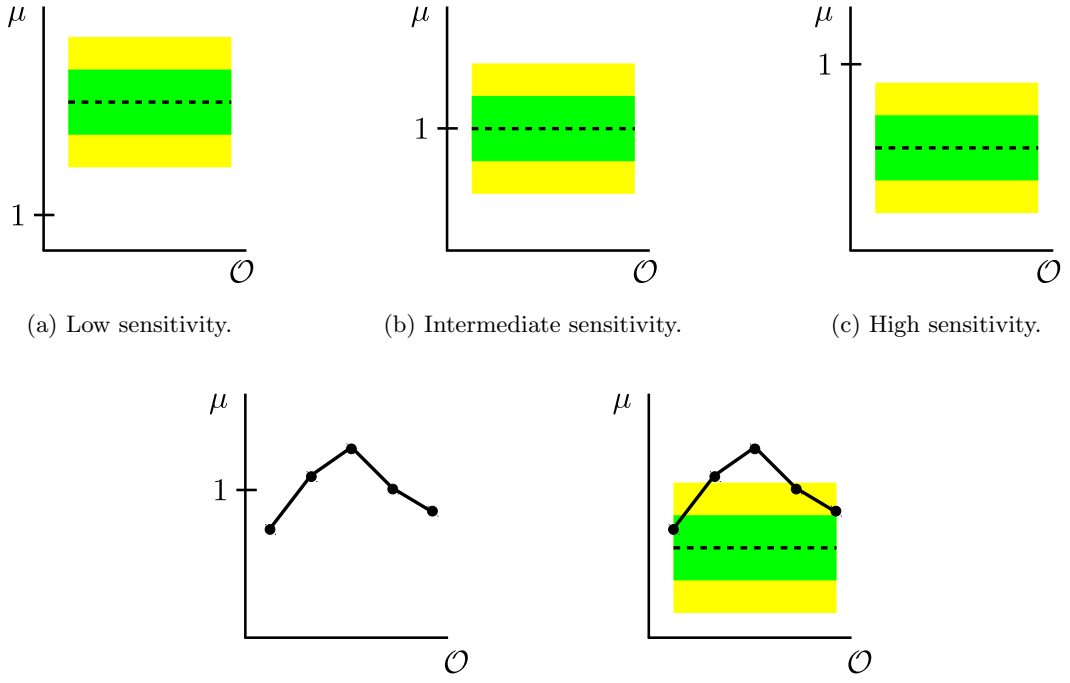
to be valid in the absence (presence) of the signal process. Evaluating σ from the Asimov dataset is relevant when calculating expected significances and limits, whereas - depending on the test statistic - in many cases, observed results based on analytic approximations of the distribution of the test statistic can be calculated without using the Asimov dataset.

6.2.5.4. Interpretation of Limit Plots

In the context of a search such as the one for the Higgs boson where the new physics model corresponding to the non-background-only hypothesis already contains a prediction of the coupling strength of e.g. $\mu = 1$, the regimes depicted in Figure 6.6(a)-(c) are thinkable: In the beginning when the available dataset is still rather small, a possible signal cannot be distinguished from fluctuations of the background. Therefore, the expected sensitivity is low, which is reflected in the fact that the expected limits are larger than the predicted coupling strength of $\mu = 1$. So unless the real coupling strength is in fact much larger than predicted by the theory, there is no hope of rejecting the background-only hypothesis. As the collected dataset grows and statistical fluctuations become smaller compared to the recorded number of events, a phase of intermediate sensitivity is passed and finally, the regime of good sensitivity is reached where the odds are good that the hypothesised signal strength of $\mu = 1$ can be excluded if the corresponding theory is in fact wrong. This does not yet mean, however, that one can hope to already make a discovery since the required confidence level for a discovery by convention is 5σ and therefore much higher than the 2σ significance level usually tested for in such limit or exclusion plots.

After conducting the experiment, the observed limits are derived from the real experimental data. With an increasing dataset size the formerly large upper limits will decrease and eventually even drop below the predicted signal strength of $\mu = 1$. Then the signal hypothesis is rejected at the 95% confidence level. If the theory depends on an additional parameter $\mathcal{O}$, such as e.g. the mass of the Higgs boson m_H, this may happen only for some values of the parameter while the signal hypothesis cannot be rejected for parameter values in the vicinity of the true parameter value. This situation is sketched in Figure 6.6(d). If the signal hypothesis is in fact true, then the observed limits will eventually rise above the uncertainty bands that surround the expected limits in the background-only case as depicted in Figure 6.6(e). This indicates a growing tension between the observation and the background-only hypothesis which has to reach significance levels of 3σ and 5σ to warrant a claim of evidence and discovery, respectively.

All of the aforementioned aspects are combined in the so-called *Brazil plots*. As an example, in Figure 6.7 the upper limits at the 95% confidence level on the signal strength $\mu = \sigma/\sigma_{\mathrm{SM}}$ for the combined search for the Higgs boson in various decay channels based on collisions at 7 TeV collected during 2011 are shown as published in [153]. Based on this data, the Higgs boson hypothesis could be ruled out for the mass ranges 111.4-116.6 GeV, 119.4-122.1 GeV and 129.2-541 GeV for which observed limits smaller than 1 were found. At the same time, tension between the observed data and the background-only hypothesis was solidifying around a mass of 126 GeV. The corresponding local and global significances of 2.9σ and 1σ, respectively, were not yet large enough, however, to claim a discovery of the Higgs boson. If in reality there were no Higgs boson, an exclusion in the mass range of 120-560 GeV would have been expected as indicated by a decrease of the expected limits below 1 for those masses.



(a) Low sensitivity.

(b) Intermediate sensitivity.

(c) High sensitivity.

(d) Excluding some observable values.

(e) Tension with the background-only hypothesis.

Figure 6.6.: Elements of the so-called *Brazil plots* that usually show both the expected and observed limits on some parameter of interest, for which in this example the signal strength μ is used. The expected limits are shown as dashed lines surrounded by the 1σ and 2σ uncertainty bands in green and yellow, respectively. The observed limits are represented by solid lines. For the sake of the example, a generic observable $\mathcal{O}$ is used. In the context of searches for Higgs signals the limits are commonly derived for a number of hypothesised values of the Higgs mass as indicated by the black dots in the subfigures (d) and (e).

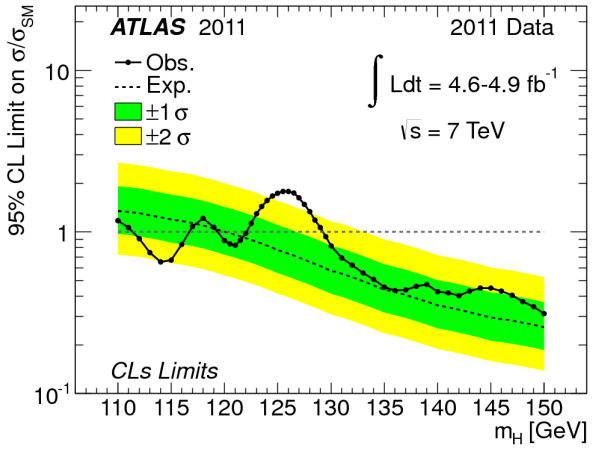


Figure 6.7.: Upper limits at the 95% confidence level on the signal strength $\mu = \sigma/\sigma_{\mathrm{SM}}$ for the combined search for the Higgs boson in various decay channels based on collisions at 7 TeV collected during 2011, published as Figure 4b in [153].

6.2.6. The Look-Elsewhere Effect

When evaluating the null hypothesis multiple times, e.g. when scanning a mass range with multiple bins, occasionally there will be a seemingly significant discrepancy between the data and the null hypothesis even if the null hypothesis actually is true. This simply is a consequence of the confidence level α of the hypothesis test and the presence of statistical fluctuations in the data. Specifically, e.g. for a 95% confidence level, i.e. $\alpha = 0.05$, a significant disagreement between the valid null hypothesis and the data will on average appear in 5% of the cases. The larger the number of observables, distributions or bins, the more likely even excesses or discrepancies purely from statistical fluctuations with significances at the 3σ- or 4σ-levels will appear. This phenomenon is known as *look-elsewhere effect*, *trials factor* or *multiple comparisons problem*. In order to compensate for it and correct the overly large local significance usually also a reduced global significance is quoted alongside any potentially large local significance seen in searches.

One way of estimating the global significance is by means of the Dunn-Šidák correction [154,155]. It is based on the assumption that all of the n different local measurements are uncorrelated. Recalling that α denotes the probability of a type-1 error in a single test, the probability of making at least one type-1 error in a range of tests is given by

$$\begin{aligned} \alpha_{\text{global}} &= P\,(\text{at least one type-1 error}) \\ &= 1 - P\,(\text{no type-1 error anywhere}) \\ &= 1 - [P\,(\text{no type-1 error in a single test})]^n = 1 - (1-\alpha)^n \ . \end{aligned} \tag{6.28}$$

From this it can be seen that the probability of wrongly rejecting the null hypothesis due to one or more fluctuations somewhere in the tested range grows with number of tests performed, e.g. corresponding to an increasing number of mass bins in a search for a new particle.

In experimental analyses, the global significance is commonly evaluated using a different approach [156] which is based on toy experiments. While in principle random experiments can be used to evaluate the global significance of a local excess, this procedure may be very inefficient if implemented naively. This is due to the fact that one is interested in the very rare cases of large background fluctuations faking a signal. A huge amount of pseudo-experiments has to be generated in order to obtain just a few of these large fluctuations. Therefore, methods were developed to make such simulations more efficient. A core idea is to first evaluate how often a lower reference significance is exceeded and to then estimate the probability of exceeding the real, desired significance level by using the known closed-form approximate distribution of the test statistic in the large sample limit. This has the advantage that evaluating the frequency of excesses for a lower reference significance can be done with a greatly reduced number of random experiments.

Let $q(m)$ be the test statistic for a specific mass point m and q_{obs} the value of the test statistic at the largest observed deviation from the background-only expectation in the search range. The probability of seeing a local discrepancy at least as large as the observed one is written as $P(q(m) \geq q_{\text{obs}})$. According to [156,157], the global probability $P(q(\hat{m}) \geq q_{\text{obs}})$ of observing an excess at least as large somewhere in the full search range is given by

$$P(q(\hat{m}) \geq q_{\text{obs}}) = P(q(m) \geq q_{\text{obs}}) + \langle N(q_{\text{obs}}) \rangle \ , \tag{6.29}$$

where $\langle N(q_{\text{obs}}) \rangle$ is the expectation value of the number of upward crossings beyond the level q_{obs}. The χ^2-behaviour of a suitable test statistic q can be used to recast this in a

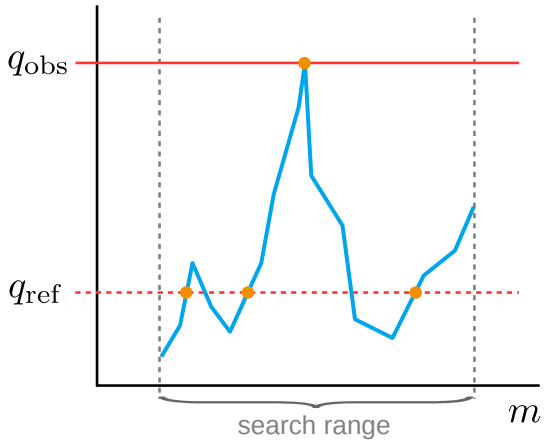


Figure 6.8.: Methods for improving the efficiency of random experiments for the estimation of the global significance of a local excess can be based on estimating the number of excesses at a lower reference level and then exploiting the analytically known asymptotic behaviour of the used test statistic to extrapolate this to the desired higher level of significance.

way that evaluates the number of upcrossings at a lower level q_{ref}:

$$\langle N(q_{\mathrm{obs}})\rangle = \left(\frac{q_{\mathrm{obs}}}{q_{\mathrm{ref}}}\right)^{\frac{s-1}{2}} \cdot \exp\left[-\frac{q_{\mathrm{obs}} - q_{\mathrm{ref}}}{2}\right] \cdot \langle N(q_{\mathrm{ref}})\rangle \ , \tag{6.30}$$

where s denotes the number of parameters of interest. As mentioned before, this can be evaluated much more efficiently as fluctuations of a reduced level by definition occur much more frequently.

Part II.

Event Generation with Herwig in ATLAS and Studies of the $t\bar{t}c\bar{c}$ Background to $t\bar{t}H(b\bar{b})$

Chapter 7

Event Generation with Herwig7 in ATLAS

The detailed simulation of particle collisions at the LHC is an immensely complex task. Many of the relevant ideas and approaches were sketched in Chapter 3. Over the last decades, these physics principles were implemented in an ever growing number of software programs and libraries, some of which are quite specialised while others are designed to cover a wider range of tasks. A few programs encompass, to some degree, all of the steps listed in the beginning of Chapter 3. These general-purpose Monte Carlo event generators are heavily used and relied on for many physics studies and analyses at the LHC. Currently, the most widely used tools are `Herwig7` [87], `Pythia8` [86], and `Sherpa` [97]. The first one of these three, `Herwig7`, will be briefly introduced in this chapter.

While in principle all of these tools can be installed and run as published and foreseen by their respective developers, the enormous scope and scale of Monte Carlo production in an LHC collaboration like ATLAS create additional requirements. Based on these, the ATLAS collaboration created a complex and integrated software framework for event generation to which the different event generators are interfaced. One of the main focuses of these efforts is the reproducibility of the results. This comprises aspects such as dealing with software versions, generator configurations and the balance of centralisation against flexibility that still ensures that users are provided with standardised and validated environments while still having enough flexibility and a reasonable amount of overhead when making local changes. Other aspects include the automation of the sample production, the distributed nature of the computing infrastructure and the associated orchestration and the bookkeeping of large arrays of event samples. Finally, Monte Carlo event generation is only the first of a number of different steps and is followed by the detector simulation and other tasks that contribute to the need for an integrated system. Therefore, in the second part of this chapter, the interface of the ATLAS software framework to the `Herwig7` event generator will be outlined.

7.1. Overview of Herwig7 and Matchbox

As mentioned before, Herwig7 is one of the main general-purpose Monte Carlo event generators used by the ATLAS and CMS collaborations at the LHC. As such, it strives to provide a full description of the proton collisions, from the incoming protons to the outgoing particles entering the detector. Therefore, it includes algorithms and models for all of the steps outlined in the beginning of Chapter 3. Herwig7 is the successor of the HERWIG [95] and the Herwig++ [96] event generators, written in Fortran and C++, respectively. One key difference between the various general-purpose event generators is the parton shower algorithm. As stated in Section 3.3, different choices are possible for the ordering variable of a parton shower. While some event generators have parton showers ordered in transverse momentum or virtuality of the emissions, the Herwig family is traditionally based on an angular ordering. This means that the first emissions generated by the shower are at a wide opening angle and subsequently narrow down to be more and more collinear to the emitting particle as the shower evolves to lower scales.

Before the release of Herwig7, the Herwig++ event generator was focused more on the parton shower and the subsequent stages of the event generation, but only had at its disposal a few hard process matrix elements each of which needed to be coded individually and separately. The more common way of generating events was to run a dedicated hard process generator, such as aMC@NLO [80, 81] or PowhegBox [83–85], and pass the resulting events on to Herwig++ by means of *Les Houches Event* (LHE) files [158]. Herwig7 goes beyond this approach and puts much more emphasis on the hard process. This is motivated by the increasing move to NLO as the default order for predictions for LHC physics processes. The combination of matrix elements and partons showers at NLO via matching and merging, as introduced in Section 3.4, becomes easier and more flexible once all relevant ingredients are available in a single tool.

The Matchbox module, that is part of Herwig7, has a few built-in matrix elements, but also contains interfaces to a number of external tools that provide MEs for virtually the entire range of technically accessible SM processes. These external tools are hereafter referred to as *matrix element providers*. Some of them, such as VBFNLO [159, 160], contain manually implemented processes while others, such as MadGraph5_aMC@NLO [82, 161] and GoSam [162], are able to dynamically generate, compile and run the code for the requested physics process at runtime. This way of interfacing the parton shower MCs and the hard process generators is distinctly different from the LHE file approach in that the parton shower Monte Carlo event generator can separately evaluate the different matrix element contributions, i.e. the Born level as well as real-emission and virtual corrections. As a consequence, for a given process Herwig7 can use a combination of MEs from different providers, e.g. tree-level amplitudes from MadGraph5_aMC@NLO and virtual corrections from OpenLoops [163].

Besides the traditional angular-ordered parton shower based on $1 \to 2$ splittings, Herwig7 also features a dipole-type parton shower [164]. The latter uses $2 \to 3$ splittings, i.e. apart from the parton undergoing the splitting, a *spectator parton* is considered, which absorbs the momentum recoil that is necessary in order to keep all participating partons on their mass shells after each splitting. The dipole shower is ordered in transverse momentum. Having the required amplitudes and matrix elements at its disposal, Herwig7 can perform a *subtractive* or a *multiplicative* matching to its angular-ordered and dipole parton showers. The subtractive and multiplicative matching algorithms are Herwig7-specific implementations of the MC@NLO and Powheg method sketched in Section 3.4. NLO multi-jet merging capabilities, released with Herwig7 version 7.1 [165], are based on the same footing.

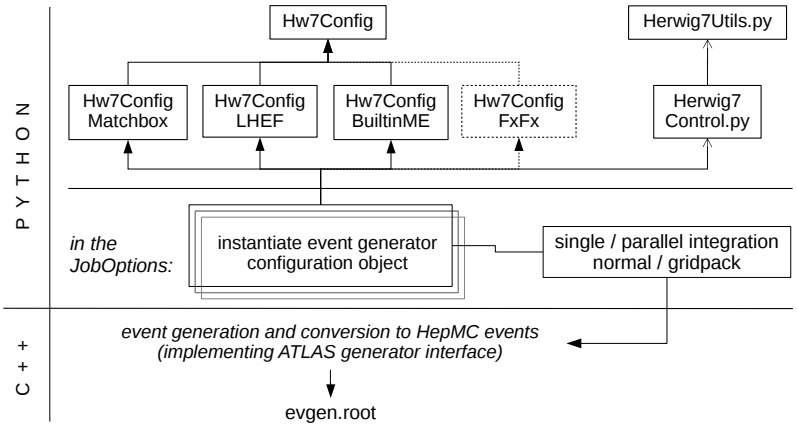


Figure 7.1.: Schematic of the Herwig7 interface to ATLAS' Athena software framework. The instantiation and configuration of the event generator is steered via the Python module Herwig7Control.py, which makes use of the different configuration classes Hw7ConfigMatchbox, Hw7ConfigLHEF, Hw7ConfigBuiltinME and Hw7ConfigFxFx as well as helper functions provided in Herwig7Utils.py. After the configuration and a possible phase space sampling and integration stage, the event generation itself is performed by the C++ part of the interface which implements the standard Athena event generator interface functionality.

While not being fully exhaustive, this list of new features of Herwig7 compared to Herwig++ is meant to motivate the necessity of substantial changes to the interface for this event generator to the ATLAS software framework, which will be presented in the following section.

7.2. The Interface for Herwig7 in Athena

As motivated in the previous section, the vastly expanded functionality of Herwig7 compared to Herwig++ necessitated substantial modifications of the Herwig7 interface in Athena. From the previous Herwig++ interface, the basic design with the generator configuration part written in Python and the event generation part written in C++ was adopted. The new interface design is sketched in Figure 7.1. The event generation part of the interface implements Athena's generator interfaces and remained largely unchanged. In the configuration stage, which is performed before the actual event generation, the generator settings and parameters are adjusted according to the details specified in the input file, e.g. the hard process that is to be generated, particle masses, the PDF set as well as jet and lepton cuts. This configuration stage, however, needed to be rewritten completely, which was required to accommodate Herwig7's improved hard process capabilities. Herwig++ was mostly used to shower LHE files containing hard process events generated with MadGraph5_aMC@NLO or PowhegBox. More rarely, Herwig++'s internal matrix elements were used. In contrast, Herwig7's Matchbox module in principle allows for the generation of arbitrary SM processes at LO and NLO. This degree of flexibility and automation requires a dedicated *warm-up stage*, during which the phase space of the hard process is sampled, the sampling strategy is optimised and the cross sections of the contributing subprocesses are estimated, where *subprocess* refers to a matrix element for a specific set of incoming

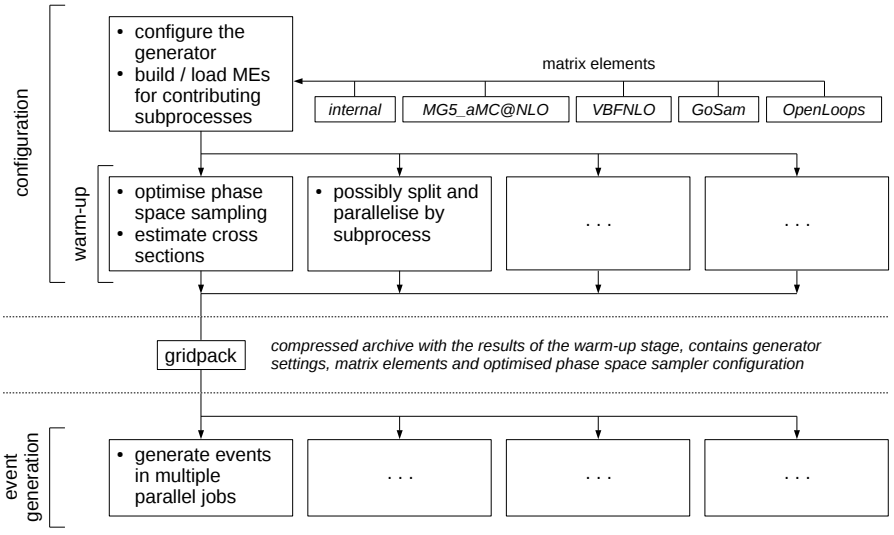


Figure 7.2.: Schematic of the warm-up and event generation stages for Herwig7 when using the Matchbox module. At the beginning of the warm-up stage, the generator is configured and the relevant matrix elements are loaded or built. The MEs may be provided internally by Matchbox itself or by a range of interfaced matrix element providers such as e.g. VBFNLO and OpenLoops. Later, opportunities for parallelisation exist during the optimisation of the phase space sampling as well as during the event generation. The results of the warm-up stage may be compressed into a gridpack, which is provided to each event generation job. This way, the warm-up stage only needs to be done once before the event generation.

and outgoing fundamental particles. This warm-up stage is performed after the configuration of the event generator and essentially puts to use variants of the methods for variance reduction sketched in Section 3.1. One of the requirements for the new Herwig7 interface was the ability to steer this warm-up stage. Besides technical details of the phase space sampling and optimisation, one important aspect in this context is parallelisation. While for sufficiently simple processes calculated at LO it may be sufficient to perform the sampling and integration step using a single CPU core, this is in general not practically feasible for NLO processes. Therefore, Herwig7 allows for the splitting of the warm-up stage according to the different contributing subprocesses. This functionality was reflected in the interface and can be used to specify the desired degree of parallelisation as well as full control over the parallel subprocesses. In particular, different subprocesses can be run independently of each other on different host machines, which enables the full exploitation of computing farms and batch clusters and allows for a level of parallelisation much beyond the number of CPU cores of a single host machine. This is schematically depicted in Figure 7.2.

Related to this is the preparation of so-called *gridpacks*. These can be created when using the hard process generation with Matchbox and they consist of compressed archives containing the relevant matrix elements for the selected hard process as well as the files with the results of the warm-up stage, i.e. the generator settings and the optimised configuration of the phase space sampler. In principle, the warm-up stage can be performed separately by each generation job that is run to create a subset of a larger sample. This is inefficient,

however, as each event generation job would perform the same tasks during the warm-up stage. Moreover, especially for increasingly complex processes, the warm-up stage may be quite demanding in terms of computing time. Therefore, it often makes sense to perform the warm-up stage only once for a single sample, run an extensive phase space sampling and integration phase and then package the resulting generator configuration into a gridpack which is used by all of the different jobs that actually generate the events for a sample. The preparation and usage of such gridpacks was implemented in the new Herwig7 interface.

Another major conceptual change in the interface was the introduction of different Herwig7 *run modes*, which describe and encode what functionality of Herwig7 is used: event generation where the hard process is calculated with the Matchbox module, showering of LHE files or event generation based on the built-in matrix elements inherited from Herwig++. The most recent addition to this list of run modes is the showering of LHE files prepared with MadGraph5_aMC@NLO using FxFx multi-jet merging [88]. As sketched in Figure 7.1, each of these four run modes is implemented in a separate Python class inheriting from the base class Hw7Config. This design was chosen in order to create some level of isolation between the different run modes and only expose the settings that are relevant in each case, while at the same time having a compact and flexible implementation with the least amount of code duplication between the different run modes. It is also straightforward to add new run modes in the future, building on the functionality provided by the base class and using the helper functions implemented in Herwig7Utils.py.

As described in Section 7.1, the Matchbox module can use amplitudes and matrix elements from a range of other tools and programs. Consequently, these other programs have to be installed and properly configured for the use with Herwig7. The number of dependencies and combinations of Herwig7 and the different programs make it extremely desirable to have available an automated testing procedure. Therefore, as part of the new Herwig7 interface, a test suite was written that is capable of running an extensive array of configured run-time tests. The tests are designed to cover all the different run modes, matrix element provider combinations and all aspects of the event generation, ranging from the warm-up stage and the preparation of gridpacks to ensuring the presence of adequate output files after the event generation. All tests or a subset of them can be run at the push of a button and upon completion, a summary is created, which gives an overview of the test results and also includes log files which can be used to track down errors. An example test summary is provided in Figure 7.3, showing the set of tests designed for Herwig7's Matchbox module. Different tests were designed for the other run modes, e.g. the showering of LHE files, covering the relevant aspects of each case. This suite of run-time tests proved of immense value when preparing new installations of newly released versions of Herwig7 and when making changes to the generator interface.

With this rewritten interface, the groundwork was laid to use the full functionality of the Herwig7 event generator in ATLAS. The availability of an extensive test suite is hoped to allow for efficient installation and testing cycles, preceding the physics validation campaigns for new releases of the Herwig7 generator, thereby contributing to a prompt adoption of bugfixes, improvements and new physics capabilities.

Test Group/Test case	Count	Pass	Fail	Error	View
tests.gridpack.Internal.TestInternal	22	22	0	0	Detail
test_read_completes			pass		
test_read_runfile			pass		
test_gz_compress			pass		
test_run_completes			pass		
test_run_outfile			pass		
test_run_logfile			pass		
test_run_hepmcfile			pass		
test_relocate_completes			pass		
test_relocate_outfile			pass		
test_relocate_logfile			pass		
test_relocate_hepmcfile			pass		
test_relocate_outfile_identical_to_run			pass		
test_relocate_logfile_identical_to_run			pass		
test_relocate_hepmcfile_identical_to_run			pass		
test_gridpack_targz_extract			pass		
test_gridpack_targz_completes			pass		
test_gridpack_targz_outfile			pass		
test_gridpack_targz_logfile			pass		
test_gridpack_targz_hepmcfile			pass		
test_run_gridpack_targz_outfile_identical_to_run			pass		
test_run_gridpack_targz_logfile_identical_to_run			pass		
test_run_gridpack_targz_hepmcfile_identical_to_run			pass		
tests.gridpack.MadGraph.TestMadGraph	22	22	0	0	Detail
tests.gridpack.OpenLoops.TestOpenLoops	22	22	0	0	Detail
tests.gridpack.GoSam.TestGoSam	22	22	0	0	Detail
tests.gridpack.VBFNLO.TestVBFNLO	22	22	0	0	Detail
tests.gridpack.NJet.TestNJet	22	22	0	0	Detail
tests.gridpack.HJets.TestHJets	22	22	0	0	Detail
tests.gridpack.MadGraphOpenLoops.TestMadGraphOpenLoops	22	22	0	0	Detail
tests.gridpack.MadGraphGoSam.TestMadGraphGoSam	22	22	0	0	Detail
tests.gridpack.MadGraphNJet.TestMadGraphNJet	22	22	0	0	Detail
Total	**220**	**220**	**0**	**0**	

Figure 7.3.: Example summary of a partial run of the run-time test suite for the Herwig7 interface in Athena. The Matchbox module is tested with internal matrix elements as well as processes provided by the different matrix element providers and possible combinations of them. For each matrix element configuration, a number of tests is performed as indicated by the corresponding *count*. A test is *passed* if it finishes and the test condition is fulfilled, whereas a test *fails* if it finishes but the test condition is not met. An *error* indicates that the test did not finish. Thus, failures and errors indicate problems with the relevant software installations and configurations. In the particular case of this example, all tests were passed successfully. The tests are designed to cover various aspects of the warm-up and event generation stages, such as the successful termination of the different stages, the creation of all relevant files and the consistency of the results when reusing existing warm-up runs and gridpacks.

Chapter 8

Studies on the Production of $t\bar{t}c\bar{c}$

The top quark is the heaviest particle of the Standard Model, which leads to two special properties: Firstly, the top quark has a very short lifetime and decays before it forms hadrons. It decays almost exclusively to a W boson and a b-quark [118]. Secondly, being the heaviest particle, it has the largest Yukawa coupling to the Higgs boson. This is a major reason for the importance of measuring the Higgs production process in association with a top quark pair, $t\bar{t}H$, which gives a direct tree-level access to this coupling. At the LHC, the top quark is predominantly produced in pairs. Besides the presence of two b-quarks, the $t\bar{t}$ decay topology is determined by decays of the two W bosons. A single W boson can either decay leptonically to a charged lepton and the corresponding neutrino or hadronically to a quark pair. As a consequence, $t\bar{t}$ decays are usually classified into three categories: The *dilepton* decay channel, where both W bosons decay leptonically, the *semi-leptonic* decay channel, where one W boson decays leptonically and the other hadronically, and the *fully hadronic* decay channel.

Top quark pair production in association with two charm quarks, which is investigated in the following study, is an interesting and important process. From an experimental viewpoint it is particularly interesting because it constitutes a sizeable background contribution for analyses targeting the Higgs decay to a pair of b-quarks in $t\bar{t}H$ production. This is due to the non-negligible probability for wrongly identifying a c-jet as a b-jet.[1] Together with the larger prevalence of c-jets compared to b-jets, this leads to a sizeable background from charm jets in $t\bar{t}H(b\bar{b})$. At the same time, it is theoretically challenging because it is a process involving two mass scales which are different by approximately two orders of magnitude. Different strategies may be adopted for the treatment of the charm quark pair production as well as the c-quark mass. Previously, the approach taken in ATLAS was to use samples of inclusive top quark pair production and rely on the parton shower for the production of c-quark pairs. Other options are to increase the multiplicity of the hard process matrix elements and add two jets - either explicitly at the Born level or via multi-jet merging. Furthermore, the charm quarks may be treated either as massless

[1] A b-jet is defined as a jet containing one or more b-hadrons, while a c-jet is defined as a jet containing one or more c-hadrons but no b-hadron. More details can be found in [166].

or as massive in the hard process. While a number of discussions and phenomenological studies are available for the treatment of the b-quark mass [167] and for the process $t\bar{t}b\bar{b}$, e.g. [168–174], no corresponding investigations were available for $t\bar{t}c\bar{c}$.

The study described in this chapter is a first look at $t\bar{t}c\bar{c}$ production and was presented in [175]. The `MadGraph5_aMC@NLO` and `Herwig++` event generators were used. The charm quark pair is generated in the hard process matrix element and the charm quarks are taken to be massive. Different scale choices are implemented and the generated samples are compared against a `MadGraph5_aMC@NLO`/`Herwig++` sample of inclusive NLO top quark pair production, where charm quarks are produced in the parton shower, and a `Sherpa` sample of multi-jet merged $t\bar{t}$ + 0/1@NLO + 2/3/4j@LO production, where massless c-quarks are included in the hard process. The studies presented in this chapter are still based on samples generated with `Herwig++` because they were carried out before the new `Herwig7` interface described in Section 7.2 was finalised and validated.

8.1. Event Generation

This section was published in almost identical form in Section 2 of [175], which was written by myself. Only minor changes were made when including it in this thesis, such as ensuring consistent style and spelling.

The cross section predictions and the hard process were generated at next-to-leading order using `MadGraph5_aMC@NLO` version 2.3.2. `MadSpin` [82] was used to simulate the decay of the top quarks and the W bosons. The resulting partons were matched to `Herwig++` version 2.7.1 to include the effects of the parton shower, hadronisation and multi-parton interactions. The UE-EE-5 `Herwig++` tune was used in conjunction with the corresponding CTEQ6L1 parton distribution function [176] in the parton shower and the underlying event. The angular-ordered parton shower was run with the default value for the strong coupling of $\alpha_s(m_Z) = 0.12$ obtained by the `Herwig++` authors from tuning to LEP data. Decays of b-quarks are handled by `EvtGen` [177]. The top quark mass was set to $m_t = 172.5$ GeV in all samples and a filter was applied during production to select events in the dilepton and single-lepton decay channels. Predictions of top quark pair production associated with two massive c-quarks were calculated using the 3-flavour scheme (3FS), i.e. the charm quarks are not considered in the PDFs and their evolution, with quark masses of $m_c = 1.55$ GeV and $m_b = 4.95$ GeV and the CT10f3 [178] PDF via `LHAPDF` [179].

In the samples, the jets produced in the hard process are required to have a p_{T} of at least 5 GeV with no requirement on η. For this, *partonic jets* were built from partons using the anti-k_t algorithm [135] with a radius of $R = 0.4$ using `FastJet` [180]. The predictions were done both with and without the minimal jet p_{T} requirement applied to the partonic jets initiated by c-quarks.[2] It was seen that applying the cut to partonic c-jets significantly reduces the rate of failures in the event generation.

The study was performed using the $H_{\mathrm{T}}/4$ and CMMPS [170] scales as defined in Table 8.1 for the renormalisation and factorisation scales μ_R and μ_F. Since the CMMPS scale, which is a generalisation of the BDDP scale introduced in [181], can become very soft, a lower bound defined as

$$\mu_F = \mu_R = \max(\mu_{\mathrm{CMMPS}}, \mu_{\mathrm{fixed}}) \tag{8.1}$$

[2] This behaviour was steered using the `maxjetflavor` parameter. `maxjetflavor=3`: jet cuts not applied to partonic jets initiated by c-quarks, `maxjetflavor=4`: jet cuts also applied to partonic jets initiated by c-quarks.

Scale name	Definition	Behaviour for $p_{T,i} \to 0$
$H_{\mathrm{T}}/4$	$\mu = \frac{1}{4}\sum_{i \in FS} E_{\mathrm{T},i}$	$\mu \to \frac{1}{4}\sum_{i \in FS} m_i \approx 87.03$ GeV
BDDP	$\mu = \sqrt{m_t \cdot \sqrt{p_{\mathrm{T},c} \cdot p_{\mathrm{T},\bar{c}}}}$	$\mu \to 0$ GeV
CMMPS	$\mu = \prod_{i=t,\bar{t},c,\bar{c}} E_{\mathrm{T},i}^{1/4}$	$\mu \to \prod_{i=t,\bar{t},c,\bar{c}} m_i^{1/4} \approx 16.35$ GeV

Table 8.1.: Scale definitions considered for the production of $t\bar{t}c\bar{c}$ events and their behaviour in the limit of $p_{\mathrm{T},i} \to 0$ for all final state particles i. The transverse energy is given by $E_{\mathrm{T},i} = \sqrt{m_i^2 + p_{\mathrm{T},i}^2}$ where m_i denotes the mass of any final state parton. The masses of the top and charm quarks are given by m_t and m_c, respectively. The BDDP scale is shown for comparison only and is not used for the studies presented in this thesis.

was introduced for this scale, where μ_{fixed} was set to values of $0, 50, 80$ and 150 GeV. The lowest value for which the event generation ran stably was 50 GeV, and to understand the effect of the lower bound, the choices of 50 and 150 GeV were studied in further detail. Given its different behaviour in the soft limit, the $H_{\mathrm{T}}/4$ scale did not necessitate any fixed lower bound.

Where possible, the recommended integration accuracy was used, corresponding to a statistical uncertainty of approximately 0.1% on the total cross section.[3,4] Different parameter settings in terms of the value of the minimal partonic jet p_{T}, the inclusion of the charm quarks in the partonic jet cut and the lower bound for the CMMPS scale were tried. For some of the tested configurations, technical problems were encountered which prevented the generation of events and motivate the setup used in the following with the c-quarks being included in the partonic jet cuts of $p_{\mathrm{T}} \geq 5$ GeV and a lower bound of at least 50 GeV for the CMMPS scale.

Scale variations were performed by simultaneously varying the hard process renormalisation and factorisation scales by factors of 0.5 and 2. Table 8.2 shows the cross sections obtained for the different scale choices along with the statistical and scale uncertainties as well as the efficiencies of the single-lepton and dilepton event filter.

Furthermore, the scale Q_{sh}, which is passed from `MadGraph5_aMC@NLO` to the `Herwig++` parton shower via the LHE event files and can be interpreted as restricting the hardness of the parton shower emissions, was varied down by factors of 0.5 and 0.25[5] for the inclusive $t\bar{t}$ sample following the procedure in [182]. The scale Q_{sh} is calculated for each event and is defined in `MadGraph5_aMC@NLO` as the partonic centre-of-mass energy multiplied by a

[3] The samples were generated from within the ATLAS software framework using the gridpack option in the interface for `MadGraph5_aMC@NLO`. In this procedure, the process libraries are generated, compiled and the integration grids are set up in one step. The resulting gridpack is subsequently used for the actual event generation which is done in parallel jobs, thus avoiding duplication of the setup step.

[4] The accuracy of the integration grids included in the gridpacks can be steered via the `req_acc` parameter in the `MadGraph5_aMC@NLO` run card. It sets the goodness of the integration done prior to the event generation and is approximately equal to the numerical precision of the total cross section calculated during the integration.

[5] Since variations of this scale via the `MadGraph5_aMC@NLO` run card are only possible starting with version 2.3.3, the variations were achieved by simultaneously multiplying the parameters `frac_low` and `frac_upp` in the file `SubProcesses/madfks_mcatnlo.inc` inside the `MadGraph5_aMC@NLO` process folder with the scale variation factors.

Scale choice	Lower bound	Cross section (stat. unc.) ± scale unc. · filter eff. [pb]
$H_{\mathrm{T}}/4$		$65.44(0.11)^{+84\%}_{-42\%} \cdot 0.54 = 35.5^{+84\%}_{-42\%}$
CMMPS	50 GeV	$135.6(0.3)^{+112\%}_{-47\%} \cdot 0.54 = 73.8^{+112\%}_{-47\%}$
CMMPS	150 GeV	$60.50(0.11)^{+84\%}_{-42\%} \cdot 0.54 = 32.9^{+84\%}_{-42\%}$

Table 8.2.: Inclusive cross sections for $t\bar{t}c\bar{c}$ production obtained with `MadGraph5_aMC@NLO` and `Herwig++` for different choices of the renormalisation and factorisation scales. Charm quarks are included in the 5 GeV jet p_{T} cut. The scale definitions are given in Table 8.1 and in Equation (8.1). The uncertainties are derived by varying simultaneously the factorisation and renormalisation scales up or down by a factor of two.

factor generated randomly in the interval between 0.1 and 1. A more complete description is given in Section 2.4.4 of [82] according to which the variation of Q_{sh} is an indication of the matching systematics related to the MC@NLO method.

An inclusive top quark pair production sample was generated at NLO with the same parameter settings as above except using the factorisation and renormalisation scale $\mu_F = \mu_R = \sqrt{m_t^2 + \frac{1}{2}(p_{T,t}^2 + p_{T,\bar{t}}^2)}$, using the 5FS, i.e. massless c-quarks, along with the CT10 PDF and jet cuts of $p_{\mathrm{T}} \geq 0.1$ GeV, corresponding to the sample in [183].

The aforementioned samples are compared to a `Sherpa` [97] sample generated using v2.1.1 with the MEPS@NLO [184] setup which was also used in [183]. The events are generated with a $t\bar{t}$ matrix element including zero or one partons simulated at NLO and 2, 3 or 4 partons at LO. The factorisation and renormalisation scales are defined as $\mu_F = \mu_R = \sqrt{m_t^2 + \frac{1}{2}(p_{T,t}^2 + p_{T,\bar{t}}^2)}$. The sample makes use of the 5FS with the CT10 PDF and the PS, hadronisation and UE are simulated using the default `Sherpa` settings.

8.2. Analysis

In this section, the aforementioned samples are analysed in a fiducial phase space typical of ATLAS $t\bar{t}$ analyses. Special attention is paid to the kinematic properties of the charm jets and the differences between the various samples are assessed. Note that no detector simulation was performed and the analysis is based on the generator-level particles, also referred to as *truth-level* particles.

The anti-k_t jet algorithm [135] with a radius parameter of $R = 0.4$ is used for jet reconstruction and is applied to stable final state particles, excluding muons and neutrinos. Only jets with a p_{T} larger than 25 GeV and within the inner detector acceptance, i.e. with $|\eta| < 2.5$, are considered. b- and c-hadrons with $p_{\mathrm{T}} > 5$ GeV are geometrically matched to the closest jet within $\Delta R = 0.4$ and the matched jet is called b- or c-jet, with the b-tag taking preference in the case of matches to both flavours. The c-hadron matching is not performed for decay products of the top quark and neither of both matchings considers b- and c-hadrons from previous b- or c-hadron decays. Therefore, the collection of b-jets includes the b-jets from the top quark decays, but c-jets will not include e.g. $W^+ \to c\bar{s}$ decays. As a result of this approach, a given b- or c-hadron is only matched to a single jet,

Sample	$tt1c$	$tt2c$	$ttcc$	Sum
$H_T/4$ (3FS)	$2.6^{+1.9}_{-1.0}$ pb	$1.6^{+1.5}_{-0.7}$ pb	$0.42^{+0.26}_{-0.15}$ pb	4.6 pb
CMMPS, 50 GeV cut-off (3FS)	$5.3^{+5.3}_{-2.4}$ pb	$2.6^{+2.9}_{-1.3}$ pb	$0.69^{+0.62}_{-0.29}$ pb	8.6 pb
CMMPS, 150 GeV cut-off (3FS)	$2.6^{+2.0}_{-1.0}$ pb	$1.6^{+1.6}_{-0.7}$ pb	$0.47^{+0.33}_{-0.19}$ pb	4.7 pb
Inclusive $t\bar{t}$ (5FS)	$2.0^{+0.3}_{-0.3}$ pb	$2.5^{+0.4}_{-0.4}$ pb	$0.25^{+0.04}_{-0.04}$ pb	4.8 pb
`Sherpa` $t\bar{t}$ + jets MEPS@NLO (5FS)	3.7 pb	3.6 pb	0.54 pb	7.8 pb

Table 8.3.: Fiducial cross sections for the $tt1c$, $tt2c$ and $ttcc$ categories. Where given, the uncertainties represent the changes of the fiducial cross sections from simultaneous variations of the renormalisation and factorisation scales by factors of 1/2 and 2. Taken from [175] and extended by mentions of the flavour schemes and the cross sections summed across all categories.

but a given jet may be matched to multiple same-flavour hadrons. If one c-hadron was matched to a jet, the latter is called a $1c$-jet, whereas in the case of two or more matching c-hadrons, the jet is called a $2c$-jet.

Based on these jet definitions, the event selection targets semi-leptonic or dilepton $t\bar{t}$ final states and requires the presence of exactly two b-jets, at least one c-jet and either a single charged lepton with $p_T > 25$ GeV or two charged leptons with $p_T > 20$ GeV each. Here, electrons and muons are considered as charged leptons. A categorisation is performed depending on the type and multiplicity of the c-jets: Events with a single $1c$-jet and no $2c$-jets are labelled $tt1c$, events with a single $2c$-jet and no $1c$-jets are labelled $tt2c$ and events with at least two c-jets of any type are labelled $ttcc$.

The fiducial cross sections of the different samples in the different analysis categories are shown in Table 8.3. Where given, the uncertainties indicate the changes of the fiducial cross sections from simultaneous variations of the renormalisation and factorisation scales by factors of 1/2 and 2. As can be seen, both the cross sections as well as the uncertainties predicted with the $H_T/4$ scale and the CMMPS scale with a lower cut-off at 150 GeV are in agreement for the $tt1c$ and $tt2c$ categories and, to a reduced extent, for the $ttcc$ category. The CMMPS scale option with a cut-off at 50 GeV predicts cross sections larger by about 50-100% in all categories, which is due to the fact that the lower cut-off leads to smaller scale values corresponding to increased values of the strong coupling α_s. The inclusive $t\bar{t}$ sample predicts a relatively large cross section in the $tt2c$ category and a reduced cross section in the $ttcc$ category. This behaviour corresponds to the parton shower preferentially generating close-by or collimated rather than hard and well-separated pairs of charm quarks, as described in Sections 3.3 and 3.4. It can also be seen that the scale uncertainties are much smaller for the inclusive $t\bar{t}$ sample than for the samples where the charm quarks are included in the matrix elements. In the latter case, the scale uncertainties amount to 35-100% of the respective fiducial cross sections. This reflects the complexity of the $t\bar{t}c\bar{c}$ process and at the same time underlines that scale variations at the level of inclusive $t\bar{t}$ production do not give realistic perturbative uncertainty estimates for $t\bar{t}c\bar{c}$ production. For the $tt1c$ and $ttcc$ categories, the fiducial cross sections predicted by the `Sherpa` multi-jet merged sample lie in between the other predictions, while for the $tt2c$ category the `Sherpa` sample predicts more events.

Figure 8.1 shows the charm p_T fraction, which is the ratio of the scalar p_T sum of the charm

hadrons in a jet over the p_{T} of the jet and therefore measures the relative hardness of the heavy flavour components in comparison to the associated jet. Observables of this type are interesting because the non-zero mass of heavy flavour quarks shields the IR divergences described in Section 3.2. Therefore, heavy flavour quarks are typically accompanied by less gluon radiation than light flavour quarks. This is known as the *dead cone effect*. As a consequence, heavy flavour quarks or hadrons retain a larger fraction of the momentum or energy of the associated jet than their light flavour counterparts. [185] As can be seen from the plots, this fraction may exceed unity since the jets are built from stable particles, i.e. particles with a mean lifetime $\tau > 3 \cdot 10^{-11}$ s, and stable decay products of the charm hadron may end up outside of the jet cone. Both the combination of all c-jets as well as a separate breakdown for the $1c$-jets and the $2c$-jets are presented. The curves reflect the differences in the cross sections, but still show similar shapes for most of the samples with a maximum located at fractions between 0.6 and 0.8. The distributions are most different below momentum fractions of 0.4 and especially in the case of $1c$-jets, for which both the inclusive $t\bar{t}$ as well as the `Sherpa` samples predict a second peak at p_{T} fractions of around 0.1-0.2. For the `Sherpa` sample, this peak is even more pronounced than the peak at higher momentum fractions. These structures are also reflected in the plot for all c-jet types. The origin of the different shapes, especially in the case of the `Sherpa` sample, could not be identified in this study. It should be noted, however, that this region of the charm p_{T} fraction was found to be sensitive to the application of a p_{T} requirement on jets built from the partons at the level of the hard process. Figure 8.2 shows the charm p_{T} distribution for two versions of the $t\bar{t}c\bar{c}$ sample with the $H_{\mathrm{T}}/4$ scale choice. In the nominal version of this sample, partonic jets initiated by c-quarks are included in the minimum p_{T} requirement of 5 GeV, as described in Section 8.1. In an alternative version of this sample, the generator-level cuts on partonic jets are not applied to c-quark-initiated jets. The comparison of both settings was motivated by technical problems with the event generation when no p_{T} requirement was applied on charm-initiated partonic jets. Both samples show increasing differences for soft charm hadrons below a p_{T} fraction of approximately 0.4. A cut at this value is applied subsequently in the analysis in order to mitigate effects of the generator-level jet cut.

