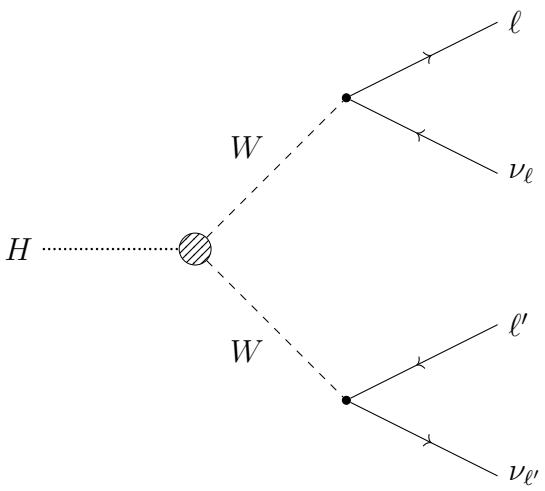
Doktorarbeit

Measurement of $H \to W^{\pm} W^{\mp *} \to \ell \bar{\nu}_{\ell} \ell' \nu_{\ell'}$ Decays at $\sqrt{s} = 13$ TeV with the ATLAS Experiment and Analysis of Effective Higgs Boson Couplings

Carsten D. Burgard



CERN-THESIS-2017-195
24/07/2017

Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

Measurement of $H \to W^{\pm} W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$ Decays at $\sqrt{s} = 13$ TeV with the ATLAS Experiment and Analysis of Effective Higgs Boson Couplings

Doktorarbeit

vorgelegt von

Carsten D. Burgard

Prof. Dr. Karl Jakobs

Fakultät für Mathematik und Physik

Albert-Ludwigs-Universität

Freiburg im Breisgau

30.01.2017

Dekan:	Prof. Dr. Gregor Herten
Betreuer:	Prof. Dr. Karl Jakobs
1. Gutachter:	Prof. Dr. Gregor Herten
2. Gutachter:	Prof. Dr. Marc Schumann
Verkündung des Prüfungsergebnisses:	24.07.2017
1. Prüfer:	Prof. Dr. Karl Jakobs
2. Prüfer:	Prof. Dr. Gregor Herten
3. Prüfer:	Prof. Dr. Stefan Dittmaier

> “If I have seen further it is by standing on ye shoulders of giants.”
>
> *Isaac Newton*

Abstract

This thesis presents and discusses measurements of the coupling of the Higgs boson to vector bosons, using data collected at $\sqrt{s} = 13\,\text{TeV}$ by the ATLAS detector. A full analysis of the first $5.8\,\text{fb}^{-1}$ of LHC Run 2 data investigating the $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay channel is presented, targeting the gluon fusion, vector boson fusion and Higgs strahlung modes of Higgs boson production with dedicated event selections. The analyses presented here include the first measurement of Higgs boson production via gluon fusion in this decay channel using ATLAS data at $\sqrt{s} = 13\,\text{TeV}$. A combined analysis of the $H\to ZZ^{*}\to\ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ and $H\to\gamma\gamma$ decay modes employing $14.8\,\text{fb}^{-1}$ and $13.3\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data, respectively, is also presented. No significant deviation from the Standard Model predictions is observed. The state-of-the-art techniques for obtaining physics predictions for such measurements are discussed, including different methods of obtaining continuous likelihood functions from a finite set of samples. This includes Matrix Element Reweighting as well as different methods of interpolation. The technique of *Effective Lagrangian Morphing*, which was developed in the context of this thesis, is derived from first principles, discussed in detail and supplemented with enhancements and methods to reduce the dimensionality and computational cost and to improve the accuracy of its predictions. A first application of this technique for a measurement of the CP-even and the CP-odd coupling parameters g_{HVV} and g_{AVV} in the $H\to ZZ^{*}\to\ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ channel is presented, and opportunities and challenges for future applications of this technique to measurements in the $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ channel are discussed.

Acknowledgements

I would like to thank my supervisor Prof. Karl Jakobs for the unique opportunity to participate in the research of his distinguished group for experimental high energy physics in Freiburg, for which I feel both grateful and honored.

I would also like to thank my adviser Dr. Karsten Köneke for his continued support throughout my thesis, for his substantial advice, for his patience with my questions, and for his openness to my ideas.

Special thanks go to my estimated colleagues Lydia Brenner, Katharina Ecker, and Adam Kaluza. The close, productive, fruitful and reliable collaboration during the development of the *Effective Lagrangian Morphing* was nothing less than an awe-inspiring experience for me, for which I am truly thankful.

For many interesting discussions that often inspired new ideas and helped to overcome many challenges, I want to thank Michael Dührssen and Nicolas Berger.

I want to thank my dear colleague Ralf Gugel for his motivation and technical skills, which were a great asset in this endeavor.

Thanks also go to all the other members of the ATLAS HWW analysis group. Special thanks go to Kathrin Becker, Ruchi Gupta, Claudia Bertella and Natalie Wieseotte for their dedication to the success of our analysis.

For invaluable help with any organizational challenge, I want to thank Mrs. Christina Skorek.

I want to thank all of my friends for their continued support in all circumstances. Special thanks go to Dimitri Scheftelowitsch and Gottfried Herold for many interesting technical discussions, often times opening new perspectives on the challenges I was facing.

Lastly and mostly, I want to thank my parents for providing me with everything that I needed, be it support or advice, during all the steps of my education up to this point, without which I would not be the person that I am today.

1 Introduction

> **It is no good to try to stop knowledge from going forward. Ignorance is never better than knowledge.**
>
> *Enrico Fermi*

The quest for knowledge about the fundamental building blocks of nature is one of the most primal aspects of science itself. Around the middle of the first millennium BCE, Greek and Indian philosophers proposed the philosophy of atomism, theorizing that nature consists of indivisible *atoms* as fundamental constituents in the void of space and time [1], long before even the word *physics* was coined in the 16th century [2]. It is certainly no coincidence that the German writer Johann Wolfgang von Goethe chose to have his most famous protagonist Faust ask to "**perceive whatever holds the world together in its inmost folds**" [3] on his quest for the deepest knowledge in 1808.

Science has come a long way since the days of philosophical atomism, continuously building on the achievements of the past, developing new theories and rigorously putting them to the test with dedicated experiments. This path, however, was not always a straight one. When Max Planck asked his physics professor Philipp von Jolly in 1874 about the prospects in the field of physics, he was told that "**in this field, almost everything is settled, and only minor holes need to be filled**", which was a common opinion among physicists at the time [4]. Planck himself helped to overturn this judgment with his solution to the black-body radiation problem [5] – one of the "minor holes" – leading to the development of quantum mechanics in the early 20th century. This, and of course Albert Einstein's theory of relativity [6], mark the beginning of the era of modern physics. Max Planck's example illustrates what is a recurring theme to the history of physics – that small "holes" in the currently accepted theory can act as gateways to a new understanding or even a new era of physics.

The Standard Model of Particle Physics was developed in the second half of the 20th century as an attempt to bring together the different theories of fundamental particles and their interactions discovered previously. The set of locally gauge invariant theories from which it originated also had a "hole": The particles these theories described were massless, in direct contradiction with experimental evidence.

Furthermore, any attempt to resolve this discrepancy *ad hoc* would result in the loss of some of the symmetry properties that were used to construct the theory in the first place.

Three independent groups of scientists were able to resolve this issue in 1964 by the mechanism of *spontaneous symmetry breaking*, a mathematical technique already used to explain superconductivity in solid-state physics. Robert Brout, François Englert [7], Peter Higgs [8,9], Gerald Guralnik, Carl Richard Hagen, and Tom Kibble [10] published their findings almost simultaneously. Today, however, this achievement is mostly known as the Brout-Englert-Higgs (BEH) mechanism. François Englert and Peter Higgs were awarded the Nobel Prize in 2013 [11]. Peter Higgs is widely quoted for explicitly stating that this mechanism also predicts the presence of a new particle, the *Higgs boson*.

Despite concerted efforts, the new particle was nowhere to be found using powerful instruments such as the Large Electron Positron Collider (LEP) at the European Organization for Nuclear Research (CERN), or using the TeVatron at Fermilab. Already being referred to as the "elusive" Higgs boson or as "the God(damn) Particle" [12], the discovery of this particle was one of the main motivations for the construction of the Large Hadron Collider (LHC) at CERN.

When the discovery was finally announced on July 4th 2012 by the ATLAS and CMS experiments, the particle physics community celebrated not only the discovery of a new particle, but of the last missing puzzle piece to the Standard Model of Particle Physics. The initial publications were cautiously titled "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC" [13] and "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC" [14]. As larger amounts of data were collected, and more and more detailed investigations could be performed, confidence grew that the newly discovered particle may indeed be the long sought-for Standard Model Higgs boson. A combined analysis published in 2016 is already titled "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" [15]. Following the principle of Ockhams razor to "**admit no more causes of natural things than such as are both true and sufficient to explain their appearances**" as Isaac Newton put it [16], it is natural to assume that the newly discovered particle is indeed the Standard Model Higgs boson predicted by the BEH mechanism of electroweak symmetry breaking (EWSB). This mechanism makes precise predictions about the properties of this Higgs boson. For example, it is expected to be a CP-even, scalar particle that couples to the electroweak bosons, and by extension of the mechanism also to all fermions, in a manner proportional to their mass. Measurements of the properties of the Higgs boson, and especially its couplings, are thus the appropriate tool to test this assumption. And while

the measurements performed on the data set collected by the LHC at $\sqrt{s} = 7$ and 8 TeV have already helped to constrain the spin and CP properties [17, 18] as well as the couplings [15] of the Higgs boson significantly, there is still much room for deviations from the Standard Model predictions.

This thesis presents a detailed analysis of pp collision data collected at $\sqrt{s} =$ 13 TeV in the years 2015 and 2016, searching for events containing Higgs boson candidates. A special focus is put on the Higgs bosons decay mode to a pair of W bosons which further decay into leptons, which is the most sensitive analysis channel with respect to the total signal strength. Different modes of producing Higgs bosons are investigated, with a special focus on Higgs boson production via gluon fusion, which is the dominant production mode at the LHC. However, Higgs boson production via vector boson fusion or in association with a W boson are also considered, and other decay modes to a pair of leptonically decaying Z bosons or to a pair of photons are discussed briefly. Dedicated analysis techniques are developed to allow precise measurements of the properties of the Higgs boson, aiming to "decipher the nature of the Higgs sector", as the title of the newest handbook of LHC cross sections eloquently claims [19]. Is it indeed the long sought-for Higgs boson completing the Standard Model of particle physics, or does this discovery once more revolutionize our understanding of nature? One thing however is certain: The discovery marks the end of one chapter of particle physics – the search for the elusive Higgs boson – and marks the beginning of a new one – the measurement of its properties.

The text of this thesis is structured as follows: Chapter 2 introduces the theoretical framework. Chapter 3 presents the Large Hadron Collider and the ATLAS Experiment and discusses the phenomenology of the Higgs boson at the LHC. Chapter 4 presents the analysis conducted to measure Higgs boson production in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay mode. Chapter 5 discusses a combined measurement of the $H \to ZZ^{*}$ and $H \to \gamma\gamma$ decay modes and presents the technique of *Effective Lagrangian Morphing* and its application for Higgs boson coupling measurements. Finally, Chapter 6 summarizes the findings of this thesis and gives an outlook on future studies.

Declaration

Of this work, the chapters 2 and 3 are mostly reviewing existing literature. Section 2.4.3 is a notable exception, featuring original research performed by the author, which has nevertheless been previously published [19, 20]. Chapters 4 and 5 constitute original research to which the author has contributed significantly, and which has partially been published [21–23]. Finally, chapters 1 and 6 are of purely supplementary nature and do not document any research *per se*.

2 Theoretical Framework

> **Each piece, or part, of the whole of nature is always merely an approximation to the complete truth, or the complete truth so far as we know it. In fact, everything we know is only some kind of approximation because we know that we do not know all the laws as yet.**
>
> *Richard P. Feynman*

This chapter outlines the Lagrange formulation of the *Standard Model of Elementary Particle Physics* (SM), a successful framework to study the fundamental laws governing the universe. Special attention is devoted to the Brout-Englert-Higgs (BEH) mechanism of electroweak symmetry breaking (EWSB). A more detailed introduction including all aspects of the model can be found in many textbooks [24–28].

All equations in this chapter use natural units, expressing physical action in units of the reduced Planck constant $\hbar$ and velocities in units of c, the speed of light in vacuum. The letter s, which is also present in the title of this thesis, refers to a systems center-of-mass energy, deriving from the Mandelstam variables s, t, and u [29]. The notations $\mathfrak{R}(x)$ and $\mathfrak{I}(x)$ denote the real and imaginary parts of some complex quantity x. As far as matrices are printed in the text, entries equal to zero are omitted.

2.1 Introduction

On the smallest scales currently accessible to observation, the fundamental building blocks of the universe are the elementary particles. Figure 2.1 shows a popular representation of the particle content of the Standard Model.

The table is separated into fermions, *i. e.*, particles with half integer spin on the left hand side, and bosons, *i. e.*, particles with integer spin on the right hand side. The first three columns show the three generations of fermions, the rows being occupied by up- and down-type quarks, charged and neutral leptons. The right half of the graphic is occupied by the various gauge bosons connected with weak, electromagnetic and strong interactions. The recently discovered Higgs boson, as well as the so-far unobserved graviton, are also shown.

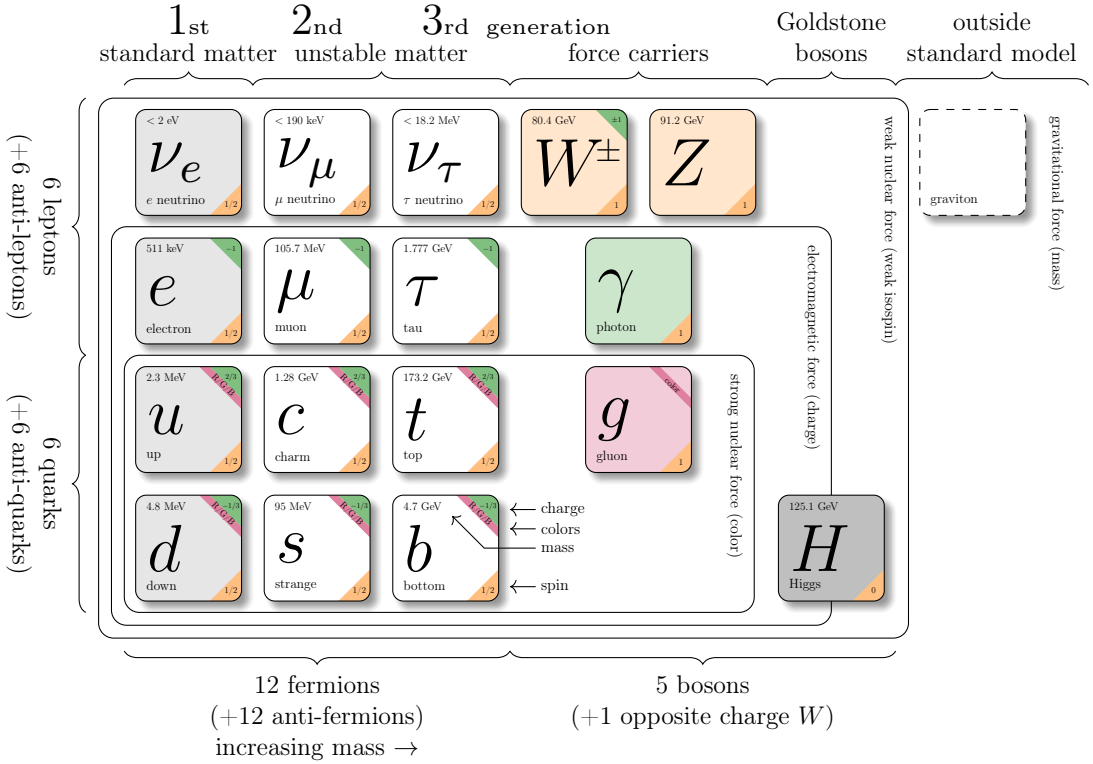


Figure 2.1.: Popular depiction of the particle content of the Standard Model of Particle Physics [30], including the respective quantum numbers as well as mass measurements and limits [31].

While it is very intuitive to imagine interactions of these particles akin to collisions of billiard balls, the most successful theoretical formulation employs *fields* as the fundamental objects, of which the particles are mere excitations. This interpretation of particles as quanta of universal fields and the underlying mathematical formulation are commonly referred to as *Quantum Field Theory* (QFT). An especially elegant formulation of QFT and the Standard Model can be achieved using the Lagrange-formalism and group theory, which are briefly revisited in the next sections to introduce the notation used and highlight some specific aspects that are of interest for the remainder of this chapter.

2.1.1 Lie groups

Lie groups are infinite groups that are also differentiable manifolds, *i. e.*, that are locally similar enough to linear space for the group operations to allow for

differential calculus. Any group element continuously connected to the identity $\mathbb{1}$ can be written as

$$U = e^{i\theta^a T^a}\mathbb{1}, \tag{2.1.1}$$

where θ^a are numbers parametrizing the group elements, and T^a are the group generators. These form a *Lie algebra* defined through its commutation relations

$$\left[T^a, T^b\right] = if^{abc}T^c, \tag{2.1.2}$$

where f^{abc} are the structure constants. A Lie group is abelian if $f^{abc} = 0\,\forall a, b, c$ and non-abelian otherwise.

Well-known examples for Lie groups include

- the unitary groups $\mathrm{U}(N)$, representable by the set of all unitary $N \times N$ matrices and multiplication as the group operation,
- the *special* unitary groups $\mathrm{SU}(N)$, representable by the set of all unitary $N \times N$ matrices with determinant $D = 1$ and multiplication as the group operation,
- the orthogonal groups $\mathrm{O}(N)$, representable by the set of all orthogonal $N \times N$ matrices and multiplication as the group operation, and
- the special orthogonal groups $\mathrm{SO}(N)$, representable by the set of all orthogonal $N \times N$ matrices with determinant $D = 1$ and multiplication as the group operation.

Notably, all of these examples are only abelian for $N = 1$, where $\mathrm{U}(1)$ is more commonly known as the circle group of all complex numbers $c \in \mathbb{C}$ that have an absolute value of $\|c\| = 1$ and are thus located on the complex unit circle.

2.1.2 Lagrangian densities

A Lagrangian density (or simply Lagrangian) is a function $\mathscr{L}$ of some continuously varying fields $\phi(x^\mu)$. Given the functional form of the Lagrangian for some physical system, the equations of motion can be generated using the Euler-Lagrange equation

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial \left(\partial_\mu \phi\right)} \right) - \frac{\partial \mathscr{L}}{\partial \phi} = 0, \tag{2.1.3}$$

where the shorthand ∂_μ is used for $\frac{\partial}{\partial x^\mu}$.

In some sense, this formulation reduces the task of understanding the physics of some particular system to coming up with the correct Lagrangian for the system at hand. In general, different choices for the Lagrangian $\mathscr{L}$ will yield different physical equations of motion.

2.1.3 Perturbation theory

Even though the fundamental objects of QFT are fields, and not their quanta (or particles) themselves, the particle interpretation of QFT interaction is still valid in some cases.

The derivation of the equations of motion for the (classical) fields using the Euler-Lagrange-Equation is usually not trivially possible. Solution of the differential equations can, however, be facilitated using Greens functions. Considering for example the Lagrangian of the ϕ^4 theory

$$\mathscr{L} = \frac{1}{2}\left(\partial\phi\right)^2 - \frac{1}{2}m^2\phi^2 + \frac{g}{4!}\phi^4, \tag{2.1.4}$$

where ϕ is a scalar field with mass m, this field exhibits a self-interaction with a coupling strength g. As long as this interaction g is small enough, one can expand all interacting Greens functions in a power series in g.

Individual contributions to these sums can be visualized as *Feynman diagrams* like the one decorating the cover of this thesis. These diagrams visualize the excitations of the fields as actual particles interacting with each other and are valuable assets in constructing and performing perturbative calculations in Quantum Field Theories.

However, individual contributions (or diagrams) may yield infinite results. In order to proceed, one can rescale the field

$$\phi \to \sqrt{Z}\phi_R, \tag{2.1.5}$$

using some scaling parameter $Z \in \mathbb{C}$. The Lagrangian now reads

$$\mathscr{L} = \frac{1}{2}Z\left(\partial\phi\right)^2 - \frac{1}{2}Zm^2\phi^2 + \frac{Z^2 g}{4!}\phi^4. \tag{2.1.6}$$

In realizing that the *bare* parameters Z, m and g might themselves be infinite, one can now rewrite the parameters in the Lagrangian in terms of *renormalized* parameters

$$Z = 1 + \delta_Z, \qquad m^2 Z = m_R^2 + \delta_m, \quad \text{and} \quad Z^2 g = g_R + \delta_g \tag{2.1.7}$$

to obtain the *renormalized* Lagrangian

$$\mathscr{L} = \frac{1}{2}\left(\partial\phi_R\right)^2 - \frac{1}{2}m^2\phi_R^2 + \frac{g}{4!}\phi_R^4 + \frac{1}{2}\delta_Z\left(\partial\phi_R\right)^2 - \frac{1}{2}\delta_m\phi_R^2 + \frac{\delta_g}{4!}\phi_R^4. \tag{2.1.8}$$

The terms containing δs are called *counterterms* and can be used to absorb the divergences. Any remaining ambiguities can be resolved in deciding for a specific

subtraction scheme, which typically introduces an unphysical constant, the *renormalization scale*, where the details depend on the chosen scheme. For examples, the coupling constant g will in perturbative calculations typically depend on the momentum or energy scale at which the interaction takes place.

However, the dependence on any unphysical parameter vanishes as more terms of the power series are included. Theories for which the parameters in the Lagrangian are sufficient to absorb all infinities appearing in the power series are called *renormalizable*.

2.1.4 Global gauge symmetries

Considering the Lagrangian

$$\mathscr{L} = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi, \tag{2.1.9}$$

one can employ Eq. 2.1.3 to obtain

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi = 0, \tag{2.1.10}$$

the Dirac equation for a free fermionic field ψ of mass m.

Eq. 2.1.10 is invariant under any global phase (or "gauge") transformation

$$\psi\left(x\right) \rightarrow U(\alpha)\,\psi(x) \tag{2.1.11}$$

where the transformations

$$U(\alpha) = e^{i\alpha} \tag{2.1.12}$$

form the abelian Lie group U(1).

Through Noether's theorem [32], this symmetry implies the existence of a conserved current. This can be seen by considering the infinitesimal transformation

$$\psi \rightarrow \left(1 + i\alpha\right)\psi. \tag{2.1.13}$$

Invariance requires the Lagrangian (Eq. 2.1.9) to stay unchanged, thus

$$0 = \delta\mathscr{L} \tag{2.1.14}$$

$$= \frac{\partial\mathscr{L}}{\partial\psi}\delta\psi + \frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\psi\right)}\delta\left(\partial_{\mu}\psi\right) + \text{h. c.} \tag{2.1.15}$$

$$= i\alpha\left[\frac{\partial\mathscr{L}}{\partial\psi} - \partial_{\mu}\left(\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\psi\right)}\right)\right]\psi + i\alpha\partial_{\mu}\left(\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\psi\right)}\psi\right) + \text{h. c.}, \tag{2.1.16}$$

where the first term vanishes by virtue of the Euler-Lagrange-Equation 2.1.3. Hence, Eq. 2.1.16 reduces to the form of a conserved current

$$\partial_\mu j^\mu = 0, \tag{2.1.17}$$

where, introducing the charge $e = 2\alpha$,

$$j^\mu = \frac{ie}{2}\left(\frac{\partial \mathscr{L}}{\partial\left(\partial_\mu \psi\right)}\psi - \bar{\psi}\frac{\partial \mathscr{L}}{\partial\left(\partial_\mu \bar{\psi}\right)}\right). \tag{2.1.18}$$

Employing Eq. 2.1.9, this leads to

$$j^\mu = -e\bar{\psi}\gamma^\mu\psi, \tag{2.1.19}$$

thus j^μ can be identified as the electromagnetic charge current density. Equation 2.1.17 hence implies that the electromagnetic charge

$$Q = \int d^3x j^0 \tag{2.1.20}$$

is a conserved quantity.

2.1.5 Local gauge symmetries and interactions

In trying to generalize this approach to transformations of the type

$$\psi \to e^{i\alpha(x)}\psi \tag{2.1.21}$$

for any arbitrary space-time dependent function $\alpha(x)$, it becomes apparent that the Lagrangian in Eq. 2.1.9 is no longer invariant under this type of *local* U(1) gauge transformation.

However, local gauge invariance can be enforced by replacing the derivative ∂_μ in Eq. 2.1.9 with the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu - ieA_\mu\left(x\right), \tag{2.1.22}$$

where $-e$ is the charge of the Dirac particle and the vector field A_μ transforms under $U((\alpha(x))$ as

$$A_\mu\left(x\right) \to A_\mu\left(x\right) + \frac{1}{e}\partial_\mu\alpha\left(x\right). \tag{2.1.23}$$

With this covariant derivative, the Lagrangian becomes

$$\mathscr{L} = i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\bar{\psi}\psi \tag{2.1.24}$$
$$= i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi + e\bar{\psi}\gamma^\mu \psi A_\mu \tag{2.1.25}$$
$$= i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - j^\mu A_\mu. \tag{2.1.26}$$

The gauge field A_μ couples to the Dirac field ψ in exactly the same way as the photon interacts with charged particles. In going from Eq. 2.1.25 to Eq. 2.1.26, a part of the last term has been identified as the electromagnetic current density introduced in Eq. 2.1.19.

Theories of interactions generated with this mechanism are called *gauge theories*. Interestingly, they are very friendly to perturbative approaches, as they are very often renormalizable when used in three space dimensions and one time dimension. The coupling constants of these theories – in this case the electron charge e – will of course depend on the renormalization scale chosen. However, this dependency can be derived from first principles in the form of the *renormalization group equation* for some concrete perturbative theory.

In order to regard A_μ as the physical photon field, the corresponding kinetic energy term needs to be added. Any gauge invariant extension can only involve the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.1.27}$$

Combining these insights yields the Lagrangian of Quantum Electrodynamics

$$\mathscr{L} = \bar{\psi}\left(i\gamma^\mu \partial_\mu \psi - m\right)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{2.1.28}$$

If one were to consider a mass term for the photon, the Lagrangian would read

$$\mathscr{L}' = \mathscr{L} - \frac{1}{2} m_\gamma^2 A_\mu A^\mu. \tag{2.1.29}$$

However, this again would violate gauge invariance, which can be easily seen by using Eq. 2.1.23 to find

$$A_\mu A^\mu \to \left(A_\mu + \frac{1}{e}\partial_\mu \alpha\right)\left(A^\mu + \frac{1}{e}\partial^\mu \alpha\right) \neq A_\mu A^\mu \tag{2.1.30}$$

As mass terms are seemingly forbidden by gauge invariance, the photon must be massless. The implications of this observation are investigated in the following sections.

2.1.6 Spontaneously broken symmetries

In order to understand how apparently symmetry-violating mass terms can be introduced to a theory, one can investigate the Lagrangian of the linear sigma model with a complex scalar field $\phi(x)$. This Lagrangian reads

$$\mathscr{L} = \frac{1}{2}\left(\partial_\mu \phi\right)\left(\partial^\mu \phi^*\right) - V(\phi)\,, \tag{2.1.31}$$

where

$$V(\phi) = \mu^2 \phi\phi^* - \frac{\lambda}{4}\phi^2\phi^{2*}. \tag{2.1.32}$$

This Lagrangian is invariant under any transformation

$$\phi \to \phi e^{i\alpha}, \tag{2.1.33}$$

with α being some angle of rotation.

The potential $V(\phi)$ from Eq. 2.1.32 is minimized when

$$|\phi| = \sqrt{\frac{2\mu^2}{\lambda}} \stackrel{\text{def}}{=} v. \tag{2.1.34}$$

Suppose now to transform to a pair of real fields σ and π using

$$\phi(x) = \left(v + \frac{1}{\sqrt{2}}\sigma(x)\right) e^{i\beta\pi(x)} \tag{2.1.35}$$

such that the minimum is reached for $\sigma(x) = 0$. In this notation, the Lagrangian reads

$$\begin{aligned}\mathscr{L} = &\frac{1}{2}\left(\partial_\mu \sigma\right)^2 \\ &+ \left(\sqrt{\frac{2\mu^2}{\lambda}} + \frac{\sigma}{\sqrt{2}}\right)^2 \beta^2 \left(\partial_\mu \pi\right)^2 \\ &+ \frac{\mu^4}{\lambda} - \mu^2\sigma^2 - \frac{\mu\sqrt{\lambda}}{2}\sigma^3 - \frac{\lambda}{16}\sigma^4\end{aligned} \tag{2.1.36}$$

and now contains a term

$$-\frac{1}{2}\left(2\mu^2\right)\sigma^2,$$

which can be interpreted as a mass term with $m = \sqrt{2}\mu$ for the field σ, whereas the field π remains massless.

The appearance of massless particles by virtue of spontaneously breaking a continuous symmetry is a general result known as *Goldstone's theorem* [33], the particle π taking the role of the *Goldstone boson* in this case.

2.1.7 The Brout-Englert-Higgs mechanism

The free Lagrangian

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.1.37}$$

with the electromagnetic field strength tensor $F_{\mu\nu}$ as introduced in Section 2.1.5 corresponds to a "universe of light", containing nothing but photons.

It is invariant under the local gauge transformation

$$A_\mu(x) \to A_\mu(x) - \frac{\partial}{\partial x_\mu} \eta(x) \tag{2.1.38}$$

for some arbitrary function $\eta(x)$.

Suppose now that the photons would be found to carry a non-zero mass in this universe. Naive addition of an explicit mass term breaks the local gauge symmetry, as the Lagrangian

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu \tag{2.1.39}$$

is no longer invariant under the transformation defined in Eq. 2.1.38.

However, to overcome this, one can add a new scalar field ϕ that

- carries a charge $-e$,
- interacts with itself as well as with photons, and
- is subject to a global potential

$$V(\phi) = -\mu^2 |\phi|^2 + \lambda^2 |\phi|^4 . \tag{2.1.40}$$

The Lagrangian including the field could read

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \left| \left(\frac{\partial}{\partial x_\mu} - ieA_\mu(x) \right) \phi \right|^2 - V(\phi) . \tag{2.1.41}$$

This is invariant under the local gauge transformation

$$A_\mu(x) \to A_\mu(x) - \frac{\partial}{\partial x_\mu}\eta(x) \tag{2.1.42}$$

$$\phi(x) \to e^{ie\eta(x)}\phi(x). \tag{2.1.43}$$

For $\mu^2 > 0$, the potential $V(\phi)$ has a minimum at

$$\langle\phi\rangle = \sqrt{\frac{\mu^2}{2\lambda}} \overset{\text{def}}{=} \frac{v}{\sqrt{2}}. \tag{2.1.44}$$

A sketch of this situation is shown in Fig. 2.2, exhibiting the characteristic "Mexican Hat" (or "lemon press") shape. By reparametrizing

$$\phi(x) = \frac{v + h(x)}{\sqrt{2}} e^{i\frac{\chi(x)}{v}}, \tag{2.1.45}$$

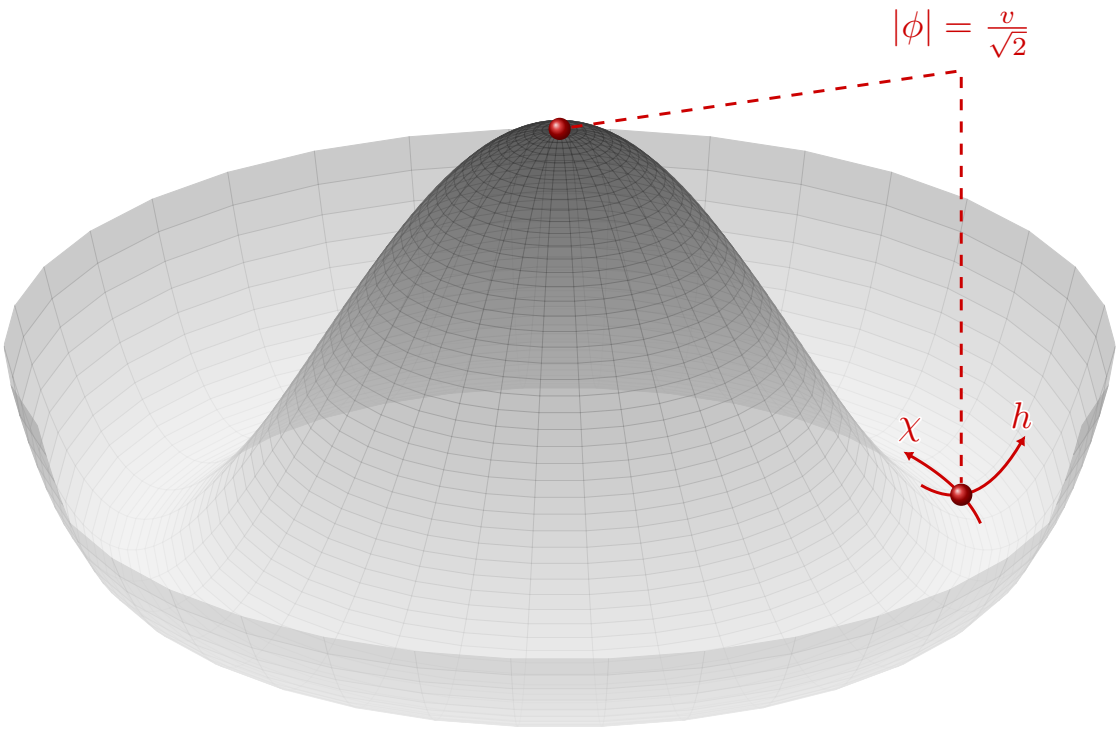


Figure 2.2.: A sketch of the potential $V(\phi)$ from Eq. 2.1.40 for $\mu^2 > 0$ and $\lambda^2 > 0$. The minimum defined in Eq. 2.1.44 and the reparametrization defined in Eq. 2.1.45 are also shown.

the Lagrangian changes to

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(ev)^2 A_\mu A^\mu - \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{2}e^2 A_\mu A^\mu h^2 + v e^2 A_\mu A^\mu h + \frac{1}{2}\partial_\mu h \partial^\mu h.$$

In this notation, it suddenly becomes apparent that the Lagrangian contains

- a photon mass term for $m = ev$, and
- a physical Higgs field $h(x)$ with mass $m = \sqrt{2\lambda}v$

while maintaining the local gauge invariance defined in Eq. 2.1.42–2.1.43. The χ degree of freedom can be eliminated by choosing a particular gauge, defined as the unitary gauge.

This mechanism of spontaneously breaking local gauge symmetries was discovered by Robert Brout and François Englert [7], by Peter Higgs [8,9], as well as by Gerald Guralnik, Carl Richard Hagen, and Tom Kibble [10]. François Englert and Peter Higgs were awarded a Nobel Prize in 2013 [11].

2.2 The Standard Model of Particle Physics

With the formalism introduced in Section 2.1, an extension of the notation and mechanisms to the full particle content of the Standard Model is now straight forward.

The complete gauge group of the Standard Model is

$$\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y\,, \tag{2.2.1}$$

where $\mathrm{SU}(3)_C$ is the gauge group of Quantum chromodynamics (QCD) acting on colour charge (subscript C), and $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ is the subgroup associated with the Glashow-Salam-Weinberg model of electroweak (EW) interactions, acting on left-handed fermions (subscript L) and hypercharge (subscript Y).

2.2.1 Quantum chromodynamics

The quarks of the Standard Model are arranged in colour triplets, that is, one triplet per quark flavour. The theory of Quantum chromodynamics is generated by an SU(3) symmetry with transformations

$$U(x) = e^{i\sum_{j=1}^{8}\gamma_j(x)\frac{\lambda_j}{2}}, \tag{2.2.2}$$

with arbitrary gauge functions $\gamma_j(x)$ acting on these triplets. As $3 \otimes 3 = 8 \oplus 1$, there are eight coloured generators λ_i. These are often written in a standard basis as the Gell-Mann-Matrices

$$\lambda^1 = \begin{pmatrix} & 1 & \\ 1 & & \\ & & \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} & -i & \\ i & & \\ & & \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} 1 & & \\ & -1 & \\ & & \end{pmatrix},$$

$$\lambda^4 = \begin{pmatrix} & & 1 \\ & & \\ 1 & & \end{pmatrix}, \quad \lambda^5 = \begin{pmatrix} & & -i \\ & & \\ i & & \end{pmatrix}, \quad \lambda^6 = \begin{pmatrix} & & \\ & & 1 \\ & 1 & \end{pmatrix},$$

$$\lambda^7 = \begin{pmatrix} & & \\ & & -i \\ & i & \end{pmatrix}, \quad \lambda^8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}.$$

As in Section 2.1.5, local gauge symmetry is obtained by substituting a covariant derivative, this time of the form

$$\mathcal{D}_\mu = \partial_\mu + ig_C \sum_{j=1}^{8} \frac{\lambda^j}{2} G_\mu^j(x), \tag{2.2.3}$$

with the gluon fields G_μ^j and the strong coupling constant g_C. Each index j corresponds to one specific type of gluon, identified by a specific colour-anticolour combination. The ninth gluon would be a colour-neutral, photon-like singlet state, corresponding to a long range strong interaction, for which there is no experimental evidence. The Lagrange density of the gluon field is

$$\mathscr{L} = -\frac{1}{4}\sum_{j=1}^{8} G_j^{\mu\nu} G_{\mu\nu}^j \tag{2.2.4}$$

with the gluon field strength tensor

$$G^j_{\mu\nu} = \partial_\mu G^j_\nu - \partial_\nu G^j_\mu - g_C \sum_{k,l=1}^{8} g^{jkl} G^k_\mu G^l_\nu, \tag{2.2.5}$$

where g^{jkl} are the structure constants of the SU(3) group.

Bound states of QCD are called hadrons. Several types of hadrons have been observed, which are most easily classified by their quark content.

- A combination of a quark and an anti-quark can form a singlet:

  $$3 \otimes \overline{3} = 8 \oplus 1$$

  These are called *mesons*. Mesons are generally unstable with mean life times in the range of 10^{-8} s to 10^{-23} s, depending on the type of interaction in their decay vertex [31].
- A combination of three quarks (or three anti-quarks) can also form a singlet:

  $$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$$

  These are called *baryons*. The two most stable baryons are the proton and the neutron, all other baryons having mean life times of 10^{-10} s or shorter [31].
- Four quarks and one anti-quark (or vice-versa) can also form a singlet:

  $$3 \otimes 3 \otimes 3 \otimes 3 \otimes \overline{3} = 35 \oplus 3 \cdot 27 \oplus 2 \cdot \overline{10} \oplus 4 \cdot 10 \oplus 8 \cdot 8 \oplus 3 \cdot 1$$

  Such *pentaquarks* have long been disputed and were discovered beyond reasonable doubt only recently [34].

No other type of bound QCD state has been observed yet, although the Standard Model would allow for bound states of gluons, so-called *glueballs*, as two or three gluons could also form a colour singlet. However, as the total fermion number of this state is zero, it is expected to mix with mesonic states.

Although no stringent proof exists until this day, it is generally believed that it is by virtue of its non-abelian nature that QCD is *confining*, that is, bound states can never be truly dissolved as the attractive colour force between them does not diminish with increasing distance. If the amount of energy stored in the colour field is sufficient, new quark-antiquark pairs are generated, allowing the bound state to split into several colour-neutral hadrons. This process is referred to as *hadronization*.

2.2.2 Electroweak interactions

The left-handed fermions of the Standard Model are arranged in doublets of weak isospin I, whereas their right-handed counterparts form isosinglets. The left- and right-handed components are obtained from the fermion fields ψ via

$$\psi_{L,R} = \frac{1}{2}\left(1 \mp \gamma_5\right)\psi. \tag{2.2.6}$$

The three generations can be generalized by a template pattern, the second and third generation only differing from the first one in mass. This template is shown in Table 2.1. Right-handed neutrinos are omitted due to the lack of experimental evidence.

The Glashow-Salam-Weinberg model of electroweak interactions constitutes the subgroup $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ of the complete Standard Model gauge symmetry. It is generated by local gauge transformations of the form

$$U\left(x\right) = \exp\left(i\frac{Y}{2}\alpha\left(x\right) + iI\sum_i \sigma_i\beta_i(x)\right), \tag{2.2.7}$$

where Y is the hypercharge, I is the weak isospin, which is $1/2$ for left-handed and zero for right-handed fermions, and σ_i are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}, \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

The $\mathrm{SU}(2)_L$ gauge boson fields are W_μ^a $(a = 1, 2, 3)$, whereas the $\mathrm{U}(1)_Y$ gauge boson is B_μ. The full Lagrangian reads

$$\mathscr{L} = -\frac{1}{4}W_{\mu\nu}^{a2} - \frac{1}{4}B_{\mu\nu}^2 + \left(\mathcal{D}_\mu\phi\right)^\dagger\left(\mathcal{D}_\mu\phi\right) - V(\phi)\,, \tag{2.2.8}$$

Table 2.1.: Arrangement of first generation quarks u & d and leptons e & ν_e in weak isodoublets and isosinglets.

	quarks	leptons
right-handed	$u_R,\ d_R$	e_R
left-handed	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$

with

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.2.9}$$

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\varepsilon^{abc} W^b_\mu W^c_\nu \tag{2.2.10}$$

with the totally antisymmetric tensor ε^{abc}, and where the covariant derivative is

$$\mathcal{D}_\mu \phi = \partial_\mu \phi - igW^a_\mu \frac{\sigma^a}{2}\phi - \frac{1}{2} ig' B_\mu \phi. \tag{2.2.11}$$

Here, g and g' are the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ coupling strengths. The field ϕ is a complex scalar doublet with $Y = \frac{1}{2}$, the Higgs multiplet, on which the potential

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi\right)^2 \tag{2.2.12}$$

acts. By virtue of this potential, ϕ acquires a vacuum expectation value

$$v = \sqrt{-\frac{\mu^2}{\lambda}}, \tag{2.2.13}$$

breaking the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry. Expanded around v, the Higgs doublet reads

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{2.2.14}$$

in unitary gauge. With this,

$$|\mathcal{D}_\mu \phi| = g^2 \frac{v^2}{8} \left[\left(W^1_\mu\right)^2 + \left(W^2_\mu\right)^2 + \left(\frac{g'}{g} B_\mu - W^3_\mu \right)^2 \right] \tag{2.2.15}$$

yields the mass terms of the gauge bosons. They can be diagonalized by the transformation

$$Z_\mu = W^3_\mu \cos\theta_w - B_\mu \sin\theta_w \tag{2.2.16}$$

$$A_\mu = W^3_\mu \sin\theta_w + B_\mu \cos\theta_w \tag{2.2.17}$$

$$W^\pm_\mu = \frac{1}{\sqrt{2}} \left(W^1_\mu \mp W^2_\mu\right) \tag{2.2.18}$$

with

$$\tan\theta_w = \frac{g'}{g}. \tag{2.2.19}$$

The masses of the physical gauge bosons can then be read off as

$$m_W = \frac{vg}{2} \tag{2.2.20}$$

$$m_Z = \frac{vg}{2\cos\theta_w} = \frac{m_W}{\cos\theta_w} \tag{2.2.21}$$

$$m_A = 0. \tag{2.2.22}$$

Notably, as the masses of the W and Z bosons are generated by the Higgs mechanism, their longitudinal polarizations absorb three of the four degrees of freedom contained in the complex scalar doublet ϕ, leaving exactly one degree of freedom for the physical Higgs field H. Its mass can be read off as

$$m_H = 2\lambda v^2 = -2\mu^2 \tag{2.2.23}$$

by plugging Eq. 2.2.14 into Eq. 2.2.8.

While fermion mass terms are not an intrinsic part of the Higgs mechanism, they can be generated by postulating Yukawa couplings of the fermions to the Higgs doublet. For a left handed fermion doublet L and the corresponding right handed singlet R, the masses for the lower component ψ_L of the doublet can then be generated by extending the Lagrangian by

$$\mathscr{L} = -g\left(\bar{L}\phi R + \bar{R}\bar{\phi}L\right). \tag{2.2.24}$$

Substituting the Higgs field after symmetry breaking from Eq. 2.2.14 and choosing the coupling g such that $vg = m_\psi\sqrt{2}$ then yields

$$\mathscr{L} = -m_\psi\bar{\psi}\psi. \tag{2.2.25}$$

The mass terms for the upper component can be generated similarly, only using $\phi_c = -i\sigma_2\phi^*$ instead of ϕ in Eq. 2.2.24.

It should be noted that while for charged leptons, the eigenstates of the weak interaction coincide with their Yukawa interaction (or *mass*) eigenstates, this is not the case for quarks or neutrinos. For quarks, the rotation between the eigenstates is performed with the Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix}. \tag{2.2.26}$$

If the Brout-Englert-Higgs Mechanism is also responsible for the neutrino masses, the same mapping can in this case be performed by the Pontecorvo-Maki-Nakagawa-

Sakata (PMNS) matrix

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{bmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} . \tag{2.2.27}$$

None of these parameters can be predicted from theory, but assuming unitarity, the total number of degrees of freedom in each of these matrices drops to four independent parameters. If neutrinos are Majorana particles, the physical *Majorana-phases* appear additionally.

2.2.3 Parameter measurements of the Standard Model

Given the above and assuming unitarity of both the CKM and the PMNS matrix, the Standard Model of Particle Physics relies on 26 independent parameters:

- the 12 fermion masses,
- three mixing angles and one CP-violating phase in the CKM matrix,
- three mixing angles and one CP-violating phase in the PMNS matrix, with possibly two additional Majorana phases,
- a possible CP-violating phase in QCD,
- three gauge couplings (or two gauge couplings and one mixing angle),
- the Higgs field vacuum expectation value (or the W boson mass), and
- the Higgs boson mass.

The charged fermion masses have been known to reasonable precision since the 1990s. The study of the CKM matrix is the main focus of the field of flavour physics, whereas the entries in the PMNS matrix and neutrino masses are measured by dedicated neutrino experiments. There is no evidence of CP violation in QCD.

Following the electroweak precision measurements of the 1990s and 2000s, the three electroweak parameters

$$\sin^2 \theta_w = 0.23155(5) \tag{2.2.28}$$

$$m_W = 80.385(15)\,\text{GeV} \tag{2.2.29}$$

$$\alpha_{\text{EM}}^{-1}\left(Q^2 = 0\right) = 4\pi g^{-2} \sin^{-2} \theta_w = 137.035999074(44) \tag{2.2.30}$$

have been measured to high accuracy, the values given here including the latest measurements [31]. This left only the mass of the Higgs boson as a free parameter to be determined. The latest measurements following its long-anticipated discovery in 2012 [35] find

$$m_H = 125.09 \pm 0.21\,(\text{stat.}) \pm 0.11\,(\text{syst.})\ \text{GeV}. \tag{2.2.31}$$

With this discovery, the last puzzle piece of the Standard Model has fallen into place, leaving open what might lie beyond.

2.3 Effective Field Theories

An Effective Field Theory (EFT) is an approach to describe the physics in some regime as the approximation or limit obtained by neglecting (or integrating out) effects beyond that regime. The boundary of the regime of validity of the EFT in question is typically denoted in units of energy as a cutoff scale Λ. There are two characteristic approaches to EFT:

- In the **top-down** approach, the full theory is known *a priori*. By integrating out high energy effects of the theory, a simplified, effective theory can be obtained to ease calculations and help build up intuition for the physics in the low-energy regime below the cutoff scale Λ.
- In the **bottom-up** approach, the full theory is not known. Instead, it is assumed that the currently accepted theory is in fact (part of) an effective theory. Attempts to measure and describe deviations from that theory then allow to gradually infer and obtain information about the full theory.

These two approaches are complementary in the sense that an EFT description originally used bottom-up for discovery of new physics beyond the currently accepted model can *a posteriori* serve in a top-down fashion as an educational tool. One prominent example for this pattern of events is the Fermi theory of weak interactions [36].

2.3.1 Fermi theory

The β-decay $n \to p^+ e^- \bar{\nu}_e$ was discovered at the dawn of the 20th century, but the neutrino and its role in this process were only revealed in the 1930s. Enrico Fermi succeeded at describing this process as a four-fermion local interaction as shown in Fig. 2.3a, long before the discovery of the W and Z bosons introduced in Section 2.2.2 occurred in the 1980s.

Fermi's model can be described with the Lagrangian

$$\mathscr{L} = \frac{G_F}{\sqrt{2}} J^\mu J^\dagger_\mu, \tag{2.3.1}$$

where the current J^μ can be written in modern notation as

$$J^\mu = \bar{u}\gamma^\mu \left(1 - \gamma^5\right) \left(d \cos\theta_c + s \sin\theta_c\right) + \bar{\nu}_e \gamma^\mu \left(1 - \gamma^5\right) e \tag{2.3.2}$$

to exhibit the $V-A$ form characteristic to weak processes, with the term proportional only to γ^μ corresponding to the vector current and the one proportional to $\gamma^\mu\gamma^5$ corresponding to the axial vector current. Here, θ_c is the Cabibbo-angle, and G_F is the Fermi constant. Notably, it is not dimensionless as coupling constants of renormalizable QFTs ought to be, but instead

$$G_F = \frac{\sqrt{2}g^2}{8M_W^2} \approx 1.16637 \cdot 10^{-5}\,\mathrm{GeV}^{-2}. \tag{2.3.3}$$

Only later it became apparent that Fermi's four-point interaction vertex is indeed a contraction of a W-boson mediating this weak decay, shown in Fig. 2.3b.

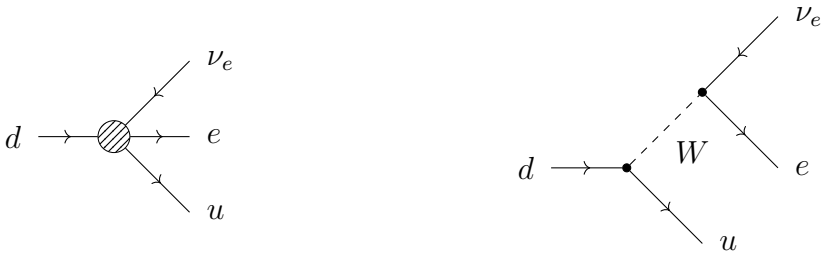
In order to understand, from a top-down perspective, why Fermi's theory was able to accurately model the β decay, it is important to note that the involved W-boson is deeply virtual, the momentum transfer mediated in this interaction being much smaller than its mass. Therefore, the vector-boson propagator reduces to a contact interaction

$$\frac{-g_{\mu\nu} + \frac{q_\mu q_\nu}{M_W^2}}{q^2 - M_W^2} \xrightarrow{\frac{q_\mu q_\nu}{M_W^2}\to 0} \frac{g_{\mu\nu}}{M_W^2}. \tag{2.3.4}$$

This of course yields the explanation for why Eq. 2.3.3 contains the W-bosons mass, which is responsible for the dimensionality of G_F and plays the role of the cutoff-scale Λ for this effective field theory, which remains valid as long as the premise of the W-boson being deeply virtual holds true.

Fermi theory provides an example for several properties of effective field theories:

1. EFT can help provide a framework for new physics discoveries and facilitate studies of effects beyond the energy regime currently accessible to observation.
2. EFT descriptions work best when physics regimes are separated by several orders of magnitude in energy.



(a) Four-fermion interaction proposed by Enrico Fermi 1933.

(b) Standard Model tree level diagram for β decay.

Figure 2.3.: Feynman diagrams for β decay in Fermi theory and in the Standard Model.

3. The appearance of an energy scale in coupling constants is a tell-tale sign for a hidden energy scale like the mass of some heavy, previously undiscovered particle.

Notably, the non-renormalizability of an EFT is no conceptual problem, as the explicitly limited validity range allows for the theory to be embedded in a larger, renormalizable theory that is not required to be known at the inception of the EFT.

2.3.2 Effective Lagrangians

Searches for physics beyond the Standard Model commonly assume that new physics effects appear in the form of some previously undiscovered particle. Not being stable, these will first appear as internal lines of Feynman diagrams for observable processes. These can then be contracted and treated as effective vertices, as seen in the previous section.

In general, any effective Lagrangian extending the Standard Model will have the form

$$\mathscr{L} = \mathscr{L}_{\text{SM}} + \sum_i g_i \mathcal{O}_i, \tag{2.3.5}$$

where the $\mathcal{O}_i$ are operators constructed from combinations of SM fields, and the g_i are the corresponding coupling constants, commonly referred to as *Wilson coefficients*.

It is useful and customary to investigate and classify these effective operators with dimensional analysis, counting powers of electron volts. Establishing

$$[\mathscr{L}] = 4 \quad \text{and} \quad [\mathcal{O}_i] = d_i \tag{2.3.6}$$

in natural units, it follows immediately that

$$[g_i] = 4 - d_i \tag{2.3.7}$$

and thus

$$g_i \propto \Lambda^{4-d_i}, \tag{2.3.8}$$

as Λ is the characteristic energy scale of the EFT. As long as this scale is much larger than the momentum transfer $q^2 \ll \Lambda$, contributions by these effective operators are suppressed by powers of Λ. Using this power counting technique, operators are classified in one of three categories:

- Operators $\mathcal{O}_i$ for which $d_i < 4$ are called **relevant**. These operators are not suppressed by the scale Λ and are an essential part of the effective Lagrangian. They might receive large quantum corrections at low energies.

- Operators $\mathcal{O}_i$ for which $d_i = 4$ are called **marginal**. They contribute equally at all energy scales, but their scaling behavior might be modified by quantum effects.
- Operators $\mathcal{O}_i$ for which $d_i > 4$ are called **irrelevant**. They are suppressed by powers of Λ and can be neglected within the validity range of the EFT, but play an increasingly important role near the boundary. Theories with *irrelevant* operators are non-renormalizable.

In the framework of EFT, the search for physics beyond the Standard Model can be seen as the attempt to measure the Wilson coefficients of not previously observed and not necessarily renormalizable effective interactions between known particles that are *irrelevant* within the validity range of the EFT that is the Standard Model.

2.3.3 Extensions of the Standard Model

Adopting the terminology from the previous section, any imaginable operator can be categorized by its *dimensionality* d. Operators of different dimensionality generally behave very differently. This section briefly discusses classes of possible operators extending the Standard Model and their properties.

Even restricting the imagination to the fields contained in the Standard Model, the number of conceivable operators is very large, but many can be ruled out due to symmetry considerations already. It is customary and useful to sort the remaining ones by dimensionality.

For the power counting, the dimensionality of the individual components of the operators needs to be established:

$$[\psi] = \frac{3}{2}, \quad [V_\mu] = 1, \quad [V_{\mu\nu}] = 2, \quad [\phi] = 1, \quad [\partial_\mu] = 1, \quad [\mathcal{D}_\mu] = 1. \tag{2.3.9}$$

Here, ψ denotes any fermion; V_μ the gauge bosons G^i_μ, W^i_μ and B_μ; $V_{\mu\nu}$ the corresponding field strength tensors; and $\mathcal{D}_\mu$ a covariant derivative combining a partial derivative with couplings to gauge bosons.

There is only one conceivable $d = 5$ operator, the *Weinberg operator* [37]

$$\mathcal{O}_{\mathrm{W}\alpha\beta} = \frac{1}{2}\left(\bar{\ell}^C_{L\alpha}\widetilde{\phi}^*\right)\left(\widetilde{\phi}^\dagger \ell_{L\beta}\right) + \mathrm{h.\,c.}, \tag{2.3.10}$$

where the superscript C denotes a charge conjugated field, and where the shorthand notation

$$\widetilde{\phi} = i\sigma_2\phi^* \tag{2.3.11}$$

is used. The symbol ℓ_L here denotes left-handed leptons with Greek letters denoting flavour indices, and ϕ is the Higgs doublet. This operator requires right-handed neutrinos to be added to the Standard Model and generates a Majorana mass term for neutrinos. However, its effects are likely unobservable at the LHC and will not be discussed further here. In general, any operator with odd dimension will violate conservation of $B-L$, the difference between baryon and lepton number, and will thus be neglected henceforth.

A first complete set of 80 operators with $d = 6$ was published in 1985 [38]. It was soon discovered that not all of these are independent: they can be linked through equations of motion of the SM fields, integration by parts, or field redefinitions. A complete, minimal basis consists of 59 operators, where several conventions exist: The *Warsaw* basis [39], the SILH convention [40], and the HISZ basis [41]. The total number of independent coefficients, taking into account all flavour combinations, is 2499 [19].

Beyond $d = 6$, the number of conceivable operators increases quickly to several thousand at $d = 8$ and tens of thousands at $d = 10$. So far, no comprehensive experimental studies including all of these operators exist.

2.4 Predictions for Measurements

After the groundwork of the Standard Model has been established in the preceding sections, the remainder of this chapter gives a short introduction into how this theoretical framework can be used to derive predictions that can be tested experimentally, specifically using particle detectors such as ATLAS at hadron colliders like the LHC.

2.4.1 Cross sections and matrix elements

When looking at some quantum mechanical transition $i \to f$, the transition probability $\mathcal{P}_{i\to f}$ is generally defined via the *scattering matrix* $\mathcal{S}$ as

$$d\mathcal{P}_{i\to f} = \frac{|\langle f \mid \mathcal{S} \mid i\rangle|^2}{\langle i \mid i\rangle \langle f \mid f\rangle} \tag{2.4.1}$$

The *scattering amplitude* or *matrix element* $\mathcal{M}_{i\to f}$ is commonly defined via

$$\mathcal{S}_{i\to f} = \delta_{if} - i\,(2\pi)^4\,\delta^{(4)}\,(p_f - p_i)\,\mathcal{M}_{i\to f}, \tag{2.4.2}$$

where p are the four-momenta and energies of the particles in the process. A cross section for this transition is then given as

$$\sigma_{i\to f} = \int \mathcal{P}\, d\Omega \propto \int d\Omega\, |\mathcal{M}|^2\,. \tag{2.4.3}$$

The precise derivation of the matrix element and thus the cross section for some specific process can be performed by virtue of the Feynman rules, which depend on the Lagrangian of the underlying theory.

After writing down some Lagrangian, the derivation of the Feynman rules of this theory can be achieved using standard prescriptions as can be found in the literature [42, 43], or by using software solutions like FEYNRULES [44]. These can then be used to calculate the matrix elements $\mathcal{M}$ and thus the (total or differential) cross sections σ for some physics processes in that theory.

Assuming a Lagrangian with the structure

$$\mathscr{L} = \sum_i g_i \mathcal{O}_i, \tag{2.4.4}$$

a generic amplitude will at tree level take the form

$$\mathcal{M} = \prod_{\mathcal{V}} \sum_{j \in \mathcal{V}} g_j \mathcal{M}_j, \tag{2.4.5}$$

where the product runs over all vertices $\mathcal{V}$ in the process, and where each partial matrix element $\mathcal{M}_j$ corresponds to exactly one $\mathcal{O}_i$ contributing to that vertex. Using Eq. 2.4.3 and expanding the product of sums into a sum of products, one can write

$$\sigma_{i \to f} = c \cdot \int d\Omega \sum_i \mathfrak{R} \left(\prod_{k(i)} g_k \mathcal{M}_k \right), \tag{2.4.6}$$

where c is some constant encoding the combinatorics of the process, and the $\mathcal{M}_k$ incorporate the partial matrix elements introduced in Eq. 2.4.5 and their complex conjugates. As the phase space integration is far from trivial, it is typically performed using Monte Carlo integrators that work by sampling the phase space with a finite set of randomly generated events.

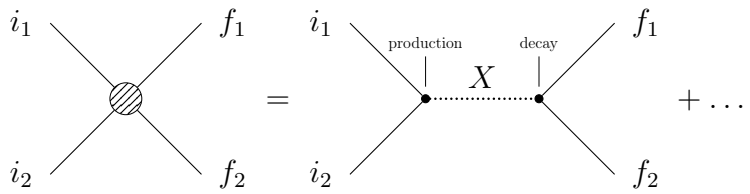


Figure 2.4.: Tree-level Feynman diagram for a general $2 \to 2$ process.

Considering a process of the type $2 \to 2$, as shown in Fig. 2.4, the tree-level cross section generally has the form

$$\sigma = c \cdot \int d\Omega_p d\Omega_d \quad \mathcal{M}_p\left(q^2\right) \mathcal{M}_d\left(q^2\right) D_X\left(q^2\right), \tag{2.4.7}$$

where $\mathcal{M}_p$ and $\mathcal{M}_d$ are the amplitudes for the production and the decay vertex, and where $D_X(q^2)$ is the propagator of the internal particle X. The propagator is generally of the form

$$D_X\left(q^2\right) = \frac{1}{\left(q^2 - m_X^2\right)^2 - \Gamma_X^2 m_X^2}, \tag{2.4.8}$$

where m_X and Γ_X are the mass and width of the particle X. In the *narrow* or *zero width approximation* (NWA) of $\Gamma_X \to 0$, this Breit-Wigner resonance becomes a delta-function, such that

$$D_X^{\text{NWA}}\left(q^2\right) = \frac{\pi}{\left(q^2 - m_X^2\right)} \delta\left(q^2 - m_X^2\right). \tag{2.4.9}$$

In this limit, the cross section factorizes into a production and a decay part, which can be computed separately

$$\sigma^{\text{NWA}} = c \cdot \int d\Omega_p \mathcal{M}_p\left(m_X^2\right) \cdot \int \frac{dq^2}{2\pi} D_X\left(q^2\right) \cdot \int d\Omega_d \mathcal{M}_d\left(m_X^2\right) \tag{2.4.10}$$

$$= c' \cdot \int d\Omega_p \mathcal{M}_p\left(m_X^2\right) \cdot \int d\Omega_d \mathcal{M}_d\left(m_X^2\right). \tag{2.4.11}$$

However, before predictions from such calculations can be compared to real measurements, additional effects need to be taken into account. An illustration of this chain of effects for a simple $2 \to 2$ s-channel process at leading order with two baryons in the initial state and hadrons in the final state is shown in Fig. 2.5. The individual stages are further explained in the remainder of this section.

Parton density functions

If the initial state contains bound states of QCD, a usual approximation is that for each hadron, only one of the *partons* takes part in the *hard scattering* for which the matrix element is being computed, whereas the others act as *spectators*. It is customary to make the assumption that the probability to interact with some parton of type p carrying a fraction x of the longitudinal momentum of some hadron h is only dependant on the momentum scale q^2 of the interaction.

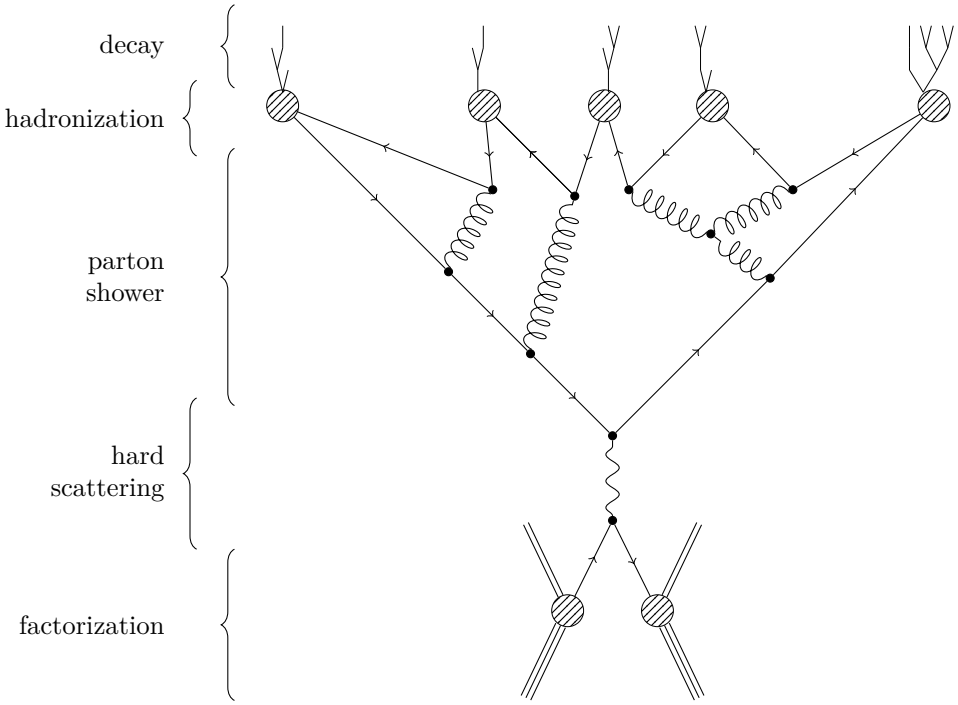


Figure 2.5.: Illustration showing the effects complicating the calculation of a simple $2 \to 2$, s-channel process with two baryons in the initial state (bottom) and hadrons in the final state (top).

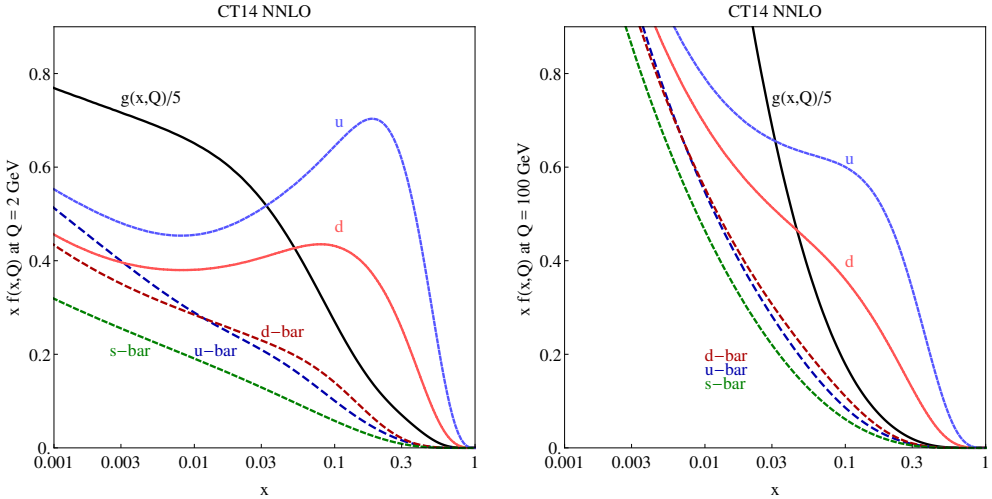


Figure 2.6.: The CT14 parton distribution functions for the proton at $Q = 2\,\text{GeV}$ (left) and $Q = 100\,\text{GeV}$ (right) for u, $\bar{u}$, d, $\bar{d}$, s and $\bar{s}$ quarks as well as gluons.

Under this assumption, the physical cross section factorizes into the parton-level cross section $\widehat{\sigma}$ and the parton densitiy function (PDF) $p_h(x, q^2)$, evaluated at the *factorization scale* q^2.

Using this approach, the cross section σ of two colliding hadrons h_1 and h_2 can be obtained from the parton-level cross section $\widehat{\sigma}$ involving a parton a from h_1 and a parton b from h_2 via

$$\sigma = \sum_{a,b} \int dx_1\, dx_2\, a_{h1}\left(x_1, q^2\right)\, b_{h2}\left(x_2, q^2\right)\, \widehat{\sigma},$$

where the sum runs over the contributing parton types a and b.

These PDFs cannot currently be predicted from first principles with reasonable precision, but are instead obtained for common hadron types from a combined fit to a multitude of individual experimental measurements. The most recent results of the proton PDFs computed by the CTEQ collaboration are shown in Fig. 2.6 as an example.

Underlying event

The parton density functions are instrumental in computing the physical cross section of some parton-level process. However, the remaining spectator partons are generally still present in the final state. This contribution is usually referred to as the *underlying event*, and is generally estimated using Monte Carlo generators.

Parton shower

Highly energetic final state particles may take part in additional hard interactions. This process involves the appearance of more particles in the course of a *parton shower*. This parton shower is usually modeled with a Monte Carlo generator on an event-by-event basis, often in conjunction with the underlying event.

Hadronization

Because of the confining nature of QCD, final states involving particles carrying colour charge are further subject to the formation of hadrons. This process is called *hadronization* and falls into the "soft" regime of QCD that is inaccessible to perturbative calculations due to the large values of the strong coupling constant for small momenta. Thus, it is modeled with fragmentation functions. Like the parton density functions, they cannot currently be calculated from first principles, but instead need to be obtained from experiment.

Decays

The particles procuded in high-energy collisions are often unstable. While some of them (like muons) have sufficiently long mean life times to be generally regarded as stable for the purposes of their detection, others (like b-hadrons) will very likely decay before they can be detected. These decays are often modeled by dedicated pieces of software.

Detector simulation

Any real-world detector will always suffer from detection inefficiencies. The behavior of the detector response to the Monte Carlo generated events needs to be modeled by dedicated detector simulation software, factoring in particle measurement and identification inefficiencies resulting from the detector geometry and material. The resulting simulation data can then be passed through the same reconstruction software that also the real experimental data is subject to in order to provide comparable results.

Luminosity

The cross section σ and the number of expected events N arising from a given process in some experiment are connected by

$$N = \sigma\mathcal{L} = \sigma \int L dt,$$

where $\mathcal{L}$ is the integrated luminosity collected by the experiment. For a collider experiment with two colliding beams of particle bunches, the instantaneous luminosity L can be calculated from first principles using

$$L = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y},$$

where n_1 and n_2 are the respective numbers of particles in the colliding bunches, σ_x and σ_y denote the root mean square deviations of the transverse beam size in horizontal and vertical direction, and f is the frequency at which bunches collide at the interaction point. In practice, however, the luminosity is often measured by the detector apparatus itself.

2.4.2 Frameworks of interpretation

Within the Standard Model, with all parameters assumed to be known, all observable quantities can in principle be predicted from theory calculations. When searching for effects beyond the Standard Model, it is thus required to find a parametrization that introduces new degrees of freedom within which deviations from Standard Model predictions can be interpreted.

While it is always possible to investigate and test specific theory models for their compatibility with experimental data, the extreme variety of such models quickly renders attempts for a comprehensive falsification campaign of all but the best motivated of these models unfeasible.

One rather simplistic case is the search for a new particle, where it is common to parametrize its presence or absence by virtue of a *signal strength* parameter

$$\mu = \frac{\sigma_{\text{obs.}}}{\sigma_{\text{theo.}}} \tag{2.4.12}$$

in terms of the observed cross section $\sigma_{\text{obs.}}$ with respect to the theory prediction $\sigma_{\text{theo.}}$ for some specific measurement channel.

The signal strength contains a convolution of information from production and decay of the particle searched for. Hence, it is not easily generalizable over different search channels. Consequently, it is important to employ sensible alternatives once a discovery has been established, as has happened for the Higgs boson.

This section presents two different model-independent approaches of interpreting experimental evidence from measurements of the Higgs sector in the search for effects beyond the Standard Model.

The κ framework

Using the narrow width approximation introduced in Section 2.4.1, the production cross section σ_i and partial decay width Γ_f of an individual process $i \to H \to f$ factorize completely at leading order. The partial width Γ_f is hereby defined as the product of total width Γ and the branching fraction $\mathcal{B}_f$.

In order to model possible deviations from Standard Model predictions of these quantities, one can introduce *coupling modifiers* κ [45, 46], such that

$$\kappa_i^2 = \frac{\sigma_j}{\sigma_j^{\text{SM}}} \quad \text{or} \quad \kappa_j^2 = \frac{\Gamma_j}{\Gamma_j^{\text{SM}}}. \tag{2.4.13}$$

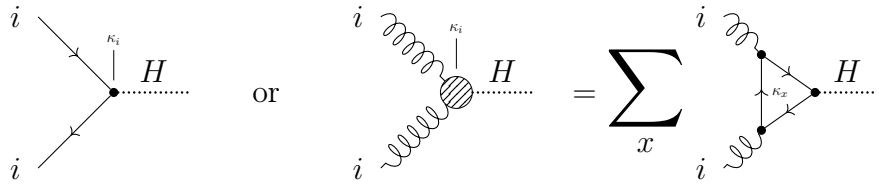


Figure 2.7.: Feynman diagrams for Higgs boson production via a Standard Model coupling (left) or an effective coupling (right).

It is also possible to define a coupling modifier for an effective vertex containing a loop of virtual particles. Feynman diagrams exemplifying this are shown in Fig. 2.7.

Additionally, one defines

$$\kappa_H^2 = \sum_j \mathcal{B}_j \kappa_j^2 \tag{2.4.14}$$

to model variations in the total Higgs boson decay width.

While this interpretation framework has proven useful in the past [15], its applicability for future precision measurements is limited. One of its main limitations is that it can only model changes in total rate (or cross section) and has no handle over possible shape differences in differential cross sections.

Higgs effective field theory

Using the mechanisms introduced in Section 2.3.3, one can spell out an effective Lagrangian containing any conceivable interaction in and beyond the Standard Model up to some order. As several equivalent parametrizations (or bases) exist up to $d = 6$, it remains a matter of taste which one of these may be used. However, as there is no Monte Carlo event generator currently available that includes a complete $d = 6$ basis, the most usable alternative is the *Higgs Characterization Model* [47]. The effective Lagrangian of this model reads

$$\begin{aligned}
\mathscr{L} = \Bigg\{ & \cos\alpha\kappa_{\text{SM}}\big[\frac{1}{2} g_{HZZ}\, Z_\mu Z^\mu + g_{HWW}\, W_\mu^+ W^{-\mu}\big] \\
& - \frac{1}{4}\big[\cos\alpha\kappa_{H\gamma\gamma} g_{H\gamma\gamma}\, A_{\mu\nu}A^{\mu\nu} + \sin\alpha\kappa_{A\gamma\gamma} g_{A\gamma\gamma}\, A_{\mu\nu}\widetilde{A}^{\mu\nu}\big] \\
& - \frac{1}{2}\big[\cos\alpha\kappa_{HZ\gamma} g_{HZ\gamma}\, Z_{\mu\nu}A^{\mu\nu} + \sin\alpha\kappa_{AZ\gamma} g_{AZ\gamma}\, Z_{\mu\nu}\widetilde{A}^{\mu\nu}\big] \\
& - \frac{1}{4}\big[\cos\alpha\kappa_{Hgg} g_{Hgg}\, G^a_{\mu\nu}G^{a,\mu\nu} + \sin\alpha\kappa_{Agg} g_{Agg}\, G^a_{\mu\nu}\widetilde{G}^{a,\mu\nu}\big] \\
& - \frac{1}{4}\frac{1}{\Lambda}\big[\cos\alpha\kappa_{HZZ}\, Z_{\mu\nu}Z^{\mu\nu} + \sin\alpha\kappa_{AZZ}\, Z_{\mu\nu}\widetilde{Z}^{\mu\nu}\big] \\
& - \frac{1}{2}\frac{1}{\Lambda}\big[\cos\alpha\kappa_{HWW}\, W^+_{\mu\nu}W^{-\mu\nu} + \sin\alpha\kappa_{AWW}\, W^+_{\mu\nu}\widetilde{W}^{-\mu\nu}\big] \\
& - \frac{1}{\Lambda}\cos\alpha\big[\kappa_{H\partial\gamma}\, Z_\nu\partial_\mu A^{\mu\nu} + \kappa_{H\partial Z}\, Z_\nu\partial_\mu Z^{\mu\nu} + \big(\kappa_{H\partial W}\, W^+_\nu\partial_\mu W^{-\mu\nu} + \text{h. c.}\big)\big]\Bigg\} X_0\,,
\end{aligned} \tag{2.4.15}$$

where the (reduced) field strength tensors are defined as

$$V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \quad (V = A, Z, W^\pm)\,, \tag{2.4.16}$$

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu\,, \tag{2.4.17}$$

and the dual tensor is

$$\widetilde{V}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}V^{\rho\sigma}\,. \tag{2.4.18}$$

The framework of EFT provides powerful handles over observable quantities like differential cross sections, embodied by the Wilson coefficients, which are (somewhat unfortunately) named κ in this model. However, as these parameters directly control the fundamental interactions, their effects on the final observables are in principle non-trivial. The usage of this framework for interpretations of measurements would rely on a vast number of sampled points in configuration space to be used for hypothesis testing, unless an efficient mechanism to accurately interpolate between different EFT scenarios was available.

2.4.3 Effective Lagrangian Morphing

Starting from Eq. 2.4.6, one can write

$$\sigma(\vec{g}) = \int d\Omega \sum_i c_i P_i(\vec{g}) P_i(\vec{\mathcal{M}}) \tag{2.4.19}$$

$$= \sum_i c_i P_i(\vec{g}) \int d\Omega P_i(\vec{\mathcal{M}}), \tag{2.4.20}$$

where the P_i are polynomials of the coupling parameters $\vec{g} = (g_1, \ldots, g_n)$ of the form

$$P(\vec{x}) = \mathfrak{R}\left(\prod_\alpha x_\alpha\right). \tag{2.4.21}$$

Notably, the

$$\chi_i = \int d\Omega P_i(\vec{\mathcal{M}}) \tag{2.4.22}$$

form a basis of a vector space containing σ. If one was to pick a new, different basis ξ_i, one would transform between the two with

$$\chi_i = \sum_j b_{ij}\xi_j \tag{2.4.23}$$

for some set of coefficients b_{ij}. With this transformation,

$$\sigma(\vec{g}) = \sum_i c_i P_i(\vec{g}) \chi_i \tag{2.4.24}$$

$$= \sum_i c_i P_i(\vec{g}) \sum_j b_{ij} \xi_j \tag{2.4.25}$$

$$= \sum_{i,j} P_i(\vec{g}) A_{ij} \xi_j \tag{2.4.26}$$

for appropriately chosen coefficients A_{ij}. If the ξ_j form a basis that can be expressed as a physical cross section $\sigma(\vec{g}_j)$, one can calculate the coefficients A_{ij} from the closure condition

$$\xi_k = \sigma(\vec{g}_k) = \sum_{i,j} P_i(\vec{g}_k) A_{ij} \xi_j. \tag{2.4.27}$$

This is a simple system of linear equations. It can be rewritten as

$$\delta_{jk} = \sum_i P_i(\vec{g}_k) A_{ij}. \tag{2.4.28}$$

Alternatively, in matrix notation,

$$A \cdot M = \mathbb{1} \qquad \text{with} \qquad M_{ik} = P_i(\vec{g}_k). \tag{2.4.29}$$

Equation 2.4.27 has a solution precisely if the matrix M is invertible. In analogy to linear algebra terminology, a set of parameter points $\vec{g}_k$ fulfilling this condition is called a *morphing basis*.

With this result, it is now easily possible to interpolate between any set of precomputed cross sections ξ_k at parameters $\vec{g}_k$ in such a way as to obtain a prediction σ at any other parameter point $\vec{g}$.

2.5 Statistical Data Analysis

An attempt to measure properties of a newly discovered particle, such as the couplings or the parity of the newly discovered Higgs boson, typically proceeds by means of *hypothesis testing*. Two competing hypotheses are defined:

- The **null hypothesis**, denoted H_0, is typically the hypothesis that the Standard Model prediction is true and any observed deviation is a result of pure chance.

- The **alternative hypothesis**, denoted H_1, is the hypothesis that some non-random cause produces the observed deviations.

A test statistic t is a quantity that measures the agreement between the observed data and the prediction obtained under either hypothesis, and for which the probability distribution under either hypothesis can be computed. The choice of definition of the test statistic is in principle arbitrary, but several conventional choices exist, which are discussed in the following paragraphs. The p-value under either of the two hypotheses is defined as the probability to measure deviations from the predictions at least as large as the deviations that were observed. Here, two different scenarios are of particular interest.

- If the p-value under the null hypothesis is found to be smaller than a predefined threshold of confidence, one can discard the null hypothesis. This is considered a *discovery*. A commonly used threshold in particle physics is $p < 2.87 \cdot 10^{-7}$. For a Gaussian probability distribution, this corresponds to a deviation of 5σ with respect to the expectation value.
- If the p-value under the alternative hypothesis is found to be smaller than a predefined threshold of confidence, the alternative hypothesis needs to be discarded. This is considered an *exclusion*. A commonly used threshold is $p < 0.05$. For a Gaussian probability distribution, this corresponds to a deviation of about 2σ with respect to the expectation value.

For non-parametric hypotheses, this strategy is very straight-forward. If, however, the hypotheses are parametric, the strategy becomes more involved. A typical procedure for such cases is to perform a *likelihood fit*. The likelihood

$$\mathcal{L}(\mu, \theta) = P(x; \mu, \theta) \tag{2.5.1}$$

is a function yielding the probability P to observe the given data set x given the parameters μ and θ, where μ is taken to be a set of *parameters of interest*, and θ is taken to be a set of *nuisance parameters*. The distinction between the two is purely conceptual:

- A *parameter of interest* is typically a free parameter of the theory comprising the alternative hypothesis, that is either not present or fixed to some specific value under the null hypothesis. The signal strength, coupling modifiers, or Wilson coefficients introduced in Section 2.4.2 are pristine examples of this type of parameter.
- A *nuisance parameter* is a parameter that is typically present under the alternative and the null hypothesis. Often times, prior knowledge is available to constrain these parameters within certain bounds, but also unconstrained nuisance parameters are common. Typical examples would be parameters

of the experimental setup, or the normalization of a background prediction with respect to the observed data set.

It is important to note that $\mathcal{L}$ is itself not a probability, as it is not normalized to unity and its arguments are not statistical samples. However, it is a *likelihood* in the sense that large values of $\mathcal{L}$ correspond to scenarios under which the observation of the given data set is more likely.

The **likelihood ratio** $\Lambda(\mu)$ is then defined as

$$\Lambda(\mu) = \frac{\mathcal{L}\Big(\mu, \widehat{\widehat{\theta}}(\mu)\Big)}{\mathcal{L}\big(\widehat{\mu}, \widehat{\theta}\big)}, \tag{2.5.2}$$

where $\widehat{\mu}$ and $\widehat{\theta}$ denote the unconditional maximum likelihood estimators, that is, those values of μ and θ that maximize the likelihood in absolute terms. On the other hand, $\widehat{\widehat{\theta}}(\mu)$ denotes the maximum likelihood estimator of θ conditional on μ, that is, the function $\widehat{\widehat{\theta}}(\mu)$ yields the values of θ that maximize the likelihood for a given value of μ. It is customary to refer to the function $\mathcal{L}(\mu, \widehat{\widehat{\theta}}(\mu))$ as the *profile likelihood*, the practice of *profiling* generally referring to the replacement of parameters by their corresponding maximum likelihood estimators.

In most cases, the maximization cannot be performed analytically. For the technical reason of facilitating this type of computation in a finite precision calculation, it is common to use the *negative log likelihood ratio*

$$t_\mu = -\log \Lambda(\mu) \tag{2.5.3}$$

as a test statistic. Its probability distribution can either be evaluated over some randomly generated set of samples (or "toys") in the range expected for the data, or in certain cases by using some analytic approximation [48]. Asymptotically, this test statistic is distributed as a χ^2 distribution with n degrees of freedom, where n is number of parameters of interest.

2.6 Higgs Physics at the LHC

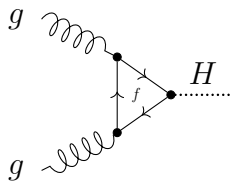
The Large Hadron Collider (LHC) is the largest and most powerful particle collider ever built, capable of colliding proton pairs with center-of-mass energies up to $\sqrt{s} = 13\,\text{TeV}$. It houses a number of large particle detectors, concentrating on specific aspects of the highly energetic proton collisions. Two of these experiments, ATLAS and CMS, are often referred to as general-purpose detectors, being capable of performing a wide range of physics measurements, with a special focus on the physics of the Higgs boson. The Large Hadron Collider and its experiments are presented in greater detail in Chapter 3.

2.6.1 Higgs boson production

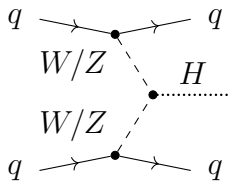
Higgs bosons can be produced via various mechanisms at the LHC. Figure 2.8 shows a collection of Feynman diagrams identifying the main modes of Higgs boson production at the LHC. The relative contributions of these mechanisms at different center-of-mass energies are displayed in Fig. 2.10. The displayed cross sections have been calculated to varying degrees of accuracy, as indicated by the labels. Further details on the individual modes of Higgs boson production are provided in the following.

Gluon fusion

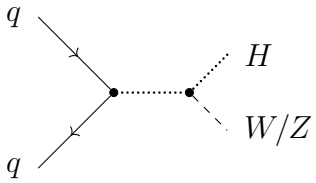
Gluon fusion (ggF) via a fermion loop is the leading mechanism for Higgs boson production at the LHC. The most important contributions come from top quark



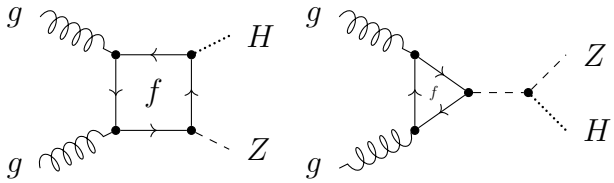
(a) Gluon fusion (ggF)



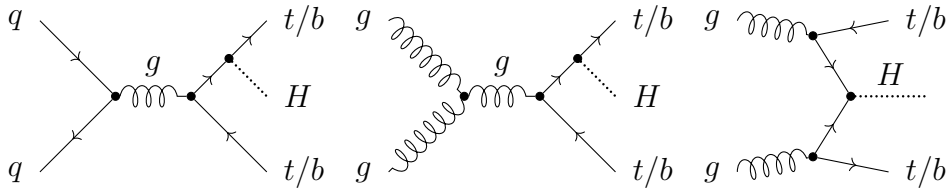
(b) Vector Boson fusion (VBF)



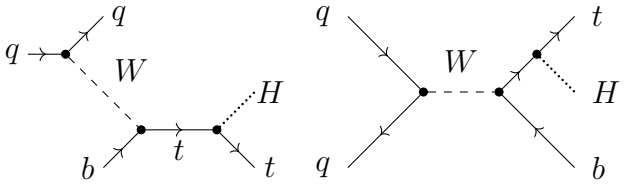
(c) Higgs strahlung (WH or ZH)



(d) Higgs strahlung (ZH only)



(e) Associated production (ttH/bbH)



(f) Associated production (tH) in s-channel (left) and t-channel (right)

Figure 2.8.: Example Feynman diagrams for Higgs boson production mechanisms.

