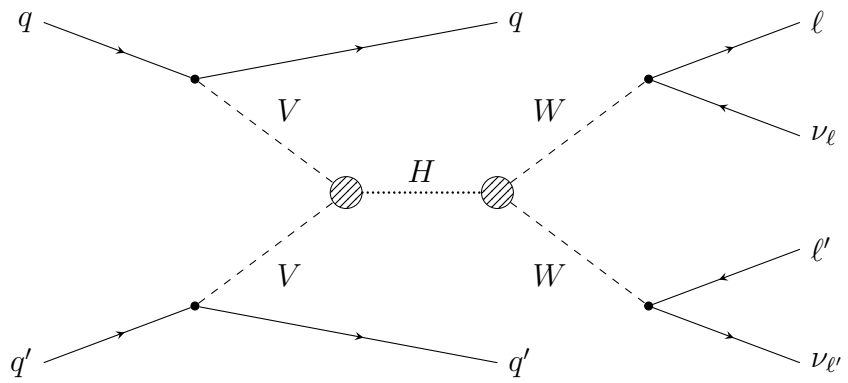
Doktorarbeit

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

Adam Kaluza



Fachbereich Physik, Mathematik und Informatik der
Johannes Gutenberg Universität Mainz

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

DISSERTATION

zur Erlangung des Grades
Doktor der Naturwissenschaften

vorgelegt von

Adam Kaluza

geboren in Tarnowskie Góry, Polen

Fachbereich Physik, Mathematik und Informatik der

Johannes Gutenberg Universität Mainz

Abstract

The first part of this thesis describes and discusses the technique of Effective Lagrangian Morphing, which was developed in the context of this dissertation. For coupling-like parameters in the model, this new method enables to describe the entire parameter space analytically exact exploiting the Lagrangian structure of the underlying theory. Thereby data samples, that are generated at just a few parameter points, are needed as input. The method is derived from first principles, discussed in detail and supplemented with enhancements and techniques to reduce the dimensionality and computational cost, and to improve the accuracy of its predictions. Other state-of-the-art techniques for obtaining physics predictions for measurements are discussed, including different methods of obtaining continuous likelihood functions from a finite set of samples, including Matrix Element reweighting, as well as different methods of interpolation.

In the second part, a first application of this technique for measurements with different sets of coupling parameters derived from effective field theory affecting the coupling of the Higgs boson to vector bosons is presented. The analysis focuses thereby on the Higgs bosons decay mode to a pair of W-bosons which further decay into leptons. Data recorded by the ATLAS detector at $\sqrt{s} = 13\,\text{TeV}$ are investigated using $36.1\,\text{fb}^{-1}$ of LHC Run-2 data. Custom Monte Carlo samples have been generated at different coupling paramters suitable to the Effective Lagrangian Morphing. Based on impacts of the effective couplings on physics observables a dedicated event selection using multivariate techniques is developed. No significant deviation from the Standard Model predictions is observed and limits on the coupling parameters are set. Opportunities and challenges for future applications of this technique to measurements are discussed.

Kurzfassung

Der erste Teil dieser Arbeit beschreibt und diskutiert die Methode des Effektiven Lagrangian Morphings, welche im Rahmen dieser Dissertation entwickelt wurde. Mithilfe dieser neuen Methode können für kopplungsartige Parameter im Modell der ganze Parameterraum unter Ausnutzung der Lagrange-Struktur der zugrundeliegenden Theorie analytisch exakt beschrieben werden. Dabei werden an nur wenigen Parameterpunkten generierte Datensätze als Input benötigt. Die Methode ist von Grund auf hergeleitet, im Detail diskutiert, und durch Verbesserungen und Techniken, um die Dimensionalität und den Rechenaufwand zu verringern und die Genauigkeit seiner Vorhersagen zu verbessern, ergänzt. Andere neueste Methoden, um physikalischer Vorhersagen für Messungen zu erhalten, werden diskutiert, einschließlich verschiedener Verfahren zur Berechnung kontinuierlicher Wahrscheinlichkeitsfunktionen aus einer endlichen Anzahl von Datensätzen, oder der Umgewichtung von physikalischen Ereignissen unter Ausnutzung der Matrixelemente, sowie verschiedener Interpolationsverfahren.

Im zweiten Teil wird eine erste Anwendung des Effektiven Lagrangian Morphings für Messungen mit verschiedenen Sets von Kopplungsparametern vorgestellt, die aus der effektiven Feldtheorie abgeleitet wurden und die Kopplung des Higgs-Bosons an Vektorbosons beeinflussen. Die Analyse konzentriert sich dabei auf den Higgs-Bosonen-Zerfallskanal zu einem Paar von W-Bosonen, die weiter in Leptonen zerfallen. Hierbei werden $36.1\,\mathrm{fb}^{-1}$ des LHC Run-2 Datensatzes, welche vom ATLAS Detektor bei einer Schwerpunktsenergie von $\sqrt{s} = 13\,\mathrm{TeV}$ aufgezeichnet wurden, untersucht. Speziell zugeschnittene Monte-Carlo-Datensätze wurden bei verschiedenen Kopplungsparametern generiert, die für das Effektive Lagrangian Morphing geeignet sind. Basierend auf den Auswirkungen der effektiven Kopplungen auf physikalische Observablen wurde eine dedizierte Ereignisselektion unter Verwendung multivariater Methoden entwickelt und angewendet. Keine signifikante Abweichung von den Standardmodellvorhersagen konnte beobachtet werden und Ausschlussgrenzen an die Kopplungsparameter wurden gesetzt. Möglichkeiten und Herausforderungen für zukünftige Anwendungen des Effektiven Lagrangian Morphings auf Messungen werden diskutiert.

1. Introduction **1**

2. Theoretical Foundations **3**
2.1. The Standard Model of Particle Physics 3
2.1.1. Strong Interaction and Quantum Chromodynamics 6
2.1.2. The Gauge Theory of the Electroweak Interaction 8
2.1.3. The Spontaneous Symmetry Breaking and the Higgs Mechanism 10
2.2. SM Higgs Boson at the LHC 12
2.2.1. Higgs Boson Production 13
2.2.2. Higgs Boson Decay Modes and Search Channels 15
2.2.3. Status of Observations and Predictions 16
2.3. Effective Field Theory 18
2.3.1. Fermi's Interaction 18
2.3.2. Effective Lagrangian 20
2.3.3. Higher Dimension Operators in a Higgs EFT 22
2.3.4. The Higgs Characterization Model 23
2.4. Phenomenology of Proton Collisions 24
2.4.1. From a Lagrangian to a Cross Section 24
2.4.2. Descripton of Proton Collisions 26
2.4.3. Parton Distribution Functions 27
2.4.4. Factorization Theorem 29
2.4.5. Luminosity 30

3. ATLAS at the LHC **31**
3.1. The Large Hadron Collider 31
3.2. The ATLAS Detector 34
3.2.1. The ATLAS Coordinate System 35
3.2.2. The Inner Detector 37
3.2.3. The Calorimenter System 39
3.2.4. The Muon Spectrometer 42
3.2.5. The Trigger System 43

4. Physics Object Definition and Reconstruction **45**
4.1. Tracks 45
4.2. Vertices 46

4.3. Electrons . . . 48
4.4. Muons . . . 51
4.5. Jets . . . 53
4.6. Missing Transverse Energy . . . 58

5. Signal Modelling **61**
5.1. Statistical Data Analysis . . . 61
5.2. Modelling Techniques . . . 63
5.2.1. Matrix Element Reweighting . . . 65
5.2.2. Integral Morphing . . . 68
5.2.3. Moment Morphing . . . 70
5.3. Effective Lagrangian Morphing . . . 72
5.3.1. Derivations . . . 73
5.3.2. Simple Examples . . . 76
5.3.3. Dimensionality . . . 80
5.3.4. Uncertainty Propagation . . . 81
5.4. Validations and Applications . . . 85
5.4.1. ggF $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ . . . 86
5.4.2. VBF $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ . . . 89
5.4.3. ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ . . . 95
5.5. Improvements . . . 96
5.5.1. The Curse of Dimensionality . . . 97
5.5.2. Dimensionality Reduction . . . 103
5.5.3. Coupling Interferences . . . 112
5.5.4. Morphing Basis Optimization . . . 115

6. Effective VBF $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ Coupling Analysis **129**
6.1. Characteristics of the Signal Process . . . 130
6.2. Data and MC Samples . . . 136
6.2.1. Data . . . 136
6.2.2. Signal and Higgs Processes . . . 138
6.2.3. Background . . . 141
6.3. Object Identification and Selection . . . 149
6.3.1. Electrons and Muons . . . 149
6.3.2. Jets and Missing Transverse Energy . . . 151

6.4. Observables . . . 151
6.5. Signal Region Selection . . . 153
6.6. Boosted Decision Tree Application . . . 160
6.6.1. Application . . . 163
6.6.2. Results . . . 165
6.7. Background Estimation . . . 173
6.7.1. Control Regions . . . 173
6.7.2. Misidentified Lepton . . . 180
6.8. Final Selections . . . 186
6.9. Systematic Uncertainties . . . 189
6.9.1. Experimental Uncertainties . . . 191
6.9.2. Theoretical Uncertainties . . . 194
6.9.3. Impact of Systematics Uncertainties . . . 196
6.10. Statistical Treatment . . . 202
6.10.1. Likelihood Construction . . . 202
6.10.2. Statistical Evaluation . . . 206
6.10.3. Model Diagnostics . . . 208
6.11. Results . . . 209
6.11.1. Measurement of SM and AVV . . . 210
6.11.2. Measurement of SM, HVV and H∂V . . . 217

7. Conclusions and Outlook **225**
7.1. Summary . . . 225
7.1.1. Effective Lagrangian Morphing . . . 225
7.1.2. Effective Coupling Measurement . . . 226
7.2. Outlook . . . 226

A. BDT Supplementary Distributions **229**
A.1. BDT Validation Regions . . . 229
A.2. BDT Inputs in the Control Regions . . . 232
A.2.1. Top Control Regions . . . 232
A.2.2. $Z \to \tau\tau$ Control Region . . . 235

List of Figures **237**

List of Tables **243**

Bibliography **245**

1 Introduction

One of the most primal aspects of science itself is the quest for knowledge about the fundamental building blocks of nature. For generations scientists and philosophers have been wondering about the origin of the universe and the exploration of nature. In the last 100 years, modern science has made progress in the description of particle physics and astronomy. While astrophysics deals with the unimaginably large dimensions of the universe, particle physics aspires towards smaller and smaller scales to explore the fundamental building blocks of matter. Nevertheless, both disciplines are closely linked. At high energies, particle physics describes the state of the early universe shortly after the Big Bang, when all matter was condensed in compact space. According to current knowledge, this matter is composed of quarks and leptons whose interactions are mediated by bosons. All findings are summarized in the Standard Model (SM) of Particle Physics, which forms the theoretical foundations of this thesis. It brings together the fundamental forces of electromagnetic, weak and strong interactions. Only in the course of baryogenesis, the elementary particles formed bound states. By the means of gravity, those are responsible for the structures of stars and solar systems observed today.

In the 20th century early experiments lead to the unveiling of an expanding zoo of elementary and bound state particles, whereby the SM managed to introduce an elegant ordering scheme, launching its huge success story. Particles could be predicted that were discovered in later experiments, filling in the gaps in the symmetric nature of the theoretical construct. Yet the model was not complete without one keystone which persistently defied detection - the Higgs boson. Its discovery at the Large Hadron Collider (LHC) with the two detectors ATLAS and CMS in 2012 answered many unresolved questions, particulary how the W- and Z-bosons that mediate the weak force, and all other massive elementary particles gain their masses with the Higgs mechanism in the electroweak symmetry breaking.

Nontheless, the SM can not incorporate all observations made so far. Deficiencies, such as the strong CP problem, neutrino oscillations, matter–antimatter asymmetry, and the nature of dark matter and dark energy cannot be explained with the present state of the model. The compatibility of today's high-precision measurements with the SM raises the expectation that only at high energies the SM will collapse, which classifies it as a low-energy effective field theory (EFT) for a comprehensive theory

yet to be explored. These energies might be beyond the TeV scale of the LHC, supported by the lack of additional new state discoveries in recent years. However, in the framework of an EFT, discoveries beyond the LHC energy scale could be made indirectly by finding deviations in data recorded at the present available high energy limits with high precision measurements.

Simultaneously, with the discovery of the Higgs boson a new door in particle physics has been opened for potential new explorations in processes inluding the Higgs boson. Especially in this context, questions like the origin of the dark matter could be solved by precisely measuring the couplings of the Higgs boson to other particles where new particle states may cause deviations from the model in higher order processes.

The presented thesis studies the effects of such an extended SM with higher dimensional operators affecting the coupling of the Higgs boson to the electroweak W- and Z-bosons in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay channel mainly focusing at the vector boson fusion (VBF) Higgs production, exploiting the benefits of a high production cross section, high Higgs branching fraction to vector bosons, and the presence of two $H \to VV$ ($V = W$ or Z) vertices with potential enhanced effects on EFT observables. The main goal is to measure effects of the most promising higher order coupling parameters of the EFT for proton collisions at the LHC. The dataset has been recorded by the ATLAS detector at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$.

In this context, a main focus of this dissertation addresses the newly developed modelling technique of the *Effective Lagrangian Morphing*, which is an interpolation method for coupling-like parameters, such as coefficients arising from EFTs, applicable to pre-computed predictions for total or differential cross sections and branching fractions. Compared to other interpolation techniques, the new Effective Lagrangian Morphing comprises several advantages. It is fast and efficient enough to be performed within a likelihood fit of the parameters, and models accurately the underlying physics without making any additional approximations.

This thesis is structured as follows. It starts with an introduction into the theoretical foundations in Chapter 2, discussing the SM of particle physics in general, and in the framework of an EFT. Additionaly, the basic phenomenology of high energy proton collisions are summarized. Chapter 3 follows with the experimental composition of the LHC and the ATLAS detector. The reconstruction of the recorded data and the definition of the analysis objects is presented in Chapter 4. Subsequently, the main parts of the thesis begins. Chapter 5 presents the newly developed signal modelling technique of Effective Lagrangian Morphing and its application on Higgs boson coupling measurements. The main analysis in Chapter 6 measures couplings derived from the EFT, which constuitutes a first used case for the Effective Lagrangian Morphing.

2 Theoretical Foundations

This chapter provides a brief introduction to the Standard Model of particle physics. After an overview of the gauge symmetries and particle content, the discussion turns towards the mathematical formalism of the fundamental forces, the Lagrangian of the Standard Model and the Higgs mechanism in Section 2.1. More details on the Higgs boson and its production and decay at a hadron collider can be found in Section 2.2. The introduction of the effective field theory, especially in the Higgs sector, in Section 2.3 will be needed for the interpretation of the analysis later on. Finally, Section 2.4 elaborates on the phenomenology of proton collisions at a hadron collider like the LHC.

While these descriptions are condensed to a comprehensive minimum, more thorough explanations may be found in the text books [1,2] as well as the articles [3,4].

Throughout this thesis, natural units will be used, i.e. $\hbar = c = 1$.

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics [1,3] is the fundamental theory of the elementary particles and their interactions. It has been tested in the last decades with many experiments and found to be in outstanding consistency with most of the measurements. Although one has to assume that the Standard Model does not provide a full explanation of all fundamental forces of the elementary particles[1], all extended theories should include the Standard Model as an effective theory in the low energy limit and should confirm experimental predictions on this energy scale. In the following all important theoretical concepts of the Standard Model are introduced.

As a relativistic quantum field theory the Standard Model includes two of the most substantial achievements in physics of the 20th century: quantum mechanics and special relativity. Therefore all states of Standard Model particles can be classified according to the irreducible representations of the Poincaré group for

[1] The gravitational force is not included and many questions like the origin of dark matter and dark energy or the hierarchy problem are unsolved.

a fixed mass and spin. Particles with half integer spin are called fermions and obey Fermi-Dirac statistics and the Pauli principle, particles with integer spin are called bosons and obey Bose-Einstein statistics. They can be ordered by their interactions as can be seen in Table 2.1.

Furthermore the Standard Model is a local gauge theory and fulfills the local gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$[2]. The invariance of the Lagrange density of the Standard Model regarding these gauge transformations leads to the fundamental interactions between the elementary particles and also, because of Noether's theorem, to the associated conserved charges. The electromagnetic and the weak interactions combined into the unified electroweak interaction are based on the symmetry group $SU(2)_L \times U(1)_Y$, while the strong interaction is described by the symmetry group of quantum chromodynamics, $SU(3)_C$. The Abelian symmetry group $U(1)_{em}$ of quantum electrodynamics with the conserved electric charge Q is contained as a subgroup in the symmetry group of the electroweak interaction.

Due to the renormalizability of the gauge theory it is ensured that also definitive predictions involving higher orders of perturbation theory can be made. Therefore divergences are absorbed into coupling constants and masses, which causes a dependence of them at a given energy scale.

Vector fields act as mediators of the interactions in the Standard Model. They are connected to the generators of the algebra of the symmetry group in the adjunct description as followed:

$$\begin{aligned} SU(3)_C &\rightarrow G_\mu^\alpha, \quad \alpha = 1, \ldots, 8 \\ SU(2)_L &\rightarrow W_\mu^\alpha, \quad \alpha = 1, \ldots, 3 \\ U(1)_Y &\rightarrow B_\mu \end{aligned} \tag{2.1.1}$$

The resulting mass Eigen states from these vector fields are the photon, as the force carrier of the electromagnetic interaction, the vector bosons $W^\pm$ and Z for the weak interaction, plus eight gluons, which mediate the strong force. Together with the Higgs particle, which will be explained in more detail in the Sections 2.1.3 and 2.2, these particles form the bosonic sector of the Standard Model.

The fermionic sector of the Standard Model consists of matter particles with spin $\frac{1}{2}$ and can be subdivided into three generations with two leptons and two quarks each. The particles from different generations have identical interactions and quantum numbers and differ only in their masses. Fermions that take part in the strong interaction are the quarks and are classified in the **3** or $\bar{\mathbf{3}}$ representation of $SU(3)_C$, whereas leptons do not couple to the gauge bosons of the strong interaction.

[2]The indices refer to the couplings of the underlying interactions. The indices C and L stand for *color* and *left*, respectively, whereas the index Y represents the weak hypercharge.

All fermions couple to the neutral currents of the electroweak interaction, but only left-handed fermions or right-handed[3] anti-fermions participate in the interaction with charged currents. This means that right-handed fermions transform with the trivial representation **1** of $SU(2)_L$, while left-handed fermions align in weak isospin doublets under the fundamental representation **2** of $SU(2)_L$. The Eigen states of the down-type quarks, regarding the weak interaction, are denoted with an apostrophe and arise from the mass Eigen states using the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM} according to $d'_i = \sum_j V_{ij} d_j$. A similar mixing is observed for neutrinos, but omitted here for brevity and clarity.

Table 2.1 includes the quantum numbers electric charge Q, weak isospin T_3 and weak hypercharge Y, which are related to each other via the Gell-Mann-Nishijima relation $Q = T_3 + Y/2$. In addition the corresponding adjunct or fundamental representations of the gauge groups $SU(3)_C$, $SU(2)_L$ and $U(1)$ are listed. Not taking neutrino masses into account, the dynamics in the most general Lagrangian depends on 19 parameters whose numerical values have to be established by experiment [5]. These parameters are summarized in Table 2.2.

Table 2.1.: The content of particles in the Standard Model subdivided into fermions, gauge bosons and scalar boson with quantum numbers electric charge Q, weak isospin T_3 and corresponding representations of the gauge groups $SU(3)_C$, $SU(2)_L$ and weak hypercharge of $U(1)_Y$.

Fermions (Spin 1/2)						
Name	Fields			Q	T_3	$SU(3)_C$, $SU(2)_L$, $U(1)_Y$
Leptons	$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}_L$	0 -1	$+1/2$ $-1/2$	$(\mathbf{1},\mathbf{2},-1)$
	e_R	μ_R	τ_R	-1	0	$(\mathbf{1},\mathbf{1},-2)$
Quarks	$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	$+2/3$ $1/3$	$+1/2$ $-1/2$	$\left(\mathbf{3},\mathbf{2},\frac{1}{3}\right)$
	u_R	c_R	t_R	$+2/3$	0	$\left(\bar{\mathbf{3}},\mathbf{1},\frac{4}{3}\right)$
	d'_R	s'_R	b'_R	$-1/3$	0	$\left(\bar{\mathbf{3}},\mathbf{1},-\frac{2}{3}\right)$
Gauge Bosons (Spin 1)						
Symmetry Group	Gauge Fields			Q	T_3	$SU(3)_C$, $SU(2)_L$, $U(1)_Y$
$SU(3)_C$	$G^\alpha_\mu, \alpha = 1, \cdots, 8$			0	0	$(\mathbf{8},\mathbf{1},0)$
$SU(2)_L$	$W^\alpha_\mu, \alpha = 1, \cdots, 3$			$(\pm1,0)$	$(\pm1,0)$	$(\mathbf{1},\mathbf{3},0)$
$U(1)_Y$	B_μ			0	0	$(\mathbf{1},\mathbf{1},0)$
Scalar Boson (Spin 0)						
	Gauge Fields			Q	T_3	$SU(3)_C$, $SU(2)_L$, $U(1)_Y$
	H			0	0	$(\mathbf{1},\mathbf{2},1)$

[3] A particles is right-handed, if its spin is oriented in the opposite direction as its momentum. In contrast to the chirality, the handedness is a Lorentz invariant quantity only for massless particles.

Table 2.2.: Free parameters of the Standard Model [5]. Taking the neutrino masses into account, one gets additionally three neutrino mass parameters and four mixing parameters. The figures in parentheses after the values give the one-standard-deviation uncertainties in the last digits. The coupling constants are given in $\overline{MS}$ scheme at an energy scale at the Z-boson mass ($Q = 91.2\,\text{GeV}$) [2, 5].

Parameter	Description	Value
m_e	electron mass	$0.5109989461(31)\,\text{MeV}$
m_μ	muon mass	$105.6583745(24)\,\text{MeV}$
m_τ	tau mass	$1776.86(12)\,\text{MeV}$
m_u	up quark mass	$2.2^{+0.6}_{-0.4}\,\text{MeV}$
m_d	down quark mass	$4.7^{+0.5}_{-0.4}\,\text{MeV}$
m_s	strange quark mass	$96^{+8}_{-4}\,\text{MeV}$
m_c	charm quark mass	$1.28(3)\,\text{GeV}$
m_b	bottom quark mass	$4.18^{+0.04}_{-0.03}\,\text{GeV}$
m_t	top quark mass	$173.1(6)\,\text{GeV}$
$\sin^2\theta_{12}$	CKM 12-mixing angle	0.307(13)
$\sin^2\theta_{23}$	CKM 23-mixing angle	0.51(4)
$\sin^2\theta_{13}$	CKM 13-mixing angle	0.0210(11)
δ	CP-violating phase	1.02(22)
g'	U(1) gauge coupling	0.313397(17)
g	SU(2) gauge coupling	0.59756(20)
g_s	SU(3) gauge coupling	1.2198(36)
θ_{QCD}	QCD vacuum angle	~ 0
v	Higgs vacuum expectation value	$246.2206(11)\,\text{GeV}$
m_H	Higgs mass	$125.09(24)\,\text{GeV}$

2.1.1 Strong Interaction and Quantum Chromodynamics

The theory of the strong force describes the interactions between the color-charged quarks and gluons, and is called quantum chromodynamics (QCD). The underlying non-Abelian gauge symmetry group is $SU(3)_C$. Quarks exist in three different colors (red, green and blue, and the respective anti-colors) and interact via an octet of vector gauge bosons, the gluons, which are binding the quarks into hadrons (baryons or mesons). They can be very well classified in the quark picture [6] by ordering them into irreducible representations of a symmetry group (e.g. the $SU(4)$ multiplets composed of u-, d-, s- and c-quarks pictured in Figure 2.1).

Quantum chromodynamics provides a quantum field theoretical description of the strong force. The related renormalizable Lagrange density, which is invariant

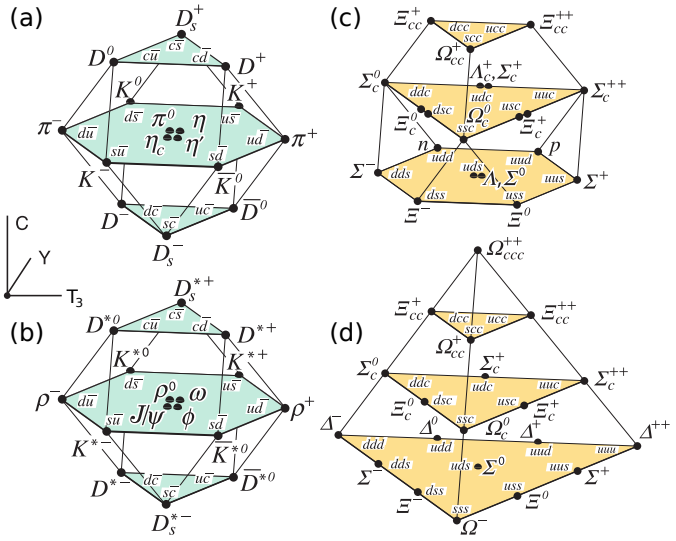


Figure 2.1.: $SU(4)$ Hadron Multiplets for (a) pseudo-scalar (spin 0) and (b) vector (spin 1) mesons and baryons with (c) spin 1/2 and (d) spin 3/2 made of u-, d-, s- and c-quarks as function of isospin T_3, charmness C and hypercharge Y [5].

under local $SU(3)_C$ transformations in form of $U(\boldsymbol{\theta}) = e^{iT^\alpha\theta_\alpha}$, is written in terms of the quarks $q(x)$ and their covariant derivatives and contains in addition the kinetic term for the gluon fields $G^{\mu\nu}_\alpha$:

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{q}(i\not{D} - m_q)q - \frac{1}{4}G^{\mu\nu}_\alpha G^\alpha_{\mu\nu}, \tag{2.1.2}$$

with the Feynman slash notation $\not{D} = \gamma^\mu D_\mu$ and the covariant derivative

$$D_\mu = \partial_\mu - ig_s T_\alpha G^\alpha_\mu \tag{2.1.3}$$

containing the coupling constant g_s and including the gluon field strength tensor

$$G^\alpha_{\mu\nu} = \partial_\mu G^\alpha_\nu - \partial_\nu G^\alpha_\mu + g_s T_\alpha G^\alpha_\mu. \tag{2.1.4}$$

The eight generators $T^\alpha = \lambda^\alpha/2$, where λ^α are the Gell-Mann matrices, satisfy the commutator relations

$$[T^\alpha, T^\beta] = if^{\alpha\beta\gamma}T_\gamma. \tag{2.1.5}$$

The kinetic term $G^{\mu\nu}_\alpha G^\alpha_{\mu\nu}$ contains three and four point self interactions of the gluon fields G^α_μ as a direct consequence of the non-Abelian nature of the gauge group $SU(3)_C$ and indicates that gluons carry combinations of one color and one

anti-color charge that are ordered in a color octet. Related to the charged gluons is the effect of *asymptotic freedom* that is the decrease of the strong coupling at high energies and small distances. Opposite, for low energies and large distances the coupling enlarges and therefore quarks and gluons cannot exist in free form. This phenomenon is called *confinement*. The outcome is compositions of quarks and gluons in colorless hadrons. In collider experiments this accounts for the so-called *jets*. These are narrow cones of hadrons and other particles, produced by the hadronization and fragmentation of a strong interacting particle.

On a bigger scale (about 1 to 3 fm) the strong force is responsible for the cohesion of protons and neutrons in atomic nuclei.

2.1.2 The Gauge Theory of the Electroweak Interaction

The electroweak interaction forms the basis of a unified theory of quantum electrodynamics (QED) and the weak interaction. It was developed, starting with Fermi's theory of weak interactions, which succeeded to describe theoretically the β-decay of the neutron (cf. Section 2.3.1). Later the V-A theory of charged currents could explain parity violation in weak processes. Finally the Glashow-Weinberg-Salam theory unified the weak and electromagnetic interactions. Containing quantum electrodynamics, it describes also the electromagnetism from Maxwell's equations to the magnetic moment of charged leptons, which provides the most accurately verified prediction.

It is based on the gauge group $SU(2)_L \times U(1)_Y$ with the Lagrangian

$$\mathcal{L}_{\mathrm{EW}} = \mathcal{L}_{\mathrm{G}} + \mathcal{L}_{\mathrm{f}} + \mathcal{L}_{\mathrm{SB}} + \mathcal{L}_{\mathrm{YW}}. \tag{2.1.6}$$

The Lagrangian for the gauge fields is given by

$$\mathcal{L}_{\mathrm{G}} = -\frac{1}{4} W^i_{\mu\nu} W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.1.7}$$

including the field strength tensors

$$\begin{aligned} W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g \epsilon_{ijk} W^j_\mu W^k_\nu \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned} \tag{2.1.8}$$

with gauge fields W^i_ν $(i = 1, 2, 3)$ of $SU(2)_L$ and B_μ of $U(1)_Y$. Similar to the case of QCD, three and four point self interaction terms for the $SU(2)$ fields exist. The

physical mass Eigen states arise as a result of mixing the gauge fields according to

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \\ Z_\mu &= \cos(\theta_w)W_\mu^3 - \sin(\theta_w)B_\mu \\ A_\mu &= \sin(\theta_w)W_\mu^3 + \cos(\theta_w)B_\mu, \end{aligned} \tag{2.1.9}$$

with the Weinberg angle θ_w. The coupling of the photon A_μ to charged leptons should be the QED coupling e, which yields to a relation between the coupling constant g of $SU(2)_L$, g' of $U(1)_Y$ and the electric charge e

$$e = g\sin(\theta_w) = g'\cos(\theta_w). \tag{2.1.10}$$

The fermionic term in Equation 2.1.6 can be written as

$$\mathcal{L}_\mathrm{f} = \sum_{L=l,q} \bar{L} i \not{D} L + \sum_{R=l,q} \bar{R} i \not{D} R \tag{2.1.11}$$

using the covariant derivative

$$D_\mu = \partial_\mu - ig\vec{T}\cdot\vec{W}_\mu - ig'\frac{Y}{2}B_\mu. \tag{2.1.12}$$

The left-handed fermions $L = (\nu_l, l)_L^T$ or $L = (q_u, q_d)_L^T$ transform as $SU(2)$ doublets, while the right-handed fermions $R = l_R$ and $R = (q_{u,d})_R$ represent $SU(2)$ singlets. The different behavior in transformation for right- and left-handed fermions is the origin of maximal violation of parity in the electroweak sector. It can also be seen in the shape of the coupling of the charged currents

$$-i\frac{g}{\sqrt{2}}\gamma^\mu\frac{1}{2}\left(1-\gamma^5\right), \tag{2.1.13}$$

which has the well-known vector-axial-vector structure. The three generators $T_i = \sigma_i/2$ (with σ_i as the Pauli matrices) of $SU(2)_L$ and the generator Y of $U(1)_Y$ fulfill the commutator relations

$$[T_i, T_j] = i\epsilon_{ijk}T_k, \quad [T_i, Y] = 0, \quad i,j = 1,2,3. \tag{2.1.14}$$

It is experimentally confirmed that the gauge bosons of weak interaction, $W^\pm$ and Z, have a non-zero mass. But the gauge invariance of $SU(2)_L \times U(1)_Y$ forbids terms in the Lagrangian in form of $M^2 W_\mu W^\mu$. The Higgs mechanism solves this problem and will be discussed in the following chapter. There, the last two terms of Equation 2.1.6 $\mathcal{L}_\mathrm{SB}$ and $\mathcal{L}_\mathrm{YW}$ play an important role.

2.1.3 The Spontaneous Symmetry Breaking and the Higgs Mechanism

The transfer of the concept of spontaneous symmetry breaking to the electroweak sector contains the construction of a Lagrangian, which is invariant under local transformations from $SU(2)_L \times U(1)_Y$, but which has a non-vanishing vacuum expectation value, i.e. the vacuum does not feature this symmetry. This is achieved by introducing an additional complex scalar field

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.1.15}$$

with a non-vanishing weak isospin ($T_3 = -1/2$) and a weak hypercharge of $Y = 1$, and the Lagrangian term

$$\mathcal{L}_{\mathrm{SB}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \tag{2.1.16}$$

with the Higgs potential

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2. \tag{2.1.17}$$

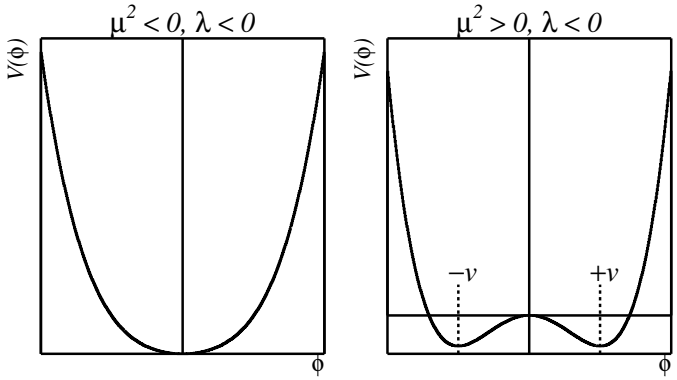


Figure 2.2.: One dimensional representation of the Higgs potential $V(\Phi)$, where v indicates the vacuum expectation value.

For $\mu^2 > 0$ the potential V has a $SU(2)$ set of degenerated minima at (cf. Figure 2.2)

$$|\Phi| = \frac{1}{2} \sum_{i=1}^{4} \phi^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \tag{2.1.18}$$

The spontaneous breaking of the $SU(2)$ symmetry results in the selection of one of these minima as the ground state.

Writing Φ in terms of four fields $\phi_{1,2,3,4}(x)$ and expanding ϕ_3 around one of the minima v with $H(x)$ at first order, one derives

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ (v + H) + i\phi_4 \end{pmatrix} = e^{i\phi_a(x)\tau^a(x)/v} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ (v + H) \end{pmatrix}. \tag{2.1.19}$$

By using the gauge freedom scalar fields transform to

$$\Phi(x) \to e^{-i\phi_a(x)\tau^a(x)/v} \Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ (v + H) \end{pmatrix}. \tag{2.1.20}$$

This choice of gauge, for which only the physical particles are left in the Lagrangian, is called the unitary gauge. Now Φ is the complex scalar field in the low energy limit. From muon lifetime calculations the vacuum expectation value can be determined experimentally to

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{1}{2v^2} \quad \Rightarrow \quad v = (\sqrt{2}\, G_F)^{-\frac{1}{2}} = 246\,\text{GeV}, \tag{2.1.21}$$

with the Fermi coupling constant G_F. This is the energy scale at which the symmetry breaking occurs.

It should be noted that the $U(1)_{em}$ symmetry is conserved, and therefore the photon remains massless. So the spontaneous symmetry breaking reduces the gauge symmetry of the Standard Model to the form of

$$SU(3)_C \times SU(2)_L \times U(1)_Y \to SU(3)_C \times U(1)_{em}.$$

The real field H from Equation 2.1.19 is referred to as the Higgs field. Its excitation corresponds to the Higgs particle. The remaining three degrees of freedom would become Goldstone bosons in a global symmetry. But, since it is a local (gauge) symmetry, these degrees of freedom are absorbed as longitudinal polarization modes for the now massive weak gauge bosons. Their masses can be calculated on tree level to

$$M_W = \frac{g\, v}{2}, \quad M_Z = \frac{\sqrt{g^2 + g'^2}\, v}{2}, \tag{2.1.22}$$

which is obtained by a substitution of the expanded field Φ around the minimum into the Lagrangian $\mathcal{L}_{\text{SB}}$. The fundamental mass of the Higgs particle $m_H = \sqrt{2\lambda}\, v$ (also on tree level) depends on the parameter λ and v, so it remains as free parameter of the Standard Model.

In addition to the weak gauge bosons, also the fermions get their masses from the Higgs mechanism by a Yukawa coupling that can be described with the Lagrangian

$$\mathcal{L}_{\mathrm{YW}} = -g_f(\bar{L}\Phi R + \bar{R}\Phi^\dagger L) = -\frac{g_f v}{\sqrt{2}}\bar{f}f - \frac{g_f}{\sqrt{2}}\bar{f}fH. \tag{2.1.23}$$

So the coupling of a fermion to the Higgs particle is proportional to its mass $m_f = g_f v/\sqrt{2}$. The Standard Model does not predict a value for the coupling constant g_f, so the fermion masses remain as free parameters.

By rewriting the symmetry breaking Lagrangian $\mathcal{L}_{\mathrm{SB}}$ and the Yukawa Lagrangian $\mathcal{L}_{\mathrm{YW}}$ after applying the Higgs mechanism, the kinetic and interaction terms of the Higgs sector read

$$\begin{aligned}
\mathcal{L}_H^{\mathrm{free}} &= \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 H^2 \\
\mathcal{L}_H^{\mathrm{int}} &= \frac{m_H^2}{2v}H^3 - \frac{m_H^2}{8v^2}H^4 - \frac{m_f}{v}\bar{f}fH + \\
&\quad (M_W^2 W_\mu^+ W^{\mu-} + \frac{1}{2}M_Z^2 Z_\mu Z^\mu)(1 + \frac{2}{v}H + \frac{2}{v^2}H^2).
\end{aligned} \tag{2.1.24}$$

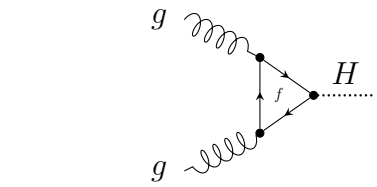
Besides the already discussed Higgs coupling to fermions, three- and four-point self interaction vertices of the Higgs boson can be identified, as well as three- and four-point vertices between two same gauge bosons and the Higgs particle. With $\lambda = m_H^2/(2v^2)$ the four-fold coupling of the Higgs boson is directly proportional to λ.

2.2 SM Higgs Boson at the LHC

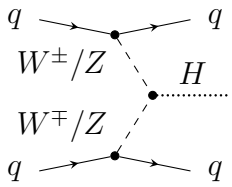
The Large Hadron Collider (LHC) is the up to the present world's largest particle collider ever build and capable of accelerating and colliding proton beams so far up to a center of mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. One of its major purposes was to find the Higgs boson to finally confirm the mechanism that generates the masses of the other elementary particles and to complete the Standard Model. The discovery was achieved by the two major multi-purpose detectors ATLAS and CMS in 2012. The collider and its experiments are presented in greater detail in Chapter 3. This part summarizes the main production mechanisms for the Higgs boson at a pp collider and the important decay channels used to analyze the recorded data from the experiments.

2.2.1 Higgs Boson Production

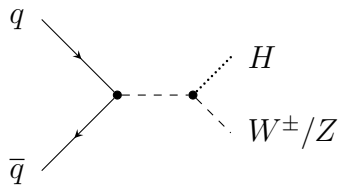
The Higgs bosons can be produced via various mechanisms at the LHC. Figure 2.3 shows a collection of the representative Feynman diagrams for the main production modes. The corresponding cross section for these mechanisms as function of the center of mass energy of the collider can be seen in Figure 2.4. They have been calculated to various degrees of precision as indicated by the labels.



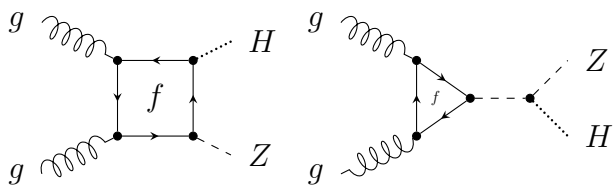
(a) Gluon fusion (ggF)



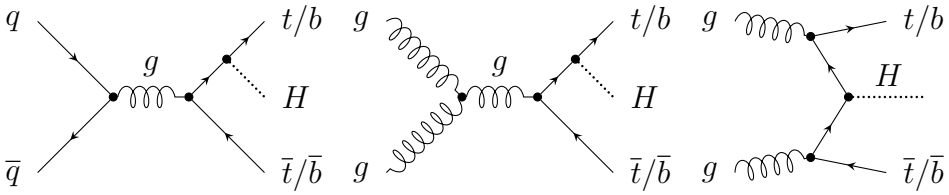
(b) Vector Boson fusion (VBF)



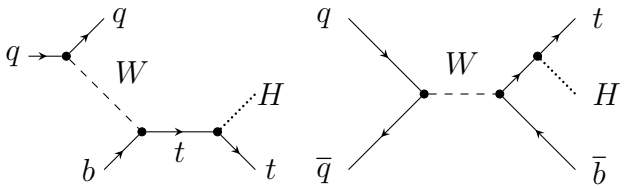
(c) Higgs strahlung (WH or ZH)



(d) Higgs strahlung (ZH only)



(e) Associated production (ttH/bbH)



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

Figure 2.3.: Example Feynman diagrams for Higgs boson production mechanisms.

Gluon Fusion

The most likely way to produce a Higgs particle at the LHC is if two gluons from each proton combine to form a loop of virtual quarks. Since the coupling of particles to the Higgs boson is proportional to their mass, this process is more likely for heavy

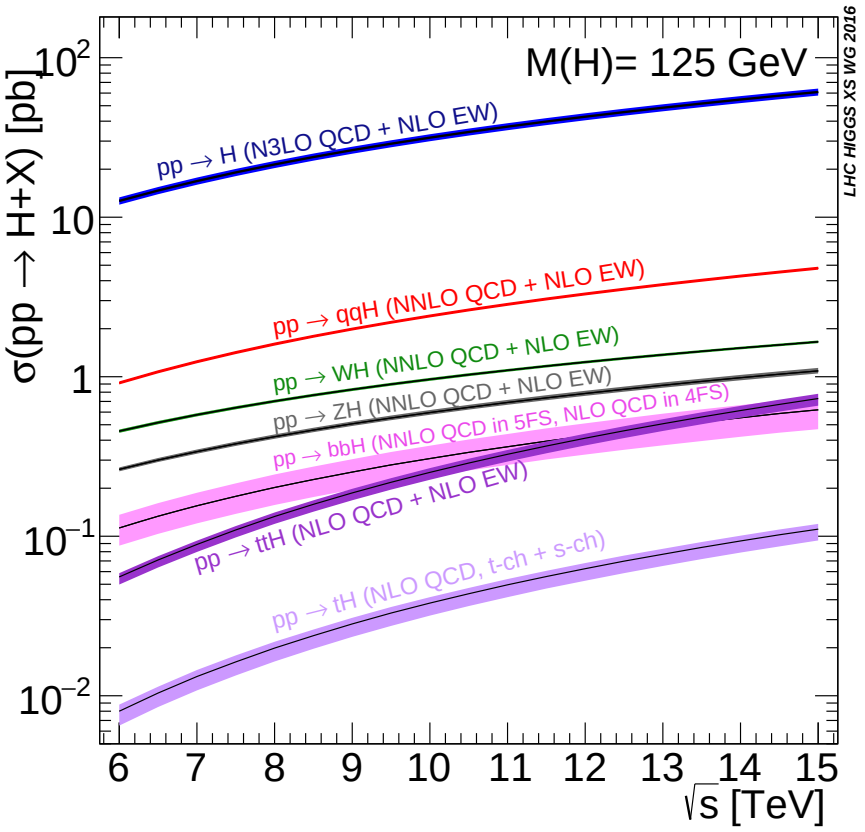


Figure 2.4.: The SM Higgs boson production cross sections as a function of the LHC center of mass energy. Line widths encode the respective theory uncertainties [7].

particles with the largest contribution of the top quark and a $5-10\%$ correction from the bottom quark and the negative top-bottom interference effect.

The inclusive gluon fusion Higgs boson production cross section has a slowly convergent perturbative expansion in QCD with large corrections at next-to-leading (NLO) and next-to-next-to-leading order (NNLO). Therefore, uncertainties due to missing higher orders have always been large and comparable to PDF uncertainties. The most precise calculations include N^{3}LO QCD corrections in the effective theory where the top-quark is assumed to be infinitely heavy, while all other Standard Model quarks are massless, as well as NLO corrections for electroweak (EW) contributions [8].

Vector Boson Fusion

The second most common Higgs production process is via fusion of two weak vector bosons (VBF). The two associated hard quark jets in the forward and backward regions of the detector provide valuable information on the couplings of Higgs bosons

to the massive gauge bosons. The cross section predictions for $pp \to H + 2$ jets are calculated up to NNLO in QCD and NLO EW corrections [9].

Higgs Strahlung

Production of Higgs bosons in association with a W- or Z-vector boson (VH), the so-called Higgs strahlung, 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.

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. The ttH process has been calculated to NLO in QCD as well as to NLO in EW corrections. The calculations of the tH process only employ NLO QCD predictions, featuring s-channel and t-channel diagrams. The bbH cross section prediction combines results from an NNLO QCD calculation in a five-flavor-scheme, as well as from an NLO QCD calculation in a four-flavor-scheme. In the four-flavor-scheme, bottom quarks do not appear in the initial state, but rather via gluon splitting into b-quark pairs.

2.2.2 Higgs Boson Decay Modes and Search Channels

Once produced in a proton-proton collision, the Higgs boson predominantly decays immediately into the heaviest particles that are kinematically available. The search channels at the LHC therefore are focusing on the different decay modes as they lead to different signatures in the final state. With a mass of approximately 125 GeV the coupling of the Higgs boson to gauge bosons and fermions are fully determined by the SM and so are the cross sections, the decay widths and the branching ratios. The branching ratios are defined as the ratio of the decay width of the process to the full decay width to all possible particles,

$$\mathrm{BR}(H \to XX) = \frac{\Gamma(H \to XX)}{\sum_i \Gamma(H \to X_i X_i)}. \tag{2.2.1}$$

Figure 2.5 shows the branching ratios as a function of the Higgs mass between 120 and 130 GeV. The dominant mode is $H \to b\bar{b}$. However measurements in this decay channel are challenging at hadron colliders, as the final state is purely hadronic. In order to be able to separate signal events from the background, analyses need

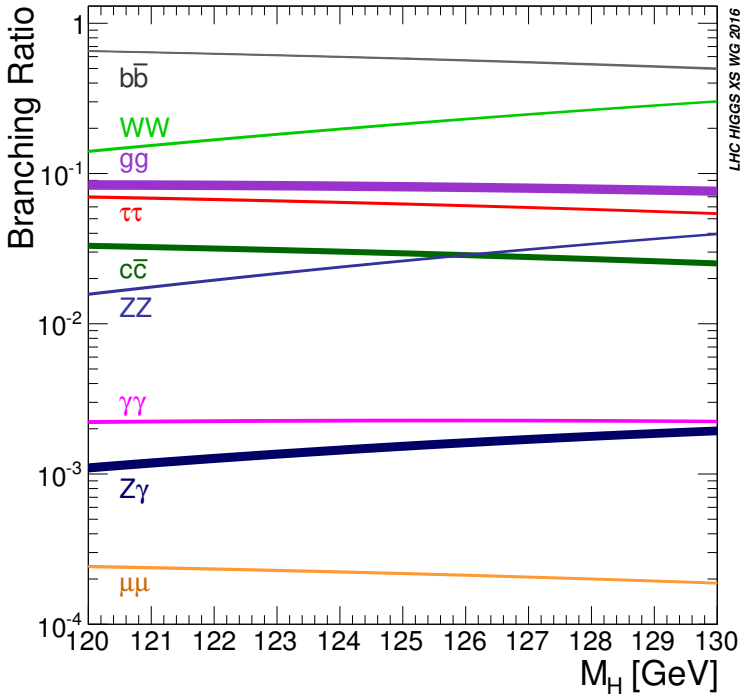


Figure 2.5.: SM predictions for the leading decay branching fractions of a Higgs boson with a mass in the range of 120-130 GeV, with the line widths encoding the respective theory uncertainties [7].

to rely on additional features of the final state and cannot use the dominant ggF production mode. A first evidence was found using the associated production mode [10]. 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. Therefore different production modes from gluon fusion [11] to VBF and associated production [12] can be exploited. Similarly, the decay to a pair of τ-leptons can be used for dedicated analyses [13, 14]. Decays to gluons and charm quarks are so far too difficult to distinguish from the QCD background. More promising are the channels $H \to ZZ$ and $H \to \gamma\gamma$. They are of particular interest because the Higgs boson mass can be fully reconstructed from the final state leptons and photons [15–17]. 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 [18–21].

2.2.3 Status of Observations and Predictions

On 4 July 2012 at CERN, the ATLAS and CMS Collaborations [22, 23] announced the observation of a new particle that, within the present accuracy, does indeed

look like the SM Higgs boson. The latest combined result from both experiments ATLAS and CMS measured a mass of $m_H = 125.09 \pm 0.24\,\text{GeV}$ [24]. The observed value of m_H is a bit too low for the SM to be valid up to the Planck mass with an absolutely stable vacuum [25], but it corresponds to a meta-stable value with a lifetime longer than the age of the universe, so that the SM may well be valid up to the Planck mass. The spin and parity measurements strongly favor the SM Higgs boson hypothesis, corresponding to the quantum numbers $J^P = 0^+$, against all other tested scenarios [26]. Also combined measurement of the Higgs boson production and decay rates [27, 28] agree very well with the SM Higgs particle as shown in Figure 2.6.

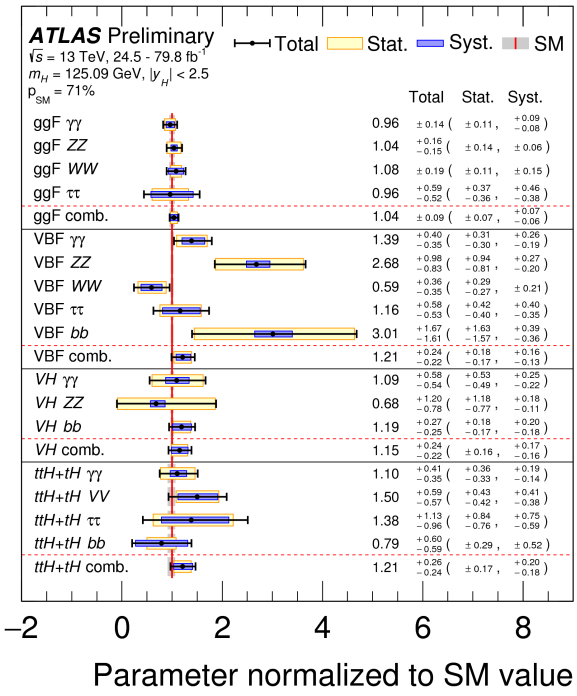


Figure 2.6.: Results of the cross section times branching ratio $\sigma_i \cdot B^f$ for each specific channel $i \to H \to f$ for the combined ATLAS measurements. The values are obtained from a simultaneous fit to all channels. The cross sections of the production processes are fixed to their SM predictions. Combined results for each production mode are also shown, assuming SM values for the branching ratios into each decay mode. The black error bars, blue boxes and yellow boxes show the total, systematic, and statistical uncertainties in the measurements, respectively. The gray bands show the theory uncertainties in the predictions. [28].

2.3 Effective Field Theory

To describe phenomena in nature at a certain energy scale one does not need to know the physics at other scales. Such theories are connected to higher energy scales in the sense that they can be treated as effective field theories (EFT) approximated from a full theory valid at higher energies by integrating out or neglecting higher order effects which are experimentally not relevant at these low energies below a certain cut-off scale Λ. One can distinguish two characteristic contemplations depending on the a-priori knowledge of the theory (cf. Figure 2.7).

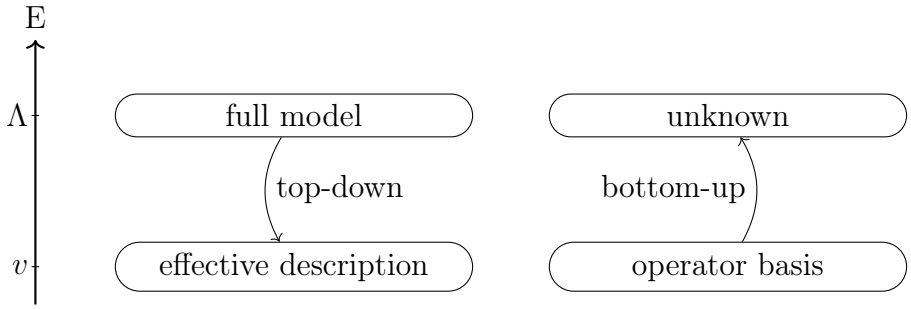


Figure 2.7.: Top-down vs. bottom-up picture of effective field theories.

If the full high-energy theory is unknown the current experimentally confirmed theory can be used as an EFT to learn about the full theory. By systematically expanding the EFT and measuring deviations from it one can gradually expand the knowledge to obtain information about the full theory above the cut-off scale Λ. This approach is called *bottom-up* since one tries to discover new physics going from a low to a high energy regime.

On the other hand in the *top-down* approach a fully known theory can be approximated to an EFT to simplify calculations and learn more about the physics below a certain cut-off scale.

A famous example for an EFT that was first used as a discovery tool in the *bottom-up* approach and can be utilized now in an educational sense is Fermi's interaction describing the β-decay in atomic nuclei.

2.3.1 Fermi's Interaction

To solve the puzzle of the continuous energy spectrum of the electrons emitted in the β-decay $n \to pe^-\bar{\nu}_e$ of the nuclei, Enrico Pauli had suggested that along with the electron, an almost massless neutral particle was emitted which he named neutrino. He described the interaction for this process with a four point fermion vertex as seen in Figure 2.8a, which is now known as a weak interaction mediated by the W-boson.

The modern form of his theory is called V-A theory which incorporates the totally CP violating nature of the weak interaction and uses the quarks as fundamental particles instead of the nucleons. The Lagrangian

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

with the total current

$$J^\mu = \bar{\nu}_e \gamma^\mu (1 - \gamma^5) e + \bar{u} \gamma^\mu (1 - \gamma^5) d' \tag{2.3.2}$$

includes a vector-like current component proportional to γ^μ minus the axial-vector-like current proportional to $\gamma^\mu \gamma^5$ between the left-handed lepton and quark doublets. These particles couple via the Fermi coupling constant G_F that was already introduced in Equation 2.1.21. The dimensionality of the coupling constant hints at an intermediate particle in the four-point Fermi interaction which was later discovered to be the W-boson seen in 2.8b.

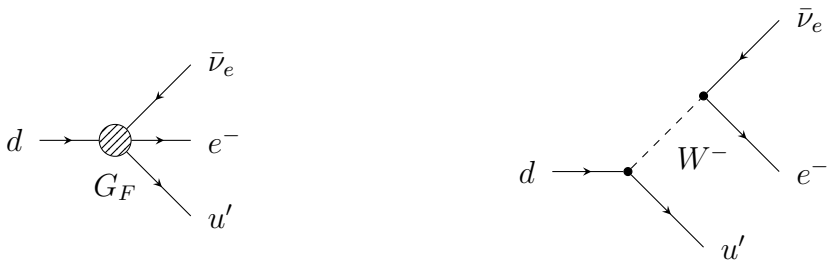
Energies involved in β decays are in the order of $\sim$MeV which is orders of magnitude lower compared to the W on-shell mass at 80.4 GeV. The W-boson involved in the decay is therefore highly virtual with a much higher mass compared to the momentum transfer. Using this approximation the propagator can be reduced to

$$\frac{g_{\mu\nu} - \frac{q_\mu q_\nu}{M_W^2}}{q^2 - M_W^2} \quad \xrightarrow{q^2 \ll M_W^2} \quad \frac{g_{\mu\nu}}{M_W^2}, \tag{2.3.3}$$

which results into a four fermion contact interaction vertex. The W-boson mass dependence corresponds to the same dependence in the Fermi constant and explains it's dimensionality. It acts in addition as cut off scale Λ. Only sufficiently below this mass scale the approximation holds valid within the EFT for the weak interaction. The typical energy scale for a muon decay is it's mass $E = m_\mu$ with which we can estimate the error using the EFT rather than the full model:

$$\Delta_{\text{EFT}} = \frac{\sigma_{\text{EFT}}}{\sigma_{\text{full}}} \sim \frac{E^2}{\Lambda^2} \sim \frac{m_\mu^2}{m_W^2} \approx 10^{-6}.$$

In proton collisions at the LHC the same interaction takes place, but at potentially much larger momentum transfer $E < 13$ TeV. The EFT error increases with E. For $E \gtrsim m_W$, the full model allows on-shell W production, a feature entirely missing in the EFT. Here the two descriptions diverge and Fermi theory is no longer a valid approximation.



(a) Effective four-point interaction for fermions.

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

Figure 2.8.: Feynman diagrams for β decay in the effective Fermi interaction and in the Standard Model.

2.3.2 Effective Lagrangian

EFTs are useful and straightforward to construct especially in the context of quantum field theories (QFT). A first step is a dimension analysis of the objects appearing in a QFT. The basic objects describing the theory is the action

$$S = \int d^4x \mathcal{L}. \tag{2.3.4}$$

It can appear e.g. in an exponent e^{iS} therefore it doesn't have a mass dimension: $[S] = 0$. Looking at the energy dimensions of the space-time integral it is clear with $[x^\mu] = -1$ that a Lagrangian has the dimension $[\mathcal{L}] = 4$. Canonical dimensions of quantum fields follow from the free Lagrangian. E.g. for a scalar field ϕ

$$\mathcal{L}_{\text{free}}^{\text{scalar}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 \tag{2.3.5}$$

one gets $[m] = [\partial_\mu] = [\phi] = 1$. Likewise, from the free gauge field Lagrangian

$$\mathcal{L}_{\text{free}}^{\text{gauge}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.3.6}$$

one derives that the field strength tensor has to have dimension $[F_{\mu\nu}] = 2$ and the gauge field itself $[A_\mu] = 1$.

The mass dimension of an operator in the Lagrangian shows directly it's importance at lower energies looking at the renormalization group flow, i. e. the running of the couplings between different energy scales. One distinguishes *relevant*, *irrelevant* and *marginal* operators. Relevant operators have mass dimension $D < 4$ and receive large quantum corrections going from high to low energies and are therefore needed to describe the low energy behavior of a process. These operators causes

several finetuning problems such as the hierarchy problem or the cosmological constant problem. Irrelevant operators have dimensions $D > 4$ and are typically suppressed when going to lower energies. Operators with dimensions $D = 4$ are called marginal. They contribute equally at all energy scales, but their scaling behavior might be modified by quantum effects.

Another consequence of the mass dimensions of an operator affects the renormalizability of a theory. Relevant operators are non-renormalizable, which means that calculations of 2 particles loop diagrams lead to infinities for energies of the particles $E \to \infty$ that cannot be hidden in a renormalization of the parameters.

For the Standard Model, an EFT can be constructed in a bottom-up approach by expanding the low energy Lagrangian systematically with higher dimension operators. This extended Lagrangian is given by

$$\mathscr{L}_{\text{SM-EFT}} = \mathscr{L}_{\text{SM}}^{(4)} + \sum_{d=5}^{\infty} \sum_{i} \frac{c_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)}. \tag{2.3.7}$$

$\mathscr{L}_{\text{SM}}^{(4)}$ is the usual renormalizable part of the Standard Model and contains dimension-two and -four operators only[4]. The operators $\mathcal{O}_i^{(d)}$ have mass dimension d and their couplings are split into a dimensionless coupling, the Wilson coefficient c_i, and $d-4$ powers of the new physics scale Λ, such that the full Lagrangian term hold the mass dimension four. Once the underlying high-energy theory is specified, all the Wilson coefficients can be determined by integrating out the heavy fields. How the operators $\mathcal{O}_i^{(d)}$ look like might be clear in a top-down situation where the underlying theory is known. In a bottom-up approach, however, we need a recipe to construct a list of operators in a model-independent way. Three ingredients constraint the list of useful operators in a straightforward way: the particle content, the symmetries, and a counting scheme that decides which operators are relevant at the scale of interest.

1. **Particle content**: The dynamical content of the EFT that can form either external legs or internal propagators in the Feynman diagrams has to be defined. At least all particles with masses $m \ll \Lambda$ should be included. The operators are build from combinations of these fields and derivatives. In the SM case the full field content is contained in the EFT. However one distinction can be made: Either the $SU(2)_L$ Higgs doublet field Φ can be taken as fundamental building block for the EFT or the Higgs boson and the Goldstone bosons after symmetry breaking in the electroweak chiral Lagrangian. The

[4] Canonical dimensions of operators are determined from the field contents alone, excluding possible dimensionful coupling constants. The only dimension-two operator in $\mathscr{L}_{\text{SM}}^{(4)}$ is $\Phi^\dagger\Phi$ in the Higgs mass term (cf. Equation 2.1.17).

first type is called *linear* or decoupled SM EFT, whereas the latter is named *non-linear*.

2. **Symmetries**: Some symmetry properties in nature have been measured with high precision, and one can expect that a violation of these symmetries has to be extremely small or happens at very high energies. For the SM all operators have to be invariant under Lorentz transformations and under the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. They also should conserve lepton and baryon number.
3. **Counting scheme**: Taking the particle content and the symmetries into account an infinitive set of operators can be constructed. Therefore a rule is needed which operators can be neglected and which should be kept. From the dimensionality of the operator $d > 4$ follows directly its suppression by the factor $\frac{1}{\Lambda^{d-4}}$. Operators of higher mass dimension are therefore more strongly suppressed. Setting a maximal operator dimension is thus a way of limiting the EFT to a finite number of operators that should include the leading effects at energies $E \ll \Lambda$. For the SM for current measurements it is sufficient to take operators up to dimension 6 into account.

2.3.3 Higher Dimension Operators in a Higgs EFT

Indirect signs of new physics at the electroweak scale will very likely appear in the Higgs sector. Therefore, an effective field theoretical model independent extension of the SM could be very useful for high energy measurements at the LHC.

Following the scheme described in the previous section using the linear Higgs EFT, only one dimension-five operator exists, the so-called Weinberg operator:

$$\mathcal{O}_{W\alpha\beta} = \frac{1}{2}\left(\overline{\ell}^C_{l\alpha}\tilde{\Phi}^*\right)\left(\tilde{\Phi}^\dagger \ell_{l\beta}\right) + \text{h.c.}, \tag{2.3.8}$$

where the symbol ℓ_l denotes left-handed leptons with α and β as their flavor indices, and where Φ is the Higgs doublet. The superscript C denotes a charge conjugated field, and the short-hand notation $\tilde{\Phi} = i\sigma_2\Phi^*$ is used. An explicit model that generates this operator is given by adding heavy, right-handed neutrinos to the SM. Their quantum numbers allow a Yukawa interaction and a Majorana mass term for the heavy right handed neutrinos as well as for the left handed neutrinos. Even though the operator is only suppressed by a single power of the new-physics scale Λ, the experimental bounds on neutrino masses indicate a strong suppression by $\Lambda \sim 10^{14}\,\text{GeV}$ [4]. In general, any operator with odd dimension will violate conservation of the lepton number and $B - L$, the difference between baryon and lepton number, and will thus be neglected henceforth.

A first complete set of dimension-six operators was published in 1985 consisting of 80 operators [29]. It was soon discovered that not all of these operators are independent: they can be linked through equations of motion of the SM fields, integration by parts, or field redefinitions. A complete, non-redundant basis consists of 59 operators, assuming baryon number conservation and barring flavor structure and Hermitian conjugations. Several conventions exist: The Warsaw basis [30], the SILH convention [31, 32], and the HISZ basis [33]. The total number of independent operators, taking into account all flavor combinations and baryon and lepton conservation, is 2499 [34]. For even higher dimensions the number of operators grows very quickly. Dimension-eight in three generations has 44 807 operators and dimension-ten already over 2 million independent operators [35].

So far, no comprehensive experimental studies including all of these operators exist.

2.3.4 The Higgs Characterization Model

A framework based on the EFT approach that allows to perform characterization studies in the Higgs sector is the *Higgs Characterization model* introduced in Reference [36]. It is integrated in the Monte Carlo (MC) generator[5] MADGRAPH5_aMC@NLO [36]. The effective Lagrangian of this model reads

$$
\begin{aligned}
\mathscr{L} = \Bigg\{ & \cos\alpha\kappa_{\mathrm{SM}}\big[\frac{1}{2} g_{HZZ}\, Z_\mu Z^\mu + g_{HWW}\, W^+_\mu W^{-\mu}\big] \\
& - \frac{1}{4}\big[\cos\alpha\kappa_{H\gamma\gamma} g_{H\gamma\gamma}\, A_{\mu\nu}A^{\mu\nu} + \sin\alpha\kappa_{A\gamma\gamma} g_{A\gamma\gamma}\, A_{\mu\nu}\widetilde{A}^{\mu\nu}\big] \\
& - \frac{1}{2}\big[\cos\alpha\kappa_{HZ\gamma} g_{HZ\gamma}\, Z_{\mu\nu}A^{\mu\nu} + \sin\alpha\kappa_{AZ\gamma} g_{AZ\gamma}\, Z_{\mu\nu}\widetilde{A}^{\mu\nu}\big] \\
& - \frac{1}{4}\big[\cos\alpha\kappa_{Hgg} g_{Hgg}\, G^a_{\mu\nu}G^{a,\mu\nu} + \sin\alpha\kappa_{Agg} g_{Agg}\, G^a_{\mu\nu}\widetilde{G}^{a,\mu\nu}\big] \\
& - \frac{1}{4}\frac{1}{\Lambda}\big[\cos\alpha\kappa_{HZZ}\, Z_{\mu\nu}Z^{\mu\nu} + \sin\alpha\kappa_{AZZ}\, Z_{\mu\nu}\widetilde{Z}^{\mu\nu}\big] \\
& - \frac{1}{2}\frac{1}{\Lambda}\big[\cos\alpha\kappa_{HWW}\, W^+_{\mu\nu}W^{-\mu\nu} + \sin\alpha\kappa_{AWW}\, W^+_{\mu\nu}\widetilde{W}^{-\mu\nu}\big] \\
& - \frac{1}{\Lambda}\cos\alpha\big[\kappa_{H\partial\gamma}\, Z_\nu\partial_\mu A^{\mu\nu} + \kappa_{H\partial Z}\, Z_\nu\partial_\mu Z^{\mu\nu} + \big(\kappa_{H\partial W}\, W^+_\nu\partial_\mu W^{-\mu\nu} + \mathrm{h.\,c.}\big)\big]\Bigg\} H\,,
\end{aligned}
\tag{2.3.9}
$$

where the (reduced) field strength tensors are defined as

$$
V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \quad (V = A, Z, W^\pm)\,, \tag{2.3.10}
$$

[5] Monte Carlo generators are software packages, which primarily simulate collider interactions with the Monte Carlo Method [37].

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu \,, \tag{2.3.11}$$

and the dual tensor is

$$\widetilde{V}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma} \,. \tag{2.3.12}$$

Each coupling in the Lagrangian 2.3.9 is parametrized with a dimensionless coupling parameter κ that enables to scale the contribution of every operator individually. To recover the SM case all κ's have to be set to 0 except for $\kappa_{\text{SM}} \cdot \cos\alpha$ which has to be equal to 1. The mixing angle α allows for a complete general description of CP-mixed states. CP-even operators are multiplied with $\cos\alpha$ whereas CP-odd operators have a factor $\sin\alpha$. Operators in Equation 2.3.9 that are suppressed by the cut-off scale Λ are derived from EFT using the SILH operator basis including all possible three-point Higgs interactions with vector bosons that are generated by gauge-invariant dimension-six operators and calculating into the mass Eigen state basis [31, 36]. An overview how the Wilson coefficients translate into the parametrization of this framework can be found in [36] Table 1. In addition the Lagrangian includes CP-odd state couplings typical for SUSY or for generic two-Higgs-doublet models ($A\gamma\gamma$, $AZ\gamma$ and Agg). The operators $H\gamma\gamma$, $HZ\gamma$ and Hgg describe the effective CP-even Higgs coupling to the respective vector bosons.

Contributions of the different BSM operators have individual influences on cross sections and kinematic distributions for distinct physics processes. The characteristics of the most important operators for the VBF $H{\to}W^{\pm}W^{\mp *}{\to}\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ process of interest for this analysis can be found in Section 6.1.

2.4 Phenomenology of Proton Collisions

The transfer from purely theoretical model predictions to actually measurable quantities in experiments such as ATLAS at hadron colliders like the LHC involves many ingredients which are summarized in the following.

2.4.1 From a Lagrangian to a Cross Section

The transition probability P of an initial state $|i\rangle$ transformed by an interaction or scattering matrix S to a final state $|f\rangle$ is given by the absolute square of the transition amplitude

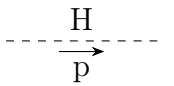
$$P_{i\to f} = \frac{|\langle i|\, S\, |f\rangle|^2}{\langle i \mid i\rangle \langle f \mid f\rangle} \tag{2.4.1}$$

The S-matrix always contains a diagonal piece representing no interaction between initial and final state and an interaction part with the matrix element $\mathcal{M}$ including

the global momentum conservation δ-distribution between two initial particles with four-momentum p_1 and p_2 and the n-body final state with momenta q_f

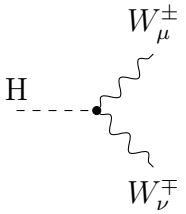
$$S_{fi} = \mathbb{1} + (2\pi)^4 \delta^4 \left(p_1 + p_2 - \sum_{f=1}^{n} q_f \right) i\mathcal{M}_{fi}. \tag{2.4.2}$$

The matrix element can be computed via Feynman rules that connects the Lagrangian of the underlying theory to specific physics processes. A prescription of how Feynman rules are calculated can be found in e.g. [2,38] or automated software packages can be used such as FEYNRULES [39]. These rules contain propagators of particles e.g. the Higgs propagator



$$\hat{=} \quad \frac{i}{p^2 - m_H^2 + i\epsilon} \tag{2.4.3}$$

and fundamental vertices like the SM Higgs interaction with two vector bosons



$$\hat{=} \quad igm_W g_{\mu\nu}. \tag{2.4.4}$$

With these building blocks, pictorial representations of specific processes, the *Feynman diagrams* (cf. e.g. Figures 2.3 or 6.1), can be directly translated to mathematical expressions to calculate the matrix element. If more than one Feynman diagram exists for the same process, the corresponding amplitudes have to be added, following the Feynman-rules, to take interferences into account. In addition there exists initial or final state radiation or various loop diagrams, including a larger number of vertices and hence coupling factors. These diagrams are of higher order and are suppressed by powers of the coupling factor accordingly. But the number of higher order diagrams grows quickly when more loops or additional radiation is included which can result in a sizeable contribution to the total cross section. Many processes have nowadays been calculated to next-to-leading (NLO) and even next-to-next-to-leading (NNLO), to make the theory predictions more precise.

The cross section, which represents the effective area for a reaction to happen, is calculated with the matrix element squared in a $2 \to n$ scattering process illustrated in Figure 2.9 with

$$\sigma = \frac{C}{4\sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}} \int |\mathcal{M}|^2 (2\pi)^4 \delta^4 \left(p_1 + p_2 - \sum_{f=1}^{n} q_f \right) \times \prod_{f=1}^{n} 2\pi \delta(q_f^2 - m_f^2) \Theta(q_f^0) \frac{d^4 q_f}{(2\pi)^4}. \quad (2.4.5)$$

The factor C corrects for double counting in case of identical particles in the final state, the first δ-distribution ensures again momentum conservation between ingoing and outgoing particles, the second δ-distribution forces each outgoing particle to be on its mass shell and the energies of the outgoing particles have to be positive as provided by the Θ-function.

The cross section is very convenient to use as an observable since it is a pure physics quantity and independent of the experimental setup details. However, before predictions from such calculations can be compared to real measurements, additional effects need to be taken into account which are summarized hereafter.

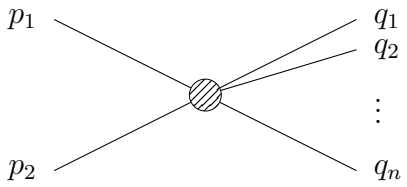


Figure 2.9.: Feynman diagram for a general $2 \to n$ particles process.

2.4.2 Descripton of Proton Collisions

The compositeness of hadrons complicates the description of a proton collision with respect to events at a lepton collider, where elementary particles are colliding. In a typical pp collision primarily two partons interact in a hard scattering process, while the remnants of the initial protons give rise to additional activities. This process can be decomposed into different stages as illustrated in Figure 2.10.