The differential cross section as a function of the number of charm jets not produced as part of the top quark decay chain is shown in Figure 8.3. The samples with the $H_{\mathrm{T}}/4$ scale and the CMMPS scale with a lower cut-off at 150 GeV mostly show very similar behaviour. Due to their larger cross sections, the sample generated with the CMMPS scale with a lower cut-off at 50 GeV and the `Sherpa` sample predict more events for all c-jet multiplicities. While the prediction of the inclusive $t\bar{t}$ sample lies much above the $t\bar{t}c\bar{c}$ samples for the zero c-jet bin, it drops below the other samples for two or more c-jets. At the time when this study was performed, the predictions of top quark pair plus charm jet production in the $t\bar{t}H(b\bar{b})$ analysis [186] were based on the very same inclusive $t\bar{t}$ sample.

The shapes of the p_{T} distributions of the leading c-jet are shown in the left pane of Figure 8.4. The curves are normalised to the fiducial cross sections in the different categories and indicate that the `Sherpa` and the inclusive $t\bar{t}$ samples have a softer leading c-jet distribution, while the curve corresponding to the sample with the CMMPS scale and a lower cut-off at 150 GeV is the hardest. The right pane of Figure 8.4 shows the differential cross section as a function of the ΔR separation between both c-jets for events containing exactly two charm jets. With the exception of the CMMPS/50 GeV scale choice, most of the predictions are comparable for separations below $\Delta R = 0.8$. Notably, the `Sherpa` sample comes closest to the inclusive $t\bar{t}$ prediction, which drops below the other samples for separations larger than $\Delta R = 0.8$. This behaviour is expected due to the nature of

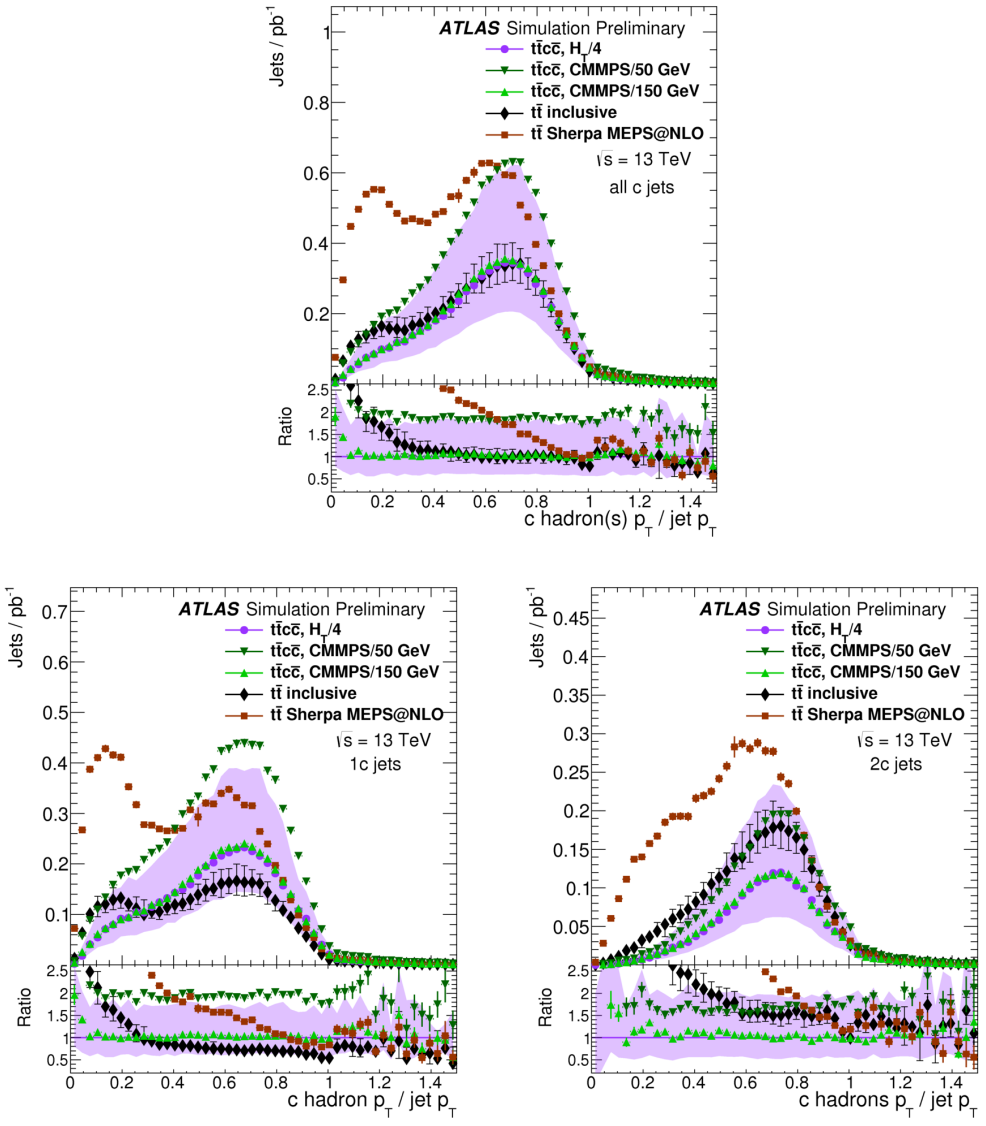


Figure 8.1.: The fraction of the charm hadron p_{T} w.r.t. the p_{T} of the corresponding charm jet. The upper shows all charm jets, whereas only jets with exactly one associated c-hadron ($1c$-jets) are shown in the lower left plot and only jets with two or more associated c-hadrons ($2c$-jets) are shown in the lower right figure. The band around the sample generated with the $H_{\mathrm{T}}/4$ scale option (lilac circle markers) and the bars on the inclusive $t\bar{t}$ sample (black diamond markers) show the simultaneous variations of the renormalisation and factorisation scales by factors of 1/2 and 2. For the other samples, only the statistical uncertainties are shown. All samples are normalised to the respective fiducial cross sections and the ratios in the lower panes are given w.r.t. the calculation with the $H_{\mathrm{T}}/4$ scale option. The charm p_{T} fraction may exceed unity since the jets are built from stable particles, i.e. particles with a mean lifetime $\tau > 3 \cdot 10^{-11}$ s, and stable decay products of the charm hadron may end up outside of the jet cone. Taken from [175].

the parton shower and matrix element descriptions as described in Sections 3.3 and 3.4.

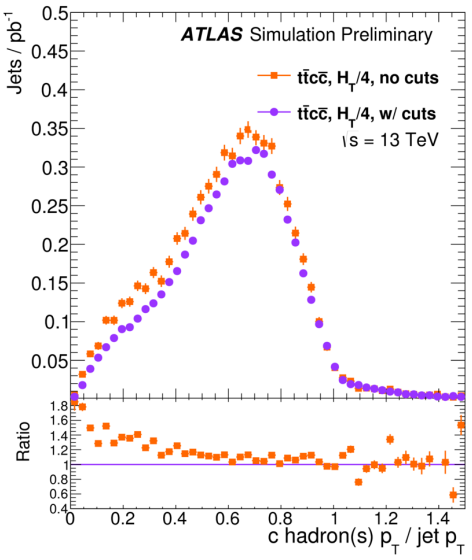


Figure 8.2.: The fraction of the charm hadron p_{T} w.r.t. the p_{T} of the corresponding charm jet. The distributions are shown for samples with the $H_{\mathrm{T}}/4$ scale option, which were generated with (lilac circle markers) and without (orange square markers) a minimal requirement of $p_{\mathrm{T}} > 5$ GeV on partonic jets initiated by c-quarks at the level of the hard process. This cut was configured using the `maxjetflavour` parameter as described in Section 8.1. Both samples are normalised to the respective fiducial cross sections and the ratio in the lower pane is given w.r.t. the calculation with the p_{T} requirement. Taken from [175].

For these intermediate and larger separations, the `Sherpa` sample lies closer to the $t\bar{t}c\bar{c}$ calculations, indicating that the multi-jet merging interpolates between parton shower-like descriptions of close-by objects and matrix element-like descriptions of separated objects, which is precisely the desired behaviour.

Finally, the fiducial cross sections of the different categories and the differential cross section as a function of the p_{T} of the leading c-jet are shown again in the left and right pane of Figure 8.5, respectively. Here, the effects of variations of the hard shower scale Q_{sh} introduced in Section 8.1 by factors of 1/2 and 1/4 are shown. While these are negligible for the cross sections in the different categories, an effect can be seen in the transverse momentum distribution of the leading c-jet, where lower values of Q_{sh} lead to a softer p_{T} spectrum.

8.3. Conclusion

In this chapter, a study comparing different predictions for the $t\bar{t}c\bar{c}$ process was presented. Monte Carlo samples with massless and massive c-quarks, c-quarks from the hard process matrix elements and the parton shower, NLO matched and multi-jet merged calculations, different parton showers, different functional forms and cut-off values of the hard process renormalisation and factorisation scales, different settings of the hard parton shower cut-off scale and different hadronisation models were compared and fiducial cross sections and differential distributions of various observables were studied. The Monte Carlo samples

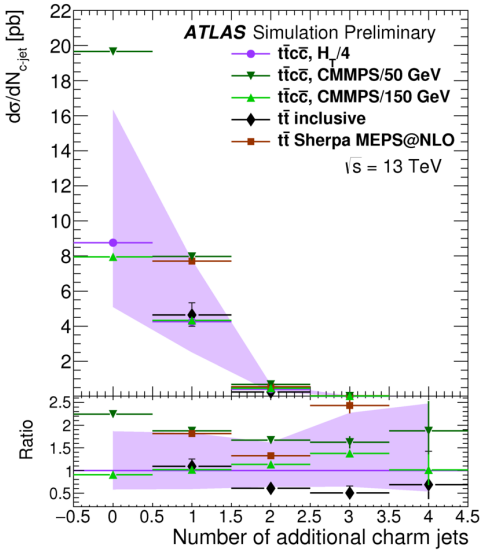


Figure 8.3.: The differential cross section as a function of the number of the charm jets not produced as part of the top quark decay chain. A cut at 0.4 was applied on the fraction of the charm hadron p_{T} w.r.t. the p_{T} of the corresponding jet. The predictions of the zero c-jet bin from the inclusive $t\bar{t}$ and `Sherpa` samples lie above the range of the axis. The band around the sample generated with the $H_{\mathrm{T}}/4$ scale option and the bars on the inclusive $t\bar{t}$ sample show the simultaneous variations of the renormalisation and factorisation scales by factors of 1/2 and 2. For the other samples, only the statistical uncertainties are shown. All samples are normalised to the respective fiducial cross sections and the ratio in the lower pane is given w.r.t. the calculation with the $H_{\mathrm{T}}/4$ scale option. Taken from [175].

were generated with `MadGraph5_aMC@NLO` matched to the `Herwig++` parton shower and with `Sherpa`.

While it should be noted that this study should be considered a starting point for further investigations, it already shows that uncertainties for the $t\bar{t}c\bar{c}$ final state based on hard process scale and hard parton shower cut-off scale variations for inclusive $t\bar{t}$ production do not reflect the large perturbative uncertainties obtained with $t\bar{t}c\bar{c}$ matrix elements. Conversely, systematic uncertainties derived from inclusive $t\bar{t}$ matrix elements underestimate the uncertainties of the $t\bar{t}c\bar{c}$ process. As a consequence, this study formed the basis for the evaluation of systematic uncertainties related to the $t\bar{t}c\bar{c}$ background to $t\bar{t}H(b\bar{b})$ production presented in [186].

The fiducial cross sections in the different analysis categories show large differences between the different hard process scale choices and similarly large uncertainties from variations of these scales. In both cases, the differences reach levels of up to 50-100%, reflecting the high complexity of this process due to the presence of two very different scales given by the t- and c-quark masses. For the inclusive $t\bar{t}$ sample, the uncertainties from scale variations are much smaller, which is expected and underlines that hard process scale variations at the level of the $t\bar{t}$ process do not provide probes of the uncertainties in the $t\bar{t}c\bar{c}$ final state. In most cases, the differences in the shapes of kinematic observables are less dramatic than for the fiducial cross sections. A notable exception is the charm p_{T} fraction, where the

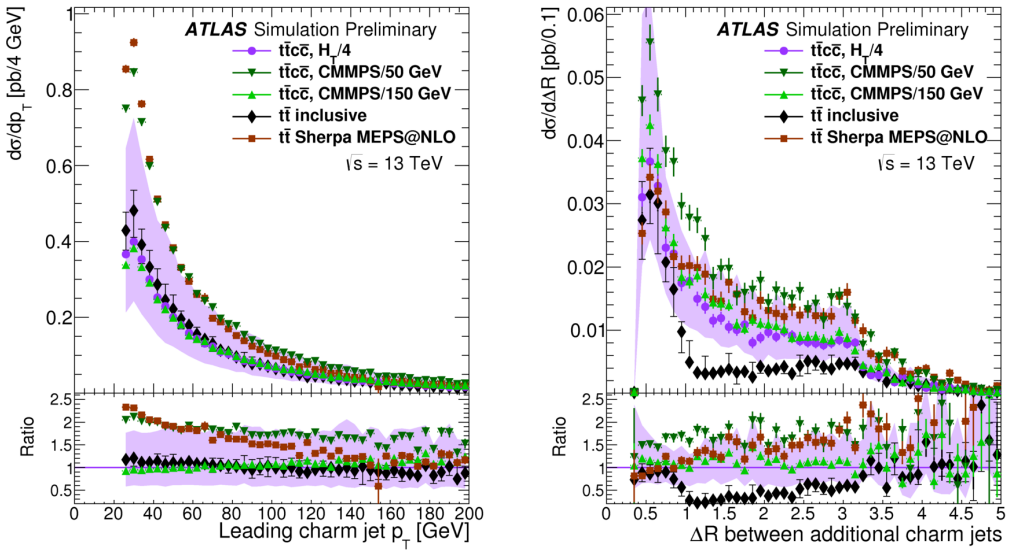


Figure 8.4.: Differential cross sections as a function of the transverse momentum of the leading charm jet (left) and the separation ΔR between two c-jets (right). For the latter, only events with exactly two c-jets are considered. The band around the sample generated with the $H_{\mathrm{T}}/4$ scale option and the bars on the inclusive $t\bar{t}$ sample show the simultaneous variations of the renormalisation and factorisation scales by factors of 1/2 and 2. For the other samples, only the statistical uncertainties are shown. All samples are normalised to the respective fiducial cross sections and the ratios in the lower panes are given w.r.t. the calculation with the $H_{\mathrm{T}}/4$ scale option. Taken from [175].

`Sherpa` sample and to a much lesser degree the inclusive $t\bar{t}$ sample show a second peak at lower momentum fractions in the spectrum of charm jets containing a single c-hadron. The origin of this difference with respect to the $t\bar{t}c\bar{c}$ matrix element calculations could not be established within the scope of this study. In the differential cross section as a function of the ΔR between two charm jets, it can be seen that the multi-jet merged calculation from `Sherpa` interpolates between the inclusive $t\bar{t}$ sample for small separations and the samples with $t\bar{t}c\bar{c}$ matrix elements for larger separations. In contrast, the inclusive $t\bar{t}$ prediction drops below the matrix element calculations for larger separations, which is expected as the parton shower approximation holds for the soft and collinear regions of phase space.

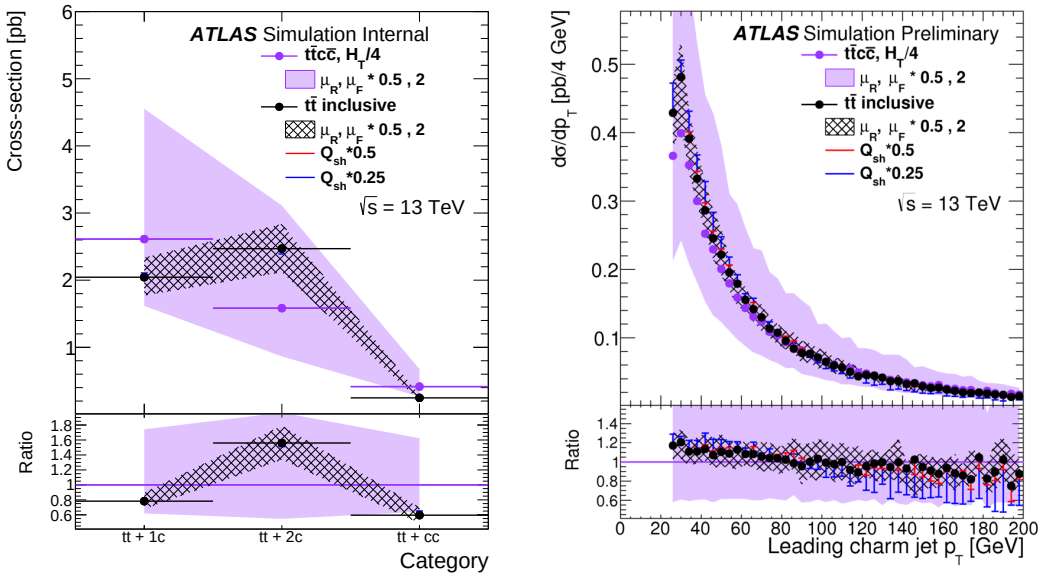


Figure 8.5.: Fiducial cross sections in the different analysis categories (left) and differential cross section as a function of the leading charm jet p_{T} (right). The solid and hatched bands were obtained by simultaneously varying the renormalisation and factorisation scales by factors of 1/2 and 2 for the $t\bar{t}c\bar{c}$ sample with the $H_{\mathrm{T}}/4$ scale and the inclusive $t\bar{t}$ sample. Variations of the hard shower scale Q_{sh} by factors of 1/2 and 1/4 in the inclusive $t\bar{t}$ sample are indicated by red and blue bars. They are too small to be visible in the left plot. All samples are normalised to the respective cross sections and the ratios in the lower panes are given w.r.t. the calculation with the $H_{\mathrm{T}}/4$ scale option. Taken from [175].

Part III.

Search for the $Z\gamma$ Decay Mode of the Higgs Boson and for New High-Mass Resonances in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector

Chapter 9

Introduction

The $H \to Z\gamma$ decay is one of the rare decay modes of the Higgs boson. By itself, the $H \to Z\gamma$ branching fraction is comparable to its $H \to \gamma\gamma$ counterpart, with $\mathcal{B}(H \to Z\gamma) = 1.54 \cdot 10^{-3}$ and $\mathcal{B}(H \to \gamma\gamma) = 2.27 \cdot 10^{-3}$, respectively, for a Higgs mass of $m_H = 125.09$ GeV [33]. However, branching ratios of $\mathcal{B}(Z \to ee) = 3.363\%$ and $\mathcal{B}(Z \to \mu\mu) = 3.366\%$ reduce the number of events in the leptonic decay channels [118]. While more signal events may be selected using the hadronic Z boson decays, the leptonic final states offer the advantages of a good invariant mass resolution and manageable backgrounds. Because of the aforementioned branching ratios in combination with the signal-to-background ratio, the $Z\gamma$ decay of the Higgs boson has not yet been observed, whereas the $\gamma\gamma$ decay mode was among the first channels to be discovered. Searches for the $Z\gamma$ decay mode are still interesting because modifications of the expected branching ratios or exclusion limits would hint at physics deviating from the SM. Possible scenarios investigated in the literature include a neutral scalar particle acting as a Higgs impostor [187,188], composite models [189] and models featuring additional charged particles in the $H \to Z\gamma$ loop [190–192].

Searches for the $Z(\ell\ell)\gamma$ decay ($\ell = e, \mu$) of the Higgs boson were previously performed by both ATLAS and CMS. No significant excesses over the background were found. Using datasets of 4.5 fb^{-1} at 7 TeV and 20.3 fb^{-1} at 8 TeV and assuming a Higgs boson mass of $m_H = 125.5$ GeV, the ATLAS Collaboration observed an upper limit at the 95% CL of 11 times the SM expectation [193]. In the presence (absence) of the $H \to Z\gamma$ decay, limits of 10 (9) were expected, corresponding to hypothesised signal strengths of $\mu = 1$ ($\mu = 0$). Based on datasets of 5.0 fb^{-1} at 7 TeV and 19.6 fb^{-1} at 8 TeV, the CMS Collaboration reported observed (expected) exclusion limits of 9.5 (10) times the SM expectation for $m_H = 125$ GeV [194].

The $Z\gamma$ final state was also previously exploited in different analyses to search for new resonances with masses above the Higgs boson. Using a dataset of 20.3 fb^{-1} at 8 TeV collected in 2012 and $Z\gamma$ decays in the electron and muon channels, ATLAS excluded spin-1 Low Scale Technicolor (LSTC) resonances with masses between 200 and 700 GeV and between 750 and 890 GeV as well as neutral composite scalar resonances in most of the search

range between 200 and 1180 GeV for certain benchmark sets of model parameters [195]. Based on 3.2 fb^{-1} of data at 13 TeV collected in 2015, ATLAS published 95% CL exclusion limits on $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ between 295 and 8.2 fb for a narrow scalar resonance X produced in gluon-gluon fusion for a mass range between 250 GeV and 2.75 TeV [196]. For this analysis, the electron and muon channels as well as the hadronic channel of the $Z\gamma$ decays were used. Adding 10.1 fb^{-1} of data at 13 TeV collected in 2016 and restricting the analysis to the electron and muon final states, ATLAS reported observed (expected) limits at the 95% $\mathrm{CL_s}$ between 215 and 5 fb (103 and 5 fb) on $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ for a mass range between 250 GeV and 2.4 TeV [197]. An analysis of hadronic $Z\gamma$ decays using 36.1 fb^{-1} of pp collisions at 13 TeV was performed by ATLAS in [198] and found upper limits at the 95% CL on $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ between 10 and 0.1 fb for resonance masses between 1 and 6.8 TeV. A CMS analysis of 19.7 fb^{-1} at 8 TeV and 2.7 fb^{-1} at 13 TeV, using both leptonic and hadronic $Z\gamma$ decays, found 95% CL exclusion limits on $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ between 300 and 2.5 fb for scalar resonances with masses between 200 GeV and 3 TeV and two different width hypotheses.

The analysis presented in the following chapters was published in [60]. It includes both a search for the leptonic $Z\gamma$ decay mode of the Higgs boson, also referred to as the *low-mass search*, as well as a search for resonances of higher masses decaying to the same final state, correspondingly termed *high-mass search*. In comparison to previous searches, the analysis benefits from a larger dataset. The search for the $Z\gamma$ decay of the Higgs boson profits from the increased Higgs production cross section due to the higher centre-of-mass energy of 13 TeV in LHC Run 2. A category sensitive to VBF Higgs production was added to the low-mass search and models for spin-2 resonances produced from gluon-gluon and quark-antiquark initial states were added to the high-mass search. The presentation of the analysis proceeds as follows: The event selection and categorisation are described in Chapter 10. In Chapter 11, the simulated event samples, the likelihood function used to describe the analysis as well as the signal and background modelling are detailed. This is followed by a description of the systematic uncertainties in Chapter 12. Finally, the results are presented in Chapter 13. My contributions to this analysis were the development of the scheme for the theoretical uncertainties and their evaluation as well as general contributions and cross-checks related to the event selection and categorisation.

Based on the results of the $H \to Z\gamma$ search presented in [60] and described in this thesis, an extrapolation predicts a significance of 4.9σ for a dataset of 3000 fb^{-1} [199].

Chapter 10

Event Selection and Categorisation

In this chapter, the event selection and categorisation are described. To a large part, the event selection is identical for the search of the $Z\gamma$ decay of the SM Higgs boson and for the search for high-mass resonances. A common preselection includes requirements on the lepton candidates and loosened requirements on the photon candidate. The latter are then replaced by tightened requirements in the final selection steps, which are different for both mass ranges. This approach allows for the use of data-driven methods for the estimation of the photon-related backgrounds based on control regions.

10.1. Event Selection

The analysis is based on pp collision data collected by the ATLAS detector in 2015 and 2016 at a centre-of-mass energy of 13 TeV and a bunch spacing of 25 ns. Large parts of the event selection are identical for the low- and the high-mass searches, with only the final steps being different between both parts of the analysis. The event selection begins by considering the so-called *good-run list* (GRL), which marks time slots of one minute during which the detector was sufficiently functional. Only events from accordingly flagged time slots are analysed. This criterion leads to a dataset of 36.1 fb^{-1}. For these events, the presence of at least one primary vertex candidate is required. As described in Section 5.1, the PV with the largest sum of squared momenta of the associated tracks is selected as the PV of the interaction of interest in a given collision event. Events passing single lepton and dilepton triggers are considered. The relevant triggers and their p_{T} thresholds are summarised in Table 10.1.

From the resulting dataset, events are preselected that contain at least one photon candidate and at least two opposite-charge, same-flavour lepton candidates ℓ, where $\ell = e, \mu$, with an invariant mass $m_{\ell\ell} > 45$ GeV. Muons are required to have $p_{\text{T}} > 10$ GeV, $|\eta| < 2.7$ and satisfy the *Medium* ID requirements [121]. Likewise, electrons are required to have $p_{\text{T}} > 10$ GeV, $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ and to satisfy the *Medium Likelihood* ID

Trigger type	Trigger p_{T} thresholds per year 2015	2016	Final analysis p_{T} thresholds
$1e$	24 GeV	26 GeV	27 GeV
	60 GeV	60 GeV	61 GeV
	120 GeV	140 GeV	141 GeV
$2e$	12 GeV, 12 GeV	17 GeV, 17 GeV	18 GeV, 18 GeV
1μ	26 GeV	26 GeV	27 GeV
	50 GeV	50 GeV	51 GeV
2μ	22 GeV, 8 GeV	22 GeV, 8 GeV	24 GeV, 10 GeV

Table 10.1.: Summary of the different triggers and the corresponding final lepton p_{T} thresholds in the analysis. Single lepton and dilepton triggers are used. For dilepton triggers, the first threshold refers to the leading and the second threshold to the subleading lepton. It should be noted that the triggers with different p_{T} thresholds may also differ in other aspects, in particular in particle ID requirements. The final analysis p_{T} thresholds are larger than the corresponding trigger thresholds to avoid trigger turn-on effects.

requirements [141]. In order to ensure that both lepton tracks are compatible with the selected primary vertex, requirements on the transverse and longitudinal impact parameters are applied. Specifically, $|d_0|/\sigma_{d_0} < 3\,(5)$ is required for the muon (electron) tracks, where σ_{d_0} denotes the uncertainty on the transverse impact parameter d_0. Additionally, the lepton tracks must satisfy $|\Delta z_0 \sin\theta| < 0.5$ mm, where Δz_0 describes the distance along the z-direction between the PV and the point of closest approach of the track to the beamline and θ the polar angle of the track. For both electrons and muons, the *LooseTrackOnly* isolation working point is required, which only makes use of track isolation, but not of calorimeter isolation. Specifically, for muons, the cone size is $\Delta R = 0.3$ for $p_{\mathrm{T}} < 33.3$ GeV and $\Delta R = 10/(p_{\mathrm{T}}/\mathrm{GeV})$ for $p_{\mathrm{T}} > 33.3$ GeV and for electrons $\Delta R = 0.2$ for $p_{\mathrm{T}} < 50$ GeV and $\Delta R = 10/(p_{\mathrm{T}}/\mathrm{GeV})$ for $p_{\mathrm{T}} > 50$ GeV. For the *LooseTrackOnly* working point, the allowed track isolation $p_{\mathrm{T}}^{\mathrm{iso}}$ is chosen in such a way that an efficiency of 99% is reached. For photons, the *Loose* ID working point [137] is used along with the requirements of $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. No isolation requirements are applied for photons at the preselection level.

After the preselection stage, an overlap removal procedure is performed. In case different electrons share a track or have clusters with separations of $|\Delta\eta| < 0.075$ and $|\Delta\phi| < 0.125$, only the electron with the highest p_{T} is kept. Furthermore, only electrons that are at least $\Delta R = 0.02$ away from muons are considered. Similarly, photons within $\Delta R = 0.3$ around an electron or muon candidate are discarded. This serves to reject photons from final-state radiation. Finally, jets are removed that are within $\Delta R = 0.2$ around a photon or lepton candidate.

For this as well as for the categorisation, which will be described in Section 10.2, jets reconstructed with the anti-k_t algorithm with a radius parameter of $R = 0.4$ are used. The reconstruction and calibration of jets proceeds as described in Section 5.2. Jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 4.4$. Additional requirements are imposed on jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ in order to reject jets from pile-up interactions. This is done by means of the jet-vertex-tagger (JVT) algorithm [200, 201], which for a given jet relates the scalar p_{T} sum of the tracks associated to the hard scatter PV to both the jet

p_{T} and the total scalar p_{T} sum of all tracks of the jet, including a correction aiming at reducing the pile-up dependence.

In the next stages of the event selection, the final analysis objects are defined and the relevant Z and either H or X candidates are selected. The leptons are required to match the relevant triggers, which means, that the reconstructed and selected leptons should be compatible with the objects that activated a given trigger for an event. In order to avoid turn-on effects near the trigger thresholds, additional p_{T} requirements are imposed on the leptons, depending on which trigger was active for a given event. An overview of the different triggers and the corresponding final p_{T} thresholds is given in Table 10.1. A correction for energy losses from collinear final-state radiation is performed for $Z \to \mu\mu$ candidates, in which ECal clusters with $p_{\mathrm{T}} > 1.5$ GeV lying in close proximity of $\Delta R < 0.15$ to a muon track are added if the resulting corrected invariant mass $m_{\mu\mu} < 100$ GeV [202]. This results in a 3% improvement of the $m_{Z\gamma}$ resolution in the low-mass search and a similar improvement across the high-mass search range. Subsequently, the four-momenta of the leptons forming the Z boson candidate are adjusted using a constrained kinematic fit [203]. This is done in both the electron and the muon channel. The Z boson lineshape is parameterised with a Breit-Wigner distribution and used as a constraint and the lepton energies and momenta are varied according to Gaussian distributions for which the width parameters are taken to be the corresponding expected resolutions. The kinematic fit improves the $m_{Z\gamma}$ resolution by 13% (7%) in the electron (muon) channel of the low-mass search. In the high-mass search, the resolution improvements range between 10% (9%) at 200 GeV and about 0% (33%) at 2.5 TeV in the electron (muon) channel. The resulting Z boson candidate is required to have a mass between 76.18 GeV and 106.18 GeV. In case more than one Z boson candidate fulfils this requirement, the candidate with the mass closest to $m_Z = 91.18$ GeV [118] is selected.

The Z boson candidate is then combined with the highest p_{T} photon to form H boson and X candidates. Subsequently, the *Tight* photon ID requirement is used along with *FixedCutLoose* photon isolation [137]. The basic workings of isolation were described in Section 5.3.2. Here, a cone size of $\Delta R = 0.2$ is used for both the track and the calorimeter isolation and the track (calorimeter) isolation $p_{\mathrm{T}}^{\mathrm{iso}}$ ($E_{\mathrm{T}}^{\mathrm{iso}}$) is required to be less than 5% (6.5%) of the photon p_{T}. The very last steps of the event selection are specific to the low- and high-mass searches of the analysis. In the search for the $Z\gamma$ decay of the Higgs boson, the final requirements are $p_{\mathrm{T}} > 15$ GeV for the photon and 115 GeV $< m_{Z\gamma} <$ 170 GeV. In the high-mass part of the analysis, a relative criterion of $p_{\mathrm{T}}/m_{Z\gamma} > 0.3$ is used in addition to an increased absolute photon p_{T} threshold of 50 GeV and the three-body invariant mass $m_{Z\gamma}$ is required to be between 200 GeV and 2.5 TeV. The main aspects of the event selection are summarised in Table 10.2.

For the low-mass search, the expected signal event selection efficiencies range between 20% and 24%, depending on the Higgs production mode. For the high-mass search, the expected signal event selection efficiencies vary between approximately 20% and 55%, depending on the resonance mass, spin and production mode, which can be seen from Figure 10.1. In general, the event selection efficiencies increase towards higher resonance masses, which is due to the increased transverse momenta of the final state photon and leptons and the resulting higher efficiencies for passing the various p_{T} requirements. In the spin-0 case, the event selection efficiency in the electron channel is slightly worse than in the muon channel at a resonance mass of $m_X = 300$ GeV, but eventually exceeds the latter for resonance masses above $m_X = 1$ TeV. For spin-2 resonances, the event selection efficiency is larger in the quark-induced production mode. The reason for this is that the photon spectrum

	$H \to Z\gamma$	$X \to Z\gamma$
Preselection	• GRL, primary vertex, triggers • ≥ 2 leptons passing the preselection ◦ electrons: $p_{\mathrm{T}} > 10$ GeV, ID: *Medium Likelihood*, ISO: *LooseTrackOnly* ◦ muons: $p_{\mathrm{T}} > 10$ GeV, ID: *Medium*, ISO: *LooseTrackOnly* • leptons have opposite sign and $m_{ll} > 45$ GeV • 1 photon passing the preselection ◦ $p_{\mathrm{T}} > 10$ GeV, ID: *Loose*, ISO: no requirements • overlap removal and ≥ 1 H/X candidate present afterwards • trigger matching and final p_{T} thresholds • 76.18 GeV $\leq m_{\ell\ell} \leq$ 106.18 GeV	
Final selection	• photon ID: *Tight* • photon ISO: *FixedCutLoose* • $p_{\mathrm{T}}^{\gamma} \geq 15$ GeV • 115 GeV $\leq m_{Z\gamma} \leq 170$ GeV	• photon ID: *Tight* • photon ISO: *FixedCutLoose* • $p_{\mathrm{T}}^{\gamma} > 50$ GeV and $p_{\mathrm{T}}^{\gamma}/m_{Z\gamma} \geq 0.3$ • $m_{Z\gamma} \geq 200$ GeV

Table 10.2.: Overview of the event selection strategies in the $H \to Z\gamma$ and $X \to Z\gamma$ searches. The common preselection includes momentum and quality requirements on the leptons, but does not yet apply the final set of criteria on the photons. These are left to the final selection stage so that control regions can be defined.

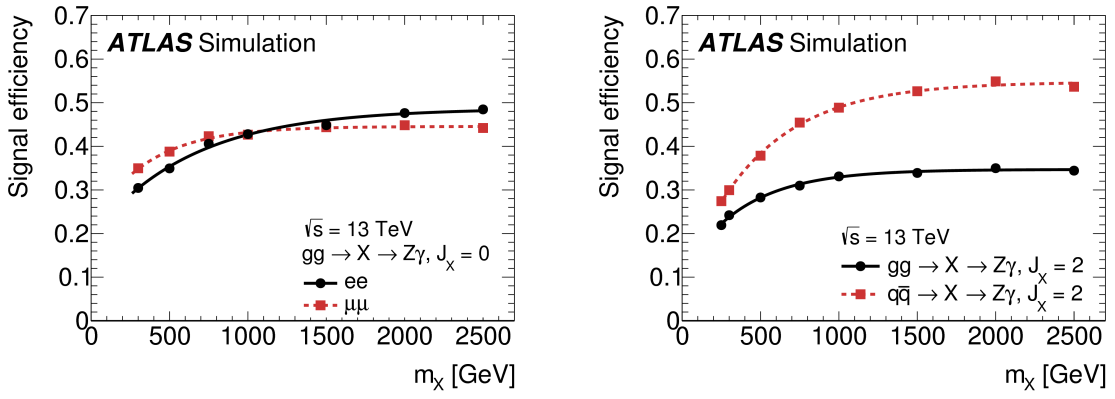
is harder in the quark-induced than in the gluon-induced production mode, which can be seen in Figure B.2. A selection of event selection cutflow tables is given in Appendix B.1.

10.2. Categorisation

In order to maximise the sensitivity of the analysis, the selected events are divided into different categories. In the search for the $Z\gamma$ decay of the Higgs boson, six different categories were defined based on the criteria sketched in Figure 10.2. At first, a *Boosted Decision Tree* (BDT) is used to isolate events with kinematic properties typical of the VBF production mode. This is typically characterised by the presence of two forward jets and by the decay products of the Higgs boson dominating the central rapidity region.

Symbol	Definition
$m_{j_1 j_2}$	invariant dijet mass
$\Delta\phi(Z\gamma, j_1 j_2)$	azimuthal separation between the $Z\gamma$ and the dijet system
$p_{\mathrm{T}t}$	$(\boldsymbol{p}_T^{\gamma} + \boldsymbol{p}_T^{Z}) \times (\boldsymbol{p}_T^{\gamma} - \boldsymbol{p}_T^{Z})/\lvert\boldsymbol{p}_T^{\gamma} - \boldsymbol{p}_T^{Z}\rvert$
$\eta_{\mathrm{Zeppenfeld}}$	$\lvert\eta_{Z\gamma} - (\eta_{j_1} + \eta_{j_2})/2\rvert$
$\Delta R_{\mathrm{min}}(Z/\gamma, j)$	minimum ΔR between the photon or the Z boson and either of the jets
$\Delta\eta(j_1, j_2)$	$\lvert\eta_{j_1} - \eta_{j_2}\rvert$

Table 10.3.: Variables entering the BDT used to define a category enriched in VBF events.



(a) Spin-0. (b) Spin-2.

Figure 10.1.: Event selection efficiency, including detector acceptance, for the high-mass resonance in terms of the spin J_X of the resonance X and either the Z boson decay channel (for spin-0) or the production mode (for spin-2). The markers denote the event selection efficiencies determined from Monte Carlo signal samples at the corresponding mass points. The curves show the parameterisation with an exponentiated second-order polynomial as a function of the resonance mass, which is used as part of the signal model. Published in [60].

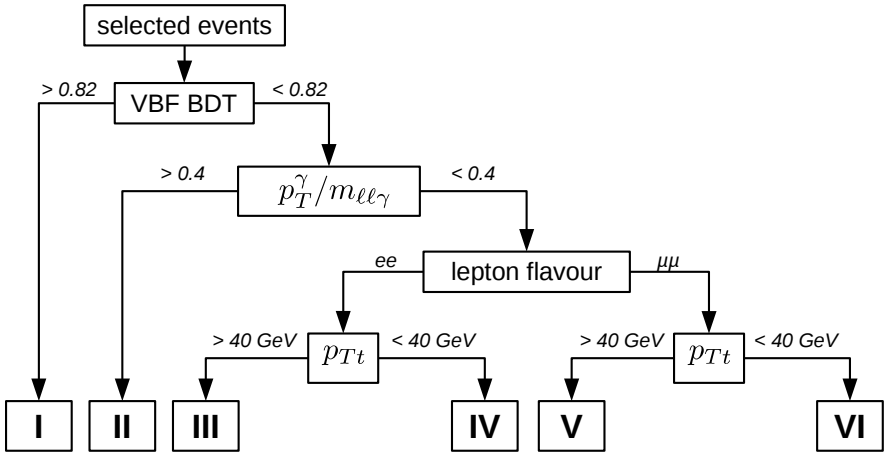


Figure 10.2.: Schematic of the categorisation of the selected events in the search for the $Z\gamma$ decay of the Higgs boson.

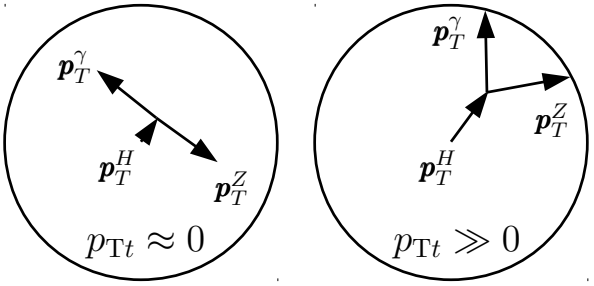


Figure 10.3.: Illustration of the $p_{\mathrm{T}t}$ variable defined in Equation (10.1) which is correlated with the transverse momentum of the Higgs boson.

To achieve the desired separation, the variables listed in Table 10.3, that are sensitive to VBF-like event topologies, were used in the BDT, which was applied to events preselected according to the requirements $N_{\mathrm{jets}} \geq 2$ and $\Delta\eta(j_1, j_2) > 2$. j_1 and j_2 denote the jets with the largest and second-largest transverse momentum, respectively. From the events that failed either the BDT preselection or did not pass the cut on the BDT discriminant, those with a relative photon momentum $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} > 0.3$ are separated. The remaining events are classified according to the flavour of leptons coming from the Z boson decay and eventually by the variable

$$p_{\mathrm{T}t} = (\boldsymbol{p}_{\mathrm{T}}^{\gamma} + \boldsymbol{p}_{\mathrm{T}}^{Z}) \times \frac{\boldsymbol{p}_{\mathrm{T}}^{\gamma} - \boldsymbol{p}_{\mathrm{T}}^{Z}}{|\boldsymbol{p}_{\mathrm{T}}^{\gamma} - \boldsymbol{p}_{\mathrm{T}}^{Z}|} = 2\frac{\boldsymbol{p}_{\mathrm{T}}^{Z} \times \boldsymbol{p}_{\mathrm{T}}^{\gamma}}{|\boldsymbol{p}_{\mathrm{T}}^{\gamma} - \boldsymbol{p}_{\mathrm{T}}^{Z}|} \ . \tag{10.1}$$

The latter is correlated with the transverse momentum of the Higgs boson which is illustrated in Figure 10.3 [204, 205].

The resulting signal event fractions $\xi_c^{(m)}$ for each category c and for each Higgs production mode m as well as the composition $\zeta_c^{(m)}$ of each category in terms of the different Higgs production modes are given in Table 10.4. These results are derived from Monte Carlo simulations which are described in Section 11.1.1. Approximately 27% of the VBF signal is separated into the VBF-enriched category, leading to a purity of about 68% of the VBF signal in that category. The largest individual contributions from the ggF production mode enter the low $p_{\mathrm{T}t}$ categories. Together with the fact that ggF is by far the most abundant production mode, this results in purities of 93% for ggF in those categories. The event fractions of VH in the non-VBF categories range between 15 and 20%, leading to a slight accumulation of VH signal in the high relative p_{T}^{γ} category and, to a lesser extent, the high $p_{\mathrm{T}t}$ categories.

As can be seen from Table 10.5, the low $p_{\mathrm{T}t}$ categories contain the largest number of data events with almost 12000 events in the electron channel and about 16500 events in the muon channel in the $m_{Z\gamma}$ invariant mass range between 115 and 150 GeV. The VBF category is the least populated one and contains approximately 90 data events. The invariant mass resolution is rather similar for all categories. The half-width of the $m_{Z\gamma}$ mass window expected to contain 90% of the SM Higgs signal events for a Higgs boson mass of $m_H = 125.09$ GeV ranges between 3.7 and 3.9 GeV for all categories but the ee low $p_{\mathrm{T}t}$ category, for which it is expected to be 4.2 GeV. While the signal event yield in the mass window, S_{90}, is derived from simulated events, the number of background events in the mass window, B_{90}, is determined in a data-assisted way. Specifically, the number of events and the fraction of $Z\gamma$ fakes from Zj pairs were determined from data in the

	ggF			VBF			WH			ZH		
Category	$\xi_c^{(m)}$ [%]	$\epsilon^{(m)}\xi_c^{(m)}$ [%]	$\zeta_c^{(m)}$ [%]	$\xi_c^{(m)}$ [%]	$\epsilon^{(m)}\xi_c^{(m)}$ [%]	$\zeta_c^{(m)}$ [%]	$\xi_c^{(m)}$ [%]	$\epsilon^{(m)}\xi_c^{(m)}$ [%]	$\zeta_c^{(m)}$ [%]	$\xi_c^{(m)}$ [%]	$\epsilon^{(m)}\xi_c^{(m)}$ [%]	$\zeta_c^{(m)}$ [%]
1 / VBF-enriched	1.2	0.25	30.5	27.3	6.5	67.5	1.7	0.34	1.3	1.2	0.24	0.6
2 / high relative p_{T}^{γ}	5.1	1.1	71.5	10.9	2.6	14.3	20.3	4.0	8.3	19.3	4.1	5.3
3 / ee high $p_{\mathrm{T}t}$	8.1	1.7	80.8	11.7	2.8	11.0	16.0	3.2	4.7	16.9	3.6	3.3
4 / ee low $p_{\mathrm{T}t}$	32.8	7.1	93.2	15.3	3.6	4.1	18.4	3.7	1.5	19.8	4.2	1.1
5 / $\mu\mu$ high $p_{\mathrm{T}t}$	10.1	2.2	80.4	15.1	3.6	11.3	20.5	4.1	4.8	19.8	4.2	3.1
6 / $\mu\mu$ low $p_{\mathrm{T}t}$	42.6	9.2	93.4	19.7	4.7	4.1	19.8	4.6	1.5	23.0	4.8	1.0
Total efficiency $\epsilon^{(m)}$ [%]		21.5			23.8			20.2			21.0	
Expected events		35			3.3			1.0			0.7	

Table 10.4.: The expected signal event category fractions $\xi_c^{(m)}$ and event selection efficiencies times category fractions $\epsilon^{(m)}\xi_c^{(m)}$ per category c and production mode m as well as the expected relative contribution $\zeta_c^{(m)}$ of the different Higgs production modes to each category of the low-mass analysis. The expected total signal event selection efficiencies $\epsilon^{(m)}$ before the categorisation and the expected number of signal events are also given for each production mode m. The results are based on simulated SM Higgs boson events assuming $m_H = 125$ GeV and the corresponding Monte Carlo samples are described in Section 11.1.1. Taken and adapted from [60].

Category	Events	S_{90}	w_{90} [GeV]	S_{90}/B_{90} [%]	$S_{90}/\sqrt{S_{90}+B_{90}}$
1 / VBF-enriched	88	1.2	3.9	9.5	0.32
2 / high relative p_{T}^{γ}	443	2.3	3.9	3.0	0.26
3 / ee high $p_{\mathrm{T}t}$	1053	3.3	3.9	1.1	0.19
4 / ee low $p_{\mathrm{T}t}$	11707	11.2	4.2	0.3	0.18
5 / $\mu\mu$ high $p_{\mathrm{T}t}$	1413	4.0	3.7	1.2	0.22
6 / $\mu\mu$ low $p_{\mathrm{T}t}$	16529	14.5	3.8	0.3	0.21

Table 10.5.: The number of events in data in each category in the $m_{Z\gamma}$ invariant mass range between 115 and 150 GeV. The central mass window expected to contain 90% of the SM signal is indicated by the subscript 90. For this window, the expected SM Higgs signal yield S_{90} for a Higgs boson mass of $m_H = 125.09$ GeV, the half-width of the mass window w_{90}, the ratio of expected signal over background events S_{90}/B_{90} as well as the expected significance calculated as $S_{90}/\sqrt{S_{90}+B_{90}}$ are given. B_{90} is the number of background events in the mass window as determined in a hybrid approach using both data and simulated events. Taken from [60].

inclusive $m_{Z\gamma}$ mass region between 115 and 170 GeV. Using this information, the $Z+\gamma$ and $Z+$ jet Monte Carlo event samples were scaled and used to determine the number of background events in the 90% signal mass window. More information on the purity measurement will be given in Chapter 11. The signal-over-background ratios S_{90}/B_{90} and significances $S_{90}/\sqrt{S_{90}+B_{90}}$ are largest in the less populated categories and the VBF-enriched category is the most sensitive one.

In the search for high-mass resonances, the events are divided into two categories based on the flavour of the leptons coming from the Z boson decays. The motivations for this are the better $m_{Z\gamma}$ invariant mass resolution for $Z \to e^+e^-$ decays towards high resonance masses and differences in the systematic uncertainties. Specifically, the muon momentum is only determined from the track curvature, which is increasingly difficult towards higher p_{T} values arising from heavier resonances. In contrast, the energy of electrons is measured in the ECal.