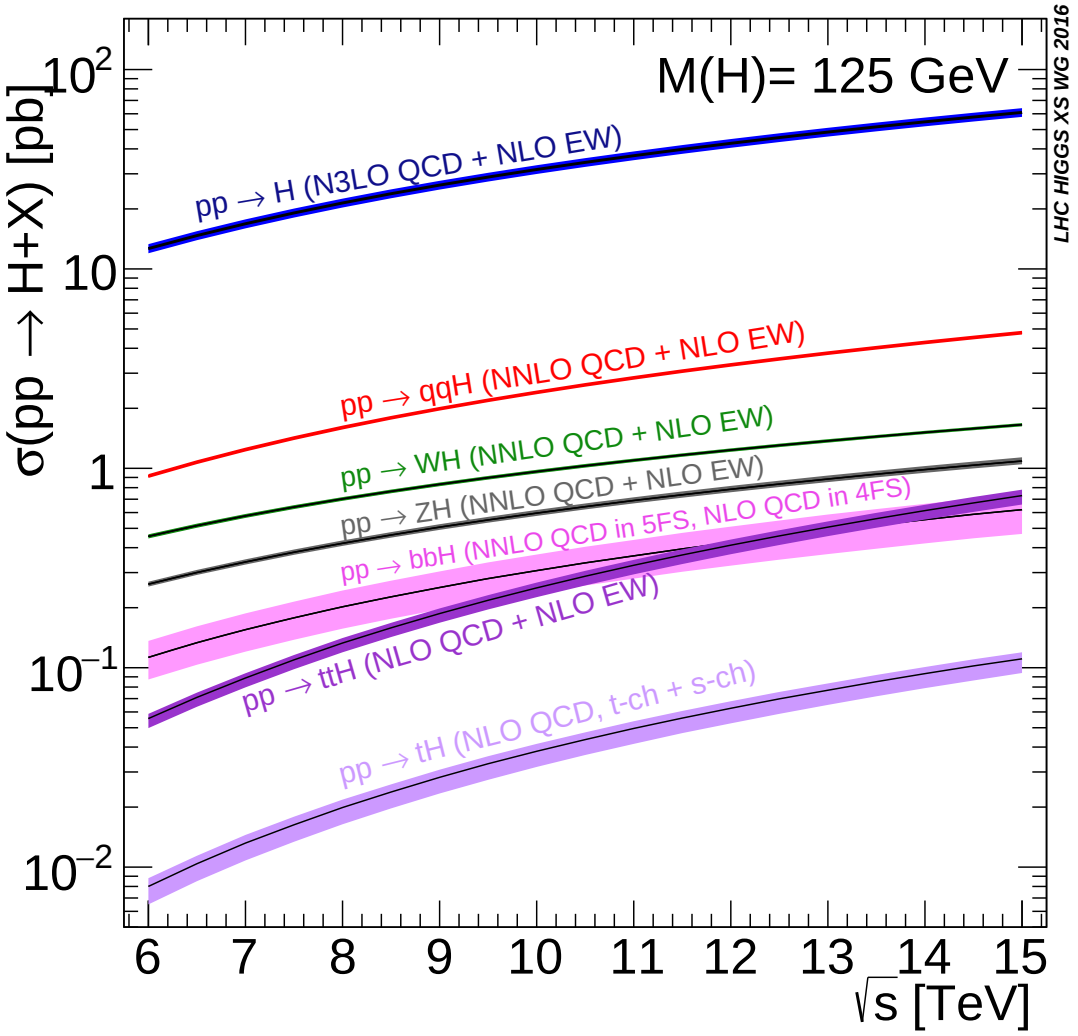


Figure 2.10.: Predicted Standard Model Higgs boson production cross sections for different production modes at the LHC. Line widths encode the respective theory uncertainties [49].

loops, whereas bottom quark loops and the negative top-bottom interference can contribute 5–10%.

For the $pp \to H$ gluon Fusion process, the convergence of the perturbative expansion in QCD is slow, with large corrections occurring at NLO and NNLO, comparable in size or even exceeding the uncertainties introduced by the parton density functions and QCD scale variations [19]. Figure 2.10 includes the recent results calculated to NNNLO in QCD [50] using an effective theory with the top-quark integrated out, as well as electroweak corrections calculated to NLO precision.

Vector boson fusion

Higgs boson production via fusion of two weak vector bosons (VBF) is the second most important contribution at the LHC. This production mechanism can be targeted specifically in the experiment by exploiting the two accompanying highly energetic quark jets in the final state qqH to select event candidates.

The VBF process $pp \to qqH$ has been calculated to NNLO in QCD, including NLO EW corrections. The calculation employs the approximation of double deep inelastic scattering attached to the colourless pure electroweak vector boson fusion into a Higgs boson [51].

Higgs strahlung

Higgs strahlung, or production of Higgs bosons in association with a W or Z vector boson (VH), is still in the process of being established experimentally as a mechanism for Higgs boson production at the LHC. It can be specifically targeted by searching for events which exhibit evidence for the presence of an additional vector boson in the final state.

The VH processes $pp \to WH$ and $pp \to ZH$ have also been calculated to NNLO in QCD with NLO EW corrections, where the calculation has been split into Drell-Yan induced, top-loop induced, photon-induced, and an additional $gg \to ZH$ component for the ZH mode which is calculated to NLO accuracy with NLL corrections added on top.

Quark associated production

Higgs boson production in association with top quarks (ttH or tH) or with bottom quarks (bbH) is significantly harder to measure than the above processes due to the comparably low production cross section and the less accessible final state. However, efforts are being made to establish this mechanism for Higgs boson production at the LHC nevertheless.

The ttH process has been calculated to NLO accuracy in QCD as well as to NLO accuracy in EW corrections. The calculations of the tH process only employ NLO QCD predictions, featuring s-channel and t-channel diagrams, as shown in Fig. 2.9f.

The bbH cross section prediction combines results from an NNLO QCD calculation in a five-flavour-scheme, as well as from an NLO QCD calculation in a four-flavour-scheme. In the four-flavour-scheme, bottom quarks do not appear in the initial state, but rather via gluon splitting into b-quark pairs.

Observations

The ggF and VBF production processes have already been established beyond the commonly required threshold of 5σ, the latest combination of measurements performed by the ATLAS and CMS collaborations observing VBF Higgs boson production at 5.4σ confidence level [15]. The other production modes are still being investigated to gather further evidence.

2.6.2 Decay channels

The phenomenology of the Standard Model Higgs boson is rich and diverse. Figure 2.11 shows the leading branching fractions of Higgs boson decays. For a Higgs boson mass of $m_H \approx 125\,\text{GeV}$, the decay mode $H \to bb$ is the dominant one. However, measurements of Higgs boson decays to this final state is difficult at hadron colliders, as the final state is purely hadronic. In order to be able to select these events at the trigger level and separate signal events from the background, analyses need to rely on additional features of the final state and cannot use the dominant gluon fusion production mode, instead concentrating on the VBF, VH and ttH production modes [52–57]. The second most probable decay mode to two W bosons does not share this problem, as the W bosons can decay leptonically, providing an effective handle to select these events. A dedicated analysis for the ggF, VBF and VH modes of Higgs boson production in this decay channel is presented in Chapter 4. Similarly, the decay to a pair of τ-leptons can be exploited for dedicated analyses [58, 59]. Decays to gluons or to charm quarks are too difficult to distinguish from the background at a hadron collider. Decays to pairs of Z bosons or to photons can be very effectively selected due to excellent control over the backgrounds. A corresponding combined analysis is presented in Section 5.1. The decays $H \to Z\gamma$ and $H \to \mu\mu$ finally are so rare that they could not be established experimentally so far. Dedicated analyses for these final states are being conducted, but will need to rely on more data than is currently available in order to reach a definite conclusion [60–63].

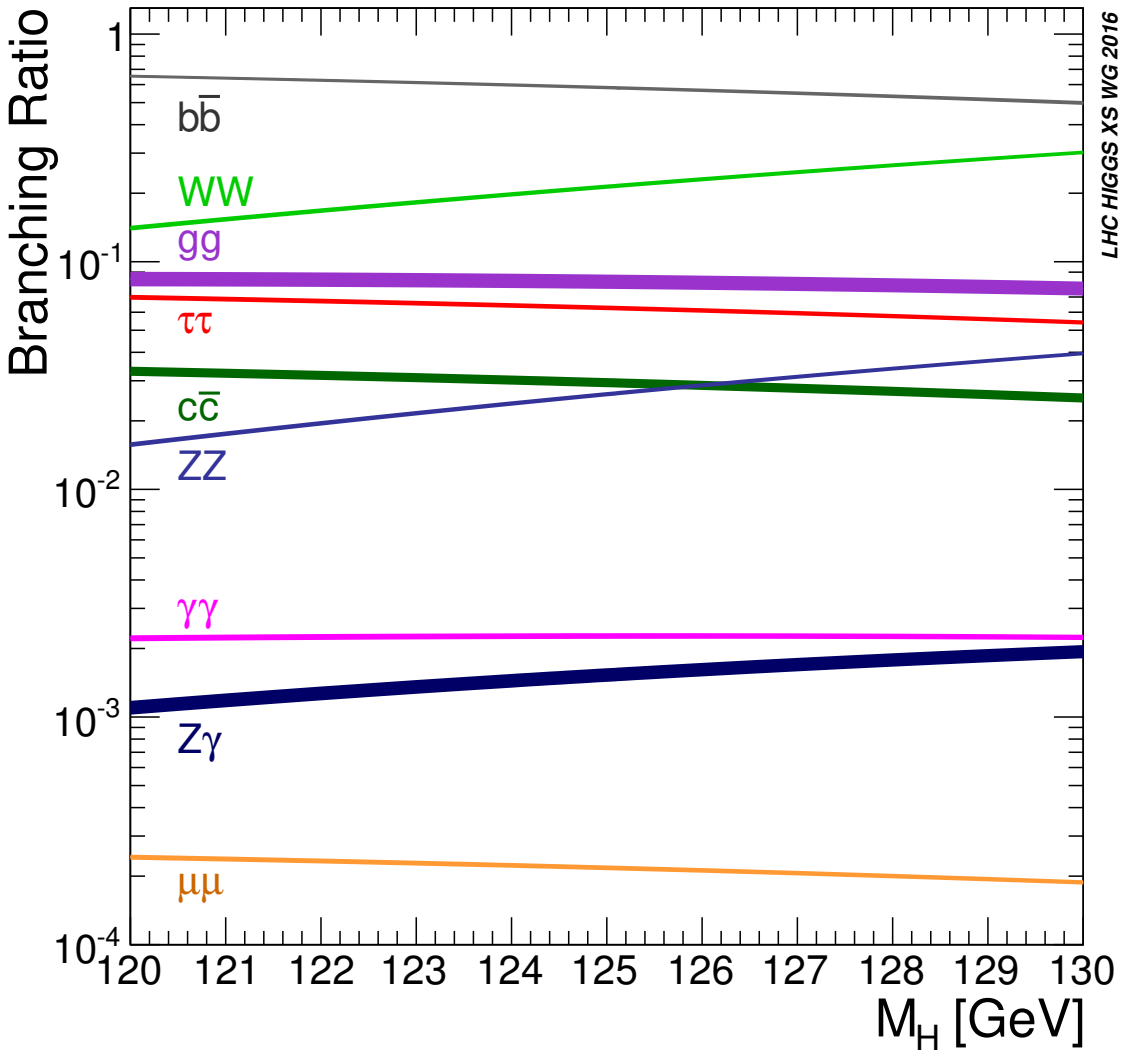


Figure 2.11.: Standard Model predictions for the leading decay branching fractions of a Higgs boson with a mass in the range of 120-130 GeV, with the line widths encoding the respective theory uncertainties [49].

3 ATLAS at the LHC

> **If speculative ideas cannot be tested, they're not science; they don't even rise to the level of being wrong.**
>
> *Wolfgang Pauli*

The most irrefutable evidence for the existence of a new particle is always its direct observation. As the vast majority of particles are unstable, they need to be produced in an experiment in order to observe them. *Collider experiments* have proven extremely successful in achieving this task, colliding beams of particles at a high center-of-mass energy at dedicated *interaction points* surrounded with a wealth of measurement devices to observe short-lived, exotic particles at a high rate under controlled experimental conditions.

The Large Hadron Collider (LHC), shown in Fig. 3.1, is the largest and most advanced particle collider ever built, capable of achieving center-of-mass energies of $\sqrt{s} = 13\,\text{TeV}$ in proton-proton collisions. It is operated by the European Council for Nuclear Research (CERN, *Conseil Européen pour la Recherche Nucléaire*), is located near Geneva (Switzerland) and hosts various experiments of high-energy physics. The ATLAS detector is located at one of its interaction points. It was built and is operated by a worldwide collaboration of physicists to test the predictions of the Standard Model to extreme accuracy and possibly discover deviations.

This chapter gives an overview over the design and functionality of both the LHC and the ATLAS detector. A much more detailed description of the hardware aspects can be found in the original technical design reports [64–68], in the journal of instrumentation [69, 70] and in the form of reports on installed upgrades [71].

The last section of this chapter also provides a summary of the measurements of the Standard Model at the LHC and summarizes the accomplishments achieved by the concerted efforts of the ATLAS and CMS experiments since large-scale data taking of collisions at unprecedented center-of-mass energy started in 2011. Particular attention is paid to the properties and the discovery of the Higgs boson.

3.1 The Large Hadron Collider

The Large Hadron Collider is a two-ring, superconducting hadron accelerator and storage ring, installed in a 27 km long tunnel at CERN. It is located in the area of Geneva, between the Jura mountains in the north west, Lake Geneva in the east and the Rhône river in the south. It is capable of colliding proton pairs with an unprecedented center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ or pairs of heavy ions such as lead nuclei at center-of-mass energies of about 5 TeV per nucleon. It hosts four large and a collection of smaller high-energy physics experiments.

- ATLAS (see Section 3.2) and CMS [74] are two general purpose detectors, aiming for precise measurements of electroweak physics and the discovery of new particles and interactions beyond the predictions of the Standard Model. They were instrumental for the discovery of the Higgs boson in July 2012 [13, 14].
- LHCb [75] is an experiment dedicated to measurements of CP violation and rare decays of b-hadrons.

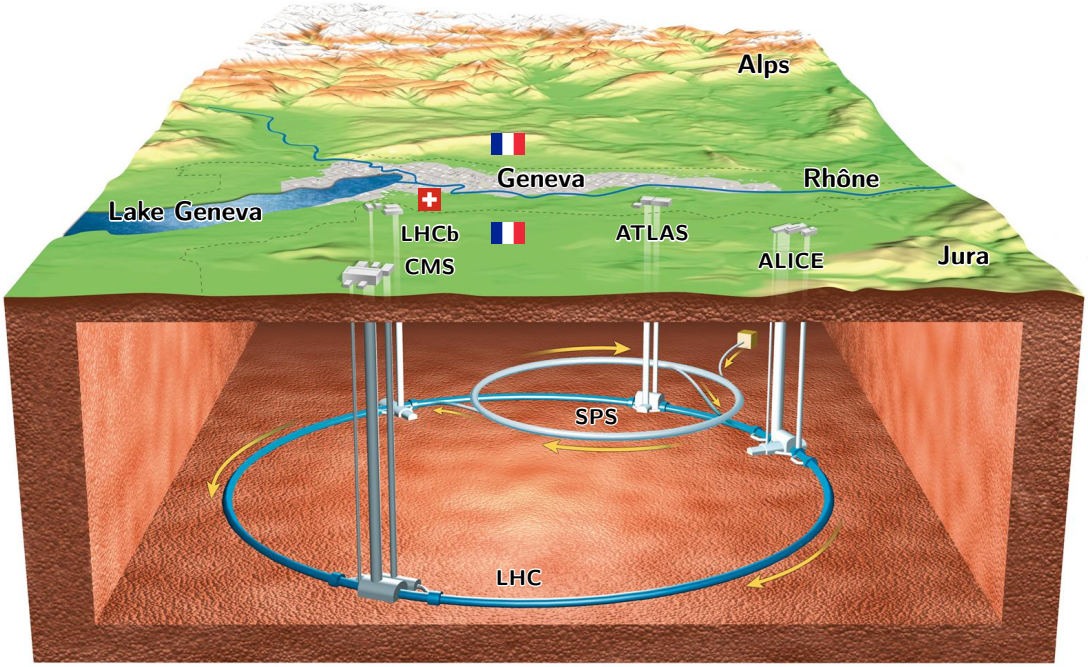


Figure 3.1.: Schematic display of the location of the LHC underground tunnel as well as the SPS pre-accelerator in the Geneva area. The four main LHC experiments ATLAS, CMS, LHCb and ALICE are also shown [72].

- ALICE [76] is a heavy-ion experiment, designed to address physics of strongly interacting matter and the quark-gluon-plasma at extreme values of energy density and temperature.
- TOTEM [77] is an experiment dedicated to measuring the total cross section in pp collisions and study elastic and diffractive scattering. It shares an interaction point with the larger CMS experiment.
- LHCf [78] collects data to improve the hadron interaction models used in the study of extremely high energy cosmic rays. It shares an interaction point with the larger ATLAS experiment.
- MoEDAL [79] is a dedicated experiment searching for magnetic monopoles and other massive meta-stable particles beyond the Standard Model. It shares an interaction point with the larger LHCb experiment.

The LHC is located in the same tunnel that previously housed the Large Electron Positron Collider (LEP) [80], which was dismantled around the year 2001 after completing its scientific agenda.

Earlier hadron colliders such as the TeVatron [81] were using a shared magnet and vacuum system to collide protons with anti-protons. However, the high beam intensities implied by instantaneous luminosities in the range of $L = 10^{34}\text{cm}^{-2}s^{-1}$ make the use of anti-protons impractical. Thus, the LHC was designed as a proton-proton collider with separate magnetic fields and vacuum systems for each beam.

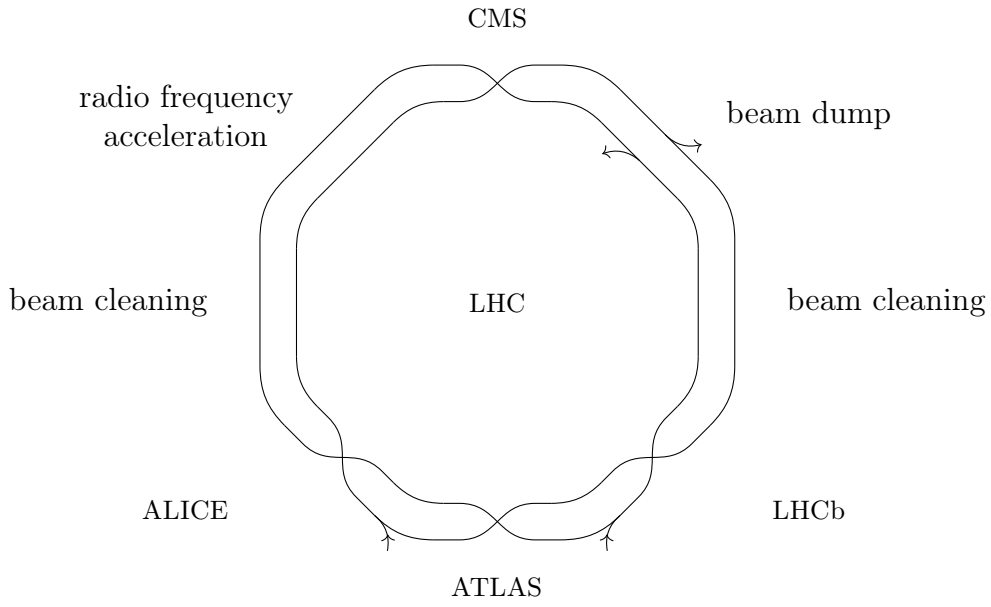


Figure 3.2.: Schematic display of the octagonal structure of the LHC with interaction points and facilities.

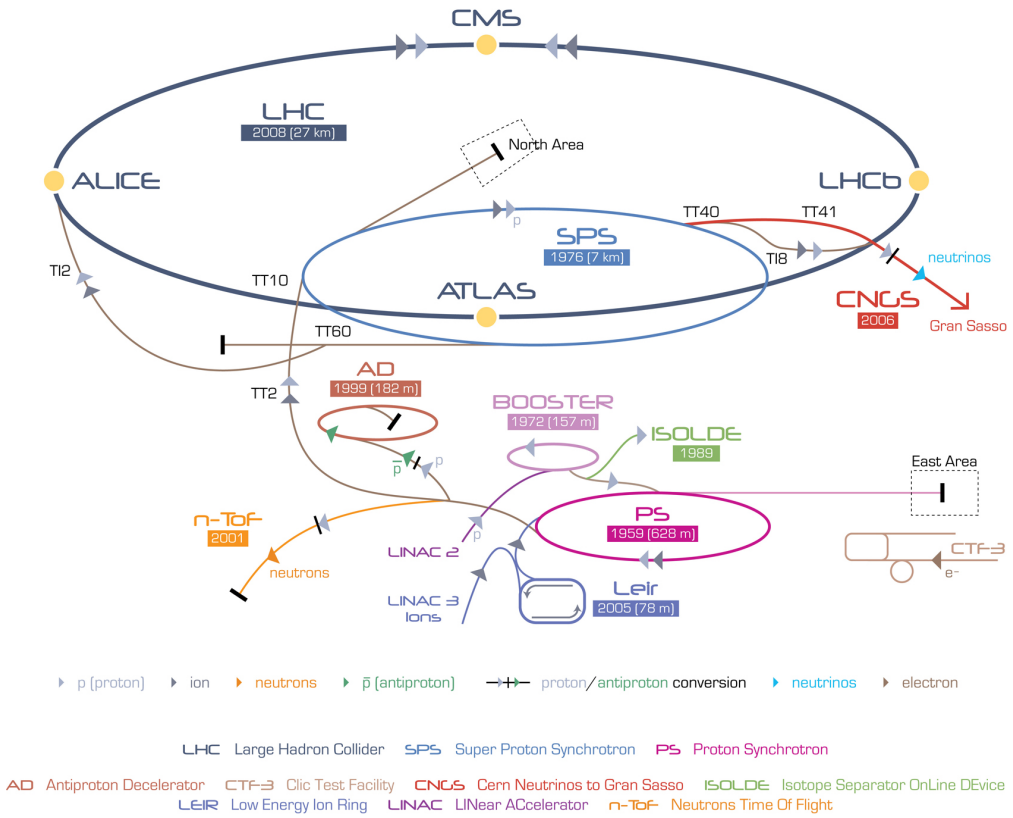


Figure 3.3.: The CERN accelerator complex, showing the Linear accelerator 2 (Linac 2), the Booster, Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) which act as pre-accelerators for protons to be injected into the Large Hadron Collider (LHC), as well as many other experimental facilities located at CERN [73].

The LHC is based on twin-bore, superconducting Niob-Titanium magnets, cooled down to 1.9 K with superfluid Helium and able to reach magnetic fields of up to 8.33 T. More than 1200 dipole magnets force the hadron beams into their near circular orbit, and smaller quantities of various dipole, quadrupole, sextupole and octupole magnets provide the beam optics necessary to maintain, focus and collide the beams at the four interaction points in straight sections of the octagonal structure housing the experiments. The other four straight sections house technical facilities for beam cleaning, radio frequency acceleration and beam dumping. A schematic display of the structure can be seen in Fig. 3.2.

Up to 2808 bunches of approximately 100 billion protons each are injected into the LHC from the Super Proton Synchrotron (SPS) at an energy of 450 GeV in 12 batches of up to 288 bunches each. The entire accelerator complex used in this process as well as other experimental facilities located at CERN can be seen

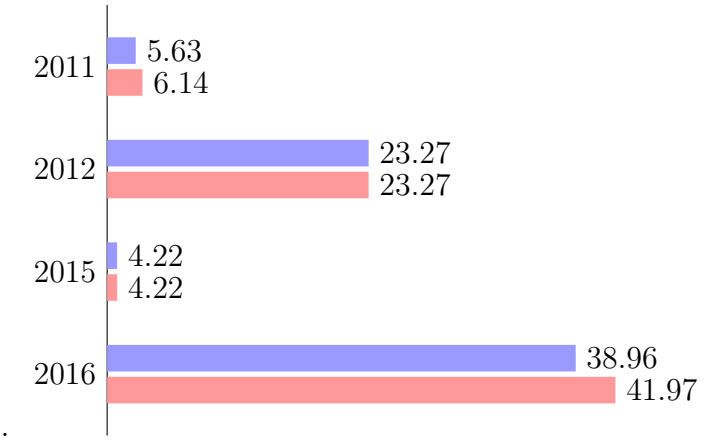


Figure 3.4.: Total luminosity in inverse femtobarn (fb^{-1}) delivered by the LHC to the ATLAS (blue, top) and CMS (red, bottom) experiments by year.

in Fig. 3.3. Ramping up the beams to an energy of 6.5 TeV takes approximately 25 minutes. Combined with around 20 minutes time required for the injection sequence to complete, this yields a total minimum of 45 minutes for the LHC to obtain stable beams. Collisions are then ideally taken for 10 hours or more, until the beam deteriorates so much that it is more efficient to dump and refill.

The LHC has been actively recording *pp* collision data since the year 2010, the center-of-mass energy increasing over time. One generally distinguishes two phases of data taking:

- The phase employing $\sqrt{s} = 7\,\text{TeV}$ in the year 2011 and $\sqrt{s} = 8\,\text{TeV}$ in the year 2012 is generally referred to as **LHC Run 1**.
- The phase employing $\sqrt{s} = 13\,\text{TeV}$ in the years 2015 and 2016 is generally referred to as **LHC Run 2**.

During the 2016 proton run, the LHC was able to deliver stable beams to the experiments for about 60% of its operational time, exceeding its design instantaneous luminosity by 30% and featuring collisions at a rate of one bunch crossing every 25 ns, delivering a total of about $40\,\text{fb}^{-1}$ to the ATLAS and CMS experiments each in the course of the year [82]. A bar chart showing the luminosity delivered by the LHC in different years of data taking is shown in Fig. 3.4.

3.2 The ATLAS detector

The ATLAS (**A** **T**oroidal **L**HC **A**pparatu**s**) detector is a general purpose particle detector located at the LHC interaction point closest to the main CERN site in Meyrin near Geneva, Switzerland, as can be seen in Fig. 3.1. An artists depiction showing a cut-away view can be seen in Fig. 3.5.

The entirety of the detector is around 44 m long and 25 m in diameter, with a total weight of around 7000 t. It is close to hermetic, providing almost omnidirectional single particle identification as well as precision measurements of energy and momentum for extremely high rates of high-energy particles emitted from the interaction region in the center. Like most general-purpose detectors, it is built in a multi-layered, onion-like structure, each layer fulfilling a dedicated task. This section gives an overview of the individual layers, their function, design and characteristics as well as the mechanisms involved in data taking.

The ATLAS detector is operated by a diverse collaboration of around 3000 scientists and engineers from approximately 180 institutions, representing 38 countries from all populated continents [83]. The agenda of the ATLAS experiment encompasses a wide range of topics, including

- measurement of the properties of the Higgs boson,

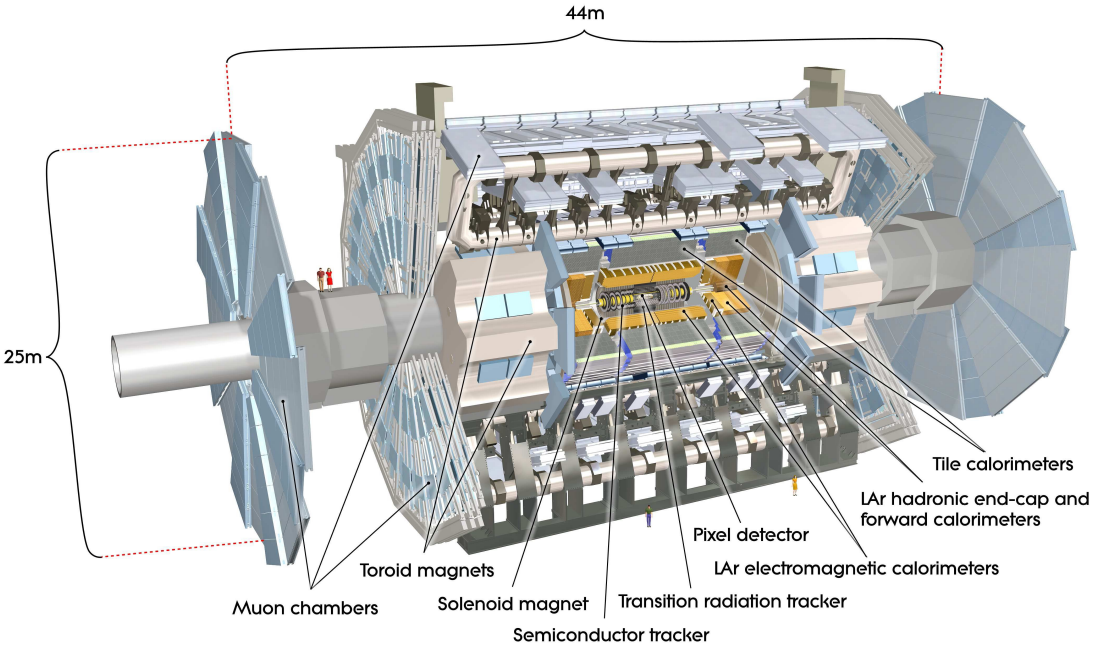


Figure 3.5.: Schematic view of the ATLAS detector, showing the largest components as well as length scales and humans for comparison [69].

- precision measurements of the Standard Model predictions, and
- search for effects beyond the Standard Model.

The ATLAS detector uses a right-handed coordinate system with its origin placed at the nominal interaction point. The z-axis is a tangent to the LHC ring, pointing from the Jura mountains towards Geneva. The x-axis is perpendicular to that, pointing towards the center of the LHC ring, while the y axis is pointing upwards. The azimuthal angle $\phi \in [0, 2\pi]$ is measured around the beam axis against the x axis. The polar angle $\theta \in [0, \pi]$ is measured against the beam axis, but it is customary to use the *pseudorapidity* $\eta = -\ln \tan \theta/2$ instead. Transverse momenta and energies are defined in the x-y-plane. Distances are typically measured as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.2.1 The inner detector

The inner detector is the detector section closest to the interaction point. It is devoted to tracking, that is, measuring the trajectories or "tracks" of individual charged particles emerging from the collision. Figure 3.6 shows a schematic view of the inner detector. It is immersed in a 2 T magnetic field generated by a superconducting solenoid coil and is divided in two main sections: the silicon tracker and the transition radiation tracker.

The silicon tracker

The silicon tracker covers the range of $|\eta| < 2.5$ and exploits the semiconductor properties of the silicon. Charged particles passing through the doped silicon will set free valence electrons, generating electron-hole-pairs in the conduction band. Under the influence of an electric field, these charge carriers travel to the electrodes, where the electric signals can be recorded.

The **insertable *b*-layer** (IBL) shown in Fig. 3.7 was inserted in the long LHC shutdown 2013–2015, providing an additional barrel layer. The closest sensors are located a mere 3.1 cm from the beam, yielding an additional 12 million readout channels. The main function of this additional layer is to improve the precision of vertex reconstruction. This is especially useful to identify decay vertices of heavy hadrons including b-quarks, as their typical decay lengths are around 450 μm.

The next layers form the **silicon pixel detector**, shown in Fig. 3.8. It consists of three barrel layers and three end-cap layers of semiconductor material with a very fine ϕ-z-segmentation down to $50 \times 400\,\mu\text{m}^2$, starting around 5 cm from the beam line in the barrel region. The number of readout channels exceeds 80 million.

The bulk of the silicon tracker is taken up by the **silicon microstrip** detector, consisting of four barrel layers and nine end-cap layers. The first modules in the barrel region start around 30 cm from the beam line. Each module has a twin

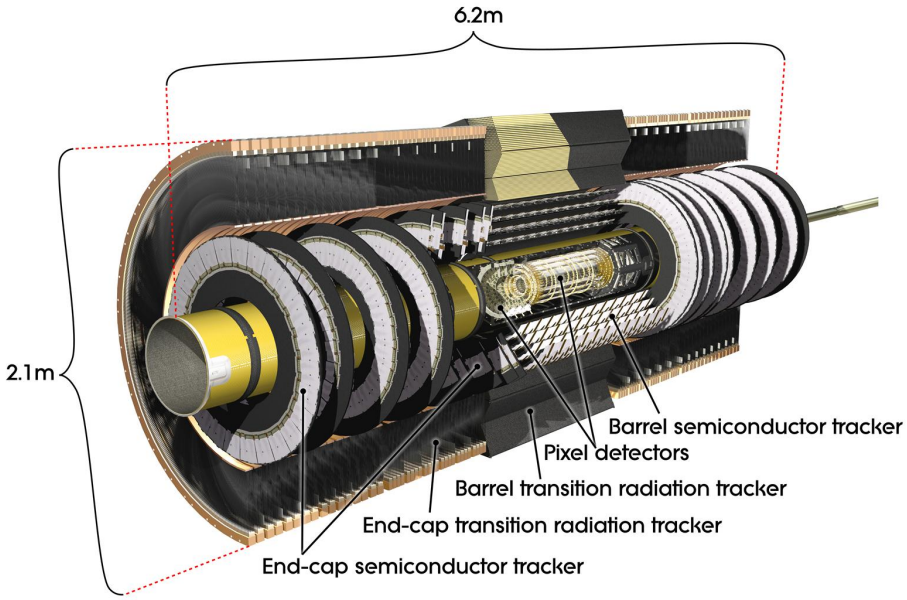


Figure 3.6.: Schematic view of the ATLAS inner detector [69] as installed during LHC Run 1 (2008-2013).

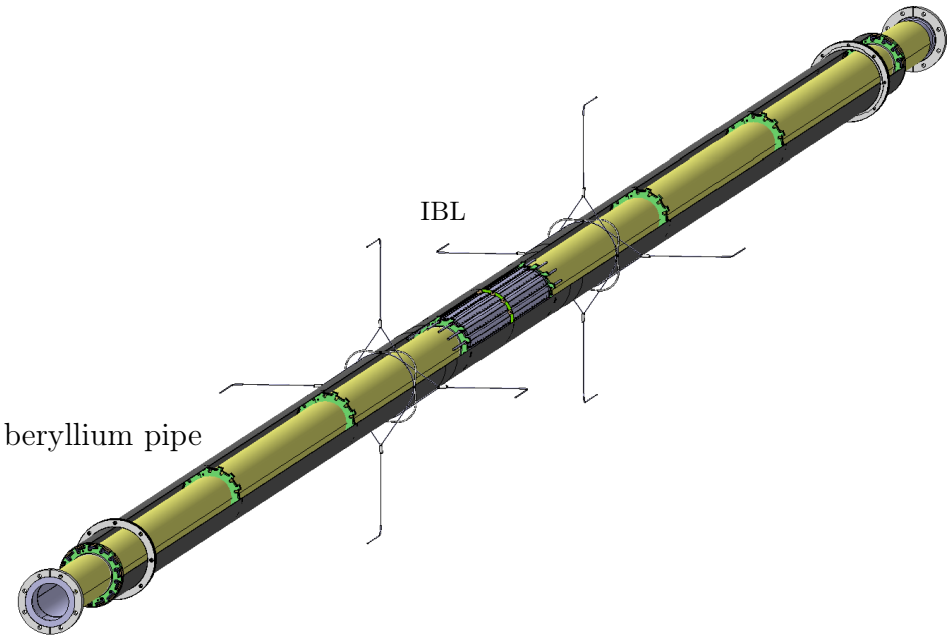


Figure 3.7.: Schematic view of the ATLAS insertible *b*-layer (IBL) and the new beryllium beam pipe [71] installed in 2015.

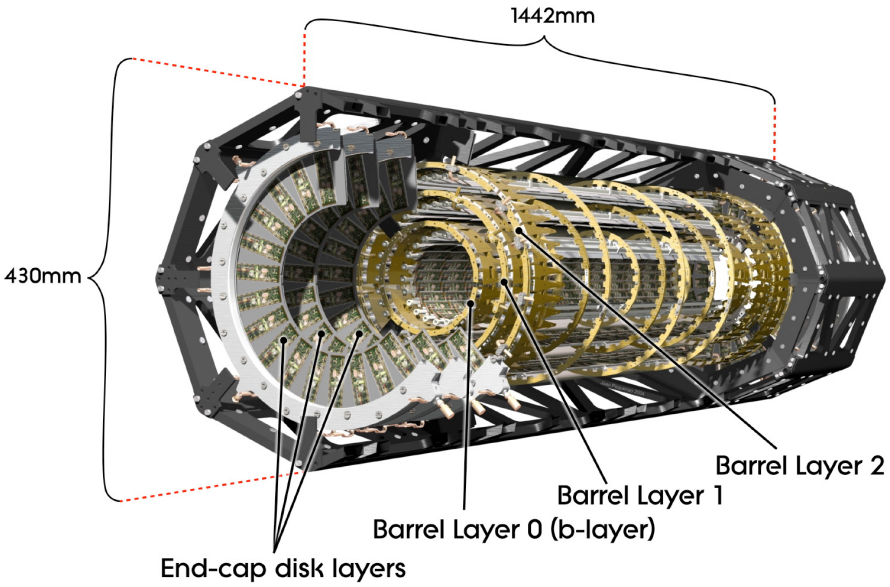


Figure 3.8.: Schematic view of the ATLAS pixel detector, not yet including the insertible *b*-layer (IBL) [71] installed in 2015.

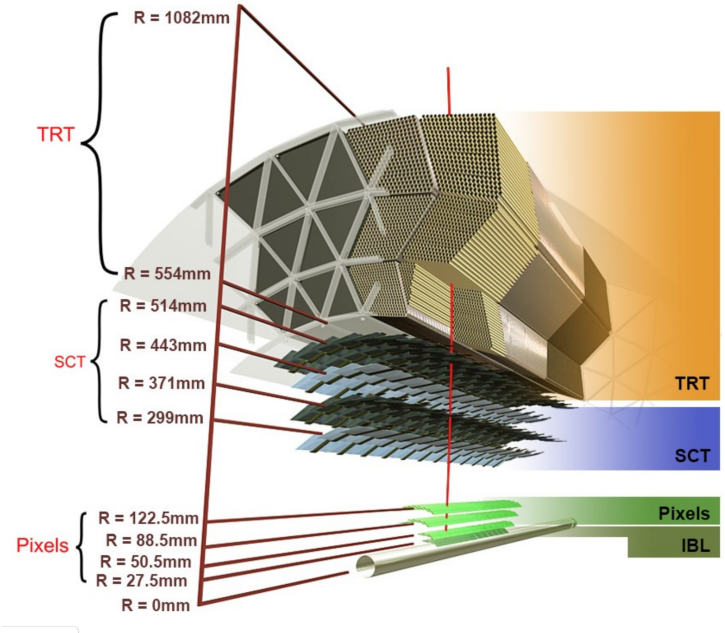


Figure 3.9.: Schematic view of the structure of the inner detector barrel, including the IBL [84].

layer of silicon strips,between 6−13 cm long and 80 μm wide, which are askew by 40 mrad with respect to each other to allow for two-dimensional readout, providing a total of 6.3 million readout channels.

This fine granularity allows to combine information from the multitude of measurement layers and reconstruct the track splines of individual charged particles curved by the Lorentz force of the magnetic field.

The transition radiation tracker

When a charged particle passes through a boundary between two media with different dielectric constants, it emits electromagnetic radiation with intensity logarithmically proportional to the Lorentz factor of that particle. This transition radiation is exploited by the transition radiation tracker (TRT), covering the range of $|\eta| < 2.0$.

The TRT consists of 4 mm diameter polyimide straw tubes coated with aluminum and graphite-polyimide on the inside and polyurethane on the outside. The 298 thousand tubes are filled with a xenon-based gas mixture. They do not provide any longitudinal segmentation.

The transition radiation intensity provides a measurement of the Lorentz factor of charged particles. Given a momentum measurement, this allows to directly measure the mass of a particle, thus providing additional sensitivity to distinguish electrons from charged pions, which is the primary function of this detector component. At 80% detection efficiency for electrons, the misidentification rate for pions is roughly 10% in the range $4\,\text{GeV} < p_\text{T} < 20\,\text{GeV}$ [85].

Combined performance

A cut-away view of the inner detector barrel can be seen in Fig. 3.9. With the new IBL, the $\delta x \times \delta y \times \delta z$ vertex resolution was improved from $15 \times 15 \times 34\,\mu\text{m}^3$ down to $11 \times 11 \times 24\,\mu\text{m}^3$ [71]. The misidentification rate for a very aggressive working point of 85% b-jet identification efficiency is approximately 3.1 for c-jets and approximately 33 for light jets [86].

The ATLAS inner detector reaches a relative momentum resolution of $\sigma_p/p \approx 0.48 \cdot p_\text{T}/\text{TeV}$ for highly energetic charged particles [87]. The relative uncertainty is expected to be proportional to the momentum.

3.2.2 The calorimeter system

The ATLAS calorimeter system shown in Fig. 3.10 consists of two main components and covers a range of $|\eta| < 4.9$. The electromagnetic calorimeter is a highly granular inner layer, providing precision energy measurements for electrons and photons, whereas the hadronic calorimeter forms the outer layer, providing sufficient stopping

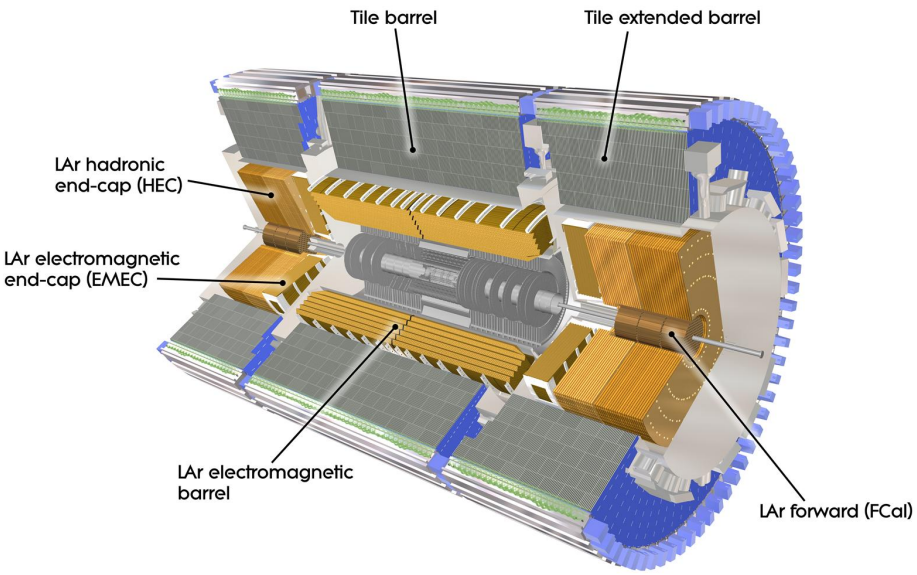


Figure 3.10.: Schematic view of the ATLAS calorimeter system [69].

power even for highly energetic hadronic jets and subsequently allows to measure the total energy-momentum-balance of the collision.

Both components of the ATLAS calorimeter system are sampling calorimeters, alternating layers of active medium and passive layers of high-density material to initiate and enforce full deposition of the energy of the traversing particle. In such sampling calorimeters, a typically fluid or gaseous active medium allows ionization charges created by the traversing particle to travel to oppositely charged electrodes for measurement. However, a scintillating solid converting the energy of the particle into light can also be used as an active medium.

The relative uncertainty on energy measurements using calorimeters has the form

$$\sigma_E^{\text{rel.}} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{3.2.1}$$

where a, b and c are η-dependent parameters. The sampling term with the coefficient a dominates at low energies. It is induced by the statistical nature of particle showers and heavily dependent on the fraction of active medium in the calorimeter. The noise term with coefficient b is dependent on the occupancy and thus on the average number of interactions per bunch crossing and also includes electronic noise. The constant term c originates from instrumental non-uniformities and dominates at high energies. The total uncertainty on jet energy measurements in the barrel region is around 2–5% over a broad range of the p_{T} spectrum [88].

The electromagnetic calorimeter

The barrel of the electromagnetic calorimeter stretches up to $|\eta| < 1.475$, the inner and outer wheel of the end-cap up to $|\eta| < 2.5$ and $|\eta| < 3.2$, respectively.

The materials used for active and passive layers are liquid argon and lead. The layers are arranged in an accordion-like geometry, providing complete ϕ symmetry without azimuthal cracks.

The total thickness exceeds $22X_0$ in the barrel region and $24X_0$ in the end-cap region, where X_0 is the *electromagnetic radiation length* over which a highly energetic electron loses on average all but $1/e$ of its energy due to bremsstrahlung.

The electromagnetic calorimeter consists of three designated layers. The first layer, the *presampler*, has a very high granularity of $\delta\eta \times \delta\phi = 0.003 \times 0.1$, allowing for a distinction between photons and neutral pions as well as between electrons and charged pions. The second layer is with $\delta\eta \times \delta\phi = 0.025 \times 0.025$ much coarser and with a thickness of $16X_0$ sufficient to contain the bulk of the energy deposition. The third layer has a thickness of only $2X_0$, a coarse resolution of $\delta\eta \times \delta\phi = 0.05 \times 0.025$ and is used to assess the leftover deposition intensity and thus the depth to which the shower has penetrated the calorimeter. The total number of readout channels arising from this segmentation is 175 thousand, with 110 thousand originating from the barrel.

Quoting the design values [89] of individual uncertainty terms as given in Eq. 3.2.1,

- the sampling coefficient is about $a \approx 0.1/\sqrt{\text{GeV}}$ at low $|\eta|$, and is expected to worsen in the forward and backward regions,
- the noise coefficient is about $b \approx 350 \cosh \eta \,\text{MeV}$ for the design configuration of around 20 interactions per bunch crossing, and
- the constant term has a design value of $c = 0.7\%$.

The hadronic calorimeter

The bulk of the hadronic calorimeter is the **tile calorimeter**. The tile barrel covers the region $|\eta| < 1.0$, and the tile extended barrels cover $0.8 < |\eta| < 1.7$. In this component, steel is used as the absorber, whereas scintillating plastic tiles play the role of the active medium.

The **hadronic end-cap** extends out to $|\eta| < 3.2$, consisting of two independent wheels per cap. It shares the liquid argon supply with the electromagnetic calorimeter for the active medium, but uses copper as absorber material.

The **forward calorimeter** ranges up to $|\eta| < 4.9$, consisting of three independent wheels in each hemisphere. It also shares the liquid argon supply with the electromagnetic calorimeter, and while the first layer uses copper as absorber material similar to the hadronic end-cap, the other two layers use tungsten.

The total thickness of the ATLAS calorimeter system for hadronic interactions is around 9.7λ in the barrel region and 10λ in the end-caps, where λ is the *nuclear interaction length* over which the number of relativistic hadrons reduces by a factor $1/e$.

With the nearly hermetic calorimeter system, it is possible to measure momentum imbalances in collisions up to a precision of 5-10%, depending on the algorithm used [90]. The total number of readout channels for the tile calorimeter is 10 thousand.

Quoting the design values of individual uncertainty terms as given in Eq. 3.2.1, obtained for pions in a test beam of 20 GeV around $\eta = 0.35$ [69],

- the sampling coefficient is (in units of GeV) about $a \approx 56\%$ at low $|\eta|$ for the tile calorimeter, about $a \approx 71\%$ for the hadronic end-caps, and about $a \approx 94\%$ for the forward calorimeter,
- the noise coefficient is consistent with $b \approx 0$ everywhere, and
- the constant term has a design value of $c = 5.5\%$ for the tile calorimeter, $c = 5.8\%$ for the hadronic end-cap, and $c = 7.5\%$ for the forward calorimeter.

3.2.3 The muon spectrometer

The muon spectrometer of ATLAS is perhaps the most iconic detector component. It occupies by far the largest share of volume and relies on its own magnet system, featuring the large superconducting toroid magnets eponymous to ATLAS, of which eight are located in the barrel and two in the forward region, cooled with liquid helium and providing $0.5 - 2\,\mathrm{T}$ of magnetic field strength in the detection volume.

The muon chambers are structured in an octagonal symmetry and feature a variety of different designs, all based on ionization in gas with the charges drifting to oppositely charged electrodes to be measured, mostly in the form of gas-filled tubes with immersed wires.

Eight layers of Monitored Drift Tubes (MDT, 354 240 tubes) with anode wires reach up to $|\eta| < 2.7$, except for the innermost end-cap layer ending at $|\eta| < 2.0$. In the forward region $2.0 < |\eta| < 2.7$, cathode strip chambers (CSC, 70 000 channels) are used – multiwire proportional chambers with the cathodes segmented into strips in orthogonal directions. These components both use an argon-based gas mixture and provide precision tracking capability for muons.

As a complement, Resistive Plate Chambers (RPC, 380 000 channels) and Thin Gap Chambers (TGC, 400 000 channels) have been chosen to provide fast readout for trigger purposes. The RPCs are wireless chambers with anode and cathode plates and are based on an organic gas mixture, covering a range up to $|\eta| < 1.05$. The TGCs are multiwire proportional chambers where the wire-to-cathode distances are smaller than the wire-to-wire distances. They range $1.05 < |\eta| < 2.4$ and use a mixture of n-pentane and CO_2.

The total uncertainty for the measurement of transverse momentum for muons is around 3% over a wide range of p_{T} and up to 10% for muons with 1 TeV of transverse momentum [91].

3.2.4 Trigger

The rate at which pp collisions are produced in the ATLAS interaction region is far larger than what could possibly be recorded. The ATLAS trigger system forms a highly optimized network of hardware and software solutions to this problem. A brief introduction is given in this section.

With bunch crossings happening every 25 ns, the instantaneous luminosity exceeds $L = 10^{34}\mathrm{cm}^{-2}s^{-1}$ and is continuously measured by dedicated luminosity monitors. Photomultipliers with quartz fiber bundles to observe Cherenkov light, emitted when a charged particle passes through a medium with a speed exceeding the

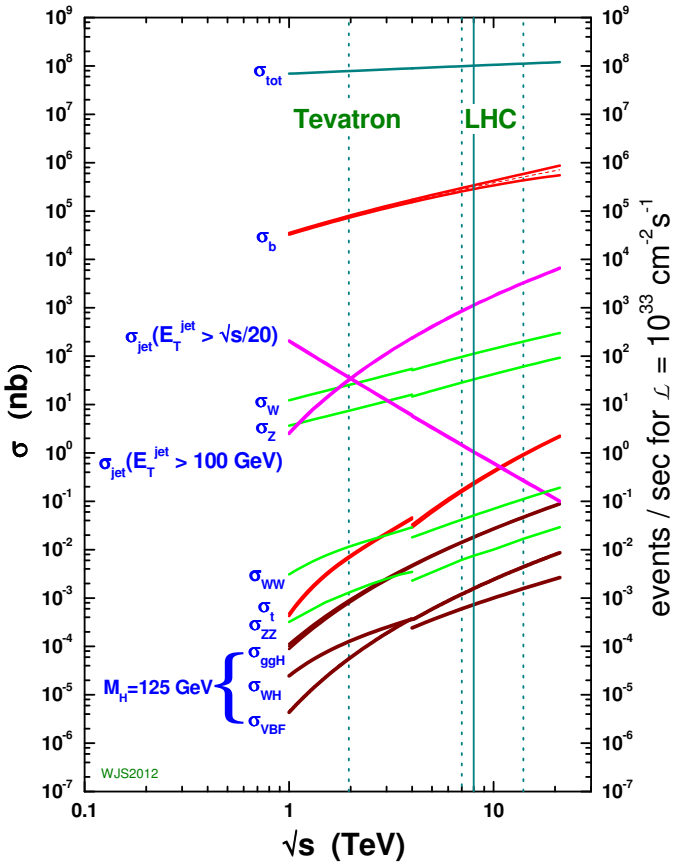


Figure 3.11.: Cross sections for proton collisions as a function of center-of-mass energy $\sqrt{s}$. The discontinuity between TeVatron and LHC energies reflects the switch from $p\bar{p}$ to pp collisions [92].

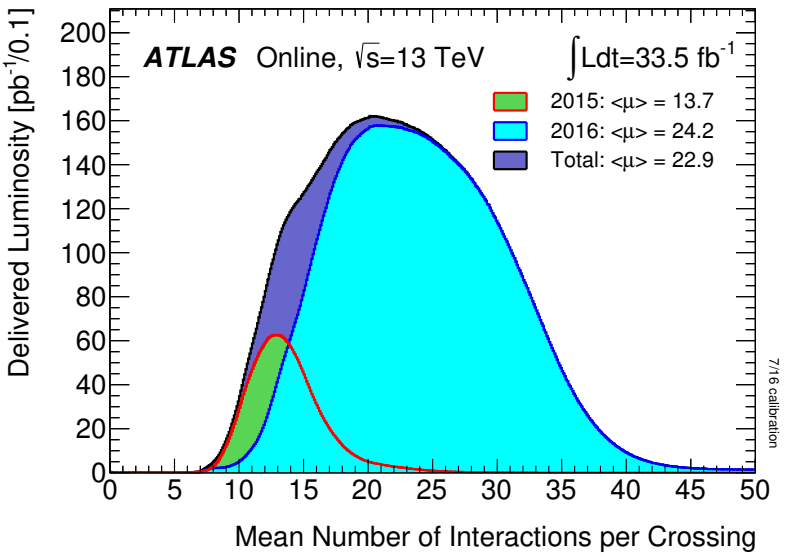


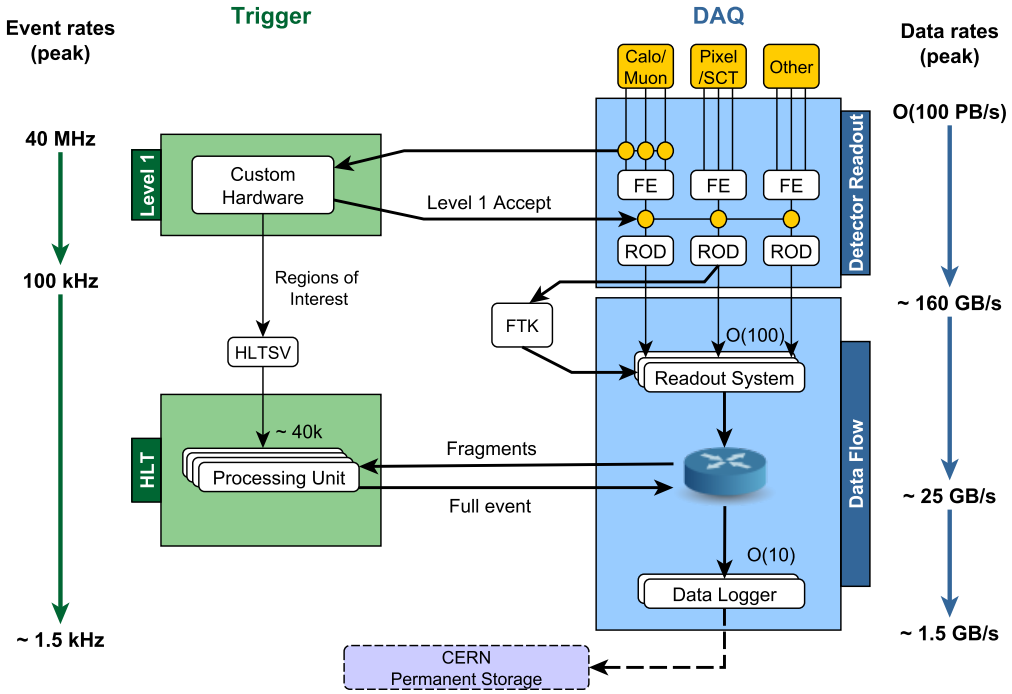
Figure 3.12.: Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the combined pp collision data at $\sqrt{s} = 13\,\text{TeV}$ recorded in 2015 and 2016, calculated from the instantaneous per-bunch luminosity, where the inelastic pp collisions cross section is taken as 80 mb [93].

speed of light in that medium, are arranged around the beam pipe 17 m from the interaction point for this purpose.

At this rate, a continuous extraction of collision data from the millions of readout channels is not feasible. Events need to be selected efficiently, especially since the cross section is massively dominated by QCD interactions such as multijet production, whereas processes like Higgs boson production only contribute few *parts per trillion* to the total cross section. This can be seen in Fig. 3.11, which shows cross section predictions for different physics processes, obtained for pp and $p\bar{p}$ collisions for center-of-mass energies typical for the LHC.

The *trigger system* is designed to make fast and informed decisions about which events should be recorded, and which should not. The average number of interactions per bunch crossing as observed in data during LHC operation, a quantity often loosely referred to as *pile-up*, is shown in Fig. 3.12. The mean value of $\mu = 24.2$ achieved during LHC Run 2 is significantly higher than the maximum number of $\mu = 20$ envisioned in the original design considerations. The extremely high rates put tough constraints on the hardware and software used in the trigger system. The ATLAS trigger system used during LHC Run 2 consists of two layers:

- The **Level 1** (L1) trigger consists of fast, custom-made electronics that can determine regions of interest (RoI) from calorimeter and muon system signals by selecting an $\eta \times \phi$ region around accepted L1 objects. Using integrated circuits, the L1 trigger can perform event selections using complex quantities



HLT: High Level Trigger – FTK: Fast Tracker – DAQ: Data Acquisition
ROD: Read-Out Device – HLTSV: HLT Supervisor – FE: Front End

Figure 3.13.: Data Acquisition and Trigger diagram, showing data flow and event rates achieved during LHC Run 2 [94].

such as missing transverse energy. It is capable of reducing the data rate from around 40 MHz down to 100 kHz.

- The **High-Level Trigger** (HLT) is a software solution that can make use of these RoIs or perform high-quality, full event reconstructions. It reduces the event rates from around 100 kHz down to roughly 1 kHz.

The trigger system features a list of triggers selecting events for different criteria. Common criteria involve requiring leptons passing certain identification criteria as well as some transverse momentum threshold. Events selected by any trigger are then written to disk and reconstructed at a designated computing farm at CERN, which serves as the root node for a distributed computing network, the *Worldwide LHC Computing Grid* used for creation of simulation data and large-scale data analysis. Figure 3.13 shows a diagram of the data flow and event rates achieved by the ATLAS trigger system during LHC Run 2.

3.2.5 Reconstruction

The ATLAS reconstruction software uses a wide variety of different algorithms to identify and disentangle energy depositions and tracks from individual particles or particle jets recorded by the detector. As these algorithms greatly influence the efficiency and performance of the reconstruction process, they are under constant development. This section gives a short overview over the state-of-the-art algorithms used to reconstruct data events recorded by ATLAS during LHC Run 2.

Tracking

Track reconstruction is the process of identifying individual charged particles emerging from the interaction and measuring their traces in the inner detector. Such charged particles can be measured up to $|\eta| < 2.5$ within ATLAS. With the radius of their curvature and the known magnetic field, it is ultimately possible to calculate the transverse momenta of particles. This procedure is subdivided in three stages [95]:

1. Individual cells from the silicon detector are clustered by connecting adjacent hits. Cluster splitting is performed using a neural network. Timing information from the TRT is converted into drift-circles.
2. Track seeds are identified from a combination of pixel and strip layer hits, requiring a total of three hits. They are processed by applying a set of thresholds as well as a combinatorial Kalman filter to form a combined particle track.
3. Ambiguities are resolved using a scoring mechanism taking into account double usage of hits in several tracks of the same event as well as layers missing a hit for a specific track.
4. Tracks within the TRT coverage that survive the ambiguity solver are extended and combined with the TRT tracking information. This significantly increases the momentum resolution due to the increased lever.

After tracks have been identified, splines are fit and combined with data about the magnetic field strength to measure the momentum of the particle associated to the track.

Vertexing

A *primary vertex* is an identified interaction point of two beam protons. Vertices are reconstructed by matching up intersecting track groups, which proceeds in three main steps: seeding, track assignment, and fitting. A rough outline of the procedure is as follows:

- All reconstructed tracks satisfying certain criteria are extrapolated to the beam axis. The impact parameter z_0, defined as the longitudinal distance at the point of closest approach, is computed for every track with respect to the nominal beam spot. This information is used to generate seeds, for which an iterative likelihood maximization method is used.
- Tracks compatible with a seed are grouped together for fitting.
- The adaptive vertex fitting algorithm [96] is used to estimate the position and uncertainty of the vertex.
- Incompatible tracks not used in a previous vertex are used to repeat the procedure starting from the creation of a new seed.

The identification of vertices becomes more challenging as the average number of interactions per bunch crossing (the *pile-up*) increases, and thus new algorithms are being developed to increase the resilience of ATLAS vertexing [97].

Calorimeter clusters

As the calorimeters are subject to a significant amount of background noise, the mean noise level is constantly measured and subtracted from measurements. For reconstruction of calorimeter clusters, a topological clustering algorithm is used [98]:

- First, seed cells are identified as cells exceeding the noise level by a factor of $S = 4$.
- Seeded clusters are iteratively grown to neighboring cells exceeding the noise level by a factor of $N = 2$.
- Finally, any cell with a positive energy contribution directly neighboring a cluster is added.

If overlapping clusters be identified, they are merged. If a cluster is found to have several local maxima, an attempt is made to split the cluster.

Electrons and photons

The primary tool for reconstruction of electromagnetic objects such as electrons and photons is the electromagnetic calorimeter. Roughly speaking, energy depositions with an associated track are regarded as electrons, such depositions without an associated track are regarded as photons. As tracking information is crucial, identification is most accurate in the area $|\eta| < 2.47$ covered by the inner detector acceptance, and excluding the $1.37 < |\eta| < 1.52$ transition region between barrel and end-cap of the electromagnetic calorimeter.

Electromagnetic energy depositions are reconstructed using a *sliding-window* approach acting directly on the calorimeter cells. A window is slid across $\delta\eta \times \delta\phi$ surface of the middle layer of the electromagnetic calorimeter, measuring 3×5

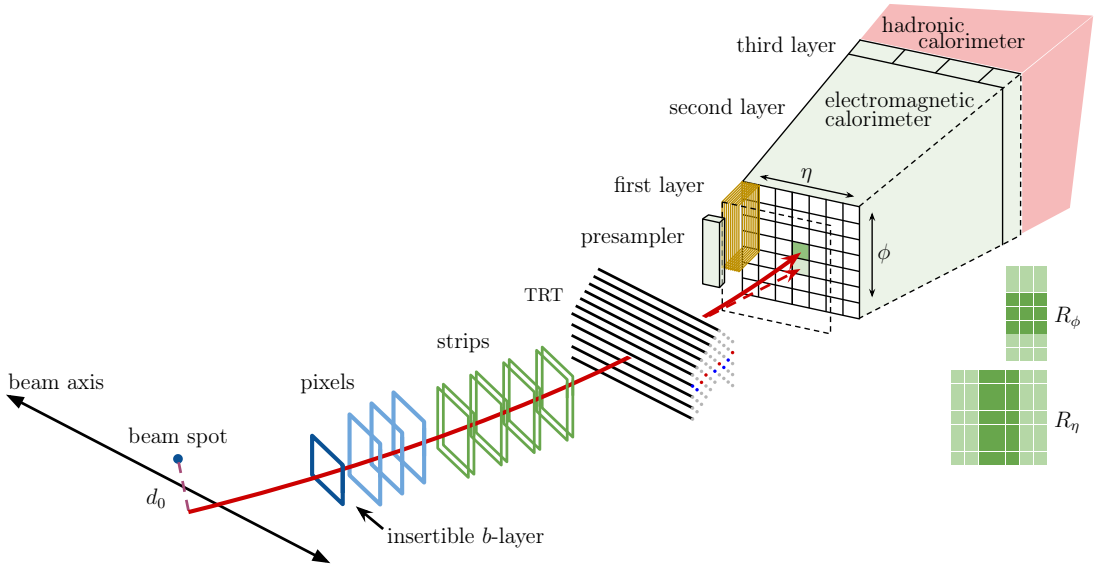


Figure 3.14.: Illustrative sketch of an electron reconstruction [99].

units of 0.025×0.025. A cluster is seeded if the transverse energy in the window exceeds $E_{\mathrm{T}} > 2.5\,\mathrm{GeV}$.

These clusters are then matched geometrically to extrapolated tracks from the inner detector to identify electrons, taking into account the radiative losses suffered while traversing the silicon tracker by virtue of a Gaussian-Sum-Filter [100]. A requirement of $E/p < 10$ is imposed to reject hadronic jets. When a match is found, the energy measurement is refined using a larger cluster size of 3×7 in the barrel region or 5×5 in the end-cap regions. The track is then refit and the momentum recomputed including the calorimeter information to form the four-momentum of the electron candidate. The entire procedure is sketched in Fig. 3.14.

In order to accommodate the needs of different analyses, several working points are defined for electron identification. For this purpose, a likelihood-based method taking into account various discriminating variables including shower shapes, track properties and the track-cluster-matching is used, providing three levels of identification, corresponding to different identification efficiencies at a benchmark transverse energy of $E_{\mathrm{T}} = 100\,\mathrm{GeV}$ with increasing background rejection [101]:

- `LooseLH` corresponds to 96% identification efficiency,
- `MediumLH` corresponds to 94% identification efficiency, and
- `TightLH` corresponds to 88% identification efficiency.

Muons

About 96% of muons are reconstructed by fitting hits from inner detector and muon spectrometer tracks, taking into account energy loss in the calorimeters,

producing *combined muons*. The remainder of reconstructed muons are formed by tagging tracks identified by the inner detector with muon signatures in the calorimeter or the muon spectrometer [102].

Muons are identified using a set of requirements developed to reject secondary muons coming from pion and kaon decays. Four muon selections are provided:

- The `Loose` identification criteria are designed to maximize muon reconstruction efficiency while ensuring reliable muon tracks.
- The `Medium` criteria, representing the default selection for muons in ATLAS, is designed to minimize systematic uncertainties associated with muon reconstruction and calibration.
- The `Tight` criteria minimize the fake muon rate to optimize sample purity.
- Finally, the `High-`p_{T} selection aims to maximize momentum resolution for tracks with transverse momentum above 100 GeV.

Lepton isolation

As leptons originating from decays of W or Z bosons are expected to be isolated, which is not necessarily the case for leptons originating from secondary processes, lepton isolation criteria are also often imposed to reject these contributions. These criteria can encompass track-based isolation by placing a cut on the scalar momentum sum over all particles originating from the same primary vertex in a certain cone size around the lepton candidate, or calorimeter-based isolation by placing a cut on the sum of calorimeter energy depositions around the cluster associated with the lepton candidate. Also here, different working points are provided:

- `Loose` and `Tight` isolation provide isolation efficiencies of 99% and 95%, respectively, independent of the lepton p_{T}.
- `Gradient` and `GradientLoose` isolation provide isolation efficiencies linear in lepton p_{T}. The `Gradient` working point is tuned to produce a nominal efficiency of 90% for leptons with p_{T} of 25 GeV, rising to 99% at 60 GeV. The `GradientLoose` working point is tuned to produce a nominal efficiency of 95% for a p_{T} of 25 GeV, rising to 99% at 60 GeV. The efficiency functions are

$$\begin{aligned} \epsilon &= 0.1143 \cdot p_{\mathrm{T}}/\mathrm{GeV} + 92.14 && \text{for } \texttt{Gradient} \\ \epsilon &= 0.057 \cdot p_{\mathrm{T}}/\mathrm{GeV} + 95.57 && \text{for } \texttt{GradientLoose} \end{aligned}$$

  for both track and calorimeter isolation.
- `LooseTrackOnly` isolation has an efficiency of 99% for all leptons independent of their p_{T} and applies only track isolation criteria.

Jet finding

Jets are reconstructed starting from topological calorimeter clusters. For this purpose, it is useful to define the distance measure d_{ij} between two objects i and j or d_{iB} between the object i and the beam B as

$$d_{ij} = \min\left(p_{T,i}^{z}, p_{T,j}^{z}\right) \cdot \frac{R_{ij}^{2}}{R^{2}} \tag{3.2.2}$$

$$d_{iB} = p_{T,i}^{z}, \tag{3.2.3}$$

where $p_{T,i}$ is the transverse momentum of object i, R is some fixed distance measure to be used as a parameter by the algorithm, and the parameter z governs the power to which the momentum enters the calculation.

With these distance measures defined, the jet finding then proceeds by iteratively finding the object pair with the smallest distance, and

- if both are clusters, merge the clusters i and j, or
- if either is a beam, identify the other as a jet and remove the cluster from the list

until no unmatched clusters are left.

A typical value often used for the distance measure is $R = 0.4$. Different values of z correspond to different, widely used jet finding algorithms:

- The case $z = 2$ corresponds to the k_T algorithm,
- the case $z = -2$ corresponds to the anti-k_T algorithm, and
- the case $z = 0$ corresponds to the Cambridge-Aachen jet finding.

It has been found that the anti-k_T variant produces clusters that are most resilient with respect to soft radiation [103].

The four-momenta of the jets are calculated as the sum of the four-momenta of their constituents, which are assumed to be massless. The momenta of jets are corrected for losses in passive material, the non-compensating response of the calorimeter, and contributions from pile-up [104].

As pile-up is a particular challenge for the reconstruction of jets from the hard scatter, dedicated algorithms exist to identify pile-up jets. One such technique used to suppress jets from pile-up events that is based on a multi-variable likelihood method is the *jet vertex tagger* (JVT) [105].

Heavy flavour tagging

Jets with heavy flavour content, most notably b-quarks, are very interesting for many physics analyses. These *b-jets* can be identified by virtue of the extremely precise vertexing. The mean life time of heavy hadrons is often sufficient to allow them to fly a considerable distance – often on the scale of millimeters – before decaying. If the vertexing can detect the distance between the decay vertex from which the jet then seems to originate as being different from the primary vertex, the jet can be tagged as a b-jet.

Very sophisticated multivariate techniques are used to perform the b-tagging, providing a variety of working points. These working points are typically denoted with some identifier `MV2c`X as well as the efficiency of the working point. The working point `MV2c10` 85% for example corresponds to a complicated multivariate algorithm trained on a sample with 10% c-jets with a threshold placed such that 85% of b-jets are successfully selected [86, 106].

Momentum and energy balance

As the LHC is a hadron collider, the remnants of the colliding hadrons often escape along the beam pipe. This part of the *underlying event* can carry away arbitrary amounts of longitudinal momentum. Thus, the longitudinal momenta of the detected objects in any given event are not necessarily balanced. Hence, longitudinal quantities such as p_Z are seldomly used. It is noteworthy, however, that rapidity differences Δy and approximately also pseudorapidity differences $\Delta\eta$ are Lorentz invariant with respect to boosts along the z-direction.

The transverse momentum balance is always well-defined, as the initial state particles do not carry any significant amount of transverse momentum. In fact, the vectorial sum of all transverse momenta should always equal zero, unless undetectable particles such as neutrinos escape the event. This yields an effective mechanism to measure invisible particles via *missing transverse momentum*. This quantity is a transverse vector and is fully determined by its magnitude and its azimuthal angle. However, there exist several different definitions for this quantity.

All of these use inner detector tracks with $|\eta| < 2.5$ that satisfy the following selection criteria:

- A transverse momentum of $p_{\mathrm{T}}^{\mathrm{track}} > 500\,\mathrm{MeV}$,
- at least 7 hits in the silicon semiconductor tracker,
- no more than two missing hits in the silicon layers or one missing hit in the pixel layers, and
- a tansverse impact parameter of $d_0 < 1.5\,\mathrm{mm}$, with an additional requirement of $d_0/\sigma_{d_0} < 3$.

The **track soft term missing transverse energy**, with magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ and angle $\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}$, is calculated as the negative vectorial sum of the transverse momenta of all calibrated selected objects, such as electrons, muons, and jets. This contrbution is generally referred to as the *hard term*. Tracks compatible with the primary vertex and not matched to any of those objects, comprising the *track soft term*, are also included in the reconstruction [107, 108].

A different definition, the **track-based missing transverse momentum** $p_{\mathrm{T}}^{\mathrm{miss}}$ [108], takes advantage of the ATLAS excellent vertex resolution to enhance the resilience with respect to pile-up. It is calculated as the negative vectorial sum of the transverse momenta of all inner detector tracks. Only for identified electrons or muons, which are more precisely measured using the calorimeter than using the track momentum, the calorimeter cluster measurement is used instead of the measured p_{T} of the track. On the down side, this definition fails to take into account contributions from neutral particles, which lack tracks in the inner detector. An additional shortcoming is the limited η coverage of $|\eta| < 2.5$, corresponding to the range of the inner detector.

Simulation

In order to compare the measured data with theory predictions, Monte Carlo samples are generated using the worldwide LHC grid. A wide variety of different Monte Carlo generators exists, specializing in different types of physics processes. These generated events are then subject to various processing steps, taking into account initial and final state radiation, hadronization, and showering. The interaction of the particles produced in the event with the detector material is simulated with a model of the ATLAS detector using the simulation software Geant4 [109]. Finally, the detector response to these simulated interactions is modeled in the digitization step, after which the reconstruction of simulated events proceeds equivalently to the reconstruction of data events.

3.3 The Standard Model at the LHC

In the years 2011–2016, the ATLAS and CMS experiments (and likewise the other LHC experiments) were able to record a remarkable amount of data for pp collisions. The discovery of the long-sought Higgs boson was the largest, but by far not the only striking success achieved by analysis of this data. A few selected results from the area of electroweak physics will be presented in this section to outline and highlight the achievements and performance of the experiments and to introduce existing measurements for important background processes relevant to the analyses presented in this thesis.

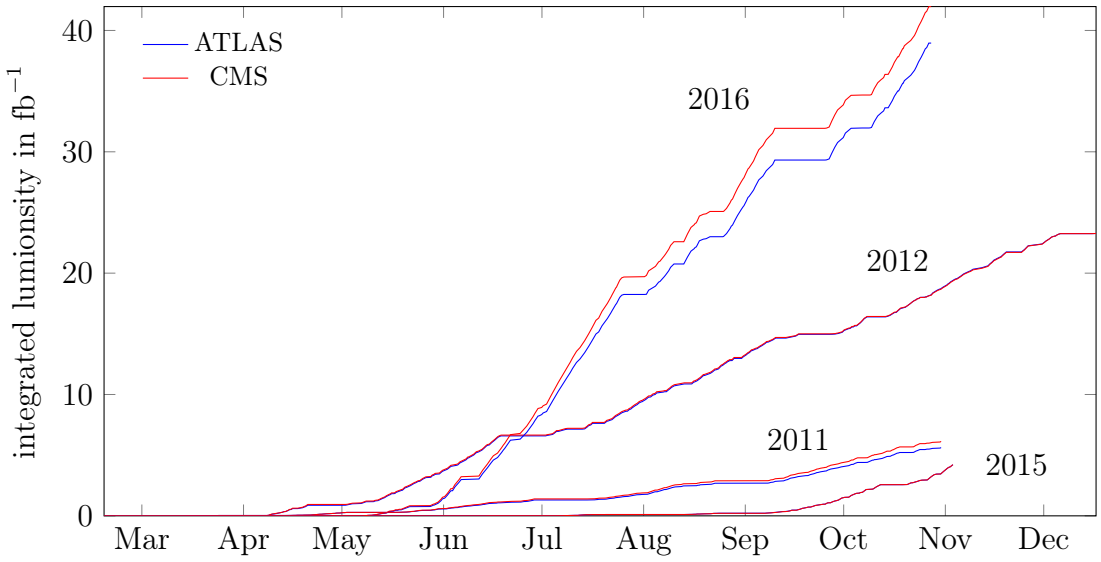


Figure 3.15.: LHC delivered luminosity to the ATLAS and CMS experiments by year.

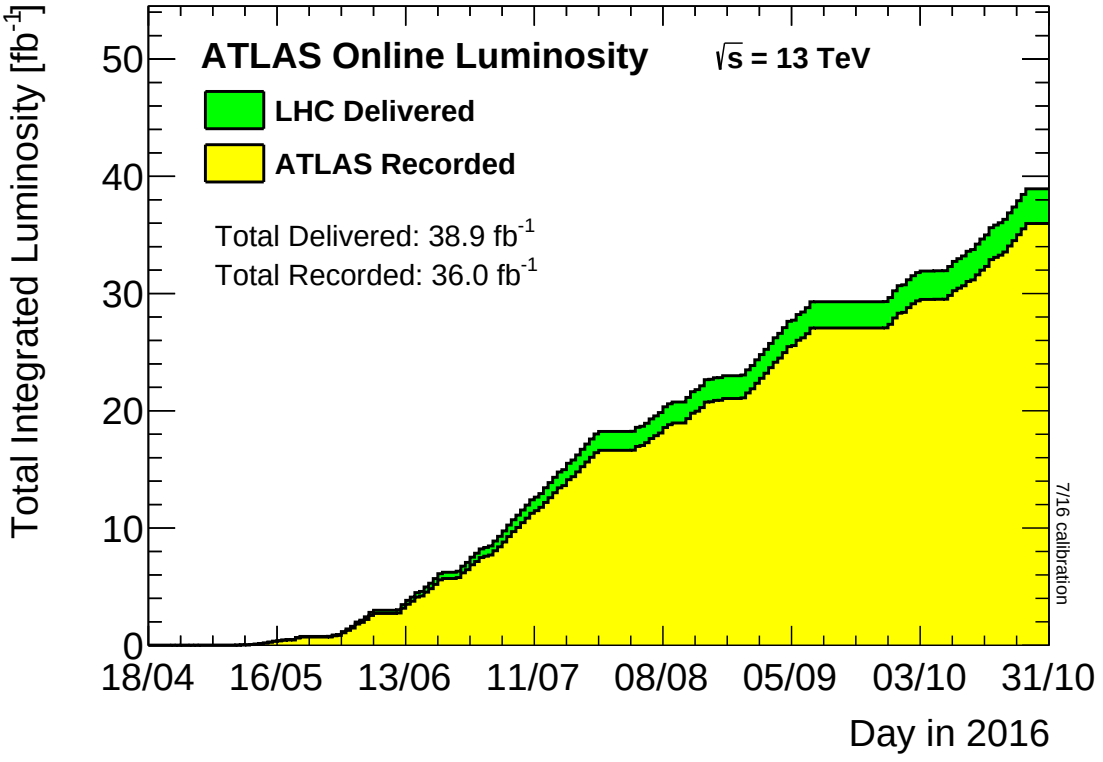


Figure 3.16.: ATLAS recorded luminosity during the year 2016 [93].

3.3.1 Data taking

The luminosity delivered by the LHC has been outstanding and quickly superseded the total integrated luminosity of $12\,\text{fb}^{-1}$ delivered by the TeVatron $p\bar{p}$ collider at $\sqrt{s} = 1.96\,\text{TeV}$. The development of the recorded integrated luminosity over time for the individual years of pp collision data taking is shown in Fig. 3.15 for the ATLAS and CMS experiments. A comparison of the luminosity recorded by the ATLAS experiment to the luminosity delivered to the corresponding interaction point by the LHC is shown in Fig. 3.16 for the 2016 data taking period. It is clearly visible that the performance of the LHC and the efficiency of the data taking are remarkable.

3.3.2 Standard Model measurements

With this data, the Standard Model has been tested to a remarkable precision. The precise measurement of a vast spectrum of Standard Model particles is a document of the tremendous success of the LHC and the experiments, but also substantiates the validity of the theory at collisions with unprecedented center-of-mass energies.

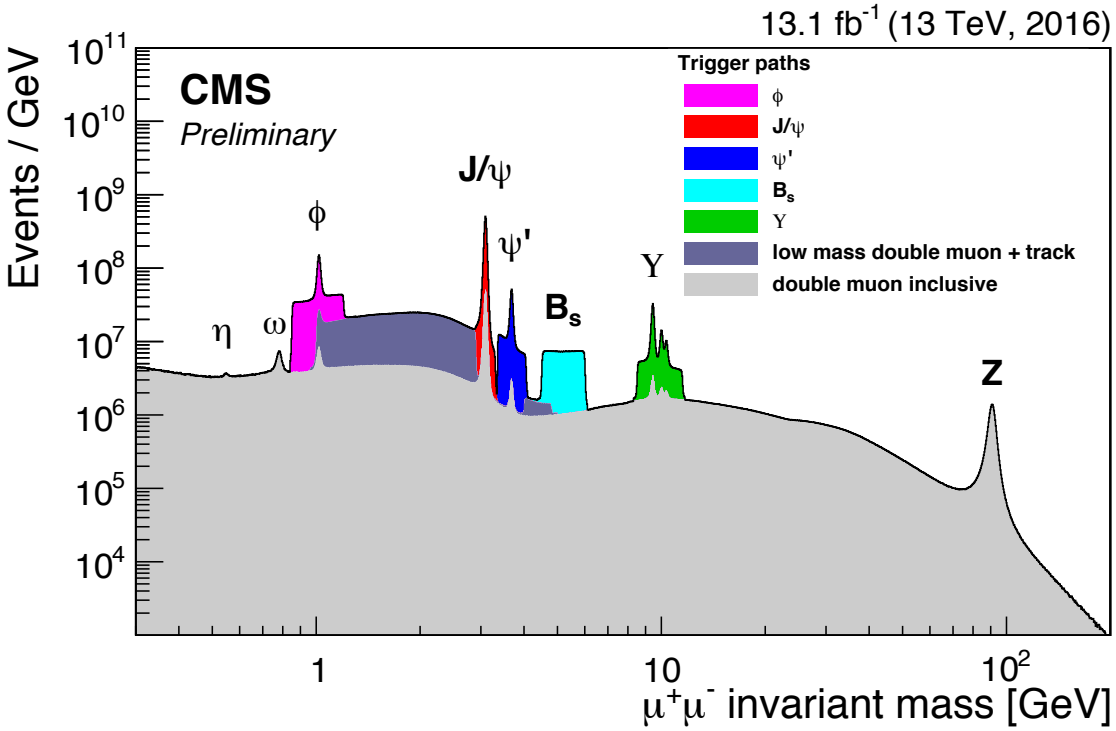


Figure 3.17.: Dimuon mass distribution collected with various dimuon triggers (represented by the different colours) in $13.1\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data collected in 2016 by the CMS experiment [110].

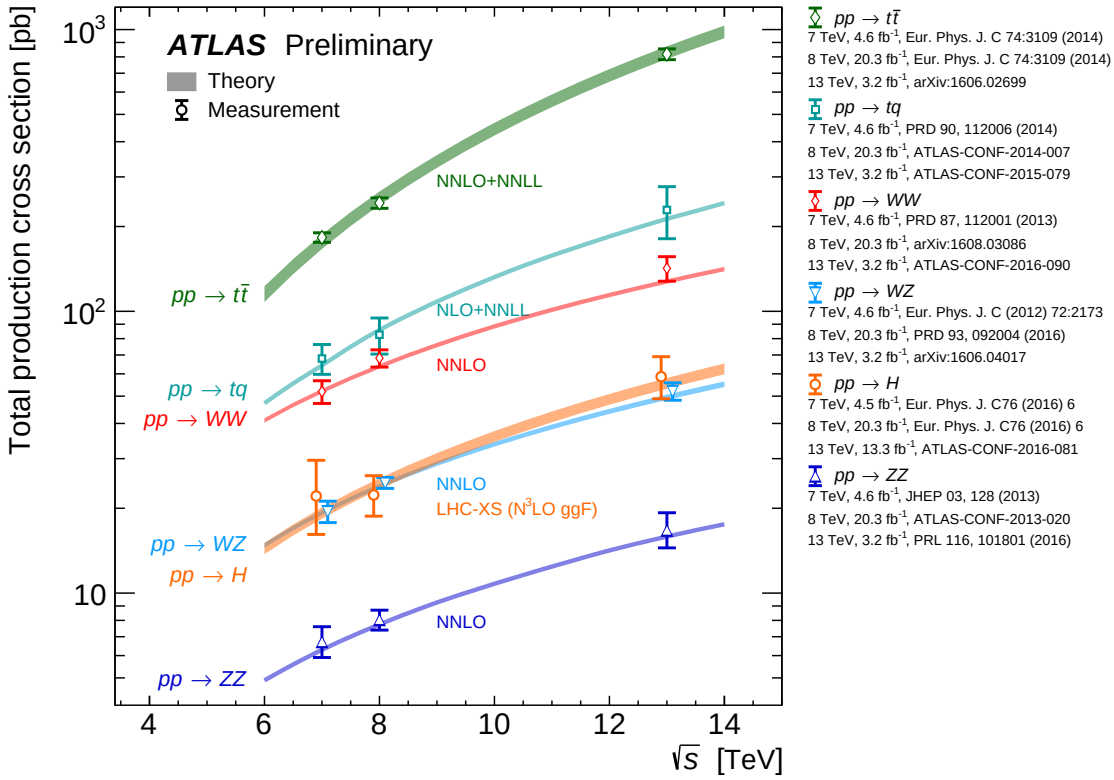


Figure 3.18.: Summary of total production cross-section measurements performed by the ATLAS experiment, presented as a function of center-of-mass energy $\sqrt{s}$ for 7, 8 and 13 TeV for a few selected processes. The line widths encode the corresponding uncertainties on the theory predictions. Some markers are displaced horizontally for better visibility. The legend on the right hand side also lists the publications in which the respective measurements have been published [111].

The number of $\mu^+\mu^-$ events recorded in pp collisions by the CMS experiment is shown in Fig. 3.17 as a function of the invariant mass of the di-muon system. Clearly visible are the localized peaks corresponding to the η, ω, ϕ, J/ψ, ψ', and Y hadronic resonances, as well as the Z boson peak. The colours indicate the contributions from events originating from trigger paths dedicated to specific measurements.

Figure 3.18 shows a summary of cross sections of different Standard Model processes measured by the ATLAS experiment for different values of $\sqrt{s}$, overlayed with the respective theory predictions. Shown are the inclusive production cross sections for $t\bar{t}$ production, tq production, WW, WZ and ZZ production as well as Higgs boson production. The analysis referenced for the $\sqrt{s} = 13\,\text{TeV}$ data point for Higgs boson production is presented in greater detail in Section 5.1.

In total, the agreement between the measurements and the respective theory predictions reveals the impressive performance of the LHC and the experiments it supports, as well as the predictive power of the Standard Model. So far, no measurement has shown a significant deviation from its predictions.

3.3.3 Higgs boson measurements

As the Higgs boson was discovered only recently, its properties are still under thorough investigation – a task which the LHC experiments ATLAS and CMS are extremely well prepared to take on. Measurements [15] including data corresponding to about $20\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$ as well as about $5\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$ find an impressive agreement with the predictions of the Standard Model also in the Higgs sector. An updated combination for $\sqrt{s} = 13\,\text{TeV}$, including only the $H{\to}ZZ^*{\to}\ell^+\ell^-\ell'^+\ell'^-$ and $H{\to}\gamma\gamma$ final states measured in data sets corresponding to $14.8\,\text{fb}^{-1}$ and $13.3\,\text{fb}^{-1}$, respectively, is presented in Section 5.1.