The two incoming protons are symbolized by the large dark-green ellipses and the three green lines indicating the three valence quarks of each proton. The main interaction of these protons is the *hard scattering process* depicted as the large red circle. The colliding partons can undergo QCD radiation before the main process as *initial state radiation* (ISR) or after as *final state radiation* (FSR) either by gluon splitting ($g \to gg$ or $g \to q\bar{q}$) or gluon radiation ($\overline{q} \to \overline{q}\, g$). The hard scatter products (red lines), in this example a Higgs boson and two top quarks, decay further (small red circles) into the final state particles. Strong interacting particles radiate and decay further in so-called *parton showers* (PS), and form

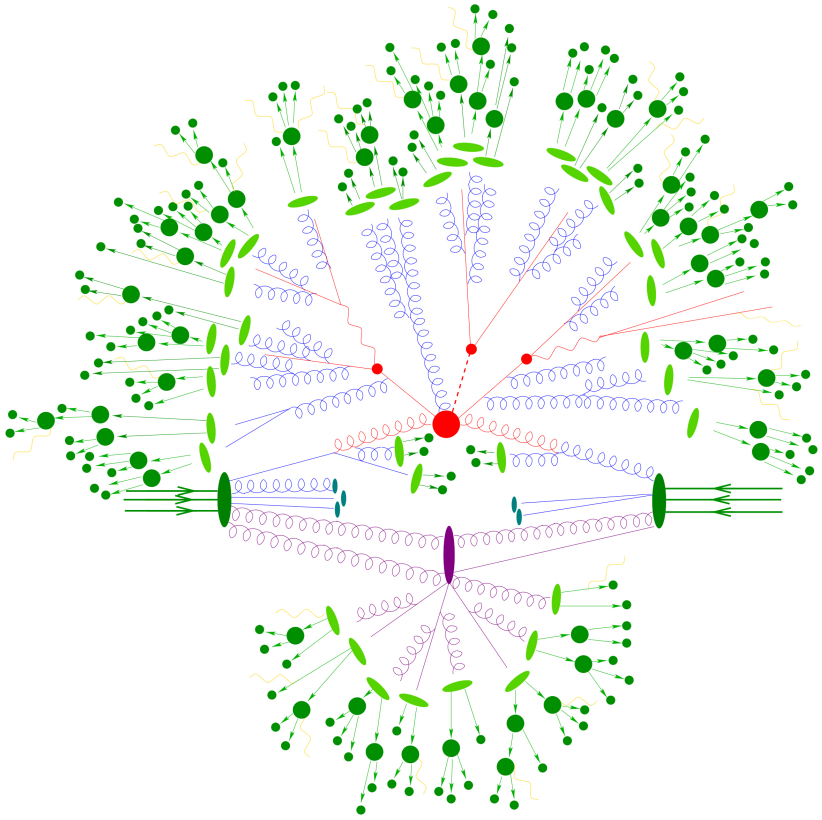


Figure 2.10.: Pictorial representation of a typical pp collision event [40].

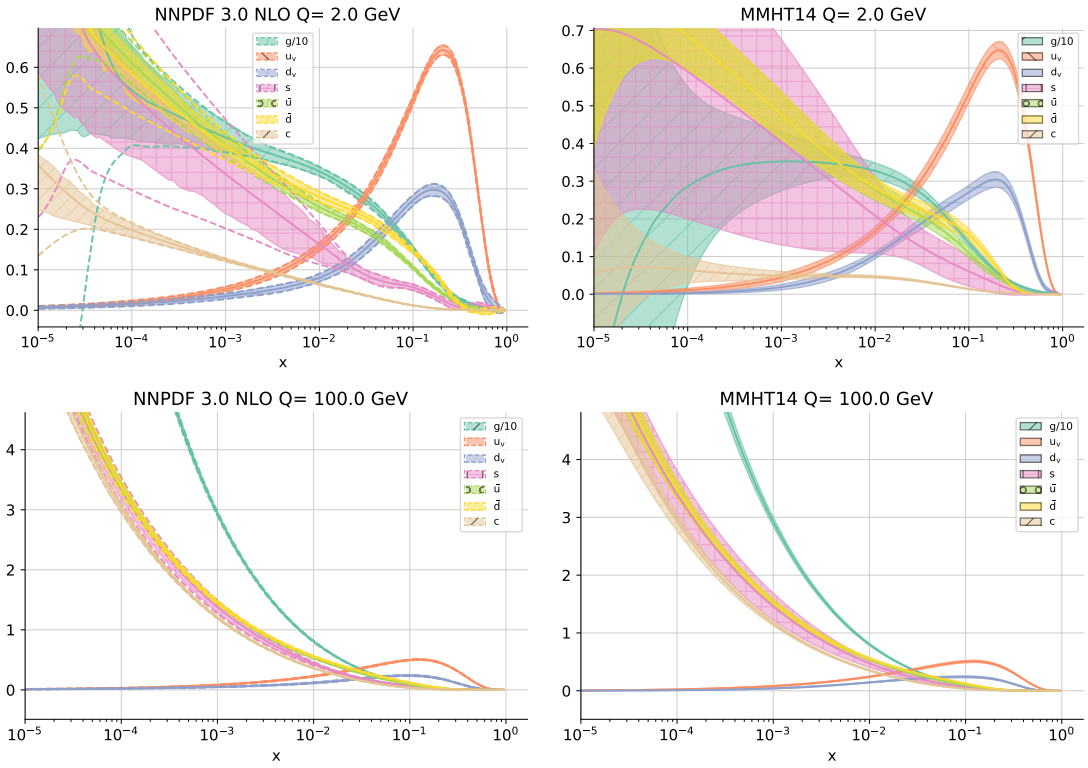
colorless hadrons (green ellipses), a process called *hadronization*. Further these hadrons decay into stable particles (green circles). The resulting collimated bundles of hadrons are called *jets*.

In addition, a secondary process between remnants of the protons (purple ellipse) can occur shown in the lower part of Figure 2.10. Again, a parton shower is produced (purple lines), resulting in hadronization and decay into stable particles. Typically this process is much softer than the primary interaction and, together with the other remnants of the proton (blue ellipses), is part of the *underlying event* (UE).

Electromagnetic radiation (yellow lines) can be emitted by charged particles at any stage of the event development.

2.4.3 Parton Distribution Functions

Since the fundamental scattering does not happen among the protons themselves but their constituents, which can be either the gluons or the quarks inside the protons, one needs to translate the momentum of the proton to the colliding partons. Assuming no rescattering, the momenta of the partons $\hat{p}_i$ are parallel and



(a) The NNPDF3.0 PDFs at NLO [41]. **(b)** The MMHT2014 PDFs [41].

Figure 2.11.: Comparison of different PDF sets, evaluated at $Q = 2\,\text{GeV}$ (top) and $Q = 100\,\text{GeV}$ (bottom).

a fraction of the total momentum of the proton p_i. The parton momentum fraction is expressed by the Bjorken variable $x_i = \frac{\hat{p}_i}{p_i}$. Not the total center of mass energy $\sqrt{s}$ of the pp collision is available but instead only the center of mass energy between the partons $\sqrt{\hat{s}}$. It can be expressed with $\hat{s} = x_1 x_2 s$ by neglecting the parton masses, which is a good approximation at high energies. The probability density to find a parton a with a certain momentum fraction x_a at a given momentum transfer Q^2 is given by the parton density function (PDF), $f_{a/A}(x_a, Q^2)$.

In context of QCD these PDFs can not be calculated perturbatively and need to be determined by experiment. Only the evolution for different energy scales can be determined with the DGLAP equations. There are many sets of PDFs available which rely on different experimental inputs and on a variety of mechanisms for fitting, interpolation and uncertainty estimation. Examples of these sets are NNPDF [42], MMHT [43] or CT14 [44] and a comparison of two of them at different energy scales are shown in Figure 2.11. The PDFs can be evaluated from discretised data files with the software LHAPDF [45].

2.4.4 Factorization Theorem

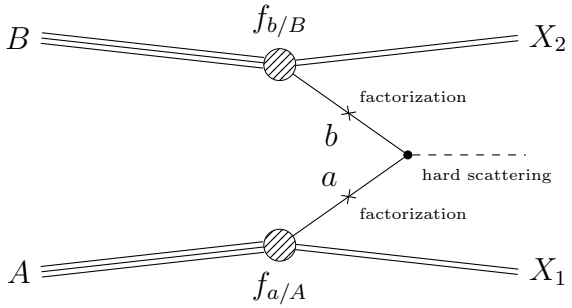


Figure 2.12.: Illustration of the factorization theorem of a hadron-hadron collision.

Using these PDFs the cross sections of proton collisions can be calculated by utilizing the *factorization theorem*, which separates the non-perturbative long distance behavior from the perturbative short-distance part in a systematic fashion, illustrated in Figure 2.12. The cross section of a collision of two hadrons A and B denoted by σ_{AB} is calculated by the convolution of two PDFs $f_{a/A}$ and $f_{b/B}$, describing the probability density to find a parton of flavor a, b inside the hadron A, B carrying a momentum fraction $x_{a/b}$:

$$\sigma_{AB} = \sum_{a,b\in(q,\bar{q},g)} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \hat{\sigma}_{ab\to X}. \tag{2.4.6}$$

The factorization scale μ_F can be thought of as the scale which separates the long and short-distance physics. Infrared singularities from the partonic cross section calculation arising from collinear or soft partons are absorbed in the PDFs consistently at all orders in the perturbative expansion. Roughly speaking, a parton with a transverse momentum less than μ_F is considered to be part of the hadron structure and is absorbed in the parton distribution. Partons with larger transverse momenta participate in the hard scattering process with a short-distance partonic cross section $\hat{\sigma}$. At each order in α_S the partonic cross section $\hat{\sigma}_{ab\to X}$ contains ultraviolet infinities that are renormalized. A remnant of the subtraction point is left at each perturbative order as a renormalization scale dependence μ_R

$$\hat{\sigma}_{ab\to X} = \alpha_S^k(\mu_R) \sum_{m=0}^{n} \hat{\sigma}_{ab\to X}^{(m)} \alpha_S^m(\mu_R). \tag{2.4.7}$$

Formally, the cross section calculated to all orders in perturbation theory is invariant under changes of the scales μ_F^2 and μ_R^2, e.g. $\hat{\sigma}_{ab\to X}^{(1)}$, exactly compensates the explicit scale dependence of the parton distributions and the coupling constant. This

compensation becomes more exact as more terms are included in the perturbation series. In the absence of a complete set of higher order corrections, it is necessary to make a specific choice for the two scales in order to make cross section predictions. To avoid unnaturally large logarithms reappearing in the perturbation series it is sensible to assume $\mu_F = \mu_R$ and choose μ_F and μ_R values of the order of the typical momentum scales of the hard scattering process.

2.4.5 Luminosity

The number of events per second generated in a collider R is given by the cross section σ of the event under study and the machine *luminosity* L

$$R = \frac{dN}{dt} = L\sigma, \qquad [L] = \mathrm{s}^{-1}\mathrm{cm}^{-2}. \tag{2.4.8}$$

The machine luminosity depends only on the beam parameters and can be written for a Gaussian beam distribution with equal beam parameters for both circulating beam as

$$L = \frac{n_b N_b^2 f_{\mathrm{rev}} \gamma_r}{4\pi\epsilon_n \beta^*} F, \tag{2.4.9}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency and γ_r the relativistic gamma factor. The normalized transverse beam emittance ϵ_n express the beam size. It measures the average spread of particle coordinates in position-and-momentum phase space. In order to preserve its value as function of the beam energy, it is usful to normalize the emittance with the relativistic factors via $\epsilon_n = \gamma_r \beta_r \epsilon$. The beta function, which value at the collision point is defined as β^*, is related to the transverse size of the particle beam at the location along the nominal beam trajectory. The additional factor F accounts for a geometrical correction due to the crossing angle with which the beams are brought into collision. Apart from the instantaneous luminosity, the (time-) integrated luminosity, $\mathcal{L} = Ldt$, is also of interest for the experiments. It is a measure for the amount of data produced in a certain period of time in the units of inverse cross section, i.e. in 1/pb, 1/fb, etc.

3 ATLAS at the LHC

For the most basic insights into the dynamics and fundamental structure of matter, physicists seek the simplest kinds of interactions at the highest possible energies. These typically entail interactions of one of the simpler kinds of particles like protons surrounded with a wealth of measurement devices to observe many types of particles at a high rate under controlled experimental conditions.

The largest and highest energy particle accelerator used for elementary particle physics is the Large Hadron Collider which is introduced in this Section 3.1. First, the main characteristics of the LHC and its main experiments are briefly established with an overview of the running conditions until 2016 and the corresponding dataset delivered. In Section 3.2 the ATLAS experiment, whose recorded physics data is used in this analysis, will be discussed in more detail, focussing on the relevant subdetectors and the trigger system.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [46] is a two-ring-superconducting-hadron accelerator and storage ring. It is located at the European Laboratory for Particle Physics (CERN[1]) near Geneva between the swiss french border and is installed in a 26.7 km long tunnel, which lies on average 100 m below ground level and previously housed the Large Electron Positron Collider (LEP) [47] that was dismantled around the year 2001 after completing its scientific agenda. The LHC was designed to collide proton pairs with an unprecedented center-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ and a luminosity of $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$ or pairs of heavy ions such as lead nuclei at center-of-mass energies of about $\sqrt{s} = 5\,\text{TeV}$ per nucleon and a luminosity of $10^{27}\,\text{cm}^{-2}\text{s}^{-1}$.

To accelerate two proton beams in counter-rotating directions, two separate beam pipes with antipodal magnetic dipole fields share one twin-bore tube hosting the vacuum vessel and the iron yoke for the superconducting NbTi magnets due to the limited space in the tunnel diameter of 3.7 m. The 1232 dipole magnets are cooled down to 1.9 K with superfluid Helium and are able to reach magnetic

[1] Conseil Européen pour la Recherche Nucléaire

fields of up to 8.33 T. In addition to the dipole magnets for steering, there are 392 quadrupole magnets for focusing the beams and magnets with higher multipole order to provide the beam optics necessary to maintain and collide the beams at the four interaction points in the straight sections of the octagonal structure where the experiments are located. The other four straight sections house technical facilities for beam cleaning, radio frequency acceleration and beam dumping.

Each proton beam at full intensity consists of 2808 bunches per beam and each bunch is separated by 25 ns and contains $1.15 \cdot 10^{11}$ protons at the start of a nominal fill. After a sequence of preceding accelerators shown in Figure 3.1, the bunches are fed into the LHC by the Super Proton Synchroton (SPS) at an energy of 450 GeV. After approximately 20 min injection time the beams are accelerated to the operational energy in around 25 min. Collisions are then ideally taken for 10 hours or more with stable beams, until the beam deteriorates so much that it is more efficient to dump and refill.

The storage ring has an eight-fold symmetry with eight arc sections and eight straight sections. Four large experiments are located at the collisions points at the straight sections (see Figure 3.2) and a number of smaller experiments situated around the LHC ring.

- ATLAS (cf. Section 3.2) and CMS (**C**ompact **M**uon **S**olenoid) are two general purpose detectors with the aim to cover a range of physics measurements and searches as wide as possible in order to be able to take full advantage of the discovery potential of the LHC. They were instrumental for the discovery of the Higgs boson in July 2012.
- LHCb focuses on physics involving bottom quarks dedicated to measurements of CP violation and rare decays of B-hadrons.
- ALICE (**A** **L**arge **I**on **C**olliding **E**xperiment) focusing on the physics of strongly interacting matter at extreme energy densities. It is primarily designed to study heavy-ion collisions to produce quark–gluon plasma.
- TOTEM (**Tot**al **E**lastic and diffractive cross section **M**easurement) aims to measure the total pp cross section and study elastic and diffractive scattering. It shares an interaction point with the CMS experiment.
- ALFA (**A**bsolute **L**uminosity **F**or **A**TLAS) is located at 240 m from the ATLAS interaction point and intends to measure the elastic pp-scattering at small angles in the Coulomb-Nuclear Interference region.
- LHCf is designed to study the particles generated in the forward region of collisions, those almost directly in line with the colliding proton beams to explain the origin of ultra-high-energy cosmic rays. It is located at the ATLAS interaction point.

- MOeDAL (**Mo**nopole and **E**xotics **D**etector **a**t the **L**HC) is dedicated to directly search for the magnetic monopoles and other highly ionizing stable massive particles and was implemented at the LHCb cavern.

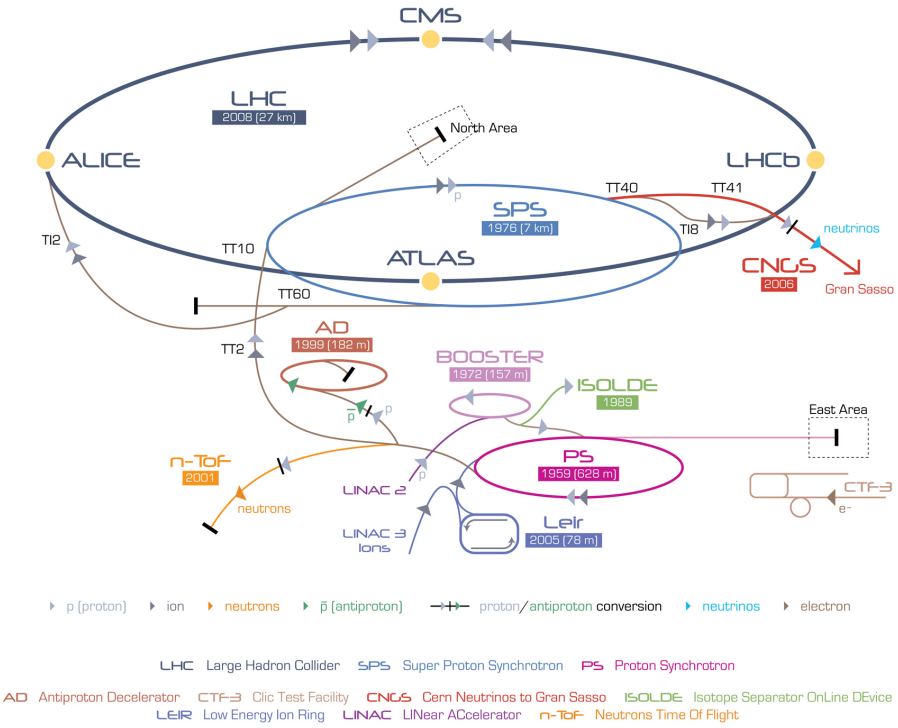


Figure 3.1.: The CERN accelerator complex, showing the Linear accelerator 2 (Linac 2), the Booster, Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) which act as pre-accelerators for protons to be injected into the Large Hadron Collider (LHC), as well as many other experimental facilities located at CERN [48].

Since the year 2010 the LHC has been actively recording data from pp collisions with increasing center-of-mass energy over time. One distinguishes two phases of data taking: Run 1 in the year 2011 with at $\sqrt{s} = 7\,\text{TeV}$ and in 2012 at $\sqrt{s} = 8\,\text{TeV}$, and Run 2 started in the years 2015 and 2016 with a center-of-mass energy at $\sqrt{s} = 13\,\text{TeV}$. The delivered integrated total luminosities in fb^{-1} for these periods were

year	ATLAS	CMS
2011	5.63	6.14
2012	23.27	23.27
2015	4.22	4.22
2016	38.96	41.97

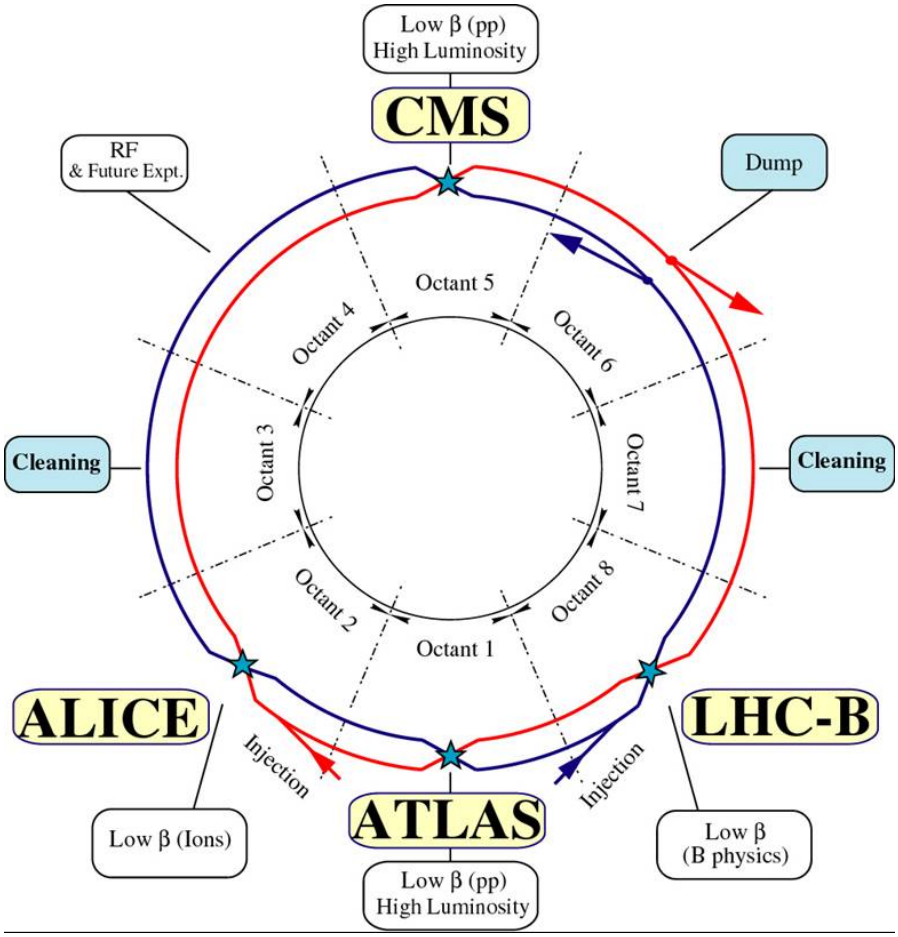


Figure 3.2.: Overview schematic display of the octagonal structure of the LHC with interaction points and facilities [49].

Differences of integrated luminosities between ATLAS and CMS result from inefficiency in the data aquisition due to e.g. unavailability or malfunctions of the different sub-detectors, magnets or trigger system.

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 about 30% to a peak value of $1.38 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

3.2 The ATLAS Detector

The ATLAS detector (**A** **T**oroidal **LHC** **A**pparatu**s**) is one of the two general purpose detectors at the LHC with a total size of $46\,\mathrm{m} \times 26\,\mathrm{m} \times 26\,\mathrm{m}$. Its design was driven by the goal of covering a broad physics spectrum with emphasis on the discovery

of the Higgs boson and searches for new physics beyond the SM. It is structured in layers, which are made up of detectors of different types, each of which is designed to observe specific types of particles. The different traces that particles leave in each layer of the detector allow for effective particle identification and accurate measurements of energy and momentum. Figure 3.3 shows an overview of the ATLAS layout, consisting of the Inner Detector for tracking and vertexing, electromagnetic and hadronic calorimeters for particle energy measurements and the muon spectrometer, specially designed to detect muons traversing the detector. The Inner Detector contains a superconducting solenoid magnet providing magnetic fields of up to 2 T, bending the tracks of charged particles to allow for precise momentum measurements. Superconducting air-core magnets surrounding the calorimeters provide another toroidal magnetic field to allow for muon momentum measurements. The detector components as well as the geometry of the detector are described in more detail in the following subsections.

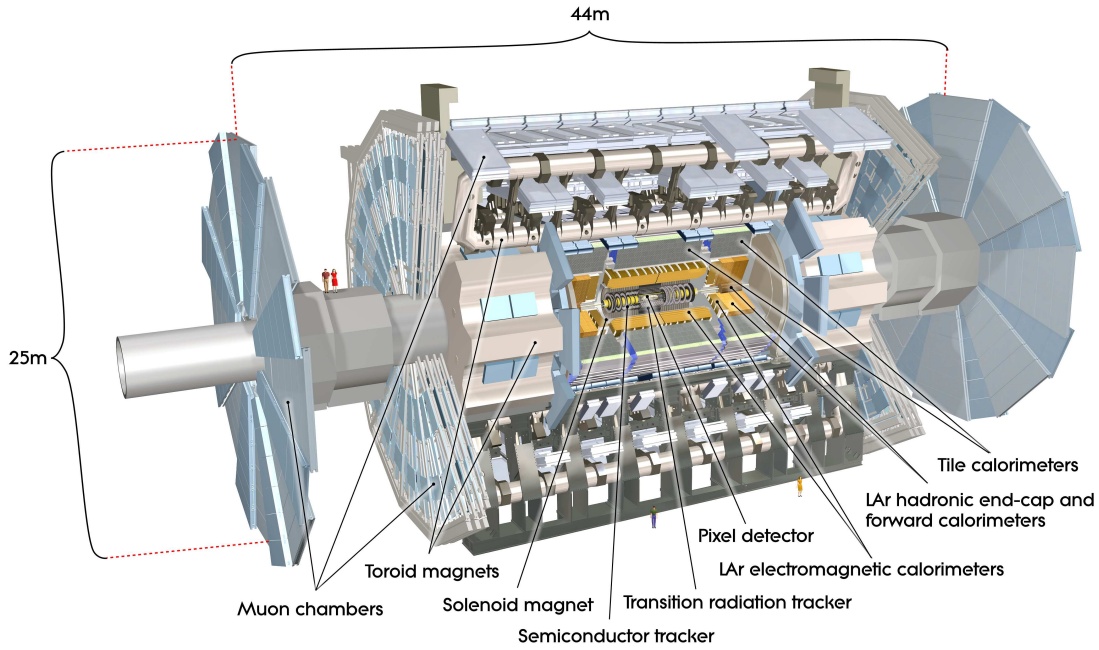


Figure 3.3.: Cut-away view of the ATLAS detector with different sub detectors and magnet components highlighted [50].

3.2.1 The ATLAS Coordinate System

The nominal collision or interaction point within ATLAS defines the origin of the coordinate system. The x-axis is pointed radially towards the center of the LHC ring, while the y-axis points upwards, and the z-axis points along the beam pipe

such that the coordinate system is right-handed. In the transverse x-y-plane the azimuthal angle $\phi \in [0, 2\pi]$ is measured relative to the x-axis. The polar angle $\theta \in [0, \pi]$ is measured with respect to the z-axis (see Figure 3.4a). Since differences $\Delta\theta$ are not Lorentz-invariant, the pseudorapidity η, defined as

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{3.2.1}$$

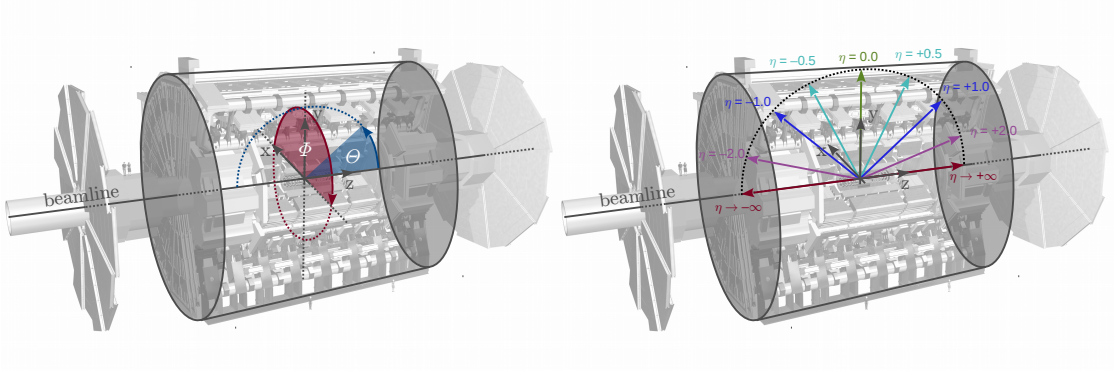
is mostly used for position specification. It is positive in the forward region, negative in the backward region and diverges to infinity close to the beamline (see Figure 3.4b). In the relativistic limit where the momentum of a massive particle $\vec{p}$ is much larger than its mass m, the pseudorapidity can be derived from the rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.2.2}$$

with $E = \sqrt{\vec{p}^2 + m^2}$ being the energy of the particle and p_z the momentum along the z-axis. The distance between two objects is given by

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}, \tag{3.2.3}$$

where the $\Delta\eta$ and $\Delta\phi$ label the difference of the two objects coordinates in η and ϕ respectively.



(a) Depiction of the azimuthal angle ϕ and polar angle θ.

(b) Illustration of the pseudorapidity η.

Figure 3.4.: Coordinate systems of the ATLAS detector [51].

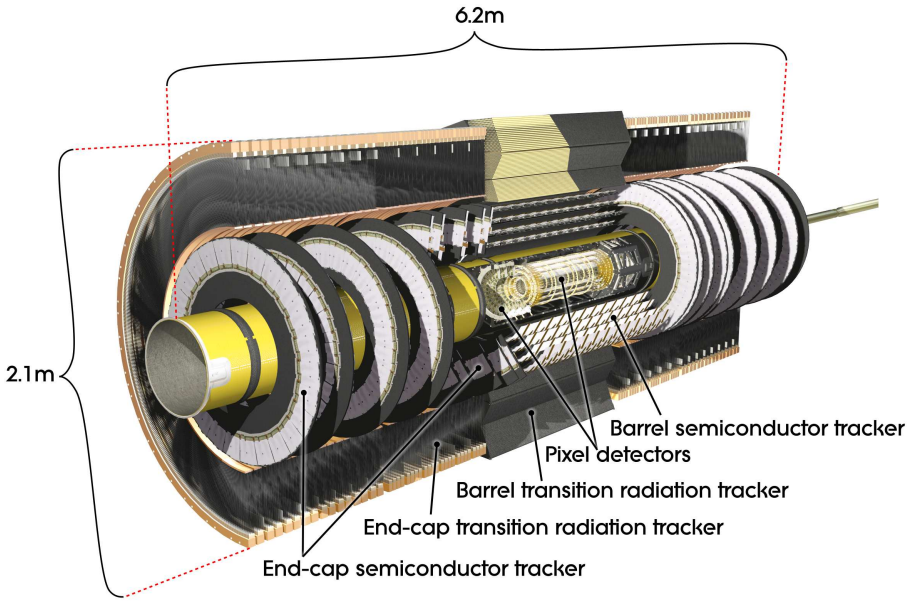


Figure 3.5.: Cut-away view of the Inner Detector [50].

3.2.2 The Inner Detector

The Inner Detector shown in Figures 3.5 and 3.6 is built as close as possible to the interaction point to provide robust pattern recognition, precise tracking and accurate momentum and charge measurement. It is operating in a nearly homogeneous magnetic field of 2 T created by a superconducting solenoid magnet forcing charged particle trajectories to bend depending on their charge and momentum. Made up of three independent subcomponents with decreasing spacial resolution, the Inner Detector consists of the pixel detector followed by the semi-conductor tracker (SCT) and the transition radiation tracker (TRT). Each component is arranged on concentric cylinders around the beam axis while in the end-cap regions they are located on disks perpendicular to the beam axis.

The Pixel Detector

The innermost component is the Pixel Detector consisting of pixel modules arranged in four concentric layers around the beampipe and three additional disks on either side of the end-caps covering the region $|\eta| < 2.5$. The pixel modules consist of silicon sensors and the corresponding front-end readout electronics and exploit the semiconductor properties of the silicon. Charged particles passing through the doped silicon pixels generates electron-hole-pairs in the conduction band by setting free valence electrons. Under the influence of an electric field, these charge carriers travel to the electrodes, where the electric signals can be recorded. The innermost layer, the insertable b-layer (IBL) was implemented in 2014 before Run-2 at a radius

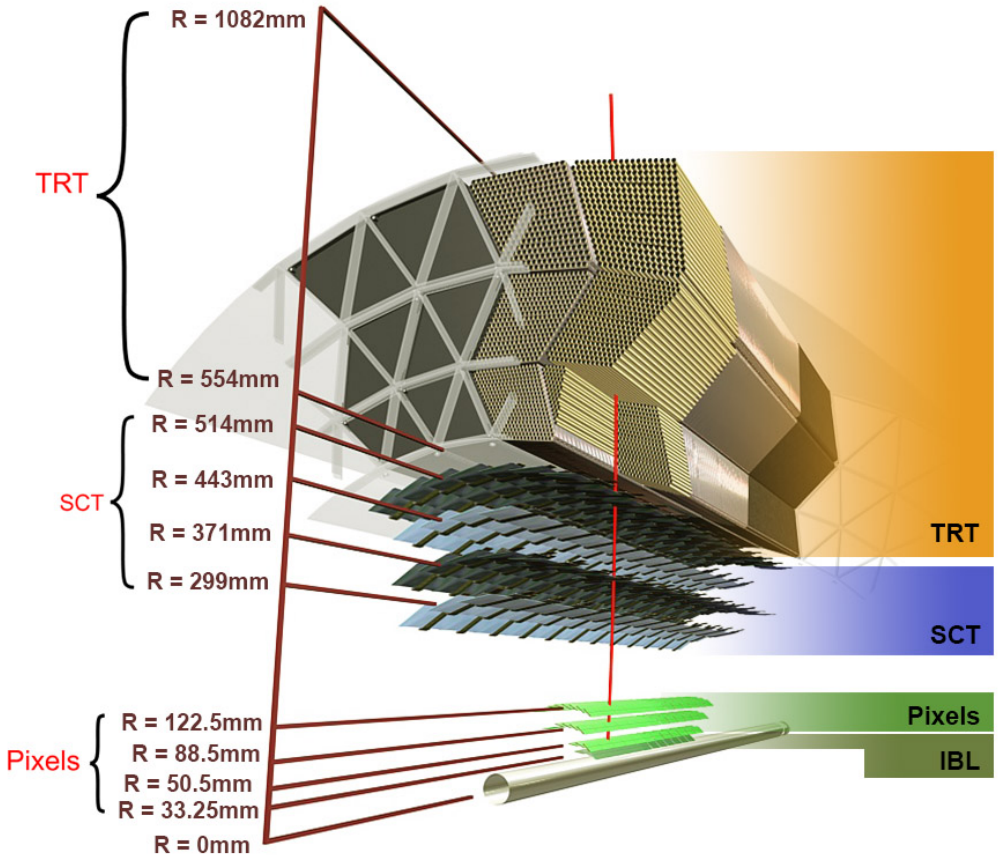


Figure 3.6.: Sensors and structural elements of the ATLAS Inner Detector including the IBL [52].

of 3.3 cm between the existing Pixel Detector and a new smaller radius beam-pipe. It consists of 12 million pixels with a typical size of 50 μm$\times$250 μm. The three outer barrel layers, each starting around 5 cm from the beam line, and the end-caps discs provide in total approximately 80 million readout channels of $50\times400\,\mu\text{m}^2$. The high precision of the pixel detector is vital for the measurement of impact parameters and for the reconstruction of primary interaction and secondary decay vertices especially in an environment with multiple interactions per bunch crossing (see Figure 4.2). The latter are needed in order to tag heavy-flavor quarks like b-quarks and τ-leptons via their decays.

The Semi-Conductor Tracker

The SCT is a silicon strip detector which contains four cylindrical layers in the barrel region and nine end-cap disks on each side covering the same region as the pixel detector up to $|\eta| < 2.5$. The innermost layer is at a radial distance of about 30 cm from the z-axis, the outermost layer at 51.4 cm. Each layer consists of two

strips build back to back and rotated against each other by 40 μrad with one set of strips running parallel to the beam axis. In the disks, the same arrangement is made with one set of strips running radially. The spatial resolution is 17 μm in $R-\phi$ and 580 μm in z for the barrel as well as in R for the end caps. Although the granularity is generally lower with about 6.3 million readout channels, the SCT provides better momentum resolution compared to the Pixel Detector due to the stronger curvature of charged particles. These particles are usually creating eight hit points in the stereo layers resulting in four three-dimensional track points. The SCT and the Pixel Detector sensors are operating at low temperatures from -5°C to -10°C to maintain the noise performance after damage due to radiation.

The Transition Radiation Tracker

The last and outermost component of the Inner Detector is the TRT [53] which contains around 300 thousand thin-walled straw-tubes providing on average 30 two-dimensional space-points with an intrinsic resolution of 120 μm in $(R-\phi)$ for charged particle tracks in a pseudorapidity region $|\eta| < 2$ for $p_T > 0.5$ GeV. The drift tubes have a diameter of 4 mm and are filled with a xenon based gas mixture. When a charged particle traverses the TRT it ionizes the gas inside the straws. The resulting free electrons drift towards the centrally tensioned wire where they are amplified and read out. In the barrel region the tubes are aligned parallel to the beam line and radially in the end caps. The TRT is an important component for the momentum measurement since the high number of hits and the larger track length compensate for the lower precision per point compared to the silicon detectors. In addition the tubes are interleaved with polypropylene fibers (barrel) and foil (end-caps) to create transition radiation, which may be emitted by highly relativistic charged particles as they traverse the material boundary. The number of transition-radiation X-ray photons depends on the Lorentz factor of the traversing particle and therefore its mass. Lower mass particles emmit more photons resulting in higher intensity. Thus, by applying two different thresholds, radiation from the very light electrons can be distinguished from, e.g. radiation emitted due to a pion, enabling the discrimination between different types of particles. At 80% detection efficiency for electrons, the misidentification rate for pions is roughly 10% in the range $4\,GeV < p_T < 20\,GeV$.

3.2.3 The Calorimenter System

The main purpose of the calorimeter system (cf. Figure 3.7) is to measure the energy of electrons, photons and jets, and to determine the missing transverse energy E_T^{miss}. It is located outside the solenoidal magnet that surrounds the Inner Detector. Since the interaction of electrons and photons with matter differs from that of hadronic particles, there are two basic calorimeter systems: an inner electromagnetic (EM) calorimeter and an outer hadronic calorimeter which cover a range of $|\eta| < 4.9$.

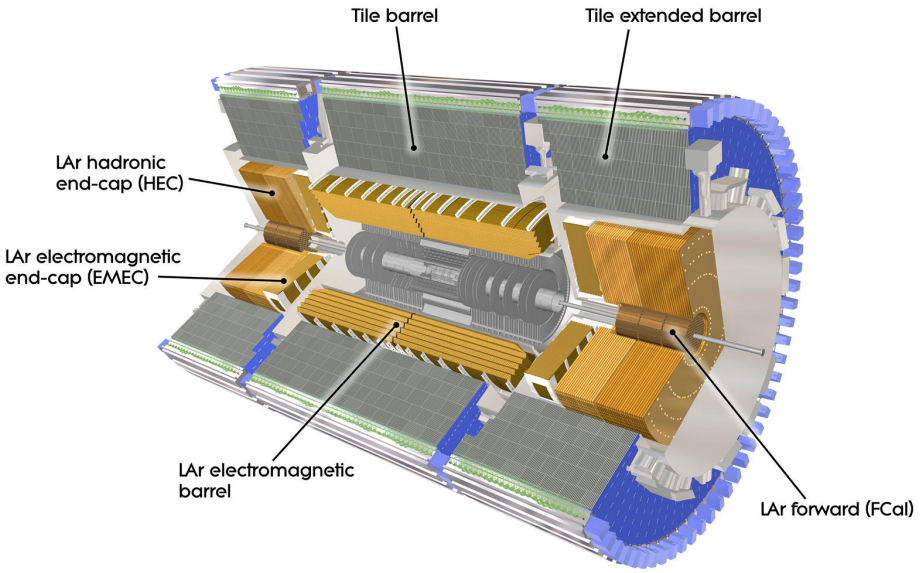


Figure 3.7.: Schematic view of the ATLAS calorimeter system [46].

The EM calorimeter with its high granularity is ideally suited for precision measurements of electrons and photons whereas the hadronic calorimeter has a coarser granularity which is sufficient to reconstruct hadronic jets and measure the missing transverse energy of the collision.

In general the resolution of a sampling calorimeter is a function of the energy and can be parameterized as

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C, \tag{3.2.4}$$

where N, S and C are η-dependent parameters. At low energies the noise term N dominates the resolution which is dependent on the occupancy of the calorimeter and thus on the average number of interactions per bunch crossing and also includes electronic noise. In an intermediate energy regime the stochastic or sampling term S becomes dominant. It contains stochastical fluctuations of the shower and depends mainly on the thickness of the absorber layers. The constant term C contains the contributions that is not dependent on the energy of the particle such as the signal loss in passive material and dominates the high energy regime. The total uncertainty on jet energy measurements in the barrel region is around 2-5% over a broad range of the p_T spectrum.

The Electromagnetic Calorimeter

The EM calorimeter is divided into a barrel part and two end cap wheels enclosing the Inner Detector. The barrel is split into two halves of 3.2 m length separated by a 4 mm gap at $z = 0$, covering $|\eta| < 1.475$. The total thickness of 53 cm exceeds $22X_0$, where X_0 is the electromagnetic radiation length over which a high-energy electron loses all but $1/e$ of its energy mainly by bremsstrahlung. The two end caps expand the acceptance within $1.35 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$, each wheel having a thickness of 63 cm which translates to $24X_0$ and a radius of 2 m. The material used for the absorbers is lead constructed in a so called accordion shape, which provides optimal ϕ-symmetry without azimuthal cracks. As active material liquid argon (LAr) is used, chosen for its radiation hardness, signal speed and linearity. For the extensive survey of the shower shapes, the barrel EM calorimeter is composed of three longitudinal layers with decreasing granularity. The first layer is called the presampler and is made up of very fine slices in η with a granularity of $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$, allowing for a distinction between photons and neutral pions as well as electrons and charged pions. The second layer collects most of the energy of electrons and photons with a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ and a thickness of $16X_0$. The outer third layer is only $2X_0$ thick and slightly coarser with a granularity of $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ and is used to measure the leftover deposition intensity of the highest energy electrons and thus the depth to which the shower has penetrated the calorimeter. The end caps consist of only two such layers with reduced granularity. Overall there are about 175 thousand readout channels.

The number of lead/argon interfaces define the sampling term of the energy resolution and is 8 to $11\%/\sqrt{E[\mathrm{GeV}]}$ depending on the rapidity. The noise term is about $350 \times \cosh\eta$ MeV for a typical cluster in the barrel for a mean number of interactions per bunch crossing of $\langle\mu\rangle = 20$. At high energies, the relative energy resolution approaches the constant term, whose design value is 0.7%.

The Hadronic Calorimeter

The hadronic calorimeter is divided into different subcomponents. The tile calorimeter in the barrel region covers $|\eta| < 1$ and has an extended barrel covering $0.8 < |\eta| < 1.7$. As a sampling calorimeter it uses steel as absorbing material and scintillating tiles as the active material. The light of the scintillators is collected at the edges of each tile and read out into two photomultiplier tubes by wavelength-shifting fibers. The LAr hadronic end-cap calorimeter consists of two independent wheels on both detector sides extending to $|\eta| < 3.2$. It shares the same cryostat and liquid argon supply of the EM calorimeter but uses copper as absorber material. Also the LAr forward calorimeter which reaches up to $|\eta| < 4.9$, shares the liquid argon supply with the EM calorimeter. It consists of three layers, the first using copper as absorbing material is mainly used for electromagnetic interaction and the two others using tungsten are optimized for the detection of hadronic showers.

The thickness of the hadronic calorimeter reaches approximately 10λ, where the hadronic interaction length λ is the mean path length required to reduce the numbers of relativistic charged particles by the factor $1/e$ as they pass through matter.

The expected resolution is $\sigma_E/E = 10\%\sqrt{E[\text{GeV}]}\oplus 3\%$ for the barrel and endcaps and $\sigma_E/E = 100\%\sqrt{E[\text{GeV}]} \oplus 10\%$ for the forward calorimeter.

3.2.4 The Muon Spectrometer

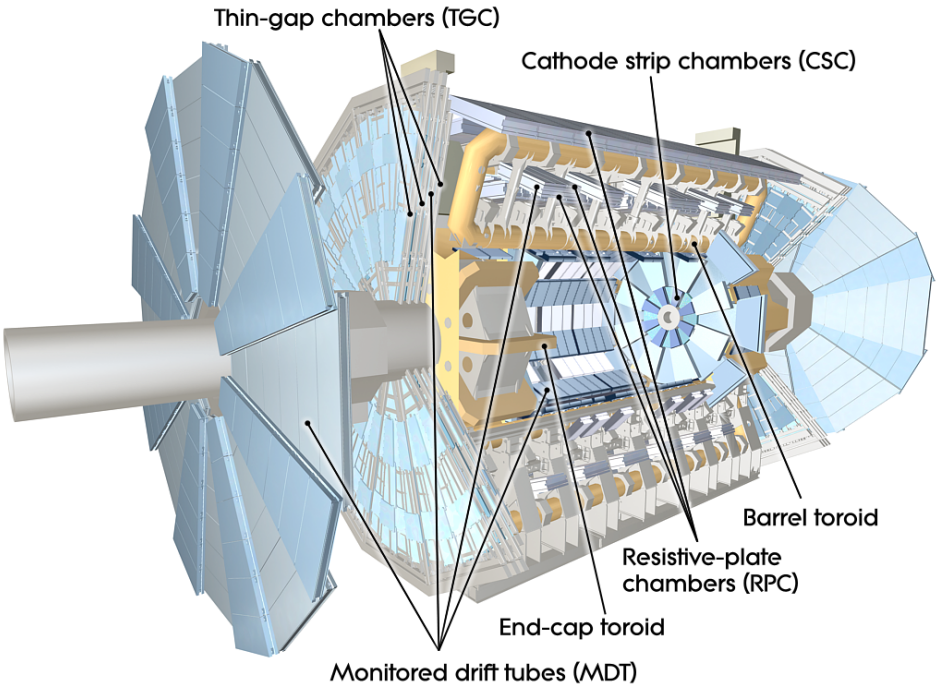


Figure 3.8.: Schematic layout of the muon spectrometer [46].

Due to the high interaction length of the calorimeters, only muons that are minimally ionizing particles are expected to pass through the detectors. To measure the transverse momentum of the muons precisely their trajectory is deflected by a magnetic field and their tracks are measured in at least three chambers. The magnetic field is generated by a system of three superconducting air-core toroid magnets capable of reaching magnetic fields up to 4 T. Each of the three toroids consists of 8 coils assembled radially and symmetrically around the beam axis. The large barrel toroid provides a magnetic field up to $|\eta| < 1.4$. The two smaller end-cap magnets bend the muon tracks in $1.6 < |\eta| < 2.7$ and are inserted in the barrel toroid, lined up with the central solenoid and rotated by 22.5° with respect to the barrel toroid coils to optimize bending power due to radial overlap.

The transition region $1.4 < |\eta| < 1.6$ is covered by the combination of the barrel and the end-cap toroid.

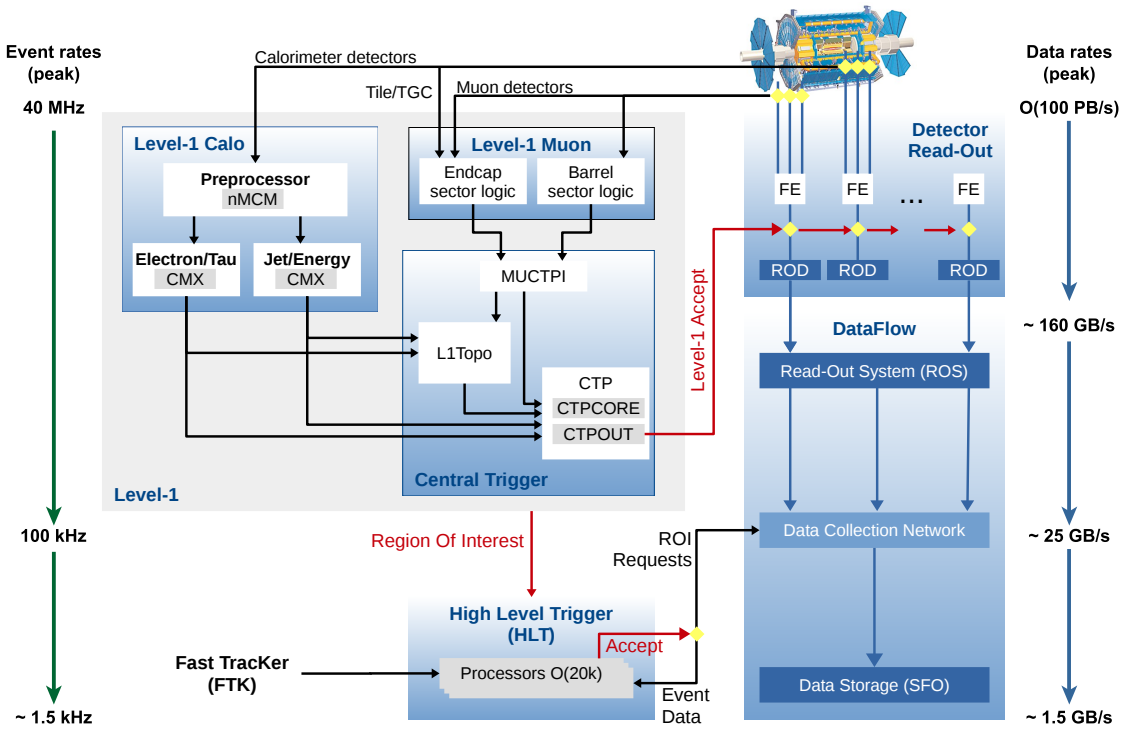
The precise measurement of the muon track coordinates are performed by two different detector parts depicted in Figure 3.8. The main part covering the barrel region up to $|\eta| < 2.7$ are the Monitored Drift Tubes (MDT). 1088 drift chambers are arranged perpendicular to the beam-axis, each chamber carrying two multi-layers of three to four drift tube layers. Those tubes are 3 cm in diameter and filled with an argon based gas mixture at a pressure of 3 bar and a 50 μm diameter tungsten rhenium wire serving as anode. When the passing muons are ionizing the gas, the induced electron avalanche is collected at the anode and the signal timing as function of the drift radius gives the muon's closest approach to the wire providing the position in one dimension with an average resolution of 35 μm per chamber. The range $2.0 < |\eta| < 2.7$ of the muon system is covered by 32 four-layered high granularity Cathode Strip Chambers (CSC) which are multiwire proportional chambers. The wires are radially oriented and the cathodes are divided into orthogonal strips to measure both coordinates simultaneously with a precision of 40 μm in the bending and 5 mm in the transverse plane.

In addition to the tracking chambers, fast muon chambers used for triggering are installed, that deliver signals within 15-25 ns after the passage of a particle. In the barrel region $|\eta| < 1.05$ a total of 544 resistive plate chambers (RPC) are located below and above the MDTs, which are wireless chambers with anode and cathode plates filled with an organic gas mixture. The end caps in $1.05 < |\eta| < 2.7$ are covered by 3 588 Thin Gap Chambers (TGC), which are multiwire proportional chambers where the wire-to-cathode distance is smaller than the wire-to-wire distance. Those trigger chambers are providing not only trigger information but also bunch crossing information and azimuthal muon coordinates, orthogonal to the tracking chambers.

3.2.5 The Trigger System

With a collision rate of 40 MHz and approximately 100 million readout channels which sums up to the order of 100 PB data per second, it is impossible to store every collision event. Therefore, a trigger system is responsible for deciding whether or not to keep a given event for later studies. It consists of a hardware-based first level trigger (Level-1 or L1) and a software based high level trigger (HLT). Figure 3.9 shows a schematic overview of the ATLAS Trigger and Data Acquisition (DAQ) system used in Run-2.

The Level-1 trigger takes calorimeter and muon detector information in a coarse granularity and uses custom made electronics to find regions of interest (RoIs) in the detector. It reduces the event rate from 40 MHz to about 100 kHz and



ROI: region of interest – FE: front end – ROD: readout driver
nMCM: new multi chip module – CMX: common extended merger module
CTP: central trigger processor – MUCTPI: muon interface to CTP

Figure 3.9.: The ATLAS trigger and data acquisition system (DAQ) in Run-2 (modified from [54]).

determines its decision in only 2.5 μs. The L1 system in Run-2 consists of the L1 calorimeter trigger system (L1Calo), the L1 muon trigger system (L1Muon), new L1 topological trigger modules (L1Topo) and the Central Trigger Processors (CTP). L1Calo reconstructs energy depositions in the calorimeters and finds candidates for electrons, photons, taus, jets and missing transverse energy. Muon candidates are reconstructed in L1Muon and the L1Topo modules combines information of both subsystems into variables that are used for additional selection criteria such as topological or angular selections, kinematic selections or sums of objects. The final trigger decision of the L1 trigger provides the CTP.

The L1 RoIs are forwarded to the HLT where fast algorithms are used to select events using either the RoIs or the full detector information of the event. There, a further reduction of the event rate from 100 kHz to ca. 1.5 kHz is achieved within a processing time of about 200 ms. Two types of algorithms are used to either build objects like tracks and clusters or apply selection cuts on different kinematic variables like transverse momentum of the tracks or invariant masses.

4 Physics Object Definition and Reconstruction

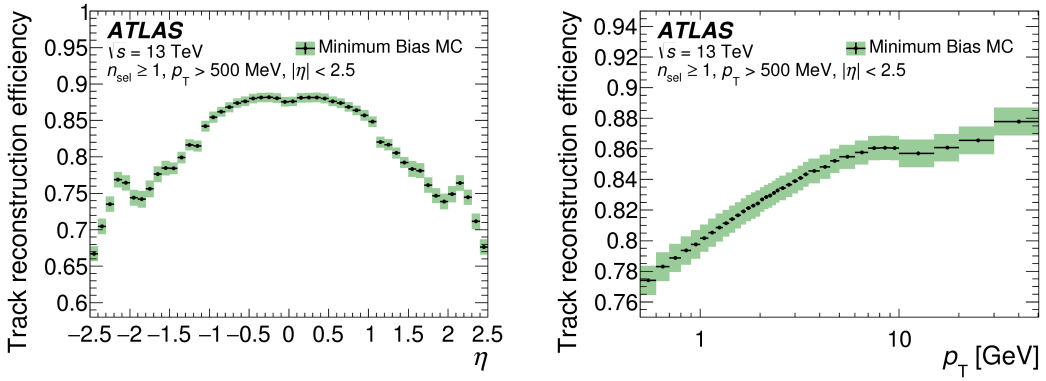
The transition from raw detector information to physical objects is a major prerequisite for any physics analysis. Therefore, the methods and algorithms in the reconstruction stage have to be reliably robust to maintain optimal performance. The ATLAS reconstruction software uses a wide variety of different 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. While the first part deals with the reconstruction of tracks and vertices, the second part passes on to the identification of the physics objects, including their four-momenta and the complete event kinematics.

4.1 Tracks

Charged particles produced in pp collisions traversing the Inner Detector deposit small fractions of energy, the so-called hits, in the Pixel, SCT and TRT detectors. From these hits the trajectories of the particles (tracks) are reconstructed up to a pseudorapidity region of $|\eta| < 2.5$ and in an environment of up to 1000 tracks per bunch crossing[1]. Analyzing the circular deflection in the transverse plane due to the axial magnetic field the transverse momentum of the particle can be determined. The track reconstruction algorithm is outlined in the following [55,56]. First clusters are assembled that consists of groups of pixels and strips in the Pixel and SCT detectors respectively, where the energy deposit yields a charge above the threshold. These clusters form three dimensional space points. Then a combinatorial Kalman filter using track seeds formed from sets of three space-points is executed to build track candidates by incorporating additional space-points which are compatible with the preliminary trajectory. Also momentum and impact parameter requirement

[1] For heavy ion collision the tracks per bunch crossing is in the order of $\mathcal{O}(10\,000)$.

are applied. This results in a very high efficiency for reconstructing primary particles and for removing non-physical, purely random tracks. An ambiguity solver processes these individual track candidates in order of their track score, which expresses the likelihood to correctly represent the trajectory of a charged primary particle. This algorithm dissolves clusters assigned to multiple track candidates with the help of an artificial neural network and rejects candidates that fail to meet basic quality criteria, such as minimal transverse momentum, impact parameters or certain numbers of hits in the different subdetectors. Tracks that pass successfully the ambiguity solving stage and are within the coverage of the TRT are then extended to this subdetector which increases the momentum resolution significantly by exploiting the longer lever arm. At the end a high-resolution fit is performed for each track using all available information, where additional neural networks determine the position and uncertainty of each cluster. In Figure 4.1 the track reconstruction efficiency as function of the p_T and η of the track is shown.



(a) Reconstruction efficiency vs. η. **(b)** Reconstruction efficiency vs. p_T.

Figure 4.1.: Inner Detector track reconstruction efficiency defined as number of matched tracks compared to number of generated charged particles versus track η and track p_T as predicted by single-particle simulation. The statistical uncertainties are shown as black vertical bars, the total uncertainties as green shaded areas [57].

4.2 Vertices

An identified interaction point of two colliding protons is called primary vertex. Usually within one bunch crossing more than one interaction takes place simultane-

ously which is referred to as *pile-up*[2] (cf. Figure 4.2) and results in several primary vertices. The reconstruction of these vertices is organized in two steps [58]:

First quality requirements are applied on tracks to remove tracks originating from secondary vertices. A vertex seed is found at the global maximum in the distribution of the longitudinal impact parameter z_0. Using the seed position and the associated tracks an adaptive vertex fitting algorithm finds the position of the primary vertex. Tracks incompatible with this vertex are used to seed a new primary vertex. In a second step a vertex fitting algorithm reconstructs the vertex positions and also refits associated tracks constraining them to originate from the reconstructed interaction point.

The efficiency of this procedure for low number of interactions per bunch crossing is shown in Figure 4.3. Raising number of simultaneous collisions increases the number of reconstructed vertices (cf. Figure 4.4). To ensure continued performance in vertex reconstruction new algorithms are being developed [59].

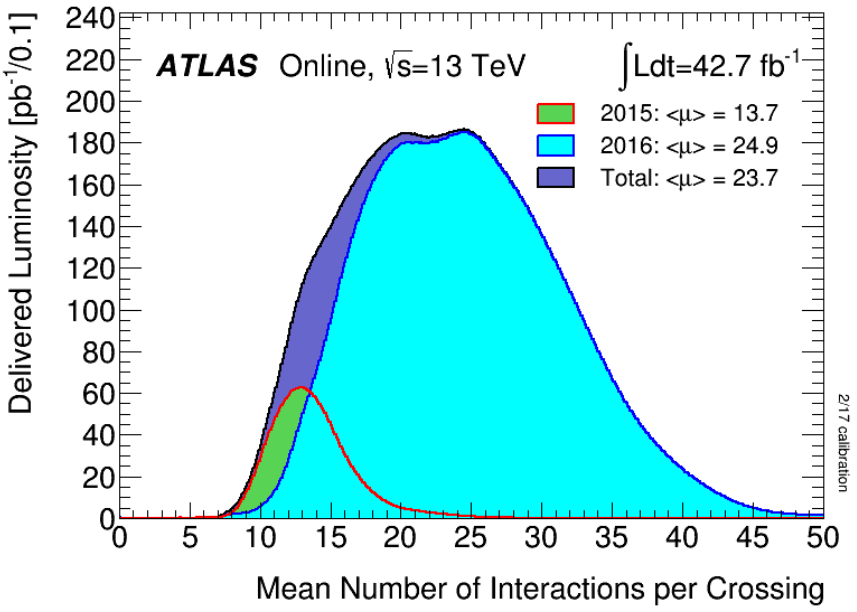


Figure 4.2.: Average number of interactions per bunch crossing for ATLAS in 2015 and 2016 data taking [60].

[2]One differentiates between in-time and out-of-time pile-up. While in-time pile-up arises from additional *pp* interactions in the current bunch-crossing, out-of-time pile-up refers to energy deposits in the calorimeter from previous and following bunch crossings relative to the triggered event.

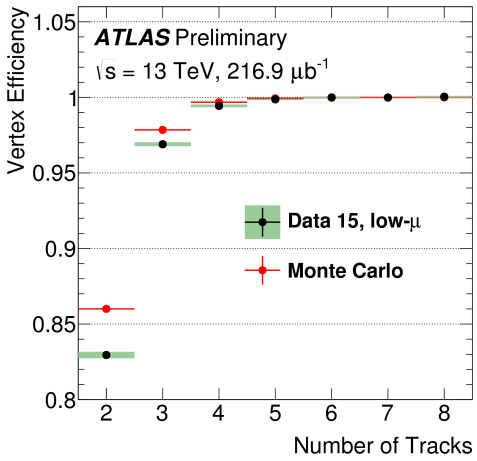


Figure 4.3.: Vertex reconstruction efficiency as a function of the number of tracks in data with low number of interaction per bunch-crossing (low-μ) [61].

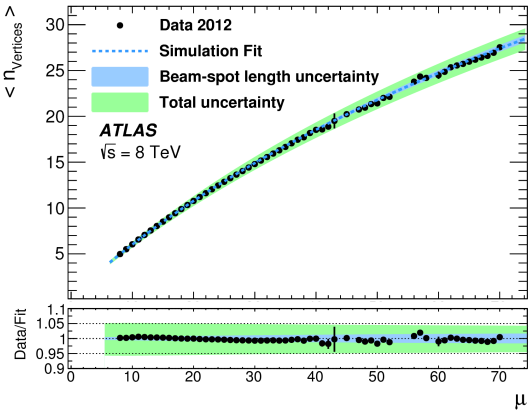


Figure 4.4.: Distribution of the average number of reconstructed vertices $\langle n_{\text{vertices}} \rangle$ as a function of the number of interactions per bunch crossing μ [62].

4.3 Electrons

A good understanding of how to reconstruct and identify electrons is crucial especially for the analysis presented in Chapter 6 where electrons are one of the main objects in the final state. A brief overview over the different steps for electron reconstruction and identification is given in the following [63, 64].

Electron Reconstruction

Electrons are reconstructed in the central region $|\eta| < 2.47$ by matching energy deposits in the electromagnetic calorimeter to tracks in the Inner Detector. First, a sliding window algorithm [65] with a window size of 3×5 in units of 0.025×0.025 in $\eta \times \phi$ searches for electron cluster seeds. A cluster is seeded when the total transverse energy in the energy sum of the longitudinal tower over the calorimeter layers exceeds a transverse energy of $E_T > 2.5\,\text{GeV}$. These clusters are then loosely matched to extrapolated tracks from the Inner Detector using the distance in η and ϕ. A Gaussian Sum Filter refits the tracks that have a significant number of precision hits in the silicon detector to account for energy-loss due to bremsstrahlung of the electron. If no tracks are associated with the energy cluster the electron candidate is removed and considered to be a photon. Alternatively if several tracks fulfill the matching condition, one of the tracks is chosen to be the primary track based on the hits in the pixel detector and the ΔR distance between cluster and track. This primary track is used to determine the kinematics and charge of the

electron. Finally the electron cluster is refit using 3×7 (5×5) longitudinal towers of cells in the barrel (endcaps) of the electromagnetic calorimeter and calibrated with multivariate techniques to determine the total electron energy. The entire procedure is sketched in Figure 4.5 and the reconstruction efficiency as function of η and E_T is shown in Figure 4.6.

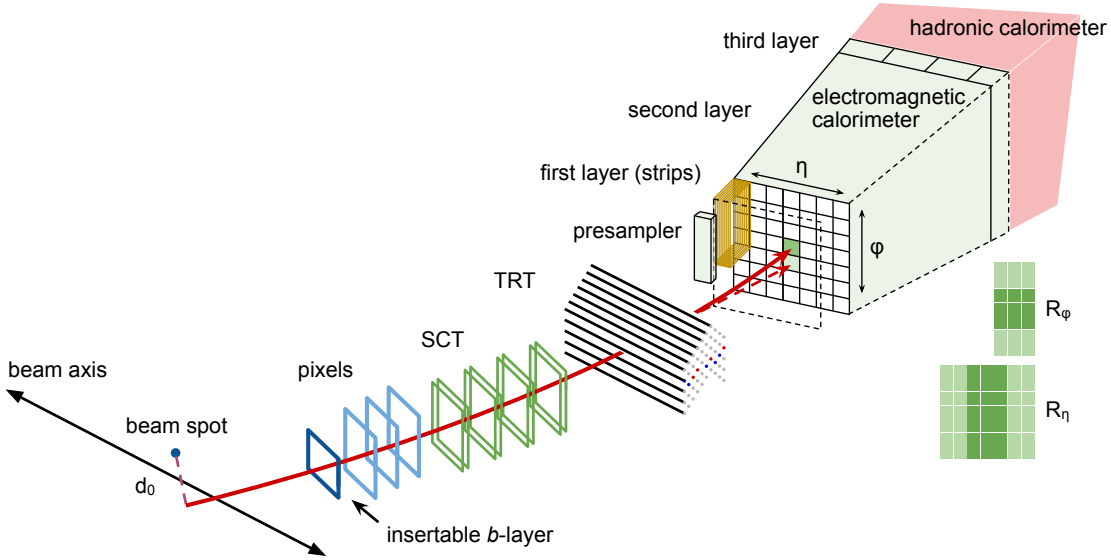
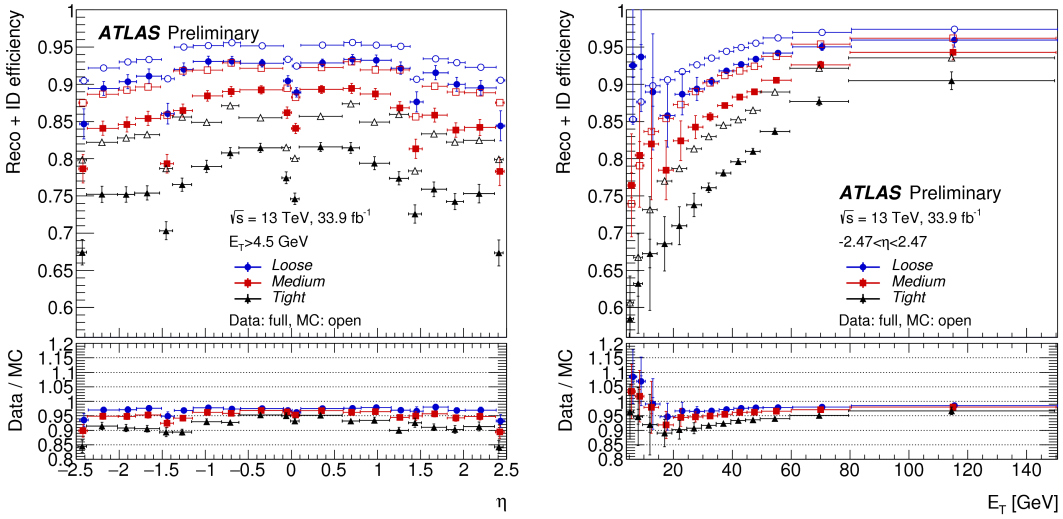


Figure 4.5.: Schematic view of the electron reconstruction and identification [64].

Electron Identification

Reconstructed electron candidates may not be definitely real signal electrons but can originate e.g. from hadronic jets or non-prompt electrons descending predominately from photon conversions and heavy flavor hadron decays. To discriminate against these backgrounds, different sets of identification criteria are defined based on variables describing the transverse and longitudinal shower profiles, information from the TRT, track-cluster matching related quantities, track properties, and variables measuring bremsstrahlung effects. The baseline identification algorithms are likelihood-based (LH) multivariate analysis methods with typically three operating points at different level of background rejection and signal efficiency. These are referred to, in order of increasing background rejection, as `LooseLH`, `MediumLH` and `TightLH`. At $E_T = 25\,\text{GeV}$ the identification efficiencies for the signal range from 78% to 90% depending on the operating point and increase with E_T, whereas the background efficiency ranges from 0.3% to 0.8% and decrease for higher transverse energies (cf. Figure 4.7).



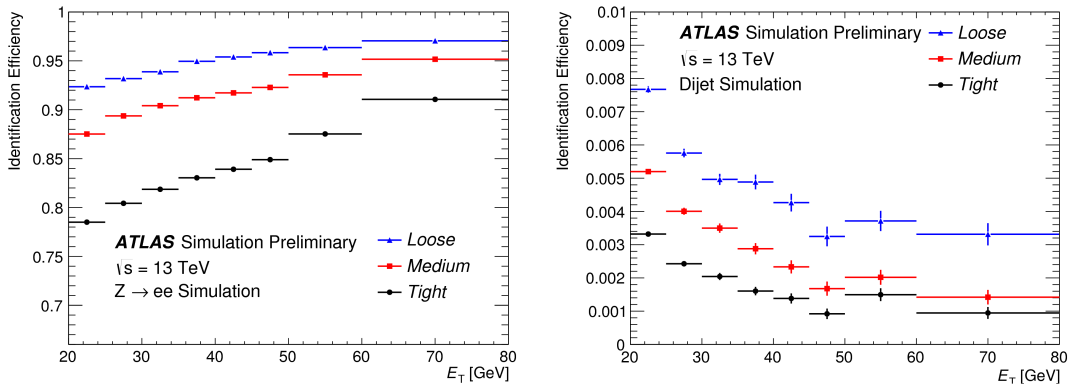
(a) Reconstruction efficiency vs. η. **(b)** Reconstruction efficiency vs. E_T.

Figure 4.6.: Electron reconstruction and identification efficiency measurements using an integrated luminosity of 33.9 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data recorded in 2016 [66].

Electron Isolation

To further discriminate between signal and background, electron isolation critera can be applied [64]. Specific isolation variables quantify the energy of particles measured around the electron candidate and allow to disentangle prompt electrons (e.g. from $W \to e\nu$ or $Z \to ee$ decays) from other, non-isolated electron candidates. Calorimeter and track based quantities can be distinguished:

- The sum of transverse energies of topological clusters within a cone of $\Delta R = 0.2$ around the candidate electron cluster is denoted as the *calorimeter isolation energy* $E_T^{\text{cone0.2}}$. Only positive-energy clusters are considered and the E_T around the rectangular cluster of size $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ centered around the electron cluster barycentre is subtracted. Corrections are applied to account for electron energy leakage outside this cluster and pile-up.
- The sum of all tracks satisfying quality requirements within a cone of $\Delta R = \min(0.2, 10\,\text{GeV}/E_T)$ around the candidate electron track origintated from the primary vertex of the hard collision is used as *track isolation* $p_T^{\text{varcone0.2}}$. The associated track from the electron and additional tracks from converted bremsstrahlung photons are subtracted and certain quality requirements are applied, such as $E_T > 1$ GeV, specific number of hits in the silicon detectors and $|\Delta z_0 \sin\theta| < 3$ mm.



(a) Signal identification efficiency in simulated $Z \to ee$ events.

(b) Background identification efficiency in simulated dijets events.

Figure 4.7.: The efficiency to identify electrons from $Z \to ee$ decays and the efficiency to identify hadrons as electrons (background rejection) estimated using simulated dijet samples [64].

Efficiency targeted operation points are defined with varying requirements on $E_T^{\text{cone0.2}}/E_T$ and $p_T^{\text{varcone0.2}}/E_T$ which are summarized in Table 4.1.

Table 4.1.: Efficiency targeted electron isolation operation points. [64].

	Efficiency		
Operation point	calorimeter isolation	track isolation	total efficiency
`LooseTrackOnly`	-	99%	99%
`Loose`	99%	99%	∼98%
`Tight`	96%	99%	∼95%
`Gradient`	$0.1143\% \times E_T/\text{GeV} + 92.14\%$	$0.1143\% \times E_T/\text{GeV} + 92.14\%$	90/99% at 25/60 GeV
`GradientLoose`	$0.057\% \times E_T/\text{GeV} + 95.57\%$	$0.057\% \times E_T/\text{GeV} + 95.57\%$	95/99% at 25/60 GeV

4.4 Muons

A second important ingredient in the analysis presented in Chapter 6 are muons. As minimal ionizing particles, they deposit only a very small fraction of their energy in the calorimeters. Therefore, approximately 96% of muons are reconstructed [67,68] as *combined muons* by fitting hits from the muon spectrometer with tracks from the Inner Detector taking the small energy loss in the calorimeter into account. The remainder are formed by tagging Inner Detector tracks with muon signatures in the calorimeter or the muon spectrometer which is mostly relevant for low-p_T

muons or muons at $\eta \sim 0$, where the muon detector is only partially equipped because of the needed space for the services of the inner subdetectors.

The identification of muons is performed by applying quality requirements to suppress background and to select prompt muons with high efficiency and a robust momentum measurement. The main source of background are in-flight decays of charged hadrons like pions or kaons which feature a characteristic kink in the reconstructed track. Quantities comparing the momentum and charges from the Inner Detector and the muon spectrometer and requiring sufficient hits in the single subdetectors are utilized to separate background from signal muons. Four identification selections are provided to address the specific needs of different physics analyses, that are `Loose`, `Medium`, `Tight` and `High-`$\mathtt{p_T}$. The efficiencies in two different transversal momentum regions can be seen in Table 4.2.