Chapter 11

Signal and Background Modelling

The statistical analysis of the searches for the $Z\gamma$ decay of the SM Higgs boson and of high-mass resonances decaying to $Z\gamma$ is based on an unbinned likelihood formalism. This procedure necessitates the formulation of analytic models for the invariant mass distributions of the signal and background contributions. This chapter is dedicated to the derivation of these models. The description of the simulated event samples in Section 11.1 is followed by an introduction of the likelihood function in Section 11.2. Finally, the signal and background models are discussed in Sections 11.3 and 11.4.

11.1. Simulated Event Samples

The design and optimisation of an analysis should not involve data from regions expected to contain signal events. On the other hand, it is necessary to quantify how a hypothesised signal is expected to manifest itself, in order to set up an analysis and determine and maximise its sensitivity. Therefore, Monte Carlo simulations are often an indispensable tool. In this section, the simulated signal and background event samples used in the $H/X \to Z\gamma$ analysis are described. They are used for the optimisation of the analysis for maximum sensitivity, the evaluation of selection efficiencies and category compositions and the construction of signal and background models. All samples described in this section are made of fully exclusive Monte Carlo events generated with the steps outlined in Chapter 3 and subsequently passed through a simulation of the ATLAS detector with `Geant4`. Unless otherwise noted, the FullSim detector simulation is used. Beside the primary collision, additional inelastic pp interactions were added. These were simulated with `Pythia8` version 8.186 with the MSTW2008LO PDF set [206] and the A2 tune [207] and account for pile-up effects.

11.1.1. Signal Samples

Monte Carlo event samples for SM Higgs production with a mass of $m_H = 125$ GeV and a width of $\Gamma_H = 4.1$ MeV were generated for the ggF, VBF and VH production modes.

The total cross sections for these processes, which were published in [33] and are given in Table 2.1, are known with better accuracy at higher orders of the relevant couplings than available in the Monte Carlo generators. Therefore, constant factors were applied to reweight the samples to the corresponding higher-order total cross sections. Unless otherwise noted, the `PowhegBox` v2 generator [83–85] was used for all listed signal processes together with the PDF4LHC PDF set [208]. For ggF Higgs production, the NNLOPS approach [209] was used. It is based on the MiNLO method [210], which provides NLO accuracy for both the inclusive $H+0j$ as well as the $H+1j$ cross sections. On top of this, full NNLO accuracy for the inclusive process is achieved by means of a reweighting procedure, dependent on the rapidity y_H of the Higgs boson. The renormalisation and factorisation scales, μ_R and μ_F, were chosen as $\mu_R = \mu_F = m_H/2$ and finite top and bottom quark masses are accounted for up to NLO [211]. The events were reweighted to the total cross section of an N^{3}LO QCD and NLO EW calculation, which is described in more detail in Section 12.2.1. The VBF production mode was generated at NLO and reweighted to the total cross section calculated at approximate NNLO QCD and NLO EW. VH production was generated with the MiNLO approach, providing NLO accuracy for both the $H + 0j$ and $H + 1j$ multiplicities, and reweighted to the total VH cross section calculated at NNLO QCD and NLO EW. To all of these samples, parton showering, hadronisation and multiple parton interactions, simulated with `Pythia8` version 8.186 using the CTEQ6L1 PDF set [176] and the AZNLO tune [212], were added. Additional ggF samples were generated and used for the evaluation of specific theoretical uncertainties. These samples are described in Section 12.2.2.2.

For the high-mass analysis, ggF (VBF) samples were generated for a spin-0 resonance for 9 (4) mass points in the range 300 GeV $\leq m_X \leq$ 2.5 TeV and for a resonance width of $\Gamma_X = 4$ MeV. Less important than the exact value of the width is the fact, that it is much smaller than the experimental resolution on the invariant three-body mass $m_{\ell\ell\gamma}$. For these samples, the `PowhegBox` v1 event generator was used in conjunction with the CT10 PDF set [178]. The samples are not based on the NNLOPS approach, but rather a simpler NLO calculation of Higgs production including quark mass effects and electroweak corrections as described in [213]. The parton shower, hadronisation and underlying event simulations were added with `Pythia8` version 8.186 using the AZNLO tune and the CTEQ6L1 PDF set. Spin-2 resonances for 8 masspoints in the range 250 GeV $\leq m_X \leq$ 2.5 TeV and a width of $\Gamma_X = 4$ MeV were generated with `MadGraph5_aMC@NLO` 2.3.3 [82] for gluon-gluon and quark-antiquark initial states at LO using the Higgs Characterisation Model [214,215]. These hard process calculations were interfaced with `Pythia8` version 8.186 using the A14 tune [216] and the NNPDF2.3 PDF set [217].

11.1.2. Background Samples

There are two main sources for the background: Non-resonant $Z+\gamma$ production and Z+jet production, where a jet is misidentified as a photon.

For non-resonant $Z + \gamma$ production, two sets of samples were generated. The first sample was generated with `Sherpa` version 2.1.1 and is based on the hard process $pp \to \ell^+\ell^-\gamma + 0/1/2/3j$ with $\ell = e, \mu$, where the partons are described at the LO and the matrix elements for the different partonic multiplicities are merged with the MEPS@LO strategy [218] using a merging scale of 20 GeV. The CT10 PDF set was used. The entire sample is made of 12 million events, generated separately in the ee and $\mu\mu$ channels and in different intervals of the transverse momentum of the photon so as to ensure a sufficient number of simulated events in the high p_{T}^{γ} regions. For this sample, the FullSim detector simulation was used.

The second sample comprises a total of 30 million events and was therefore based on the FastSim detector simulation. It was also generated with `Sherpa` version 2.1.1 and has the hard process $pp \to \ell^+\ell^-\gamma + 0/1/2j$ with $\ell = e, \mu$. The merging algorithm and the scale choices are the same as for the FullSim sample, but it was generated in intervals of the invariant mass $m_{\ell\ell\gamma}$ with 20 million events in 110 GeV $\leq m_{\ell\ell\gamma} \leq$ 175 GeV and 10 million events in 175 GeV $\leq m_{\ell\ell\gamma} \leq$ 6.5 TeV. This sample was used for evaluating the functional form of the background model, which will be described in Section 11.4.

The Z+jets process was simulation with `Sherpa` 2.2.0 using the hard process $pp \to \ell^+\ell^- + 0/1/2/3/4j$ with $\ell = e, \mu$. The MEPS@NLO merging algorithm [219] was used, the two leading partons are described at NLO and the other two partons at the LO. The sample was generated separately in intervals of the transverse momentum of the dilepton system and for light-, charm- and bottom-flavoured hadrons. The NNPDF3.0 PDF set was used.

11.2. Likelihood Function

In order to illustrate how the signal and background modelling described in the following sections enters the analysis and to motivate the approach for evaluating the theoretical uncertainties as described in Section 12.2, it is instructive to first describe the likelihood function used in this analysis.

An extended unbinned likelihood procedure was used and the likelihood function is given by

$$L\left(\mu, \boldsymbol{\nu} | \left\{m_{Z\gamma}^{(i)}\right\}_{i=1}^{n}\right) = \frac{1}{n!} e^{-N(\mu,\boldsymbol{\nu})} N^n(\mu, \boldsymbol{\nu}) \times \left(\prod_{i=1}^{n} f_{\text{tot}}\left(m_{Z\gamma}^{(i)}, \mu, \boldsymbol{\nu}\right)\right) \times G(\boldsymbol{\nu}) \ , \tag{11.1}$$

where the signal strength μ is the parameter of interest, $\boldsymbol{\nu}$ denotes the nuisance parameters and $m_{Z\gamma}^{(i)}$ the invariant three-body mass of the Higgs boson candidate as reconstructed from the lepton and photon final state momenta. The first factor in Equation (11.1) is the Poissonian probability that relates the expected number of events N to the actually observed number of events n. f_{tot} describes the combined signal and background PDF as a function of $m_{Z\gamma}$ and $G(\boldsymbol{\nu})$ the penalty terms that constrain the nuisance parameters. f_{tot} is given by

$$\begin{aligned} &f_{\text{tot}}\left(m_{Z\gamma}^{(i)}, \mu, \boldsymbol{\nu}\right) \\ &= \frac{1}{N} \sum_{c} \left[\left(N_{\text{sig}}^{(c)}(\mu, \boldsymbol{\nu}) + N_{\text{spur}}^{(c)} \cdot \nu_{\text{spur}}^{(c)}\right) \times f_{\text{sig}}^{(c)}\left(m_{Z\gamma}^{(i)}, \boldsymbol{\nu}\right) + N_{\text{bkg}}^{(c)} \times f_{\text{bkg}}^{(c)}\left(m_{Z\gamma}^{(i)}, \boldsymbol{\nu}\right)\right] \ , \end{aligned} \tag{11.2}$$

where the sum with index c runs over the different categories of the analysis. $f_{\text{sig}}^{(c)}$ and $f_{\text{bkg}}^{(c)}$ are the signal and background PDFs and $N_{\text{sig}}^{(c)}$, $N_{\text{spur}}^{(c)}$ and $N_{\text{bkg}}^{(c)}$ the expected number of signal, spurious signal and background events for category c. The spurious signal refers to an apparent signal contribution induced by any potential bias from choosing a specific background model. This will be discussed in more detail in Section 11.4. $\nu_{\text{spur}}^{(c)}$ are the nuisance parameters related to the spurious signal in the different categories. The expected number of signal events can be written in further detail as

$$\begin{aligned} N_{\text{sig}}^{(c)}(\mu, \boldsymbol{\nu}) &= \mu \cdot \mathcal{L} \cdot \sigma_{\text{SM}}(pp \to H) \cdot \mathcal{B}_{\text{SM}}(H \to Z\gamma) \cdot \mathcal{B}_{\text{SM}}(Z \to \ell\ell) \cdot \epsilon \cdot \xi_c(\boldsymbol{\nu}) \\ &= \sum_{m} \mu \cdot \mathcal{L} \cdot \sigma_{\text{SM}}^{(m)}(pp \to H) \cdot \mathcal{B}_{\text{SM}}(H \to Z\gamma) \cdot \mathcal{B}_{\text{SM}}(Z \to \ell\ell) \cdot \epsilon^{(m)} \cdot \xi_c^{(m)}(\boldsymbol{\nu}) \ , \end{aligned} \tag{11.3}$$

where $\mathcal{L}$ denotes the integrated luminosity, ϵ the detector acceptance and event selection efficiency and ξ_c the relative event yield in category c. By construction $\sum_c \xi_c = 1$. $\sigma_{\mathrm{SM}}(pp \to H)$ denotes the SM Higgs production cross section for all production modes considered, $\mathcal{B}_{\mathrm{SM}}(H \to Z\gamma)$ the SM branching ratio of the $H \to Z\gamma$ decay and $\mathcal{B}_{\mathrm{SM}}(Z \to \ell\ell)$ the branching ratio of the Z boson decaying to electrons and muons. The index m runs over the different Higgs production modes.

Besides the signal strength μ, as illustrated in the notation above, the Higgs production cross section times branching ratio, $\sigma(pp \to H) \cdot \mathcal{B}(H \to Z\gamma)$, and the branching ratio $\mathcal{B}(H \to Z\gamma)$ are used individually as alternative parameters of interest in the search for the $Z\gamma$ decay of the Higgs boson. In the search for high-mass resonances, $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ was chosen as the parameter of interest. More details on the statistical procedures used to derive the results of the analysis will be presented later on in Section 13.1.

11.3. Signal Modelling

The signal model comprises the signal PDF as a function of the three-body invariant mass $m_{\ell\ell\gamma}$ and the cross sections, the branching ratios, the detector acceptance and event selection efficiencies as well as the relative event yields in the different categories appearing in Equation (11.3). The cross sections are given in Table 2.1, the branching ratios in Chapter 9 and the event selection efficiencies $\epsilon^{(m)}$ and category fractions $\xi_c^{(m)}$ in Table 10.4 for the low-mass region and in Figure 10.1 for the high-mass region. For the latter, the event selection efficiencies were parameterised with an exponentiated second-order polynomial. Tables with the event selection efficiencies for different production modes and mass points are given in Appendix B.1.

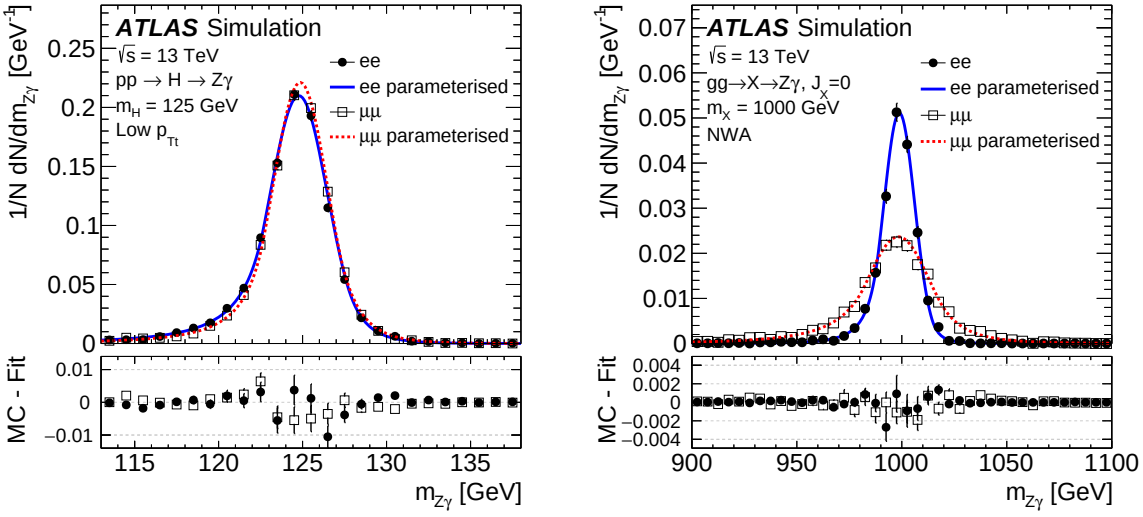
Just like the event selection efficiencies and the category fractions, also the signal PDFs $f_{\mathrm{sig}}^{(c)}$ for each category c were derived from the signal Monte Carlo event samples. For both the search for the $Z\gamma$ decay mode of the Higgs boson as well as for the high-mass resonance search, the Double-Sided Crystal Ball (DSCB) [220, 221] function was chosen. The DSCB function consists of a Gaussian core and power-law tails at either side. Up to an overall normalisation factor, its analytic form is given by

$$\begin{cases} \frac{\exp\left(-\frac{\alpha_l^2}{2}\right)}{\left[\frac{\alpha_l}{n_l}\left(\frac{n_l}{\alpha_l}-\alpha_l-t\right)\right]^{n_l}} & \text{for } t < -\alpha_l \\ \exp\left(-\frac{t^2}{2}\right) & \text{for } -\alpha_l \leq t \leq \alpha_h \\ \frac{\exp\left(-\frac{\alpha_h^2}{2}\right)}{\left[\frac{\alpha_h}{n_h}\left(\frac{n_h}{\alpha_h}-\alpha_h+t\right)\right]^{n_l}} & \text{for } t > \alpha_h \end{cases} \tag{11.4}$$

with

$$t = \frac{m_{\ell\ell\gamma} - m_{H/X} - \mu_{\mathrm{CB}}}{\sigma_{\mathrm{CB}}} \,,$$

where μ_{CB} and σ_{CB} correspond to the peak position and width of the central Gaussian, $-\alpha_l$ and α_h describe the transition points between the different piecewise functional forms and n_l and n_h give the exponents of the left and right power-law tails. For the low-mass region, a signal model is built for the combined ggF, VBF and VH signal separately in each category. Since the corresponding Monte Carlo samples were generated with a



(a) low-mass region.

(b) high-mass region.

Figure 11.1.: Example signal models for the low-$p_{\mathrm{T}t}$ categories of the low-mass region and for the $m_X = 1$ TeV mass point of the high-mass region in both the electron and muon channels. The DSCB function is used for all signal models and the models are built from fits to the corresponding signal Monte Carlo event samples. The Monte Carlo data points are indicated by the markers and the analytic models by the curves. The lower parts show the differences between the Monte Carlo data and the analytic models. Taken from [60].

Higgs boson mass of $m_H = 125$ GeV, a shift of 90 MeV was manually applied to the peak position parameter μ_{CB}. For the high-mass region, simultaneous fits were done for all mass points for which signal Monte Carlo samples were available. More specifically, the DSCB parameters themselves were parameterised as first- or second-order polynomial functions of the resonance mass and the corresponding coefficients were adjusted in the fits. This way, analytic signal models are available for the entire invariant mass search range. Separate signal models were built for ggF-induced spin-0 and the gluon-gluon- and quark-antiquark-initiated spin-2 processes. As an example, the signal models for the low-$p_{\mathrm{T}t}$ categories of the low-mass region and for the $m_X = 1$ TeV mass point of the high-mass region are shown in Figure 11.1 in both the electron and muon channels. As can be seen from the corresponding residual plots, the signal models are compatible with the Monte Carlo data points within the statistical uncertainties in most mass bins. In the search for the $Z\gamma$ decay mode of the SM Higgs boson, the invariant mass resolutions, as given by the σ_{CB} parameter, range between 1.66 and 1.86 GeV. In the search for narrow high-mass resonances, the invariant mass resolution in the electron channel is better than in the muon channel and this difference in resolution increases towards higher resonance masses. This is due to the fact that determining the muon momentum from the curvature of the muon tracks becomes less precise for increasing momenta, whereas the electron energy measurement is accomplished with the calorimeter and therefore not impaired by decreasing curvatures of high-momentum tracks.

11.4. Background Modelling

In both mass regions, the searches look for a localised excess of events above a smoothly falling background distribution, which is made of two relevant contributions: An irreducible component from non-resonant $Z\gamma$ production and a reducible component from Zj production, where a jet is misidentified as a photon.

The composition of the background, i.e. the relative size of both contributions, was determined with the so-called *isolation template fit* method. It relies on the calorimeter isolation transverse energy $E_{\mathrm{T}}^{\mathrm{iso}}$, introduced in Section 5.3.2, being shaped differently for prompt photons and jets faking photons. A cone size of $R = 0.2$ is used. After correcting for leakage and effects from the underlying event and pile-up, the isolation transverse energy distribution is nominally independent of the photon candidate E_{T} [222]. A control region (CR) is created by requiring that photon candidates fail the *Tight* identification requirements used for the signal region (SR), but pass a modified set of loose ID requirements. These modified loose ID requirements, referred to as the *LoosePrime4* working point, are formed by removing from the *Tight* photon ID working point the requirements on those four shower shape variables that are the least correlated with the isolation energy [222]. This aims at keeping the $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions as similar as possible in shape between the signal and control regions. In a first step, the $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions of prompt photons in both the signal and the control region are constructed using photons from the $Z\gamma$ background Monte Carlo sample. The shapes of both distributions are encoded by the signal and leakage factors f_i and l_i, where the index i runs over bins in $E_{\mathrm{T}}^{\mathrm{iso}}$. The signal factors f_i describe the shape of the photon component in the signal region, while the product $f_i l_i$ of the signal and leakage factors gives both the shape of the photon component in the control region as well as its relative normalisation with respect to the signal region. These factors are used in a second step, in which a simultaneous binned fit to the data $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions in both the signal and control region, $b_i^{(\mathrm{SR})}$ and $b_i^{(\mathrm{CR})}$, is performed to determine the following free parameters:

- the overall photon yield in the signal region, $B_{Z\gamma}^{(\mathrm{SR})}$,
- the overall yield of photon fakes from jets in the signal region, $B_{Zj}^{(\mathrm{SR})}$ and
- the number of events in each bin i of the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution for the jet component in the control region is given by $B_{i,Zj}^{(CR)}$.

Assuming that the $E_{\mathrm{T}}^{\mathrm{iso}}$ shapes of the Zj background component are identical in the signal and control regions, the numbers of events are given by the relations

$$b_i^{(\mathrm{CR})} = f_i l_i B_{Z\gamma}^{(\mathrm{SR})} + B_{i,Zj}^{(\mathrm{CR})} \tag{11.5}$$

$$b_i^{(\mathrm{SR})} = f_i B_{Z\gamma}^{(\mathrm{SR})} + \frac{B_{i,Zj}^{(\mathrm{CR})}}{\sum_k B_{k,Zj}^{(\mathrm{CR})}} B_{Zj}^{(\mathrm{SR})} \ . \tag{11.6}$$

This procedure is sketched in Figure 11.2. Based on these results, the background composition ϱ was found to be

$$\varrho = \frac{B_{Z\gamma}^{(\mathrm{SR})}}{B_{Z\gamma}^{(\mathrm{SR})} + B_{Zj}^{(\mathrm{SR})}} = \begin{cases} 0.838 \pm 0.005(\mathrm{stat.}) \pm 0.031(\mathrm{syst.}) & \text{in the low-mass region} \\ 0.916 \pm 0.009(\mathrm{stat.})^{+0.013}_{-0.019}(\mathrm{syst.}) & \text{in the high-mass region} \ . \end{cases} \tag{11.7}$$

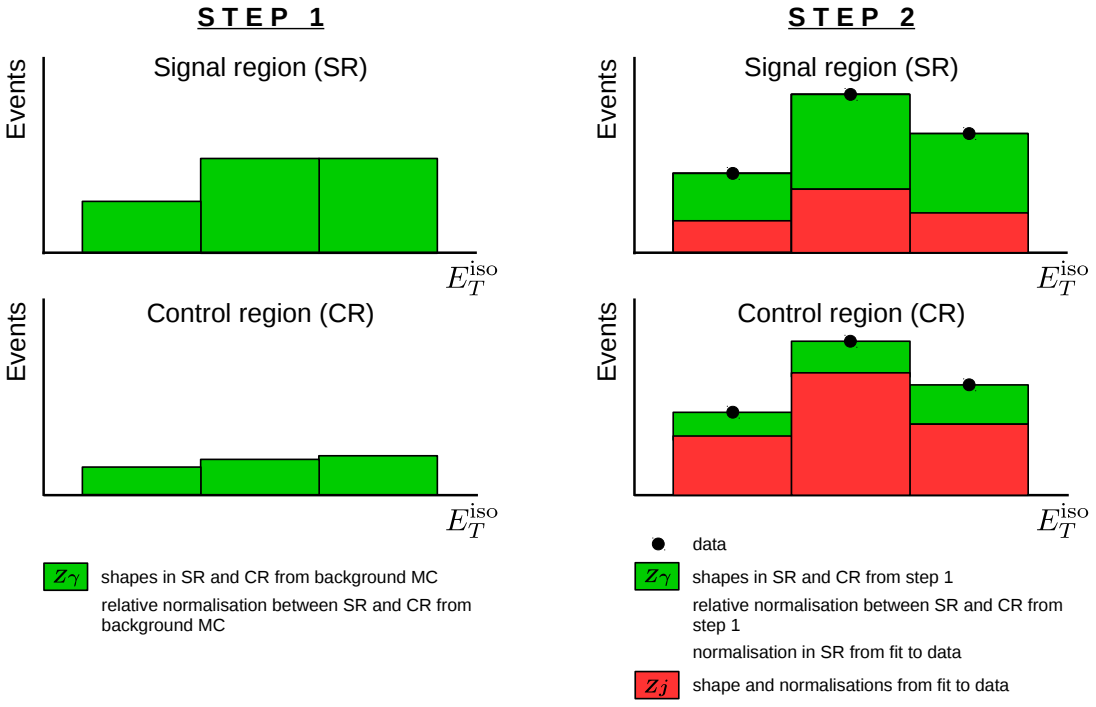


Figure 11.2.: The isolation template fit method proceeds in two steps: First, the shapes of the $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions of the $Z\gamma$ background component in the signal and control regions and their relative normalisation are determined from a fit to the $Z\gamma$ background Monte Carlo event sample. In the second step, the remaining parameters, describing the Zj shape and normalisations as well as the $Z\gamma$ normalisation in the signal region, are determined from a binned fit to the $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions in the signal and control regions. The background composition ϱ is then calculated from the total $Z\gamma$ and Zj event yields in the signal region.

The systematic uncertainties are based on variations of the modified loose photon ID requirements. Specifically, different sets of shower shape requirements are removed to form these alternative modified loose photon ID working points, removing a range of two to all five shower shape variables calculated from energy depositions in the ECal strip layer. These results for the background compositions in both mass regions were cross-checked and confirmed with the two-dimensional sideband method, based on control regions defined from combinations of nominal and inverted photon identification and isolation requirements [222].

The background compositions in the low- and high-mass regions are used to construct a Monte Carlo template for the determination of the functional form of the background model. In order to avoid restrictions due to the limited size of the Zj Monte Carlo sample, a hybrid approach was taken. By inverting the photon identification and isolation requirements, control regions enriched in Zj data events were created for each category. The ratio $R_{c,i}$ of the invariant mass distribution in this control region to the distribution of the FastSim $Z\gamma$ Monte Carlo events in the signal region was evaluated, i.e. $R_{c,i} = b^{(Zj)}_{c,i,\mathrm{CR}}/B^{(Z\gamma)}_{c,i,\mathrm{SR}}$, where the index c runs over the different categories and the index i over bins of the invariant mass $m_{\ell\ell\gamma}$ within a given category. b and B indicate the number of events in data and Monte Carlo, respectively. In each category, the ratio values $R_{c,i}$ were fitted with first-order polynomials and combined with the overall background composition ϱ to determine the expected number of Zj events in each category c and bin i of the $m_{\ell\ell\gamma}$ distribution in the signal region. This can be written as

$$B^{(Zj)}_{c,i,\mathrm{SR}} = \alpha R^{(\mathrm{fit})}_{c,i} B^{(Z\gamma)}_{c,i,\mathrm{SR}} \quad \text{and} \quad \varrho \stackrel{!}{=} \frac{\sum_{c,i} B^{(Z\gamma)}_{c,i,\mathrm{SR}}}{\sum_{c,i} B^{(Z\gamma)}_{c,i,\mathrm{SR}} + B^{(Zj)}_{c,i,\mathrm{SR}}} = \frac{\sum_{c,i} B^{(Z\gamma)}_{c,i,\mathrm{SR}}}{\sum_{c,i} B^{(Z\gamma)}_{c,i,\mathrm{SR}} \left(1 + \alpha R^{(\mathrm{fit})}_{c,i}\right)} , \tag{11.8}$$

from which α is determined so as to absorb the arbitrary normalisation of the Zj event sample in the control region and ensure the appropriate background composition ϱ in the signal regions. The invariant mass distribution obtained from the $Z\gamma$ background Monte Carlo in the signal region is then scaled accordingly by $1 + \alpha R^{(\mathrm{fit})}_{c,i}$ and the resulting hybrid expected background invariant mass distributions were then used in the determination of the functional form of the background models in the different mass regions and categories.

The functional forms of the background models are determined from studies of the signal bias, also referred to as the *spurious signal*, which describes the apparent signal yield resulting from a combined $S+B$-fit to the B-only invariant mass distributions constructed as described above. In other words, the spurious signal is caused by a specific choice for the background model, but affects the fitted signal yield. Therefore, the product of the expected number of spurious signal events $N^{(c)}_{\mathrm{spur}}$ and the corresponding nuisance parameter $\nu^{(c)}_{\mathrm{spur}}$ is added to the number of signal events in a given category c in the likelihood function, as can be seen from Equation (11.2). The aforementioned $S + B$ fits were done for different hypothesised values of m_H between 121 GeV and 129 GeV and the largest spurious signal yield from that range was used as an estimate of the potential signal bias. This procedure serves as a safeguard against accidental underestimation of the spurious signal yield at a single masspoint due to statistical fluctuations. In the search for the $Z\gamma$ decay of the Higgs boson, also different fit ranges are studied so as to select the option with the smallest signal bias. In order for a functional form to be eligible for the background parameterisation, the resulting spurious signal yield should be less than 40% of the expected statistical uncertainty in the signal yield in the low-mass region and less than 20% in the high-mass region. In case multiple functional forms are eligible in a

given category, the one with the fewest degrees of freedom is selected. As a result of these studies, the fit range in the search for the $Z\gamma$ decays of the Higgs boson is determined to be 115 GeV $\leq m_{\ell\ell\gamma} \leq$ 150 GeV. For the VBF-enriched and high relative p_{T} categories, i.e. categories 1 and 2, second-order Bernstein polynomials are chosen and fourth-order Bernstein functions for the remaining four categories. The spurious signal yield ranges between 1.7 events in the VBF-enriched category and 25 events in the low $p_{\mathrm{T}t}$ muon category. In the search for high-mass resonances, functions of the type

$$N(1 - x^{1/3})^b x^{\sum_{j=0}^{k} a_k \ln^j(x)} \tag{11.9}$$

were chosen with $k = 0$ in both the electron and the muon category. N is the normalisation factor, $x = m_{Z\gamma}/\sqrt{s}$, k denotes the order of the exponent and b and a_k are free parameters. The resulting spurious signal yield ranges between 3.6 events at 250 GeV and 0.01 events at 2.4 TeV in the muon channel and 6.1 and 0.005 events in the electron channel.

Chapter 12

Systematic Uncertainties

In this chapter, the relevant uncertainties in the $H/X \to Z\gamma$ analyses are discussed. They are related to quantities entering the likelihood function described in Section 11.2. The experimental uncertainties, discussed in Section 12.1, can be divided into two groups: On the one hand uncertainties affecting the signal and background modelling, e.g. uncertainties from energy and momentum scales, and on the other hand uncertainties affecting the expected event yield, e.g. uncertainties from trigger and reconstruction efficiencies. Similarly, the theoretical uncertainties, presented in Section 12.2, can be split into uncertainties affecting the expected event yield, e.g. uncertainties on the signal cross sections and the branching ratio, and uncertainties related to how the events are distributed between the different categories.

12.1. Experimental Uncertainties

The first group of experimental systematic uncertainties is related to the signal and background models. The assumed SM Higgs boson mass of $m_H = 125.09$ GeV has an uncertainty of 0.24 GeV [35], corresponding to a 0.2% uncertainty on μ_{CB} in the $H \to Z\gamma$ search. The uncertainties on the photon and electron energy scales and the muon momentum scale affect the reconstructed invariant mass $m_{Z\gamma}$ and therefore also the peak position μ_{CB} of the DSCB signal modelling function. The width σ_{CB} of the Gaussian core of the DSCB function is influenced by both the energy and momentum scales and resolutions for the photons and leptons. These uncertainties are evaluated by varying the scales and resolutions by their uncertainties and refitting the resulting signal $m_{Z\gamma}$ distributions in each category. The modified peak position and width parameters are then compared to the parameters obtained in the nominal case. The uncertainties on μ_{CB} due to the energy and momentum scales were found to be smaller than 0.6%. The uncertainties on the peak width σ_{CB} are mostly smaller than 4%, with only the uncertainty from the electron and photon energy resolution reaching up to 30% for $m_{Z\gamma} > 1$ TeV in the electron channel of the high-mass region. The fact that the relative uncertainty is larger in the electron

channel is due to the smaller nominal σ_{CB} in the $e^+e^-\gamma$ compared to the $\mu^+\mu^-\gamma$ final state. Details on the determination of the uncertainties on the muon momentum scale and resolution can be found in [121] and information on the energy scale and resolution uncertainties for electrons and photons in [117]. As detailed in Section 11.4, the choice of a specific functional form for the background model can potentially induce an apparent signal even when fitting background-only invariant mass distributions. These spurious signal event yields in the different categories are multiplied with their respective nuisance parameters and added to the expected signal event yields as indicated in Equation (11.2).

The second group of experimental systematic uncertainties affect the event yield. They can be further classified into *yield uncertainties*, which are related to the overall selected event yield, and *migration uncertainties*, which are related to the distribution of the events among the different categories. In the latter case, the uncertainties are anticorrelated between the different affected categories, otherwise they are correlated. Besides the uncertainty of 3.2% on the integrated luminosity and uncertainties on the signal efficiency from the ratio of the predicted and measured inelastic cross sections of 0.03% (0.2%) in the low-mass (high-mass) region, a number of uncertainties on the signal efficiency due to aspects of the lepton and photon reconstruction and selection have to be considered. The trigger, reconstruction, identification and isolation efficiencies are estimated from Monte Carlo simulations, and corrections, so-called *scale factors*, accounting for differences between efficiencies estimated from simulations and efficiencies measured in data are included. The scale factors are studied using $J/\psi \to \ell^+\ell^-$, $Z \to \ell^+\ell^-$ and $Z \to \ell^+\ell^-\gamma$ decays as well as inclusive photon samples. The methodologies, results and uncertainties are discussed in [121] for muons, [141] for electrons and [137] for photons. In the search for the $Z\gamma$ decay mode of the Higgs boson and in the high-mass resonance search, most of the corresponding uncertainties are at the level of no more than 2%, with only uncertainties due to muon trigger efficiencies and electron identification and isolation efficiencies reaching up to approximately 4% in some of the categories. All of the aforementioned uncertainties are taken to be correlated between all categories. The definition of the VBF-enriched category in the $H \to Z\gamma$ search is based on the presence of at least two jets and their kinematic properties. Therefore, uncertainties on the jet energy scale and resolution [136] and on the jet vertex tagging, which is used for pile-up suppression, have to be taken into account. Since they are not related to the total selected event yield, but only to how the events are distributed between the VBF-enriched and the remaining categories, the jet-related uncertainties must be anticorrelated between these categories. They reach up to 10 % in the VBF-enriched category and are much smaller in the other categories of the $H \to Z\gamma$ search.

A list of the relevant experimental systematic uncertainties and their respective sizes is given in Table 12.1.

12.2. Theoretical and Modelling Uncertainties

The theoretical expectations and predictions are subject to uncertainties that also need to be included in the analysis. Since the parameters of interest of the analysis are the Higgs production signal strength, cross section and branching ratio, theoretical uncertainties related to properties of the Higgs boson were considered. As can be seen from Equation (11.3), the relevant quantities are

- the Higgs production cross section $\sigma_{\mathrm{SM}}^{(m)}(pp \to H)$ in each production mode m,

Sources	$H \to Z\gamma$	$X \to Z\gamma$
Uncertainties related to the luminosity [%]		
Luminosity	3.2	3.2
Uncertainties related to the signal efficiency [%]		
Modelling of pile-up interactions	0.02–0.03	< 0.01–0.2
Photon identification efficiency	0.7–1.7	2.0–2.6
Photon isolation efficiency	0.07–0.4	0.6–0.6
Electron identification efficiency	0.0–1.6	0.0–2.6
Electron isolation efficiency	0.0–0.2	0.0–3.5
Electron reconstruction efficiency	0.0–0.4	0.0–1.0
Electron trigger efficiency	0.0–0.1	0.0–0.2
Muon selection efficiency	0.0–1.6	0.0–0.7
Muon trigger efficiency	0.0–3.5	0.0–4.2
MC statistical uncertainty	–	1.2–2.0
Jet energy scale, resolution, and pile-up	0.2–10	–
Total (signal efficiency)	2.1–10	4.0–6.3
Uncertainties related to the signal modelling - σ_{CB} [%]		
Electron and photon energy scale	0.6–3.5	1.0–4.0
Electron and photon energy resolution	1.1–4.0	4.0–30
Muon momentum scale	0.0–0.5	0.0–3.0
Muon ID resolution	0.0–3.7	0.0–2.0
Muon MS resolution	0.0–1.7	0.0–4.0
Uncertainties related to the signal modelling - μ_{CB} [%]		
Electron and photon energy scale	0.1–0.2	0.2–0.6
Muon momentum scale	0.0–0.03	0.0–0.03
Higgs mass	0.2	–
Uncertainties related to the background modelling [events]		
Spurious signal	1.7–25	0.005–6.1

Table 12.1.: Relevant experimental uncertainties for the $H/X \to Z\gamma$ analyses. The uncertainty estimates are based on the ggF signal samples for $m_H = 125$ GeV and $m_X = [300\text{-}2500]$ GeV. The uncertainties are given as ranges across the different categories and m_X resonance masses. The sizes of the uncertainties are mostly given in relation to the nominal values of the quantities they are affecting. Only the spurious signal yields are given as event numbers. Taken and adapted from [60].

- the Higgs decay branching ratio $\mathcal{B}_{\mathrm{SM}}(H \to Z\gamma)$,
- the detector acceptance and event selection efficiency ϵ and
- the event fractions per category ξ_c.

Uncertainties due to the truncation of the perturbative series, modelling uncertainties, uncertainties from the PDFs and the strong coupling α_s as well as uncertainties for the modelling of the underlying event were taken into account and will be described in this section.

12.2.1. Uncertainties on the Production Cross Section, Branching Ratio and Acceptance

The cross sections of the different SM Higgs boson production modes are given in Table 2.1. For the uncertainties on the production cross sections and the branching ratio for the $H \to Z\gamma$ channel, the recommendations given by the LHC Higgs Cross Section Working Group (HXSWG) as published in [33] were adopted. Two types of uncertainties are taken into account for the production cross sections: theoretical uncertainties, encoding limitations of the methods and calculations used, and parametric uncertainties, describing incomplete knowledge of the exact numerical values of the relevant parameters. For all production modes, the parametric uncertainties are given by the combined PDF+α_s uncertainties. The exact meaning and definition of the different theoretical uncertainties depend on the details of the calculations performed for the various production modes. In general, the theoretical uncertainties encode the lack of knowledge of the fully exact perturbative expansions. A common component of the theoretical uncertainties for all production modes are factorisation and renormalisation scale variations. The relevant accuracies and features of the various calculations are as follows:

- **ggF:** N^{3}LO QCD in effective theory ($m_t \to \infty$), heavy quark mass effects at NLO QCD for the t, b and c-quarks and subleading top quark mass effects at NNLO QCD, NLO EW and approximate mixed QCD/EW corrections
- **VBF:** approximate NNLO QCD with NLO EW corrections
- **VH:** $q\bar{q} \to VH$ at NNLO QCD with NLO EW corrections, $gg \to ZH$ at NLO QCD

All details of these calculations as well as the corresponding uncertainty prescriptions can be found in [33]. The theoretical and the combined PDF+α_s uncertainties provided by the LHC HXSWG for a centre-of-mass energy of $\sqrt{s} = 13$ TeV and a Higgs boson mass of $m_H = 125.09$ GeV for the different production mechanisms as given in [223, 224] are listed in Table 12.2. These values were used in the analysis.

Similar considerations apply to the uncertainties on the Higgs branching ratio for the $H \to Z\gamma$ decay channel. An additional parametric uncertainty related to the values of the quark masses is taken into account. The uncertainty values provided by the LHC HXSWG in [33, 225] are given in Table 12.3.

Finally, an uncertainty on the detector acceptance and event selection efficiency of 5.3% was derived for the modelling of the underlying event. The details are given in Section 12.2.2.2, where the corresponding uncertainties on the event fractions for the different categories are given.

Theoretical uncertainties on the Higgs production cross sections				
Production mechanism		THU		PDF+α_s
Gluon-gluon fusion	ggF	+3.9%	−3.9%	±3.2%
Vector boson fusion	VBF	+0.4%	−0.3%	±2.1%
W-associated production	WH	+0.5%	−0.7%	±1.9%
Z-associated production	ZH	+3.8%	−3.0%	±1.6%

Table 12.2.: Theoretical uncertainties (THU) and PDF+α_s uncertainties for the different Higgs production mechanisms for a Higgs boson mass of $m_H = 125.09$ GeV as given in [33, 223, 224].

Theoretical uncertainties on the $H \to Z\gamma$ decay						
$\mathcal{B}(H \to Z\gamma)$	THU		PU(m_Q)		PU(α_s)	
$1.541 \cdot 10^{-3}$	+5.71%	−5.71%	+0.91%	−1.00%	+0.58%	−0.64%

Table 12.3.: Uncertainties on the branching ratio for the $H \to Z\gamma$ decay for a Higgs boson mass of $m_H = 125.09$ GeV as given in [33, 225]. The listed components are theoretical uncertainties (THU) from missing higher orders in the calculations which are estimated from scale variations and parametric uncertainties (PU) that express the inexact knowledge of the charm, bottom and top quark masses and of the strong coupling α_s.

12.2.2. Uncertainties on the Event Fractions Per Category

12.2.2.1. Uncertainties from Perturbation Theory

Uncertainties arising from the truncation of the perturbation series at a certain order are in general evaluated by variations of the factorisation and renormalisation scales. Such variations lead to a modification of the event weights with respect to the nominal scale choices. Ideally, these modified event weights are included in the Monte Carlo samples, in which case they can be used to obtain the variations of differential distributions and event yields in the different analysis categories. In the version of the analysis described in this part of the thesis, such event weights from scale variations were not available for the signal Monte Carlo samples. Therefore, a different procedure had to be adopted which will be introduced and described in the following.

An additional motivation for the adopted procedure is that the order and formal accuracy of the calculation underlying the Monte Carlo events was lower than that of state of the art fixed-order calculations at that time. Yet, fully exclusive Monte Carlo events were needed as well and the approach taken aims at combining the Monte Carlo events with information from more accurate fixed-order calculations. Finally, it is known that pure scale variations can underestimate the perturbative uncertainties in exclusive regions of phase space defined by cuts on jet-related observables. For all these reasons, a modification of the Stewart-Tackmann (ST) method [226,227] was used in this analysis. In the following, the fundamentals of the ST procedure will be presented, followed by a generalisation for more than two jet bins and the extension and application to the $H \to Z\gamma$ analysis.

The Stewart-Tackmann Procedure

The ST method aims at calculating robust and realistic perturbative uncertainties in an analysis with categories based on the number of jets, the so-called *jet bins*, or, more generally, categories sensitive to the presence and hardness of jets. In the following, the basic structure of the ST method is sketched for the example case of two bins. A generalisation for more bins will be presented in the next section.

In order to enhance the sensitivity of an analysis, the events are often divided into different, mutually exclusive categories. Among other aspects, this entails the necessity of assigning proper theoretical uncertainties for each of the categories, to consider possible correlations and to ensure that the proper total uncertainty is restored when combining the different categories. The ST procedure is designed to address these questions.

Consider two jet bins which are defined by a cut $p_{\mathrm{T}}^{\mathrm{cut}}$ on the transverse momentum of a jet. The total inclusive cross section is the cross section for having at least zero jets, i.e. $\sigma_{\mathrm{tot}} = \sigma_{\geq 0}$. Events with no jet above the threshold $p_{\mathrm{T}}^{\mathrm{cut}}$ are classified into the *exclusive* 0-jet bin, while events that have at least one jet with a transverse momentum larger than $p_{\mathrm{T}}^{\mathrm{cut}}$ are classified into the *inclusive* 1-jet bin. The corresponding cross sections are $\sigma_{=0}$ and $\sigma_{\geq 1}$, respectively. The term *inclusive* means that the required number of jets is not bounded, i.e. an inclusive 1-jet category may contain events with two or even more jets. Correspondingly, the term *exclusive* encodes that additional jet activity is rejected or vetoed and the category just and only contains events of the given jet multiplicity. Note that the method also works in principle for categories of higher jet multiplicity and categories defined by other cuts that are related to jet activity. The case presented here is simply chosen as an example. Since both categories defined above are mutually exclusive and complete, the total cross section is given by the sum of the cross sections of both categories. It is worth noting that the cross sections in both categories will depend on the jet cut $p_{\mathrm{T}}^{\mathrm{cut}}$, which is not the case for the total inclusive cross section, i.e. $\sigma_{\mathrm{tot}} = \sigma_{=0}(p_{\mathrm{T}}^{\mathrm{cut}}) + \sigma_{\geq 1}(p_{\mathrm{T}}^{\mathrm{cut}})$. This motivates the *migration uncertainty* Δ_m, which is anti-correlated between both categories and which encodes that modifying the jet cut would make events shift between the two categories. The remaining correlated *yield uncertainties* are given by Δ_y.

A fully general parameterisation for a 2×2 covariance matrix, that explicitly shows the correlated and anti-correlated contributions, is given by

$$C = \begin{pmatrix} \left(\Delta^y_{=0}\right)^2 & \Delta^y_{=0}\Delta^y_{\geq 1} \\ \Delta^y_{=0}\Delta^y_{\geq 1} & \left(\Delta^y_{\geq 1}\right)^2 \end{pmatrix} + \begin{pmatrix} \Delta_m^2 & -\Delta_m^2 \\ -\Delta_m^2 & \Delta_m^2 \end{pmatrix} . \tag{12.1}$$

This specific form is motivated by the fact that it encodes the desired behaviour, i.e. it reproduces the correct per-category uncertainties as well as the correct total uncertainty:

$$\begin{aligned} \Delta^2_{=0} &= \begin{pmatrix} 1 & 0 \end{pmatrix} C \begin{pmatrix} 1 \\ 0 \end{pmatrix} = {\Delta^y_{=0}}^2 + \Delta_m^2 \\ \Delta^2_{\geq 1} &= \begin{pmatrix} 0 & 1 \end{pmatrix} C \begin{pmatrix} 0 \\ 1 \end{pmatrix} = {\Delta^y_{\geq 1}}^2 + \Delta_m^2 \\ \Delta^2_{\geq 0} &= \begin{pmatrix} 1 & 1 \end{pmatrix} C \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \left(\Delta^y_{=0} + \Delta^y_{\geq 1}\right)^2 . \end{aligned} \tag{12.2}$$

As can be seen, this parameterisation correctly reflects that the dependence on the jet cut vanishes when both categories are combined.

The three parameters $\Delta^y_{=0}$, $\Delta^y_{\geq 1}$ and Δ_m can be estimated from perturbative scale variations Δ_μ, i.e. calculations where the relevant factorisation and normalisation scales are typically varied by factors of 2 and 1/2 around the given nominal scale choices. Specifically, scale variations are used to estimate the uncertainties of the inclusive cross sections, i.e. $\Delta_{\geq 0} = \Delta^\mu_{\geq 0}$ and $\Delta_{\geq 1} = \Delta^\mu_{\geq 1}$. There are, however, only two independent cross section variations since the variation of the total inclusive cross section is given by the sum of the variations in both categories, i.e.

$$\Delta^\mu_{\geq 0} = \Delta^\mu_{=0} + \Delta^\mu_{\geq 1} \ . \tag{12.3}$$

This means there is some freedom of choice when building the covariance matrix C.

In the ST method the migration uncertainty Δ_m is kept and in the yield uncertainty $\Delta^y_{\geq 1}$ is set to zero. The result is

$$\Delta_{\geq 0} = \Delta^y_{=0} = \Delta^\mu_{\geq 0} \qquad \text{and} \qquad \Delta_{\geq 1} = \Delta_m = \Delta^\mu_{\geq 1} \tag{12.4}$$

and

$$\Delta^2_{=0} = \left(\Delta^y_{=0}\right)^2 + \Delta^2_m = \left(\Delta^\mu_{\geq 0}\right)^2 + \left(\Delta^\mu_{\geq 1}\right)^2 = \Delta^2_{\geq 0} + \Delta^2_{\geq 1} \ . \tag{12.5}$$

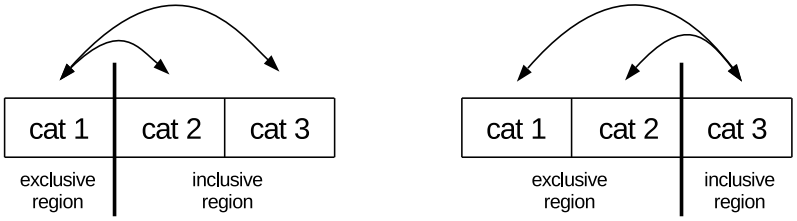
This choice is motivated by regarding the inclusive cross sections of different jet multiplicities, $\sigma_{\geq 0}$ and $\sigma_{\geq 1}$ as approximately uncorrelated. The reason for this is that the perturbative series of both quantities start at different orders in α_s and that they are not modified in a directly related manner by the jet cut. As a consequence, the uncertainty on the exclusive cross section $\sigma_{=0}$ is given by the sum in quadrature of the uncertainties of the inclusive categories, as can be seen in Equation (12.5). Similarly, often times the cross section $\sigma_{\geq 1}$ is dominated by the logarithms stemming from the underlying jet cuts and therefore the scale variations $\Delta^\mu_{\geq 1}$ will probe the logarithms associated with the definitions of the categories, corresponding to the migration uncertainty, which is reflected in Equation (12.4). In summary, the full decomposed covariance matrix is given by

$$C = \Delta^2_y \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \Delta^2_m \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix} \qquad \text{with} \qquad \begin{matrix} \Delta_y = \Delta^\mu_{\geq 0} \\ \Delta_m = \Delta^\mu_{\geq 1} \end{matrix} \ . \tag{12.6}$$

Extension for More Than Two Categories

In experimental analyses, the event categorisation can be quite complex, and often, there are more than two categories defined by cuts on kinematic variables. This is also the case for the $H \to Z\gamma$ analysis, and therefore, in this section, the ST method is extended beyond the simplest case of two categories, described previously. There are two possible basic scenarios, which are sketched in Figure 12.1:

- Multiple inclusive categories: In this case more than one category exists on the inclusive side of the cut under investigation. As an example consider three categories with a cut separating events with less than two jets (exclusive category) and two or more jets (inclusive categories), where the latter are further divided into two categories according to a different observable.
- Multiple exclusive categories: In this case, more than one category exists on the exclusive side of the cut under investigation. In the aforementioned example, this would mean that the events with less than two jets (exclusive categories) are further divided into two subcategories, while the events with two or more jets (inclusive category) are collected into one single category.