As displayed in Fig. 3.18, the cross section for Higgs boson production is much smaller than the cross section for the WW and $t\bar{t}$ processes. Factoring in the branching ratio of the Higgs boson to any specific final state (as displayed in Fig. 2.11 in Section 2.6.1), it is also much smaller than the cross section for ZZ production. These Standard Model processes act as backgrounds for many analyses targeting Higgs boson production. It is worthwhile to note that the cross sections of these processes increase at very different rates towards higher energies due to the changes in the parton density functions. While the ratio of cross sections between $\sqrt{s} = 13\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$ is approximately 2.5 for Higgs boson production [15,23], it is about 2.0 for WW pair production [112–114] and about 3.3 for $t\bar{t}$ production [115,116]. This causes not only the background composition, but also the signal-to-background ratio to change for many analyses. The implications of this observation for the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ analysis are discussed in greater detail in Chapter 4.

Figure 3.19 shows the combined measurements from the ATLAS and CMS experiments for $20\,\text{fb}^{-1}$ of data recorded at $\sqrt{s} = 8\,\text{TeV}$ and $5\,\text{fb}^{-1}$ of data recorded at $\sqrt{s} = 7\,\text{TeV}$ in terms of signal strengths μ, disentangled for all production and decay processes. While all values are in good agreement to the Standard Model predictions, the measurements in the $H \to bb$ channel seem to find a somewhat smaller signal strength than expected – a trend that continues in the first $\sqrt{s} = 13\,\text{TeV}$ measurement performed by the ATLAS collaboration [53].

The coupling modifiers κ, introduced in Section 2.4.2, quantify the coupling strength of the Higgs boson to individual particles. If the Higgs boson observed at $m_H \approx 125\,\text{GeV}$ is indeed responsible for the BEH mechanism of electroweak symmetry breaking, the coupling of the Higgs boson to any one particle must be

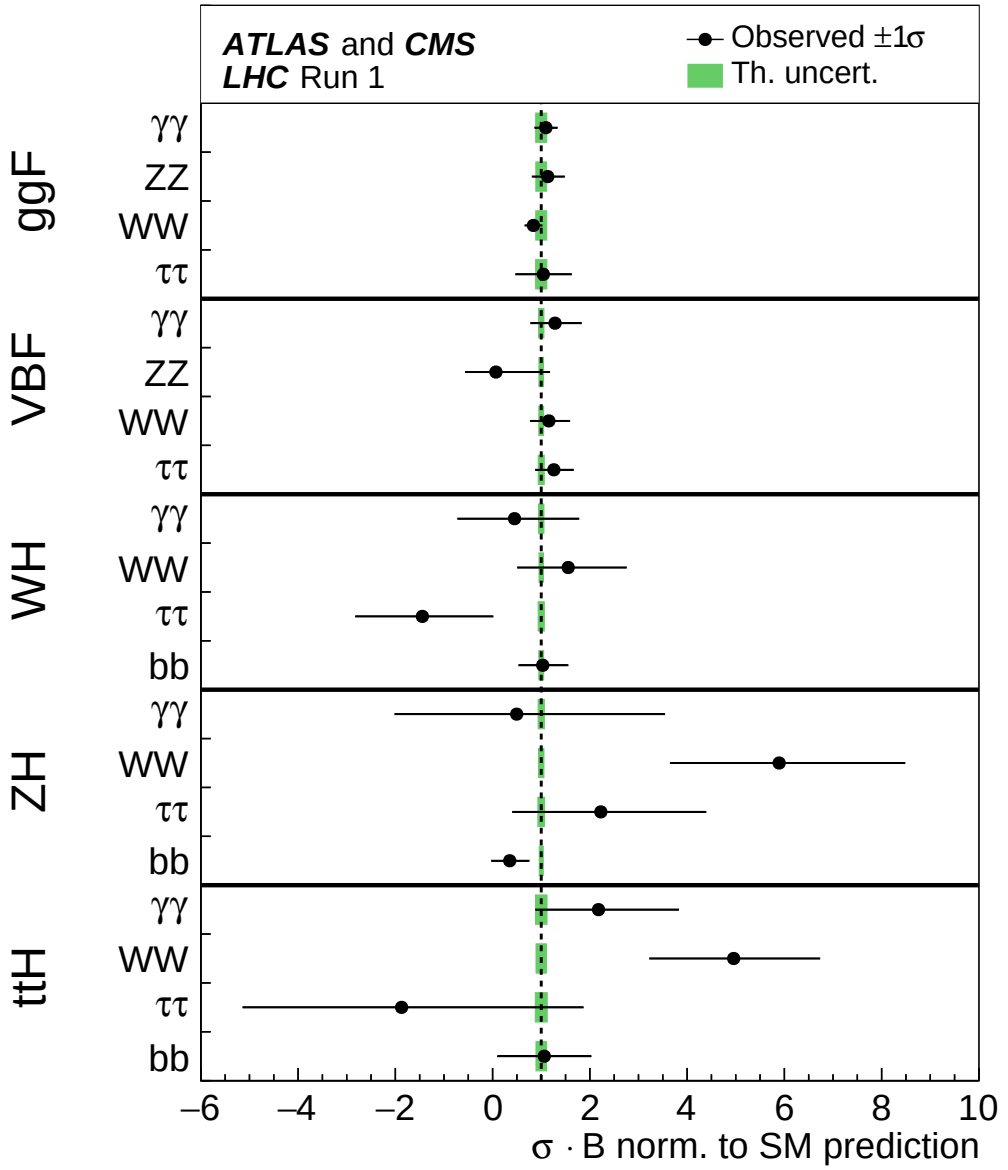


Figure 3.19.: Best fit values of $\sigma_i \cdot \mathcal{B}_f$ for each specific channel $i \to H \to f$ for the combination of the results published by the ATLAS and CMS collaborations. The error bars indicate the 1σ intervals. The fit results are normalized to the SM predictions for the various parameters and the shaded bands indicate the theoretical uncertainties in these predictions. Only parameters measured to meaningful precision are shown [15].

proportional to its mass. This prediction is tested in Fig. 3.20. The uncertainties are still quite large, especially for the measurement of the small $H \to \mu\mu$ coupling, which currently only constitutes an upper limit. So far, no significant deviation from the Standard Model prediction could be observed in Run 1 data. Also shown is a fit to the $[M, \epsilon]$ phenomenological model [117], where the coupling modifiers κ_V and κ_f are parametrized as

$$\kappa_V = v\frac{m_V^{2\epsilon}}{M^{1+2\epsilon}} \tag{3.3.1}$$

$$\kappa_f = v\frac{m_f^{\epsilon}}{M^{1+\epsilon}}. \tag{3.3.2}$$

Here, v is the Higgs field vacuum expectation value, m_V and m_f are the masses of the corresponding particles, and M and ϵ are the fit parameters. The Standard Model prediction corresponds to the values $M = v$ and $\epsilon = 0$, such that $\kappa_V = \kappa_f = 1$.

In summary, the LHC experiments have met and greatly exceeded their design performances, the LHC itself providing collisions at center-of-mass energies close to the intended design value of $\sqrt{s} = 14\,\text{TeV}$, and have accumulated an impressive amount of evidence fortifying the foundations of the Standard Model. No significant deviations from the Standard Model predictions have been found. However, the collection of new data in the years 2015 and 2016 signifies a substantial increase in the size of the available data set and provides new insights in the physics at unprecedented center-of-mass energies. The remainder of this thesis is dedicated to a detailed analysis of this data and the search for effects beyond the Standard Model in the Higgs sector.

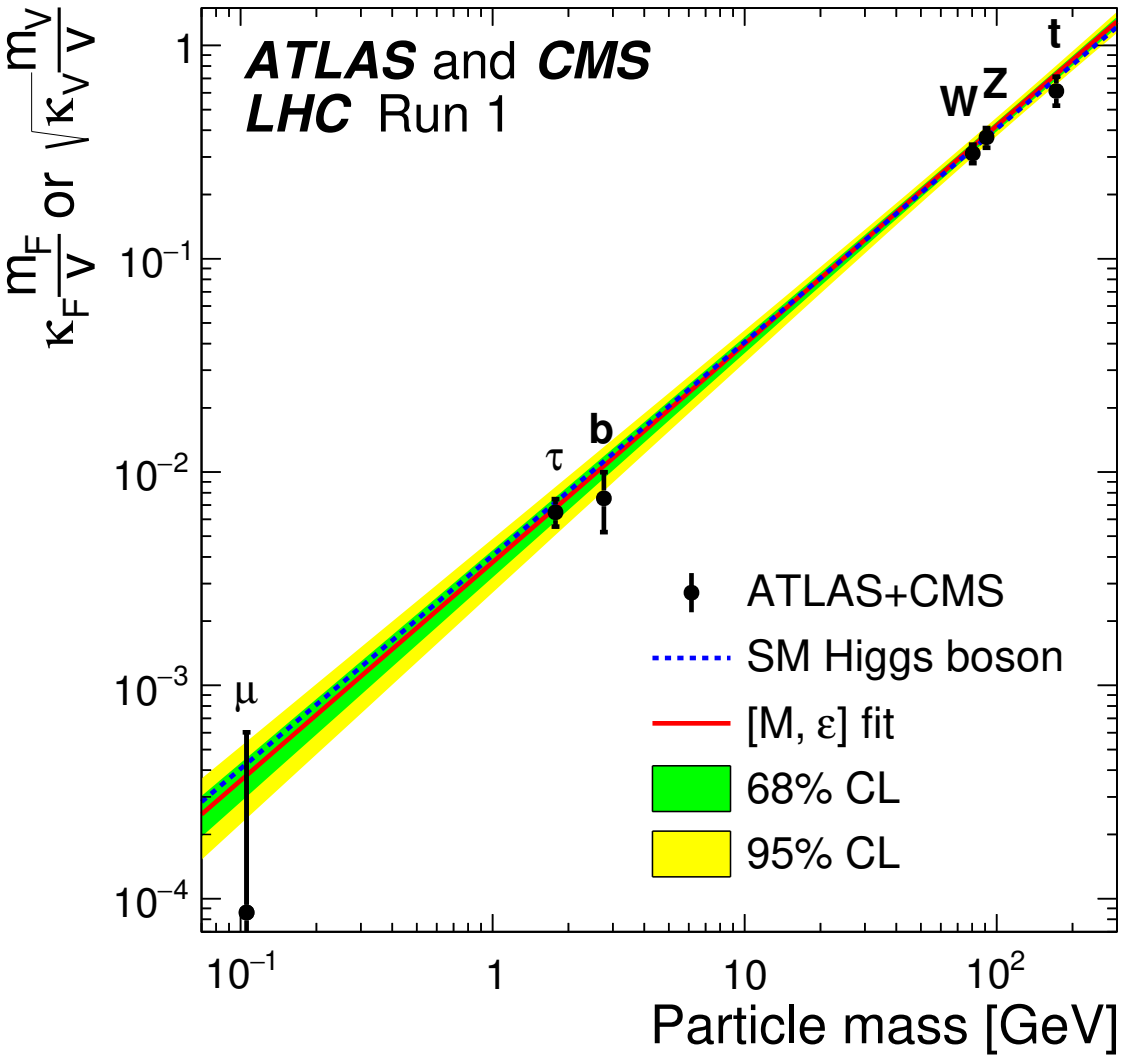


Figure 3.20.: Best fit values as a function of particle mass for the combination of ATLAS and CMS measurements, employing the coupling modifiers κ introduced in Section 2.4.2 and the vacuum expectation value $v = 246\,\text{GeV}$ of the Higgs field. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The solid (red) curve indicates the best fit result to the $[M, \epsilon]$ phenomenological model [117] with the corresponding 68% and 95% confidence level bands in green and yellow [15].

4 $H \to W^{\pm} W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$

> “Physics is actually too hard for physicists.”
>
> *David Hilbert*

With the Higgs boson mass known as $125.09 \pm 0.24\,\text{GeV}$ [35], precise predictions are possible for production cross sections and decay branching fractions into Standard Model particles. Measurements in different channels, differentiating in production as well as in decay mode, provide comprehensive insights into the physics of the Higgs boson sector and the mechanism of electroweak symmetry breaking as outlined in Chapter 2. As detailed in Section 2.6, the $H \to W^{\pm} W^{\mp *}$ decay mode has the second highest branching fraction, featuring 22.9% of all Higgs boson decays [45].

The only decay channel with a branching fraction exceeding that of $H \to W^{\pm} W^{\mp *}$ is the decay to a pair bottom quarks. Due to the purely hadronic final state, identification of this type of event is extraordinarily difficult. Existing analyses rely on the specific signatures of individual production modes like VH or VBF to measure this process [53], while the events stemming from the dominant ggF production mode are too difficult to disentangle from the multijet background. This constraint does not exist for events decaying via $H \to W^{\pm} W^{\mp *}$, making this decay channel together with the $H \to ZZ^{*}$ and $H \to \gamma\gamma$ decay modes a prime candidate for measuring Higgs boson production via the dominant gluon fusion process.

As W bosons are very short-lived, only their decay products can be measured. The branching fractions of all relevant decay modes are displayed in Fig. 4.1. Notably, the total branching fraction to leptons is $\mathcal{B}_{W \to e\nu/\mu\nu/\tau\nu} = 32.4\%$ [31]. However, while electrons are stable and muons are at least sufficiently stable to be measured by the detector directly, τ leptons generally decay further. Their decay modes are displayed in Fig. 4.2, the total branching fraction to lighter leptons being roughly $\mathcal{B}_{\tau \to e\nu/\mu\nu} = 35.2\%$ [31]. It is for this reason that the lepton symbol ℓ is taken to mean *light lepton*, that is, electrons e or muons μ, in the context of this analysis. The decays $W \to \tau\nu_{\tau} \to \ell\nu_{\tau}\nu_{\tau}\nu_{\ell}$ are thus included implicitly.

The total chances of finding a Higgs boson decay via a pair of W bosons to a double lepton, single lepton or fully hadronic final state can be calculated easily. They are displayed in Fig. 4.3.

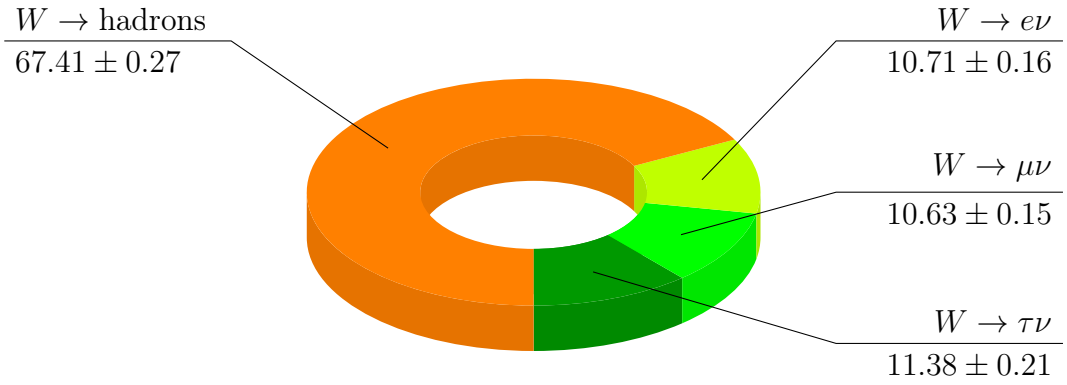


Figure 4.1.: Measured decay branching fractions of W bosons in percent with their respective uncertainties [31].

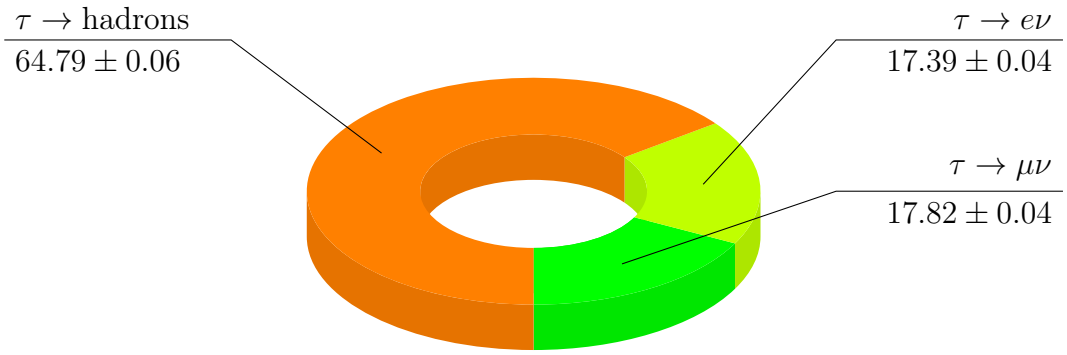


Figure 4.2.: Measured decay branching fractions of τ leptons in percent with their respective uncertainties [31].

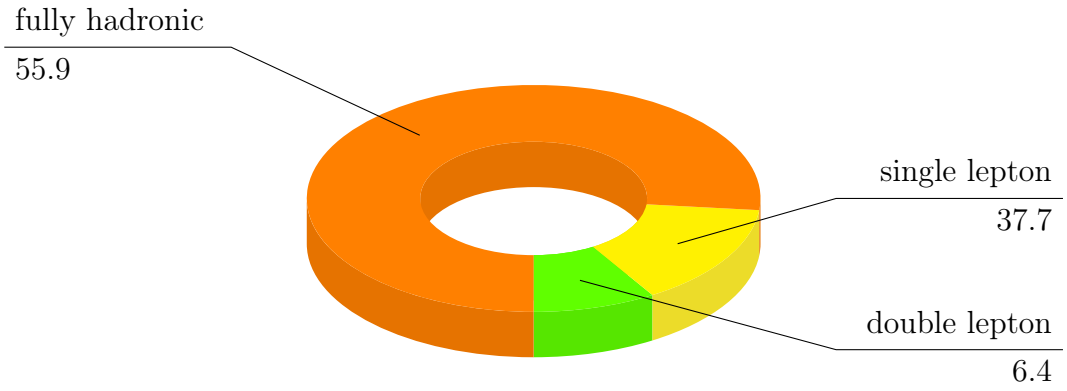


Figure 4.3.: Approximate effective decay branching fractions of W boson pairs by number of primary light leptons in the final state, including indirect decays via $W \to \tau\nu_{\tau} \to X$. Values are given in percent.

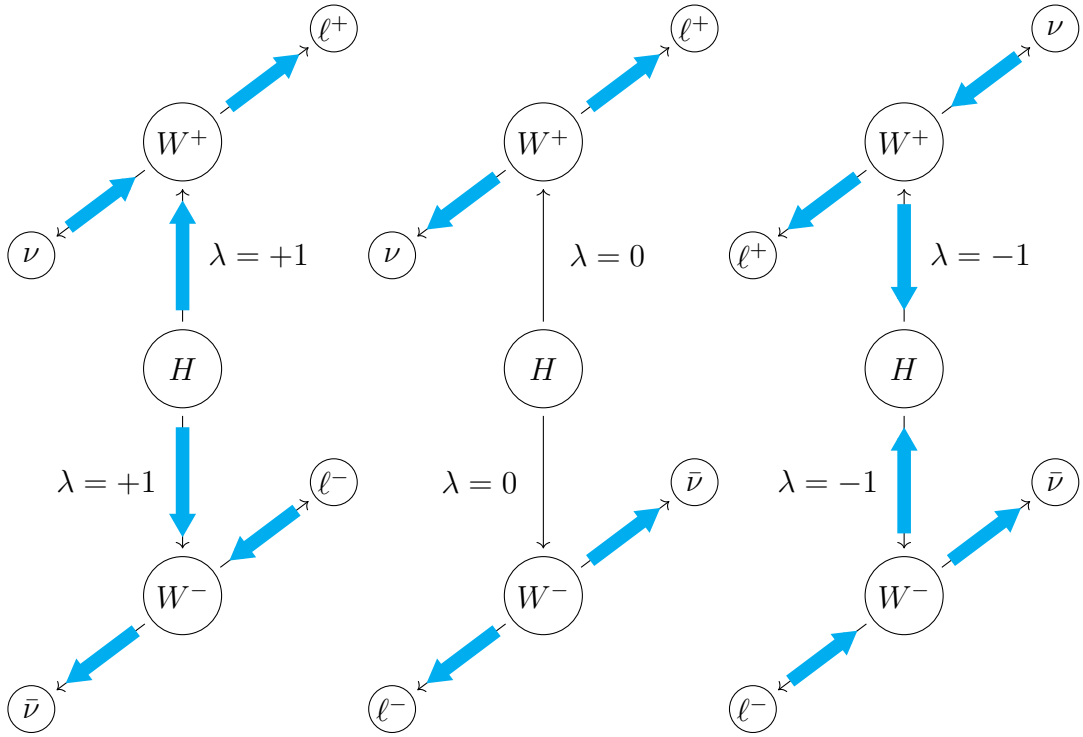


Figure 4.4.: Spin correlations in the $H \to W^{\pm} W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ decay enhance the fraction of events where the two charged leptons are emitted in one and the two neutrinos in the opposite hemisphere. The left (right) case exhibits a positive (negative) helicity λ, as spin ▲ and direction of flight ↑ are parallel (antiparallel), resulting in the prescribed topology. Only the third (middle) case is isotropic.

The presence of two highly energetic and oppositely charged leptons provides a powerful handle to select events using dedicated lepton triggers and reject hadronic background from various sources. The strong background rejection however comes at the price of sacrificing a significant amount of statistical power, which is suitable for precision measurements of the Standard Model Higgs boson. The one-lepton final state on the other hand provides more statistical power and a weaker background rejection, and has proven useful for the search for high-mass resonances such as a second, heavier Higgs boson or new heavy vector bosons [118].

For the analyses featuring two leptons in the final state, it is customary to enumerate the leptons in the order of their transverse momenta p_{T}^{ℓ}, with ℓ_1 denoting the lepton with the highest or "leading" transverse momentum, and ℓ_2 denoting the lepton with the second highest or "subleading" transverse momentum. The analysis thus differentiates between four lepton flavour configurations,

- the two same-flavour (SF) configurations ee and $\mu\mu$, and

- the two different-flavour (DF) configurations $e\mu$ and μe,

where the latter two are only distinguished by the p_{T} ordering of the leptons. As events with two same-flavour, opposite-charge leptons can easily arise from pair production via a Z or γ boson, the different-flavour mode provides a much higher sensitivity and is thus focused on the analysis of early Run 2 data [21], whereas an analysis of the same-flavour mode has only been conducted on Run 1 data so far [119].

Additional spin and kinematic properties of the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay can be exploited to enhance the separation of the signal from the background. Figure 4.4 illustrates how spin correlations induce an enhanced fraction of events where the two charged leptons are emitted in the same hemisphere, whereas the two neutrinos are emitted in the opposite one. This allows to exploit observables such as the dilepton invariant mass $m_{\ell\ell}$, the transverse momentum of the dilepton system $p_{\mathrm{T}}^{\ell\ell}$, the azimuthal angle between the two leptons $\Delta\phi_{\ell\ell}$ and the missing transverse momentum to discriminate between signal and background events, as many Standard Model processes do not exhibit the same characteristic.

One can further differentiate and categorize the events by additional leptons and jets recorded in the final state to distinguish between the different production mechanisms ggF, VBF and VH.

This chapter details a measurement of Higgs boson production and decay to a pair of W bosons, employing dedicated event selections for all three of these production mechanisms. However, as the VBF and VH measurements have recently been published [21], special attention is devoted to the analysis dedicated to the ggF production mode. For this specific channel, there is no published measurement at $\sqrt{s} = 13\,\mathrm{TeV}$ yet, as latest measurements still only use the $\sqrt{s} = 7$ and $8\,\mathrm{TeV}$ data recorded by the LHC in the years 2011-2012 [119].

4.1 Data and Monte Carlo Samples

These analyses use a data set corresponding to an integrated luminosity of $5.8\,\mathrm{fb}^{-1}$ of pp collision data recorded at $\sqrt{s} = 13\,\mathrm{TeV}$ by the ATLAS detector in the year 2015 as well as the first months of 2016 with a bunch spacing of 25 ns. The triggers and associated transverse momentum requirements used to select data events for this analysis are listed in Table 4.1. All three analyses, dedicated to the three dominant Higgs boson production modes ggF, VBF and VH, use the same set of single-lepton triggers.

The lepton trigger efficiencies are measured using leptonic decays of Z bosons as a function of p_{T} and η of the lepton. The single-lepton trigger efficiencies are approximately 70% for muons with $|\eta| < 1.05$, 90% for muons in the range

$1.05 < |\eta| < 2.40$, and exceed 90% for electrons in the range $|\eta| < 2.40$. They are shown in Fig. 4.5 for the lepton triggers with the lowest p_{T} thresholds used in this analysis.

Table 4.1.: Minimum p_{T} requirements used at the different trigger levels, with values given in GeV. Letters "T", "M" and "L" next to the threshold value stand for the `tight`, `medium` and `loose` electron identification requirement, respectively. The letter "i" indicates an isolation requirement, which helps to improve the rejection power against non-prompt leptons (see Section 4.3) while maintaining a high efficiency.

Year	Lepton	Level-1	High-level
2015	e	20	24M $\wedge$ 60M $\wedge$ 120L
	μ	15	20i $\wedge$ 50
2016	e	20	24Ti $\wedge$ 60M $\wedge$ 120L
	μ	15	24i $\wedge$ 50

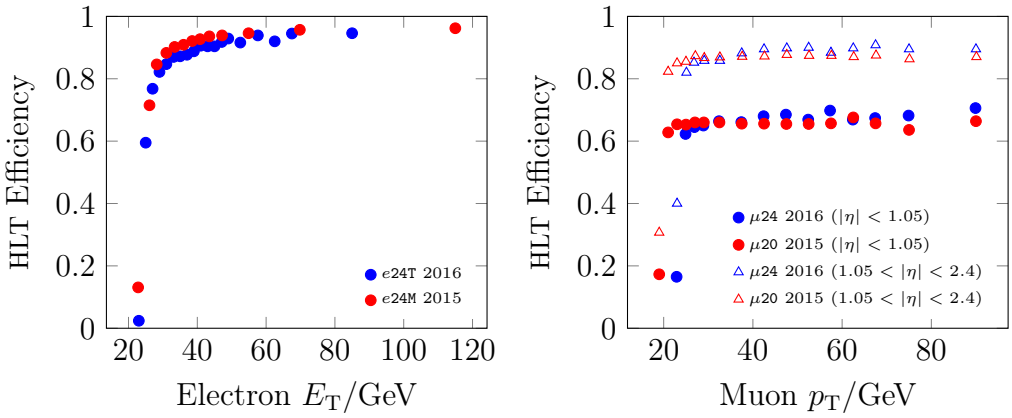


Figure 4.5.: Trigger efficiencies measured in the years 2015 and 2016 for the high-level triggers (HLT) with the lowest p_{T}-thresholds used for single leptons used in this analysis [120, 121]. Numbers in the legend indicate the minimum p_{T} requirement of the trigger, with values given in GeV. The letters "T", "M" and "L" next to the threshold value stand for the `tight`, `medium` and `loose` electron identification requirement, respectively.

Signal

The signal contributions considered for this analysis include the dominant gluon fusion production process (ggF), the vector boson fusion production process (VBF), and the Higgs-strahlung process from a W boson (WH). Contributions from the Higgs-strahlung process from a Z boson are included in the analysis as a background.

Other Higgs boson production processes like ttH, bbH and tHx are neglected because their contributions are small, whereas a small contamination of $H \to \tau\tau$ decays has is considered as a background. Only the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay mode is considered as signal, with final states featuring two charged leptons.

The signal cross section is computed to NNNLO accuracy in QCD [122] for the ggF process, with additional NLO EW corrections being applied. Full NNLO QCD and NLO EW corrections are used to calculate the cross sections for the VBF signal production. The cross sections of the associated WH/ZH production processes are calculated up to NNLO QCD corrections and NLO EW corrections, while the cross section for $gg \to ZH$ is calculated at NLO QCD precision. The numerical values and uncertainties are summarized in Table 4.2.

Background

The main sources of Standard Model backgrounds to the analyses presented here include events from the production of top-quarks, pairs and triplets of electroweak bosons, W or Z bosons produced in association with hadronic jets, and pure QCD multijet events.

The cross section of the inclusive WW background process is known at NNLO accuracy [123] and includes contributions from $q\bar{q}/g \to WW$ and $gg \to WW$. Since the processes $q\bar{q}/g \to WW$ and $gg \to WW$ are modeled with different MC samples and a higher order calculation is available for $gg \to WW$, cross sections for both processes are handled separately. In the inclusive prediction, $gg \to WW$ is only calculated to LO precision and does not interfere with $q\bar{q}/g \to WW$ production. Therefore the $gg \to WW$ contribution to the inclusive cross section is subtracted, yielding $\sigma_{q\bar{q}/g \to WW} = 111.9\,\text{pb}^{+2.6\%}_{-2.2\%}$. The non-resonant $gg \to WW$ production cross section has been calculated at NLO accuracy [124] and is $\sigma_{gg \to WW} = 6.82\,\text{pb}^{+5.3\%}_{-7.4\%}$.

Monte Carlo generators

The Monte Carlo generators used to model signal and background processes are listed in Table 4.3. For most processes, separate programs are used to generate the

Table 4.2.: Predicted signal process cross sections. The uncertainties quoted here correspond to standard Gaussian 68% confidence levels. However, no signal cross section uncertainty is applied in this analysis, as its effect is negligible with respect to the total uncertainty.

process	cross section in pb
ggF	$48.58 \pm 3.9\%\,(\text{QCD scale}) \pm +3.2\%\,(\text{PDF} + \alpha_{\text{S}})$
VBF	$3.782 \pm 0.2\%(\text{QCD scale}) \pm 2.1\%\,(\text{PDF} + \alpha_{\text{S}})$
WH	$1.373 \pm 0.4\%\,(\text{QCD scale}) \pm 1.9\%\,(\text{PDF} + \alpha_{\text{S}})$
ZH	$0.8839^{\pm}2.2\%(\text{QCD scale}) \pm 1.6\%\,(\text{PDF} + \alpha_{\text{S}})$

hard scattering process and to model the parton showering (PS), hadronization, and the underlying event (UE). PYTHIA8.210, PYTHIA8.186 [125] or PYTHIA6.428 [126] are used for the latter three steps for the signal and some of the background processes. HERWIG [127] is used for PS and hadronization in estimating systematic uncertainties with UE being modeled with JIMMY [128]. The CT10 PDFs are used for the hard scattering process in POWHEG-BOX v2 [129], while for MADGRAPH5 (version 2.2.1 and 2.2.2) [130] the NNPDF23LO [131] PDF set is used. The CTEQ6L1 [132] PDF set is used for the PYTHIA showering, with the AZNLO [133] tune when interfaced to POWHEG-BOX v2 and with the A14 [134] tune when interfaced to MADGRAPH5. The NLO prediction from SHERPA2.1.1 [135] is calculated using CT10 PDF set in conjunction with a dedicated set of tuned parameters from the parton shower developed by the SHERPA authors [136].

Table 4.3.: Predictions used for estimation of signal and background in the analyses, computed for the Standard Model including a Higgs boson with $m_H = 125\,\text{GeV}$. Listed are the Monte Carlo generators used to model the kinematic properties of the events and the corresponding Standard Model cross section predictions obtained from theory as well as the precision level of the underlying calculation. The last column identifies which analyses use the corresponding samples, with an empty cell denoting that the sample is used by all analyses.

Process	Generator	$(\sigma\mathcal{B})$ (pb)	Precision $\sigma_{\text{incl.}}$	analysis
ggF, $H \to WW$	POWHEG +PYTHIA8	10.4	NNNLO	
VBF, $H \to WW$	POWHEG +PYTHIA8	0.808	NNLO	
WH, $H \to WW$	POWHEG +PYTHIA8 (MiNLO [137])	0.293	NNLO	
ZH, $H \to WW$	POWHEG +PYTHIA8 (MiNLO [137])	0.189	NNLO	
inclusive $W \to \ell\nu$	POWHEG +PYTHIA8	6.02×10^4	NNLO	WH
inclusive $Z/\gamma^\star \to \ell\ell$ ($m_{\ell\ell} \geq 40\,\text{GeV}$)	MADGRAPH	6.04×10^3	NNLO	
inclusive $Z/\gamma^\star \to \ell\ell$ ($40 \geq m_{\ell\ell} \geq 10\,\text{GeV}$)	MADGRAPH	8.01×10^3	NNLO	
$(W \to \ell\nu)\gamma$ ($p_T^\gamma > 10\,\text{GeV}$)	SHERPA	453	LO	
$(Z \to \ell\ell)\gamma$ ($p_T^\gamma > 10\,\text{GeV}$)	SHERPA	175	LO	
$t\bar{t}$ leptonic	POWHEG +PYTHIA6	87.6	NNLO+NNLL	
Wt leptonic	POWHEG +PYTHIA6	7.55	NLO	
$t\bar{t}W/Z$	MADGRAPH	0.62	NLO	WH
tZ non-all-hadronic	MADGRAPH	0.24	LO	WH
$q\bar{q}/g \to WW \to \ell\nu\ell\nu$	POWHEG +PYTHIA8	11.2	NNLO	ggF, VBF
$Z^{(*)}Z^{(*)} \to 2\ell 2\nu$ ($m_{\ell\ell} \geq 4\,\text{GeV}$)	POWHEG +PYTHIA8	0.925	NLO	VBF
$q\bar{q}/g \to 2\ell 2\nu$	SHERPA	12.8	NLO	WH
$gg \to 2\ell 2\nu$	SHERPA	0.72	NLO	
$q\bar{q}/g \to \ell\nu\ell\ell$	SHERPA	11.9	NLO	
$q\bar{q}/g, gg \to \ell\ell\ell\ell$	SHERPA	11.5	NLO	
$WZ \to qq\ell\ell$	SHERPA	3.76	NLO	WH
$ZZ \to qq\ell\ell$	SHERPA	2.36	NLO	WH
EW $WW + 2$ jets ($\ell\nu\ell\nu$)	SHERPA	0.012	LO	VBF
EW $WZ + 2$ jets ($\ell\nu\ell\ell$)	SHERPA	0.038	LO	WH
EW $ZZ + 2$ jets ($\ell\ell\ell\ell$)	SHERPA	0.116	LO	WH
EW $q\bar{q} \to (Z \to \tau\tau)q\bar{q}$	SHERPA	2.54	LO	VBF
WWW ($3\ell 3\nu$)	SHERPA	8.34×10^{-3}	NLO	WH
WWZ ($4\ell 2\nu, 2\ell 4\nu$)	SHERPA	5.16×10^{-3}	NLO	WH
WZZ ($5\ell 1\nu, 3\ell 3\nu$)	SHERPA	1.07×10^{-3}	NLO	WH
ZZZ ($6\ell 0\nu, 4\ell 2\nu, 2\ell 4\nu$)	SHERPA	3.16×10^{-4}	NLO	WH

The $q\bar{q}$-initiated diboson processes are modeled either with POWHEG-BOX v2+PYTHIA8.210 or SHERPA2.1.1. The VBF analysis uses POWHEG-BOX v2+PYTHIA8.210 for $q\bar{q}/g \to \ell\nu\ell\nu$ since it provides a better modeling of the observed transverse momentum imbalance distribution. For the other processes, the VBF analysis uses SHERPA2.1.1. The WH analysis uses SHERPA2.1.1 as the default generator for all the diboson processes, as it includes the single-resonant $Z^{(*)}Z^{(*)}$ process, which is the dominant $Z^{(*)}Z^{(*)}$ component in the analysis, the observed difference being covered by a designated uncertainty. Contributions to WW production from double-parton scattering ($q\bar{q} \to W + q\bar{q} \to W$) are not included, but are not expected to contribute significantly.

The loop-induced gg-initiated diboson processes are simulated by SHERPA2.1.1 with less than two additional jets and normalized to the inclusive NLO predicted cross section [124]. For $q\bar{q}/g \to Z^{(*)}Z^{(*)} \to \ell\nu\ell\nu$ by POWHEG-BOX v2, a requirement on the invariant mass of the two charged leptons $m_{\ell\ell} > 4\,\text{GeV}$ is imposed. For all $q\bar{q}/g$ initiated diboson processes generated by SHERPA2.1.1, at least two charged leptons with $p_{\text{T}} > 5\,\text{GeV}$ are required, additionally imposing $m_{\ell\ell} > 2m_{\ell} + 250\,\text{MeV}$ for any same-flavour, opposite-charge lepton pairs, where m_{ℓ} is the mass of the charged lepton. For gg initiated diboson processes, $m_{\ell\ell} > 2\,\text{GeV}$ is required for $\ell\nu\ell\nu$, and $m_{\ell\ell} > 10\,\text{GeV}$ for $\ell\ell\ell\ell$ final states. SHERPA2.1.1 is also used for the modeling of diboson process at LO with no $\mathcal{O}(\alpha_S)$ terms for the $\ell\ell\ell\ell$, $\ell\nu\ell\ell$ and $\ell\nu\ell\nu$ plus two jets final states as well as for the $q\bar{q} \to Zq\bar{q}$ processes, requiring a decay via $Z \to \tau\tau$ by demanding an invariant mass of the τ lepton pair of $m_{\tau\tau} > 40\,\text{GeV}$.

Samples modeling the production of top quarks, either as $t\bar{t}$ pairs or singular as Wt, are generated with POWHEG-BOX v2.0 interfaced to PYTHIA6.428 for parton showering, using the Perugia 2012 [138] tune. EVTGEN 1.2.0 [139] is used to model properties of the b- and c-hadron decays. The predicted $t\bar{t}$ production cross section is calculated with TOP++ 2.0 to NNLO precision in perturbative QCD, including soft-gluon resummation to NNLL order [140], and assuming a top-quark mass of 172.5 GeV. Both the $t\bar{t}$ and the Wt samples are required to have at least two charged leptons in the final state. Other top quark processes such as $t\bar{t}W/Z$ and tZ are generated at LO using MADGRAPH5 version 2.2.2 interfaced with PYTHIA8.186 ($t\bar{t}W/Z$) and MADGRAPH5 version 2.2.1 interfaced with PYTHIA6.428 (tZ). Events from the tZ samples are required to have three charged leptons in the final state, in which the lepton pair from the Z boson decay is required to have $m_{\ell\ell} > 10\,\text{GeV}$.

Events with $Z\gamma$ and $W\gamma$ as well as the triboson backgrounds WWW, WWZ, ZZW, and ZZZ with up to six genuine leptons in the final state are modeled using SHERPA2.1.1 at NLO accuracy. For $Z\gamma$ and $W\gamma$ processes, the p_{T} of the photon is required to be larger than 10 GeV, additionally requiring the distance to any selected lepton in the $\eta - \phi$ plane to be $\Delta R > 0.1$. The lepton pair originating from the Z boson in the $Z\gamma$ final state is required to have $m_{\ell\ell} > 2\,\text{GeV}$.

The production of Z bosons in association with jets for the invariant mass range of the two charged leptons $m_{\ell\ell} > 10\,\text{GeV}$ is modeled by MadGraph5 version 2.2.2 at LO interfaced with Pythia8.186.

The event overlap between $Z\gamma$ and Z+jets samples is removed by rejecting events from the Z+jets sample that contain a photon with $p_{\text{T}} > 10\,\text{GeV}$ and $|\eta| < 2.5$ radiated by a lepton in the final state.

The modeling of the W+jets processes is based on a purely data driven method described in Section 4.3. However, this background contribution is neglected in the WH analysis.

The effect of additional pp interactions in local or temporal proximity to the primary vertex (pile-up) is included by overlaying each signal or background event with additional events obtained from Pythia8 employing minimum-bias requirements. The number of overlaid events is chosen to reflect the conditions in the recorded data, that is, an average number of interactions per bunch crossing of $\widehat{\mu} = 13$ for 2015 and $\widehat{\mu} = 21$ for 2016. A correction to the actual pile-up conditions observed in data is performed using an event-by-event reweighting. All generated events are subject to a Geant4-based detector simulation [109, 141], and the standard ATLAS reconstruction for collision data is used.

4.2 Object Identification and Selection

Events are required to have at least one primary vertex with no less than two associated tracks with transverse momenta $p_{\text{T}} > 400\,\text{MeV}$. If more than one such vertex is reconstructed in any single event, the one with the largest sum of square transverse track momenta is selected.

4.2.1 Electrons and muons

The electrons used in these analyses are required to have a transverse energy exceeding $E_{\text{T}} > 15\,\text{GeV}$, and pass the `MediumLH` or `TightLH` selection defined in Section 3.2.5 for electrons with a transverse energy greater or smaller than 25 GeV, respectively. The pseudorapidity is required to be within the range of $|\eta| < 2.47$, excluding the transition region $1.37 < |\eta| < 1.52$ between the barrel and end-caps in the liquid argon calorimeter.

The muons are obtained via the *combined muon* definition, as introduced in Section 3.2.5. The muon candidates used in these analysis are required to pass the `Medium` selection and to satisfy $p_{\text{T}} > 15\,\text{GeV}$ and $|\eta| < 2.5$.

All leptons are required to originate from the primary vertex by requiring the absolute value of the longitudinal impact parameter to satisfy $|z_0 \sin\theta| < 0.5\,\text{mm}$.

Furthermore, the significance of the transverse impact parameter is required to be less than $|d_0|/\sigma_{d_0} < 5$ ($|d_0|/\sigma_{d_0} < 3$) for electrons (muons). The isolation criteria `Gradient` and `GradientLoose` as defined in Section 3.2.5 are used. At least one of the lepton candidates is required to match a trigger object for the event to be selected.

Table 4.4 summarizes the lepton requirements in each analysis. The ggF and VBF analyses require exactly one electron and one muon in an event, whereas the WH analysis requires three leptons regardless of their flavour. In order to ensure the mutual orthogonality between the analyses targeting final states with different numbers of leptons, a *veto lepton* definition is chosen. Each analysis rejects events with additional leptons that satisfy the *veto lepton* definition.

Table 4.4.: Lepton selections used in each analysis.

Analysis	p_{T} range	electron ID	muon ID	isolation	leptons
ggF	$< 25\,\mathrm{GeV}$	`TightLH`	`Tight`	`Gradient`	$1\mu + 1e$
	$> 25\,\mathrm{GeV}$	`MediumLH`	`Medium`	`Gradient`	$1\mu + 1e$
VBF	$< 25\,\mathrm{GeV}$	`TightLH`	`Tight`	`Gradient`	$1\mu + 1e$
	$> 25\,\mathrm{GeV}$	`MediumLH`	`Medium`	`Gradient`	$1\mu + 1e$
WH	$< 25\,\mathrm{GeV}$	`TightLH`	`Medium`	`Gradient`	3ℓ
	$> 25\,\mathrm{GeV}$	`MediumLH`	`Medium`	`Gradient`	3ℓ
Veto lepton		`MediumLH`	`Medium`	`GradientLoose`	

4.2.2 Jets and missing transverse energy

Jets are reconstructed using the anti-k_t algorithm with a distance parameter of $R = 0.4$, as defined in Section 3.2.5. They are required to have $p_T > 25\,\mathrm{GeV}$ for jets with $|\eta| < 2.4$, and $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ for jets outside this range. Any jets with $|\eta| > 4.5$ are discarded.

For jets with $p_T < 50\,\mathrm{GeV}$ and $|\eta| < 2.4$, the calibrated JVT variable, defined in Section 3.2.5, is required to be larger than 0.64 to suppress pile-up.

Leptons depositing significant amounts of energy in the calorimeter may be reconstructed as jets. Also, jets may be reconstructed as genuine leptons originating from the hard scatter. Thus, jets are discarded if they lie within a cone of size $\Delta R < 0.2$ of an electron candidate, or if they have less than three associated tracks and are within a cone of size $\Delta R < 0.2$ of a muon candidate. However, if a jet with three or more associated tracks is within a cone of size $\Delta R < 0.4$ of a muon candidate, or any jet is within $0.2 < \Delta R < 0.4$ of an electron candidate, the corresponding lepton candidate is discarded instead.

Jets containing b-hadrons are identified using the MV2c10 b-tagging algorithm introduced in Section 3.2.5, using the 85% working point. Top-quark background is suppressed in all analyses by vetoing events with any b-tagged jet passing a reduced p_{T} threshold of 20 GeV.

This analysis uses the track soft term missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ as well as the track-based missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$, as defined in Section 3.2.5.

4.3 Misidentified Leptons

Despite vigorous efforts to optimize the identification of objects, the rate of misidentification can never be reduced to zero. This is an especially pressing problem with the identification of leptons, as even though the chance of a jet being misidentified as a lepton is extremely small (on the level of 1 : 100 000 [142]), the overwhelmingly large cross section for multijet production at a hadron collider still causes the fraction of objects selected as leptons for any given analysis to be a cause of concern. These objects are generally referred to as *fake leptons*. Notably, real leptons originating from secondary processes such as heavy- (or light-) flavour hadrons decaying to muons, or electrons originating from pair production by photons originating from bremsstrahlung as well as from initial and final state radiation are also included in this broad category, as they are no true indicators of the high-energy electroweak processes investigated by the analysis. In contrast, a primary lepton originating from a hard electroweak scattering event is referred to as a *prompt lepton*. All of this applies to both electrons and muons, but the rate of *fake electrons* is much higher than that of *fake muons*, mostly because muons are identified with an additional measurement in the dedicated muon system. The composition of events with fake leptons can vary dramatically depending on the phase space investigated. A comparison between the composition of dilepton pairs with opposite charge and same charge can be seen in Fig. 4.6.

For example, collisions with one leptonically decaying W boson in association with one or more jets are accepted by the signal selection that requires two identified leptons in cases where a jet is misidentified as a prompt lepton. Similarly, also pure QCD processes can contribute to the analysis in cases where two jets are misidentified as prompt leptons. However, this contribution is much smaller – not only because the chance of two objects being misidentified in the same event is quadratically smaller, but also because the additional requirement of a large missing transverse momentum and other selections applied in this analysis strongly reject events without high-energy neutrinos in the final state.

As accurate knowledge of the object misidentification rate relies on an extremely precise modeling of the detector effects, the difficulties to accurately describe

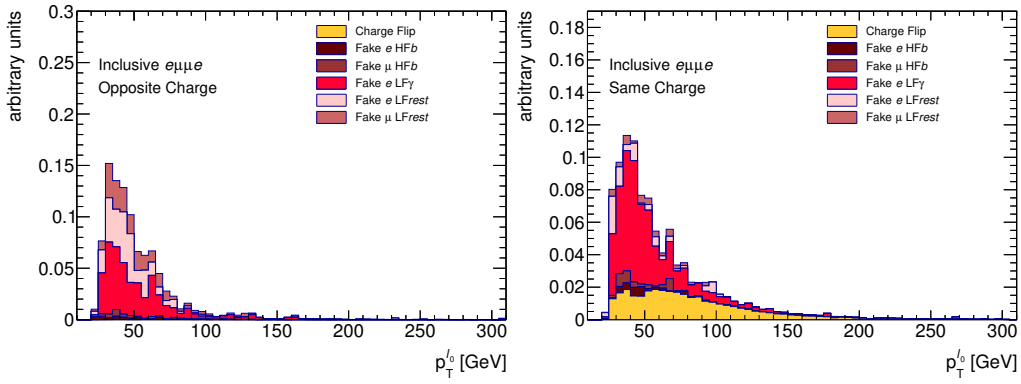


Figure 4.6.: Origin of fake leptons as a function of the leading lepton p_{T} in events with opposite-charge (left) and same-charge (right) lepton pairs after a loose pre-selection (described in Section 4.5), predicted by Monte Carlo simulation using POWHEG+PYTHIA8 +EVTGEN. HFb denotes leptons from heavy flavour jets (b-quark decays), LFγ denotes leptons originating from photons, LF*rest* denotes decays of light, strange and charm hadrons as well as τ-leptons.

contributions from misidentified objects in the analysis are notorious. Thus, it is common practice to use data-driven methods to estimate these contributions from the measured data itself instead. In this analysis, the *fake factor method* is used to estimate these "fakes" originating from W+jets processes, also sparing the need to generate large Monte Carlo samples to model these highly prevalent contributions.

4.3.1 The fake factor method

The total number of events in the selected signal region with two objects that have been identified as prompt leptons (✓) can be written as

$$N_{\checkmark\checkmark} = N^{\mathrm{f}}_{\checkmark\checkmark} + N^{\mathrm{p}}_{\checkmark\checkmark}, \tag{4.3.1}$$

where the subscript denotes two identified leptons, and the superscript denotes the true type of the identified object, that is, whether any of the leptons is fake ("f"), or whether both are prompt ("p").

In order to estimate the background contribution from fake leptons $N^{\mathrm{f}}_{\checkmark\checkmark}$, a control sample is defined with an enhanced rate of such fake leptons. Events in the control sample are required to have one lepton satisfying the signal selection criteria and one "anti-identified" lepton (✗) failing the signal selection, but satisfying less restrictive criteria.

The criteria defining the less restrictive anti-identified categories in comparison with the standard identified categories are listed in Table 4.5 and 4.6 for both,

Table 4.5.: Requirements imposed for fully identified and anti-identified muons, as well as for muon fake candidates.

identified	anti-identified	fake candidate
$p_{\mathrm{T}} > 15$ GeV		
$\lvert\eta\rvert < 2.45$		
$\lvert z_0 \sin\theta\rvert < 0.5$ mm		
Pass `Tight` if $p_{\mathrm{T}} < 25\,\mathrm{GeV}$		
Pass `Medium` if $p_{\mathrm{T}} > 25\,\mathrm{GeV}$		
$\lvert d_0\rvert / \sigma(d_0) < 3$	$\lvert d_0\rvert / \sigma(d_0) < 6$	
Pass `Gradient` Isolation	Veto against identified muon	–

Table 4.6.: Requirements imposed for fully identified and anti-identified electrons, as well as for electron fake candidates.

identified	anti-identified	fake candidate
$p_{\mathrm{T}} > 15\,\mathrm{GeV}$		
$\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$		
$\lvert z_0 \sin\theta\rvert < 0.5\,\mathrm{mm}$		
Pass `TightLH` if $p_{\mathrm{T}} < 25\,\mathrm{GeV}$ Pass `MediumLH` if $p_{\mathrm{T}} > 25\,\mathrm{GeV}$	Pass `LooseLH`	
$\lvert d_0\rvert / \sigma(d_0) < 5$		
Pass `Gradient` isolation	Veto against identified electron	–

muons and electrons, respectively. The isolation cuts are removed and events with fully identified leptons are rejected. For electrons, the less restrictive identification criterion corresponds to the `LooseLH` identification, as opposed to the `MediumLH` and `TightLH` criteria used for signal leptons. For muons, the d_0 significance cut is loosened to enhance the fake contribution.

Apart from this change, the events in the control sample are required to pass the full event selection, where the anti-identified lepton is treated as fully identified. As for the signal sample, the composition of the control sample can be expressed as

$$N_{\checkmark\times} = N^{\mathrm{f}}_{\checkmark\times} + N^{\mathrm{p}}_{\checkmark\times}, \tag{4.3.2}$$

where the superscript again refers to the true nature of the objects identified as leptons.

The fake contamination in the signal region is estimated by scaling the number of events in the control sample by the *fake factor* $f^{\checkmark}_{\times}$, defined as the ratio between

fully identified leptons $N_{\checkmark}$ and anti-identified leptons $N_{\times}$, that is,

$$f_{\times}^{\checkmark} = \frac{N_{\checkmark}}{N_{\times}}. \tag{4.3.3}$$

The fake contribution in the signal sample can be determined as

$$N_{\checkmark\checkmark}^{\mathrm{f}} = N_{\checkmark\times}^{\mathrm{f}} \times f_{\times}^{\checkmark} \tag{4.3.4}$$

$$= \left(N_{\checkmark\times} - N_{\checkmark\times}^{\mathrm{p}}\right) \times f_{\times}^{\checkmark}. \tag{4.3.5}$$

Under the assumption that the fake factor is independent of the details of the event topology, it can be measured in some fake-enriched data sample, using the same definitions of identified and anti-identified as for the control sample.

Despite requiring one identified object, the control sample will of course also contain events with no prompt leptons at all in cases where the identified lepton is itself a fake. Here, an additional complication arises due to the different combinations of identified and anti-identified objects. If the leptons are ordered, *i. e.*, by transverse momentum p_{T} or by flavour, the total number of events in the signal region can be written as

$$N_{\checkmark\checkmark} = N_{\checkmark\checkmark}^{\mathrm{pp}} + \overline{N}_{\checkmark\checkmark} \quad \text{where} \quad \overline{N}_{\checkmark\checkmark} = N_{\checkmark\checkmark}^{\mathrm{pf}} + N_{\checkmark\checkmark}^{\mathrm{fp}} + N_{\checkmark\checkmark}^{\mathrm{ff}}, \tag{4.3.6}$$

where the subscripts and superscripts are now ordered, the first one referring to the first lepton, the second one referring to the second lepton. Now, the fake component $\overline{N}_{\checkmark\checkmark}$ is estimated using the fake factor $f_{\times}^{\checkmark}$ as

$$\overline{N}_{\checkmark\checkmark} = \left(N_{\checkmark\times} + N_{\times\checkmark} - N_{\checkmark\times}^{\mathrm{pp}} - N_{\times\checkmark}^{\mathrm{pp}}\right) \times f_{\times}^{\checkmark} \tag{4.3.7}$$

$$= \left(N_{\checkmark\times}^{\mathrm{pf}} + N_{\checkmark\times}^{\mathrm{fp}} + N_{\checkmark\times}^{\mathrm{ff}} + N_{\times\checkmark}^{\mathrm{pf}} + N_{\times\checkmark}^{\mathrm{fp}} + N_{\times\checkmark}^{\mathrm{ff}}\right) \times f_{\times}^{\checkmark} \tag{4.3.8}$$

The components $N_{\checkmark\times}^{\mathrm{fp}}$ and $N_{\times\checkmark}^{\mathrm{pf}}$ require a fake lepton to pass identification and a prompt lepton to fail identification in the same event. This is very unlikely, hence these components are neglected. The estimate can then be decomposed as

$$N_{\checkmark\checkmark}^{\mathrm{pf}} = N_{\checkmark\times}^{\mathrm{pf}} \times f_{\times}^{\checkmark} \tag{4.3.9}$$

$$N_{\checkmark\checkmark}^{\mathrm{fp}} = N_{\times\checkmark}^{\mathrm{fp}} \times f_{\times}^{\checkmark} \tag{4.3.10}$$

$$N_{\checkmark\checkmark}^{\mathrm{ff}} = N_{\times\checkmark}^{\mathrm{ff}} \times f_{\times}^{\checkmark} = N_{\checkmark\times}^{\mathrm{ff}} \times f_{\times}^{\checkmark}, \tag{4.3.11}$$

where it becomes apparent that Eq. 4.3.8 double-counts the component with two fake leptons.

Earlier publications [119] applied a dedicated correction to amend this shortcoming. However, this contribution can be neglected in this analysis, as the higher requirements on the lepton p_{T} in comparison to these previously conducted analyses reduce the overall size of the fake background significantly by sacrificing the statistical power of the low-p_{T} component of the data set. In order to estimate the contribution from such events where both leptons are fakes, an explicit double-anti-identified region was defined and the fake-factor was applied to both such leptons. With this test, it was shown that the source of double fakes is at most 3% (6%) of the total fake contribution in the $n_{\mathrm{jet}} = 0$ ($n_{\mathrm{jet}} = 1$) gluon fusion signal region.

4.3.2 Measurement of the fake factors in di-jet events

For this analysis, the fake factors used in the estimate are measured in a di-jet data sample. They are computed as a function of p_{T} and $|\eta|$ of the anti-identified object. In order to measure suitable fake factors, a data sample created from the low-p_{T}, single-lepton prescaled trigger streams listed in Table 4.7 is used.

Most of the time, the anti-identified lepton will be

- the lepton with the lower transverse momentum p_{T}, and
- not the lepton that fired the trigger.

The fake factors applicable to all events where the leading lepton was selected by the trigger are referred to as the *nominal fake factors*.

However, it can happen that the anti-identified lepton is the only triggered lepton in the event and thus was selected by one of the un-prescaled single-lepton triggers used in the analysis. If the nominal fake factor is applied to these events, a small trigger bias is introduced in the background estimation, as the trigger selection cuts into the definition of the fake candidate object. In order to avoid this trigger bias, separate fake factors are extracted for these events using the data set selected by the standard triggers. In the following, these fake factors are called *triggered fake factors*.

The events selected by the single lepton triggers are required to have exactly one *fake candidate object*, *i. e.*, a loosely identified lepton satisfying criteria matching the definition of *anti-identified* objects, but dropping the veto against being selected as *identified*. In order to enforce the di-jet-like character of the events and to suppress background contributions from electroweak processes the following selections are applied:

Table 4.7.: Single-lepton prescaled trigger streams used for the fake factor determination.

electrons	`HLT_e12_lhvloose_nod0_L1EM10VH`
muons	`HLT_mu14_L1_MU10`

- Require exactly one fake candidate object with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, as well as
- at least one jet with $p_{\mathrm{T}} > 22\,\mathrm{GeV}$,
- both being separated by an azimuthal angle of $\Delta\phi > 2.5$, and finally
- $p_{\mathrm{T}}^{\mathrm{miss}} < 30\,\mathrm{GeV}$ as well as
- $m_{\mathrm{T}} < 60\,\mathrm{GeV}$.

This *fake candidate* definition is listed alongside the *identified* and *anti-identified* definitions in Tables 4.5 and 4.6.

The fake-enriched sample obtained with these selections will have small contributions from hard electroweak scattering processes, which could contaminate it with real prompt leptons. These contributions are estimated from Monte Carlo for the processes W+jets, Z+jets and $t\bar{t}$ and are then subtracted from the sample. The statistical uncertainties arising from the finite size of these Monte Carlo samples are then attached to the resulting fake factors.

As detector conditions and prescales may change over time, so may the fake factors. Thus, the data events used to derive the fake factors are weighted according to the prescale of the trigger used to collect them. This effectively removes the prescale *a posteriori*, ensuring that the fake factors match the un-prescaled data in the region to which they are applied.

After the selection criteria are applied, the remaining events are categorized as *identified* or *anti-identified*. After the electroweak background has been subtracted, these two selections constitute the numerator and denominator in the fake factor, respectively.

The fake factors are finally computed as a binned ratio in p_{T} and $|\eta|$ using four bins in p_{T} with boundaries $[15, 20, 25, 35, 1000]$ in GeV, and two bins in η with boundaries $[0, 1.5, 2.5]$ for electrons and $[0, 1.1, 2.5]$ for muons. For muons, the two $|\eta|$ bins are merged in the highest-p_{T} bin to reduce the statistical uncertainty of the estimate.

As the data-taking conditions changed from 2015 to 2016, different fake factors are derived for these data-taking periods. The fake factors are applied to the events in the control sample, based on the trigger configuration of the event and the p_{T} and $|\eta|$ of the anti-identified lepton. The triggered fake factors are only applied to events where the anti-identified lepton is triggered and the identified lepton is not triggered. The nominal fake factors are used for all other events.

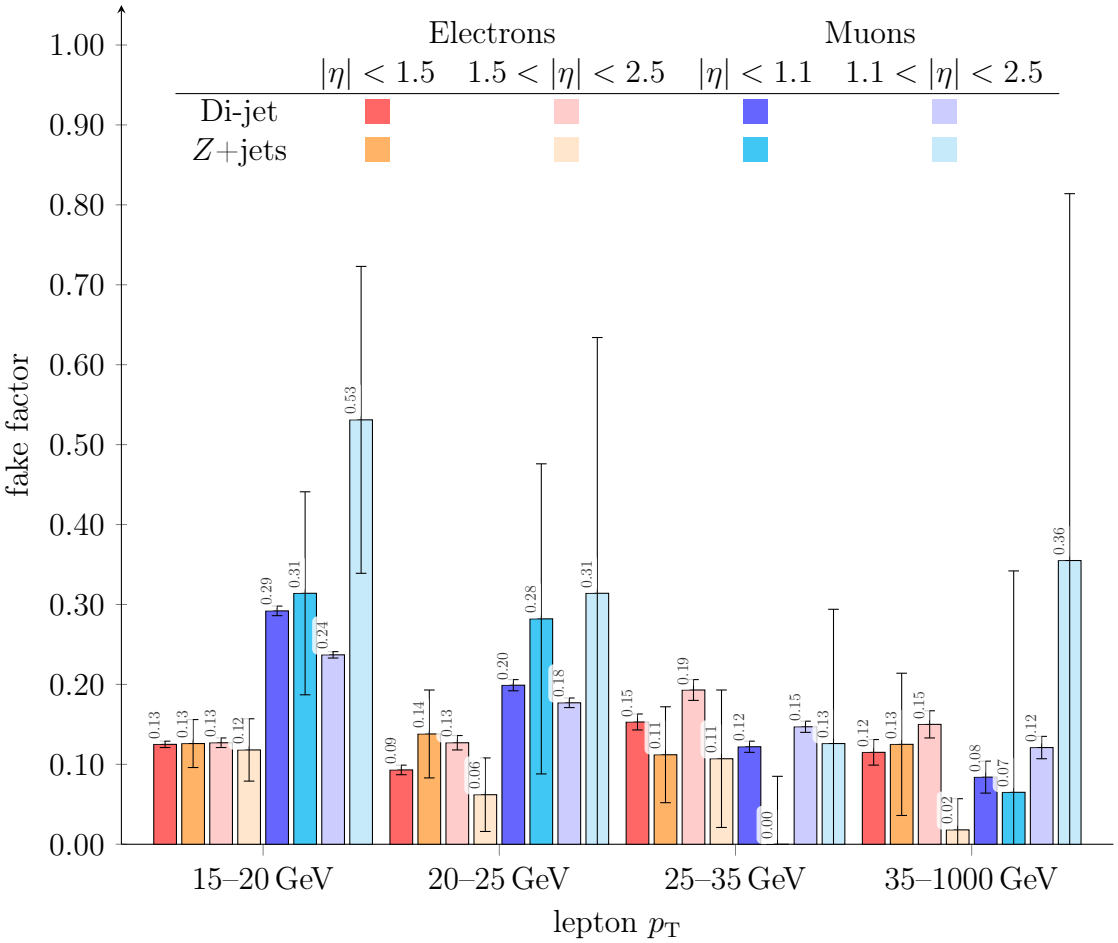


Figure 4.7.: Results of the fake factor calculations, including statistical uncertainties. Results that are negative, but compatible with zero within their uncertainties, have been shifted to zero.

4.3.3 Measurement of the fake factors in Z+jets events

In collisions producing Z bosons, the object recoiling against this Z boson is very likely to be a jet, due to the significantly higher cross section of Z+jets with respect of events with real additional leptons, such as diboson processes. Thus, such events pose an excellent data set to measure the fake factor. This second measurement of the fake factor is used to assess systematic uncertainties linked to this background estimation technique.

First, an inclusive Z selection is performed, requiring

- two same-flavour, opposite-charge leptons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, passing the identification criteria for electrons and muons, and
- an invariant dilepton mass in the range $70\,\mathrm{GeV} < m_{\ell\ell} < 110\,\mathrm{GeV}$.

The data set used for the measurement of the fake factor is then built

- by requiring exactly one additional loosely identified lepton
  - with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and
  - within $|\eta| < 2.5$ for muons and within $|\eta| < 2.47$, excluding the crack region $1.37 < |\eta| < 1.52$, for electrons,
- by requiring the additional lepton to be back-to-back with the Z candidate by imposing $\Delta\phi_{Z\ell} > 2.5$, and
- by applying a ZW veto, rejecting events if the transverse mass of the system of the third lepton and the missing transverse energy exceeds $m_{\mathrm{T}} > 30\,\mathrm{GeV}$, thus excluding events with a W candidate.

Based on this additionally selected lepton, the data set is split in an identified and an anti-identified category, with the fake factor estimate proceeding as outlined in the previous section, including a Monte Carlo based subtraction of diboson events.

The fake factor is then derived as the ratio between the number of identified and anti-identified leptons as a function of the transverse momentum and pseudorapidity. Figure 4.7 shows the comparison of the fake factors derived using Z+jets events with those estimated in di-jet events.

4.3.4 Uncertainty estimation

In order to evaluate the uncertainty on the data-driven fake estimate, several sources need to be taken into account:

- The statistical uncertainty on the fake factor measurement is included.
- The uncertainty associated to the real lepton contamination from electroweak processes in the di-jet data sample can be estimated by varying the subtracted Monte Carlo predictions.
- The difference between the fake factor derived from the di-jet data sample and from the Z+jets data sample can be used to estimate the uncertainty connected to the varying flavour composition of the jets in the different samples.
- As the composition of the contributions entering the analysis due to lepton misidentification depends on the charge of the lepton pair, the corresponding variation is applied to the same-charge region.

In the following sections, more details are provided on how these individual components are assessed.

Table 4.8.: Summary of the systematic uncertainties on the fake factors as defined in the text, given in percent, for the individual contributions summed in quadrature (Total): the sample composition (Comp.), the subtraction of electroweak processes (EW) and the statistical uncertainty (Stat.). Here, OC denotes the systematic uncertainty in the opposite-charge region, and SC denotes the systematic uncertainty in the same-charge region. The additional systematic uncertainty in the same-charge region for the electron (muon) fake factor is ±15% (±13%), respectively.

Kinematic region ($\|\eta\|$ and p_T range)	Total		Comp.		EW	Stat.
	OC	SC	OC	SC		
Muon:						
$0.0 < \|\eta\| < 1.05$						
$15 - 20\,\text{GeV}$	63	65	63	65	0	2
$20 - 25\,\text{GeV}$	63	65	63	65	1	4
$25 - 35\,\text{GeV}$	64	66	63	65	6	6
$1.05 < \|\eta\| < 2.5$						
$15 - 20\,\text{GeV}$	63	65	63	65	0	2
$20 - 25\,\text{GeV}$	63	65	63	65	2	3
$25 - 35\,\text{GeV}$	63	65	63	65	6	5
$0 < \|\eta\| < 2.5$						
$35 - 1000\,\text{GeV}$	73	75	63	65	36	11
Electron:						
$0.0 < \|\eta\| < 1.5$						
$15 - 20\,\text{GeV}$	40	43	40	43	0	3
$20 - 25\,\text{GeV}$	40	43	40	43	2	6
$25 - 35\,\text{GeV}$	41	44	40	43	3	7
$35 - 1000\,\text{GeV}$	44	47	40	43	11	14
$1.5 < \|\eta\| < 2.5$						
$15 - 20\,\text{GeV}$	40	43	40	43	0	4
$20 - 25\,\text{GeV}$	41	44	40	43	1	7
$25 - 35\,\text{GeV}$	41	44	40	43	2	7
$35 - 1000\,\text{GeV}$	42	45	40	43	5	11

Subtraction of electroweak processes in di-jet events

The contribution of the electroweak processes is dominated by W/Z+jets production. The shapes of the distributions are predicted by Monte Carlo generators and then normalized to the theory prediction.

As no efficiency corrections are available for event selections without tightly identified isolated leptons, the normalization is subject to a considerable uncertainty, which is estimated to be $\pm 20\%$. This estimate also covers the cross section uncertainties on W/Z+jets production and Monte Carlo modeling uncertainties on W/Z decays.

For each systematic variation of the electroweak contribution, the nominal and triggered fake factors are recomputed coherently, *i. e.*, the nominal fake factor variation is fully correlated with the triggered fake factor variation. The difference between values obtained using the variation and the central values is used as the uncertainty estimate.

Sample composition

The composition of the fake background varies depending on the precise event selection. These differences might affect the fraction of gluon-induced jets with respect to quark-induced jets, as well as the fraction of heavy-flavour jets in the selected phase space. The data samples used to determine the fake factor will also have a fake composition different from the one in the signal region. In order to estimate the systematic uncertainty arising from this variation in sample composition, the alternative derivation of the fake factors in Z+jets events is used.

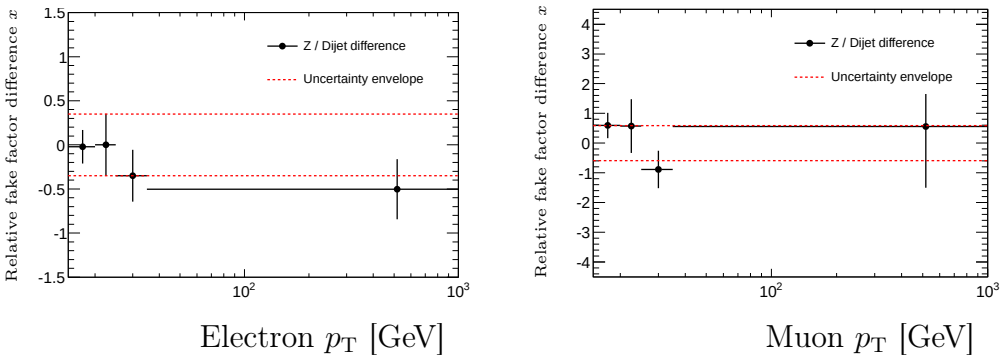


Figure 4.8.: Resulting systematic uncertainty related to sample composition for electron (left) and muon fakes (right), determined by comparing the di-jet with the Z+jets estimates for the fake factors.

This is done by considering the relative difference x of the fake factors derived from the di-jet sample ($f_{\text{✓✗}}$) and from the Z+jets sample ($f_{\text{✓✗}}'$) as

$$x(p_{\mathrm{T}}, \eta) = \frac{\left| f_{\text{✓✗}}(p_{\mathrm{T}}, \eta) - f_{\text{✓✗}}'(p_{\mathrm{T}}, \eta) \right|}{f_{\text{✓✗}}(p_{\mathrm{T}}, \eta)}. \tag{4.3.12}$$

The numbers used for this calculation are the ones presented in Fig. 4.7. The differences in Eq. 4.3.12 are then averaged over p_{T} and η using the weighted average

$$\mu = \frac{\sum_{ij} x_{ij}/\sigma_{ij}^2}{\sum_{ij} 1/\sigma_{ij}^2} \quad \text{and} \quad \sigma^2 = \frac{1}{\sum_{ij} 1/\sigma_{ij}^2}, \tag{4.3.13}$$

where the indices i and j run over bins of transverse momentum and pseudorapidity, with x_{ij} denoting the central value of the difference and σ_{ij} its statistical uncertainty. In this way, bins with large statistical uncertainty have less weight in the determination of the systematic uncertainty, resulting in an uncertainty of $\pm(35 \pm 12)\%$ for electrons and $\pm(60 \pm 30)\%$ for muons. A comparison of the systematic uncertainty estimated using this procedure with the relative differences between the fake factors in Z+jets and di-jet events can be seen in Fig. 4.8, clearly showing that the differences are covered by the uncertainty.

However, the difference in composition between the Z+jets sample and the W+jets sample chosen to model the fakes in the signal region also has to be taken into account. As there is no data-driven way to estimate this uncertainty, it has to be evaluated on Monte Carlo samples. Studies performed for previous publications [119] comparing different Monte Carlo generators find this uncertainty to be around $\pm 20\%$, which is added in quadrature. Thus, the total systematic uncertainty corresponds to $\pm 40\%$ for electrons and to $\pm 63\%$ for muons.

Statistical uncertainty

The statistical uncertainties on the nominal and triggered fake factors are the ones indicated as error bars in Fig. 4.7. For muons, the two bins in $|\eta|$ are merged for the highest bin in muon p_{T} to reduce the statistical uncertainty. The resulting uncertainties are applied as systematic variation to the fake estimate. All the bins are varied independently, *i. e.*, the uncertainties are treated as uncorrelated across all bins and different sets of fake factors.

Lepton charge dependence

Depending on whether the leptons have the same or opposite charge, the fake composition changes. Especially, the contribution from charm quark production is enhanced in the opposite-charge case. Hence, an additional uncertainty is applied to the fake estimate for event selections with same-charge leptons.

The uncertainty is estimated using the difference in fake factors between the same-charge and opposite-charge case, as derived from W+jets Monte Carlo (POWHEG with PYTHIA8). The fake factors are derived from events with two loosely identified leptons, where the leading lepton is required to originate from the W-boson decay and the preselection criteria from the ggF and VBF analyses are applied. The fake factor is then computed using the subleading lepton

In order to reduce the statistical uncertainties, the fake factors are recomputed after integrating over the full p_{T} and $|\eta|$ range to obtain a single-bin, global fake factor. The difference between the opposite-charge and same-charge global fake factors are $\pm 15\%$ for electrons and $\pm 13\%$ for muons. These numbers are used as systematic uncertainties.

Summary

Table 4.8 summarizes the percental systematic uncertainties on the fake factors. The sample composition uncertainty is generally the dominating contribution. Only for the high-p_{T} bin of the muon fake factor, the electroweak subtraction uncertainty is of comparable size.

4.4 Other Data-Driven Methods

Even though Monte Carlo generators are indispensable assets for conducting analyses in high energy physics, their applicability and predictive power is often limited by the sheer volume of simulated data required for a reliable estimate, by the accuracy of the cross section predictions available from theory, and by the precision at which the experimental conditions are known. As already presented in Section 4.3, data-driven methods are often used to overcome these deficiencies. A few simple yet effective and thus commonly used methods are briefly outlined in this section.

4.4.1 Single process normalization

In an analysis with a signal process s and a background process b, one usually attempts to define a signal region SR which is enriched in signal events with respect to the background. However, one will usually not be able to define a signal region that is entirely void of any background, such that a reliable estimate of this background is an important asset for measuring the signal. One could, of course, try to measure the number of background events from the same signal region. However, lacking a localized mass peak that can help to distinguish signal from background, this will spoil any sensitivity on the signal.

Instead, it is customary to define a control region CR from a subset of the events that were rejected from the signal region such that it is enriched in background

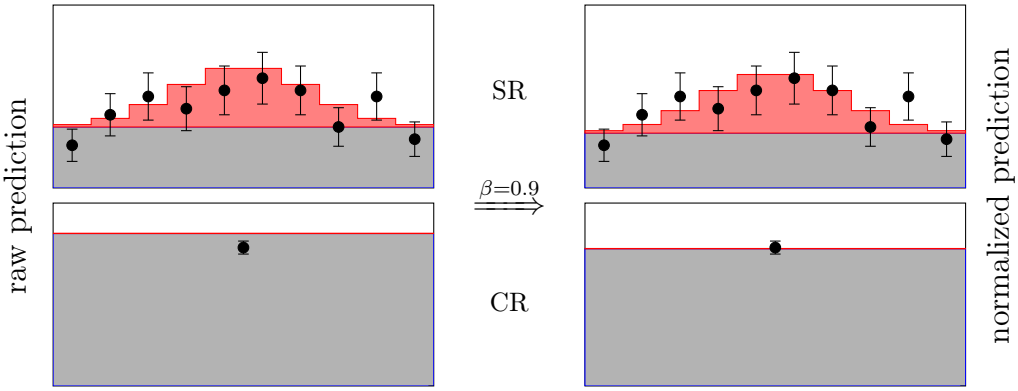


Figure 4.9.: Schematic setup with one signal contribution (red), one background contribution (gray) and some data points (black). The normalization factor β is measured in the control region CR and is applied to the background yield in the control region CR and the signal region SR, while the signal prediction remains unchanged.

with respect to signal. As the background is generally much more prominent than the signal, it is often possible to define this control region such that the signal contribution in that region is negligible. The number of events $n^{\text{CR}}_{\text{obs.}}$ observed in the control region and the corresponding number $n^{\text{CR}}_{b,\text{theo.}}$ of background events predicted by theory for that region can then be used to measure a normalization factor β of the background, that is,

$$\beta = \frac{n^{\text{CR}}_{\text{obs.}}}{n^{\text{CR}}_{b,\text{theo.}}}, \tag{4.4.1}$$

and apply this normalization factor to the signal region. The number of background events in the signal region n^{SR}_b can then be estimated from the corresponding theory prediction $n^{\text{SR}}_{b,\text{theo.}}$ as

$$n^{\text{SR}}_b = \beta \cdot n^{\text{SR}}_{b,\text{theo.}}. \tag{4.4.2}$$

Of course, an estimate of the statistical uncertainty on β and subsequently on n^{SR}_b can easily be obtained with simple Gaussian uncertainty propagation.

The applicability of this method relies on some assumptions:

- This method only corrects the total normalization. Any kinematic mismodeling within the signal region itself cannot be amended, and any difference in kinematic properties between the signal and the control region will bias the results.
- The control region needs to be similar enough to the signal region such that the extrapolation is justifiable. If the background composition in the signal

region is different from the one in the control region or the Monte Carlo generator mismodels the ratio between the control and the signal region, a single normalization factor β will not be able to model this change accurately.

- If it is not possible (or feasible) to define a control region with a negligible signal contribution, the signal contribution needs to be subtracted from the control region prediction first. This will then induce an inconsistency, as the presence and properties of the signal are assumed in the background estimate for its measurement.

Nevertheless, this type of method has proven very successful, and this section outlines some of its more complex variants. Figure 4.9 shows how this technique can improve the agreement between the prediction and the measured data for a simple toy model.

4.4.2 ABCD method

One possible alternative to the normalization factor method is the ABCD method. For this method, one defines four different regions, labeled A, B, C and D, defined by cuts in two independent variables x and y, splitting the data at the values $x = \widetilde{x}$ and $y = \widetilde{y}$. One usually chooses A to be the signal region. A sketch of the situation is shown in Fig. 4.10.

If the variables x and y are independent, the event yield of region A must relate to the yield in region B just as the yield in region C relates to the yield in region D, or

$$n_A = n_B \cdot \alpha \quad \text{with} \quad \alpha = \frac{n_C}{n_D}, \tag{4.4.3}$$

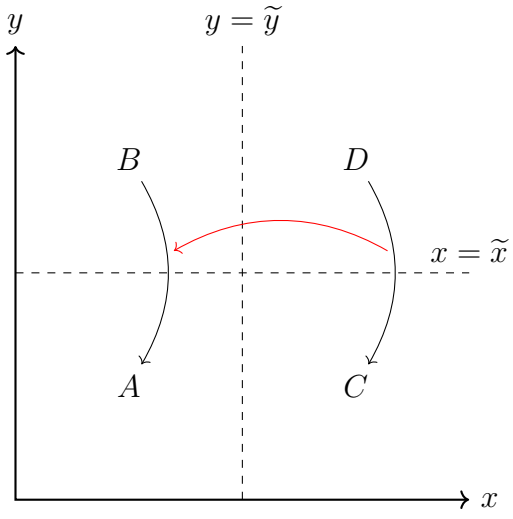


Figure 4.10.: Sketch of the ABCD method.

where α is referred to as an *extrapolation factor* from region D to region C, which is then applied to extrapolate from region B to region A.

This gives an entirely independent estimate of the event yield n_A in the signal region. In principle, one could even employ this method to estimate the shape of a kinematic distribution in region A by replacing n_B with a histogram of the distribution in region B. In this mode, the method can be considered a pristinely data-driven template method, as no theory prediction whatsoever is required for the signal region A. The propagation of statistical uncertainties via this method can be estimated using Gaussian uncertainty propagation.

However, in practice, the applicability of this method is often limited by the requirement that the variables should be independent, and that the statistical power in the regions B, C and D needs to be sufficient.

4.4.3 Matrix method

One further development of the simple normalization factor method is the matrix method. In a scenario with N different background processes labeled $1, \ldots, N$, one would be interested in the total background prediction $n_{\text{bkg}}^{\text{SR}}$ for the signal region. However, when attempting to estimate all N processes with the simple normalization

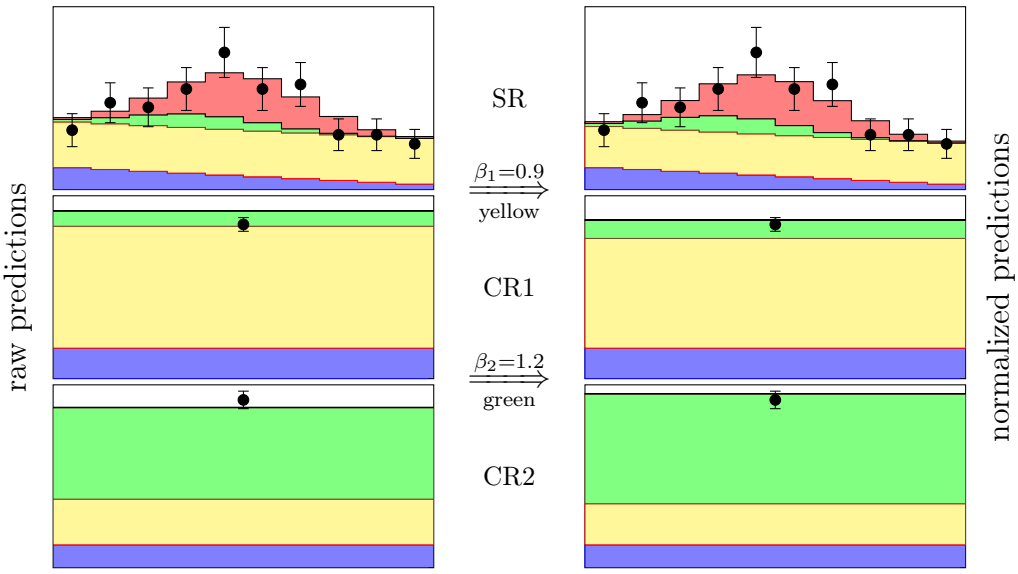


Figure 4.11.: Schematic setup with one signal contribution (red) and several background contributions (blue, green, and yellow), featuring one signal region SR and two control regions CR1 and CR2 with differing purity. Normalization factors β_i are measured for the green and yellow background contributions in the two control regions simultaneously and then applied to all regions, changing the respective background yields.

factor method presented in Section 4.4.1 from N different control regions, results will be biased by the fact that either control region will be contaminated by other backgrounds, which should themselves be estimated by the method. This problem can be cast in matrix notation as

$$n_{\text{bkg}}^{\text{SR}} = \begin{pmatrix} n_1^{\text{SR}} & \cdots & n_N^{\text{SR}} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}. \tag{4.4.4}$$

The same normalization factors β_i obviously connect the estimates in the control regions, that is,

$$\begin{pmatrix} n_{\text{obs.}}^{\text{CR1}} \\ \vdots \\ n_{\text{obs.}}^{\text{CR}N} \end{pmatrix} = \begin{pmatrix} n_1^{\text{CR1}} & \cdots & n_N^{\text{CR1}} \\ \vdots & \ddots & \vdots \\ n_1^{\text{CR}N} & \cdots & n_N^{\text{CR}N} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}, \tag{4.4.5}$$

such that

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix} = \begin{pmatrix} n_1^{\text{CR1}} & \cdots & n_N^{\text{CR1}} \\ \vdots & \ddots & \vdots \\ n_1^{\text{CR}N} & \cdots & n_N^{\text{CR}N} \end{pmatrix}^{-1} \cdot \begin{pmatrix} n_{\text{obs.}}^{\text{CR1}} \\ \vdots \\ n_{\text{obs.}}^{\text{CR}N} \end{pmatrix}. \tag{4.4.6}$$

Thus, the consistency problem can be solved with a simple matrix inversion. However, one caveat arising from this solution is the fact that uncertainty propagation is no longer trivially possible, as the quantities entering the matrix inversion have uncertainties attached. One simple and reliable way to estimate the uncertainty is the use of toys, that is, randomly generated matrices within the uncertainties of their inputs are inverted and the mean and standard deviation of the resulting distributions in the normalization factors β are then used for further calculations. Similarly, systematic uncertainties can be propagated through the method by evaluating the method for every variation separately and comparing the results. A toy example featuring a 2×2 matrix is shown in Fig. 4.11.

4.4.4 Likelihood fit

Ultimately, most results are extracted from the data by using the likelihood fitting methods introduced in Section 2.5. Thus, one of the most straight-forward methods of deriving the correct normalization for some background process is the inclusion of a free normalization parameter for this background in the final likelihood fit in the form of an unconstrained *nuisance parameter*. This, of course, has the additional

feature that the normalization factors derived are automatically consistent with the results obtained from the signal extraction.

However, just like the aforementioned methods, this approach relies on the availability of additional information used to derive the corresponding normalization factors, usually in form of dedicated control regions. In some cases, construction of such a control region is not possible. In these cases, more intricate methods of estimation need to be developed and integrated with the signal extraction strategy. Several such examples are discussed in Section 4.5.2.

4.5 The Gluon Fusion Analysis

The gluon fusion process produces jets only through parton radiation from the initial state partons. Therefore, the largest signal yield for gluon fusion Higgs boson production is expected for lower jet multiplicity ($n_{\text{jet}} < 2$). Numerous background processes contribute to these signal regions, including the production of top ($t\bar{t}$ and Wt), WW, WZ, $W\gamma^{(*)}$, ZZ, W+jets, QCD, $Z{\to}\tau\tau$, and $Z{\to}ee/\mu\mu$ events. These can be grouped into categories based on the final state properties which allow these events to pass the signal region selection.

- All of $t\bar{t}$, Wt, and WW have two W bosons in the final state, similar to the signal. The presence of jets, especially those tagged as b-jets, is used to discriminate against processes containing top quarks. The spin correlation kinematics described in Fig. 4.4 are used to reject WW.
- $Z/\gamma^* \to \tau\tau$, and the "Non-WW diboson" processes WZ, $W\gamma^{(*)}$, and ZZ, collectively referred to as VV, have small cross sections, but more signal-like kinematics because of the softer subleading lepton, particularly in the case of $W\gamma$ and $W\gamma^*$.
- W+jets and multijet production via QCD processes have a high cross section and pass the signal region selection under the rare condition that a jet produces an object reconstructed as an isolated lepton. The kinematics of W+jets events are similar to those of signal events because leptons produced by jets tend to be soft.

Many background processes are normalized using control regions. These are designed to be orthogonal to the signal regions, but otherwise as similar as possible to the signal regions to minimize the impact of theoretical and experimental uncertainties on the background predictions.