Table 4.2.: Muon identification efficiency for the four criteria in two p_T regions for muon candidates in the central region $|\eta| < 2.5$ determined from W-decays for prompt muons and hadrons decaying in flight and misidentified as prompt muons using a $t\bar{t}$ MC sample [68].

	$4 < p_T < 20\,\text{GeV}$		$20 < p_T < 100\,\text{GeV}$	
Selection	ϵ_μ^{MC} [%]	$\epsilon_{\text{Hadrons}}^{\text{MC}}$ [%]	ϵ_μ^{MC} [%]	$\epsilon_{\text{Hadrons}}^{\text{MC}}$ [%]
`Loose`	96.7	0.53	98.1	0.76
`Medium`	95.5	0.38	96.1	0.17
`Tight`	89.9	0.19	91.8	0.11
`High-`$\mathtt{p_T}$	78.1	0.26	80.4	0.13

To seperate muons from other particles in the event, similar quantities are used as in the electron isolation: the calorimeter-based isolation variable E_T^{cone20} and the track-based isolation variable $p_T^{\text{varcone30}}$, which is identically defined as $p_T^{\text{varcone20}}$ used in the electron isolation but using a cone size of $\Delta R = 0.3$. Several operation points are defined with different efficiencies or fixed criteria. They are summarized in Table 4.3.

Table 4.3.: Definition of the muon isolation operation points [68].

Isolation	Discriminating variable(s)	Definition
`LooseTrackOnly`	$p_T^{\text{varcone30}}/p_T^\mu$	99% efficiency constant in η and p_T
`Loose`	$p_T^{\text{varcone30}}/p_T^\mu$, $E_T^{\text{cone20}}/p_T^\mu$	99% efficiency constant in η and p_T
`Tight`	$p_T^{\text{varcone30}}/p_T^\mu$, $E_T^{\text{cone20}}/p_T^\mu$	96% efficiency constant in η and p_T
`Gradient`	$p_T^{\text{varcone30}}/p_T^\mu$, $E_T^{\text{cone20}}/p_T^\mu$	$\geq 90(99)\%$ efficiency at 25 (60) GeV
`GradientLoose`	$p_T^{\text{varcone30}}/p_T^\mu$, $E_T^{\text{cone20}}/p_T^\mu$	$\geq 95(99)\%$ efficiency at 25 (60) GeV
`FixedCutTightTrackOnly`	$p_T^{\text{varcone30}}/p_T^\mu$	$p_T^{\text{varcone30}}/p_T^\mu < 0.06$
`FixedCutLoose`	$p_T^{\text{varcone30}}/p_T^\mu$, $E_T^{\text{cone20}}/p_T^\mu$	$p_T^{\text{varcone30}}/p_T^\mu < 0.15$, $E_T^{\text{cone20}}/p_T^\mu < 0.30$

4.5 Jets

Jets have been introduced in Section 2.4.2 as collimated bundles of hadrons emerging from the fragmentation of partons. These parton showers deposit most of their energy in the hadronic calorimeter and are usually accompanied by hits in the electromagnetic calorimeter and Inner Detector due to charged particles. They have a characteristic conical shape and are grouped to so-called *topoclusters* that are used as input in the reconstruction of jets.

Topoclusters

Topoclusters [69] are build from energy deposits in the calorimeter to extract the significant signal from a background of electronic noise and other sources of fluctuation such as pile-up. They are not expected to contain the entire response to a single particle all the time, but can also represent only a fractional response to a single particle, several particles or a combination of merged showers, depending on the incoming particle types, energies, spatial separations and cell signal formation. Seeds for topoclusters are found if the energy of a calorimeter cell, that is measured in the EM scale, exceeds the expected average noise by a factor of 4. Direct neighboring cells are included into a topocluster as long as the signal over noise ratio is over a threshold of 2. The first neighboring cells with a smaller signal-to-background ratio are still included in the cluster, but the iteration stops there. Cluster with several local maximum cells that have an energy above 500 MeV are split to account for the presence of two or more particles injecting energy into the calorimeter in close proximity. Around these maxima new clusters are formed using only the cells that are present in the original cluster. Cells that end up in more than one cluster are included in both clusters with weights that depends on the distance to the centroid and the cluster energy.

The topoclusters are corrected using the *local hadronic cell weighting* (LCW) calibration, which uses so-called *cluster moments*. These cluster moments are reconstructed observables describing the shape of a topocluster and its internal signal distribution like the shower axis, center of gravity of the cluster, or geometrical center of a calorimeter cell in the cluster. The calibration is applied for each topocluster separately, and corrects for the effects of the non-compensating calorimeter response to hadrons, accidental signal losses due to the clustering strategy, and energy lost in inactive material in the vicinity of the topocluster.

Jet Reconstruction

These positive-energy topoclusters are used as inputs in the reconstruction of jets [70]. A proper definition of a jet should fulfill several requirements. First such a jet algorithm should be collinear safe, that is insensitive to a collinear splitting of a hard particle, and infrared safe, i.e. independent of soft gluon radiation. In

addition it should not be sensitive to non-perturbative effects like hadronization or the underlying event due to multiple interactions per bunch crossing. There are several algorithms available which can be divided into sequential clustering algorithms (k_T, anti-k_T and Cambridge Aachen) and cone algorithms, which have rigid circular boundaries but generally suffer from infrared or collinear safety issues. All sequential clustering algorithms are based on the distance measure d_{ij} between two particles i and j

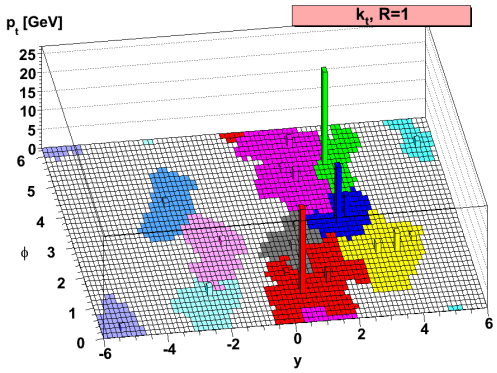
$$d_{ij} = \min\left(k_{T,i}^{2p}, k_{T,j}^{2p}\right) \frac{\Delta R_{ij}}{R}, \tag{4.5.1}$$

where $k_{T,i}$ is the transverse momentum of particle i, $\Delta R_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, with y_i the rapidity and ϕ_i the azimuthal angle of particle i, and R the radius parameter. The distance between any particle i and the beam axis d_{iB} is given by

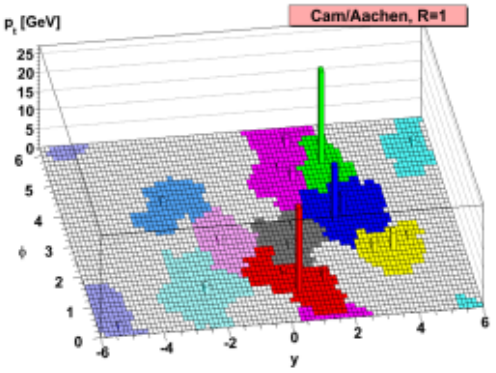
$$d_{iB} = k_{T,i}^{2p}. \tag{4.5.2}$$

First, the algorithm finds the smallest distance d_{ij} and combines these two particles by summing their four momenta and then updates the distances with the new combined particle. If the smallest distance is d_{iB} then the particle i is removed and associated to a jet. This process is iterated and repeated until all particles are clustered into jets. The Parameter R scales the distances d_{ij} with relation to d_{iB} such that any pair of final jets a and b are at least separated by $\Delta R_{ab} = R^2$. The Parameter p governs the relative power of energy versus geometrical scales. The k_T algorithm has $p = 1$, Cambridge/Aachen $p = 0$ and the main jet algorithm used in this analysis, the anti-k_T algorithm [71] uses $p = -1$. The behaviors of these jet algorithms are illustrated in Figure 4.8. Here, a sample parton-level event was modified with random soft particles and then clustered with the different algorithms. The jagged borders of the jets derived from the k_T and Cambridge/Aachen algorithms shows the dependence of the randomness of the added soft particles. In the anti-k_T algorithm, the hard jets are all conical shaped with a radius $R = 1.0$ and less dependent on the added ghosts. Only the softer jets have more complex shapes. For ATLAS the typical distance parameter is set to $R = 0.4$. A well performing software implementation for these jet finding algorithms is provided by FASTJET [72].

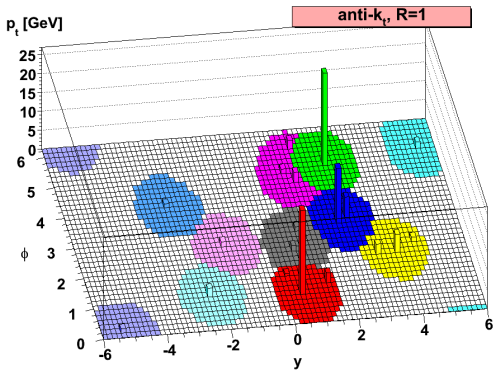
The reconstructed jets are then corrected to account for additional effects that influence the jet kinematics. First, a jet is corrected to point back to the identified hard-scatter vertex. Next, the effect of pile-up is removed using an area-based subtraction procedure and residual corrections. The jet energy is then calibrated by applying a p_T and η-dependent correction derived from MC simulation. Further corrections are applied to the jets which reduce the difference in response between



(a) k_T jet reconstruction algorithm.



(b) Cambridge/Aachen jet reconstruction algorithm.



(c) Anti-k_T jet reconstruction algorithm.

Figure 4.8.: A comparison of the resulting jet-shape of the most common jet reconstruction algorithms used at ATLAS, showing an event reconstructed with a jet cone radius of $R = 1.0$ for the k_T, Cambridge/Aachen and anti-k_T jet reconstruction algorithm in the $(y - \phi)$-plane [71].

gluon and quark initiated jets and also correct for jets which are not fully contained in the calorimeter.

Jet Vertex Tagging

The suppression of pile-up jets from hard scatter jets can be achieved using a multivariate combination of two variables in the *jet-vertex-tagger* (JVT) [73]. An efficient track based variable is the *jet vertex fraction* (JVF) which is defined as the scalar p_T sum of the track that are associated with the jet and originate from

the hard-scatter vertex (PV_0) divided by the scalar p_T sum of all associated tracks from primary vertices (PV_n) due to pile-up interactions in the same bunch crossing:

$$\mathrm{JVF} = \frac{\sum_k p_T^{\mathrm{track}_k}(\mathrm{PV}_0)}{\sum_l p_T^{\mathrm{track}_l}(\mathrm{PV}_0) + \sum_{n\geq 1}\sum_l p_T^{\mathrm{track}_l}(\mathrm{PV}_n)}. \tag{4.5.3}$$

The average scalar p_T sum from pile-up tracks is dependent on the number of pile-up tracks per event $n_{\mathrm{track}}^{\mathrm{PU}}$ and its linear increase is corrected in the variable

$$\mathrm{corrJVF} = \frac{\sum_k p_T^{\mathrm{track}_k}(\mathrm{PV}_0)}{\sum_l p_T^{\mathrm{track}_l}(\mathrm{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_T^{\mathrm{track}_l}(\mathrm{PV}_n)}{k\cdot n_{\mathrm{track}}^{\mathrm{PU}}}}, \tag{4.5.4}$$

where the scaling factor k is set to 0.01 to roughly reflect the slope of the dependence on $n_{\mathrm{track}}^{\mathrm{PU}}$.

The second variable used in the JVT is R_{pT} which is defined as the scalar p_T sum of all tracks associated with the jet and originate from the hard-scatter vertex divided by the fully calibrated and pile-up corrected jet p_T:

$$R_{pT} = \frac{\sum_k p_T^{\mathrm{track}_k}(\mathrm{PV}_0)}{p_T^{\mathrm{jet}}}. \tag{4.5.5}$$

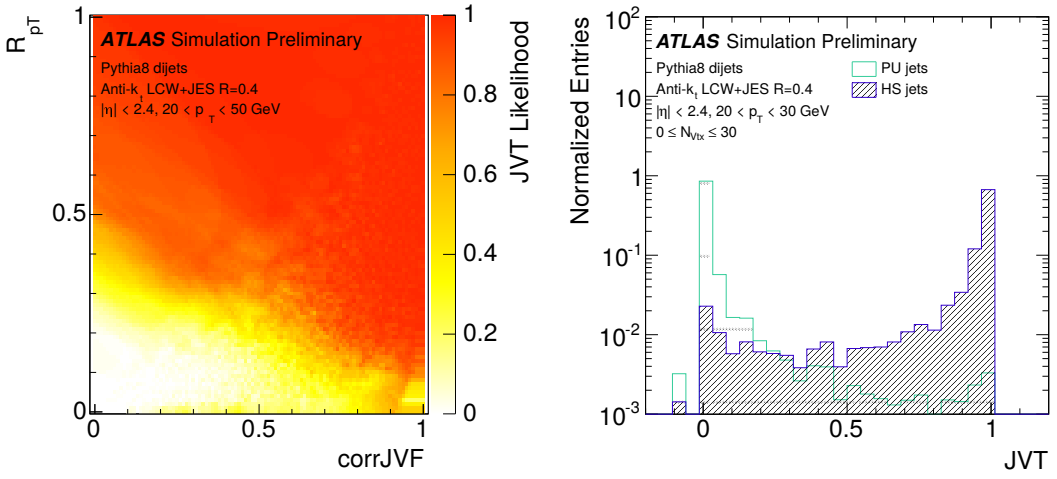
Pile-up jets peak at $R_{pT} = 0$ and steeply fall for larger values whereas for hard-scatter jets R_{pT} has the meaning of a charged p_T fraction and its mean value and spread is larger than for pile-up jets.

The JVT is constructed using both variables as a two-dimensional likelihood, which can be seen in Figure 4.9 using dijet MC events.

Furthermore, in the forward direction $|\eta| > 2.5$ where no tracking and vertex informations are available the ForwardJVT helps additionally to reduce pile-up jets [74] exploiting jet shapes and topological jet correlations in pile-up interactions. The most important inputs are the maximum JVT value, $\mathrm{JVT}_{\mathrm{max}}$, to reject central hard-scatter jets and the minimum ΔR_{pT} requirement, to ensure the selected pile-up jets are QCD pile-up jets.

B-tagging

Jets that arise from bottom-quark hadronization and decay (b-jets) are present in a wide range of physics processes of interest, such as the decay of top quarks. The ability to accurately identify b-jets is vital in separating these channels from processes involving jets from gluons and light-flavor quarks (u, d, s) and from c-quark fragmentation. Properties of the bottom and, to a lesser extent, the charm quark can be used to identify the hadronic jets into which they fragment. These



(a) 2-dimensional JVT likelihood as function of corrJVF and R_{pT}.

(b) Distribution of JVT for pile-up and hard-scatter jets with $20 < p_T < 30\,\text{GeV}$. A value of -0.1 is assigned to jets with no associated tracks.

Figure 4.9.: JVT likelihood as function of the input variables and distribution for pile-up and hard-scatter jets [73].

properties include their hard fragmentation functions and the relatively large mass and long lifetime of the heavy flavor hadrons. b-flavored hadrons have a lifetime of 1.5 ps, which results in a flight path of e.g. 4 mm for a transverse momentum of $p_T = 50\,\text{GeV}$. This property is used either by reconstructing explicitly the displaced secondary vertex, or by measuring the impact parameters of charged particle tracks.

Several algorithms exists to identify jets containing b-hadrons for ATLAS [75, 76] which can be classified in three strategies: impact parameter-based algorithms, an inclusive secondary vertex reconstruction algorithm and a decay chain multi-vertex reconstruction algorithm. The output of these b-tagging algorithms are combined in a multivariate discriminant (MV2) that is the output of the default algorithm, which provides the best separation between the different jet flavors.

The impact parameter based algorithms use several inputs: the transverse impact parameter d_0 defined as the distance of closest approach in the $r - \phi$ plane of the track to the primary vertex, the longitudinal impact parameter $z_0 \sin\theta$, which is the distance of the track to the primary vertex in the longitudinal plane at the point of closest approach in $r - \phi$, as well as the number of hits in the pixel detector. The secondary vertex finding algorithm explicitly reconstructs an inclusive displaced secondary vertex within the jet, testing all track pairs for a two-track

vertex hypothesis. Finally, the chain multi vertex algorithm exploits the topological structure of weak b- and c-hadron decays inside the jet and tries to reconstruct the full b-hadron decay chain using a Kalman filter.

These basic algorithms serve as input variables to the combined multivariate algorithm MV2 using a boosted decision tree (BDT). Three variants can be used that differ by c-jet fraction in the training: MV2c00 without any c-jets contribution, MV2c10 with 7% c-jets in the background and MV2c20 with 15%. For each variant different operating points are defined with different b-jet efficiencies and rejection rates. In Table 4.4 these operating points are listed for MV2c10 that is also used in this analysis.

Table 4.4.: Operating points for the MV2c10 b-tagging algorithm, including benchmark numbers for the efficiency and rejections rates, extracted from $t\bar{t}$ events with jets requirement of $p_T > 20\,\text{GeV}$ [76].

BDT cut value	b-jet efficiency [%]	c-jet rejection	light-jet rejection	τ rejection
0.9349	60	34	1538	184
0.8244	70	12	381	55
0.6459	77	6	134	22
0.1758	85	3.1	33	8.2

4.6 Missing Transverse Energy

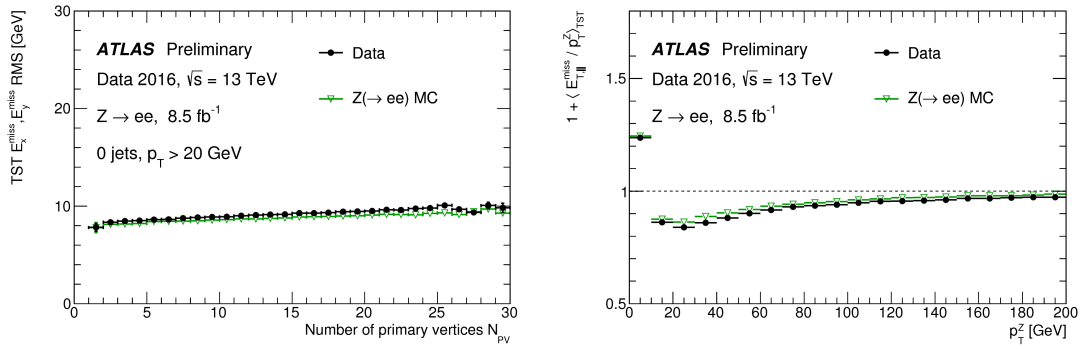
Neutral weakly interacting stable particles, such as neutrinos, escape from typical collider detectors without producing any direct response in the detector elements. The presence of such particles must be inferred from the imbalance of total momentum. The vector momentum imbalance in the plane perpendicular to the beam direction is particularly useful in hadron colliders, and is known as missing transverse momentum, here denoted $\vec{E}_{\mathrm{T}}^{\,\mathrm{miss}} = \left(E_{\mathrm{x}}^{\mathrm{miss}}, E_{\mathrm{y}}^{\mathrm{miss}}\right)$. Its magnitude is called missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{E}_{\mathrm{T}}^{\,\mathrm{miss}}|$. Missing transverse energy is one of the most important observables for discriminating leptonic decays of W-bosons and top quarks from background events which do not contain neutrinos, such as multijet and Drell–Yan events. The reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ [77] is challenging because it involves all detector subsystems and requires the most complete and definite representation of the hard interaction of interest by calorimeter and tracking signals. This representation is obscured by limitations introduced by the detector acceptance and by signals and signal remnants from additional pp interactions from pile-up, or by particle momentum mismeasurements, particle misidentification, detector malfunctions or cosmic-ray particles.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction process uses reconstructed, calibrated objects to estimate the transverse momentum imbalance in an event. The missing transverse energy of an event is calculated as the sum of a number of components:

$$E_{\mathrm{x/y}}^{\mathrm{miss}} = E_{\mathrm{x/y}}^{\mathrm{miss},e} + E_{\mathrm{x/y}}^{\mathrm{miss},\gamma} + E_{\mathrm{x/y}}^{\mathrm{miss},\tau} + E_{\mathrm{x/y}}^{\mathrm{miss,jets}} + E_{\mathrm{x/y}}^{\mathrm{miss},\mu} + E_{\mathrm{x/y}}^{\mathrm{miss,soft}}. \tag{4.6.1}$$

The terms for jets, charged leptons, and photons are the negative sum of the momenta for the respective calibrated objects. To avoid potential double counting of the same signal, when e.g. the same calorimeter signal used to reconstruct an electron is also used to reconstruct a jet, the hard objects contributions are reconstructed in sequence. The order starts with electrons (e), followed by photons (γ), hadronically decaying tau-leptons (τ), jets and finally muons (μ). Lower-priority objects are rejected if they share calorimeter signal with a higher priority object that has already entered the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. The *soft term* is reconstructed from the transverse momentum deposited in the detector but not associated with any reconstructed hard object. It may be reconstructed either by calorimeter-based methods, known as the Calorimeter Soft Term (CST), or track-based methods, resulting in the Track Soft Term (TST). The choice of soft-term algorithm influences the performance of and uncertainties in $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. Pile-up dependence is strongly suppressed by including only fully calibrated objects and using only tracks from the hard-scatter vertex for the soft term.

The resolution and scale of the TST $E_{\mathrm{T}}^{\mathrm{miss}}$ for a selection of $Z \to ee$ events in the 2016 ATLAS dataset compared to a MC simulated sample can be seen in Figure 4.10.



(a) TST $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the number of primary vertices requiring no primary-vertex jet activity.

(b) TST $E_{\mathrm{T}}^{\mathrm{miss}}$ scale, as quantified by the projection of TST $E_{\mathrm{T}}^{\mathrm{miss}}$ along the direction of the Z boson p_T^Z.

Figure 4.10.: TST $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution for a selection of $Z \to ee$ events in the 2016 ATLAS dataset compared to a Powheg+Pythia8 MC simulated sample [78].

5 Signal Modelling

Deviation in the parameters of the underlying theory - be it an effective field theory or otherwise - can have complicated and not trivially predictable consequences on the final predictions arising from that theory. The modelling of these predictions can consume a lot of resources or may be inconceivable in terms of computing time, which applies especially for models with a high dimensional parameter space. Many techniques have been developed to tackle these challenges, but often approximations or restrictions in the phase space that is modelled have to be tolerated. For coupling parameters that are spelled out in the Lagrangian, the technique of the Effective Lagrangian Morphing approaches these problems by modelling the signal processes in an analytical way using the underlying theory.

In this chapter a brief introduction into statistical data analysis is given first in Section 5.1. More insights into the statistics used in high energy physics can be found in [79, 80]. Section 5.2 provides some example of other techniques to model the parameter dependencies of signal processes like the Matrix Element reweighting or different morphing methods. Then, in Section 5.3 the Effective Lagrangian Morphing is introduced, giving some examples and validations in Section 5.4 and optimization procedures in Section 5.5, where methods to improve the utilization of the Effective Lagrangian Morphing are presented.

5.1 Statistical Data Analysis

A statistical analysis quantifies experimental observations from theoretical predictions. Measuring new properties of a particle such as its couplings or the parity for example, typically involves a *hypothesis test* where a null hypothesis is tested against an alternative hypothesis. The background-only or null hypothesis H_0 generally assumed to be true until evidence indicates otherwise and typically incorporates the SM case. Deviations arising from new physics phenomena are modelled in the signal-plus-background or alternate hypothesis H_1, which are e.g. non-zero values for EFT coupling parameters as probed in the analysis outlined in Chapter 6.

Next, a test statistic $T(\mathcal{D}) \rightarrow \mathbb{R}$ has to be selected in such a way as to quantify, within observed data $\mathcal{D}$, behaviors that would distinguish the null from the alternative hypothesis. An *acceptance region* is defined so that if the test statistic is

smaller than a threshold $T(\mathcal{D}) < k_\alpha$ one accepts the null hypothesis. The *size of the test* is given as the probability α that the null hypothesis will be rejected when it is true (Type-I error), i.e. $\alpha = P(T(\mathcal{D}) \geq k_\alpha | H_0)$. In contrast, the probability β to accept the null hypothesis when the alternative is true (Type-II error) is given by $\beta = P(T(\mathcal{D}) < k_\alpha | H_1)$. One calls $1 - \beta$ the *power* of the test.

Using a *probability model* $f(\mathcal{D}|\boldsymbol{\alpha})$ for the observed data given model parameters $\boldsymbol{\alpha}$, one can calculate with the test statistic the p-value with

$$p(\boldsymbol{\alpha}) = P(T \geq T_0 | \boldsymbol{\alpha}). \tag{5.1.1}$$

T_0 is the value of the test statistic based on the observed data. 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 observed. Small p-values correspond to strong evidence. Two scenarios are of particular interest:

- A *discovery* is claimed, if the p-value under the null hypothesis is found to be smaller than a predefined threshold of confidence, and therefore the null hypothesis can be discarded. This corresponds to the size of the test, with the conventional threshold of a 5σ criterion for a Gaussian probability distribution, which translates to $p < 2.87 \cdot 10^{-7}$.
- An *exclusion* of the alternative hypothesis is declared, if the p-value under the alternative hypothesis is found to be smaller than a predefined threshold of confidence. Commonly the threshold is set to be $p < 0.05$.

The parameters of the model $\boldsymbol{\alpha} = \{\boldsymbol{\mu}, \boldsymbol{\theta}\}$ generally represent parameters of a physical theory or an unknown property of the detector's response. There are two types of model parameters:

- The *parameters of interest* $\boldsymbol{\mu}$ typically characterize the alternative hypothesis and are fixed or not present at all in the null hypothesis. Examples are the mass of a hypothetical particle, a signal strength or coupling parameters such as Wilson coefficients.
- The *nuisance parameters* $\boldsymbol{\theta}$ are parameters that are not of immediate interest, but which must be accounted for. They can be unconstrained but often prior knowledge is available to constrain them within certain bounds. These parameters are present in both the alternative and the null hypothesis. Typical examples are parameters of the experimental setup, or the normalization of a background prediction with respect to the observed data set.

Numerically equivalent to the probability model is the *likelihood function*, but is interpreted as a function of the parameters of interest and hence not normed to 1:

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\theta}; \mathcal{D}) = f(\mathcal{D}|\boldsymbol{\mu}, \boldsymbol{\theta}) \tag{5.1.2}$$

The maximum likelihood estimate $\widehat{\boldsymbol{\mu}}$, $\widehat{\boldsymbol{\theta}}$ are the values of the parameters that maximize the likelihood function $\mathcal{L}$ or, equivalently, minimize the $-\log\mathcal{L}$. The conditional maximum likelihood estimate $\widehat{\widehat{\boldsymbol{\theta}}}$ is the value of $\boldsymbol{\theta}$ that maximizes the likelihood function with fixed $\boldsymbol{\mu}$, which is referred to as *profiling*. Using these definition the profile likelihood ratio is defined as

$$\lambda(\boldsymbol{\mu}) = \frac{\mathcal{L}\left(\boldsymbol{\mu}, \widehat{\widehat{\boldsymbol{\theta}}}\right)}{\mathcal{L}\left(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\theta}}\right)}, \tag{5.1.3}$$

which depends explicitly on the parameters of interest $\boldsymbol{\mu}$, implicitly on the data $\mathcal{D}$, and is independent of the nuisance parameters $\boldsymbol{\theta}$, which have been eliminated via profiling. The presence of the nuisance parameters broadens the profile likelihood as a function of $\boldsymbol{\mu}$ relative to what one would have if their values were fixed. This reflects the loss of information about the parameters of interest due to the systematic uncertainties.

Hypothesis testing usually can not be analytically formulated but follows from a chain of MC simulation processes. The negative logarithm of the profile likelihood

$$t(\boldsymbol{\mu}) = -2\ln\lambda(\boldsymbol{\mu}) \tag{5.1.4}$$

is therefore used as a convenient choice for the test statistic to simplify and stabilize numerical calculations, since the magnitude of the likelihoods can be very small and doing a logarithmic transformation converts these small numbers to larger negative values which a finite precision machine can handle better.[1]

5.2 Modelling Techniques

A data analysis in high energy physics usually needs MC simulated events. The production consists of a chain of computational steps. First the hard scattering process is generated. Then the soft physics of the proton interaction is modelled and the full event is simulated inside the ATLAS detector. Together with real recorded physics data, the MC simulated events are reconstructed to finally get processed in an analysis. Finally, a discriminating observable is build to probe the different hypotheses. This procedure is normally very cost intensive in terms of CPU time. A fully simulated event takes up to 15 min per event. Faster approaches with pre-simulated electromagnetic showers and sacrifices of a certain level of accuracy can speed up the event simulation up to a factor of 25 [81]. Most of the available

[1] A factor of 2 is included in the definition to make it similar to the χ^2-expression of the method of least squares, which is discussed in more detail in Section 6.10.2.

CPU time is spent to simulate the large amount of SM processes which are used for Standard Model measurements as well as for BSM searches, where these samples are needed to estimate the background constributions. Signal samples can usually be generated more cost-effectively with individual filters customized to the specific needs of a dedicated analysis. However, this virtue is rendered invalid once the theories subject to testing are highly parametric.

Simulated signal events are only produced for one point $(\boldsymbol{\mu}, \boldsymbol{\theta})$ of the model parameter. The likelihood $\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\theta}; \mathcal{D})$ does therefore not have a continuous description for each value of the unknown model parameters. The simplest solution to approximate the likelihood in a certain parameter space is a *grid scan* of individual parameter points. Unfortunately, a satisfying precision especially in a highly parametric analysis is unachievable due to the high demand of CPU time to generate many MC samples for each parameter grid point. The preferred solution would be a procedure to turn such a grid scan into a continuous description with as little grid points needed and as closely to the analytically correct solution as possible (cf. Figure 5.1). The next sections show first attempts to achieve these targets.

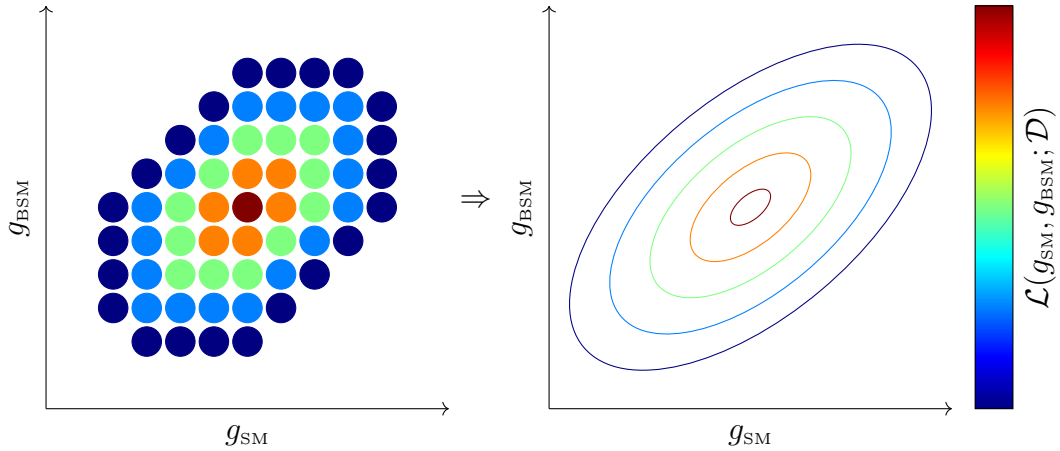


Figure 5.1.: Depiction of a likelihood with two parameters of interest g_{SM} and g_{BSM} with a discrete set of points used in a grid scan and a continuous representation from e.g. an interpolation technique.

5.2.1 Matrix Element Reweighting

A powerful technique to cover a large set of parameter points for the likelihood without investing a lot of computing power once a MC sample is generated is the *Matrix Element reweighting*.

In a MC event generation usually weights are assigned to the events corresponding to the associated probability that the event occurs in the specific phase space given the underlying physics process. To avoid losing time in some latter part of the simulation (e.g. detector simulation) to events that possess only a very small weight, methods exists to equalize the weights of different events uniformly at least in leading order. In NLO simulations, the real and virtual part is integrated separately and the partonic predictions to parton showers include counterterms whose contributions have to add up to zero but are individually negative and positive which yield to varying and even negative event weights [82].

With the Matrix Element reweighting it is possible to assign new event weights for a MC sample that were generated under a certain theoretical hypothesis, i.e. a model and its parameters with given values, associated to a new theoretical hypothesis. Both the original and the additional weights are based solely on matrix element computations. These additional weight can be propagated through or calculated after the simulation chain, and saves one from performing a full simulation on an additional event sample for different model parameters. The method works only if both the original and the new hypothesis give non-negligible contributions to the same parts of the phase space. In LO the new weights are calculated via the matrix elements $\mathcal{M}_{\text{new/orig}}$ for the new and original set of model parameters respectively and the original event weight w_{orig} with

$$w_{\text{new}} = \frac{|\mathcal{M}_{\text{new}}|^2}{|\mathcal{M}_{\text{orig}}|^2} \cdot w_{\text{orig}}. \tag{5.2.1}$$

Also NLO Matrix Element reweighting already exists but is more complicated due to Born, real and virtual contributions to the matrix element [83].

A drawback of the method are potential large event weights. If the probabilities for events to occur in a certain phase space region for the original model parameters are much higher for alternative model parameters then large weights are assigned to these events, accordingly. Hence, a few events that are in the tails of kinematic distributions might dominate and lead to large statistical uncertainties in single bins in distributions. The only solutions would be larger statistics for the MC samples or generating the samples closer to the model parameters of interest, but these strategies would undo the advantages of this method.

A feature of reweighted samples is that they are statistically dependent which leads to more complicated handling of uncertainties on the predictions but may

also be beneficial to have exact degree of agreement e.g. in developing new methods and having closure tests.

Many generators have already implemented the Matrix Element reweighting technique like MADGRAPH5_aMC@NLO [84] since only the matrix element calculator of the generator has to be linked to the original event information. Calculating weights for different configurations after the full simulation chain is more complicated. The computation of the matrix elements needs the original four-momenta of the initial and final state particles of the hard scatter process. The implemented simulation framework complicates this task. After the parton shower a large number of particles are generated from consecutive decays of the final state particles as well as initial and final state radiation and jet hadronization processes which renders the identification of the hard scatter particles more difficult. Also initial and final state radiation shifts the center-of-mass system of parts of the hard scatter process and change the original four-momenta of these particles. Hence, even correctly identified hard scatter particles can lead to different results in the matrix element calculations compared to the results obtained directly from the starting hard scatter event.

These complications could be avoided by either just keeping the original hard scatter events in the event record such that it can be used in the matrix element calculator or use a unique identifier to afterwards match the event with the separately saved original generator record. As samples are often split into smaller parts to parallelize subsequent computations and are not necessarily merged in the same order due to e.g. failed jobs, the identification often proves to be more problematic.

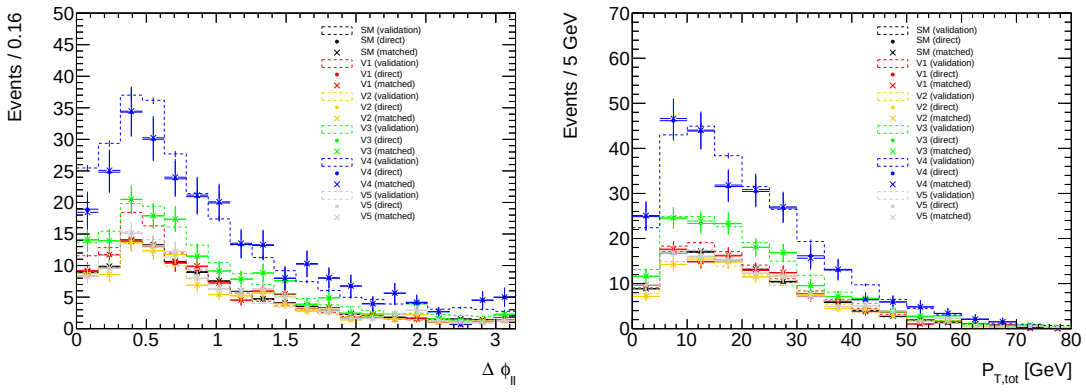
Therefore to circumvent these difficulties two approaches were implemented:

- The *direct reweighting* uses the recorded truth parton information from the showering as input. Here, the four-momenta may slightly vary due to the showering effects mentioned above.
- The *matched reweighting* uses the showered particles to identify the event with the original hard scatter record which have to be saved during generation of the sample. This gives exactly the same results as the matrix element generator would calculate in an a priori Matrix Element reweighting.

Figure 5.2 shows two kinematic distributions of the azimuthal angle between both leptons $\Delta\phi_{ll}$ and the scalar sum of all final state objects $p_{T,\text{tot}}$ for the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ process produced via VBF or VH production modes for different EFT scenarios. The parameters used for the different samples are listed in Table 5.1 utilizing to the Higgs Characterization model introduced in Section 2.3.4. First, samples for all different parameter sets were generated and afterwards the SM sample is reweighted to the parameters of the other samples via direct and matched reweighting. As expected, the uncertainties of the reweighted distributions are

larger than for the directly generated validation samples but still in good agreement. The larger the difference between the EFT parameter set and the SM parameters the larger the statistical uncertainties of the reweighted distributions become.

The impact of both strategies can be seen in Figure 5.3 where the difference between the weights from direct and matched reweighting is shown. The most probable and expected value is zero but also large tails are visible due to boosting and other effects from the showering. But it is however reassuring to see that the tails are smeared out roughly symmetrical which motivates the good agreement seen in the kinematic distributions.



(a) Distribution of the azimuthal angle between the two leptons $\Delta\phi_{ll}$.

(b) Distribution of the total scalar sum of the final state objects $p_{\mathrm{T}}^{\mathrm{tot}}$.

Figure 5.2.: Two kinematic distributions from samples with different EFT parameter settings according to Table 5.1 overlayed with reweighted distributions of the SM sample to the parameters of the validation samples V1 to V5 using both *direct* and *matched reweighting* strategies.

Table 5.1.: Parameter settings used to generate MC samples for the kinematic distributions seen in Figure 5.2. The model at hand is the Higgs Characterization model introduced in Section 2.3.4. For all samples other κ couplings are set to zero and the following parameters were fixed: $\Lambda = 1\,\text{TeV}$ and $\cos\alpha = 1/\sqrt{2}$.

Sample	κ_{SM}	κ_{HWW}
SM	$\sqrt{2}$	0
V1	$\sqrt{2}$	-4.762
V2	$\sqrt{2}$	-2.381
V3	$\sqrt{2}$	2.381
V4	$\sqrt{2}$	4.762
V5	1.447	-2.741

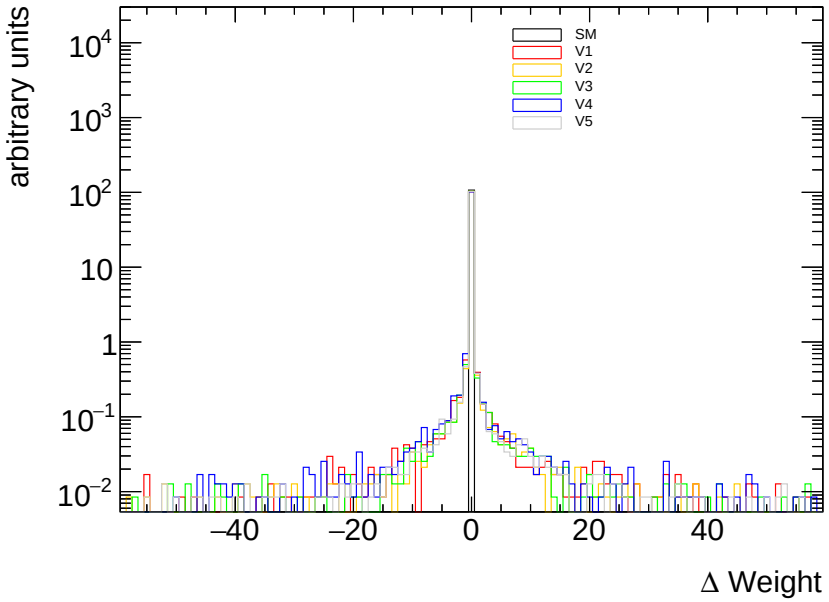


Figure 5.3.: Distribution of the difference between the weights predicted by the two Matrix Element reweighting strategies *direct* and *matched reweighting*. The matched reweighting is subtracted event-by-event from the direct reweighting and displayed in the ΔWeight axis.

5.2.2 Integral Morphing

Many techniques exists to obtain a continuous description of the model parameters for a likelihood from a grid scan. One of the earliest developed methods is the so-called *integral morphing* [85] which is based on the linear interpolation of the inverses of the cumulative distribution functions. Consider a target function $f(x|m)$

depending on a single model parameter m and describing the observable x. The corresponding cumulative distribution function reads

$$F(x|m) = \int_{-\infty}^{x} f(x'|m)\,\mathrm{d}x'. \tag{5.2.2}$$

If two distributions at m_1 and m_2 are known the goal is to obtain a continuous distribution between these parameters with the corresponding cumulative distribution function

$$\bar{F}(x|m) = \int_{-\infty}^{x} \bar{f}(x'|m)\,\mathrm{d}x'. \tag{5.2.3}$$

To do so, the first step is to find x_1 and x_2 such that

$$F(x_1|m_1) = F(x_2|m_2) = y \tag{5.2.4}$$

for a given cumulative probability y. The cumulative probability for the new distribution $\bar{F}$ is set to the same value y at a linearly interpolated position x

$$\bar{F}(x|m) = y \tag{5.2.5}$$

with

$$x = \frac{m_1}{m_1 + m_2} x_1 + \frac{m_2}{m_1 + m_2} x_2. \tag{5.2.6}$$

By inverting the cumulative distributions in Equations 5.2.4 and 5.2.5 and substituting these results into Equation 5.2.6,

$$\bar{F}^{-1}(y|m) = \frac{m_1}{m_1 + m_2} F^{-1}(y|m_1) + \frac{m_2}{m_1 + m_2} F^{-1}(y|m_2), \tag{5.2.7}$$

one can solve this equation with respect to y and retrieve the interpolated or *morphed* distribution

$$\bar{f}(x|m) = \frac{f(x_1|m_1) f(x_2|m_2)}{\frac{m_1}{m_1+m_2} f(x_2|m_2) + \frac{m_2}{m_1+m_2} f(x_1|m_1)}. \tag{5.2.8}$$

Especially for normal distributions the interpolated distribution is also normal with linear interpolated mean and variance. For exponential decays the decay constant is linearly interpolated.

Figure 5.4 shows an example of the integral morphing. Here, two invariant Higgs mass distributions with masses at 50 and 70 GeV in the decay channel $ZH \to e^+e^-H$ simulated for the LEP collider are used as input and morphed to the distribution at a Higgs mass of 60 GeV.

This method is very well suited for models with rapidly shifting means but is computationally intense due to numeric integration and root-finding steps and is restricted only to the description of one-dimensional distributions.

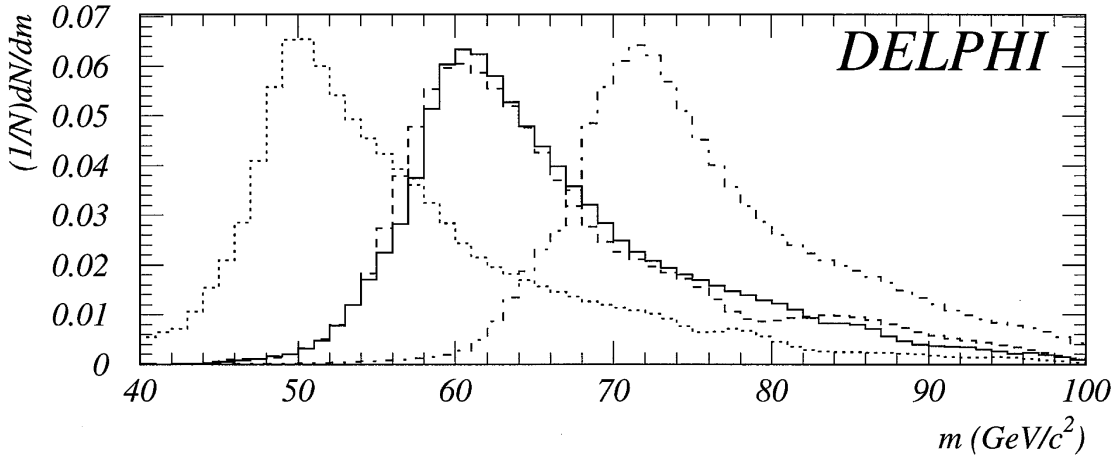


Figure 5.4.: Example of integral morphing. Invariant Higgs mass distributions for the $ZH \to e^+e^-H$ decay channel simulated for the LEP collider at 50 and 70 GeV are used as input and interpolated to the dashed distribution at 60 GeV. The simulated result at 60 GeV is overlayed (dashed line) [85].

5.2.3 Moment Morphing

Another approach is the so-called *moment morphing* [86] capable to interpolate also in a multi-dimensional parameter space. In the following, the procedure is outlined for a single parameter.

The distribution $f(x|m)$ of an observable x and model parameter m is known for n discrete values of m. Given these sampling points the target function can be expanded in a Taylor series up to an order of $n-1$ around a reference value m_0

$$f(x|m_i) \approx \sum_{j=0}^{n-1} (m_i - m_0)^j \frac{1}{j!} \frac{\mathrm{d}^{(j)} f(x|m_0)}{\mathrm{d}m^{(j)}}. \tag{5.2.9}$$

Substituting $M_{ij} = (m_i - m_0)^j$ and $f'_j(x|m_0) = \frac{1}{j!} \frac{\mathrm{d}^{(j)} f(x|m_0)}{\mathrm{d}m^{(j)}}$ the Taylor expansion can be written in matrix form,

$$f(x|m_i) \approx \sum_{j=0}^{n-1} M_{ij} f'_j(x|m_0). \tag{5.2.10}$$

Inverting the $n \times n$ transformation matrix gives the values for the n derivatives

$$f'_j(x|m_0) = \sum_{i=0}^{n-1} (M^{-1})_{ji} f(x|m_i). \tag{5.2.11}$$

Reinserting the derivatives into Equation 5.2.9 can be used to predict the template shape at any new value of the model parameter m

$$\bar{f}(x|m) = \sum_{i=0}^{n-1} c_i(m) f(x|m_i), \tag{5.2.12}$$

with coefficients

$$c_i(m) = \sum_{j=0}^{n-1} (m - m_0)^j (M^{-1})_{ji}. \tag{5.2.13}$$

At the input templates the coefficients yield $c_i(m_j) = \delta_{ij}$ and they sum up $\sum_i c_i(m) = 1$.

The simplest case with $n = 2$ results in the weighted average

$$\bar{f}(x|m) = \frac{m_1 - m}{m_1 - m_0} f(x|m_0) + \frac{m - m_0}{m_1 - m_0} f(x|m_1). \tag{5.2.14}$$

This interpolation is called *vertical morphing*. An example is shown in Figure 5.5a for input templates that have the same mean values and leads to convincing results, in contrast to Figure 5.5b with shifted means and width in x and unsatisfying interpolation results.

To accounts for shifts of the mean and width in x all input distributions have first to be translated to the same mean $\bar{\mu}(m)$ and width $\bar{\sigma}(m)$ such that their location match up and apply vertical morphing afterwards. With the means $\mu(m_i)$ and width $\sigma(m_i)$ of the input templates one derives

$$\begin{aligned} \bar{\mu}(m) &= \sum_{i=0}^{n-1} c(m) \mu(m_i) \\ \bar{\sigma}(m) &= \sum_{i=0}^{n-1} c(m) \sigma(m_i). \end{aligned} \tag{5.2.15}$$

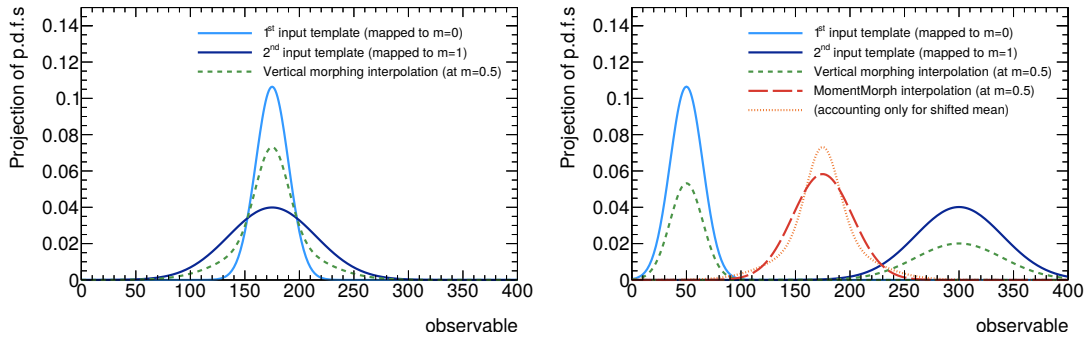
The linear transformation of the input observable is performed with

$$\widetilde{x}_i(x|m) = \frac{\sigma(m_i)}{\bar{\sigma}(m)} \left(x - \bar{\mu}(m)\right) + \mu(m_i) \tag{5.2.16}$$

and can be substituted into Equation 5.2.12

$$\widetilde{f}(x|m) = \sum_{i=0}^{n-1} c_i(m) f(\widetilde{x}_i|m_i). \tag{5.2.17}$$

This interpolation technique is referred to as *moment morphing*. Figure 5.5b shows the interpolated distribution using moment morphing with suitable outcome compared to vertical morphing. For a Gaussian probability model with linearly changing mean and width, moment morphing of two normal templates is the exact solution.



(a) Vertical morphing interpolation demonstrated using two normal input templates with same mean value.

(b) Vertical and moment morphing interpolation demonstrated using two normal input templates with different mean values.

Figure 5.5.: Example of vertical and moment morphing using two normal distributions with same and shifted mean values [86].

Still, moment morphing is an interpolation technique derived from empirical properties of the input distributions. A better and preferred way would include a physics inspired morphing technique at best with analytical exact solutions taking quantum mechanical effects into account, e.g. interferences of different contributions, like a matrix element based event generator. The newly developed method of *Effective Lagrangian Morphing* is constructed exactly to include these features for the case of modelling Lagrangians depending on coupling parameters as for EFT scenarios with Wilson coefficients.

5.3 Effective Lagrangian Morphing

For coupling parameters that are spelled out in the Lagrangian of a physics model, one can construct an interpolation procedure that is based on the underlying physics. The novel technique is called *Effective Lagrangian Morphing* [7, 87]. It is not only a simple interpolation method, but gives in principle analytically exact solutions without any additional approximations for the entire coupling phase space only with just a few pre-computed input samples required. Not only effective couplings derived from e.g. an EFT approach can be used but also non-effective operators. The interpolation is applicable to total and differential cross sections, branching fractions or even individual reweighted events. Since the detector simulation and reconstruction does not influence the structure of the underlying physics model, this morphing can be used on parton level as well as on samples with full reconstructed information.

However, it cannot replace other interpolation techniques. Non-coupling-like parameters can not be interpolated, which are for example particle masses, widths or other model parameter like renormalization and factorization scales or PDF variations.

Since Effective Lagrangian Morphing can be applied on distribution level, it is computational very fast and can perform directly a likelihood fit to physics data, which would be unimaginably using e.g. just Matrix Element reweighting, since this works on an event-by-event basis.

In the following section the derivation of the method is presented followed by two exemplary morphing scenarios to get a better understanding.

5.3.1 Derivations

Consider a Lagrangian that is structured as a linear combination of coupling parameters of interest $g_i \in \boldsymbol{g}$ and associated operators $\mathcal{O}_i$,

$$\mathcal{L} = \sum_i g_i \mathcal{O}_i. \tag{5.3.1}$$

For an arbitrary physics process with vertices $\mathcal{V}$ at which these operators act, the amplitude can be written in a general form with

$$\mathcal{M}(\boldsymbol{g}) = \prod_{\mathcal{V}} \sum_{j \in \mathcal{V}} g_j \mathcal{M}_j, \tag{5.3.2}$$

where $\mathcal{M}_j \in \boldsymbol{\mathcal{M}}$ are the partial matrix elements each corresponding to one of the operators $\mathcal{O}_i$ in the Lagrangian. These partial matrix elements include not only the

vertex term that is related to the coupling g_j, but also all propagators, vertices or other terms of the physics process, which do not depend on any coupling parameters.

In the squared amplitude the product of sums can be multiplied out to a sum of products,

$$\begin{aligned} |\mathcal{M}(\boldsymbol{g})|^2 &= \prod_{\mathcal{V}} \left| \sum_{j \in \mathcal{V}} g_j \mathcal{M}_j \right|^2 \\ &= \sum_i c_i \, \mathrm{Re} \left(\prod_{k(i)} g_k \mathcal{M}_k \right) \\ &= \sum_i c_i \underbrace{\prod_{k(i)} g_k}_{P_i(\boldsymbol{g})} \underbrace{\mathrm{Re} \left(\prod_{k(i)} \mathcal{M}_k \right)}_{P_i(\boldsymbol{\mathcal{M}})} . \end{aligned} \tag{5.3.3}$$

Here, $k(i)$ denotes the combinatorics that results from the expansion of the product over the vertices and c_i represents the number of identical terms. The polynomials $P_i(\boldsymbol{g})$ of the coupling constant are of the order $2v$, if v vertices are affected by the coupling constants of interest. $P_i(\boldsymbol{\mathcal{M}})$ are the corresponding polynomials of the partial matrix elements, which are multiplied in a complex conjugated way to become a real quantity, e.g. $\mathrm{Re}(\mathcal{M}_1^* \mathcal{M}_2)$.

Using Equation 2.4.5 for the cross section calculation for a $2 \to n$ process in a compact form, where the distributions are included in the phase space integral $\mathrm{d}\Omega$ over the final state particle momenta and the constant C is absorbed by the partial matrix elements, the amplitude squared can be included into the following form

$$\begin{aligned} \sigma(\boldsymbol{g}) = &\int \mathrm{d}\Omega \; |\mathcal{M}(\boldsymbol{g})|^2 \\ &\sum_i c_i P_i(\boldsymbol{g}) \int \mathrm{d}\Omega \, P_i(\boldsymbol{\mathcal{M}}). \end{aligned} \tag{5.3.4}$$

The integrals containing the cross section

$$\chi_i = \int \mathrm{d}\Omega \, P_i(\boldsymbol{\mathcal{M}}) \tag{5.3.5}$$

span a basis for a vector space, where the coupling polynomials represent the coordinates. This basis can be mapped to another basis ξ_i with

$$\chi_i = \sum_j A'_{ij}\xi_j \tag{5.3.6}$$

for an appropriate chosen set of coefficients A'_{ij}. The cross section with the couplings $\boldsymbol{g}$ in the other basis is expressed as

$$\sigma(\boldsymbol{g}) = \sum_i c_i P_i(\boldsymbol{g})\chi_i \tag{5.3.7}$$

$$\sum_i c_i P_i(\boldsymbol{g}) \sum_j A'_{ij}\xi_j \tag{5.3.8}$$

$$\sum_{i,j} P_i(\boldsymbol{g}) A_{ij}\xi_j, \tag{5.3.9}$$

where the numbers c_i are absorbed in the coefficients A'_{ij}.

Let now choose sets of parameter points $\boldsymbol{g}_k$ such that the new basis ξ_k is expressed as physical cross sections at these parameters. With Equation 5.3.9 the closure condition

$$\xi_k = \sigma(\boldsymbol{g}_k) = \sum_{i,j} P_i(\boldsymbol{g_k}) A_{ij}\xi_j \tag{5.3.10}$$

gives linear equations to compute the unknown coefficients A_{ij}.

To satisfy $\xi_k = \sum_j \delta_{jk}\xi_j$ one can write

$$\delta_{jk} = \sum_{i,j} P_i(\boldsymbol{g_k}) A_{ij}, \tag{5.3.11}$$

equivalent in matrix form

$$A \cdot M = \mathbb{I} \qquad \text{with} \qquad M_{ik} = P_i(\boldsymbol{g_k}). \tag{5.3.12}$$

Thus, by inverting the matrix M Equation 5.3.10 has as a definite solution the *morphing matrix* A. A set of coupling points $\{\boldsymbol{g_k}\}$ that yield to an invertible matrix M is called *morphing basis*.

Having the cross sections $\sigma(\boldsymbol{g}_i)$ and the solution for the morphing matrix A a cross section at an arbitrary parameter point $\sigma(\boldsymbol{g})$ can be calculated with Equation 5.3.9,

$$\sigma(\boldsymbol{g}) = \sum_i \underbrace{\sum_j P_j(\boldsymbol{g}) A_{ji}}_{w_i(\boldsymbol{g}, \{\boldsymbol{g}_k\})} \sigma(\boldsymbol{g}_i), \tag{5.3.13}$$

and comprise a linear combination of the weights w_i with the cross sections $\sigma(\boldsymbol{g}_i)$.

5.3.2 Simple Examples

In the following, two simple examples will illustrate the calculations and help to get a better understanding.

One Vertex with two Couplings

Let's consider a process where two different couplings $\boldsymbol{g} = \{g_{\text{SM}}, g_{\text{BSM}}\}$ act on just one vertex. One exemplary process is a Higgs boson produced via gluon fusion decaying into $H \to W^{\pm} W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ illustrated in Figure 5.6, where in the decay vertex the SM and a BSM coupling from e.g. the Higgs Characterization model (cf. Section 2.3.4) contribute.

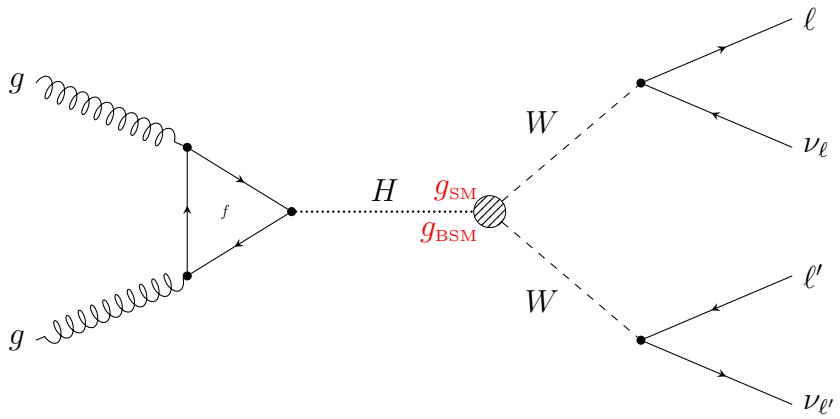


Figure 5.6.: Feynman diagram of an exemplary process where two coupling parameters appear in one vertex. The process depicted here is the production of a Higgs boson via gluon fusion decaying into two leptonically decaying W-bosons. The two couplings acting on the Higgs decay vertex are g_{SM} and g_{BSM}.

The matrix element in this case can be written as

$$\mathcal{M}(\boldsymbol{g}) = g_{\text{SM}} \mathcal{M}_{\text{SM}} + g_{\text{BSM}} \mathcal{M}_{\text{BSM}} \tag{5.3.14}$$

$$|\mathcal{M}(\boldsymbol{g})|^2 = g_{\text{SM}}^2 \left|\mathcal{M}_{\text{SM}}\right|^2 + 2 g_{\text{SM}} g_{\text{BSM}} \,\text{Re}(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{BSM}}) + g_{\text{BSM}}^2 \left|\mathcal{M}_{\text{BSM}}\right|^2 . \tag{5.3.15}$$

Three different coupling polynomials of order 2 are identified:

$$P(\boldsymbol{g}) = \{g_{\mathrm{SM}}^2, g_{\mathrm{SM}} g_{\mathrm{BSM}}, g_{\mathrm{BSM}}^2\} \tag{5.3.16}$$

In principle three arbitrary coupling parameter points could be chosen as morphing basis as long as the morphing matrix

$$M = \begin{pmatrix} g_{\mathrm{SM},1}^2 & g_{\mathrm{SM},1} \cdot g_{\mathrm{BSM},1} & g_{\mathrm{BSM},1}^2 \\ g_{\mathrm{SM},2}^2 & g_{\mathrm{SM},2} \cdot g_{\mathrm{BSM},2} & g_{\mathrm{BSM},2}^2 \\ g_{\mathrm{SM},3}^2 & g_{\mathrm{SM},3} \cdot g_{\mathrm{BSM},3} & g_{\mathrm{BSM},3}^2 \end{pmatrix} \tag{5.3.17}$$

is still invertible. For simplicity, let's choose the parameters listed in Table 5.2 corresponding to a pure SM, a pure BSM and a mixed sample.

Table 5.2.: Set of coupling parameter points for a morphing example with two couplings of interest contribute to one vertex.

Sample	g_{SM}	g_{BSM}
SM	1	0
Mix	1	1
BSM	0	1

With these parameters the resulting morphing matrix and its inversion is given by

$$M = \begin{matrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \begin{matrix} g_{\mathrm{SM}}^2 & g_{\mathrm{SM}} g_{\mathrm{BSM}} & g_{\mathrm{BSM}}^2 \\ \end{matrix} \begin{matrix} \mathrm{SM} \\ \mathrm{Mix} \\ \mathrm{BSM} \end{matrix} \tag{5.3.18}$$

$$M^{-1} = \begin{matrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \begin{matrix} \mathrm{SM} & \mathrm{Mix} & \mathrm{BSM} \\ \end{matrix} \begin{matrix} g_{\mathrm{SM}}^2 \\ g_{\mathrm{SM}} g_{\mathrm{BSM}} \\ g_{\mathrm{BSM}}^2. \end{matrix} \tag{5.3.19}$$

Multiplied with the coupling polynoms in Equation 5.3.13 one obtains

$$\begin{aligned}\sigma(g_{\mathrm{SM}}, g_{\mathrm{BSM}}) = \quad & \underbrace{(g_{\mathrm{SM}}^2 - g_{\mathrm{SM}} \cdot g_{\mathrm{BSM}})}_{w_{\mathrm{SM}}} && \cdot\sigma(1,0) \\ & + \underbrace{g_{\mathrm{SM}} \cdot g_{\mathrm{BSM}}}_{w_{\mathrm{Mix}}} && \cdot\sigma(1,1) \\ & + \underbrace{(g_{\mathrm{BSM}}^2 - g_{\mathrm{SM}} \cdot g_{\mathrm{BSM}})}_{w_{\mathrm{BSM}}} && \cdot\sigma(0,1).\end{aligned} \tag{5.3.20}$$

With this formula the cross section of an arbitrary parameter point can be calculated just using three pre-computed cross sections $\sigma_{\mathrm{SM}} = \sigma(1,0)$, $\sigma_{\mathrm{Mix}} = \sigma(1,1)$ and $\sigma_{\mathrm{BSM}} = \sigma(0,1)$ and the corresponding weights.

This simple morphing case can be understood very intuitively by looking at the matrix element squared of Equation 5.3.15, which is proportional to the cross section and can directly be compared to the morphing formula 5.3.20. Three terms must be summed up: the SM, the BSM and an interference term. The interference term is calculated by subtracting the pure SM and BSM cross sections from the cross section of the mixed sample. Replacing the partial matrix elements by the cross sections

$$\begin{aligned} |\mathcal{M}_{\mathrm{SM}}|^2 &\rightarrow \sigma_{\mathrm{SM}} \\ 2\,\mathrm{Re}(\mathcal{M}_{\mathrm{SM}}^* \mathcal{M}_{\mathrm{BSM}}) &\rightarrow \sigma_{\mathrm{Mix}} - \sigma_{\mathrm{SM}} - \sigma_{\mathrm{BSM}} \\ |\mathcal{M}_{\mathrm{BSM}}|^2 &\rightarrow \sigma_{\mathrm{BSM}} \end{aligned}$$

returns the morphing formula 5.3.20. This calculation is illustrated in Figure 5.7.

The morphing formula is generalized for an arbitrary morphing basis by simply choosing variable coupling parameters $\boldsymbol{g}_1$, $\boldsymbol{g}_2$ and $\boldsymbol{g}_3$ for the three input samples.

Two Vertices with shared Couplings

The case, where two vertices are affected by the couplings g_{SM} and g_{BSM}, is a bit more complicated, however the derivation of the morphing function is analogous. An exemplary physics process is shown in Figure 5.8, where the Higgs is produced via vector boson fusion and decays again in $H \rightarrow W^{\pm} W^{\mp *} \rightarrow \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$. Since the production and decay vertex both connect the Higgs and two W-bosons, the couplings act on both vertices.

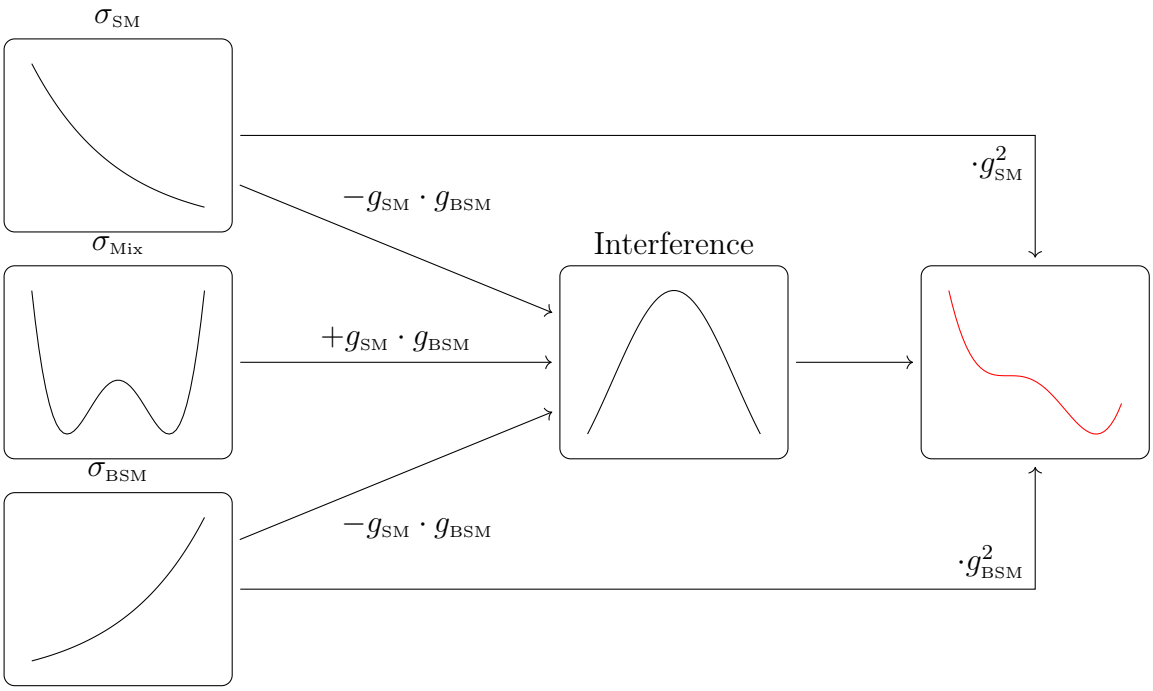


Figure 5.7.: Schematic depiction of a simple Effective Lagrangian Morphing with two couplings of interest contribution to one vertex with fixed morphing basis listed in Table 5.2.

The matrix element as function of the couplings is written as

$$
\begin{aligned}
\mathcal{M}(\boldsymbol{g}) = {} & \underbrace{(g_{\mathrm{SM}}\mathcal{M}_{\mathrm{SM},p} + g_{\mathrm{BSM}}\mathcal{M}_{\mathrm{BSM},p})}_{\text{production}} \cdot \underbrace{(g_{\mathrm{SM}}\mathcal{M}_{\mathrm{SM},d} + g_{\mathrm{BSM}}\mathcal{M}_{\mathrm{BSM},d})}_{\text{decay}} \\
|\mathcal{M}(\boldsymbol{g})|^2 = {} & g_{\mathrm{SM}}^4 \cdot |\mathcal{M}_{\mathrm{SM},p}|^2 |\mathcal{M}_{\mathrm{SM},d}|^2 \\
& + g_{\mathrm{SM}}^3 g_{\mathrm{BSM}} \cdot 2\big[|\mathcal{M}_{\mathrm{SM},p}|^2 \operatorname{Re}(\mathcal{M}_{\mathrm{SM},d}^* \mathcal{M}_{\mathrm{BSM},d}) \\
& \qquad + \operatorname{Re}(\mathcal{M}_{\mathrm{SM},p}^* \mathcal{M}_{\mathrm{BSM},p}) |\mathcal{M}_{\mathrm{SM},d}|^2\big] \\
& + g_{\mathrm{SM}}^2 g_{\mathrm{BSM}}^2 \cdot \big[|\mathcal{M}_{\mathrm{SM},p}|^2 |\mathcal{M}_{\mathrm{BSM},d}|^2 + |\mathcal{M}_{\mathrm{BSM},p}|^2 |\mathcal{M}_{\mathrm{SM},d}|^2 \\
& \qquad + 4 \operatorname{Re}(\mathcal{M}_{\mathrm{SM},p}^* \mathcal{M}_{\mathrm{BSM},p}) \operatorname{Re}(\mathcal{M}_{\mathrm{BSM},p}^* \mathcal{M}_{\mathrm{SM},p})\big] \\
& + g_{\mathrm{SM}} g_{\mathrm{BSM}}^3 \cdot 2\big[|\mathcal{M}_{\mathrm{BSM},p}|^2 \operatorname{Re}(\mathcal{M}_{\mathrm{SM},d}^* \mathcal{M}_{\mathrm{BSM},d}) \\
& \qquad + \operatorname{Re}(\mathcal{M}_{\mathrm{SM},p}^* \mathcal{M}_{\mathrm{BSM},p}) |\mathcal{M}_{\mathrm{BSM},d}|^2\big] \\
& + g_{\mathrm{BSM}}^4 \cdot |\mathcal{M}_{\mathrm{BSM},p}|^2 |\mathcal{M}_{\mathrm{BSM},d}|^2 .
\end{aligned}
\tag{5.3.21}
$$

Since in this case two vertices are modified by the parameters of interest, the coupling polynoms are of order four. In total five polynoms are identified in the matrix element squared and therefore the same number of input samples are

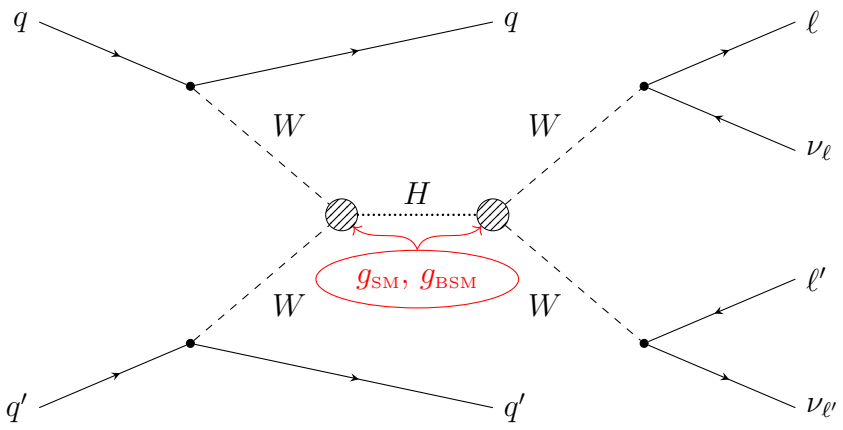


Figure 5.8.: Feynman diagram of an exemplary process where two coupling parameters appear in two vertices. The process depicted here is the production of a Higgs boson via vector boson fusion decaying into two leptonically decaying W-bosons. The two couplings acting on the $H \to WW$ vertex are g_{SM} and g_{BSM}.

needed to exert the morphing. Now, the same procedure as shown in the previous example has to be applied. A morphing basis is defined by choosing five coupling parameter points. Then the morphing matrix is formulated, inverted and finally the morphing formula can be derived.

5.3.3 Dimensionality

As is clearly apparent from the examples in the previous section, the number of input samples that have to be generated for the Effective Lagrangian Morphing is always exactly equal to the number of different coupling polynoms that appear in the matrix element squared. It has to be noted that this number is fixed by the number of parameters and in which vertices they appear. It is not possible to use fewer samples or to include additional generated samples. In principle there is no limitation on how many parameters can be modelled simultaneously.