(a) Multiple inclusive categories (b) Multiple exclusive categories

Figure 12.1.: Visualisation of event migrations for the case of more than two categories.

In both cases, the distribution of the migration part of the uncertainty across multiple categories on one side of the cut has to be accounted for. In this section, only the case of multiple exclusive regions is discussed. This is motivated by the details of the event categorisation of the $H \to Z\gamma$ analysis as sketched in Figure 10.2.

As detailed in Equation (12.6), the yield uncertainty enters the part of the covariance matrix that corresponds to the exclusive category. Therefore, in the case of multiple exclusive categories, not only the migration but also the yield uncertainty has to be distributed across the exclusive categories. Since the size of the yield and migration uncertainties are still calculated from the same inputs, i.e. the corresponding inclusive scale variations, the different uncertainty contributions for the different exclusive categories are considered to be fully correlated. This results in the following covariance matrix, corresponding to the situation sketched in Figure 12.1(b):

$$C = \Delta_y^2 \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2 & 0 \\ \alpha_1\alpha_2 & \alpha_2^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \Delta_m^2 \begin{pmatrix} +\beta_1^2 & +\beta_1\beta_2 & -\beta_1 \\ +\beta_1\beta_2 & +\beta_2^2 & -\beta_2 \\ -\beta_1 & -\beta_2 & +1 \end{pmatrix} , \tag{12.7}$$

with $\alpha_1 + \alpha_2 = 1$ and $\beta_1 + \beta_2 = 1$. This ensures that, just like in the case of two categories, the migration uncertainty contributes to the uncertainties for each of the individual categories, but vanishes once all categories are combined. In particular,

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} C \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} &= \Delta_y^2 + \Delta_m^2 \\ \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} C \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} &= \Delta_m^2 \\ \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} C \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &= \Delta_y^2 \, . \end{aligned} \tag{12.8}$$

The factors α_c describe the relative populations of the two exclusive categories and can therefore be expressed as $\alpha_c = n_c/(n_1 + n_2)$ for $c = 1, 2$, where n_c is the event yield in category c. The factors β_c describe how the migration across the cut separating the inclusive and exclusive region is distributed across the two exclusive categories. Therefore, they are estimated by removing this cut from the categorisation, which effectively forces the events that would normally be counted in the inclusive category into both exclusive categories. The β_c factors then encode the fraction of these migrating events that end up in each of the exclusive categories, i.e. $\beta_c = (n'_c - n_c)/n_3$ for $c = 1, 2$, where n_3 is the event

yield in the third category before removing the categorisation cut and n'_c is the event yield in category c after removing the categorisation cut. By construction, $n_1+n_2+n_3 = n'_1+n'_2$, which ensures that $\beta_1 + \beta_2 = 1$.

Application to Likelihood Fits

In the previous sections as well as in the original publications, the ST covariance matrices were written down explicitly. In the context of the analysis at hand, however, the corresponding theoretical uncertainties are treated on the same footing as all other uncertainties. This is done by means of a likelihood function and likelihood fits to determine the values and distributions of the relevant test statistics as described in Chapter 6. Rather than covariance matrices, this procedure involves nuisance parameters and impact matrices, which describe the uncertainty contributed to each category by each source or nuisance parameter. For the common case of relative uncertainties, this can be written componentwise as

$$n_c^{(v)}(\nu_v) = n_c^{(0)} \cdot (1 + \delta_c^{(v)} \cdot \nu_v) \ , \tag{12.9}$$

where $n_c^{(0)}$ is the nominal event yield in category c, $n_c^{(v)}$ the event yield in category c when varied by the nuisance parameter ν_v corresponding to the uncertainty source v. The relative uncertainty is described by $\delta_c^{(v)}$.

In this analysis, however, not the event yields, but rather the event fractions per category ξ_c appear in the likelihood function. Therefore, first, the varied event yields are calculated as

$$\boldsymbol{n}_y^{(v)} = \boldsymbol{n}^{(0)} + \Delta_y \cdot \boldsymbol{\alpha} \cdot \nu_y^{(v)} \qquad \text{and} \qquad \boldsymbol{n}_m^{(v)} = \boldsymbol{n}^{(0)} + \Delta_m \cdot \boldsymbol{\beta} \cdot \nu_m^{(v)} \ , \tag{12.10}$$

where the vectors $\boldsymbol{n}^{(0)}$, $\boldsymbol{n}_y^{(v)}$ and $\boldsymbol{n}_m^{(v)}$ represent the nominal and varied event yields due to the yield and migration variations, respectively, in the different categories. The scalar quantities Δ_y and Δ_m denote the yield and migration uncertainties and the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ contain the relevant yield and migration factors for the different categories that previously entered the covariance matrices. In a second step, the varied event fractions per category and their relative uncertainties are calculated as

$$\xi_c^{(v)} = \frac{n_c^{(v)}}{\sum_{c'} n_{c'}^{(v)}} \qquad \text{and} \qquad \delta_c^{(v)} = \frac{\xi_c^{(v)} - \xi_c^{(0)}}{\xi_c^{(0)}} \ . \tag{12.11}$$

The $\delta_c^{(v)}$ are the entries of the impact matrices.

Implementation for the Analysis and Results

In the $H \to Z\gamma$ analysis, the impact of renormalisation, factorisation and resummation scale variations on the categorisation of the events was evaluated by means of uncertainty factors, which were derived in bins of the transverse momentum p_{T}^H of the Higgs boson. They are based on predictions from `HRes` 2.3 [228, 229], which include QCD corrections up to NNLO and resummation effects at NNLL. The central renormalisation scale μ_R, the factorisation scale μ_F and the two resummation scales Q_1 and Q_2 were chosen as

$$\mu_R = \mu_F = \frac{1}{2}\sqrt{m_H^2 + {p_{\mathrm{T}}^H}^2} \qquad Q_1 = \frac{m_H}{2} \qquad Q_2 = m_b \ . \tag{12.12}$$

Besides the nominal calculation, the envelopes of the 7-point perturbative scale variations, consisting of individual and simultaneous variations of μ_R and μ_F while avoiding simultaneous variations into opposite directions, i.e. up/down and down/up, and individual

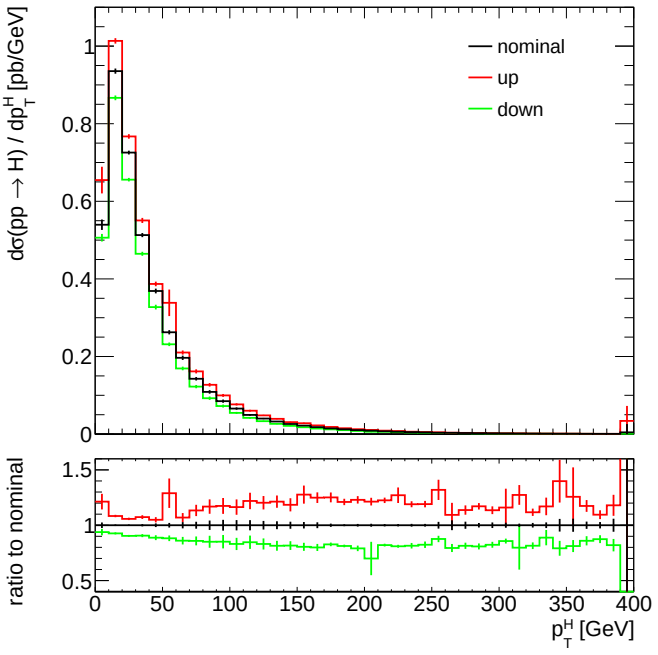


Figure 12.2.: Differential generator-level Higgs production cross section from ggF as a function of the transverse momentum of the Higgs boson and uncertainty factors obtained from `HRes` by comparing the scale variation envelopes to the nominal prediction with the central scale choices.

variations of both resummation scales were considered. The resulting uncertainty factors are shown in the ratio plot in Figure 12.2. They are taken to be the ratios of the scale variation envelopes and the nominal prediction corresponding to the central scale choices and were applied on top of the fully exclusive ggF `NNLOPS` MC sample with $m_H = 125$ GeV at the reconstruction level in order to calculate the QCD scale uncertainties on the cross sections in the different categories. The modified ST procedure was then applied to calculate separate yield and migration variations for the following uncertainties:

- the uncertainty due to the cut on the relative transverse momentum of the photon, $p_{\mathrm{T}}^{\gamma}/m_{\ell\ell\gamma}$, separating category 2 from categories 3-6
- the uncertainties due to the cuts on $p_{\mathrm{T}t}$, separating category 3 from category 4 and category 5 from category 6.

The applicability of the ST method to these cuts is justified by the fact that both observables are correlated with the jet activity in the event, as can be seen in Figure 12.3, where breakdowns in terms of the jet multiplicity are shown.

For the case of the cut on the relative photon p_{T}, category 2, the high relative photon p_{T} category, corresponds to the inclusive region, whereas categories 3-6 constitute the

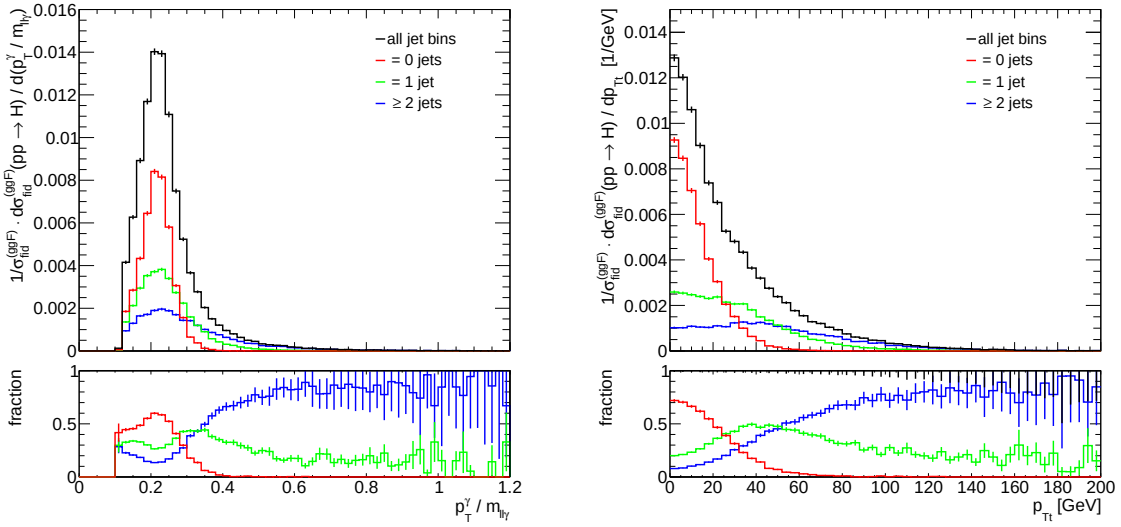


Figure 12.3.: Normalised differential reconstruction-level Higgs production cross section from ggF as a function of the relative transverse momentum of the photon and of $p_{\mathrm{T}t}$ of the Higgs boson candidate. As can be seen from the breakdowns in terms of the jet multiplicity, both observables are correlated with the jet activity in the event.

exclusive region. Therefore, the underlying ST covariance matrix is

$$\Delta_y^2 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_3^2 & \alpha_3\alpha_4 & \alpha_3\alpha_5 & \alpha_3\alpha_6 \\ 0 & 0 & \alpha_3\alpha_4 & \alpha_4^2 & \alpha_4\alpha_5 & \alpha_4\alpha_6 \\ 0 & 0 & \alpha_3\alpha_5 & \alpha_4\alpha_5 & \alpha_5^2 & \alpha_5\alpha_6 \\ 0 & 0 & \alpha_3\alpha_6 & \alpha_4\alpha_6 & \alpha_5\alpha_6 & \alpha_6^2 \end{pmatrix} + \Delta_m^2 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\beta_3 & -\beta_4 & -\beta_5 & -\beta_6 \\ 0 & -\beta_3 & \beta_3^2 & \beta_3\beta_4 & \beta_3\beta_5 & \beta_3\beta_6 \\ 0 & -\beta_4 & \beta_3\beta_4 & \beta_4^2 & \beta_4\beta_5 & \beta_4\beta_6 \\ 0 & -\beta_5 & \beta_3\beta_5 & \beta_4\beta_5 & \beta_5^2 & \beta_5\beta_6 \\ 0 & -\beta_6 & \beta_3\beta_6 & \beta_4\beta_6 & \beta_5\beta_6 & \beta_6^2 \end{pmatrix}, \tag{12.13}$$

resulting in the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ vectors

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & 0 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \end{pmatrix}^T \quad \text{and} \quad \boldsymbol{\beta} = \begin{pmatrix} 0 & +1 & -\beta_3 & -\beta_4 & -\beta_5 & -\beta_6 \end{pmatrix}^T . \tag{12.14}$$

As described in the previous section, the factors α_c account for the distribution of the yield uncertainty Δ_y between categories 3-6 and the β_c factors encode how events migrating out of category 2 populate the different $p_{\mathrm{T}t}$ categories. The β_c factors were derived by classifying events failing the previous BDT cut into the different $p_{\mathrm{T}t}$ categories without applying the high relative photon p_{T} categorisation cut. The resulting absolute and relative variations of the event fractions per category are given in Table 12.4. Despite the vanishing yield uncertainty components for the first two categories in Equation (12.13), the corresponding yield uncertainties are non-zero in Table 12.4. This is due to the fact, that in the table, the yield uncertainties on the event fractions per category $\xi_c^{(m)}$ are given. By definition, their sum across all categories is unity, i.e. $\sum_c \xi_c^{(m)} = 1$. Therefore, when the event yields in a subset of the categories change, also the event fractions of the remaining categories have to change in order to retain the normalisation to unity. The same applies to the uncertainties on the event fractions due to the $p_{\mathrm{T}t}$ categorisation, which will be discussed in the following and summarised in Tables 12.5 and 12.6.

From the same `HRes` scale variations, ST uncertainties were calculated for the $p_{\mathrm{T}t}$ categories

ST uncertainty due to the relative photon p_{T} cut							
		Category					
		1	2	3	4	5	6
Nominal event fractions		1.16%	5.14%	8.13%	32.84%	10.10%	42.64%
Absolute uncertainties							
Statistical uncertainty		0.04%	0.08%	0.11%	0.24%	0.12%	0.28%
Yield uncertainties	up	-0.11%	-0.48%	+0.05%	+0.20%	+0.06%	+0.27%
	down	+0.13%	+0.58%	-0.06%	-0.25%	-0.08%	-0.33%
Migration uncertainties	up	-0.00%	+0.89%	-0.33%	-0.07%	-0.41%	-0.08%
	down	+0.00%	-0.89%	+0.33%	+0.07%	+0.41%	+0.08%
Relative uncertainties							
Statistical uncertainty		3.27%	1.58%	1.35%	0.74%	1.23%	0.65%
Yield uncertainties	up	-9.25%	-9.25%	+0.62%	+0.62%	+0.62%	+0.62%
	down	+11.35%	+11.35%	-0.76%	-0.76%	-0.76%	-0.76%
Migration uncertainties	up	-0.00%	+17.38%	-4.00%	-0.23%	-4.06%	-0.19%
	down	+0.00%	-17.38%	+4.00%	+0.23%	+4.06%	+0.19%

Table 12.4.: Absolute and relative ST uncertainties on the event fractions per category for the high relative photon p_{T} categorisation cut. The uncertainties are evaluated applying uncertainty factors from `HRes` Higgs p_{T} scale variations on top of the `NNLOPS` reco-level events. For comparison, the nominal event fractions as well as the statistical uncertainties in all of the categories are given as well.

with the underlying ST covariance matrix taken to be

$$\Delta_y^2 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} + \Delta_m^2 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & -1 & 0 & 0 \\ 0 & 0 & -1 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{12.15}$$

for the electron categories and accordingly for the muon categories. Since the $p_{\mathrm{T}t}$-based categories are not divided further, the α_c or β_c vectors are simply given by

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}^T \quad \text{and} \quad \boldsymbol{\beta} = \begin{pmatrix} 0 & 0 & +1 & -1 & 0 & 0 \end{pmatrix}^T \tag{12.16}$$

for the electron channel. The resulting absolute and relative systematic uncertainties on the event fractions per category are given in Tables 12.5 and 12.6 for the electron and muon channels, respectively.

Even though the ggF production mode only contributes about a third of the event yield to the first category, which is enriched in VBF, it is responsible for the dominant uncertainty due to its perturbative uncertainties, which are much larger than for the VBF process. Therefore, an additional and separate perturbative uncertainty on the ggF contamination of the VBF-enriched category was derived using a procedure and a tool developed for similar studies in Run 1 [230]. It is based on a theory input of `MCFM` $H + 2j$ calculations, parameterised in bins of the azimuthal separation of the Higgs boson and the dijet system $\Delta\phi(H, jj)$. The scale uncertainties on this spectrum were evaluated and corresponding

ST uncertainty due to the $p_{\mathrm{T}t}$ cut in the electron channel							
		Category					
		1	2	3	4	5	6
Nominal event fractions		1.16%	5.14%	8.13%	32.84%	10.10%	42.64%
Absolute uncertainties							
Statistical uncertainty		0.04%	0.08%	0.11%	0.24%	0.12%	0.28%
Yield uncertainties	up	-0.05%	-0.20%	-0.32%	+2.64%	-0.40%	-1.68%
	down	+0.05%	+0.22%	+0.35%	-2.86%	+0.43%	+1.82%
Migration uncertainties	up	+0.00%	+0.00%	+1.15%	-1.15%	+0.00%	+0.00%
	down	+0.00%	+0.00%	-1.15%	+1.15%	+0.00%	+0.00%
Relative uncertainties							
Statistical uncertainty		3.27%	1.58%	1.35%	0.74%	1.23%	0.65%
Yield uncertainties	up	-3.93%	-3.93%	-3.93%	+8.04%	-3.93%	-3.93%
	down	+4.26%	+4.26%	+4.26%	-8.72%	+4.26%	+4.26%
Migration uncertainties	up	+0.00%	+0.00%	+14.11%	-3.49%	+0.00%	+0.00%
	down	+0.00%	+0.00%	-14.11%	+3.49%	+0.00%	+0.00%

Table 12.5.: Absolute and relative ST uncertainties on the event fractions for the electron $p_{\mathrm{T}t}$ categorisation cut. The uncertainties are evaluated applying uncertainty factors from `HRes` Higgs p_{T} scale variations on top of the `NNLOPS` reco-level events. For comparison, the nominal event fractions as well as the statistical uncertainties in all of the categories are given as well.

ST uncertainty due to the $p_{\mathrm{T}t}$ cut in the muon channel							
		Category					
		1	2	3	4	5	6
Nominal event fractions		1.16%	5.14%	8.13%	32.84%	10.10%	42.64%
Absolute uncertainties							
Statistical uncertainty		0.04%	0.08%	0.11%	0.24%	0.12%	0.28%
Yield uncertainties	up	-0.06%	-0.25%	-0.40%	-1.63%	-0.50%	+2.84%
	down	+0.06%	+0.28%	+0.45%	+1.81%	+0.56%	-3.15%
Migration uncertainties	up	+0.00%	+0.00%	+0.00%	+0.00%	+1.42%	-1.42%
	down	+0.00%	+0.00%	+0.00%	+0.00%	-1.42%	+1.42%
Relative uncertainties							
Statistical uncertainty		3.27%	1.58%	1.35%	0.74%	1.23%	0.65%
Yield uncertainties	up	-4.95%	-4.95%	-4.95%	-4.95%	-4.95%	+6.67%
	down	+5.50%	+5.50%	+5.50%	+5.50%	+5.50%	-7.40%
Migration uncertainties	up	+0.00%	+0.00%	+0.00%	+0.00%	+14.03%	-3.32%
	down	+0.00%	+0.00%	+0.00%	+0.00%	-14.03%	+3.32%

Table 12.6.: Absolute and relative ST uncertainties on the event fractions for the muon $p_{\mathrm{T}t}$ categorisation cut. The uncertainties are evaluated applying uncertainty factors from `HRes` Higgs p_{T} scale variations on top of the `NNLOPS` reco-level events. For comparison, the nominal event fractions as well as the statistical uncertainties in all of the categories are given as well.

Uncertainty on the ggF contamination in VBF							
		Category					
		1	2	3	4	5	6
Nominal event fractions		1.16%	5.14%	8.13%	32.84%	10.10%	42.64%
Absolute uncertainties							
Statistical uncertainty		0.04%	0.08%	0.11%	0.24%	0.12%	0.28%
Systematic variations	up	+0.52%	-0.03%	-0.04%	-0.17%	-0.05%	-0.22%
	down	-0.52%	+0.03%	+0.04%	+0.17%	+0.05%	+0.23%
Relative uncertainties							
Statistical uncertainty		3.27%	1.58%	1.35%	0.74%	1.23%	0.65%
Systematic variations	up	+44.46%	-0.52%	-0.52%	-0.52%	-0.52%	-0.52%
	down	-44.93%	+0.53%	+0.53%	+0.53%	+0.53%	+0.53%

Table 12.7.: Absolute and relative theoretical uncertainties from perturbative scale variations for the ggF contribution to the VBF-enriched category. The uncertainties were evaluated with a tool based on MCFM $H+2j$ theory inputs calculated for Run 1 and the ggF NNLOPS Monte Carlo sample for $m_H = 125$ GeV. For comparison, the nominal event fractions as well as the statistical uncertainties in all of the categories are given as well.

uncertainties were derived using the Stewart-Tackmann procedure and the procedures described in Sections 8.3 of [231] and 3.4.4 of [232]. These theory inputs are then combined with the reconstruction-level spectrum of $\Delta\phi(H, jj)$ as predicted by the ggF NNLOPS MC sample in order to evaluate the exclusive perturbative uncertainty on the ggF event yield in the VBF category. The uncertainty was found to be 49.75%. Due to the definition of the event fractions per category as being normalised to unity when summing over all categories, any variation in the first category has to be compensated by opposite variations in the other categories. The results for these uncertainties are given in Table 12.7.

12.2.2.2. Other Theoretical and Modelling Uncertainties

Modelling of Observables for the VBF BDT

An uncertainty on the modelling of observables that are exploited in the BDT for the VBF categorisation was derived based on a MadGraph5_aMC@NLO 2.3.3 [82] sample with FxFx multi-jet merging [88] of $H + 0/1j$ at NLO using the NNPDF3.0 PDF set [233]. The decays of the Higgs boson to $Z(\ell^+\ell^-)\gamma$ were handled with Pythia8 version 8.186, which also added the parton shower, hadronisation and underlying event simulations. The A14 tune [216] was used together with the NNPDF2.3 PDF set [217]. A Rivet routine was implemented in order to study the differences between the NNLOPS ggF sample and the aforementioned FxFx sample at the generator level. This is done in a phase space similar to the event selection strategy described in Section 10.1. However, since this study was performed at the generator level, no particle identification and isolation requirements were used. Resulting differences between the ggF NNLOPS and the FxFx samples were propagated from the generator level to the reconstruction level by means of uncertainty factors. These were derived separately for all of the observables entering the BDT except for the azimuthal separation of the $Z\gamma$ and the dijet system $\Delta\phi(Z\gamma, jj)$ as this observable shows quite good agreement in the VBF-dominated limit $\Delta\phi(Z\gamma, jj) \to \pi$. The impact on the cross section in the VBF-enriched category (category 1) was evaluated and the

Uncertainties on the modelling of observables for the VBF BDT							
		Category					
		1	2	3	4	5	6
Nominal event fractions		1.16%	5.14%	8.13%	32.84%	10.10%	42.64%
Absolute uncertainties							
Statistical uncertainty		0.04%	0.08%	0.11%	0.24%	0.12%	0.28%
Systematic variations	up	+0.18%	-0.01%	-0.01%	-0.06%	-0.02%	-0.08%
	down	-0.18%	+0.01%	+0.01%	+0.06%	+0.02%	+0.08%
Relative uncertainties							
Statistical uncertainty		3.27%	1.58%	1.35%	0.74%	1.23%	0.65%
Systematic variations	up	+15.30%	-0.18%	-0.18%	-0.18%	-0.18%	-0.18%
	down	-15.36%	+0.18%	+0.18%	+0.18%	+0.18%	+0.18%

Table 12.8.: Absolute and relative uncertainties on the event fractions per category from modelling differences of the observables exploited by the BDT for the separation of a VBF-enriched category. The numbers are calculated from uncertainty propagation of the generator-level differences between the ggF `NNLOPS` sample and an FxFx $H+0/1j$@NLO sample. For comparison, also the nominal event fractions per category and their statistical uncertainties are given.

contributions from uncertainty propagation for the different observables to this category were added in quadrature before calculating the varied event fractions per category. The Rivet plots from which the uncertainty factors were derived are given in Figure 12.4 and the resulting absolute and relative uncertainties on the event fractions per category are given in Table 12.8.

PDF and α_s Uncertainties

The categorisation uncertainties due to combined PDF and α_s variations were assessed from the PDF weights from the `NNLOPS` ggF sample with $m_H = 125$ GeV. These weights are calculated by the `PowhegBox` event generator during the hard process generation and allow for a consistent estimation of the PDF uncertainties using the recommendations given in [208]. The `PDF4LHC15_nlo_30_pdfas` PDF set and its corresponding error members were used. The resulting PDF and α_s uncertainties of the different event fractions per category for the low-mass region are given in Table 12.9. They were found to be smaller than or equal to the statistical uncertainty in each category. However, since the calculation of the PDF-related uncertainties is based on different weights specified for each and every single event, they are not affected by statistical fluctuations.

Uncertainties from the Underlying Event Modelling

Two separate ggF Higgs production samples were used to study the uncertainties from the UE modelling. These samples were generated with `PowhegBox` v1 and describe ggF at NLO QCD including finite top and bottom quark mass effects at NLO [213]. Both samples use the same `Pythia8` setup as the main Higgs signal samples, but for one of the two, the MPI / UE simulation was deactivated. The resulting variations of the event fractions per category of the low-mass region are given in Table 12.10. Besides the variations of the event fractions, an uncertainty on the acceptance times event selection efficiency is considered. As can be seen from Table B.1, the respective efficiencies, with the piecewise

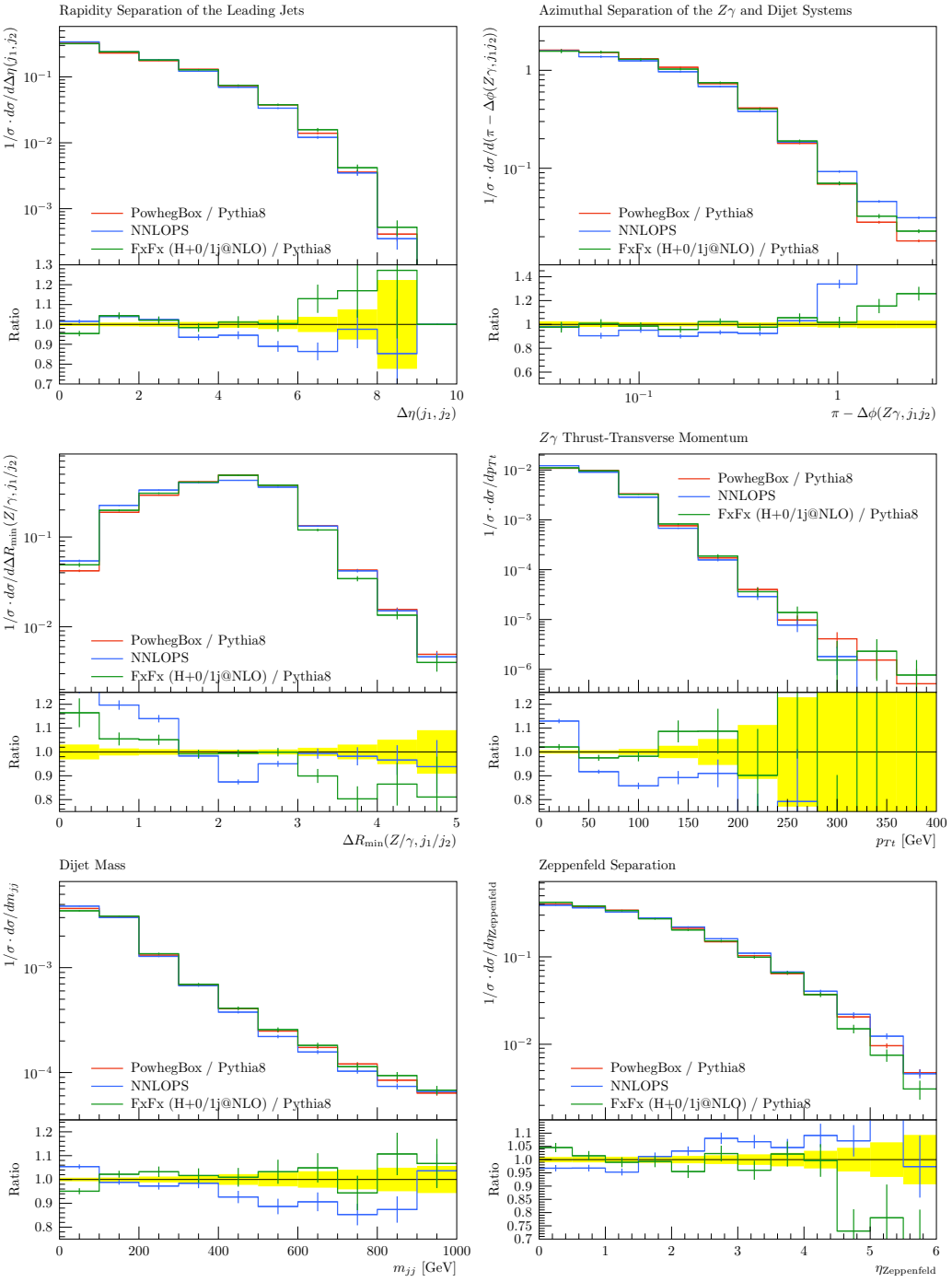


Figure 12.4.: Generator-level comparison between different ggF Monte Carlo generator predictions with $m_H = 125$ GeV for the observables exploited in the BDT used for the separation of a VBF-enriched category in the low-mass search. The ratio shows the relative differences with respect to the `PowhegBox/Pythia8` prediction and the yellow band and the vertical error bars indicate the statistical uncertainties of the different samples. The `PowhegBox/Pythia8` sample is only displayed for comparison as it was superseded by the NNLOPS sample. From these plots, uncertainty factors were derived from the differences between the FxFx and the NNLOPS samples for all observables except for the azimuthal separation of the $Z\gamma$ and the dijet system $\Delta\phi(Z\gamma, jj)$ as this observable shows quite good agreement in the VBF-dominated limit $\Delta\phi(Z\gamma, jj) \to \pi$.

PDF and α_s uncertainties							
		Category					
		1	2	3	4	5	6
Nominal event fractions		1.16%	5.14%	8.13%	32.84%	10.10%	42.64%
Absolute uncertainties							
Statistical uncertainty		0.04%	0.08%	0.11%	0.24%	0.12%	0.28%
Systematic variations	up	+0.02%	+0.10%	+0.09%	-0.08%	+0.11%	-0.12%
	down	-0.02%	-0.10%	-0.10%	+0.08%	-0.11%	+0.12%
Relative uncertainties							
Statistical uncertainty		3.27%	1.58%	1.35%	0.74%	1.23%	0.65%
Systematic variations	up	+1.64%	+1.89%	+1.14%	-0.24%	+1.07%	-0.28%
	down	-1.72%	-1.98%	-1.19%	+0.26%	-1.13%	+0.29%

Table 12.9.: Absolute and relative theoretical uncertainties on the event categorisation in the low-mass region due to the PDF error members and alternative central PDF sets. For comparison, also the nominal event fractions per category and their statistical uncertainties are given.

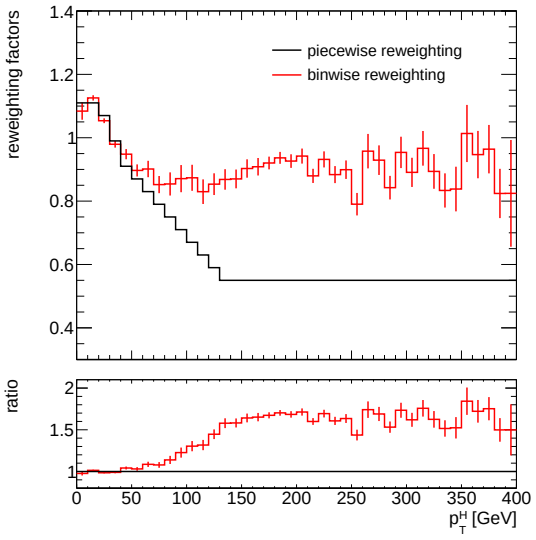
p_{T}^{H}-reweighting described below being applied, are 21.65% (MPI On) and 22.80% (MPI Off) resulting in a relative uncertainty of 5.3%.

For the calculation of the UE uncertainties the transverse momentum spectrum of the Higgs boson predicted by the `PowhegBox/Pythia8` Monte Carlo samples was reweighted using the piecewise defined function

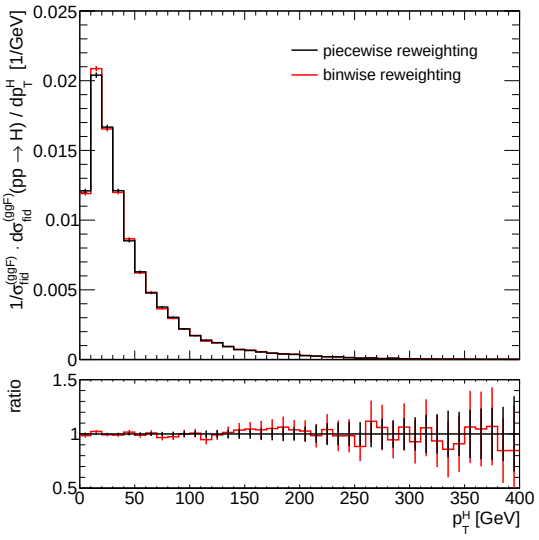
$$\begin{cases} 1.11 & p_{\mathrm{T}}^{H} < 20 \text{ GeV} \\ 1.11 - 0.2 \cdot \frac{p_{\mathrm{T}}^{H} - 20 \text{ GeV}}{25 \text{ GeV}} & 20 \text{ GeV} < p_{\mathrm{T}}^{H} < 45 \text{ GeV} \\ 0.91 - 0.36 \cdot \frac{p_{\mathrm{T}}^{H} - 45 \text{ GeV}}{90 \text{ GeV}} & 45 \text{ GeV} < p_{\mathrm{T}}^{H} < 135 \text{ GeV} \\ 0.55 & p_{\mathrm{T}}^{H} > 135 \text{ GeV} \end{cases} \tag{12.17}$$

proposed in [223]. It uses the generator-level p_{T}^{H} and is only applied to events with less than two jets, where generator-level jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta < 4.5|$ are considered. This reweighthing was derived for a centre-of-mass energy of 13 TeV. It is motivated by the fact that the Higgs boson p_{T}-distribution predicted by `PowhegBox/Pythia8` is slightly too hard compared to the formally more accurate `HRes` calculation. Apart from this approach, a cross-check was done with a naive jet-bin-ignorant reweighting to the central `HRes` prediction in bins of the Higgs boson p_{T}, referred to as binwise reweighting. A comparison of the reweighting factors obtained from both the piecewise and the binwise reweighting procedures is shown in Figure 12.5(a). Despite the different reweighting factors obtained from both procedures, the reweighted `PowhegBox/Pythia8` predictions in the fiducial volume, which are shown in Figure 12.5(b), show good agreement. As can be seen from 12.5(c), this is due to the fact that for $p_{\mathrm{T}}^{H} > 100$ GeV, where the reweighting methods differ the most, the spectrum is dominated by 2-jet-configurations which are not subject to the piecewise reweighting procedure.

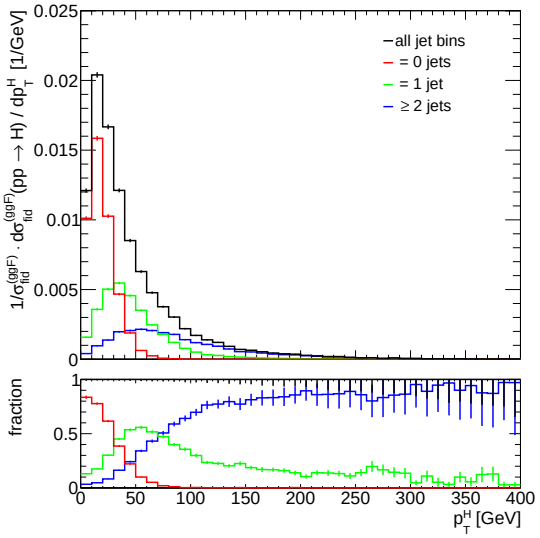
A summary of all the relative theory uncertainties on the event fractions per category is given in Table 12.11.



(a) Comparison of the reweighting factors



(b) Comparison of the reweighted p_{T}^{H} distributions in the fiducial volume



(c) Decomposition of the reweighted p_{T}^{H} spectrum in terms of the individual jet multiplicity configurations

Figure 12.5.: Comparison of the piecewise and binwise Higgs p_{T} reweighting strategies in the low-mass search. The error bars indicate the statistical uncertainties.

Underlying event uncertainties							
		Category					
		1	2	3	4	5	6
Nominal	MPI on	1.46%	5.18%	7.81%	33.03%	9.85%	42.67%
event fractions	MPI off	1.41%	6.09%	8.67%	31.41%	10.98%	41.44%
Absolute uncertainties							
Statistical	MPI on	0.05%	0.09%	0.11%	0.24%	0.12%	0.27%
uncertainties	MPI off	0.10%	0.26%	0.30%	0.53%	0.31%	0.63%
Systematic	up	-0.05%	+0.90%	+0.87%	-1.62%	+1.13%	-1.23%
variations	down	+0.07%	-1.31%	-1.26%	+2.35%	-1.64%	+1.78%
Relative uncertainties							
Statistical	MPI on	3.28%	1.69%	1.36%	0.71%	1.21%	0.63%
uncertainties	MPI off	7.35%	4.19%	3.40%	1.70%	2.83%	1.53%
Systematic	up	-3.25%	+17.41%	+11.09%	-4.90%	+11.46%	-2.88%
variations	down	+4.72%	-25.26%	-16.09%	+7.12%	-16.63%	+4.18%

Table 12.10.: Absolute and relative theoretical uncertainties on the event categorisation in the low-mass region due to the underlying event simulation. For comparison, also the nominal event fractions per category and their statistical uncertainties are given.

Source	Type	Var	Category 1	2	3	4	5	6
Cut on $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma}$	yield	up	-9.3%	-9.3%	+0.6%	+ 0.6%	+0.6%	+0.6%
		down	+11.4%	+11.4%	-0.8%	-0.8%	-0.8%	-0.8%
	migration	up	0.0%	+17.4%	-4.0%	-0.2%	-4.1%	-0.2%
		down	0.0%	-17.4%	+4.0%	+0.2%	+4.1%	+0.2%
Cut on $p_{\mathrm{T}t}$ in e^+e^- channel	yield	up	-3.9%	-3.9%	-3.9%	+8.0%	-3.9%	-3.9%
		down	+4.3%	+4.3%	+4.3%	-8.7%	+4.3%	+4.3%
	migration	up	0.0%	0.0%	+14.1%	-3.5%	0.0%	0.0%
		down	0.0%	0.0%	-14.1%	+3.5%	0.0%	0.0%
Cut on $p_{\mathrm{T}t}$ in $\mu^+\mu^-$ channel	yield	up	-5.0%	-5.0%	-5.0%	-5.0%	-5.0%	+6.7%
		down	+5.5%	+5.5%	+5.5%	+5.5%	+5.5%	-7.4%
	migration	up	0.0%	0.0%	0.0%	0.0%	+14.0%	-3.2%
		down	0.0%	0.0%	0.0%	0.0%	-14.0%	+3.3%
ggF contamination of VBF category		up	+44.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%
		down	-44.9%	+0.5%	+0.5%	+0.5%	+0.5%	+0.5%
Modelling of observables used in the BDT for the VBF category		up	+15.3%	-0.2%	-0.2%	-0.2%	-0.2%	-0.2%
		down	-15.4%	+0.2%	+0.2%	+0.2%	+0.2%	+0.2%
PDF and α_s		up	+1.6%	+1.9%	+1.1%	-0.2%	+1.1%	-0.3%
		down	-1.7%	-2.0%	-1.2%	+0.3%	-1.1%	+0.3%
UE / MPI		up	-3.3%	+17.4%	+11.1%	-4.9%	+11.5%	-2.9%
		down	+4.7%	-25.3%	-16.1%	+7.1%	-16.6%	+4.2%
Statistical uncertainty			3.3%	1.6%	1.4%	0.7%	1.2%	0.7%

Table 12.11.: Summary of the theory uncertainties on the event fractions per category for the low-mass region. The numbers represent the relative uncertainties for each category. For comparison, also the statistical uncertainties in each category are given.

Chapter 13

Statistical Analysis and Results

Finally, in this chapter, after briefly detailing the adopted statistical procedures, the results of the analysis are presented in the form of the p_0-values and upper limits on the parameters of interest. As described in Section 6.2, the p_0-values indicate the probability that a background fluctuation causes a signal at least as large as the one observed. The upper limits indicate how large the respective parameters of interest could be while still being compatible with the observed numbers of events.

13.1. Statistical Procedure

In order to quantify the compatibility of the observed data with the background-only hypotheses, i.e. the absence of the $Z\gamma$ decay of the SM Higgs boson and the absence of any high-mass resonance X decaying to the same final state, respectively, the corresponding p_0-values were calculated using the q_0 test statistic introduced in Section 6.2.3.1. In the case of the search for the $Z\gamma$ decay mode of the SM Higgs boson, only the results for a single masspoint, $m_H = 125.09$ GeV, are presented. For the high-mass region, a search was performed over the mass range 250 GeV $\leq m_X \leq$ 2.4 TeV. As discussed in Section 6.2.6, this necessitates taking into account the look-elsewhere effect. This is accomplished using the trial factors method [156], resulting in a global significance for the largest observed deviation from the background-only expectation.

In the absence of significant deviations from the background-only expectation, upper limits are calculated using the $\tilde{q}_\mu$ test statistic presented in Section 6.2.3.3, to which the CL_s method, introduced in Section 6.2.4.2, is applied. These upper limits are mainly derived from closed-form asymptotic formulae, describing the distribution of the test statistic in an analytic manner. However, the low number of expected and observed events in the tail of the invariant mass distribution of the high-mass region questioned the validity of the large sample limit, on which the closed-form approximation is based, in this region. Therefore, for $m_X > 1.6$ TeV, results based on ensemble tests were also derived.

13.2. Results

The observed distributions of events selected in the search for the $Z\gamma$ decay of the SM Higgs boson are shown as a function of the invariant mass $m_{Z\gamma}$ for the different categories in Figure 13.1 along with the background fitted in the range 115 GeV $\leq m_{Z\gamma} \leq$ 150 GeV and the expected signal for a SM Higgs boson with a mass of $m_H = 125$ GeV, scaled to 20 times the SM rate. As can be seen from the figures, there is no obvious localised excess of events around the 125 GeV masspoint. This is confirmed by the evaluation of the p_0-values: The observed p_0-value is 0.16 for $m_H = 125.09$ GeV, corresponding to a local significance of 1.0σ. The expected p_0-value is 0.33, corresponding to 0.5σ significance. The observed upper limit at the 95% CL on the signal strength μ is 6.6, which is slightly weaker than the expected limit on μ of 4.4 (5.2) assuming the absence (presence) of the $Z\gamma$ decay of the SM Higgs boson with $m_H = 125.09$ GeV. The observed limits on $\sigma(pp \to H) \cdot \mathcal{B}(H \to Z\gamma)$ and on $\mathcal{B}(H \to Z\gamma)$ are 547 fb and 1.0%, respectively. For the derivation of the limit on the branching fraction, the expected SM Higgs production cross section was assumed.

The observed distribution of events selected in the search for high-mass resonances is shown as a function of the invariant mass $m_{Z\gamma}$ for both categories in Figure 13.2 along with the background fitted in the range 200 GeV $\leq m_{Z\gamma} \leq$ 2.5 TeV. The largest deviation of the observed number of events from the background-only fit is located at $m_X = 960$ GeV and has a local significance of 2.7σ, corresponding to a global significance of 0.8σ when taking into account the entire search range of 250 GeV $\leq m_{Z\gamma} \leq$ 2.4 TeV. As no significant excess of events was observed, upper limits on $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ were derived at the 95% CL. As can be seen from Figure 13.3, the observed (expected) limits in the search region range between 88 fb (61 fb) and 2.8 fb (2.7 fb) for a spin-0 resonance produced via ggF. Corresponding plots of the expected and observed upper limits for a spin-2 resonance produced from gluon-gluon or quark-antiquark initial states are shown in Figure B.6 in Appendix B.2. For gluon-induced production the observed (expected) limits range between 117 fb (82 fb) and 3.7 fb (3.6 fb), while the corresponding values for quark-induced production lie between 94 fb (66 fb) and 2.3 fb (2.2 fb).