This analysis solely focuses on the different-flavour channel and defines two disjoint signal regions corresponding to different jet multiplicities, $n_{\text{jet}} = 0$ and $n_{\text{jet}} = 1$. These signal regions are further split into eight signal region categories

Table 4.9.: Signal region categories of the gluon fusion analysis. There are eight categories per signal region ($n_{\text{jet}} = 0$ and $n_{\text{jet}} = 1$). The product operator $\otimes[x_1, x_2, \ldots, x_{n+1}]$ denotes a cut and split of the data set in n disjoint regions into bins separated by the boundaries x_1 to x_{n+1} in some variable, or into a set of discrete categories like in the case of $\ell_2 = e$ or $\ell_2 = \mu$.

n_{jet}	$\otimes m_{\ell\ell}$	$\otimes p_{\text{T}}^{\ell_2}$	$\otimes \ell_2$
$n_{\text{jet}} = 0$	$\otimes[10, 30, 55]$	$\otimes[15, 20, \infty]$	$\otimes[e, \mu]$
$n_{\text{jet}} = 1$	$\otimes[10, 30, 55]$	$\otimes[15, 20, \infty]$	$\otimes[e, \mu]$

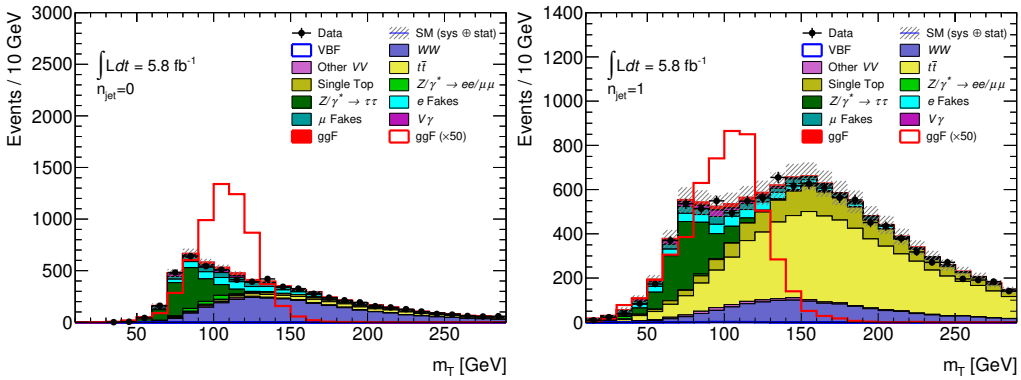


Figure 4.12.: Distributions of transverse mass in the $n_{\text{jet}} = 0$ (left) and $n_{\text{jet}} = 1$ (right) category after the preselection (see Section 4.5.1) has been applied. The normalization factors from Table 4.14 have been applied.

each to extract the results, as shown in Table 4.9, splitting in $e\mu$ and μe as well as in two bins each of $m_{\ell\ell}$ and $p_{\text{T}}^{\ell_2}$.

For all signal region categories, the transverse mass, defined as

$$m_{\text{T}} = \sqrt{(E_{\text{T}}^{\ell\ell} + E_{\text{T}}^{\text{miss}})^2 - \left|\vec{p}_{\text{T}}^{\,\ell\ell} + \vec{E}_{\text{T}}^{\,\text{miss}}\right|^2} \tag{4.5.1}$$

with

$$E_{\text{T}}^{\ell\ell} = \sqrt{|\vec{p}_{\text{T}}^{\,\ell\ell}|^2 + m_{\ell\ell}^2}, \tag{4.5.2}$$

is used as a discriminating variable, as Higgs boson candidates are expected to peak in the distribution of the transverse mass. The shape of the distributions of signal and backgrounds in transverse mass after a loose preselection (defined in the upcoming section) can be seen in Fig. 4.12. A likelihood fit method (see Section 2.5) is used on a data set binned in m_{T} to extract the signal strength parameter μ (see Section 2.4.2), which can be directly translated into a cross section.

4.5.1 Event selection

Events are selected by requiring exactly one electron and one muon with opposite charge, a dilepton invariant mass $m_{\ell\ell}$ greater than 10 GeV, and $p_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV. The variable $p_{\mathrm{T}}^{\mathrm{miss}}$ as opposed to other definitions of missing transverse momentum is used because it best separates the signal from the Drell-Yan background, especially in scenarios with high pile-up. This is demonstrated by Fig. 4.14. At least one of the two leptons is required to have $p_{\mathrm{T}} > 25$ GeV. If the event stems from the 2015 data set (or the associated share of Monte Carlo generated events), the transverse momentum threshold is lowered to $p_{\mathrm{T}}^{\ell_1} > 22$ GeV for muons to increase the size of the available data set. The subleading lepton is required to satisfy $p_{\mathrm{T}}^{\ell_2} > 15$ GeV regardless.

Two disjoint signal regions are defined following that selection, distinguished by the number of reconstructed jets n_{jet} as $n_{\mathrm{jet}} = 0$ or $n_{\mathrm{jet}} = 1$. The sub-categorization in the number of reconstructed jets within the ggF analysis helps to accommodate the different background compositions in the individual jet bins, as can be seen

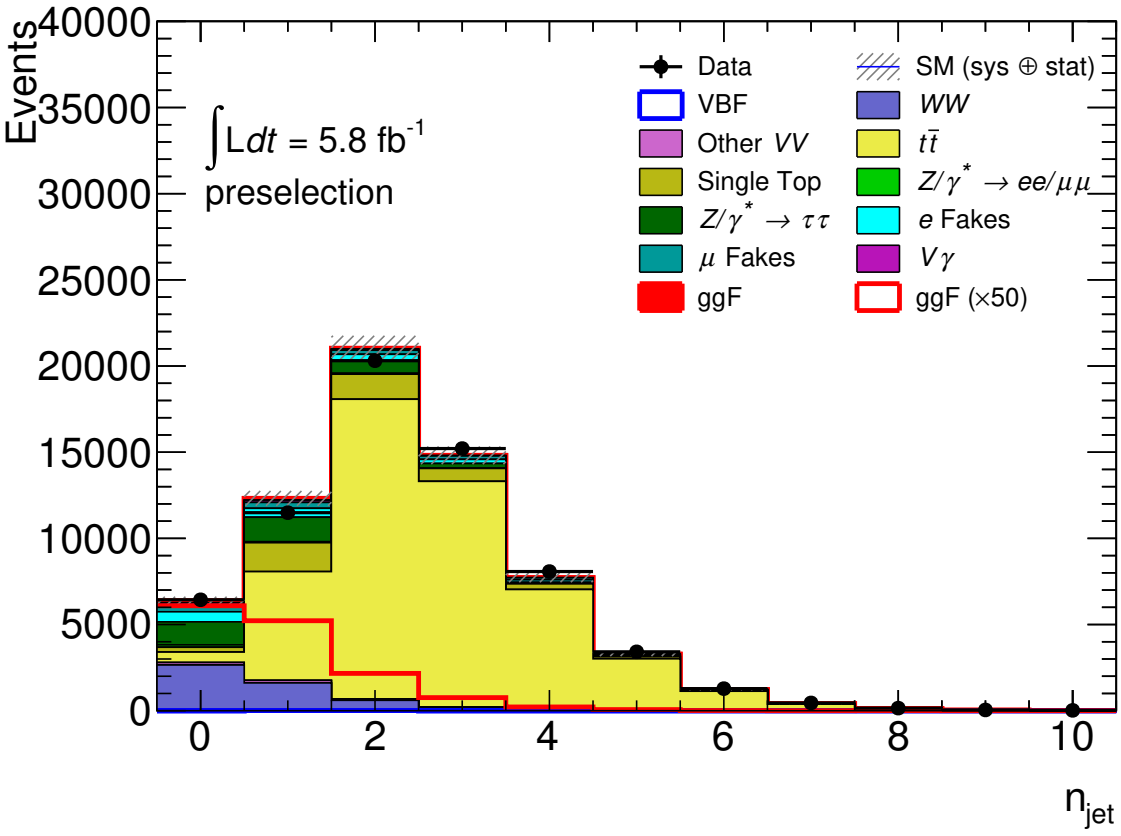


Figure 4.13.: Distribution of events by number of identified jets. The normalization factors from Table 4.14 have been applied.

Table 4.10.: Event selection criteria used to define the signal regions in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ ggF analysis. Definitions including the p_{T} thresholds for jet counting are given in the text.

Category	$n_{\mathrm{jet}} = 0$	$n_{\mathrm{jet}} = 1$
Preselection	Opposite-charge isolated lepton pair	
	$p_{\mathrm{T}}^{\ell_1} > 25\,\mathrm{GeV}$ ($p_{\mathrm{T}}^{\ell_1} > 22\,\mathrm{GeV}$ for muons in 2015)	
	$p_{\mathrm{T}}^{\ell_2} > 15\,\mathrm{GeV}$	
	$m_{\ell\ell} > 10\,\mathrm{GeV}$	
	$p_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$	
background rejection	–	$n_{b\text{-jet}} = 0$
	$\Delta\phi_{\ell\ell, E_{\mathrm{T}}^{\mathrm{miss}}} > 1.57$	$\max_i m_{\mathrm{T}}^{\ell_i} > 50\,\mathrm{GeV}$
	$p_{\mathrm{T}}^{\ell\ell} > 30\,\mathrm{GeV}$	$m_{\tau\tau} < m_Z - 25\,\mathrm{GeV}$
$H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ topology		$m_{\ell\ell} < 55\,\mathrm{GeV}$
		$\Delta\phi_{\ell\ell} < 1.8$

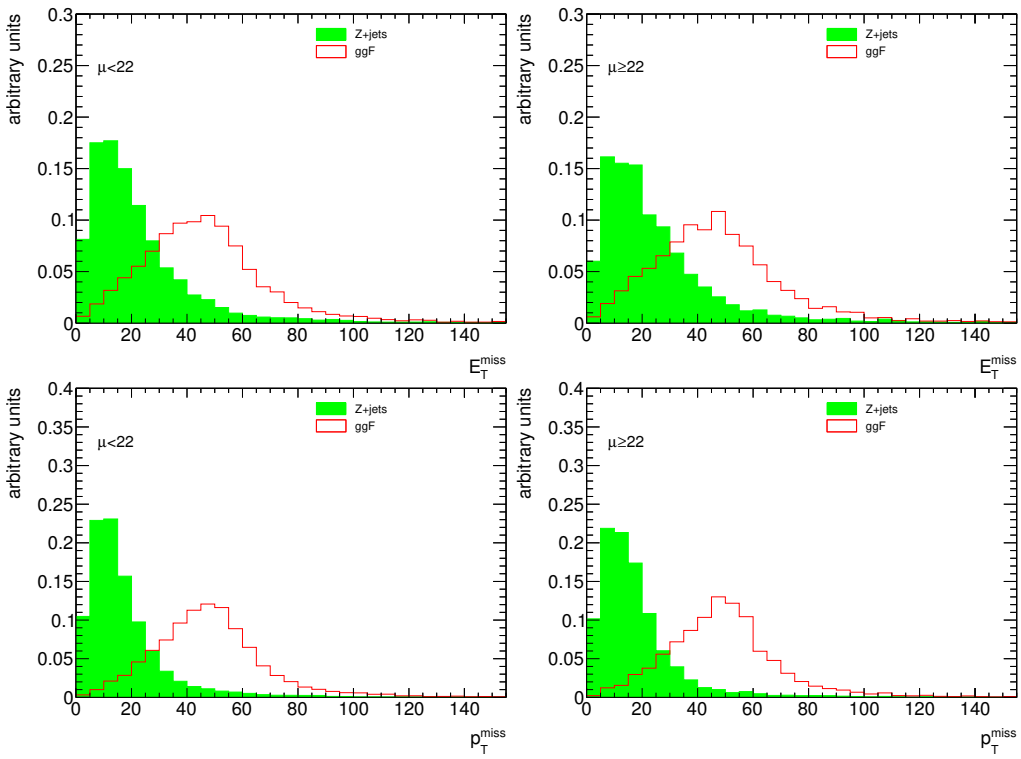


Figure 4.14.: Variants $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ of missing transverse momentum and their separation power between the gluon fusion signal (red) and the Drell-Yan background (green), for events with low pile-up ($\mu < 22$, left) and high pile-up ($\mu \geq 22$, right).

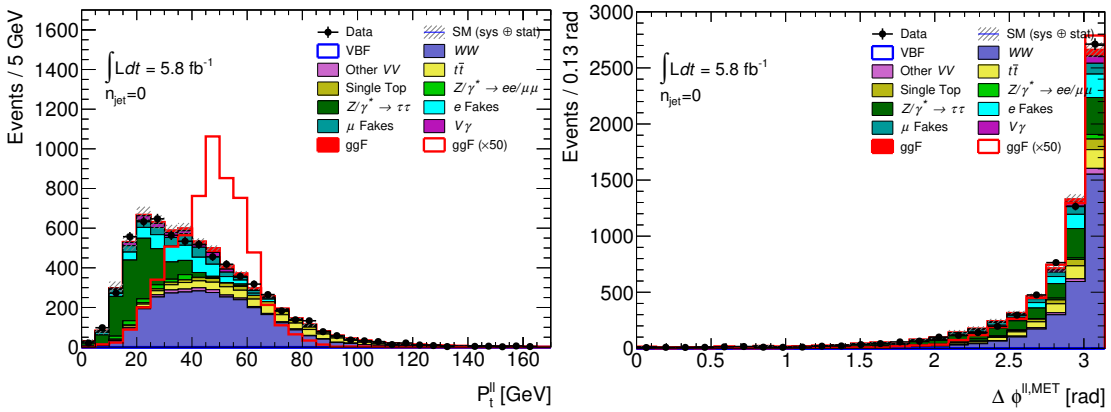


Figure 4.15.: Distributions of events in transverse momentum of the dilepton system $p_{\mathrm{T}}^{\ell\ell}$ (left) and in angular separation of the dilepton system against the missing transverse momentum vector $\Delta\phi_{\ell\ell,E_{\mathrm{T}}^{\mathrm{miss}}}$ (right). These quantities are used in the event selection of the $n_{\mathrm{jet}} = 0$ category. The normalization factors from Table 4.14 have been applied.

in Figure 4.13. Events with higher jet multiplicities are excluded and used in the analysis dedicated to VBF Higgs boson production presented in Section 4.6 instead. The two resulting jet multiplicity categories comprise a very different background composition. The $n_{\mathrm{jet}} = 0$ category is dominated by non-resonant WW production, whereas a large share of the events falling into the $n_{\mathrm{jet}} = 1$ category originates from Wt and $t\bar{t}$ processes. Thus, the two categories employ different event selections. The full selection is summarized in Table 4.10.

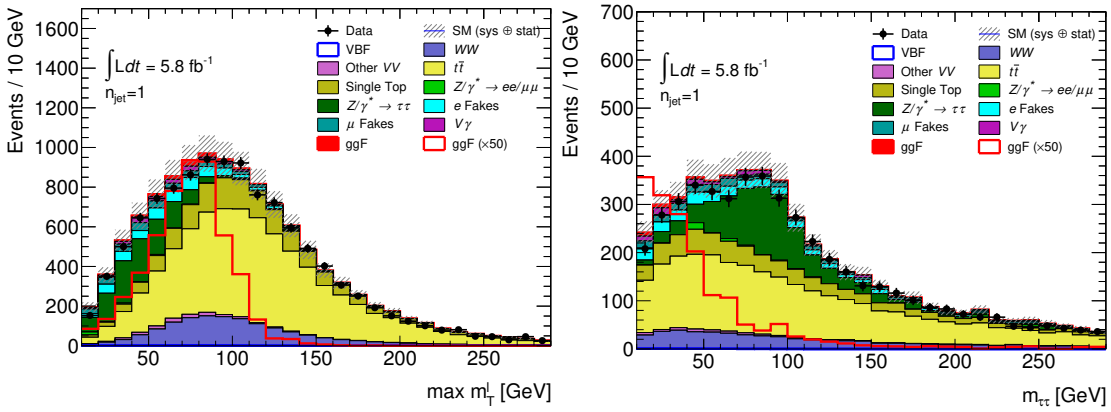


Figure 4.16.: Distributions of events in maximum transverse lepton mass $\max_i m_{\mathrm{T}}^{\ell_i}$ and reconstructed invariant mass $m_{\tau\tau}$ of the $\tau\tau$ system. These quantities are used in the event selection of the $n_{\mathrm{jet}} = 1$ category. The normalization factors from Table 4.14 have been applied.

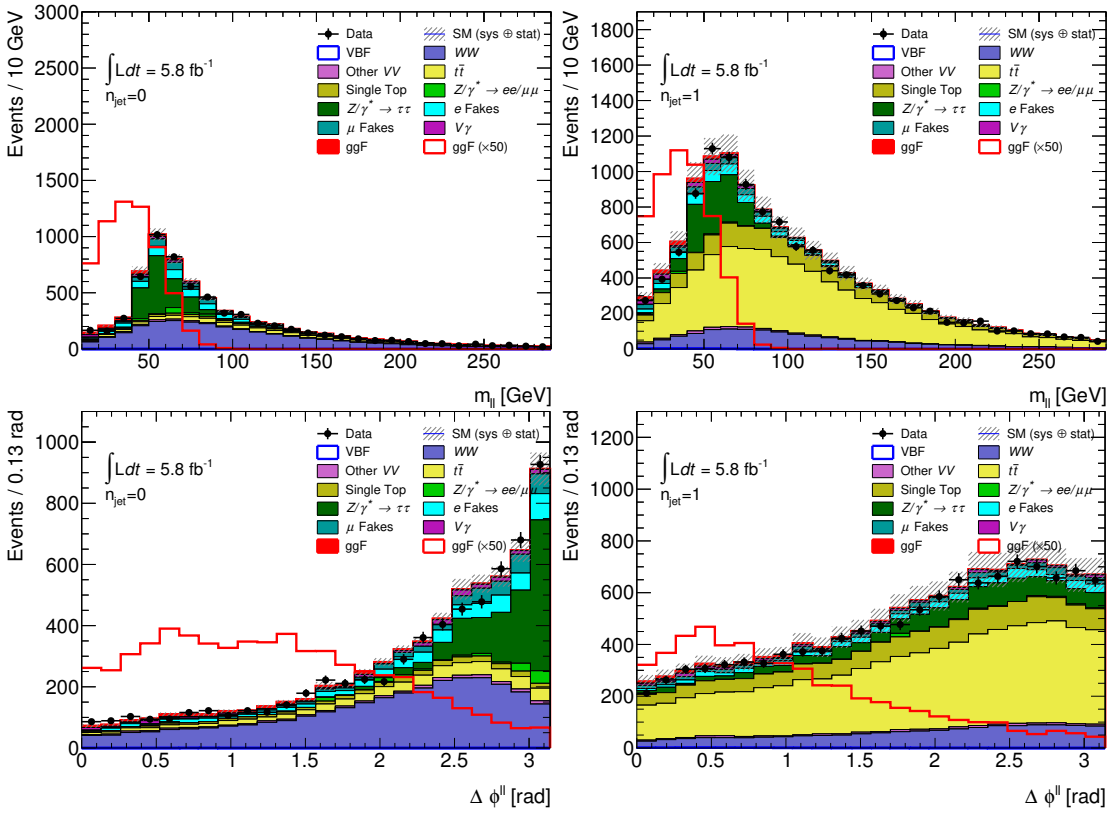


Figure 4.17.: Distributions of events by invariant dilepton mass $m_{\ell\ell}$ (top) and azimuthal angular spread $\Delta\phi_{\ell\ell}$ (bottom) of the dilepton system in the $n_{\text{jet}} = 0$ (left) and $n_{\text{jet}} = 1$ (right) categories. The normalization factors from Table 4.14 have been applied.

In the $n_{\text{jet}} = 0$ category, the Higgs boson is not boosted. Thus, the transverse momentum of the dilepton system $p_{\text{T}}^{\ell\ell}$ provides a handle over the unique $H \to W^{\pm}W^{\mp *}$ topology, large values of $p_{\text{T}}^{\ell\ell}$ corresponding to a scenario when the leptons and the neutrinos are emitted in opposite hemispheres. This topology can further be enhanced by requiring a large angular distance between the dilepton system and the missing transverse momentum $\Delta\phi_{\ell\ell,E_{\text{T}}^{\text{miss}}}$. Distributions of events in both of these quantities are shown in Fig. 4.15.

In the $n_{\text{jet}} = 1$ category, background from $Z \to \tau\tau$ is rejected using the invariant mass $m_{\tau\tau}$ of the hypothetical $\tau\tau$ system using the collinear approximation [143], employing the assumption that the τ leptons are sufficiently boosted to force the products of the τ decay to light leptons to be collinear. Background from $Z \to \tau\tau$ and multijets is further reduced with a requirement on the single-lepton transverse

mass m_{T}^{ℓ}. This is calculated for both of the charged leptons as

$$m_{\mathrm{T}}^{\ell} = \sqrt{2\left(E_{\mathrm{T}}^{\mathrm{miss}} p_{\mathrm{T}}{}^{\ell} - \vec{p}_{\mathrm{T}}{}^{\ell} \cdot \vec{E}_{\mathrm{T}}{}^{\mathrm{miss}}\right)}. \tag{4.5.3}$$

At least one of the two leptons is required to satisfy $m_{\mathrm{T}}^{\ell} > 50\,\mathrm{GeV}$. Distributions of both of these quantities are shown in Fig. 4.16.

The upper bounds on $m_{\ell\ell}$ and the azimuthal angle between the leptons $\Delta\phi_{\ell\ell}$ take advantage of the spin and helicity correlations unique to the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay. Distributions of both of these quantities are shown in Fig. 4.17. The b-jet veto rejects b-tagged jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$.

4.5.2 Background estimation

The diverse spectrum of background processes impacting the analysis is tackled by a wide range of different approaches. Many processes, particularly involving VV processes, have a small cross section and are modeled sufficiently well by the Monte Carlo generator and theoretical cross section predictions. On the other hand, backgrounds which contribute significantly to the analysis are estimated from data as much as possible. Separate normalization factors are derived in the $n_{\mathrm{jet}} = 0$ and $n_{\mathrm{jet}} = 1$ categories.

This analysis makes use of several control regions in order to estimate the normalization factors of important background contributions from data. Table 4.11 gives an overview of these control regions. Their definitions are summarized in Table 4.12, and their resulting composition from the various sources of background processes contributing to this analysis is summarized in Fig. 4.18. The total event yields of the various processes contributing to the control and signal regions are shown in Table 4.13. The general strategy is heavily influenced by earlier studies [119]. The W+jets background contribution is estimated using the fake factor method explained in Section 4.3.

Table 4.11.: Background estimation strategies in the ggF analysis. This table differentiates between fully data-driven methods ("Data"), control regions used to measure the normalization by full inclusion to the likelihood fit ("CR"), and normalization regions used in estimation methods specifically tailored for the respective cases ("NR").

Category	WW	Top	$Z{\to}ee/\mu\mu/\tau\tau$	W+jets/QCD	$V\gamma$
$n_{\mathrm{jet}} = 0$	CR	NR	CR	Data	NR
$n_{\mathrm{jet}} = 1$	CR	CR	CR	Data	NR

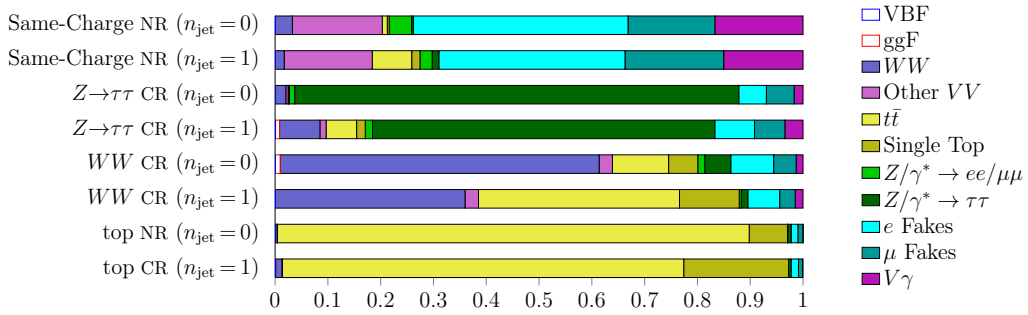


Figure 4.18.: Compositions of all control and normalization regions. The normalization factors from Table 4.14 have been applied.

Table 4.12.: Event selection criteria used to define the control regions. Every control region starts from the preselection defined in Table 4.10, unless indicated otherwise.

CR	$n_{\text{jet}} = 0$	$n_{\text{jet}} = 1$
WW	$55 < m_{\ell\ell} < 110\,\text{GeV}$ $\Delta\phi_{\ell\ell} < 2.6$	$m_{\ell\ell} > 80\,\text{GeV}$ $\lvert m_{\tau\tau} - m_Z \rvert > 25\,\text{GeV}$ b-jet veto $m_{\text{T}}^{\ell} > 50\,\text{GeV}$
top quark	no n_{jet} requirement $\Delta\phi_{\ell\ell} < 2.8$ $p_{\text{T}}^{\ell\ell} > 30\,\text{GeV}$	$n_{b\text{-jet}} = 1$ $\max(m_{\text{T}}^{\ell}) > 50\,\text{GeV}$ $m_{\tau\tau} < m_Z - 25\,\text{GeV}$
top quark aux.	no n_{jet} requirement ≥ 1 b-jet required $\Delta\phi_{\ell\ell} < 2.8$ $p_{\text{T}}^{\ell\ell} > 30\,\text{GeV}$	$n_{\text{jet}} = 2$ ≥ 1 b-jet required
$Z \to \tau\tau$	no $p_{\text{T}}^{\text{miss}}$ requirement $m_{\ell\ell} < 80\,\text{GeV}$ $\Delta\phi_{\ell\ell} > 2.8$	no $p_{\text{T}}^{\text{miss}}$ requirement $m_{\ell\ell} < 80\,\text{GeV}$ $m_{\tau\tau} > m_Z - 25\,\text{GeV}$ b-jet veto
VV	same-charge leptons all SR cuts	same-charge leptons all SR cuts

Table 4.13.: Event yields in the various control regions, normalization regions and signal regions of the gluon fusion analysis. The normalization factors from Table 4.14 have been applied.

	VBF	ggF	WW	Other VV	$t\bar{t}$	Single Top	$Z/\gamma^* \to ee/\mu\mu$	$Z/\gamma^* \to \tau\tau$	e Fakes	μ Fakes	$V\gamma$	Total Bkg.	Data/Bkg.	Data
Signal Region ($n_{\text{jet}} = 0$)	0.920 ± 0.060	85.0 ± 1.5	528.0 ± 5.0	32.9 ± 2.0	74.7 ± 2.7	37.9 ± 1.7	3.5 ± 3.5	0.90 ± 0.50	91.0 ± 6.0	39.0 ± 4.0	46.0 ± 9.0	854 ± 13	1.090 ± 0.040	930
Signal Region ($n_{\text{jet}} = 1$)	4.82 ± 0.14	60.0 ± 1.2	208.8 ± 2.8	28.1 ± 2.3	183.0 ± 4.0	65.8 ± 2.4	0.70 ± 0.70	13.0 ± 4.0	50.0 ± 5.0	22.9 ± 3.1	36.0 ± 6.0	607 ± 11	1.080 ± 0.050	657
Same-Charge NR ($n_{\text{jet}} = 0$)	0.0020 ± 0.0014	0.330 ± 0.080	25.8 ± 1.1	136.0 ± 4.0	7.60 ± 0.90	2.80 ± 0.50	34 ± 10	2.7 ± 1.5	324 ± 14	131.0 ± 6.0	133 ± 12	796 ± 22	0.900 ± 0.040	718
Same-Charge NR ($n_{\text{jet}} = 1$)	0.047 ± 0.013	0.55 ± 0.13	14.10 ± 0.70	142.0 ± 4.0	63.6 ± 2.6	13.3 ± 1.1	20.0 ± 6.0	11.0 ± 4.0	300 ± 12	159.0 ± 7.0	128 ± 13	851 ± 20	0.860 ± 0.040	736
$Z \to \tau\tau$ CR ($n_{\text{jet}} = 0$)	0.077 ± 0.019	6.70 ± 0.40	162.2 ± 2.9	36.2 ± 2.4	9.7 ± 0.1	6.60 ± 0.80	86 ± 19	6980 ± 200	434 ± 20	436 ± 30	140 ± 13	8300 ± 200	1.004 ± 0.027	8325
$Z \to \tau\tau$ CR ($n_{\text{jet}} = 1$)	0.620 ± 0.050	5.50 ± 0.40	55.7 ± 1.5	8.6 ± 1.1	42.1 ± 2.0	11.9 ± 1.0	10.0 ± 6.0	472 ± 32	55.0 ± 6.0	42.0 ± 6.0	25.0 ± 6.0	723 ± 34	1.000 ± 0.060	726
WW CR ($n_{\text{jet}} = 0$)	0.161 ± 0.023	13.60 ± 0.60	840.0 ± 6.0	34.9 ± 2.0	148.0 ± 4.0	76.9 ± 2.5	18 ± 10	69 ± 20	113.0 ± 6.0	59.0 ± 5.0	18.0 ± 4.0	1377 ± 26	1.002 ± 0.033	1380
WW CR ($n_{\text{jet}} = 1$)	0.031 ± 0.011	0.39 ± 0.10	678.0 ± 5.0	48.1 ± 2.5	718.0 ± 8.0	213.0 ± 4.0	7.0 ± 4.0	24.0 ± 6.0	114.0 ± 7.0	55.0 ± 4.0	28.0 ± 5.0	1886 ± 16	1.005 ± 0.025	1896
top NR ($n_{\text{jet}} = 0$)	2.54 ± 0.01	12.50 ± 0.60	136.1 ± 2.4	25.5 ± 2.1	35920 ± 60	2920 ± 16	10.0 ± 4.0	263 ± 19	515 ± 19	373 ± 14	17.5 ± 3.5	40180 ± 70	1.0130 ± 0.0050	40700
top CR ($n_{\text{jet}} = 1$)	0.325 ± 0.035	4.50 ± 0.40	45.3 ± 1.4	3.20 ± 0.60	2973 ± 18	776.0 ± 8.0	0	19.0 ± 6.0	55.0 ± 7.0	30.0 ± 4.0	3.8 ± 2.0	3905 ± 22	0.965 ± 0.017	3770

In this analysis, the contributions from VBF and VH production are treated as background processes. Their respective cross sections, branching ratios and the corresponding acceptance are assumed to have their nominal Standard Model values, assuming a Higgs boson mass of $m_H = 125\,\text{GeV}$. Contributions from $H \to \tau\tau$ decays are negligible, with an expected yield of less than 0.01 events in the final signal region.

Standard control regions

Standard control regions are used to normalize the non-resonant WW background. They are constructed mainly altering the signal region cuts specifically selecting the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ topology. In the $n_{\text{jet}} = 1$ category, additional cuts are imposed to protect the WW control region from the dominant top quark background.

The contribution from $Z \to \tau\tau$ events in the signal region is estimated with dedicated control regions. The contributions from $Z \to ee$ and $Z \to \mu\mu$ are much smaller due to the requirement of two different-flavour leptons. Hence, no dedicated control region is constructed for these contributions. Instead, they are included in the $Z \to \tau\tau$ control region, and the normalization factor is calculated for and applied to the combination of all $Z \to ee/\mu\mu/\tau\tau$ events.

For the top-quark background, different strategies are employed for the $n_{\text{jet}} = 0$ and $n_{\text{jet}} = 1$ categories. For the $n_{\text{jet}} = 1$ category, a standard control region can be defined by inverting the b-jet veto. The corresponding normalization factor can then be derived from the likelihood fit. However, this is conceptually impossible in the $n_{\text{jet}} = 0$ channel. Hence, a different strategy is employed for the $n_{\text{jet}} = 0$ top normalization factor derivation.

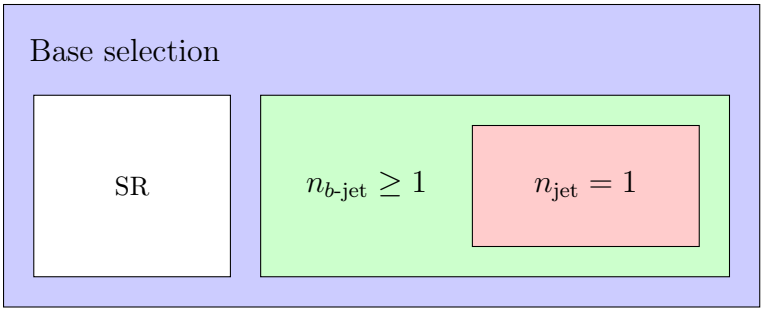


Figure 4.19.: Sketch of the regions used for the top-quark estimate in the $n_{\text{jet}} = 0$ category. The jet veto survival probability α_{JVSP} is the inverse ratio between the $n_{b\text{-jet}} \geq 1$ selection (green) and its $n_{\text{jet}} = 1$ subset (red).

Top estimate in the $n_{\text{jet}} = 0$ category

As a first step, a very inclusive *base region* is defined, including all events passing the preselection and the $\Delta\phi_{\ell\ell} < 2.8$ requirement, but imposing no requirements on the number of jets. Notably, this region includes the signal region as well, but is heavily dominated by events originating from $t\bar{t}$ processes, *i. e.*, events with two true heavy-flavour jets.

Events that have acquired at least one b-tag are thus very likely to have another true b-jet, even if that second jet was not successfully reconstructed or tagged. In such a scenario, a *tag-and-probe* method can be used to determine the probability of a true b-jet to be reconstructed *in situ*. Considering the number $N_{\geq 1\text{j}\geq 1\text{b}}$ of events with at least one b-tag, and comparing it to the number $N_{1\text{j}1\text{b}}$ of events with exactly one jet where this lone jet has been b-tagged, one can estimate the probability for a b-jet to fail being selected as a jet by the analysis requirements. This quantity is the *jet veto survival probability*

$$\alpha_{\text{JVSP}} = \frac{N_{1\text{j}1\text{b}}}{N_{\geq 1\text{j}\geq 1\text{b}}}. \tag{4.5.4}$$

A sketch of the regions included in this estimate can be seen in Fig. 4.19.

In order to model the failure to select two heavy-flavour jets in a single event in the inclusive base region, the square of the jet veto survival probability needs to be employed, such that

$$N_{\text{top}}^{\text{base}} = N^{\text{base}} \cdot \alpha_{\text{JVSP}}^2. \tag{4.5.5}$$

The estimate can be expressed as a top quark normalization factor for the $n_{\text{jet}} = 0$ category by dividing the prediction obtained in this way on observed data by the corresponding prediction based on Monte Carlo only, *i. e.*,

$$\beta_{\text{top}}^{n_{\text{jet}}=0} = \frac{N_{\text{obs.}}^{\text{base}}}{N_{\text{MC}}^{\text{base}}} \cdot \frac{\alpha_{\text{JVSP,obs.}}^2}{\alpha_{\text{JVSP,MC}}^2}. \tag{4.5.6}$$

Here, the event yields from non-top-quark events are subtracted from event yields in data using the respective estimates.

Top estimate in the $n_{\text{jet}} = 1$ category

A similar method has been developed to estimate the top contribution in the $n_{\text{jet}} = 1$ signal region. However, this method is only used as a cross check and does not contribute to the final result.

Again, an inclusive base selection is defined, this time requiring exactly two jets, at least one of which is b-tagged.

The efficiency ϵ for the kth b-tag in an event can then be derived as

$$\epsilon_k = \frac{N_{kjkb}}{N_{kjkb} + N_{kj(k-1)b}}, \tag{4.5.7}$$

where N_{kjlb} denotes the number of events with k jets, l of which are b-tagged.

This can be used to obtain an estimate for the *in-situ* b-tagging efficiency $\widetilde{\epsilon}$ by comparing Monte Carlo simulated samples to the corresponding value in data, that is,

$$\widetilde{\epsilon} = \frac{\epsilon_1^{\mathrm{MC}}}{\epsilon_2^{\mathrm{MC}}} \cdot \epsilon_2^{\mathrm{obs.}}. \tag{4.5.8}$$

This estimate can then be employed to extrapolate from the data control region with one jet and one b-tag to the contributions in the corresponding $n_{\mathrm{jet}} = 1$ signal region without any b-tag as

$$N_{\mathrm{top}}^{\mathrm{SR1j}} = N_{\mathrm{1j1b}}^{\mathrm{obs.}} \cdot \alpha \quad \text{where} \quad \alpha = \frac{1 - \widetilde{\epsilon}}{\widetilde{\epsilon}}, \tag{4.5.9}$$

or expressed as a normalization factor,

$$\beta_{\mathrm{top}}^{n_{\mathrm{jet}}=1} = \frac{N_{\mathrm{1j1b}}^{\mathrm{obs.}}}{N_{\mathrm{1j0b}}^{\mathrm{MC}}} \cdot \alpha. \tag{4.5.10}$$

This data-driven method in principle provides the capability to avoid or reduce systematic uncertainties connected to the b-tagging efficiency, as it effectively provides an *in-situ* data-driven estimate. However, incorporation of this complicated technique into the standard likelihood fitting procedure is technically challenging, hence it is only used as a cross check in the analysis presented here.

Same-charge normalization region

The samples modeling $V\gamma$ processes in this analysis (with V denoting the heavy weak bosons W and Z) have only been generated to leading order precision. The control region used to normalize this process is identical to the signal region, but requires the two leptons to have the same instead of opposite charge. As this control region only captures very few events, it is not useful for the normalization of $V\gamma^*/VZ$, as the corresponding cross section is known to NLO accuracy, which provides a smaller uncertainty than the statistical power of the control region would allow to derive. The normalization derived from this control region is hence only applied to $V\gamma$.

The normalization factor derived from this control region is calculated separately, as the different fake composition in the same-charge region due to the lower fraction

of Wc events combined with the low statistical power of the region pose a challenge for the consistent treatment of uncertainties in the likelihood fit.

Summary

An overview over all normalization factors derived with the matrix method for the classical control regions as well as the dedicated methods presented here is shown in Table 4.14, including the statistical uncertainties.

The combination of different estimates is not trivially possible, as the individual estimates can influence each other via the impurities of the respective control regions. Thus, the following strategy has been chosen: The top quark estimate in the respective jet multiplicity category is performed first, as the associated control regions are highly pure. Then, a 2×2 matrix method is performed on the WW and $Z \to ee/\mu\mu/\tau\tau$ control regions, incorporating the top estimate. Finally, a simple normalization factor is computed for $V\gamma$ from the same-charge normalization region, the rationale being that the limited statistical power of that region makes it highly prone to fluctuations, and any cross-talk to the much more strongly constrained normalization factors for the other backgrounds will only artificially enhance the associated uncertainties.

The uncertainties on the normalization factors have been derived using a set of randomized toy data sets. As the combination of different methods in the same estimate is challenging without the use of a full likelihood fit, an iterative approach has been chosen to arrive at a consistent estimate, repeating the derivation of correction factors with any method on an input data set taking into account the already derived normalization factors for other processes until the technique converges to a consistent set of stable normalization factors. These numbers are used to produce the event distributions shown in this section, unless indicated otherwise.

Of the values presented above, however, only the values derived for $V\gamma$ and for the top contribution in the $n_{\text{jet}} = 0$ category are used in the final signal extraction. The normalization factors for the top background in the $n_{\text{jet}} = 1$ category as well as the normalization factors for WW and $Z \to ee/\mu\mu/\tau\tau$ are left floating as unconstrained nuisance parameters during the likelihood maximization.

Table 4.14.: Background normalization factors obtained from the control and normalization regions using the methods explained in the text, including statistical uncertainties.

Control Regions	WW	Top	$Z \to ee/\mu\mu/\tau\tau$	$V\gamma$
$n_{\text{jet}} = 0$	1.110 ± 0.060	0.911 ± 0.025	0.952 ± 0.030	1.41 ± 0.19
$n_{\text{jet}} = 1$	0.880 ± 0.070	1.028 ± 0.031	0.780 ± 0.070	1.34 ± 0.24

4.5.3 Experimental systematic uncertainties

The handling of experimental systematic uncertainties is very similar between all the analyses presented in this chapter. This section thus contains a general summary of the treatment, as well as a dedicated section detailing the special choices made in the gluon fusion analysis.

The experimental systematic uncertainties are applied as $\pm 1\sigma$ variations to the four-momenta of an object, dubbed *four-vector* or *P4 systematics*, or to the event or particle weight, dubbed *scale factor* or *SF systematics*.

Lepton-related uncertainties

In order to estimate the uncertainties related to the reconstruction and identification of leptons as well as to the associated momentum and energy scale and resolution uncertainties and isolation criteria, decays of $J/\psi \to \ell^{+}\ell^{-}$, $W^{\pm} \to \ell^{\pm}\nu$ and $Z \to \ell^{+}\ell^{-}$ are studied [101, 144].

Table 4.15.: Experimental systematic uncertainties connected to muons. Whether the respective uncertainty is modeled as a scale factor or four-momentum variation is indicated by the last column.

resolution	MUONS_ID, MUONS_MS	track momentum smearing in the inner detector and muon spectrometer	P4
scale	MUONS_SCALE	variations in muon momentum scale	P4
identification	MUON_EFF_STAT, MUON_EFF_SYS	statistical and systematic variations in identification efficiency	SF
isolation	MUON_ISO_STAT, MUON_ISO_SYS	statistical and systematic variations in isolation efficiency	SF
trigger	MUON_EFF_TrigUncertainty_STAT, MUON_EFF_TrigUncertainty_SYS	statistical and statistical variations in trigger efficiency	SF

The muon momentum resolution and scale calibration are derived from a simulation template fit, comparing the mass of $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ candidates in data and Monte Carlo simulation [145]. All parameters modeling these uncertainties are listed in Table 4.15 for muons and in Table 4.16 for electrons.

Jet-related uncertainties

The uncertainties on the jet energy scale and resolution are derived as functions of the transverse momentum and pseudorapidity of the jet, as well as the pile-up conditions and jet flavour composition. They are determined using a combination of simulated and data samples, as measurements of the jet-response balance in di-jet, Z+jets and γ+jet events, as well as for high-p_{T} multi-jet events [146].

Table 4.16.: Experimental systematic uncertainties connected to electrons. Whether the respective uncertainty is modeled as a scale factor or four-momentum variation is indicated by the last column.

resolution	EG_RESOLUTION_ALL	track smearing in the inner detector	P4
scale	EG_SCALE_ALLCORR, EG_SCALE_E4SCINTILLATOR, EG_SCALE_LARCALIB_EXTRA2015PRE, EG_SCALE_LARTEMPERATURE_EXTRA2015PRE, EG_SCALE_LARTEMPERATURE_EXTRA2016PRE	variations in momentum scale originating from different sources	P4
identification	EL_EFF_ID_CorrUncertaintyNP[0:14], EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP[0:15]	correlated and uncorrelated variations in identification efficiency	SF
isolation	EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	variations in isolation efficiency	SF
reconstruction	EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	variations in reconstruction efficiency	SF
trigger	EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	variations in trigger efficiency	SF

Table 4.17.: Jet-related systematic uncertainty parameters. The last column indicates whether the individual uncertainty is modeled using four-momentum variations or scale factor variations.

JER	Single nuisance parameter covering jet energy resolution uncertainties	P4
JES_AFII	Uncertainty arising from the use of the ATLFast-II fast simulation framework	P4
JES_BJES	Heavy-flavour jet energy scale uncertainty.	P4
JES_EffectiveNP_[1-6]	Linear decomposition of jet energy scale uncertainties.	P4
JES_EtaInter_Model, JES_EtaInter_NonClosure, JES_EtaInter_Stat	Uncertainties covering η-dependence of the jet energy scale	P4
JES_Flavor_Comp, JES_Flavor_Comp_nonTop, JES_Flavor_Comp_top, JES_Flavor_Resp	Uncertainties covering flavour-dependence of the jet energy scale.	P4
JES_HighPt	High-p_{T} jet energy scale uncertainty	P4
JES_PU_OffsetMu, JES_PU_OffsetNPV, JES_PU_PtTerm, JES_PU_Rho, JES_PU_Rho_nonTop, JES_PU_Rho_top	Uncertainties covering the effects of pile-up on the jet energy scale.	P4
JES_PunchThrough	Uncertainty covering effects of the calorimeter not covering the entirety of the shower due to insufficient thickness	P4
JVT	Uncertainty on jet vertex tagging	SF

The total uncertainty of the jet energy scale is split into 19 individual nuisance parameters to be determined by the likelihood fit *in-situ*, providing full treatment of bin-to-bin correlations. The jet energy resolution on the other hand is modeled by a single nuisance parameter. All parameters are listed in Table 4.17.

Flavour tagging uncertainties

Events which are likely to originate from events involving top quarks are suppressed in this analysis by vetoing events with at least one b-tagged jet. Conversely, events with b-tagged jets are used to construct the top control regions.

Table 4.18.: Experimental systematic uncertainties connected to flavour tagging. All of these uncertainties are modeled as scale factor uncertainties.

`FT_EFF_Eigen_B_[0:2]`	Eigen-vector decomposition of b-tagging uncertainties
`FT_EFF_Eigen_C_[0:3]`	Eigen-vector decomposition of c-tagging uncertainties
`FT_EFF_Eigen_Light_[0:4]`	Eigen-vector decomposition of light flavour tagging uncertainties
`FT_EFF_extrapolation`	uncertainty covering extrapolation from Run 1 to Run 2
`FT_EFF_extrapolation_from_charm`	charm-quark specific extrapolation uncertainty

The connected uncertainties are modeled by a set of parameters corresponding to the eigen-vector decomposition of the corresponding parameter space [106], as well as some additional parameters. The parameters modeling these flavour-tagging related uncertainties are listed in Table 4.18.

Missing transverse momentum

The missing transverse momentum is one of the most important variables in this analysis. Not only is it used in the construction of the signal region to remove events without high-energy neutrinos, it is also one component entering the calculation of the transverse mass, the final discriminant used during the likelihood fit.

Table 4.19.: The systematic variations connected to the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term. All of these uncertainties are modeled as scale factor uncertainties.

`MET_SoftTrk_ResoPara, MET_SoftTrk_ResoPerp`	Soft term resolution
`MET_SoftTrk_Scale`	Soft term scale

The uncertainties related to the measurement of the track-soft-term missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ have been evaluated using different Monte Carlo generators [146]. They are listed in Table 4.19 and are dependent on high-p_{T} activity in the event, quantized by the $\vec{p}_{\mathrm{T}}^{\,\mathrm{hard}}$ variable defined as the missing transverse momentum calculated only taking into account high-p_{T} objects from the event. The direction of this vector is used to discriminate between effects affecting the scale and the

resolution, the first being primarily longitudinal to the $\vec{p}_{\mathrm{T}}^{\,\mathrm{hard}}$ axis and the latter being both longitudinal and transverse. These uncertainties were measured in $Z \to \mu\mu$ events, and take into account uncertainties on detector material effects.

Other sources of uncertainty

The uncertainties connected to the estimate of background arising from misidentified leptons are discussed in Section 4.3.4. The individual parameter names and their physical interpretation are listed in Table 4.20.

Table 4.20.: Systematic uncertainty parameters connected to the fake factor estimate as presented in Section 4.3.4.

HWW_FakeFactor_[el/mu]_EWSUBTR	Uncertainty arising from the electroweak subtraction for fake electrons/muons
HWW_FakeFactor_[el/mu]_SAMPLECOMPOSITION	Uncertainty arising from the sample composition variance for fake electrons/muons
HWW_FakeFactor_[el/mu]_WJETSSIGN	Additional uncertainty covering lepton charge dependency of the fake composition for fake electrons/muons
HWW_FakeFactor_[el/mu]_STAT_[2015/2016]_[i]_[j]	Statistical uncertainty on the fake factor in p_{T}-bin i and $\lvert\eta\rvert$-bin j for fake electrons/muons in the 2015/2016 data set

The Monte Carlo samples are generated with a generic spectrum of average interactions per bunch crossing. As this does usually not reflect the data taking conditions, a reweighting is applied to alter the pile-up distribution in Monte Carlo to coincide with the one measured in data. The uncertainty connected to this reweighting is incorporated in the form of a scale factor uncertainty, yielding uncertainties of $+6\%/-14\%$.

The integrated luminosity of the data set is estimated from a preliminary calibration [147] of the luminosity scale using *van der Meer* beam separation scans [148] performed during 2015 and 2016. The relative uncertainties on the luminosity are $\pm 2.1\%$ and $\pm 3.7\%$ for the two years, respectively. This results in an uncertainty of $\pm 2.0\%$ on the total luminosity.

4.5.4 Theoretical uncertainties

For the signal and the backgrounds normalized from theory predictions, uncertainties on the absolute yields in each control and signal region are taken into account, including migrations between the different signal regions. For the backgrounds which are normalized by control regions, theoretical uncertainties on the extrapolation from the control to the signal region are taken into account instead.

The theoretical uncertainties considered include QCD scale variations, uncertainties on modeling of the parton shower and underlying event (PS/UE), as well as

variations of the PDF set. Some other uncertainties specific to certain contributions are also included, like the uncertainty introduced to the gluon fusion prediction in the signal region by the categorization in the number of jets. An overview is given in Table 4.21.

For the QCD scale uncertainties, the renormalization and factorization scales are varied independently by a factor of 2.0 and 0.5 relative to the nominal value.

For the determination of PDF model uncertainties, an envelope of the 68% confidence level CT10 PDF Eigen vector variations and the differences of the CT10 to the MSTW2008 and the NNPDF3.0 PDF sets are added in quadrature. All these computations are performed at Monte Carlo truth level using event weights only.

In the derivation of these uncertainties, some calculations suffer heavily from the limited size of the Monte Carlo samples. Whenever the systematic variation was not significant with respect to the statistical power of the sample in question, the uncertainty was dropped from the calculation.

Higgs boson production via gluon fusion

The signal cross section uncertainties for the gluon fusion analysis are calculated separately for the $n_{\text{jet}} = 0$ and $n_{\text{jet}} = 1$ categories. Uncertainties on the predictions are evaluated by applying variations to the QCD scales and the PDF set used in the signal sample generation. The PDF uncertainties are determined using POWHEG+PYTHIA8.

The PS/UE uncertainty in the ggF sample, which is showered with PYTHIA8 using the AZNLO tune, is estimated using four up- and down-variations of parameter sets in that tune. Additional evaluations include comparisons between POWHEG+PYTHIA8 and POWHEG+HERWIG7 [149] using the UEEE5 tune [150]. In order to correct the acceptance for resolution effects in the m_{T} distribution, a folding matrix is applied.

Additionally, an uncertainty on the jet-exclusive cross section is assigned. This uncertainty is evaluated using the Steward-Tackmann method [151], where the QCD scale uncertainties on the inclusive cross sections are assumed to be independent between the jet multiplicity categories.

Continuum WW production

The WW background is normalized using control regions separated from the signal regions in $m_{\ell\ell}$. The QCD scale uncertainties and PDF uncertainties are evaluated similar to the strategy used for gluon fusion samples, using POWHEG.

The PS/UE uncertainty is again computed by comparing the nominal POWHEG+PYTHIA8 sample with the AZNLO tune to the predictions from POWHEG interfaced to HERWIG++ using the UEEE5 tune. SHERPA and POWHEG+PYTHIA8 are compared to obtain estimates of the uncertainty of the NLO parton shower matching [152]. Both these computations were evaluated on Monte Carlo truth

Table 4.21.: Overview over the theory uncertainties included in the gluon fusion analysis. Most uncertainties have been specifically evaluated for this analysis. If the evaluation was adopted from an earlier publication, the source given in the corresponding column. The last column indicates whether the uncertainty is used in the likelihood fit.

process	uncertainty	source	included
ggF	jet veto uncertainty		✓
	matching		✓
	PS/UE		✓
	PDF		✓
	QCD scale		✓
WW	generator/matching		✓
	PS/UE		✓
	PDF		✓
	QCD scale		✓
	EW correction	[119]	✓
	$gg \to WW$ fraction	[119]	✓
$WZ/W\gamma^*$	QCD and merging scales		✓
$W\gamma$	NLO correction, QCD scale		✗
top	radiation		✓
	PS/UE		✓
	matching		✓
	Wt diagram removal		✓
	single top cross-section	[119]	✗
DY	alternative generators		✓

level, and combining the $e\mu$ and μe categories to enhance the statistical power. Additionally, the $m_{\ell\ell}$ categories were merged for $n_{\text{jet}} = 1$ in the estimate of the generator uncertainty.

The estimates for the systematic uncertainty on the electroweak corrections from theory and the for the extrapolation uncertainty due to the NNLO $gg \to WW$ calculation were adopted from older publications [119].

Other diboson production

The QCD renormalization, factorization and resummation scale were varied by a factor of two in either direction, and the matching scale was varied from 20 GeV up to 30 GeV and down to 15 GeV using Sherpa. As the statistical power of this estimate is very limited, all signal regions with the same jet multiplicity are merged. As these processes are normalized to the corresponding theoretical predictions, cross section uncertainties are also included.

The $W\gamma$ process is normalized by the same charge control region and assigned a normalization uncertainty accordingly.

Top backgrounds

The systematic uncertainties connected to QCD renormalization and factorization scales as well as to the PDF model are derived analogously to the other background processes. For the PS/UE uncertainties, POWHEG+PYTHIA6 is compared with POWHEG+HERWIG++. The NLO matching uncertainty is evaluated comparing POWHEG+HERWIG++ with MADGRAPH5_aMC@NLO+HERWIG++. Additionally, variations in radiation are evaluated using POWHEG+PYTHIA6. The uncertainty on the interference between $t\bar{t}$ and Wt contributions is assessed by comparing two different overlap subtraction strategies.

These estimates suffer from the limited statistical power of the Monte Carlo samples. For the $n_{\text{jet}} = 1$ signal region, fully simulated samples can be used. However, the $n_{\text{jet}} = 0$ top background estimate is split in two components. The uncertainty on the correction factor itself is estimated from samples that have undergone the full detector simulation to take into account flavour tagging uncertainties. The extrapolation uncertainty to the signal region is then measured using truth information Monte Carlo.

Z/Drell-Yan

The uncertainties arising from generator and PS/UE modeling for the Z/Drell-Yan background contribution are estimated by comparing the nominal, leading-order multi-leg MADGRAPH+PYTHIA8 sample, with an NLO POWHEG+PYTHIA8 sample and an NLO multi-leg SHERPA sample.

4.5.5 Statistical treatment

The event selection described in Section 4.5.1 selects two disjoint signal regions, one in the $n_{\text{jet}} = 0$ category, and one in the $n_{\text{jet}} = 1$ category. The distributions of events in transverse momentum p_{T} and pseudorapidity η of the leading and the subleading lepton observed in data and predicted for the signal and background processes by the various methods presented above are shown in Fig. 4.20.

The events passing the selection in either of these regions are further categorized into 16 signal regions in total, eight for every jet category. The categorization employs the transverse momentum of the subleading lepton, displayed in Fig. 4.20, and the invariant dilepton mass, displayed in Fig. 4.21. Five control regions are taken into account by the fit. Additionally, three normalization factors are precomputed from corresponding dedicated control regions.

In the control regions, only the total event yield is considered, whereas each signal region is subdivided in eight bins in the final discriminant variable, the

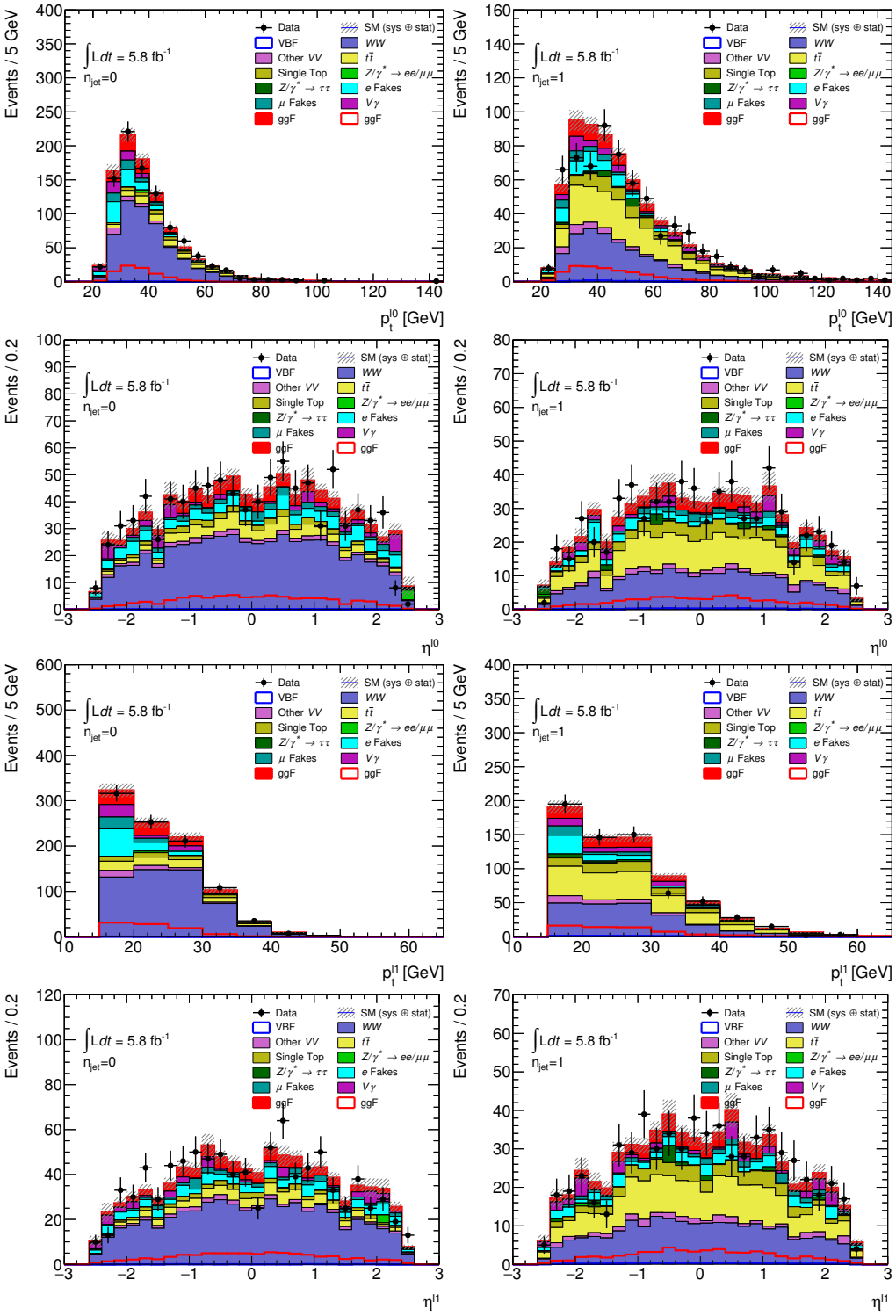


Figure 4.20.: Distributions of events in transverse momentum p_{T} and pseudorapidity η of the leading lepton ℓ_0 and subleading lepton ℓ_1 for the $n_{\mathrm{jet}} = 0$ (left) and the $n_{\mathrm{jet}} = 1$ (right) signal regions. The normalization factors from Table 4.14 have been applied.

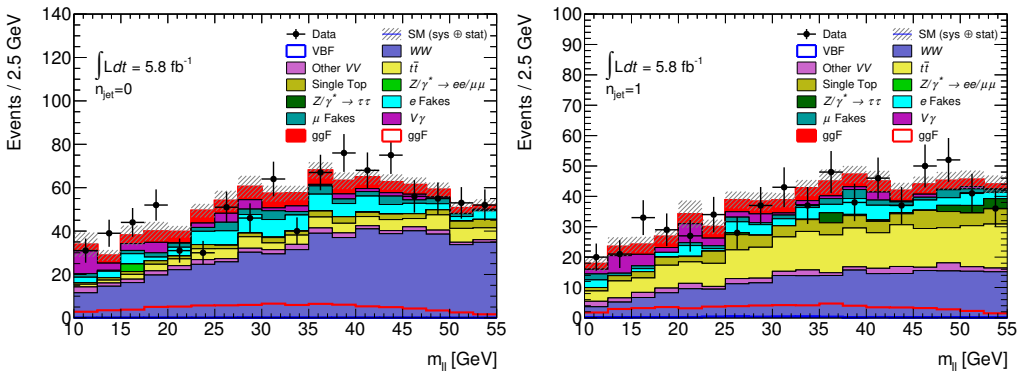


Figure 4.21.: Distributions of events in invariant dilepton mass $m_{\ell\ell}$ for the $n_{\text{jet}} = 0$ (left) and the $n_{\text{jet}} = 1$ (right) signal regions. The normalization factors from Table 4.14 have been applied.

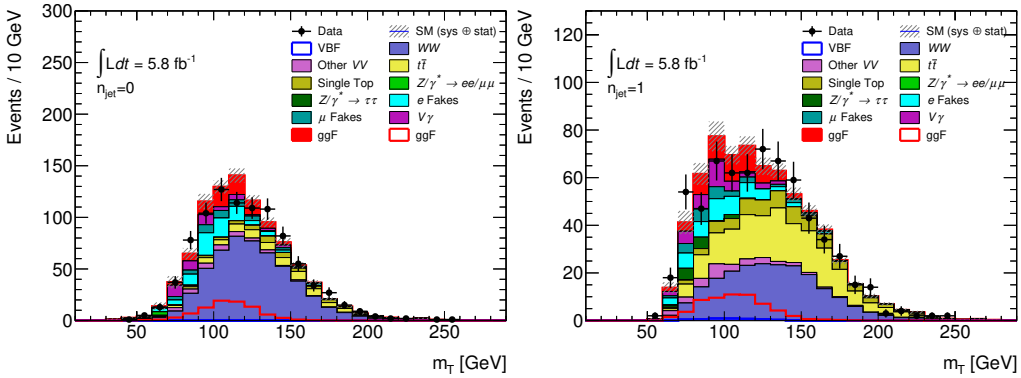


Figure 4.22.: Distributions of events in transverse mass m_{T} for the $n_{\text{jet}} = 0$ (left) and the $n_{\text{jet}} = 1$ (right) signal regions. The normalization factors from Table 4.14 have been applied.

transverse mass m_{T}. Distributions of events in transverse mass m_{T} including all data-driven corrections, but preceding the likelihood minimization procedure, are shown in Fig. 4.22. The original input histograms have 50 uniform bins, spanning the range $80\,\text{GeV} < m_{\text{T}} < 130\,\text{GeV}$, as well as two additional bins covering the underflow $m_{\text{T}} < 50\,\text{GeV}$ and the overflow $m_{\text{T}} > 80\,\text{GeV}$. In each signal region, these 52 bins are then remapped to eight bins, including underflow and overflow, using a heuristic tree-search algorithm such that the total signal yield is as flat as possible across the bins. The resulting bin boundaries in m_T for all signal regions are shown in Fig. 4.23.

For each systematic uncertainty, $\pm 1\sigma$ variations are considered where available, with the nuisance parameter interpolating between the corresponding effects using the method of vertical interpolation, which is discussed in greater detail in Section 5.2.3. Additional parameters, notably the background normalization

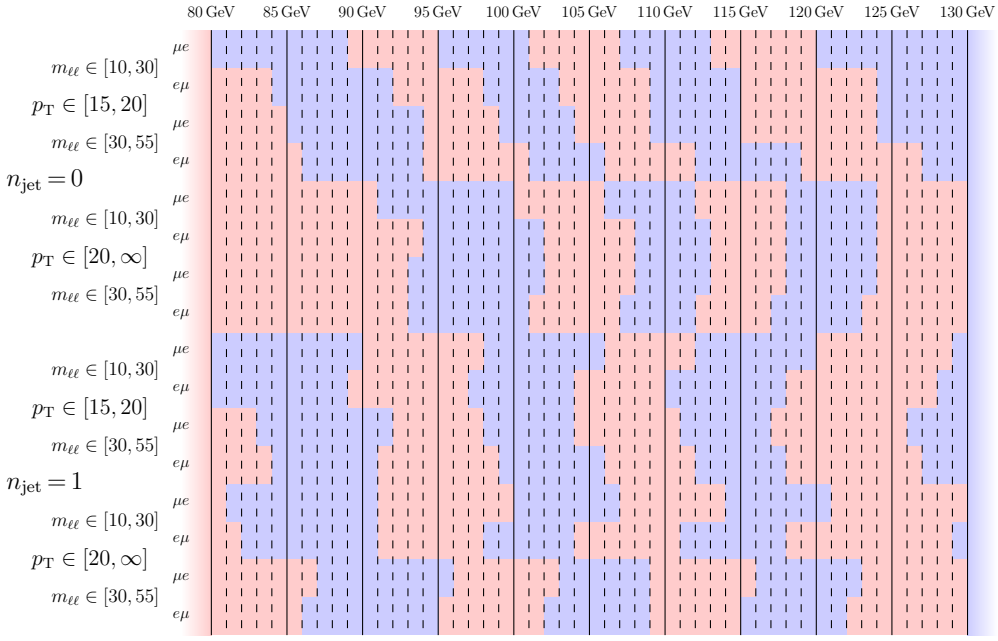


Figure 4.23.: Bin boundaries in transverse mass m_T as determined by the heuristic optimization algorithm described in the text. The under- and overflow areas are indicated by the shading on the sides.

parameters and the parameter of interest, the signal strength parameter μ, are acting on their associated contributions as event yield multipliers.

The likelihood function is then constructed as the product of Poisson terms over all bins in all signal regions as well as the control regions, and additional Gaussian or log-normal shaped terms are multiplied to model constraints on the nuisance parameters. For most parameters, Gaussian shaped terms are used, whereas log-normal shaped terms are used for parameters that have large relative variances, but for which only positive values are physical. With products running over the bins in the categories denoted in the superscript, the likelihood can be written as

$$\mathcal{L}\left(\mu, \vec{\theta}\right) = \prod_{l=e,\mu}^{\ell_2} \prod_{p=1,2}^{p_{\mathrm{T}}} \prod_{m=1,2}^{m_{\ell\ell}} \prod_{j=0,1}^{n_{\mathrm{jet}}} \prod_{i=1,\ldots,8}^{m_{\mathrm{T}}} P\left(N_{lpmji} \middle| \mu \cdot s_{lpmji}\left(\vec{\theta}\right) + \sum_{\Phi} \beta_{\Phi} b^{\Phi}_{lpmji}\left(\vec{\theta}\right)\right) \prod_{\theta \in \vec{\theta}} C(\theta)\,.$$

Here, s and b denote the predicted signal and background yields in any specific category, and Φ denotes a specific background process with the predicted yield b^{Φ} and the normalization factor β_{Φ}. For convenience, normalization factors have been included for all background processes, but they are fixed to unity for most processes. Further, $P(x \mid y)$ denotes the Poisson probability of observing x events

when expecting y, and $C(\theta)$ denotes a constraint term acting on the nuisance parameter θ.

Handling of systematic uncertainties

In case only the event weights are changed for some systematic variation, the event yields are fully correlated to the nominal yields. Systematic variations changing the four-momentum, however, can cause migrations into or out of a signal or control region. For some samples with small statistical power, this can create artificially huge and unphysical variations.

This problem is especially prominent when individual Monte Carlo events acquire very large (or very small) event weights. The latter effect can be found for the variation of the pile-up scale factor, which can effectively remove events from a region by changing their weight to zero.

In order to overcome the issues of limited statistical power of the Monte Carlo predictions, the effects of systematic variations of experimental nuisance parameters in the signal regions are determined based on the $qq \to WW$ contribution, which provides good statistical power, while being topologically similar to most of the other background processes. The relative changes of the normalization of each signal region calculated from $qq \to WW$ are then applied to the other background processes to model the effects of systematic uncertainties. Exempt from this treatment are the data-driven W+jets background, as well as the statistically powerful and topologically different top quark related backgrounds in the $n_{\text{jet}} = 1$ category.

As systematic variations of experimental nuisance parameters affecting the four-momenta of objects may be subject to large fluctuations due to bin migrations, they are purged if they cannot be determined with sufficient confidence. Parameters are removed from a combination of region and sample if neither of the two variational templates would exhibit a variation of at least 0.2σ with respect to the statistical uncertainty.

Systematic variations of the normalization are neglected and removed from the model in any specific region for any specific sample if the corresponding effect on the yield of this sample in that region is smaller than 0.5%. This threshold is only applied for experimental systematics, for which only normalization effects are considered. Systematic uncertainties from theory, which can also effect the shapes, have no such threshold applied.

For some variations, however, the effect was found to be exactly zero. These parameters were removed from the model to avoid problems with the numerical minimization.

The effects of most nuisance parameters are modeled by templates corresponding to the $\pm 1\sigma$ variations of the respective parameter. The effect on the respective region is then modeled by using the method of vertical interpolation, which is

introduced in Section 5.2.3. Here, differences between the variational templates and the nominal prediction are computed. The variational templates are usually expected to have opposite effects on the yield in the corresponding region, that is, if the *up*-variation of one parameter increases the yield of a process in some region, the *down*-variation is expected to decrease it. However, statistical fluctuations in the variation templates can cause these variations to be highly imbalanced, such that the apparent effect on the yield can be zero or even have the wrong sign. In such cases, the templates are symmetrized using the following strategy:

- The relative effect of the nuisance parameter is conservatively determined as the envelope of the up- and the down-variation.
- The sign of the variations is assigned such that the variation for which the deviation from the nominal prediction is more significant in terms of its statistical power is kept, whereas the inverse sign is chosen for the other variation.

The full summary of treatments of all systematic uncertainties in all regions and for all constrained nuisance parameters is shown in Appendix A.1.

4.5.6 Results

The analysis presented here represents the first measurement of gluon fusion Higgs boson production in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay channel using pp collision data collected at $\sqrt{s} = 13\,\text{TeV}$ by the ATLAS experiment.

A signal has been measured with a significance of 1.01σ, where 2.78σ was expected based on the Standard Model prediction. The observed significance corresponds to a p-value of $p_0 = 0.156$. The signal strength parameter μ was measured as

$$\begin{aligned}\mu &= 0.37\,^{+0.39}_{-0.37}\,(\text{tot.}) \\ &= 0.37\,^{+0.32}_{-0.29}\,(\text{exp. sys.})\,\pm 0.23\,(\text{stat.})\,^{+0.16}_{-0.13}\,(\text{theo.}) \\ &\qquad {}^{+0.13}_{-0.14}\,(\text{norm.}) \quad {}^{+0.12}_{-0.10}\,(\text{fake}) \quad \pm 0.10\,(\text{MC stat.}),\end{aligned}$$

where the individual contributions to the uncertainties have been disentangled into the component stemming from experimental systematic uncertainties (exp. sys.), from statistical uncertainties on the data (stat.), from systematic uncertainties from theory (theo.), from the floating background normalization factors (norm.), from the fake estimate (fake), and from statistical uncertainties arising from the finite size of the Monte Carlo samples (MC stat.).

The final distributions in transverse mass m_{T} in all signal and control regions directly included in the likelihood fit are shown in Fig. 4.24. Here, the signal and background normalization factors and all systematic uncertainty parameters have been fixed to their best-fit values, and the uncertainty band includes all statistical

and systematic uncertainties. The corresponding numerical values of the event yields in each region are shown in Table 4.22.

A listing of the results for the individual parameters obtained in the fit is shown in Fig. 4.25. The total uncertainty has been broken down to the impacts of each individual nuisance parameter (and for certain groups of nuisance parameters). Shown are the breakdowns for the fit to the observed data, as well as for a control fit to an Asimov data set generated with the assumption of $\mu = 1$. The post-fit values and uncertainties of the individual nuisance parameters are shown as well, where the constrained nuisance parameters have been normalized to their respective constraints.

The correlations between the nuisance parameters, observed during the fit to data, are shown in Fig. 4.26. However, parameters for which no individual correlation exceeded a value of 10% are not shown.

The background normalization factors observed in the fit are shown in Table 4.23. This table contains the explicit values of the free normalization factor parameters used in the fit. Hence, the top quark estimate for the $n_{\text{jet}} = 0$ category and the $V\gamma$ normalization are not shown. These values do not include the effects of the constrained nuisance parameters shown in Fig. 4.25. Table 4.24 instead lists the *effective* background normalization factors observed in the signal region, *i. e.*, the ratio of yields of the individual samples before and after the fit, including indirect effects of constrained nuisance parameters. The normalization factors are compatible with the values shown in Table 4.14 derived from the independent normalization factor calculation within the (sometimes large) uncertainties. The differences between the different normalization factors can be attributed to the indirect influence of and correlations with constrained nuisance parameters.

Figure 4.27 shows a *profile likelihood scan* of the signal strength parameter μ. Here, the parameter has been artificially fixed to the value indicated on the horizontal axis, and a dedicated fit using this constraint has been performed for every point. The respective difference of the minimum with respect to the unconstrained fit is then shown on the vertical axis. In this representation, the limits on the parameter μ corresponding to 1σ, 2σ, and 3σ observed exclusion can be read off easily, as

$$Z(\mu) = \sqrt{2\Delta \log \mathcal{L}(\mu)}. \tag{4.5.11}$$

The likelihood curve indicates that the observation is compatible with the Standard Model prediction of $\mu = 1$ within 1.6σ.

Finally, Fig. 4.28 shows the final distribution of the transverse mass, with the full post-fit normalization factors applied, including the signal strength measurement. While the statistical fluctuations are still quite large from bin to bin, the emergence of the signal is clearly visible.

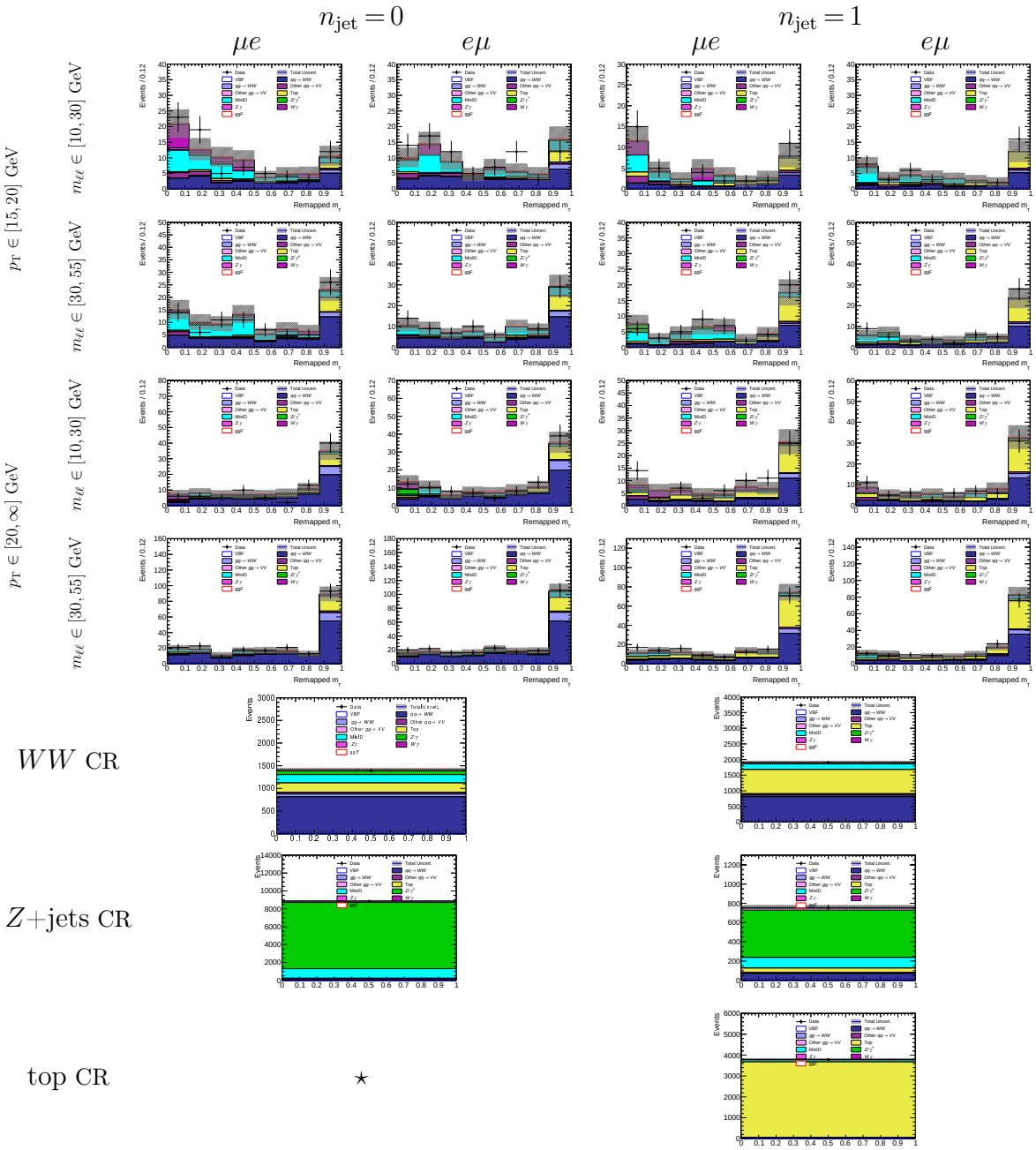


Figure 4.24.: Overview showing all regions included in the fit with their event yields as observed in data, and their composition as predicted by the various methods of background estimation. The observed signal is also shown and stacked on the background prediction.

Table 4.22.: Expected event yields in the signal regions for the different signal and background processes as predicted by the Standard Model in comparison with the corresponding observed values. The yields include the full post-fit normalization, as described in Table 4.23.

	ggF	Data	Total Bkg.	VBF	$qq \to WW$	$gg \to WW$	Other $qq \to VV$	Other $gg \to VV$	Top	$W\gamma$	$Z\gamma$	Z/γ^*	MisID
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = e$	3.0 ± 1.7	78	75.0 ± 9.0	0.09 ± 0.30	23.0 ± 5.0	2.1 ± 1.5	4.8 ± 2.2	0.000 ± 0.060	4.4 ± 2.1	14.0 ± 4.0	2.2 ± 1.5	0	24.0 ± 5.0
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = \mu$	3.2 ± 1.8	82	69.0 ± 8.0	0.08 ± 0.29	30.0 ± 5.0	3.0 ± 1.7	4.6 ± 2.1	0.010 ± 0.080	6.2 ± 2.5	4.5 ± 2.1	0.40 ± 0.60	0	20.0 ± 5.0
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = e$	4.0 ± 2.0	87	88.0 ± 9.0	0.10 ± 0.32	50.0 ± 7.0	8.6 ± 2.9	3.8 ± 1.9	0.000 ± 0.070	13.0 ± 4.0	5.0 ± 2.2	0.50 ± 0.70	0	6.7 ± 2.6
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = \mu$	4.9 ± 2.2	101	96 ± 1	0.20 ± 0.40	54.0 ± 7.0	8.1 ± 2.9	2.8 ± 1.7	0.01 ± 0.11	14.0 ± 4.0	5.1 ± 2.3	0.40 ± 0.60	2.8 ± 1.7	8.6 ± 2.9
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = e$	2.7 ± 1.6	85	88.0 ± 9.0	0.08 ± 0.29	36.0 ± 6.0	3.0 ± 1.7	2.9 ± 1.7	0.01 ± 0.10	8.7 ± 3.0	4.2 ± 2.0	1.3 ± 1.1	0.60 ± 0.80	31.0 ± 6.0
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = \mu$	3.3 ± 1.8	88	85.0 ± 9.0	0.12 ± 0.34	43.0 ± 7.0	4.3 ± 2.1	3.6 ± 1.9	0.01 ± 0.10	11.0 ± 3.3	0.70 ± 0.80	0.10 ± 0.31	0.30 ± 0.50	22.0 ± 5.0
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = e$	5.8 ± 2.4	215	198 ± 14	0.20 ± 0.40	131 ± 11	15.0 ± 4.0	6.5 ± 2.6	0.01 ± 0.11	27.0 ± 5.0	3.2 ± 1.8	0.20 ± 0.50	0	14.0 ± 4.0
SR $n_{\text{jet}} = 0$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = \mu$	6.5 ± 2.6	239	226 ± 15	0.20 ± 0.40	149 ± 12	18.0 ± 4.0	5.0 ± 2.2	0.02 ± 0.14	30.0 ± 5.0	2.4 ± 1.6	1.3 ± 1.1	0	19.0 ± 4.0
WW $n_{\text{jet}} = 0$ CR	5.0 ± 2.2	1389	1410 ± 40	0.20 ± 0.40	813 ± 29	62.0 ± 8.0	34.0 ± 6.0	0.07 ± 0.27	211 ± 15	13.0 ± 4.0	4.9 ± 2.2	86.0 ± 9.0	184 ± 14
Z+jets $n_{\text{jet}} = 0$ CR	2.5 ± 1.6	8859	8850 ± 90	0.08 ± 0.28	175 ± 13	5.1 ± 2.3	37.0 ± 6.0	0.11 ± 0.33	17.0 ± 4.0	23.0 ± 5.0	122 ± 11	7410 ± 90	1062 ± 33
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = e$	1.4 ± 1.2	42	39.0 ± 6.0	0.30 ± 0.50	9.4 ± 3.1	0.9 ± 0.1	3.6 ± 1.9	0.000 ± 0.070	7.4 ± 2.7	6.7 ± 2.6	0.08 ± 0.28	0	10.3 ± 3.2
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = \mu$	1.7 ± 1.3	40	39.0 ± 6.0	0.20 ± 0.50	11.7 ± 3.4	1.2 ± 1.1	3.4 ± 1.8	0.010 ± 0.090	10.1 ± 3.2	1.0 ± 1.0	0.010 ± 0.070	0.10 ± 0.40	11.3 ± 3.4
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = e$	3.3 ± 1.8	78	68.0 ± 8.0	0.90 ± 0.90	25.0 ± 5.0	3.1 ± 1.8	3.4 ± 1.8	0.02 ± 0.14	23.0 ± 5.0	7.5 ± 2.7	2.6 ± 1.6	0	2.9 ± 1.7
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [10, 30]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = \mu$	3.6 ± 1.9	75	74.0 ± 9.0	0.90 ± 0.90	28.0 ± 5.0	3.4 ± 1.8	3.2 ± 1.8	0.02 ± 0.13	27.0 ± 5.0	4.1 ± 2.0	0.20 ± 0.50	0.70 ± 0.80	5.6 ± 2.4
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = e$	1.4 ± 1.2	56	52.0 ± 7.0	0.30 ± 0.50	15.0 ± 4.0	1.2 ± 1.1	2.4 ± 1.6	0.010 ± 0.080	13.0 ± 4.0	1.3 ± 1.1	0.40 ± 0.60	3.3 ± 1.8	15.0 ± 4.0
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [15, 20]$ $\ell_2 = \mu$	1.8 ± 1.3	62	55.0 ± 7.0	0.30 ± 0.50	21.0 ± 5.0	2.2 ± 1.5	2.5 ± 1.6	0.01 ± 0.11	18.0 ± 4.0	0.04 ± 0.20	0.06 ± 0.25	1.5 ± 1.2	10.4 ± 3.2
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = e$	4.5 ± 2.1	161	152 ± 12	0.9 ± 0.1	63.0 ± 8.0	6.9 ± 2.6	5.4 ± 2.3	0.03 ± 0.19	55.0 ± 7.0	2.9 ± 1.7	1.6 ± 1.3	1.4 ± 1.2	15.0 ± 4.0
SR $n_{\text{jet}} = 1$ $m_{\ell\ell} \in [30, 55]$ $p_{\text{T}}^{\ell_2} = [20, \infty]$ $\ell_2 = \mu$	5.1 ± 2.3	157	163 ± 13	1.1 ± 1.1	67.0 ± 8.0	7.8 ± 2.8	6.1 ± 2.5	0.04 ± 0.20	60.0 ± 8.0	0.90 ± 0.90	0.80 ± 0.90	5.0 ± 2.2	13.0 ± 4.0
WW $n_{\text{jet}} = 1$ CR	0.10 ± 0.40	1905	1910 ± 40	0.03 ± 0.18	812 ± 28	44.0 ± 7.0	49.0 ± 7.0	0.20 ± 0.40	774 ± 28	18.0 ± 4.0	7.4 ± 2.7	29.0 ± 5.0	181 ± 13
Z+jets $n_{\text{jet}} = 1$ CR	2.0 ± 1.4	756	753 ± 27	0.60 ± 0.80	70.0 ± 8.0	2.6 ± 1.6	9.3 ± 3.0	0.06 ± 0.24	44.0 ± 7.0	8.2 ± 2.9	17.0 ± 4.0	489 ± 22	111 ± 11
top $n_{\text{jet}} = 1$ CR	1.7 ± 1.3	3792	3790 ± 60	0.30 ± 0.60	51.0 ± 7.0	4.2 ± 2.1	3.4 ± 1.8	0.02 ± 0.12	3620 ± 60	2.7 ± 1.7	0.60 ± 0.80	18.0 ± 4.0	92 ± 1

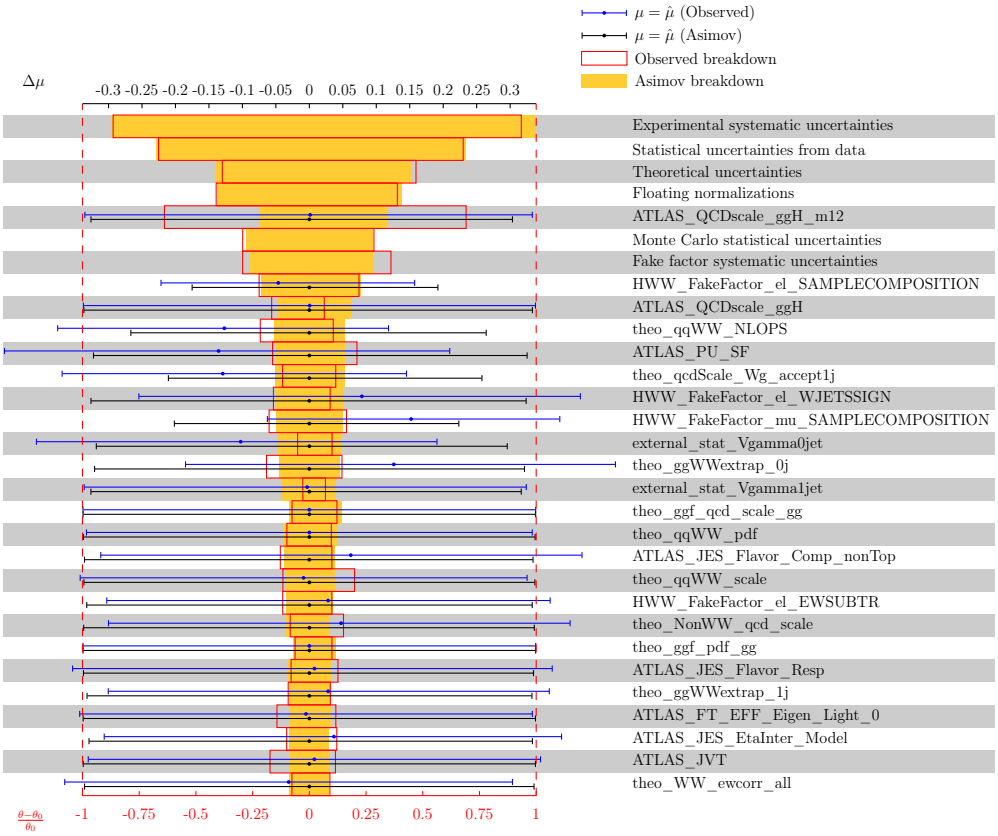


Figure 4.25.: Breakdown of the total uncertainty on the signal strength parameter μ into the individual components and nuisance parameters, shown for the fit to observed data (yellow) and the Asimov data set (red). Only the leading sources of uncertainty are shown. The post-fit values and uncertainties of the individual parameters are also shown in either case, normalized to their respective prefit constraint where applicable.