For a generic $2 \to 2$ s-channel process at tree level only one or two vertices are modified by coupling parameters of interest. The coupling polynoms are in this case of the order four and the formula for the number of required samples

is calculated with

$$\begin{aligned} N = & \frac{n_p\,(n_p+1)}{2} \cdot \frac{n_d\,(n_d+1)}{2} + \binom{4+n_s-1}{4} \\ & + \left(n_d \cdot n_s + \frac{n_s\,(n_s+1)}{2} \right) \cdot \frac{n_p\,(n_p+1)}{2} \\ & + \left(n_p \cdot n_s + \frac{n_s\,(n_s+1)}{2} \right) \cdot \frac{n_d\,(n_d+1)}{2} \\ & + \frac{n_s\,(n_s+1)}{2} \cdot n_p \cdot n_d + (n_p + n_d) \binom{3+n_s-1}{3}, \end{aligned} \tag{5.3.22}$$

where n_p and n_d are the number of parameters that contribute exclusively in the production or decay vertex respectively, and n_s is the number of shared couplings appearing in both production and decay vertex. Each line in the equation counts the number of terms for different categories of polynoms. The first line sums terms pure in production and decay, or containing only shared couplings. Production coupling polynoms of second order are counted in the second line and respectively decay coupling polynoms of second order in the third line. The last line counts the terms which are only of first order in either production or decay or both.

For cases where only one vertex is modified, like in the example shown in Figure 5.6 for the gluon fusion process, the expression can be simplified to

$$N_{\text{ggF}} = \frac{n(n+1)}{2}. \tag{5.3.23}$$

Similar, for studies with processes having the same production and decay vertex and looking only at shared parameters like the VBF process shown in Figure 5.8 the formula reduces to

$$N_{\text{VBF}} = \binom{4+n-1}{4}. \tag{5.3.24}$$

As apparent from these equations the number of required samples grow rapidly as function of the number of samples especially when parameters are shared between vertices. Compared to a minimalistic grid scan where each dimension is segmented into three samples, the morphing technique nevertheless requires far fewer samples, which is visualized in Figure 5.9.

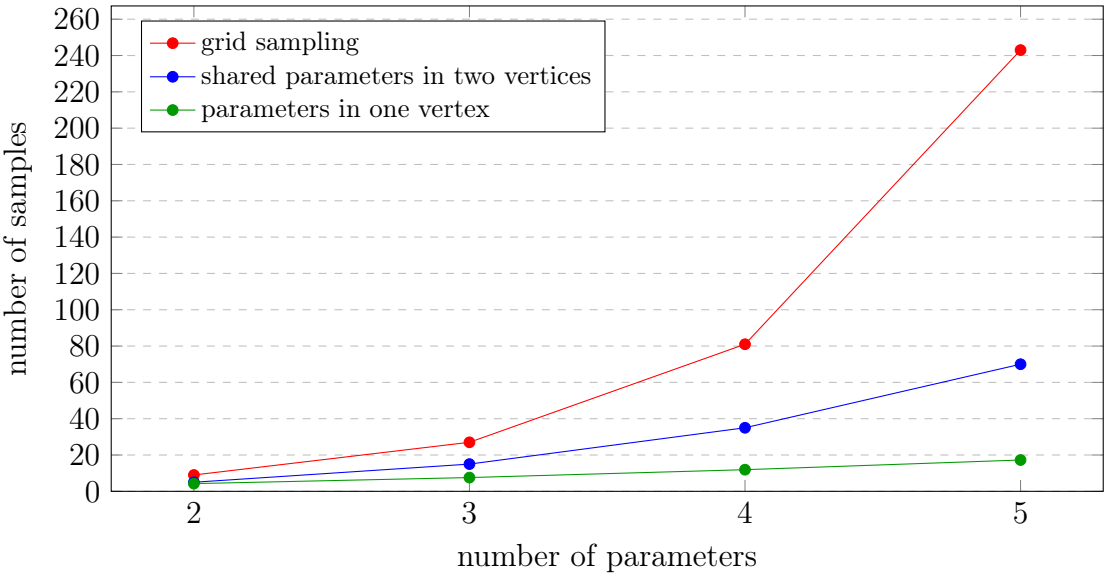


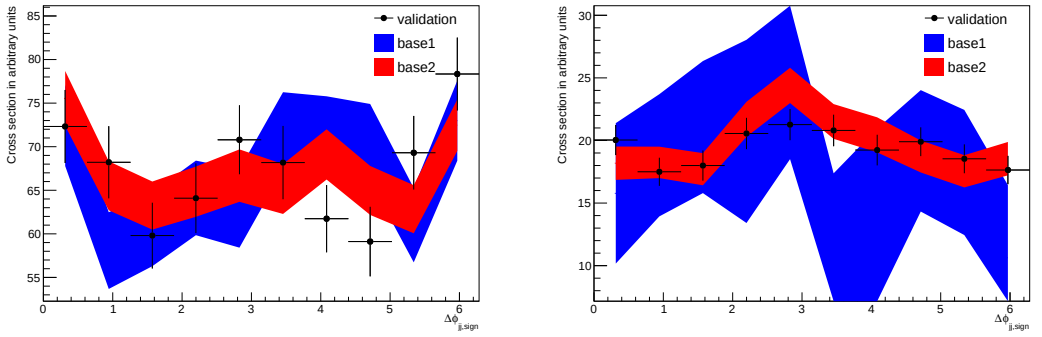
Figure 5.9.: Number of input samples required for Effective Lagrangian Morphing as function of the number of free coupling parameters contributing all to one or two vertices and compared to a grid scan of the parameter space with three samples per dimension.

5.3.4 Uncertainty Propagation

For a specific model with coupling parameters of interest and a given physics process, the Effectie Lagrangian Morphing is completely deterministic. The only freedom remains in choosing the morphing basis, meaning the parameter points at which the input samples are generated. The choice of these input parameters can have a huge impact on the uncertainty from e.g. cross section calculations of the input samples propagated through the morphing function, increasingly for larger number of free coupling parameters.

Figure 5.10 shows two examples for different processes with different sets of coupling parameters using the Higgs Characterization model. In each examples two morphing bases are used and the inputs are morphed to the SM parameters. The resulting distributions for the observable $\Delta\phi_{jj,\text{sign}}$ (cf. Equation 6.1.1) are superimposed with an independently generated distribution at the SM target parameters. In the first example seen in Figure 5.10a the process at hand is ggF with $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay and the free coupling parameters are κ_{SM} and κ_{AWW} which apply both in the decay. The production coupling κ_{Hgg} is set constant to the SM coupling for all inputs. The parameters used for both morphing bases are listed in Table 5.3a. It's clearly visible that the blue distribution has larger uncertainties than the red one despite that the only difference are larger κ_{AWW} values

for the inputs of the blue distribution. Nevertheless, the effect is not substantial. Looking at the second example, the effect of the morphing basis choice on the uncertainty becomes more apparent. Here, the events are produced in the process VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ with free couplings κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$[2]. Again, two different morphing bases are used whose parameter settings are listed in Table 5.3b. Despite that only one input sample differs for both bases the uncertainty of the blue distribution is significant larger than the red one. It can be stated that the higher the dimensionality is - that is the more free couplings are analyzed - and therefore the more inputs are necessary, the larger is the impact of the morphing basis on the uncertainties, as can be nicely seen in these two examples.



(a) The inputs are generated with the process ggF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ using κ_{SM} and κ_{AWW} parameters.

(b) The inputs are generated with the process VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ using κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$.

Figure 5.10.: Impact on the uncertainty for different morphing bases. The corresponding input coupling parameters are listed in Table 5.3. In each example two different bases are used and the corresponding distributions for the $\Delta\phi_{jj,sign}$ observable are morphed to the SM parameters. An independently generated SM distribution is overlayed as validation.

Since the Effective Lagrangian Morphing is based on the matrix elements of the underlying physics, no model uncertainties are introduced in the interpolation. Only two possible sources of uncertainty remain:

1. Due to finite precision in computational calculations, numerical inaccuracies are introduced especially in the inversion of the morphing matrix which is often bad conditioned.

[2]The coupling κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ refer to identically set coupling to WW and ZZ bosons, such that the parameters κ_{HWW} and κ_{HZZ} are set equal as well as $\kappa_{\mathrm{H}\partial\mathrm{WR}}$, $\kappa_{\mathrm{H}\partial\mathrm{WI}}$ and $\kappa_{\mathrm{H}\partial\mathrm{Z}}$. This has only an effect on the production vertex since here the Higgs can couple to both W- and Z-boson (cf. Equation 6.1.2).

Table 5.3.: Input parameters used to generate the distribution of Figure 5.10. Constant parameters are $\Lambda = 1000$ and $\cos\alpha = 1/\sqrt{2}$.

(a) Input parameter corresponding to Figure 5.10a. The production parameter κ_{ggH} is set constant to $\sqrt{2}$.

Sample	κ_{SM}	κ_{AWW}
base1	0	2.402
	$\sqrt{2}$	24.023
	$\sqrt{2}$	-24.023
base2	0	2.402
	$\sqrt{2}$	2.402
	$\sqrt{2}$	-2.402
validation	$\sqrt{2}$	0

(b) Input parameter corresponding to Figure 5.10b. Only one sample is varied between both bases.

Sample	κ_{SM}	κ_{HVV}	$\kappa_{\mathrm{H}\partial\mathrm{V}}$
base1 and base2	0	2.28	0
	0	0	0.26
	$\sqrt{2}$	22.8	0
	$\sqrt{2}$	-22.8	0
	$\sqrt{2}$	0	2.58
	$\sqrt{2}$	0	-2.58
	$\sqrt{2}$	22.8	2.58
	$\sqrt{2}$	-22.8	2.58
	$\sqrt{2}$	11.4	0
	$\sqrt{2}$	0	1.29
	$\sqrt{2}$	11.4	-2.58
	$\sqrt{2}$	22.8	3.87
	$\sqrt{2}$	-11.4	-2.58
	$\sqrt{2}$	-22.8	-3.87
base1	$\sqrt{2}$	0	0.26
base2	$\sqrt{2}$	2.28	0
validation	$\sqrt{2}$	0	0

2. Only a limited number of events per samples are generated which results in an uncertainty on the cross section calculation proportional to the square root of the sum of weights squared, $\sqrt{\sum w_i^{\mathrm{event}}}$, where w_i^{event} is the weight of the i-th generated event of the MC sample.

To avoid finite precision problems in the matrix inversion the `C++` implementation of the method uses `uBLAS` and `multiprecision` from the `boost` libraries [88] to arbitrarily increase the number of significant digits and therefore achieve a negligible effect for the numerical accuracy. Hence, the uncertainty estimate is based purely on propagation of the uncertainties of the input values between which the interpolation is performed.

In case the input samples are totally uncorrelated, i.e. independently generated at coupling parameters $\{\boldsymbol{g}_i\}_{i=1}^{N}$, the uncertainty $\Delta\sigma_i$ on the N computed cross

sections $\sigma_i = \sigma(\boldsymbol{g}_i)$ are propagated in the morphing function with

$$\Delta\sigma(\boldsymbol{g})^2 = \sum_{i=1}^{N} \left(w_i(\boldsymbol{g})\,\Delta\sigma_i)\right)^2. \tag{5.3.25}$$

The weights w_i are functions not only of the target parameters $\boldsymbol{g}$ but also on the morphing basis $\{\boldsymbol{g}_i\}_{i=1}^N$. The uncertainty propagation is here shown for cross sections but is equivalent for other quantities like branching ratios or differential cross sections. The relative uncertainty is calculated with

$$\Delta_{\text{rel}}\sigma(\boldsymbol{g}) = \frac{\sqrt{\sum_{i=1}^{N} \left(w_i(\boldsymbol{g})\,\Delta\sigma_i\right)^2}}{\sum_{i=1}^{N} w_i(\boldsymbol{g})\,\sigma_i}. \tag{5.3.26}$$

Assuming the same relative uncertainty $\delta = \Delta\sigma_i/\sigma_i$ on the cross sections this equation simplifies to

$$\Delta_{\text{rel}}\sigma(\boldsymbol{g}) = \delta \cdot \frac{\sqrt{\sum_{i=1}^{N} \left(w_i(\boldsymbol{g})\,\sigma_i\right)^2}}{\sum_{i=1}^{N} w_i(\boldsymbol{g})\,\sigma_i}. \tag{5.3.27}$$

This may be true in some approximation for total cross sections but can change significantly in specific signal regions with different kinematic acceptances for individual samples especially in differential distributions with low statistics. Still, it helps to understand how to choose the morphing basis to have as small uncertainties as possible in a parameter space region of interest.

The other interesting case consists of totally correlated samples, where e.g. one MC sample is reweighted to other coupling parameters via Matrix Element reweighting. Here, the propagated morphing uncertainty is calculated with

$$\Delta_{\text{corr}}\sigma(\boldsymbol{g})^2 = \left(\sum_{i=1}^{N} w_i(\boldsymbol{g})\,\Delta\sigma(\boldsymbol{g}_i)\right)^2, \tag{5.3.28}$$

the relative uncertainty is derived to

$$\Delta_{\text{corr,rel}}\sigma(\boldsymbol{g}) = \frac{\sum_{i=1}^{N} w_i(\boldsymbol{g})\,\Delta\sigma_i}{\sum_{i=1}^{N} w_i(\boldsymbol{g})\,\sigma_i} \tag{5.3.29}$$

and simplifies assuming again equal relative uncertainties of all input samples to

$$\Delta_{\text{corr,rel}}\sigma(\boldsymbol{g}) = \delta. \tag{5.3.30}$$

Unfortunately, the assumption of equal uncertainties in the correlated case is rarely true, since reweighted samples have often huge variations in the uncertainties when the parameters of the source sample differ a lot from the target parameters. Then few events in the kinematic tails dominate with large weights obtaining thus large uncertainties.

In most analyses constraints already exists on the target parameters. Usually one expects only small deviations from the SM case and can limit the target parameters around the SM coupling point. In this region the propagated uncertainty in the morphing function can be minimized by choosing a proper morphing basis. This challenge is tackled in Section 5.5.4 where an iterative method is presented to obtain an optimal morphing basis for the coupling parameter space of interest.

5.4 Validations and Applications

Many validations of the method have been performed to prove the concept and show its capabilities. The usual strategy consists of several steps. First, a physics process and coupling parameters of interest are chosen, which fixes the number of input samples. A suitable morphing basis is selected and corresponding samples are generated as well as validation samples with different parameters. The Effective Lagrangian Morphing is used with the input samples to obtain the predictions at the validation parameters which are finally compared to the distributions of the independently generated validation samples. In addition, fits can be performed to the validation distributions to check if the original coupling parameters can be extracted.

Here, different applications are presented for the ggF and VBF production mode and via $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ and $H\to ZZ^{*}\to \ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ decay with different parameters settings for truth and reconstructed events. Three different studies are presented:

- truth level ggF $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ events at $\sqrt{s} = 13\,\text{TeV}$ with κ_{SM} and κ_{AWW} couplings,
- truth level VBF $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ events at $\sqrt{s} = 13\,\text{TeV}$ with κ_{SM}, κ_{HVV} and $\kappa_{\text{H}\partial\text{V}}$ couplings,
- reconstructed ggF $H\to ZZ^{*}\to \ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ events at $\sqrt{s} = 7\,\text{TeV}$ with κ_{SM}, κ_{HZZ} and κ_{AZZ} couplings.

For the first two validations dedicated samples are produced using the MADGRAPH5_aMC@NLO 2.3.3.p1 generator with the Higgs Characterization model introduced in Section 2.3.4. Each sample consists of 100 000 events with varying parameter settings. The events are generated in NLO precision in QCD and LO in electroweak couplings. In the ggF case the vertex between the gluons and the Higgs boson is modelled as an effective coupling where as approximation the mass of the heavy quark in the loop are taken to infinity. The NLO PDF NNPDF3.0 [89] with strong coupling $\alpha_s = 0.118$ is used for all samples. The parton shower and hadronization was produced with Pythia8 8.212 [90] with the A14 tune [91]. NLO parton-jet matching in merging between the generator and the parton showering for ggF events is achieved by the FxFx-merging technique [84].

Reconstructed events are subject to a GEANT4-based detector simulation [92,93], and the standard ATLAS reconstruction used for collision data. On truth level samples no detector simulation has been performed and no pile-up has been taken into account. Jet candidates are reconstructed using the anti-k_t algorithm with a distance parameter of $R = 0.4$.

5.4.1 ggF $H \to W^{\pm} W^{\mp *} \to \ell \bar{\nu}_{\ell} \ell' \nu_{\ell'}$

This study represents the minimal example shown in Section 5.3.2 since only the SM and one additional BSM parameter that is κ_{AWW} are considered. These parameters apply exclusively in the decay vertex. The κ_{Hgg} coupling that appears in the production vertex is set constant to the SM value for all samples and drop out in the calculation of the morphing function, such that just second order coupling polynomials emerge. Three samples are necessary as inputs for the morphing function. The chosen morphing basis are listed in Table 5.4 and consists of one SM sample (`s0`) and two mixed samples with the same SM coupling and opposite sign values for the AWW coupling (`s1` and `s2`). In addition, two validation samples `v0` and `v1` are generated, at which parameters the functionality of the morphing method is demonstrated. The events are produced on truth level and the object and event selection follows closely the selection of the analysis presented in Chapter 6 up to the 2-jets cut requirement.

The observable used in this example is the total transverse momentum $p_{\text{T}}^{\text{tot}}$, defined as $p_T^{\ell_1} + p_T^{\ell_2} + E_{\text{T}}^{\text{miss}} + \sum p_T^{\text{jets}}$, where the sum over jets runs over all jets present in the respective event. In Figure 5.11 the distributions of the tree input samples `s0-s2` are shown as well as the distribution of the validation samples `v0` and `v1`.

Two different test are carried out. First, the morphing method is used to reproduce the distribution at certain parameter values. Here, the parameters at which the validation samples has been produced, are provided as input to the morphing and the distribution generated by the morphing function are compared to

Table 5.4.: Morphing basis and parameters for validation samples for the ggF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. Constant parameters are $\Lambda = 1000\,\text{GeV}$, $\cos\alpha = 1/\sqrt{2}$ and $\kappa_{\text{Hgg}} = \sqrt{2}$. All remaining Wilson coefficients are set to 0.

Sample	κ_{SM}	κ_{AWW}
s0	$\sqrt{2}$	0
s1	$\sqrt{2}$	24.02
s2	$\sqrt{2}$	-24.02
v0	$1.05{\cdot}\sqrt{2}$	4.80
v1	$0.95{\cdot}\sqrt{2}$	-16.82

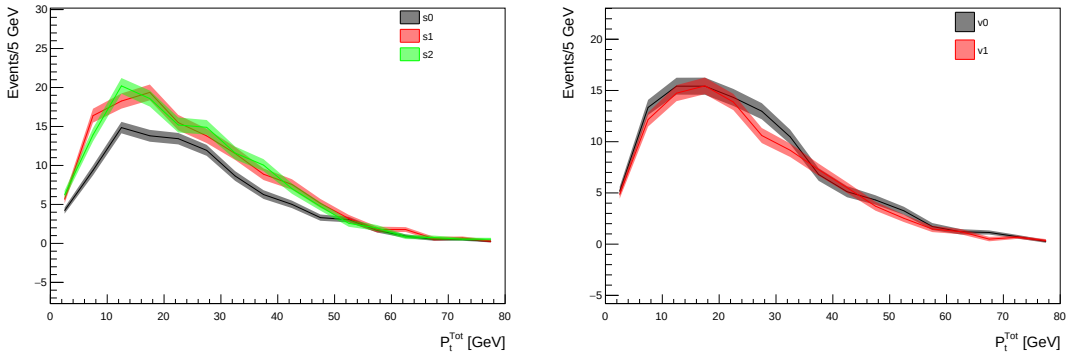


Figure 5.11.: Distribution of $p_{\text{T}}^{\text{tot}}$ for the three input samples s0-s2 (left) and the validation samples v0 and v1 generated at parameters listed in Table 5.4.

the independently generated validation distribution which should agree inside the statistical uncertainties. The second test is to not provide the original parameters of the validation distribution and directly fit the morphing function to the validation distribution to extract the coupling parameters. Since the distributions are often degenerated in the parameter space, that is at two or a set of different parameter points the same cross section and shape are expected for a distribution of a certain observable, one does not certainly expect a perfect match to the original parameters of the validation sample. The common case are observables which are not sensitive to the sign of a coupling and identical distributions are expected at plus and minus the respective parameter. Ideally, in this case, the fit will result in parameters which are preferred due to the same statistical fluctuation of the morphing and validation distribution. Such a parameter fit reflects a real case analysis in which the nature of the EFT couplings are unknown and need to be determined from physics data.

Figure 5.12 shows the results of both types of tests for both validation samples. The agreement between the prediction of the morphing (black) and the validation distribution (blue) are in excellent agreement inside the statistical uncertainty.

The fit parameters listed inside the subfigures are in perfect agreement to the original validation parameters, although especially the large uncertainties on the κ_{AVV} parameter show the degeneracy of the distributions throughout the parameter space for this particular analysis.

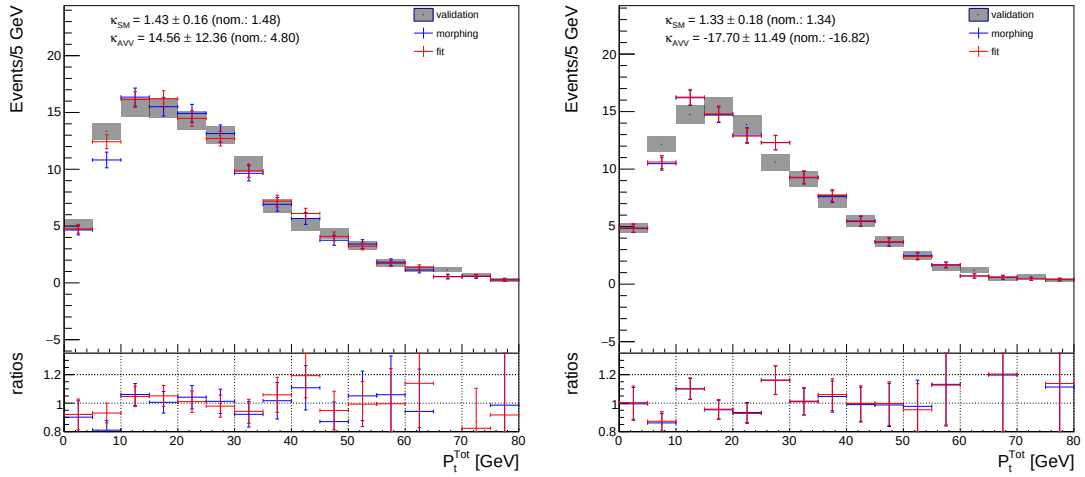


Figure 5.12.: Comparison of the morphing predictions (blue) to the validation distribution (black) and the morphing distribution generated by a fit to the validation distribution (red) for both validations `v0` (left) and `v1` (right). The respective nominal parameters as well as the parameters with uncertainties resulting from the fit are listed inside the respective subfigure. The ratios of the morphing and fit prediction are with respect to the validation distribution.

This degeneracy at least for the total cross section is nicely shown in the left distribution in Figure 5.13. Here, the cross section is plotted as function of both parameters κ_{SM} and κ_{AVV}. It is symmetrical for positive and negative AVV parameters and only have small changes for κ_{SM} at fixed κ_{AVV}, which explains the large uncertainties for κ_{AVV} in the fits in Figure 5.12.

In the right distribution of Figure 5.13 the relative uncertainty on the cross section in the parameter space is shown and additionally the parameter points at which the input samples for the morphing has been produced. Naively, one would expect that the uncertainties are especially small the closer one considers target parameters to one of the input parameter points as it would be valid for simple interpolation techniques, because the smaller the difference between the parameters are the better the respective input distribution emulates the distribution at these target parameters. But since the Effective Lagrangian Morphing is not a usual interpolation method this is not the case. It is clearly visible that the relative

uncertainty are even smaller in between the input parameters compared with the cross section uncertainty at the input parameters. The choice of the morphing basis is therefore crucial for the resulting uncertainties all over the parameter space.

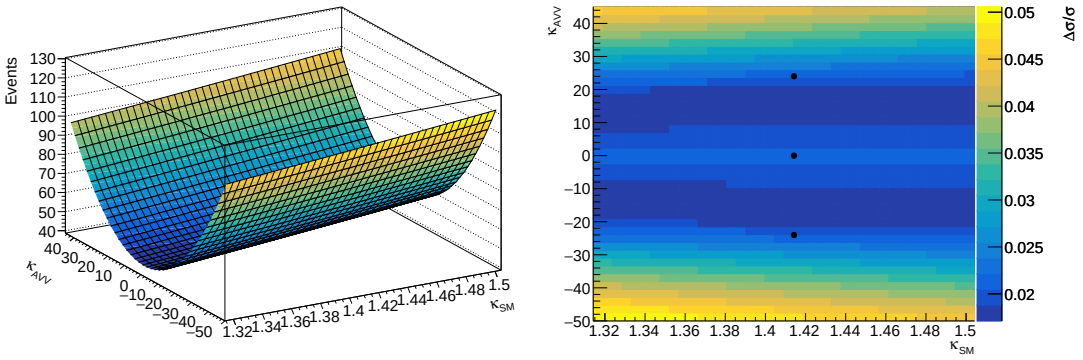


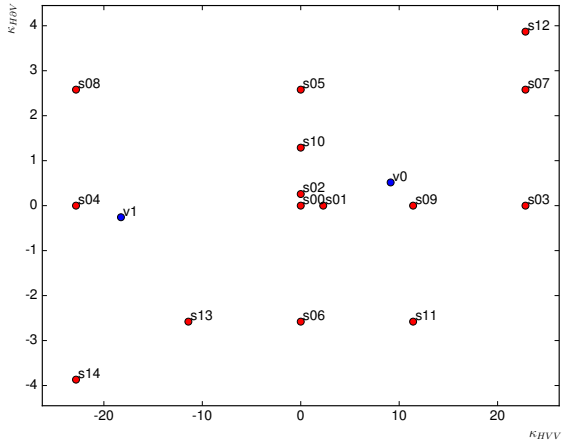
Figure 5.13.: Expected number of events as function of the free parameters κ_{SM} and κ_{AVV} (left) and the respective relative uncertainty on the cross section (right). The luminosity is set to $10\,\mathrm{fb}^{-1}$. The black dots in the right distribution are the positions at which the input samples for the morphing have been generated.

5.4.2 VBF $H \to W^{\pm}W^{\mp *} \to \ell \bar{\nu}_{\ell} \ell' \nu_{\ell'}$

The second study comprises a more complicated case where three couplings κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ are taken into account. Since the process VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$ has similar production and decay vertices the parameters affect both vertices which leads to 15 required morphing inputs according to Equation 5.3.24. A list of the morphing basis as well as the parameters for the validation benchmarks can be found in Figure 5.14.

Similar to the ggF $H \to W^{\pm}W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$ study in the previous section, the events are produced on truth level and the object and event selection follows closely the selection of the analysis presented in Chapter 6 up to the 2-jets cut requirement. At this cut stage the 15 input and two validation distributions for the observable $\Delta\phi_{jj,\mathrm{sign}}$ are generated (cf. Figure 5.15). Compared to the previous analysis, clear cross section and shape differences are visible for the different distributions such that one can expect much smaller uncertainties on fitted parameters.

Again, both tests are performed to both validation distributions shown in Figure 5.16. The predictions from the morphing (blue) to the validation distributions (black) are in excellent agreement, as well as the distribution at the resulting fit parameters (red) to the validation distribution. Also the fitted parameters are in good agreement with the original validation parameters and as indicated before, the



Sample	κ_{SM}
s01 & s02	0
s00 & s03-s14	$\sqrt{2}$
v0	$1.05{\cdot}\sqrt{2}$
v1	$0.95{\cdot}\sqrt{2}$

Figure 5.14.: Morphing basis and parameters for the validation samples for the VBF $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. The projection of the parameters in the κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ space are draw on the left, where the red dots represent the morphing basis and the blue dots the validation parameters. The SM parameters are listed right. Constant parameters are $\Lambda = 1000\,\mathrm{GeV}$ and $\cos\alpha = 1/\sqrt{2}$. The parameters κ_{HWW} and κ_{HZZ} are set equal labeled as κ_{HVV} as well as $\kappa_{\mathrm{H}\partial\mathrm{WR}}$, $\kappa_{\mathrm{H}\partial\mathrm{WI}}$ and $\kappa_{\mathrm{H}\partial\mathrm{Z}}$ defined as $\kappa_{\mathrm{H}\partial\mathrm{V}}$. All remaining Wilson coefficients are set to 0.

uncertainties on these fitted parameters are much smaller compared to the previous study, as expected. Performing a χ^2-test between the validation distribution v0 and the prediction of the morphing gives a corresponding p-value of 0.80, which indicate a good agreement. An even better result gives a χ^2-test between the distribution for the fitted values to the validation distribution with a p-value of 0.96, which indicates that the fit finds more suitable parameters then the nominal validation parameters. That is of course due to statistical fluctuations. For the validation v1 the p-values evince an even better agreement with both at 0.98.

Furthermore, one can extract the correlation matrices from the fits which are shown in Table 5.5. High (anti-)correlations are present. For example $\kappa_{\mathrm{H}\partial\mathrm{V}}$ and κ_{HVV} are totally anti-correlated at the parameters from validation sample v0. This can be potentially used to reduce the number of free parameters by going into a diagonalized parameter space without loosing analytical power. This task is examined in more details in Section 5.5.2.

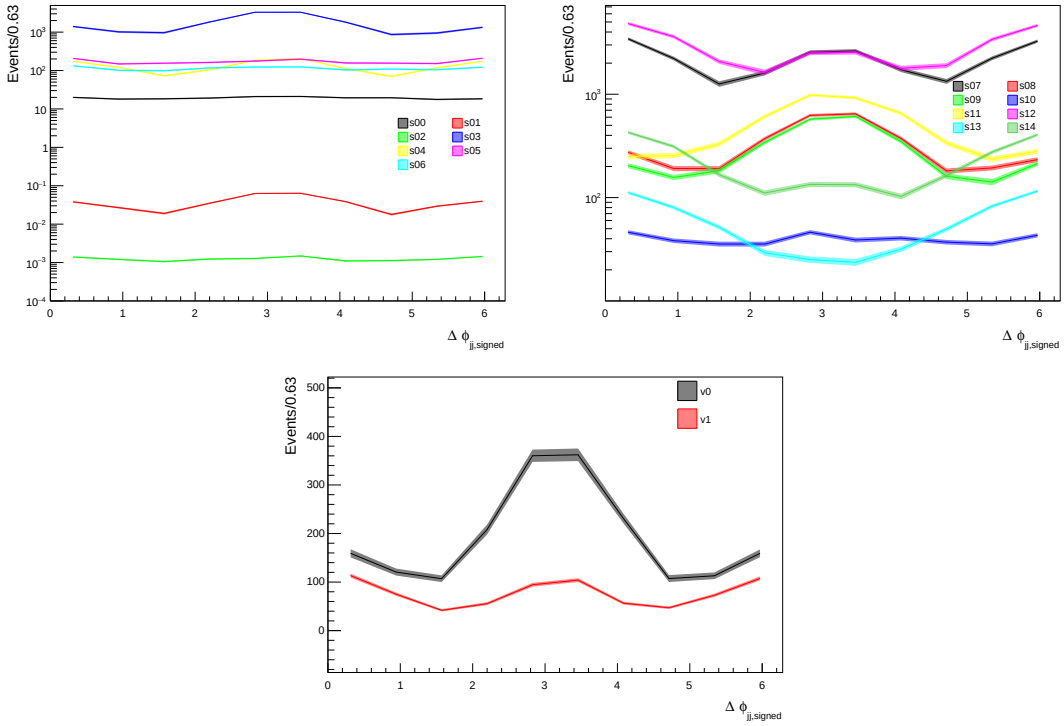


Figure 5.15.: Input (`s00`-`s14`) and validation (`v0` and `v1`) distributions for the $\Delta\phi_{jj,\text{sign}}$ observable. The corresponding parameters are shown in Figure 5.14.

Table 5.5.: Correlation matrices for the fit of the morphing function to the validation distributions `v0` (left) and `v1` (right) shown in Figure 5.16.

	κ_{SM}	κ_{HVV}	$\kappa_{\text{H}\partial\text{V}}$
κ_{SM}	1.00	-0.87	0.87
κ_{HVV}	-0.87	1.00	-1.00
$\kappa_{\text{H}\partial\text{V}}$	0.87	-1.00	1.00

	κ_{SM}	κ_{HVV}	$\kappa_{\text{H}\partial\text{V}}$
κ_{SM}	1.00	-0.70	-0.70
κ_{HVV}	-0.70	1.00	0.91
$\kappa_{\text{H}\partial\text{V}}$	-0.70	0.91	1.00

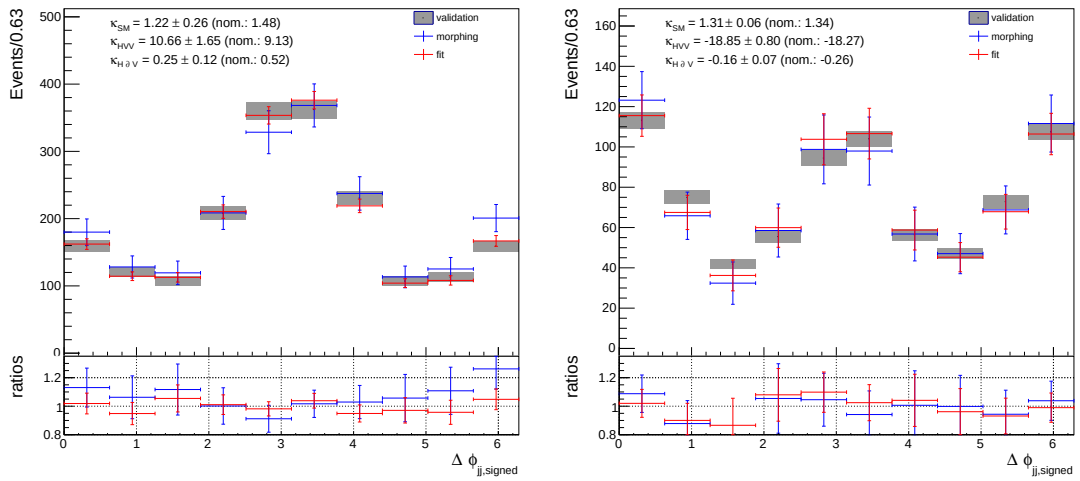


Figure 5.16.: Comparison of the morphing predictions (blue) to the validation distribution (black) and the morphing distribution generated by a fit to the validation distribution (red) for both validations `v0` (left) and `v1` (right). The respective nominal parameters as well as the parameters with uncertainties resulting from the fit are listed inside the respective subfigure. The ratios of the morphing and fit prediction are with respect to the validation distribution.

In Figure 5.17 the cross section predictions as function of the parameters are shown as predicted by the morphing in two dimensional distributions keeping the respective third parameter at the SM case value, i.e. 0 for BSM parameters and $\sqrt{2}$ for κ_{SM}. In general a tendency is discernible to have higher event yields for increasing values of any coupling parameter. Also, the relative uncertainties on the cross section predictions of the morphing are displayed in Figure 5.18 with an equivalent slicing of the parameter space as in Figure 5.17. Mostly in the parameter space of interest the uncertainty stays small, except for an anomaly at ca. $\kappa_{\mathrm{HVV}} = -10$, $\kappa_{\mathrm{H}\partial\mathrm{V}} = -4$ and $\kappa_{\mathrm{SM}} = \sqrt{2}$ where the uncertainty peaks to a value of $\sigma/\Delta\sigma \approx 25$. Making only a small change to the morphing basis by exchanging the input sample `s14` with the validation `v0` already reduces the relative uncertainty to a value of approximately 1 (cf. Figure 5.19) which shows how significant the impact of the input parameters to the uncertainty is.

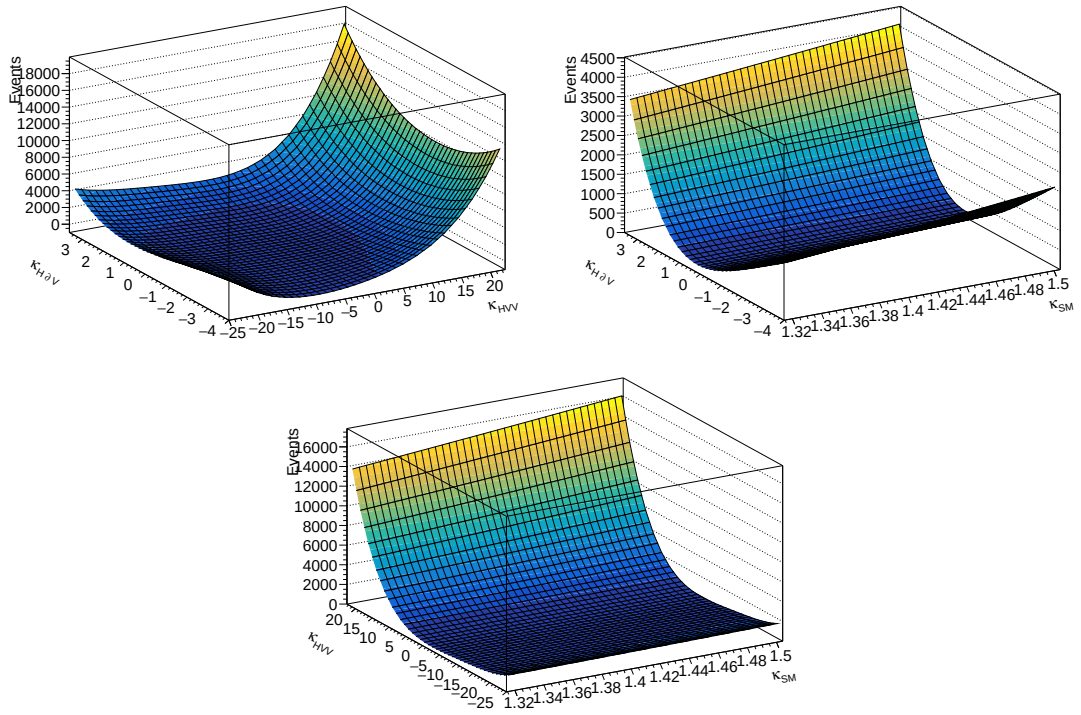


Figure 5.17.: Expected number of events as function of the free parameters κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ (top left), κ_{SM} and κ_{HVV} (top right), and κ_{SM} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ (bottom) where the respective third parameter is set to the SM value,i.e. $\kappa_{\mathrm{BSM}} = 0$ and $\kappa_{\mathrm{SM}} = \sqrt{2}$. The luminosity is set to $10\,\mathrm{fb}^{-1}$.

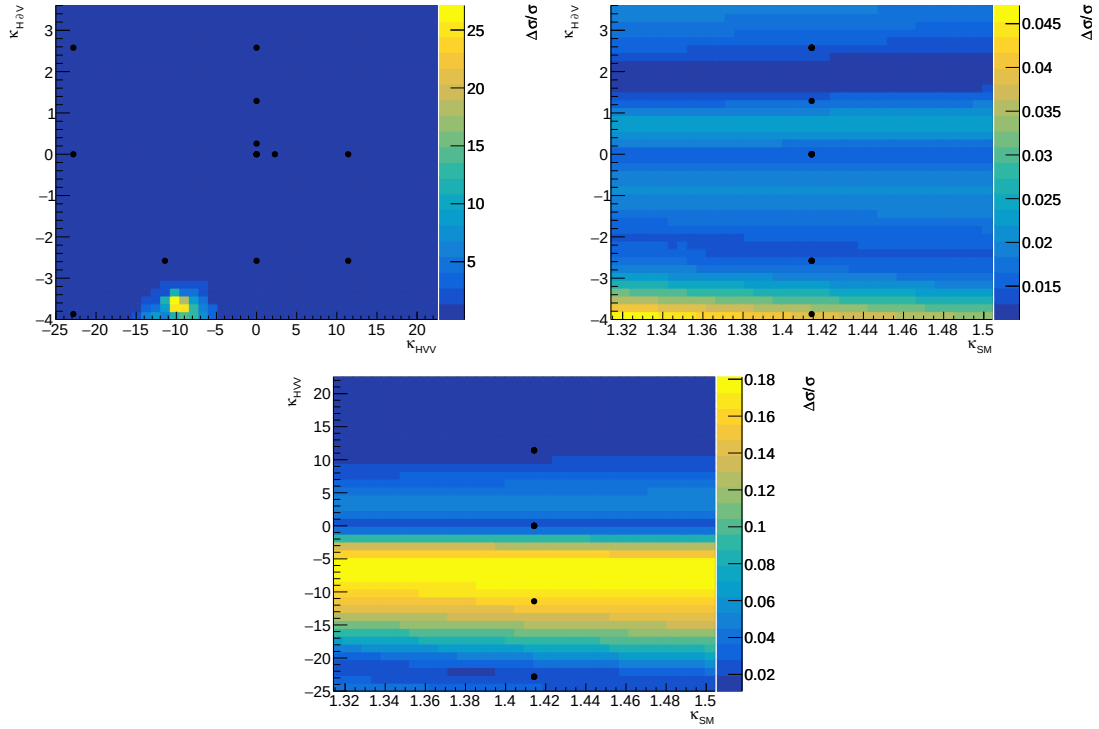


Figure 5.18.: Relative uncertainty on the cross section as function of the free parameters κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ (top left), κ_{SM} and κ_{HVV} (top right), and κ_{SM} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ (bottom) where the respective third parameter is set to the SM value,i.e. $\kappa_{\mathrm{BSM}} = 0$ and $\kappa_{\mathrm{SM}} = \sqrt{2}$. The black dots in the right distribution are the projected positions at which the input samples for the morphing have been generated.

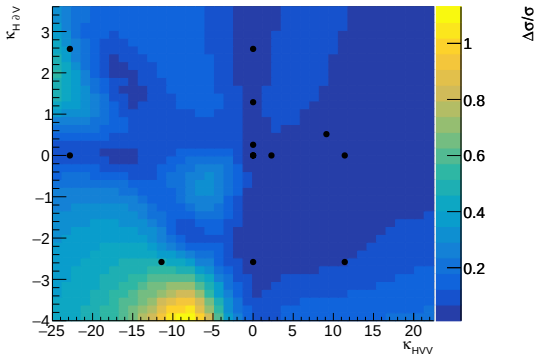


Figure 5.19.: Relative uncertainty on the cross section as function of the free parameters κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ after exchanging input sample `s14` with validation sample `v0`.

5.4.3 ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$

The last validation study shown here has been already presented in [87] using fully reconstructed MC samples of the Run-1 $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ tensor structure analysis [26] at $\sqrt{s} = 8\,\text{TeV}$. Only the dominating ggF production mode is considered. In this analysis two separate coupling mixing have been investigated, i.e. the SM coupling with either the CP-even κ_{HZZ} or CP-odd κ_{AZZ} coupling using the parametrization of the Higgs Characterization model with redefined couplings according to

$$\tilde{\kappa}_{\text{HZZ}} = \frac{1}{4} \cdot \frac{v}{\Lambda} \cdot \kappa_{\text{HZZ}} \quad \text{and} \quad \tilde{\kappa}_{\text{AZZ}} = \frac{1}{4} \cdot \frac{v}{\Lambda} \cdot \kappa_{\text{AZZ}}, \tag{5.4.1}$$

with the vacuum expectation value v and the cut-off scale Λ.

The MC events are produced in two parts. First the SM Higgs boson production via gluon fusion is simulated using the POWHEG-BOX [94] generator with a Higgs mass $m_H = 125.5\,\text{GeV}$. The Higgs boson decay for the non-SM signals are simulated, according to the Higgs boson parity assumptions, using the JHU [95,96] MC generator at leading order. The event selections is applied according the tensor structure analysis of [26]. The morphing basis as well as the parameters chosen for the validation samples are shown in Table 5.6.

Table 5.6.: Morphing basis and parameters for validation samples for the ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ study.

	$\tilde{\kappa}_{\text{SM}}$	$\tilde{\kappa}_{HZZ}$	$\tilde{\kappa}_{AZZ}$
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 ($\text{TO}_2(\tilde{\kappa}_{\text{AZZ}}, \alpha)$)	1	0	5
Validation sample 4 ($\text{TO}_2(\tilde{\kappa}_{\text{AZZ}}, \alpha)$)	1	0	3.25

Two observables are probed. The first one is the spin and parity sensitive angle Φ between the decay planes of two lepton pairs matched to the two Z-boson decays expressed in the four-lepton rest frame (cf. Figure 5.20). The second discriminant is

the transformed optimal observable $\mathrm{TO}_2(\widetilde{\kappa}_{AZZ}, \alpha)$ derived from the corresponding optimal observable $\mathrm{O}_2(\widetilde{\kappa}_{\mathrm{AZZ}}, \alpha)$ defined as

$$\mathrm{O}_2(\widetilde{\kappa}_{\mathrm{AZZ}}, \alpha) = \frac{\left|\mathcal{M}\left(\widetilde{\kappa}_{\mathrm{AVV}} \neq 0, \kappa_{\mathrm{SM}} = \widetilde{\kappa}_{\mathrm{HVV}} = 0, \alpha = \frac{\pi}{2}\right)\right|^2}{\left|\mathcal{M}\left(\widetilde{\kappa}_{\mathrm{SM}} \neq 0, \kappa_{\mathrm{AVV}} = \widetilde{\kappa}_{\mathrm{HVV}} = 0, \alpha = \frac{\pi}{2}\right)\right|^2}, \tag{5.4.2}$$

which is then normally distributed to the SM scenario. By construction the TO_2 observable is sensitive to the modulus of the $(\tilde{\kappa}_{\mathrm{AZZ}}/\tilde{\kappa}_{\mathrm{SM}}) \cdot \tan\alpha$ ratio. The single samples used for the validation are not statistically independent but rather received by using the Matrix Element reweighting on a single sample with large statistics. Therefore, as seen in Figure 5.21, a perfect agreement of the morphing result with the validation distributions is expected.

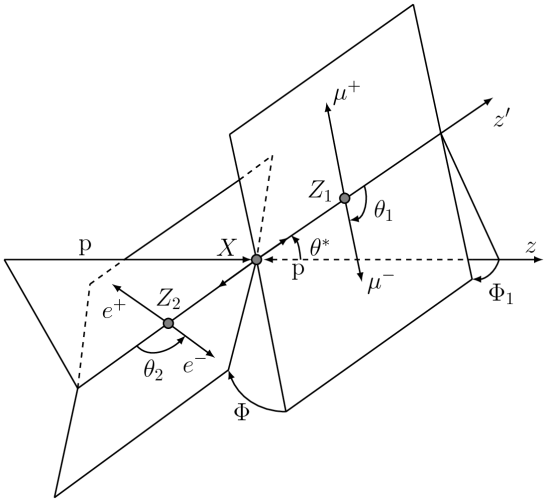


Figure 5.20.: Definitions of the angular observables sensitive to the spin and parity of the resonance in the $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ decay [26].

This validation is a perfect showcase of how the Matrix Element reweighting can be utilized to generate the input sample for the Effective Lagrangian Morphing. In addition, it can be noted that since a different generator as MADGRAPH has been used, opposing to the previous validation studies, this validation shows the independence of the morphing technique on the choice of the MC generator.

More validation studies can be found in Reference [87] with similar satisfying results.

5.5 Improvements

An optimal analysis would incorporate all coupling parameters that the model provide to obtain a complete description of the underlying physics, with the SM and

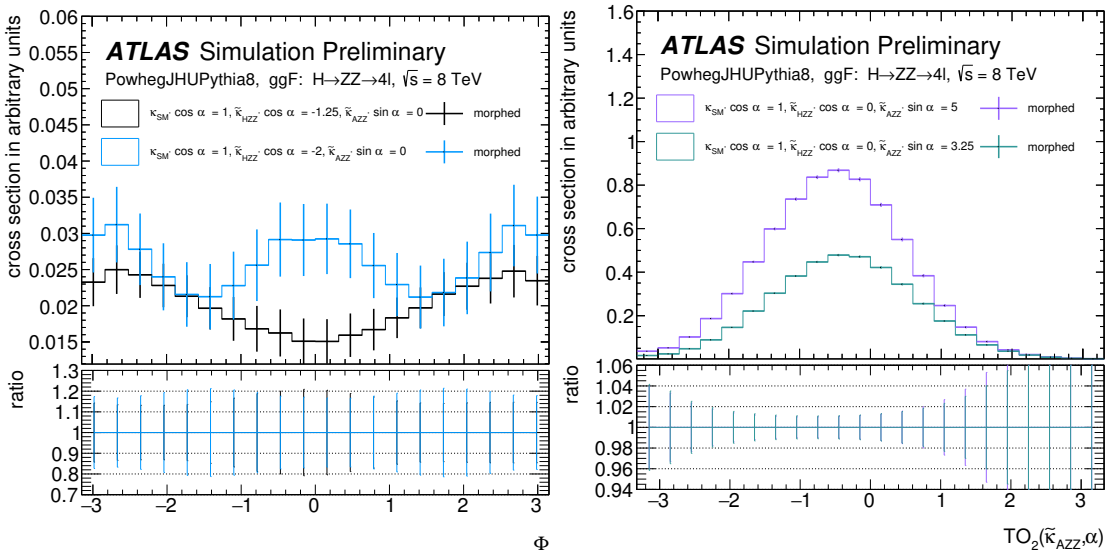


Figure 5.21.: Distributions of ggF $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ events generated at $\sqrt{s} = 8\,\text{TeV}$ [26], comparing the shapes of physical observables Φ and $\text{TO}_2(\tilde{\kappa}_{AZZ}, \alpha)$ for two validation benchmarks samples with the prediction acquired via Effective Lagrangian Morphing. The perfect agreement is due to the fact that all input and validation samples have been derived from the same original sample via Matrix Element reweighting [87].

every BSM contributions and all interference effects not only between the SM and the BSM couplings but also between the BSM couplings themselves. But intuitively one expects that a higher dimensionality, i.e. more free coupling parameters renders the optimizations increasingly harder. This can be observed in the case of the Effective Lagrangian Morphing, too.

In the next section the dimensionality problem is investigated in more detail and in the subsequent sections different methods are presented to find solutions either to reduce the dimensionality without short-cuts to the physics description or to find optimal morphing bases with small uncertainties in a parameter configuration space of interest.

5.5.1 The Curse of Dimensionality

For low dimensionalities it is fairly easy to find a suitable morphing basis which gives reasonable results in terms of propagated uncertainties of the morphing inputs in the configuration space of interest. Mostly, a sufficient method for problems with two or three free coupling parameters is to evenly distribute the parameter points for the base samples inside the predefined region. The chance of randomly generate a morphing matrix which is not invertible is practically zero and judging

the quality of the morphing basis is not straightforward since the uncertainties remain often very stable and small for low dimensional analyses.

But already for four free parameters a basis with evenly distributed parameter points in the region of interest can lead to very large propagated uncertainties and even negative number of events in distributions as Figure 5.22 demonstrates. Here, events for VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ for EFT configurations with four couplings κ_{SM}, κ_{HVV}, κ_{AVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$, visualized in the parameter space in Figure 5.23 have been generated and morphed to the validation parameters shown in the same figure (blue dot). The actual shape of the distribution is not visible anymore due to the large uncertainties. Although the yields in every bin is compatible with a positive cross section, some of them have negative predictions.

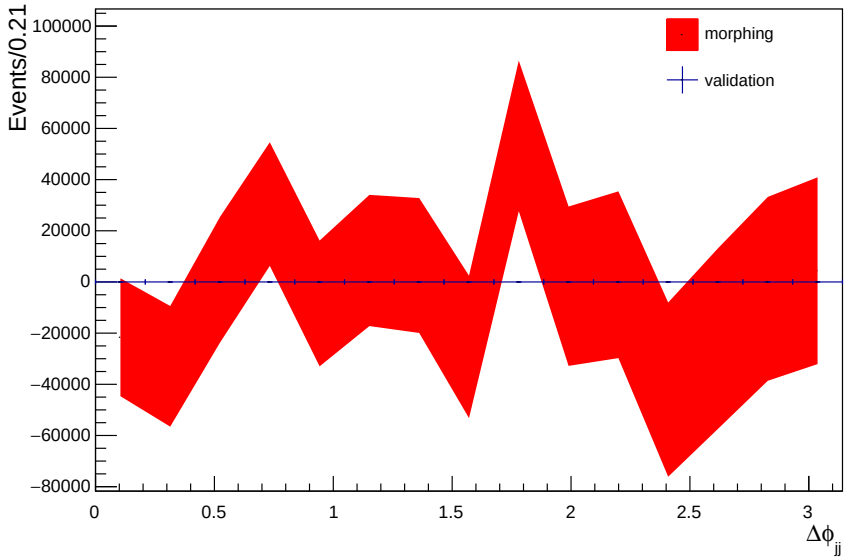


Figure 5.22.: Example of a morphed distribution with large uncertainties. Inputs are events for VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ using four couplings κ_{SM}, κ_{HVV}, κ_{AVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$. The morphing basis and the parameters for the validation benchmark are shown in Figure 5.23.

This shows an intuitive misconception to assume that base samples chosen only in the configuration space of interest would decrease the propagated uncertainty of the morphing. The Effective Lagrangian Morphing is not a typical interpolation technique like the Integral morphing presented in Section 5.2.2 or the Moment morphing presented in Section 5.2.3 in the sense that a sampling near the parameters, at which a morphing result is desired, gives a more precise outcome. The iterative optimization algorithm presented in the coming Section 5.5.4 demonstrates even

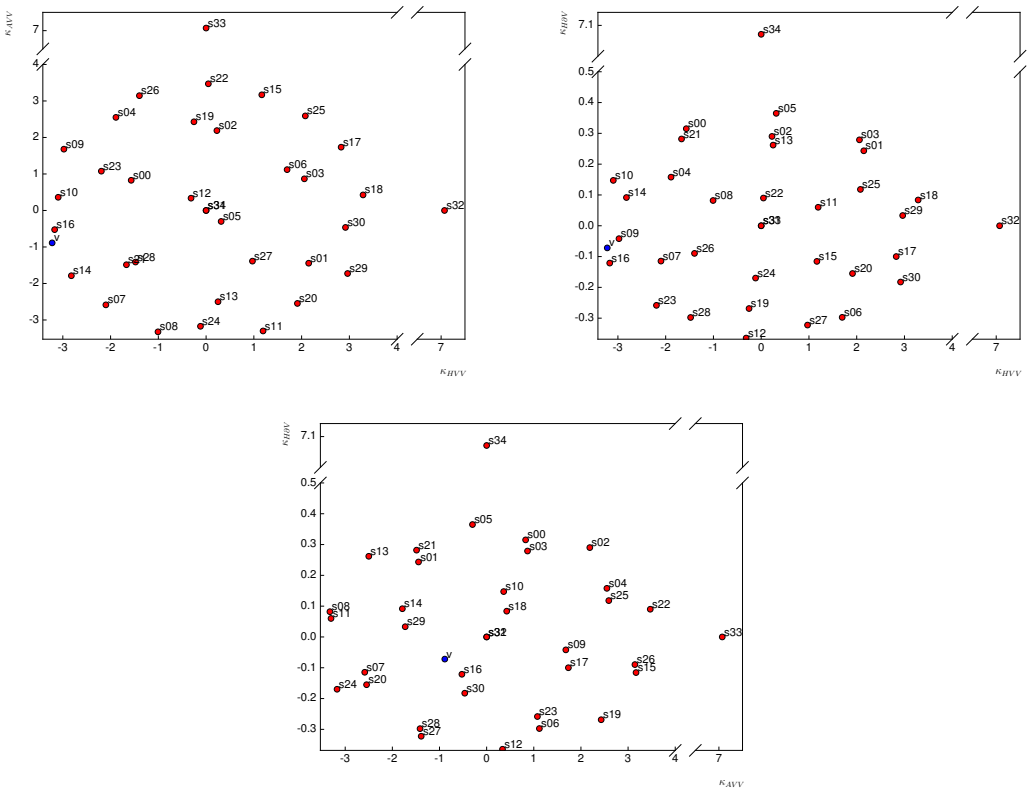


Figure 5.23.: Morphing basis s00-s34 (red dots) which was used for Figure 5.22 and parameters for the respective validation sample v (blue dot). Constant parameters are $\Lambda = 1\,\text{TeV}$, $\cos\alpha = 1/\sqrt{2}$ and κ_{SM} is set to $\sqrt{2}$ for all samples except s32-s34 for which this parameter is 0.

the contrary, i.e. base samples far outside the parameter region of interest stabilize the uncertainties inside this parameter region.

To demonstrate the difficulties going to analyses with higher dimensionalities a showcase study is constructed in the following way. First, the cross section as function of the coupling parameters are modelled as a fourth order polynomial as they would appear in a $2 \to 2$ actual process. The g_{SM} parameter and exemplary for all additional BSM contributions the g_{HVV} parameters are used and real cross sections at four different parameters are calculated by MADGRAPH in the process VBF $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ listed in Table 5.7.[3] With these cross sections, the general formula for a fourth order polynomial $\sigma(g) = ag + bg^2 + cg^3 + dg^4$ is solved

[3]The couplings g represent the full expression before the respective operator term in the Higgs Characterization model, i.e. $g_{\text{SM}} = \kappa_{\text{SM}} \cos\alpha$ and $g_{\text{HVV}} = \kappa_{\text{HVV}} \cos\alpha/\Lambda$.

by matrix inversion. The resulting full cross section formula reads

$$\sigma(g_{\text{SM}}, \{g_{\text{BSM}_i}\}) = \big(-2.16 g_{\text{SM}} + 4.44 g_{\text{SM}}^2 - 3.01 g_{\text{SM}}^3 + 0.82 g_{\text{SM}}^4 + \sum_i (-3.25 \cdot 10^{-4} g_{\text{BSM}_i} + 0.67 g_{\text{BSM}_i}^2 - 453 g_{\text{BSM}_i}^3 + 125145 g_{\text{BSM}_i}^4) \big) \text{ pb}. \tag{5.5.1}$$

This formula describes by no means the physically correct cross sections throughout the parameter phase space, since for example the interferences between SM and BSM are not taken into account, but they depict an approximately realistic behavior along the axes and are sufficient for this purpose.

Table 5.7.: Calculated cross sections by MADGRAPH in the process VBF $H \to W^{\pm} W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ for different parameters g_{SM} (left) and g_{HVV} (right). In each case all other couplings are set to 0.

g_{SM}	Cross section [pb]
1	7.911×10^{-2}
2	1.259
3	6.313
4	2.019×10^{1}

g_{HVV}	Cross section [pb]
1×10^{-3}	1.193×10^{-5}
2×10^{-3}	1.932×10^{-4}
3×10^{-3}	9.696×10^{-4}
4×10^{-3}	3.092×10^{-3}

Then 100 benchmarks are determined by finding equally distributed points in a n-dimensional parameter space using a 100-body system $1/r^2$-potential, where r is the distance between two benchmarks. Limits are set as follow

$$\begin{aligned} g_{\text{SM}} &= 1 \pm \frac{1}{50} \\ g_{\text{BSM}} &= \pm \frac{1}{50}, \end{aligned} \tag{5.5.2}$$

such that for the benchmarks the maximal cross section is approximately 25% above the SM cross section, e.g. $\sigma(1 + \frac{1}{50}, 0) \approx \sigma(1, \frac{1}{50}) \approx 1.25\sigma(1, 0) = 1.25\sigma_{\text{SM}}$. The respective cross sections at these benchmark parameter points are calculated using Equation 5.5.1.

For different dimensionalities n, i.e. different numbers of BSM parameters in addition to the SM parameter, many morphing bases are randomly generated in the same parameter limits as the benchmarks and the corresponding cross sections are again calculated with Equation 5.5.1. For all the cross section calculations a relative uncertainty of 5% is assumed. The relative cross section uncertainty is calculated at the benchmark points for each morphing setup. Since the identical cross section formula, which is taken to be exact, is used for the benchmarks and

the inputs for the morphing, the prediction of the morphing is identical to the benchmark cross sections, but the propagated 5% uncertainties of the morphing inputs vary dependent on the choices of the morphing basis. To be independent on the number of benchmarks and number of randomly picked morphing bases the propagated relative uncertainties are summed and averaged out.

A second case is studied to demonstrate also the influence of the cross section magnitudes, where the parameters are assumed to be all BSM like and equal such that there is always the same suppression from the Λ parameter unlike it would be the case including the SM parameter. Here, the limits for the benchmark and morphing bases parameters are set as in Equation 5.5.2 using only the BSM limit and the cross sections are evaluated as in Equation 5.5.1 using only the BSM polynoms.

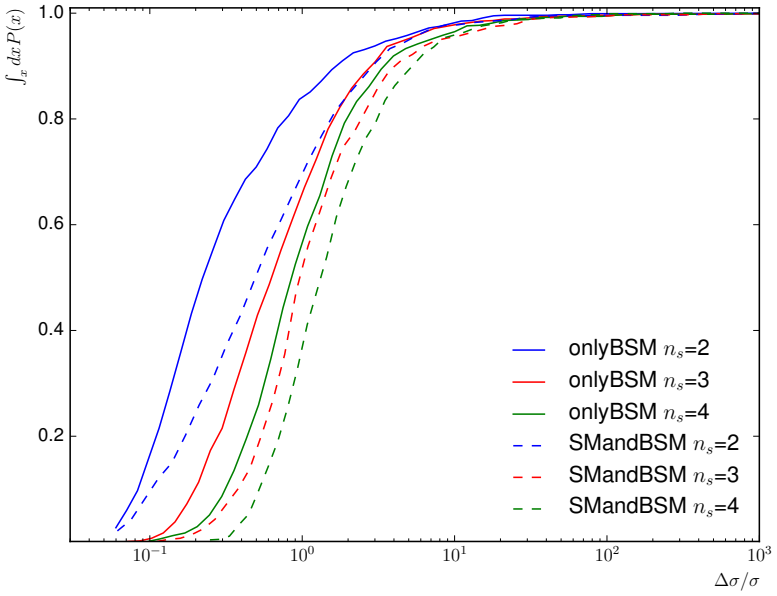


Figure 5.24.: Cumulative probability distributions of finding a randomly chosen morphing basis as function of the relative uncertainties on the Effective Lagrangian prediction evaluated at equally distributed benchmark points for different number of model parameter shared in both production and decay vertex. The limits of the parameter space is given by Equation 5.5.2 and the calculation of the cross section by Equation 5.5.1. Two cases are presented: `SMandBSM` includes a SM parameter and `onlyBSM` incorporate only BSM parameters.

The cumulative probability distributions of finding a randomly chosen morphing basis as function of the relative uncertainties shows how difficult it is to find a proper

morphing basis by chance for different dimensionalities. Figure 5.24 demonstrates the cases for dimensions 2 to 4, where the parameters contribute to the production and decay vertices, for both cases including the SM parameter (`SMandBSM`) and without (`onlyBSM`). It is firstly evident that the more parameter are included in the model the increasingly harder it is to find randomly a good basis, i.e. a basis with a small relative cross section uncertainty.

A second observation of this showcase is, that also differences in the cross section magnitude of the morphing inputs exacerbate the difficulty of finding a good morphing basis. This happens by having a SM like parameter which is not suppressed by a factor Λ compared to all other BSM like parameters (`SMandBSM`). If all parameters are setup equally and no parameter is enhanced (`onlyBSM`), a randomly chosen morphing basis gives on average better results.

Examining the uncertainty formula gives a better insight. One could e.g. select coupling points at which the cross sections are equal for all input samples, $\sigma = \sigma_i \,\forall\, i$. So not only the contributions for the cross section from all parameters are in the same order of magnitude, but in addition the input cross sections itself are equal. Since the cross sections are just polynomials of the couplings, this is certainly possible by taking coupling points along the contours of equal cross sections. The formula for the relative uncertainty 5.3.27 reduces to

$$\Delta_{\text{rel}}\sigma(\boldsymbol{g}) = \delta \cdot \frac{\sqrt{\sum_{i=1}^{N} w_i^2(\boldsymbol{g})}}{\sum_{i=1}^{N} w_i(\boldsymbol{g})}, \tag{5.5.3}$$

having the relative uncertainty only dependent on the morphing weights.

Alternatively, the morphing basis could be chosen such that for fixed target parameters $\boldsymbol{g}$ all the weights are equal, $w = w_i(\boldsymbol{g}) \,\forall\, i$. The uncertainty in this case

$$\Delta_{\text{rel}}\sigma(\boldsymbol{g}) = \delta \cdot \frac{\sqrt{\sum_{i=1}^{N} \sigma_i^2}}{\sum_{i=1}^{N} \sigma_i} \tag{5.5.4}$$

is smaller than the individual contributions. 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 cross section, that determines the absolute value of the uncertainty.

In a sense, the *weight* with which the relative uncertainty on σ_i propagates to the final prediction is the product of $w_i\sigma_i$. Therefore, the optimal basis needs to be determined by taking into account both, the values of the physical observables and the entries of the morphing matrix and both need to be balanced in a way that they contribute equally to the full relative uncertainty. An analytic approach finding the optimal solution is not yet found but an iterative numerical method has been developed with excellent results presented in Section 5.5.4.

5.5.2 Dimensionality Reduction

Analyses with low dimensionality are easier to conduct not only because they require less input samples but also a suitable morphing basis is easier to find as has been shown in the previous section. Therefore, investigating which couplings can be safely neglected in specific studies and still exploit the full underlying physics model is of great interest.

The following generator level study targets specifically the $VV \to H$ production vertex by utilizing the Higgs Characterization model introduced in Section 2.3.4. It describes this vertex with 13 coupling parameters that are listed in Table 5.8. Specifically the production vertex is examined and is therefore decoupled from the decay vertex by just requesting the Higgs boson to decay to two muons. With 13 parameters contributing just in the production vertex, 91 samples are necessary as inputs for the Effective Lagrangian Morphing as can be calculated using Equation 5.3.23. Each of these samples are generated with 30 000 events at $\sqrt{s} = 13\,\text{TeV}$ with MADGRAPH5_aMC@NLO 2.2.2.PL [84]. To preserve gauge invariance not only the VBF production mode was included but also diagrams where the Higgs boson is produced with a hadronically decaying associated vector boson. Parton showering and hadronization are simulated with PYTHIA8 8.186 with A14 Tune [91].

The parameter configurations of the samples are chosen under certain requirements. First, the couplings values are limited such that they reflect the current experimental limits of the Run-1 Higgs production cross section measurement $\sigma_{\text{VBF}} \lesssim 1.25 \cdot \sigma_{\text{VBF,SM}}$ [27]. The used coupling values for the different parameters listed in Table 5.8 reproduce the SM cross section for a pure BSM sample. Since no significant deviations from the SM are found, the SM parameter is included in each sample in addition to one or two BSM couplings with the respective values from the table. Furthermore, the parameters of the samples are chosen in such a way that it is possible to construct flexibly subsets of parameters with lower dimensionality in the morphing.

The events are preselected. Two opposite charged muons with $p_T > 6\,\text{GeV}$ and $|\eta| < 2.7$ are required as well as at least two anti-k_T ($R = 0.4$) truth jets with $p_T > 20\,\text{GeV}$, $|\eta| < 5.0$ and $\Delta\eta_{jj} > 5$ in the final state, which results in a VH suppressed phase space. The jets with the highest p_T are chosen to be the jets emerging from the VBF production vertex, if more than two jets fulfill the jet requirements.

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 $p_{\text{T}}^{j_1}$ are highly sensitive to the effects of non-SM couplings, as shown exemplary in Figure 5.25 for the samples with non-zero HWW coupling contribution. 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}, $\kappa_{\mathrm{HZ}\gamma}$ and κ_{HWW}, while the distributions $p_{\mathrm{T}}^{j_1}$ and $\Delta\eta_{jj}$ are sign-sensitive for $\kappa_{\mathrm{H}\partial\mathrm{Z}}$ and $\kappa_{\mathrm{H}\partial\mathrm{WR}}$.

Table 5.8.: The listed values of the coupling parameters reproduce the SM cross section at $\sqrt{s} = 13\,\mathrm{TeV}$. 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 for the presented VBF production vertex study. [87]

Parameter	Value
κ_{SM}	1
$\kappa_{\mathrm{H}\gamma\gamma}$	203.22
$\kappa_{\mathrm{A}\gamma\gamma}$	408.62
$\kappa_{\mathrm{HZ}\gamma}$	109.13
$\kappa_{\mathrm{AZ}\gamma}$	986.88
κ_{HZZ}	5.75
κ_{AZZ}	6.96
κ_{HWW}	3.36
κ_{AWW}	3.92
$\kappa_{\mathrm{H}\partial\mathrm{WR}} = \mathfrak{R}(\kappa_{\mathrm{H}\partial\mathrm{W}})$	0.76
$\kappa_{\mathrm{H}\partial\mathrm{WI}} = \mathfrak{I}(\kappa_{\mathrm{H}\partial\mathrm{W}})$	0.84
$\kappa_{\mathrm{H}\partial\mathrm{A}}$	1.77
$\kappa_{\mathrm{H}\partial\mathrm{Z}}$	1.37

To maximize the sensitivity of the VBF production vertex to the BSM effects a combined observable $\mathrm{var}_{4\mathrm{d}}$ is constructed in the following manner

$$\mathrm{var}_{4\mathrm{d}} = 3^3 \cdot i_{\Delta\eta_{jj}} + 3^2 \cdot i_{m_{jj}} + 3 \cdot i_{p_{\mathrm{T}}^{j_1}} + i_{\Delta\phi_{jj}}, \tag{5.5.5}$$

where the values of each of the four jet variables is divided into three bins

$$i_x = \begin{cases} 0 & x < \left(\min_x + \frac{1}{3} \cdot (\max_x - \min_x)\right) \\ 1 & \left(\min_x + \frac{1}{3} \cdot (\max_x - \min_x)\right) \leq x < \left(\min_x + \frac{2}{3} \cdot (\max_x - \min_x)\right) \\ 2 & x \geq \left(\min_x + \frac{2}{3} \cdot (\max_x - \min_x)\right). \end{cases}$$

In this way all bin-by-bin correlations of the four variables are taken into account. The minimum and maximum values of each variable are listed in Table 5.9 and the distribution of events in this combined observable is shown in Figure 5.26.

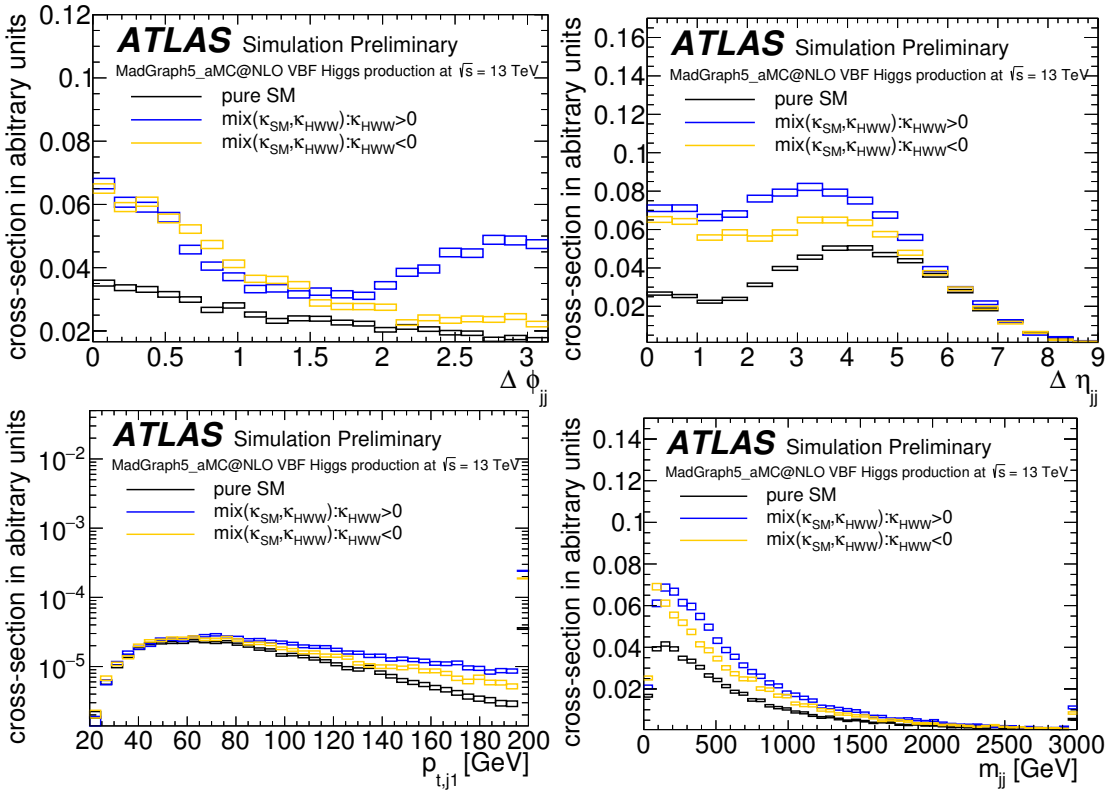


Figure 5.25.: VBF $H \to \mu\mu$ distributions for $\Delta\phi_{jj}$ (top left), m_{jj} (top right), $p_{\mathrm{T}}^{j_1}$ (bottom left), and $\Delta\eta_{jj}$ (bottom right) with mixing of SM and BSM CP-even HWW couplings. The box sizes correspond to the MC statistical uncertainties. [87]

Using this observable a likelihood is constructed and fit to pseudo-data generated at SM parameters assuming a statistical uncertainty of 8% on the total cross section. The resulting uncertainties on the fit parameters then allow to assess the influence of the respective parameter on the kinematic properties of events arising from VBF Higgs boson production, and thus on the principle capacity to measure these parameters in such events. Table 5.10 summarizes the post-fit values and uncertainties of all coupling parameters.