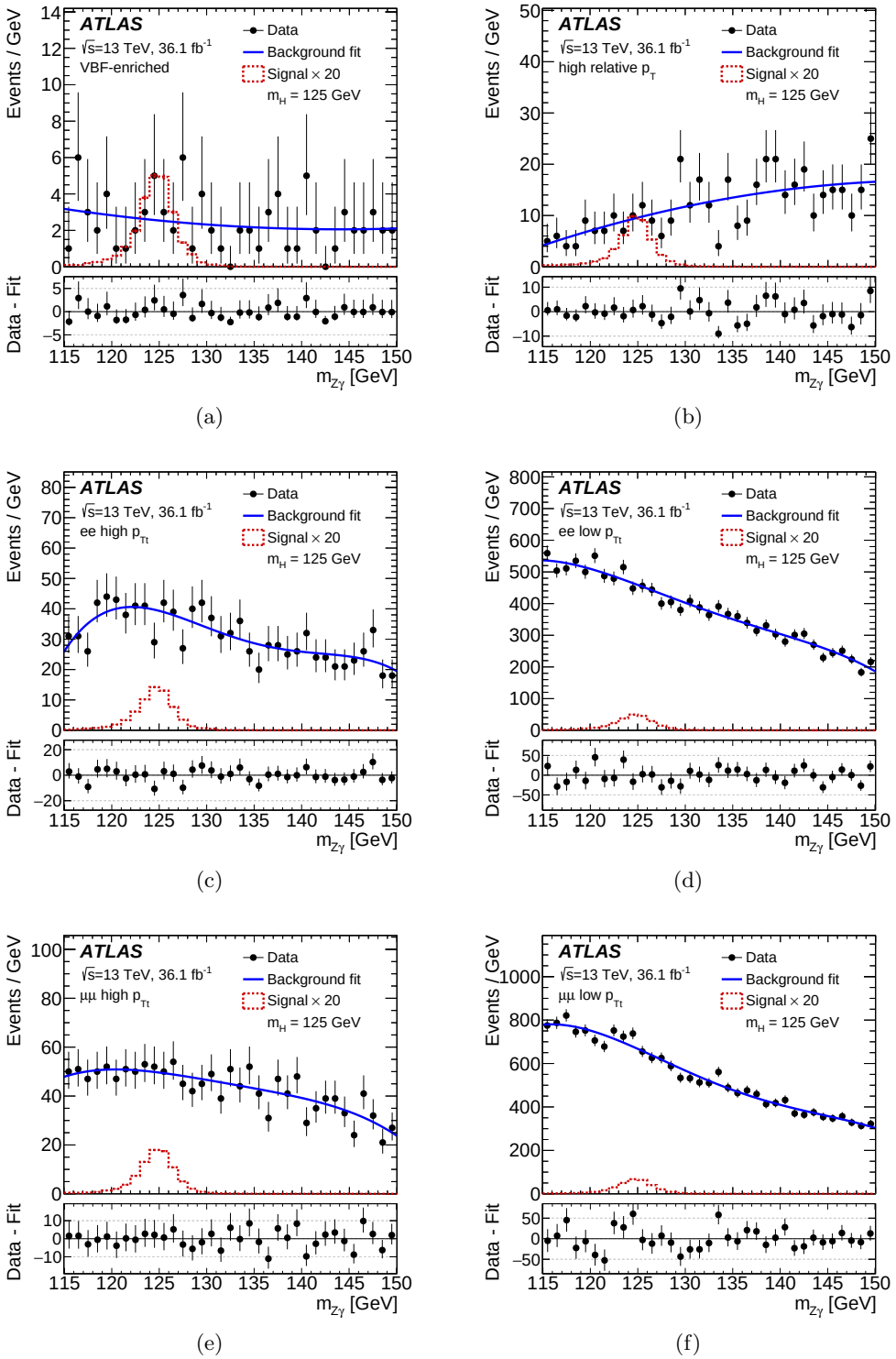
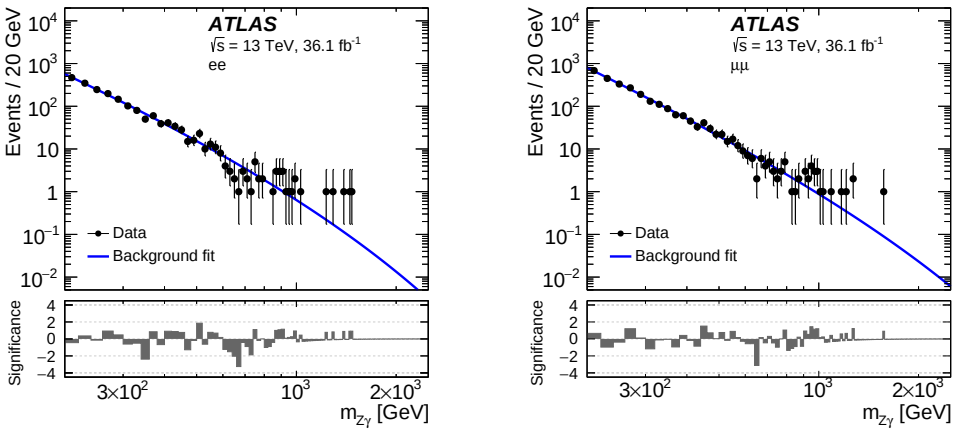


Figure 13.1.: Invariant $Z\gamma$ mass distributions in the search for the $Z\gamma$ decay of the Higgs boson. The data points show the selected events in the fit range 115 GeV $\leq m_{Z\gamma} \leq$ 150 GeV for the different categories as indicated in the respective plots. The solid blue line indicates the fitted background and the dashed red line shows the Higgs boson signal corresponding to a rate of 20 times the SM expectation for a mass of $m_H = 125$ GeV. The differences between the numbers of observed data events and the fitted backgrounds are shown in the bottom parts of the respective figures. Taken from [60].



(a) (b)

Figure 13.2.: Invariant $Z\gamma$ mass distributions in the search for high-mass resonances decaying to $Z\gamma$. The data points show the selected events in the fit range 200 GeV $\leq m_{Z\gamma} \leq$ 2500 GeV for the different categories as indicated in the respective plots. The solid blue line indicates the fitted background and the differences between the numbers of observed data events and the fitted backgrounds divided by the statistical uncertainty of the data points are shown in the bottom parts of the respective figures. Taken from [60].

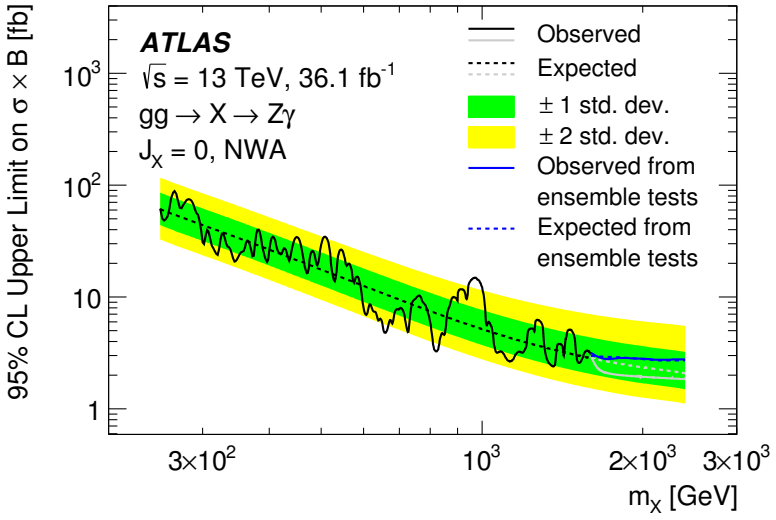


Figure 13.3.: Upper limits at the 95% CL on $\sigma(pp \to X)\cdot\mathcal{B}(X \to Z\gamma)$ for a spin-0 resonance produced via ggF as a function of the resonance mass m_X. The solid lines indicate the observed limits and the dashed line the expected limits. The green and yellow uncertainty bands indicate the regions corresponding to ± 1 and ± 2 standard deviations around the expected limits. The results represented with the black and grey lines as well as the uncertainty bands were derived from closed-form asymptotic formulae. For masses above 1.6 TeV, additional results, indicated in blue, were derived using ensemble tests. Taken from [60].

Part IV.

Search for the $\gamma^*(\ell^+\ell^-)\gamma$ Decay Mode of the Higgs Boson

Chapter 14

Introduction

The same final state as in the $H \to Z(\ell^-\ell^+)\gamma$ analysis presented in Part III can also be produced via the Higgs decay to two photons, one of which is off-shell and decays to a pair of charged leptons. A Born-level diagram shared by both processes is depicted in Figure 14.1. The relative size of the Z and γ^* contributions to the cross section varies with the invariant mass of the dilepton pair. Therefore, it is a natural complement of the $Z\gamma$ analysis to investigate the lower dilepton mass region as well, which in the following will be referred to as the $H \to \gamma^*\gamma$ channel.

The studies presented in the following chapters represent the first steps towards an analysis of the $H \to \gamma^*(\ell^+\ell^-)\gamma$ process in ATLAS. These efforts were aimed at a preliminary working version including a first estimate of the expected sensitivity. Thus, it was not possible to fully work out or optimise the different aspects and parts of the analysis. For the same reason, no systematic uncertainties could be studied. Improvements and optimisations as well as the incorporation of systematic uncertainties are undertaken by the remaining analysis group working towards a publication.

The presentation of the search for the $\gamma^*(\ell^+\ell^-)\gamma$ decay mode of the SM Higgs boson is organised as follows: After a brief overview in Section 14.1, some basic kinematic properties of the $\gamma^* \to \ell^+\ell^-$ splitting and the $H \to \gamma^*(\ell^+\ell^-)\gamma$ events are discussed in Section 14.2. This is followed by a presentation of the identification of close-by electrons stemming from the γ^* decays and the event selection and categorisation in Chapter 15. The statistical procedure, the resulting expected sensitivity and the prospects for larger datasets are given in Chapter 16.

14.1. Overview

Analyses covering the low dilepton mass region were carried out by CMS in Run 1 [67] and Run 2 [61, 234]. In Run 1, both the final states with electrons and muons were considered up to invariant dilepton masses of 20 GeV in the muon channel and 1.5 GeV in the electron

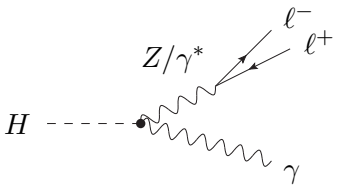


Figure 14.1.: Born-level diagram for $H \to \ell^+\ell^-\gamma$ via an intermediate Z/γ^* interference. The effective coupling of the Higgs boson to the photons and the Z boson, which proceeds via a loop, is indicated by the black dot. The relative size of the Z and the γ^* contributions to the production of the lepton pair depends on the dilepton invariant mass as will be shown later in Figure 14.4.

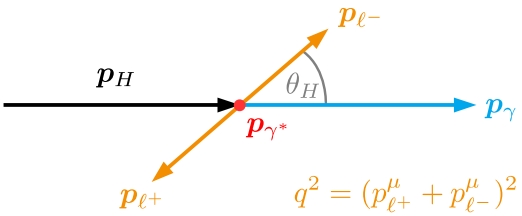


Figure 14.2.: Higgs and lepton momenta and the decay angle θ_H in the dilepton rest frame. θ_H is defined as the angle between the lepton and the Higgs momenta in that frame. Based on the discussion in [235].

channel and observed (expected) upper limits of 6.7 (5.9) times the SM expectation were found based on 19.7 fb^{-1} of data at 8 TeV. In the Run 2 analysis, only the final state with muons was included up to invariant dilepton masses of 50 GeV. The reason for discarding the electron channel is the difficulty of properly identifying the collimated electron-positron pairs from the decay of the off-shell photon. Using 35.9 fb^{-1} of data at 13 TeV, observed upper limits of 4 times the SM expectation were found. Assuming the absence (presence) of this decay of the SM Higgs boson, i.e. a signal strength of $\mu = 0$ ($\mu = 1$), the expected upper limit is 2.2 (3.1) times the SM expectation. Besides the $H \to \gamma^*\gamma \to \mu^+\mu^-\gamma$ channel, the $H \to Z\gamma \to \ell^+\ell^-\gamma$ channel was also investigated with $\ell = e, \mu$ in [61, 234]. The sensitivity of the analysis mainly comes from the $H \to \gamma^*\gamma$ channel, though. So far, no analysis of the $H \to \gamma^*\gamma$ channel has been performed by ATLAS and the studies presented in this part of the thesis at hand represent the first steps towards designing this search in ATLAS. As can be seen from the CMS results quoted above, so far no evidence or observation of this decay channel of the Higgs boson has been accomplished.

From a theory viewpoint, studying photon decays to lepton pairs gives a handle on CP-sensitive observables for $H \to \gamma\gamma$ via the angular distributions of the leptons. More commonly, such CP-sensitive observables are based on the momenta of four reconstructed particles, e.g. in $H \to 4\ell$ or $H \to \tau\tau \to$ hadrons. It is, however, also possible to define CP-sensitive observables based on three reconstructed particle momenta [235, 236]. In particular, the forward-backward asymmetry based on the angle θ_H between the lepton and the Higgs momenta in the dilepton rest frame, as sketched in Figure 14.2, is non-zero in the SM. Any deviations would indicate CP-violating couplings of the Higgs boson to $\gamma\gamma$ and $Z\gamma$. Studies presented in [236] estimate that there is a chance of observing the forward-backward asymmetry with the dataset expected from the HL-LHC.

An example subset of the relevant Feynman diagrams is given in Figure 14.3 in order to illustrate the different possibilities of how the $\ell^+\ell^-\gamma$ final state can be produced [67,

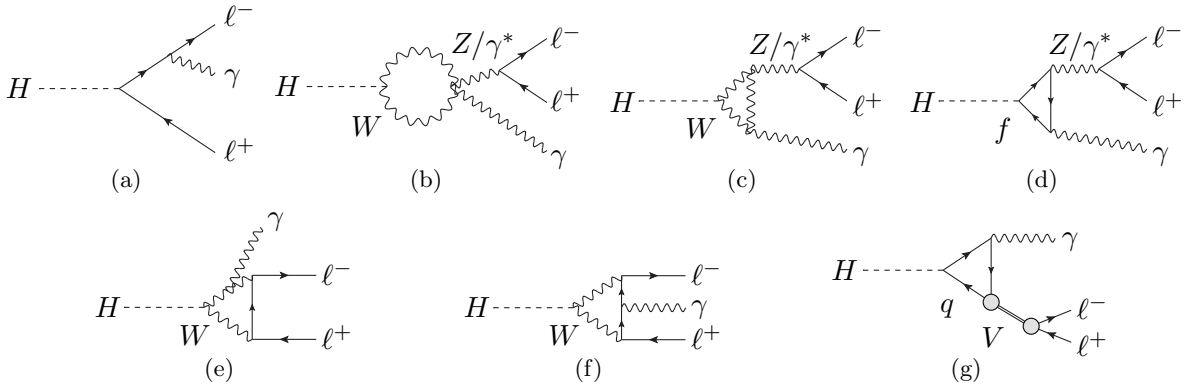
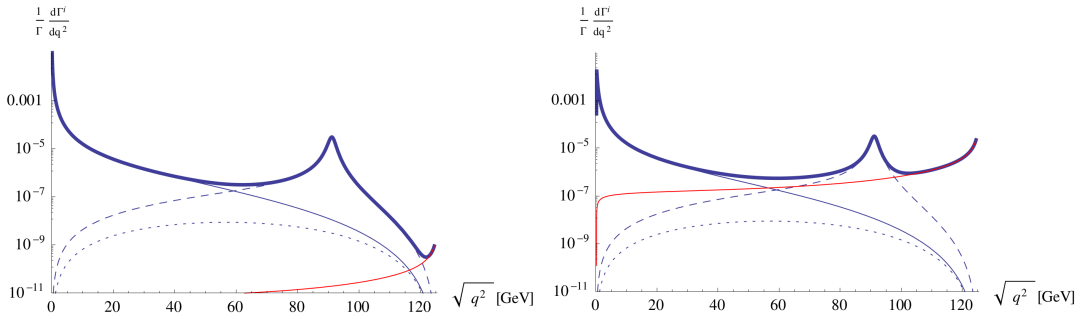


Figure 14.3.: Feynman diagrams for the Higgs Dalitz decay $H \to \ell^+\ell^-\gamma$. The tree-level process depicted in (a) contributes less than the pole diagrams shown in (b)-(d). Also smaller than the pole diagrams is the contribution from the four-point box diagrams, shown in (e) and (f). Another possibility is the production via an intermediate meson as sketched in (g).

235]. The diagrams corresponding to the aforementioned process $H \to \gamma^*(\ell^+\ell^-)\gamma$ are sketched in Figures 14.3(b)-(d) and are called *pole diagrams*. Since the photon pair is produced via W boson and heavy fermion loops, these diagrams are not suppressed by small Yukawa couplings of the Higgs boson to the final state leptons and are indeed the largest contributions to this final state [235, 237]. In the tree-level process, shown in Figure 14.3(a), the Higgs directly decays to a lepton pair and the additional photon is produced via FSR. This contribution is proportional to the lepton mass squared, however, and is therefore small compared to the pole diagram contributions, especially in the final state with electrons. The next type of diagrams are the *four-point box diagrams*, sketched in Figure 14.3(e)-(f), which contribute much less than the pole diagrams as well. A fourth option is the production of a photon and a meson which decays to a lepton pair, shown in Figure 14.3(g). These branching fractions of the Higgs boson to a meson and a photon are very small, however, and the branching ratios of the subsequent meson decays further reduce the magnitude of these contributions.[1]

The first three of these four classes of contributions are depicted individually in Figure 14.4, which shows the normalised differential decay width as a function of the invariant mass $\sqrt{q^2}$ of the lepton pair. The normalisation factor is the total Higgs to diphoton decay width $\Gamma(H \to \gamma\gamma)$. The contributions from the tree-level diagrams is shown by the red line, the thin solid and dashed blue lines give the contributions from the γ and Z pole diagrams and the dotted blue line the contributions from the four-point box diagrams. As can be seen from these plots, the tree-level diagrams are completely irrelevant in the final state with electrons. They are the subleading contribution in the final state with muons, but still at least two orders of magnitude smaller than the dominating γ pole diagram contributions for invariant dilepton masses smaller than 20 GeV. The Z pole diagrams start to dominate only for invariant dilepton masses of at least about 60 GeV and are tiny in the region of low invariant mass. Even smaller than that is the contribution from the four-point box diagrams. All this motivates why the analysis presented in the following chapters is based on the production via the γ pole diagrams only.

[1] Relevant mesons are the J/ψ and the Υ with $\mathcal{B}(H \to J/\Psi\gamma) \approx 3 \cdot 10^{-6}$, $\mathcal{B}(H \to \Upsilon\gamma) \approx 2 - 6 \cdot 10^{-10}$ and $\mathcal{B}(J/\Psi \to \ell^+\ell^-) \approx 6\%$ for $\ell = e, \mu$ each. A search for these decay modes was performed by ATLAS and presented in [238].



(a) $H \to e^+e^-\gamma$. (b) $H \to \mu^+\mu^-\gamma$.

Figure 14.4.: Normalised differential Higgs decay width as a function of the invariant mass $\sqrt{q^2}$ of the dilepton system. The normalisation factor is the total Higgs to diphoton decay width $\Gamma(H \to \gamma\gamma)$. The red line denotes the contributions from the tree-level diagrams, the thin solid and dashed lines the γ and Z pole diagrams and the dotted line the four-point box diagrams. The thick solid line gives the total decay width. As can be seen, the γ pole diagrams dominate the lower invariant mass region, especially in the final state with electrons. The tree-level diagrams are the subleading but still tiny contribution in the final state with muons, while they are completely irrelevant in the electron final state. Taken from [235].

The branching ratios relevant for this channel are [67, 231, 234, 239][2]

$$\frac{\Gamma(H \to \gamma^*\gamma \to e^+e^-\gamma)}{\Gamma(H \to \gamma\gamma)} \approx 3.5\% \qquad \frac{\Gamma(H \to \gamma^*\gamma \to \mu^+\mu^-\gamma)}{\Gamma(H \to \gamma\gamma)} \approx 1.7\% \tag{14.1}$$

and because of $\mathcal{B}(H \to \gamma\gamma) = 0.228\%$

$$\mathcal{B}(H \to \gamma^*\gamma \to e^+e^-\gamma) \approx 7.98 \cdot 10^{-5} \quad \text{and} \quad \mathcal{B}(H \to \gamma^*\gamma \to \mu^+\mu^-\gamma) \approx 3.88 \cdot 10^{-5} \,. \tag{14.2}$$

14.2. Splitting and Event Kinematics

Some key properties of the $\gamma^* \to \ell^+\ell^-$ splitting can be derived from basic kinematic considerations already. Let E_{γ^*} be the energy and m_{γ^*} the virtuality of the photon that splits into a pair of leptons of mass m_ℓ each. In a convenient kinematic parameterisation, the photon momentum is oriented along the z-direction and given by k and a photon energy fraction z is defined so that the z-component of the lepton momentum is given by $p_-^z = zE_{\gamma^*}$. A more detailed treatment with the corresponding mathematical expressions is presented in App. C.1.

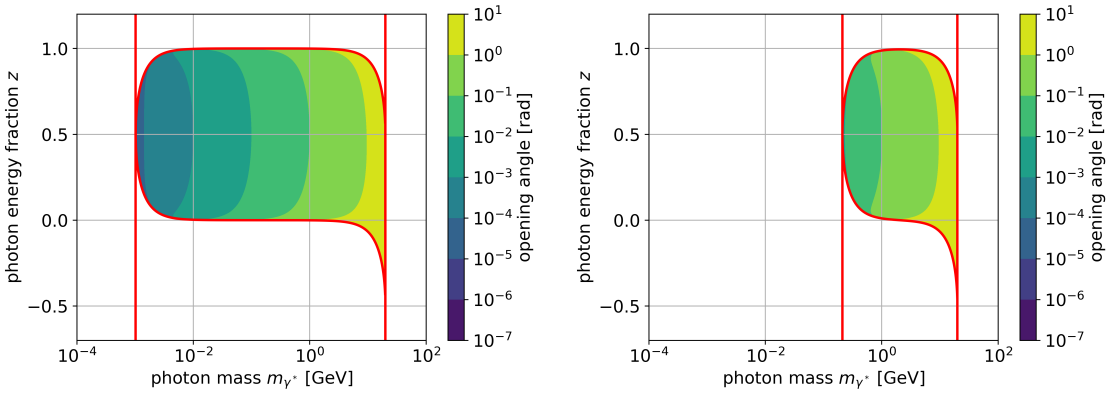
The opening angle between the two leptons is shown in Figure 14.5 as a function of the photon virtuality m_{γ^*} and the photon energy fraction z for different photon energies and lepton masses. Specifically, decays to pairs of electrons and muons are depicted. The red lines in these plots denote the kinematic bounds on the allowed splitting phase space. The

[2] Further phenomenological studies with different sets of experimental cuts are presented in [240] for $H \to \ell^+\ell^-\gamma$ and in [241] for $H \to e^+e^-\gamma$.

photon virtuality is bounded from above by the available photon energy. On the other hand, the masses of the leptons lead to a lower bound on the photon virtuality. This is the reason why lower photon virtualities and, as a consequence, also lower opening angles are accessible in decays to pairs of electrons compared to pairs of muons. For a given energy, a lower virtuality leads to a larger momentum which increases the boost effect and leads to reduced opening angles. The same logic also holds when increasing the photon energy E_{γ^*}. This can be seen by comparing the plot in the lower row against the ones in the upper row of Figure 14.5. The remaining bounds, which limit z, are dictated by energy conservation. All the different kinematic boundaries are also depicted and labelled in Figure 14.6, which also shows the regions where soft and collinear splittings occur.

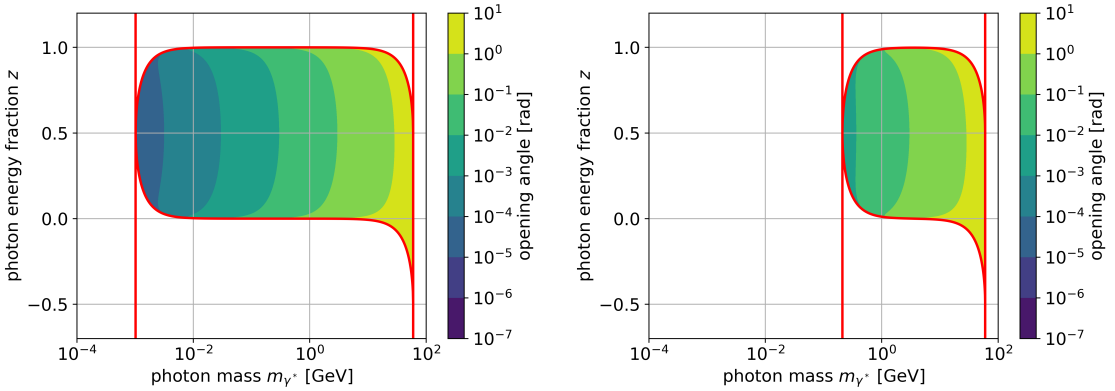
These soft and collinear regions are interesting because, just like in parton showers, they are expected to be associated to an increased splitting probability. The soft region is defined by letting either the energy of the lepton $E_- \to 0$ or of the antilepton $E_+ \to 0$, which corresponds to the red curve bounding the values of z for $z \to 0$ or $z \to 1$. The collinear region is defined by a vanishing lepton transverse momentum and both lepton momenta being oriented in the same direction. This translates to the red curve bounding z between values of $k/(2E_{\gamma^*}) \pm \sqrt{m_{\gamma^*}^2 - 4m_\ell^2}/(2m_{\gamma^*})$.

These basic kinematic considerations are reflected in the generator-level distributions of the invariant dilepton mass $\sqrt{q^2}$ and the separation ΔR between the leptons from the γ^* decay which are given in Figure 14.7. The shapes of these distributions are influenced by the logarithmically spaced bin edges and by the fact that the bin content was not scaled to account for the unequal bin sizes. This form of presentation, however, has the benefit of illustrating the similar behaviour of electrons and muons in the region of large invariant mass and separation as well as showing the larger phase space towards low invariant masses and separations in the electron channel. Finally, the generator-level transverse momentum distributions of the real photon γ, the leptons as well as the Higgs boson and the virtual photon γ^* are depicted in Figure 14.8. These distributions were derived from Monte Carlo event samples of the ggF production mode of the Higgs boson, simulated with `PowhegBox` using the NNLOPS approach. More details on the Monte Carlo event samples will be given in Section 15.1. No kinematic cuts were applied in these plots.



(a) electrons from photon with $E_{\gamma^*} = 20$ GeV

(b) muons from photon with $E_{\gamma^*} = 20$ GeV



(c) electrons from photon with $E_{\gamma^*} = 60$ GeV

(d) muons from photon with $E_{\gamma^*} = 60$ GeV

Figure 14.5.: Opening angle in radians between pairs of leptons stemming from the decay of a photon, as a function of photon virtuality m_{γ^*} and photon energy fraction z for different photon energies E_{γ^*}. Since $z = p_-^z / E_{\gamma^*}$, it can be negative. The red lines denote the kinematic bounds on the allowed splitting phase space. A lower bound on the photon virtuality m_{γ^*} is dictated by the mass of the leptons. Therefore, lower photon virtualities and splitting angles are accessible in decays to electron pairs as opposed to muon pairs. Similarly, the larger the photon energy the more pronounced the boost effect. Smaller opening angles become accessible in the low virtuality region when the photon energy increases, leading to an increased photon momentum, as can be seen by comparing the lower against upper set of plots.

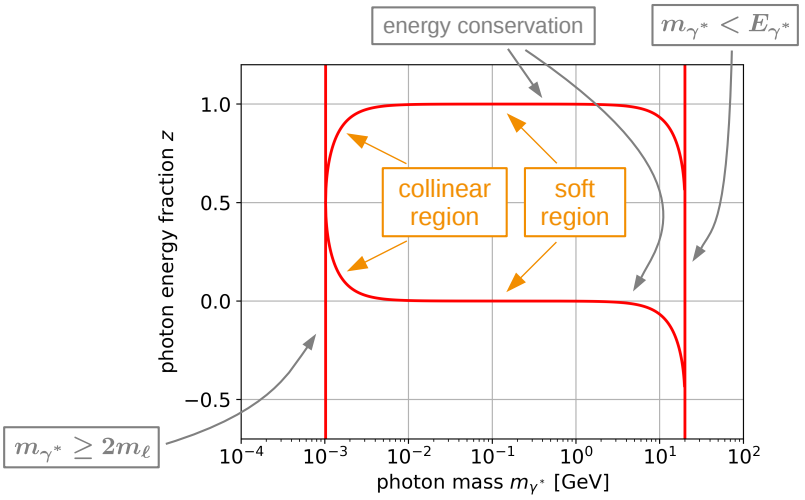


Figure 14.6.: Kinematic boundaries and regimes for electrons from the decay of a virtual photon with $E_{\gamma^*} = 20$ GeV.

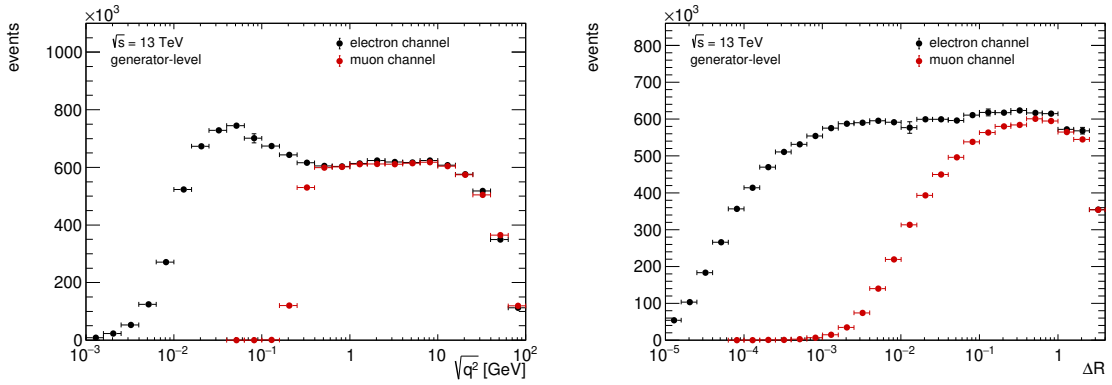


Figure 14.7.: Generator-level distributions of the invariant dilepton mass $\sqrt{q^2}$ (left) and the separation ΔR between both leptons from the $\gamma^* \to \ell^+\ell^-$ decay (right) for both the electron and the muon channel. The shapes of the distributions are influenced by the logarithmically spaced bin edges and the fact that the bin content is not compensated for the unequal bin sizes. The benefit of this presentation is, however, that the similar behaviour of electrons and muons for sufficiently large invariant masses and separations can be seen. When the kinematic thresholds are approached, the production of the respective lepton flavour diminishes towards smaller invariant masses and separations. These distributions demonstrate the larger phase space in the electron channel already mentioned in Section 14.2.

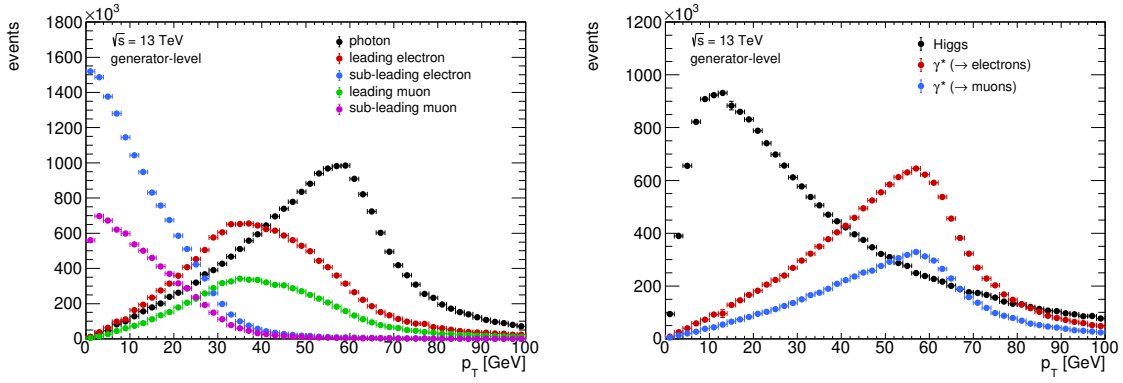


Figure 14.8.: Generator-level transverse momentum distributions of the real photon γ and the leptons (left) as well as the Higgs boson and the virtual photon γ^* as reconstructed from the two leptons (right).

Chapter 15

Event Selection and Categorisation

The lower the invariant mass of the virtual photon, the more collimated the outgoing leptons. This is especially true in the final state with electrons, as was shown in Section 14.2. This close proximity can lead to configurations where the ECal clusters stemming from both electrons overlap and are therefore merged into only one reconstructed electron object. The standard electron identification algorithms were not designed to deal with these cases and therefore cannot be used for such merged configurations. In order to overcome this limitation and use such events in the analysis, a dedicated electron identification algorithm was developed. After a brief discussion of the signal and background Monte Carlo event samples used for this analysis in Section 15.1, the dedicated merged electron ID algorithm will be presented in Section 15.2. This is followed by a discussion of the event selection and categorisation in Section 15.3, on which the expected results presented in Chapter 16 are based.

15.1. Monte Carlo Event Samples

The studies of the signal are based on the ggF production mode of the Higgs boson only. For the Monte Carlo simulation of Higgs production in this mode, the `PowhegBox` v2 event generator [83–85] was used with the NNLOPS [209] approach and the same setup already described in Section 11.1.1. The Higgs boson decay as well as the parton shower, hadronisation and underlying event simulation were performed with `Pythia8` verion 8.230 using the CTEQ6L1 PDF set [176] and the AZNLO tune [212]. Since `Pythia8` does not include a direct $H \to \gamma^*(\ell^+\ell^-)\gamma$ decay chain, the $H \to \gamma\gamma$ decay was selected and the parton shower was relied upon to perform a QED $\gamma \to \ell^+\ell^-$ splitting. Due to the smallness of the electromagnetic coupling α compared to the strong coupling α_s, this QED splitting occurs much less frequently than the QCD splitting processes. To make the best use of the available hard process events, the parton shower was repeated for a given event until the desired splitting was applied to one of the photons from the Higgs decay.

The background is modelled with a range of LO processes generated simultaneously with `Pythia8` using the NNPDF2.3 LO PDF set [217] and the A14 tune [216]. The simulated processes include:

- hard QCD $2 \to 2$ production
- jet + photon production
- single weak gauge boson production
- top quark pair production

A filter was applied to select events with at least one jet with a transverse momentum of more than 17 GeV reconstructed in a grid in $\Delta\eta \times \Delta\phi$ of size 0.12×0.12. Such samples with different p_{T} thresholds are commonly used for electron identification studies in ATLAS.

For both the signal and the background Monte Carlo samples the FullSim detector simulation was used.

15.2. Identification of Merged Electrons

A preliminary event selection, described in the following in Section 15.2.1, was used for the optimisation of the merged electron identification. The ID menu itself is then presented in Section 15.2.2. In the merged electron ID studies presented in this section, the MiNLO event weights of the ggF signal sample were used. The expected results presented in Chapter 16, however, are based on the full accuracy given by the NNLOPS weights.

15.2.1. Event Selection for the Optimisation

For both signal and background, a common selection based on the properties of reconstructed objects was applied. For the signal component, additional generator-level requirements were applied.

In this generator-level event selection, the decay products of the Higgs boson are analysed. The presence of exactly two opposite-charge leptons, either electrons or muons, and of at least one photon is required. One of the photons should come directly from the Higgs boson decay, whereas others may be radiated off the leptons. This ensures that only one of the photons from the Higgs decay was split into a lepton pair by the parton shower. Finally, transverse momentum thresholds of 20 GeV (5 GeV) are applied to the leading (subleading) lepton.

In the event selection applied to both the signal and the background, requirements are applied to the reconstructed objects. Photons are required to be within $|\eta| < 2.37$ and to satisfy the conditions of the *Loose* ID working point [242]. In order to make good use of the limited sample size, only a very loose p_{T} cut of 1 GeV is applied. Only events with at least one such photon are considered. Electrons are required to be within $|\eta| < 2.47$ and have a transverse momentum of at least 5 GeV. No electron ID requirements are used. The leading electron is then required to have a p_{T} of at least 21 GeV. The electron reconstruction allows for an electron candidate to have multiple tracks associated to it and the same track may be associated to different electron candidates. Across all reconstructed electron objects, there should be at least two different reconstructed tracks. For this condition, it does not matter whether both tracks are associated to the same or to different electron objects.

tracking variables	
n_{Pix}	number of hits in the PIX detector
n_{Si}	number of hits in the PIX and SCT detectors
eProbabilityHT	likelihood probability based on transition radiation in the TRT [141]
d_0	transverse impact parameter
$\lvert d_0\rvert/\sigma_{d_0}$	significance of the transverse impact parameter
σ_p/p	relative track momentum resolution
track-cluster matching variables	
$\Delta\eta_1$	difference in η between the cluster position and the extrapolated track impact position in the ECal strip layer
$E_{\text{cluster}}/p_{\text{tracks}}$	energy of the ECal cluster divided by the scalar sum of the two track momenta

Table 15.1.: Tracking and track-cluster matching variables considered in the optimisation of the merged electron identification. The tracking variables are evaluated for each of the two tracks individually.

Subsequently, quality requirements are imposed on the electron tracks: Track candidates should have at least two hits in the PIX and at least seven hits in the PIX and SCT detectors. Finally, the presence of at least one track of a positive charge and one track of a negative charge is required.

γ^* candidates are formed from pairs of opposite-charge electron tracks. Candidates not satisfying the condition $|z_0(\text{track 1}) - z_0(\text{track 2})| \leq 2$ mm are discarded. This serves to reject combinations of tracks originating from different primary vertices. From the valid candidates, the track pair with the largest combined vector-summed p_{T} is chosen. Based on this choice, the category is determined: If both selected tracks are associated to the same electron object, the event falls in the *merged* category, whereas if both selected tracks are associated to two different electron objects, the event is marked as belonging to the *resolved* category. Next, an overlap removal procedure is applied. In the merged category, events where the highest p_{T} photon candidate is within ΔR of 0.4 of the γ^* candidate are removed and in the resolved category, events where the highest p_{T} photon is within ΔR of 0.4 of one of the two electrons are removed. Finally, the γ^* is required to have a transverse momentum of at least 21 GeV.

15.2.2. Optimisation of the Merged Electron Identification

The algorithm for the identification of merged electrons is formulated in bins of the transverse momentum p_{T} and the pseudorapidity η. In each p_{T} or η-bin i, the signal and the background components are analysed as distributions of a variable x, which can be a shower shape, a tracking or a track-cluster matching variable. Besides the shower shape variables listed in Table 5.1, the variables given in Table 15.1 were considered.

From the normalised histograms of the distributions, the *separation* $\langle S_i^2 \rangle$ is calculated in each p_{T}/η-bin i as [143]

$$\langle S_i^2 \rangle = \frac{1}{2} \sum_j \frac{(s_{ij} - b_{ij})^2}{s_{ij} + b_{ij}} \ , \tag{15.1}$$

where s_{ij} and b_{ij} describe the signal and background content in bin j of the histogram of the variable x. If the signal and background distributions are identical, the separation $\langle S_i^2 \rangle$ will be zero, whereas if there is no overlap between both components, $\langle S_i^2 \rangle$ will be equal to one. From the separations $\langle S_i^2 \rangle$ in each p_{T} or η-bin i, a combined total separation $\langle S^2 \rangle$ is calculated as

$$\langle S^2 \rangle = \sum_i w_i \langle S_i^2 \rangle \qquad \text{with} \quad w_i = \frac{S_i B_i}{\sum_k S_k B_k} \ . \tag{15.2}$$

The sum over k runs over all p_{T} or η-bins and S_k and B_k denote the signal and background event yields in bin k. This way, more weight is given to bins containing larger fractions of both signal and background.

The merged electron identification algorithm is loosely based on the ranking in terms of the combined separations. It consists of a series of sequential cuts in each bin of one variable at a time. For R_{had} and f_3, the cut value depends on the bin indices in both p_{T} as well as η. For these two variables, the cuts are at least as tight as in the merged electron trigger. The subsequent cuts on some of the other variables are one-dimensional, i.e. the cut value only depends on either the p_{T}-bin or the η-bin but not on both of them at the same time. The cuts are applied in two stages and the corresponding efficiencies for the 2015/16 period for both the signal and background component are given in Table 15.2. The corresponding numbers for the conditions of 2017 are given in Table C.9 in Appendix C.2. The first-stage cuts (above the double horizontal line) are based on the separations evaluated after the event selection, whereas the second-stage cuts (below the double horizontal line) are based on the separations evaluated after the first stage cuts. The combined and per-bin separations in p_{T} and η before and after the first-stage cuts are given in Tables C.5 through C.8 in Appendix C.2. The exact locations of the cuts were chosen in a way so as to retain a substantial fraction of the signal component.

Some example signal and background distributions are given in Figures 15.1-15.3. In Figure 15.1, the distributions of R_{had} are shown in bins of p_{T}. As mentioned before, the values of R_{had} at which the cuts are located depend on both p_{T} and η. For visualization, however, the distributions are only shown in bins of p_{T} here. R_{had} is defined as the ratio of the cluster E_{T} deposited in the first layer of the HCal over the cluster E_{T} in the ECal. As can be seen, the signal is concentrated close to zero, which means that the cluster energy is contained within the ECal. This is expected since the signal is from electrons. The background extends to larger values of R_{had}. The locations of the cuts are indicated by the vertical red lines. Since the cuts depend on both p_{T} and η, there are multiple cut values for a single p_{T} bin. In Figure 15.2, the signal and background are shown as a function of the variable R_η, which is the lateral energy ratio along the η-direction. As described in Figure 5.1 and Table 5.1, it is defined as the ratio of the energy in 3×7 cells in $\eta \times \phi$ over the energy in 7×7 cells in the second layer of the ECal. Since the magnetic field in the inner detector separates tracks of different electric charges in the ϕ-direction, the width of the shower in the ECal is not affected in the η-direction. As electromagnetic showers are typically narrower than jets, more of their energy is contained in the smaller 3×7 grid already and therefore, the signal is concentrated at larger values of R_η than the background. The region of $1.37 < |\eta| < 1.52$ is the crack region at the transition between the barrel and the endcaps, which is not fully instrumented. Events with γ^* candidates in this region of the detector were not considered. Finally, the signal and background are shown as a function of E/p in Figure 15.3. Introduced in Table 15.1, E/p describes the ratio of the cluster energy in the ECal over the scalar sum of the momenta of the tracks

variable	cut in bins of	efficiency signal		background	
		ϵ_{rel}	ϵ_{abs}	ϵ_{rel}	ϵ_{abs}
R_{had}	$p_{\text{T}} \times \eta$	92.9%	92.9%	32.5%	32.5%
f_3	$p_{\text{T}} \times \eta$	99.5%	92.4%	78.0%	25.4%
R_η	η	95.6%	88.3%	24.4%	6.17%
E_{ratio}	η	97.7%	86.3%	58.2%	3.59%
$w_{s,\text{tot}}$	η	93.8%	80.9%	71.5%	2.57%
$w_{\eta,2}$	η	99.6%	80.6%	90.4%	2.32%
$\Delta\eta_1$	p_{T}	96.3%	77.6%	56.9%	1.32%
R_{had}	η	98.9%	76.7%	78.2%	1.03%
eProbabilityHT(track$_2$)	η	98.9%	75.9%	95.6%	0.98%
eProbabilityHT(track$_1$)	η	98.9%	75.1%	96.4%	0.95%
E/p	η	97.8%	73.4%	96.6%	0.92%
R_ϕ	η	88.5%	65.0%	80.0%	0.74%
$\lvert d_0\rvert/\sigma_{d_0}$(track$_1$)		98.4%	64.0%	89.3%	0.66%
$\lvert d_0\rvert/\sigma_{d_0}$(track$_2$)		96.6%	61.8%	90.6%	0.60%

Table 15.2.: The cut sequence of the merged electron identification algorithm along with the corresponding cut efficiencies for the signal and background components for the conditions of the 2015/16 period. The corresponding numbers for the conditions in 2017 are given in Table C.9 in Appendix C.2.

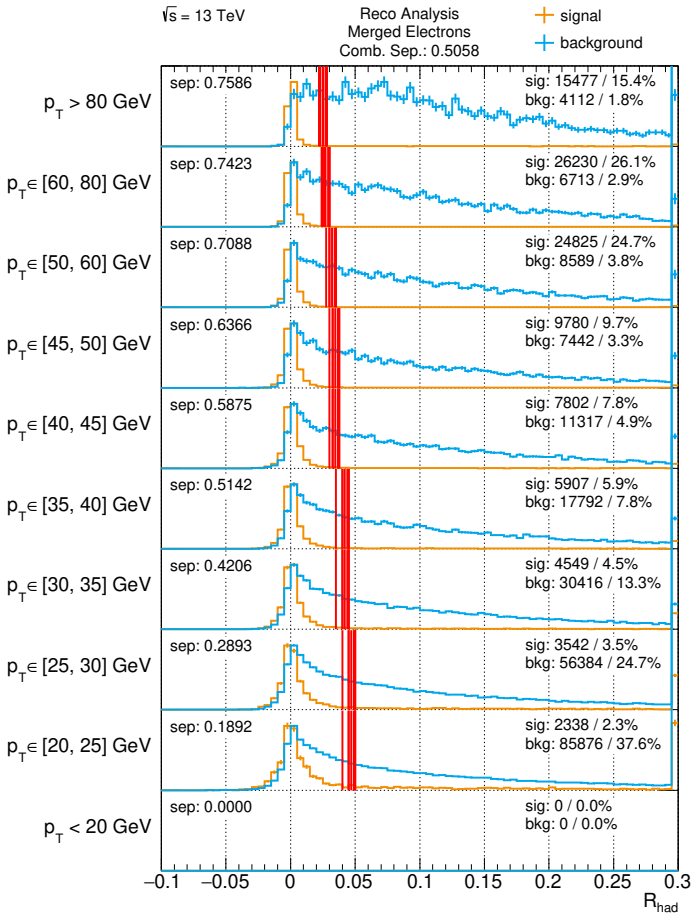


Figure 15.1.: Distribution of the signal and background as a function of R_{had} in bins of p_{T}. The γ^* signal is shown in orange and the background in blue. The vertical red lines show the locations of the cut values. Since the cuts on R_{had} depend on both the p_{T} and the η bin, there are multiple cut values in every single p_{T} bin. The signal and background event yields and their fractions in each bin are given along with the separations $\langle S_i^2 \rangle$ between the signal and background.

which were selected as belonging to the γ^* candidate. Therefore, this variable peaks around one for signal events, whereas it extends over a broader range for the background. The tail towards $E/p > 1$ in the signal is assumed to be due to momentum losses from bremsstrahlung radiated from the electron tracks. While there is still some separation between signal and background at increasingly large values of η, it is not possible anymore to reject background while at the same time retaining the full signal. Therefore, the cuts on E/p were limited to the central region of the detector.

For pile-up conditions corresponding to the years 2015 and 2016, a signal efficiency of 61.8% and a background efficiency of 0.60% were found. For the conditions of 2017, the results are similar, with a signal efficiency of 60.9% and a background efficiency of 0.57%. The signal and background efficiencies are shown as a function of the pile-up in Figure 15.4. The signal efficiency decreases only mildly with increasing pile-up, whereas the background efficiency appears approximately constant across the range. Finally, the signal efficiency is shown as a function of the separation ΔR of the electron tracks at their extrapolated impact location in the strip layer of the ECal, the reconstructed dielectron mass m_{ee}

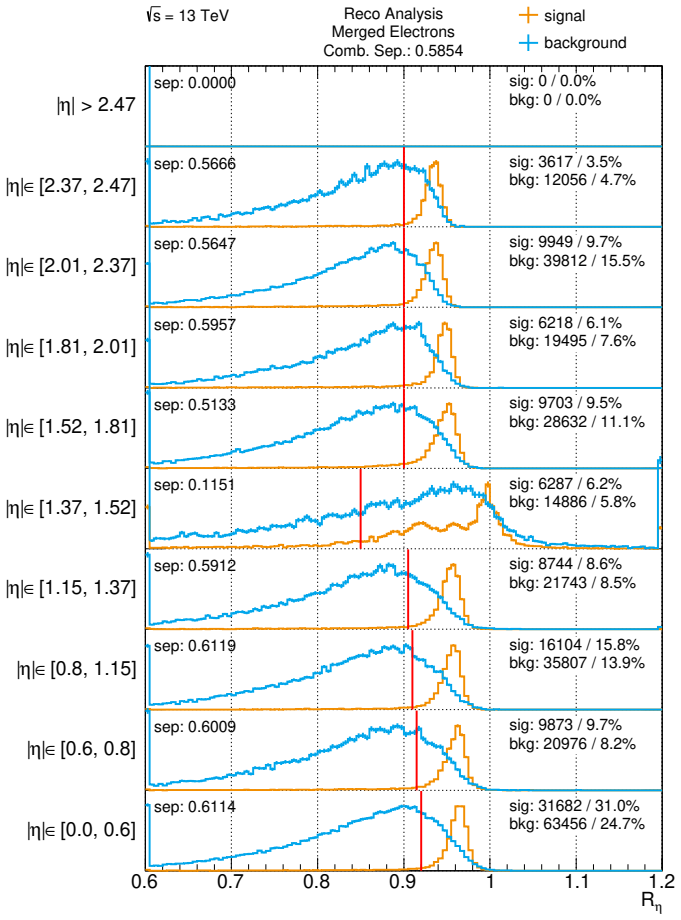


Figure 15.2.: Distribution of the signal and background as a function of R_η in bins of η. The γ^* signal is shown in orange and the background in blue. The vertical red lines show the locations of the cut values. The signal and background event yields and their fractions in each bin are given along with the separations $\langle S_i^2 \rangle$ between the signal and background.