Table 4.23.: Background normalization factors obtained from the likelihood fit, including the total uncertainty propagated from all sources.

Control Regions	WW	Top	$Z{\to}ee/\mu\mu/\tau\tau$	$V\gamma$
$n_{\text{jet}} = 0$	$1.19^{+0.13}_{-0.12}$	–	$0.95^{+0.08}_{-0.07}$	–
$n_{\text{jet}} = 1$	$1.07^{+0.25}_{-0.23}$	$0.74^{+0.10}_{-0.09}$	$0.92^{+0.18}_{-0.15}$	–

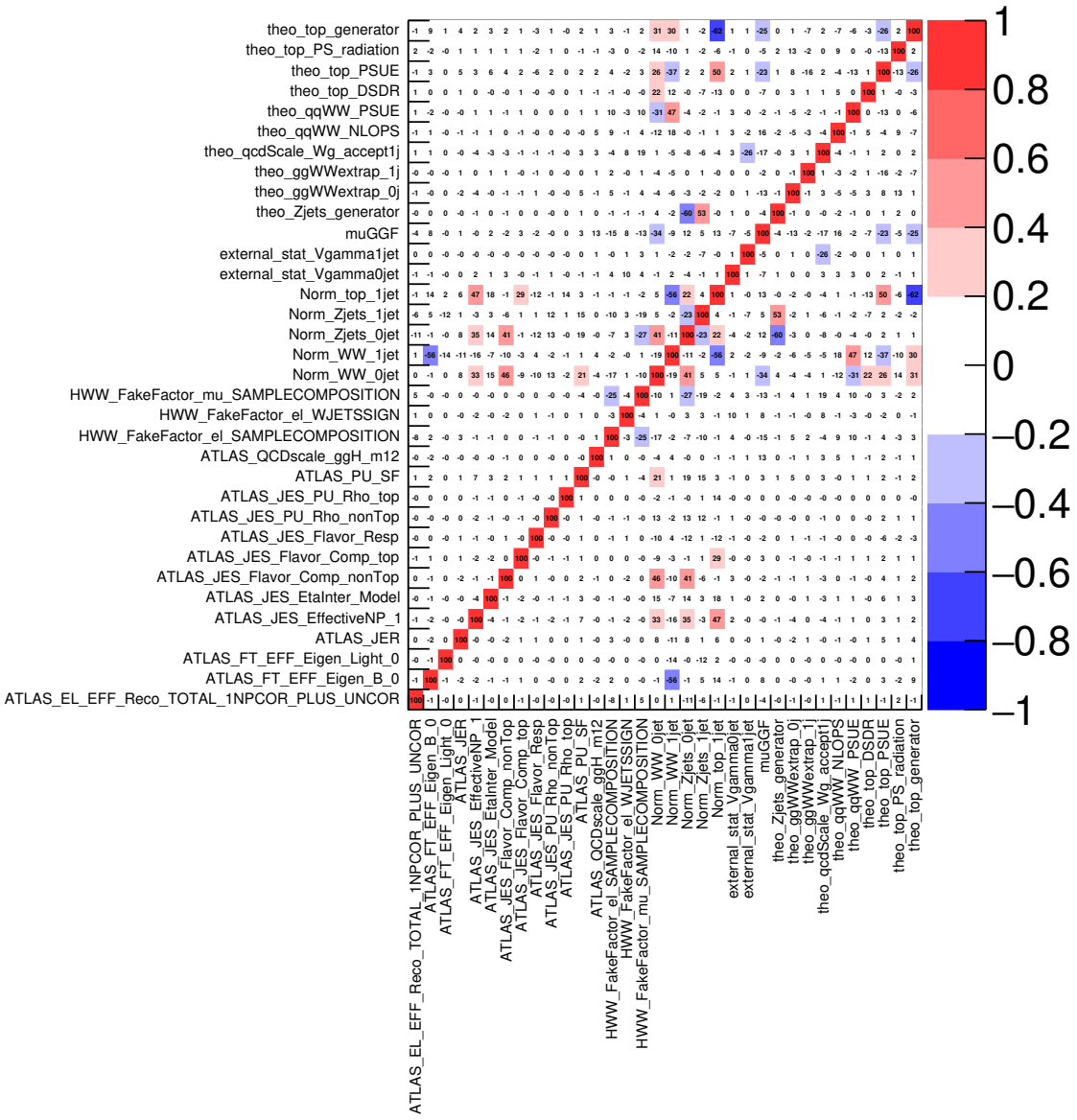


Figure 4.26.: Correlations between nuisance parameters observed in the likelihood fit. Only parameters with correlations above 10% are shown.

Table 4.24.: Effective background normalization factors obtained from the likelihood fit, derived as ratios of the respective contributions before and after the fit. The uncertainties shown are upper bounds that have been derived under the assumption of statistical independence of the event yields before and after the fit, which is not given. Instead, these yields will be positively correlated, which will decrease the true uncertainty on the ratio.

Control Regions	WW	Top	$Z \to ee/\mu\mu/\tau\tau$	$V\gamma$
$n_{\text{jet}} = 0$	1.190 ± 0.050	1.00 ± 0.01	0.80 ± 0.80	0.83 ± 0.19
$n_{\text{jet}} = 1$	1.110 ± 0.070	0.870 ± 0.060	0.70 ± 0.29	0.70 ± 0.18

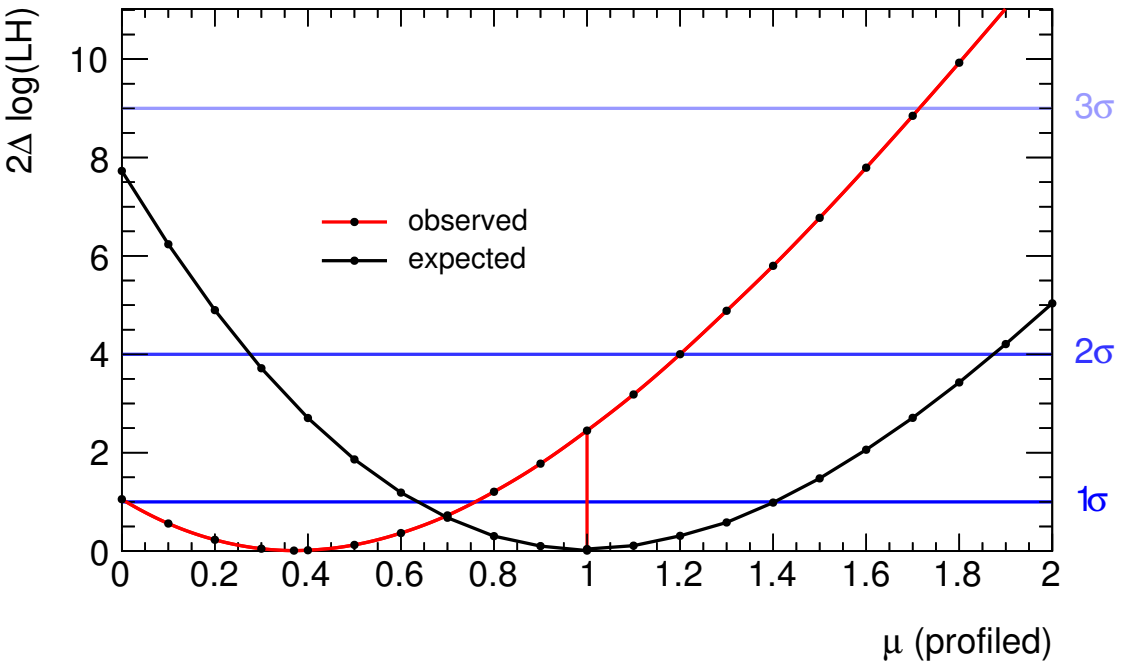


Figure 4.27.: Scan of the logarithmic difference between likelihood values for different values of the signal strength. Shown are the curves corresponding to the observed shape of the likelihood (black) and for the shape expected for the Standard Model prediction of $\mu = 1$. The right hand axis shows the corresponding exclusion confidence levels in terms of the Z-value.

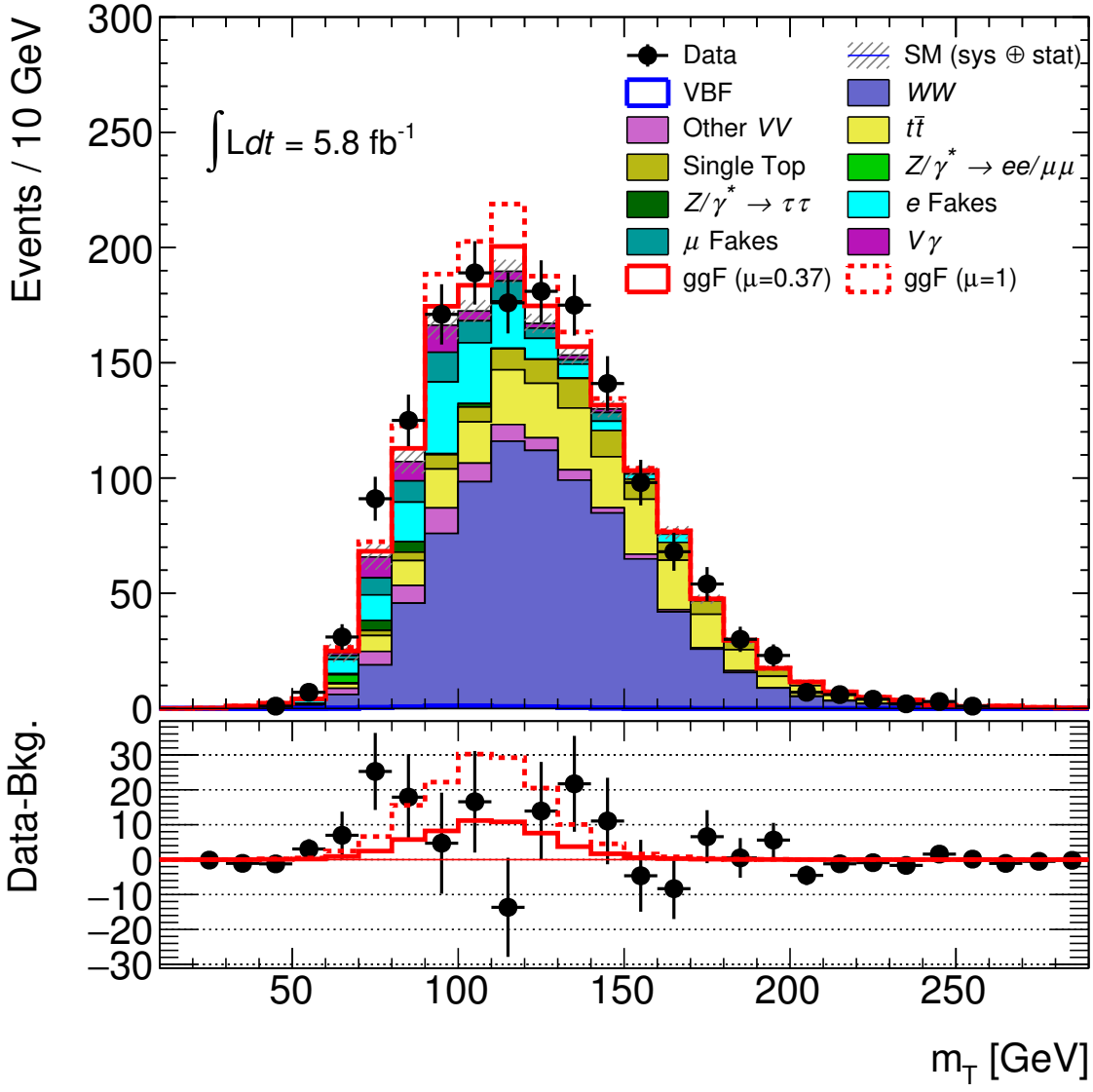


Figure 4.28.: Distribution of transverse mass m_{T} in the final signal region, combined over all categories. The full post-fit normalization factors from Table 4.24 and the measured signal strength parameter $\mu = 0.37$ are applied to the Monte Carlo predictions. The uncertainty band includes experimental systematic and statistical uncertainties on the background. The sub-panel shows the data with the background subtracted and the signal overlayed.

The experimental systematic uncertainties are already the dominating contribution to the total uncertainty of the result. However, many of the systematic uncertainties are themselves determined from data, such that the addition of the remaining $30.7\,\text{fb}^{-1}$ of the 2016 data set will nevertheless significantly boost the sensitivity of the measurement. A prime example for a highly sensitive measurement facilitated through careful study of the contributing systematic uncertainties has been set with the Run 1 measurement at $\sqrt{s} = 7$ and 8 TeV. The experience gained and the methods developed in the course of this analysis can be used to improve upon the results obtained here. Future studies will need to not only optimize the event selection, but also consider ways to reduce the impact of systematic uncertainties on the measurement.

4.6 The Vector Boson Fusion Analysis

A characteristic feature of Higgs boson production via vector boson fusion is the presence of two highly energetic jets with large separation in rapidity. A boosted decision tree (BDT [153–155]) is used to fully exploit the correlations among various discriminating variables.

Events are preselected by requiring exactly one electron and one muon satisfying the criteria in Table 4.4. Light mesonic resonances and Drell-Yan events are suppressed by imposing a requirement of $m_{\ell\ell} > 10\,\text{GeV}$. Orthogonality to the ggF analysis is ensured by requiring $n_{\text{jet}} \geq 2$, and a b-jet veto is applied to reject background from top-quark production.

After the preselection, three additional cuts are applied to improve the background rejection:

Table 4.25.: Event selection criteria used to define the signal and control regions in the VBF analysis.

	Signal region	$Z\to\tau\tau$ CR	Top-quark CR
	Two isolated leptons ($\ell = e, \mu$) with opposite charge		
Preselection	$p_{\text{T}}^{\ell_1} > 25\,\text{GeV}$ ($p_{\text{T}}^{\ell_1} > 22\,\text{GeV}$ for muons in 2015), $p_{\text{T}}^{\ell_2} > 15\,\text{GeV}$		
	$m_{\ell\ell} > 10\,\text{GeV}$, $n_{\text{jet}} \geq 2$		
	$n_{b\text{-jet}} = 0$	$n_{b\text{-jet}} = 0$	$n_{b\text{-jet}} = 1$
Selection	$m_{\tau\tau} < m_Z - 25\,\text{GeV}$	$\lvert m_{\tau\tau} - m_Z\rvert < 25\,\text{GeV}$	–
	–	$m_{\ell\ell} < 80\text{GeV}$	–
	OLV applied, CJV applied, BDT > -0.8		
	SR1: $-0.8 <$ BDT ≤ 0.7	–	–
	SR2: $0.7 <$ BDT ≤ 1.0	–	–

- The "outside-lepton-veto" (OLV) requires $\eta_{\ell}^{*} < 1$, that is, the leptons to be within the rapidity gap spanned by the jets. A precise definition of this quantity is given in the next section.
- The "central-jet-veto" (CJV) rejects events with additional jets with a transverse momentum of $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ in the rapidity gap of the two leading jets.
- The $Z\to\tau\tau$ veto eliminates most of the $Z\to\tau\tau$ background by imposing a requirement on the invariant mass of the $\tau\tau$ system of $m_{\tau\tau} < m_Z - 25\,\mathrm{GeV}$, calculated by using the collinear approximation [143]. This employs the assumption that the τ leptons are sufficiently boosted to force the products of the τ decay to light leptons to be collinear.

Table 4.25 gives a summary of the full VBF selection.

4.6.1 Boosted Decision Tree

The BDT is constructed based on eight discriminating variables, analogously to the construction used for earlier publications [119]. The variables have been selected because they exploit properties of the $H\to W^{\pm}W^{\mp *}$ decay, as is the case for $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$, and m_{T}, as already outlined for the ggF analysis, or of the characteristic VBF signature, as is the case for Δy_{jj}, m_{jj}, $p_{\mathrm{T}}^{\mathrm{tot.}}$, $\sum_{\ell,j} m_{\ell j}$, and η_{ℓ}^{*}.

- The di-jet observables Δy_{jj} and m_{jj} provide powerful handles to selected events exhibiting the characteristic VBF topology of two high-energy jets with a large separation in pseudorapidity.
- The lepton η-centrality

$$\eta_{\ell}^{*} = \frac{|2\eta_{\ell} - \eta_{j_1} - \eta_{j_2}|}{\eta_{j_1} - \eta_{j_2}} \tag{4.6.1}$$

  quantifies the longitudinal positions of the leptons with respect to the jets, with small values corresponding to configurations where the leptons are emitted near the center of the di-jet system, which is characteristic for VBF events.
- For the same reason, the sum of the invariant masses of all four possible lepton-jet pairs $\sum_{\ell,j} m_{\ell j}$ is larger for VBF events.
- The total transverse momentum, defined as

$$p_{\mathrm{T}}^{\mathrm{tot.}} = \sum_{\ell} \vec{p}_{\mathrm{T}}^{\,\ell} + \vec{E}_{\mathrm{T}}^{\,\mathrm{miss}} + \sum_{j} \vec{p}_{\mathrm{T}}^{\,j}, \tag{4.6.2}$$

  is larger for backgrounds with significant soft gluon radiation, and can thus be exploited to suppress these backgrounds, specifically $t\bar{t}$ production.

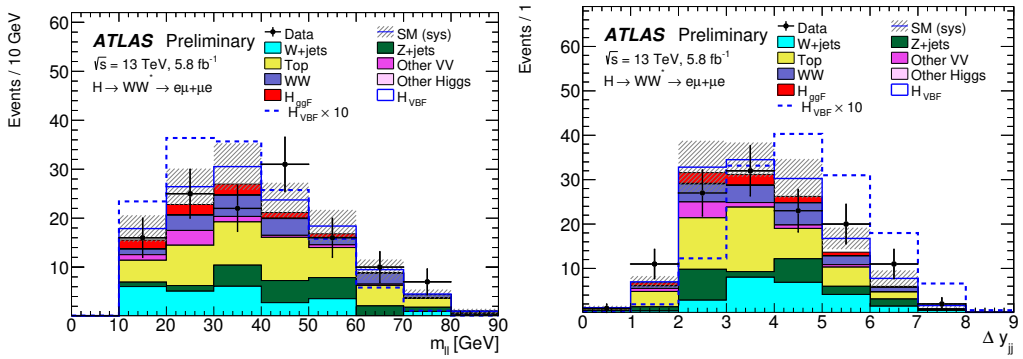


Figure 4.29.: Distributions of $m_{\ell\ell}$ (left) and Δy_{jj} (right) after the preselection. The VBF signal (dashed) is scaled by a factor of 10. The hatched uncertainty band includes MC statistical, experimental, and theory systematic uncertainties [21].

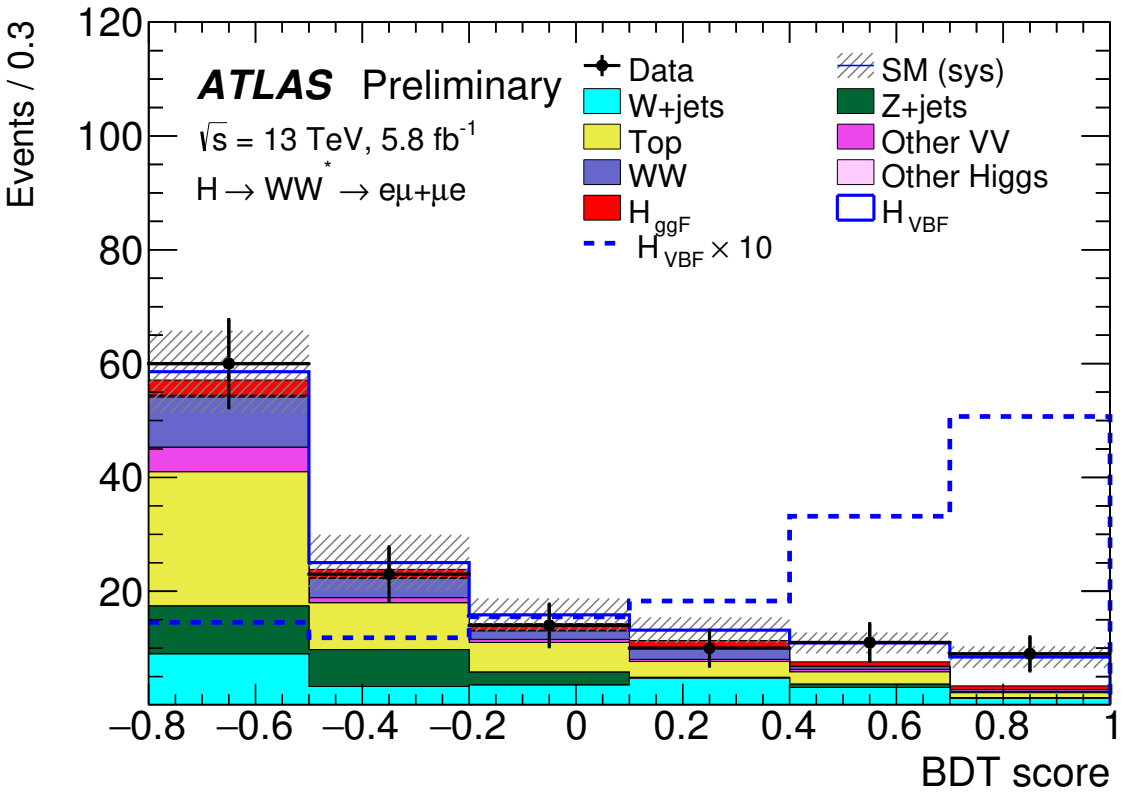


Figure 4.30.: Distribution of the BDT output score in the full VBF selection. The VBF signal (dashed) is scaled by a factor of 10. The hatched uncertainty band includes MC statistical, experimental, and theory systematic uncertainties [21].

The training of the boosted classifier is performed on a data set where the preselection has been applied, including possible contamination from the gluon fusion signal. The variables ranking highest in the training are $m_{\ell\ell}$ and Δy_{jj}. Distributions of these variables after the preselection are shown in Fig. 4.29.

The scoring of the multivariate classifier ranges from -1 for background-like events to $+1$ for signal-like events. The binning has been chosen with the boundaries $[-1, -0.8, 0.7, 1]$ as a result of optimizing for the significance expected from the statistical analysis while requiring every bin to be sufficiently populated. The lowest bin is dominated by background and is not used for the signal extraction, the middle and highest bins are henceforth denoted as SR1 and SR2, respectively. The distribution of the BDT output score in the signal region is shown in Fig. 4.30.

Background estimation

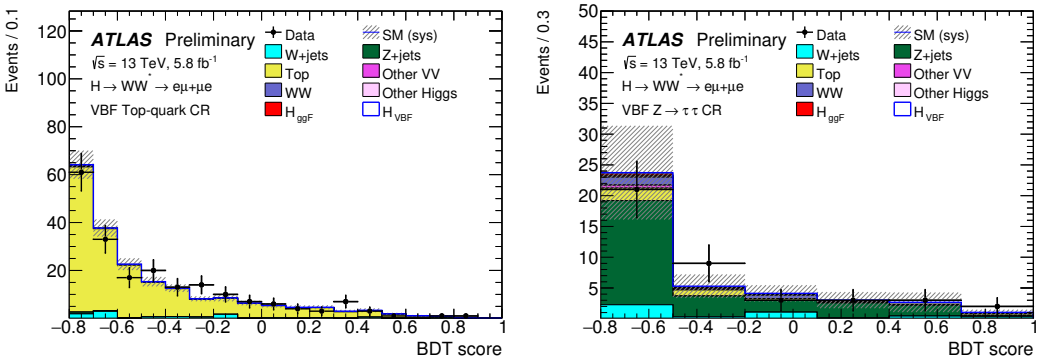
The background processes contributing to the VBF analysis are grouped into several categories, namely

- top-quark production, including $t\bar{t}$ and Wt,
- non-resonant WW production,
- $Z \to \tau\tau$ production,
- W+jets production,
- ggF Higgs boson production, and
- other VV processes, such as $W\gamma^*$, $W\gamma$, WZ, and ZZ.

Most predictions are obtained directly from theory, including the kinematic and detector acceptance as modeled by the respective MC generator and simulation software. The statistical analysis includes associated systematic uncertainties. The estimate of the fake lepton contribution arising from W+jets is performed with the fake factor method detailed in Section 4.3.

The normalizations of the top-quark and $Z \to \tau\tau$ backgrounds are fixed by the use of data control regions. The normalizations shown here have been derived from the combined likelihood including the signal region. Table 4.25 also lists the definitions of these control regions.

The control region dedicated to measure the normalization of top-quark induced contributions is defined identical to the signal region, the only difference being the requirement of exactly one b-tagged jet. The resulting purity in $t\bar{t}$ and Wt events is 92%, and the resulting normalization factor is $\beta_{\text{top}} = 0.91 \pm 0.06\,(\text{stat.})^{+0.44}_{-0.30}\,(\text{syst.})$, where the systematic uncertainties are dominated by the modeling accuracy of the MC sample, and the experimental uncertainties associated to energy scale and resolution of jets and the b-tagging. Figure 4.31a shows the BDT score distribution in this control region.



(a) Top-quark CR **(b)** $Z\to\tau\tau$ CR

Figure 4.31.: Distributions of the BDT output score in the VBF top control region and $Z\to\tau\tau$ control region, including normalization factors, statistical and systematic uncertainties [21].

The control region dedicated to measure the normalization of $Z\to\tau\tau$ contributions is again defined identical to the signal region, but requiring the invariant mass of the $\tau\tau$ system to be within $|m_{\tau\tau} - m_Z| < 25\,\text{GeV}$ and the invariant dilepton mass to suffice $m_{\ell\ell} < 80\,\text{GeV}$. The resulting purity in $Z\to\tau\tau$ events is 61%, and the resulting normalization factor is $\beta_{Z\to\tau\tau} = 0.87 \pm 0.20\,(\text{stat.})^{+0.21}_{-0.18}\,(\text{syst.})$, where the systematic uncertainties are dominated by the limited statistical power of the MC samples and the uncertainties on the energy scale and resolution of jets. Figure 4.31b shows the BDT score distribution in this control region.

4.6.2 Signal extraction

The statistical treatment used in the VBF analysis is very similar to the one performed in the ggF analysis, as introduced in Section 4.5.5. This section contains a brief overview over the differing treatment of systematic uncertainties, as well as the final results of the measurement.

Systematic uncertainties

The leading systematic uncertainties in the measurements of the coupling strength and cross section are summarized in Table 4.26. The treatment of systematic uncertainties is very similar to the strategy performed in the gluon fusion analysis, as described in Section 4.5.3 and Section 4.5.4.

The nominal signal samples are generated using Powheg+Pythia8. The PS/UE uncertainties for the VBF signal are evaluated by comparing the nominal Powheg prediction interfaced with Pythia8 and Herwig7 using the UEEE5 tune. An uncertainty is assigned to the choice of generator, obtained from comparing Powheg+Pythia8 to MadGraph5_aMC@NLO+Pythia8.

The gluon fusion signal, which is treated as a background, is subject to an uncertainty on the acceptance of the central jet veto, which is estimated using the Stewart-Tackmann method [151]. The inclusive H+2-jet and H+3-jet cross sections are treated as uncorrelated and their uncertainties are determined at NLO and LO respectively, using MadGraph5_aMC@NLO. This procedure results in an uncertainty of ±28% for H+2-jet and ±32% for the H+3-jet cross sections. An additional uncertainty is assigned to the choice of generator, estimated by comparing Powheg +Pythia8 to MadGraph5_aMC@NLO+Pythia8, which has NLO accuracy for up to two additional jets. In each measurement range of the BDT output score, QCD scale uncertainties are evaluated to be ±3% in BDT SR1, and ±32% in BDT SR2.

For the background processes normalized using data control regions, the theoretical uncertainties encompass uncertainties on the extrapolation from the control to the signal region.

The QCD scale uncertainties on the non-resonant WW background are evaluated with Sherpa, including non-resonant WW+2 jets production in the matrix element. A comparison of Powheg+Pythia8 and Sherpa results in a ±50% uncertainty attributed to generator choice, which is the leading theoretical uncertainty, driven by a mismodeling of forward jets in Sherpa2.1 with data and the higher precision in the prediction of non-resonant WW+2 jets events compared to Powheg.

For the top-quark background, the renormalization and factorization scales are varied simultaneously using Powheg+Pythia6, with the Powheg h_{damp} parameter regulating high-p_{T} radiation, and varying the corresponding UE tune [156]. The dominant source of uncertainty in the VBF and WH analyses is the generator uncertainty of ±35% and ±21%, respectively, obtained by comparing Powheg+Herwig++ with MadGraph5_aMC@NLO+Herwig++.

Results

The final event yields observed in the data as well as the contributions from all signal and background processes normalized according to the likelihood fit results are shown in Fig. 4.32. The observed and expected signal significance at m_H=125.0 GeV are determined to be 1.9σ and 1.2σ, respectively. The signal strength measurement yields

$$\mu_{\text{VBF}}=1.7^{+1.1}_{-0.9}, \tag{4.6.3}$$

corresponding to

$$(\sigma\mathcal{B})^{H\to W^{\pm}W^{\mp *}}_{\text{VBF}}=1.4^{+0.9}_{-0.7}\,\text{pb}. \tag{4.6.4}$$

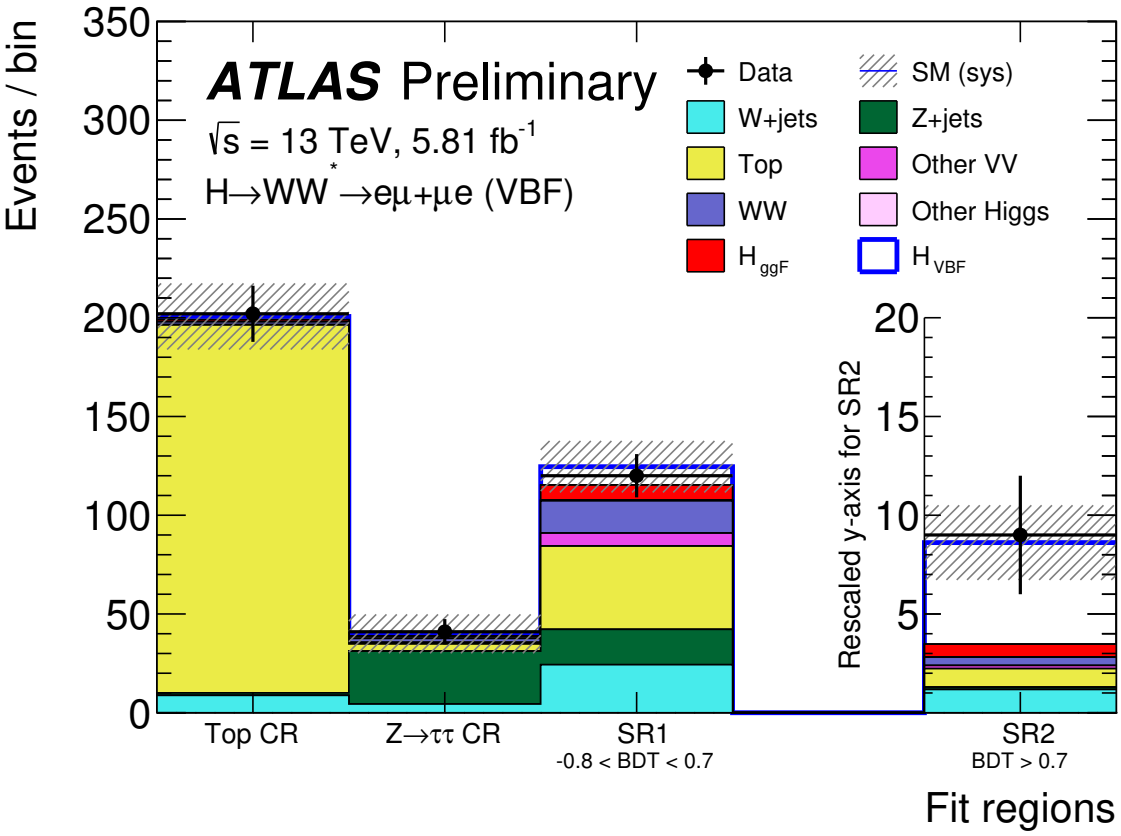


Figure 4.32.: Number of events and sample composition in the various fit regions of the VBF analysis, including a measured signal strength of $\mu = 1.7$ and the background normalization factors derived in the likelihood fit. The most sensitive signal region SR2 is displayed with a different vertical axis. The hatched band includes all statistical and systematic uncertainties [21].

The 95% confidence level upper limit is 3.0 pb, in good agreement with the Standard Model prediction of 0.808 ± 0.021 pb.

4.7 The WH Analysis

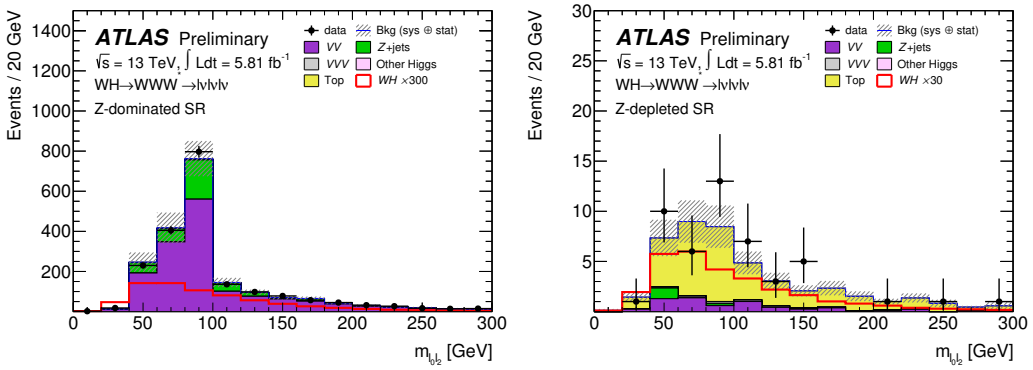
Higgs boson production in association with a W boson results in final states containing the decay products of both the W and the Higgs bosons. Assuming a $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay and requiring the additional W boson to decay leptonically results in a final state containing three isolated leptons with total net charge of $\pm e$.

Events are preselected for this analysis by requiring exactly three such leptons with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$. The leptons are then labeled as follows:

- The lepton ℓ_0 is the one carrying the unique charge.
- The lepton ℓ_1 is the one closer to ℓ_0 in ΔR.
- The remaining lepton is ℓ_2.

The most relevant background processes for this analysis are $WZ/W\gamma^*$, $Z\gamma$, Z+jets, and top-quark production. The categories $Z\gamma$ and Z+jets are distinguished by their kinematic properties, utilizing the invariant mass of the three lepton system. Events with $|m_{\ell\ell\ell} - m_Z| < 15\,\text{GeV}$ are categorized as $Z\gamma$, whereas other events are categorized as Z+jets. The data set is subdivided into two categories:

- The subset of events containing a same-flavour, opposite-charge (SFOC) lepton pair contains roughly $^3/_4$ of signal events and is enriched in background processes containting Z bosons. It is labeled *Z-dominated*.
- The complementary subset is labeled *Z-depleted* and is dominated by backgrounds arising from lepton misidenification.



(a) Z-dominated category **(b)** Z-depleted category

Figure 4.33.: The $m_{\ell_0\ell_2}$ distribution in the Z-dominated and the Z-depleted category after the jet cut and b-jet-veto. The VV category includes $WZ/W\gamma^*$, $ZZ^{(*)}$, WW, $Z\gamma$, $W\gamma$, the VVV category includes WWW, WWZ, WZZ, the *Other Higgs* category includes ggF, VBF, ZH, VH($\tau\tau$). The error band includes all statistical and systematic uncertainties [21].

Background events arising from top-quark production are suppressed by rejecting events with two or more jets or any b-tagged jet with $p_{\text{T}} > 20\,\text{GeV}$ and $|\eta| < 2.4$. Background from $WZ/W\gamma^*$ events is reduced by requiring the largest opposite-sign dilepton invariant mass to suffice $m_{\ell\ell}^{\text{max}} < 200\,\text{GeV}$. Events with non-prompt leptons arising from heavy-flavour meson decays are rejected by requiring the smallest invariant mass of any opposite-charge dilepton pair to exceed $m_{\ell\ell}^{\text{min}} > 12\,\text{GeV}$ or $m_{\ell\ell}^{\text{min}} > 6\,\text{GeV}$ in the Z-dominated or Z-depleted category, respectively. Additionaly, a requirement is imposed on the angular separation between ℓ_0 and ℓ_1 to suffice

Table 4.26.: Leading ten sources of systematic uncertainties and the Monte Carlo statistical uncertainty in the VBF analysis, given in percent. The last row shows the post-fit uncertainty, including all correlations.

Source	$\Delta\mu_{\mathrm{VBF}}/\mu_{\mathrm{VBF}}$ [%]
Statistical	+60 / −50
Fake factor, sample composition	+18 / −15
MC statistical	±15
VBF generator	+14 / −5
WW generator	+11 / −7
QCD scale for ggF signal for $n_{\mathrm{jet}} \geq 3$	+8 / −7
Jet energy resolution	+8 / −7
b-tagging	+8 / −6
Pile-up	+8 / −6
QCD scale for ggF signal for $n_{\mathrm{jet}} \geq 2$	±6
JES flavour composition	+6 / −4
WW renormalisation scale	±5
Total systematic	+33 / −26
Total uncertainty	+70 / −50

Table 4.27.: Event selection criteria and categorization in the WH analysis.

Category	Z-dominated SR	Z-depleted SR
	≥ 1 SFOC pair	no SFOC pair
Preselection	Three isolated leptons with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$	
	total charge $= \pm 1$	
	≥ 1 lepton matches to the trigger	
Background Rejection	$N_{\mathrm{jet}} \leq 1$, $n_{b\text{-jet}} = 0$	
	$E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV	–
	$\lvert m_{\ell^+\ell^-} - m_Z \rvert > 25$ GeV	$Z \to ee$ veto
	$m_{\ell^+\ell^-}^{\mathrm{max}} < 200$ GeV	
	$m_{\ell^+\ell^-}^{\mathrm{min}} > 12\,\mathrm{GeV}$	$m_{\ell^+\ell^-}^{\mathrm{min}} > 6\,\mathrm{GeV}$
$H \to W^{\pm}W^{\mp *} \to \ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ topology	$\Delta R_{\ell_0\ell_1} < 2.0$	

$\Delta R_{\ell_0,\ell_1} < 2$ in order to enhance the Higgs boson signal contribution with respect to $WZ/W\gamma^*$ events.

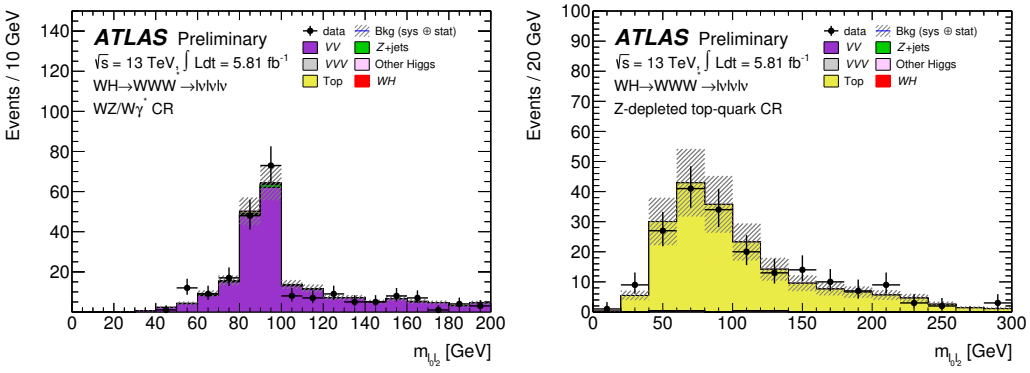
In the Z-dominated category, additional requirements are put in place. The presence of neutrinos from W decays is enforced by requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$. Events with leptons originating from Z decays are suppressed by rejecting events where the SFOC lepton pair lies within $|m_{\ell\ell} - m_Z| < 25\,\mathrm{GeV}$.

In the Z-depleted category, $WZ/W\gamma^*$ events can also be present due to a charge mismeasurement. Such events are suppressed by a $Z \to ee$ veto, rejecting events with same-charge electron pairs with $|m_{\ell\ell} - m_Z| < 10\,\mathrm{GeV}$ where at least one of the electrons is within $|\eta| > 0.8$.

The event selection criteria are summarized in Table 4.27. The distribution of the invariant mass of the opposite-charge lepton pair is shown in Fig. 4.33 after jet cuts have been applied.

Table 4.28.: Event selection criteria used to define the control regions in the WH analysis given with respect to the signal region.

CR	Process	Reference SR	Changes w.r.t. reference SR
CRa	$WZ/W\gamma^*$	Z-dominated	≥ 1 SFOC pair with $\lvert m_{\ell\ell} - m_Z\rvert < 25\,\mathrm{GeV}$
CRb	$Z\gamma$	Z-dominated	no Z-mass veto $\lvert m_{\ell\ell\ell} - m_Z\rvert < 15\,\mathrm{GeV}$ $E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$ only $eee, \mu\mu e$
CRc	Z+jets	Z-dominated	≥ 1 SFOC pair with $\lvert m_{\ell\ell} - m_Z\rvert < 25\,\mathrm{GeV}$ $E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$ $\lvert m_{\ell\ell\ell} - m_Z\rvert > 15\,\mathrm{GeV}$ one lepton without an isolation requirement NFs are derived for e-fake sample and μ-fake sample separately
CRd	Top-quark	Z-dominated	no $m_{\ell^+\ell^-}^{\mathrm{max}}$ and $\Delta R_{\ell_0\ell_1}$ cuts at least 1 jet one b-jet one lepton without an isolation requirement
CRe	Top-quark	Z-depleted	no $m_{\ell^+\ell^-}^{\mathrm{max}}$ and $\Delta R_{\ell_0\ell_1}$ cuts at least 1 jet one b-jet one lepton without an isolation requirement



(a) $WZ/W\gamma^*$ CR **(b)** Z-depleted top-quark CR

Figure 4.34.: The $m_{\ell_0\ell_2}$ distributions for two selected control regions. The VV category includes $WZ/W\gamma^*$, $ZZ^{(*)}$, WW, $Z\gamma$, $W\gamma$, the VVV category includes WWW, WWZ, WZZ, the *Other Higgs* category includes ggF, VBF, ZH, VH($\tau\tau$). The error band includes all statistical and systematic uncertainties [21].

Background estimation

The background normalizations for the most important background processes in this analysis are estimated from data using control regions. The corresponding event selections are summarized in Table 4.28.

The control region targeting top-quark contributions is split between the Z-dominated and the Z-depleted region to accomodate any differences induced by the changed $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold. All other control regions are shared between the two categories. All control regions are generally based on the event selections of the signal region, but control regions targeting backgrounds with fake leptons, *i. e.*, the Z+jets and the top-quark control region, drop the isolation requirement for one of the leptons, significantly increasing the statistical power and purity of these regions.

Even though triboson production, in particular $WWW^{(*)}$, is an irreducible background, it is difficult to define a control region for this process, hence the prediction for this process is taken from theory.

Events in the top control regions are required to have at least one jet and exactly one b-tag. In order to enhance the statistical power of the regions, any requirements on the minimum or maximum invariant mass of the opposite-charge lepton pair and the topological cut on ΔR are removed. The purity is 98% in the Z-dominated and 99% in the Z-depleted regions. The measured normalization factors are $\beta = 0.97 \pm 0.06\,(\mathrm{stat.}) \pm 0.04\,(\mathrm{syst.})$ and $\beta = 0.89 \pm 0.06\,(\mathrm{stat.}) \pm 0.04\,(\mathrm{syst.})$, respectively, where the systematic uncertainty is dominated by the b-tagging uncertainty.

A control region for $WZ/W\gamma^*$ events is defined based on the Z-dominated category. The Z veto is inverted, resulting in a purity of 92% and a normalization factor of $\beta = 1.24 \pm 0.09\,(\text{stat.}) \pm 0.07\,(\text{syst.})$, with the systematic uncertainty being dominated by contributions associated to the jet energy scale and pile-up.

The control region for $Z\gamma$ is identical to the one for $WZ/W\gamma^*$, but only contains $\mu\mu e$ and eee events, has an inverted cut on $E_{\text{T}}^{\text{miss}}$ and additionally requires $|m_{\ell\ell\ell} - m_Z| < 15\,\text{GeV}$. The resulting purity is 83% and the extracted normalization factor is $\beta = 0.83 \pm 0.08\,(\text{stat.}) \pm 0.20\,(\text{syst.})$.

The Z+jets control region is again identical to the one targeting $Z\gamma$ events, but split for events with the lepton flavour combinations $\mu\mu e + eee$ and $\mu\mu\mu + ee\mu$. The respective purities are 58% and 55%, the obtained normalization factors are $\beta = 0.69 \pm 0.07\,(\text{stat.}) \pm 0.14\,(\text{syst.})$ and $\beta = 0.70 \pm 0.07\,(\text{stat.}) \pm 0.13\,(\text{syst.})$.

The distribtion of $m_{\ell_0\ell_2}$ is shown in Fig. 4.34 for the $WZ/W\gamma^*$ and Z-depleted control region, where all normalization factors obtained from the combined fit including the signal region have been applied, exhibiting good agreement between the predictions and the data.

4.7.1 Signal extraction

The statistical treatment used in the WH analysis is again very similar to the one used in the VBF analysis.

Systematic uncertainties

The treatment of systematic uncertainties is similar to the one performed for the VBF analysis presented in Section 4.6.2. An overview is given in Table 4.29. Again, the nominal signal samples are generated using POWHEG+PYTHIA8, and also the uncertainty estimates are evaluated in the same way as in the VBF analysis. Also here, the theoretical uncertainties on the extrapolation from the control to the signal region are included in the theoretical uncertainties.

In this analysis, an additional uncertainty is assigned to the Z+jets e-fake normalization factor to model the difference in flavour composition between the signal regions and the corresponding control region which is caused by the diffeent lepton isolation definitions, amounting to a normalization uncertainty of ±18%.

For the evaluation of theoretical uncertainties on the $WZ/W\gamma^*$ production, the effects of scale variations for QCD renormalization, factorization and resummation are evaluated using SHERPA. A generator uncertainty is assigned from comparing SHERPA with POWHEG+PYTHIA8, yielding ±4% and ±8% uncertiainty in the Z-dominated and Z-depleted regions, respectively. The predictions by SHERPA and VBFNLO [157–159] are compared to estimate uncertainties connected to WWW^* production. The total normalization difference after the application of the

Table 4.29.: Leading ten sources of systematic uncertainties and the Monte Carlo statistical uncertainty in the WH analysis, given in percent. The last row shows the post-fit uncertainty, including all correlations.

Source	$\Delta\mu_{WH}/\mu_{WH}$ [%]
Statistical	+120 / −100
MC statistical	+60 / −70
Pile-up	+22 / −26
Jet energy resolution	+22 / −23
Top-quark generator	+17 / −20
b-tagging	+10 / −11
Top-quark PS/UE	+7 / −8
JES flavour composition	+8 / −5
JES η intercalibration	+7 / −6
$WZ/W\gamma^*$ generator	+7 / −6
Top-quark QCD scales	+6 / −7
$WZ/W\gamma^*$ resummation scale	±5
Total systematic	+70 / −80
Total uncertainty	+140 / −130

NLO k-factor to the VBFNLO prediction is ±20%, while the acceptance varies by ±13% and ±4% in the Z-dominated and Z-depleted regions, respectively. These uncertainties are directly applied to the predicted event yields, as no control region is defined for this process.

Results

The final event yields observed in data as well as the contributions from all signal and background processes normalized according to the likelihood fit results are shown in Fig. 4.35. The observed and expected signal significances at m_H=125.0 GeV were determined as 0.77σ and 0.24σ. The signal strength measurement yields

$$\mu_{\mathrm{WH}}=3.2^{+4.4}_{-4.2}, \tag{4.7.1}$$

corresponding to

$$(\sigma\mathcal{B})^{H\to W^{\pm}W^{\mp *}}_{WH}=0.9^{+1.3}_{-1.2}\,\mathrm{pb}, \tag{4.7.2}$$

with a 95% confidence level upper limit at 3.3 pb. While this result does not provide strong constraints compared to the Standard Model prediction of 0.293 ± 0.007 pb, it is still mostly limited by the statistical uncertainty arising from the limited size of the data set. Future measurements including the full 2016 data set, which

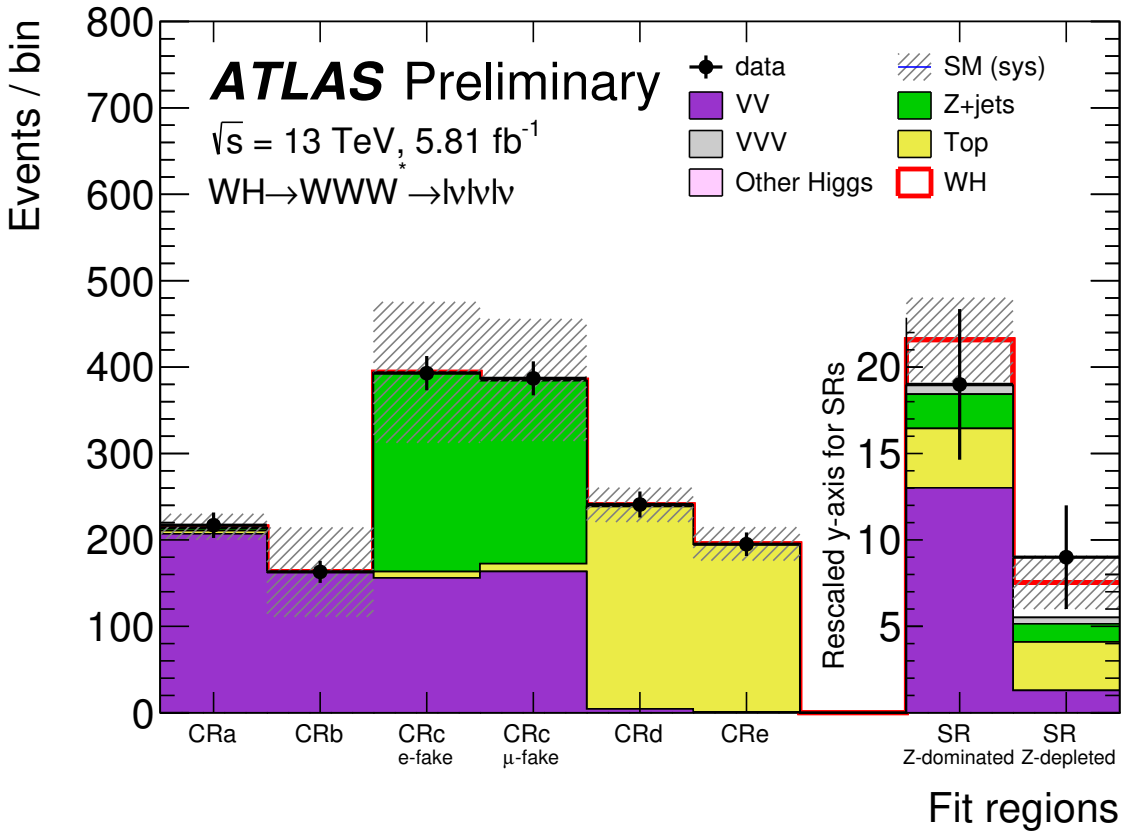


Figure 4.35.: Number of events and sample composition in the various fit regions of the WH analysis, including the normalization factors derived in the likelihood fit. The two signal regions are displayed with a different vertical axis. The hatched band includes all statistical and systematic uncertainties [21].

is almost eight times the size of what is presented here, are thus expected to be much more sensitive.

4.8 Conclusions

Higgs boson production has been measured in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay mode in gluon fusion, vector boson fusion and WH associated production, using $5.8\,\text{fb}^{-1}$ of pp collision data collected by the ATLAS experiment at $\sqrt{s} = 13\,\text{TeV}$ in the years 2015 and 2016.

4.8.1 Summary

The respective measurements of the signal strength parameters and observed as well as expected significances yield

$$\mu_{\text{ggF}} = 0.37\,^{+0.39}_{-0.37} \qquad Z_{\text{obs.}} = 1.0\sigma \qquad Z_{\text{exp.}} = 2.8\sigma \tag{4.8.1}$$

$$\mu_{\text{VBF}} = 1.7\,\,^{+1.1}_{-0.9} \qquad Z_{\text{obs.}} = 1.9\sigma \qquad Z_{\text{exp.}} = 1.2\sigma \tag{4.8.2}$$

$$\mu_{\text{WH}} = 3.2\,\,^{+4.4}_{-4.2} \qquad Z_{\text{obs.}} = 0.77\sigma \qquad Z_{\text{exp.}} = 0.24\sigma. \tag{4.8.3}$$

These results are compatible with the Standard Model prediction of $\mu = 1$. The uncertainties are large compared to those of earlier measurements [119], which is however mostly a result of the much smaller data set available for analysis at $\sqrt{s} = 13\,\text{TeV}$ up to this point. A more detailed discussion of this comparison is presented in the next section.

The results of the three individual measurements can be combined to a measurement of the global signal strength parameter μ, assuming fixed ratios between the individual production rates corresponding to the predictions of the Standard Model.

For the combined fit, all experimental nuisance parameters as well as the signal strength parameter μ have been treated as fully correlated between all three individual analysis channels, while all parameters modeling statistical or theory uncertainties have been treated as fully uncorrelated. This combined fit yields a global signal strength of

$$\mu = 0.67\,^{+0.20}_{-0.20}\,(\text{stat.})\,^{+0.28}_{-0.23}\,(\text{syst.}) \tag{4.8.4}$$

$$= 0.67\,^{+0.34}_{-0.31}. \tag{4.8.5}$$

A scan of the combined likelihood profile is shown in Fig. 4.36. The combined evidence for Higgs boson production measured in $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decays in $5.8\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data amounts to 2.2σ. The expected sensitivity was 3.4σ.

The analysis of $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decays is one of the most complex and yet one of the most sensitive analyses of Higgs boson production and decay. Earlier measurements have observed Higgs boson production in $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decays in a combined data set corresponding to $25\,\text{fb}^{-1}$ of $\sqrt{s} = 7$ and $8\,\text{TeV}$ data [119]. The analysis presented here does not yet include all data collected by

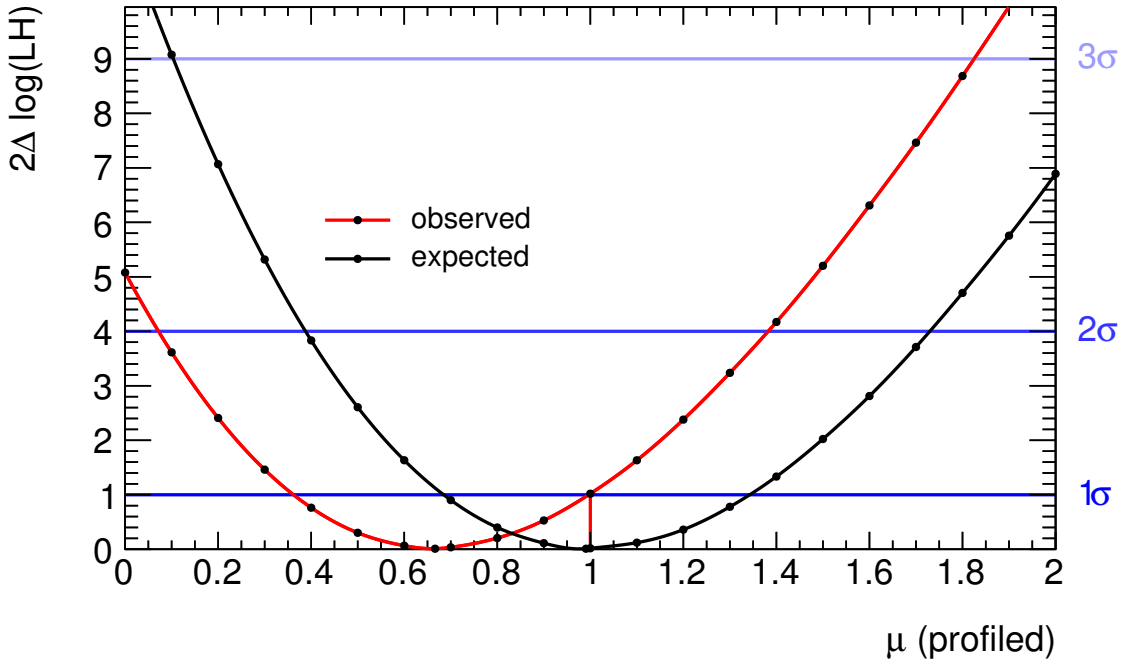


Figure 4.36.: Scan of the logarithmic difference between likelihood values for different values of the signal strength. Shown are the curves corresponding to the observed shape of the likelihood (black) and for the shape expected for the Standard Model prediction of $\mu = 1$. The right hand axis shows the corresponding exclusion confidence levels in terms of the Z-value.

the ATLAS experiment at $\sqrt{s} = 13\,\text{TeV}$ in the years 2015 and 2016 during LHC Run 2, and future studies will be able to improve the results presented here by utilizing a larger data set.

The corresponding result made public by the CMS collaboration [160] includes $2.7\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 13\,\text{TeV}$ during 2015 and targets the gluon fusion and the vector boson fusion modes of Higgs boson production. The analysis strategy is very similar to the one discussed here, but differs in a few significant aspects. For example, only events in the $n_{\text{jet}} = 0$ and $n_{\text{jet}} = 1$ categories are used, whereas a dedicated VBF signal region is not constructed, given the smaller data set. The estimate for contributions originating from misidentified leptons for this analysis is obtained from a control region where both leptons are required to have the same electric charge. The combined expected sensitivity of the analysis presented by the CMS collaboration is 2.0σ, whereas only 0.7σ have been observed. The signal strength has been measured as $\mu = 0.3 \pm 0.5$, which is in good agreement with the Standard Model prediction of $\mu = 1$, with the absence of a Higgs boson signal, and with the results presented here.

The sensitivity of the VBF and the WH analyses are still dominated by statistical uncertainties arising from the size of the data set. They will significantly improve with the inclusion of the full 2016 data set with a roughly eight-fold size as compared to what has been analyzed here, as well as the data set to be recorded in 2017 and beyond. The WH analysis will additionally benefit from making use of multivariate analysis techniques, as was done for the Run 1 analysis of this channel [119].

4.8.2 Comparison to Run 1 results

The gluon fusion analysis of the full Run 1 data set of $25\,\text{fb}^{-1}$ collected in the years 2012 and 2013 [119] had an expected sensitivity corresponding to 4.3σ. If the background composition and signal-to-background ratio were to remain unchanged, simply scaling this result down to $5.8\,\text{fb}^{-1}$ by assuming that sensitivity is roughly proportional to the square root of the luminosity yields the observation that the corresponding sensitivity of the Run 1 analysis on the data set analyzed here would be roughly 2.1σ. However, as introduced already in Section 3.3.3, the background composition and signal-to-background ratio are very different at $\sqrt{s} = 13\,\text{TeV}$ compared to $\sqrt{s} = 8\,\text{TeV}$. Including the respective cross section factors of 2.5 for Higgs boson production, 2.0 for WW production and 3.3 for $t\bar{t}$ production into such a prediction is not straight-forward, as the various analysis categories have very different contributions from the respective processes. Using a uniform scaling factor of 2.5 for all signal and background processes, the projected sensitivity of the Run 1 analysis on the Run 2 data set corresponding to $5.8\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data would correspond to 3.3σ. The performance of the analysis is thus comparable to the one presented in Run 1.

There are, however, many differences to the Run 1 analysis, which was taking into account a much larger phase space, featuring a dedicated low-p_T signal region, a gluon fusion analysis in the $n_\text{jet} \geq 2$ category, and a comprehensive analysis of same-flavour events in the ee and $\mu\mu$ final states. Including these phase space regions into the analysis will enhance the sensitivity in the future.

The cross section for one of the leading background processes, top quark pair production, increases by a factor of roughly 3.3 from $\sqrt{s} = 8$ to $\sqrt{s} = 13\,\text{TeV}$, which is more than the increase of the signal cross section. This poses a challenge for the analysis especially in the higher jet multiplicity bins, as the rejection of top quark background becomes more and more important.

4.8.3 Outlook

The sensitivity of the analysis in the gluon fusion category is already dominated by the various systematic uncertainties. Significantly improving the sensitivity in this analysis will require reducing the leading sources of systematic uncertainty.

For the estimate of the background contribution from misidentified ("fake") leptons, the increased data set will allow to use the alternative measurement of the fake factor from Z+jets events. This will reduce the systematic uncertainty covering the variations in composition of the contribution from misidentified leptons, which is among the leading uncertainties in the gluon fusion analysis. An optimization of lepton identification and isolation criteria will also allow to more effectively remove the fake lepton background events from the signal region without compromising the sensitivity. This will reduce the total impact of the uncertainties associated with this method. Finally, a larger data set will allow to choose tighter lepton identification criteria, sacrificing some of the newly gained statistical power to improve upon the systematic uncertainties.

The normalization factor derived for the $V\gamma$ background contribution is also still severely limited by the statistical power of the data set. Inclusion of more data will not only improve the accuracy of this estimate, but will also allow it to be fully included in the likelihood fit due to the increased robustness of the prediction with respect to statistical fluctuations.

The alternative estimation of the top quark background contribution in the $n_{\text{jet}} = 1$ category, which is only used as a cross-check at the moment, can also be included in the likelihood fit. This procedure provides the opportunity to reduce the systematic uncertainties related to flavour tagging, as they would largely cancel in this data-driven estimate.

Many of the experimental systematic uncertainties that are currently estimated and constrained by additional measurements will be able to improve with the increased data set, as the corresponding supplementary measurements will also benefit from the increased statistical power.

The Monte Carlo statistical uncertainty, which is also one of the largest single contributors to the total uncertainty on the gluon fusion estimate, will also improve as additional (and larger) Monte Carlo samples will be generated over time. The use of improved generator filters will enhance the selection efficiency of these samples and greatly reduce the statistical uncertainty on the respective predictions derived for the signal regions. To name only one example, a larger sample of the $Z \to \tau\tau$ background with an optimized generator filter is currently being produced, enhancing the effective luminosity by a factor of ten.

Finally, the cross section predictions are continuously being improved. Advancements in this area will also contribute to a reduction of the systematic uncertainty in the future. For example, a switch from the current POWHEG+PYTHIA8 event generator towards NNLOPS will allow a better description of signal shapes.

5 Coupling Measurements

> “Science walks forward on two feet, namely theory and experiment, ... but continuous progress is only made by the use of both.”
>
> *Robert Millikan*

This chapter describes a combination of measurements performed for $\sqrt{s} = 13\,\text{TeV}$ in the $H\to\gamma\gamma$ and $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$ channels, the latter analysis also using the novel technique of *Effective Lagrangian Morphing* introduced in Section 2.4.3 to measure the Wilson coefficients. Furthermore, the prospects of applying this technique to measurements in the $H\to W^\pm W^{\mp *}\to\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ decay channel are investigated, presenting approaches to conquer the challenges arising from the high dimensionality of the available parameter space.

5.1 Measuring $H\to ZZ^*$ and $H\to\gamma\gamma$ Couplings

While measurements of Higgs boson production and decay processes are typically conducted one channel at a time, combined measurements including data from many different analyses provide the highest amount of sensitivity by combining constraints obtained from different parts of the data set. Such a measurement has been conducted [23], combining the $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$ [22] and $H\to\gamma\gamma$ [161] analyses, using data collected at $\sqrt{s} = 13\,\text{TeV}$ in the years 2015 and 2016. The total integrated luminosities used here correspond to $14.8\,\text{fb}^{-1}$ for the $H\to ZZ^*$ analysis and to $13.3\,\text{fb}^{-1}$ for the $H\to\gamma\gamma$ analysis, respectively, of which $3.2\,\text{fb}^{-1}$ have been collected in the year 2015, and the rest in 2016. Unfortunately, the analysis presented in Chapter 4 was not completed in time to be included in this combined measurement, such that a combination of all bosonic decay channels has to be postponed to future publications.

5.1.1 $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$

The Higgs boson decay to four leptons provides an excellent and characteristic signature for measurements at hadron colliders, so much so that it has been famously dubbed the "gold-plated channel" [162]. Four isolated leptons in the final state allow to efficiently separate these events from the background, such that even though the dileptonic decay branching ratio of the Z boson to any lepton flavour is only about 3.3% each [31], signal region for this decay mode can be constructed with a signal-to-background ratio exceeding two-to-one.

Events are selected for the analysis by requiring at least four isolated leptons forming at least two same-flavour opposite-charge (SFOC) pairs and meeting some general preselection criteria [22]. Transverse momentum thresholds of $p_{\mathrm{T}} > 20/15/10\,\mathrm{GeV}$ are imposed on the leading three leptons in the quadruplet, and $p_{\mathrm{T}} > 7/5\,\mathrm{GeV}$ is required for the fourth lepton for electrons or muons, respectively. Among the quadruplet, the SFOC lepton pair closer to the Z boson mass is regarded as the *leading* pair, the other one being referred to as *subleading*. If several quadruplets can be built in the same event, the one with the leading (subleading) SFOC lepton pair invariant mass $m_{\ell\ell}$ (second) closest to the Z boson mass is given preference. The event selection proceeds by imposing $50\,\mathrm{GeV} < m_{12} < 106\,\mathrm{GeV}$ and a sliding mass window cut of $x < m_{34} < 115\,\mathrm{GeV}$ with x linearly interpolating between $12\,\mathrm{GeV} \leq x \leq 50\,\mathrm{GeV}$ in the range $140\,\mathrm{GeV} \leq m_{4\ell} \leq 190\,\mathrm{GeV}$, and $x = 12\,\mathrm{GeV}$ or $x = 50\,\mathrm{GeV}$ beyond the respective boundaries. Furthermore, $m_{\ell\ell} > 5\,\mathrm{GeV}$ is required for all SFOC lepton pairs to protect against low-mass dilepton resonances, and $\Delta R_{\ell\ell} > 0.1\,(0.2)$ is required for all same (opposite) flavour pairs. A signal region is then constructed from events with an invariant four-lepton mass close to the mass of the Higgs boson, that is, $118\,\mathrm{GeV} < m_{4\ell} < 129\,\mathrm{GeV}$.

Four different final states are considered for the analysis, $\mu^+\mu^-\mu^+\mu^-$, $\mu^+\mu^-e^+e^-$, $e^+e^-\mu^+\mu^-$ and $e^+e^-e^+e^-$, where the ordering of the leptons is chosen such that the leading lepton pair is the first one listed, and the lepton with positive charge is listed first within each pair.

The analysis is further subdivided by jet multiplicity similar to the strategy employed by the $H \to W^\pm W^{\mp *} \to \ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ analysis presented in Chapter 4, the $n_{\mathrm{jet}} = 0$ and $n_{\mathrm{jet}} = 1$ categories being used to measure Higgs boson production through gluon fusion, while the $n_{\mathrm{jet}} \geq 2$ category focuses on Higgs boson production via vector boson fusion and VH associated production with the additional vector boson decaying hadronically. The subdivision between these two modes is made by placing a cut at $m_{jj} = 120\,\mathrm{GeV}$ to separate between the VBF-like events above, and the VH-like events below that threshold. An additional category is created by demanding a fifth lepton with $p_{\mathrm{T}} > 8\,\mathrm{GeV}$ to capture VH induced events where the additional vector boson decays leptonically.

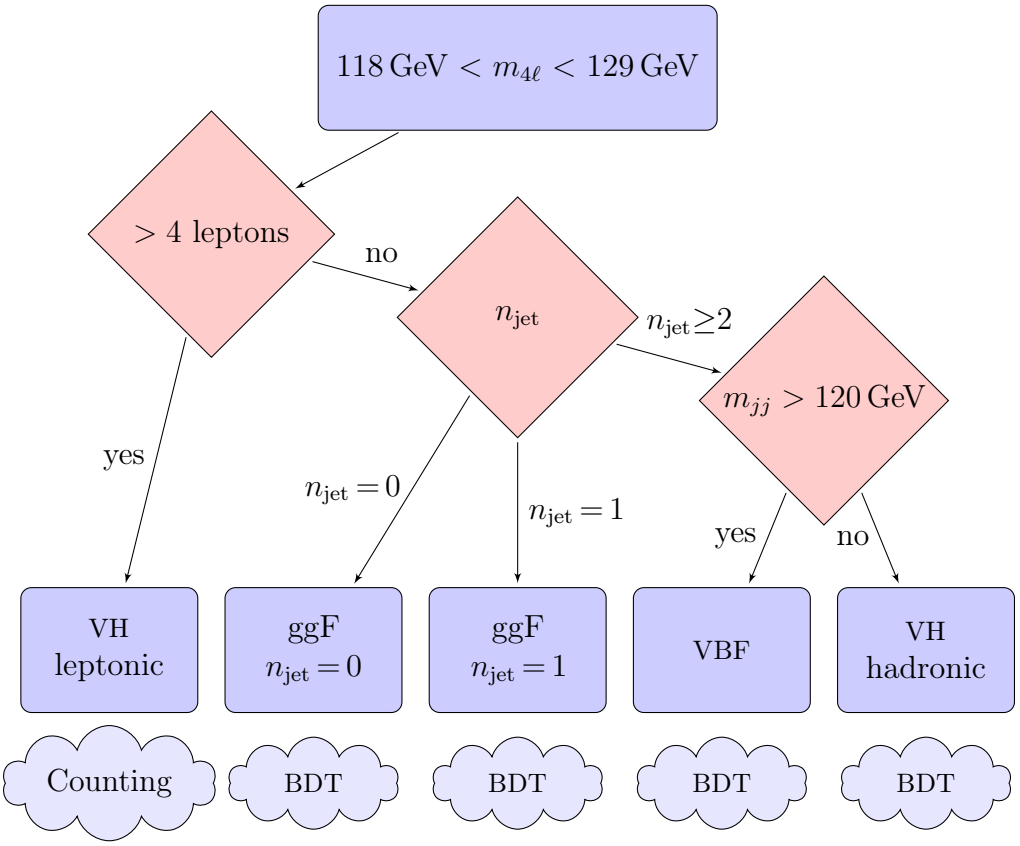


Figure 5.1.: Schematic view of the event categorization in the $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ analysis.

Table 5.1.: The expected and observed yields and uncertainties for the signal mass region in all categories for the different production modes, including all lepton flavour combinations and assuming $m_H = 125$ GeV [22].

	VH leptonic	ggF $n_{\text{jet}} = 0$	ggF $n_{\text{jet}} = 1$	VBF	VH hadronic	Total
ggF+bbH+ttH	0.055 ± 0.004	11.2 ± 1.4	5.7 ± 2.4	1.9 ± 0.9	1.1 ± 0.5	20 ± 4
VBF	< 0.01	0.120 ± 0.019	0.59 ± 0.05	0.92 ± 0.07	0.084 ± 0.009	1.71 ± 0.14
WH	0.067 ± 0.004	0.047 ± 0.007	0.137 ± 0.012	0.074 ± 0.007	0.143 ± 0.012	0.47 ± 0.04
ZH	0.011 ± 0.001	0.060 ± 0.006	0.091 ± 0.008	0.052 ± 0.005	0.101 ± 0.009	0.315 ± 0.027
ZZ^*	0.016 ± 0.002	6.2 ± 0.6	1.62 ± 0.21	0.22 ± 0.05	0.166 ± 0.035	8.2 ± 0.9
Z+jets, $t\bar{t}$	0.012 ± 0.010	0.84 ± 0.12	0.44 ± 0.07	0.24 ± 0.11	0.088 ± 0.011	1.62 ± 0.07
expected	0.16 ± 0.01	18.4 ± 1.6	8.5 ± 2.4	3.4 ± 0.9	1.6 ± 0.5	32 ± 4
observed	0	21	12	9	2	44

Within each 4ℓ category, a dedicated Boosted Decision Tree (BDT [153–155]) is trained to maximize the discriminatory power, whereas the analysis in the leptonic VH category is performed by simply counting the events. A schematic view of the strategy is shown in Fig. 5.1.

The largest background for this analysis is non-resonant production of pairs of Z^*/γ^* bosons, generally referred to as ZZ^*. This contribution is accurately modeled by the Monte Carlo generated samples [22]. Below $m_{4\ell} < 160\,\text{GeV}$, there are also non-negligible contributions from Z+jets and $t\bar{t}$ events with two leptons, where the additional lepton candidates arise from decays of heavy flavour hadrons, from π/K in-flight decays, from photon conversions or from lepton misidentification. For these contributions, control regions are used to derive the normalization.

The final expected and observed event yields in each signal region are shown in Table 5.1. The signal extraction proceeds by means of a likelihood fit binned in the BDT output score.

5.1.2 $H\to\gamma\gamma$

The analysis of $H\to\gamma\gamma$ employs an event categorization into 13 exclusive signal regions targeting different Higgs boson production modes. These are

- two signal regions enriched in ttH production,
  1. one targeting single-lepton and double-lepton decays of the $t\bar{t}$ system, requiring at least one lepton, at least two jets, one b-tag, and either large missing transverse momentum or additional b-tags,
  2. one targeting fully hadronic decays of the $t\bar{t}$ system, requiring at least five jets with at least one b-tag.
- five signal regions enriched in VH production,
  3. a dilepton category with a same-flavour, opposite-sign lepton pair consistent with a Z-boson decay,
  4. a single lepton category targeting WH production, requiring exactly one lepton and placing a minimum requirement on the ratio between the missing transverse momentum and the scalar sum of all its contributions,
  5. a missing transverse momentum category with no leptons with a minimum requirement on that same ratio and the missing transverse momentum itself, targeting ZH production with $Z \to \nu\nu$ decays,
  6. one hadronic VH category with two jets consistent with a hadronic vector boson decay and a high score of a specifically trained Boosted Decision Tree,
  7. and a second hadronic VH category similar to the first one selecting events scoring less in the same BDT,

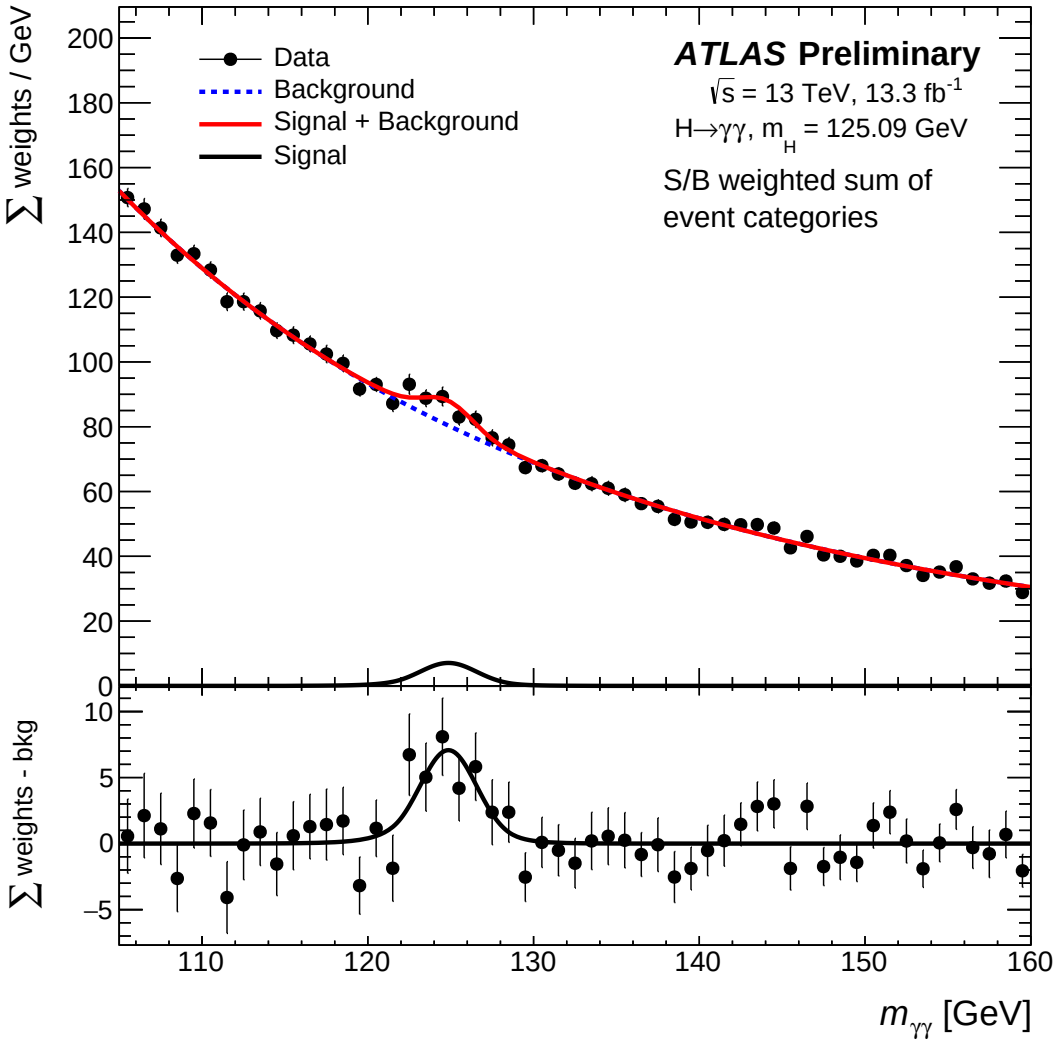


Figure 5.2.: The invariant mass spectrum combining all production mode categories. The data points show the measured distribution, with each event weighted by the expected signal-to-background ratio of its category. The red and blue curves show the distributions retrieved from the likelihood fit. The black curve shows the signal component, and the bottom panel displays the residuals of the data with respect to the fitted background component [161].

- two VBF-enriched signal regions with two jets consistent with VBF topology, featuring large transverse momenta and a wide separation in pseudorapidity, based on a second BDT,
  8. one selecting the events with higher scores to form a tight category,
  9. and a second one selecting events with lower scores for a loose category,
- and four additional categories capturing events depending on pseudorapidity and the momentum-based variable $p_{\mathrm{T}t}$, defined as the orthogonal component of the diphoton momentum with respect to the axis connecting the momentum vectors of the two photons, which separates events with different diphoton invariant mass resolution and further discriminates VBF and ggF production, the categories being
  10. a central high-$p_{\mathrm{T}t}$ category, requiring both photons within $|\eta| < 0.95$ and $p_{\mathrm{T}t} > 70\,\mathrm{GeV}$,
  11. a central low-$p_{\mathrm{T}t}$, requiring both photons within $|\eta| < 0.95$ and $p_{\mathrm{T}t} < 70\,\mathrm{GeV}$,
  12. a forward high-$p_{\mathrm{T}t}$ category, with at least one photons with $|\eta| > 0.95$ and $p_{\mathrm{T}t} > 70\,\mathrm{GeV}$,
  13. and finally a forward low-$p_{\mathrm{T}t}$ category, with at least one photons with $|\eta| > 0.95$ and $p_{\mathrm{T}t} < 70\,\mathrm{GeV}$.

The background in this analysis is mostly comprised of non-resonant $\gamma\gamma$, γj and jj production and is modeled using a two-dimensional sideband method or ABCD method (introduced in Section 4.4.2). The signal distribution is modeled as a double-sided Crystal Ball function (see Section 5.2.4), centered on the previously measured Higgs boson mass of $m_H = 125.09\,\mathrm{GeV}$. The signal extraction is performed with a likelihood fit. Figure 5.2 shows the result of this procedure, combined over all categories.

5.1.3 Combination procedure

The extraction of the parameters of interest relies on the method of likelihood ratio maximization, as described in Section 2.5, using the standard implementations thereof [163–165], and is based on earlier analyses conducted similarly [15]. The parameters of interest correspond to Higgs boson production cross sections or decay branching ratios, or reparametrizations thereof.

The five sets of production modes considered in this analysis encompass

- the ggF mechanisms of Higgs boson production,
- the VBF mechanisms of Higgs boson production,
- Higgs boson production in association with a vector boson decaying leptonically (VHlep),

- Higgs boson production in association with a vector boson decaying hadronically (VHhad), and
- Higgs boson production in association with top quarks (*top*), either singular in the form of tH or pairwise in the form of $t\bar{t}H$.

As the analysis presented here is not sensitive to the Higgs boson coupling to bottom-type quarks, it is hence grouped together with the dominant ggF production mode. All production modes grouped together in this way are assumed to be related by their nominal Standard Model ratio. Some of the reparametrizations presented here perform a stronger grouping of the individual processes.

The likelihood function constructed treats the number of events found in any specific category as an independent measurement following a Poisson distribution, or is based on the unbinned distribution of the diphoton invariant mass in the case of some $H\to\gamma\gamma$ categories with a low number of expected events.

In addition to the aforementioned parameters of interest, the likelihood model includes around 200 nuisance parameters, modeling systematic uncertainties of experimental or theoretical origin. Wherever possible, these are constrained by independent control measurements. The constraints themselves are modeled by multiplicative terms in the likelihood of a Gaussian or log-normal shape. Individual uncertainties are either treated as fully correlated, or as fully uncorrelated. Fully correlated uncertainties are modeled by a single nuisance parameter.

Sources of experimental uncertainties include the electromagnetic and jet energy scale systematic uncertainties, luminosity uncertainties, and others. All experimental uncertainties shared between the two analyses are fully correlated in the combined fit.

The sources of theoretical uncertainties include the PDF uncertainties [166, 167], the uncertainties in the modeling of the underlying event and parton shower (UE/PS), which is treated sharedly between the two channels, and the uncertainties related to the acceptance of the individual categories due to higher-order corrections in QCD. The corresponding parameter for the inclusive $n_{\text{jet}}\geq 2$ VBF categories is shared between $H\to\gamma\gamma$ and $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$, and so is the parameter for the splitting in the $p_{\text{T}t}$ categories [168, 169] in $H\to\gamma\gamma$ and the $n_{\text{jet}}=0$ and $n_{\text{jet}}=1$ categories [151, 170, 171] in $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$.

For the measurement of the signal strength, uncertainties on the predicted cross section are taken into account, including uncertainties from missing higher-order QCD corrections, as well as uncertainties on the PDFs and α_s. These uncertainties are fully correlated between the different decay channels. In addition, theoretical and parametric uncertainties on the decay branching ratios are included [19, 172–175]. The correlation scheme for these systematic uncertainties follows the strategy used for earlier combined measurements [15].

Table 5.2.: Predicted acceptances of the fiducial volume. The uncertainties arising from the finite size of the Monte Carlo samples are on the order of 5‰, and are not included.

Process	Contributing to	$\sigma_i(\lvert y_H \rvert < 2.5)/\sigma_i$
$gg \to H$	ggF	0.907
$qq' \to qq'H$	VBF	0.932
$q\bar{q}' \to WH(W \to \text{had})$	VHhad	0.870
$q\bar{q}/gg \to ZH(Z \to \text{had})$	VHhad	0.900
$q\bar{q}' \to WH(W \to \text{lep})$	VHlep	0.869
$q\bar{q} \to ZH(Z \to \text{lep})$	VHlep	0.900
$gg \to ZH(Z \to \text{lep})$	VHlep	0.965
$q\bar{q}/gg \to t\bar{t}H$	*top*	0.985

The measurement is performed in a "central" fiducial volume of $\lvert y_H \rvert < 2.5$ and extrapolated to the total phase space with acceptances calculated for a Standard Model scenario with simulated Monte Carlo samples and are shown in Table 5.2. The ratio between the central and the total cross section for $b\bar{b}H$ and tH are assumed to have a negligible impact on the calculation and are set to unity.

5.1.4 Products of cross sections and branching fractions

In this first parametrization, the parameters of interest are the products of cross sections σ_i for Higgs boson production via some mechanism i, and the branching fraction $\mathcal{B}^f$ to some final state f, which is either $\gamma\gamma$ or ZZ^*. These products $(\sigma\mathcal{B})_i^f$ are measured independently for each production and decay channel.

While the five production mechanisms and two final states investigated here would allow for ten independent measurements of $(\sigma\mathcal{B})_i^f$, three of these are not significantly constrained by the data and are fixed to their respective Standard Model prediction, namely $(\sigma\mathcal{B})_{\text{VHhad}}^{ZZ^*}$, $(\sigma\mathcal{B})_{\text{VHlep}}^{ZZ^*}$, and $(\sigma\mathcal{B})_{top}^{ZZ^*}$, as the $H{\to}ZZ^*{\to}\ell^+\ell^-\ell'^+\ell'^-$analysis does not have a dedicated event category to measure Higgs boson production in association with top quarks, and the VH categories provide very low statistical power. This is summarized in Table 5.3. For the parameters fixed to their Standard Model values, the corresponding theoretical uncertainties have been taken into account.

The resulting values of the products of cross sections and branching fractions $(\sigma\mathcal{B})_{\mathtt{i}}^f$ are shown in Fig. 5.3 and Table 5.4. Small differences to the measurements in the individual analyses may arise due to the nuisance parameter correlation scheme and the grouping of production modes.