The uncertainties provide insights about the sensitivities of the signal distribution to the parameters for a given scenario, where large uncertainties correspond to small sensitivity. As could be expected, the influence of the Higgs boson couplings to photons $\kappa_{\mathrm{A}\gamma\gamma}$, $\kappa_{\mathrm{AZ}\gamma}$, $\kappa_{\mathrm{H}\gamma\gamma}$ and $\kappa_{\mathrm{H}\partial\gamma}$ have very little influence on VBF events. One major result of this study 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.

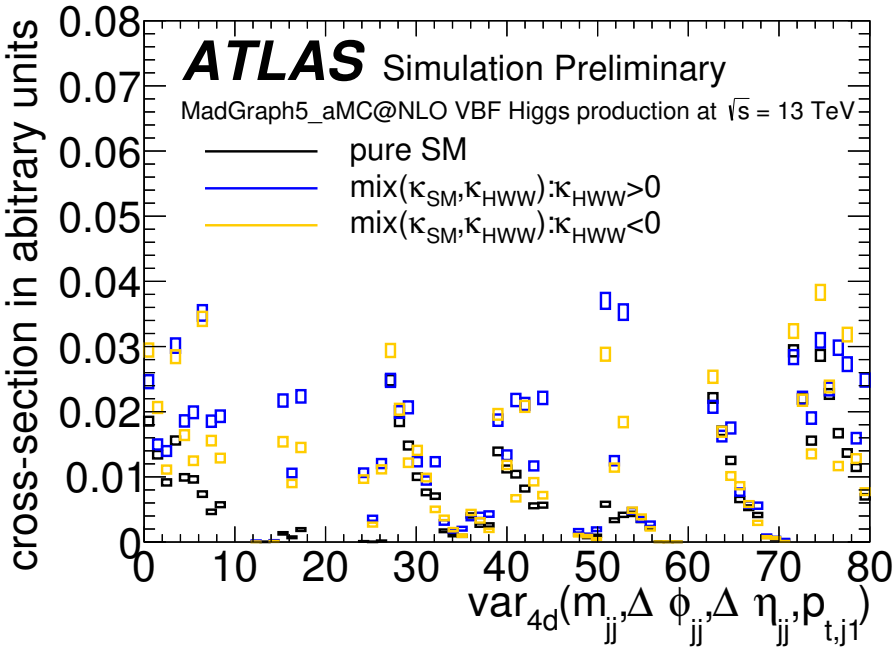


Figure 5.26.: VBF $H \to \mu\mu$ distribution of $\text{var}_{4\text{d}}$, which is a combination of $\Delta\eta_{jj}$, dphijj, m_{jj} and $p_{\text{T}}^{j_1}$ into one discriminating variable, shown for SM and mixing of SM with BSM CP-even HWW couplings. The box sizes correspond to the MC statistical uncertainties. [87]

Table 5.9.: Minimum and maximum values of the VBF production jet variables $\Delta\eta_{jj}$, m_{jj}, $p_{\text{T}}^{j_1}$ and $\Delta\phi_{jj}$. [87]

	$\Delta\eta_{jj}$	m_{jj} [GeV]	$p_{\text{T}}^{j_1}$ [GeV]	$\Delta\phi_{jj}$
min	0	0	20	0
max	6	1000	200	π

However, it would be desirable to reduce the dimensionality of the model even further. Looking at the post-fit correlations in Figure 5.27 some of the parameters are either highly correlated or anti-correlated as for example the parameters $\kappa_{\text{H}\partial\text{Z}}$ and $\kappa_{\text{H}\partial\text{WR}}$. One can imagine that a variable transformation can further reduce the number of free parameters. New parameters can be constructed by rotating into the Eigen basis of the covariance matrix and thus eliminate correlated parameters as illustrated in Figure 5.28.

In order to understand more clearly what happens during the diagonalization, exemplary a reduced model with only three parameters, $\kappa_{\text{H}\gamma\gamma}$, $\kappa_{\text{A}\gamma\gamma}$ and κ_{SM} is constructed and all the other parameters are set to zero. Figure 5.29 shows the post-fit profile likelihood between all three parameter combinations and the

Table 5.10.: Values of the coupling parameters and their respective uncertainties after fitting to SM pseudo-data with 8% cross section uncertainty. Values without uncertainties are fixed to their nominal values during the fit.

parameter	post-fit value	+	−
Λ	1000.		
$\cos\alpha$	0.71		
$\kappa_{\mathrm{H}\ell\ell}$	1.41		
$\kappa_{\mathrm{A}\gamma\gamma}$	0	+219	−441
κ_{AWW}	0	+3	−2.6
$\kappa_{\mathrm{AZ}\gamma}$	0	+441	−398
κ_{AZZ}	0	+2.7	−1.3
$\kappa_{\mathrm{H}\gamma\gamma}$	0	+236	−91
$\kappa_{\mathrm{H}\partial\gamma}$	0	+0.3	−0.6
$\kappa_{\mathrm{H}\partial\mathrm{WI}}$	0	+1.6	−0
$\kappa_{\mathrm{H}\partial\mathrm{WR}}$	0	+0.5	−0.3
$\kappa_{\mathrm{H}\partial\mathrm{Z}}$	0	+1.2	−0.5
κ_{HWW}	0	+1.5	−3
$\kappa_{\mathrm{HZ}\gamma}$	0	+38	−49
κ_{HZZ}	0	+8	−2.5
κ_{SM}	1.41	+0.22	−0.11

respective correlation matrix which is calculated at the minimum of the fit. The Hessian approximation of the likelihood is overlayed as green contours in the figures. It is obvious that the Hessian approximation does not fit the real likelihood behavior between the SM and the two BSM parameters, and that a sign ambiguity is present, which will render a wrong diagonalization transformation. Still, ignoring the likelihood structure the diagonalization algorithm can be applied which results in

$$\begin{aligned}
\kappa_{\mathrm{A}\gamma\gamma} &= -1.000000 \cdot \mathrm{EV}_0 - 0.054000 \cdot \mathrm{EV}_1 - 0.000004 \cdot \mathrm{EV}_2 \\
\kappa_{\mathrm{H}\gamma\gamma} &= -0.054000 \cdot \mathrm{EV}_0 + 1.000000 \cdot \mathrm{EV}_1 + 0.000022 \cdot \mathrm{EV}_2 \\
\kappa_{\mathrm{SM}} &= -0.000003 \cdot \mathrm{EV}_0 - 0.000022 \cdot \mathrm{EV}_1 + 1.000000 \cdot \mathrm{EV}_2
\end{aligned}$$

After the transformation, a fit of the diagonalized model to the SM pseudo-data is performed and the results are shown in Figure 5.30. As expected, the rotation into another parameter basis does not resolve the sign ambiguities and as can be seen in the post-diagonalization correlation matrix in the same figure, the correlations are at the same order of magnitude, concluding that the diagonalization does not

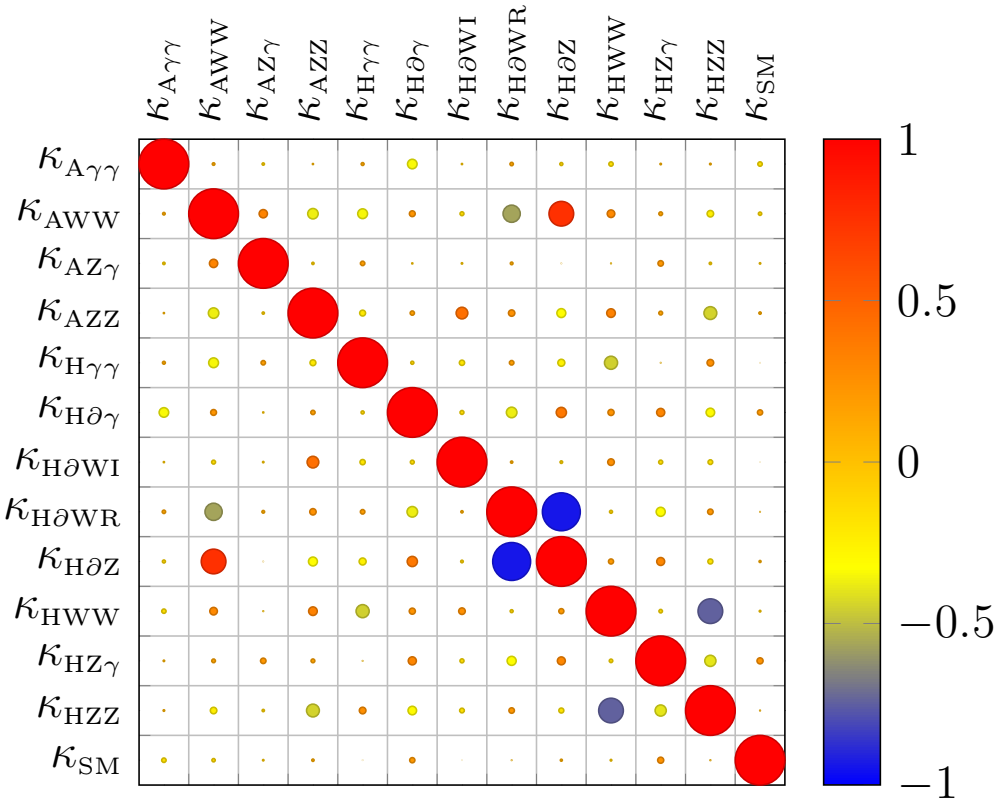
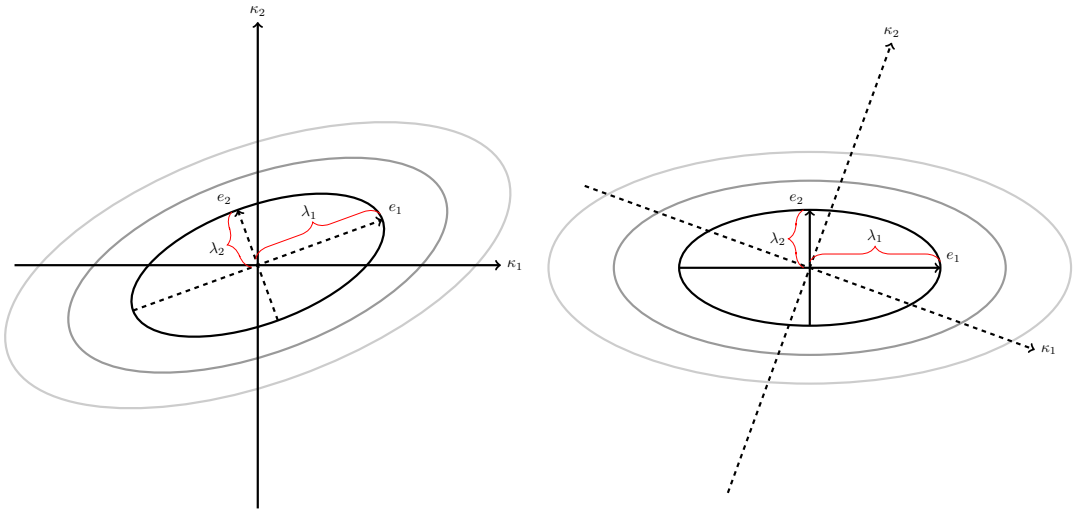


Figure 5.27.: Correlation matrix of the coupling parameters after fitting to SM pseudo-data with 8% cross section uncertainty. The area of the marker size is proportional to the correlation.

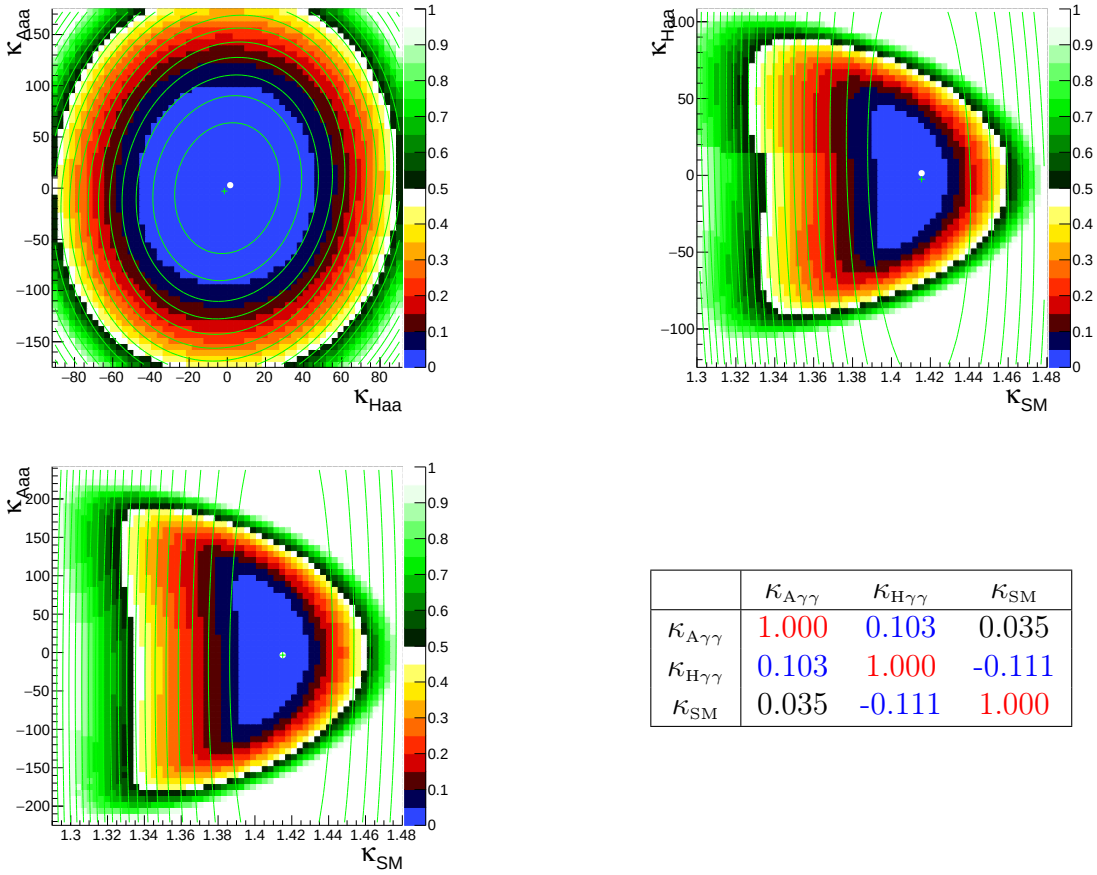
work in this specific case. A modification e.g. in the fit input distribution may resolve the sign ambiguities.

Thus, reducing the parameter space by rotating into another Eigen basis is not trivial and must be analyzed individually in each analysis for every parameter combinations and in each considered signal region. Nevertheless, a closer look into the likelihood distributions leads to worthwhile insights how the parameters are connected to each other and how to possibly distinguish them in a better way.



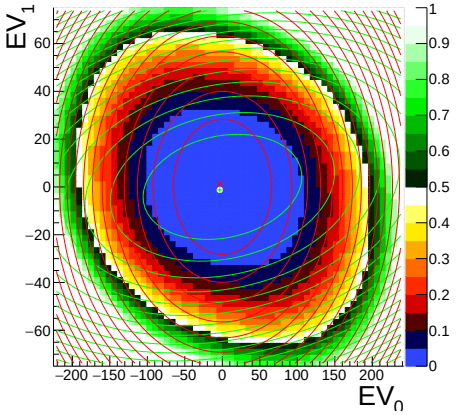
(a) Basis before the rotation. **(b)** Basis after the rotation.

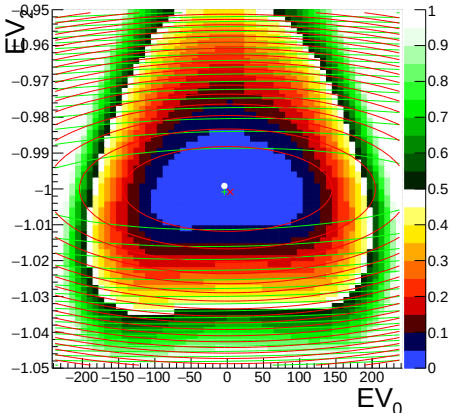
Figure 5.28.: Illustration of the re-parametrization into the Eigen basis of the covariance matrix. The ellipses indicate the likelihood contours coming from a fit. The e_i vectors represent the Eigen directions with Eigen value λ_i.

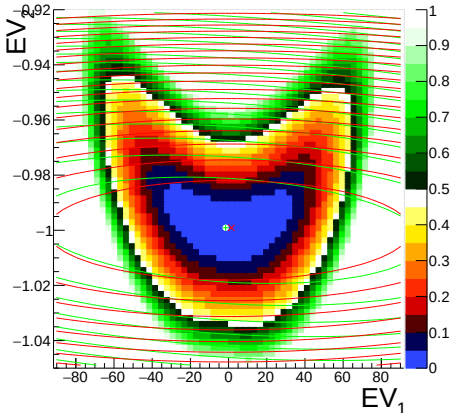


	$\kappa_{A\gamma\gamma}$	$\kappa_{H\gamma\gamma}$	κ_{SM}
$\kappa_{A\gamma\gamma}$	1.000	0.103	0.035
$\kappa_{H\gamma\gamma}$	0.103	1.000	-0.111
κ_{SM}	0.035	-0.111	1.000

Figure 5.29.: The profiled likelihood plots after a fit is performed on the reduced model for $\kappa_{H\gamma\gamma}$ vs. $\kappa_{A\gamma\gamma}$ (top left), $\kappa_{H\gamma\gamma}$ vs. κ_{SM} (top right) and $\kappa_{A\gamma\gamma}$ vs. κ_{SM} (bottom left), and the corresponding correlation matrix (bottom right). The z-axis indicates the likelihood and the overlaid green lines show the Hessian approximation of the likelihood as calculated in the fit minimum.







	EV_0	EV_1	EV_2
EV_0	1.000	0.193	0.069
EV_1	0.193	1.000	-0.214
EV_2	0.069	-0.214	1.000

Figure 5.30.: The likelihood plots after a fit on the diagonalized reduced model for EV_0 vs. EV_1 (top left), EV_0 vs. EV_2 (top right) and EV_1 vs. EV_2 (bottom left), and the corresponding correlation matrix after diagonalization (bottom right). The z-axis indicates the likelihood and the overlaid green lines show the Hessian approximation of the likelihood as calculated in the fit minimum. The overlaid red lines show the shape of the likelihood when the transformation is applied to the Hessian approximation before diagonalization.

5.5.3 Coupling Interferences

In many analyses one knows already from physics insights that some interference terms are zero. This might be the case when for example CP-even and CP-odd couplings appear in the same vertex. The Lagrangian describing an anomalous interaction between the Higgs boson H and a vector boson V with both CP contributions takes the form

$$\mathscr{L} = \left\{ a_1 V_\mu V^\mu + a_2 V_{\mu\nu} V^{\mu\nu} + b V_{\mu\nu} \widetilde{V}^{\mu\nu} \right\} H. \tag{5.5.6}$$

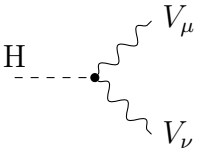
Here,

$$V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \tag{5.5.7}$$

is the reduced field strength tensor, and

$$\widetilde{V}^{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma} \tag{5.5.8}$$

is the dual tensor. The three couplings are the SM coupling a_1 the two BSM couplings a_2, which is CP-even and b which is CP-odd. The tensor structure of the $H \to VV$ vertex on tree level derives from this Lagrangian to



$$\propto \quad a_1 g^{\mu\nu} + a_2 (q_1^\mu q_2^\nu - q_1 q_2 g^{\mu\nu}) + b \epsilon^{\mu\nu\rho\sigma} q_{1\rho} q_{2\sigma}, \tag{5.5.9}$$

which will be a factor in the matrix element. However, the contraction of the totally anti-symmetric 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.

Looking at the case of the process VBF $H \to W^\pm W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ and assuming that the SM operator is CP-even and the BSM operator is CP-odd, the matrix element squared of Equation 5.3.21 reduces to only three terms. All terms including factors with $\mathrm{Re}(\mathcal{M}^*_{\mathrm{SM}} \mathcal{M}_{\mathrm{BSM}})$ have a contraction of $g_{\mu\nu}$ from the SM matrix element term and a $\epsilon^{\mu\nu\rho\sigma}$ from $\mathcal{M}_{\mathrm{BSM}}$. Therefore, the number of polynoms reduce from five to three resulting in fewer base samples that are required in total as input, which simplifies the morphing calculations. Especially the size of the morphing matrix gets smaller and the inversion of the matrix simplifies which leads to reduces sensitivity to large weight factors in the morphing functions and reduced propagated uncertainties.

To visualize these interference effects, a set of MC samples with different parameter settings are generated summarized in Table 5.11. The basis including the interference terms is required to have five samples `s0`-`s5` and the reduced basis excluding the interference terms need only three samples `s0`, `s4` and `s5`. Figure 5.31

Table 5.11.: Morphing bases to visualize the absent interference effects between a CP-even (κ_{SM}) and a CP-odd (κ_{AVV}) coupling in the process VBF $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$. The parameters for the validation sample are shown in the bottom. Constant parameters are $\cos\alpha = 1/\sqrt{2}$ and $\Lambda = 1\,\text{TeV}$.

Sample	κ_{SM}	κ_{AVV}
basis with interference		
s0	$\sqrt{2}$	0
s1		24.02
s2		−24.02
s3		12.01
s4		−12.01
basis without interference		
s0	$\sqrt{2}$	0
s4		−12.01
s5	0	2.40
validation		
v0	$1.05 \cdot \sqrt{2}$	4.80

shows for two different observables, the azimuthal angle between the two p_T-leading jets $\Delta\phi_{jj}$ and the invariant mass of these two jets m_{jj}, the agreement of the morphing prediction in both bases with an independently generated validation sample, which parameters are as well listed in Table 5.11. The agreement for both the full and the reduced basis is excellent and the uncertainties are very similar despite all samples have the same statistics of 100 000 events each and the reduced basis uses two fewer samples.

Nevertheless, one needs to be very cautious about the physics insights. Although it is true that for almost all differential cross sections the interference terms have no contribution in this showcase, there exist observables for which this is not the case. Figure 5.32a shows the same test for the variable $\Delta\phi_{jj,\text{sign}}$. It is obvious that the reduced basis gives completely wrong predictions and that the interference terms cannot be neglected. Figure 5.32b shows the $\Delta\phi_{jj,\text{sign}}$ distributions for two samples, both with $\kappa_{\text{SM}}\cos\alpha = 1$ but different signs for the κ_{HVV} parameter at ± 2.4. Since all the terms of Equation 5.3.21 without the interference terms have an even number as exponent for the g_{BSM} coupling one would expect both distributions to agree within the statistical uncertainty. But it is apparent from the distributions that the differential cross section has a dependency in the form

$$\frac{\partial\sigma}{\partial\Delta\phi_{jj,\text{sign}}} \propto g_{\text{AVV}} \cdot \sin(\Delta\phi_{jj,\text{sign}}). \tag{5.5.10}$$

Therefore one can conclude that for this observable the interference terms cannot be neglected.

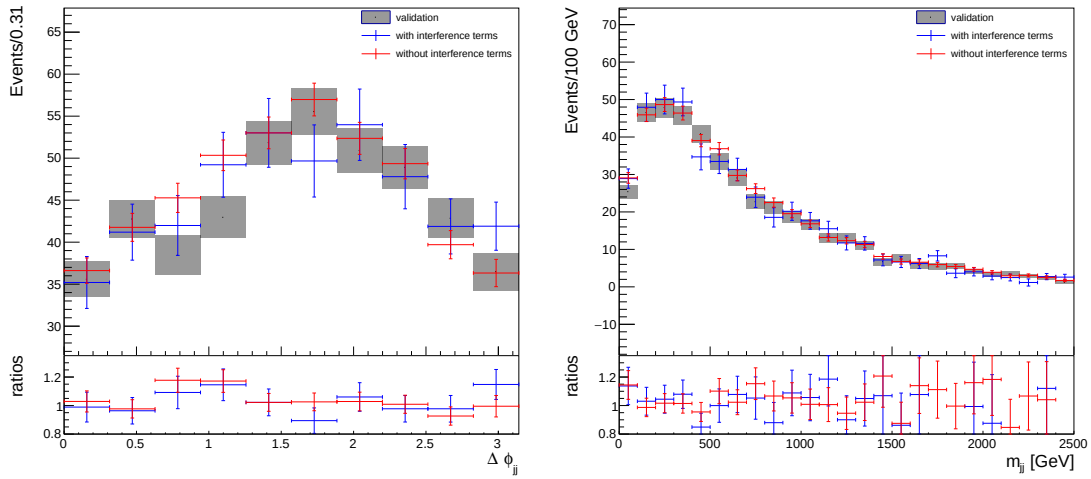
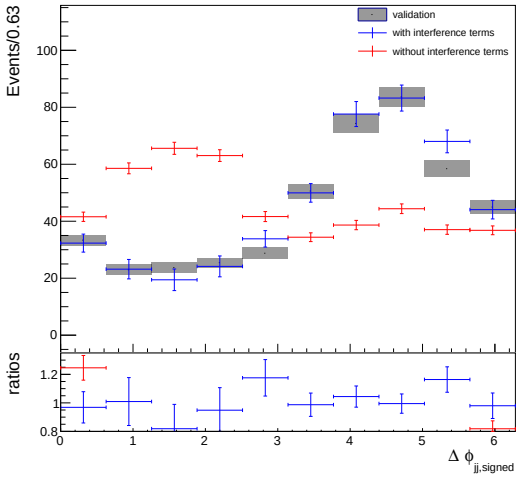
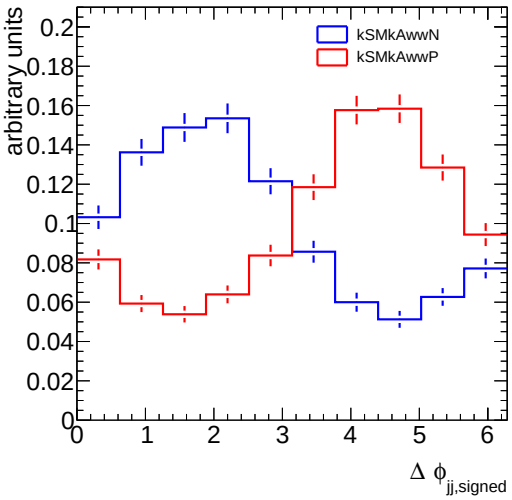


Figure 5.31.: Prediction of the morphing of $\Delta\phi_{jj}$ (left) and m_{jj} (right) compared to the validation distribution (black) for morphing bases including the interference terms (blue) and without the interference terms (red). The coupling parameters of the inputs and validation is listed in Table 5.11.



(a) Prediction of the morphing of $\Delta\phi_{jj,\text{sign}}$ compared to the validation distribution (black) for morphing bases including the interference terms (blue) and without the interference terms (red). The coupling parameters of the inputs and validation is listed in Table 5.11.



(b) Distribution of $\Delta\phi_{jj,\text{sign}}$ for two parameter settings, both with $\kappa_{\text{SM}} \cos\alpha = 1$, `kSMkAwwN` (blue) with $\kappa_{\text{HVV}} = +2.4$ and `kSMkAwwP` (red) with $\kappa_{\text{HVV}} = -2.4$.

Figure 5.32.: Illustration of the interference effects in the observable $\Delta\phi_{jj,\text{sign}}$ for the process VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ between the CP-even coupling κ_{SM} and the CP-odd coupling κ_{AVV}. The interference terms cannot be neglected in the morphing function.

5.5.4 Morphing Basis Optimization

As shown in Section 5.5.1, an analysis with a high-dimensional parameter space poses a challenge for finding a usable morphing basis. Often the physical properties of the system help to reduce the parameter dimensionality, but this is not always possible without losing the generality of the analysis. A preferable solution is to have a generic procedure to finding an optimal basis.

At first, it is necessary to quantify a good usable morphing basis $\overline{\boldsymbol{g}} = \{\boldsymbol{g}_i\}_{i=1}^{N}$, which consists of $n \times N$ free parameters, where N are the number of required morphing inputs and n is the dimensionality of the parameter space. The fundamental figure of merit is of course the uncertainty $\Delta\sigma(\boldsymbol{g}; \overline{\boldsymbol{g}})$ of the prediction obtained from the morphing shown in Equation 5.3.25. Still, it would not be sufficient to test just one specific target parameter point $\boldsymbol{g} \in \mathbb{R}^n$ as it does not guarantee that the uncertainty around this target remain as well sufficiently small. Parameters of a

model are usually already limited by other measurements or theoretical insights of the model itself. In the case of an EFT measurement only small deviations from the SM case are expected. Consequently, it is reasonable to define a parameter region of interest $\mathcal{A} \subset \mathbb{R}^n$, where e.g. the cross sections stay below a 25% increased SM cross section. The objection is then to find a morphing basis which leads to small propagated uncertainties in this specific region. A first idea to quantify the quality of a morphing basis is the maximal uncertainty in the region of interest,

$$f(\overline{\boldsymbol{g}}) = \max_{\boldsymbol{g}\in\mathcal{A}} \left(\Delta\sigma(\boldsymbol{g};\overline{\boldsymbol{g}})\right), \tag{5.5.11}$$

but as seen in Figure 5.13 this figure of merit would mostly depend on the chosen boundary of $\mathcal{A}$, as uncertainties usually diverge strongly at the bounds of the region. A better choice seems to incorporate the integral over the uncertainty in the region of interest,

$$f(\overline{\boldsymbol{g}}) = \int_{\boldsymbol{g}\in\mathcal{A}} d^n\boldsymbol{g}\, \Delta\sigma(\boldsymbol{g};\overline{\boldsymbol{g}}). \tag{5.5.12}$$

Unfortunately, calculating such an integral is not feasible in an analytical approach and would be very time consuming to do so in a numerical MC integration. Instead, a practicable solution is to evaluate the uncertainty at a finite set $\mathcal{B} = \{\boldsymbol{b}_i\}_{i=1}^{M} \subset \mathcal{A}$ of M benchmark parameter points,

$$f(\overline{\boldsymbol{g}}) = \sum_{\boldsymbol{b}\in\mathcal{B}} \Delta\sigma(\boldsymbol{b};\overline{\boldsymbol{g}}). \tag{5.5.13}$$

In principle no samples need to be generated at these benchmark parameters, since only the parameters themselves are used in the figure of merit. However, if they are generated as real samples, a bias term can be included

$$B_b = \sigma(\boldsymbol{b};\overline{\boldsymbol{g}}) - \sigma_b, \tag{5.5.14}$$

with the benchmark cross section σ_b to avoid convergence to regions of parameter space that are subject to large systematic biases.

In addition, the morphing predictions at the benchmark parameters are normalized to the respective MC uncertainty of the benchmark cross sections to account for e.g. regions of parameter space that are difficult to model by the MC generator or the concrete implementation of the model in question and thus associated to large uncertainties.

Accordingly, the final figure of merit used in this analysis reads

$$f(\overline{\boldsymbol{g}}) = \sqrt{\sum_{\boldsymbol{b}\in\mathcal{B}} \left(\frac{\Delta\sigma^2(\boldsymbol{b};\overline{\boldsymbol{g}})}{\sigma_b^2} \cdot \left(\sigma(\boldsymbol{b};\overline{\boldsymbol{g}}) - \sigma_b\right)^2 \right)}. \tag{5.5.15}$$

To find an optimal morphing basis for a fixed configuration many bases need to be tested and the figure of merit minimized. The problem for such a minimization is to have the predictions of the cross sections and their uncertainties at changing parameters and it is of course not sensible to generated new samples for each iteration in the minimization when a new basis is tested. Thus, it is necessary to have a technique to estimate the cross section at any given point in the parameter space. Luckily, the technique of Effective Lagrangian Morphing does exactly that. Therefore, 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 and uncertainties 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. Hence, it would be best to start the minimization of the figure of merit with an already optimized cross section basis. This is obviously the desired final outcome of the optimization and not feasable to provide. But an incremental procedure leads to such an optimized basis for the cross section morphing by introduction a superordinate optimization iteration, where the optimized test basis is applied as the new basis of the cross section morphing in the next iteration. However, one should note that the iteratively replacement of the bases is only partially beneficial, as samples of the optimized test basis do not need to lie in the region of interest covered by the benchmarks, in which the basis was optimized to have small uncertainties. The resulting algorithm is schematically displayed in Figure 5.33.

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.

However, concatenating two morphing functions have the property to be *transitive*, which is proven in a quick derivation. The morphing function from Equation 5.3.13 in matrix form is written as

$$\sigma(\boldsymbol{g}) = \boldsymbol{P}^T(\boldsymbol{g}) A(\overline{\boldsymbol{g}}) \boldsymbol{\sigma}(\overline{\boldsymbol{g}}), \tag{5.5.16}$$

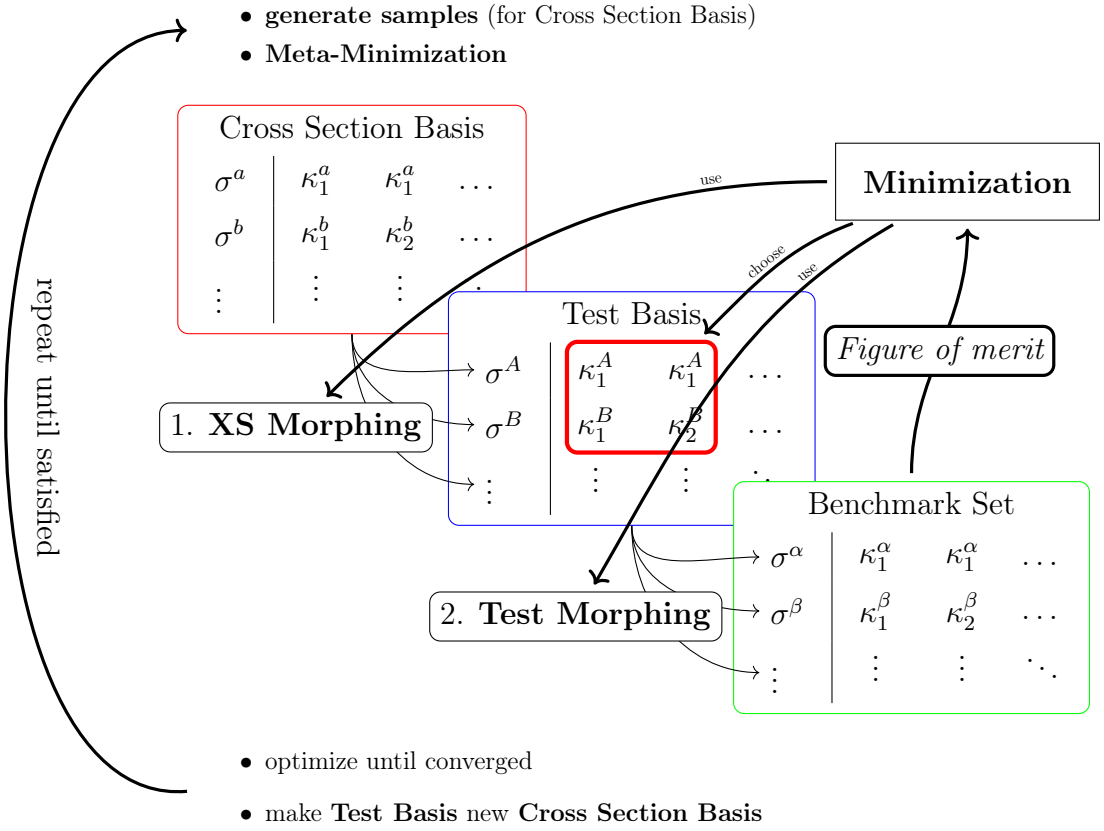


Figure 5.33.: Depiction of the morphing basis optimization algorithm. First, samples are generated at the parameters of the *cross section basis*, which is chosen arbitrary in the first iteration. This morphing provides cross sections and uncertainties for the *test basis*, which calculates the cross section uncertainties at the benchmark parameter points used in the *figure of merit*. The meta-minimization of the figure of merit leads to a better test morphing basis, which is used as the new cross section basis in the next iteration, at which new samples are generated. The superordinated iteration is repeated until satisfactorily small uncertainties are reached.

where

$$\boldsymbol{P}^T(\boldsymbol{g}) = (P_1(\boldsymbol{g}), \cdots, P_N(\boldsymbol{g})) \tag{5.5.17}$$

is a vector with all polynomials evaluated at the target parameters $\boldsymbol{g}$ and

$$\boldsymbol{\sigma}^T = (\sigma(\boldsymbol{g}_1), \cdots, \sigma(\boldsymbol{g}_N)) \tag{5.5.18}$$

is a vector of all input cross sections determined at the morphing basis parameter points $\overline{\boldsymbol{g}}$. Considering now two morphing formulas, where the first provides the

input cross sections for the second,

$$\begin{aligned} \sigma(\boldsymbol{g}) &= \boldsymbol{P}^T(\boldsymbol{g}) A(\overline{\boldsymbol{g}}^{(2)}) \boldsymbol{\sigma}(\overline{\boldsymbol{g}}^{(2)}) \\ \sigma(\boldsymbol{g}_i^{(2)}) &= \boldsymbol{P}^T(\boldsymbol{g}_i^{(2)}) A(\overline{\boldsymbol{g}}^{(1)}) \boldsymbol{\sigma}(\overline{\boldsymbol{g}}^{(1)}), \quad i = 1, \ldots, N \end{aligned} \tag{5.5.19}$$

Writing the overlying morphing function to a vector $\boldsymbol{\sigma}(\overline{\boldsymbol{g}}^{(2)}) = \{\sigma(\boldsymbol{g}_i^{(2)})\}_{i=1}^N$ and plugging it into the first morphing leads to

$$\sigma(\boldsymbol{g}) = \boldsymbol{P}^T(\boldsymbol{g}) A(\overline{\boldsymbol{g}}^{(2)}) \left[\boldsymbol{P}^T(\overline{\boldsymbol{g}}^{(2)}) A(\overline{\boldsymbol{g}}^{(1)}) \boldsymbol{\sigma}(\overline{\boldsymbol{g}}^{(1)})\right] \tag{5.5.20}$$

From Equation 5.3.11 it is clear that $\boldsymbol{P}^T(\overline{\boldsymbol{g}}^{(2)})$ is by construction the inverted matrix $A(\overline{\boldsymbol{g}}^{(2)})$ and Equation 5.5.20 simplifies to a morphing function using directly the first basis,

$$\sigma(\boldsymbol{g}) = \boldsymbol{P}^T(\boldsymbol{g}) A(\overline{\boldsymbol{g}}^{(1)}) \boldsymbol{\sigma}(\overline{\boldsymbol{g}}^{(1)}). \tag{5.5.21}$$

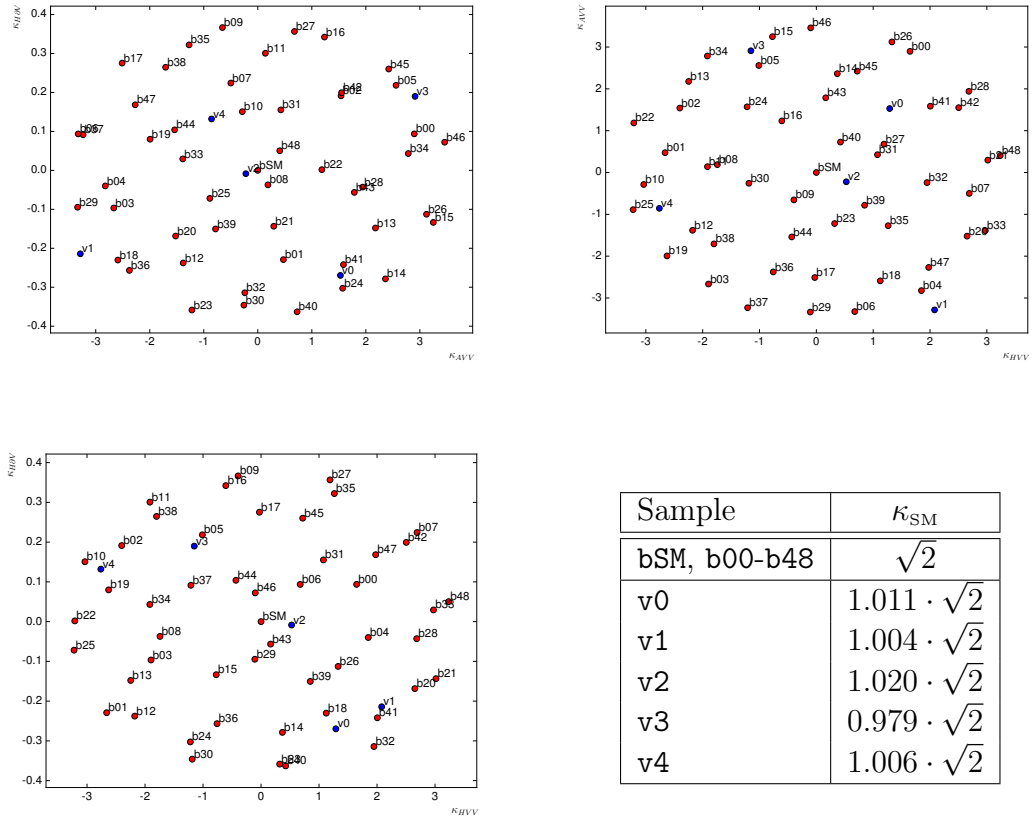
Consequently, the cross section predicted from the test morphing is always exactly the same as the cross section morphing would have predicted. Only numerical instabilities can lead to deviations from the predictions of the cross section morphing, when the morphing matrix of the currently tested basis is near a singularity. This instability is penalized by the bias term in the figure of merit. In the same way the cross section uncertainties are transitive if all correlations between the interluded morphing functions are taken into account. But in the optimization algorithm it would be unreasonable to assume correlated inputs of the test basis and because these input cross section uncertainties are taken to be uncorrelated the transitivity does not apply.

The performance of the iterative optimization algorithm is shown for the process VBF $H{\to}W^{\pm}W^{\mp *}{\to}\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ with four independent variables, κ_{SM}, κ_{HVV}, κ_{AVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ executed at the 2-jet requirement in the selection presented in Section 6.5 using generator level MC samples each produced with a statistic of 100 000 events. In this configuration usually 35 inputs are necessary for a morphing basis, but in this exemplary study the interference terms between κ_{AVV} and the other three couplings are neglected as described in Section 5.5.3 which leads to only 22 inputs.

First, the parameter space of interest is defined with $|\kappa_{\mathrm{HVV}}| < 2.28$, $|\kappa_{\mathrm{AVV}}| < 2.40$ and $|\kappa_{\mathrm{AVV}}| < 0.258$. At these individual BSM parameter values the cross sections stays approximately below the 25% increased SM cross section with fixed parameters $\kappa_{\mathrm{SM}} = \sqrt{2}$, $\cos\alpha = 1/\sqrt{2}$ and $\Lambda = 1\,\mathrm{TeV}$.

Fifty benchmarks and five additional validation samples are generated at the parameters shown in Figure 5.34. The benchmarks parameters are equally distributed inside the parameter space of interest. The validation samples are not

used in the optimization algorithm and utilized only to validate at independent parameters inside the parameter region of interest with varying κ_{SM} values that the uncertainties are reasonable small.



Sample	κ_{SM}
bSM, b00-b48	$\sqrt{2}$
v0	$1.011 \cdot \sqrt{2}$
v1	$1.004 \cdot \sqrt{2}$
v2	$1.020 \cdot \sqrt{2}$
v3	$0.979 \cdot \sqrt{2}$
v4	$1.006 \cdot \sqrt{2}$

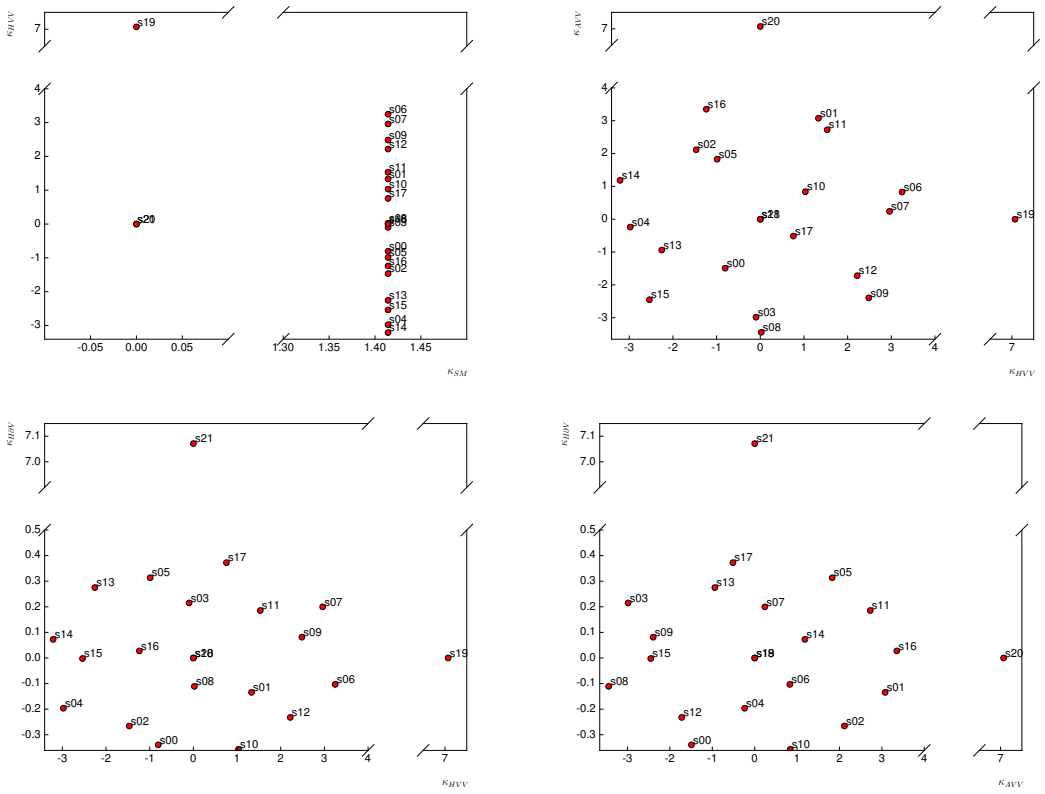
Figure 5.34.: Fifty benchmark bSM, b00-b48 (red) and five validation v0-v4 (blue) parameter configuration shown in the parameter space used in the morphing basis optimization algorithm.

A first set of input samples for the *cross section morphing* is generated at the parameters shown in Figure 5.35a. Here, samples s00-s18 are equally distributed in the parameter space of interest and samples s19-s21 are set at larger BSM values with $\kappa_{\mathrm{SM}} = 0$ to pre-optimize the morphing basis as previous experience has shown that additional pure BSM samples help to reduce the uncertainties. Still, as Figure 5.36a exposes for the validations v0-v3 for the $\Delta\phi_{jj}$ observable the uncertainties are very large.

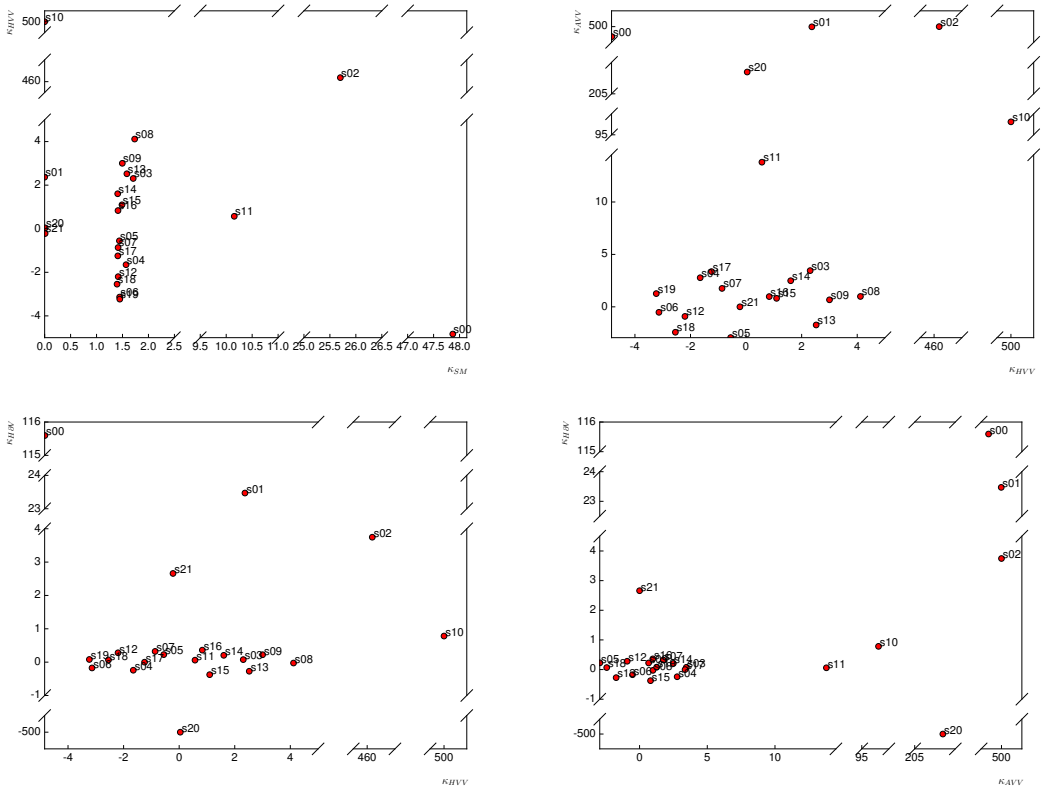
In the meta-minimization this first set of input samples provides the cross sections and uncertainties for the test basis. All $n \cdot N = 4 \cdot 22 = 88$ free parameters of the morphing basis are optimized with MINUIT to minimize the figure of merit.

The result is shown in Figure 5.35b. Most of the parameter points remain close to the region of interest but compared to the initial morphing basis more spread out with more outliners far away from the central region. New samples are generated at these parameters and utilized as input for the cross section morphing in the second iteration. Already in the first iteration a significant improvement is observed comparing Figure 5.36a to 5.36b.

In total four iterations are conducted. The development of the optimized basis parameters are shown in the Figures 5.35a to 5.35e and the improvements of the uncertainty for the validation $\Delta\phi_{jj}$ distributions `v0-v3` in Figures 5.36a to 5.36e.

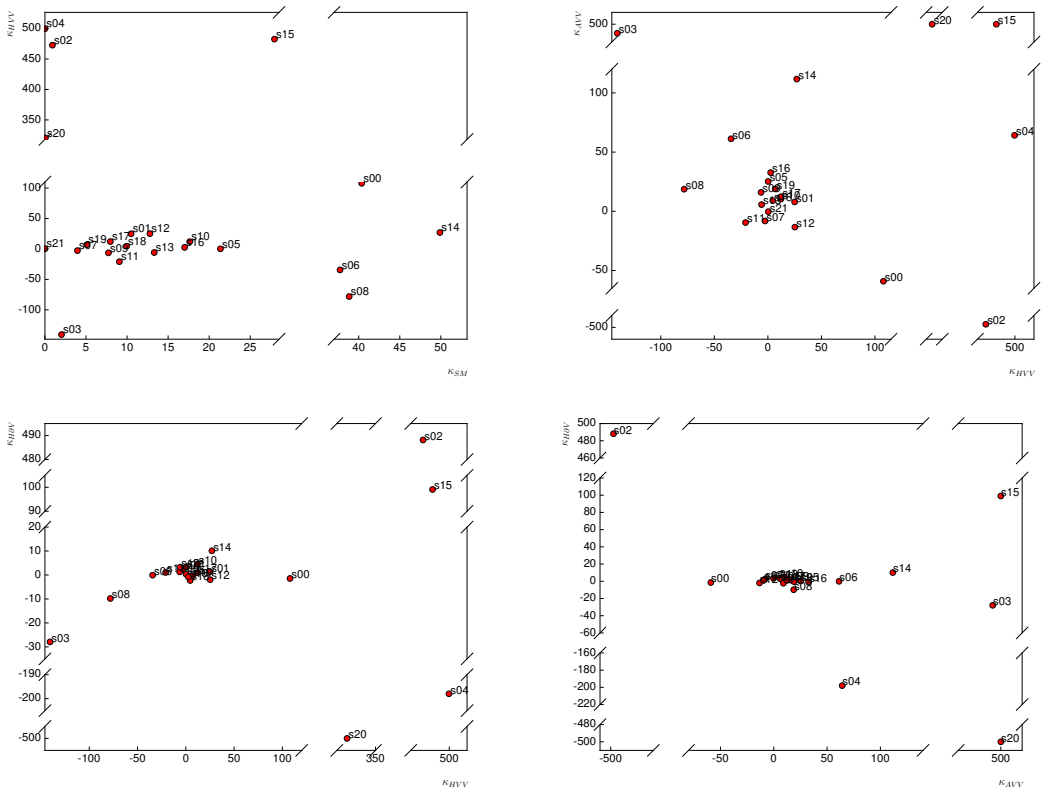


(a) Initial configuration of the morphing basis.

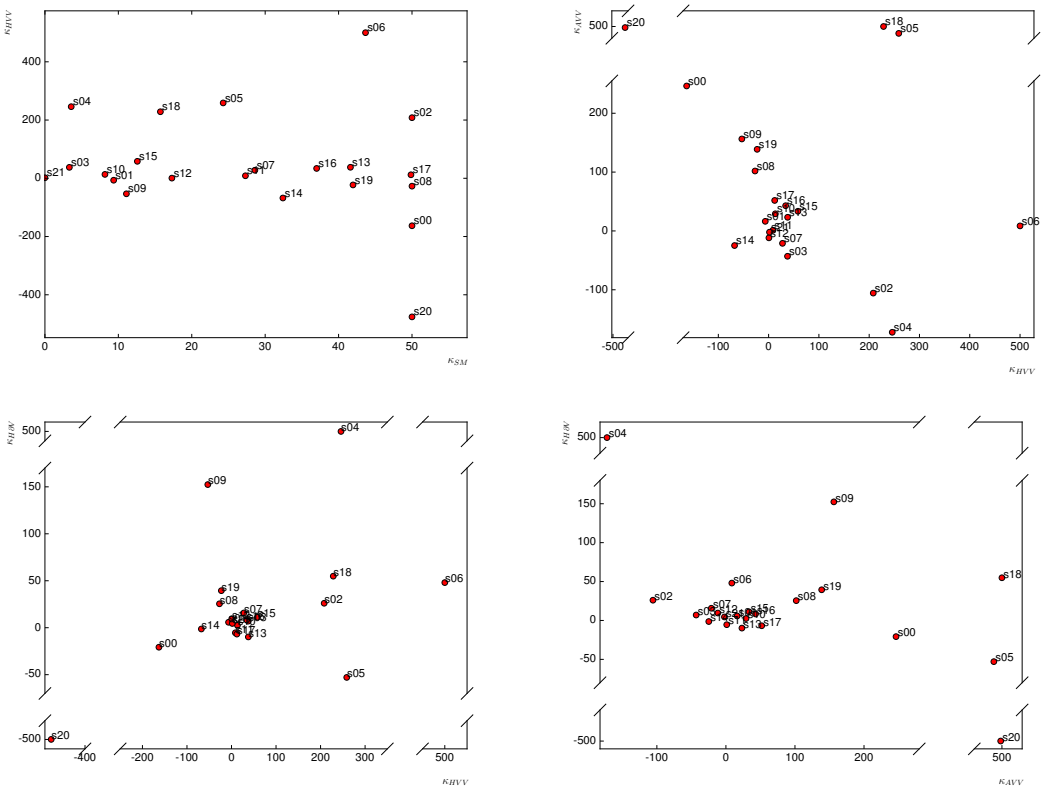


(b) Optimized morphing basis in the first iteration.

Figure 5.35.: Development of the morphing bases during the interaction of the morphing basis optimization algorithm.

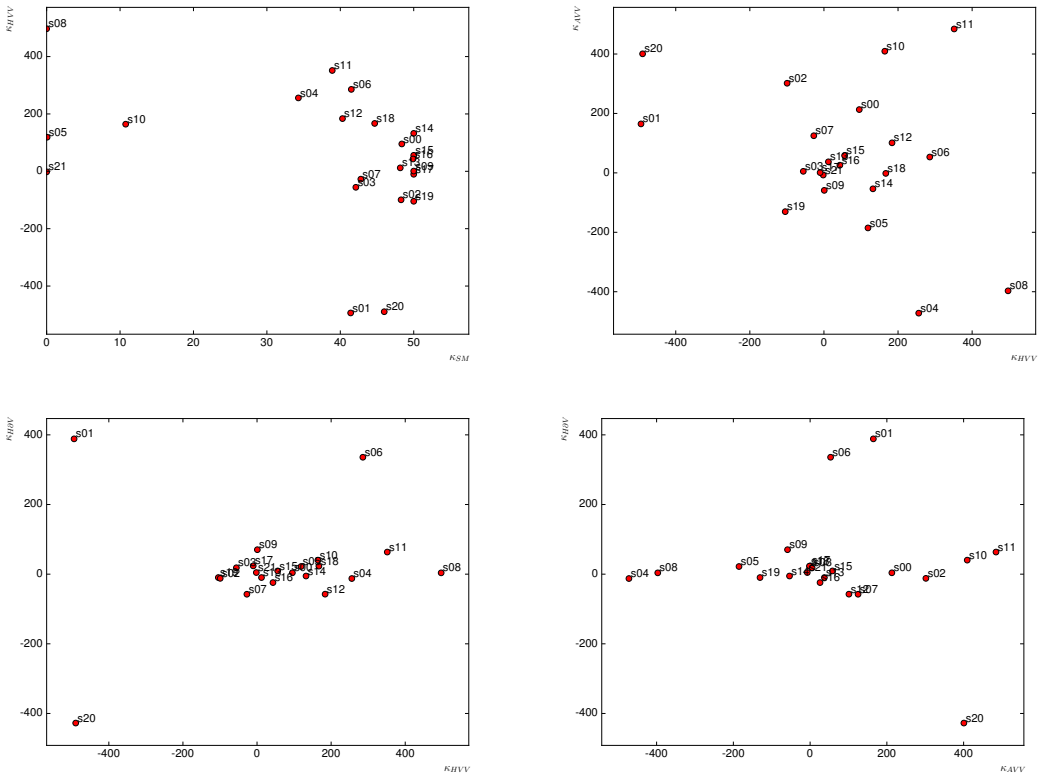


(c) Optimized morphing basis in the second iteration.



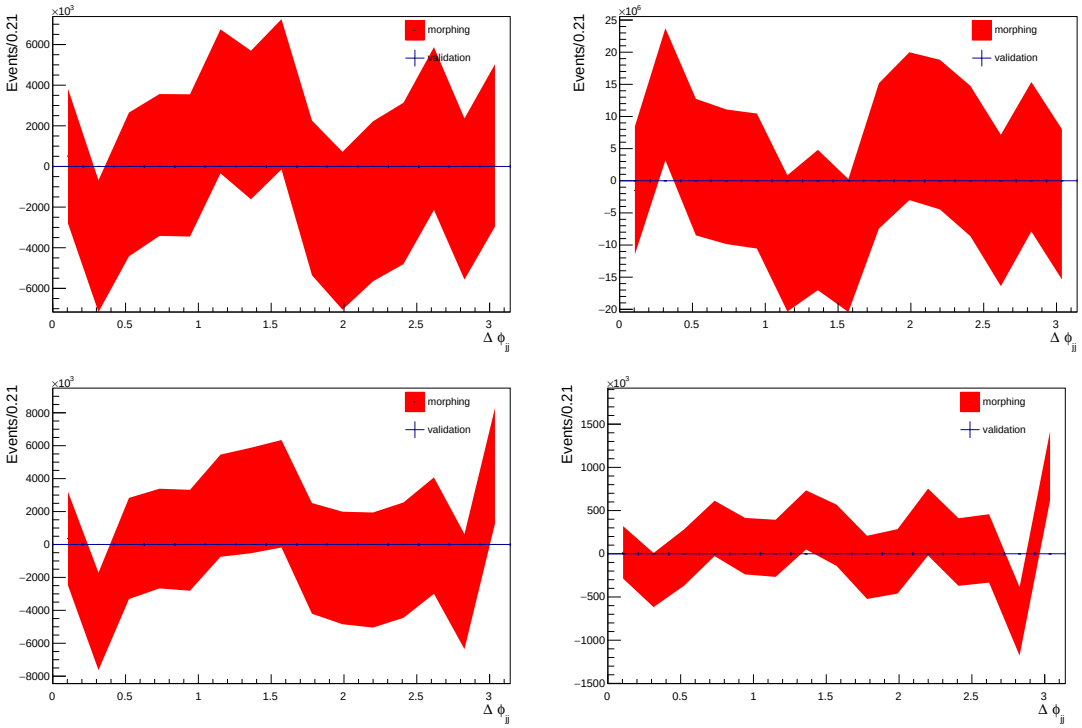
(d) Optimized morphing basis in the third iteration.

Figure 5.35.: Development of the morphing bases during the interaction of the morphing basis optimization algorithm. (cont.)

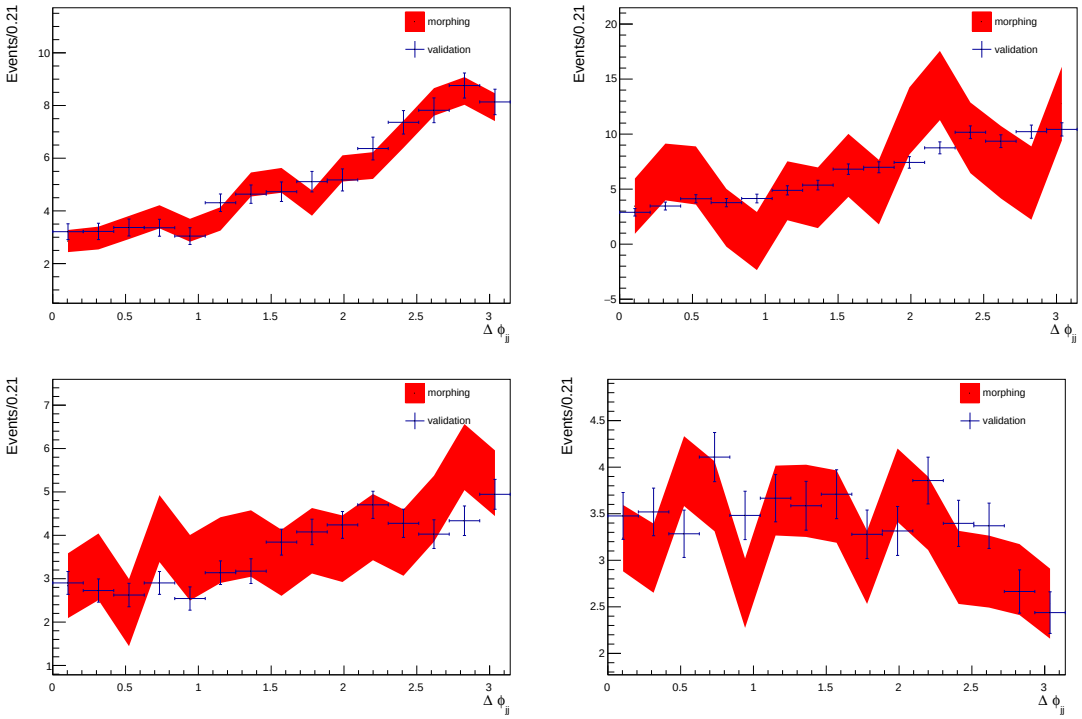


(e) Optimized morphing basis in the fourth iteration.

Figure 5.35.: Development of the morphing bases during the interaction of the morphing basis optimization algorithm. (cont.)

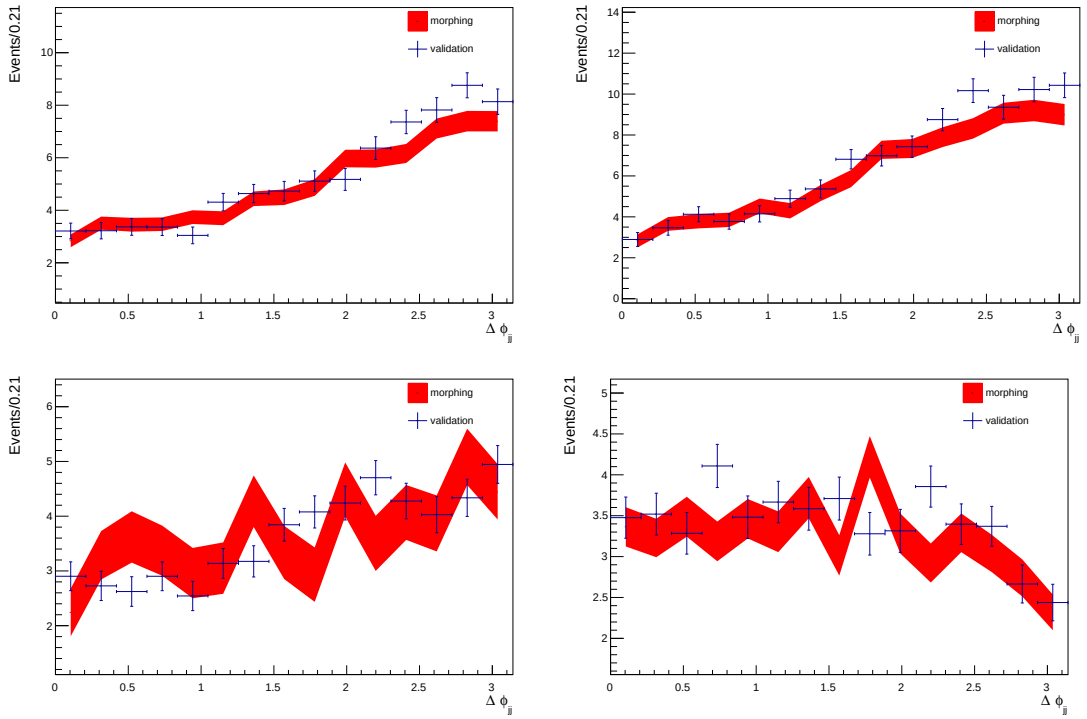


(a) Validations using the initial morphing basis shown in Figure 5.35a.

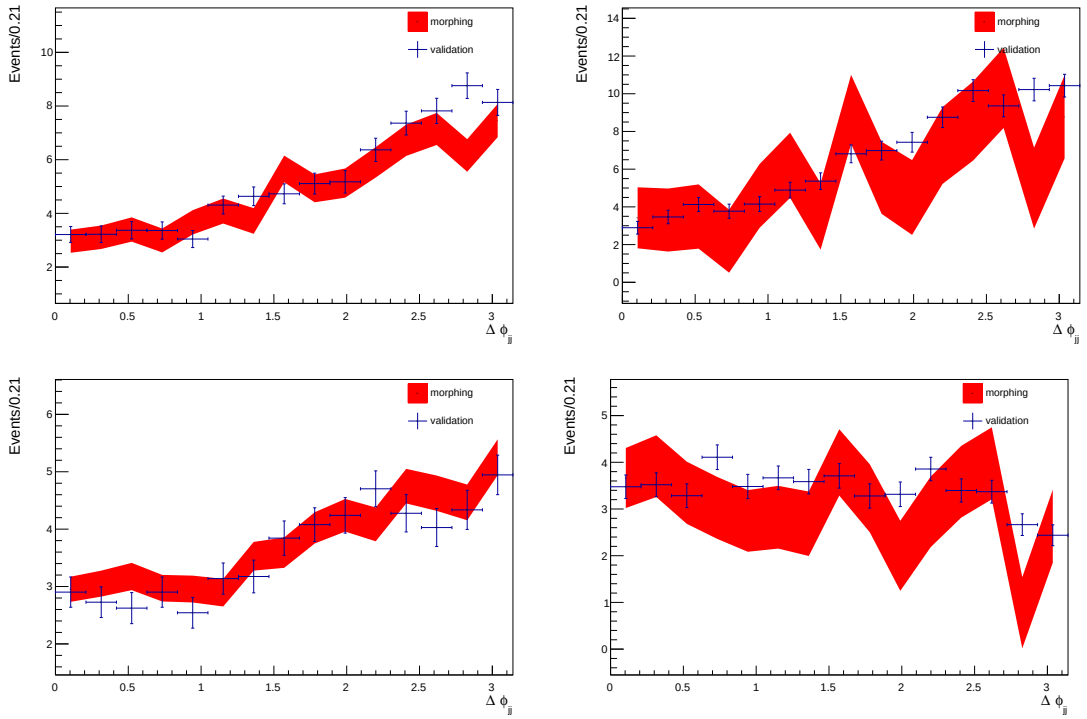


(b) Validations using the optimized morphing basis from the first iteration shown in Figure 5.35b.

Figure 5.36.: $\Delta\phi_{jj}$ distribution of the validation samples (blue) `v0` (top left), `v1` (top right), `v2` (bottom left) and `v3` (bottom right) compared to the prediction of the morphing (red) for the single iterations in the morphing basis optimization algorithm. The parameters of the validations are summarized in Figure 5.34.

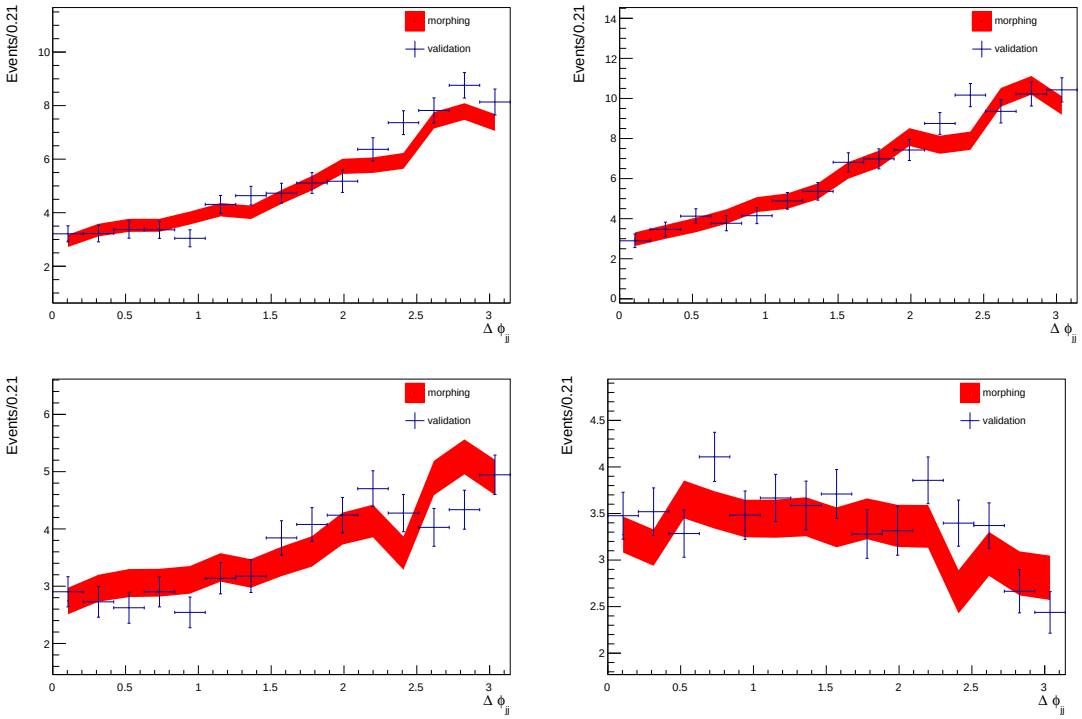


(c) Validations using the optimized morphing basis from the second iteration shown in Figure 5.35c.