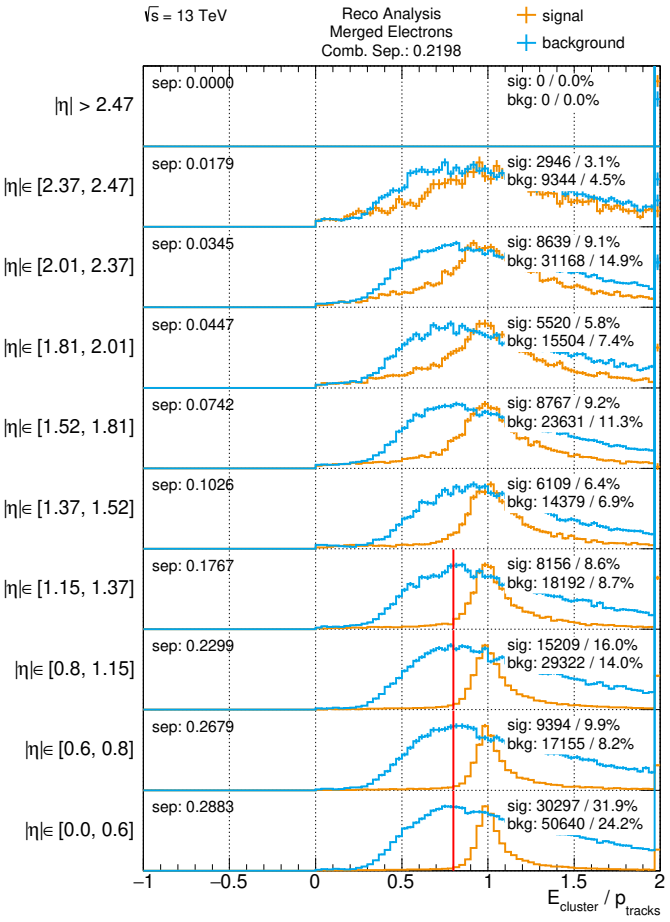
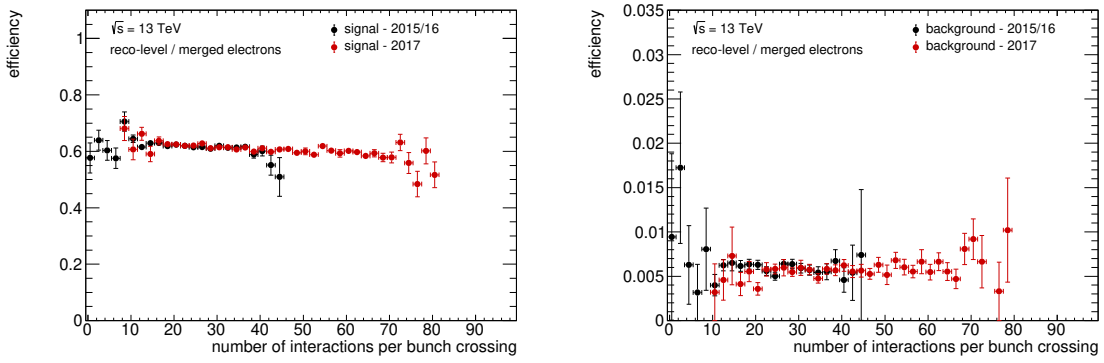


Figure 15.3.: Distribution of the signal and background as a function of E/p in bins of η. The γ^* signal is shown in orange and the background in blue. The vertical red lines show the locations of the cut values. The signal and background event yields and their fractions in each bin are given along with the separations $\langle S_i^2 \rangle$ between the signal and background.



(a) Signal.

(b) Background.

Figure 15.4.: Efficiency of the merged electron identification as a function of pile-up for the ggF signal and the background MC samples.

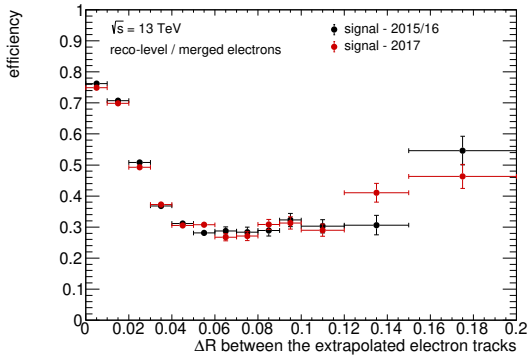
and the reconstructed transverse momentum $p_{\mathrm{T}}(\gamma^*)$ in the left column of Figure 15.5. The normalised distributions of the signal component in the same variables are given for reference in the right column of the same figure. It is worth noting that the efficiency does not decrease towards the regions of sizeable differential cross sections.

15.3. Event Selection and Categorisation

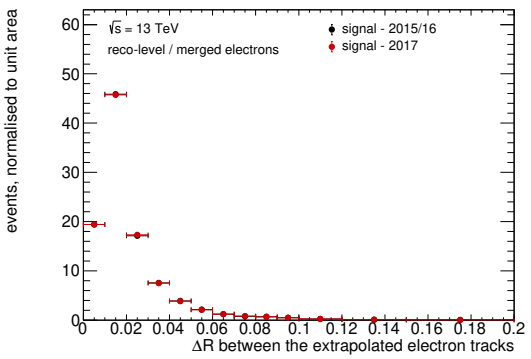
The merged electron ID algorithm presented in the previous section is applied as part of the full event selection and categorisation which is detailed in the following and which is the basis of the expected results given in Chapter 16. In the version of the analysis presented here, no detailed categorisation based on particular kinematic properties is in place yet. Instead, the categorisation is solely based on the flavour and type of the final state, differentiating between two resolved muons, two resolved electrons and the merged electron configurations.

Data collected in the years 2015 through 2017 is considered, leading to a dataset of 79.8 fb^{-1} after requiring a functional detector as encoded in the GRL. The presence of at least one PV candidate is required and the vertex with the largest sum of squared momenta of the associated tracks is selected as PV of the primary hard interaction. Single and dilepton triggers as well as lepton-photon and diphoton triggers are used. The full list of triggers and p_{T} thresholds, broken down by year, is given in Table 15.3. Among the different triggers is a new HLT electron and photon trigger, which was created for this analysis and has been in operation since the beginning of 2017. It targets collimated electron configurations in combination with a photon and therefore has relaxed electron ID requirements, using the variables R_{had}, f_1 and f_3 as well as requirements on track hits. It has p_{T} thresholds of 25 GeV for electrons and 35 GeV for photons.

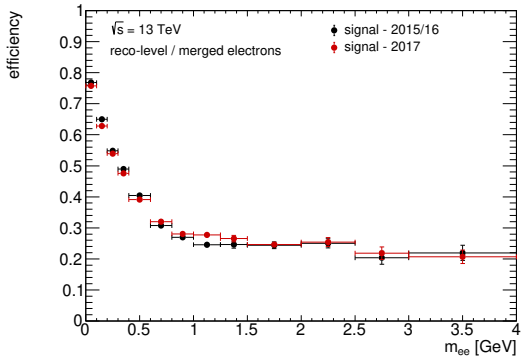
In the next step, the physics objects are preselected. Muons with a transverse momentum of at least 3 GeV, within $|\eta| < 2.7$ and satisfying the *Medium* ID criteria [121] are considered. Electrons are required to have a p_{T} of at least 4.5 GeV and to be within $|\eta| < 2.47$. Furthermore, GSF tracks with a p_{T} of at least 300 MeV, within $|\eta| < 2.47$, with at least 2 PIX and 7 combined PIX and SCT hits and which match a preselected electron are considered. Photons are required to have a transverse momentum of at least



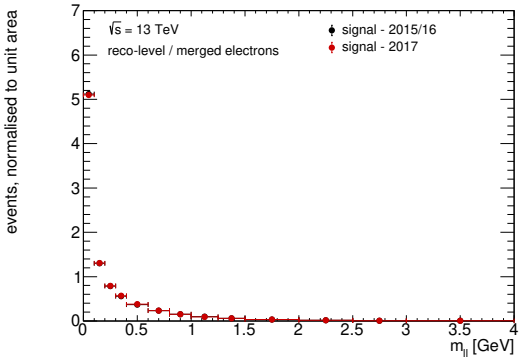
(a) Efficiency as a function of the separation ΔR between the electron tracks extrapolated to the ECal.



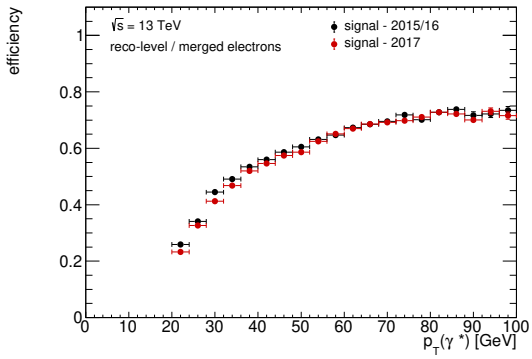
(b) Separation ΔR between the electron tracks extrapolated to the ECal.



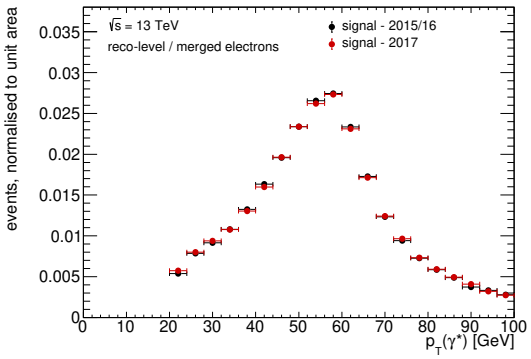
(c) Efficiency as a function of the dielectron mass m_{ee}.



(d) dielectron mass m_{ee}.



(e) Efficiency as a function of the transverse momentum of the electron pair p_{T} (γ^*).



(f) Transverse momentum of the electron pair p_{T} (γ^*).

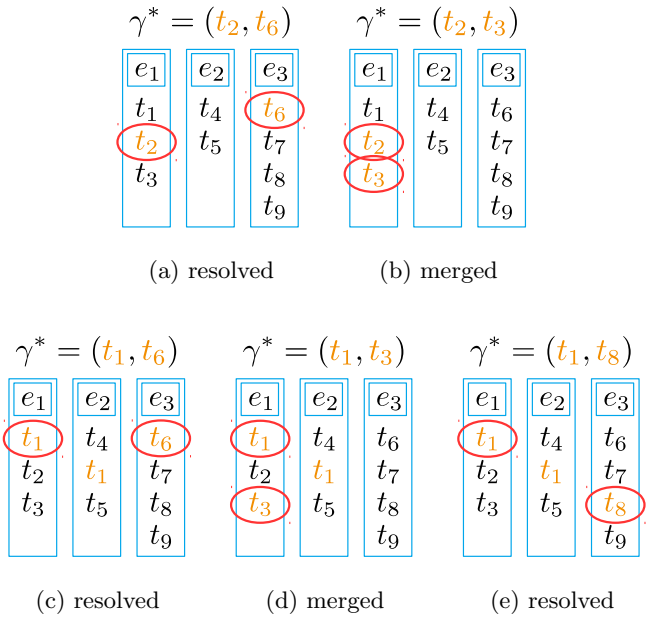
Figure 15.5.: Efficiency of the merged electron identification as a function of the separation ΔR between the tracks extrapolated to the ECal, the dielectron mass m_{ee} and the transverse momentum $p_{\mathrm{T}}(\gamma^*)$ (left column) and of the distributions of the corresponding observables (right column) for the ggF signal. The distributions in the right column are normalised to unit area. All observables are at the reconstruction level.

Trigger type	Trigger p_{T} thresholds per year [GeV] 2015	2016	2017	Final analysis p_{T} thresholds [GeV]
$1e$	24	26	26	27
	60	60	60	61
	120	140	140	121/141
$2e$	12, 12	17, 17	24, 24	25, 25
$e\gamma$	20, 35	20, 35	25, 35	21/26, 36
$\gamma\gamma$	35, 25	35, 25	35, 25	36, 26
1μ	20	26	26	27
	40	50	50	51
2μ	10, 10	14, 14	14, 14	15, 15
	18, 8	22, 8	22, 8	23, 9

Table 15.3.: Summary of the different triggers and the corresponding final lepton and photon p_{T} thresholds in the analysis. Single lepton and dilepton triggers as well as lepton-photon and diphoton triggers are used. For two-object triggers, the first threshold refers to the first and the second threshold to the second object. It should be noted that some of the triggers with different p_{T} thresholds also differ in other aspects, in particular in particle ID requirements. The final analysis p_{T} thresholds are larger than the corresponding trigger thresholds to avoid trigger turn-on effects.

10 GeV, $|\eta| < 2.37$ and to fulfil the *Loose* ID criteria [242]. Subsequently, the presence of at least two opposite-sign muons or tracks as well as at least one photon is required. The highest vector-sum-p_{T} opposite-sign muon or track pair is preselected as the γ^* candidate. If both muon and track pairs are present, preference is given to the muon pairs. Similarly, the highest-p_{T} photon is preselected.

Next, given the γ^* candidate, the event category is determined. Besides the *muon* final state, the categorisation differentiates between *resolved* and *merged* electrons. This differentiation is based on a matching of the tracks to the electron clusters. More specifically, the tracks are extrapolated to the ECal and their impact locations are compared against the electron's cluster barycentre position in the second layer of the ECal [141]. A single electron object may have multiple tracks associated to it, ranked according to this geometrical matching criterion as well as the number of hits in the PIX and SCT. It should also be pointed out that the extrapolation of the tracks to the ECal is not based on the track momenta but rather on the cluster energy in order to reduce the impact of bremsstrahlung. If the two tracks forming the γ^* candidate are each only associated to one single electron, the classification is straightforward: If the tracks are associated to different electrons, the event is classified as *resolved* (ref. Figure 15.6(a)) and if they are associated to the same electron object, the event is classified as *merged* (ref. Figure 15.6(b)). Otherwise, i.e. if at least one of the tracks is associated to more than one electron, it is checked if the respective track is the best-ranked, i.e. primary, track for some of the electrons it is associated to. In case a given track is the primary track of more than one electron, the electron with the largest p_{T} is chosen. If both tracks are the primary tracks of some electrons, the event is classified as *resolved* (ref. Figure 15.6(c)). If only one of the two tracks is the primary



(a) resolved (b) merged

(c) resolved (d) merged (e) resolved

Figure 15.6.: Schematic of the categorisation. Each electron may have multiple tracks associated to it, ranked according to a geometrical matching between the track impact location in the ECal and the electron cluster barycentre as well as the number of hits in the PIX and SCT. The tracks forming the γ^* candidate are indicated in orange. The tracks from which the category is determined are circled in red. A more detailed description of the categorisation and the different cases is given in the text.

track of an electron, then it is checked if the other track is among the lower-ranked tracks associated to the same electron. In that case, the event is classified as *merged* (ref. Figure 15.6(d)). If the other track is among the lower-ranked tracks of just one single other electron, then the event is classified as *resolved* (ref. Figure 15.6(e)). If none of these criteria are met, the event is discarded, which happens for less than 1% of the events.

Once the category is determined, additional preselection cuts are applied to the γ^* candidate: For resolved electrons, the *VeryLoose Likelihood* ID working point [242] is required for both electrons. Merged electrons need to have a transverse momentum of at least 20 GeV, the energy fraction in the hadronic calorimeter, R_{had}, should be smaller than 0.1, and at least one of the two tracks should have hits in the two innermost PIX layers. For γ^* candidates built from muons, no additional cuts are applied. If the γ^* candidate fails these criteria, it is discarded and the next track pair candidate is considered in the same way.

In the next step, an overlap removal procedure is applied: Electrons and jets are removed if they are within $R < 0.4$ of a photon. Jets are removed if they are within $R < 0.2$ of an electron, and electrons if they are within $R < 0.4$ of a jet. Muons are removed if they are within $R < 0.4$ of a photon or a jet. Further ID, isolation and track-to-vertex association criteria are applied in the final selection: For muons, the *medium* ID and the *CloseByCorrected*[1] *GradientLoose* ISO working points [121] are required. The criteria on

[1] In the close-by correction, contributions from the second lepton are removed from the isolation variables of the first lepton and vice versa.

- GRL, primary vertex, triggers
- preselection
  - muons: $p_{\mathrm{T}} > 3$ GeV, $|\eta| < 2.7$, ID: *Medium*
  - electrons: $p_{\mathrm{T}} > 4.5$ GeV, $|\eta| < 2.47$
  - tracks: GSF tracks, $p_{\mathrm{T}} > 0.3$ GeV, $|\eta| < 2.47$, ≥ 2 PIX hits, ≥ 7 PIX+SCT hits
  - photons: $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.37$, ID: *Loose*
- ≥ 2 opposite-sign muons or electron tracks, ≥ 1 photon
  - preselect the highest vector-sum-p_{T} opposite-sign muon or track pair as the γ^* candidate
  - preselect the highest-p_{T} photon
- preselection cuts on the γ^* candidate
  - muons: no additional cuts
  - resolved electrons: ID: *VeryLoose Likelihood*
  - merged electrons: $p_{\mathrm{T}} > 20$ GeV, $R_{\mathrm{had}} < 0.1$, hits in the innermost PIX layers
  - if the preselection cuts are not passed, a new γ^* candidate is picked
- overlap removal
- final selection
  - muons: ID: *Medium*, ISO: *CloseByCorrected GradientLoose*, $|d_0|/\sigma_{d_0} < 3$, $|z_0 \sin\theta| < 0.5$ mm
  - resolved electrons: ID: *Medium*, ISO: *CloseByCorrected Loose*, $|d_0|/\sigma_{d_0} < 5$, $|z_0 \sin\theta| < 0.5$ mm
  - merged electrons: ID: *Merged*, ISO: *Loose*, $|d_0|/\sigma_{d_0} < 5$, $|z_0 \sin\theta| < 0.5$ mm
  - photons: ID: *Tight*, ISO: *FixedCutLoose*
- $m_{\ell\ell} < 45$ GeV and 105 GeV $\leq m_{\ell\ell\gamma} \leq 160$ GeV

Table 15.4.: Overview of the event selection strategy in the $H \to \gamma^*(\ell^+\ell^-)\gamma$ search. The electron and photon ID and ISO working points can be found in [242] and the muon working points in [121].

the interaction point are $|d_0|/\sigma_{d_0} < 3$ and $|z_0 \sin\theta| < 0.5$ mm. For resolved electrons, the *Medium Likelihood* ID and *CloseByCorrected*[1] *Loose* ISO working points [242] are used along with the IP criteria $|d_0|/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm. The same IP criteria are also used for merged electrons, but for them, the *Merged* electron ID presented in Section 15.2 is used in combination with the *Loose* ISO working point [242]. For photons, the *Tight* ID and *FixedCutLoose* ISO working points are required [242]. The event selection is concluded by requiring $m_{\ell\ell} < 45$ GeV and 105 GeV $\leq m_{\ell\ell\gamma} \leq 160$ GeV.

Chapter 16

Statistical Analysis and Results

As mentioned in Chapter 6, an important question before carrying out a full analysis on real data is to quantify the expected sensitivity. In the case of a search, the expected significance and the expected limits on the parameter of interest, such as the signal strength or the cross section times branching ratio, can be determined. This is not only interesting in its own right, but can also be used as a baseline when further optimising the analysis and as a cross-check when determining the observed results based on a real dataset. In this chapter, the strategy for the statistical analysis in the search for the $H \to \gamma^*\gamma$ decay will be discussed. In accordance with the $H/X \to Z\gamma$ analysis described in Part III, an unbinned profile likelihood method is used. In the studies presented in the following, no systematic uncertainties were considered. These will be added to the analysis at a later point in time. However, while it is important to eventually include the systematic uncertainties, they are not expected to play a major role and hence, the statistics-only results presented in this thesis should still be a reasonable approximation. In Section 16.1 and 16.2, the signal and background modelling is discussed, followed by the determination of the expected sensitivity and upper limits on the signal strength in Section 16.3. The chapter is concluded with a discussion of the projected results for larger future datasets in Section 16.4.

16.1. Signal Modelling

An analytic signal model is built based on the Monte Carlo simulation of the ggF production mode with the `NNLOPS` method, described in Section 15.1.[1] Separate models are built for the different categories, i.e. the muon, the resolved and the merged electron channels, and for the 2015-2016 and 2017 run periods. The event samples are normalised to the integrated luminosities of both periods in order to derive the expected signal event yields,

[1] The limitation to the ggF production mode is imposed by the availability of Monte Carlo event samples. Simulations of additional production modes are expected to become available and will be added in the future.

which are given in Table 16.1. Furthermore, the N^{3}LO cross section from [33] quoted in Table 2.1 and the branching ratios given in Equations (14.1) and (14.2) are assumed.

An unbinned likelihood function is used to fit individual signal models to the simulated ggF sample for all categories and run periods considered. A fixed Higgs mass value of $m_H = 125.0$ GeV is used. Different functional forms were tried and the double-sided crystal ball function was found to provide a good description of the expected $m_{\ell\ell\gamma}$ signal distributions. These distributions, as well as the DSCB fits to them and the resulting parameter values, are shown separately for the different categories and run periods in Figure 16.1.

When comparing the 2015-2016 and 2017 run periods, the signal mass resolutions, described by the σ_{CB} parameters, deteriorate by approximately 4% in the muon and resolved electron categories and by about 6% in the merged electron category. This is likely caused by the increase in the number of pile-up interactions over the aforementioned run periods. The peak positions, encoded by the μ_{CB} parameters, are stable for both periods, but are closer to the Higgs mass in the muon and resolved electron categories. In the version of the analysis described here, the normal electron calibration was used for merged electrons. It was later found that using the calibration for converted photons results in an improvement of both the peak position and the mass resolution. Finally, it can be seen that the fitted functions overshoot the Monte Carlo predictions in the high invariant mass tails above 132 GeV in the muon categories. The reason for this is not yet known, but only a vanishingly small amount of the signal falls into this region.

16.2. Background Modelling

Data events in the three-body invariant mass range 105 GeV $\leq m_{\ell\ell\gamma} \leq$ 160 GeV are considered, excluding events in a mass window from 120 GeV to 130 GeV so as to blind the region containing the expected signal. An extended unbinned likelihood fit is performed separately in the muon, resolved and merged electron categories and for the 2015-2016 and 2017 run periods. Falling exponential functions of the form $ne^{am_{\ell\ell\gamma}}$ with $a < 0$ are found to provide a good description of the data sidebands in all categories. All data mass distributions as well as the fitted background models and the corresponding normalisation and slope parameters n and a are shown in Figure 16.2. The observed data event yields in the sidebands as well as the expected background event yields in the full invariant mass range are also summarised in Table 16.1.

The careful assessment of the background is of great importance since, due to the small expected signal to background ratio, already a small mismodelling of the background can have a large impact on the inferred observed signal yield. A satisfactory fit result in the data sidebands is not sufficient to ensure an unbiased signal estimation. For this reason, the spurious signal was evaluated in the $H \to Z\gamma$ analysis as described in Section 11.4. Therefore, as the $H \to \gamma^*(\ell^+\ell^-)\gamma$ analysis evolves further, the procedure described in this section will be replaced by a more sophisticated approach.

16.3. Expected Sensitivity

The expected results are summarised in Table 16.1. The significances Z are calculated from Asimov datasets using the uncapped test statistic $q_0^{(\text{uncapped})}$ as described in Section 6.2.3.1.

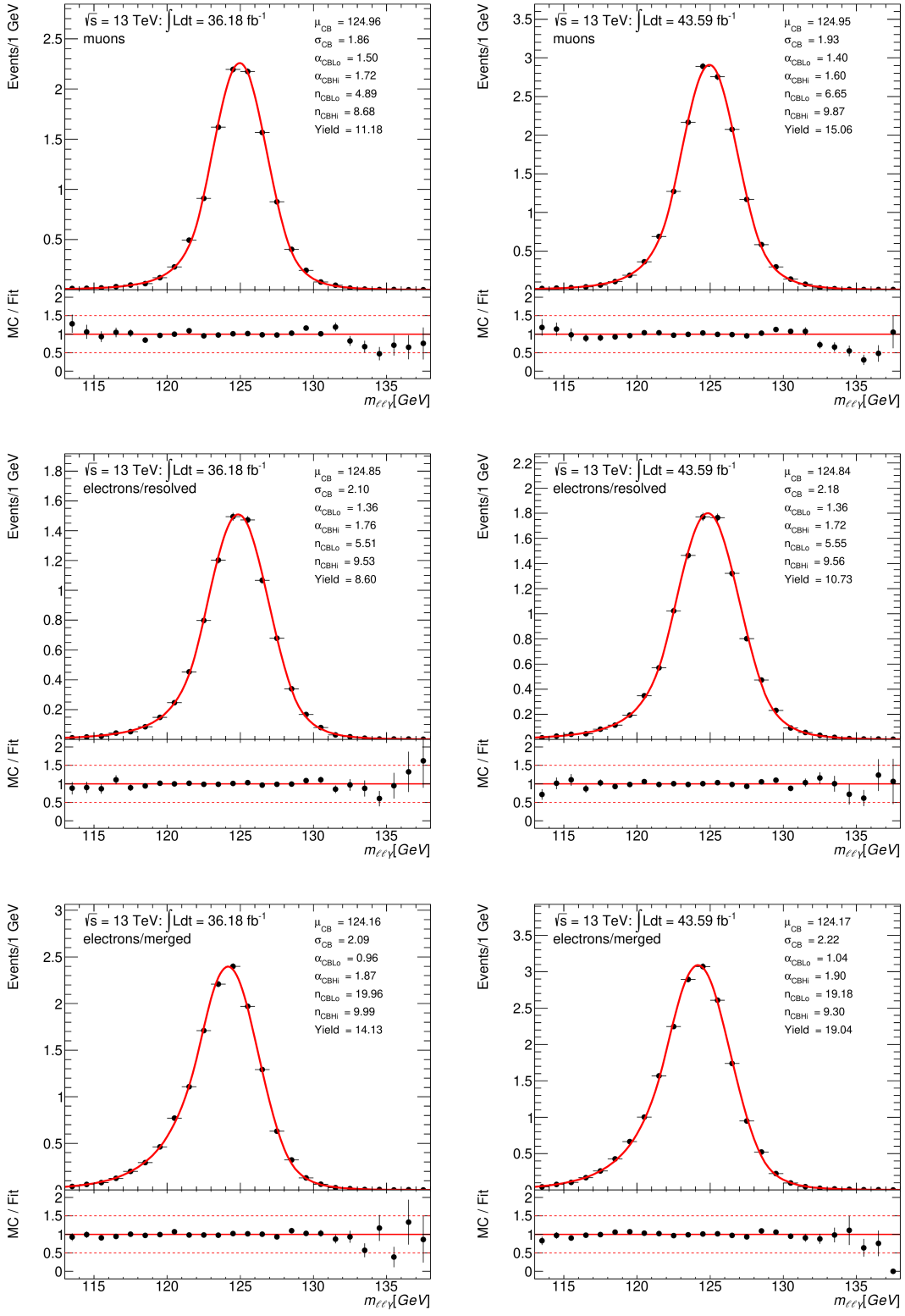


Figure 16.1.: Signal parameterisation for the 2015-2016 (left column) and 2017 (right column) run period for the muon, resolved and merged electron categories. The chosen signal model is a double-sided crystal ball function fitted separately in each category. The model is fitted to MC events using an unbinned likelihood fit. The events are only binned for display purposes and are shown as black data points, with the error bars indicating the statistical uncertainty. The obtained signal model curves are shown with solid red lines.

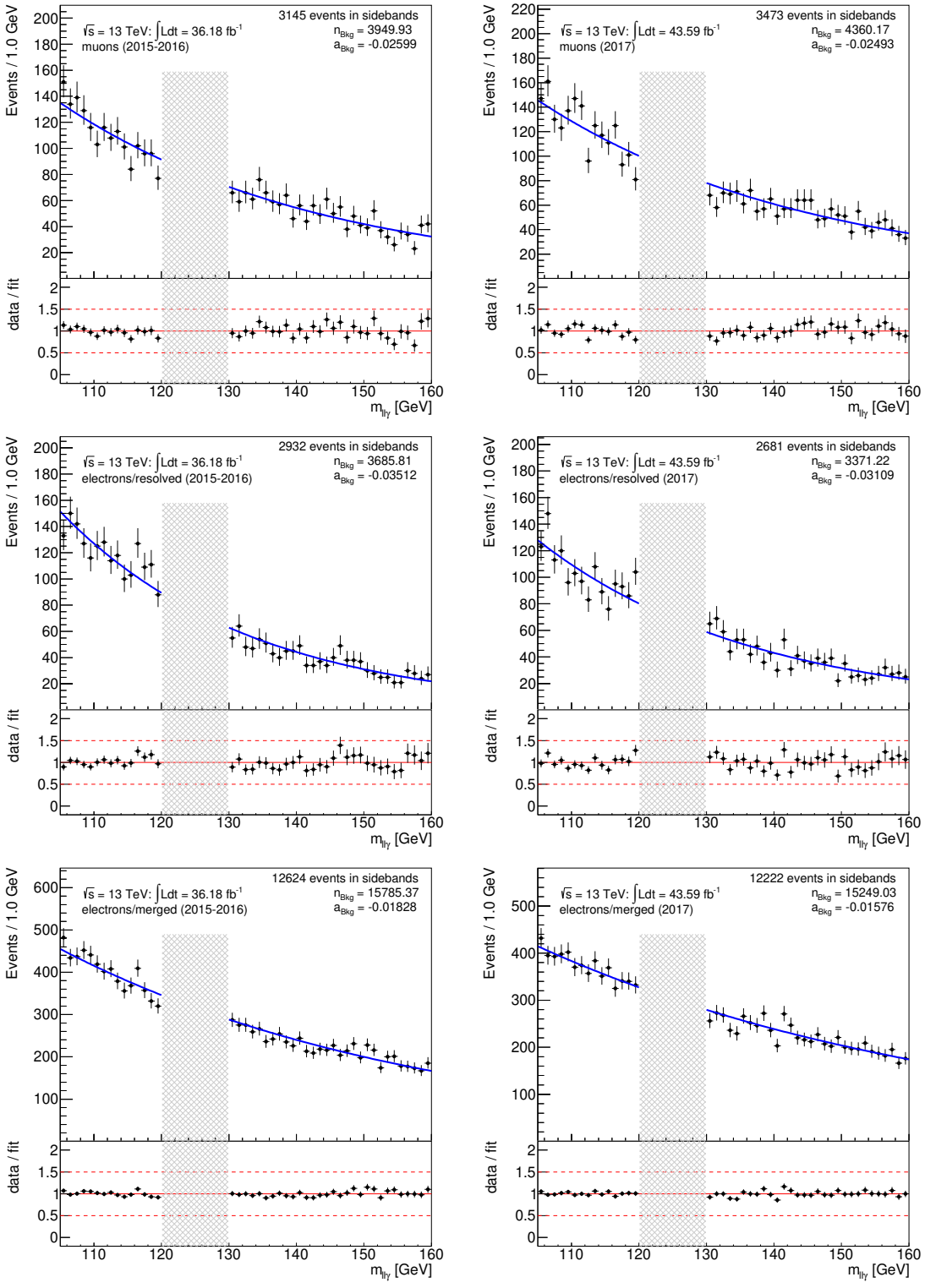


Figure 16.2.: Background parameterisation for the 2015-2016 (left column) and 2017 (right column) run period for the muon, resolved and merged electron categories. The chosen background model is a falling exponential function fitted separately in each category. The model is fitted to data events using an extended unbinned likelihood fit. Events with three-body invariant masses $m_{\ell\ell\gamma}$ between 120 and 130 GeV are excluded. The events are only binned for display purposes and are shown as black data points, with the error bars indicating the statistical uncertainty. The obtained background model curves are shown with solid blue lines.

They are cross-checked with the ratios $S_P/\sqrt{B_P}$, where the signal and background event yields S_P and B_P are calculated from the corresponding analytic models and evaluated in the invariant mass region for $m_{\ell\ell\gamma}$ from 122 GeV to 128 GeV. Depending on the category, this window contains between 72% and 84% of the total signal yields. As can be seen from Table 16.1, the ratios $S_P/\sqrt{B_P}$ and the proper significances Z agree well in each category. The merged electron categories have the largest event yields, but suffer from worse signal-to-background ratios than the other categories. As a consequence, the muon categories have the highest significance, followed by the resolved electron categories. For the combined results for each run period and for the total combined results, the ratios $S_P/\sqrt{B_P}$ are calculated from the corresponding sums of the signal and background event yields. Therefore, a comparison of the total $S_P/\sqrt{B_P}$ and significance Z indicates that the categorisation results in a gain of almost 15% in sensitivity. The overall combined expected significance is 0.9σ.

Since no significant excess over the background-only hypothesis is expected in the SM, the expected upper limits $\mu_{95\%\,\mathrm{CL}}$ on the signal strength μ were calculated. They are determined from Asimov datasets using the $\tilde{q}_\mu$ test statistic introduced in Section 6.2.3.3 with the $\mathrm{CL_s}$ method applied. Assuming no SM $H \to \gamma^*(\ell^+\ell^-)\gamma$ decay, a combined expected upper limit at the 95% CL of 2.1 was found. Assuming the presence of this SM Higgs decay, an expected upper limit of 2.9 was found.

run period	integrated luminosity [fb^{-1}]	category	data event yields in sidebands	event yields S	event yields B	$\frac{S}{B}$	expected event yields S_P	expected event yields B_P	$\frac{S_P}{B_P}$	$\frac{S_P}{\sqrt{B_P}}$	significance Z	upper limits $\mu^{(\mu=0)}_{95\%\,\mathrm{CL}}$	upper limits $\mu^{(\mu=1)}_{95\%\,\mathrm{CL}}$
2015-2016	36.18	muons	3145	11.18	3950	0.283%	9.43	482	1.96%	0.43	0.43	4.76	5.50
		electrons/resolved	2932	8.60	3686	0.233%	6.82	451	1.51%	0.32	0.31	6.48	7.21
		electrons/merged	12624	14.13	15785	0.090%	10.34	1895	0.55%	0.24	0.25	8.45	9.15
		combined	18701	33.91	23421	0.145%	26.59	2828	0.94%	0.50			
2017	43.59	muons	3473	15.06	4360	0.345%	12.42	531	2.34%	0.54	0.53	3.82	4.57
		electrons/resolved	2681	10.73	3371	0.318%	8.31	413	2.01%	0.41	0.39	5.10	5.84
		electrons/merged	12222	19.04	15249	0.125%	13.77	1816	0.76%	0.32	0.31	6.20	6.91
		combined	18376	44.83	22980	0.195%	34.50	2760	1.25%	0.66			
total	79.77		37077	78.74	46402	0.170%	61.09	5588	1.09%	0.82	0.94	2.13	2.91

Table 16.1.: Expected signal and background event yields, significances Z and upper limits at 95% CL on the signal strength μ for the different channels and run periods. The signal and background event yields S and B are given for a mass window in $m_{\ell\ell\gamma}$ from 105 GeV to 160 GeV, the event yields observed in data are determined in the sidebands when excluding a central mass window for $m_{\ell\ell\gamma}$ from 120 GeV to 130 GeV. The signal and background event yields S_P and B_P are given for a mass window for $m_{\ell\ell\gamma}$ from 122 GeV to 128 GeV. The expected signal event yields S and S_P are obtained from ggF MC samples scaled to the integrated luminosities of the different run periods. The expected background event yields B and B_P are determined from a fit to the data sidebands of the $m_{\ell\ell\gamma}$ invariant mass distribution. The significances Z are calculated from Asimov datasets using the uncapped test statistic $q_0^{(\mathrm{uncapped})}$ as described in Section 6.2.3.1. In each category, the ratio $S_P/\sqrt{B_P}$ is a good approximation of the proper significance Z. For the combined results for each run period and for the total combined results, the ratios $S_P/\sqrt{B_P}$ are calculated from the corresponding sums of signal and background yields. Therefore, a comparison of the total $S_P/\sqrt{B_P}$ and significance Z indicates that the categorisation results in a gain of almost 15% in sensitivity. The upper limits $\mu_{95\%\,\mathrm{CL}}$ on the signal strength are likewise determined from Asimov datasets using the $\tilde{q}_\mu$ test statistic with the $\mathrm{CL_s}$ method applied in addition.

run period	integrated luminosity per period [fb^{-1}]	integrated luminosity accumulated [fb^{-1}]	LHC CM energy $\sqrt{s}$ [TeV]	remarks
2015-2016	36.18	36.18	13	
2017	43.59	79.77	13	
2018	62.20	141.97	13	projection based on 2017 conditions
end of Run 3	158.03	300.00	13 / 14	projection based on 2017 conditions
HL-LHC	858.03	1000.00	13 / 14	
	2858.03	3000.00	13 / 14	

Table 16.2.: Scenarios for the projections for the $H \to \gamma^*(\ell^+\ell^-)\gamma$ analysis. The projections are based on the conditions and expected results for the year 2017. Integrated luminosities of up to 3000 fb^{-1} and LHC centre-of-mass energies of 13 and 14 TeV are considered.

16.4. Projections for Larger Datasets

Based on the expected results presented in the previous section and using the same statistical procedures, projected results were derived for total integrated luminosities of up to 3000 fb^{-1}. This corresponds to the dataset expected by the end of the LH-LHC [243]. Projections for intermediate datasets were also derived for integrated luminosities of 300 fb^{-1}, expected by the end of LHC Run 3 in 2023 [243], and 1000 fb^{-1}. Due to restrictions in the availability of simulated datasets, the projections are based on the conditions and expected results for the year 2017. LHC centre-of-mass energies of both 13 and 14 TeV were considered for the time after 2018. The different scenarios are summarised in Table 16.2. The signal and background event yields were scaled according to the increased total integrated luminosities, while all remaining parameters describing the signal and background models were kept constant, which corresponds to assuming constant efficiencies and resolutions. In the past, such assumptions have often proven reasonable because challenges from increased pile-up could be compensated by improvements in the object reconstruction and the analyses. For an increased centre-of-mass energy of 14 TeV, the signal event yields were also scaled by the ratio of the Higgs production cross sections at 14 and 13 TeV, giving rise to an additional increase of 12.2%. No detailed studies were carried out for the background event yields, which were kept fixed at the yields expected for an energy of 13 TeV. While this on its own may be considered overly optimistic, this assumption was deemed reasonable because of expected improvements to the analysis design and categorisation. Furthermore, the differences in the expected upper limits between the scenarios at 13 and 14 TeV were found to be smaller than the 1σ statistical uncertainty on these limits. Likewise, changes in acceptance due to different event kinematics at 13 and 14 TeV were not considered, but are not expected to play a major role.

The projection for the expected significance is given in Figure 16.3. The data points for which projected significances were calculated are interpolated using square root functions. Depending on the scenario, evidence (observation), i.e. an expected significance of 3σ (5σ), of the $\gamma^*(\ell^+\ell^-)\gamma$ decay mode of the SM Higgs boson is expected for a dataset corresponding to a total integrated luminosity of approximately 700 fb^{-1} (1700 fb^{-1} – 2000 fb^{-1}). The projection for the expected upper limits at the 95% CL is shown in Figure 16.4. Inverse square root functions were used for the interpolation between the results expected at the different integration luminosities considered. In the absence of

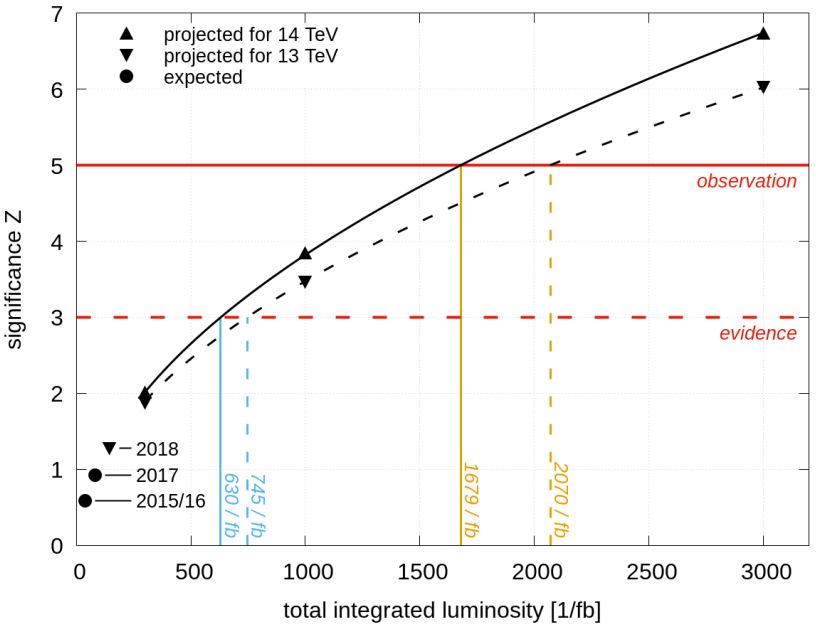


Figure 16.3.: Projections for the expected significance of the $H \to \gamma^*(\ell^+\ell^-)\gamma$ analysis, based on the conditions and expected results for the year 2017. The labels beside some of the data points indicate the year when the corresponding integrated luminosity was reached. Thus, e.g., the 2018 data point refers to the integrated luminosity of the dataset collected in the years 2015 through 2018. LHC centre-of-mass energies of 13 and 14 TeV were considered and evidence (observation) is expected to be reached for approximately 700 fb^{-1} (1700 fb^{-1} – 2000 fb^{-1}).

this decay mode of the SM Higgs boson ($\mu = 0$), the expected upper limit on the signal strength μ reaches one for a total integrated luminosity of approximately 300 fb^{-1}, which would lead to an exclusion of this decay mode at the 95% CL. For all of this it should be noted that the projections are based on the same simple categorisation into the muon, resolved and merged electron final states as the expected results presented in the previous section. In the future, improvements are expected from a more detailed categorisation.

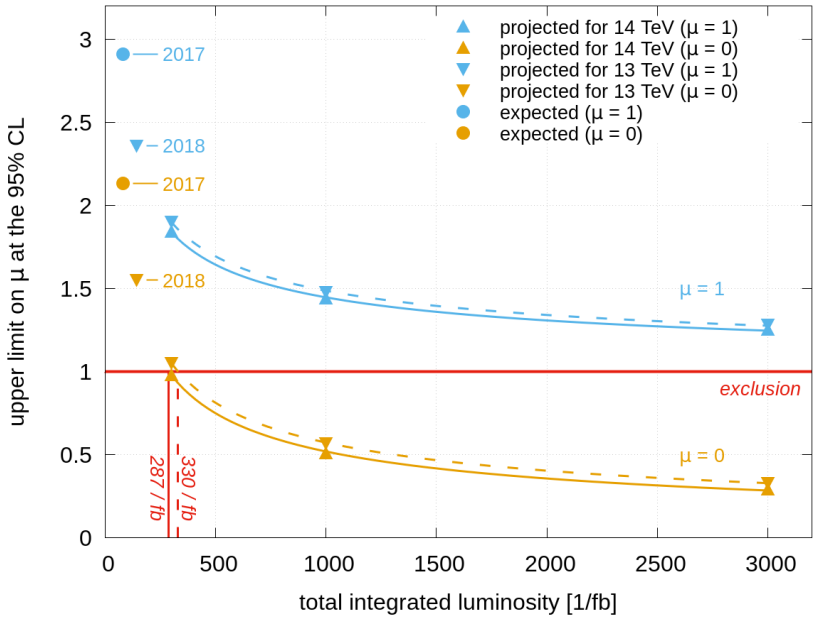


Figure 16.4.: Projections for the expected upper limits in the $H \to \gamma^*(\ell^+\ell^-)\gamma$ analysis, both assuming the absence (presence) of this decay mode of the SM Higgs boson, i.e. assuming $\mu = 0$ ($\mu = 1$). The projections are based on the conditions and expected results for the year 2017. LHC centre-of-mass energies of 13 and 14 TeV were considered. The labels beside some of the data points indicate the year when the corresponding integrated luminosity was reached. Thus, e.g., the 2018 data point refers to the integrated luminosity of the dataset collected in the years 2015 through 2018.

Chapter 17

Conclusions

The discovery of the Higgs boson was one of the main goals for the Large Hadron Collider at CERN. In July 2012, the observation of a new particle in the searches for the Higgs boson was announced by the ATLAS and CMS collaborations. Ever since then, studies of the new particle's properties have been conducted and shown consistency with the expectation for the Standard Model Higgs boson. At a proton-proton collider such as the LHC, the Higgs boson can be created in a variety of production processes ranging from the dominant gluon-gluon fusion mode to the much rarer production in association with heavy quarks. Similarly, due to its large mass and its couplings to massive bosons and fermions, the Higgs boson features a wide range of decay channels. Many of its decay modes have been observed experimentally and exploited in property measurements, whereas some remain undetected as of today. This thesis is dedicated to aspects of rare production and decay processes of the Higgs boson.

Concerning Higgs production, a Monte Carlo study of the $t\bar{t}H(b\bar{b})$ background process $pp \to t\bar{t}c\bar{c}$ were carried out. NLO event generation with massive charm quarks in the matrix elements was explored and different options for the renormalisation and factorisation scales were tested. The resulting predictions were compared against a sample of inclusive $t\bar{t}$ production where the charm quarks are produced in the parton shower and against a multi-jet merged sample with matrix elements of different parton multiplicities including massless charm quarks. The study shows that there are sizeable differences between the different samples and that the hard process scale variations for inclusive $t\bar{t}$ production used in some experimental analyses do not reflect the much larger perturbative uncertainties of $t\bar{t}c\bar{c}$ matrix elements. As a consequence of this study, an additional uncertainty on the $t\bar{t}c\bar{c}$ background was added to the $t\bar{t}H(b\bar{b})$ analysis.

Apart from a small production cross section, a rare Higgs signal with a given signature can stem from a low branching fraction. In this thesis, the rare decay modes $H \to Z(\ell^+\ell^-)\gamma$ and $H \to \gamma^*(\ell^+\ell^-)\gamma$ were studied. A search for the process $H \to Z(\ell^+\ell^-)\gamma$ was performed with a dataset of 36.6 fb^{-1}. No excess of events above the background was observed and an upper limit at the 95% CL on the cross section times branching fraction of 6.6 times the expectation for the Standard Model Higgs boson with a mass of 125.09 GeV was derived.