The measurements concerning ggF and VBF Higgs boson production achieved the highest precision. The respective event categories for either of these measurements

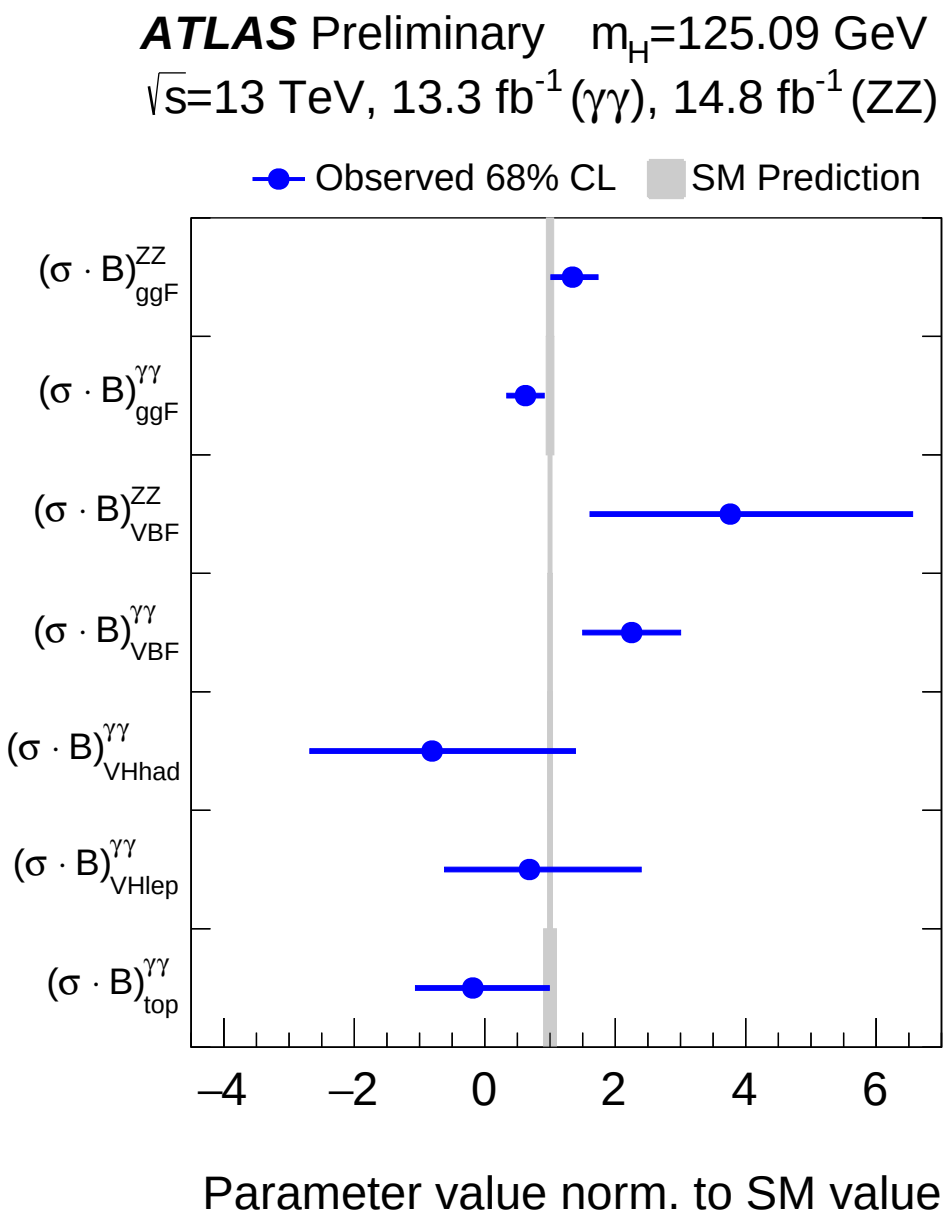


Figure 5.3.: Measurements of products of production cross sections and branching fractions $(\sigma\mathcal{B})_i^f$. The fit results are normalized to the respective SM prediction. The gray bands indicate the theoretical uncertainties. The blue error bars show the full uncertainty, including experimental and theoretical uncertainties [23].

Table 5.3.: Parameters of interest for the measurement of $(\sigma\mathcal{B})_i^f$.

Production & Decay	ggF	VBF	VHhad	VHlep	*top*
$H\to\gamma\gamma$	$(\sigma\mathcal{B})^{\gamma\gamma}_{\text{ggF}}$	$(\sigma\mathcal{B})^{\gamma\gamma}_{\text{VBF}}$	$(\sigma\mathcal{B})^{\gamma\gamma}_{\text{VHhad}}$	$(\sigma\mathcal{B})^{\gamma\gamma}_{\text{VHlep}}$	$(\sigma\mathcal{B})^{\gamma\gamma}_{top}$
$H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$	$(\sigma\mathcal{B})^{ZZ}_{\text{ggF}}$	$(\sigma\mathcal{B})^{ZZ}_{\text{VBF}}$	SM	SM	SM

Table 5.4.: Best fit values of $(\sigma\mathcal{B})^f_{\mathtt{i}}$ for each specific channel $\mathtt{i}\to H\to f$, as obtained from the generic parametrization with 7 parameters as given in Table 5.3. The SM predictions [19] are shown for each $(\sigma\mathcal{B})^f_{\mathtt{i}}$.

		$H\to ZZ^*$	$H\to\gamma\gamma$
ggF	Best fit value (pb)	$1.58\,^{+0.46}_{-0.39}$	$0.063\,^{+0.030}_{-0.029}$
	SM prediction (pb)	1.18 ± 0.07	0.101 ± 0.006
VBF	Best fit value (fb)	350^{+260}_{-200}	$18\,^{+6}_{-6}$
	SM prediction (fb)	93.0 ± 2.8	8.00 ± 0.29
VHhad	Best fit value (fb)	fixed to SM	$-2.5\,^{+6.8}_{-5.8}$
	SM prediction (fb)	36.0 ± 1.2	3.09 ± 0.12
VHlep	Best fit value (fb)	fixed to SM	$1.0^{+2.5}_{-1.9}$
	SM prediction (fb)	17.0 ± 0.5	1.46 ± 0.05
top	Best fit value (fb)	fixed to SM	$-0.3^{+1.6}_{-1.2}$
	SM prediction (fb)	15.9 ± 1.5	1.36 ± 0.13

typically have substantial contributions from the other. Figure 5.4 shows the likelihood contours of the two measurements in the $(\sigma\mathcal{B})^f_{\text{ggF}}$–$(\sigma\mathcal{B})^f_{\text{VBF}}$ plane divided by their respective Standard Model value.

Among the experimental uncertainties, the parameters modeling the electromagnetic energy resolution and photon identification efficiency are the dominant contributions. The most prominent theory uncertainties are the ones modeling acceptance of gluon fusion events into categories depending on the number of jets. The compatibility with the Standard Model prediction corresponds to a p-value of 11%.

Cross sections

The second parametrization focuses on the measurement of the individual production cross sections σ_i under the assumption of Standard Model Higgs boson decay branching fractions, taking into account the theory uncertainties on the corresponding predictions.

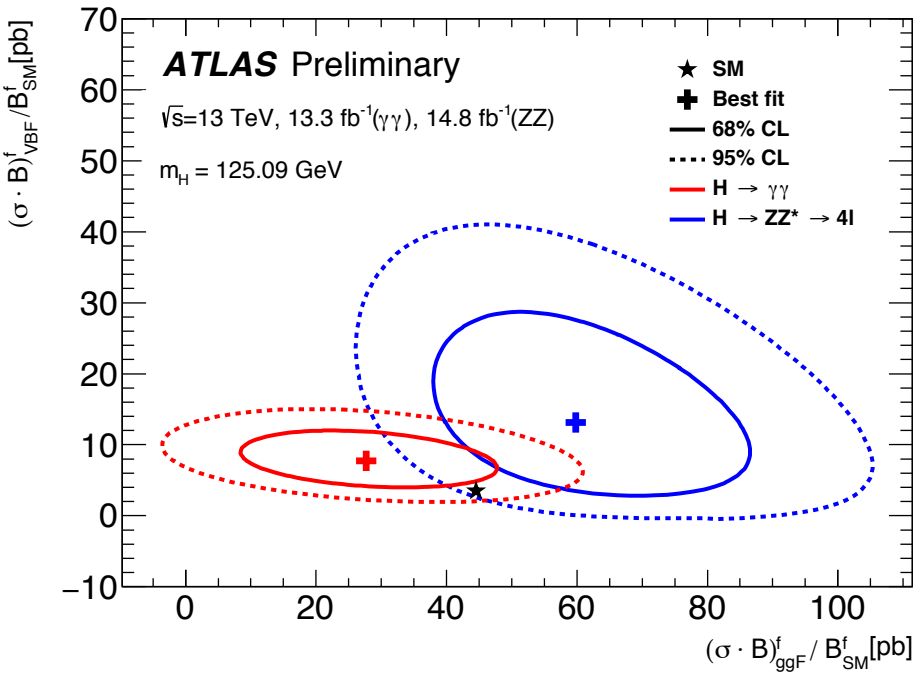


Figure 5.4.: Contours in the $(\sigma\mathcal{B})^f_{\mathrm{ggF}}/\mathcal{B}^f_{\mathrm{SM}}$–$(\sigma\mathcal{B})^f_{\mathrm{VBF}}/\mathcal{B}^f_{\mathrm{SM}}$ plane as measured in $H \to \gamma\gamma$ and $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ and the respective Standard Model predictions [23].

Table 5.5.: Best fit values of the production cross sections σ_i assuming Standard Model Higgs boson decay branching fractions [19].

	Best fit value (pb)	SM prediction (pb)
σ_{ggF}	$47.8\,^{+9.8}_{-9.4}$	44.5 ± 2.3
σ_{VBF}	$7.9\,^{+2.8}_{-2.4}$	3.52 ± 0.07
σ_{VHhad}	$-2.5\,^{+2.9}_{-2.6}$	1.36 ± 0.03
σ_{VHlep}	$0.32\,^{+1.07}_{-0.79}$	0.64 ± 0.02
σ_{top}	$-0.11\,^{+0.67}_{-0.54}$	0.60 ± 0.06

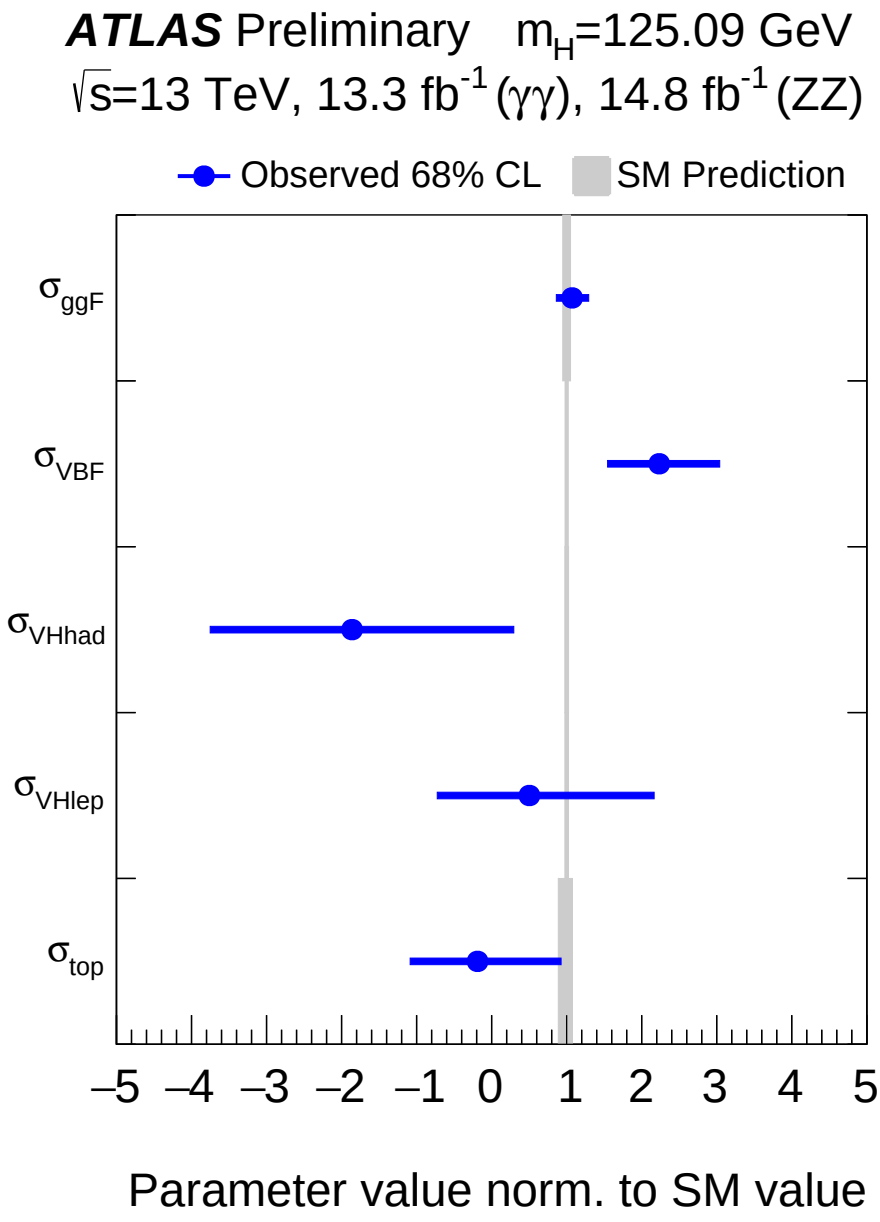


Figure 5.5.: Production cross section measurements assuming Standard Model branching fractions [19]. The gray bands indicate the theoretical uncertainties on the predictions. The blue error bars show the full uncertainty, including experimental uncertainties and theoretical uncertainties [23].

Figure 5.5 and Table 5.5 show the corresponding fit results. The compatibility between the measurement and the Standard Model prediction corresponds to a p-value of 21%.

Ratios of cross sections and of branching fractions

Ratios of cross sections and branching fractions are extracted from a combined fit to the data by normalizing the production cross section to σ^{ggF} and the decay branching fraction to $\mathcal{B}^{ZZ^*}$. The product of cross section and branching fraction can then be expressed as

$$(\sigma\mathcal{B})_i^f = (\sigma\mathcal{B})_{\text{ggF}}^{ZZ^*} \cdot \frac{\sigma_i}{\sigma_{\text{ggF}}} \cdot \frac{\mathcal{B}^f}{\mathcal{B}^{ZZ^*}}. \tag{5.1.1}$$

The present analysis is only sensitive to

$$(\sigma\mathcal{B})_{\text{ggF}}^{ZZ}, \quad \sigma_{\text{VBF}}/\sigma_{\text{ggF}} \quad \text{and} \quad \mathcal{B}^{\gamma\gamma}/\mathcal{B}^{ZZ}, \tag{5.1.2}$$

and the remaining ratios are profiled. The results of the measurement are shown in Fig. 5.6 and compared to their Standard Model expectation, showing compatibility corresponding to a p-value of 5%.

The evidence for Higgs boson production via vector boson fusion measured in events collected at $\sqrt{s} = 13\,\text{TeV}$ amounts to 4.0σ, where the Standard Model expectation with the given data set was 1.9σ using the asymptotic approximation [48].

Table 5.6.: Best-fit values of the cross section $(\sigma\mathcal{B})_{\text{ggF}}^{ZZ}$ and of the ratios $\sigma_{\text{VBF}}/\sigma_{\text{ggF}}$ and $\mathcal{B}^{\gamma\gamma}/\mathcal{B}^{ZZ}$. The remaining ratios between production cross sections and ggF are profiled in the combined fit. The SM predictions [19] are shown in the last column.

Parameter	Best-fit value	SM prediction
$(\sigma\mathcal{B})_{\text{ggF}}^{ZZ}$ (pb)	$1.67\,^{+0.41}_{-0.37}$	1.18 ± 0.07
$\sigma_{\text{VBF}}/\sigma_{\text{ggF}}$	$0.25\,^{+0.15}_{-0.10}$	0.079 ± 0.004
$\mathcal{B}^{\gamma\gamma}/\mathcal{B}^{ZZ}$	$0.041\,^{+0.015}_{-0.013}$	0.086 ± 0.003

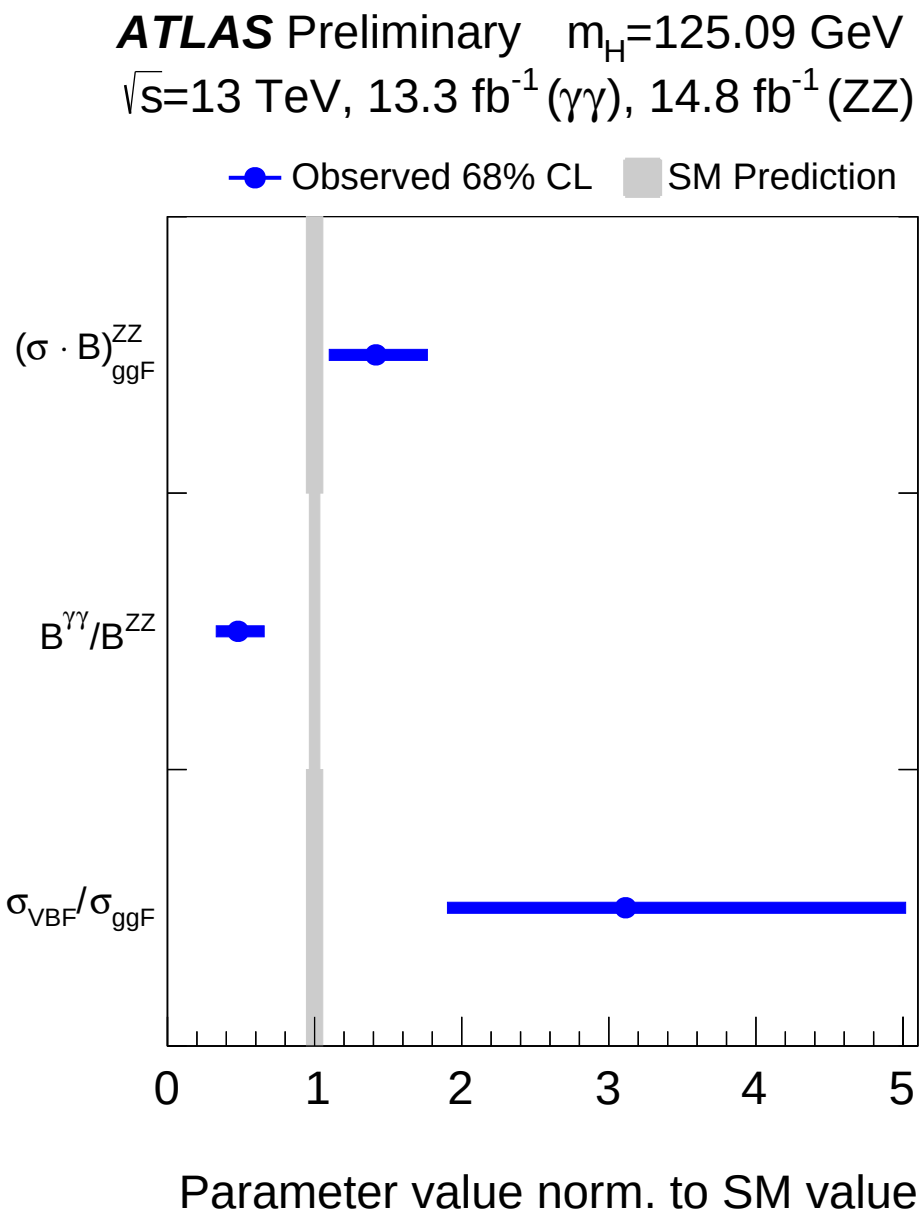


Figure 5.6.: Measurement of $(\sigma\mathcal{B})^{ZZ}_{\mathrm{ggF}}$, $\mathcal{B}^{\gamma\gamma}/\mathcal{B}^{ZZ}$ and $\sigma_{\mathrm{VBF}}/\sigma_{\mathrm{ggF}}$. The gray bands indicate the theoretical uncertainties. The remaining ratios between production cross sections and ggF are profiled in the combined fit. The blue error bars show the full uncertainty, including experimental and theoretical uncertainties [23].

5.1.5 Signal strength measurements

The simplest parametrization is that of a total signal strength parameter μ, as introduced in Section 2.4.2, defined as the total observed yield with respect to the Standard Model expectation. The signal strength parameter

$$\mu = \frac{(\sigma\mathcal{B})}{(\sigma\mathcal{B})^{\mathrm{SM}}} \tag{5.1.3}$$

acts as a single scaling factor for all modes of Higgs boson production and decay after extrapolation to the full phase space including $|y_H| \geq 2.5$.

The observed evidence for Higgs boson production amounts to a local significance of about 10σ (8.6σ expected). The global signal strength measurement finds $\mu = 1.13^{+0.18}_{-0.17}$. The compatibility to the Standard Model corresponds to a p-value of 43%.

5.1.6 Combination of total cross sections

Table 5.7 and Fig. 5.7 show the Standard Model predictions as well as the measured values of the total $pp \to H + X$ cross sections at center-of-mass energies of $\sqrt{s} = 7$ and 8 TeV as presented in Section 3.3 [15] as well as $\sqrt{s} = 13$ TeV, measured in $H \to \gamma\gamma$ and $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ and their combination.

The breakdown of the uncertainties into the individual sources is obtained by fixing all other nuisance parameters to their best-fit values and calculating the quadratic difference of the total uncertainty with respect to the result of the full unconditional fit. For the two final states presented here, the systematic uncertainties are smaller than the statistical ones for all three center-of-mass energies, and the results are compatible with each other and with the Standard Model prediction.

Table 5.7.: Total $pp \to H + X$ cross sections $\sqrt{s} = 7$ and 8 TeV [15] and 13 TeV [23]. The Standard Model predictions [19] assume a Higgs boson mass of 125.09 GeV [35].

	$H \to \gamma\gamma$	$H \to ZZ^*$	Combination	SM predictions
$\sqrt{s} = 7$ TeV	35^{+13}_{-12} pb	33^{+21}_{-16} pb	$34 \pm 10\,(\mathrm{stat.})^{+4}_{-2}\,(\mathrm{syst.})$ pb	19.2 ± 0.9 pb
$\sqrt{s} = 8$ TeV	$30.5^{+7.5}_{-7.4}$ pb	37^{+9}_{-8} pb	$33.3^{+5.5}_{-5.3}\,(\mathrm{stat.})^{+1.7}_{-1.3}\,(\mathrm{syst.})$ pb	24.5 ± 1.1 pb
$\sqrt{s} = 13$ TeV	37^{+14}_{-13} pb	81^{+18}_{-16} pb	$59.0^{+9.7}_{-9.2}\,(\mathrm{stat.})^{+4.4}_{-3.5}\,(\mathrm{syst.})$ pb	$55.5^{+2.4}_{-3.4}$ pb

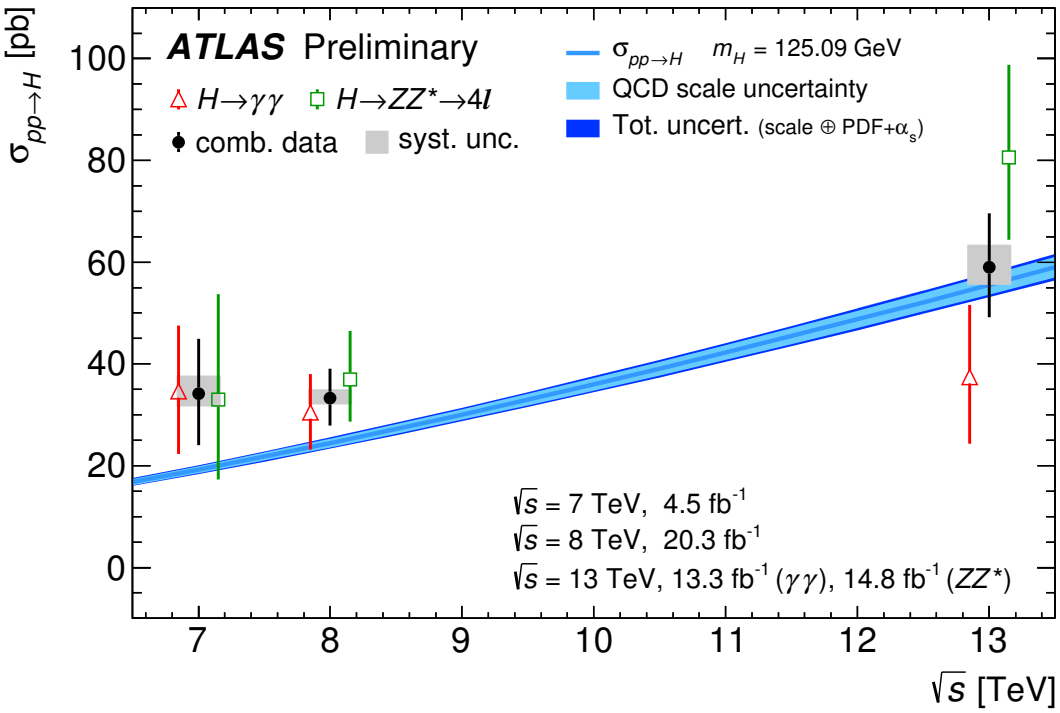


Figure 5.7.: Total $pp \to H + X$ cross section measurements compared to Standard Model predictions at up to N^{3}LO in QCD, assuming a Higgs boson mass of 125.09 GeV [35], shown as a smooth curve with the light (dark) blue band indicating the uncertainty from missing higher-order QCD corrections (the total uncertainty). The red triangles show the measurements from the $H\to\gamma\gamma$ channel, the green rectangles show the measurements from the $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$ channel, and the black dots show their combinations. The gray bands represent the systematic uncertainty of the combined measurement, while the lines denote the total uncertainty [23].

5.1.7 Combination results

Higgs boson production is observed at $\sqrt{s} = 13\,\text{TeV}$ with a local significance of 10σ (8.6σ expected) in the decay modes $H\to\gamma\gamma$ and $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$, with the evidence for vector boson fusion amounting to about 4σ (1.9σ expected). In all parametrizations entertained, including differentiation between the two decay modes and up to five different modes of Higgs boson production, no significant deviation from the Standard Model predictions has been observed. The global signal strength was measured as $\mu = 1.13^{+0.18}_{-0.17}$. The total cross section of $pp \to H + X$ in the full phase space was determined from fiducial cross section measurements as $59.0^{+9.7}_{-9.2}\,(\text{stat.})^{+4.4}_{-3.5}\,(\text{syst.})$ pb.

5.2 Modeling Techniques

Any deviation in the parameters of the underlying theory – be it an effective field theory or otherwise – has complicated and not trivially predictable consequences on the final predictions arising from that theory. The standard procedure of computing these predictions involves a multi-step, resource intensive process, including calculation of

1. the Matrix Elements involved in the hard scatter,
2. the parton-level cross section, and
3. the parton shower and hadronization

as well as the simulation of

4. additional interactions from the same or different bunch crossings,
5. interaction of the particles with the detector material,
6. the data acquisition by the detector, including the trigger,

and finally the reconstruction of the physical objects used for analysis.

In the parlance of experimental high energy physics, the items 1-3 are typically referred to as the "generation" of events by some Monte Carlo generator, whereas the second block of items is usually referred to as "simulation".

5.2.1 Computational cost

It is worthwhile to note that the simulation stage typically consumes several orders of magnitude more computing time than the generation stage, as apparent from the comparison of Fig. 5.8 with Table 5.8. The simulation of Monte Carlo events is the most costly task of the worldwide LHC computing grid, as apparent from Fig. 5.9. Samples, events or particles that have not been processed by the simulation and reconstruction are usually referred to as "truth", whereas the final physics objects that can be directly compared to measured data events are referred to as "reconstructed".

While the bulk of this computing time is spent for the creation of large Standard Model samples that are central to Standard Model measurements or serve as background estimates for many searches for physics beyond the Standard Model (BSM), the creation of dedicated signal samples for such searches is often much less costly, as the higher efficacy allows them to be produced in much smaller sizes. However, this virtue is rendered invalid once the theories subject to testing are highly parametric.

In this case, the method of likelihood fitting introduced in Section 2.5 would in principle call for the creation of a dedicated Monte Carlo sample, fully generated and simulated for this precise parameter configuration for the evaluation of the

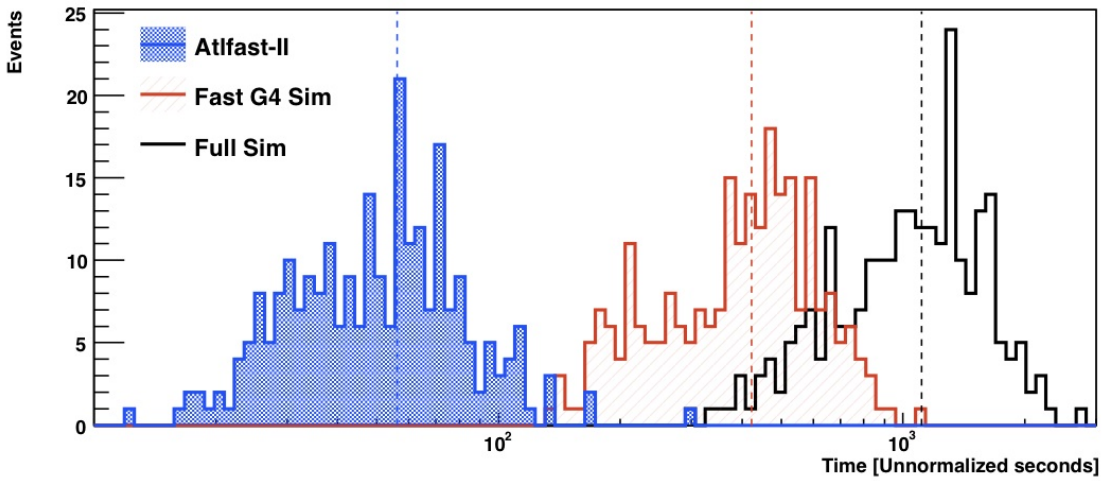


Figure 5.8.: Distributions of CPU time for 250 $t\bar{t}$ events in full, Fast GEANT, and ATLFAST-II simulations. Vertical dotted lines denote the averages of the distributions [141].

Table 5.8.: Run times for generating 10,000 events using MADGRAPH5. For all processes, $p = j = g/u/\bar{u}/c/\bar{c}/d/\bar{d}/s/\bar{s}$, $l^{\pm} = e^{\pm}/\mu^{\pm}$. The run times for 0-, 1- and 2-jet processes are for a laptop computer with 1.06 GHz Intel Core Duo CPU, while the 3-, 4- and 5-jet run times (marked by $\star$) are for a 128-core computer cluster with Intel Xeon 2.50 GHz CPUs [176].

Process	CPU time	in	
$pp \to W^+j$	1:55	min	
$pp \to W^+jj$	4:19	min	
$pp \to W^+jjj$	8:14	min	$\star$
$pp \to W^+jjjj$	46:50	min	$\star$
$pp \to W^+jjjjj$	11:39	h	$\star$
$pp \to l^+l^-j$	3:00	min	
$pp \to l^+l^-jj$	11:52	min	
$pp \to l^+l^-jjj$	16:38	min	$\star$
$pp \to l^+l^-jjjj$	2:45	h	$\star$
$pp \to t\bar{t}$	1:55	min	
$pp \to t\bar{t}j$	3:52	min	
$pp \to t\bar{t}jj$	32:37	min	
$pp \to t\bar{t}jjj$	23:15	min	$\star$

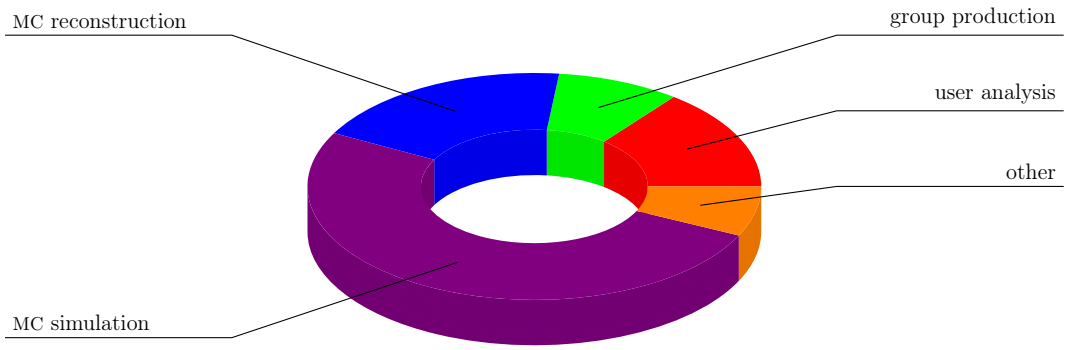


Figure 5.9.: ATLAS usage of worldwide LHC computing grid resources in 2012 [177].

likelihood at some different new physics parameter set. Seeing that a likelihood fit of a complicated, combined model such as the ones considered in Section 3.3.3 or Section 5.1 can easily call for hundreds of thousands of evaluations, this is clearly unfeasible. The techniques presented in the remainder of this chapter focus on overcoming this challenge, outlining existing and novel strategies how complicated models can be built such that they

- are computationally feasible, and
- model the underlying physics as accurately as possible.

5.2.2 Matrix Element Reweighting

One common practice to obtain a large quantity of sampling points in the BSM parameter space is the technique of *Matrix Element Reweighting*.

Each individual Monte Carlo event is usually produced with some *event weight*. These event weights are technical artifacts allowing the Monte Carlo generators to more easily and efficiently sample the phase space by assigning individual points a larger weight than others – and while some generators like MADGRAPH produce samples with uniform weight at least at leading order, others can produce large (or even negative) event weights. Weights are also at the analysis stage commonly applied in the context of reweighting the Monte Carlo generated samples to reflect the pile-up distribution or other physical observables in data in the hope of improving the general quality of the physics modeling.

Method description

The idea of Matrix Element Reweighting takes this concept one step further: The weight of each event is necessarily proportional to the Matrix Element modulus squared, evaluated at the corresponding phase space point. If one were to change the parameters of the underlying theory, the Matrix Element and thus the weight would change accordingly. As weights are multiplicative and do not alter any of

the observable quantities of the event, they can be applied at any processing stage. Thus, one can *reweight* an event from the theory parameters it was originally generated for to a different set of theory parameters, using

$$w_o = w_i \cdot \frac{|\mathcal{M}_o|^2}{|\mathcal{M}_i|^2}, \tag{5.2.1}$$

where w_i and w_o stand for the input and output weights of the event, and $\mathcal{M}_i$ and $\mathcal{M}_o$ for the Matrix Elements accordingly.

This technique can be applied on readily simulated and reconstructed samples and hence provides an extremely efficient approach of obtaining simulated signal samples, as the calculation of the Matrix Element is extremely resource-efficient compared to simulating and reconstructing the event.

Discussion

Matrix Element Reweighting is an extremely powerful and useful tool that can be used for many different applications, not only reweighting between different BSM physics scenarios, but also for *a posteriori* changing empiric parameters of the Monte Carlo generator to assess systematic uncertainties, for example examining the effects of the PDF set used for sample generation. For this type of task, Matrix Element Reweighting is an invaluable mechanism to save computing resources, and for many applications, there is no real alternative. However, some properties of the reweighted samples pose challenges for the analyses using them:

- If the target parameter set is very different from the initial one, the weights of individual events can become very large. The resulting samples are effectively dominated by a few high-weight events from the tails of the kinematic distributions, reducing the statistical power significantly. Strategies to overcome this challenge by simulating more events or having samples generated closer to the target parameters undo the benefits of the method.
- All samples obtained from reweighting the same source sample are statistically dependent. Taking this interdependence into account properly can be challenging in some cases and will in general increase the uncertainty on the prediction.

It should be noted, however, that the last item can also be seen as a benefit in some scenarios, providing the analyst with an extremely high degree of reproducibility and opening the possibility to observe exact measures of agreement when assessing newly developed methods and performing closure checks.

Implementation

The implementation of the Matrix Element Reweighting technique is principally straight forward in that the analysis software just needs to be interfaced back to a Matrix Element calculation functionality present in any Monte Carlo generator – some generators like MadGraph5 even provide methods to automatically generate code to perform this task for a specific physics process.

However, the details are often more complicated. As the Matrix Element calculator needs access to the four-momenta of the initial particles produced in the hard scatter, they need to be identified and accessed. The principles of operation of the computing infrastructure pose several challenges to this task.

The technical representation of the event typically contains a very large number, often several thousand, of truth particles originating from the parton shower, including several stages of intermediate unstable particles, as well as additional particles from initial or final state radiation. Not only can it be difficult to identify which particle instances correspond to the results of the hard scatter calculated by the Monte Carlo generator itself (as opposed to, say, the parton shower calculation), but also the initial and final state radiation can boost the center-of-mass frame and actually change the four-momenta of these particles. Hence, even if the identification of particles to be passed to the Matrix Element calculator is successful, the result of its calculation may not be the same as if applied directly on the original hard scatter event.

In order to circumvent these difficulties, one might try to pass along the original hard scatter event, such that it can be used for the Matrix Element calculation. However, also this approach faces difficulties. As samples are often split into smaller chunks for processing and not necessarily merged in the same order, and unique event identifiers are not always available, it can be difficult to identify the reconstructed events with their original versions *a posteriori*.

These two approaches are referred to as *direct reweighting* in the case where the post-parton-shower truth particles are passed to the Matrix Element calculation, and as *matched reweighting* in the case where the truth content of the event is matched to some original hard scatter event which is then used to recalculate the Matrix Element.

Figure 5.10 shows distributions of two kinematic variables, the transverse mass m_{T} and the azimuthal angle between the dilepton pair $\Delta\phi_{\ell\ell}$, for $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ events produced via VBF or VH production modes for different EFT scenarios. Here, an event selection equivalent to the preselection used for the analysis presented in Section 4.6 has been applied. The parameter configurations of these scenarios in terms of the Higgs Characterization Model introduced in Section 2.4.2 are shown in Fig. 5.11. Notably, the predictions obtained via the two different methods of

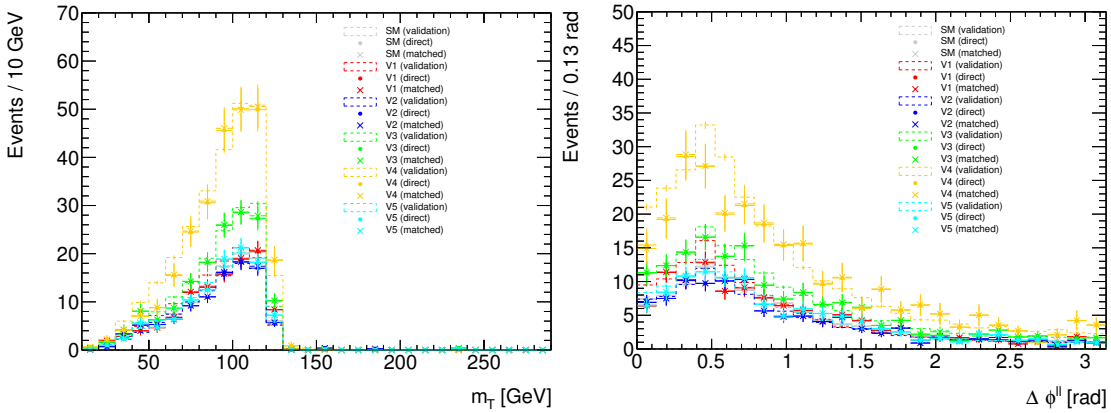


Figure 5.10.: Distributions of events in transverse mass m_{T} and azimuthal angular separation between the lepton pair $\Delta\phi_{\ell\ell}$ in $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ events with Higgs bosons produced via the VBF and VH production modes, as predicted with Matrix Element Reweighting. For each configuration tested, three data rows are depicted: One independent validation sample obtained with a Monte Carlo generator, one obtained with direct reweighting, and one obtained with matched reweighting. The configurations of the different samples are displayed in Fig. 5.11.

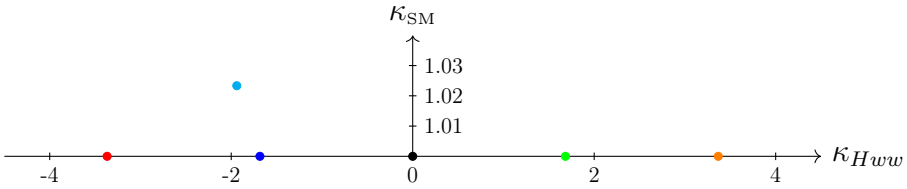


Figure 5.11.: Validation samples used for the Matrix Element Reweighting, generated with MADGRAPH5_aMC@NLO 2.3.3.p1, showered with PYTHIA8.2.12 using the Higgs Characterization model introduced in Section 2.4.2. All samples have been generated with $\cos\alpha = 1/\sqrt{2}$ and $\Lambda = 1\,\mathrm{TeV}$.

Matrix Element Reweighting as well as from independent validation samples are in good agreement.

This can also be seen in Fig. 5.12, which shows the distributions of differences in event weights between the two methods for all of the above scenarios. While the most probable value of the difference is zero in all cases, the distributions do have far-reaching tails. This is a result of the fact that the *direct* reweighting method uses the four-momenta of the particles after initial and final state radiation, whereas the *matched* variant uses the original four-momenta as produced by the Monte Carlo generator. It is however reassuring to see that the tails are roughly symmetric, which motivates the good agreement visible for the kinematic variables.

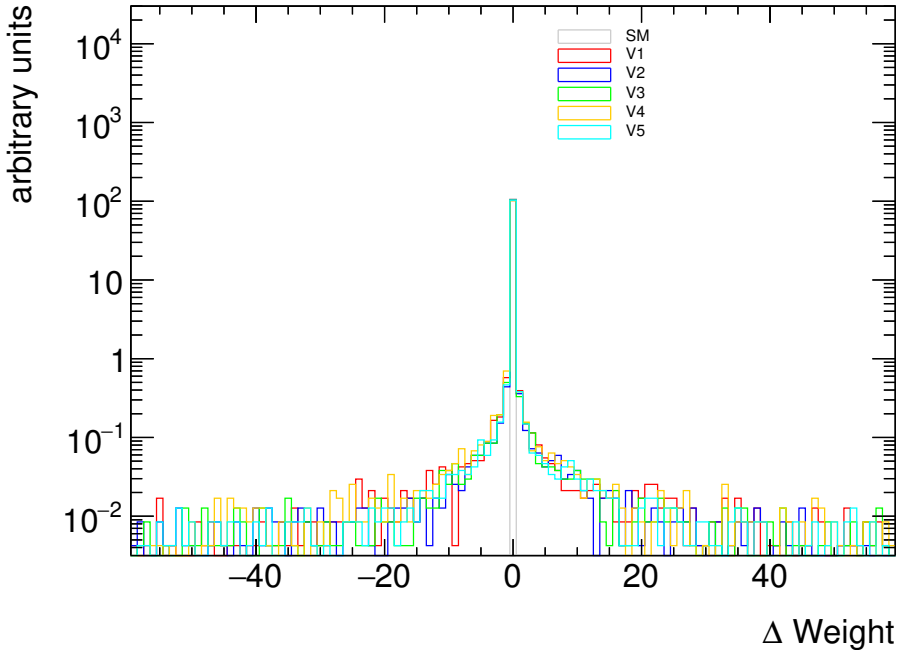


Figure 5.12.: Distribution of difference in event weight as predicted by direct and matched reweighting. The ΔWeight displayed on the horizontal axis corresponds to the event-by-event difference in weight, where the event weight obtained by matched reweighting is subtracted from the one obtained by direct reweighting.

5.2.3 Morphing

As established in the previous section, it is not feasible to use Monte Carlo generators to evaluate a likelihood during a numeric minimization process. The general strategy in this case is to evaluate the prediction on some appropriately chosen samples, and then interpolate or "morph" between them during the minimization. An illustration of this process can be seen in Fig. 5.13. A wide variety of different interpolation techniques is available and commonly used for different scenarios.

Vertical morphing

If the target function $f(x \mid m)$ is known for n values of m, the Taylor expansion of f up to order $n-1$ around some reference value m_0 reads

$$f(x \mid m_i) \approx \sum_{j=0}^{n-1} (m_i - m_0)^j \, \frac{1}{j!} \frac{d^{(j)}}{dm^{(j)}} f(x \mid m_0) \,. \tag{5.2.2}$$

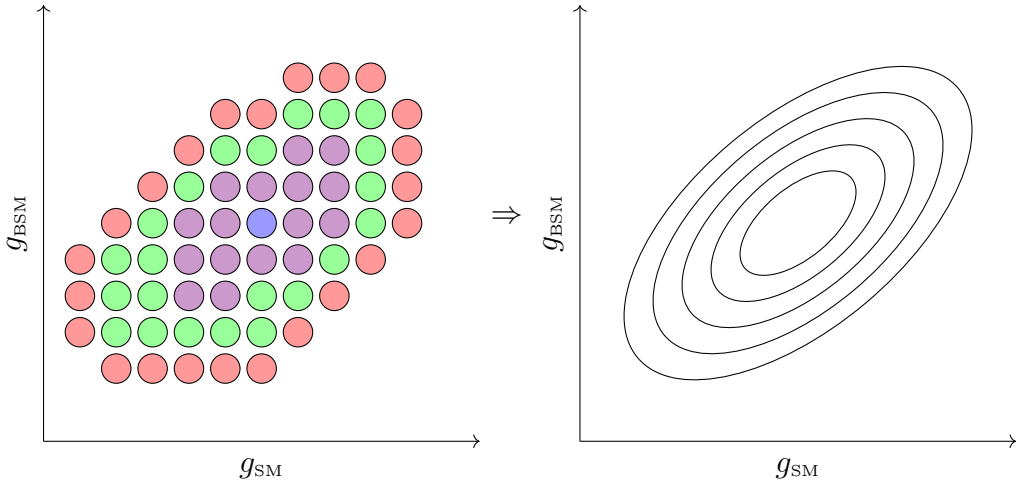


Figure 5.13.: Likelihood function of two parameters of interest, g_{SM} and g_{BSM}. A discrete representation with a finite set of samples (left) and a continuous representation obtained by some interpolation technique (right) are shown.

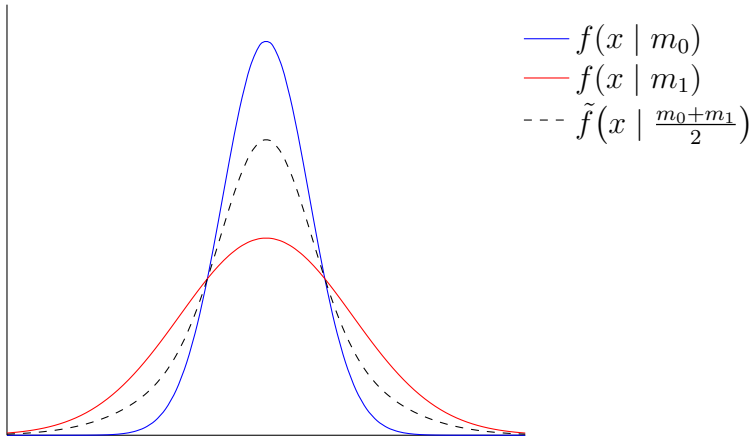


Figure 5.14.: Simple example for vertical interpolation, using $n = 2$ samples for a Gaussian curve.

Introducing

$$M_{ij} = (m_i - m_0)^j \quad \text{and} \quad f'_j(x \mid m_0) = \frac{1}{j!}\frac{d^{(j)}}{dm^{(j)}} f(x \mid m_0)\,, \tag{5.2.3}$$

this can be recast in matrix notation

$$f(x \mid m_i) \approx M_{ij} f'_j(x \mid m_0)\,, \tag{5.2.4}$$

such that

$$f'_j(x \mid m_0) \approx M^{-1}_{ij} f(x \mid m_i)\,. \tag{5.2.5}$$

The prediction of f at any value m' using vertical interpolation is then given as

$$\tilde{f}(x \mid m') = \sum_{i,j=0}^{n-1} (m' - m_0)^j \, M^{-1}_{ij} f(x \mid m_i)\,. \tag{5.2.6}$$

The behavior of this kind of interpolation can best be understood for the simplest case of $n = 2$, in which case the interpolation formula is exactly the weighted average

$$\tilde{f}(x \mid m') = f(x \mid m_0) + \frac{m' - m_0}{m_1 - m_0} f(x \mid m_1) + \frac{m' - m_1}{m_1 - m_0} f(x \mid m_0)\,. \tag{5.2.7}$$

An example for this type of vertical interpolation can be seen in Fig. 5.14.

This purely vertical interpolation technique is very useful as long as the mean of the input distributions is identical. If this is not the case, the results are not very satisfying, as apparent from Fig. 5.15.

Moment morphing

In order to improve the morphing technique such that differences in statistical moments are taken into account, the vertical interpolation technique can be supplemented with a coordinate transformation.

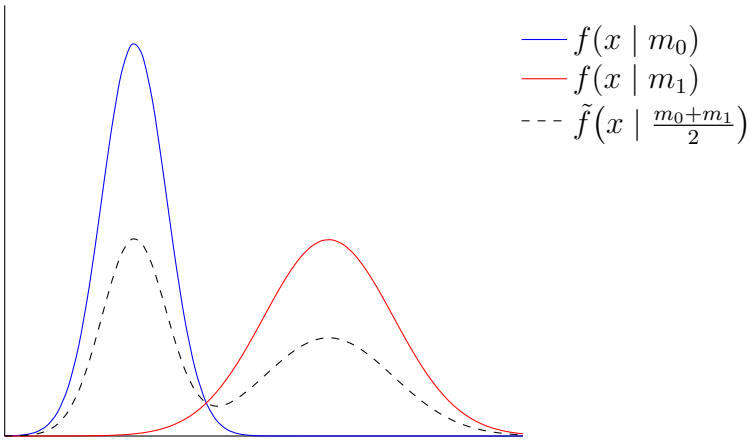


Figure 5.15.: Example for vertical interpolation in a simple $n = 2$ case where the input distributions have very different mean values.

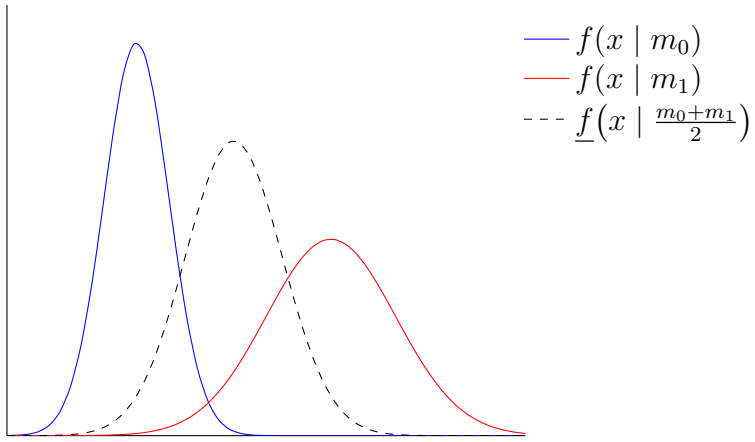


Figure 5.16.: Simple example for moment morphing, using $n = 2$ samples for a Gaussian curve.

If the distribution $f(x \mid m_i)$ has mean $\mu(m_i)$ and standard deviation $\sigma(m_i)$, then the estimated mean and standard deviation of the output distribution are

$$\tilde{\sigma}(m') = \sum_{i,j=0}^{n-1} (m' - m_0)^j \, M_{ij}^{-1} \sigma(m_i) \tag{5.2.8}$$

$$\tilde{\mu}(m') = \sum_{i,j=0}^{n-1} (m' - m_0)^j \, M_{ij}^{-1} \mu(m_i) \,. \tag{5.2.9}$$

Using these, one can employ the coordinate transformations

$$\xi_i(x \mid m') = \frac{x - \tilde{\mu}(m')}{\tilde{\sigma}(m')} \cdot \sigma(m_i) + \mu(m_i) \tag{5.2.10}$$

to map the predicted range of values in x to an interval corresponding to the statistical moments of the distribution corresponding to m_i. Notably, vertical interpolation is used here on the statistical momenta themselves. In this interval, each input distribution can be evaluated. Afterwards, the final prediction can be obtained by vertically interpolating between these remapped distributions.

The prediction of f at any value m' using moment morphing is then given as

$$\underline{f}(x \mid m') = \sum_{i,j=0}^{n-1} (m' - m_0)^j \, M_{ij}^{-1} f(\xi_i(x \mid m') \mid m_i) \,. \tag{5.2.11}$$

Figure 5.16 shows the same example as Fig. 5.15, but using moment morphing instead of vertical interpolation. Clearly, the results are much more satisfactory.

5.2.4 Considerations

The vertical interpolation and moment morphing are just two examples of a wide range of interpolation techniques. They are frequently used for data analysis in high energy physics and have been developed in this community [178], but are ultimately agnostic of the underlying physics. Rather, their ability to accurately model the physical distribution relies on a rather fine sampling of the available parameter space, requiring a grid of pre-generated samples to interpolate between. It is obvious that any approach that does not contain additional information about the intricate details of QFT can never hope to accurately model effects like quantum mechanical interference.

In some cases, it might be justifiable to take a predicted distribution and reparametrize it using some analytical function. One example for this approach

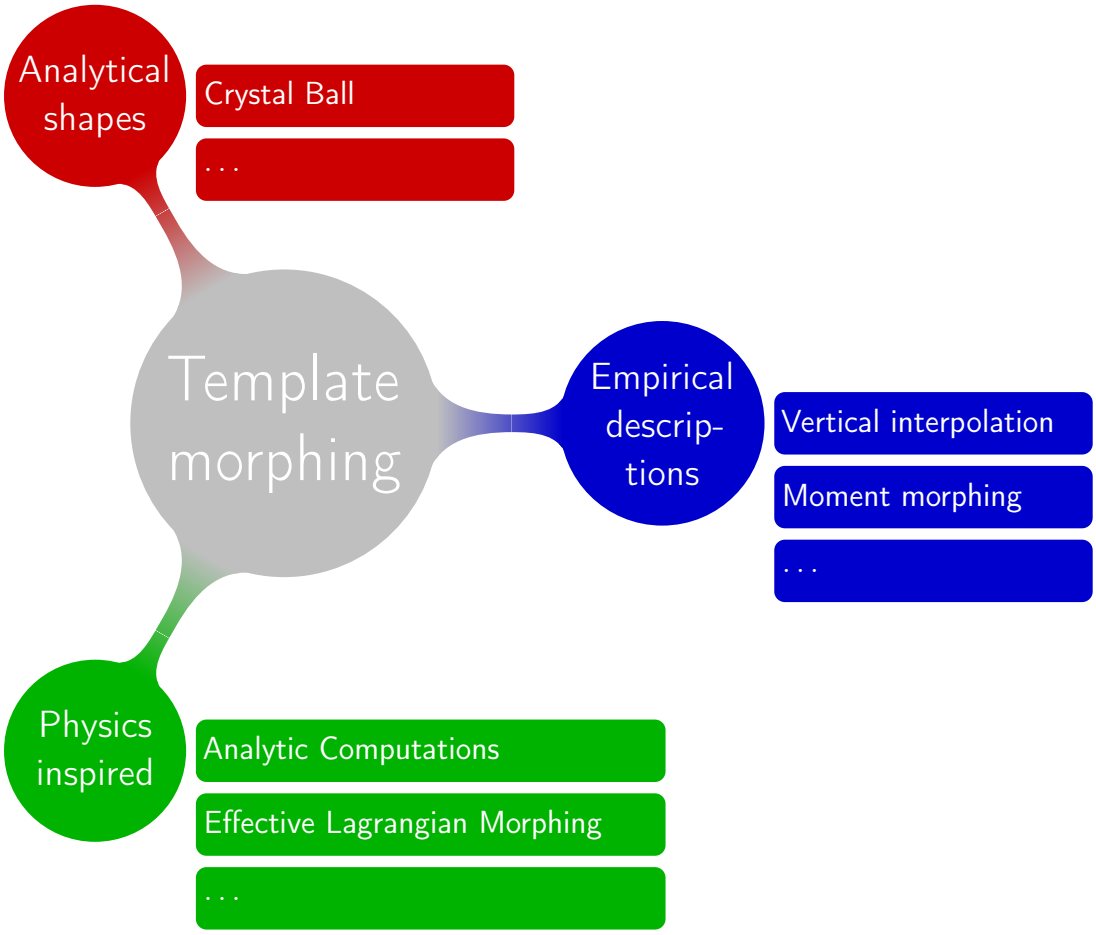


Figure 5.17.: Mind map of common morphing techniques.

would be the double-sided Crystal Ball function used in the $H \to \gamma\gamma$ analysis [179]. This function is defined as

$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n} & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases}, \tag{5.2.12}$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \tag{5.2.13}$$

$$B = \frac{n}{|\alpha|} - |\alpha| \tag{5.2.14}$$

$$N = \frac{1}{\sigma(C+D)} \tag{5.2.15}$$

$$C = \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \tag{5.2.16}$$

$$D = \sqrt{\frac{\pi}{2}}\left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right) \tag{5.2.17}$$

and where N is a normalization factor and α, n, $\bar{x}$ and σ are parameters which are fitted using the data. The symbol erf refers to the error function.

Even though the function is very well suited for the case at hand, the concrete choice of parametrization for any specific application is often arbitrary and hard to justify, and the accuracy of the result ultimately relies on the assumption that the chosen parametrization models the prediction well within the parameter range.

In rare cases, the dependency of some individual physical observable on the parameters might even have an analytic form, at least to some approximation. However, short of finding such an analytic representation for a wide variety of physics processes, there is no apparent "silver bullet" for efficiently and accurately modeling physics distributions.

In summary, the three available approaches are:

- Empiric modeling, using some interpolation method akin to vertical or moment morphing.
- Analytic parametrizations, like the Crystal Ball parametrization of the background used in the $H \to \gamma\gamma$ analysis [179].
- Physics inspired computations, exploiting the specific mathematical form of the cross section at hand.

A visual representation of the above summary is also given in Fig. 5.17.

Any modeling method that could be derived from first principles and would be capable of taking into account the quantum mechanical effects arising from cancellations and interference of different contributions with comparable accuracy as a Matrix Element based event generator would be nothing short of a "holy grail".

The novel method of *Effective Lagrangian Morphing* introduced in Section 2.4.3 rises to this challenge for the case of modeling the dependence of physical cross sections on Wilson coefficients in an EFT scenario.

5.3 Effective Lagrangian Morphing

All measurements conducted thus far find an excellent agreement between the Standard Model predictions including the Higgs sector and the observed data, with the maybe notorious exception of finding a somewhat low $H \to bb$ coupling [53], that is, however, still within the realm of a plausible statistical fluctuation. However, even a significant deviation in signal strength in any one channel is in and of itself no receipe on how to construct a quantum field theory that describes the observed data better than the Standard Model. Instead, the type of excess or deficit or the change in shape that was observed needs to be analyzed carefully to understand its theoretical implications. For this purpose, different frameworks of interpretation exist, some of which are briefly introduced in Section 2.4.2. This section focuses on the EFT interpretation of possible deviations in the Higgs sector and provides a detailed discussion of the technique of *Effective Lagrangian Morphing*, introduced in Section 2.4.3.

Employing this technique for a tree level calculation of a $2 \to 2$ s-channel process like the production and decay of a Higgs boson at the LHC at leading order, the cross section prediction can be expressed as

$$\sigma(\vec{g}) = \vec{P}(\vec{g}) \cdot A \cdot \vec{\sigma}, \tag{5.3.1}$$

where $\vec{P}(\vec{g})$ is a vector of fourth-order polynomials in the couplings g, $\vec{\sigma}$ is a vector of precomputed cross sections at the input parameter points, and A is the inverted morphing matrix.

5.3.1 A simple example

The simplest possible case is a scenario where any BSM coupling would only affect the decay vertex of the Higgs boson. One example for this would be Higgs boson production with decay to a pair of oppositely charged muons, where one would be interested only in two possible contributions to this vertex. These are

- a Standard Model coupling g_{SM}, and
- a BSM coupling g_{BSM}.

In this case, the polynomials appearing in the calculation are only of second order, as the constant couplings of the production vertex drop out of the calculation. The Matrix Element computation takes the form

$$\mathcal{M} = g_{\mathrm{SM}}\mathcal{M}_{\mathrm{SM}} + g_{\mathrm{BSM}}\mathcal{M}_{\mathrm{BSM}} \tag{5.3.2}$$

$$|\mathcal{M}|^2 = g_{\mathrm{SM}}^2 \left|\mathcal{M}_{\mathrm{SM}}\right|^2 + g_{\mathrm{BSM}}^2 \left|\mathcal{M}_{\mathrm{BSM}}\right|^2 + 2g_{\mathrm{SM}}g_{\mathrm{BSM}}\mathfrak{R}(\mathcal{M}_{\mathrm{SM}}\mathcal{M}_{\mathrm{BSM}}^*)\,, \tag{5.3.3}$$

where the notation $\mathfrak{R}(\mathcal{M})$ denotes the real-valued component of the Matrix Element $\mathcal{M}$.

As Eq. 5.3.3 has three independent components proportional to g_{SM}^2, g_{BSM}^2 and $g_{\mathrm{SM}}g_{\mathrm{BSM}}$, respectively, the construction of the morphing function will need the same number of independent input samples. For the sake of simplicity, one can choose the values given in Table 5.9 and call the samples "SM", "BSM" and "Mix".

In this case, the morphing matrix and its inverse are given as

$$\begin{array}{c} \\ \text{SM} \\ \text{Mix} \\ \text{BSM} \end{array}\!\!\begin{array}{c} \begin{array}{ccc} g_{\mathrm{SM}}^2 & g_{\mathrm{SM}}g_{\mathrm{BSM}} & g_{\mathrm{BSM}}^2 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \xrightarrow{\text{inversion}} \begin{array}{c} \\ g_{\mathrm{SM}}^2 \\ g_{\mathrm{SM}}g_{\mathrm{BSM}} \\ g_{\mathrm{BSM}}^2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} \text{SM} & \text{Mix} & \text{BSM} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array}. \tag{5.3.4}$$

Table 5.9.: Coupling value assignments for three input samples called "SM", "BSM" and "Mix".

Sample	g_{SM}	g_{BSM}
SM	1	0
Mix	1	1
BSM	0	1

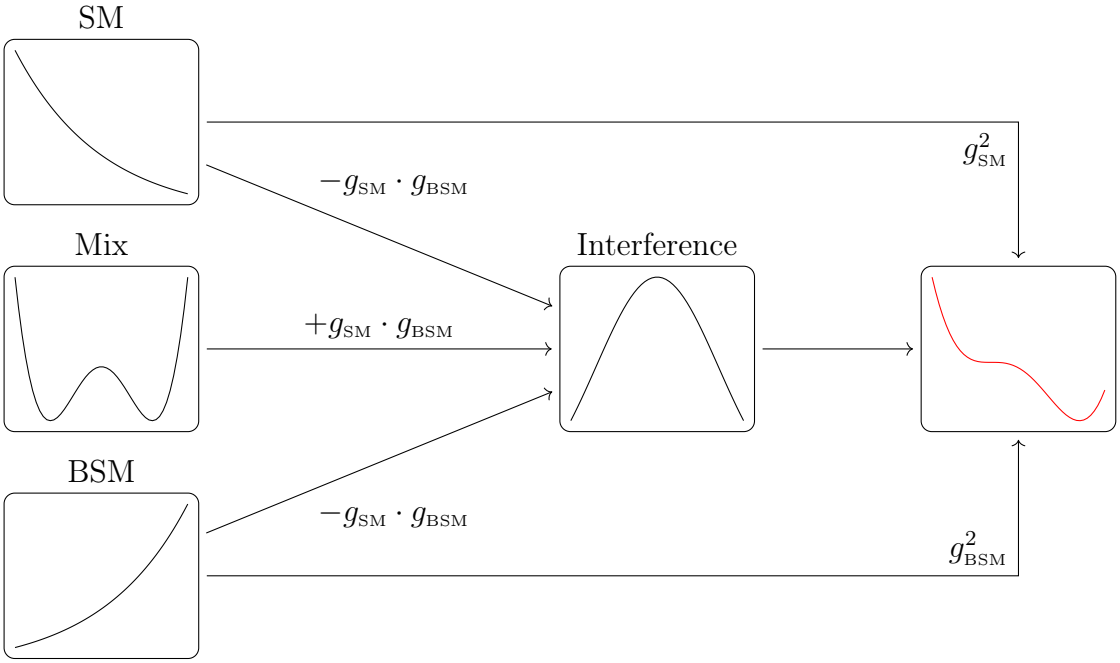


Figure 5.18.: Illustration of the simplest case of *Effective Lagrangian Morphing*.

From the inverted matrix, one can directly read off the coefficients of the morphing function. In this case, the morphing function is given as

$$\begin{aligned}\sigma(g_{\mathrm{SM}}, g_{\mathrm{BSM}}) = & \underbrace{\left(g^2_{\mathrm{SM}} - g_{\mathrm{SM}} \cdot g_{\mathrm{BSM}}\right)}_{w_{\mathrm{SM}}} \sigma_{\mathrm{SM}} \\ & + \underbrace{g_{\mathrm{SM}} \cdot g_{\mathrm{BSM}}}_{w_{\mathrm{Mix}}} \sigma_{\mathrm{Mix}} \\ & + \underbrace{\left(g^2_{\mathrm{BSM}} - g_{\mathrm{SM}} \cdot g_{\mathrm{BSM}}\right)}_{w_{\mathrm{BSM}}} \sigma_{\mathrm{BSM}},\end{aligned} \tag{5.3.5}$$

where the effective weights used for the samples in the linear combination have been named w_{SM}, w_{BSM} and w_{Mix}.

In this case, the structure of the morphing function can be understood intuitively. The two pure samples "SM" and "BSM" enter the calculation with the corresponding pure second-order polynomials as prefactors. The "Mix" sample is essentially a sum of the "SM" and the "BSM" sample, plus the interference term. The interference term can thus be isolated by subtracting the pure samples from the mixed sample. Afterwards, the interference term can be added to the morphing function with the mixed second order polynomial as a prefactor. Figure 5.18 illustrates this intuitive understanding.

5.3.2 Applicability

The method of Effective Lagrangian Morphing can be applied to interpolate arbitrarily between physics scenarios that can be continuously transformed into one another by adjusting a coupling-like parameter. While this is definitely true for any Wilson coefficients appearing in effective field theories, the technique is not limited to that case. Indeed, couplings of non-effective operators can be used just as well.

No assumption is made on the physics quantity subject to the interpolation other than that it is a linear combination of the individual terms arising from the pure and the interfering contributions to the process in question, with the coefficients being polynomials in the parameters of interest. Notably, this applies to any parton-level cross-section, total or differential, as well as branching fractions. However, under the assumption that the new physics effects do not alter the factorization, the hadronization, the parton shower or the interaction of the particles with the detector, it also trivially applies to predicted cross sections on the fully reconstructed information level. This same assumption is also required for Matrix Element Reweighting on reconstructed events to be a viable option for the production of BSM signal samples.

The only way in which Effective Lagrangian Morphing falls short of the capabilities of Matrix Element Reweighting is that Matrix Element Reweighting can of course also model non-coupling-like parameters, such as masses, widths and functional inputs such as variations in parton density functions. Effective Lagrangian Morphing will thus not be able to replace Matrix Element Reweighting completely, as the latter will remain an indispensable tool for the estimation of systematic uncertainties connected to the limited accuracy of the parton density functions, for non-EFT approaches to deriving limits on the masses of new particles, and to studies concerning the decay widths of particles. However, in so far as an EFT captures the parameters of the underlying theory in coupling-like parameters, Effective Lagrangian Morphing still allows to model their effects.

On the other hand however, Effective Lagrangian Morphing is computationally several orders of magnitude faster as compared to Matrix Element Reweighting. For example, performing Matrix Element Reweighting within the context of a likelihood fit for every step of the minimization is unimaginable, as it needs to be performed on an event-by-event basis and does not work on final distributions directly. This realization exhibits the true strengths of the method that it

- provides an interpolation method between precomputed predictions for total or differential cross sections and branching fractions, that it
- is fast and efficient enough to be performed in the context of a likelihood fit of the parameters, and that it

- accurately models the underlying physics without making any additional approximations.

It is important to stress at this point that the method of Effective Lagrangian Morphing does not introduce any additional uncertainties to the prediction. The resulting uncertainty estimate will be based purely on propagation of the uncertainties of the input values between which the interpolation is performed. The morphing matrix itself is derived purely as the inverse of a matrix of polynomials of arbitrarily chosen parameters and is thus immaculate, save for the numerical inaccuracy of the finite-precision arithmetic of computer algorithms. The calculation and reduction of uncertainties on the predictions obtained via Effective Lagrangian Morphing is, however, an intricate subject to which Section 5.3.6 is devoted.

5.3.3 Dimensionality

There is no conceptual limitation as to how many parameters the method of Effective Lagrangian Morphing can model simultaneously, as the addition of more parameters will just increase the size of the morphing matrix. The size of this square matrix will always need to be identical to the number of independent terms appearing in the Lagrangian, which then translates into the number of input samples required for the method.

For a general $2 \to 2$ s-channel process at leading order with n_p couplings appearing exclusively in the production vertex, n_d couplings appearing exclusively

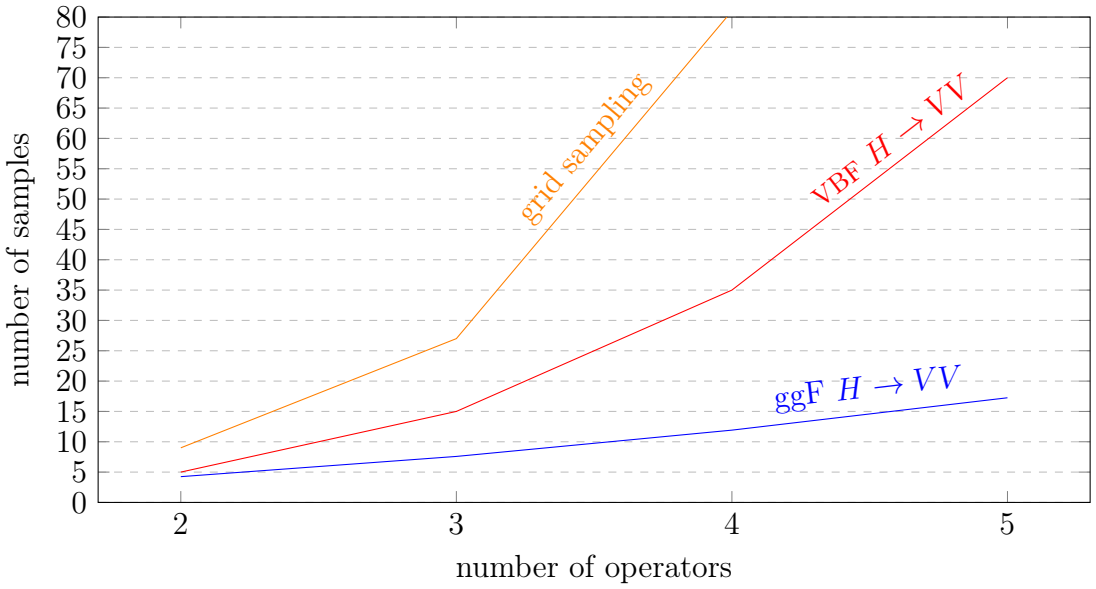


Figure 5.19.: Number of input samples required for Effective Lagrangian Morphing in comparison to a simple $n = 3$ grid sampling.

in the decay vertex, and n_s shared couplings appearing in both the production and the decay vertex, this number is

$$\begin{aligned} N = &\frac{n_p\,(n_p+1)}{2}\cdot\frac{n_d\,(n_d+1)}{2} + \binom{4+n_s-1}{4} \\ &+ \left(n_p\cdot n_s + \frac{n_s\,(n_s+1)}{2}\right)\cdot\frac{n_d\,(n_d+1)}{2} \\ &+ \left(n_d\cdot n_s + \frac{n_s\,(n_s+1)}{2}\right)\cdot\frac{n_p\,(n_p+1)}{2} \\ &+ \frac{n_s\,(n_s+1)}{2}\cdot n_p\cdot n_d + (n_p+n_d)\binom{3+n_s-1}{3}. \end{aligned} \tag{5.3.6}$$

This formula can be obtained by simple counting. In this representation, the individual contributions can be identified, where

- the first line counts the terms only containing production couplings, only containing decay couplings, or only containing shared couplings,
- the second line counts the terms which are of second order in decay-only couplings,
- the third line counts the terms which are of second order in production-only couplings,
- and the last line counts the terms which are only of first order in either production or decay or both.

Particularly relevant for the studies presented here are the cases of ggF and VBF Higgs boson production with a decay to two vector bosons in the Higgs effective Lagrangian presented in Section 2.4.2. In the particular cases of $H{\to}W^{\pm}W^{\mp *}$ and $H{\to}ZZ^{*}$, the parameters of primary interest are the Wilson coefficients coupling the Higgs field to the $W^{\pm}$ and Z fields. The corresponding reduced forms of the above formula for n free $H \to VV$ couplings are

$$N_{\text{ggF}} = \frac{n\,(n+1)}{2} \tag{5.3.7}$$

$$N_{\text{VBF}} = \binom{4+n-1}{4}, \tag{5.3.8}$$

respectively, as displayed in Fig. 5.19. Here, one coupling is assumed for the ggF production vertex, and all couplings are assumed shared between the two HVV vertices in the VBF case.

From this, it becomes immediately apparent that the number of required samples grows very quickly with increasing dimensionality of the space spanned by the morphing. This is especially true in the VBF case where the simultaneous increase

of the number of operators in both the production and the decay vertex accelerates the growth with respect to the ggF scenario, where only the number of operators in the decay vertex is effectively increasing. However, as can be seen from that same graphic, the number of samples is still much smaller than what would be required for even a minimalistic sampling with one Standard Model sample and samples for all $\pm 1\sigma$ combinations for every Wilson coefficient, resulting in a simple grid with three grid lines along every direction.

5.3.4 Applications

In order to prove experimentally the validity of the method and its implementation and to showcase its capabilities, several studies have been performed. They all share a common strategy:

1. A physical mode of Higgs boson production and decay is chosen.
2. A number of Monte Carlo samples with coupling parameters suitable for use as a basis for Effective Lagrangian Morphing is selected.
3. Additional samples are used as validation benchmarks.
4. Effective Lagrangian Morphing is used to obtain predictions at ("morph to") these validation benchmarks.
5. The resulting predictions of differential cross sections are studied for significant discrepancies.

The studies presented here have been previously published [20] and encompass investigations of

- truth level ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ events at $\sqrt{s} = 13\,\text{TeV}$,
- truth level VBF $H \to W^\pm W^{\mp *} \to \ell^-\bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ events at $\sqrt{s} = 13\,\text{TeV}$, and
- simulated ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ events at $\sqrt{s} = 8\,\text{TeV}$,

with varying numbers of parameters coupling the Higgs field to the respective weak bosonic field. For the first two studies, dedicated samples have been produced, whereas the study on simulated samples uses the samples used for earlier publications of the $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ analysis [17].

Various distributions of differential cross sections are shown after applying the respective object and event selection. These selections are briefly introduced in the text, but are ultimately of little interest for the studies performed here, as they only affect the shape or normalization of the distributions, not the agreement between the distributions predicted by the validation sample and the morphing. As shape and normalization effects affecting all distributions simultaneously cancel and do not affect the validation, no conscious effort has been made to normalize or reweight the distributions to any calculation other than the predictions of cross section and acceptance obtained directly by the Monte Carlo generator.

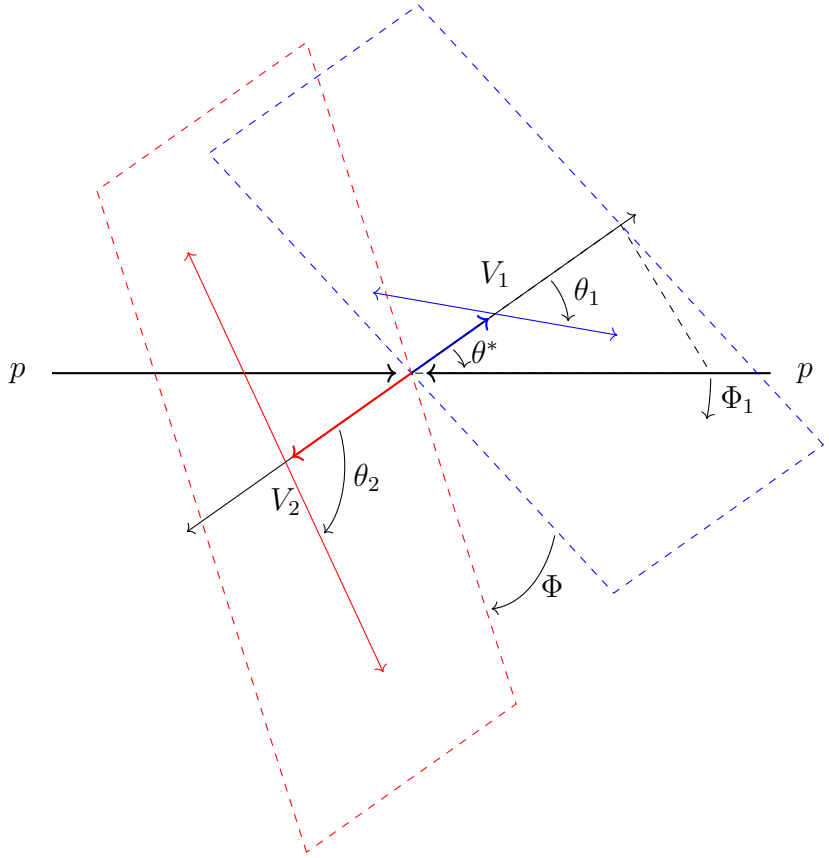


Figure 5.20.: Angles in the system of a Higgs boson decaying to four fermions.

The $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ studies employ the angular variables Φ, Φ_1, $\cos\theta^\star$ and $\cos\theta_{1/2}$ that can be measured in events with four charged leptons in the final state. They are commonly used in the analysis of $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ events and are known to have discriminatory power between different BSM scenarios, including a CP-odd Higgs boson [180]. The definitions of these variables are shown in Fig. 5.20.

For the VBF studies, the azimuthal angle between the tagging jets $\Delta\phi_{jj}$ is used, which also provides discriminatory power between scenarios with CP-odd and CP-even Higgs bosons [181].

The studies are based on the Higgs characterization model [47] introduced in Section 2.4.2, using a reduced parameter set that consists of

- the Standard Model coupling $g_{\text{SM}} = \kappa_{\text{SM}} \cos\alpha$,
- one additional non-SM CP-even coupling $g_{HVV} = \frac{1}{\Lambda}\kappa_{HVV} \cos\alpha$,
- and one CP-odd coupling $g_{AVV} = \frac{1}{\Lambda}\kappa_{AVV} \sin\alpha$,

where the cutoff scale has been chosen as $\Lambda = 1\,\text{TeV}$ for all studies presented here.

Monte Carlo samples at $\sqrt{s} = 13\,\text{TeV}$

Samples of events incorporating gluon fusion and vector boson fusion Higgs boson production with subsequent decays into W and Z bosons decaying into purely leptonic final states have been produced for several EFT scenarios. For each study, exactly one Standard Model sample has been generated with the Standard Model coupling set to unity and all BSM Wilson coefficients set to zero. All other samples have the Standard Model coupling set to unity as well, but the non-SM couplings chosen such that a pure non-SM sample with that coupling would have the same cross section as the pure SM sample.

The samples have been generated for proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with MADGRAPH5_aMC@NLO 2.2.2.p1 [130] using the Higgs characterization model [47], introduced in Section 2.4.2.

The base and validation samples were generated with 50.000 events each. The gluon fusion samples were generated using a version of that model including an effective vertex for the coupling of gluons to the Higgs boson, effectively contracting the fermion loop. All samples have been generated using LO precision in Standard Model and BSM electroweak effects, but using NLO precision for QCD effects. The leading-order PDF set NNPDF23LO [131] was used for all samples. The parton shower and hadronization were produced with PYTHIA8.186 [125] with the A14 tune [134]. No simulation of detector effects has been performed, and no pile-up has been taken into account. Jet candidates are reconstructed using the anti-k_t algorithm [103, 182] with a distance parameter of $R = 0.4$.

Validation in ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$

This validation study is minimal in the sense that it reflects the setup chosen for the simple example discussed in Section 5.3.1. It incorporates only two parameters of interest, one Standard Model coefficient κ_{SM} and one non-SM coefficient κ_{AZZ}. The third coupling $g_{Hgg} = \kappa_{Hgg} \cos\alpha$ employing the coefficient κ_{Hgg} only appears in the production vertex and has the same value for all samples. It thus drops out of the calculation, effectively resulting in a morphing matrix that is only second order in the couplings.

The precise values of the parameters used for the samples in this study are listed in Table 5.10. The samples have been chosen such that there is exactly one Standard Model sample without any non-SM contribution, exactly one pure BSM sample with no Standard Model contribution at all, and one mixed sample that combines the value of the SM coupling from the SM sample with the value of the CP-odd coupling from the BSM sample. Save for the precise numerical values, the situation is thus identical to the one discussed in Section 5.3.1.

The distributions shown here are derived from an event selection inspired by the one applied in the nominal $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ analysis [22], as presented in

Section 5.1.1. Electrons (muons) are required to pass $E_{\mathrm{T}} > 7\,\mathrm{GeV}$ ($p_{\mathrm{T}} > 6\,\mathrm{GeV}$) and to be measured in the pseudorapidity range $|\eta| < 2.47$ ($|\eta| < 2.7$). Further, the dilepton invariant mass of the leading (subleading) lepton pair is required to lie within $50\,\mathrm{GeV} < m_{\ell\ell} < 106\,\mathrm{GeV}$ ($c < m_{\ell\ell} < 115\,\mathrm{GeV}$, where c interpolates between 12 GeV for $m_{4\ell} < 140\,\mathrm{GeV}$ and 50 GeV for $m_{4\ell} > 190\,\mathrm{GeV}$), and $m_{\ell\ell} > 5\,\mathrm{GeV}$ is required for SFOC lepton pairs, and $\Delta R_{\ell\ell} > 0.1\,(0.2)$ is required for all same (opposite) flavour lepton pairs.

Figure 5.21 shows the comparison between the morphing output and the predicted distributions of the two validation benchmarks as listed in Table 5.10. Distributions in $\cos\theta_1$ and Φ, which are sensitive to the presence of non-SM couplings, are shown for validation.

The distributions show the angle θ_1 between the on-shell Z boson and its negatively charged lepton, and the angle ϕ between the decay planes of the two Z bosons, all calculated in the rest frame of the Higgs boson. Geometric representations of these angles can be seen in Fig. 5.20.

In both cases, the prediction of the morphing and the distribution generated from a statistically independent Monte Carlo sample are in excellent agreement.

Validation in VBF $H\to W^{\pm}W^{\mp *}\to \ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$

For this study, two non-SM parameters, the CP-even coefficient κ_{HWW} and the CP-odd coefficient κ_{AWW} are taken into account in addition to the Standard Model coefficient κ_{SM}. When assuming equality of the $H \to WW$ and $H \to ZZ$ couplings, all three parameters appear in both vertices of the VBF $H \to VV$ process. Employing Eq. 5.3.8 with $n_p = 0$, $n_d = 0$, and $n_s = 3$, the required number of samples for this case is 15.

The parameter choices for all samples generated for this study, including two validation samples dubbed `v0` and `v1`, can be seen in Fig. 5.22. The value of $\cos\alpha = 1/\sqrt{2}$ has been chosen for all samples uniformly, as the parameter is redundant. The parameters for all base samples have been chosen arbitrarily within ranges such that a pure BSM sample at the limit would have the same cross section as the Standard Model sample. A conscious attempt was made to distribute the

Table 5.10.: Overview of the $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$ samples used in Section 5.3.4.

		κ_{SM}	κ_{Azz}	κ_{Hgg}	$\cos\alpha$
Input Sample 0	"SM"	1	0	1	1
Input Sample 1	"BSM"	0	13.938	$\sqrt{2}$	$1/\sqrt{2}$
Input Sample 2	"Mix"	$\sqrt{2}$	13.938	$\sqrt{2}$	$1/\sqrt{2}$
Validation Sample 1		1	0.25	$\sqrt{2}$	$1/\sqrt{2}$
Validation Sample 2		$\sqrt{2}$	-2	$\sqrt{2}$	$1/\sqrt{2}$

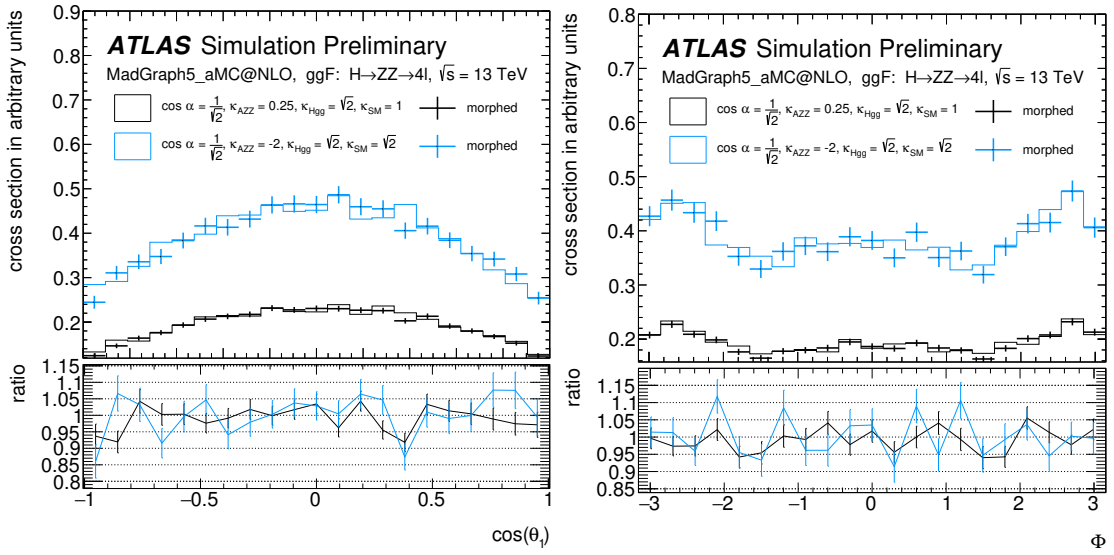


Figure 5.21.: Distributions of $\cos(\theta_1)$, where θ_1 is the angle between the on-shell Z boson and its negatively charged lepton (left) and the angle ϕ between the decay planes of the two Z bosons (right). The validation benchmarks (solid) as well as predictions calculated via morphing (dashed) are shown. The ratios between the morphing output and the validation distributions are shown in the bottom panels [20].

base samples evenly within the available parameter space. The parameters of the validation samples have been chosen randomly within the available ranges, employing a pseudo-random number generator.

On these samples, a loose event selection has been applied, requiring two selected leptons and two selected jets on truth level. For muons or electrons to be selected, requirements of $|\eta| < 2.7$ and $p_{\mathrm{T}} > 6\,\mathrm{GeV}$ or $|\eta| < 2.47$ and $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ are imposed, respectively. Jets are selected above $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ in the range of $|\eta| < 2.4$ and above $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ in the range of $|\eta| < 4.5$. This object selection is loosely based on the one imposed in the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ analysis presented in Chapter 4.

The cross section predictions as functions of the two BSM coefficients, the morphing technique and their associated relative uncertainties are shown in Fig. 5.23. It is worthwhile to note that these are no approximations, but exact predictions with uncertainties relying solely on the finite statistical power of the Monte Carlo generated input samples.