(d) Validations using the optimized morphing basis from the third iteration shown in Figure 5.35d.

Figure 5.36.: $\Delta\phi_{jj}$ distribution of the validation samples (blue) `v0` (top left), `v1` (top right), `v2` (bottom left) and `v3` (bottom right) compared to the prediction of the morphing (red) for the single iterations in the morphing basis optimization algorithm. The parameters of the validations are summarized in Figure 5.34. (cont.)



(e) Validations using the optimized morphing basis from the fourth iteration shown in Figure 5.35e.

Figure 5.36.: $\Delta\phi_{jj}$ distribution of the validation samples (blue) `v0` (top left), `v1` (top right), `v2` (bottom left) and `v3` (bottom right) compared to the prediction of the morphing (red) for the single iterations in the morphing basis optimization algorithm. The parameters of the validations are summarized in Figure 5.34. (cont.)

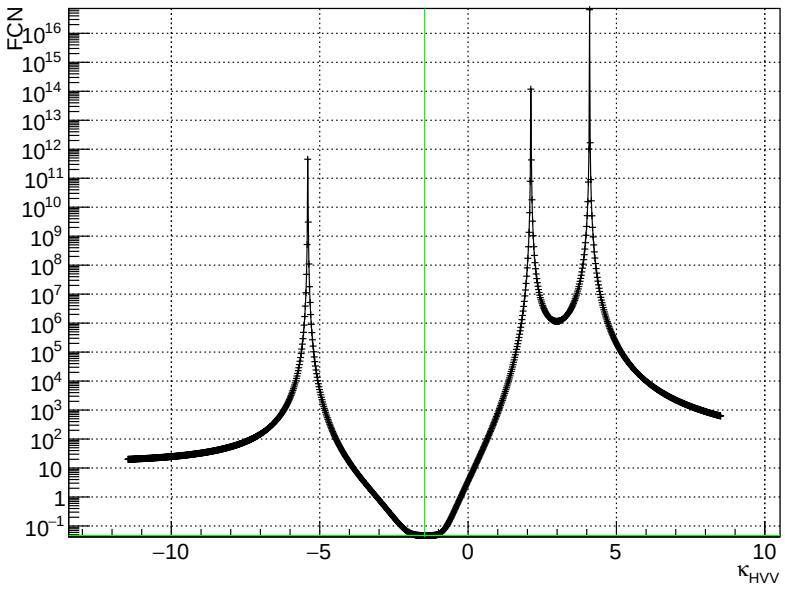


Figure 5.37.: Optimization score as a function of the parameter κ_{HVV} for one of the samples in the last interaction. The green lines indicate the minimum. The spikes correspond to singularities in the morphing matrix.

Figure 5.37 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, sharp spikes can be seen, which correspond to a non-invertible configuration of the morphing matrix, which is connected to high uncertainties on the result.

It is interesting to note though that, as can be seen clearly in Figure 5.35e, 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 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 a few iterations, which is a significant success.

6 Effective VBF $H \to W^{\pm} W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$ Coupling Analysis

After the discovery of a new boson at the LHC, studies of its properties to gain insights of the Higgs boson sector and the mechanism of electroweak symmetry breaking have become the first priority in high energy physics [7,97]. The published Run-1 results of the ATLAS and CMS experiments cover the main production and decay channels of the SM Higgs boson, along with the spin and CP properties of the new particle [98], precision measurements of its mass [24], and a recent combined analysis [27]. Overall, the results are consistent with the predictions of the SM, as shown in Figure 2.6. Additionally, searches for an extended Higgs sector were performed covering many BSM scenarios, like different supersymmetrical models or more generic models featuring charged and neutral Higgs bosons. But no significant evidence of any deviations from the SM was observed. Besides using such specific models a different approach to extend the SM Lagrangian is in a model independent way through an effective descriptions of sufficiently high-mass degrees of freedom referred to as effective field theory. In this analysis such an approach is attempted in the channel, where the Higgs boson is produced via VBF and decays into $H \to W^{\pm} W^{\mp *} \to \ell^{-} \bar{\nu}_{\ell} \ell'^{+} \nu_{\ell'}$ focusing in the anomalous coupling of the Higgs boson to vector bosons (cf. Figure 6.1).

First, in Section 6.1 the characteristics of the signal process in the EFT expansion of the SM are presented. The physics data used in this analysis and the MC samples are described in Section 6.2. Next, the object definitions are given in Section 6.3, followed in Section 6.4 by the definitions of observables used to discriminate signal and background processes or select topological features of the signal process. Thereafter, the signal region selection is shown in Section 6.5 with the Boosted Decision Tree application described in Section 6.6. The background estimation with its control regions and data-driven approaches is outlined in Section 6.7. The final selection of all categories are summarized in Section 6.8 and the listings of incorporated systematic uncertainties are presented in Section 6.9. Finally, the statistical treatment is described in Section 6.10 and the results of the measurements are reported in Section 6.11.

6.1 Characteristics of the Signal Process

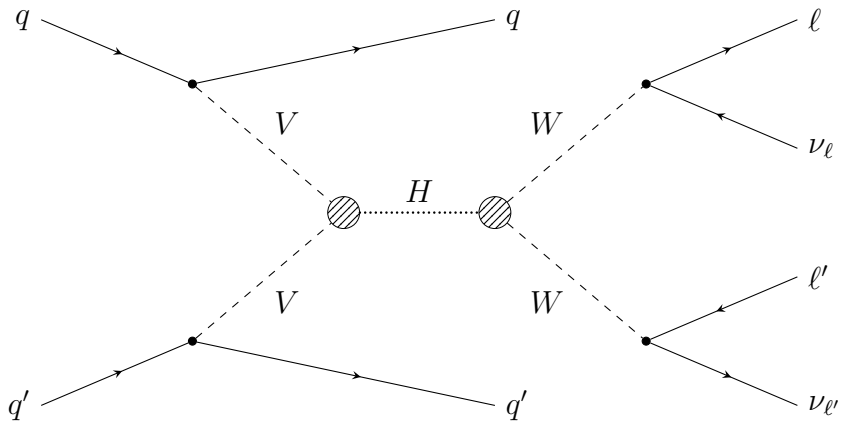


Figure 6.1.: Leading order Feynman diagram of a Higgs boson produced via VBF ($V = W$ or Z) generating two jets in the final state. The decay into two W-bosons which decay leptonically themselves leads to the special signature of two isolated, oppositely charged leptons and large missing transverse energy. The additional EFT couplings act on the VBF production vertex and $H \to WW$ decay depicted as large circles.

One of the main decay channels which contributed by large to the Higgs discovery, besides $H \to ZZ$ and $H \to \gamma\gamma$, is $H \to W^{\pm}W^{\mp *}$, which appears to be very useful for studying its properties. This channel has the second highest branching fraction (21.4%) and unique properties of the final state decay particles. Only the decay to two b-quarks has a higher branching fraction but is very difficult to analyze due to the large QCD background. For the same reason only leptonically decaying W-bosons are considered for the $H \to W^{\pm}W^{\mp *}$ decay channel, such that one has a good handle to both trigger events in the ATLAS detector and select them properly against the huge hadronic background. The total branching fraction for W-bosons to leptons is $\mathcal{B}_{W \to e/\mu/\tau} = 32.4\,\%$. However, while electrons are stable and muons exist at least sufficiently long enough to be measured by the detector directly, τ leptons generally decay further. Therefore, 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 branching fraction for the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay computes to 1.29%.

Having two leptons in the final state it is handy to sort them according to their transverse momenta p_{T}^{ℓ}, denoting the lepton with the highest or *leading* transverse momentum as ℓ_1 and the second highest or *subleading* transverse momentum ℓ_2. Another differentiation is provided by the lepton flavors, where the same flavor (SF) configuration ee and $\mu\mu$ is distinguished from the two different flavor (DF)

configurations $e\mu$ and μe. The difference between the latter two configurations is by p_T ordering of the leptons. As events with two same-flavor, opposite-charge events can easily arise from pair production via a Z- or γ-boson, the different flavor mode provides a much higher sensitivity and is thus focused on this analysis.

Furthermore, the spin correlation shown in Figure 6.2 introduces a unique feature in the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ channel, which leads to an enhanced fraction of events where the two charged leptons are emitted in the same hemisphere, and respectively the two neutrinos are emitted in the opposite one. These spin and kinematic properties can be exploited with 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 $E_{\mathrm{T}}^{\mathrm{miss}}$ to discriminate between signal and background events, as many SM processes do not exhibit the same characteristic.

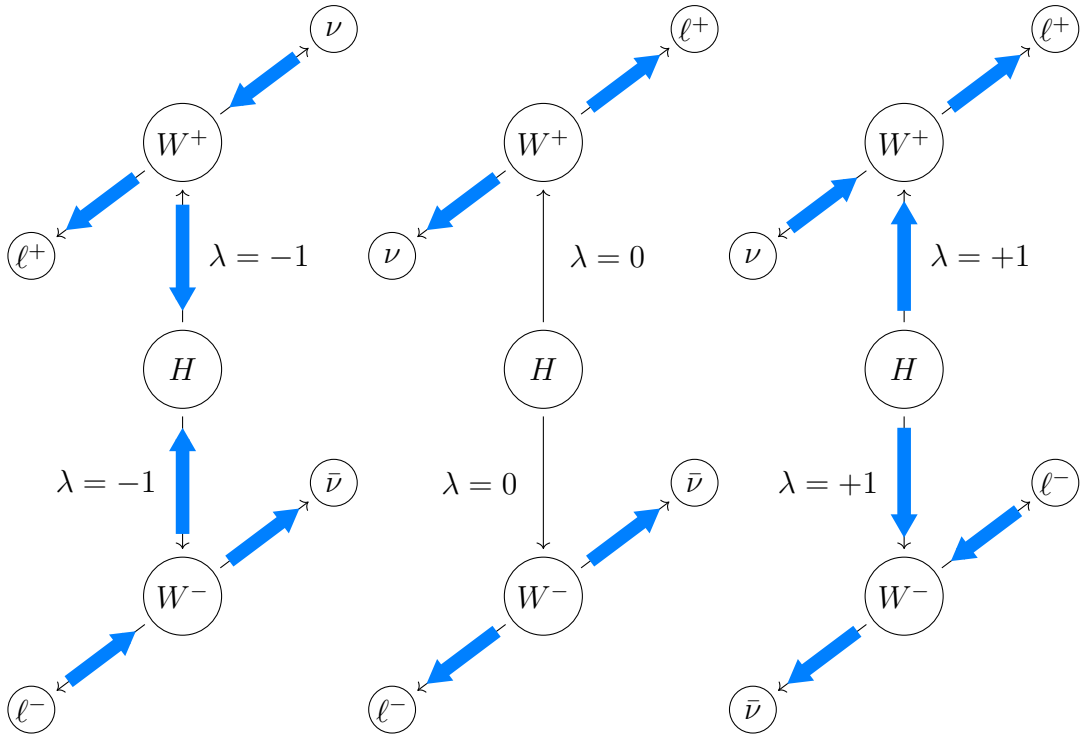


Figure 6.2.: Spin correlation in the $H{\to}W^{\pm}W^{\mp *}{\to}\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay increases the fraction of events where the two charged leptons are emitted into the same hemisphere and the neutrinos in the opposite one. The left (right) case shows a negative (positive) helicity λ, as spin ▲ and direction of flight ↑ are anti-parallel (parallel), resulting in the prescribed topology. Only the third (middle) case where the spin and momentum of the W-bosons are perpendicular is isotropic.

Investigating the Higgs production modes in terms of measuring effective couplings between the Higgs boson and vector bosons, the VBF mode is prominent due to several properties. First, it has after the gluon fusion the second highest production cross section which enhances the available statistics of signal events. Opposing to the ggF production, it introduces no mixing with fermion Higgs couplings and includes a second Higgs to vector boson vertex, enhancing the sensitivity of EFT contributions to the events. Additionally, the VBF features two energetic jets in the final state, the so-called *tagging* jets, with large separation in rapidity, and no color activity between them and the central system consisting of the Higgs boson and its decay products, due to the fact that the Higgs boson is created by fusion of a pair of electroweak bosons. They provide valuable information to the BSM couplings with observables such as the invariant dijet mass m_{jj}, the azimuthal angle between the jets $\Delta\phi_{jj}$, the pseudorapidity between the jets $\Delta\eta_{jj}$ or the transverse momenta of the jets p_{T}^{j}. Similar to the leptons, the jets are enumerated according to their transverse momenta. The jet with the highest or *leading* transverse momentum is denoted as j_1, whereas the second-highest or *subleading* jet is defined as j_2. A particularly convenient observable to distinguish diverse EFT contributions to the $H \to VV$ vertex is the azimuthal angle between the tagging jets, where the leading and subleading jets are sorted by pseudorapidity. It is defined in $[0, 2\pi]$ as

$$\Delta\phi_{jj,\mathrm{sign}} = \begin{cases} \Delta\phi_{jj} + 2\pi\,(\ \ \Delta\phi_{jj} < 0) & \text{for } \eta_{j_1} > \eta_{j_2} \\ -\Delta\phi_{jj} + 2\pi\,(-\Delta\phi_{jj} < 0) & \text{for } \eta_{j_1} \leq \eta_{j_2} \end{cases}, \tag{6.1.1}$$

where $\Delta\phi_{jj} = \phi_{j_1} - \phi_{j_2}$ and the azimuthal angle of the jets are defined in $\phi_{j_{1,2}} \in [-\pi, \pi]$.

Ideally suited to measure EFT contributions is the in MADGRAPH implemented Higgs Characterization model introduced in Section 2.3.4 consisting of an effective Lagrangian for the interaction of scalar and pseudo-scalar states with vector bosons. The most sensitive parameters contributing to the VBF production vertex are besides the κ_{SM} coupling, the CP-even κ_{HWW} and κ_{HZZ}, the CP-even derivative couplings $\kappa_{\mathrm{H}\partial\mathrm{WR}}$, $\kappa_{\mathrm{H}\partial\mathrm{WI}}$ and $\kappa_{\mathrm{H}\partial\mathrm{Z}}$, and the CP-odd couplings κ_{AWW} and κ_{AZZ}, which have been analyzed in details in Section 5.5.2 (cf. Table 5.8). To minimize the degrees of freedom in this analysis, the coupling of the Higgs boson to W- and Z-bosons are assumed to be the same. This results in identifying the coupling as

$$\begin{aligned} \kappa_{\mathrm{HVV}} &:= \kappa_{\mathrm{HWW}} = \kappa_{\mathrm{HZZ}} \\ \kappa_{\mathrm{H}\partial\mathrm{V}} &:= \kappa_{\mathrm{H}\partial\mathrm{WR}} = \kappa_{\mathrm{H}\partial\mathrm{WI}} = \kappa_{\mathrm{H}\partial\mathrm{Z}} \\ \kappa_{\mathrm{AVV}} &:= \kappa_{\mathrm{AWW}} = \kappa_{\mathrm{AZZ}}. \end{aligned} \tag{6.1.2}$$

The impact of these various EFT contributions have already been analyzed in [99]. A number of benchmarks with different parameter settings, summarized in Table 6.1

Table 6.1.: Benchmark scenarios used to illustrate the impact of different EFT parameters on typical VBF observables in [99].

Scenario	Parameters	
0^+ (SM)	$\kappa_{\mathrm{SM}} = 1$	$(\cos\alpha = 1,\ \Lambda = 1\,\mathrm{TeV})$
0^+ (HD)	$\kappa_{\mathrm{HVV}} = 1$	$(\cos\alpha = 1,\ \Lambda = 1\,\mathrm{TeV})$
0^+ (HDder)	$\kappa_{\mathrm{H}\partial\mathrm{V}} = 1$	$(\cos\alpha = 1,\ \Lambda = 1\,\mathrm{TeV})$
0^+ (SM+HD)	$\kappa_{\mathrm{SM,HVV}} = 1$	$(\cos\alpha = 1,\ \Lambda = 246\,\mathrm{GeV})$
0^- (HD)	$\kappa_{\mathrm{AVV}} = 1$	$(\cos\alpha = 0,\ \Lambda = 1\,\mathrm{TeV})$
$0^\pm$ (HD)	$\kappa_{\mathrm{HVV,AVV}} = 1$	$(\cos\alpha = 1/\sqrt{2},\ \Lambda = 1\,\mathrm{TeV})$

are probed for typical VBF production observables at generator level. The transverse momentum of the Higgs boson p_{T}^H and the azimuthal angle between the tagging jets $\Delta\phi_{jj}$ are shown in Figure 6.3, where jet requirements on the transverse momentum $p_{\mathrm{T}}^j < 25\,\mathrm{GeV}$ and the pseudorapidity $|\eta_j| < 4.5$, as well as an additional VBF cut on the invariant dijet mass $m_{jj} > 500\,\mathrm{GeV}$ are applied. Two main characteristics are apparent. For all BSM benchmarks the events tend to have a higher transverse momentum of the Higgs boson and very different behavior in the $\Delta\phi_{jj}$ distribution depending which EFT parameter contributes to the events.

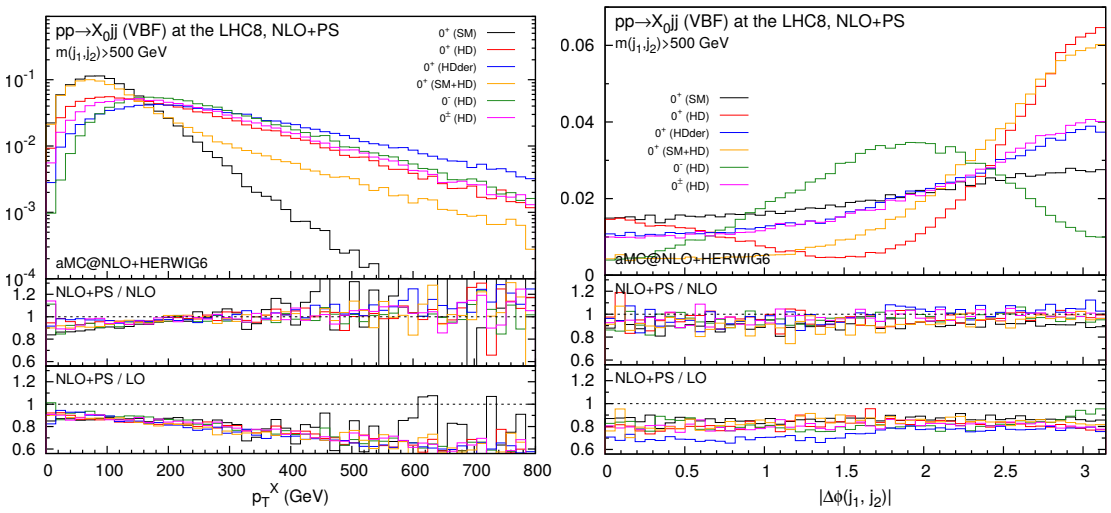


Figure 6.3.: Distribution of p_{T}^H (left) and $\Delta\phi_{jj}$ (right) on generator level with acceptance cuts for the jets and additional VBF selection for different BSM scenarios [99]. The parameter configurations for the samples are listed in Table 6.1.

Table 6.2.: Parameters used for the benchmarks to show the signal characteristics. Values are taken such that the BSM samples have approximately a 25% increased SM cross section without any acceptance cuts. For each samples the fixed parameters are $\kappa_{\mathrm{SM}} = \sqrt{2}$, $\cos\alpha = 1/\sqrt{2}$ and $\Lambda = 1\,\mathrm{TeV}$. All parameter that are not listed are set to 0.

Sample	Parameters
SM	$\kappa_{\mathrm{BSM}} = 0$
SM+HVV	$\kappa_{\mathrm{HVV}} = 2.28$
SM−HVV	$\kappa_{\mathrm{HVV}} = -2.28$
SM+AVV	$\kappa_{\mathrm{AVV}} = 2.40$
SM−AVV	$\kappa_{\mathrm{AVV}} = -2.40$
SM+HdV	$\kappa_{\mathrm{H}\partial\mathrm{V}} = 0.258$
SM−HdV	$\kappa_{\mathrm{H}\partial\mathrm{V}} = -0.258$

An even better choice for an observable to distinguish different EFT contributions to the $H \to VV$ vertices appears to be $\Delta\phi_{jj,\mathrm{sign}}$. To illustrate these variations, generator level MC samples have been generated with MADGRAPH5_aMC@NLO [84] including the decay to $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ using MADSPIN [100] at different parameter settings summarized in Table 6.2. The parameter choices are determined in such a way that the cross section of the SM+BSM samples have approximately an increased SM cross section by 25% without any acceptance cuts. The EFT samples include always the SM parameter with additionally one BSM constribution with positive or negative parameter value. The largest impact on the cross section in terms of numerical values of the parameters appears to have the derivative operator $\kappa_{\mathrm{H}\partial\mathrm{V}}$ since the parameter has a value about one order of magnitude smaller, whereas the κ_{HVV} and κ_{AVV} have roughly similar impact on the cross section enhancement.

The same jet requirements and m_{jj} cut as in [99] are applied with additionally conditions for the leptons. Only events with two different flavor, opposite sign leptons and transverse momentum for the leading lepton of $p_{\mathrm{T}}^{\ell_1} > 22\,\mathrm{GeV}$ and subleading lepton of $p_{\mathrm{T}}^{\ell_2} > 15\,\mathrm{GeV}$ are accepted. Figure 6.4 shows the $p_{T,jj}$ and $\Delta\phi_{jj,\mathrm{sign}}$ observables after applying these cuts. The $p_{T,jj}$ distribution shows an enhanced event yield in the high-p_{T} region for most of the BSM benchmarks similar to the p_{T}^{H} distribution in Figure 6.3 visible in the overflow bin. The three other distributions show the impact of the three EFT parameters κ_{HVV}, κ_{AVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$. While in the SM case the distribution peaks marginally at π, the SM+HVV distribution has a much enlarged peak at π and for negative HVV contribution the distribution is inverted with a drop at π. The AVV parameter shows as well different behaviors for positive and negative contributions with a peak at $3\pi/2$ and $\pi/2$ respectively, such that one can distinguish these cases very well from contributions of the HVV operator. This feature indicates that the $\Delta\phi_{jj,\mathrm{sign}}$ observable appears

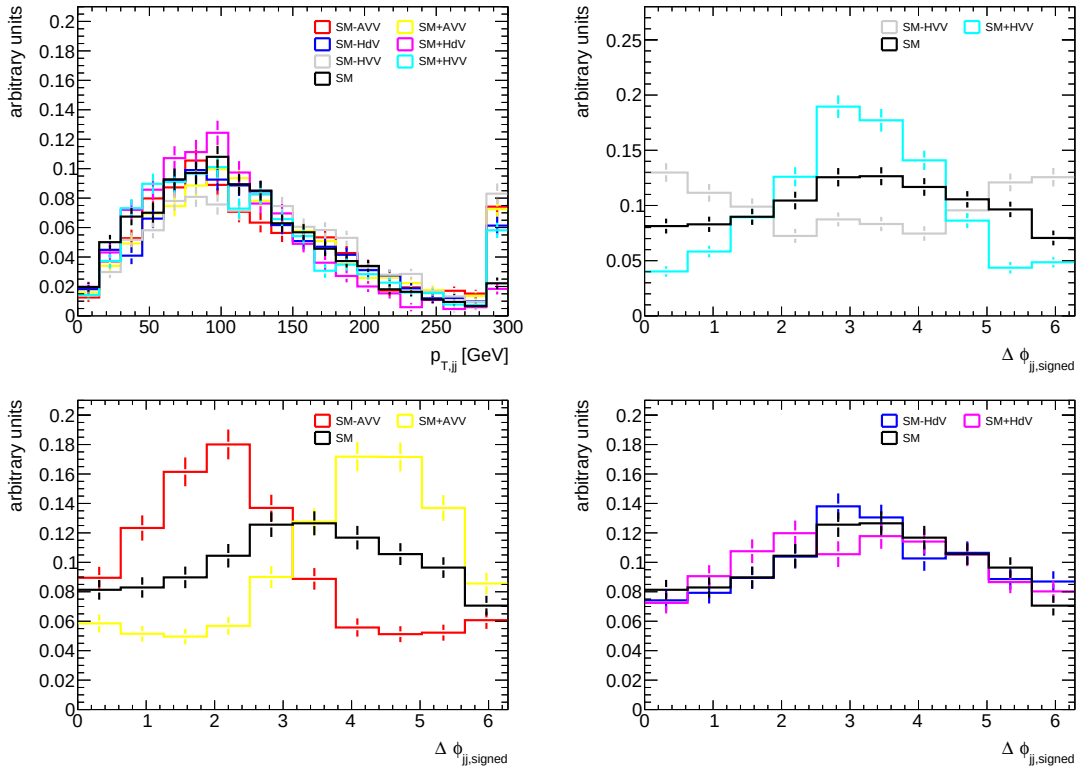


Figure 6.4.: $p_{T,jj}$ (top left) and $\Delta\phi_{jj,\text{sign}}$ (top right for κ_{HVV}, bottom left for κ_{AVV} and bottom right for $\kappa_{\text{H}\partial\text{V}}$) distributions for benchmarks at parameters listed in Table 6.2 showing characteristic features of EFT contributions to the $H \to VV$ vertex. The distributions are normalized to unity.

to be in additions sensitive to interference effects between CP-odd and CP-even terms as already shown in Figure 5.32b. The derivative operator H∂V has nearly no impact to the $\Delta\phi_{jj,\text{sign}}$ observable and the corresponding benchmark distributions are similar to the SM slope. This parameter is best to distinguish from the other EFT parameters by the cross section differences, where an interference effect is clearly visible between SM and H∂V. Investigating the $p_{T,jj}$ distribution shows, that a positive H∂V coupling has a lower event yield in the high-p_{T} region than for the pure SM case, whereas a negative coupling has a much larger event yield. These characteristics of the signal process are exploited in the analysis presented in the next sections.

Since the ggF production mode has an approximately ten times larger cross section than the VBF mode, it is worthwhile investigating the impact of the EFT contributions on the same observables. Dedicated samples at the same parameters as in the VBF case listed in Table 6.2 has been generated including two jets in NLO precision in the matrix element of the MadGraph generator and the same

acceptance cuts have been applied. Figure 6.5 shows the distributions comparable to the ones in Figure 6.4. No significant deviation can be observed either in the high-p_T region for $p_{T,jj}$, nor differences in the slopes of the $\Delta\phi_{jj,\text{sign}}$ distributions, where in all cases the distribution peaks at approximately π. Significant effects are anyway not expected because unlike for VBF the jets in ggF emerge only from initial state radiation and have not a direct connection to the $H \to VV$ vertex.

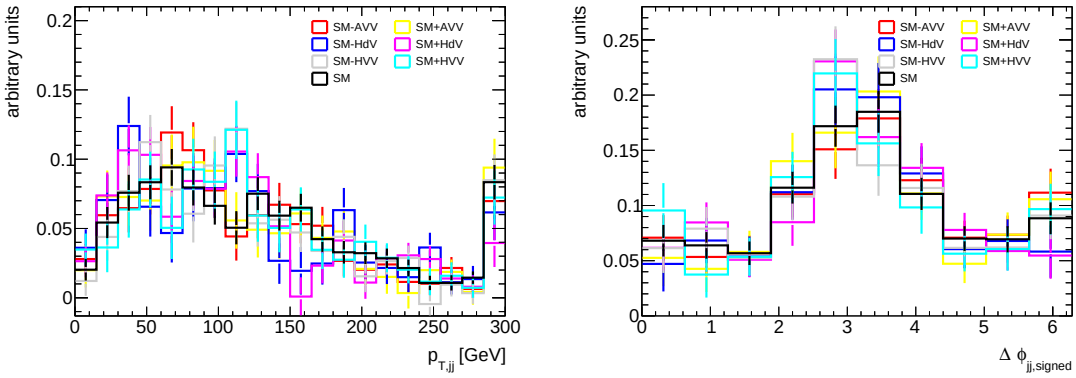


Figure 6.5.: The $p_{T,jj}$ (left) and $\Delta\phi_{jj,\text{sign}}$ (right) distributions for the benchmarks generated at parameters listed in Table 6.2 showing low impact of the EFT contributions to the $H \to VV$ vertex for the ggF production mode. The distributions are normalized to unity.

6.2 Data and MC Samples

Besides the VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ signal process also other signal processes including a $H \to VV$ vertex, and primarily many other background processes contribute to the final state. Most of these additional processes contribute by much larger event yields and therefore have to be simulated or included in a data-driven approach to be able to compare the simulation to real physics data.

6.2.1 Data

This analysis uses data recorded at $\sqrt{s} = 13\,\text{TeV}$ by the ATLAS detector in 2015 and 2016 with a bunch spacing of 25 ns corresponding to an integrated luminosity of $39.5\,\text{fb}^{-1}$ of pp collision data. To ensure that the collected data is ready for a reliable physics analysis, the calibration, the alignment and the general condition of all subdetectors have to be confirmed. Only if all relevant components have been working properly under nominal conditions, the events are considered. The ATLAS data quality [101] is monitored online and offline by a dedicated working group.

While the states of the different detector components is monitored online, the offline data quality monitoring provides quick feedback on the prompt reconstruction of the data. In that way any irregularities that undermine the quality and could lead to problems in the later analysis stages are uncovered. The data quality information is condensed into a so called *Good Run List* (GRL) which filters the problematic data blocks. After the data quality examination, the dataset collected by the ATLAS detector is reduced from $39.1\,\text{fb}^{-1}$ to the final amount of $(36.1 \pm 0.8)\,\text{fb}^{-1}$ available for physics analysis [102]. The uncertainty is derived, following a methodology similar to that detailed in Reference [102], from a preliminary calibration of the luminosity scale using a mini x-beam-separation scans performed in August 2015 and May 2016.

Events are triggered using dedicated single and di-lepton triggers, summarized in Table 6.3, together with the minimum transverse momentum requirements at different levels. In order to correct the MC simulation description of the data, per-event scale factors are applied, which are calculated with $\text{SF} = \epsilon_{\text{data}}/\epsilon_{\text{MC}}$ as

Table 6.3.: The minimum p_{T} requirements used at the different levels of the trigger, with values given in GeV. The values for the L1 trigger are for data. Letters 'T', 'M' and 'L' next to the threshold value stand for the `Tight`, `Medium` and `Loose` electron identification requirement, respectively. Letter 'i' indicates an isolation requirement which helps to increase the rejection power against non-prompt leptons. The single-lepton trigger with higher-p_{T} thresholds are more efficient at high lepton p_{T} than the lower-p_{T} triggers because of this isolation requirement. The di-lepton trigger `e17_lhloose_mu14` require $p_T^e > 17\,\text{GeV}$ with `LooseLH` identification and $p_T^\mu > 14\,\text{GeV}$. Due to increasing pile-up environment during the 2016 data taking, the lepton p_{T} thresholds has been increased starting with period D4.

Lepton	L1	HLT
Year 2015		
e	20	`24M` $\vee$ `60M` $\vee$ `120L`
μ	15	`20i` $\vee$ `50`
$e\mu$	e: 15, μ: 10	`e17_lhloose_mu14`
Year 2016: Periods A-D3		
e	20	`24Ti` $\vee$ `60M` $\vee$ `120L`
μ	15	`24i` $\vee$ `50`
$e\mu$	e: 15, μ: 10	`e17_lhloose_mu14`
Year 2016: Periods D4-L		
e	20	`26Ti` $\vee$ `60M` $\vee$ `120L`
μ	15	`26i` $\vee$ `50`
$e\mu$	e: 15, μ: 10	`e17_lhloose_mu14`

function of η and p_{T} of the leptons. Here, ϵ_{data} is the efficiency obtained on real data and ϵ_{MC} on MC, respectively on $Z \to \mu\mu$ and $Z \to ee$ data and simulated events. In both cases, muon and electron trigger efficiencies are obtained by means of the Tag-and-Probe method. The single-lepton trigger efficiencies on the plateaus are approximately 70% for muons with $|\eta| < 1.05$, 90% for muons in the range $1.05 < |\eta| < 2.40$, and $\geq 90\%$ for electrons in the range $|\eta| < 2.40$. The trigger efficiencies for the single lepton triggers and the combination of the single and the di-lepton trigger is shown in Figure 6.6 as function of the lepton transverse momentum. Since the di-lepton triggers are able to have lower p_{T}-thresholds the lepton p_{T} cuts in the pre-selection can have lower values at high lepton identification efficiencies. Inclusion of `e17_lhloose_mu14` trigger increases the gain in the $e\mu$ channel by 8% and 3.8% in the μe channel.

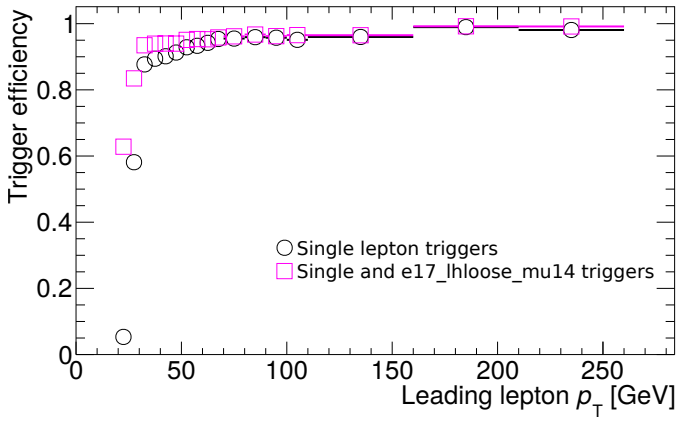
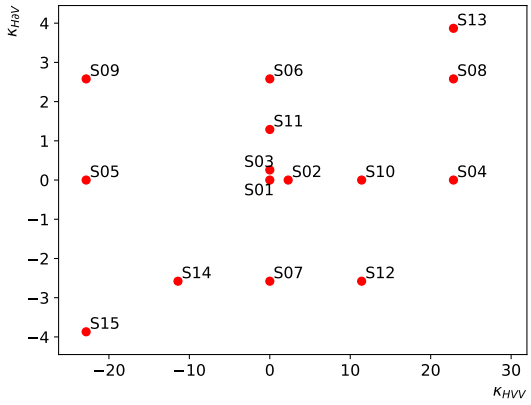


Figure 6.6.: Trigger efficiencies for the single lepton triggers (○) and single lepton triggers combined with the dilepton trigger `e17_lhloose_mu14` (□) after pre-selection, requiring at least two jets and a leading lepton transverse momentum of $p_{\mathrm{T}}^{\ell_1} > 22\,\mathrm{GeV}$.

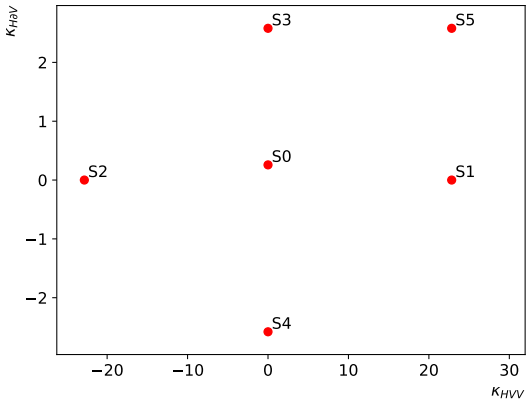
6.2.2 Signal and Higgs Processes

The signal contributions considered in this analysis include the vector boson fusion production process (VBF) and the dominant gluon fusion production process (ggF). All samples are generated with MADGRAPH5_aMC@NLO v2.3.3 in NLO precision in QCD with the Higgs Chracterization model. For ggF the heavy quark (top) loop is approximated by taking the mass of the heavy quark in the loop to infinity resulting in an effective $gg \to H$ vertex, which is for a not too heavy Higgs ($m_H = 125\,\mathrm{GeV}$) a good approximation. The $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay is executed by MADSPIN. Parton showering and hadronization are simulated with PYTHIA 8.212 with A14 tune [91] using the parton distribution functions from NNPDF3.0 NLO [103]



Sample	κ_{SM}
S01-02	0
S03-15	$\sqrt{2}$

(a) VBF parameter configurations for κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$.



Sample	κ_{SM}
S1	0
S2-6	$\sqrt{2}$

(b) ggF parameter configurations for κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$.

Sample	κ_{SM}	κ_{AVV}
S1	$\sqrt{2}$	0
S2	$\sqrt{2}$	12.01
S3	$\sqrt{2}$	−12.01
S4	$\sqrt{2}$	24.02
S5	$\sqrt{2}$	−24.02

(c) VBF parameter configurations for κ_{SM} and κ_{AVV}.

Sample	κ_{SM}	κ_{AVV}
S1	$\sqrt{2}$	0
S2	$\sqrt{2}$	24.02
S3	$\sqrt{2}$	−24.02

(d) ggF parameter configurations for κ_{SM} and κ_{AVV}.

Figure 6.7.: Parameter configurations for VBF and ggF used to generate samples as inputs for the Effective Lagrangian Morphing to scan over the parameter space for κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$, and over κ_{SM} and κ_{AVV}. Fixed parameter are $\Lambda = 1000\,\mathrm{GeV}$ and $\cos\alpha = 1/\sqrt{2}$ and for the ggF cases $\kappa_{\mathrm{Hgg}} = \sqrt{2}$.

with $\alpha_S = 0.018$. In case of the ggF production channel the merging of the jet multiplicities in NLO accuracy between the generator and the parton shower is performed by the `FxFx` merging scheme [84]. Generator level filters on the leptons are applied to increase the MC statistics in the final selection regions. At least one electron and one muon is required with the leading lepton to have a transverse momentum of $p_{\mathrm{T}}^{\ell_1} > 18\,\mathrm{GeV}$ and the subleading lepton an transverse momentum of $p_{\mathrm{T}}^{\ell_1} > 8\,\mathrm{GeV}$, both within a pseudorapidity of $|\eta| < 2.8$.

Samples for different EFT scenarios are generated to be able to scan over the parameter space with the Effective Lagrangian Morphing for ggF and VBF. Two

seperate morphing functions are constructed with the CP-even parameters κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$, and the CP-odd parameter κ_{AVV} together with κ_{SM}. The number of samples needed, calculated with Equation 5.3.22, for VBF (ggF) scanning over the three CP-even parameters are 15 (6), and for κ_{SM} and κ_{AVV} are 5 (3), respectively. All parameter configurations for the morphing inputs are shown in Figure 6.7.

Additional VBF samples are generated to have a set of inputs for the signal selection optimization with multivariate techniques (cf. Section 6.6). The parameter configurations are listed in Table 6.4 together with the corresponding cross sections calculated by MADGRAPH. The SM VBF cross section has already been calulated to NNLO QCD and NLO EW corrections and for ggF in NNNLO in QCD with additional NLO EW corrections being applied. The numerical values and uncertainties are summarized in Table 6.5. Using the branching fraction of 2.27% for the $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay the cross sections calculated by MADGRAPH are corrected by applying the appropriate k-factor $\sigma_{\mathrm{NLO,MadGraph}}/\sigma_{\mathrm{N^xLO}}$ of 1.059 for VBF and 1.439 for ggF. Since no cross sections with higher order corrections for any EFT scenarios are available, these k-factors are applied to all VBF and ggF samples, respectively.

Contributions from the Higgs-strahlung process from a vector boson $V = W$ or Z (VH) are included in the analysis as a background due to the small contribution and are simulated only at SM parameters. Other Higgs boson production processes like ttH and bbH are neglected because their contributions are small, whereas a small contamination of $H \to \tau\tau$ decays has been included as well as a background. For all samples including the Higgs boson only the $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay

Table 6.4.: Parameter configuration of the benchmark EFT VBF samples. Fixed parameters are $\kappa_{\mathrm{SM}} = \sqrt{2}$, $\Lambda = 1000\,\mathrm{GeV}$ and $\cos\alpha = 1/\sqrt{2}$.

Sample	κ_{HVV}	$\kappa_{\mathrm{H}\partial\mathrm{V}}$	κ_{AVV}	$\sigma \cdot \mathrm{Br}$ [pb]	Filter efficency
SM	0	0	0	0.0810	0.194
SM+HVV	2.28	0	0	0.1161	0.198
SM−HVV	−2.28	0	0	0.0718	0.203
SM+H∂V	0	0.258	0	0.0762	0.196
SM−H∂V	0	−0.258	0	0.0906	0.196
SM+AVV	0	0	2.40	0.0913	0.201
SM−AVV	0	0	−2.40	0.0912	0.203
SM+HVV+H∂V	2.28	0.258	0	0.1069	0.197
SM+HVV+AVV	2.28	0	2.40	0.1271	0.199
SM+AVV+H∂V	0	0.258	2.40	0.0858	0.203
SM+HVV+AVV+H∂V	2.28	0.258	2.40	0.1193	0.199
SM−HVV−AVV−H∂V	−2.28	−0.258	−2.40	0.0859	0.206

mode is considered, with final states featuring two charged leptons. The cross sections of the associated VH production processes are calculated up to NNLO QCD corrections [104] and NLO EW corrections [105], while the cross section for $gg \to ZH$ is calculated at NLO QCD precision, listed in Table 6.5.

The VH MC samples are fully generated with POWHEG-BOX V2 using MiNLO [106], including showering, modelling of multiple parton interactions and hadronization with one jet in NLO in the matrix element calculation. The VH predictions are normalized to the corresponding cross sections in Table 6.5. The PDF set used is PDF4LHC15 NLO [107].

Table 6.5.: Predicted signal process cross sections. The uncertainties quoted here correspond to standard Gaussian 68% confidence levels. Compared to the total uncertainty of this analysis, the signal cross section uncertainties are negligible and thus not applied.

Process	cross section [pb]
ggF	$48.58 \pm 3.9\%$ (QCD scale) $\pm 3.2\%$ (PDF + α_S)
VBF	$3.782 \pm 0.2\%$ (QCD scale) $\pm 2.1\%$ (PDF + α_S)
WH	$1.373 \pm 0.4\%$ (QCD scale) $\pm 1.9\%$ (PDF + α_S)
ZH	$0.8839 \pm 2.2\%$ (QCD scale) $\pm 1.6\%$ (PDF + α_S)

6.2.3 Background

Although the signature of the $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ final state is quite rare, it is unfortunately not unique. There are various background processes that either share the same final state particles or mimic them due to misidentification, additional pile-up remnants or the limited detector coverage. Since the latter of those typically show obvious differences, they are called reducible backgrounds. However, if the final state is exactly the same as the signal, refined analysis techniques need to be developed. This kind of background processes is called irreducible.

The main sources of backgrounds 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. These backgrounds have large cross sections and have to be removed by selecting signal dominated phase space. They are presented in the following sections.

The MC generators used to model the background processes are listed in Table 6.6. For most processes, separate programs are used to generate the hard scattering process and to model the parton showering, hadronisation and the underlying event. PYTHIA 8.210, PYTHIA 8.186 [108] or PYTHIA 6.428 [109] are used for the latter three steps for the signal and some of the background processes.

The CT10 and NNPDF3.0 PDF set [44] are used for the hard scattering process in POWHEG-BOX V2. The AZNLO [110] tune is used for the diboson processes while the A14 tune [91] for other processes.

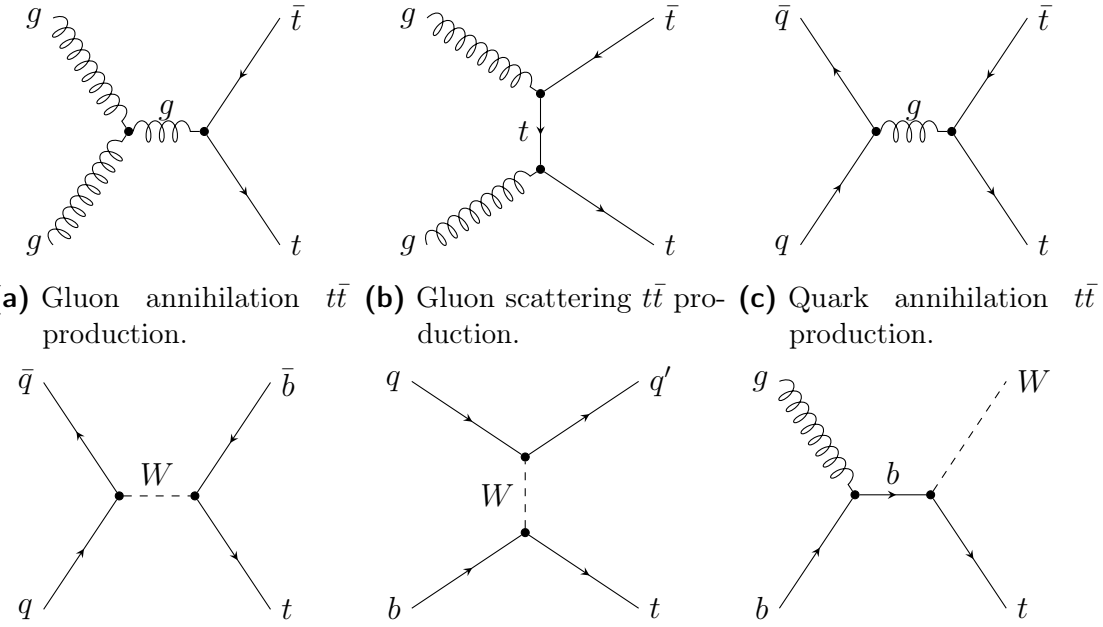
The hard scattering NLO predictions from SHERPA 2.2.1 [111] are calculated using NNPDF3.0 NNLO PDF set in conjunction with a dedicated set of tuned parameters from the parton shower developed by the Sherpa authors [112].

Production of $t\bar{t}$ and Single Top

For events with at least two jets the largest background contribution comes from the production of a top-quark pair ($t\bar{t}$). In pp collisions at the LHC, top-antitop pairs are produced via gluon-gluon fusion, and gluon-gluon and quark-antiquark annihilation shown in Figure 6.8 **(a)-(c)**. Mostly the gluon induced production dominates with its 90% contribution to the total $t\bar{t}$ cross section. Single top-quarks are produced via three different mechanisms. Either in the decay of a virtual W-boson (s-channel), in the exchange of a W-boson (t-channel) or in association with a W-boson (Wt). The Feynman diagrams for those modes are shown in Figure 6.8 **(d)-(f)**. Since the top-quarks almost exclusively decay into a bottom-quark and a W-boson, the top-related backgrounds have the same signature as the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ signal process, when the W-bosons decay leptonically. While the $t\bar{t}$ and Wt processes lead to signatures containing a $W^{+}W^{-}$-pair plus additional jets, the other single top channels can mimic the signal only due to mismeasurements. In any case, the identification of the top-quark background relies on the associated bottom-quarks. By determining events that contain b-tagged jets, the contribution can be reduced significantly. Therefore, a dedicated control region is used to estimate and normalize the top background described in Section 6.7.1.

The top-quark pair production is simulated using POWHEG-BOX V2 using the NNPDF3.0 NLO PDFs with NLO matrix element precision and assuming a top-quark mass of 172.5 GeV. The generated events are then interfaced to PYTHIA 8 using NNPDF2.3 LO PDFs for parton showering with the A14 tune [91] and a matching paramter of $h_{\text{damp}} = 1.5m_t$. The $t\bar{t}$ samples generated include a filter to require that the W-bosons decays leptonically. All the three charged leptons are considered for the W-decay in POWHEG. The τ leptons are then decayed by TAUOLA [113] in either leptonically or hadronically mode. The $t\bar{t}$ production cross section is normalised to the predictions calculated with the TOP++2.0 program to NNLO in perturbative QCD, including soft-gluon resummation in NNLL order [114].

The associated production of a single top-quark and a W-boson (Wt) is generated with POWHEG-BOX V1 interfaced to PYTHIA 6.428 for parton showering, using PERUGIA2012 tune [115] and the CT10 [44] PDF set. EVTGEN 1.2.0 [116] is used for properties of the bottom and charm hadron decays. The Wt sample is required to have at least two chraged leptons in the final state. Overlaps with



(a) Gluon annihilation $t\bar{t}$ production.

(b) Gluon scattering $t\bar{t}$ production.

(c) Quark annihilation $t\bar{t}$ production.

(d) Single top-quark production in the s-channel.

(e) Single top-quark production in the t-channel.

(f) Single top-quark production associated with a W-boson.

Figure 6.8.: Leading order Feynman diagram top-antitop pair production and single-top production.

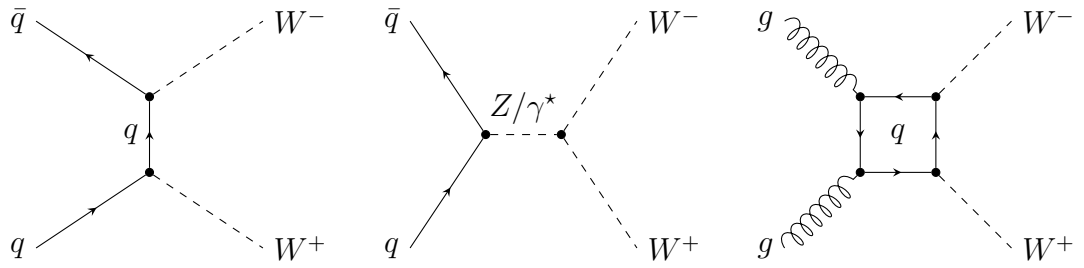
the top-quark pair production process that occur at NLO calculations in QCD are removed using the diagram removal scheme.

WW Production

The dominant background contribution to $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ with no additional jets arises from the WW production, where both W-bosons decay leptonically. This results in the same final state as the signal process and is therefore an irreducable background. Figure 6.9 shows the leading order Feynman diagrams of the dominant production modes. At the LHC the leading mode is the t-channel quark-antiquark annihilation. The s-channel contribution involves the triple gauge coupling and makes up approximately 10% of the total production, while the non-resonant $gg \to WW$ production accounts for only ca. 3%.

The $q\bar{q}/g \to WW$ MC sample is generated using SHERPA 2.2.2 interfaced with NNPDF3.0 NNLO PDFs having 0 and 1 jet at NLO accuracy and 2 and 3 jet at LO accuracy with lepton requirements $m_{\ell\ell} > 4\,\text{GeV}$ and $p_{\text{T}}^{\ell_{1,2}} > 5\,\text{GeV}$. The $gg \to WW$ process is simulated by SHERPA 2.1.1 with zero or one additional jets.

These samples are normalized to the cross section of the inclusive WW background process which is known at NNLO accuracy [117] and includes contributions



(a) Quark initiated t-channel production.

(b) Quark initiated s-channel production.

(c) Gluon fusion via quark loop.

Figure 6.9.: Leading order Feynman diagram of the WW background production.

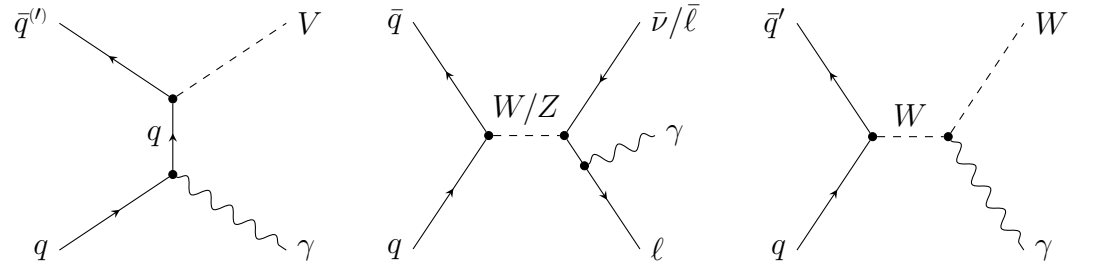
from $q\bar{q}/g \to WW$ and $gg \to WW$. Since the processes $q\bar{q}/g \to WW$ and gluon fusion WW production are modelled with different MC samples and a higher order calculation is available for $gg \to WW$, cross sections for both processes are handled separately. In the inclusive WW prediction, $gg \to WW$ is calculated to LO precision and does therefore 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 $gg \to WW$ production cross section has been calculated at NLO accuracy [118] to $\sigma_{gg \to WW} = 6.82\,\text{pb}^{+5.3\%}_{-7.4\%}$.

SHERPA 2.1.1 is used for the modelling of diboson process at LO with no $O(\alpha_S)$ terms including vector boson scattering for the $\ell\nu\ell\nu$ plus two jets in the final states.

$WZ/ZZ/V\gamma^\star$ Production

A related set of background processes in this context is called diboson production, referring to WZ, ZZ and $V\gamma^\star$ processes. The WW production of the previous section is explicitly excluded from this set since it stands out as one dominant background process, despite it actually belongs to the same category. Although these diboson production processes contain isolated leptons and missing transverse momentum emerging from leptonic W-decays, they differ from the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ signal by e.g. not having exactly two reconstructed charged leptons. A simple veto on a third or more charged leptons already reduces their contribution significantly. Therefore, these processes are expected to be small and are determined purely by MC simulation.

Several ZZ and WZ samples are generated with POWHEG-BOX v2. Depending on the final state, different requirements are set on the invariant dilepton masses for any same-flavor opposite-charge lepton pairs and for the invariant di-quark masses at either $m_{\ell\ell} > 4\,\text{GeV}$ or $m_{qq} > 20\,\text{GeV}$ and $m_{\ell\ell} > 20\,\text{GeV}$ for one electroweak boson decaying hadronically. The WZ sample with both electroweak bosons decaying leptonically (Z-boson decaying to two charged leptons) is generated with



(a) Quark initiated t-channel production.

(b) Bremsstrahlung.

(c) Quark initiated s-channel production (trilinear coupling).

Figure 6.10.: Leading order Feynman diagram of the $V\gamma$ background production.

SHERPA 2.1.1, using the CT10 PDFs, at NLO accurary for 0 and 1 jet and LO for 2 and 3 jets. These samples include also the $\gamma^\star$ process and are produced with at least two charged leptons with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ and additionally imposing $m_{\ell\ell} > 2 \times m_\ell + 250\,\mathrm{MeV}$ for any same-flavor, opposite-charge lepton pairs, where m_ℓ is the mass of the charged lepton.

$V\gamma$ Production

A class of reducible background processes are the production of a W- or Z-vector boson with a real emmited photon in the final state labeled as $V\gamma$. A misidentification e.g. of the photon as electron leads to the same final state as the signal process. Leading order Feynman diagrams of the most common production modes are shown in Figure 6.10 including the quark initiated t-channel production, the quark initiated s-channel production with the trilinear vector boson coupling and the bremsstrahlung process.

The MC samples are modelled using SHERPA 2.2.2 at the NLO accuracy for 0- and 1-jet and LO for 2- and 3-jet interfaced with NNPDF3.0 NNLO PDFs. For both $Z\gamma$ and $W\gamma$ processes the p_T of the photon are required to be larger than 7 GeV, and the distance in the $\eta - \phi$ plane $\Delta R > 0.1$. In addition the leptons from Z-boson in the $Z\gamma$ final state are required to have $m_{\ell\ell} > 2\,\mathrm{GeV}$.

$Z/\gamma^\star$+jets Production

The Drell-Yan process has a large cross section at hadron colliders. With a subsequent leptonic decay of the Z-boson, a background of two isolated, oppositely charged leptons arises. True missing transverse momentum appears only in the case of $Z/\gamma^\star \to \tau\tau$ decays, though. But mismeasurement of the leptons and especially of additional associated jets can also result in a wrong reconstruction. And pile-up contributions can lead to a significant degradation of the energy measurement. Figure 6.11 illustrates the production mechanisms of the $Z/\gamma^\star$+jets background

(a) Drell Yank production without associated jets.

(b) Drell Yank production with one associated jets.

Figure 6.11.: Leading order Feynman diagram of the Drell-Yan background production with and without associated jets.

process with and without associated jets. The $Z/\gamma^\star$ decays into two same flavor leptons and therefore mostly affects the ee and $\mu\mu$ channel. However, the decay via two τ leptons leads to a non-negligible contribution to the $e\mu$ channel as well. With stringent requirements on the missing transverse energy, the fraction of the background can be reduced significantly. The predicted rate of events is normalized in a dedicated control region, as explained in detail in Section 6.7.1.

The $Z/\gamma^\star$ production is simulated with SHERPA 2.2.1 having up to two jets in NLO and the forth and fifth jet in LO accuracy using the NNPDF3.0 NNLO PDFs with dedicated parton shower tuning developed by the Sherpa authors. In order to generate sufficient statistics at high transverse momenta of the Z-boson, the samples are split according to $\max(H_T, p_T^V)$, where H_T is defined as the scalar sum of all the final state transverse momenta. Additionally, to obtain sufficient heavy-flavor final state statistics, the samples are generated applying c- and b-quark filters.

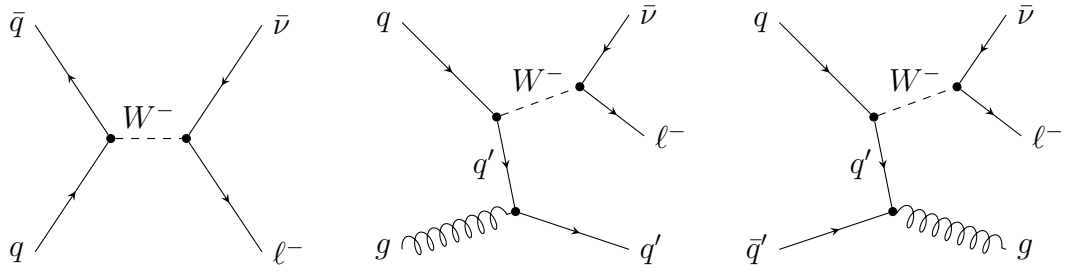
Moreover, for the $Z \to \tau\tau$ samples an additional filter on the lepton or hadron transverse momenta has been added for the samples up to $\max(H_T, p_T^V) < 280\,\text{GeV}$ to better populate the analysis phase space. Two filter requirements are applied, where the leptons originated from a $\tau \to \ell\ell$ decay require to have transverse momenta of 7 and 13 GeV if both τ leptons decay leptonically. In case of one τ decaying hadronically and the other leptonically the leptons transverse momenta have to be larger than 15 GeV and the hadron transverse momenta larger than 20 GeV.

Samples are normalised using cross sections calculated at NNLO accuracy [119].

The $q\bar{q} \to Zq\bar{q}$ process is modelled using SHERPA 2.1.1 at LO with no $O(\alpha_S)$ terms with leptonically decaying Z-boson demanding an invariant dilepton mass of $m_{\ell\ell} > 40\,\text{GeV}$.

W+jets and QCD Production

The production of a W-boson in association with jets can also contribute to the background of the $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ channel, if the W-boson decays



(a) W-boson production without an associated jet.

(b) W-boson production with an associated quark jet.

(c) W-boson production with an associated gluon jet.

Figure 6.12.: Leading order Feynman diagrams for the production of a W-boson.

leptonically and the additional jet gets misidentified as a second lepton. These events contain real missing transverse energy accounting for the neutrino of the W-decay. Figure 6.12 shows the leading order Feynman diagrams for the W+jets background.

The probability to misidentify a jet as a high-p_{T} lepton is expected to be quite small, at the order of 10^{-5}. However, the total production cross section of the W+jets background is so large that it compensates the low misidentification rate. The W+jets background and the signal process are expected to have similar size. Since they also share similar kinematics, the key to suppress the W+jets contribution lies in the misidentified second lepton. Strong requirements on the track and calorimeter isolation help to distinguish the falsely reconstructed hadronic jets or real non-isolated leptons from the signal process. Another contribution to the background processes is QCD dijet production. In this case, both signature leptons are the result of the misidentification of the two jets. Again, the large cross section does compensate the unlikliness to wrongly reconstruct both high-p_{T} objects in the event. The MC simulation is not expected to model such misidentification problems sufficiently well. That is why in this analysis, both rate and shape of the W+jets background are extracted with a data-driven approach and a separate control region as described in Section 6.7.2.

In order to calculate a correction factor in the high-p_{T} signal region due to the modified lepton selection (cf. Section 6.7.2) and to include W+jets MC events in the BDT training (cf. Section 6.6), W+jets MC samples are used. These samples are generated in the same way as the Sherpa 2.2.1 Z+jets samples that are sliced in $\max(H_T, p_T^V)$ with additional heavy quark filters. Additionally, the $WW \to \ell\nu qq$ process is included, where a misidentified jet as lepton is mimicing the same final state as the signal process. This sample is generated using Powheg-Box v2 requiring the lepton p_{T} to be above 5 GeV.

Simulation, Reconstruction and Pile-up

All events are generated at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. They are passed through the ATLAS detector simulation [92, 93], and the standard ATLAS reconstruction used for collision data. EVTGEN 1.2.0 [116] is used for all event samples that are produced with PYTHIA or HERWIG [120] as the parton shower generator in order to unify the properties of the bottom and charm hadron decays.

The effect of additional inelastic 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 PYTHIA 8 [108] 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 $\hat{\mu} = 13.7$ for 2015 and $\hat{\mu} = 24.9$ for 2016.

Table 6.6.: MC generators used to model the background processes, and corresponding cross sections. When the leptonic decay filter is applied on W/Z bosons the quoted cross sections include the branching ratios and are summed over lepton flavors. The column *Precision* $\sigma_{\text{incl.}}$ gives the precision of the inclusive cross section applied to the sample.

Process	Matrix Element	PDF	Parton Shower	Precision $\sigma_{\text{incl.}}$	$\sigma \cdot \text{Br}$ [pb]
VH	POWHEG-BOX V2 (MiNLO)	PDF4LHC15 NLO	PYTHIA 8	NNLO QCD + NLO EW	
$WH\ H \to WW$					0.136
$ZH\ H \to WW$					0.027
$H \to \tau\tau$	POWHEG-BOX V2	CT10	PYTHIA 8		
ggF $WH\ H \to \tau\tau$				N^3LO QCD + NLO EW	0.34
VBF $WH\ H \to \tau\tau$				NNLO QCD + NLO EW	0.029
top					
$t\bar{t}$ di-leptonic (e, μ, τ)	POWHEG-BOX V2	NNPDF3.0 NLO	PYTHIA 8	NNLO+NNLL	87.6
Wt leptonic	POWHEG-BOX V1	CT10	PYTHIA 6.482	NLO	7.55
WW					
$qq \to WW \to \ell\nu\ell\nu$	SHERPA 2.2.2	NNPDF3.0 NNLO	SHERPA 2.2.2	NLO	12.5
$gg \to WW \to \ell\nu\ell\nu$	SHERPA 2.1.1	CT10	SHERPA 2.1	NLO	0.87
EW $WW + 2$ jets $(\ell\nu\ell\nu)$	SHERPA 2.1.1	CT10	SHERPA 2.1	LO	0.075
WZ/Vγ^*/ZZ					
WZ/ZZ	POWHEG-BOX V2	CT10	PYTHIA 8	NLO	26.39
$WZ/\gamma^\star \to \ell\nu\ell\ell$	SHERPA 2.1.1	CT10	SHERPA 2.1	NLO	11.88
Vγ	SHERPA 2.2.2	NNPDF3.0 NNLO	SHERPA 2.2.2	NLO	
$(W \to \ell\nu)\gamma$ $(p_T^\gamma > 7\,\text{GeV})$					1065
$(Z \to \ell\ell)\gamma$ $(p_T^\gamma > 7\,\text{GeV})$					297
Z+jets					
inclusive $Z/\gamma^\star \to \ell\ell$	SHERPA 2.2.1	NNPDF3.0 NNLO	SHERPA 2.2.1	NNLO	11 555
EW $Z + 2$ jets $(\ell\ell)$	SHERPA 2.1.1	CT10	SHERPA 2.1	LO	7.63
W+jets	SHERPA 2.2.1	NNPDF3.0 NNLO	SHERPA 2.2.1	NNLO	59 625
$WW \to \ell\nu qq$	POWHEG-BOX V2	CT10	PYTHIA 8	NLO	44.17

6.3 Object Identification and Selection

Events should at least have one primary vertex with as many as two associated tracks with transversal momenta $p_{\mathrm{T}} > 400\,\mathrm{MeV}$. If more than one such vertex is reconstructed in one single event, the one with the largest sum of square transverse track momenta is chosen as the hard-scatter primary vertex and is subsequently used for calculation of the main physics objects in this analysis: electrons, muons, jets and missing transverse momentum.

6.3.1 Electrons and Muons

All leptons have to originate from the primary vertex by requiring the absolute value of the longitudinal impact parameter to satisfy $|z_0 \sin\Theta| < 0.5\,\mathrm{mm}$. Furthermore, the significance of the transverse impact parameter $|d_0|/\sigma_{d_0}$ have to be less than three (five) for muons (electrons). At least one of the lepton candidates is required to match a trigger object for the event to be selected.

The electrons are need to have a transverse energy E_T greater than 15 GeV and pass the `MediumLH` or `TightLH` selection defined in Section 4.3 for electrons with greater or smaller E_T than 25 GeV, respectively. The pseudorapidity of electrons have to be within the range of $|\eta| < 2.47$, excluding the transition region between the barrel and endcaps in the liquid argon calorimeter. Moreover, only electrons are selected that are reconstructed exclusively as electrons and no other object. Such selection reduces radically the $W\gamma$ background by more than 40% while giving a signal loss of about 1%.

The isolation criteria for electrons are optimized separately for two signal regions, the low-p_{T} region where the leading jet is required to have a transverse momentum of $p_{\mathrm{T}} < 200\,\mathrm{GeV}$ and the high-p_{T} region for events with a transverse momentum of the leading jet higher than 200 GeV.

In the low-p_{T} signal region the isolation is further subdivided for electrons with a transverse momentum lesser and greater 25 GeV. Electrons with $p_{\mathrm{T}} < 25\,\mathrm{GeV}$ need to have a track isolation, where the scalar sum of the p_{T} of the tracks within a $R = 0.4$ size cone around the electron (excluding its own track), $p_T^{\mathrm{varcone0.4}}$, is less than 6% of the lepton p_{T}. In addition, the sum of the transverse energy of the calorimeter energy clusters in a cone of $\Delta R = 0.2$ around the electron (excluding the deposit from the electron itself), $E_T^{\mathrm{cone0.2}}$, must be less than 11% of the electron p_{T}. For the isolation for electrons with a transverse momentum greater than 25 GeV the `Gradient` working point defined in Section 4.3 is used.

In the high-p_{T} region the leptons tend to be closer together (cf. Figure 6.13). Hence, the track isolation quantity $p_T^{\mathrm{varcone0.4}}$ is corrected by subtracting the transverse momentum of the muon track, if it is situated within the electron cone

$\Delta R = \min(0.4, 10\,\text{GeV}/E_T)$. Otherwise, the same requirements are applied as for the electrons with $p_{\text{T}} < 25\,\text{GeV}$ in the low-p_{T} signal region but with the corrected $p_T^{\text{varcone0.4}}$.

The muons are obtained via the *combined muon* definition, as introduced in Section 4.4. The muon candidates are required to pass the `Tight` selection and to satisfy $p_T > 15\,\text{GeV}$ and $|\eta| < 2.5$. Muon tracks are isolated by requiring $p_T^{\text{varcone0.3}}/p_{\text{T}} < 0.06$ and the calorimeter isolation of the muons selects $E_T^{\text{cone0.2}}/p_{\text{T}} < 0.09$.

Similar to the electrons in the high-p_{T} signal region both quantities $p_T^{\text{varcone0.3}}$ and $E_T^{\text{cone0.2}}$ are corrected. All electron tracks within the cone $\Delta R = \min(0.3, 10\,\text{GeV}/E_T)$ are subtracted from $p_T^{\text{varcone0.3}}$ and the electron transverse momentum is subtracted from $E_T^{\text{cone0.2}}$ if the electron is within the muon cone $\Delta R < 0.2$. The selection criteria are summarized in Table 6.7.

Additionally an overlap removal between an electron and muon is applied. If a combined muon shares an ID track with an electron, the electron is removed, whereas if a calo-tagged muon shares an ID track with an electron, the muon is removed.

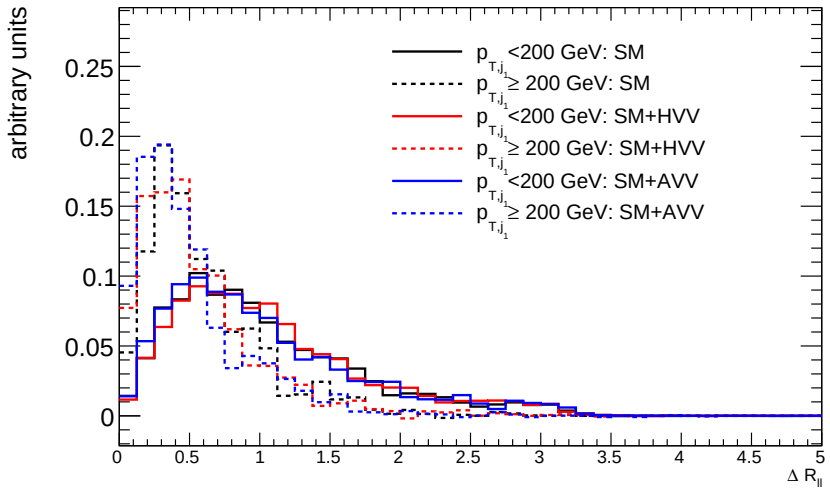


Figure 6.13.: The angular separation between the leading and subleading lepton $\Delta R_{\ell\ell}$ for events with a transverse momentum of the leading jet larger and smaller than 200 GeV for SM and different EFT scenarios. The leptons are much closer if the leading jet has a high transverse momentum.

Table 6.7.: Summary of the lepton selection. The corrected isolation quantities are used in the high-p_{T} signal region and are described in the text.

p_{T} range	ID	Track Isolation	Calo Isolation	Impact Parameters
Electrons in the low-p_{T} signal region				
$[15, 25]$ GeV	`TightLH`	$p_T^{\mathrm{varcone0.4}}/p_{\mathrm{T}} < 0.06$	$E_T^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.11$	$\lvert z_0 \sin\Theta\rvert < 0.5\,\mathrm{mm}$, $\lvert d_0\rvert/\sigma_{d_0} < 5$
> 25 GeV	`MediumLH`	`Gradient`		
Electrons in the high-p_{T} signal region				
$[15, 25]$ GeV	`TightLH`	$p_T^{\mathrm{varcone0.4,corr.}}/p_{\mathrm{T}} < 0.06$	$E_T^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.11$	$\lvert z_0 \sin\Theta\rvert < 0.5\,\mathrm{mm}$, $\lvert d_0\rvert/\sigma_{d_0} < 5$
> 25 GeV	`MediumLH`			
Muons				
> 15 GeV	`Tight`	$p_T^{\mathrm{varcone0.3,(corr.)}}/p_{\mathrm{T}} < 0.06$	$E_T^{\mathrm{cone0.2,(corr.)}}/p_{\mathrm{T}} < 0.09$	$\lvert z_0 \sin\Theta\rvert < 0.5\,\mathrm{mm}$, $\lvert d_0\rvert/\sigma_{d_0} < 3$

6.3.2 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 4.5. They are required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 4.5$. For jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$, the calibrated JVT variable, defined in Section 4.5, is required to be larger than 0.59 to suppress jets from pile-up events. Furthermore, for jets with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$ and $|\eta| > 2.5$, additionally the ForwardJVT is applied to reduce the pile-up in the forward region. The working point has an efficiency of 90%.

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 with containing b-hadrons are identified using the `MV2c10` b-tagging algorithm introduced in Section 4.5 using the 85% working point. Top-quark background is suppressed by vetoing events with any b-tagged jet passing a reduced p_{T} threshold of 20 GeV.

The missing transverse energy is used to build variables sensitive to the VBF signal, like $m_{\tau\tau}$ and $p_{\mathrm{T}}^{\mathrm{tot}}$. For this purpose the track-based TST $E_{\mathrm{T}}^{\mathrm{miss}}$ is used defined in Section 4.6.

6.4 Observables

Besides the already in Section 6.1 introduced observable $\Delta\phi_{jj,\mathrm{sign}}$ several different quantities are used to either reduce background contributions or to select specific

topological features of the signal process in single selection requirements or as input quantity to the multivariant analysis described in Section 6.6. The definitions of these observables are given in the following.

$m_{\ell\ell}$ The invariant mass of the dilepton system originating from the hard scattering interaction is especially effective at rejecting Drell-Yan production of the Υ-meson resonance and $\gamma^{\star}$ contributions. The VBF signal tends to be at low values.

m_{jj} The invariant mass of the dijet system rejects contributions from the VH production, especially if the jets both are decay products of the associated gauge boson and thus m_{jj} peaks at its mass. But also by cutting at lower phase space regions the background is rejected collectively, while keeping most of the VBF signal, since a high invariant mass is a distinctive signature of the signal events.