The same final state was exploited to look for narrow high-mass resonances. No localised excess of events above the background was observed and upper limits on the production cross section times branching ratio were derived for resonance masses between 250 GeV and 2.4 TeV for different spin hypotheses and production mechanisms. Focusing on the region of lower invariant dilepton masses of the $\ell^+\ell^-\gamma$ final state, a preliminary, blinded search for the process $H \to \gamma^*(\ell^+\ell^-)\gamma$ was set up for the first time with the ATLAS detector. A dedicated identification algorithm was developed for close-by electrons and, as a result, both the electron and the muon final state were used to derive expected upper limits at the 95% CL on the cross section times branching fraction of 2.9 (2.1) times the Standard Model expectation for a dataset of 80 fb^{-1} assuming the presence (absence) of this Higgs boson decay mode. These expected results were derived without taking systematic uncertainties into account. The projected results indicate sensitivity for evidence (observation), corresponding to a significance of 3σ (5σ), for a dataset with an integrated luminosity of approximately 700 fb^{-1} (2000 fb^{-1}) when assuming the SM decay rate.

After the current shutdown, the LHC is scheduled to resume operation in 2021, likely at an increased centre-of-mass energy of 14 TeV. In the following years of operation, and in particular in the High-Luminosity LHC phase, the collected datasets will increase dramatically, providing excellent opportunities for exploring further the rich field of Higgs physics. The $H \to Z\gamma$ and $H \to \gamma^*\gamma$ analyses will benefit from the larger amounts of data available, but further improvements and optimisations are still desirable to make the most of the collected collision events. For example, the currently used photon identification procedure is based on a sequence of cuts that do not take into account correlations between the different variables considered. Using multivariate or machine learning techniques instead could be benefitial to both analyses. While many searches and measurements have been performed already, revealing and constraining the properties of the Higgs boson, the field of Higgs physics will continue to be very interesting for years to come.

Bibliography

[1] The 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. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[2] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. Phys.Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[3] *CERN LHC sees high-energy success*, https://home.cern/news/press-release/cern/lhc-sets-new-world-record. Retrieved on August 23rd, 2019.

[4] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*. Phys. Rev. Lett. **13** 9 (1964) 321–323.

[5] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*. Phys. Rev. Lett. **13** 16 (1964) 508–509.

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*. Phys. Rev. Lett. **13** 20 (1964) 585–587.

[7] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*. Chin. Phys. **C40** (2016) no. 10, 100001.

[8] F. Tanedo, *Helicity, Chirality, Mass, and the Higgs*, https://www.quantumdiaries.org/2011/06/19/helicity-chirality-mass-and-the-higgs/. Retrieved on Oct 25th, 2018.

[9] *Standard Model of Elementary Particles*, https://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg. Retrieved on July 16th, 2018.

[10] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Addison-Wesley, Reading, 1995.

[11] *Renormalised vertex*, https://upload.wikimedia.org/wikipedia/commons/2/26/Renormalized-vertex.png. Retrieved on July 18th, 2018.

[12] A. Deur, S. J. Brodsky, and G. F. de Teramond, *The QCD Running Coupling*. Prog. Part. Nucl. Phys. **90** (2016) 1–74, arXiv:1604.08082 [hep-ph].

[13] *Coupling constant - QCD scale*, https://en.wikipedia.org/wiki/Coupling_constant#QCD_scale. Retrieved on July 17th, 2018.

[14] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte, *Parton distribution functions, α_s, and heavy-quark masses for LHC Run II*. Phys. Rev. **D96** (2017) no. 1, 014011, arXiv:1701.05838 [hep-ph].

[15] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*. Phys. Rev. **D93** (2016) no. 3, 033006, arXiv:1506.07443 [hep-ph].

[16] The H1 and ZEUS Collaborations, *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*. Eur. Phys. J. **C75** (2015) no. 12, 580, arXiv:1506.06042 [hep-ex].

[17] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*. Eur. Phys. J. **C75** (2015) no. 5, 204, arXiv:1412.3989 [hep-ph].

[18] The NNPDF Collaboration, R. D. Ball *et al.*, *Parton distributions from high-precision collider data*. Eur. Phys. J. **C77** (2017) no. 10, 663, arXiv:1706.00428 [hep-ph].

[19] F. Krauss and D. Napoletano, *Towards a fully massive five-flavor scheme*. Phys. Rev. **D98** (2018) no. 9, 096002, arXiv:1712.06832 [hep-ph].

[20] S. L. Glashow, *Partial-symmetries of weak interactions*. Nucl. Phys. **22** (1961) no. 4, 579 – 588.

[21] The UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*. Phys. Lett. **B122** (1983) 103–116.

[22] The UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider*. Phys. Lett. **B126** (1983) 398–410.

[23] The UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider*. Phys. Lett. **B122** (1983) 476–485.

[24] The UA2 Collaboration, *Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider*. Phys. Lett. **B129** (1983) 130–140.

[25] J. Schwinger, *Gauge Invariance and Mass*. Phys. Rev. **125** 1 (1962) 397–398.

[26] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*. Phys. Rev. **130** 1 (1963) 439–442.

[27] S. Weinberg, *A Model of Leptons*. Phys. Rev. Lett. **19** 21 (1967) 1264–1266.

[28] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *A Historical Profile of the Higgs Boson*. arXiv:1201.6045 [hep-ph].

[29] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*. Phys. Rept. **457** (2008) 1–216, arXiv:hep-ph/0503172.

[30] Y. Nambu, *Quasi-Particles and Gauge Invariance in the Theory of Superconductivity*. Phys. Rev. **117** 3 (1960) 648–663.

[31] J. Goldstone, *Field theories with "Superconductor" solutions*. Il Nuovo Cimento (1955-1965) **19** (1961) no. 1, 154–164.

[32] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*. Phys. Rev. **127** 3 (1962) 965–970.

[33] The LHC Higgs Cross Section Working Group, D. de Florian *et al.*, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. arXiv:1610.07922 [hep-ph].

[34] The LHC Higgs Cross Section Working Group, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG#Latest_plots. Retrieved on February 11th, 2019.

[35] The 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** (2015) 191803, arXiv:1503.07589 [hep-ex].

[36] The ATLAS Collaboration, *Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*. arXiv:1806.00242 [hep-ex].

[37] The CMS Collaboration, *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV*. JHEP **11** (2017) 047, arXiv:1706.09936 [hep-ex].

[38] The ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*. Eur. Phys. J. **C75** (2015) no. 10, 476, arXiv:1506.05669 [hep-ex]. [Erratum: Eur. Phys. J.C76,no.3,152(2016)].

[39] The CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*. Phys. Rev. **D92** (2015) no. 1, 012004, arXiv:1411.3441 [hep-ex].

[40] The ATLAS Collaboration, *Determination of spin and parity of the Higgs boson in the $WW^* \to e\nu\mu\nu$ decay channel with the ATLAS detector*. Eur. Phys. J. **C75** (2015) no. 5, 231, arXiv:1503.03643 [hep-ex].

[41] The 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** (2016) 045, arXiv:1606.02266 [hep-ex].

[42] The 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*. arXiv:1802.04146 [hep-ex].

[43] The 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* Tech. Rep. ATLAS-CONF-2018-028, CERN, 2018. https://cds.cern.ch/record/2628771.

[44] The ATLAS Collaboration, *Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*. JHEP **03** (2018) 095, arXiv:1712.02304 [hep-ex].

[45] The ATLAS Collaboration, *Measurements of the Higgs boson production, fiducial and differential cross sections in the 4ℓ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector* Tech. Rep. ATLAS-CONF-2018-018, CERN, 2018. https://cds.cern.ch/record/2621479.

[46] The ATLAS Collaboration, *Combined measurement of differential and total cross sections in the $H \to \gamma\gamma$ and the $H \to ZZ^* \to 4\ell$ decay channels at $\sqrt{s} = 13$ TeV with the ATLAS detector*. arXiv:1805.10197 [hep-ex].

[47] The CMS Collaboration, *Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*. arXiv:1807.03825 [hep-ex].

[48] The ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*. arXiv:1806.00425 [hep-ex].

[49] The CMS Collaboration, *Observation of $\mathrm{t\bar{t}}H$ production*. Phys. Rev. Lett. **120** (2018) no. 23, 231801, arXiv:1804.02610 [hep-ex].

[50] The ATLAS Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*. Eur. Phys. J. **C76** (2016) no. 1, 6, arXiv:1507.04548 [hep-ex].

[51] The CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*. Eur. Phys. J. **C75** (2015) no. 5, 212, arXiv:1412.8662 [hep-ex].

[52] The ATLAS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*. arXiv:1808.08238 [hep-ex].

[53] The CMS Collaboration, *Observation of Higgs boson decay to bottom quarks*. Phys. Rev. Lett. **121** (2018) no. 12, 121801, arXiv:1808.08242 [hep-ex].

[54] The ATLAS Collaboration, *Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector*. Phys. Rev. **D92** (2015) no. 1, 012006, arXiv:1412.2641 [hep-ex].

[55] The CMS Collaboration, *Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV*. Phys. Lett. **B791** (2019) 96, arXiv:1806.05246 [hep-ex].

[56] The ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Phys. Rev. **D99** (2019) 072001, arXiv:1811.08856 [hep-ex].

[57] The CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector*. Phys. Lett. **B779** (2018) 283–316, arXiv:1708.00373 [hep-ex].

[58] The ATLAS Collaboration, *Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment*. Phys. Rev. Lett. **120** (2018) no. 21, 211802, arXiv:1802.04329 [hep-ex].

[59] The ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*. Phys. Lett. **B726** (2013) 88–119, arXiv:1307.1427 [hep-ex]. [Erratum: Phys. Lett.B734,406(2014)].

[60] The ATLAS Collaboration, *Searches for the $Z\gamma$ decay mode of the Higgs boson and for new high-mass resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. JHEP **10** (2017) 112, arXiv:1708.00212 [hep-ex].

[61] The CMS Collaboration, *Search for the decay of a Higgs boson in the $\ell\ell\gamma$ channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*. JHEP **11** (2018) 152, arXiv:1806.05996 [hep-ex].

[62] The ATLAS Collaboration, *A search for the rare decay of the Standard Model Higgs boson to dimuons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector* Tech. Rep. ATLAS-CONF-2018-026, CERN, 2018. https://cds.cern.ch/record/2628763.

[63] The CMS Collaboration, *Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Phys. Rev. Lett. **122** (2019) no. 2, 021801, arXiv:1807.06325 [hep-ex].

[64] The ATLAS Collaboration, *Combination of searches for invisible Higgs boson decays with the ATLAS experiment* Tech. Rep. ATLAS-CONF-2018-054, CERN, 2018. https://cds.cern.ch/record/2649407.

[65] The CMS Collaboration, *Searches for invisible decays of the Higgs boson in pp collisions at $\sqrt{s} = 7$, 8, and 13 TeV*. JHEP **02** (2017) 135, arXiv:1610.09218 [hep-ex].

[66] The ATLAS Collaboration, *Searches for exclusive Higgs and Z boson decays into $J/\psi\gamma$, $\psi(2S)\gamma$, and $\Upsilon(nS)\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector*. arXiv:1807.00802 [hep-ex].

[67] The CMS Collaboration, *Search for a Higgs boson decaying into $\gamma^*\gamma \to \ell\ell\gamma$ with low dilepton mass in pp collisions at $\sqrt{s} = 8$ TeV*. Phys. Lett. **B753** (2016) 341–362, arXiv:1507.03031 [hep-ex].

[68] The ATLAS Collaboration, *Search for exclusive Higgs and Z boson decays to $\phi\gamma$ and $\rho\gamma$ with the ATLAS detector*. JHEP **07** (2018) 127, arXiv:1712.02758 [hep-ex].

[69] GEANT4, S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*. Nucl. Instr. Meth. **A506** (2003) 250–303.

[70] The ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim* Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, 2010. https://cds.cern.ch/record/1300517.

[71] The ATLAS Collaboration, *The ATLAS Simulation Infrastructure*. Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[72] S. Gieseke, *Simulation of jets at colliders*. Progr. Part. Nucl. Phys. **72** (2013) 155 – 205.

[73] D. Rauch, *Automated NLO Calculations with Massive Quarks in Herwig++*, www.itp.kit.edu/prep/diploma/PSFiles/ms-thesis-daniel-rauch.pdf, 2014.

[74] A. Buckley *et al.*, *General-purpose event generators for LHC physics*. Phys. Rept. **504** (2011) 145–233, arXiv:1101.2599 [hep-ph].

[75] S. Höche, *Introduction to parton-shower event generators*. arXiv:1411.4085 [hep-ph].

[76] S. Weinzierl, *Introduction to Monte Carlo methods*. arXiv:hep-ph/0006269 [hep-ph].

[77] T. Kinoshita, *Mass Singularities of Feynman Amplitudes*. J. Math. Phys. **3** (1962) no. 4, 650–677.

[78] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*. Phys. Rev. **133** 6B (1964) B1549–B1562.

[79] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996.

[80] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*. JHEP **06** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[81] S. Frixione and B. R. Webber, *The MC@NLO 3.3 Event Generator*. arXiv:hep-ph/0612272 [hep-ph].

[82] 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** (2014) 079, arXiv:1405.0301 [hep-ph].

[83] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*. JHEP **11** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

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

[85] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*. JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[86] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*. Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[87] J. Bellm *et al.*, *Herwig 7.0/Herwig++ 3.0 release note*. Eur. Phys. J. **C76** (2016) no. 4, 196, arXiv:1512.01178 [hep-ph].

[88] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*. JHEP **12** (2012) 061, arXiv:1209.6215 [hep-ph].

[89] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*. Physics Reports **97** (1983) no. 2, 31 – 145.

[90] B. Andersson, *The Lund model*. Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **7** (1997) 1–471.

[91] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*. JHEP **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[92] S. Wolfram, *Parton and hadron production in e^+e^- annihilation*. Largely based on a talk given at the 15th Rencontres de Moriond, Les Arcs, France, Mar 9-21, 1980.

[93] G. C. Fox and S. Wolfram, *A model for parton showers in QCD*. Nucl. Phys. B **168** (1980) no. 2, 285 – 295.

[94] R. D. Field and S. Wolfram, *A QCD model for e^+e^- annihilation*. Nucl. Phys. B **213** (1983) no. 1, 65 – 84.

[95] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*. JHEP **01** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[96] M. Bähr *et al.*, *Herwig++ Physics and Manual*. Eur. Phys. J. **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[97] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*. JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[98] D. Amati and G. Veneziano, *Preconfinement as a property of perturbative QCD*. Physics Letters B **83** (1979) no. 1, 87 – 92.

[99] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report Vol.1: The LHC Main Ring*, 2004. CERN-2004-003-V1.

[100] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report Vol.2: The LHC Infrastructure and General Services*, 2004. CERN-2004-003-V2.

[101] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report Vol.3: The LHC Injector Chain*, 2004. CERN-2004-003-V3.

[102] *CERN FAQ - LHC the guide*. https://cdsweb.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf. Retrieved on February 27th, 2019.

[103] The ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*. Eur. Phys. J. **C76** (2016) no. 12, 653, arXiv:1608.03953 [hep-ex].

[104] The ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*. Eur. Phys. J. **C73** (2013) no. 8, 2518, arXiv:1302.4393 [hep-ex].

[105] *ATLAS Experiment Public Results - Luminosity Public Results Run 2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Retrieved on February 27th, 2019.

[106] *ATLAS Experiment Public Results - Data Quality Information for Data*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010. Retrieved on February 27th, 2019.

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

[108] Z. Marshall for the ATLAS Collaboration, *Simulation of Pile-up in the ATLAS Experiment*. J. Phys.: Conf. Ser. **513** (2014) no. 2, 022024.

[109] *Track Reconstruction Performance of the ATLAS Inner Detector at* $\sqrt{s} = 13$ *TeV* Tech. Rep. ATL-PHYS-PUB-2015-018, CERN, 2015. `https://cds.cern.ch/record/2037683`.

[110] The ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, 1998. `https://cds.cern.ch/record/381263`.

[111] The ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report* Tech. Rep. CERN-LHCC-2010-013, ATLAS-TDR-19, 2010. `https://cds.cern.ch/record/1291633`.

[112] The ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report Addendum* Tech. Rep. CERN-LHCC-2012-009, ATLAS-TDR-19-ADD-1, 2012. `https://cds.cern.ch/record/1451888`. Addendum to CERN-LHCC-2010-013, ATLAS-TDR-019.

[113] A. Ahmad *et al.*, *The silicon microstrip sensors of the ATLAS semiconductor tracker*. Nucl. Instr. Meth. **A578** (2007) no. 1, 98 – 118.

[114] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance*. JINST **3** (2008) no. 02, P02013–P02013.

[115] B. Aubert *et al.*, *Performance of the ATLAS electromagnetic calorimeter barrel module 0*. Nucl. Instr. Meth. **A500** (2003) no. 1, 202 – 231.

[116] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*. Eur. Phys. J. **C74** (2014) no. 10, 3071, `arXiv:1407.5063 [hep-ex]`.

[117] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*. JINST **14** (2019) no. 03, P03017, `arXiv:1812.03848 [hep-ex]`.

[118] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*. Phys. Rev. D **98** 3 (2018) 030001.

[119] E. Diehl for the ATLAS Collaboration, *Calibration and Performance of the ATLAS Muon Spectrometer*. `arXiv:1109.6933 [physics.ins-det]`.

[120] S. Palestini, *The muon spectrometer of the ATLAS experiment*. Nucl. Phys. B (Proc. Suppl.) **125** (2003) 337 – 345.

[121] The ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*. Eur. Phys. J. **C76** (2016) no. 5, 292, `arXiv:1603.05598 [hep-ex]`.

[122] F. Bauer *et al.*, *Construction and Test of MDT Chambers for the ATLAS Muon Spectrometer*. Nucl. Instr. Meth. **A461** (2001) 17–20, `arXiv:1604.02000 [physics.ins-det]`.

[123] T. Argyropoulos *et al.*, *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*. IEEE Trans. Nucl. Sci. **56** (2009) 1568–1574.

[124] G. Aielli *et al.*, *The RPC first level muon trigger in the barrel of the ATLAS experiment.* Nucl. Phys. B (Proc. Suppl.) **158** (2006) 11–15. [,11(2006)].

[125] S. Artz *et al.*, *Upgrade of the ATLAS Central Trigger for LHC Run-2.* JINST **10** (2015) no. 02, C02030–C02030.

[126] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015.* Eur. Phys. J. **C77** (2017) no. 5, 317, arXiv:1611.09661 [hep-ex].

[127] The ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1.* Eur. Phys. J. **C77** (2017) 490, arXiv:1603.02934 [hep-ex].

[128] The ATLAS Collaboration, *Athena.* https://doi.org/10.5281/zenodo.2641997.

[129] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)* Tech. Rep. ATL-SOFT-PUB-2007-007, CERN, 2007. https://cds.cern.ch/record/1020106.

[130] The ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments* Tech. Rep. ATL-PHYS-PUB-2015-006, CERN, 2015. https://cds.cern.ch/record/2002609.

[131] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting.* Nucl. Instr. Meth. **A262** (1987) no. 2, 444 – 450.

[132] T. G. Cornelissen, M. Elsing, I. Gavrilenko, J.-F. Laporte, W. Liebig, M. Limper, K. Nikolopoulos, A. Poppleton, and A. Salzburger, *The global χ^2 track fitter in ATLAS.* J. Phys.: Conf. Ser. **119** (2008) no. 3, 032013.

[133] The ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC.* Eur. Phys. J. **C77** (2017) no. 5, 332.

[134] 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* Tech. Rep. ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, 2008. https://cds.cern.ch/record/1099735.

[135] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm.* JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[136] The ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.* Phys. Rev. **D96** (2017) no. 7, 072002, arXiv:1703.09665 [hep-ex].

[137] The 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. **C79** (2019) no. 3, 205, arXiv:1810.05087 [hep-ex].

[138] The ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung* Tech. Rep. ATLAS-CONF-2012-047, CERN, 2012. https://cds.cern.ch/record/1449796.

[139] The ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data.* Eur. Phys. J. **C76** (2016) no. 12, 666, arXiv:1606.01813 [hep-ex].

[140] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at* $\sqrt{s} = 13$ TeV Tech. Rep. ATL-PHYS-PUB-2016-015, CERN, 2016. https://cds.cern.ch/record/2203514.

[141] The 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*. Eur. Phys. J. **C79** (2019) no. 8, 639, arXiv:1902.04655 [physics.ins-det].

[142] ATLAS e/γ Combined Performance Group, *Calorimeter shower shapes of photons from* $Z \to ll\gamma$ *decays in 20.3* fb^{-1} *of pp collisions collected by ATLAS at* $\sqrt{s} = 8$ *TeV in 2012* Tech. Rep. ATL-COM-PHYS-2013-600, CERN, 2013. https://cds.cern.ch/record/1545866.

[143] A. Hoecker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*. arXiv:physics/0703039 [physics.data-an]. http://tmva.sourceforge.net/.

[144] K. Brendlinger and I. J. Kroll, *Physics with Electrons in the ATLAS Detector*. PhD thesis, University of Pennsylvania, 2018. https://cds.cern.ch/record/2228644. Presented 21 Jun 2016.

[145] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*. Philos. Trans. Royal Soc. A **231** (1933) no. 694-706, 289–337.

[146] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*. Annals Math. Statist. **9** (1938) no. 1, 60–62.

[147] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*. Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an]. [Erratum: Eur. Phys. J.C73,2501(2013)].

[148] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*. Trans. Am. Math. Soc. **54** (1943) no. 3, 426–482.

[149] L. Smestad, *The Search for the Standard Model Higgs Boson in* $H \to \gamma\gamma$ *Decays with the ATLAS Detector in 4.9* fb^{-1} *of 2011 Data at* $\sqrt{s} = 7$ *TeV*. PhD thesis, University of Oslo, 2013. CERN-THESIS-2013-039.

[150] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*. Phys. Rev. **D57** (1998) 3873–3889, arXiv:physics/9711021 [physics.data-an].

[151] A. L. Read, *Presentation of search results: The CL(s) technique*. J. Phys. **G28** (2002) 2693–2704.

[152] The ATLAS and CMS Collaborations, The LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011* Tech. Rep. ATL-PHYS-PUB-2011-11, CMS-NOTE-2011-005, CERN, 2011. https://cds.cern.ch/record/1379837.

[153] The ATLAS Collaboration, *Combined search for the Standard Model Higgs boson in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*. Phys. Rev. **D86** (2012) 032003, arXiv:1207.0319 [hep-ex].

[154] Z. Šidák, *Rectangular Confidence Regions for the Means of Multivariate Normal Distributions*. J. Am. Stat. Assoc. **62** (1967) no. 318, 626–633.

[155] *Šidák correction*, https://en.wikipedia.org/wiki/%C5%A0id%C3%A1k_correction. Retrieved on April 27th, 2017.

[156] E. Gross and O. Vitells, *Trial factors or the look elsewhere effect in high energy physics*. Eur. Phys. J. **C70** (2010) 525–530, arXiv:1005.1891 [physics.data-an].

[157] R. B. Davies, *Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternatives*. Biometrika **74** (1987) no. 1, 33–43.

[158] J. Alwall *et al.*, *A Standard format for Les Houches event files*. Comput. Phys. Commun. **176** (2007) 300–304, arXiv:hep-ph/0609017 [hep-ph].

[159] K. Arnold *et al.*, *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*. Comput. Phys. Commun. **180** (2009) 1661–1670, arXiv:0811.4559 [hep-ph].

[160] J. Baglio *et al.*, *Release Note - VBFNLO 2.7.0*. arXiv:1404.3940 [hep-ph].

[161] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5: Going Beyond*. JHEP **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[162] G. Cullen *et al.*, *GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond*. Eur. Phys. J. **C74** (2014) no. 8, 3001, arXiv:1404.7096 [hep-ph].

[163] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*. Phys. Rev. Lett. **108** (2012) 111601, arXiv:1111.5206 [hep-ph].

[164] S. Plätzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*. JHEP **01** (2011) 024, arXiv:0909.5593 [hep-ph].

[165] J. Bellm *et al.*, *Herwig 7.1 Release Note*. arXiv:1705.06919 [hep-ph].

[166] The ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*. arXiv:1907.05120 [hep-ex].

[167] F. Maltoni, G. Ridolfi, and M. Ubiali, *b-initiated processes at the LHC: a reappraisal*. JHEP **07** (2012) 022, arXiv:1203.6393 [hep-ph]. [Erratum: JHEP04,095(2013)].

[168] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, R. Pittau, and M. Worek, *Assault on the NLO Wishlist: $pp \to t\bar{t}b\bar{b}$*. JHEP **09** (2009) 109, arXiv:0907.4723 [hep-ph].

[169] G. Bevilacqua and M. Worek, *On the ratio of $t\bar{t}b\bar{b}$ and $t\bar{t}jj$ cross sections at the CERN Large Hadron Collider*. JHEP **07** (2014) 135, arXiv:1403.2046 [hep-ph].

[170] F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini, and F. Siegert, *NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks*. Phys. Lett. **B734** (2014) 210–214, arXiv:1309.5912 [hep-ph].

[171] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD corrections to $t\bar{t}b\bar{b}$ production at the LHC: 1. Quark-antiquark annihilation*. JHEP **08** (2008) 108, arXiv:0807.1248 [hep-ph].

[172] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD corrections to $pp \to t\bar{t}b\bar{b} + X$ at the LHC*. Phys. Rev. Lett. **103** (2009) 012002, arXiv:0905.0110 [hep-ph].

[173] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD Corrections to $t\bar{t}b\bar{b}$ Production at the LHC: 2. full hadronic results*. JHEP **03** (2010) 021, arXiv:1001.4006 [hep-ph].

[174] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *Production of $t\bar{t}b\bar{b}$ at the LHC at NLO QCD*. Nucl. Phys. B (Proc. Suppl.) **205-206** (2010) 80–85, arXiv:1006.2653 [hep-ph].

[175] The ATLAS Collaboration, *Studies of tt+cc production with MadGraph5_aMC@NLO and Herwig++ for the ATLAS experiment* Tech. Rep. ATL-PHYS-PUB-2016-011, CERN, 2016. https://cds.cern.ch/record/2153876.

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

[177] D. J. Lange, *The EvtGen particle decay simulation package*. Nucl. Instr. Meth. **A462** (2001) no. 1–2, 152–155.

[178] H.-L. Lai *et al.*, *New parton distributions for collider physics*. Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[179] A. Buckley *et al.*, *LHAPDF6: parton density access in the LHC precision era*. Eur. Phys. J. **C75** (2015) 132, arXiv:1412.7420 [hep-ph].

[180] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*. Eur. Phys. J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[181] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *Production of $t\bar{t}b\bar{b}$ at the LHC at NLO QCD*. Nucl. Phys. B (Proc. Suppl.) **205-206** (2010) 80–85, arXiv:1006.2653 [hep-ph].

[182] M. Wiesemann *et al.*, *Higgs production in association with bottom quarks*. JHEP **02** (2015) 132, arXiv:1409.5301 [hep-ph].

[183] The ATLAS Collaboration, *Simulation of top quark production for the ATLAS experiment at $\sqrt{s}$ = 13 TeV* Tech. Rep. ATL-PHYS-PUB-2016-004, CERN, 2016. https://cds.cern.ch/record/2120417.

[184] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*. JHEP **04** (2013) 027, arXiv:1207.5030 [hep-ph].

[185] D. Goncalves, F. Krauss, and R. Linten, *Distinguishing b-quark and gluon jets with a tagged b-hadron*. Phys. Rev. **D93** (2016) no. 5, 053013, arXiv:1512.05265 [hep-ph].

[186] The ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. Phys. Rev. **D97** (2018) no. 7, 072016, arXiv:1712.08895 [hep-ex].

[187] I. Low, J. Lykken, and G. Shaughnessy, *Singlet scalars as Higgs imposters at the Large Hadron Collider*. Phys. Rev. **D84** (2011) 035027, arXiv:1105.4587 [hep-ph].

[188] I. Low, J. Lykken, and G. Shaughnessy, *Have We Observed the Higgs (Imposter)?* Phys. Rev. **D86** (2012) 093012, arXiv:1207.1093 [hep-ph].

[189] A. Azatov, R. Contino, A. Di Iura, and J. Galloway, *New Prospects for Higgs Compositeness in $h \to Z\gamma$*. Phys. Rev. **D88** (2013) no. 7, 075019, arXiv:1308.2676 [hep-ph].

[190] C.-W. Chiang and K. Yagyu, *Higgs boson decays to $\gamma\gamma$ and $Z\gamma$ in models with Higgs extensions*. Phys. Rev. **D87** (2013) no. 3, 033003, arXiv:1207.1065 [hep-ph].

[191] C.-S. Chen, C.-Q. Geng, D. Huang, and L.-H. Tsai, *New Scalar Contributions to $h \to Z\gamma$*. Phys. Rev. **D87** (2013) 075019, arXiv:1301.4694 [hep-ph].

[192] M. Carena, I. Low, and C. E. M. Wagner, *Implications of a Modified Higgs to Diphoton Decay Width*. JHEP **08** (2012) 060, arXiv:1206.1082 [hep-ph].

[193] The ATLAS Collaboration, *Search for Higgs boson decays to a photon and a Z boson in pp collisions at $\sqrt{s}$=7 and 8 TeV with the ATLAS detector*. Phys. Lett. **B732** (2014) 8–27, arXiv:1402.3051 [hep-ex].

[194] The CMS Collaboration, *Search for a Higgs boson decaying into a Z and a photon in pp collisions at $\sqrt{s}$ = 7 and 8 TeV*. Phys. Lett. **B726** (2013) 587–609, arXiv:1307.5515 [hep-ex].

[195] The ATLAS Collaboration, *Search for new resonances in $W\gamma$ and $Z\gamma$ final states in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*. Phys. Lett. **B738** (2014) 428–447, arXiv:1407.8150 [hep-ex].

[196] The ATLAS Collaboration, *Search for heavy resonances decaying to a Z boson and a photon in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Phys. Lett. **B764** (2017) 11–30, arXiv:1607.06363 [hep-ex].

[197] The ATLAS Collaboration, *Search for new resonances decaying to a Z boson and a photon in 13.3 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector* Tech. Rep. ATLAS-CONF-2016-044, CERN, 2016. https://cds.cern.ch/record/2206125.

[198] The ATLAS Collaboration, *Search for heavy resonances decaying to a photon and a hadronically decaying $Z/W/H$ boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Phys. Rev. **D98** (2018) no. 3, 032015, arXiv:1805.01908 [hep-ex].

[199] The ATLAS Collaboration, *Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC* Tech. Rep. ATL-PHYS-PUB-2018-054, CERN, 2018. http://cds.cern.ch/record/2652762.

[200] The ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector* Tech. Rep. ATLAS-CONF-2014-018, CERN, 2014. https://cds.cern.ch/record/1700870.

[201] The 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. **C76** (2016) no. 11, 581, arXiv:1510.03823 [hep-ex].

[202] The ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*. Phys. Rev. **D90** (2014) no. 5, 052004, arXiv:1406.3827 [hep-ex].

[203] The ATLAS Collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*. Phys. Rev. **D91** (2015) no. 1, 012006, arXiv:1408.5191 [hep-ex].

[204] The OPAL Collaboration, *Search for anomalous production of dilepton events with missing transverse momentum in e+ e- collisions at $\sqrt{s} = 161$ GeV and 172 GeV*. Eur. Phys. J. **C4** (1998) 47–74, arXiv:hep-ex/9710010 [hep-ex].

[205] M. Vesterinen and T. R. Wyatt, *A Novel Technique for Studying the Z Boson Transverse Momentum Distribution at Hadron Colliders*. Nucl. Instr. Meth. **A602** (2009) 432–437, arXiv:0807.4956 [hep-ex].

[206] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*. Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[207] The ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes* Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, 2012. https://cds.cern.ch/record/1474107.

[208] J. Butterworth *et al.*, *PDF4LHC recommendations for LHC Run II*. J. Phys. **G43** (2016) 023001, arXiv:1510.03865 [hep-ph].

[209] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, *NNLOPS simulation of Higgs boson production*. JHEP **10** (2013) 222, arXiv:1309.0017 [hep-ph].

[210] K. Hamilton, P. Nason, C. Oleari, and G. Zanderighi, *Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*. JHEP **05** (2013) 082, arXiv:1212.4504 [hep-ph].

[211] K. Hamilton, P. Nason, and G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*. JHEP **05** (2015) 140, arXiv:1501.04637 [hep-ph].

[212] The 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** (2014) 145, arXiv:1406.3660 [hep-ex].

[213] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*. JHEP **02** (2012) 088, arXiv:1111.2854 [hep-ph].

[214] P. Artoisenet *et al.*, *A framework for Higgs characterisation*. JHEP **11** (2013) 043, arXiv:1306.6464 [hep-ph].

[215] P. de Aquino and K. Mawatari, *Characterising a Higgs-like resonance at the LHC*. 2013. arXiv:1307.5607 [hep-ph].

[216] The ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes* Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, 2014. https://cds.cern.ch/record/1966419.

[217] R. D. Ball *et al.*, *Parton distributions with LHC data*. Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[218] S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*. JHEP **05** (2009) 053, arXiv:0903.1219 [hep-ph].

[219] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*. JHEP **04** (2013) 027, arXiv:1207.5030 [hep-ph].

[220] M. Oreglia, *A Study of the Reactions $\psi' \to \gamma\gamma\psi$*, 1980. https://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf.

[221] The ATLAS Collaboration, *Search for resonances in diphoton events at $\sqrt{s}$=13 TeV with the ATLAS detector*. JHEP **09** (2016) 001, arXiv:1606.03833 [hep-ex].

[222] The 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. **D83** (2011) 052005, arXiv:1012.4389 [hep-ex].

[223] The ATLAS Higgs Working Group, *Theoretical Predictions and systematics*. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HiggsPropertiesTheorySystematics.

[224] The LHC Higgs Cross Section Working Group, *CERN Report 4*, 2016. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV.

[225] The LHC Higgs Cross Section Working Group, *CERN Report 4*, 2016. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR#Higgs_2_gauge_bosons.

[226] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*. Phys. Rev. **D85** (2012) 034011, arXiv:1107.2117 [hep-ph].

[227] S. Gangal and F. J. Tackmann, *Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion*. Phys. Rev. **D87** (2013) no. 9, 093008, arXiv:1302.5437 [hep-ph].

[228] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, *Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to \gamma\gamma$, $H \to WW \to l\nu l\nu$ and $H \to ZZ \to 4l$ decay modes*. JHEP **06** (2012) 132, arXiv:1203.6321 [hep-ph].

[229] M. Grazzini and H. Sargsyan, *Heavy-quark mass effects in Higgs boson production at the LHC*. JHEP **09** (2013) 129, arXiv:1306.4581 [hep-ph].

[230] F. Bernlochner, S. Gangal, D. Gillberg, and F. Tackmann.

[231] The LHC Higgs Cross Section Working Group, J. R. Andersen *et al.*, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*. arXiv:1307.1347 [hep-ph].

[232] S. Gangal, *Precision Predictions of Exclusive Jet Cross Sections at the LHC*. PhD thesis, Universität Hamburg, 2015. https://bib-pubdb1.desy.de/record/275847.

[233] The NNPDF Collaboration, R. D. Ball *et al.*, *Parton distributions for the LHC Run II*. JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[234] The CMS Collaboration, *Search for the standard model Higgs boson in the dilepton plus photon channel in* pp *collisions at $\sqrt{s} = 13$* TeV Tech. Rep. CMS-PAS-HIG-17-007, CERN, 2018. https://cds.cern.ch/record/2308435.

[235] Y. Sun, H.-R. Chang, and D.-N. Gao, *Higgs decays to $\gamma l^+ l^-$ in the standard model*. JHEP **05** (2013) 061, arXiv:1303.2230 [hep-ph].

[236] Y. Chen, A. Falkowski, I. Low, and R. Vega-Morales, *New Observables for CP Violation in Higgs Decays*. Phys. Rev. **D90** (2014) no. 11, 113006, arXiv:1405.6723 [hep-ph].

[237] A. Abbasabadi, D. Bowser-Chao, D. A. Dicus, and W. W. Repko, *Radiative Higgs boson decays $H \to$ fermion anti-fermion gamma*. Phys. Rev. **D55** (1997) 5647–5656, arXiv:hep-ph/9611209 [hep-ph].

[238] The ATLAS Collaboration, *Search for Higgs and Z Boson Decays to $J/\Psi\gamma$ and $\Upsilon(nS)\gamma$ with the ATLAS Detector*. Phys. Rev. Lett. **114** (2015) no. 12, 121801, arXiv:1501.03276 [hep-ex].

[239] A. Firan and R. Stroynowski, *Internal conversions in Higgs decays to two photons*. Phys. Rev. **D76** (2007) 057301, arXiv:0704.3987 [hep-ph].

[240] L.-B. Chen, C.-F. Qiao, and R.-L. Zhu, *Reconstructing the 125 GeV SM Higgs Boson Through $\ell\bar{\ell}\gamma$*. Phys. Lett. **B726** (2013) 306–311, arXiv:1211.6058 [hep-ph].

[241] D. A. Dicus and W. W. Repko, *Calculation of the decay $H \to e\bar{e}\gamma$*. Phys. Rev. **D87** (2013) no. 7, 077301, arXiv:1302.2159 [hep-ph].

[242] The ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data*. arXiv:1908.00005 [hep-ex].

[243] *Longer Term LHC Schedule*. https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm. Retrieved on November 13th, 2019.

Appendix

A. Statistical Methods

A.1. Probability Distributions

Some of the most commonly encountered probability distribution functions are:

- **Gauss distribution**
  - 1-dimensional

$$\mathcal{G}(x;\mu,V=\sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \tag{17.1}$$

  - m-dimensional

$$\mathcal{G}(\boldsymbol{x};\boldsymbol{\mu},V) := \frac{1}{\sqrt{(2\pi)^m \det(V)}} \cdot \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T V^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \tag{17.2}$$

- **Standard normal distribution**

$$\mathcal{N}(x) := \mathcal{G}(x;0,1) \tag{17.3}$$

- **Poisson distribution**

$$\mathcal{P}(k;\lambda) := \frac{\lambda^k e^{-\lambda}}{k!} \tag{17.4}$$

  and

$$\mathcal{P}(k;\lambda) \to \mathcal{G}(k;\lambda,\lambda) \text{ for } \lambda \to \infty \tag{17.5}$$

- **Binomial distribution**

$$\mathcal{B}(k;n,p) := \binom{n}{k} p^k (1-p)^{n-k} \tag{17.6}$$

  and

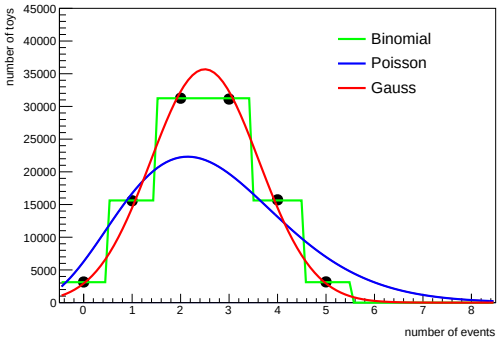
$$\mathcal{B}(k;n,p) \to \begin{cases} \mathcal{G}(k;np,np(1-p)) & \text{for } n \to \infty \\ \mathcal{P}(k;np) & \text{for } n \to \infty \text{ while } np = \text{const} \end{cases} \tag{17.7}$$

- **Chi-squared distribution**

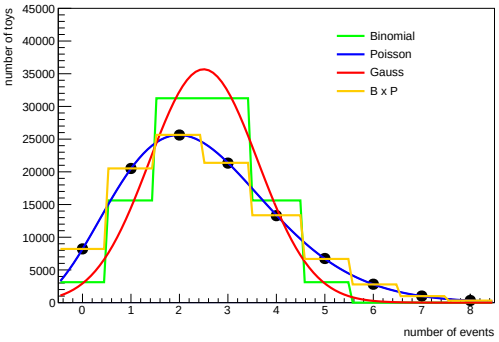
$$\chi^2(x;k) := \frac{x^{\frac{k}{2}-1} e^{-\frac{x}{2}}}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} \tag{17.8}$$

  and

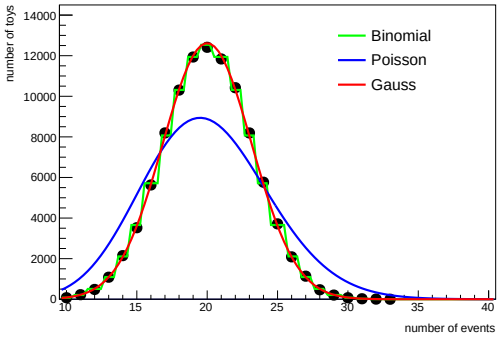
$$\chi^2(x;k) \to \mathcal{G}(x;k,2k) \text{ for } k \to \infty \tag{17.9}$$



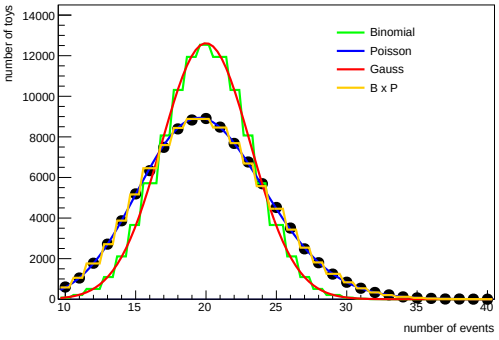
(a) Fixed total number of 5 events.



(b) Poissonian fluctuating total number of events with a mean of 5.



(c) Fixed total number of 40 events.



(d) Poissonian fluctuating total number of events with a mean of 40.

Figure A.1.: Distribution of the number of events in a specific histogram bin due to statistical fluctuations. For each plot the number of events in a bin was obtained from 10^5 toy experiments.

A.2. The Shape of Statistical Uncertainties

A.2.1. Fixed Number of Events in a Histogram

From the perspective of a bin, booking events into a histogram is a binomial process. There are only two possible outcomes for each and every event: An event either enters a specific bin or it does not. At the same time the probability for doing so, i.e. the probability p for success, remains constant for all events and can be calculated from the true distribution of the events. Therefore, the probability of finding k out of a total of N events in a specific bin is given by the binomial distribution $\mathcal{B}(k; N, p)$.

This can be seen in Figure A.1. For m toy experiments, a histogram with two bins is filled with N events drawn from a uniform distribution, i.e. $p = 0.5$. For each toy experiment, the number of events in the first of the two bins is counted and stored in a second histogram. This second histogram is shown in Figures A.1(a) and (c) for $N = 5$ and $N = 40$ events per toy experiment, respectively. Also plotted are the binomial distribution $m \cdot \mathcal{B}(k; N, p)$, the Poisson distribution $m \cdot \mathcal{P}(k; Np)$ and the Gaussian distribution $m \cdot \mathcal{G}(k; Np, Np(1-p))$.

As can be seen, the binomial distribution indeed fits best and the Gaussian distribution becomes a reasonable approximation for a sufficiently large number of events N per toy.

A.2.2. Fluctuating Number of Events in a Histogram

The typical situation in high energy physics measurements is different, however. Instead of measuring until a given number of events is reached, the analyses use datasets collected during certain periods of time. This has the consequence that the number of events n recorded during those periods is a random variable itself which is distributed according to a Poisson distribution with mean N. It is this N that is determined by the integrated luminosity, the relevant cross sections and acceptances as well as adequate efficiencies.

Therefore, the probability $P(k; N, p)$ for finding k out of N events in a bin while each event has a probability p to enter the given bin is described by

$$P(k; N, p) = \sum_{n=0}^{\infty} \mathcal{B}(k; n, p) \cdot \mathcal{P}(n; N) \ . \tag{17.10}$$

This distribution is shown as an orange line in Figures A.1(b) and (d) for $N = 5$ and $N = 40$, respectively, and is equivalent to the Poisson distribution $m \cdot \mathcal{P}(k; Np)$. Both the binomial and the Gauss distributions fail to describe the observed distribution of the number of events in this case.

This is the reason why statistical uncertainties corresponding to the Poisson distribution are an adequate choice in typical high energy physics experiments and measurements.

A.3. From p-Values to Standard Deviations and Significances

It is customary to express the significance of an excess or a deviation not only in terms of the p-value but also to provide the number of standard deviations Z of a Gauss distribution that this p-value corresponds to.

Based on the Gauss or normal distribution $\mathcal{G}(x; \mu, \sigma^2)$, the standard normal distribution is described by the probability density function $\mathcal{N}(x) = \mathcal{G}(x; 0, 1)$. The corresponding cumulative distribution function is thus given by

$$\Phi(x) = \int_{-\infty}^{x} \mathcal{N}(t) \mathrm{d}t \ . \tag{17.11}$$

A related function convenient for the evaluation of $\Phi(x)$ is the error function

$$\mathrm{erf}(x) = \frac{1}{\pi}\int_{-x}^{x} e^{-t^2}\mathrm{d}t \ . \tag{17.12}$$

Using the normalisation of the Gauss distribution, one can derive the relation between $\Phi(x)$ and $\mathrm{erf}(x)$:

$$\Phi(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t = \frac{1}{2}\left(1 + \frac{1}{2\pi}\int_{-x}^{x} e^{-t^2/2}\mathrm{d}t\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right] \ . \tag{17.13}$$

The relation between the p-value and the corresponding significance Z depends on whether one- or two-sided tails are considered:

- For the one-sided case the significance Z is given by

$$1 - p = \Phi(Z) \qquad \text{and} \qquad p = 1 - \Phi(Z) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{Z}{\sqrt{2}}\right)\right] \ . \tag{17.14}$$

- For the two-sided case the significance Z is given by

$$1 - \frac{p}{2} = \Phi(Z) \qquad \text{and} \qquad p = 2\left(1 - \Phi(Z)\right) = 1 - \mathrm{erf}\left(\frac{Z}{\sqrt{2}}\right) \ . \tag{17.15}$$

B. Miscellanea on $H/X \to Z\gamma$

B.1. Event Selection and Kinematic Distributions

Event Selection Cutflow (Low-Mass Region)	Monte Carlo Signal Samples ggF / MPI On absolute	relative	ggF / MPI Off absolute	relative
all events	100.00%		100.00%	
GRL	100.00%	100.0%	100.00%	100.0%
primary vertex	100.00%	100.0%	100.00%	100.0%
event quality	100.00%	100.0%	100.00%	100.0%
triggers	58.27%	58.3%	58.14%	58.1%
after initial cuts	58.27%	100.0%	58.14%	100.0%
≥ 2 good leptons	56.19%	96.4%	56.33%	96.9%
OS & $m_{ll} > 45$ GeV	55.19%	98.2%	55.40%	98.4%
1 *Loose* γ passing preselection	38.41%	69.6%	39.07%	70.5%
after precuts	38.41%	100.0%	39.07%	100.0%
≥ 1 Higgs candidate	35.35%	92.0%	35.96%	92.0%
trigger matching	34.97%	98.9%	35.60%	99.0%
76.18 GeV $\leq m_{ll} \leq$ 106.18 GeV	33.28%	95.2%	33.88%	95.1%
γ ID (*Tight*)	27.51%	82.6%	28.15%	83.1%
γ isol (*FixedCutLoose*)	23.00%	83.6%	24.21%	86.0%
$p_{\mathrm{T}}^{\gamma} \geq 15$ GeV	22.02%	95.7%	23.22%	95.9%
115 GeV $\leq m_{ll\gamma} \leq$ 170 GeV	21.65%	98.3%	22.80%	98.2%

Table B.1.: Absolute and relative event selection efficiencies for the low-mass search. Monte Carlo numbers are given in the NLO ggF channel with $m_H = 125$ GeV for the MPI On/Off samples with applying the p_{T}^{H}-reweighting (note: the signal efficiencies here are expressed with respect to the $Z \to ee/\mu\mu$ channels).