As expected, the physical cross sections rise as the production via the additional BSM couplings becomes more viable at larger values of the corresponding couplings. The uncertainties on the prediction are very small in the range of $2-3\%$ within large fractions of the considered parameter space. And while the relative uncertainty rises dramatically outside the parameter space covered by the input samples, the

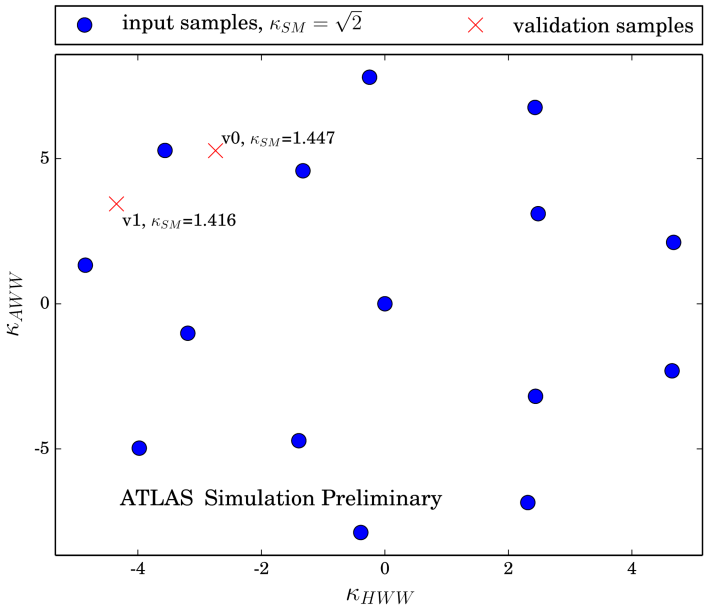


Figure 5.22.: Overview of all samples used in Section 5.3.4. The SM coupling is set to $g_{\mathrm{SM}} = \kappa_{\mathrm{SM}} \cos\alpha = 1$ for all input samples, with $\cos\alpha = 1/\sqrt{2}$ and $\kappa_{\mathrm{SM}} = \sqrt{2}$ [20].

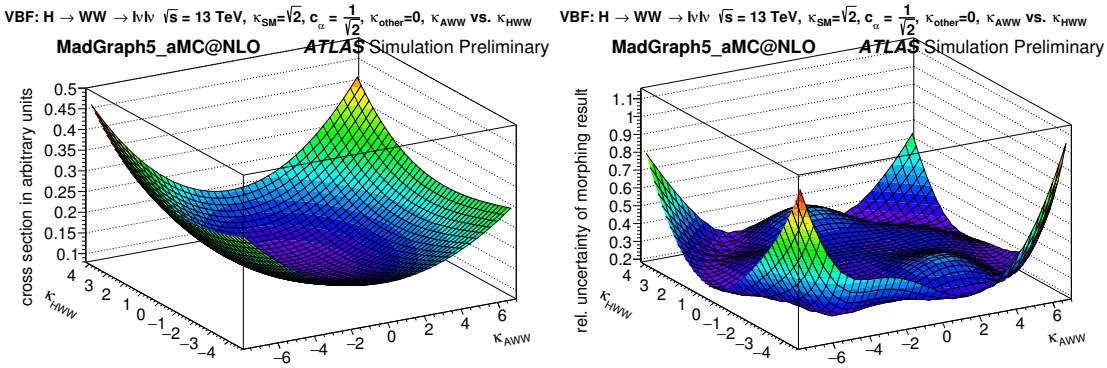


Figure 5.23.: Cross sections (left) and associated relative uncertainties (right) for the VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ process as a function of the non-SM CP-even and CP-odd coupling parameters κ_{HWW} and κ_{AWW} as obtained by Effective Lagrangian Morphing [20].

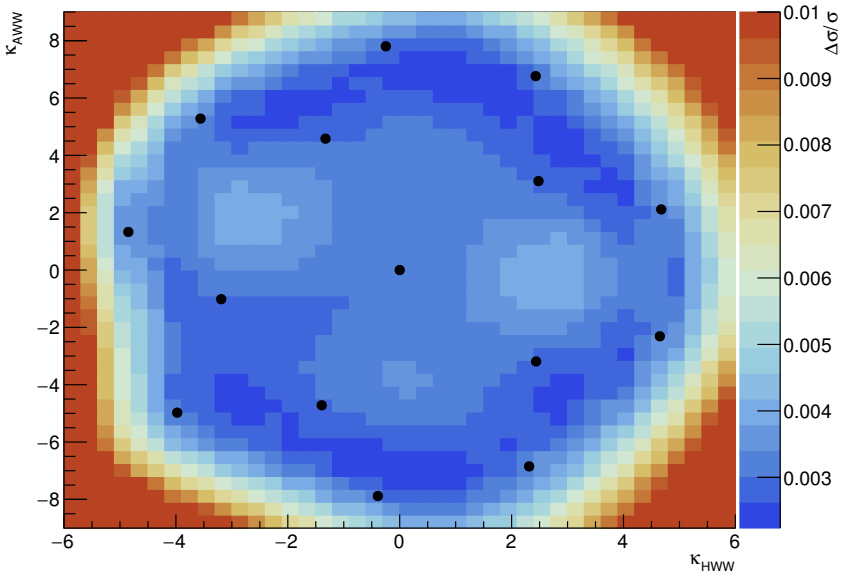


Figure 5.24.: Relative uncertainty on the morphing prediction in colour-code. Blue colours (inner) correspond to smaller relative uncertainties, red colours (outer) correspond to larger relative uncertainties. The positions of the base samples are indicated with black markers.

local maxima in the relative uncertainty do not coincide with the distribution of the input samples in any obvious way. This is more clearly illustrated in Fig. 5.24, which shows the same distribution in a purely colour-coded display, with the location of the base samples added as black points. This interesting observation is further studied in Section 5.3.6.

The kinematic observable used in this study is the azimuthal angle between the two VBF tagging jets $\Delta\phi_{jj}$. All input distributions from the morphing and the validation samples are shown in Fig. 5.25.

Table 5.11.: Correlation matrix from the fit to the `v0` (left) and `v1` (right) samples as defined in Fig. 5.22, corresponding to the fits shown in Fig. 5.26.

	κ_{SM}	κ_{HWW}	κ_{AWW}
κ_{SM}	1.00	0.20	-0.95
κ_{HWW}	0.20	1.00	0.09
κ_{AWW}	-0.95	0.09	1.00

	κ_{SM}	κ_{HWW}	κ_{AWW}
κ_{SM}	1.00	0.49	-0.93
κ_{HWW}	0.49	1.00	-0.16
κ_{AWW}	-0.93	-0.16	1.00

An additional type of closure check is performed by building a likelihood fit with the coefficients as free parameters, and fitting to either any base or validation sample as pseudo-data. The fit is generally able to extract the nominal values of

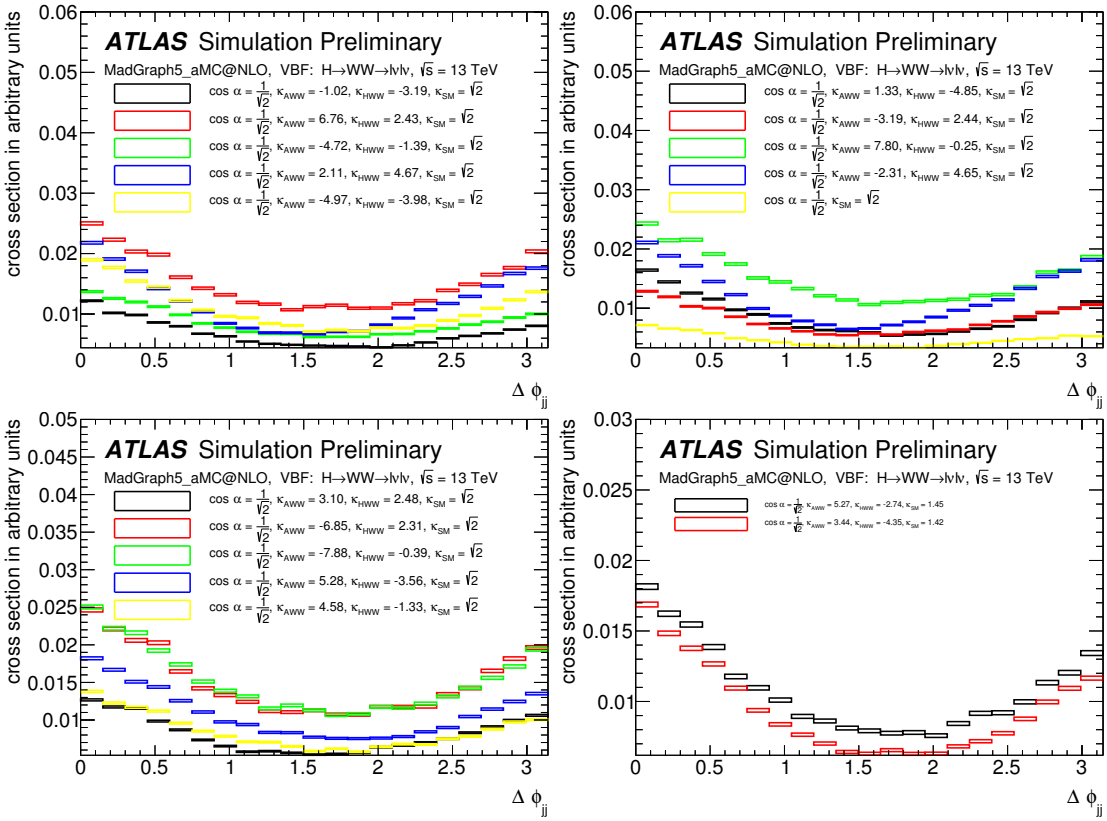


Figure 5.25.: Distributions of the azimuthal angle between the two VBF tagging jets $\Delta\phi_{jj}$ for all base and validation samples used in Section 5.3.4 [20].

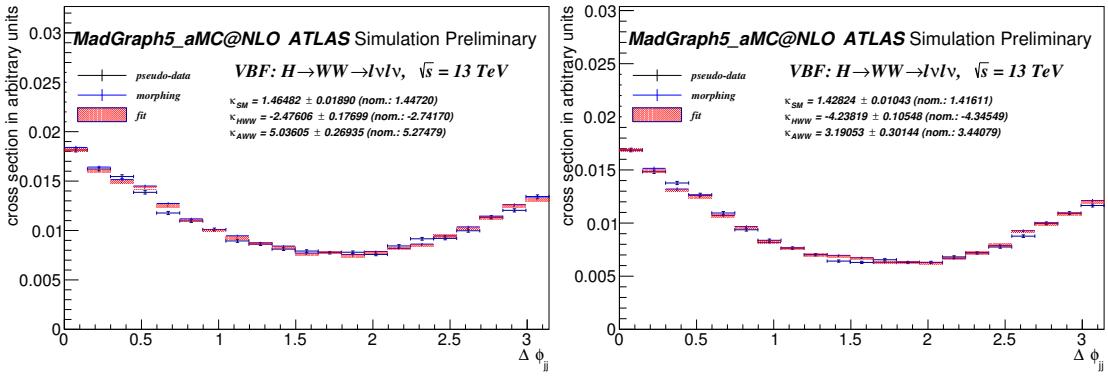


Figure 5.26.: Distributions of the azimuthal angle between the two VBF tagging jets $\Delta\phi_{jj}$ as obtained by morphing to the nominal parameters of the validation samples `v0` (left) and `v1` (right), as obtained from the best-fit-values obtained by a likelihood fit to the validation samples, and from the validation samples themselves [20].

the coefficients from the pseudo-data with high accuracy even for the statistically independent validation samples, as shown in Fig. 5.26.

The post-fit parameter correlations obtained from the fit to the validation samples are shown in Table 5.11. There is a strong correlation observed between the SM coefficient κ_{SM} and the CP-odd coefficient κ_{AWW}, as they both have a large effect on the total cross section.

5.3.5 Validation on reconstructed samples

In order to provide a convincing experimental proof that the technique of Effective Lagrangian Morphing works not only on truth-level predictions, but just as well on distributions of simulated and reconstructed events, a subset of simulated signal samples used in the ATLAS tensor structure analysis of $H{\to}ZZ^*{\to}\ell^+\ell^-\ell'^+\ell'^-$ events at $\sqrt{s} = 8\,\text{TeV}$ [17] has been used to perform a validation study similar to the ones presented above.

This study employs the non-SM CP-even coupling $\widetilde{\kappa}_{HZZ}$ and the CP-odd coefficient $\widetilde{\kappa}_{AZZ}$, defined as

$$\widetilde{\kappa}_{HZZ} = \frac{1}{4}\cdot\frac{v}{\Lambda}\cdot\kappa_{HZZ} \quad \text{and} \quad \widetilde{\kappa}_{AZZ} = \frac{1}{4}\cdot\frac{v}{\Lambda}\cdot\kappa_{AZZ}, \tag{5.3.9}$$

where v is the Higgs field vacuum expectation value. Only the dominant ggF Higgs boson production process is considered.

The first observable used in this study is the angle Φ between the decay planes of the two intermediate Z bosons. The second is the transformed optimal observable $\text{TO}_2(\widetilde{\kappa}_{AZZ}, \alpha)$, which has been introduced as one of the final discriminating variables of the $H{\to}ZZ^*{\to}\ell^+\ell^-\ell'^+\ell'^-$ tensor structure measurement [17]. The optimal observable $\text{O}(\widetilde{\kappa}_{AZZ}, \alpha)$ is defined as the ratio

$$\text{O}(\widetilde{\kappa}_{AZZ}, \alpha) = \frac{\left|\mathcal{M}(\widetilde{\kappa}_{AVV} \neq 0, \kappa_{\text{SM}} = \widetilde{\kappa}_{HVV} = 0, \alpha = \pi/2)\right|^2}{\left|\mathcal{M}(\widetilde{\kappa}_{\text{SM}} \neq 0, \kappa_{AVV} = \widetilde{\kappa}_{HVV} = 0, \alpha = \pi/2)\right|^2}, \tag{5.3.10}$$

where the Matrix Element $\mathcal{M}$ is computed *in situ* given the kinematic properties of the final state leptons. The transformation is chosen such that the observable follows a standard Gaussian normal distribution in the Standard Model scenario. This observable is very sensitive to the amplitude of CP-odd admixtures [183] in the HZZ coupling structure.

The Standard Model Higgs boson production via gluon fusion at the mass $m_H = 125.5\,\text{GeV}$ is simulated using the Powheg-Box [129] generator. For the non-SM signals, the decays of the generated Higgs bosons are simulated at leading order using the JHU Monte Carlo generator [180,184]. The parameters of the samples

Table 5.12.: Parameter configuration of the base and validation samples used in Section 5.3.5.

	$\tilde{\kappa}_{\mathrm{SM}}$	$\tilde{\kappa}_{HZZ}$	$\tilde{\kappa}_{AZZ}$
Input sample 1	1	0	0
Input sample 2	0	1	0
Input sample 3	0	0	1
Input sample 4	1	1	0
Input sample 5	0	1	1
Input sample 6	1	0	1
Validation sample 1 (Φ)	1	-2	0
Validation sample 2 (Φ)	1	-1.25	0
Validation sample 3 ($\mathrm{TO}_2(\tilde{\kappa}_{AZZ}, \alpha)$)	1	0	5
Validation sample 4 ($\mathrm{TO}_2(\tilde{\kappa}_{AZZ}, \alpha)$)	1	0	3.25

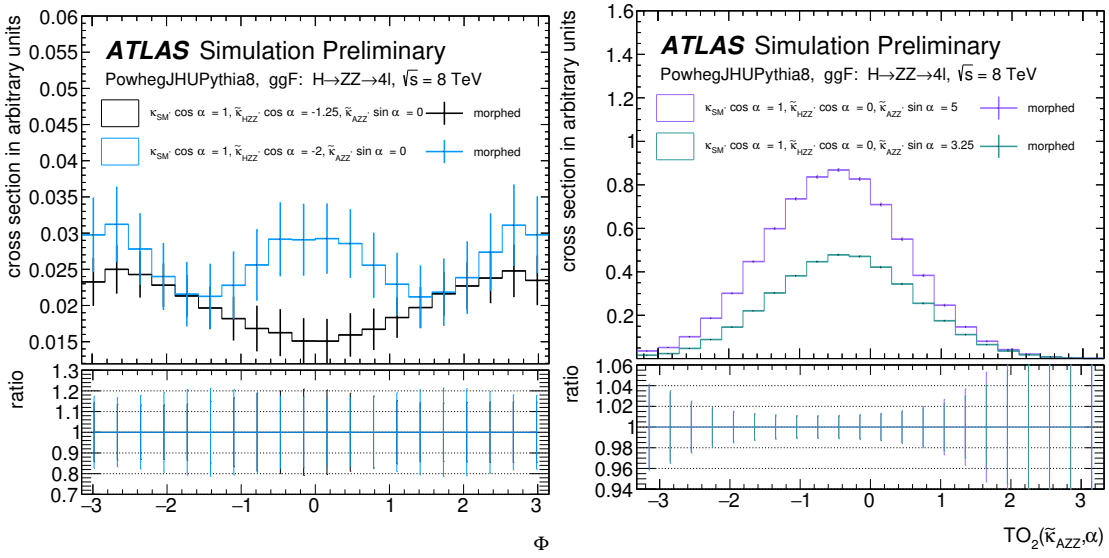


Figure 5.27.: Distributions of ggF $H \to ZZ^* \to 4\ell$ events at $\sqrt{s} = 8\,\mathrm{TeV}$ [17], comparing the shapes of physical observables Φ and $\mathrm{TO}_2(\tilde{\kappa}_{AZZ}, \alpha)$ for two validation benchmarks with the prediction obtained via Effective Lagrangian Morphing. The perfect agreement is due to the fact that all base and validation samples used here have been derived from the same original sample via Matrix Element Reweighting [20].

selected for this study are shown in Table 5.12. It is worthwhile to note at this point that the input samples are not statistically independent, but have been derived from the same base sample using the technique of Matrix Element Reweighting. It thus becomes apparent that the two methods of Effective Lagrangian Morphing and Matrix Element Reweighting are in fact not competing, but rather complementary in the sense that Matrix Element Reweighting can be employed to generate the input samples required by the technique of Effective Lagrangian Morphing, as has been done in this case.

Figure 5.27 shows distributions of the two observables in the 4ℓ final state for the four different validation benchmarks. The agreement between the predictions of the morphing and of the validation sample is perfect, as the input samples are not statistically independent. Clearly, the technique of Effective Lagrangian Morphing is capable of perfectly reproducing distributions of complicated physical observables like the ones studied here even at the level of reconstructed and simulated events.

5.3.6 Uncertainty estimation

As already alluded to, the uncertainty treatment in the context of Effective Lagrangian Morphing can have surprising and interesting effects – and while many aspects are still subject to investigative research, this section attempts to provide a comprehensive overview of the current state of research.

With the method itself being completely deterministic, the only choice to be made is the set of parameters of the base samples. Different choices of base samples can have sizable effects on the uncertainty of the prediction, as demonstrated by Fig. 5.28, comparing a "good" and a "bad" choice of base samples for the morphing of the Φ distribution in ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ events, as introduced in Section 5.3.4. The parameters of the samples are shown in Fig. 5.29. Within this section, the uncertainty estimate of the prediction obtained via Effective Lagrangian Morphing is formalized and quantified.

Formalization

The morphing function m is a mapping from n input parameters g_i (with the index i running trough $1, \ldots, n$.), typically couplings in the Lagrangian, to a single number – a cross section or branching ratio, which is denoted as X here. This is technically not true in cases where the prediction is a differential cross section, but the generalization to a list of values is trivial and will not be considered here for the sake of notational simplicity. Thus,

$$m : \quad \mathbb{R}^n \to \mathbb{R} \quad : \quad \vec{g} \to X. \tag{5.3.11}$$

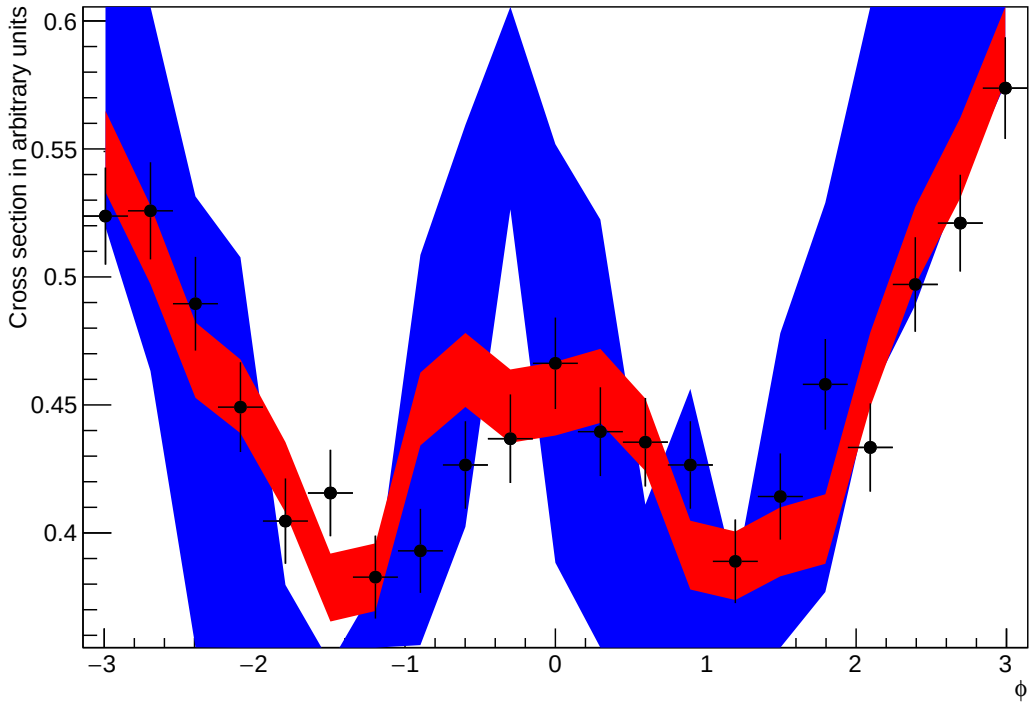


Figure 5.28.: Morphing of ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ kinematic distribution in the observable Φ, as introduced in Section 5.3.4, with two different sets of base samples (red and blue) to the parameters of a statistically independent benchmark (data points). The band widths indicate the propagated Monte Carlo statistical uncertainties. The parameter choices for all samples used here are shown in Fig. 5.29.

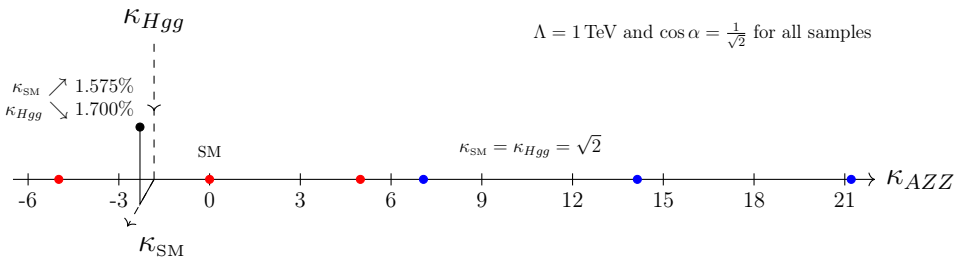


Figure 5.29.: Base and validation samples for the morphing displayed in Fig. 5.28. The samples have been generated with MADGRAPH5_aMC@NLO2.2.2.p1 [130] in the Higgs Characterization Model introduced in Section 2.4.2 using the leading-order PDF set NNPDF23LO [131] and showered with PYTHIA8.186 [125] with the A14 tune [134]. The blue and red points indicate the locations of two different sets of base samples, and the black dot indicates the validation benchmark.

The morphing function m itself is fully determined by the morphing matrix M, a square matrix of size $N \times N$, and the predictions ξ_j at the input samples. Here, the index j is meant to run through $1, \ldots, N$.

The mapping $d : n \to N$ is not entirely straight-forward and will depend on the physical process considered, but has been laid out for the case of a $2 \to 2$ process in Eq. 5.3.6. Knowing the size N of the morphing matrix then allows to freely choose the n coupling parameters for the N different base sample, the sole constraint being that the morphing matrix must remain invertible. Then, some Monte Carlo phase space integrator $\mathcal{F}$ can be used to predict the N values of ξ_j at the parameter configurations $\vec{g}_j$. The construction $\star$ of the morphing function itself thus proceeds using

$$d: \quad \mathbb{N} \to \mathbb{N} \quad : \quad n \to N \tag{5.3.12}$$

$$\mathcal{F}: \quad \mathbb{R}^n \to \mathbb{R} \quad : \quad \vec{g} \to \xi \tag{5.3.13}$$

$$\star: \quad \mathbb{R}^{n\times N} \to [\mathbb{R}^n \to \mathbb{R}] \quad : \quad \{\vec{g}\} \to m = [M, \{\xi\}] \tag{5.3.14}$$

There are only two principle sources of uncertainty in this process:

1. The Monte Carlo integrator $\mathcal{F}$ will only produce a limited set of k data points for every one of the N configurations. The result will be an uncertainty $\Delta\xi$ on the predictions of the input values ξ. In general, the relative uncertainty will be proportional to the square root of the number of events produced, that is, $\Delta\xi/\xi \sim \sqrt{k}$
2. During the construction $\star$ of the morphing function m, the matrix M is obtained via a matrix inversion. Using finite precision arithmetic, this introduces numeric inaccuracies to the calculation.

Matrix condition and numerics

While it is in theory always possible to find the exact inverse of an invertible matrix, this inverse is not necessarily representable using finite precision. Considering, for example, the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \epsilon & \epsilon^2 \\ 0 & 0 & \epsilon^2 \end{pmatrix} \quad \text{with the inverse} \quad A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{\epsilon} & \frac{1}{\epsilon} & -\frac{1}{\epsilon} \\ 0 & 0 & \frac{1}{\epsilon^2} \end{pmatrix}, \tag{5.3.15}$$

the matrix A is not invertible for $\epsilon \to 0$. If, however, the computation only takes into account the first k decimal digits, then the matrix A will only be invertible if $\epsilon > 10^{-k}$. Of course, modern day computers use *floating point* arithmetic, that is, the available precision will affect the *number of significant digits*, not the total number of digits. For example, the commonly used `double` precision will produce

results accurate to approximately 15-17 significant digits in a simple floating point operation, with the maximum exponent representable being 308. However, matrix inversions are no simple floating point operations.

The most commonly used algorithm for matrix inversion is the well-known Gaussian elimination procedure. For a matrix of size N, this technique encompasses approximately $^2/_3N^3$ floating point operations. It is thus not entirely straightforward to understand how strongly finite numeric precision will affect the result of a matrix inversion.

Some modern computer algebra systems support analytic computations. However, the analytic computation of the inverse of a matrix encompasses a very peculiar detail, which can best be seen in the simple $N = 2$ case of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{5.3.16}$$

where the denominator of the fraction is simply the determinant of the matrix. Considering the seemingly innocuous case of

$$\begin{pmatrix} 1 & \epsilon \\ \epsilon & 1 \end{pmatrix}^{-1} = \frac{1}{1 - \epsilon^2} \begin{pmatrix} 1 & -\epsilon \\ -\epsilon & 1 \end{pmatrix}, \tag{5.3.17}$$

the determinant is no longer computable with k significant digits of precision if $\epsilon < 10^{\sqrt{k}}$. Hence, even analytic computations are prone to numeric imprecisions. As a matter of fact, this type of computation is often less precise than Gaussian elimination, as specifically the matrix determinant is very vulnerable to this type of problem.

One commonly used measure for the numerical problems arising from the inversion of a given matrix is the *condition number*

$$C_X(A)_X = \left\|A^{-1}\right\|_X \cdot \|A\|_X , \tag{5.3.18}$$

where $\|A\|_X$ denotes matrix norm. One common choice is the L_2 norm

$$\|A\|_{L_2} = \sqrt{\sum_{i,j} A_{i,j}^2}. \tag{5.3.19}$$

Another common choice is the operator norm

$$\|A\|_{\text{op.}} = \inf \{\lambda \geq 0 : Ax = \lambda x \forall x\} , \tag{5.3.20}$$

that is, the modulus of the smallest eigenvalue $\lambda_{\min}$ of A. However, if λ is an eigenvalue of A and A is invertible, then λ^{-1} is an eigenvalue of A^{-1}. Hence,

the inverse of the largest eigenvalue λ_{max} of A will be the smallest eigenvalue of A^{-1}. From this, it follows that

$$C_{\text{op.}} = \frac{|\lambda_{\text{max}}|}{|\lambda_{\text{min}}|}, \tag{5.3.21}$$

that is, the condition number is the ratio of the largest and the smallest eigenvalue.

While determining the eigenvalues is often computationally expensive for large matrices, upper and lower bounds for the largest and smallest eigenvalues can often be efficiently determined by use of the Gershgorin Circle Theorem [185]. This theorem states that each eigenvalue λ of a real-valued matrix A is contained within an interval of $[A_{ii} - R_i, A_{ii} + R_i]$, where i is some column or row of the matrix, A_{ii} is the corresponding diagonal entry, and R_i is the sum of absolute values of all non-diagonal entries in that row or column. A mathematical formulation and brief proof of this is presented in Appendix B.1.

It can be shown [186] that for a matrix of condition $c \approx 10^k$, one can expect to loose at least k significant digits of precision. This proof is briefly sketched in Appendix B.2.

The matrix from Eq. 5.3.15 has the eigenvalues $\{1, \epsilon, \epsilon^2\}$, such that the condition number is ϵ^{-2} for $\epsilon \leq 1$. Incidentally, the matrix is identical to the simple example from Section 5.3.1 for $\epsilon = 1$, resulting in a condition number of $c = 1$, as well as to the matrix used in Section 5.3.4 with $\epsilon \approx 10^{-2}$, resulting in a condition number of $c = 10^4$. The condition number of the morphing matrix is $c < 10^9$ for all studies presented here, resulting in a remaining numerical accuracy of approximately 10^{-7} when using `double` precision. This is much smaller than the approximately $2 - 5\%$ uncertainty resulting from the finite Monte Carlo sample size.

However, the implementation used to produce the studies presented here in fact uses the `boost uBLAS` and `boost multiprecision` libraries [187] to achieve around 100 significant digits accuracy, to which end the numerical effects are absolutely negligible.

Uncertainty propagation

On the surface, the propagation of the Monte Carlo statistical uncertainties on the result of the morphing is straight-forward. Given the morphing matrix M, the morphing function as introduced in Eq. 2.4.26 reads

$$X(\vec{g}) = \sum_j w_j(\vec{g})\xi_j \quad \text{with} \quad w_j(\vec{g}) = \sum_i P_i(\vec{g})A_{ij}, \tag{5.3.22}$$

where $A = M^{-1}$, $P_i(\vec{g})$ are the polynomials of the coupling parameters evaluated at the target point $\vec{g}$, and ξ_j are the predictions obtained at the base samples.

Assuming the matrix itself to be free of any significant uncertainty, as has been extensively justified in the previous section, only the uncertainties $\Delta\xi_j$ on the predictions ξ_j remain. There are, in principle, two interesting cases:

1. All ξ_j are fully uncorrelated. This is the case for statistically independent base samples.
2. All ξ_j are fully correlated. This is the case for base samples that have been generated with Matrix Element Reweighting from the same original samples.

The respective uncertainties ΔX on the output X compute as

$$\Delta X_0^2(\vec{g}) = \sum_j \left(w_j(\vec{g})\Delta\xi_j\right)^2 \tag{5.3.23}$$

$$\Delta X_1^2(\vec{g}) = \left(\sum_j w_j(\vec{g})\Delta\xi_j\right)^2, \tag{5.3.24}$$

where the subscripts 0 and 1 allude to the values of the corresponding correlation coefficients, ΔX_0 denoting the uncertainty for the uncorrelated case and ΔX_1 denoting the uncertainty for the fully correlated case.

The respective relative uncertainties compute as

$$\Delta X_{0,\mathrm{rel}}(\vec{g}) = \frac{\sqrt{\sum_j \left(w_j(\vec{g})\Delta\xi_j\right)^2}}{\sum_j w_j(\vec{g})\xi_j} \tag{5.3.25}$$

$$\Delta X_{1,\mathrm{rel}}(\vec{g}) = \frac{\sum_j w_j(\vec{g})\Delta\xi_j}{\sum_j w_j(\vec{g})\xi_j}. \tag{5.3.26}$$

Under the assumption that all input values ξ_j have the same relative uncertainty q, this yields

$$\Delta X_{0,\mathrm{rel}}(\vec{g}) = q \cdot \frac{\sqrt{\sum_j \left(w_j(\vec{g})\xi_j\right)^2}}{\sum_j w_j(\vec{g})\xi_j} \tag{5.3.27}$$

$$\Delta X_{1,\mathrm{rel}}(\vec{g}) = q. \tag{5.3.28}$$

It should be noted, however, that this assumption is rarely truly satisfied, especially for the correlated case where the samples have been obtained with Matrix Element Reweighting. The uncertainties of such reweighted samples can have extreme variations if the parameters chosen for the original sample are very different to the ones being reweighted to, as very few events from the tails of the respective kinematic distributions begin to dominate the sample.

Also for the uncorrelated case, the assumption is not always justified, as different kinematic acceptances of the signal region for the different scenarios can affect the statistical power, especially when looking at binned differential distributions with very few events per bin.

In Section 5.3.4, it was pointed out that the uncertainties on the prediction are not always minimal around the input samples, which was made visible by Fig. 5.24. Looking at the formulas for the uncorrelated case, it now becomes apparent why this happens. Considering a hypothetical case of equal weights for all base samples when morphing to a certain parameter configuration, that is, $w_j(\vec{g}) = c \; \forall j$, something interesting happens. In that case, Eq. 5.3.22 is then just the arithmetic average over all ξ_j save for a constant, and by definition has an uncertainty that is smaller than any individual contribution. Then, Eq. 5.3.27 becomes

$$\Delta X_{0,\mathrm{rel}} = q \cdot \frac{\sqrt{\sum_j \xi_j^2}}{\sum_j \xi_j}. \tag{5.3.29}$$

Interestingly, the same could be achieved for the predictions as well – it could be possible to achieve a configuration where all input values, be they cross sections for the sake of the argument, are equal, and hence

$$\Delta X_{0,\mathrm{rel}}(\vec{g}) = q \cdot \frac{\sqrt{\sum_j w_j(\vec{g})^2}}{\sum_j w_j(\vec{g})}. \tag{5.3.30}$$

The latter is certainly possible, given that the cross section is a polynomial of the couplings. Hence, the base samples would only need to be distributed along a contour line of the cross section.

It might even be possible to achieve both at the same time, yielding

$$\Delta X_{0,\mathrm{rel}}(\vec{g}) = \frac{q}{\sqrt{N}}. \tag{5.3.31}$$

While it is not naively clear whether this optimum can be achieved under any remotely realistic circumstances, it is still interesting to note that there are conditions under which the prediction obtained by Effective Lagrangian Morphing can be more precise than any of the input values, and that the accuracy is apparently optimal when each individual contribution $w_i \xi_i$ is of roughly the same size. Notably, negative contributions will deteriorate the quality significantly, causing large cancellations in the predicted values, but not their uncertainties, thus dramatically increasing the relative uncertainty.

In summary, the uncertainties on the predictions obtained via Effective Lagrangian Morphing

- can fluctuate heavily depending on the choice of base sample,
- are dependent on what benchmark point is considered, and
- rely on the assumptions of correlations between the individual base samples.

5.3.7 The curse of dimensionality

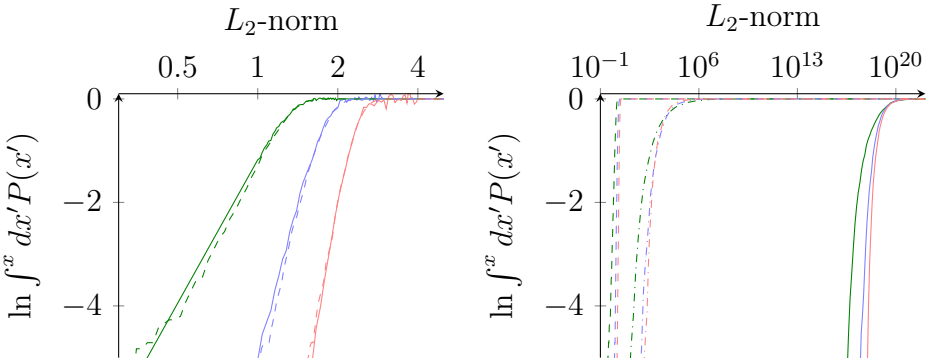
It is an intuitive truth that most optimization problems get harder with increasing dimension of the available configuration space, a problem generally known as the "curse of dimensionality". However, proving that this is also true for the case of finding a good basis for Effective Lagrangian Morphing is somewhat involved.

For low dimensionalities, it is easy to find a good basis. In fact, randomly distributing the base samples in the region of interest is a sufficient strategy for two or three free parameters. The chance of randomly generating a non-invertible matrix is virtually zero, and judging the quality of a basis independent of any specific benchmark scenario is difficult. Appendix B.3 details a couple of common misapprehensions that are capable of leading ones intuition astray on this issue.

One possibility to obtain a rough estimate of the general quality of a basis is the norm of the associated morphing matrix, as motivated in Section 5.3.6, where low values of the norm would correspond to generally "good" basis sets. This is certainly no accurate method of measuring the actual performance, and better measures will be developed in the following sections, but it can still provide an intuition on the issue that is difficult to come by otherwise. Fig. 5.30 shows how the L_2-norm, defined by Eq. 5.3.19, is distributed. Here, several scenarios and computational methods are compared.

The L_2 norm has the appealing feature that it can be interpreted geometrically: An $N \times N$ matrix with random entries A_{ij} chosen in the range $[-R, R]$ can be viewed as a point in an N^2-dimensional hypercube. The probability of this matrix having an L_2-norm smaller than some value x is then given as the volume of an N^2-dimensional hypersphere with radius x intersecting with that hypercube. However, unfortunately, no closed-form analytic solution exists for this problem. Some more details are given in Appendix B.4. Instead, analytic solutions can only be given piece-wise, such that the function has been evaluated by Monte Carlo integration in Fig. 5.30a.

Of course, the distributions can also be measured in randomly generated matrices, where different methods of randomization exist. Not every random matrix corresponds to a valid basis set. For example, it is in general not possible to find a basis that yields the unit matrix, as interference terms enforce the presence of non-diagonal matrix elements. The matrices used for Fig. 5.30b have been generated with three different methods:



(a) Solid (dashed) lines show the probability distribution predicted by geometric means (pseudo-random matrices).

(b) Probability distributions obtained from pseudo-random matrices without any constraints (dashed), with randomly generated couplings without scale suppression (dash-dotted) or with a physical EFT-scenario suppressed by a scale of $\Lambda = 1\,\text{TeV}$ (solid).

Figure 5.30.: Cumulative logarithmic probability distribution of the L_2-norm for randomly generated basis sets for a VBF $H \to VV$ scenario with 2 (green), 3 (blue) and 4 (red) shared coupling parameters.

1. Fully random matrices with uniformly distributed entries.

For the two other cases, morphing matrices have been constructed from random couplings such that their fourth powers follow a uniform distribution.

2. All coupling parameters have been chosen in the same range.
3. One coupling in each vertex has been chosen SM-like, values centered around unity, and all other couplings have been chosen BSM-like, values centered around zero and suppressed by a scale of $\Lambda = 1\,\text{TeV}$ to reflect a physical, EFT-like scenario.

It can be seen that the matrices chosen with the physical method have significantly higher L_2-norms than any more randomly generated matrices. Thus, it is reasonable to assume that the subset of matrices available as morphing matrices does not specifically cover the range of low norms. Secondly, it can be seen that the matrices with higher dimensionality have larger norms, motivating the assumption that finding a morphing basis becomes more difficult in higher dimensional scenarios.

5.3.8 Dimensionality reduction

Studies with a low dimensionality of the morphing are easier to conduct, not only because they require less base samples in total, but also because viable base

sets are easier to find. It is thus interesting to investigate which couplings can be safely neglected in a specific study, and to possibly find other methods of reducing the dimensionality.

Higgs boson production via vector boson fusion is a highly interesting subject of study, as it is sensitive to a multitude of non-SM contributions in the $VV \to H$ production vertex. In the Higgs Characterization Model, a total of 13 parameters listed in Table 5.13 contribute to this vertex. In order to decouple production and decay couplings completely, a decay mode of $H \to \mu\mu$ was chosen for this generator level study. A total of 91 Monte Carlo samples with 30.000 events each have been generated at $\sqrt{s} = 13\,\text{TeV}$ using MADGRAPH5_aMC@NLO 2.2.2.p1 [130] to form a complete basis set. In order to retain gauge invariance, the production includes VBF diagrams as well as the diagrams of the Higgs boson production in association with a hadronically decaying vector boson. Parton showers and hadronization have been calculated using PYTHIA8.186 [125] with the A14 tune [134]. The configurations of those samples were chosen with the following requirements:

Experimental sensitivity BSM couplings were chosen to reflect the current experimental limits. The Run 1 limit on the VBF Higgs boson production cross section is $\sigma_{\text{VBF}} \lessapprox 1.25 \cdot \sigma_{\text{VBF,SM}}$ [15]. The values of the coupling parameters chosen reproduce the SM cross section for pure BSM samples. The values are listed in Table 5.13.

Table 5.13.: Values of the coupling parameters to reproduce the Standard Model cross section at $\sqrt{s} = 13\,\text{TeV}$ in pure BSM scenarios. A combination of these parameters, where either the positive or the negative of the listed value is taken, is used to construct the 91 input samples.

κ_{SM}	1
$\kappa_{H\gamma\gamma}$	203.22
$\kappa_{A\gamma\gamma}$	408.62
$\kappa_{HZ\gamma}$	109.13
$\kappa_{AZ\gamma}$	986.88
κ_{HZZ}	5.75
κ_{AZZ}	6.96
κ_{HWW}	3.36
κ_{AWW}	3.92
$\kappa_{H\partial WR} = \mathfrak{R}(\kappa_{H\partial W})$	0.76
$\kappa_{H\partial WI} = \mathfrak{I}(\kappa_{H\partial W})$	0.84
$\kappa_{H\partial A}$	1.77
$\kappa_{H\partial Z}$	1.37

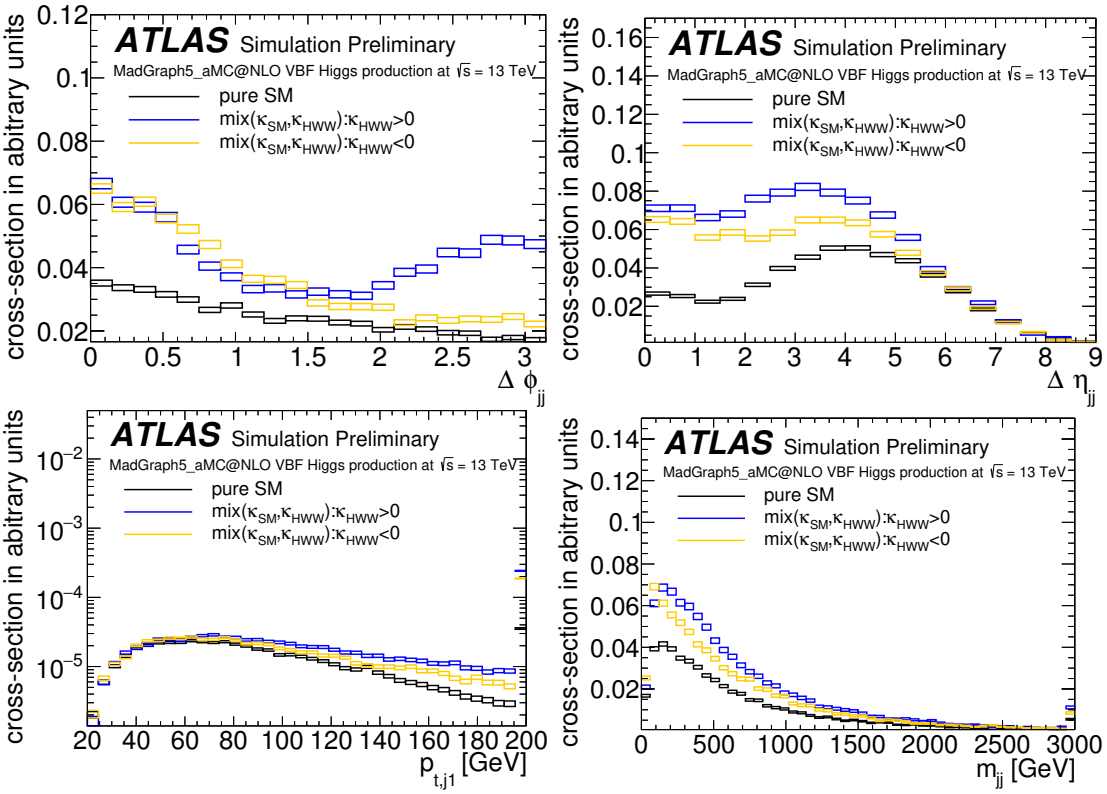


Figure 5.31.: VBF $H \to \mu\mu$ distributions for $\Delta\phi_{jj}$ (top left), m_{jj} (top right), $p_{\mathrm{T}}^{j_1}$ (bottom left), and $\Delta\eta_{jj}$ (bottom right) with mixing of SM and BSM CP-even HWW couplings. The box sizes correspond to the Monte Carlo statistical uncertainties [20].

Experimental constraints So far, no significant deviations from the Standard Model have been observed. The base samples have thus been chosen to include the Standard Model coupling with one or two additional BSM couplings.

Coverage In order to provide good coverage of the available configuration space, the samples have been chosen to include positive and negative values of the BSM couplings.

Subspace The input samples have been chosen such that it is easily possible to construct subsets that can act as morphing basis sets with lower dimensionality.

Events with two opposite-charge muons with $p_{\mathrm{T}} > 6\,\mathrm{GeV}$ and $|\eta| < 2.7$ are selected. Furthermore, at least two anti-k_t ($R = 0.4$) truth jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 5.0$ and $\Delta\eta_{jj} > 5$ in the final state are selected, which results in a VH-suppressed phase space. If several such jet pairs exist, the one including the jet highest in p_{T} is chosen as the pair of VBF tagging jets.

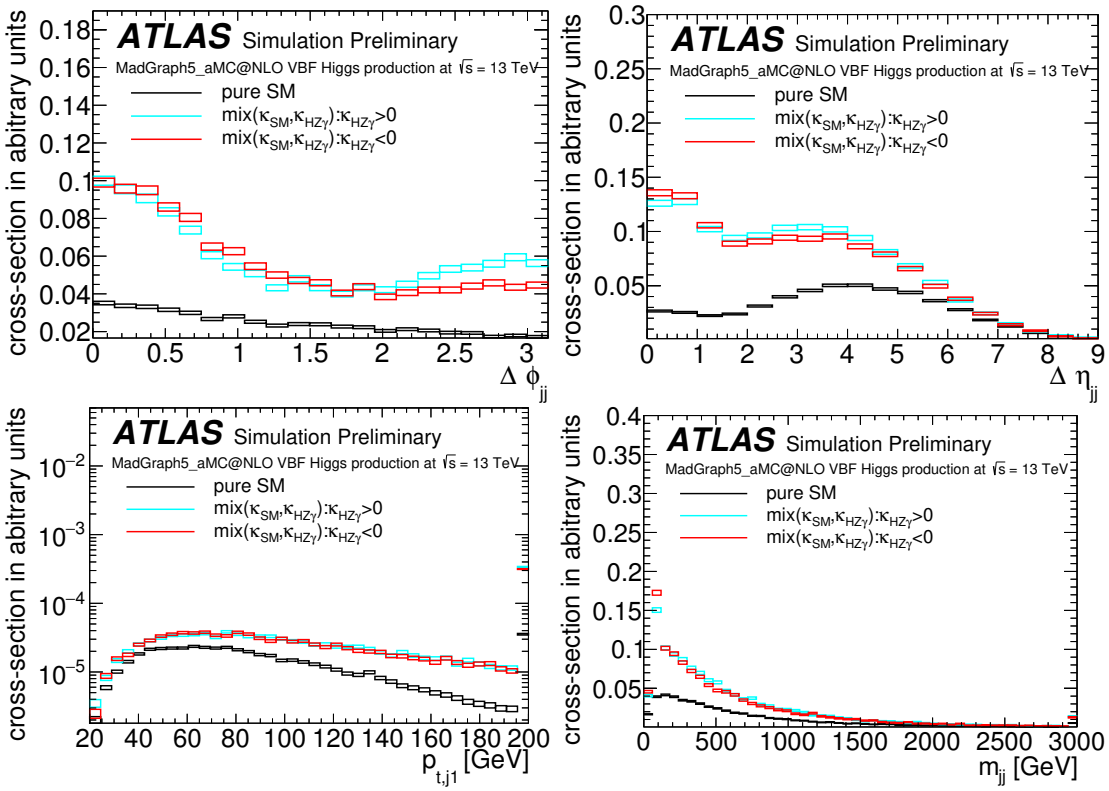


Figure 5.32.: VBF $H \to \mu\mu$ distributions for $\Delta\phi_{jj}$ (top left), m_{jj} (top right), $p_{\mathrm{T}}^{j_1}$ (bottom left), and $\Delta\eta_{jj}$ (bottom right) with mixing of SM and BSM CP-even $HZ\gamma$ couplings. The box sizes correspond to the Monte Carlo statistical uncertainties [20].

The pseudorapidity and angular distance between the two tagging jets $\Delta\eta_{jj}$ and $\Delta\phi_{jj}$, the invariant di-jet mass m_{jj} and the transverse momentum of the leading jet $p_{\mathrm{T}}^{j_1}$ are highly sensitive to the effects of non-SM couplings, as shown in Fig. 5.31, 5.32, 5.33, and 5.34 for the samples with non-zero BSM HWW coupling. The distributions in m_{jj} and $\Delta\eta_{jj}$ are strongly correlated. The observable $\Delta\phi_{jj}$ includes information on the sign of the BSM coupling parameter for κ_{HZZ}, κ_{HZA} and κ_{HWW}, while the distributions $p_{\mathrm{T}}^{j_1}$ and $\Delta\eta_{jj}$ are sign-sensitive for $\kappa_{H\partial Z}$ and $\kappa_{H\partial WR}$.

In order to maximize the sensitivity of the model to BSM effects, the observables are folded into a combined observable $\mathrm{var}_{4\mathrm{d}}$, which is constructed as

$$\mathrm{var}_{4\mathrm{d}} = 3^3 \cdot i_{\Delta\eta_{jj}} + 3^2 \cdot i_{m_{jj}} + 3 \cdot i_{p_{\mathrm{T}}^{j_1}} + i_{\Delta\phi_{jj}}, \tag{5.3.32}$$

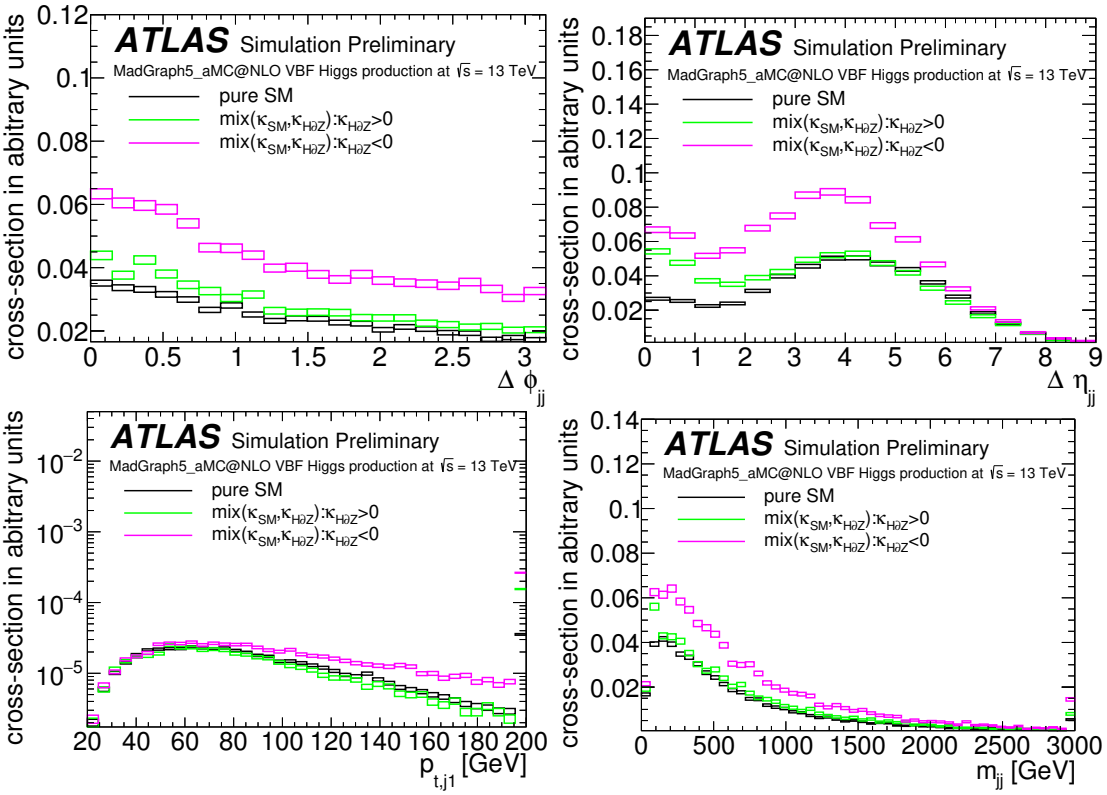


Figure 5.33.: VBF $H \to \mu\mu$ distributions for $\Delta\phi_{jj}$ (top left), m_{jj} (top right), $p_{\mathrm{T}}^{j_1}$ (bottom left), and $\Delta\eta_{jj}$ (bottom right) with mixing of SM and BSM CP-even $H\partial Z$ couplings. The box sizes correspond to the Monte Carlo statistical uncertainties [20].

where the values of each of the four variables is divided into three bins

$$i_x = \begin{cases} 0 & x < \left(\min_x + \frac{1}{3} \cdot (\max_x - \min_x)\right) \\ 1 & \left(\min_x + \frac{1}{3} \cdot (\max_x - \min_x)\right) \leq x < \left(\min_x + \frac{2}{3} \cdot (\max_x - \min_x)\right) \\ 2 & x \geq \left(\min_x + \frac{2}{3} \cdot (\max_x - \min_x)\right). \end{cases}$$

This definition is simply an "unrolling" of a four-dimensional histogram onto a single axis, correctly taking into account all bin-by-bin correlations. The minimum and maximum values of each variable are listed in Table 5.14. The distribution of events in this combined observable is shown in Fig. 5.35.

Based on this binning, a likelihood is constructed and fit to pseudo-data generated assuming the Standard Model and a statistical uncertainty of 8% on the total cross section. The resulting uncertainties on the fit parameters then allow to assess the influence the respective parameter has on the kinematic properties of events

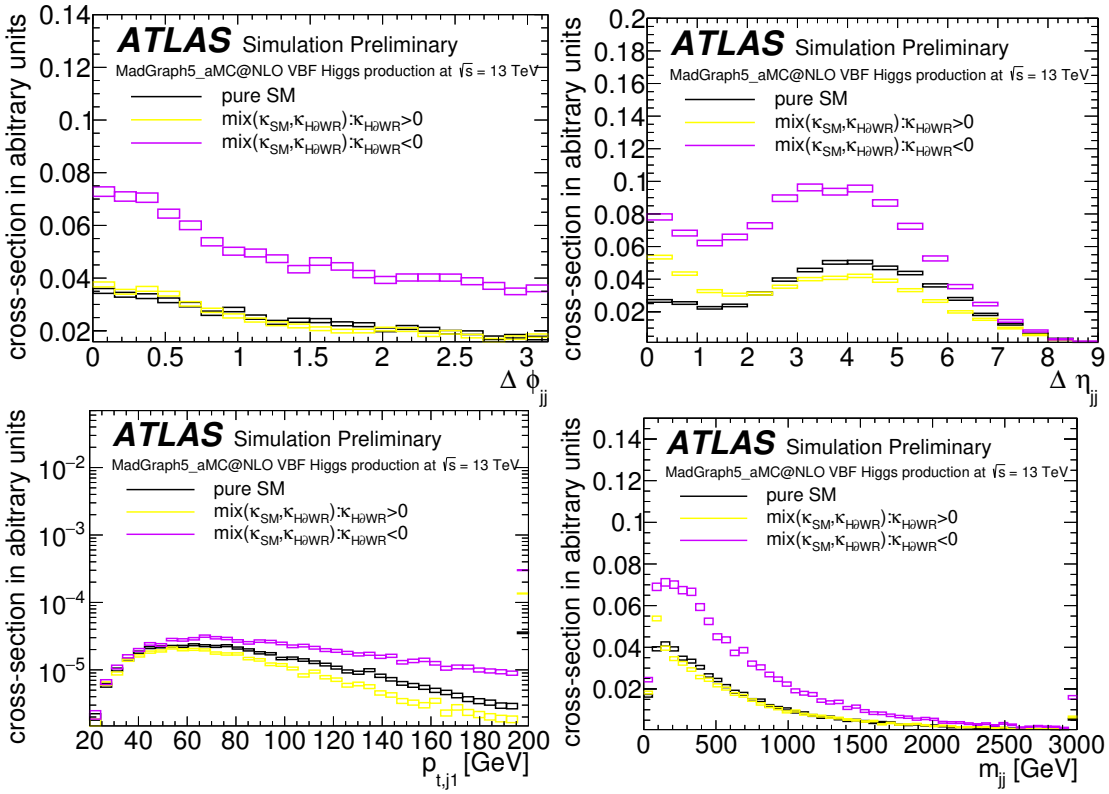


Figure 5.34.: VBF $H \to \mu\mu$ distributions for $\Delta\phi_{jj}$ (top left), m_{jj} (top right), $p_{\mathrm{T}}^{j_1}$ (bottom left), and $\Delta\eta_{jj}$ (bottom right) with mixing of SM and BSM CP-even $H\partial WR$ couplings. The box sizes correspond to the Monte Carlo statistical uncertainties [20].

arising from VBF Higgs boson production, and thus on the principle capacity to measure these parameters in such events.

Table 5.15 summarizes the resulting post-fit values and uncertainties of all coupling parameters. The value of the uncertainty provides insight about the sensitivity of the Monte Carlo signal distribution to this parameter for the given scenario, where large uncertainties correspond to small sensitivity.

As could be expected, the Higgs boson couplings to photons $A\gamma\gamma$, $AZ\gamma$, $H\gamma\gamma$ and $H\partial\gamma$ have very little influence on VBF events. One major result of this study

Table 5.14.: Minimum and maximum values of the VBF production jet variables $\Delta\eta_{jj}$, m_{jj}, $p_{\mathrm{T}}^{j_1}$ and $\Delta\phi_{jj}$.

	$\Delta\eta_{jj}$	m_{jj} [GeV]	$p_{\mathrm{T}}^{j_1}$ [GeV]	$\Delta\phi_{jj}$
min	0	0	20	0
max	6	1000	200	π

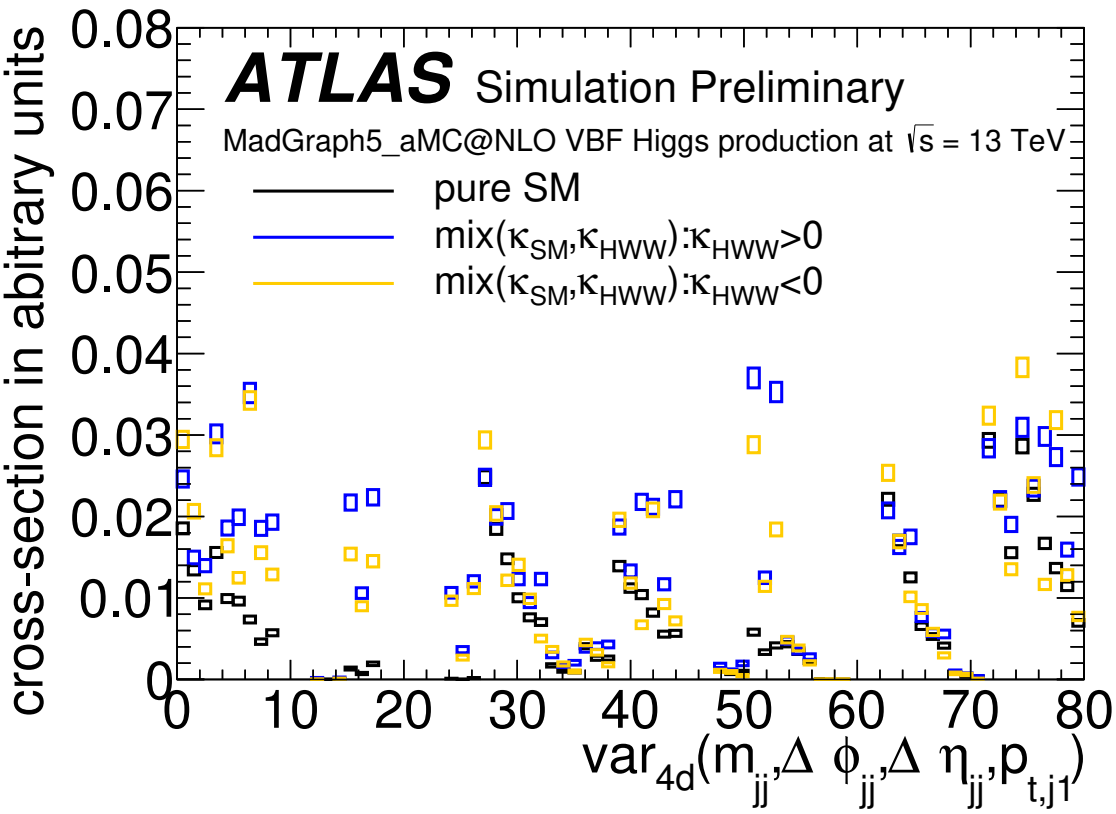


Figure 5.35.: VBF $H \to \mu\mu$ distribution of $\text{var}_{4\text{d}}$, which is a combination of $\Delta\eta_{jj}$, $\Delta\phi_{jj}$, m_{jj} and $p_{\text{T}}^{j_1}$ into one discriminating variable, shown for SM and mixing of SM with BSM CP-even HWW couplings. The box sizes correspond to the Monte Carlo statistical uncertainties [20].

is hence the conclusive proof that these couplings can be safely neglected in future studies of VBF Higgs boson production and do not need to be included in the model, allowing for a significant reduction in dimensionality.

However, it would be desirable to reduce the dimensionality of the model even further. In order to do this, it is useful to investigate the parameter correlations observed by the fit. They are shown in Fig. 5.36. From this, it becomes apparent that some of the coupling parameters like $\kappa_{H\partial Z}$ and $\kappa_{H\partial WR}$ share a significant amount of information. Hence, it might be possible to replace them with a single parameter encoding the sensitivity of both into the model.

Going one step further, one can attempt to find a transformation of the configuration space that will reduce the dimensionality of the model by removing spurious degrees of freedom.

The most straight-forward approach for this is clearly to perform a principle component analysis on the correlation matrix shown in Fig. 5.36 in order to trans-

Table 5.15.: Values of the coupling parameters and their respective uncertainties after fitting to SM pseudo-data with 8% cross section uncertainty. The top rows list the input parameters that were fixed to their nominal values during the fit.

parameter	post-fit value	+	−
Λ	1000.		
$\cos\alpha$	0.71		
$\kappa_{H\ell\ell}$	1.41		
$\kappa_{A\gamma\gamma}$	0	+219	−441
κ_{AWW}	0	+3	−2.6
$\kappa_{AZ\gamma}$	0	+441	−398
κ_{AZZ}	0	+2.7	−1.3
$\kappa_{H\gamma\gamma}$	0	+236	−91
$\kappa_{H\partial\gamma}$	0	+0.3	−0.6
$\kappa_{H\partial WI}$	0	+1.6	−0
$\kappa_{H\partial WR}$	0	+0.5	−0.3
$\kappa_{H\partial Z}$	0	+1.2	−0.5
κ_{Hww}	0	+1.5	−3
$\kappa_{HZ\gamma}$	0	+38	−49
κ_{HZZ}	0	+8	−2.5
κ_{SM}	1.41	+0.22	−0.11

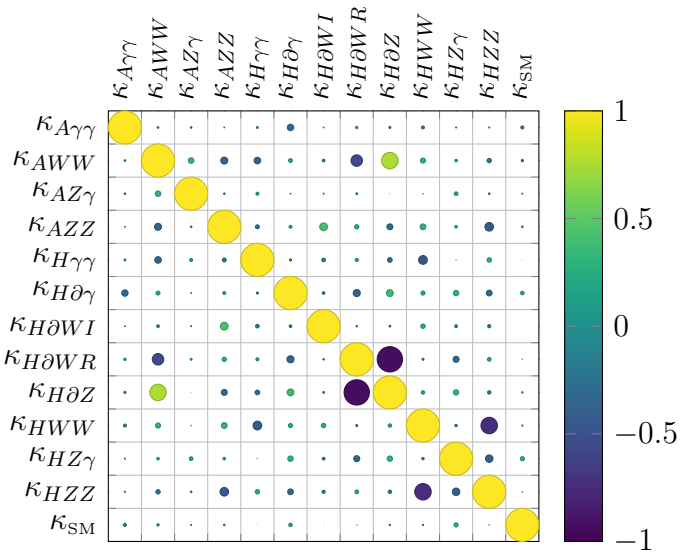


Figure 5.36.: Correlation matrix of the coupling parameters after fitting to SM pseudo-data with 8% cross section uncertainty. The area of the marker size is proportional to the correlation.

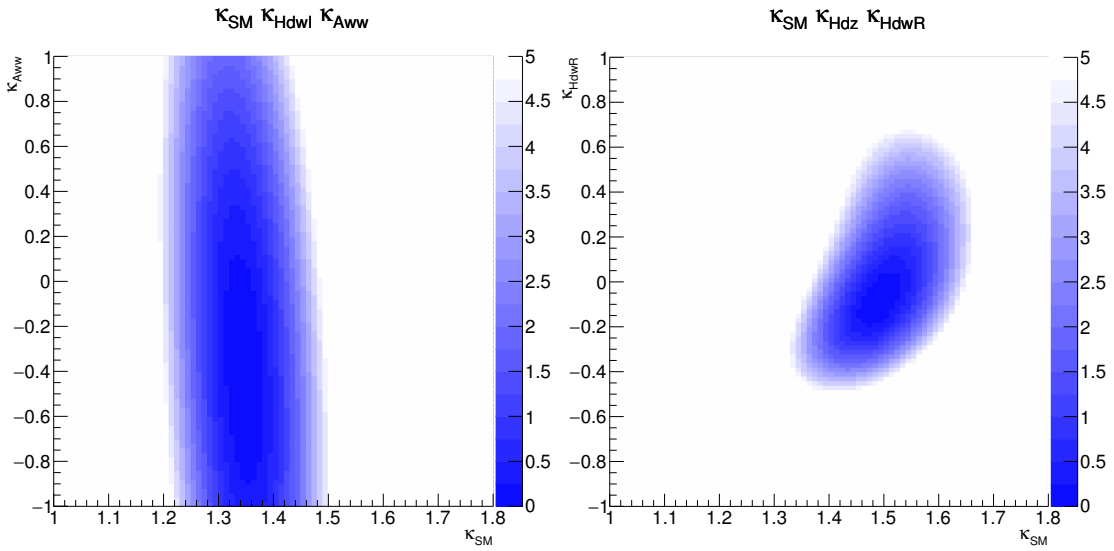


Figure 5.37.: Two-dimensional negative log-likelihood profiles of $H \to \mu\mu$ VBF models reduced to three fit parameters for the combination $\kappa_{\mathrm{SM}} \times \kappa_{H\partial WI} \times \kappa_{AWW}$ projected into the $\kappa_{\mathrm{SM}} \times \kappa_{AWW}$ plane (left) and $\kappa_{\mathrm{SM}} \times \kappa_{H\partial Z} \times \kappa_{H\partial WR}$ projected into the $\kappa_{\mathrm{SM}} \times \kappa_{H\partial WR}$ plane (right). Non-linear correlations show up as deviations from elliptical shapes.

form the model into a lower-dimensional eigen-vector space in which all linear correlations between the parameters are removed. However, first studies indicate that non-linear correlations are present between some parameters of the model, requiring a more sophisticated approach. For this insight, reduced models with only three of the coupling parameters have been built, all others set to their respective nominal Standard Model values. Then, likelihood scans are performed over the remaining parameter triplets, and the profile likelihood projections into each of the two-dimensional subspace are investigated. If only linear correlations are present, the resulting shapes are expected to be elliptical. This is not the case, as can be seen from Fig. 5.37, showing two of the more prominent examples of non-elliptical likelihood profiles. Investigations into more sophisticated methods of dimensionality reduction are currently ongoing.

5.3.9 Interference studies

The Lagrangian describing the interaction between the Higgs boson field X_0 and some vector boson V (like W or Z) takes the form

$$\mathcal{L} = \left\{ a_1 V_\mu V^\mu + a_2 V_{\mu\nu} V^{\mu\nu} + b V_{\mu\nu} \widetilde{V}^{\mu\nu} \right\} X_0, \tag{5.3.33}$$

where

$$V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \tag{5.3.34}$$

is the reduced field strength tensor, and

$$\widetilde{V}^{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}V^{\rho\sigma} \tag{5.3.35}$$

is the dual tensor. The couplings a_1 and a_2 are the SM and non-SM CP-even couplings, and b is a non-SM CP-odd coupling. Considering the tree-level $H \to VV$ decay, the tensor structure of this vertex is

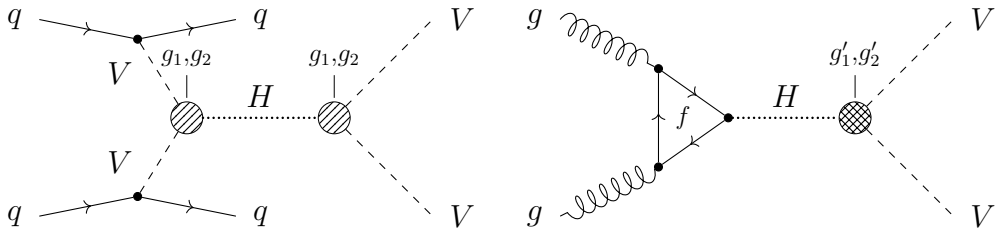
$$\mathcal{V}^{\mu\nu} \propto a_1 g^{\mu\nu} + a_2 \left(q_1^\mu q_2^\nu - q_1 q_2 g^{\mu\nu}\right) + b\epsilon^{\mu\nu\rho\sigma} q_{1\rho} q_{2\sigma}, \tag{5.3.36}$$

which will be a factor in the Matrix Element. However, the contraction of the totally antisymmetric tensor $\epsilon^{\mu\nu\rho\sigma}$ with $g^{\mu\nu}$ will make any interference terms between CP-even and CP-odd terms vanish when calculating cross sections or partial widths.

This section is to demonstrate that this is correctly modeled by *Effective Lagrangian Morphing*, and to discuss how such physical insights can be employed to use the technique of Effective Lagrangian Morphing more efficiently. One easy way to test this would of course be to manually modify the morphing function to exclude the interference terms. However, there is a more elegant solution.

Figure 5.38a shows a Feynman diagram of Higgs boson production via vector boson fusion with subsequent decay to vector bosons, with the exact same effective vertex with two available $H \to VV$ couplings g_1 and g_2 being allowed for both production and decay. The Matrix Element in this case will take the form

$$\mathcal{M} \propto \left(g_1 \mathcal{M}_1 + g_2 \mathcal{M}_2\right)^2 \tag{5.3.37}$$



(a) Production via VBF, $H \to VV$ couplings g_1 and g_2.

(b) Production via ggF, $H \to VV$ couplings g_1' and g_2'.

Figure 5.38.: Feynman diagrams for Higgs boson production via different production modes and subsequent decay $H \to VV$, employing effective $H \to VV$ couplings.

$$|\mathcal{M}|^2 \propto g_1^4 |\mathcal{M}_1|^4 + g_2^4 |\mathcal{M}_2|^4 + 2g_1^2 g_2^2 \left(|\mathcal{M}_1|^2 |\mathcal{M}_2|^2 + \mathfrak{R}(\mathcal{M}_1 \mathcal{M}_2^*)^2\right) + 4g_1^3 g_2 \left|\mathcal{M}_1^2\right| \mathfrak{R}(\mathcal{M}_1 \mathcal{M}_2^*) + 4g_1 g_2^3 \mathfrak{R}(\mathcal{M}_1 \mathcal{M}_2^*) |\mathcal{M}_2|^2 . \tag{5.3.38}$$

The number of terms is five, as expected in this case.

In contrast, Fig. 5.38b shows a Feynman diagram of Higgs boson production via gluon fusion with subsequent decay to vector bosons, with only the decay being affected by the effective $H \to VV$ couplings this time. In this case, the Matrix Element will take the form

$$\mathcal{M}' \propto (g_1' \mathcal{M}_1' + g_2' \mathcal{M}_2') \tag{5.3.39}$$

$$|\mathcal{M}'|^2 \propto g_1'^2 |\mathcal{M}_1'|^2 + g_2'^2 |\mathcal{M}_2'|^2 + 2g_1' g_2' \mathfrak{R}(\mathcal{M}_1' \mathcal{M}_2'^*) , \tag{5.3.40}$$

The number of terms is three in this case. One can now perform the replacements $g_i' \to g_i^2$ and $\mathcal{M}_i' \to |\mathcal{M}_i|^2$ $(i = 1, 2)$ to obtain

$$|\mathcal{M}'|^2 \propto g_1^4 |\mathcal{M}_1|^4 + g_2^4 |\mathcal{M}_2|^4 + 2g_1^2 g_2^2 |\mathcal{M}_1|^2 |\mathcal{M}_2|^2 . \tag{5.3.41}$$

Interestingly, this is exactly Eq. 5.3.38 without the interference terms. Hence, one can consider the two processes VBF $H \to VV$ with some couplings g and ggF $H \to VV$ with couplings g^2 "adjoint" in the sense that they produce the same structures in the Matrix Element with exception of the interference terms.

This similarity can be used to study the effect of interference terms without manually removing them from the calculation.

Physical insights like this can be used to manually reduce the dimensionality of the parameter space. For example, one can construct two different morphing functions, one only for the CP-even and another only for the CP-odd part of the Lagrangian, and add their predictions linearly, as no interference can occur.

In order to prove experimentally that this is possible, two sets of samples have been generated with MADGRAPH5_aMC@NLO 2.3.3.p1 [130] in the Higgs Characterization Model introduced in Section 2.4.2 using the leading-order PDF set NNPDF23LO [131]. Hadronization and showering were performed with PYTHIA8.212 [125] using the A14 tune [134]. One of the sets includes the CP-even BSM coupling κ_{HWW}, the other one includes the CP-odd BSM coupling κ_{AWW}. Then, two versions of the morphing function have been constructed on each sample set, one standard version including the interference terms, and another one employing the technique of adjoint processes to exclude the interference terms. The sample configurations as well as the resulting predictions in comparison with an independent validation sample are shown in Fig. 5.39. Notably, the full calculation including the interference terms requires five samples, whereas the calculation of the morphing function without the interference terms only requires three. The reduced base

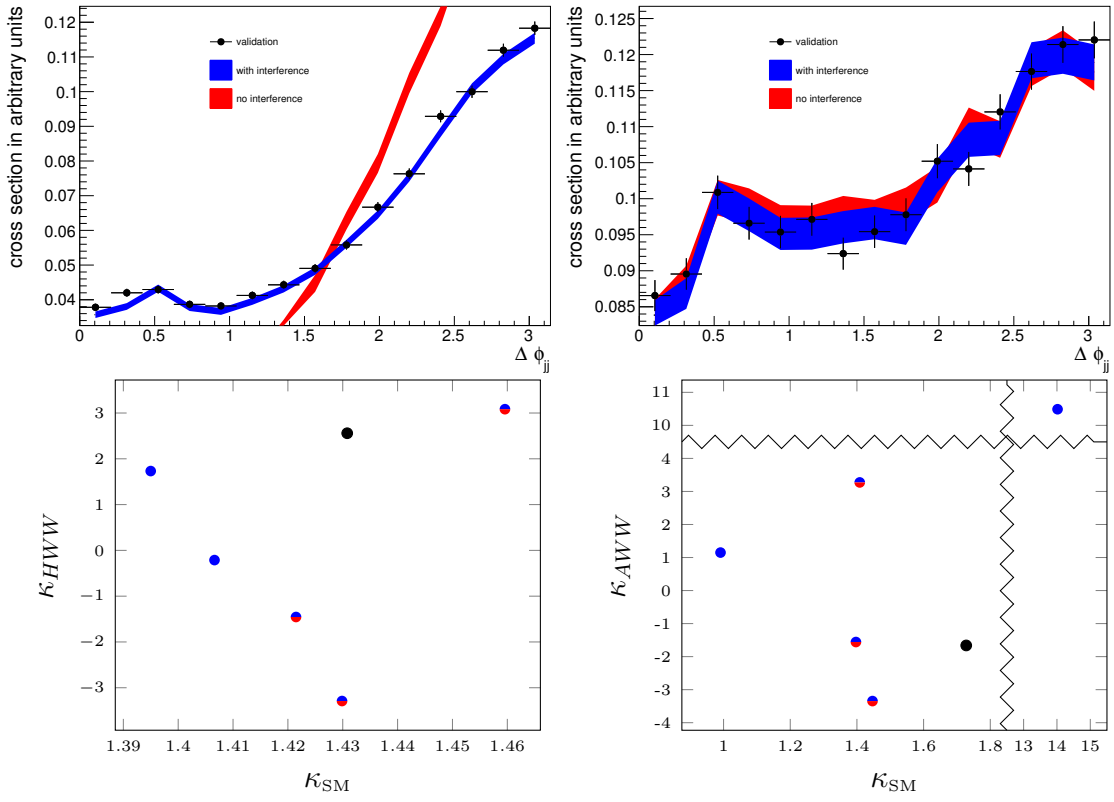


Figure 5.39.: Comparison of Effective Lagrangian Morphing for VBF $H \to W^{\pm}W^{\mp *}$, including either the CP-even HWW coupling (left) or the CP-odd AWW coupling (right) in addition to the Standard Model coupling. For each scenario, the morphing is constructed with (blue) or without (red) the interference terms. Also shown is an independent validation sample (black). The top panels show the predicted distributions in the azimuthal angular separation of the jets $\Delta\phi_{jj}$, whereas the bottom panels show the coupling parameter choices of the samples. The bi-colour markers indicate samples that are used in both cases. All samples are using $\cos\alpha = 1/\sqrt{2}$ and $\Lambda = 1\,\text{TeV}$. The zig-zag lines indicate axis discontinuities.

sets are true subsets of the full ones, hence the uncertainty bands shown are not statistically independent of each other. However, they are still useful for comparison with the statistically independent validation sample.

It is obvious that while in the CP-odd case, the modeling of the validation sample is equally good with and without the interference terms, the modeling breaks down entirely without the interference terms in the CP-even case. This is completely expected, as the $\kappa_{AWW}\kappa_{\mathrm{SM}}$ interference terms are expected to vanish either way for the CP-odd scenario, whereas the $\kappa_{HWW}\kappa_{\mathrm{SM}}$ interference terms do not vanish in the CP-even case. Hence, this result illustrates that Effective Lagrangian Morphing correctly models the interference terms, even in cases where they contribute significantly to the final result, and that physical insights into the tensor structure of the Matrix Element can be used to effectively reduce the dimensionality of the setup.

5.3.10 Base optimization

The techniques and studies described in the previous sections provide details on why a high-dimensional parameter space poses a challenge for finding a usable basis for Effective Lagrangian Morphing and provides insights how a dimensionality reduction of the parameter space can be achieved. However, ultimately, all of these approaches are related to the physical properties of the system subject to study, and not every system of interest will always provide handles for possible reduction of the dimensionality. Instead, a completely general strategy for finding the optimal basis set is required. This section now provides a prescription of such an algorithm. The results of this algorithm applied to the example of a VBF $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study employing three free, CP-even coupling parameters κ_{SM}, κ_{HWW}, and $\kappa_{H\partial WR}$ are shown in Section 5.5.2.

The previous discussion suggests that the norm of the morphing matrix or, somewhat more sophisticated, the condition number of the morphing matrix would pose adequate candidates for a figure of merit. However, both of these figures only quantify the effects of the morphing matrix itself. As highlighted in Section 5.3.6, this is not necessarily optimal. The assumption discussed there that all the base samples have comparable statistical power is very reasonable in practice, as this relates to an equal number of events and thus an equal amount of computing time for the production of all base samples. Under this assumption, the actual numerical values of the cross sections are equally as important as the entries of the morphing matrix, as both types of quantities ultimately act as scaling factors for the relative statistical uncertainties.

The pristinely accurate figure of merit is of course the uncertainty

$$\Delta\sigma^2(\vec{g}) = \sum_j^N \left(w_j(\vec{g})\Delta\xi_j\right)^2 \quad \text{with} \quad w_j(\vec{g}) = \sum_i^N P_i(\vec{g}) \tag{5.3.42}$$

of the prediction obtained from the morphing function with N samples. However, this quantity is a function of the point in parameter space that is the target of the morphing. As the underlying idea of the morphing technique is to provide full coverage of a region $\mathcal{A} \subset \mathbb{R}^{N\times N}$ of parameter space with a finite (and small) set of base samples, it is questionable which is the correct strategy to minimize the uncertainty in the entire region. One could argue that the maximum

$$f(\vec{g}_1, \ldots, \vec{g}_N) = \max_{\vec{g}\in\mathcal{A}} \sqrt{\Delta\sigma^2(g \mid \vec{g}_1, \ldots, \vec{g}_N)} \tag{5.3.43}$$

would be a viable figure of merit, but considering Fig. 5.24 back in Section 5.3.4 reveals that such a choice would make the figure of merit entirely dependent on the chosen boundary of the region of interest, as uncertainties usually diverge strongly at the bounds of the region. Instead, it seems appropriate to consider the integral

$$f(\vec{g}_1, \ldots, \vec{g}_N) = \int_{\vec{g}\in\mathcal{A}} \sqrt{\Delta\sigma^2(\vec{g} \mid \vec{g}_1, \ldots, \vec{g}_N)} d\mathcal{A} \tag{5.3.44}$$

over the uncertainty in the region of interest. The uncertainty is a function of many parameters, and it is not clear whether the calculation of the integral in a symbolic fashion is feasible. Thus, it needs to be evaluated at a finite set $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_M\}$ of M *benchmark samples*, *i. e.*,

$$f(\vec{g}_1, \ldots, \vec{g}_N) = \sum_{b\in\mathcal{B}} \sqrt{\Delta\sigma^2(\vec{g}_b \mid \vec{g}_1, \ldots, \vec{g}_N)}. \tag{5.3.45}$$

Additionally, the evaluation of the uncertainty at these benchmarks for some *test basis* requires the knowledge of the cross sections of the constituent samples, and thus their production with a Monte Carlo generator. With this approach, the optimization would essentially boil down to trying out a couple of different basis sets, and settling for the best one, which is not very satisfactory. In order to circumvent this, one would need a technique to estimate the cross sections at any given point in parameter space. Luckily, the method of Effective Lagrangian Morphing is itself such a technique. From this, it becomes apparent that in order to find the optimal set of base samples, two layers of Effective Lagrangian Morphing need to be concatenated to a chain. The underlying layer, the *cross section morphing*, provides the cross sections for the test basis. The top layer, the *test morphing*,

then calculates the predictions at the benchmarks. Of course, the cross section morphing itself needs a basis, and the precision of the prediction will depend on the choice of this basis. Any imprecisions in the prediction of the cross sections of the test basis will bias the results of the optimization.

The benchmarks themselves need not necessarily be generated as real samples, they may just be defined as points in the parameter space. However, if they are generated as real samples, this allows to estimate the bias from the imprecision of the cross section morphing with a bias term

$$B_b = \left(\sigma(\vec{g}_b \mid \vec{g}_1, \ldots, \vec{g}_N) - \sigma_b\right)^2 , \tag{5.3.46}$$

which can be included in the figure of merit to avoid convergence to regions of parameter space that are subject to large systematic biases in the calculation of the figure of merit.

It would of course be best to start the optimization with an already optimized basis. It thus seems desirable to restart the optimization, using the optimized basis as the new basis of the cross section morphing. This is, however, only partially true, as samples of the optimized basis need not themselves lie in the region of interest covered by the benchmark samples. Nevertheless, the algorithm may be restarted with the optimized basis used for the cross section inputs to obtain improved results.

Further effects might influence the result of the uncertainty calculation. For example, the benchmark might be placed at a point in parameter space that is difficult to model for the specific Monte Carlo generator or the concrete implementation of the model in question and thus associated to large uncertainties. In order to account for this effect, the uncertainty obtained from the test morphing can be normalized to the uncertainty found in the benchmark sample.

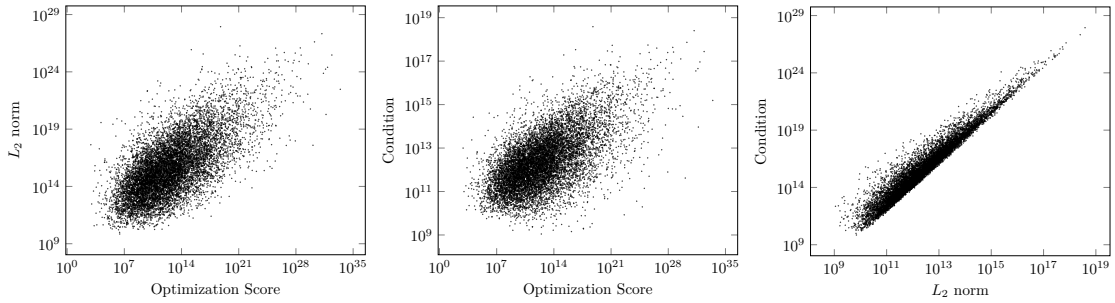


Figure 5.40.: Correlation between L_2-norm, condition number and optimization score for randomly generated basis sets for a VBF $H \to VV$ scenario with two shared coupling parameters.

Thus, the suggested figure of merit is

$$f(\vec{g}_1, \ldots, \vec{g}_N) = \sum_{b \in \mathcal{B}} \sqrt{\frac{\Delta\sigma^2(\vec{g}_b \mid \vec{g}_1, \ldots, \vec{g}_N)}{\Delta\sigma_b^2} \cdot \left(\sigma(\vec{g}_b \mid \vec{g}_1, \ldots, \vec{g}_N) - \sigma_b\right)^2}. \quad (5.3.47)$$

This optimization score has been evaluated on 10.000 test sets for a $H \to VV$ configuration with two free coupling parameters and compared with the L_2-norm and the condition number of the morphing matrices. The result is shown in Fig. 5.40. As expected, the quantities are positively correlated, but significant deviations from a one-to-one correspondence can be observed.