$m_{\tau\tau}$ In the collinear approximation [121] the charged leptons in the final state, that are the products of the decay of a pair of τ leptons, are collinear with the emitted neutrinos. Assuming these neutrinos are the only source of the observed $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event, the missing transverse energy can be computed using this approximation. Here, the TST missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss,TST}}$ is used in the $m_{\tau\tau}$ calculation. If the result is physical then a cut is imposed to suppress $Z \to \tau\tau$ background contribution and $H \to \tau\tau$ decays.

$p_{\mathrm{T}}^{\mathrm{tot}}$ The total transverse momentum $p_{\mathrm{T}}^{\mathrm{tot}}$, defined as $p_{\mathrm{T}}^{\ell_1} + p_{\mathrm{T}}^{\ell_2} + E_{\mathrm{T}}^{\mathrm{miss}} + \sum_j p_{\mathrm{T}}^{j}$, where the sum over jets runs over all jets which pass the good-jet definition criteria. This variable helps disentangling events with significant soft gluon radiation that recoils against the $\ell\ell + 2j$ system with no high-p_{T} jets.

$p_{\mathrm{T}}^{\ell\ell}$ The transverse momentum of the dilepton system $|\vec{p}_T^{\,\ell_1} + \vec{p}_T^{\,\ell_2}|$ from the leptons of the hard scattering interaction tend to be larger for the VBF signal than for the background processes, especially in a topology with large transverse recoil (high-p_{T} jets) and leptons that are close together.

$m_{\mathrm{T}}^{\ell_1}$ The transverse mass of the lepton with the highest transverse momentum is defined as

$$m_{\mathrm{T}}^{\ell_1} = \sqrt{\left(|\vec{p}_{\ell_1}| + E_{\mathrm{T}}^{\mathrm{miss}}\right)^2 - \left|\vec{p}_{\ell_1} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}}\right|^2}.$$

For this quantity the signal peaks mainly due to the opposed direction of the leading lepton and the missing transverse energy preferable at lower values than most of the background processes and can thus distinguish the VBF process from other processes.

$\Delta R_{\ell\ell}$ The angular separation $\sqrt{\Delta\eta_{\ell\ell} + \Delta\phi_{\ell\ell}}$ of the two hard scatter leptons is very small particularly in a boosted topology for the VBF process (cf. Figure 6.13).

$\sum_{\ell,j} m_{\ell j}$ The sum of the invariant masses of all four possible lepton-jet pairs peaks for the VBF signal at a higher value than the backgrounds. Its use is motivated because the jets in the VBF signal topology tend to be in the forward regions while leptons remain central, resulting in large opening angles between the leptons and the jets.

$\eta_{\text{lep}}^{\text{centrality}}$ This quantifies evaluates the exact positions of the leptons with respect to the two tag jets in the η-plane. It is defined with the so-called *outside lepton veto* (OLV) as follows

$$\text{OLV}_{l_1} = 2 \cdot \left| \frac{\eta_{l_1} - \bar{\eta}}{\eta_{j_1} - \eta_{j_2}} \right|$$
$$\text{OLV}_{l_2} = 2 \cdot \left| \frac{\eta_{l_2} - \bar{\eta}}{\eta_{j_1} - \eta_{j_2}} \right|$$

$$\eta_{\text{lep}}\,\text{centrality} = \text{OLV}_{l_1} + \text{OLV}_{l_2},$$

where $\bar{\eta} = (\eta_{j_1} + \eta_{j_2})/2$ is the average η of the two tag jets. For each lepton,

$$\text{OLV}_l \begin{cases} = 0 & : \ell \text{ is right in the middle of the rapidity gap between the two tag jets.} \\ < 1 & : \ell \text{ lies within the rapidity gap between the two tag jets.} \\ > 1 & : \ell \text{ is outside the rapidity gap between the two tag jets.} \end{cases}$$

6.5 Signal Region Selection

The selection of candidate events uses basic criteria to select the signal from the $H \to W^{\pm}W^{\mp *} \to \ell^- \bar{\nu}_\ell \ell'^+ \nu_{\ell'}$ decay and reject other final state topologies and background processes. The first requirement imposed is the presence of exactly two leptons, one electron and one muon with opposite electrical charge. Both have to exceed a p_{T} threshold, which is 15 GeV for the subleading lepton and 22 GeV for the leading one. They are chosen on the one hand as low as possible corresponding to the trigger selection described in Section 6.2.1 to enhance the signal acceptance, where low p_{T} leptons are expected, and on the other hand to reduce significantly the background contribution of misidentified leptons which has a large impact especially on low p_{T} leptons.

Figure 6.14 shows the kinematic distributions of the two leptons after the lepton transverse momenta requirements as well as the $E_{\text{T}}^{\text{miss}}$ distribution in Figure 6.15. The background processes are stacked and the SM signal process and three EFT scenarios with SM plus each considered BSM parameter are superimposed and scaled by a factor of 100 for better visibility. The majority of the background processes is derived from MC simulation. Only the contributions of the misidentified

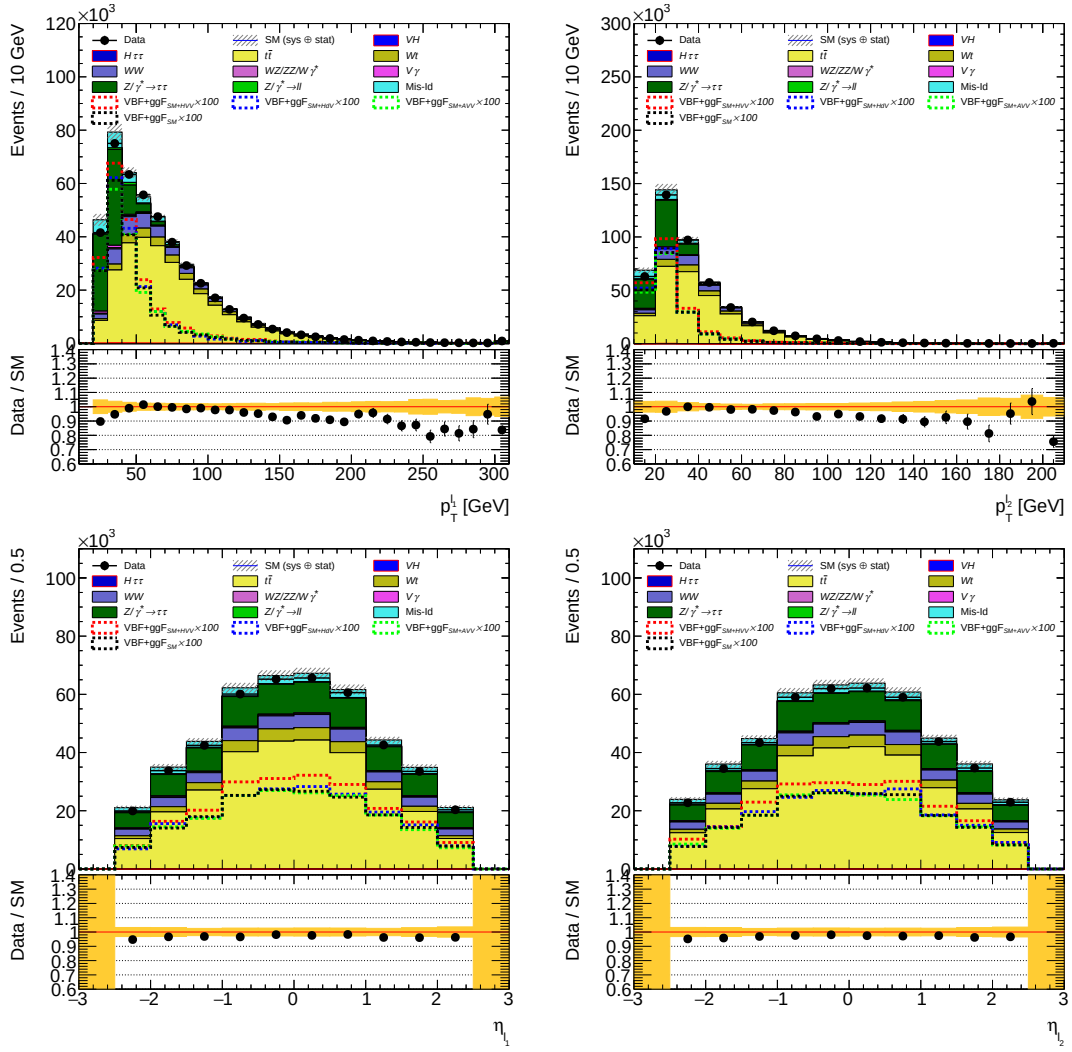


Figure 6.14.: The leading lepton (left) and subleading lepton (right) p_{T} (top) and η (bottom) distributions after applying the lepton p_{T} requirements. The SM and EFT signal scenarios are scaled by a factor of 100 for a better visibility. The yellow band represents MC statistics and all detector systematics uncertainties.

leptons are referring to dedicated data driven estimates, which is described in Section 6.7.2 later on. The large amount of selected events generally shows good agreement of data with the MC simulation, where in the data MC ratio the yellow band includes both statistical and all detector systematics uncertainties. The background contributions still surmount the signal process by several orders of magnitude, lead by $t\bar{t}$ processes. At high lepton momenta the MC prediction is slightly increased compared to the data. This region is dominated by the top background, too, which theoretical uncertainty is not applied yet and thus the uncertainty does not include the full deviation.

With two energetic and isolated leptons in the final state, a low bound on the dilepton mass of $m_{\ell\ell} > 10\,\text{GeV}$ cleans the selection from Drell-Yan production of the Υ-meson resonance and $\gamma^\star$ contributions. Figure 6.16 shows the $m_{\ell\ell}$ distributions before the application of the lower bound. Below 10 GeV the data-driven yield of misidentified leptons is over-estimated because the same lepton flavor estimates do not include the meson resonances and subtract insufficient amount of MC events from the data yields leading to an increased MC prediction in the different lepton flavor estimate. In Section 6.7.2 the data-driven method is explained in more detail.

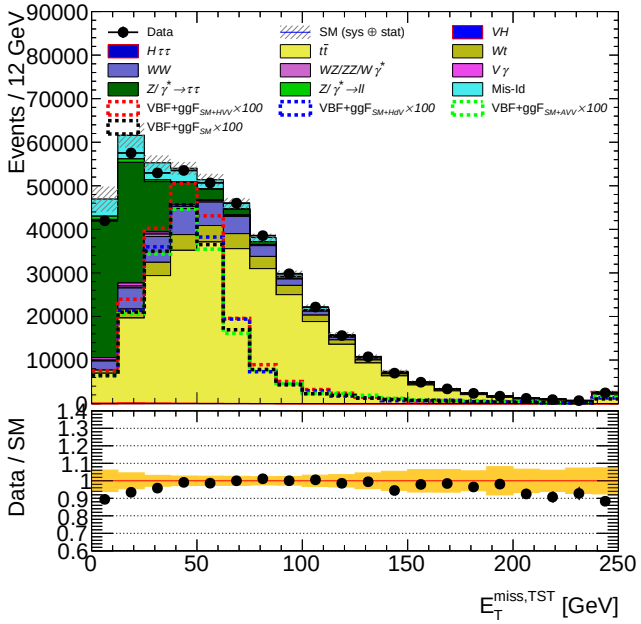


Figure 6.15.: The $E_{\text{T}}^{\text{miss}}$ distributions after applying the lepton p_{T} requirements. The SM and EFT signal scenarios are scaled by a factor of 100 for a better visibility. The yellow band represents MC statistics and all detector systematics uncertainties.

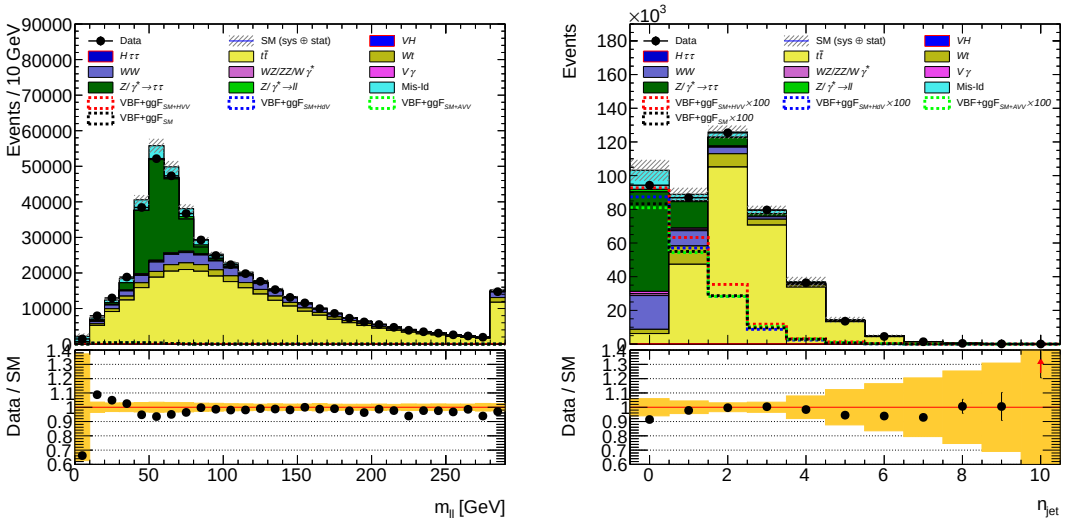


Figure 6.16.: The $m_{\ell\ell}$ (left) and n_{jet} (right) distributions before applying respective requirements on $m_{\ell\ell}$ and n_{jet}. The SM and EFT signal scenarios are scaled by a factor of 100 for a better visibility. The yellow band represents MC statistics and all detector systematics uncertainties.

Two or more jets are required to select precisely the VBF topology, where two jets from the Higgs production are expected and further jets may be produced due to e.g. initial state gluon radiation. Figure 6.16 shows the n_{jet} distribution after the $m_{\ell\ell}$ requirement. After the jet selection the dominant background contribution originates still from $t\bar{t}$. The p_{T} and η distributions of the leading and subleading jet with a good data to MC agreement are shown in Figure 6.17.

Furthermore, events with $m_{jj} > 120\,\text{GeV}$ are selected to reduce the VH contribution which is not simulated for different EFT scenarios. If both jets originate from the production vector boson in the VH process then their invariant mass peaks at the mass of the respective vector boson and is excluded by the m_{jj} requirement. Thereby approximately 60% of VH events are discarded. Additionally the $t\bar{t}$ background is rejected by 26% while keeping 92% of the SM VBF signal events. The m_{jj} distribution before applying the selection is shown in Figure 6.18.

As already discussed in Section 6.1 the main event yield variations between SM and the different EFT scenarios are present in a boosted topology, where e.g. the leading jet has a high transverse momentum. Here, the EFT samples have a much higher event yield expectation than for the SM case. Therefore, the signal region is split into a low- and a high-p_{T} signal region by selection $p_{\text{T}}^{j_1} \lessgtr 200\,\text{GeV}$. Nevertheless, the low-$p_{\text{T}}$ has higher statistics and is hence useful to conduct a differential analysis using the $\Delta\phi_{jj,\text{sign}}$ observable. Figure 6.18 shows the $p_{\text{T}}^{j_1}$ distribution before the splitting is performed. The higher event yields for the EFT samples are clearly

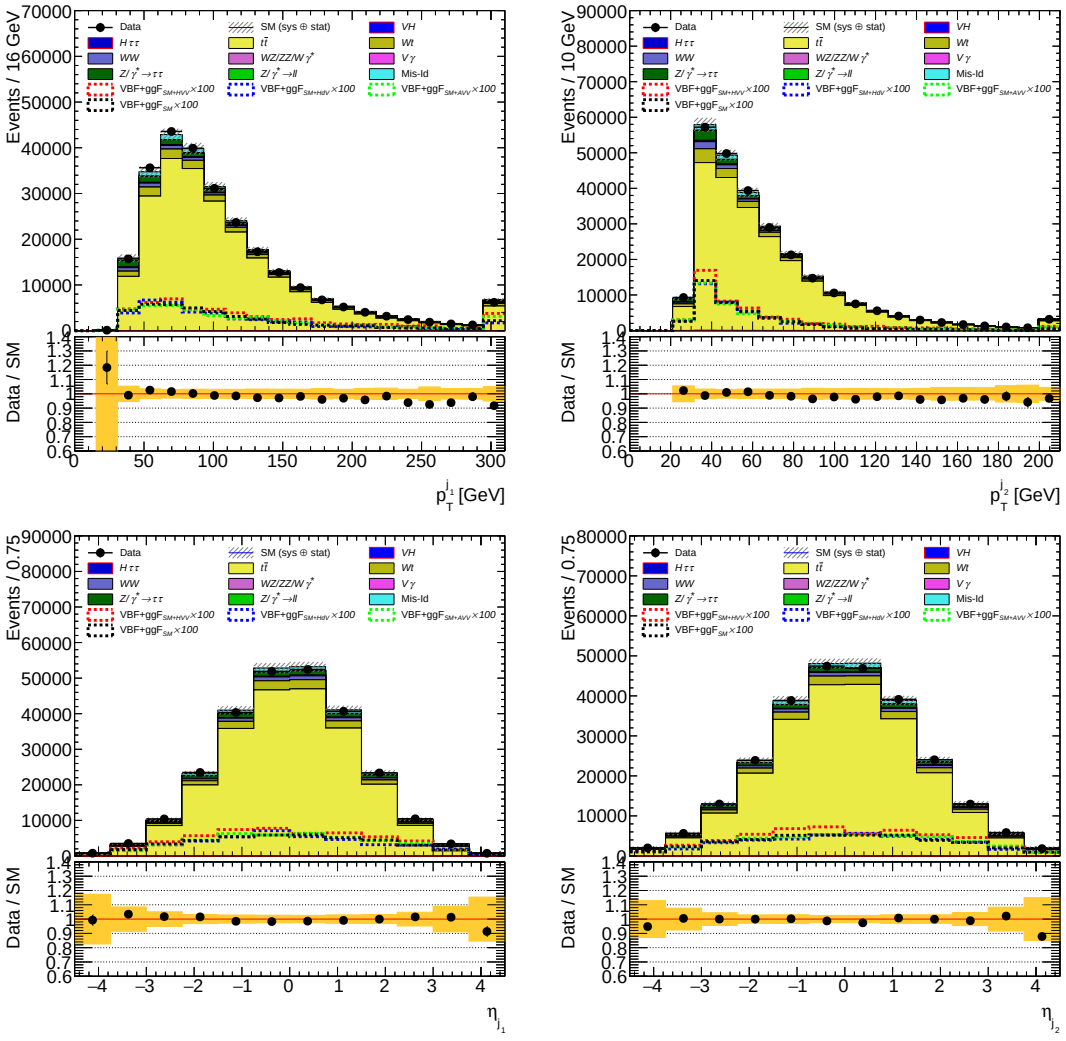


Figure 6.17.: The leading jet (left) and subleading jet (right) p_{T} (top) and η (bottom) distributions after the $n_{\mathrm{jet}} > 2$ requirement. The SM and EFT signal scenarios are scaled by a factor of 100 for a better visibility. The yellow band represents MC statistics and all detector systematics uncertainties.

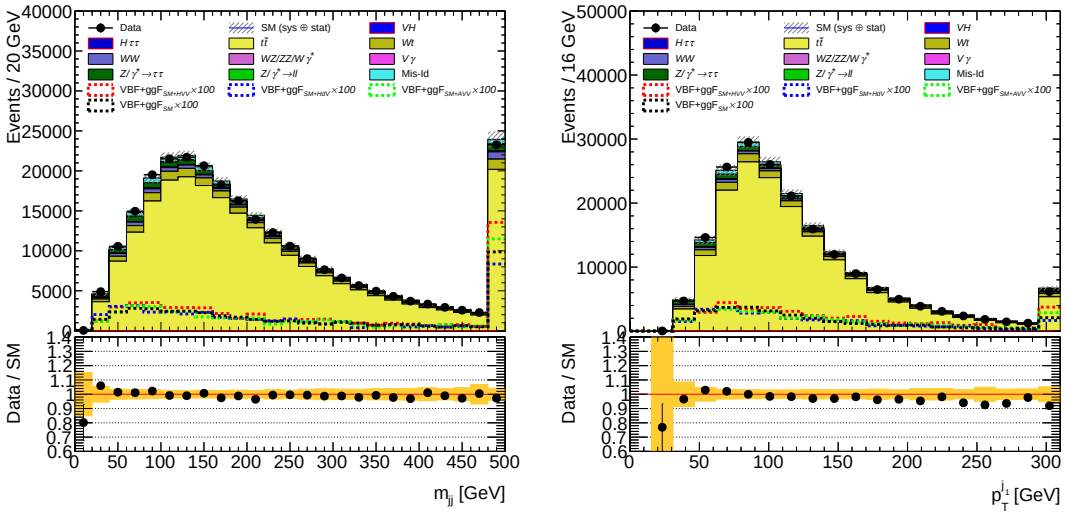


Figure 6.18.: The m_{jj} (left) and $p_{\mathrm{T}}^{j_1}$ (right) distributions before applying respective requirements on m_{jj} and $p_{\mathrm{T}}^{j_1}$. The SM and EFT signal scenarios are scaled by a factor of 100 for a better visibility. The yellow band represents MC statistics and all detector systematics uncertainties.

visible in the overflow bin of the figure for κ_{HVV} and κ_{AVV} contributions to the SM. The SM+H∂V shows lower event yields as has already been visible in Figure 6.4 for the $p_{T,jj}$ distribution, which originates from a destructive interference between both contributions. A negative H∂V parameter of the same magnitude leads to a much higher event yield compared to the pure SM case.

During the preparation and optimization of the analysis a blinding criteria after this preselection was applied using the BDT response distributions described in Section 6.6. A cut requirement was set in each signal region such that the by 25% increased VBF+ggF event yield is smaller than the total background statistical uncertainty taken as the square root of the background events at this cut stage such that it is indistinguishable from a potential statistical fluctuation. With the blinding criteria the amount of SM VBF+ggF events is less then 0.02% (0.16%) in the low(high)-p_{T} region. Anyhow, all figures shown in the thesis are unblinded, i.e. the blinding criteria is not applied anymore.

Further requirements are set to reduce more background contribution and additionally define control regions to estimate normalization factor for different backgrounds. A b-jet veto reduces the $t\bar{t}$ background by 95% (95%) and the single top background by 88% (91%) while keeping 88% (80%) of the VBF signal in the low(high)-p_{T} region. The orthogonal region requiring exactly one b-jet to estimate the top background is described in Section 6.7.1. The b-jet distribution is shown in Figure 6.19, where the normalization factors for the top related backgrounds and

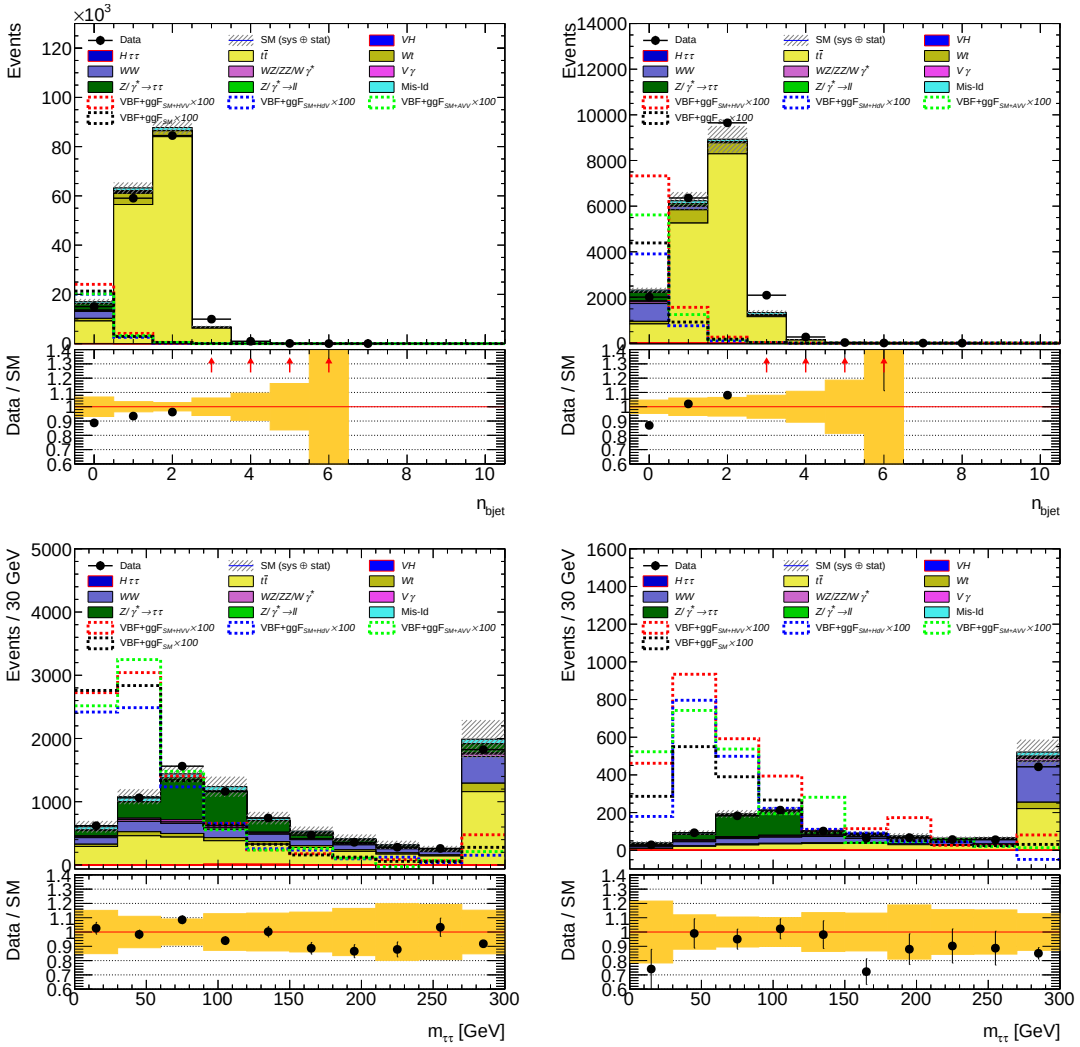


Figure 6.19.: The b-jet (top) and $m_{\tau\tau}$ (bottom) distributions before applying the respective b- and $Z \to \tau\tau$ vetos for the low- (left) and high- (right) p_{T} signal regions. The SM and EFT signal scenarios are scaled by a factor of 100 for a better visibility. The yellow band represents MC statistics and all detector systematics uncertainties.

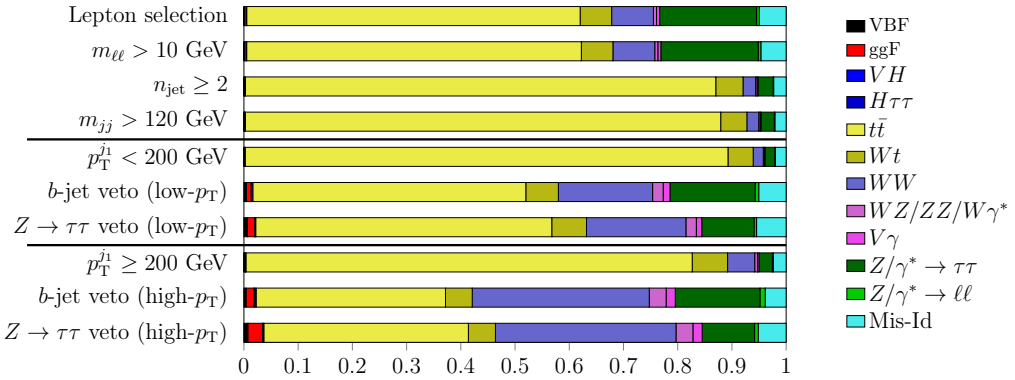


Figure 6.20.: Relative contributions of each background and the SM signals at each cutstage of the signal region selection.

the $Z \to \tau\tau$ background have already been applied. Data yield and MC estimate show a good agreement in the $n_{b\text{-jet}} = 1$ bin and a skewed distribution, where the MC has a slightly smaller expected yield in the $n_{b\text{-jet}} = 0$ bin and underestimates the data with increasing number of b-jets. The reason of the mismodelling is unknown so far and may come from interferences between $t\bar{t}$ and $b\bar{b}$ processes. Nevertheless, the region with $n_{b\text{-jet}} > 1$ is in any case not relevant for this analysis.

In order to reduce the $Z \to \tau\tau$ background a $Z \to \tau\tau$ veto is applied, specifically to remove events that are around the Z mass pole. Here, the reconstructed invariant mass of the $\tau\tau$ system $m_{\tau\tau}$ is required to be less than 25 GeV below the Z mass: $m_{\tau\tau} < m_Z - 25\,\text{GeV}$. Not only 65% (73%) of the $Z \to \tau\tau$ background contribution is reduced, but also 85% (89%) of the $H \to \tau\tau$ contamination, while keeping 86% (80%) of the SM VBF signal event yield in the low(high)-p_{T} region. The $m_{\tau\tau}$ distribution bevor applying the veto is shown in Figure 6.19. The orthogonal selection $m_{\tau\tau} > m_Z - 25\,\text{GeV}$ defines the $Z \to \tau\tau$ control region and is decribed in Section 6.7.1.

The relative contribution of each background at the different cut stages in the signal region selection is visualized in Figure 6.20. Before the b-veto the dominating process is $t\bar{t}$, after that the relative contribution of diboson processes exceed the $t\bar{t}$ processes. The event yields are presented in Section 6.8 in Table 6.16.

At this point two separate BDTs trained in the low- and high-p_{T} region to select the final signal regions, which are described in the following section.

6.6 Boosted Decision Tree Application

Different methods are available for the classification of signal events against the large amount of background events which is left after the first selection. The classical

approach is a cut based analysis where topological observables are constructed and used to separated the phase space enhancing the fraction of signal events in the final signal region while rejecting as many non-signal events as possible. Although this method has a comprehensive visual access, it is usually exceedingly outperformed by multivariate machine learning techniques, such as *boosted decision trees* (BDT) [122–124], for which reason this method is used to define the final signal regions. The advantage hereby is that the visualizability still persists at least for the single *decision trees* whereof a BDT is built from. Such a decision tree is a binary tree structured classifier similar to the one sketched in Figure 6.21. Repeated left/right (yes/no) decisions are taken on one single observable at a time until a stop criterion is fulfilled. The phase space is split this way into many regions that are eventually classified as signal or background, depending on the majority of training events that end up in the final leaf node. At each node, the split is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. The best separation is measured by calculating the impurity in each node with the *Gini index* $G = p(1-p)$, where $p = \frac{S}{S+B}$ is the fraction of signal events in the particular node. The criterion $G_{\text{parent}} - G_{\text{child}_1} - G_{\text{child}_2}$ is then maximized. These splittings are in this respect similar to optimal rectangular cuts. However, whereas a cut-based analysis is able to select only one hypercube as region of phase space, the decision tree is able to split the phase space into a large number of hypercubes, each of which is identified as either signal-like or background-like.

A shortcoming of decision trees is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. For example, if two input variables exhibit similar separation power, a fluctuation in the training sample may cause the tree growing algorithm to decide to split on one variable, while the other variable could have been selected without that fluctuation. In such a case the whole tree structure is altered below this node, possibly resulting in a substantially different classifier response.

This problem is overcome by constructing a forest of decision trees using the *adaptive boosting* technique (`AdaBoost`). The trees are derived from the same training ensemble by reweighting misidentified events. First an initial decision tree T_1 is constructed and the total weight of misclassified events $e_{\text{miss}} = \sum_{\text{miss i}} w_i$ is calculated. With the tree score $\alpha_1 = \beta \ln(\frac{1-e_{\text{miss}}}{e_{\text{miss}}})$ the event weights are updated according to $w_i \to w_i e^{\alpha_1}$ and all events are renormalized with $w_i \to w_i / \sum_i w_i$. This is repeated until the predefined number of trees are constructed. Finally, all trees are combined into a single classifier adding up the single weighted trees with $T(x_i) = \sum_{m=1}^{N_{\text{trees}}} \alpha_m T_m(x_i)$. The performance is often further enhanced by forcing a slow learning which is controlled by the leaning rate factor β. In general, boosting stabilizes the response of the decision trees with respect to fluctuations

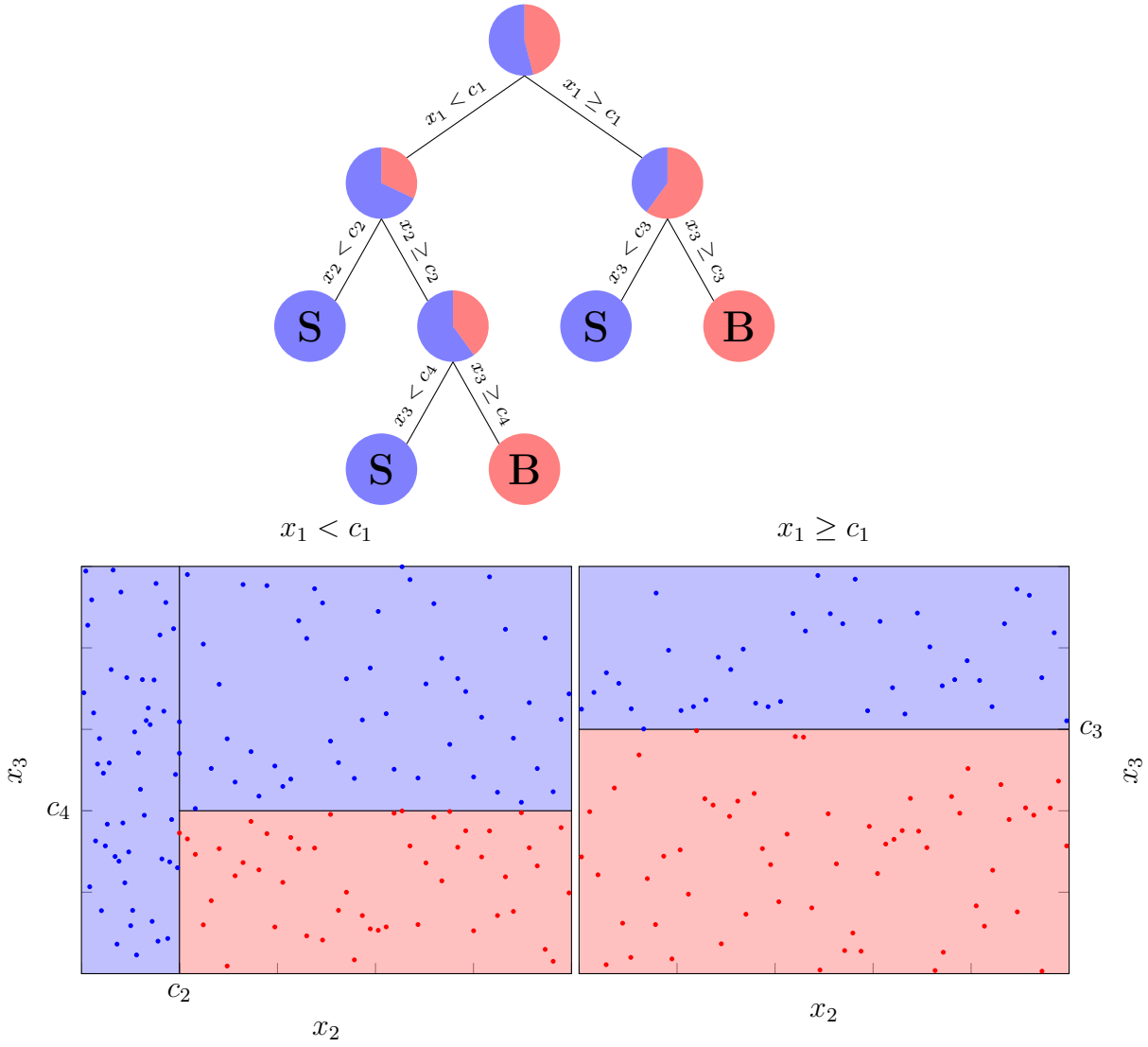


Figure 6.21.: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled `S` for signal and `B` for background depending on the majority of events that end up in the respective nodes.

in the training sample and is able to considerably enhance the performance with respect to a single tree.

6.6.1 Application

Two separate BDTs are trained and applied in the low- and high-p_{T} signal region after the $Z \to \tau\tau$ veto. To ensure that the BDTs are not trained on the statistical fluctuations in the training samples (*over-fitting*), each BDT is cross-evaluated, which means that the set of events that passes the pre-training selection is split into two statistically independent subsets, `even` and `odd`. The low- and high-p_{T} BDTs are each trained separately on both sets, $\mathrm{BDT}^{\mathrm{even}}$ and $\mathrm{BDT}^{\mathrm{odd}}$ respectively, resulting in total four trained BDTs. The discriminants used in the analysis are defined as:

$$\mathrm{BDT} = \begin{cases} \mathrm{BDT}^{\mathrm{even}} & \text{for event } \in \texttt{odd} \\ \mathrm{BDT}^{\mathrm{odd}} & \text{for event } \in \texttt{even} \end{cases} \tag{6.6.1}$$

for both low- and high-p_{T} BDT. Table 6.8 summarizes the unweighted MC statistics available for the trainings. The misidentified lepton samples (Mis-Id) are included using the W+jets MC samples described in Section 6.2.3.

Table 6.8.: MC statistics available for BDT trainings in both signal regions split in even and odd events. For the 11 VBF EFT samples the range of the raw event counts are given. The event counts for the full signal and background trainings sets are stated in *total signal* and *total background*.

Sample	low-p_{T} (even)	low-p_{T} (odd)	high-p_{T} (even)	high-p_{T} (odd)
VBF EFT	4958 – 5441	5040 – 5378	1156 – 2088 (457 SM+H∂V)	1106 – 2024 (451 SM+H∂V)
total signal	57585	57418	15699	15213
ggF SM	5387	5321	1063	1104
VH	9630	9387	1624	1714
$H \to \tau\tau$	2409	2464	244	215
$t\bar{t}$	18010	17766	1393	1352
Wt	2497	2566	219	251
WW	78546	78638	18266	18246
$WZ/ZZ/W\gamma^{\star}$	1130	1141	150	165
$V\gamma$	137	136	48	43
$Z \to \ell\ell$	184	188	179	170
$Z \to \tau\tau$	9065	9058	2092	2109
Mis-Id	133	123	114	127
total background	127128	126788	25392	25496

In order to maximize the separation between background events and VBF events while keeping the cross section and shape differences between SM and EFT samples, all 11 VBF EFT samples summarized in Table 6.4 are used as signal in the training, whereas all background samples and additionally the SM ggF sample are used as

background. First a set of input variables are defined which may be useful to discriminate between signal and background consisting of in total 93 features. A first BDT with all the features is trained in each region with each subset of events, where the maximal tree depth is set to 5, the number of trees to 100 and the minimal number of events per node to 20%. The training variables are then ranked by counting in how many nodes each variable is used and by weighting each utilization of a variable by:

- the gain in separation at each node and
- the number of events in each node.

Furthermore, KS tests [125] for each variable between the SM VBF sample and the single VBF EFT samples are performed to check which variable contains a certain amount of discrimination power between SM and EFT. Observables which are too different between EFT and SM, i.e. have a low KS test result, such as the transverse momentum of the dijet system $p_{T,jj}$, are omitted in the BDT training. In this way the BDT does not preferably select input variable phase space for a certain region of EFT parameter phase space and thus keep the EFT differences in the final selection of signal events. Comparing the KS test results and the variable ranking, the two most important variables $m_{\ell\ell}$ and m_{jj} for the low-p_{T} region and $m_{\ell\ell}$ and $\Delta R_{\ell\ell}$ for the high-pT region are detected and kept in all further BDT trainings. In addition 20 more variables are chosen with high importance and high KS test results separately in the low- and high-p_{T} region. All 8-variable combinations of the two fixed input variables and 20 additional variables are used to perform BDT trainings. Thus, in total approximately 40 000 trainings for both signal regions and both subsets of events are performed. The best set of variables is chosen by maximizing the significance value [126]

$$Z_A = \sqrt{2\left((s+b)\log\left[\frac{(s+b)(b+\sigma_b)}{b^2+(s+b)\sigma_b^2}\right] - \frac{b^2}{\sigma_b^2}\log\left[1+\frac{\sigma_b^2 s}{b(b+\sigma_b)}\right]\right)} \tag{6.6.2}$$

for the full low- and high-p_{T} BDT defined in Equation 6.6.1. Here, the background MC statistical uncertainty is explicitly taken into account, preventing to choose a BDT that selects mainly events with negative event weights. With the BDT output discriminant with values ranging between -1 and 1, the significance is scanned by cutting the BDT output at different values, calculating the number of signal events s, the number of background events b and the MC statistical uncertainty of the background σ_b and determining the respective significance. Moreover, at least one expected VBF SM event is required to remain after the BDT SR selection, such that trainings with high significance but very low VBF SM event yield are excluded and a statistical analysis with a minimal data yield is enabled. For the EFT samples even more events are expected due to their higher cross section. Finally, with the

sets of selected variables for both signal regions a hyper-parameter grid scan is performed using the parameters listed in Table 6.9. There, also the final selection of hyper-parameters are summarized.

Table 6.9.: BDT parameter used for the training and hyper-parameter grid scan values used.

Parameter	Value low-p_{T}	Value high-p_{T}	Range
Boosting algorithm	`AdaBoost`	`AdaBoost`	-
Maximum tree depth	4	4	$[1, \ldots, 6]$
Number of trees	300	300	$[50, \ldots, 400]$
Minimum number of events requires per mode	30%	20%	$[5\%, \ldots, 30\%]$
Learning rate	0.1	0.1	-

6.6.2 Results

The selected variables for both signal regions are shown in the variable ranking Figures 6.22. The rankings for both subsets `odd` and `even` are very similar, but since the variables are correlated, small variations can occur and improve the results significantly.

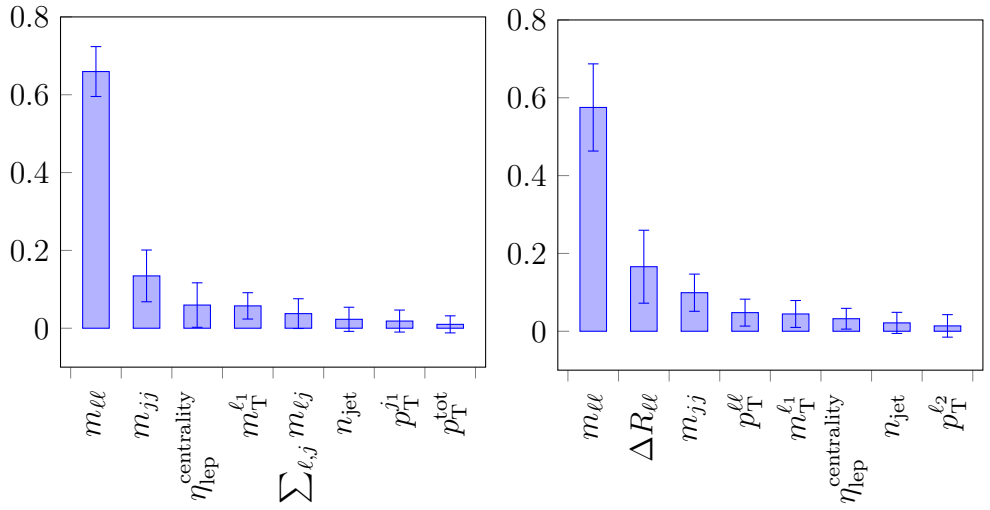


Figure 6.22.: Ranking of the input variables for $\mathrm{BDT}_{\mathrm{low}}$ (left) and $\mathrm{BDT}_{\mathrm{high}}$ (right). The ranking is taken from the `even` subset training but are similar for the `odd` subset training.

Figure 6.23 shows the BDT output distributions for signal in red and background in blue. The filled histograms are generated using the events with which the BDTs have been trained (`even/odd` events for $\mathrm{BDT}^{\mathrm{even/odd}}$), and for the dotted histograms

the opposite independent test event sets have been used as validation (`even` events for BDT^{odd} and vice versa). A good agreement is observed between both samples in each signal region thus no overtraining is present. The large first bin comprises mainly top and WW events (cf. first bin in left Figures 6.25) and include events with $m_{\ell\ell} \gtrsim 80\,\text{GeV}$ where no signal events are present.

Also the *receiver operating characteristic* (ROC) distributions in Figure 6.24 show acceptable agreement between training and test samples. Only a small overfitting is observed for the BDT_{high} which is most probably caused by the much smaller statistics that was available for the training. In any case, the method of cross-application ensures that any residual over-fitting would only result in a small degradation of sensitivity, and not in a bias, since the BDT evaluation is always performed using samples that are statistically independent from the training sample, while allowing the full MC samples to be used to develop the BDT.

The BDT output distributions shown in Figure 6.25 are then used as discriminant variable to construct the final signal regions. The most significant cuts are determined at $\text{BDT}_{\text{low}} > 0.6691$ and $\text{BDT}_{\text{high}} > 0.6305$ for both signal regions respectively using Equation 6.6.2. The same distribution is shown twice, one with a flat binning to observe the data to MC agreement which is reasonable good, especially in the high range, where the signal region is defined, and one distribution with three bins indicating the boundaries for the different regions. The first bin is used as validation region to check on potential mismodelling and the last bin defines the final signal region. Before the blinding criteria was removed the data to MC comparison could only be seen in the first validation bin which has a good agreement within the statistical and detector systematical uncertainty. After unblinding the small increased data yield is observed in both signal regions. The statistical significance of this increment is analyzed in detail in Section 6.10.

In Figure 6.26 and 6.27 the input observables are show for the BDTs in the low- and high-p_{T} signal regions respectively after applying the final BDT discriminant requirement. The increased data yields are observed in each distribution without exhibiting a specific input responsible for this deviation. Instead, all histograms show a regular data event increase throughout the bins. The distributions of the input observables in the BDT validation regions are shown in the appendix A.1. No obvious mismodelling can be observed. In addition the data to MC agreement is checked between all correlations of the input variables in Figure 6.28. For any two training variables X_i and X_j used in the particular BDT, distributions of the mean values per bin $\langle X_i \rangle$ vs. X_j and vice versa are plotted for the data and the MC model. In this figure the correlations of each pair of the training variables, as well as pairs of each variable and the BDT score are shown for the validation region. The profile plots compare the data (black) and MC model (red). The data to MC agreement is quantified using a simple χ^2 calculation. The uncertainties are statistical only.

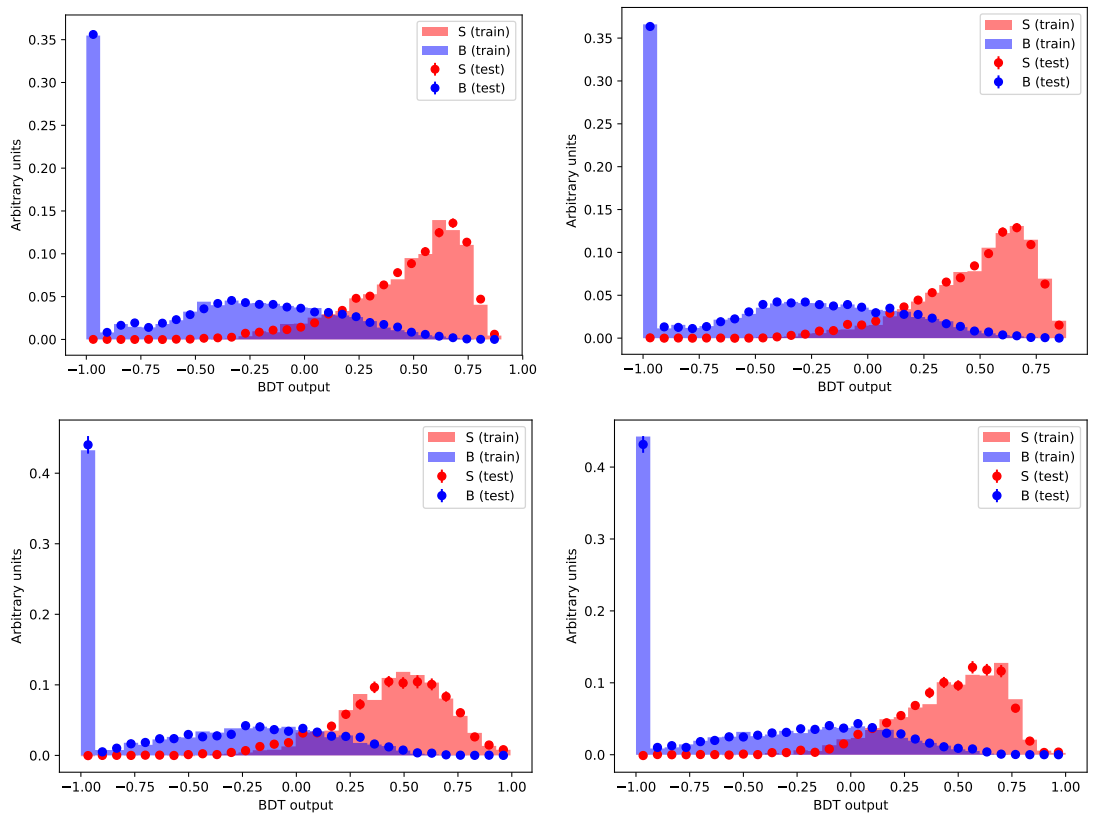


Figure 6.23.: Comparison of the BDT response between the `even` and `odd` subset of events taken as training and testing samples, respectively (left), and vice versa (right) for the low-p_{T} (top) and high-pT (bottom) signal region. The filled histograms are showing the BDT response on the same training events for signal (red) and background (blue). The pointed histograms are generated using the independent test events subset. No obvious overtraining is visible.

The canvas frame color for each plot encodes the χ^2 probability $p(\chi^2)$ from the comparison: green represents $p > 0.05$, yellow represents $0.005 < p < 0.05$ and red represents $p < 0.005$. Overall the MC models the correlations observed in data reasonably well.

In the next section the background estimation with control regions and data-driven approaches are presented.

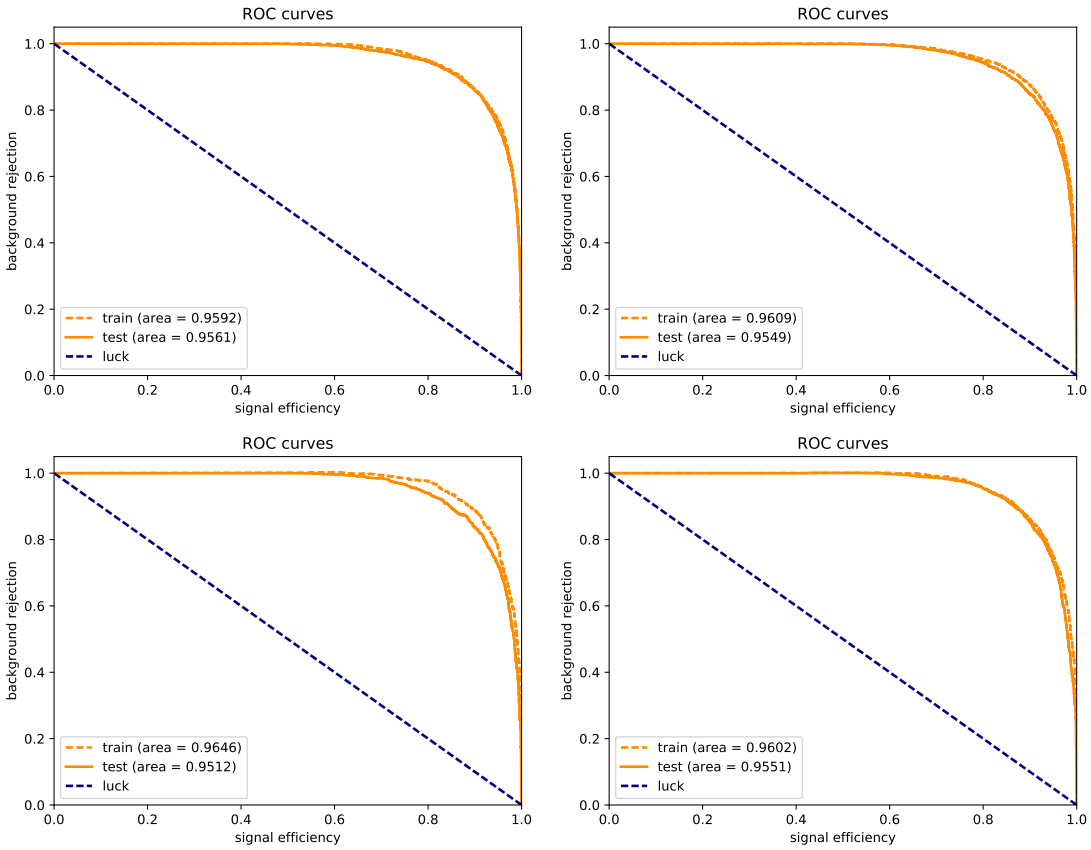


Figure 6.24.: Comparison of the ROC curves between the `even` and `odd` subset of events taken as training and testing samples, respectively (left), and vice versa (right) for the low-p_{T} (top) and high-pT (bottom) signal region. Only a small overtraining for $\mathrm{BDT}^{\mathrm{even}}_{\mathrm{high}}$ is observed.

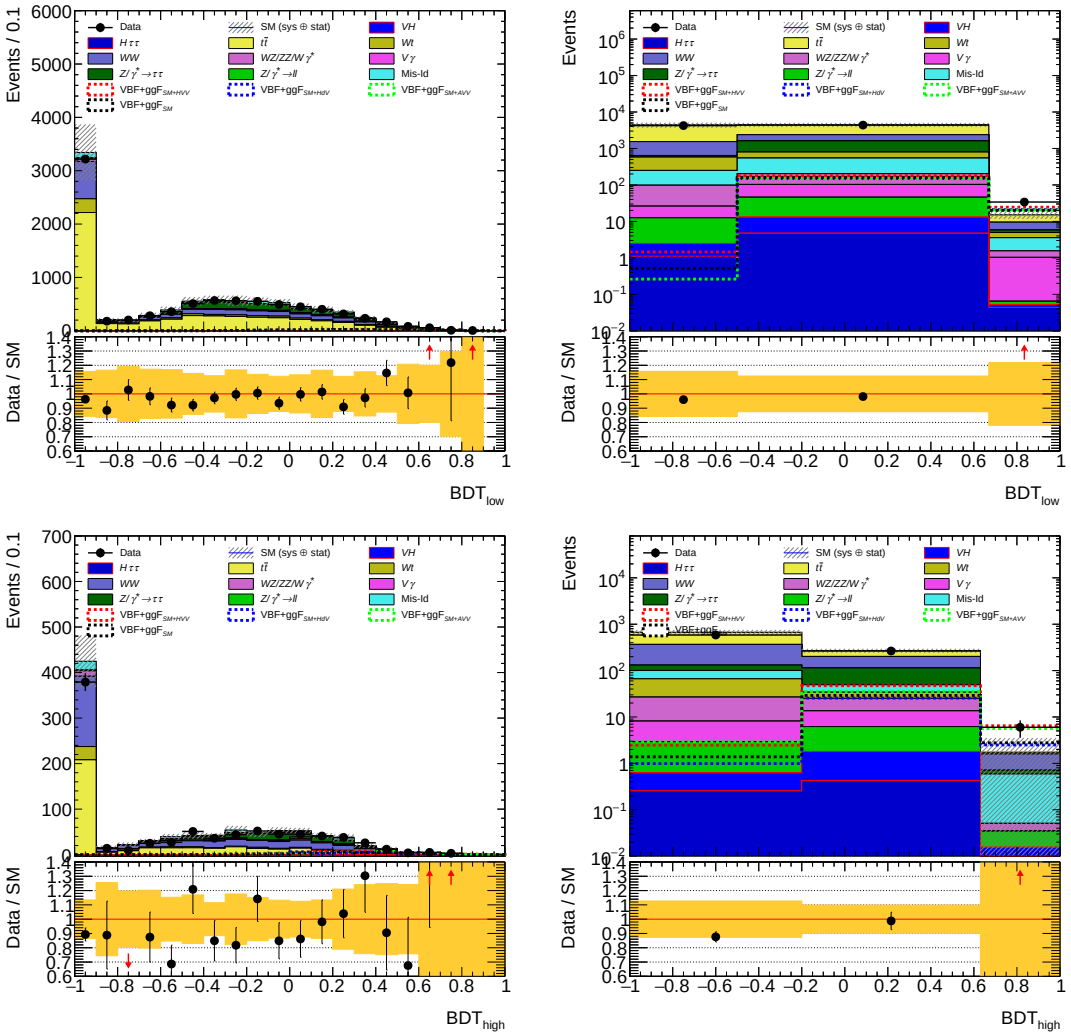


Figure 6.25.: The distributions of BDT_{low} (top) and BDT_{high} (bottom) in the respective signal regions with flat bins (left) and bins showing validation and final signal regions in logarithmic scale (right).

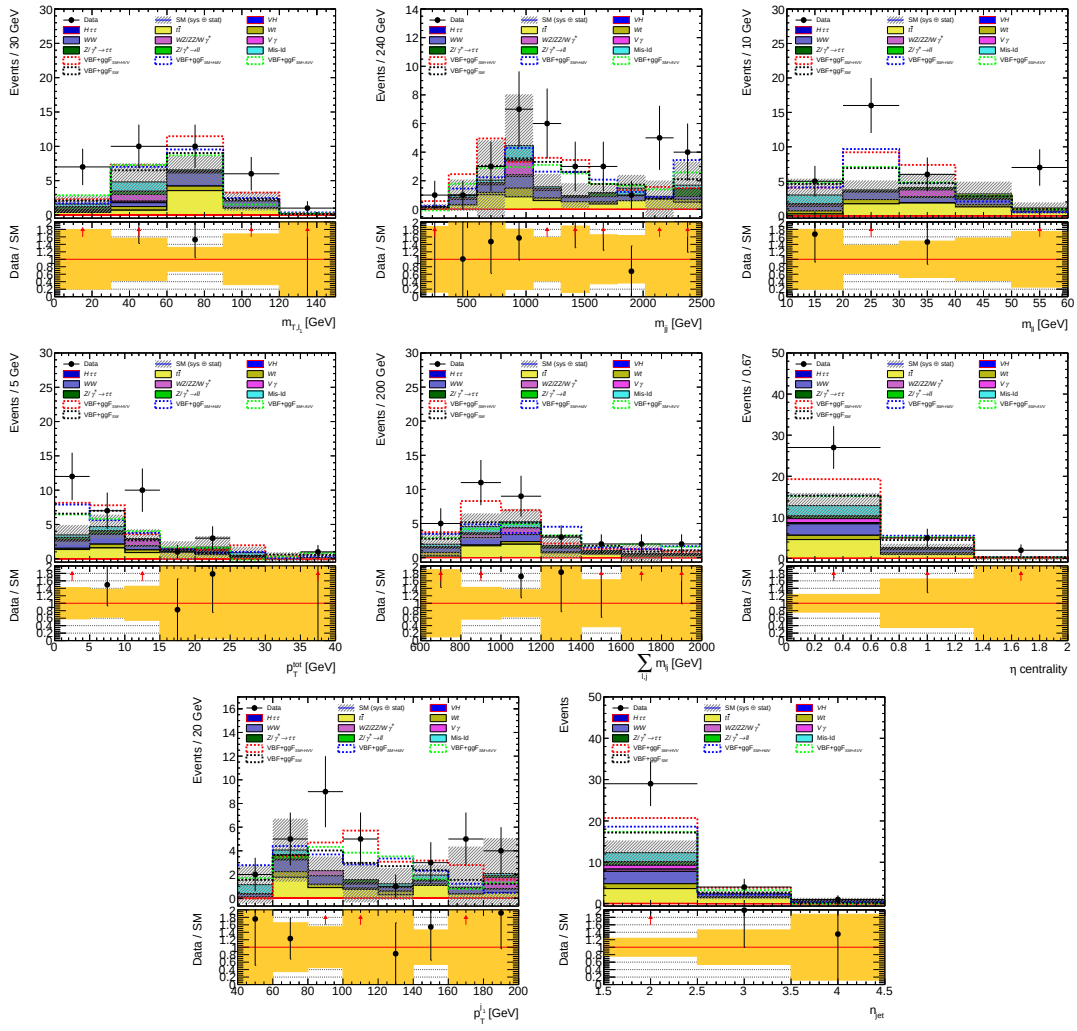


Figure 6.26.: Distributions of the $\mathrm{BDT_{low}}$ input variables in the low-p_{T} signal region: $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\mathrm{tot}}$, $\sum_{\ell,j} m_{\ell j}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{j_1}$ and n_{jet}.

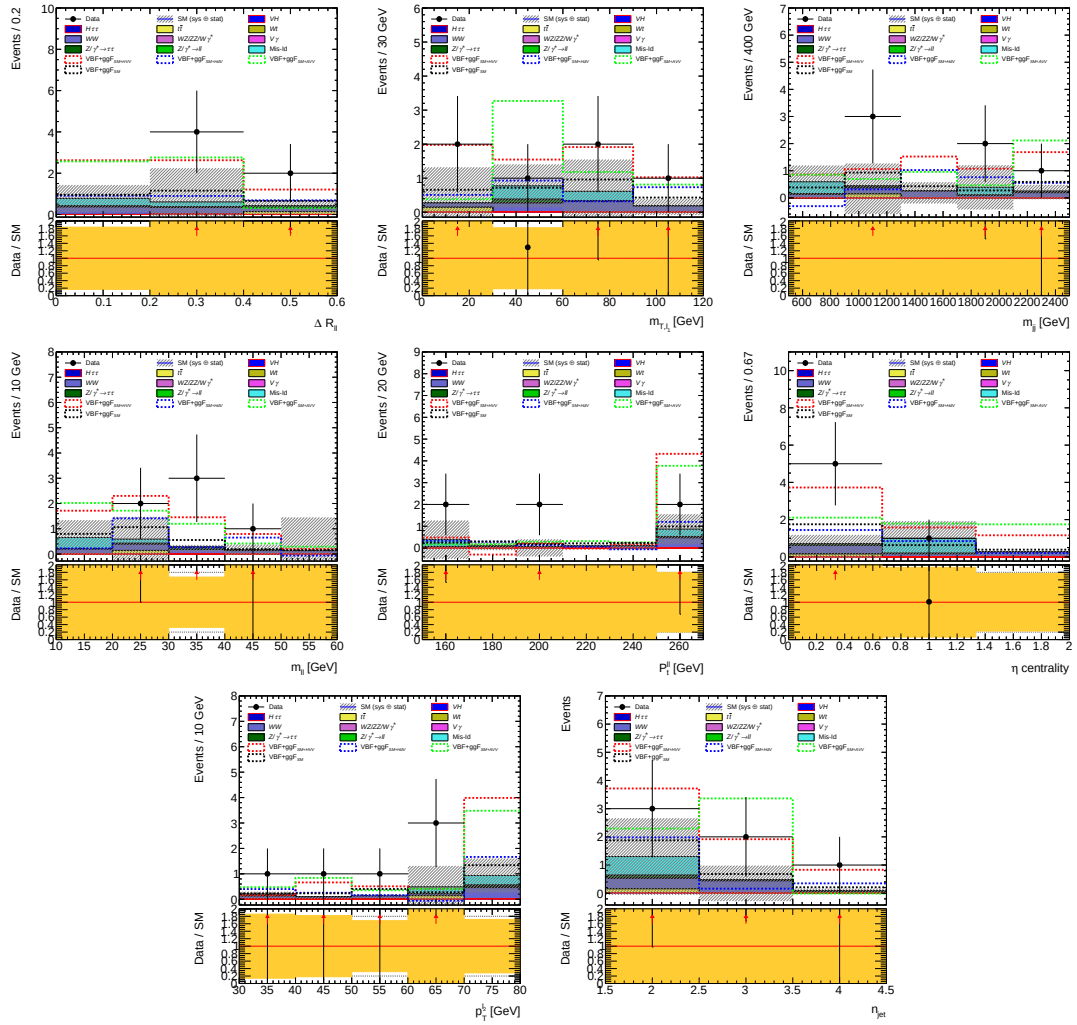


Figure 6.27.: Distributions of the $\mathrm{BDT_{high}}$ input variables in the high-p_{T} signal region: $\Delta R_{\ell\ell}$, $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{\ell_2}$ and n_{jet}.

Figure 6.28.: Correlation plots of BDT training variables for the BDT validation regions of the low-p_{T} (top) and high-pT signal regions. Distributions of $\langle X_i \rangle$ vs. X_j (as well as $\langle X_j \rangle$ vs. X_i) are shown for each pair of training variable in the BDTs as well as the correlation of each training variable with the BDT scores. The black points show the data profiles, the red points show the profiles of the MC model. The uncertainties are statistical only. The data to MC agreement is given as p-value of a χ^2 test.

6.7 Background Estimation

The diverse spectrum of background processes impacting the analysis are handled in different ways. Many processes with small effect on the final signal regions are modelled sufficiently well by the MC generator and theoretical cross section predictions. On the other hand, background, which contribute significantly are normalized by control regions if possible or derived by a data-driven approach, if the MC predictions are not reliable. Separate normalization factors are estimated in the high- and low-p_{T} categories.

After the final selection the main backgrounds contributing in the low-p_{T} region are top, WW, events with misidentified leptons and $Z{\to}\tau\tau$ events as seen in Figure 6.20. For the top and $Z{\to}\tau\tau$ events dedicated control regions are defined, whereas for the WW background no sufficiently pure control region can be constructed to receive a reliable normalization factor. Therefore, this background is modelled by the MC prediction. The contributions from $Z \to ee$ and $Z \to \mu\mu$ are much smaller due to the requirement of two different flavor 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. The misidentified leptons are estimated in a data-driven approach.

In the high-p_{T} signal region the WW background and misidentified lepton processes are mostly contributing. The same control regions and data-driven approaches as in the low-p_{T} region are constructed. The normalization factors are mainly useful to correct the dedicated background processes in validation plots, e.g. the BDT validation regions, but also contribute to correct the event yields and lower systematic uncertainties from these backgrounds in the signal region.

In the following, the control regions and the handling of events with misidentified lepton are described.

6.7.1 Control Regions

Top Control Regions

The top control regions are defined exactly as the signal regions, except that the b-jet veto is replaced by requiring one and only one b-tagged jet ($n_{b\text{-jet}} = 1$). The reason why exactly one b-jet is required instead of an inclusive b-tagged region is to bring the flavor composition of tag jets closer to the one in the b-vetoed signal regions. The $m_{\tau\tau}$ distributions before applying the $m_{\tau\tau}$ veto is shown in Figure 6.29.

Cuts on the BDT scores shown in Figure 6.30 are applied with lower thresholds compared to the signal regions to increase the top statistics in the final control regions. The respective input variables and the final BDT distributions are shown

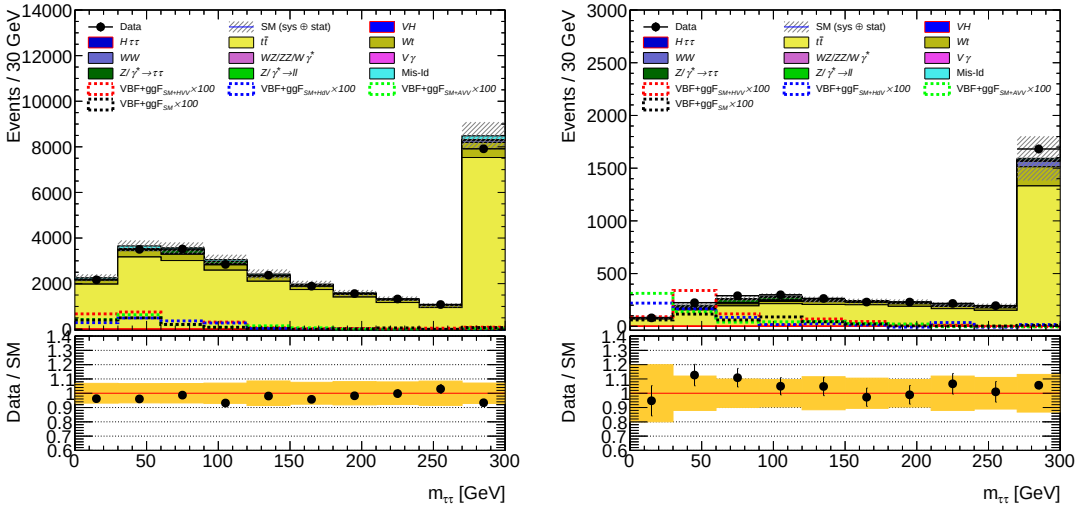


Figure 6.29.: The $m_{\tau\tau}$ distributions before applying respective requirements in the top control region in the low- (left) and high- (right) p_{T} regions. The normalization factors of Table 6.10 are applied. The yellow band represents MC statistics and all detector systematics uncertainties.

in the Appendix A.2.1. No obvious mismodelling is observed, the data to MC agreement is within the uncertainties.

The event yields at the different cut stages are summarized in Table 6.11. The top purity is about 95% (86%) in the low(high)-p_{T} region.

$Z \to \tau\tau$ Control Regions

The $Z \to \tau\tau$ control region definition follows also the cut requirements of the signal regions. Only the $Z \to \tau\tau$ veto is inverted to the requirement $|m_{\tau\tau} - m_Z| \leq 25\,\mathrm{GeV}$. To achieve a higher purity additionally a $m_{\ell\ell} < 75\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ cut are applied (cf. Figure 6.31), which results in $Z \to \tau\tau$ purities of 72% (82%) in the low(high)-p_{T} region. The event yields for cut stages defining this control region are shown in Table 6.12.

Lastly, the cuts on the BDTs are applied with the same lower thresholds as in the top control regions to increase the statistics. The modelling of the training variables are shown in the Appendix A.2.2 and the final BDT distributions in Figure 6.32 which show all a decent data to MC agreement.

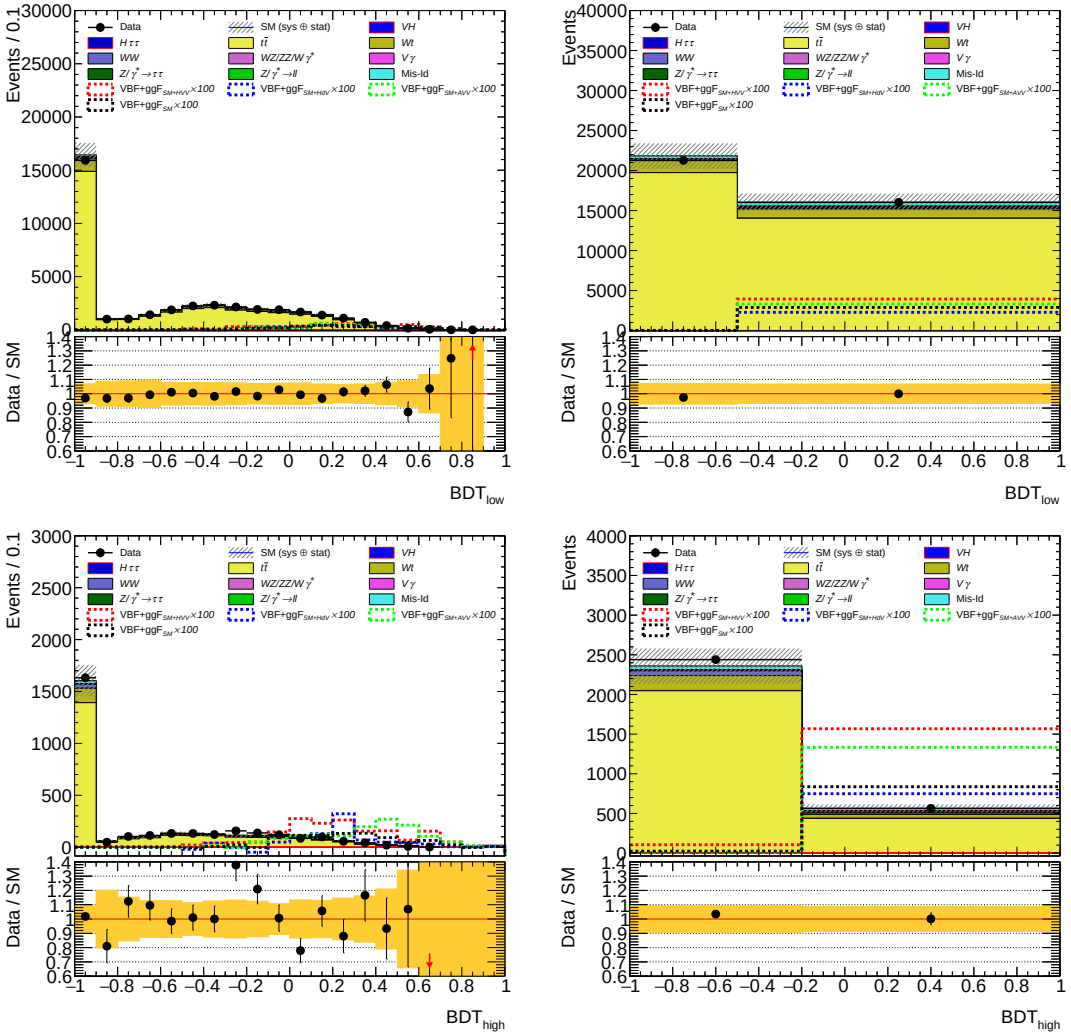


Figure 6.30.: The distributions of $\mathrm{BDT_{low}}$ (top) and $\mathrm{BDT_{high}}$ (bottom) in the respective top control regions with flat bins (left) and bins showing the top control region requirements (right). The normalization factors of Table 6.10 are applied. The yellow band represents MC statistics and all detector systematics uncertainties.