Event Selection Cutflow (High-Mass Region)	Monte Carlo Signal Samples for $m_X = 300$ GeV							
	spin-0/ggF		spin-0/VBF		spin-2/gg		spin-2/$q\bar{q}$	
	absolute	relative	absolute	relative	absolute	relative	absolute	relative
all events	100.00%		100.00%		100.00%		100.00%	
GRL	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
primary vertex	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
event quality	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
triggers	68.15%	68.2%	68.13%	68.1%	65.32%	65.3%	64.04%	64.0%
after initial cuts	68.15%	100.0%	68.13%	100.0%	65.32%	100.0%	64.04%	100.0%
≥ 2 good leptons	64.65%	94.9%	66.04%	96.9%	61.59%	94.3%	55.82%	87.2%
OS & $m_{ll} > 45$ GeV	63.09%	97.6%	64.68%	97.9%	60.50%	98.2%	54.49%	97.6%
1 *Loose* γ passing preselection	49.72%	78.8%	52.34%	80.9%	44.64%	73.8%	41.25%	75.7%
after precuts	49.72%	100.0%	52.34%	100.0%	44.64%	100.0%	41.25%	100.0%
≥ 1 Higgs candidate	46.15%	92.8%	48.76%	93.2%	41.50%	93.0%	38.41%	93.1%
trigger matching	45.90%	99.5%	48.55%	99.6%	41.30%	99.5%	38.27%	99.6%
76.18 GeV $\leq m_{ll} \leq$ 106.18 GeV	42.70%	93.0%	45.17%	93.0%	39.10%	94.7%	36.22%	94.6%
γ ID (*Tight*)	40.15%	94.0%	42.44%	94.0%	36.69%	93.8%	34.21%	94.4%
γ isol (*FixedCutLoose*)	39.24%	97.8%	41.47%	97.7%	35.32%	96.2%	33.21%	97.1%
$p_{\mathrm{T}}^{\gamma} > 50$ GeV and $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} \geq 0.3$	32.71%	83.4%	33.13%	79.9%	24.20%	68.5%	29.94%	90.2%
$m_{ll\gamma} \geq 200$ GeV	32.71%	100.0%	33.13%	100.0%	24.20%	100.0%	29.94%	100.0%

Table B.2.: Absolute and relative event selection efficiencies for the high-mass search for a resonance mass of $m_X = 300$ GeV, derived from signal Monte Carlo event samples in the ggF and VBF production modes for a spin-0 resonance and the gluon- and quark-induced production modes for a spin-2 resonance. The efficiencies are expressed with respect to the $Z \to ee/\mu\mu$ channels.

Event Selection Cutflow (High-Mass Region)	Monte Carlo Signal Samples for $m_X = 1000$ GeV							
	spin-0/ggF		spin-0/VBF		spin-2/gg		spin-2/$q\bar{q}$	
	absolute	relative	absolute	relative	absolute	relative	absolute	relative
all events	100.00%		100.00%		100.00%		100.00%	
GRL	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
primary vertex	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
event quality	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
triggers	76.46%	76.5%	76.26%	76.3%	74.22%	74.2%	77.51%	77.5%
after initial cuts	76.46%	100.0%	76.26%	100.0%	74.22%	100.0%	77.51%	100.0%
≥ 2 good leptons	73.71%	96.4%	74.72%	98.0%	69.83%	94.1%	75.77%	97.8%
OS & $m_{ll} > 45$ GeV	71.33%	96.8%	72.64%	97.2%	67.48%	96.6%	73.18%	96.6%
1 *Loose* γ passing preselection	60.71%	85.1%	62.76%	86.4%	54.23%	80.4%	62.60%	85.6%
after precuts	60.71%	100.0%	62.76%	100.0%	54.23%	100.0%	62.60%	100.0%
≥ 1 Higgs candidate	57.30%	94.4%	59.08%	94.1%	51.20%	94.4%	59.17%	94.5%
trigger matching	57.18%	99.8%	58.88%	99.7%	51.02%	99.7%	59.06%	99.8%
76.18 GeV $\leq m_{ll} \leq$ 106.18 GeV	53.16%	93.0%	55.12%	93.6%	48.59%	95.2%	55.89%	94.6%
γ ID (*Tight*)	50.09%	94.2%	51.79%	94.0%	45.65%	93.9%	52.49%	93.9%
γ isol (*FixedCutLoose*)	49.17%	98.2%	51.03%	98.5%	44.78%	98.1%	51.65%	98.4%
$p_{\mathrm{T}}^{\gamma} > 50$ GeV and $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} \geq 0.3$	42.73%	86.9%	44.86%	87.9%	33.10%	73.9%	48.86%	94.6%
$m_{ll\gamma} \geq 200$ GeV	42.73%	100.0%	44.85%	100.0%	33.10%	100.0%	48.86%	100.0%

Table B.3.: Absolute and relative event selection efficiencies for the high-mass search for a resonance mass of $m_X = 1$ TeV, derived from signal Monte Carlo event samples in the ggF and VBF production modes for a spin-0 resonance and the gluon- and quark-induced production modes for a spin-2 resonance. The efficiencies are expressed with respect to the $Z \to ee/\mu\mu$ channels.

Event Selection Cutflow (High-Mass Region)	Monte Carlo Signal Samples for $m_X = 2500$ GeV							
	spin-0/ggF		spin-0/VBF		spin-2/gg		spin-2/$q\bar{q}$	
	absolute	relative	absolute	relative	absolute	relative	absolute	relative
all events	100.00%		100.00%		100.00%		100.00%	
GRL	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
primary vertex	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
event quality	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%	100.00%	100.0%
triggers	77.67%	77.7%	78.04%	78.0%	77.47%	77.5%	78.30%	78.3%
after initial cuts	77.67%	100.0%	78.04%	100.0%	77.47%	100.0%	78.30%	100.0%
≥ 2 good leptons	77.93%	100.3%	78.68%	100.8%	74.24%	95.8%	80.58%	102.9%
OS & $m_{ll} > 45$ GeV	74.69%	95.8%	75.67%	96.2%	71.17%	95.9%	77.53%	96.2%
1 *Loose* γ passing preselection	66.19%	88.6%	67.56%	89.3%	60.20%	84.6%	69.73%	89.9%
after precuts	66.19%	100.0%	67.56%	100.0%	60.20%	100.0%	69.73%	100.0%
≥ 1 Higgs candidate	62.62%	94.6%	63.96%	94.7%	57.06%	94.8%	65.92%	94.5%
trigger matching	62.40%	99.7%	63.81%	99.8%	56.82%	99.6%	65.67%	99.6%
76.18 GeV $\leq m_{ll} \leq$ 106.18 GeV	58.66%	94.0%	59.99%	94.0%	54.39%	95.7%	63.04%	96.0%
γ ID (*Tight*)	54.40%	92.7%	55.53%	92.6%	50.46%	92.8%	58.26%	92.4%
γ isol (*FixedCutLoose*)	53.60%	98.5%	54.74%	98.6%	49.47%	98.1%	57.42%	98.6%
$p_{\mathrm{T}}^{\gamma} > 50$ GeV and $p_{\mathrm{T}}^{\gamma}/m_{ll\gamma} \geq 0.3$	46.33%	86.4%	47.31%	86.4%	34.42%	69.6%	53.67%	93.5%
$m_{ll\gamma} \geq 200$ GeV	46.33%	100.0%	47.31%	100.0%	34.42%	100.0%	53.67%	100.0%

Table B.4.: Absolute and relative event selection efficiencies for the high-mass search for a resonance mass of $m_X = 2.5$ TeV, derived from signal Monte Carlo event samples in the ggF and VBF production modes for a spin-0 resonance and the gluon- and quark-induced production modes for a spin-2 resonance. The efficiencies are expressed with respect to the $Z \to ee/\mu\mu$ channels.

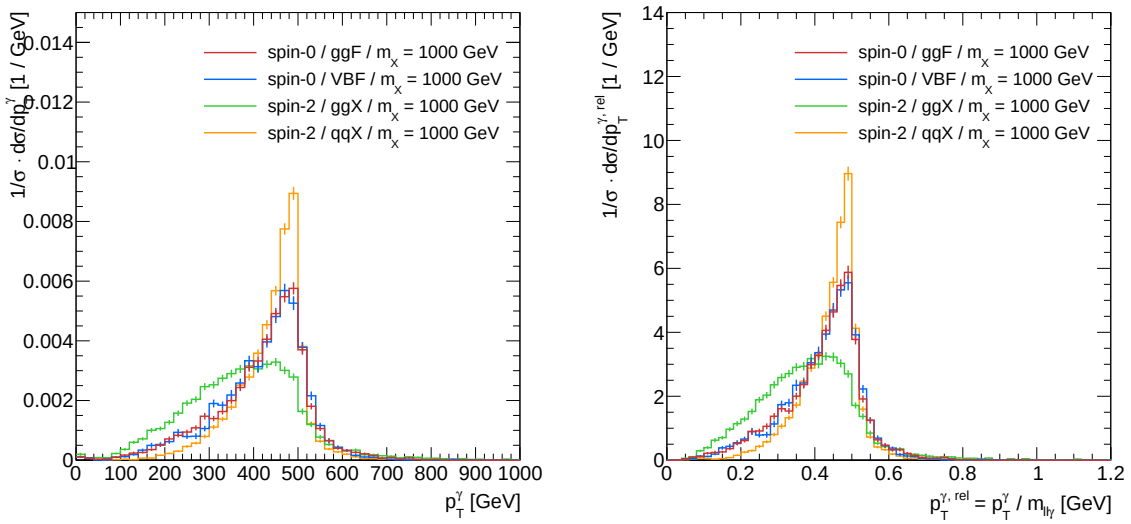


Figure B.2.: Normalised absolute and relative photon p_{T} distributions for the ggF and VBF production modes of spin-0 resonance and the gluon-induced and quark-induced production modes of a spin-2 resonance with a mass of $m_X = 1$ TeV as derived from Monte Carlo simulations after the common preselection stage.

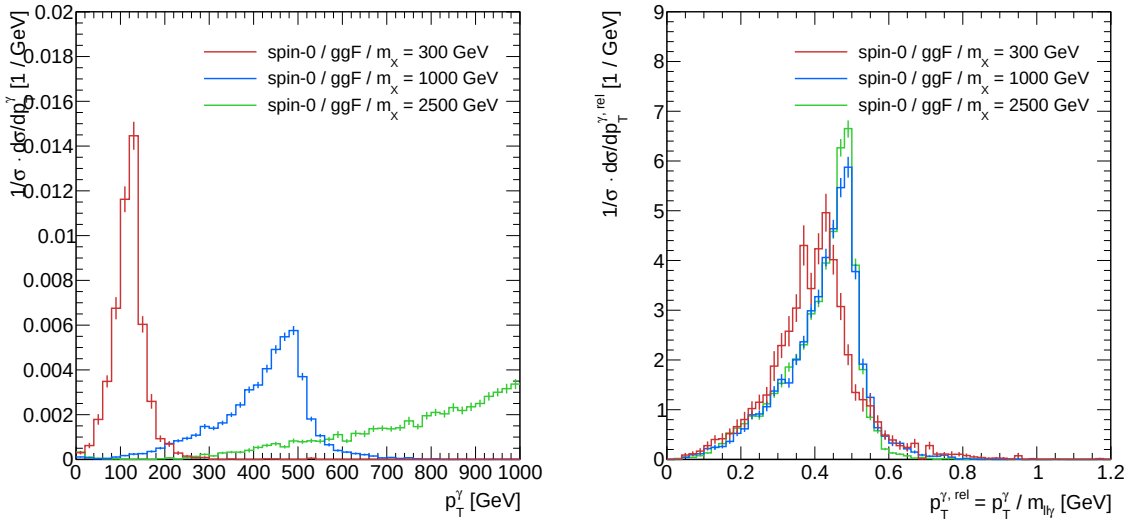


Figure B.3.: Normalised absolute and relative photon p_{T} distributions for the ggF production of a spin-0 resonance for resonance masses of $m_X = 300$ GeV, 1 TeV and 2.5 TeV as derived from Monte Carlo simulations after the common preselection stage.

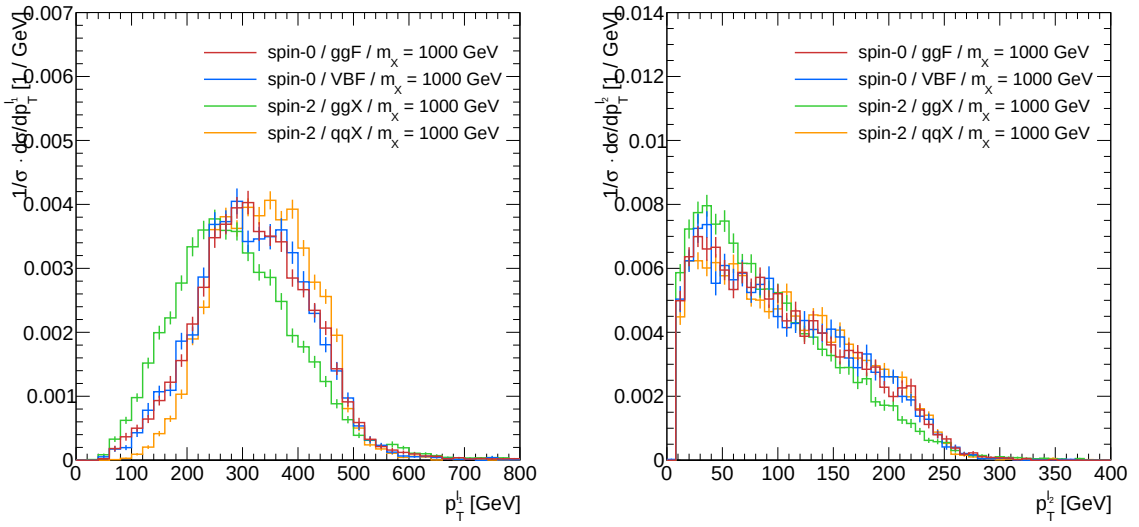


Figure B.4.: Normalised p_{T} distributions of the leading and sub-leading leptons for the ggF and VBF production of a spin-0 resonance and the gluon-induced and quark-induced production modes of a spin-2 resonance with a mass of $m_X = 1$ TeV as derived from Monte Carlo simulations after the common preselection stage.

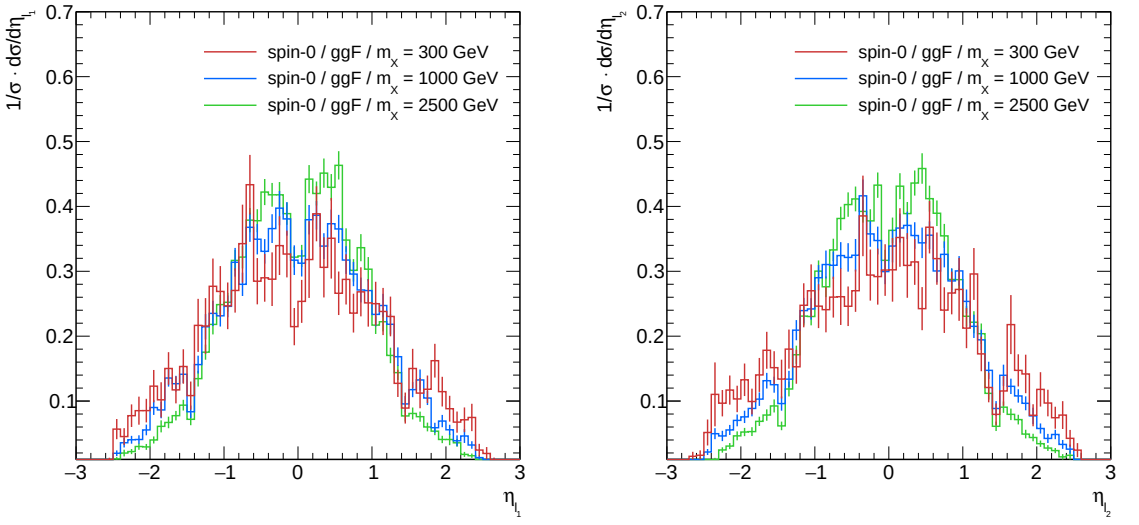


Figure B.5.: Normalised η distributions of the leptons for the ggF production of a spin-0 resonance for resonance masses of $m_X = 300$ GeV, 1 TeV and 2.5 TeV as derived from Monte Carlo simulations after the common preselection stage.

B.2. Results

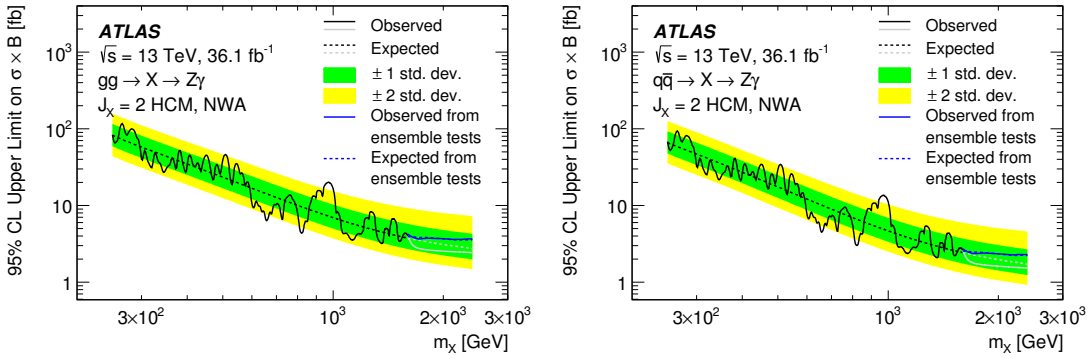


Figure B.6.: Upper limits at the 95% CL on $\sigma(pp \to X) \cdot \mathcal{B}(X \to Z\gamma)$ for a spin-2 resonance produced from (a) gluon-gluon or (b) quark-antiquark initial states as a function of the resonance mass m_X. The solid lines indicate the observed limits and the dashed line the expected limits. The green and yellow uncertainty bands indicate the regions corresponding to ± 1 and ± 2 standard deviations around the expected limits. The results represented with the black and grey lines as well as the uncertainty bands were derived from closed-form asymptotic formulae. For masses above 1.6 TeV, additional results, indicated in blue, were derived using ensemble tests.

C. Miscellanea on $H \to \gamma^*\gamma$

C.1. Splitting Kinematics

When a photon with four-momentum $p^\mu_{\gamma^*}$ splits to produce a pair of leptons ℓ^- and ℓ^+ of mass m_ℓ with four-momenta p^μ_- and p^μ_+, momentum conservation reads

$$p^\mu_{\gamma^*} = p^\mu_- + p^\mu_+ \ . \tag{17.16}$$

Without loss of generality, the z-axis can be defined along the direction of the photon momentum and the coordinate system can be rotated around this axis so that the transverse momentum of the lepton is aligned with the x-axis. Thus, the photon and lepton momenta are given by

$$p^\mu_{\gamma^*} = (E_{\gamma^*}, 0, 0, k) \qquad \text{and} \qquad p^\mu_\mp = (E_\mp, \pm p_x, 0, p^z_\mp) \ . \tag{17.17}$$

While the leptons are required to be on-shell, the photon will be off-shell with a mass of

$$m_{\gamma^*} = \sqrt{E^2_{\gamma^*} - k^2} \tag{17.18}$$

which entails the obvious kinematic boundary $m_{\gamma^*} \leq E_{\gamma^*}$. The existence of the two on-shell final state leptons ensures that $p^2_{\gamma^*} = (p_- + p_+)^2 > 0$ and thus m_{γ^*} will be a real-valued number.

Ensuring momentum conservation along the z-axis, the z-components of the lepton momenta can be parameterised in terms of the photon momentum E_{γ^*} as

$$p^z_- = zE_{\gamma^*} \qquad \text{and} \qquad p^z_+ = k - zE_{\gamma^*} \ . \tag{17.19}$$

Note that a parameterisation in terms of the z-component of the photon momentum k as

$$p^z_- = zk \qquad \text{and} \qquad p^z_+ = (1-z)k \ . \tag{17.20}$$

does not work since it would degenerate in case of extremely off-shell photons with $m_{\gamma^*} \to E_{\gamma^*}$.

The opening angle θ between the lepton and the antilepton is now given by

$$\theta = \theta_- - \theta_+ \tag{17.21}$$

with

$$\tan(\theta_-) = \frac{p_x}{p^z_-} = \frac{p_x}{zE_{\gamma^*}} \qquad \text{and} \qquad \tan(\theta_+) = -\frac{p_x}{p^z_+} = -\frac{p_x}{k - zE_{\gamma^*}} \ . \tag{17.22}$$

The relative sign between both contributions to the overall angle counters the sign obtained from using the signed inverse tangent function (arctan2).

An alternative formulation can be derived from the lepton pair four-momentum product

$$\cos\theta = \frac{E_+E_- - p_+ \cdot p_-}{|\boldsymbol{p}_+||\boldsymbol{p}_-|} \ , \tag{17.23}$$

where

$$E_+E_- - p_+ \cdot p_- = -p^2_x + zE_{\gamma^*}(k - zE_{\gamma^*}) \tag{17.24}$$

and

$$|\boldsymbol{p}_+||\boldsymbol{p}_-| = \sqrt{p_x^2 + z^2 E_{\gamma^*}^2}\sqrt{p_x^2 + (k - zE_{\gamma^*})^2} \; . \tag{17.25}$$

While the adopted parameterisation of the lepton momenta in terms of z and p_x ensures conservation of three-momentum, z and p_x are related and bounded by energy conservation. Specifically, energy conservation yields

$$E_{\gamma^*} - E_- = E_+ \quad \Rightarrow \quad p_x^2 = \frac{1}{4E_{\gamma^*}^2}\left(m_{\gamma^*}^2 + 2zkE_{\gamma^*}\right)^2 - z^2 E_{\gamma^*}^2 - m_\ell^2 \; . \tag{17.26}$$

Requiring $p_x^2 \geq 0$ one finds upper and lower bounds on z:

$$z \leq \frac{k}{2E_{\gamma^*}} + \sqrt{\frac{1}{4} - \frac{m_\ell^2}{m_{\gamma^*}^2}} \qquad \text{and} \qquad z \geq \frac{k}{2E_{\gamma^*}} - \sqrt{\frac{1}{4} - \frac{m_\ell^2}{m_{\gamma^*}^2}} \; . \tag{17.27}$$

C.2. Merged Electron Identification

p_{T}[GeV]	< 20	$[20, 25]$	$[25, 30]$	$[30, 35]$	$[35, 40]$	$[40, 45]$	$[45, 50]$	$[50, 60]$	$[60, 80]$	> 80	combined
eProbabilityHT(track 1)	0.0000	0.0956	0.1401	0.1844	0.2087	0.2214	0.2358	0.2276	0.2040	0.1274	0.1752
eProbabilityHT(track 2)	0.0000	0.0695	0.0950	0.1328	0.1555	0.1843	0.1993	0.1999	0.1806	0.1425	0.1417
n_{Pix}(track 1)	0.0000	0.0025	0.0018	0.0017	0.0024	0.0036	0.0040	0.0049	0.0074	0.0099	0.0038
n_{Pix}(track 2)	0.0000	0.0025	0.0018	0.0017	0.0024	0.0036	0.0040	0.0049	0.0074	0.0099	0.0038
n_{Si}(track 1)	0.0000	0.0014	0.0017	0.0027	0.0017	0.0031	0.0026	0.0045	0.0066	0.0095	0.0034
n_{Si}(track 2)	0.0000	0.0014	0.0017	0.0027	0.0017	0.0031	0.0026	0.0045	0.0066	0.0095	0.0034
$\Delta\eta_1$	0.0000	0.1674	0.2330	0.2789	0.3286	0.3555	0.3819	0.4022	0.4096	0.3989	0.3217
f_1	0.0000	0.0366	0.0528	0.0745	0.0887	0.0968	0.0957	0.1014	0.0975	0.0498	0.0741
f_{side}	0.0000	0.0420	0.0747	0.1153	0.1514	0.1719	0.2066	0.2258	0.2562	0.2891	0.1521
E_{ratio}	0.0000	0.1619	0.2257	0.3152	0.3612	0.4014	0.4193	0.4345	0.4217	0.3896	0.3250
$w_{s,\mathrm{tot}}$	0.0000	0.1037	0.1436	0.2051	0.2418	0.2815	0.2927	0.3223	0.3373	0.3283	0.2360
R_η	0.0000	0.1546	0.2554	0.3621	0.4387	0.4925	0.5481	0.5780	0.5893	0.5418	0.4095
R_ϕ	0.0000	0.0268	0.0353	0.0565	0.0547	0.0619	0.0683	0.0838	0.1294	0.2589	0.0735
$w_{\eta,2}$	0.0000	0.0683	0.1036	0.1692	0.2196	0.2681	0.2955	0.3245	0.3509	0.3515	0.2164
f_3	0.0000	0.0726	0.1105	0.1646	0.2061	0.2338	0.2627	0.2847	0.3055	0.3168	0.2006
R_{had}	0.0000	0.1892	0.2893	0.4206	0.5142	0.5875	0.6366	0.7088	0.7423	0.7586	0.5058
$E_{\mathrm{cluster}}/p_{\mathrm{tracks}}$	0.0000	0.0652	0.0860	0.1072	0.1193	0.1390	0.1633	0.1729	0.1915	0.1761	0.1253
d_0(track 1)	0.0000	0.0075	0.0074	0.0093	0.0157	0.0141	0.0135	0.0115	0.0132	0.0125	0.0108
d_0(track 2)	0.0000	0.0193	0.0243	0.0282	0.0271	0.0315	0.0292	0.0279	0.0215	0.0120	0.0244
$\lvert d_0\rvert/\sigma_{d_0}$(track 1)	0.0000	0.0087	0.0076	0.0085	0.0083	0.0102	0.0103	0.0104	0.0119	0.0135	0.0096
$\lvert d_0\rvert/\sigma_{d_0}$(track 2)	0.0000	0.0115	0.0100	0.0105	0.0148	0.0150	0.0161	0.0210	0.0165	0.0132	0.0142
σ_p/p(track 1)	0.0000	0.0671	0.0855	0.1061	0.1134	0.1063	0.1050	0.0935	0.0739	0.0413	0.0870
σ_p/p(track 2)	0.0000	0.0692	0.0919	0.1128	0.1135	0.1269	0.1266	0.1105	0.0913	0.0586	0.0975
ΔE	0.0000	0.0874	0.1328	0.1979	0.2300	0.2660	0.2911	0.3068	0.3011	0.2850	0.2130
$w_{s,3}$	0.0000	0.0083	0.0126	0.0261	0.0309	0.0334	0.0505	0.0607	0.0895	0.1364	0.0424

Table C.5.: Separations $\langle S_i^2 \rangle$ in each of the p_{T}-bins i as well as combined separations $\langle S^2 \rangle$ for the different variables considered in the merged electron identification before the application of the first-stage cuts. The variables include the shower shape variables defined in Table 5.1 as well as further tracking and track-cluster matching variables listed in Table 15.1. The colour coding reflects the separation values, ranging from low separation (blue) to high separation (red).

η	$[0.0, 0.6]$	$[0.6, 0.8]$	$[0.8, 1.15]$	$[1.15, 1.37]$	$[1.37, 1.52]$	$[1.52, 1.81]$	$[1.81, 2.01]$	$[2.01, 2.37]$	$[2.37, 2.47]$	$\lvert\eta\rvert > 2.47$	combined
eProbabilityHT(track 1)	0.3085	0.2124	0.2859	0.3294	0.3330	0.3447	0.2511	0.0365	0.0000	0.0000	0.2718
eProbabilityHT(track 2)	0.2712	0.1886	0.2196	0.2859	0.2826	0.3064	0.2247	0.0290	0.0000	0.0000	0.2341
n_{Pix}(track 1)	0.0059	0.0062	0.0072	0.0052	0.0031	0.0035	0.0048	0.0105	0.0041	0.0000	0.0062
n_{Pix}(track 2)	0.0059	0.0062	0.0072	0.0052	0.0031	0.0035	0.0048	0.0105	0.0041	0.0000	0.0062
n_{Si}(track 1)	0.0034	0.0043	0.0037	0.0045	0.0057	0.0041	0.0034	0.0078	0.0075	0.0000	0.0041
n_{Si}(track 2)	0.0034	0.0043	0.0037	0.0045	0.0057	0.0041	0.0034	0.0078	0.0075	0.0000	0.0041
$\Delta\eta_1$	0.3860	0.4245	0.4396	0.4331	0.2006	0.3707	0.3369	0.2663	0.2740	0.0000	0.3735
f_1	0.0609	0.0821	0.0977	0.1416	0.0205	0.1292	0.0690	0.0562	0.0242	0.0000	0.0744
f_{side}	0.3004	0.2547	0.2135	0.2085	0.0273	0.1318	0.2856	0.3483	0.1216	0.0000	0.2637
E_{ratio}	0.5055	0.4948	0.4826	0.4721	0.0753	0.4207	0.4477	0.3792	0.1322	0.0000	0.4627
$w_{s,\text{tot}}$	0.4828	0.4587	0.4161	0.3987	0.2837	0.3384	0.4172	0.3997	0.4831	0.0000	0.4449
R_η	0.6114	0.6009	0.6119	0.5912	0.1151	0.5133	0.5957	0.5647	0.5666	0.0000	0.5854
R_ϕ	0.1082	0.1043	0.0862	0.0900	0.0675	0.1144	0.1903	0.2614	0.3629	0.0000	0.1245
$w_{\eta,2}$	0.3995	0.3670	0.3749	0.3411	0.0264	0.2600	0.3157	0.3477	0.3918	0.0000	0.3690
f_3	0.2217	0.2739	0.2719	0.2119	0.0601	0.2583	0.2706	0.2467	0.2213	0.0000	0.2300
R_{had}	0.6040	0.5817	0.5231	0.4826	0.4801	0.6159	0.6146	0.5697	0.5623	0.0000	0.5783
$E_{\text{cluster}}/p_{\text{tracks}}$	0.2883	0.2679	0.2299	0.1767	0.1026	0.0742	0.0447	0.0345	0.0179	0.0000	0.2198
d_0(track 1)	0.0105	0.0148	0.0135	0.0098	0.0094	0.0059	0.0053	0.0047	0.0103	0.0000	0.0101
d_0(track 2)	0.0094	0.0105	0.0091	0.0094	0.0154	0.0091	0.0105	0.0074	0.0278	0.0000	0.0096
$\lvert d_0\rvert/\sigma_{d_0}$(track 1)	0.0207	0.0218	0.0178	0.0159	0.0140	0.0061	0.0080	0.0056	0.0158	0.0000	0.0169
$\lvert d_0\rvert/\sigma_{d_0}$(track 2)	0.0265	0.0228	0.0191	0.0125	0.0195	0.0099	0.0104	0.0070	0.0339	0.0000	0.0208
σ_p/p(track 1)	0.1999	0.1703	0.2519	0.2835	0.2866	0.2610	0.1685	0.1314	0.1463	0.0000	0.2083
σ_p/p(track 2)	0.1743	0.1600	0.2528	0.2840	0.2804	0.2564	0.1639	0.1301	0.1283	0.0000	0.1933
ΔE	0.3681	0.3552	0.2968	0.2807	0.0393	0.2756	0.3301	0.2722	0.0594	0.0000	0.3233
$w_{s,3}$	0.1089	0.0881	0.0708	0.0828	0.1189	0.0446	0.0771	0.1363	0.1403	0.0000	0.0979

Table C.6.: Separations $\langle S_i^2 \rangle$ in each of the η-bins i as well as combined separations $\langle S^2 \rangle$ for the different variables considered in the merged electron identification before the application of the first-stage cuts. The variables include the shower shape variables defined in Table 5.1 as well as further tracking and track-cluster matching variables listed in Table 15.1. The colour coding reflects the separation values, ranging from low separation (blue) to high separation (red).

p_{T}[GeV]	< 20	$[20, 25]$	$[25, 30]$	$[30, 35]$	$[35, 40]$	$[40, 45]$	$[45, 50]$	$[50, 60]$	$[60, 80]$	> 80	combined
eProbabilityHT(track 1)	0.0000	0.1728	0.1997	0.2035	0.2141	0.2140	0.2492	0.2398	0.2019	0.2636	0.2129
eProbabilityHT(track 2)	0.0000	0.1626	0.1491	0.1855	0.1687	0.2230	0.2471	0.2348	0.1960	0.2777	0.1973
n_{Pix}(track 1)	0.0000	0.0236	0.0216	0.0207	0.0226	0.0229	0.0257	0.0305	0.0200	0.0497	0.0249
n_{Pix}(track 2)	0.0000	0.0236	0.0216	0.0207	0.0226	0.0229	0.0257	0.0305	0.0200	0.0497	0.0249
n_{Si}(track 1)	0.0000	0.0065	0.0105	0.0150	0.0087	0.0263	0.0114	0.0259	0.0169	0.0666	0.0182
n_{Si}(track 2)	0.0000	0.0065	0.0105	0.0150	0.0087	0.0263	0.0114	0.0259	0.0169	0.0666	0.0182
$\Delta\eta_1$	0.0000	0.2446	0.2648	0.2674	0.2891	0.3535	0.3796	0.3007	0.3474	0.4941	0.3112
f_1	0.0000	0.2192	0.2122	0.2151	0.2142	0.2331	0.2504	0.2369	0.2165	0.2262	0.2232
f_{side}	0.0000	0.1479	0.1668	0.1701	0.1806	0.1874	0.2532	0.1909	0.2077	0.2593	0.1885
E_{ratio}	0.0000	0.1956	0.2107	0.2146	0.2569	0.2293	0.2739	0.2568	0.3011	0.4048	0.2514
$w_{s,\mathrm{tot}}$	0.0000	0.0497	0.0562	0.0559	0.0839	0.1238	0.1156	0.1370	0.1922	0.3236	0.1193
R_η	0.0000	0.1009	0.1374	0.1324	0.1882	0.1744	0.2091	0.2898	0.3170	0.4121	0.2155
R_ϕ	0.0000	0.0185	0.0151	0.0139	0.0292	0.0251	0.0634	0.0415	0.0585	0.1895	0.0413
$w_{\eta,2}$	0.0000	0.0924	0.1148	0.1275	0.1312	0.1437	0.1809	0.2316	0.2514	0.3284	0.1772
f_3	0.0000	0.1398	0.1266	0.1605	0.1479	0.1668	0.2305	0.2011	0.2680	0.2527	0.1853
R_{had}	0.0000	0.0729	0.1017	0.1155	0.1674	0.1620	0.1746	0.2640	0.2981	0.3893	0.1905
$E_{\mathrm{cluster}}/p_{\mathrm{tracks}}$	0.0000	0.1300	0.1164	0.1082	0.1707	0.1684	0.2000	0.1962	0.2143	0.4385	0.1692
d_0(track 1)	0.0000	0.0308	0.0307	0.0370	0.0495	0.0502	0.0887	0.0549	0.0745	0.1226	0.0548
d_0(track 2)	0.0000	0.0393	0.0298	0.0347	0.0398	0.0706	0.0899	0.0732	0.0848	0.1530	0.0637
$\lvert d_0\rvert/\sigma_{d_0}$(track 1)	0.0000	0.0289	0.0230	0.0277	0.0396	0.0505	0.1071	0.0941	0.0776	0.1576	0.0614
$\lvert d_0\rvert/\sigma_{d_0}$(track 2)	0.0000	0.0292	0.0184	0.0334	0.0434	0.0436	0.0872	0.0738	0.0658	0.1812	0.0552
σ_p/p(track 1)	0.0000	0.0772	0.0643	0.0759	0.0726	0.0652	0.1301	0.1003	0.1104	0.1800	0.0915
σ_p/p(track 2)	0.0000	0.1192	0.1117	0.1074	0.0703	0.1074	0.1460	0.1058	0.1287	0.1153	0.1120
ΔE	0.0000	0.0161	0.0164	0.0233	0.0310	0.0423	0.0352	0.0441	0.0734	0.1397	0.0418
$w_{s,3}$	0.0000	0.0297	0.0169	0.0284	0.0315	0.0382	0.0354	0.0530	0.0609	0.0905	0.0418

Table C.7.: Separations $\langle S_i^2 \rangle$ in each of the p_{T}-bins i as well as combined separations $\langle S^2 \rangle$ for the different variables considered in the merged electron identification after the application of the first-stage cuts. The variables include the shower shape variables defined in Table 5.1 as well as further tracking and track-cluster matching variables listed in Table 15.1. The colour coding reflects the separation values, ranging from low separation (blue) to high separation (red).

η	$[0.0, 0.6]$	$[0.6, 0.8]$	$[0.8, 1.15]$	$[1.15, 1.37]$	$[1.37, 1.52]$	$[1.52, 1.81]$	$[1.81, 2.01]$	$[2.01, 2.37]$	$[2.37, 2.47]$	$\lvert\eta\rvert > 2.47$	combined
eProbabilityHT(track 1)	0.2526	0.1732	0.2731	0.3355	0.3162	0.3263	0.1703	0.0276	0.0000	0.0000	0.2282
eProbabilityHT(track 2)	0.3369	0.2284	0.2672	0.3364	0.2866	0.3772	0.1594	0.0309	0.0000	0.0000	0.2542
n_{Pix}(track 1)	0.0045	0.0197	0.0142	0.0079	0.0029	0.0098	0.0530	0.0793	0.0290	0.0000	0.0189
n_{Pix}(track 2)	0.0045	0.0197	0.0142	0.0079	0.0029	0.0098	0.0530	0.0793	0.0290	0.0000	0.0189
n_{Si}(track 1)	0.0080	0.0148	0.0086	0.0140	0.0082	0.0143	0.0210	0.0469	0.0227	0.0000	0.0152
n_{Si}(track 2)	0.0080	0.0148	0.0086	0.0140	0.0082	0.0143	0.0210	0.0469	0.0227	0.0000	0.0152
$\Delta\eta_1$	0.2257	0.2598	0.2952	0.3498	0.2188	0.2587	0.2007	0.2044	0.1661	0.0000	0.2328
f_1	0.0760	0.0678	0.0840	0.1231	0.0364	0.1087	0.0773	0.0776	0.0272	0.0000	0.0701
f_{side}	0.1402	0.1086	0.1316	0.1181	0.0442	0.1170	0.1335	0.2327	0.0613	0.0000	0.1239
E_{ratio}	0.2778	0.2716	0.2462	0.2514	0.0560	0.2001	0.2429	0.1389	0.0420	0.0000	0.1918
$w_{s,\text{tot}}$	0.1795	0.1496	0.1144	0.1083	0.1849	0.0767	0.1652	0.2105	0.2087	0.0000	0.1618
R_η	0.2269	0.2170	0.2557	0.2792	0.0623	0.1954	0.4126	0.2545	0.1792	0.0000	0.2072
R_ϕ	0.0475	0.0575	0.0495	0.0488	0.0808	0.0618	0.0864	0.1487	0.2096	0.0000	0.0784
$w_{\eta,2}$	0.1376	0.1564	0.1127	0.1013	0.0572	0.0417	0.1268	0.2937	0.1744	0.0000	0.1453
f_3	0.1154	0.1448	0.1278	0.1125	0.0347	0.1171	0.1537	0.1094	0.0928	0.0000	0.1005
R_{had}	0.3441	0.3242	0.2063	0.2130	0.1912	0.2919	0.2731	0.2423	0.2285	0.0000	0.2660
$E_{\text{cluster}}/p_{\text{tracks}}$	0.2898	0.2666	0.1923	0.2013	0.0925	0.0837	0.0875	0.0540	0.0430	0.0000	0.1742
d_0(track 1)	0.0408	0.0839	0.0453	0.0450	0.0142	0.0470	0.0592	0.0551	0.0494	0.0000	0.0402
d_0(track 2)	0.0251	0.0587	0.0401	0.0473	0.0155	0.0520	0.0647	0.0360	0.0433	0.0000	0.0314
$\lvert d_0\rvert/\sigma_{d_0}$(track 1)	0.0427	0.0727	0.0489	0.0589	0.0186	0.0425	0.0469	0.0278	0.0331	0.0000	0.0378
$\lvert d_0\rvert/\sigma_{d_0}$(track 2)	0.0399	0.0783	0.0540	0.0370	0.0172	0.0516	0.0469	0.0198	0.0289	0.0000	0.0356
σ_p/p(track 1)	0.2233	0.1792	0.2234	0.2367	0.2766	0.2002	0.0978	0.0770	0.0577	0.0000	0.2008
σ_p/p(track 2)	0.2268	0.2510	0.3451	0.3340	0.2819	0.2930	0.1338	0.0834	0.0634	0.0000	0.2292
ΔE	0.0577	0.0772	0.0441	0.0663	0.0073	0.0769	0.0944	0.0495	0.0124	0.0000	0.0459
$w_{s,3}$	0.0536	0.0378	0.0766	0.1035	0.1300	0.0698	0.0534	0.1106	0.1008	0.0000	0.0710

Table C.8.: Separations $\langle S_i^2 \rangle$ in each of the η-bins i as well as combined separations $\langle S^2 \rangle$ for the different variables considered in the merged electron identification after the application of the first-stage cuts. The variables include the shower shape variables defined in Table 5.1 as well as further tracking and track-cluster matching variables listed in Table 15.1. The colour coding reflects the separation values, ranging from low separation (blue) to high separation (red).

variable	cut in bins of	efficiency signal		background	
		ϵ_{rel}	ϵ_{abs}	ϵ_{rel}	ϵ_{abs}
R_{had}	$p_{\text{T}} \times \eta$	92.5%	92.5%	31.8%	31.8%
f_3	$p_{\text{T}} \times \eta$	99.5%	92.0%	78.4%	24.9%
R_η	η	95.3%	87.7%	24.7%	6.16%
E_{ratio}	η	97.7%	85.7%	58.5%	3.60%
$w_{s,\text{tot}}$	η	94.4%	80.8%	74.3%	2.68%
$w_{\eta,2}$	η	99.4%	80.4%	89.6%	2.40%
$\Delta\eta_1$	p_{T}	96.1%	77.2%	55.5%	1.33%
R_{had}	η	98.2%	75.8%	74.9%	1.00%
eProbabilityHT(track_2)	η	98.9%	75.0%	96.9%	0.97%
eProbabilityHT(track_1)	η	98.9%	74.2%	96.7%	0.94%
E/p	η	97.8%	72.6%	95.7%	0.90%
R_ϕ	η	88.7%	64.4%	80.7%	0.72%
$\lvert d_0\rvert/\sigma_{d_0}(\text{track}_1)$		98.2%	63.2%	87.7%	0.63%
$\lvert d_0\rvert/\sigma_{d_0}(\text{track}_2)$		96.3%	60.9%	89.5%	0.57%

Table C.9.: The cut sequence of the merged electron identification algorithm along with the corresponding cut efficiencies for the signal and background components for the conditions of the 2017 period.

List of Symbols

Symbol	Meaning
b, B	observed, expected number of background events
$\mathcal{B}$	Born contribution to the matrix element
$\mathcal{B}$	branching ratio
C	covariance matrix
d_0	transverse impact parameter
E	energy
E_{T}	transverse energy
$E_{\mathrm{T}}^{\mathrm{iso}}$	isolation transverse energy
E_{ratio}	shower shape variable
f	generic probability distribution function, distribution of a test statistic
f_1, f_3, f_{side}	shower shape variables
$f_{a/A}$	parton distribution function for parton a and hadron A
f_{rev}	revolution frequency
h	helicity
L	likelihood
$\mathscr{L}$	instantaneous luminosity
$\mathcal{L}$	integrated luminosity
$\mathcal{M}$	scattering amplitude
$\lvert\mathcal{M}\rvert^2$	matrix element
m	mass
n, N	observed, expected number of events
$\dot{N}$	event rate
$\mathcal{O}$	generic observable
P	probability
$\mathcal{P}$	Poisson distribution
p, $\boldsymbol{p}$, p^μ	momentum, three-momentum, four-momentum
p_{T}	transverse momentum
$p_{\mathrm{T}}^{\mathrm{iso}}$	track isolation
Q_{sh}	hard parton shower cut-off scale
q	test statistic
R	jet radius parameter
R_{had}, $R_{\mathrm{had,1}}$	shower shape variables
R_η, R_ϕ	shower shape variables
$\mathcal{R}$	real emission correction
s, S	observed, expected number of signal events
$\boldsymbol{s}$	spin
$\sqrt{s}$	centre-of-mass energy
$\langle S^2 \rangle$	separation
V	covariance matrix
V	Higgs potential
$\mathcal{V}$	virtual correction
v	vacuum expectation value of the Higgs field
$w_{s,3}$, $w_{s,\mathrm{tot}}$	shower shape variables
$w_{\eta,2}$	shower shape variable
x	momentum fraction
Z	significance in multiples of Gaussian standard deviation
z_0	longitudinal impact parameter
α	electromagnetic (QED) coupling
α	type-1 error
α_s	strong (QCD) coupling
β	type-2 error
Γ	decay width
Δ	Sudakov form factor
Δ	absolute uncertainty
δ	relative uncertainty
ΔE	shower shape variable
ϵ	detector acceptance and event selection efficiency
ζ	calorimeter signal significance
$\zeta_c^{(m)}$	fractional contribution of a Higgs production mode m to a category c

The definitions of the shower shape variables are given in Table 5.1.

Symbol	Meaning
η	pseudorapidity
θ	polar angle
θ, $\boldsymbol{\theta}$	parameter(s) (of interest) in a statistical model
λ	likelihood ratio
λ^2	quartic coefficient of the Higgs potential
μ	signal strength
μ^2	quadratic coefficient of the Higgs potential
$\langle\mu\rangle$	average number of interactions per bunch crossing
μ_F, μ_R	factorisation, renormalisation scale
μ_{IR}	soft parton shower cut-off scale
ν, $\boldsymbol{\nu}$	nuisance parameter(s)
ξ_c	fraction of signal events in a category c
π	prior probability distribution
ϱ	background composition
σ	cross section
σ	uncertainty
σ_x, σ_y	horizontal, vertical width of the beam spot profile
τ	mean lifetime
Φ_n	phase space for n particles
ϕ	azimuthal angle

List of Acronyms

Acronym	Meaning
ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC Apparatus
BDT	boosted decision tree
BSM	beyond the Standard Model
CB	combined muon
CDF	cumulative distribution function
CERN	European Organization for Nuclear Research
CL	confidence level
CM	centre-of-mass
CMS	Compact Muon Solenoid
CPU	central processing unit
CR	control region
CSC	cathode strip chamber
DSCB	double-sided crystal ball
ECal	electromagnetic calorimeter
EW	electroweak
ggF, ggH	gluon-gluon fusion
GRL	good-run list
GSF	Gaussian sum filter
HCal	hadronic calorimeter
HEP	high energy physics
HLT	high-level trigger
HXSWG	(LHC) Higgs cross section working group
IBL	insertable B-layer
ID	inner detector
ID	(particle) identification
IR	infrared
ISO	(particle) isolation
JVT	jet-vertex-tagger
L1	level-1 trigger
LEP	Large Electron-Positron Collider
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LHE	Les Houches event
LL	leading-logarithmic
LO	leading order
MC	Monte Carlo
MDT	monitored drift tube
ME	matrix element
ME	extrapolated muon
MPI	multiple parton interactions
MS	muon spectrometer
NLL	next-to-leading-logarithmic
NLO	next-to-leading order
NNLL	next-to-next-to-leading-logarithmic
NNLO	next-to-next-to-leading order
N^3LO	next-to-next-to-next-to-leading order
NWA	narrow-width approximation
PDF	parton distribution function
PDF	probability distribution function
PIX	pixel detector
PU	parametric uncertainties
PV	primary vertex
QCD	quantumchromodynamics
QED	quantumelectrodynamics
RoI	region of interest
RPC	resistive plate chamber
SCT	semiconductor tracker, a.k.a. silicon strip detector
SM	Standard Model of particle physics
SR	signal region
TGC	thin gap chamber
THU	theoretical uncertainties
TRT	transition radiation tracker
UE	underlying event
UV	ultraviolet
VBF, VBFH	vector boson fusion

List of Figures

List of Tables