With this technique, it is possible to evaluate the performance of a test basis with sufficiently low computational effort to allow for a numeric optimization to take place. This is, of course, a natural consequence from the fact that Effective Lagrangian Morphing has been designed to provide the means of obtaining predictions sufficiently performant to facilitate a likelihood minimization. The resulting algorithm is schematically displayed in Fig. 5.41.

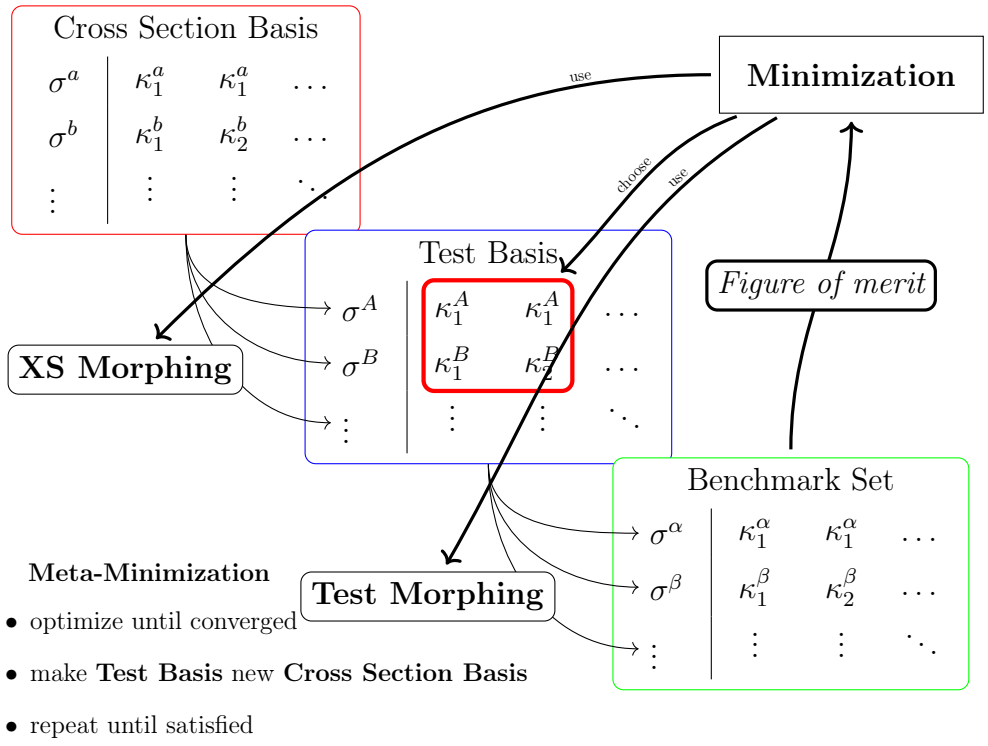


Figure 5.41.: Schematic view of the base optimization algorithm.

One very subtle point that has not yet been resolved to a satisfactory degree however is the question how exactly the uncertainties on the test basis should be defined. The morphing function predicting X at the coupling values $\vec{g}$ can be cast as

$$X(\vec{g}) = P(\vec{g})A\vec{\xi}, \tag{5.3.48}$$

where P is a vector of polynomials in $\vec{g}$, A is the morphing Matrix, and $\vec{\xi}$ are the predictions at the base samples. If the base samples are themselves results of morphing, this becomes

$$X(\vec{g}) = P(\vec{g}) \underbrace{AP(\vec{g}\prime)A'}_{\widetilde{A}} \vec{\xi}', \tag{5.3.49}$$

where the expression indicated by the underbrace can be identified as a matrix $\widetilde{A}$. Thus, it appears that effective Lagrangian morphing is *transitive* in the sense that inserting an additional layer should not change the final result, and the uncertainties of the predictions for X will only depend on the uncertainties of the cross section basis $\vec{\xi}'$. However, if the intermediate results are actually projected onto the intermediate test basis ξ_i, the correlations between the uncertainties of the respective ξ_i will vanish from the calculation. This is not an accident, but an intended effect, as the uncertainties on the test basis would also be uncorrelated in the case of actually generated samples. In principle, the statistical fluctuations on the test basis $\vec{\xi}$ propagated from the cross section basis $\vec{\xi}'$ are already accounted for by the bias term in Eq. 5.3.46. This, however, begs the question whether the uncertainty predictions obtained by the cross section morphing for the test basis $\vec{\xi}$ should at all be used, or whether they should be replaced by some fixed value corresponding to the envisioned uncertainties on the actual samples. There is no clear answer to this question yet, and research is ongoing to investigate this detail of the algorithm. Some preliminary results are presented in Section 5.5.2.

5.4 EFT measurements in $H \to ZZ^*$

The analysis introduced in Section 5.1.1 [22] has also obtained a measurement of two non-SM coupling parameters $g_{HVV} = \kappa_{HVV}$ and $g_{AVV} = \kappa_{AVV} \cdot \sin\alpha$, using the technique of Effective Lagrangian Morphing and employing the Higgs Characterization Model introduced in Section 2.4.2 [47].

Higgs boson production via the ggF, ttH and bbH modes and subsequent decay to a pair of Z bosons includes the coupling of the Higgs boson to vector bosons only in the decay vertex, whereas for the VBF and VH production modes, the same coupling is also present in the production vertex. Hence, the number of independent

base samples needed to model the effects of the non-SM coupling parameters is different in the two cases. Following Eq. 5.3.6,

$$N_{\mathrm{ggF}/ttH/bbH}^{g_{HZZ},g_{AZZ}} = 3 \tag{5.4.1}$$

$$N_{\mathrm{VBF}/\mathrm{VH}}^{g_{HZZ},g_{AZZ}} = 5. \tag{5.4.2}$$

The base samples have been produced with MADGRAPH5_aMC@NLO without allowing for variations in the total width of the Higgs boson, which is thus not modeled correctly. Hence, an additional correction corresponding to $\Gamma_{\mathrm{BSM}}/\Gamma_{\mathrm{SM}}$ is applied to the product ($\sigma\mathcal{B}$). The correction is approximately -2% for $g_{AVV} = \pm 8$, and is about $+14\%$ and -17% for $g_{HVV} = -8$ and $g_{HVV} = +8$, respectively. For the gluon fusion contribution, the acceptance is assumed to be independent of the non-SM coupling parameters and the Effective Lagrangian Morphing is just applied as a scaling of the total rate.

For this measurement, the event categories introduced in Section 5.1.1 are used, but only effects on the total yields in each category are considered, without exploiting any additional discriminant shape information. The respective numbers are on display in Table 5.1 in Section 5.1.1 in comparison to the Standard Model predictions.

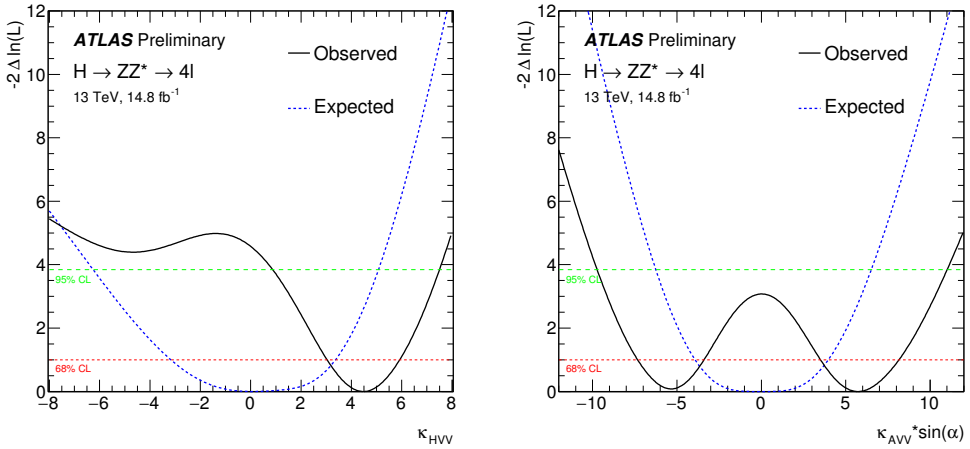


Figure 5.42.: Observed (solid black line) and expected (dashed blue line) negative log-likelihood scans for $g_{HVV} = \kappa_{HVV}$ (left) and $g_{AVV} = \kappa_{AVV} \cdot \sin\alpha$ (right). The horizontal dashed lines indicate the value of the profile likelihood ratio corresponding to the 68% (red) and 95% (green) confidence level intervals for the parameter of interest, assuming the asymptotic χ^2 distribution of the test statistic [22].

Table 5.16.: Observed and expected limits at 95% CL on κ_{HVV} and $\kappa_{AVV} \cdot \sin\alpha$ [22].

Not excluded range at 95% CL	expected	observed	central value deviation
κ_{HVV}	$[-6.3, 5.1]$	$[0.9, 7.5]$	1.8σ
$\kappa_{AVV} \cdot \sin\alpha$	$[-6.3, 6.5]$	$[-9.7, 11.0]$	2.1σ

Limits on the two non-SM couplings are derived with a likelihood fit, where the Standard Model coupling and the effective couplings modeling ggF Higgs boson production are fixed to unity. Measurements of the two couplings are conducted separately due to the low statistical power of the data set. The expected and observed negative log-likelihood scans as functions of the non-SM couplings are shown in Fig. 5.42 and listed in Table 5.16. As the observed number of events exceeds the SM predictions especially in the $n_{\text{jet}}\geq 2$ VBF category, the best-fit values are not precisely at the Standard Model point, and the observed exclusion limits are slightly weaker than the expectation.

5.5 Prospects for $H\to W^{\pm}W^{\mp *}$ measurements

The techniques developed in this chapter can be applied to facilitate a measurement of BSM couplings in the $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay channel. This is naturally most feasible for the VBF production mode, which is subject to small uncertainties from theory and will greatly profit from the inclusion of the full 2016 data set.

As a baseline for such an analysis, the Boosted Decision Tree introduced in Section 4.6 can be used to provide a clean signal region. Discriminant power between the different non-SM scenarios can then be derived either from individual kinematic distributions, or in the form of a dedicated multivariate classifier, where the latter option seems especially promising.

Figure 5.43 shows the output score of the Boosted Decision Tree used in the $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ VBF analysis presented in Section 4.6, applied to various EFT scenarios in comparison to the Standard Model background, scaled to $28\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ pp collision data. It is clearly visible that the BDT, which was trained as a binary classifier to distinguish between the SM signal and background hypotheses, is very capable of separating most of the EFT scenarios quite reasonably from the background. However, a distinction between the different EFT scenarios is hardly possible without exploitation of dedicated kinematic properties sensitive to BSM contributions.

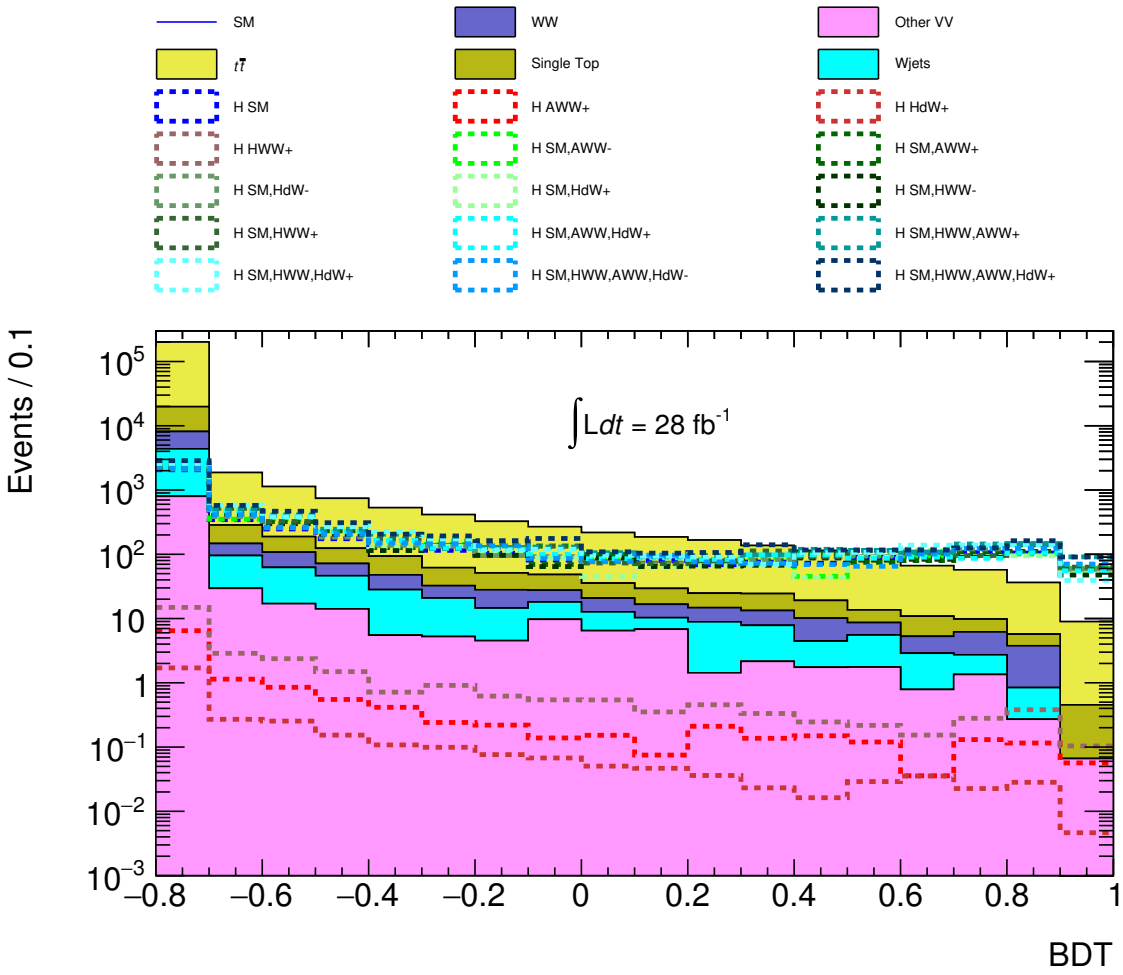


Figure 5.43.: BDT output score for the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ VBF analysis presented in Section 4.6 for various EFT scenarios in comparison to the Standard Model background expected in 28 fb^{-1} of $\sqrt{s} = 13$ TeV data. The Higgs boson signal includes ggF, VBF and VH production. All samples have been generated with MADGRAPH5_aMC@NLO2.3.3.p1 [130] in the Higgs Characterization Model introduced in Section 2.4.2 using the leading-order PDF set NNPDF23LO [131], showered with PYTHIA8.212 [125] with the A14 tune [134], and processed through the full ATLAS event simulation and reconstruction. The coupling identifiers listed in the legend encode the EFT configuration of the corresponding sample according to the following key.

coupling	label	not listed	listed	listed "+"	listed "−"
κ_{SM}	SM	0	$\sqrt{2}$		
κ_{Hgg}	Hgg	0	$\sqrt{2}$		
$\kappa_{HWW} = \kappa_{HZZ}$	HWW	0		2.28319	−2.28319
$\kappa_{AWW} = \kappa_{AZZ}$	AWW	0		2.40239	−2.40239
$\kappa_{H\partial WI} = \kappa_{H\partial WR} = \kappa_{H\partial Z}$	HdW	0		0.257933	−0.257933

5.5.1 Classifier morphing

Traditionally, multivariate classifiers have been used in High Energy Physics mainly for binary classification into a *signal* and a *background* category. Such a classifier ideally learns the regression function

$$s(x) = \frac{p(x \mid H_1)}{p(x \mid H_1) + p(x \mid H_0)}, \tag{5.5.1}$$

where $p(x \mid H)$ is the probability to observe the data x under assumption of a hypothesis H, where H_0 and H_1 denote the hypotheses assuming background only and a superposition of signal and background, respectively.

In the case at hand, however, multiple different signal hypotheses exist which need to be distinguished. In such a case, multiclass classifiers can be used to discriminate between a set of distinct benchmark scenarios.

This is, however, conceptually different to the case of an EFT coupling measurement. Techniques like the *Effective Lagrangian Morphing* allow to model the likelihood as a function of continuous coupling parameters. It is interesting to note that the likelihood ratio

$$\Lambda = \frac{p(x \mid H_1)}{p(x \mid H_0)} \tag{5.5.2}$$

has a one-to-one correspondence with the regression function s in Eq. 5.5.1. Indeed, the Neyman-Pearson lemma [188] implies that the optimal multivariate classifier will be the one approximating the likelihood ratio.

Recent publications find that discriminative classifiers can be used to approximate parametric likelihood functions by decomposing the combined classifier into a linear combination of pairwise classifiers [189]. The likelihood parametric in one or more EFT parameters θ decomposes as

$$p(x \mid \theta) = \sum_c w_c(\theta)\, p_c(x) \tag{5.5.3}$$

into a finite sum of non-parametric likelihoods over the event classes or benchmark scenarios c. The likelihood ratio then similarly decomposes as

$$\frac{p(x \mid \theta_0)}{p(x \mid \theta_1)} = \frac{\sum_c w_c(\theta_0)\, p_c(x)}{\sum_{c'} w_{c'}(\theta_1)\, p_{c'}(x)} \tag{5.5.4}$$

$$= \sum_c \left[\sum_{c'} \frac{w_{c'}(\theta_1)\, p_{c'}(x)}{w_c(\theta_0)\, p_c(x)} \right]^{-1}. \tag{5.5.5}$$

Notably, the likelihood p_c corresponds to a fixed scenario c, and the dependency on the EFT parameters is captured solely in the weight functions $w_c(\theta)$, which yield the weight associated to scenario c when evaluating the likelihood at some specific value of θ.

One can now replace the pairwise likelihood ratios with the evaluations of the corresponding classifiers [189]

$$\frac{p_{c'}(x)}{p_c(x)} = \frac{p_{c'}(s_{c,c'}(x))}{p_c(s_{c,c'}(x))}. \tag{5.5.6}$$

This construction is absolutely coherent with the derivation of Effective Lagrangian Morphing. The weights w_c are exactly the sample weights used in Effective Lagrangian Morphing. This method allows to predict the output of a multivariate classifier trained to distinguish between any two parametric hypotheses as a function of the parameters of both hypotheses.

In order to supplement Effective Lagrangian Morphing with classifier morphing, one needs a family of multivariate classifiers trained on the pairwise combinations of the basis samples of the Effective Lagrangian Morphing. After the output of the classifier optimal for the point of interest has been obtained by weighted summation over the pairwise classifier outputs for any event, the distribution of these outputs can be computed for all of the input samples. Effective Lagrangian Morphing can be then used to predict this distribution at the point of interest itself. With this technique, morphing can model the likelihood ratio, that is the optimal classifier, for any arbitrary scenario.

5.5.2 Base sample optimization

The optimization algorithm introduced in Section 5.3.10 has been employed to find an optimized set of base samples for a measurement of VBF $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$. The parameters subject to study are the Standard Model coupling κ_{SM}, the CP-even vector boson coupling $\kappa_{HVV} = \kappa_{HWW} = \kappa_{HZZ}$, and the derivative coupling $\kappa_{H\partial V} = \kappa_{H\partial WI} = \kappa_{H\partial WR} = \kappa_{H\partial Z}$. The resulting number of samples is 15.

Throughout the iterations of the optimization, several sample sets have been produced. All of the samples have been generated with MADGRAPH5_aMC@NLO 2.3.3.p1 [130] in the Higgs Characterization Model introduced in Section 2.4.2 using the leading-order PDF set NNPDF23LO [131]. Hadronization and showering were performed with PYTHIA8.212 [125] using the A14 tune [134].

Figure 5.44 shows a scan of the optimization score of the configuration as a function of the κ_{HVV} coupling parameter for one of the samples. The fact that the optimization has converged to a minimum indicated by the green lines is clearly visible. Additionally, a sharp spike can be seen, which corresponds to a

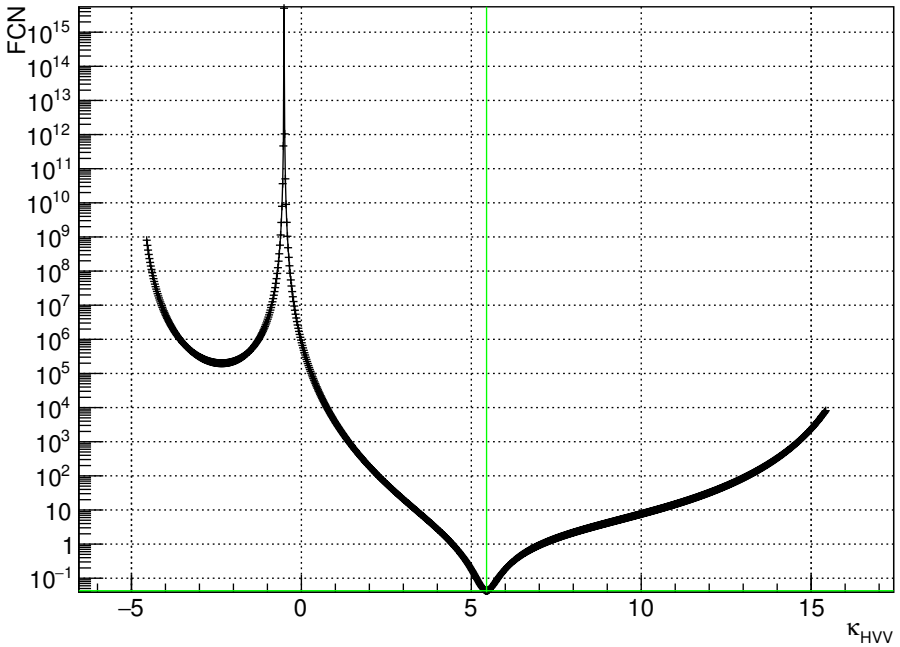


Figure 5.44.: Optimization score as a function of the parameter κ_{HVV} for one of the samples. The green lines indicate the minimum. The spike corresponds to a singular morphing matrix.

non-invertible configuration of the morphing matrix, which is connected to a large condition number and thus high uncertainties on the result.

Figure 5.45 shows distributions of the azimuthal angle between the two VBF tagging jets $\Delta\phi_{jj}$ throughout three iterations of the optimization algorithm. The positions of the samples throughout three steps of the minimization are shown in Fig. 5.46 and 5.47. In order to additionally investigate the question whether the inclusion of the propagated uncertainties on the test basis cross section is sensible, the optimization has been conducted for both scenarios – using the propagated uncertainty as well as using a fixed relative uncertainty of 5% on the test basis cross sections. The first row depicts the start values, which are shared between both variants.

The variant using the propagated uncertainty seems to converge faster, but a definite judgment is hard to make based on this type of anecdotal evidence. It is interesting to note though that, as can be seen clearly in Fig. 5.47, the optimal configuration includes a set of basis samples close to the region of interest near the Standard Model, and a few extreme BSM samples. This can be understood intuitively by realizing that the uncertainty is determined as a product of the cross section and the weight of the sample. For a region of interest near the Standard Model, the extreme BSM samples, which have large cross sections, will have a small

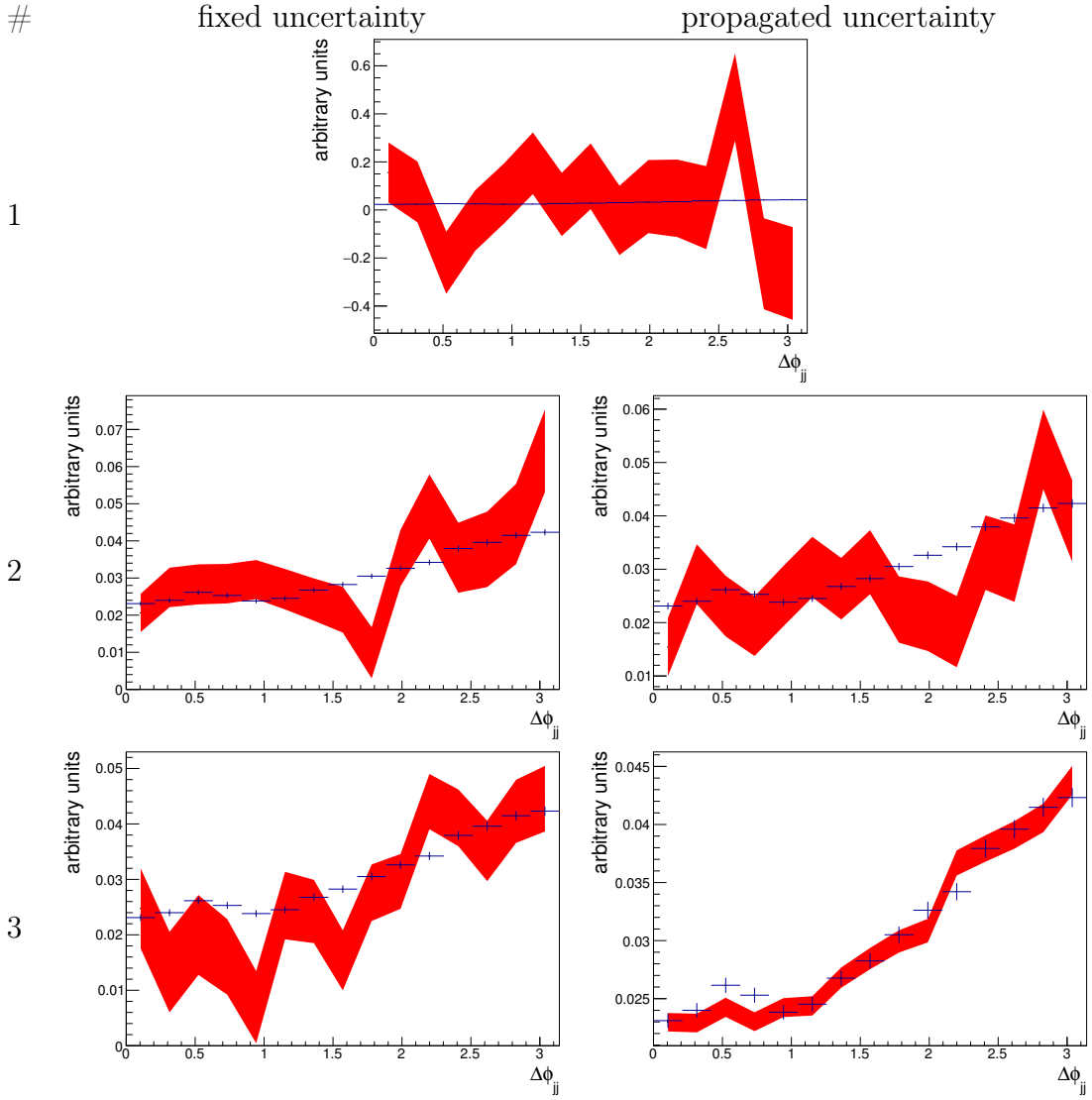


Figure 5.45.: Prediction obtained by Effective Lagrangian Morphing (red band) in comparison to a statistically independent validation sample (blue points) for the distribution of the azimuthal angle between the two VBF tagging jets $\Delta\phi_{jj}$ three iterations of the optimization algorithm for a fixed uncertainty on the cross sections of the test basis (left) and using the propagated uncertainty estimate (right).

weight. The SM-like samples on the other hand will have a large weight in the morphing function, but a relatively small cross section similar to the SM one. Thus, there is a trade-off between the two types of samples, the optimum of which is provided by the algorithm. However, in both cases, the resulting uncertainty of the prediction obtained by Effective Lagrangian Morphing reduces by several orders of magnitude in just three iterations, which is a significant success.

With these very promising results, a measurement of EFT coupling parameters in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ VBF category can certainly be conducted. However, as the physics model will also need to include the dominant ggF production mode, a similar optimization will be required to find a suitable basis for the gluon fusion signal samples before a full signal model can be constructed. Whether the analysis using the full 2016 data set and beyond can rely on a classical cut-based technique or a simple multivariate classifier, or whether the advanced techniques of classifier morphing presented in Section 5.5.1 will need to be employed to obtain a meaningful physics result in this case remains to be seen.

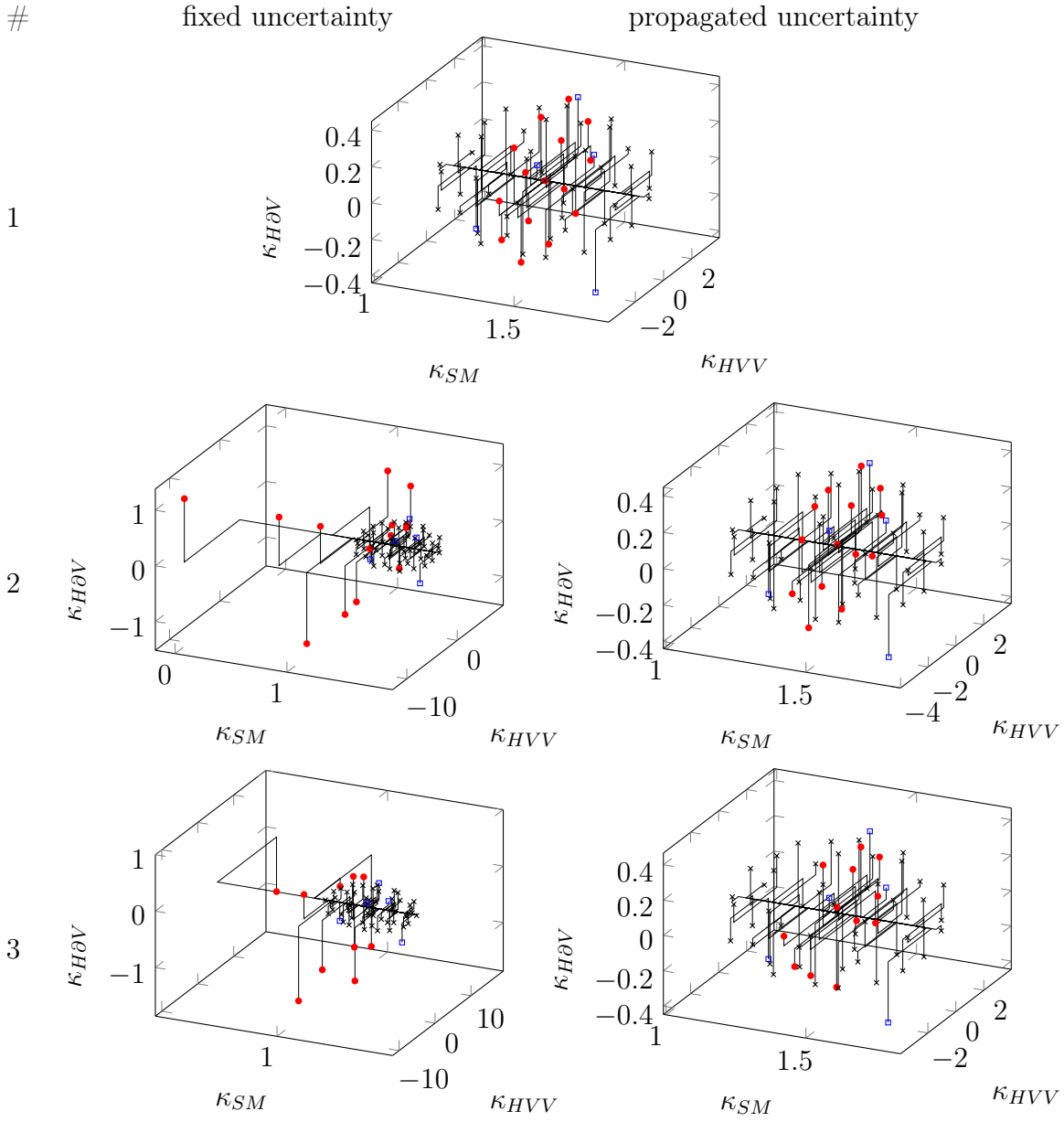


Figure 5.46.: Positions of samples in parameter space throughout the different steps of optimization. Red dots indicate base samples, black crosses indicate benchmark samples, blue squares indicate validation samples. This display shows zoomed axis ranges such that not all base samples are visible.

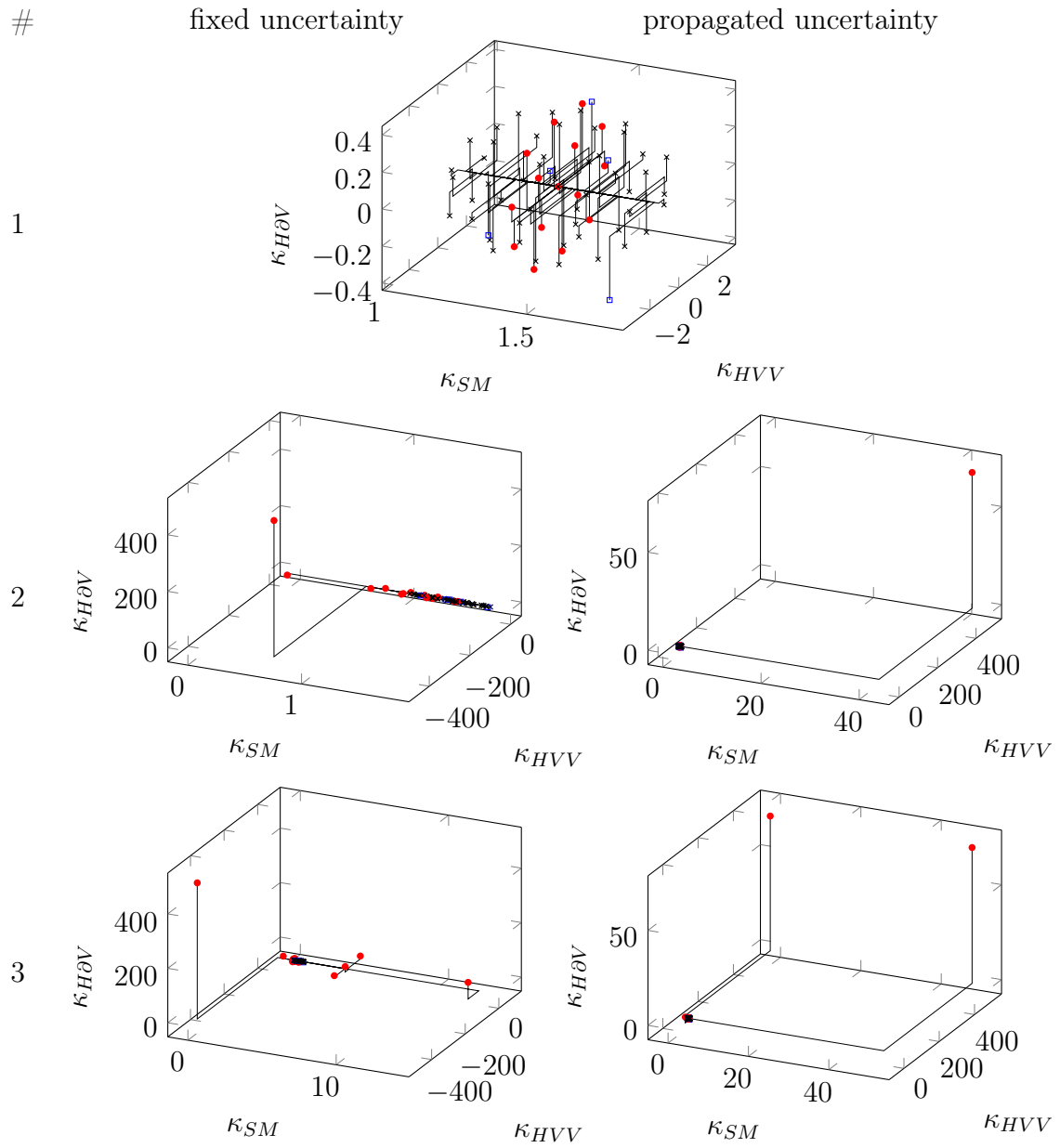


Figure 5.47.: Positions of samples in parameter space throughout the different steps of optimization. Red dots indicate base samples, black crosses indicate benchmark samples, blue squares indicate validation samples. This display shows the full axis ranges.

6 Conclusions and Outlook

> **Not only is the Universe stranger than we think, it is stranger than we can think.**
>
> *Werner Heisenberg*

With the discovery of the Higgs boson with a mass of $m_H = 125.09 \pm 0.24\,\text{GeV}$ in the year 2012 by the ATLAS and CMS collaborations, the last puzzle piece of the Standard Model has fallen into place. Hence, it is natural to assume that this Higgs boson is indeed the boson associated with the generation of masses of the fundamental particles via the Brout-Englert-Higgs mechanism of electroweak symmetry breaking. A thorough experimental test of this hypothesis, however, will require more data and the development of new, sophisticated analysis techniques.

6.1 Summary

Higgs boson decays via electroweak bosons like $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$, but also $H{\to}ZZ^{*}{\to}\ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ and $H{\to}\gamma\gamma$ have been analyzed for pp collisions at $\sqrt{s} = 13\,\text{TeV}$. The analysis in the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay mode includes the first measurement of Higgs boson production via gluon fusion in this decay mode using data collected at $\sqrt{s} = 13\,\text{TeV}$ by the ATLAS detector. State-of-the-art techniques for obtaining predictions and continuous likelihood models for physics measurements are discussed in detail.

6.1.1 $H \to W^{+}W^{-*} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$

Chapter 4 presents a full analysis of the first $5.8\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data, targeting Higgs boson decays via a pair of W bosons into a final state with two charged, different-flavour leptons. All three dominant modes of Higgs boson production, gluon fusion, vector boson fusion and production in association with a W boson were considered and targeted with mutually orthogonal event selections. Highly optimized event selections and sophisticated techniques of background estimation have been adapted, based on the strategy of the analyses conducted on data collected at $\sqrt{s} = 7$ and $8\,\text{TeV}$. Signal strength measurements have been conducted using multi-dimensional likelihood profile methods, employing physics-based quantities as well as multivariate analysis techniques. The measured values of the signal-strength parameters and the observed and expected significances are

$$\mu_{\text{ggF}} = 0.37\,^{+0.39}_{-0.37} \qquad Z_{\text{obs.}} = 1.0\sigma \qquad Z_{\text{exp.}} = 2.8\sigma \tag{6.1.1}$$

$$\mu_{\text{VBF}} = 1.7\,^{+1.1}_{-0.9} \qquad Z_{\text{obs.}} = 1.9\sigma \qquad Z_{\text{exp.}} = 1.2\sigma \tag{6.1.2}$$

$$\mu_{\text{WH}} = 3.2\,^{+4.4}_{-4.2} \qquad Z_{\text{obs.}} = 0.77\sigma \qquad Z_{\text{exp.}} = 0.24\sigma. \tag{6.1.3}$$

A combination of these results assuming the Standard Model expectation for the ratios between the contributions of the individual production modes results in a measurement of a global signal-strength parameter and observed and expected signal significances of

$$\mu = 0.67\,^{+0.34}_{-0.31} \qquad Z_{\text{obs.}} = 2.2\sigma \quad Z_{\text{exp.}} = 3.4\sigma. \tag{6.1.4}$$

The large uncertainties are a result of the relatively small data set. The sensitivity of the analysis is comparable to the one conducted on $\sqrt{s} = 7$ and $8\,\text{TeV}$ when accounting for the smaller integrated luminosity and the larger cross sections of both signal and background processes at $\sqrt{s} = 13\,\text{TeV}$.

6.1.2 Coupling measurements

In Chapter 5, a combined analysis of the $H \to ZZ^{*} \to \ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ and $H \to \gamma\gamma$ decay modes employing $14.8\,\text{fb}^{-1}$ and $13.3\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 13\,\text{TeV}$ finds

$$\mu = 1.13\,^{+0.18}_{-0.17} \qquad Z_{\text{obs.}} = 10\sigma \quad Z_{\text{exp.}} = 8.6\sigma. \tag{6.1.5}$$

Considering multiple models of parametrization, no significant deviation from the Standard Model prediction was found.

In order to facilitate in-depth measurements of the properties of the Higgs boson, new and efficient methods of obtaining predictions under a wide range of possible BSM scenarios need to be developed. State-of-the-art techniques for obtaining

such predictions are discussed, including the computational cost of obtaining accurate predictions using Monte Carlo generators as well as the shortcomings and advantages of different implementations of Matrix Element Reweighting. As measurements are generally based on likelihood formulations, different methods of obtaining continuous likelihood functions from a finite set of samples are discussed.

The novel technique of *Effective Lagrangian Morphing*, which was developed in the context of this thesis, can be used to obtain accurate continuous likelihood models based on the Lagrangian structure of the underlying theory. The technique is derived from first principles, discussed in detail and examined regarding its range of applicability and the uncertainty propagation, carefully considering the advantages of the method as well as the challenges faced during its application. Various validation studies are presented, and methods to further enhance the effectiveness, efficiency and accuracy of the method are introduced and discussed. Finally, a sophisticated algorithm to find the optimal set of base samples is presented.

As a first application of the method of Effective Lagrangian Morphing, a measurement of the CP-even and the CP-odd BSM EFT coupling parameters g_{HVV} and g_{AVV} is presented for the $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$ decay mode, observing compatibility with the Standard Model prediction at the level of 1.8σ and 2.1σ, respectively.

Future applications of the Method of Effective Lagrangian Morphing for an analysis employing the $H\to W^\pm W^{\mp *}\to\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ decay are examined with a particular focus on the VBF category, discussing the interplay with other newly developed techniques like the morphing of multivariate classifiers.

6.2 Outlook

The full data set collected by the ATLAS experiment at $\sqrt{s} = 13\,\text{TeV}$ in the years 2015 and 2016 corresponds to $36.5\,\text{fb}^{-1}$. With this, and the data set to be collected in 2017 and beyond, precision measurements of the properties of the Higgs boson will become feasible. The analysis of Higgs boson decays via $H\to W^\pm W^{\mp *}\to\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ will profit greatly from this increased data set, not only in the VBF and VH categories, where the sensitivity is currently limited by the statistical uncertainty, but also in the gluon fusion category. There, the greater statistical power of the data set will enable the use of more advanced techniques of analysis and background estimation to better control the systematic uncertainties and facilitate precise measurements of differential cross sections in the future.

The novel technique of Effective Lagrangian Morphing will, in connection with other state-of-the-art techniques, enable measurements and derivations of limits for coupling parameters in model-independent effective theories. With the statistical power of the ever increasing data set, strong limits on BSM coupling

parameters or discovery of BSM effects in the Higgs sector will be possible. With a proof-of-concept analysis employing this general approach already conducted in $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ and employing the methods of dimensionality reduction and uncertainty minimization presented here, this method can be applied for a wide variety of physics measurements in the future, deciphering the nature of the Higgs sector and opening a door to what might lie beyond.

A $H \to W^{\pm} W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$

This appendix contains some additional material for Chapter 4 that is provided for the sake of completeness.

A.1 Treatment of Systematic Uncertainties

Section 4.5.5 details the treatment of systematic uncertainty parameters in the likelihood fit used in the ggF $H \to W^{\pm} W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$ analysis. The effects of this treatment are shown in Fig. A.1 (and Fig. A.2). Here, a colour code indicates the treatment of the individual nuisance parameter, while the printed number indicates the effect of that parameter in any given region for any given sample. As clearly visible, most parameters can be effectively purged from the fit model using the approach outlined in Section 4.5.5.

Figure A.1.: Treatment of systematic uncertainties in the ggF analysis. The columns indicate individual uncertainty parameters, the rows indicate categories (signal or control regions) and samples contributing to the estimate. The cell entries indicate the relative effects the $\pm 1\sigma$ variations have on the total yield of that sample in that category. The cell colours encode the treatment of the corresponding parameter.

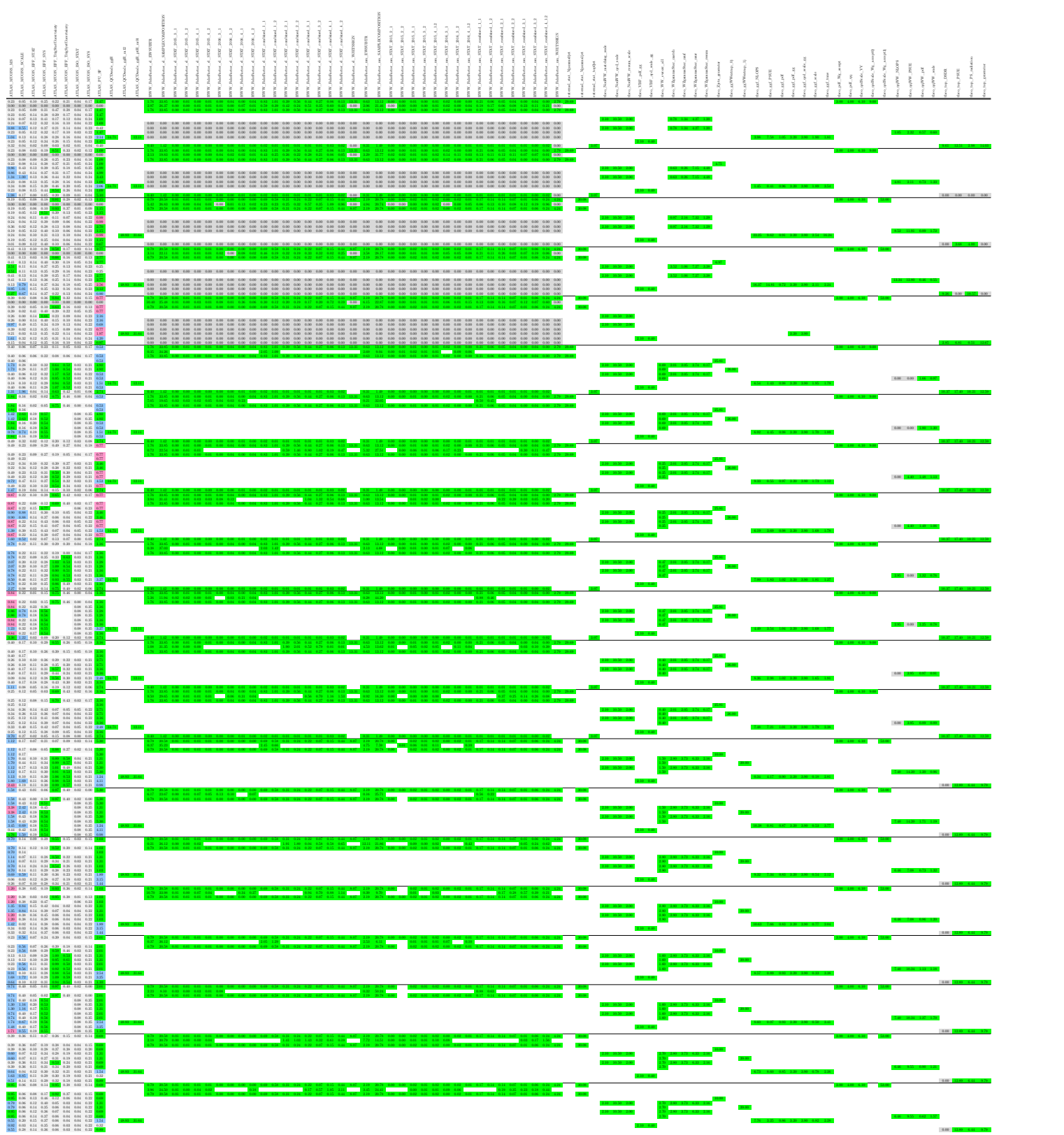


Figure A.2.: (cont.) Light (dark) gray indicates uncertainties neglected because they have little (no) effect on the yield. Blue indicates uncertainties dropped because the statistical power of the variational template is too low to conclusively determine the effect. Pink indicates cases where the templates have been symmetrized. Green finally indicates that the systematic is used as-is.

B Coupling Measurements

This appendix contains some additional material for Chapter 5 that is provided for the sake of completeness.

B.1 The Gershgorin Circle Theorem

Considering a square matrix A of size $N \times N$, choose one of its eigenvalues λ and consider the corresonding eigenvector $\vec{x}$, *i. e.*,

$$\sum_j A_{ij} x_j = \lambda x_i \qquad \forall i \in \{1, \ldots, N\} \tag{B.1.1}$$

$$\sum_{j \neq i} A_{ij} x_j = \lambda x_i - A_{ii} x_i \qquad \forall i \in \{1, \ldots, N\} \tag{B.1.2}$$

Picking i such that x_i is the largest value in $\vec{x}$, and recognizing that since $\vec{x}$ is an eigenvector of A, $|x_i| \neq 0$, one can divide by x_i and take the absolute value to find

$$|\lambda - A_{ii}| = \left| \frac{\sum_{j \neq i} A_{ij} x_j}{x_i} \right| \leq \sum_{j \neq i} \left| \frac{A_{ij} x_j}{x_i} \right| \leq \sum_{j \neq i} |A_{ij}| \stackrel{\text{def}}{=} R_i, \tag{B.1.3}$$

where the last inequality is valid because x_i is explicitly the largest value of $\vec{x}$, and thus

$$\frac{|x_j|}{|x_i|} \leq 1. \tag{B.1.4}$$

Here, R_i is the sum over all entries in row i, excluding the diagonal entry A_{ii}. Equation B.1.3 implies that every eigenvalue lies within an interval

$$\lambda \in [A_{ii} - R_i, A_{ii} + R_i] \tag{B.1.5}$$

for some row i. This is also true for matrices with complex entries, where the interval takes the form of a circle with radius R_i around A_{ii}, called a *Gershgorin disc* [185].

The theorem also applies to the columns instead of rows, which can be easily seen by replacing the matrix A with its transpose.

B.2 Condition Numbers and Uncertainties

For an invertible system of linear equations

$$Ax = b, \tag{B.2.1}$$

assume that the values x carry an uncertainty Δx. Then,

$$A\left(x + \Delta x\right) = Ax + A\Delta x = b + \Delta b, \tag{B.2.2}$$

where

$$\Delta b = A\Delta x \tag{B.2.3}$$

is the uncertainty on the outcome b.

Taking the norm of Eq. B.2.1, one obtains

$$\|b\| = \|Ax\| \leq \|A\| \, \|x\| \,, \tag{B.2.4}$$

which results in

$$\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \tag{B.2.5}$$

Empoying Eq. B.2.3, one finds

$$\delta x = A^{-1}\Delta b.. \tag{B.2.6}$$

By multiplying Eq. B.2.5 and Eq. B.2.6, one obtains

$$\frac{\|\Delta x\|}{\|x\|} \leq C(A) \cdot \frac{\|\Delta b\|}{\|b\|}, \tag{B.2.7}$$

where

$$C(A) = \left\|A^{-1}\right\| \cdot \|A\| \tag{B.2.8}$$

is the condition number of the matrix A [186].

B.3 Intuitive Misapprehensions

Intuitive appraoches to the problem of finding a good set of base samples for any concrete application of Effective Lagrangian Morphing often lead to wrong results. This section contains a set of "blind alleys" of argumentation to justify this claim, and to help develop the problem more clearly.

Proximity to base samples

Effective Lagrangian Morphing is an interpolation method, and as such, is most realiable directly at the base samples. It is thus reasonable to distribute the base samples evenly across the parameter space of interest.

Effective Lagrangian Morphing is no classical interpolation method like the vertical morphing presented in Section 5.2.3 or the moment morphing presented in Section 5.2.3. Instead, as demonstrated in Section 5.3.6, the result can even *improve* when the base samples are chosen such that they *do not coincide* with the region of interest. This can also be seen in Fig. 5.24 in Section 5.3.4.

Randomness and invertibility

Effective Lagrangian Morphing is a technique relying on a linear system of equations. Thus, the error propagation is primarily determined by the stability of the system, and thus by the determinant of the matrix encoding the system. As randomly generated matrices are virtually always invertible, randomly choosing a basis is a promising approach.

While it is true that randomly generated matrices are almost always invertible, the morphing matrices are not truly random:. The number of free parameters is $n \times N$, with n denoting the number of coupling parameters and N denoting the number of base samples. The number of entries in the morphing matrix are $N \times N$. Here, n is always smaller (and usually much smaller) than N, as argued in Section 5.3.3 and shown in Fig. 5.19. Additionally, the deteriminant (or invertibility) of the matrix is only related to the numeric stability of the matrix, which needs not coincide with how uncertainties on the base samples propagate through the method.

Base condition

The accurate measure for the quality of a basis is the condition number of the associate matrix, as this governs the uncertainty propagation.

If the base samples are all of equal size, their relative uncertainties are the same. Under these circumstances, it is actually the value of the physical observable ξ_i, usually a cross section, that determines the absolute value of the uncertainty. In a sense, the "weight" with which the relative uncertainty on ξ_i propagates to the final prediction is the product of $w_i\xi_i$, as argued in Section 5.3.6. The condition number of the matrix, however, is agnostic of the ξ_i and only takes into account the w_i. Hence, the optimal basis needs to be determined by taking into account both, the values of the physical observables ξ_i *and* the entries of the morphing matrix, wheras the condition number only takes into account the former. Putting

it differently, the condition of the matrix *does* govern the propagation of the *total* uncertainties of the ξ_i – but these uncertainties and the matrix itself both ultimately result from choice of the same parameters. Thus, only choosing a matrix with a small condition number may not result in a good basis set, as the careful blance established for the w_i might be easily spoiled by an imbalance in the ξ_i, which result from the same choice of parameters.

B.4 Hypergeometry

Let S be an n-dimensional sphere of radius R centered around the origin. Let H be an n-dimensional hypercube with edges of length $2A$ be centered around the origin. What is the volume $H \cap S$ shared between the two?

Consider the following three cases:

1. If $R < A$, the sphere is fully contained in the hypercube. The volume of the intersection is then simply the volume of the sphere itself, which is

$$V_n(R) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} R^n.$$

2. If $R > \sqrt{n}A$, the diameter of the sphere is larger than the diagonal of the cube. Thus, the hypercube is fully contained in the sphere. The volume is simply the Volume of the cube, which is

$$V_n(A) = (2A)^n$$

3. If $R > A$, but $R < \sqrt{n}A$, caps of the sphere start to punsh through the surfaces of the cube. Each cap will be centered in such a surface. These caps are non-overlapping exactly if their opening angle ϕ is smaller than $\pi/4$, or $\sin^2\phi < 1/2$. If the base of the cap has radius x, it follows that

$$\sin^2\phi = \frac{x^2}{R^2} = \frac{R^2 - A^2}{R^2}.$$

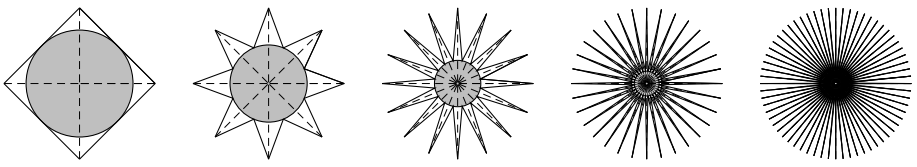


Figure B.1.: Spiked star representation of the inscription of an n-dimensional hypersphere in an n-dimensional hypercube for $n = 2, 3, 4, 5$. Clearly, the fraction of volume in the sphere (gray) becomes smaller for increasing dimensionality.

Thus, the caps are overlapping exactly if $R < \sqrt{2}A$. However, if the caps are overlapping, then the corners of the cube will be completely disjoint. Thus, one can distinguish the following cases.

a) If $R > A$, but $R < \sqrt{2}A$, the caps are disjoint. The volume of the intersection is then the volume of the sphere minus the volume of the $n \cdot 2^{n-1}$ caps [190], that is,

$$V_n(R, A) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} R^n \left(1 - n \cdot 2^{n-1} \cdot \frac{1}{2} I_{\frac{R^2-A^2}{R^2}}\left(\frac{n+1}{2}, \frac{1}{2}\right)\right)$$

where $I_{\sin^2\phi}(a, b)$ is the incomplete beta function.

b) If $R > \sqrt{2}A$, but $R < \sqrt{n}A$, then the caps are overlapping. While for $n = 3$ this implies that the corners of the cube would be disjoint and could be computed separately, this is no longer true for higher dimensions.

In the general case of n dimensions, there are $n-1$ types of extended boundaries of the hypercube, in which $1, 2, \ldots, n-1$ coordinates are maxed-out at $\pm A$. While there are only edges and faces in 3 dimensions, the hypersphere starts intersecting the boundaries of type j when its radius reaches $A\sqrt{j}$, and only fully contains them when its radius exceeds $A\sqrt{n}$, leading to a new piecewise defined section of the function in any new dimension.

It is, however, possible to derive a recursive integral formula. Let $V_n(R, A)$ be the volume of the intersection in n dimensions. Then

$$V_n(R, A) = \int_{x_1=-A}^{+A} \int_{x_2=-A}^{+A} \cdots \int_{x_n=-A}^{+A} \mathcal{I}\left(\sum_{i=1}^{n} x_i^2 < R^2\right) dx_1 dx_2 \cdots dx_n, \quad \text{(B.4.1)}$$

where

$$\mathcal{I}(\Phi) = \begin{cases} 1 \text{ if } \Phi \\ 0 \text{ otherwise} \end{cases} \quad \text{(B.4.2)}$$

The integrand is non-zero only when $|x_1| < R$, in which case

$$\mathcal{I}\left(\sum_{i=1}^{n} x_i^2 < R^2\right) = \mathcal{I}\left(\sum_{i=2}^{n} x_i^2 < R^2 - x_1^2\right), \quad \text{(B.4.3)}$$

such that

$$V_n(R, A) = \int_{-\min(A,R)}^{+\min(A,R)} V_{n-1}\left(\sqrt{R^2 - x^2}\right) dx. \quad \text{(B.4.4)}$$

The base of the recursion is

$$V_0(R, A) = 1 \tag{B.4.5}$$

or, if the 0-dimensional volume seems too contrived,

$$V_1(R, A) = 2 \min(A, R) . \tag{B.4.6}$$

As a closed form solution of this recursive integral formula is not feasible at this point, a mere numeric estimation using Monte Carlo integration methods will suffice.

List of Figures

List of Tables

Bibliography

[1] S. Berryman, "Ancient Atomism", in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, ed. Metaphysics Research Lab, Stanford University, 2016.

[2] "Random House Webster's College Dictionary", 2010.

[3] J. W. von Goethe, *Faust: Eine Tragödie. Erster Theil.* 1808.

[4] A. Lightman, *The Discoveries: Great Breakthroughs in 20th-Century Science*. Knopf Canada, 2010. ISBN 978-0-307-36986-4.

[5] M. K. E. L. Planck, "Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum", *Verhandlungen der Deutschen physikalischen Gesellschaft* **17** no. 2, (1900) 237–45.

[6] A. Einstein, "Zur Elektrodynamik bewegter Körper", *Annalen der Physik* **322** no. 10, (1905) 891–921.

[7] F. B. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Physical Review Letters* **13** (Aug., 1964) 321–323.

[8] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Physical Review Letters* **13** (Oct., 1964) 508–509.

[9] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.* **145** (May, 1966) 1156–1163.

[10] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Physical Review Letters* **13** (Nov., 1964) 585–587.

[11] Nobel Media, *The Nobel Prize in Physics 2013*. `http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/`.

[12] L. Lederman and D. Teresi, *The God Particle: If the Universe Is the Answer, What Is the Question?* Houghton Mifflin Harcourt, 2006. ISBN 978-0-547-52462-7.

[13] **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]`.

[14] **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].

[15] **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].

[16] I. Newton, F. Jacquier, and T. Seur, *Philosophiae naturalis principia mathematica*. No. 3. Typis Barrillot & Filii, 1742.

[17] **The ATLAS Collaboration**, "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector", *Eur. Phys. J.* **C75** no. 10, (2015) 476, arXiv:1506.05669 [hep-ex]. [Erratum: Eur. Phys. J.C76,no.3,152(2016)].

[18] **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** no. 1, (2015) 012004, arXiv:1411.3441 [hep-ex].

[19] **The LHC Higgs Cross Section Working Group**, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", arXiv:1610.07922 [hep-ph].

[20] **The ATLAS Collaboration**, "A morphing technique for signal modelling in a multidimensional space of coupling parameters", ATL-PHYS-PUB-2015-047, CERN, Geneva, Nov, 2015. CDS:2066980.

[21] **The ATLAS Collaboration**, "Measurements of the Higgs boson production cross section via Vector Boson Fusion and associated WH production in the $WW^* \to \ell\nu\ell\nu$ decay mode with the ATLAS detector at $\sqrt{s} = 13$ TeV", ATLAS-CONF-2016-112, CERN, Geneva, Nov, 2016. CDS:2231811.

[22] **The ATLAS Collaboration**, "Study of the Higgs boson properties and search for high-mass scalar resonances in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector", ATLAS-CONF-2016-079, CERN, Geneva, Aug, 2016. CDS:2206253.

[23] **The ATLAS Collaboration**, "Combined measurements of the Higgs boson production and decay rates in $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ final states using pp collision data at $\sqrt{s} = 13$ TeV in the ATLAS experiment", ATLAS-CONF-2016-081, CERN, Geneva, Aug, 2016. CDS:2206272.

[24] D. Griffiths, *Introduction to Elementary Particles*. Wiley-VCH, 2004. ISBN 978-3-527-40601-2.

[25] F. Halzen and A. D. Martin, *Quarks and Leptons – An Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984. ISBN 978-0-471-88741-6.

[26] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2013. ISBN 978-1-107-03473-0.

[27] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Westview Press Reading, 1995. ISBN 978-0-201-50397-5.

[28] J. F. Donoghue, E. Golowich, and B. R. Holstein, *Dynamics of the Standard Model*. Cambridge University Press, 1992. ISBN 978-0-511-52437-0.

[29] S. Mandelstam, "Determination of the Pion-Nucleon Scattering Amplitude from Dispersion Relations and Unitarity. General Theory", *Phys. Rev.* **112** (Nov, 1958) 1344–1360, APS:10.1103/PhysRev.112.1344.

[30] David Galbraith and Carsten Burgard, *Standard Model of the Standard Model*. http://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model/.

[31] **Particle Data Group**, K. A. Olive *et al.*, "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001. http://pdg.lbl.gov.

[32] E. Noether, "Invariant variation problems", *Transport Theory and Statistical Physics* **1** (Jan., 1971) 186–207, arXiv:physics/0503066.

[33] J. Goldstone, A. Salam, and S. Weinberg, "Broken Symmetries", *Phys. Rev.* **127** (Aug, 1962) 965–970, APS:10.1103/PhysRev.127.965.

[34] **The LHCb Collaboration**, "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays", *Phys. Rev. Lett.* **115** (2015) 072001, arXiv:1507.03414 [hep-ex].

[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] E. Fermi, "Versuch einer Theorie der β-Strahlen. I", *Zeitschrift fur Physik* **88** (Mar., 1934) 161–177.

[37] S. Weinberg, "Baryon- and Lepton-Nonconserving Processes", *Phys. Rev. Lett.* **43** (Nov, 1979) 1566–1570, APS:10.1103/PhysRevLett.43.1566.

[38] W. Buchmuller and D. Wyler, "Effective Lagrangian Analysis of New Interactions and Flavor Conservation", *Nucl. Phys.* **B268** (1986) 621–653, CDS:163116.

[39] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, "Dimension-Six Terms in the Standard Model Lagrangian", *JHEP* **10** (2010) 085, arXiv:1008.4884 [hep-ph].

[40] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, "The Strongly-Interacting Light Higgs", *JHEP* **06** (2007) 045, arXiv:hep-ph/0703164 [hep-ph].

[41] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, "Low energy effects of new interactions in the electroweak boson sector", *Phys. Rev. D* **48** (Sep, 1993) 2182–2203, APS:10.1103/PhysRevD.48.2182.

[42] M. Böhm, A. Denner, and H. Joos, *Gauge Theories of the Strong and Electroweak Interaction*. Vieweg+Teubner Verlag, 2001. ISBN 978-3-519-23045-8.

[43] M. Veltman, *Diagrammatica: The Path to Feynman Diagrams*. Cambridge University Press, 1994. ISBN 978-0-521-45692-0.

[44] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, "FeynRules 2.0 – A complete toolbox for tree-level phenomenology", *Comput. Phys. Commun.* **185** (2014) 2250–2300, arXiv:1310.1921 [hep-ph].

[45] **LHC Higgs Cross Section Working Group**, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", arXiv:1307.1347 [hep-ph].

[46] **LHC Higgs Cross Section Working Group**, A. David, A. Denner, M. Duehrssen, M. Grazzini, C. Grojean, G. Passarino, M. Schumacher, M. Spira, G. Weiglein, and M. Zanetti, "LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle", arXiv:1209.0040 [hep-ph].

[47] P. Artoisenet *et al.*, "A framework for Higgs characterisation", *JHEP* **11** (2013) 043, arXiv:1306.6464 [hep-ph].

[48] 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)].

[49] LHC Higgs Cross Section Working Group, *Picture Gallery*. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGCrossSectionsFigures.

[50] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, and B. Mistlberger, "High precision determination of the gluon fusion Higgs boson cross-section at the LHC", *JHEP* **05** (2016) 058, arXiv:1602.00695 [hep-ph].

[51] P. Bolzoni, M. Zaro, F. Maltoni, and S.-O. Moch, "Higgs production at NNLO in QCD: The VBF channel", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 314–319, arXiv:1006.2323 [hep-ph].

[52] **The ATLAS Collaboration**, "Search for Higgs boson production via weak boson fusion and decaying to $b\bar{b}$ in association with a high-energy photon in the ATLAS detector", ATLAS-CONF-2016-063, CERN, Geneva, Aug, 2016. CDS:2206201.

[53] **The ATLAS Collaboration**, "Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a $b\bar{b}$ pair in pp collisions at 13 TeV using the ATLAS detector", ATLAS-CONF-2016-091, CERN, Geneva, Aug, 2016. CDS:2206813.

[54] **The ATLAS Collaboration**, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", ATLAS-CONF-2016-080, CERN, Geneva, Aug, 2016. CDS:2206255.

[55] **The CMS Collaboration**, "Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$", *Phys. Rev.* **D92** no. 3, (2015) 032008, arXiv:1506.01010 [hep-ex].

[56] **The CMS Collaboration**, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev.* **D89** no. 1, (2014) 012003, arXiv:1310.3687 [hep-ex].

[57] **The CMS Collaboration**, "Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method", *Eur. Phys. J.* **C75** no. 6, (2015) 251, arXiv:1502.02485 [hep-ex].

[58] **The ATLAS Collaboration**, "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector", *JHEP* **04** (2015) 117, arXiv:1501.04943 [hep-ex].

[59] **The CMS Collaboration**, "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons", *JHEP* **05** (2014) 104, arXiv:1401.5041 [hep-ex].

[60] **The ATLAS Collaboration**, "Search for Higgs bosons decaying into di-muon in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", ATLAS-CONF-2016-041, CERN, Geneva, Aug, 2016. CDS:2206079.

[61] **The ATLAS Collaboration**, "Search for the Standard Model Higgs boson in the $H \to Z\gamma$ decay mode with pp collisions at $\sqrt{s} = 7$ and 8 TeV", ATLAS-CONF-2013-009, CERN, Geneva, Mar, 2013. CDS:1523683.

[62] **The CMS Collaboration**, "Search for a standard model-like Higgs boson in the $\mu^+\mu^-$ and e^+e^- decay channels at the LHC", *Phys. Lett.* **B744** (2015) 184–207, arXiv:1410.6679 [hep-ex].

[63] **The ATLAS 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].

[64] **The ATLAS Collaboration**, *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999. CDS:391176.

[65] **The ATLAS Collaboration**, *ATLAS detector and physics performance: Technical Design Report, 2.* Technical Design Report ATLAS. CERN, Geneva, 1999. CDS:391177.

[66] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report, 1.* Technical Design Report LHC. CERN, Geneva, 2004. CDS:782076.

[67] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report, 2.* Technical Design Report LHC. CERN, Geneva, 2004. CDS:815187.

[68] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report, 3.* Technical Design Report LHC. CERN, Geneva, 2004. CDS:823808.

[69] **The ATLAS Collaboration**, "The ATLAS Experiment at the CERN Large Hadron Collider", *Journal of Instrumentation* **3** no. 08, (2008) S08003, IOP:10.1088/1748-0221/3/08/S08003.

[70] L. Evans and P. Bryant, "LHC Machine", *Journal of Instrumentation* **3** no. 08, (2008) S08001, IOP:10.1088/1748-0221/3/08/S08001.

[71] M. Capeans, G. Darbo, K. Einsweiler, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, "ATLAS Insertable B-Layer Technical Design Report", CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Sep, 2010. CDS:1291633.

[72] CPAN, *El Gran Colisionador de Hadrones.* https://www.i-cpan.es/es/content/el-gran-colisionador-de-hadrones-lhc.

[73] CERN Low Power Converter Section, *The CERN accelerator complex.* http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/pagesources/Cern-Accelerator-Complex.jpg.

[74] **The CMS Collaboration**, "The CMS experiment at the CERN LHC", *Journal of Instrumentation* **3** (2008) S08004, IOP:10.1088/1748-0221/3/08/S08004.

[75] **The LHCb Collaboration**, "The LHCb Detector at the LHC", *Journal of Instrumentation* **3** no. 08, (2008) S08005, IOP:10.1088/1748-0221/3/08/S08005.

[76] **The ALICE Collaboration**, "The ALICE experiment at the CERN LHC", *Journal of Instrumentation* **3** (2008) S08002, IOP:10.1088/1748-0221/3/08/S08002.

[77] **The TOTEM Collaboration**, "The TOTEM Experiment at the CERN Large Hadron Collider", *Journal of Instrumentation* **3** no. 08, (2008) S08007, IOP:10.1088/1748-0221/3/08/S08007.

[78] **The LHCf Collaboration**, "The LHCf detector at the CERN Large Hadron Collider", *Journal of Instrumentation* **3** no. 08, (2008) S08006, IOP:10.1088/1748-0221/3/08/S08006.

[79] **The MoEDAL Collaboration**, "The MoEDAL Experiment at the LHC – a New Light on the Terascale Frontier", *Journal of Physics: Conference Series* **631** no. 1, (2015) 012014, IOP:10.1088/1742-6596/631/1/012014.

[80] S. Myers, *The LEP Collider, from design to approval and commissioning.* John Adams' Lecture. CERN, Geneva, 1991. CDS:226776.

[81] "Design Report Tevatron 1 project", FERMILAB-DESIGN-1984-01, FERMILAB, 1984. CDS:1478620.

[82] F. Bordry, *LHC smashes targets for 2016 run.* https://home.cern/about/opinion/2016/11/lhc-smashes-targets-2016-run-0.

[83] The ATLAS Collaboration, *The ATLAS Collaboration.* http://atlas.cern/discover/collaboration.

[84] L. Barranco Navarro, "Alignment of the ATLAS Inner Detector in the LHC Run 2", ATL-PHYS-PROC-2015-190, CERN, Geneva, Dec, 2015. CDS:2114708.

[85] **The ATLAS Collaboration**, "Particle Identification Performance of the ATLAS Transition Radiation Tracker", ATLAS-CONF-2011-128, CERN, Geneva, Sep, 2011. CDS:1383793.

[86] **The ATLAS Collaboration**, "Performance of b-Jet Identification in the ATLAS Experiment", *JINST* **11** no. 04, (2016) P04008, arXiv:1512.01094 [hep-ex].

[87] **The ATLAS Collaboration**, "The ATLAS Inner Detector commissioning and calibration", *Eur. Phys. J.* **C70** (2010) 787–821, arXiv:1004.5293 [physics.ins-det].

[88] **The ATLAS Collaboration**, "A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run 1 pp-collision data with the ATLAS detector", arXiv:1607.08842 [hep-ex].

[89] **The ATLAS Collaboration**, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data", *Eur. Phys. J.* **C74** no. 10, (2014) 3071, arXiv:1407.5063 [hep-ex].

[90] **The ATLAS Collaboration**, "Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector", arXiv:1609.09324 [hep-ex].

[91] **The ATLAS Collaboration**, "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV", *Eur. Phys. J.* **C76** no. 5, (2016) 292, arXiv:1603.05598 [hep-ex].

[92] J. Stirling, *Parton Luminosity and Cross Section Plots*. http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html.

[93] The ATLAS Collaboration, *Public Luminosity Results for Run 2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[94] P. Czodrowski, "The ATLAS Trigger System: Ready for Run 2", CDS:2039543.

[95] A. Salzburger, "Optimisation of the ATLAS Track Reconstruction Software for Run 2", *Journal of Physics: Conference Series* **664** no. 7, (2015) 072042, IOP:10.1088/1742-6596/664/7/072042.

[96] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", *J. Phys.* **G34** (2007) N343, CDS:1027031.

[97] G. Borissov, D. Casper, K. Grimm, S. P. Griso, L. E. Pedersen, K. Prokofiev, M. Rudolph, and A. Wharton, "ATLAS strategy for primary vertex reconstruction during Run 2 of the LHC", *Journal of Physics: Conference Series* **664** no. 7, (2015) 072041, IOP:10.1088/1742-6596/664/7/072041.

[98] **The ATLAS Collaboration**, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1", arXiv:1603.02934 [hep-ex].

[99] **The ATLAS Collaboration**, "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data", ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016. CDS:2157687.

[100] **The ATLAS Collaboration**, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung", ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. CDS:1449796.

[101] **The ATLAS Collaboration**, "Electron identification measurements in ATLAS using $\sqrt{s} = 13\,\text{TeV}$ data with 50 ns bunch spacing", ATL-PHYS-PUB-2015-041, CERN, Geneva, Sep, 2015. CDS:2048202.

[102] **The ATLAS Collaboration**, H. Herde, "Muon reconstruction performance in ATLAS at Run 2", ATL-PHYS-PROC-2015-121, CERN, Geneva, Oct, 2015. CDS:2059849.

[103] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[104] **The ATLAS Collaboration**, "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7\,\text{TeV}$", *Eur. Phys. J.* **C73** no. 3, (2013) 2304, arXiv:1112.6426 [hep-ex].

[105] **The ATLAS Collaboration**, "Tagging and suppression of pileup jets with the ATLAS detector", ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. CDS:1700870.

[106] **The ATLAS Collaboration**, "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run", ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016. CDS:2160731.

[107] **The ATLAS Collaboration**, "Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s} = 13\,\text{TeV}$", ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015. CDS:2037904.

[108] **The ATLAS Collaboration**, "Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13\,\text{TeV}$", ATL-PHYS-PUB-2015-023, CERN, Geneva, Jul, 2015. CDS:2037700.

[109] **The GEANT4 Collaboration**, S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[110] **The CMS Collaboration**, "Dimuon spectrum 2016", CMS-DP-2016-059. CDS:2212114.

[111] The ATLAS Collaboration, *Summary plots from the ATLAS Standard Model physics group*. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/.

[112] **The ATLAS Collaboration**, "Measurement of W^+W^- production in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector and limits on anomalous WWZ and $WW\gamma$ couplings", *Phys. Rev. D* **87** (Jun, 2013) 112001, arXiv:1210.2979.

[113] **The ATLAS Collaboration**, "Measurement of W^+W^- production in association with one jet in proton–proton collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector", *Phys. Lett.* **B763** (2016) 114–133, arXiv:1608.03086 [hep-ex].

[114] **The ATLAS Collaboration**, "Measurement of the W^+W^- production cross section in pp collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS experiment", ATLAS-CONF-2016-090, CERN, Geneva, Aug, 2016. CDS:2206811.

[115] **The ATLAS Collaboration**, "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector", *Eur. Phys. J.* **C74** no. 10, (2014) 3109, arXiv:1406.5375 [hep-ex]. [Addendum: Eur. Phys. J.C76,no.11,642(2016)].

[116] **The ATLAS Collaboration**, "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *Phys. Lett.* **B761** (2016) 136–157, arXiv:1606.02699 [hep-ex].

[117] J. Ellis and T. You, "Updated global analysis of Higgs couplings", *Journal of High Energy Physics* **2013** no. 6, (2013) 103, arXiv:1303.3879 [hep-ph].

[118] **The ATLAS Collaboration**, "Search for WW/WZ resonance production in the $\ell\nu qq$ final state at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector at the LHC", ATLAS-CONF-2015-075, CERN, Geneva, Dec, 2015. CDS:2114847.

[119] **The ATLAS Collaboration**, "Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector", *Phys. Rev.* **D92** no. 1, (2015) 012006, arXiv:1412.2641 [hep-ex].

[120] The ATLAS Collaboration, *Public e/γ Trigger Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults.

[121] The ATLAS Collaboration, *Public μ Trigger Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonTriggerPublicResults.

[122] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, "Higgs Boson Gluon-Fusion Production in QCD at Three Loops", *Phys. Rev. Lett.* **114** (2015) 212001, arXiv:1503.06056 [hep-ph].

[123] T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, and L. Tancredi, "W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD", *Phys. Rev. Lett.* **113** no. 21, (2014) 212001, arXiv:1408.5243 [hep-ph].

[124] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, "QCD corrections to W^+W^- production through gluon fusion", *Phys. Lett.* **B754** (2016) 275–280, arXiv:1511.08617 [hep-ph].

[125] T. Sjostrand, 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].

[126] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[127] 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)", *Journal of High Energy Physics* **2001** no. 01, (2001) 010, IOP:10.1088/1126-6708/2001/01/010.

[128] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, "Multiparton interactions in photoproduction at HERA", *Z. Phys.* **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph].

[129] P. Nason and C. Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG", *JHEP* **02** (2010) 037, arXiv:0911.5299 [hep-ph].

[130] 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].

[131] R. D. Ball *et al.*, "Parton distributions with LHC data", *Nucl. Phys.* **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[132] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W.-K. Tung, and C. P. Yuan, "Implications of CTEQ global analysis for collider observables", *Phys. Rev.* **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[133] **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].

[134] **The ATLAS Collaboration**, "ATLAS Run 1 Pythia8 tunes", ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014. CDS:1966419.

[135] 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].

[136] S. Schumann and F. Krauss, "A Parton shower algorithm based on Catani-Seymour dipole factorisation", *JHEP* **03** (2008) 038, arXiv:0709.1027 [hep-ph].

[137] K. Melnikov and G. Zanderighi, "W+3 jet production at the LHC as a signal or background", *Phys. Rev.* **D81** (2010) 074025, arXiv:0910.3671 [hep-ph].

[138] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes", *Phys. Rev.* **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[139] D. J. Lange, "The EvtGen particle decay simulation package", *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[140] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014) 2930, arXiv:1112.5675 [hep-ph].

[141] **The ATLAS Collaboration**, "The ATLAS Simulation Infrastructure", *Eur. Phys. J.* **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[142] **The ATLAS Collaboration**, "Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics", arXiv:0901.0512 [hep-ex].

[143] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, "A method for identifying $H \to \tau\tau \to e^{+-}\mu^{-+}$ + missing p_{T} at the CERN LHC", *Phys.Rev.* **D61** (2000) 093005, arXiv:hep-ph/9911385 [hep-ph].

[144] **The ATLAS Collaboration**, "Muon reconstruction performance in early $\sqrt{s} = 13\,\text{TeV}$ data", ATL-PHYS-PUB-2015-037, CERN, Geneva, Aug, 2015. CDS:2047831.

[145] **The ATLAS Collaboration**, M. Goblirsch-Kolb, "Muon Reconstruction Efficiency, Momentum Scale and Resolution in *pp* Collisions at 8 TeV with ATLAS", in *Proceedings, 2nd Conference on Large Hadron Collider Physics Conference (LHCP 2014): New York, USA, June 2-7, 2014*. 2014. arXiv:1408.7086 [hep-ex].

[146] "Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13\,\text{TeV}$", ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul, 2015. CDS:2037613.

[147] **The ATLAS Collaboration**, "Improved luminosity determination in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS detector at the LHC", *Eur. Phys. J.* **C73** no. 8, (2013) 2518, arXiv:1302.4393 [hep-ex].

[148] S. van der Meer, "Calibration of the effective beam height in the ISR", CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. CDS:296752.

[149] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 release note", *Eur. Phys. J.* **C76** no. 4, (2016) 196, arXiv:1512.01178 [hep-ph].

[150] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++", *Eur. Phys. J.* **C72** (2012) 2225, arXiv:1206.0041 [hep-ph].

[151] 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].

[152] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, A. Schalicke, and S. Schumann, "Matching parton showers and matrix elements", in *HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A*, pp. 288–289. 2005. arXiv:hep-ph/0602031 [hep-ph]. https://inspirehep.net/record/709818/files/arXiv:hep-ph_0602031.pdf.

[153] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

[154] "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", *Journal of Computer and System Sciences* **55** no. 1, (1997) 119 – 139.

[155] J. H. Friedman, "Stochastic gradient boosting", *Comput. Stat. Data Anal.* **38** no. 4, (2002) 367–378.

[156] **The ATLAS Collaboration**, "Simulation of top quark production for the ATLAS experiment at $\sqrt{s} = 13\,TeV$", ATL-PHYS-PUB-2016-004, CERN, Geneva, Jan, 2016. CDS:2120417.

[157] 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].

[158] K. Arnold *et al.*, "VBFNLO: A Parton Level Monte Carlo for Processes with Electroweak Bosons – Manual for Version 2.5.0", arXiv:1107.4038 [hep-ph].

[159] J. Baglio *et al.*, "Release Note – VBFNLO 2.7.0", arXiv:1404.3940 [hep-ph].

[160] **The CMS Collaboration**, "First results on Higgs to WW at $\sqrt{s} = 13$ TeV", CMS-PAS-HIG-15-003, CERN, Geneva, 2016. CDS:2161793.

[161] **The ATLAS Collaboration**, "Measurement of fiducial, differential and production cross sections in the $H \to \gamma\gamma$ decay channel with 13.3 fb^{-1} of 13 TeV proton-proton collision data with the ATLAS detector", ATLAS-CONF-2016-067, CERN, Geneva, Aug, 2016. CDS:2206210.

[162] R. Kleiss and W. Stirling, "Tagging the Higgs", *Physics Letters B* **200** no. 1–2, (1988) 193–199.

[163] W. Verkerke and D. Kirkby, "The `RooFit` toolkit for data modeling", arXiv:physics/0306116.

[164] **The ROOT Collaboration**, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, "`HistFactory`: A tool for creating statistical models for use with `RooFit` and `RooStats`", CERN-OPEN-2012-016, New York University, New York, Jan, 2012. CDS:1456844.

[165] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, "The `RooStats` project", in *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, p. 57. 2010. arXiv:1009.1003 [physics.data-an].

[166] M. Botje *et al.*, "The PDF4LHC Working Group Interim Recommendations", arXiv:1101.0538 [hep-ph].

[167] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run 2", *J. Phys.* **G43** (2016) 023001, arXiv:1510.03865 [hep-ph].

[168] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, "Higgs boson production at the LHC: transverse momentum resummation effects in the $H\to\gamma\gamma$, $H\to W^{\pm}W^{\mp *}\to\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ and $H\to ZZ^*\to\ell^+\ell^-\ell'^+\ell'^-$ decay modes", *JHEP* **06** (2012) 132, arXiv:1203.6321 [hep-ph].

[169] 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].

[170] J. M. Campbell, R. K. Ellis, and C. Williams, "Hadronic production of a Higgs boson and two jets at next-to-leading order", *Phys. Rev.* **D81** (2010) 074023, arXiv:1001.4495 [hep-ph].

[171] S. Gangal and F. J. Tackmann, "Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion", *Phys. Rev.* **D87** no. 9, (2013) 093008, arXiv:1302.5437 [hep-ph].

[172] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: A Program for Higgs boson decays in the Standard Model and its supersymmetric extension", *Comput. Phys. Commun.* **108** (1998) 56, arXiv:hep-ph/9704448 [hep-ph].

[173] A. Djouadi, M. M. Mühlleitner, and M. Spira, "Decays of supersymmetric particles: The Program SUSY-HIT", *Acta Phys. Polon.* **B38** (2007) 635–644, arXiv:hep-ph/0609292 [hep-ph].

[174] A. Bredenstein, A. Denner, S. Dittmaier, and M. Weber, "Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons", *Phys. Rev.* **D74** (2006) 013004, arXiv:hep-ph/0604011 [hep-ph].

[175] A. Bredenstein, A. Denner, S. Dittmaier, and M. Weber, "Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \to WW/ZZ \to 4$ fermions", *JHEP* **0702** (2007) 080, arXiv:hep-ph/0611234 [hep-ph].

[176] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MADGRAPH5: Going Beyond", *JHEP* **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[177] C. Debenedetti, "Concepts for fast large scale Monte Carlo production for the ATLAS experiment." CHEP, 2013.

[178] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, "Interpolation between multi-dimensional histograms using a new non-linear moment morphing method", *Nucl. Instrum. Meth.* **A771** (2015) 39–48, arXiv:1410.7388 [physics.data-an].

[179] **The ATLAS Collaboration**, "Search for Scalar Diphoton Resonances in the Mass Range $65 - 600\,GeV$ with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8\,\text{TeV}$", *Phys. Rev. Lett.* **113** no. 17, (2014) 171801, arXiv:1407.6583 [hep-ex].

[180] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, N. V. Tran, and A. Whitbeck, "Spin and parity of a single-produced resonance at the LHC", *Phys.Rev.D* **86** no. 9, (Nov., 2012) 095031, arXiv:1208.4018 [hep-ph].

[181] F. Maltoni, K. Mawatari, and M. Zaro, "Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects", *European Physical Journal C* **74** (Jan., 2014) 2710, arXiv:1311.1829 [hep-ph].

[182] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett.* **B641** (2006) 57–61, arXiv:hep-ph/0512210.

[183] **The ATLAS Collaboration**, "Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector", *Eur. Phys. J.* **C76** no. 12, (2016) 658, arXiv:1602.04516 [hep-ex].

[184] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, and N. V. Tran, "Spin determination of single-produced resonances at hadron colliders", *Phys. Rev.* **D81** (2010) 075022, arXiv:1001.3396 [hep-ph].

[185] S. Gershgorin, "Über die Abgrenzung der Eigenwerte einer Matrix", *Bull. Acad. Sci. URSS* **1931** no. 6, (1931) 749–754. http://mi.mathnet.ru/izv5235.

[186] W. Cheney and D. Kincaid, *Numerical Mathematics and Computing*. International Thomson Publishing, 4th ed., 1998. ISBN 978-0-534-35184-7.

[187] *BOOST C++ Libraries*. http://www.boost.org.

[188] J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses", *Philosophical Transactions of the Royal Society of London Series A* **231** (1933) 289–337.

[189] K. Cranmer, J. Pavez, and G. Louppe, "Approximating Likelihood Ratios with Calibrated Discriminative Classifiers", arXiv:1506.02169 [stat.AP].

[190] S. Li, "Concise Formulas for the Area and Volume of a Hyperspherical Cap", *Asian Journal of Mathematics & Statistics* **4** no. 1, (2011) 66–70, ScienceAlert:ajms.2011.66.70.