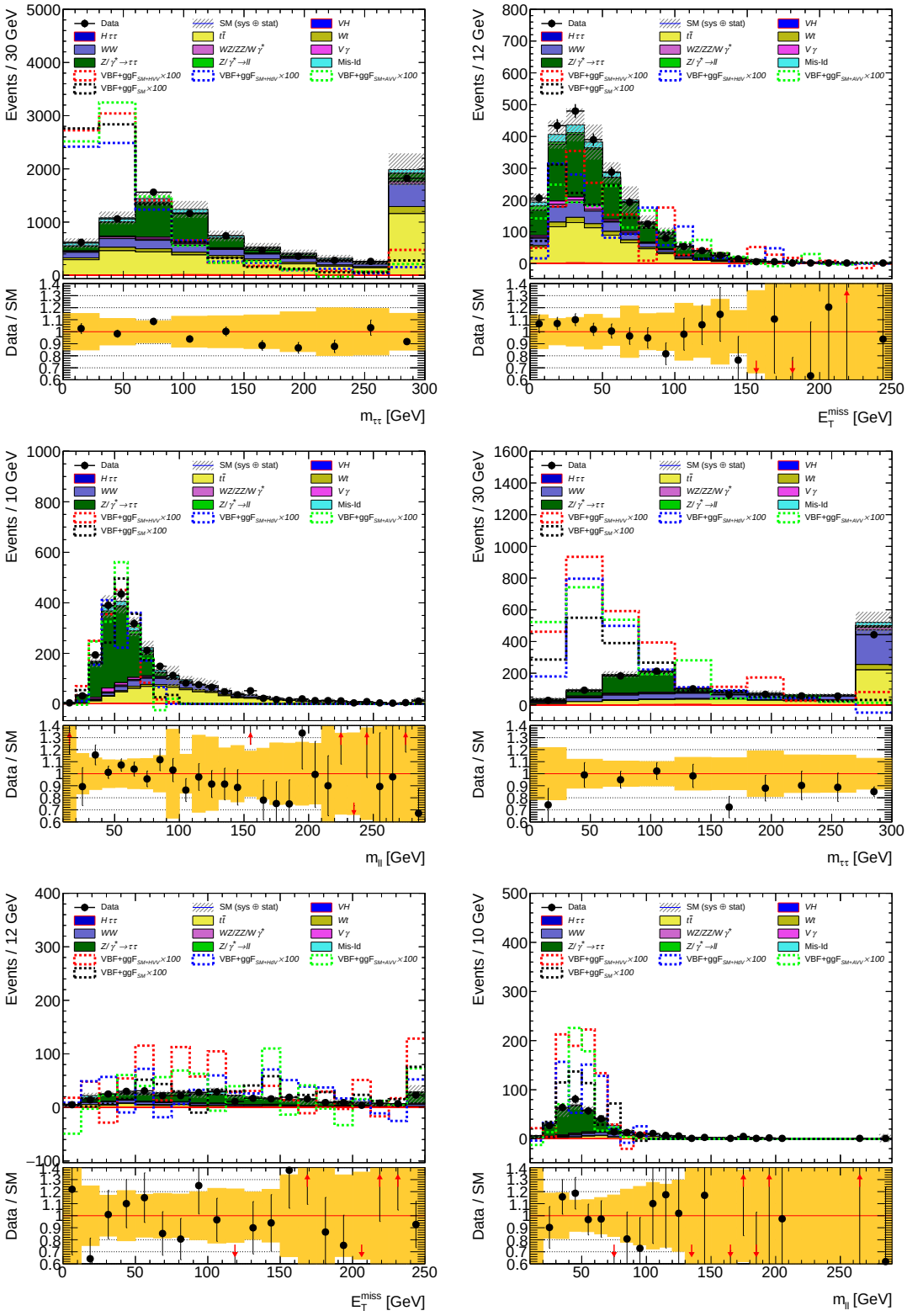


Figure 6.31.: The $m_{\tau\tau}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\ell\ell}$ distributions before applying respective requirements in the $Z \to \tau\tau$ control region in the low- (left) and high- (right) p_{T} regions. The normalization factors of Table 6.10 are applied. The yellow band represents MC statistics and all detector systematics uncertainties.

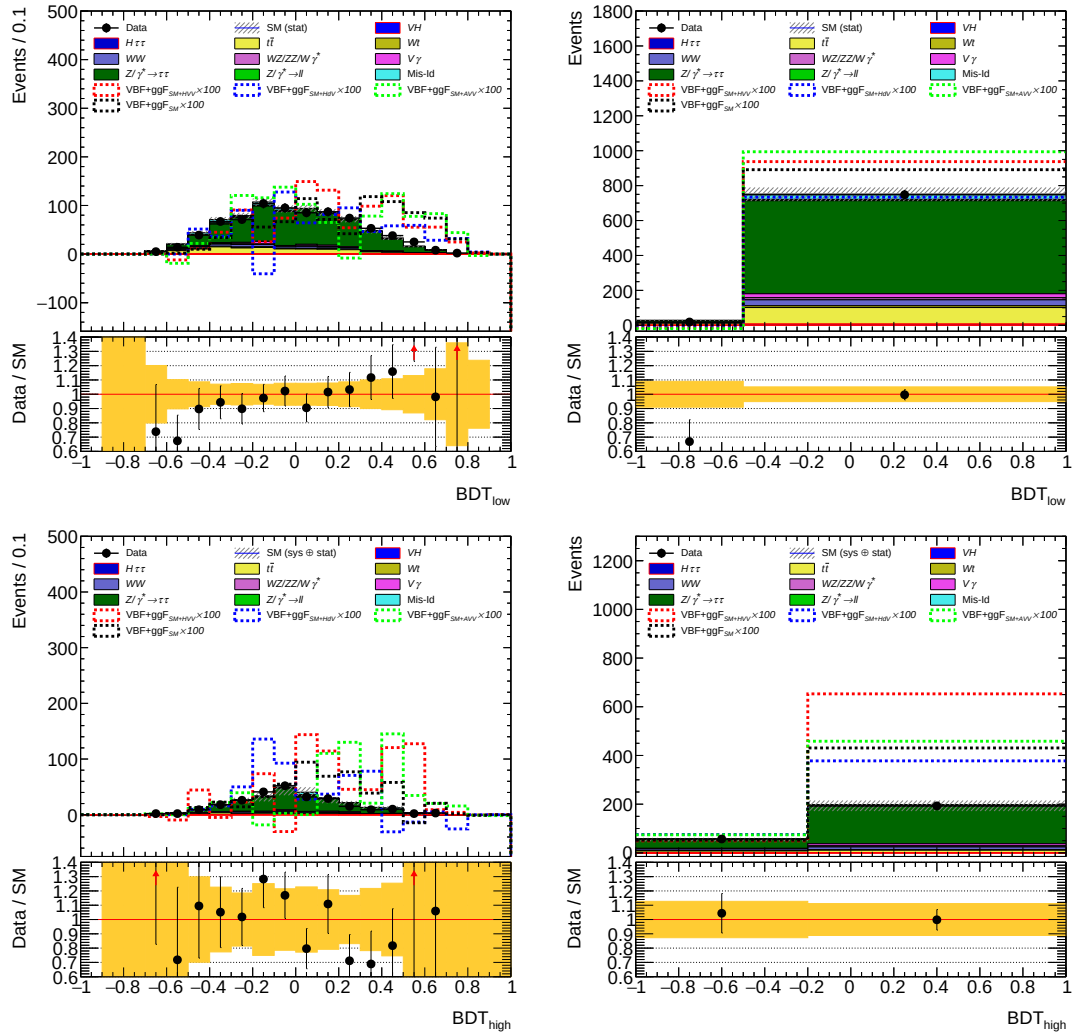


Figure 6.32.: The distributions of $\mathrm{BDT_{low}}$ (top) and $\mathrm{BDT_{high}}$ (bottom) in the respective $Z \to \tau\tau$ control regions with flat bins (left) and bins showing the $Z \to \tau\tau$ control region requirements (right). The normalization factors of Table 6.10 are applied. The yellow band represents MC statistics and all detector systematics uncertainties.

Summary

The normalization factors are derived to correct the top and $Z{\to}\tau\tau$ background in the respective signal regions. These factors are listed in Table 6.10 and are computed with the so called matrix method. It takes into account that the control regions are themselves contaminated with events of each others backgrounds to be estimated. The total number of background events from top and $Z{\to}\tau\tau$ events in the signal region $n^{\mathrm{SR}}_{\mathrm{CR-bkg}}$ can be cast in matrix notation with

$$n^{\mathrm{SR}}_{\mathrm{CR-bkg}} = \begin{pmatrix} n^{\mathrm{SR}}_{\mathrm{top}} & n^{\mathrm{SR}}_{Z\to\tau\tau} \end{pmatrix} \cdot \begin{pmatrix} \beta_{\mathrm{top}} \\ \beta_{Z\to\tau\tau} \end{pmatrix}. \tag{6.7.1}$$

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

$$\begin{pmatrix} n^{\mathrm{top\,CR}}_{\mathrm{obs.}} \\ n^{Z\to\tau\tau\,\mathrm{CR}}_{\mathrm{obs.}} \end{pmatrix} = \begin{pmatrix} n^{\mathrm{top\,CR}}_{\mathrm{top}} & n^{\mathrm{top\,CR}}_{Z\to\tau\tau} \\ n^{Z\to\tau\tau\,\mathrm{CR}}_{\mathrm{top}} & n^{Z\to\tau\tau\,\mathrm{CR}}_{Z\to\tau\tau} \end{pmatrix} \cdot \begin{pmatrix} \beta_{\mathrm{top}} \\ \beta_{Z\to\tau\tau} \end{pmatrix}, \tag{6.7.2}$$

such that

$$\begin{pmatrix} \beta_{\mathrm{top}} \\ \beta_{Z\to\tau\tau} \end{pmatrix} = \begin{pmatrix} n^{\mathrm{top\,CR}}_{\mathrm{top}} & n^{\mathrm{top\,CR}}_{Z\to\tau\tau} \\ n^{Z\to\tau\tau\,\mathrm{CR}}_{\mathrm{top}} & n^{Z\to\tau\tau\,\mathrm{CR}}_{Z\to\tau\tau} \end{pmatrix}^{-1} \cdot \begin{pmatrix} n^{\mathrm{top\,CR}}_{\mathrm{obs.}} \\ n^{Z\to\tau\tau\,\mathrm{CR}}_{\mathrm{obs.}} \end{pmatrix}. \tag{6.7.3}$$

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 deviation of the resulting distributions in the normalization factors β are then used for further calculations. These statistical uncertainties are shown in Table 6.10. Similarly, systematic uncertainties can be propagated through the method by evaluating the method for every systematic uncertainty separately and comparing the results. Nevertheless, for the final statistical derivations the normalization factors are determined by the likelihood fit, where all control regions are included. Thereby all correlations of the systematic uncertainties are taken into account in a proper way.

Table 6.10.: Background normalization factors obtained from the control regions using the methods explained in the text, including statistical uncertainties.

Control regions	Top	$Z\to\tau\tau$
low-p_{T}	1.034 ± 0.009	0.742 ± 0.050
high-p_{T}	0.869 ± 0.052	0.802 ± 0.074

Table 6.11.: Expected event yields in the top control regions for the different signal and background processes as predicted by the SM in comparison with the corresponding observed values for the selection stages. The normalization factors from 6.10 have been applied.

	VBF	ggF	VH	$H\tau\tau$	$t\bar{t}$	Wt	WW	$WZ/ZZ/W\gamma^*$	$V\gamma$	$Z/\gamma^* \to \tau\tau$	$Z/\gamma^* \to \ell\ell$	Mis-Id	Total Bkg.	Data/Bkg.	Data
$n_{b\text{-jet}} = 1$	9.08 ± 0.54	24.23 ± 1.88	3.79 ± 0.15	5.88 ± 0.21	53610.41 ± 360.33	4463.30 ± 39.59	465.89 ± 3.32	69.16 ± 3.62	35.86 ± 5.85	514.25 ± 24.10	18.12 ± 3.65	1356.19 ± 43.59	60542.84 ± 366.00	0.98 ± 0.01	59057
$Z\to\tau\tau$ veto	8.20 ± 0.50	20.92 ± 1.73	2.77 ± 0.13	0.76 ± 0.07	33772.69 ± 232.35	2697.14 ± 27.33	277.11 ± 2.61	39.81 ± 2.81	16.56 ± 3.04	186.48 ± 11.07	9.27 ± 1.81	846.38 ± 34.73	37848.99 ± 236.83	0.99 ± 0.01	37289
$\text{BDT}_{\text{low}} > -0.5$	8.20 ± 0.50	20.75 ± 1.73	2.40 ± 0.13	0.59 ± 0.06	14035.67 ± 104.29	1156.83 ± 15.59	121.68 ± 1.58	24.55 ± 2.24	13.38 ± 2.85	170.60 ± 10.49	6.67 ± 1.52	495.80 ± 23.97	16028.19 ± 108.72	1.00 ± 0.01	16025
$n_{b\text{-jet}} = 1$	2.12 ± 0.23	8.71 ± 1.00	1.08 ± 0.07	1.15 ± 0.09	5106.46 ± 218.92	573.86 ± 25.97	182.54 ± 1.44	21.28 ± 1.84	9.65 ± 2.40	96.86 ± 6.46	7.84 ± 1.48	135.12 ± 23.23	6135.84 ± 221.80	1.04 ± 0.04	6358
$Z\to\tau\tau$ veto	1.75 ± 0.22	6.86 ± 0.83	0.76 ± 0.07	0.14 ± 0.03	2483.39 ± 107.65	240.43 ± 11.74	80.79 ± 0.93	8.74 ± 1.16	2.45 ± 0.71	28.35 ± 2.30	1.69 ± 0.34	74.42 ± 15.89	2921.16 ± 109.49	1.03 ± 0.04	3003
$\text{BDT}_{\text{high}} > -0.2$	1.70 ± 0.21	6.66 ± 0.82	0.63 ± 0.06	0.08 ± 0.03	437.05 ± 20.52	49.71 ± 3.28	21.03 ± 0.47	2.62 ± 0.58	1.37 ± 0.46	19.42 ± 1.88	0.90 ± 0.29	30.48 ± 8.12	563.28 ± 22.41	1.00 ± 0.06	564

Table 6.12.: Expected event yields in the $Z\to\tau\tau$ control regions for the different signal and background processes as predicted by the SM in comparison with the corresponding observed values for the selection stages. The normalization factors from 6.10 have been applied.

	VBF	ggF	VH	$H\tau\tau$	$t\bar{t}$	Wt	WW	$WZ/ZZ/W\gamma^*$	$V\gamma$	$Z/\gamma^* \to \tau\tau$	$Z/\gamma^* \to \ell\ell$	Mis-Id	Total Bkg.	Data/Bkg.	Data
b-veto	69.64 ± 1.24	137.83 ± 3.64	12.83 ± 0.27	38.16 ± 0.49	7940.61 ± 63.97	945.89 ± 13.73	2741.62 ± 8.25	307.37 ± 11.29	198.50 ± 12.20	2478.64 ± 109.04	102.22 ± 16.82	797.42 ± 28.78	15563.28 ± 132.77	0.97 ± 0.01	15047
$m_{\tau\tau} - m_Z \leq 25$	5.76 ± 0.34	9.74 ± 1.01	1.13 ± 0.06	13.20 ± 0.28	672.58 ± 11.78	86.13 ± 3.66	242.91 ± 2.34	38.56 ± 2.27	57.78 ± 7.05	1045.21 ± 46.38	35.50 ± 9.92	109.28 ± 12.55	2302.28 ± 51.18	1.02 ± 0.03	2354
$m_{\ell\ell} < 75\,\text{GeV}$, $E_{\text{T}}^{\text{miss}} > 40\,\text{GeV}$	3.39 ± 0.27	5.56 ± 0.72	0.52 ± 0.05	6.53 ± 0.19	106.55 ± 4.50	10.99 ± 1.24	36.65 ± 0.85	12.16 ± 1.22	22.74 ± 3.01	543.75 ± 24.09	7.61 ± 2.39	31.53 ± 8.23	779.01 ± 26.21	0.98 ± 0.05	767
$\text{BDT}_{\text{low}} > -0.5$	3.38 ± 0.27	5.53 ± 0.72	0.48 ± 0.04	6.42 ± 0.19	94.69 ± 4.21	9.93 ± 1.19	34.20 ± 0.84	12.09 ± 1.22	22.82 ± 3.01	532.86 ± 23.68	7.45 ± 2.38	29.66 ± 8.15	750.60 ± 25.75	1.00 ± 0.05	748
b-veto	8.70 ± 0.43	33.26 ± 1.70	2.67 ± 0.12	6.38 ± 0.21	794.20 ± 35.75	112.17 ± 6.09	742.39 ± 2.73	70.79 ± 3.50	38.28 ± 4.87	355.17 ± 20.84	22.01 ± 2.24	87.75 ± 16.14	2231.82 ± 45.37	0.90 ± 0.03	2015
$m_{\tau\tau} - m_Z \leq 25$	1.04 ± 0.15	4.10 ± 0.76	0.35 ± 0.04	1.86 ± 0.10	50.74 ± 3.48	6.89 ± 0.94	52.55 ± 0.76	9.97 ± 1.34	9.71 ± 1.78	213.22 ± 13.10	11.51 ± 1.72	2.40 ± 7.58	359.19 ± 15.83	0.96 ± 0.07	345
$m_{\ell\ell} < 75\,\text{GeV}$, $E_{\text{T}}^{\text{miss}} > 40\,\text{GeV}$	0.95 ± 0.14	3.88 ± 0.74	0.28 ± 0.04	1.61 ± 0.10	16.85 ± 1.77	2.56 ± 0.52	18.76 ± 0.43	6.48 ± 1.15	9.30 ± 1.74	190.33 ± 11.78	7.17 ± 1.50	-5.36 ± 6.35	247.99 ± 13.76	1.01 ± 0.08	250
$\text{BDT}_{\text{high}} > -0.2$	0.89 ± 0.14	3.42 ± 0.71	0.22 ± 0.03	1.17 ± 0.08	9.55 ± 1.22	1.49 ± 0.40	11.81 ± 0.33	4.86 ± 1.07	8.33 ± 1.67	154.17 ± 9.66	5.01 ± 1.26	-3.23 ± 5.82	193.38 ± 11.59	1.00 ± 0.09	193

6.7.2 Misidentified Lepton

Despite huge efforts to optimize the identification of physical objects, the rate of misidentification can never be decreased to zero. Although the chance of identifying a jet as a lepton is extremely small (on the level of 10^{-5} [63]), the identification of leptons is a critical problem emerging from the overwhelmingly large cross section of multijet production at hadron colliders, which causes the fraction of objects selected as leptons to be a cause of concern. Generally these misidentified objects are referred to as *fake leptons*. Not only fake leptons from the fragmentation of a quark or gluon are included into this broad category but also real leptons originating from secondary processes such as heavy- (or light-) flavor hadrons decaying to muons, or electrons originating from pair production by photons originating from bremsstrahlung as well as from initial and final state radiation, as they are no true indicators of the high-energy electroweak processes under investigation. On the contrary, a primary lepton originating from a hard electroweak scattering event is referred to as a *prompt lepton*. Both, electrons and muons, are affected by misidentification, although 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.

Significant background contribution in the $H\to W^{\pm}W^{\mp *}\to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ analysis from misidentified jets arise from the production of W-bosons in association with jets. One isolated lepton and momentum imbalance due to a neutrino result from the leptonic W-decay. Together with a second fake lepton, these objects mimic the expected final state signature. The same holds for the pure QCD background, 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. The estimates of these contributions are nevertheless closely related to one another and the same methods apply for both. The processes leading to the misidentification of a jet as a lepton are hard to model in simulation. Therefore, the W+jets and the QCD estimate are derived directly from data with the so called *fake factor method* and do not rely on MC predictions.

Fake Factor Method

The signal regions have two leptons passing the lepton identification selection, referred to as *id*-leptons (✓). The total number of events in these selected signal regions can be written as

$$N_{e_{\checkmark}\mu_{\checkmark}} = N^{e_p\mu_p}_{e_{\checkmark}\mu_{\checkmark}} + N^{f}_{e_{\checkmark}\mu_{\checkmark}} \quad \text{with} \quad N^{f}_{e_{\checkmark}\mu_{\checkmark}} = N^{e_f\mu_p}_{e_{\checkmark}\mu_{\checkmark}} + N^{e_p\mu_f}_{e_{\checkmark}\mu_{\checkmark}} + N^{e_f\mu_f}_{e_{\checkmark}\mu_{\checkmark}}, \tag{6.7.4}$$

where the superscript labels the true nature of the leptons with p being a prompt lepton and f a fake lepton and the subscripts indicates if the lepton passes the identification requirements. The sub- and superscripts are ordered by flavor of the leptons. $N^{e_p\mu_p}_{e\checkmark\mu\checkmark}$ term denotes the events yield of all processes with two real prompt leptons including the signal, and $N^{f}_{e\checkmark\mu\checkmark}$ labels the number of events with at least one fake lepton. It is composed of single-fake background from W+jets processes ($N^{e_f\mu_p/e_p\mu_f}_{e\checkmark\mu\checkmark}$) and double-fake backgrounds ($N^{e_f\mu_f}_{e\checkmark\mu\checkmark}$) caused by the multijet or QCD background. While the event yield $N^{e_p\mu_p}_{e\checkmark\mu\checkmark}$ is estimated by the predictions of the respective MC samples, the fake contribution $N^{f}_{e\checkmark\mu\checkmark}$ needs to be determined by a data-driven approach. In order to estimate this background contribution from fake leptons, a control sample is defined with an enhanced rate of such fake leptons. Events in the control sample are required to have one leptons 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 in Section 6.3.1 described identified categories are listed in Table 6.13 for both, muons and electrons. The isolation requirements 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.

Table 6.13.: The requirements for fully identified and anti-identified electrons (left) and muons (right).

Id electron	Anti-id electron	Id muon	Anti-id muon
$p_{\mathrm{T}} > 15\,\mathrm{GeV}$		$p_{\mathrm{T}} > 15\,\mathrm{GeV}$	
$\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$		$\lvert\eta\rvert < 2.45$	
$\lvert z_0 \sin\theta\rvert < 0.5\,\mathrm{mm}$		$\lvert z_0 \sin\theta\rvert < 0.5\,\mathrm{mm}$	
`TightLH` if $p_{\mathrm{T}} < 25\,\mathrm{GeV}$ `MediumLH` if $p_{\mathrm{T}} > 25\mathrm{GeV}$	`LooseLH`	`Tight`	`Medium`
$\lvert d_0\rvert/\sigma_{d_0} < 5$		$\lvert d_0\rvert/\sigma_{d_0} < 3$	$\lvert d_0\rvert/\sigma_{d_0} < 15$
$p_T^{\mathrm{varcone0.4}}/p_{\mathrm{T}} < 0.06$ if $p_{\mathrm{T}} < 25\,\mathrm{GeV}$ `Gradient` isolation if $p_{\mathrm{T}} > 25\,\mathrm{GeV}$		$E_T^{\mathrm{cone0.2}}/p_{\mathrm{T}} < 0.09$ $p_T^{\mathrm{varcone0.3}}/p_{\mathrm{T}} < 0.06$	
exclusively identified as electrons			
	Veto against identified electron		Veto against identified muon

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

$$N_{e\checkmark\mu\text{✗}} = N^{e_p\mu_p}_{e\checkmark\mu\text{✗}} + N^{e_f\mu_p}_{e\checkmark\mu\text{✗}} + N^{e_p\mu_f}_{e\checkmark\mu\text{✗}} + N^{e_f\mu_f}_{e\checkmark\mu\text{✗}}\,, \tag{6.7.5}$$

for the muon passing the anti-identification criteria and

$$N_{e\text{✗}\mu\text{✓}} = N^{e_p\mu_p}_{e\text{✗}\mu\text{✓}} + N^{e_f\mu_p}_{e\text{✗}\mu\text{✓}} + N^{e_p\mu_f}_{e\text{✗}\mu\text{✓}} + N^{e_f\mu_f}_{e\text{✗}\mu\text{✓}}, \tag{6.7.6}$$

for anti-identified electrons. The components $N^{e_p\mu_p}_{e\text{✓}\mu\text{✗}}$ and $N^{e_p\mu_p}_{e\text{✗}\mu\text{✓}}$ denote processes with two real leptons which contaminate the control samples. They are evaluated using MC predictions and later subtracted from the data event yields $N_{e\text{✓}\mu\text{✗}}$ and $N_{e\text{✗}\mu\text{✓}}$, respectively, to obtain the number of events from fake backgrounds only. The terms $N^{e_f\mu_p}_{e\text{✓}\mu\text{✗}}$ and $N^{e_p\mu_f}_{e\text{✗}\mu\text{✓}}$ 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 QCD double-fake contributions $N^{e_f\mu_f}_{e\text{✓}\mu\text{✗}}$ and $N^{e_f\mu_f}_{e\text{✗}\mu\text{✓}}$ (one fake object is identified as a lepton and one fake object is anti-identified as a lepton) are typically relatively small compared to the W+jets single-fake contribution.

The extrapolation of the fake contamination in the signal region is estimated by scaling the number of events in the control sample by the fake factors $f^{\ell\text{✓}}_{\ell\text{✗}}$, for electrons ($\ell = e$) and muons ($\ell = \mu$) separately, defined as the ratio between fully identified leptons $N_{\ell\text{✓}}$ and anti-identified leptons $N_{\ell\text{✗}}$:

$$f^{\ell\text{✓}}_{\ell\text{✗}} = \frac{N_{\ell\text{✓}}}{N_{\ell\text{✗}}}. \tag{6.7.7}$$

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

The fake contribution in the signal sample can be determined with

$$\begin{aligned} N^{e_f\mu_p}_{e\text{✓}\mu\text{✓}} =& f^{e\text{✓}}_{e\text{✗}} \times N^{e_f\mu_p}_{e\text{✗}\mu\text{✓}} \\ N^{e_p\mu_f}_{e\text{✓}\mu\text{✓}} =& f^{\mu\text{✓}}_{\mu\text{✗}} \times N^{e_p\mu_f}_{e\text{✓}\mu\text{✗}}. \end{aligned} \tag{6.7.8}$$

The full fake background yield is then calculated

$$\begin{aligned} N^{f}_{e\text{✓}\mu\text{✓}} =& N^{e_f\mu_p}_{e\text{✓}\mu\text{✓}} + N^{e_p\mu_f}_{e\text{✓}\mu\text{✓}} + N^{e_f\mu_f}_{e\text{✓}\mu\text{✓}} \\ =& f^{e\text{✓}}_{e\text{✗}} \times \left(N_{e\text{✗}\mu\text{✓}} - N^{e_p\mu_p}_{e\text{✗}\mu\text{✓}} - N^{e_f\mu_f}_{e\text{✗}\mu\text{✓}}\right) + f^{\mu\text{✓}}_{\mu\text{✗}} \times \left(N_{e\text{✓}\mu\text{✗}} - N^{e_p\mu_p}_{e\text{✓}\mu\text{✗}} - N^{e_f\mu_f}_{e\text{✓}\mu\text{✗}}\right) + N^{e_f\mu_f}_{e\text{✓}\mu\text{✓}} \\ =& f^{e\text{✓}}_{e\text{✗}} \times \left(N_{e\text{✗}\mu\text{✓}} - N^{e_p\mu_p}_{e\text{✗}\mu\text{✓}}\right) + f^{\mu\text{✓}}_{\mu\text{✗}} \times \left(N_{e\text{✓}\mu\text{✗}} - N^{e_p\mu_p}_{e\text{✓}\mu\text{✗}}\right) \\ & + N^{e_f\mu_f}_{e\text{✓}\mu\text{✓}} - f^{e\text{✓}}_{e\text{✗}} \times N^{e_f\mu_f}_{e\text{✗}\mu\text{✓}} - f^{\mu\text{✓}}_{\mu\text{✗}} \times N^{e_f\mu_f}_{e\text{✓}\mu\text{✗}}. \end{aligned} \tag{6.7.9}$$

Apparently by design the QCD contribution is double subtracted and needs to be corrected. For that, an additional set of fake factors $\tilde{f}^{\ell\text{✓}}_{\ell\text{✗}}$ is estimated from

a double-fake enriched samples, to on the one hand, expand the QCD estimate in the id+anti-id region,

$$\begin{aligned} N^{e_f\mu_f}_{e_{\text{✗}}\mu_{\text{✗}}} &= \tilde{f}^{\mu_{\text{✓}}}_{\mu_{\text{✗}}} \times N^{e_f\mu_f}_{e_{\text{✗}}\mu_{\text{✓}}} \\ N^{e_f\mu_f}_{e_{\text{✗}}\mu_{\text{✗}}} &= \tilde{f}^{e_{\text{✓}}}_{e_{\text{✗}}} \times N^{e_f\mu_f}_{e_{\text{✓}}\mu_{\text{✗}}}, \end{aligned} \tag{6.7.10}$$

and on the other hand, to calculate a correction factor by extrapolating the double-fake background contribution in the signal region with

$$N^{e_f\mu_f}_{e_{\text{✓}}\mu_{\text{✓}}} = \left(f^{e_{\text{✓}}}_{e_{\text{✗}}} \tilde{f}^{\mu_{\text{✓}}}_{\mu_{\text{✗}}} + f^{\mu_{\text{✓}}}_{\mu_{\text{✗}}} \tilde{f}^{e_{\text{✓}}}_{e_{\text{✗}}} - \tilde{f}^{e_{\text{✓}}}_{e_{\text{✗}}} \tilde{f}^{\mu_{\text{✓}}}_{\mu_{\text{✗}}} \right) \times N^{e_f\mu_f}_{e_{\text{✗}}\mu_{\text{✗}}}. \tag{6.7.11}$$

The double-fake control region estimate is calculated similar to the single-fake control regions:

$$N^{e_f\mu_f}_{e_{\text{✗}}\mu_{\text{✗}}} = N_{e_{\text{✗}}\mu_{\text{✗}}} - N^{e_p\mu_p}_{e_{\text{✗}}\mu_{\text{✗}}}. \tag{6.7.12}$$

The term $N_{e_{\text{✗}}\mu_{\text{✗}}}$ is determined by data, whereas the electroweak true-lepton contamination $N^{e_p\mu_p}_{e_{\text{✗}}\mu_{\text{✗}}}$ is subtracted with the MC event yields of the corresponding samples. The small single-fake contributions are neglected.

Lastly, the final fake-background estimate in the signal region is calculated with

$$\begin{aligned} N^{f}_{e_{\text{✓}}\mu_{\text{✓}}} = \quad & f^{e_{\text{✓}}}_{e_{\text{✗}}} \times \left(N_{e_{\text{✗}}\mu_{\text{✓}}} - N^{e_p\mu_p}_{e_{\text{✗}}\mu_{\text{✓}}} \right) \\ + \quad & f^{\mu_{\text{✓}}}_{\mu_{\text{✗}}} \times \left(N_{e_{\text{✓}}\mu_{\text{✗}}} - N^{e_p\mu_p}_{e_{\text{✓}}\mu_{\text{✗}}} \right) \\ + \quad & \tilde{f}^{e_{\text{✓}}}_{e_{\text{✗}}} \tilde{f}^{\mu_{\text{✓}}}_{\mu_{\text{✗}}} \times \left(N_{e_{\text{✗}}\mu_{\text{✗}}} - N^{e_p\mu_p}_{e_{\text{✗}}\mu_{\text{✗}}} \right). \end{aligned} \tag{6.7.13}$$

Z+jets Fake Factor Estimation

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 to events with real additional leptons, such as diboson processes. Thus, such events pose an excellent data set to measure the fake factor $f^{\ell_{\text{✓}}}_{\ell_{\text{✗}}}$.

Z+jets events are selected containing exactly three loosely identified lepton candidates, all with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$. The Z-boson candidates are selected by requiring an opposite-sign ee or $\mu\mu$ lepton pair with invariant mass $70\,\mathrm{GeV} < m_{\ell\ell} < 110\,\mathrm{GeV}$. Both Z candidate leptons have to fulfill the id criteria, and at least one of them must be matched to one of the single lepton triggers listed in Table 6.3. The fake candidate is selected as the additional lepton, identified as an electron or a muon. In order to reduce the $Z + \gamma$ background in events with a fake candidate electron, the invariant mass window of the dilepton system is tightened to $80\,\mathrm{GeV} < m_{\ell\ell} < 110\,\mathrm{GeV}$. If there are two Z-boson candidates in the same event, the one with invariant mass closer to the Z mass pole is chosen. The amount of electroweak

background is reduced by vetoing an event if it contains a fake candidate lepton with $m_T = \sqrt{2p_T E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi_{\ell\nu})} > 50\,\mathrm{GeV}$. Finally, the fake candidate lepton is split into an identified and an anti-identified category according to the definitions in Table 6.13.

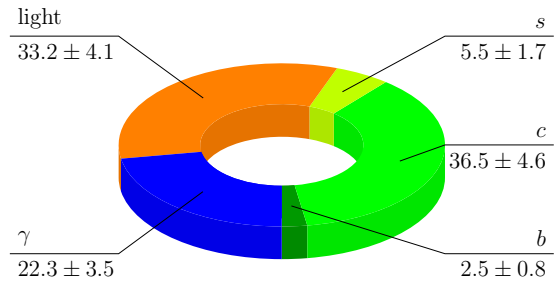
The contaminating electroweak backgrounds $V + \gamma$, diboson production processes such as WW, WZ and ZZ, as well as $t\bar{t}$ and single top production are estimated from MC simulation and subtracted from the data yields. Once the leptons in the identified and anti-identified category are selected, the fake factors are obtained from Equation 6.7.7 as a binned ratio in p_{T} using bins with boundaries [15, 20, 25, 35, 1000] ([15, 20, 25, 1000]) in GeV for electrons (muons). For electrons, the fake factor is also binned in $|\eta|$ using two bins, [0, 1.5, 2.5], excluding the crack region $1.37 < |\eta| < 1.52$, whereas for muons, there is no statistically significant difference observed between low and high pseudorapidity, thus the fake factor is merged to gain statistical precision.

Flavor Composition

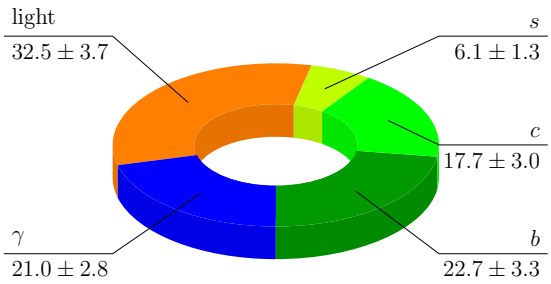
A natural complication in accurately estimating the fake background contribution is that non-prompt leptons can originate from a number of different sources whose relative abundance will vary between samples. If these fake sources have separate chances of becoming identified or anti-identified, this will in turn manifest as a discrepancy among the sample fake factors. Driven to understand this discrepancy, a flavor composition study of the fake leptons based on MC samples is performed on the W+jets and Z+jets samples. Fake leptons are selected in the W+jets samples by performing an event selection similar to the analysis selection. Additionally, the reconstructed lepton originating from the leptonically decaying W is identified by truth matching, leaving the other lepton to be classified as the fake. For the Z+jets samples, an event selection similar to the Z+jets fake factor measurement is performed, with the difference that the two Z candidate leptons are now truth matched to the Z, and the additional lepton is classified as the fake. Exemplary, the flavor composition of identified and anti-identified electrons in both samples are shown in Figure 6.33. In order to account for the difference in flavor composition between both samples when applying the fake factor, a correction factor $f_{\text{✗}}^{\text{✓}}(W\text{+jets})/f_{\text{✗}}^{\text{✓}}(Z\text{+jets})$ has been calculated using fake factors calculated from both samples. The final correction factors and corresponding systematic uncertainties which are actually applied in the analysis are then summarized in Table 6.14. They are split into low- and high-p_{T} due to their observed transverse momentum dependence.

Di-jet Fake Factor Estimation

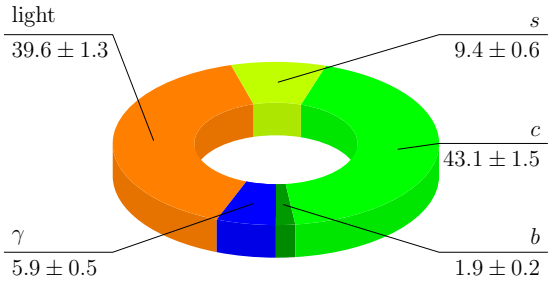
Most of the time the anti-identified lepton will be the lepton with the lower transverse momentum, and not the lepton that fired the trigger. The fake factors



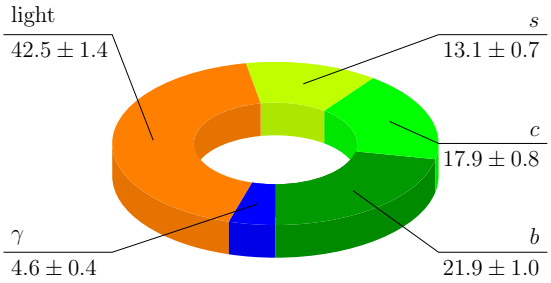
(a) Identified electrons in W+jets.



(b) Identified electrons in Z+jets.



(c) Anti-identified electrons in W+jets.



(d) Anti-identified electrons in Z+jets.

Figure 6.33.: Flavor compositions in the W+jets and Z+jets samples for identified and anti-identified electrons.

Table 6.14.: Corrections factors and corresponding systematic uncertainties. POWHEG+PYTHIA8 is used to derive the central values, while the systematic uncertainty is evaluated by comparing with ALPGEN + PYTHIA 6. The systematic uncertainty accounts for the difference in the flavor composition among samples.

	$f_{\text{✗}}^{\text{✓}}(W\text{+jets})/f_{\text{✗}}^{\text{✓}}(Z\text{+jets})$	
	$p_{\text{T}} < 25\,\text{GeV}$	$p_{\text{T}} > 25\,\text{GeV}$
electrons	0.96±0.13(stat)±0.28(syst)	1.15±0.15(stat)±0.02(syst)
muons	1.34±0.17(stat)±0.25(syst)	1.83±0.38(stat)±0.20(syst)

applying 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 a di-jet data sample and both the numerator and denominator are collected with an OR between the single lepton triggers. In the following, these fake factors are called *triggered fake factors*. In addition the nominal fake factor $\tilde{f}^{\ell\checkmark}_{\ell\times}$ from the di-jet data samples is evaluated as well and used to estimate the QCD correction factor to compensate the double counting of the double-fake event yields.

To enrich this data sample and suppress background contributions from electroweak processes the following selections are applied: At least one jet with a transverse momentum of $p_{\mathrm{T}} > 22\,\mathrm{GeV}$ is required as well as a transverse momentum of the fake lepton of at least 15 GeV. To suppress the contamination from prompt leptons from W- and Z-bosons the cut $m_T + E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$ is applied. Similar as for the Z+jets fake factor estimate the remaining electroweak background is subtracted from the MC estimation.

As the data-taking conditions changed from 2015 to 2016, different triggered 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 anti-identified leptons p_{T} and $|\eta|$. The triggered fake factors are only applied to events where the anti-identified lepton is triggered and the identified lepton is not triggered which applies only for a minority of events. The nominal fake factors are used for all other events. They are shown in Figure 6.34.

Correction Factor in High-p_{T} Signal Region

The isolation criteria of the leptons differ in the low- and high-p_{T} regions as described in Section 6.3.1 but the fake factors estimated match only precisely the lepton requirements of the low-p_{T} region (cf. Table 6.13). To correct for the different isolation criteria a correction factor is determined by comparing the W+jets MC event yields with the high- and low-p_{T} lepton isolations after the high-p_{T} requirement $p_{\mathrm{T}}^{j_1} \geq 200\,\mathrm{GeV}$. This results in a correction factor for the misidentified lepton event yield in the high-p_{T} region of 1.665. Since this is just an approximate method a 100% systematic uncertainty is applied on this correction factor, which is included in the likelihood fit calculations.

6.8 Final Selections

Table 6.15 summarizes the event selection for all signal and control regions which have been described in detail in the previous sections.

The event yields after each selection criteria of the signal regions are shown in Table 6.16. The data column labels the number of observed events, while the signal and the background processes refer to the MC prediction normalized to the

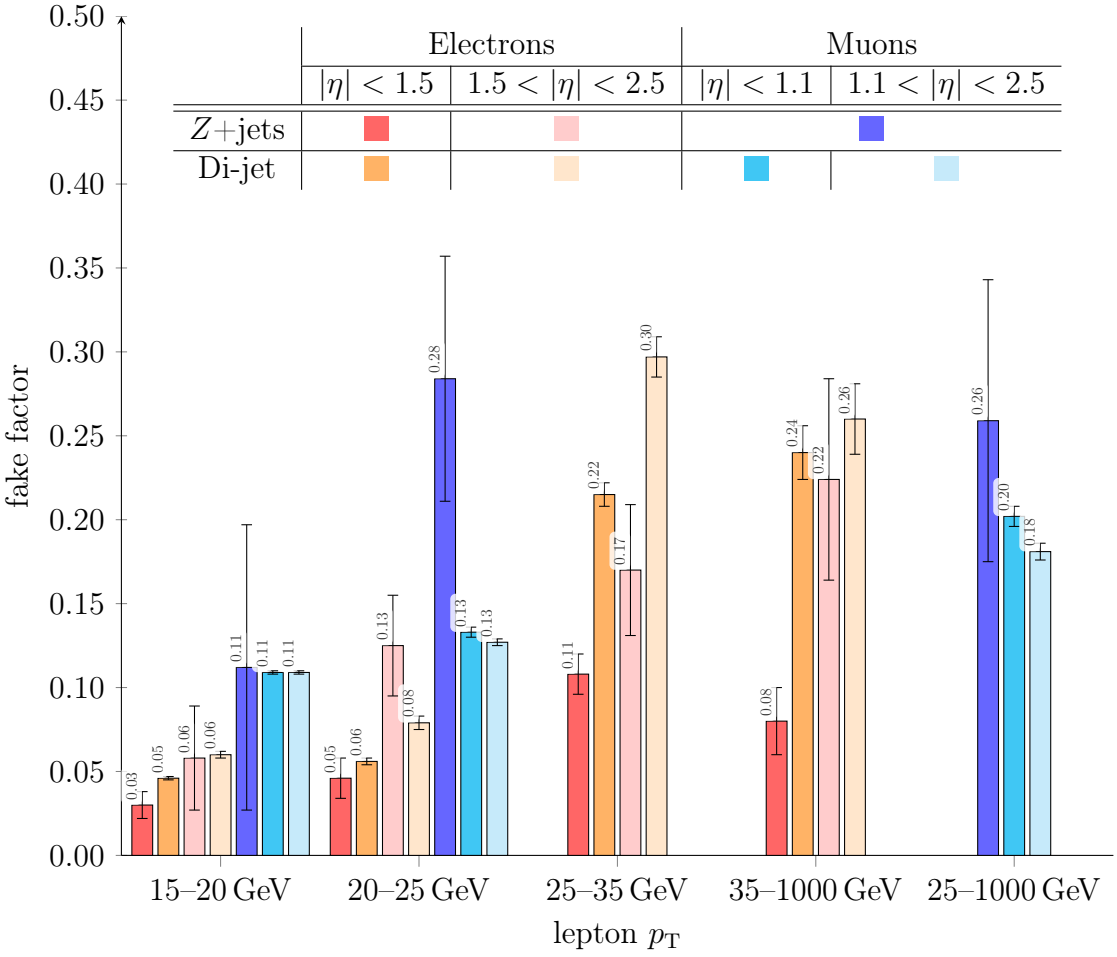


Figure 6.34.: Results of the fake factor calculations including statistical uncertainties.

integrated luminosity of the data. The normalization factors of the control regions listed in Table 6.10 are applied starting at the signal region split by the leading jet transverse momentum requirement.

Taken all together, the event yields of the final signal and control regions are listed in Table 6.17. Here, the contributions from VH and $H \to \tau\tau$ production are treated as background processes, the cross sections, branching ratios and the corresponding acceptances are assumed to have the nominal SM values and are mostly negligible.

The relative contributions of each background and SM signal in the final signal and control regions are summarized in Figure 6.35. As desired VBF events are dominant in both signal regions already at SM parameters. In the high-p_{T} region the SM VBF contribution is slightly decreased with more ggF events present than in the low-p_{T} region. But most of the benchmark VBF EFT samples (cf. Table 6.4) which have been used in the BDT trainings show especially in the high-p_{T} region

Table 6.15.: Event selection criteria used to define the signal and control regions.

	SRs	$Z \to \tau\tau$ CRs	Top CRs
Pre-selection	Two isolated leptons (one electron, one muon) with opposite charge $p_{\mathrm{T}}^{\ell_1} > 22\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\ell_2} > 15\,\mathrm{GeV}$ $m_{\ell\ell} > 10\,\mathrm{GeV}$, $n_{\mathrm{jet}} \geq 2$, $m_{jj} > 120\,\mathrm{GeV}$		
All regions are split in low-p_{T} ($p_{\mathrm{T}}^{j_1} < 200\,\mathrm{GeV}$) and high-$p_{\mathrm{T}}$ ($p_{\mathrm{T}}^{j_1} \geq 200\,\mathrm{GeV}$)			
Selection for CR definition	$n_{b\text{-jet}} = 0$ $m_{\tau\tau} < m_Z - 25\,\mathrm{GeV}$	$n_{b\text{-jet}} = 0$ $\lvert m_{\tau\tau} - m_Z \rvert \leq 25\,\mathrm{GeV}$ $m_{\ell\ell} < 75\,\mathrm{GeV}$, $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,GeV$	$n_{b\text{-jet}} = 1$ $m_{\tau\tau} < m_Z - 25\,\mathrm{GeV}$
Two BDTs are applied at this level, separate for low- and high-p_{T} events eight discriminant variables are used: for low-p_{T}: $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\mathrm{tot}}$, $\sum_{\ell,j} m_{\ell j}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{j_1}$ and n_{jet} for high-p_{T}: $\Delta R_{\ell\ell}$, $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{\ell_2}$ and n_{jet}			
BDT selection	$\mathrm{BDT}_{\mathrm{low}} > 0.669$ $\mathrm{BDT}_{\mathrm{high}} > 0.631$	$\mathrm{BDT}_{\mathrm{low}} > -0.5$ $\mathrm{BDT}_{\mathrm{high}} > -0.2$	$\mathrm{BDT}_{\mathrm{low}} > -0.5$ $\mathrm{BDT}_{\mathrm{high}} > -0.2$

much larger event yields. Additionally, some variation is observed in the low-p_{T} region which furthermore enhances the discrimination power between different EFT parameters. For these MC samples the event yields are listed in Table 6.18. The purity of the top and $Z \to \tau\tau$ control regions are as well visible in Figure 6.35.

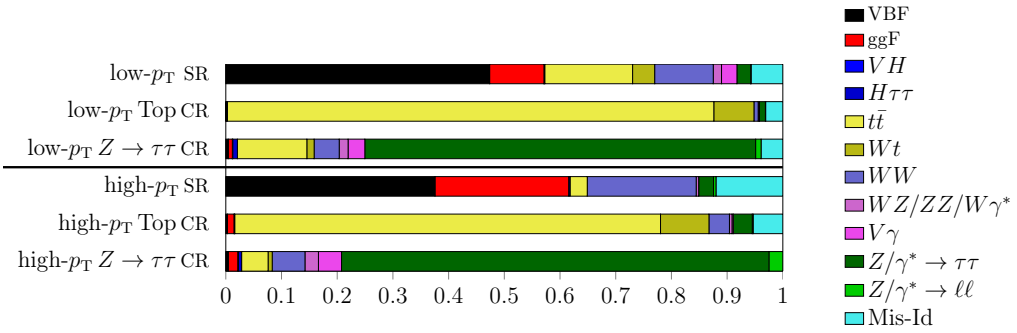


Figure 6.35.: Relative contributions of each background and the SM signals in the final regions.

As argued in Section 6.1 a particularly good choice for a variable to distinguish different EFT contributions to the $H \to VV$ vertices appears to be $\Delta\phi_{jj,\mathrm{sign}}$, which is therefore used as observable in the final fit in the low-p_{T} regions. The pre-fit distributions are shown in Figure 6.36. The data yields are increased as already observed in previous histograms. The shape differences between the EFT scenarios follow the trends as shown in the generator level distributions of Figure 6.4. The SM+HVV distribution has increased yields in at $\Delta\phi_{jj,\mathrm{sign}} \sim \pi$, a positive contribution of the AVV parameter results in a peak at $\sim 3\pi/2$, and the H∂V operator follows mostly the SM shape. At the pre-fit stage, the data seems to follow mostly the

SM+HVV shape, but due to the large uncertainties this tendency is fully compatible with the pure SM case. The small shape differences compared to the available data statistics foresee only a marginal effect of the $\Delta\phi_{jj,\text{sign}}$ shape in the fit. Still, it is included. The statistical treatment including the likelihood fit is described in Section 6.10.

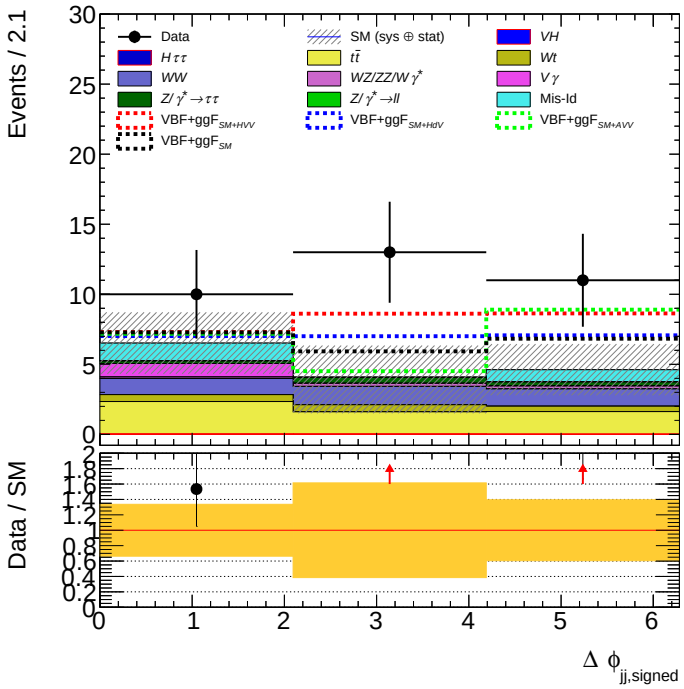


Figure 6.36.: The $\Delta\phi_{jj,\text{sign}}$ pre-fit distributions for the low-pT signal region.

6.9 Systematic Uncertainties

Several sources of systematic uncertainties affect the analysis. They are classified in theoretical uncertainties originating from the MC generation of the background and signal processes and the ones associated to the experimental measurements. The latter are associated with the reconstruction, identification and trigger efficiencies and the scale and resolution of energy and momentum.

Table 6.16.: Expected event yields in the signal regions for the different signal and background processes as predicted by the SM in comparison with the corresponding observed values for the selection stages. The normalization factors from 6.10 have been applied after the m_{jj} requirement.

	VBF	ggF	VH	$H\tau\tau$	$t\bar{t}$	Wt	WW	$WZ/ZZ/W\gamma^*$	$V\gamma$	$Z/\gamma^* \to \tau\tau$	$Z/\gamma^* \to \ell\ell$	Mis-Id	Total Bkg.	Data/Bkg.	Data
Lepton selection	161.06 ± 1.96	1678.95 ± 10.09	121.24 ± 0.81	370.05 ± 1.77	282558.99 ± 235.78	26647.44 ± 66.15	35385.04 ± 50.30	2662.09 ± 23.92	2632.32 ± 56.61	82186.52 ± 209.90	2262.63 ± 183.99	22748.62 ± 188.22	457574.92 ± 423.81	0.97 ± 0.00	444107
$m_{\ell\ell} > 10$ GeV	159.11 ± 1.95	1643.61 ± 10.01	119.97 ± 0.80	370.04 ± 1.77	282224.18 ± 235.66	26615.94 ± 66.11	35278.28 ± 50.23	2612.46 ± 23.77	2541.36 ± 55.94	82097.48 ± 207.36	2217.90 ± 182.15	21319.41 ± 183.67	455397.01 ± 419.56	0.97 ± 0.00	442671
$n_{\text{jet}} = 2$	97.82 ± 1.50	322.18 ± 5.51	53.19 ± 0.60	73.48 ± 0.72	228994.38 ± 212.06	13252.03 ± 46.54	6163.53 ± 12.02	702.71 ± 14.15	421.91 ± 18.75	7383.37 ± 61.11	236.04 ± 42.09	6082.74 ± 99.08	263363.37 ± 251.30	0.99 ± 0.00	261466
$m_{jj} > 120$ GeV	90.46 ± 1.44	211.37 ± 4.57	21.70 ± 0.35	52.68 ± 0.59	169318.22 ± 182.74	9319.29 ± 39.02	4219.69 ± 9.45	483.30 ± 12.82	295.34 ± 15.26	4760.83 ± 46.92	206.47 ± 23.04	4012.86 ± 85.14	192690.38 ± 213.03	0.99 ± 0.00	189982
$p_{\text{T}}^{j1} < 200$ GeV	79.58 ± 1.36	167.21 ± 4.11	17.65 ± 0.32	44.91 ± 0.54	156415.47 ± 1023.95	8159.58 ± 64.40	3266.28 ± 8.92	386.66 ± 12.12	245.71 ± 14.21	3092.55 ± 132.73	124.63 ± 17.79	3558.67 ± 74.33	175312.13 ± 1037.55	0.97 ± 0.01	169560
b-veto	69.64 ± 1.24	137.83 ± 3.64	12.83 ± 0.27	38.16 ± 0.49	7940.61 ± 63.97	945.89 ± 13.73	2741.62 ± 8.25	307.37 ± 11.29	198.50 ± 12.20	2478.64 ± 109.04	102.22 ± 16.82	797.42 ± 28.78	15563.28 ± 132.77	0.97 ± 0.01	15047
$Z \to \tau\tau$ veto	59.88 ± 1.15	121.68 ± 3.42	9.77 ± 0.25	5.88 ± 0.22	4958.60 ± 44.01	581.56 ± 10.44	1665.13 ± 6.74	174.26 ± 7.35	91.27 ± 7.96	875.28 ± 42.72	42.63 ± 10.21	495.43 ± 21.42	8899.81 ± 67.80	0.97 ± 0.01	8647
$\text{BDT}_{\text{low}} > 0.669$	16.61 ± 0.53	3.42 ± 0.64	0.01 ± 0.00	0.04 ± 0.01	5.53 ± 0.91	1.39 ± 0.43	3.69 ± 0.28	0.52 ± 0.23	0.97 ± 0.68	0.87 ± 0.36	0.01 ± 0.01	2.01 ± 1.06	15.05 ± 1.69	2.26 ± 0.46	34
$p_{\text{T}}^{j1} \geq 200$ GeV	10.88 ± 0.49	44.16 ± 2.00	4.05 ± 0.15	7.77 ± 0.23	15678.15 ± 667.49	1240.69 ± 54.23	953.40 ± 3.12	96.63 ± 4.18	49.63 ± 5.57	473.34 ± 27.66	30.79 ± 2.80	454.19 ± 41.52	18988.63 ± 671.59	1.08 ± 0.04	20422
b-veto	8.70 ± 0.43	33.26 ± 1.70	2.67 ± 0.12	6.38 ± 0.21	794.20 ± 35.75	112.17 ± 6.09	742.39 ± 2.73	70.79 ± 3.50	38.28 ± 4.87	355.17 ± 20.84	22.01 ± 2.24	87.75 ± 16.14	2231.82 ± 45.37	0.90 ± 0.03	2015
$Z \to \tau\tau$ veto	6.96 ± 0.39	26.33 ± 1.46	1.76 ± 0.10	0.68 ± 0.07	367.37 ± 17.77	48.62 ± 3.19	324.27 ± 1.74	30.36 ± 2.27	16.66 ± 3.23	94.14 ± 5.94	6.62 ± 0.84	50.06 ± 10.62	940.54 ± 22.21	0.91 ± 0.04	856
$\text{BDT}_{\text{high}} > 0.631$	1.69 ± 0.18	1.08 ± 0.30	0.01 ± 0.01	0.00 ± 0.00	0.14 ± 0.14		0.88 ± 0.08	0.02 ± 0.02		0.12 ± 0.06	0.02 ± 0.02	0.54 ± 0.61	1.73 ± 0.63	3.47 ± 1.90	6

Table 6.17.: Event yields in the various signal and control regions. The normalization factors from 6.10 have been applied.

	VBF	ggF	VH	$H\tau\tau$	$t\bar{t}$	Wt	WW	$WZ/ZZ/W\gamma^*$	$V\gamma$	$Z/\gamma^* \to \tau\tau$	$Z/\gamma^* \to \ell\ell$	Mis-Id	Total Bkg.	Data/Bkg.	Data
low-p_{T} SR	16.61 ± 0.53	3.42 ± 0.64	0.01 ± 0.00	0.04 ± 0.01	5.53 ± 0.91	1.39 ± 0.43	3.69 ± 0.28	0.52 ± 0.23	0.97 ± 0.68	0.87 ± 0.36	0.01 ± 0.01	2.01 ± 1.06	15.05 ± 1.69	2.26 ± 0.46	34
low-p_{T} Top CR	8.20 ± 0.50	20.75 ± 1.73	2.40 ± 0.13	0.59 ± 0.06	14035.67 ± 104.29	1156.83 ± 15.59	121.68 ± 1.58	24.55 ± 2.24	13.38 ± 2.85	170.60 ± 10.49	6.67 ± 1.52	495.80 ± 23.97	16028.19 ± 108.72	1.00 ± 0.01	16025
low-p_{T} $Z \to \tau\tau$ CR	3.38 ± 0.27	5.53 ± 0.72	0.48 ± 0.04	6.42 ± 0.19	94.69 ± 4.21	9.93 ± 1.19	34.20 ± 0.84	12.09 ± 1.22	22.82 ± 3.01	532.86 ± 23.68	7.45 ± 2.38	29.66 ± 8.15	750.60 ± 25.75	1.00 ± 0.05	748
high-p_{T} SR	1.69 ± 0.18	1.08 ± 0.30	0.01 ± 0.01	0.00 ± 0.00	0.14 ± 0.14		0.88 ± 0.08	0.02 ± 0.02		0.12 ± 0.06	0.02 ± 0.02	0.54 ± 0.61	1.73 ± 0.63	3.47 ± 1.90	6
high-p_{T} Top CR	1.70 ± 0.21	6.66 ± 0.82	0.63 ± 0.06	0.08 ± 0.03	437.05 ± 20.52	49.71 ± 3.28	21.03 ± 0.47	2.62 ± 0.58	1.37 ± 0.46	19.42 ± 1.88	0.90 ± 0.29	30.48 ± 8.12	563.28 ± 22.41	1.00 ± 0.06	564
high-p_{T} $Z \to \tau\tau$ CR	0.89 ± 0.14	3.42 ± 0.71	0.22 ± 0.03	1.17 ± 0.08	9.55 ± 1.22	1.49 ± 0.40	11.81 ± 0.33	4.86 ± 1.07	8.33 ± 1.67	154.17 ± 9.66	5.01 ± 1.26	0.00 ± 5.82	193.38 ± 11.59	1.00 ± 0.09	193

Table 6.18.: Event yields in the signal regions of the VBF EFT benchmark samples.

	SM	SM+HVV	SM−HVV	SM+H∂V	SM−H∂V	SM+AVV	SM−AVV	SM+HVV+H∂V	SM+HVV+AVV	SM+AVV+H∂V	SM+HVV+AVV+H∂V	SM−HVV−AVV−H∂V
low-p_{T} SR	16.61 ± 0.53	18.52 ± 1.23	13.94 ± 0.83	15.94 ± 0.90	16.33 ± 1.03	17.01 ± 1.09	16.15 ± 1.03	14.70 ± 1.13	18.91 ± 1.45	15.20 ± 0.98	18.46 ± 1.28	18.87 ± 1.03
high-p_{T} SR	1.69 ± 0.18	5.70 ± 0.65	3.72 ± 0.44	2.30 ± 0.33	5.41 ± 0.62	4.14 ± 0.67	5.10 ± 0.60	5.34 ± 0.58	10.72 ± 0.94	4.74 ± 0.51	9.62 ± 0.89	9.33 ± 0.82

6.9.1 Experimental Uncertainties

Two types of experimental systematic uncertainties are distinguished.

The systematic uncertainties on e.g. the scale or resolution of the energy (in case of electrons or jets) or momentum (for muons) is computed by shifting the energy/momentum of the object by a $\pm 1\sigma$ variation from the nominal value before selecting the events and by observing the effect of these variations on the number of events in the final regions. The shifting procedure is referred to as *smearing*. This type of systematic is labeled *four-vector* or *P4 systematics*.

The second type of systematic uncertainty is evaluated by comparing the nominal event yield with the one after having modified the interested quantity by applying different event weights. The nominal yield ($\sum A$) is given by the sum of nominal weights in the respective final regions, whereas the modified yield ($\sum B$) is calculated by the same raw events, but using the modified weights. The relative systematic uncertainty is evaluated as the relative difference of the nominal and modified sum of weights ($|\sum A - \sum B| / \sum A$). Comparing the yield of a $\pm 1\sigma$ variation with respect to the nominal value result in symmetric uncertainties, which are classified as *scale factor* or *SF systematics*.

The ATLAS Combined Performance groups gives recommendations, which and how the systematic uncertainties are good to calculate. The standard set of experimental systematics is made of 94 nuisance parameters listed in Table 6.19.

Table 6.19.: Experimental systematic uncertainties. Whether the respective uncertainty is modelled as a scale factor or four-momentum variation is indicated by the last column.

Event related:

`Pile-up Reweighting`	uncertainty on the pile-up reweighting	P4
`Luminosity`	uncertainty on the total integrated luminosity	SF

Electron related:

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

Muon related:

`MUON_EFF_TrigStatUncertainty`, `MUON_EFF_TrigSystUncertainty`	statistical and systematic variations in the trigger efficiency	SF
`MUON_EFF_STAT`, `MUON_EFF_SYS`, `MUON_EFF_SYS_LOWPT`	statistical and systematic variations in the identification efficiency	SF

MUON_ISO_STAT, MUON_ISO_SYS	statistical and systematic variations in the isolation efficiency	SF
MUON_TTVA_STAT, MUON_TTVA_SYS	variations of the track-to-vertex association efficiency	SF
MUON_ID	track momentum smearing in the Inner Detector	P4
MUON_MS	track momentum smearing in the muon spectrometer	P4
MUON_SCALE	variations in the muon momentum scale	P4

Jet related:

JET_EffectiveNP_[1:7], JET_EffectiveNP_8restTerm	linear decomposition of the energy scale uncertainties	P4
JET_EtaIntercalibration_Modelling, JET_EtaIntercalibration_TotalStat, JET_EtaIntercalibration_NonClosure	energy scale uncertainties covering the η-dependence	P4
JET_Pileup_OffsetMu, JET_Pileup_OffsetNPV, JET_Pileup_PtTerm, JET_Pileup_RhoTopology	energy scale uncertainties covering the effects of pile-up	P4
JET_Flavor_Composition	energy scale uncertainties on the flavor composition	P4
JET_Flavor_Response	energy scale uncertainties on the samples' flavor response	P4
JET_BJES_Response	energy scale uncertainties on the b-jets	P4
JET_PunchThrough_MC15	energy scale uncertainties on the effects of the calorimeter not covering the entirety of the shower due to insufficient thickness	P4
JET_SingleParticle_HighPt	energy scale uncertainties on the behavior of high-p_{T} jets	P4
JET_JER_SINGLE_NP	single nuisance parameter covering the jet energy resolution uncertainties	P4
JET_RelativeNonClosure_MC15	uncertainty arising from the use of the ATLFAST-II fast simulation framework	P4
JET_JvtEfficiency, JET_fJvtEfficiency	uncertainty on the (forward) jet vertex tagging	SF
FT_EFF_Eigen_B_[0:2], FT_EFF_Eigen_C_[0:3], FT_EFF_Eigen_Light_[0:4]	b-tagging efficiency uncertainties (3 components for b-jets, 4 for c-jets and 5 for light jets)	P4
FT_EFF_extrapolation	b-tagging efficiency uncertainty on the extrapolation to high-p_{T} jets	P4
FT_EFF_extrapolation_from_charm	b-tagging efficiency uncertainty on τ jets	P4

$E_{\mathrm{T}}^{\mathrm{miss}}$ related:

MET_SoftTrk_ResoPara	track-based soft term related to the longitudinal resolution uncertainty	P4
MET_SoftTrk_ResoPerp	track-based soft term related to the transverse resolution uncertainty	P4
MET_SoftTrk_Scale	track-based soft term related to the longitudinal scale uncertainty	P4

Misidentified leptons:

HWW_FakeFactor_[el/mu]_EWSUBTR	Uncertainty arising from the electroweak subtraction for fake electrons/muons	SF

HWW_FakeFactor_[el/mu]_SAMPLECOMPOSITION	Uncertainty arising from the sample composition variance for fake electrons/muons	SF
HWW_FakeFactor_[el/mu]_STAT_[2015/2016/2016D/combined]_[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 2015/2016/2016D and combined data sets	SF

Event-related Uncertainties

The MC samples are generated with a generic spectrum of average interactions per bunch crossing. As this does usually not reflect the data taking condition, a reweighting is applied to alter the pile-up distribution in MC to coincide with the one measured in data. The uncertainty connected to this reweighting is incorporated in the form of a scale factor uncertainty.

The integrated luminosity of the data set is estimated from a preliminary calibration [102] of the luminosity scale using van der Meer beam separation scans [127] performed during 2015 and 2016. The relative uncertainties on the luminosity are $\pm 2.1\%$ and $\pm 2.2\%$ for the two years, respectively. This results in an uncertainty of $\pm 2.1\%$ on the total luminosity.

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 [128, 129]. 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 MC simulation [130].

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 flavor 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 [131]. 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 modelled by a single nuisance parameter.

Flavor Tagging Uncertainties

Events which are likely to have top quarks in the hard scatter process are suppressed by vetoing events with at least one b-tagged jet. Conversely, events with b-tagged jets are used to construct the top control regions. The connected uncertainties are modelled by a set of parameters corresponding to the eigen-vector decomposition of the corresponding parameter space [76], as well as some additional parameters.

Missing Transverse Momentum

The uncertainties related to the measurement of the track-soft-term missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ have been evaluated using different MC generators [131]. They 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 both longitudinal and transverse. These uncertainties were measured in $Z \to \mu\mu$ events, and take into account uncertainties on detector material effects.

Misidentified Leptons

Processes including misidentified leptons are estimated in a data-driven approach as described in Section 6.7.2. Several systematic uncertainties are related to the data-driven fake estimate.

The subtraction of electroweak processes in di-jet events is evaluated by estimating an uncertainty on the dominating W/Z+jets production of 20% and recomputed for each systematic variation of the electroweak contribution the fake factors coherently.

The sample composition, i.e. the fraction of gluon-induced jets with respect to quark-induced jets, as well as the fraction of heavy-flavor jets, might be affected by the event selection. The data samples used to determine the fake factor will also have a fake composition different from the one in the signal region. The alternative derivation of the fake factors in Z+jets events is used to estimate the systematic uncertainty arising from this variation.

The statistical uncertainty on the fake factor measurement is evaluated in different η and p_{T} bins for the fake leptons and for different data taking periods as well as the combined data taking period.

6.9.2 Theoretical Uncertainties

For the dominant signal and the background processes normalized from theory predictions, uncertainties on the absolute yields in each control and signal region are taken into account, including migrations between the different regions. The theoretical uncertainties considered include QCD scale variations, differences of MC generators and matching schemes, uncertainties on modelling 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, which are described in the respective subsections.

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

The event yields of the nominal MC generator are primarily compared to prevailing generators like POWHEG or MADGRAPH5_aMC@NLO, dependent on availability of the particular MC samples. Typical parton shower generators used to determine theoretical uncertainties are HERWIG or PYTHIA.

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

VBF and ggF Higgs Boson Production

Theoretical uncertainties are estimated for VBF and ggF in the same way.

Generator uncertainties are evaluated by comparing the nominal MC generator MADGRAPH5_aMC@NLO+PYTHIA8 in A14 tune with events generated using POWHEG+PYTHIA8 in the AZNLO tune. Both MC samples include detector simulation. The same selection is applied on both samples and event yields in each region are compared.

PDF events weights are altered for the POWHEG sample, first for the hundred different NNPDF3.0 NLO PDF sets and additionally for the PDF sets PDF4LHC15 NLO, CT10 NLO, MMHT2014 NLO and CT14 NLO all evaluated at $\alpha_S = 0.018$.

Different event weights are used to determine the eight QCD scale variation uncertainties $\mu_{R/F} = [0.5, 1.0, 2.0]$ using the nominal MADGRAPH5_aMC@NLO sample.

On generator level, differences in modelling of the parton shower are calculated comparing the nominal PYTHIA sample with a HERWIG sample.

Only in MADGRAPH the Higgs Characterization model is available to generate MC samples with different EFT parameters. Hence, all theory uncertainties are determined at SM parameters and their relative impact on the event yields in the different regions are applied for all other EFT samples, accordingly.

Continuum WW Production

For the WW background all theoretical uncertainties are determined on generator level samples instead of reconstruction level in order to avoid artificial variations due to lack of MC statistics. The QCD scale uncertainties and PDF uncertainties are evaluated using different event weights in the nominal SHERPA MC sample. All NNPDF3.0 NLO PDF sets are included and additionally four different PDF sets, CT14 NNLO, MMHT2014 NNLO and NNPDF3.0 NNLO evaluated at $\alpha_S = 0.017$ and 0.019.

The PS/UE uncertainty is computed by comparing the nominal SHERPA sample with the A14 tune to the predictions from MADGRAPH5_aMC@NLO interfaced to PYTHIA8. For the MADGRAPH5_aMC@NLO+PYTHIA8 sample, the matrix elements are simulated in association with up to one additional parton with NLO precision, thus diagrams with two additional partons in the final state are included

with LO accuracy. These both samples are also compared to obtain estimates of the uncertainty of the NLO parton shower matching.

Uncertainties on the SHERPA WW production are evaluated varying the choice of the parameters set fixed in the generation. These parameters are the ME matching scale (CKKW [132]) which is varied from the nominal value of 20 GeV to 15 and 30 GeV and the renormalization μ_R, factorization μ_F and resummation μ_{QSF} scales that are varied by a factor 0.5 and 2 with respect to the nominal value.

Likewise, the uncertainty on the choice of the parton shower recoil scheme, which varies the recoil against the emitted gluon between the other two produced partons, is included using a similar SHERPA sample generated with a different recoil scheme.

Top Production

Generator systematics are evaluated by comparing the nominal POWHEG+PYTHIA8 samples with MADGRAPH5_aMC@NLO + PYTHIA8 for $t\bar{t}$ and MADGRAPH5_aMC@NLO + HERWIG++ for Wt. Uncertainties connected to QCD renormalization and factorization scales as well as to the PDF model are derived analogously to the other processes. The nominal parton showers are compared with HERWIG7 for $t\bar{t}$ with a POWHEG+ HERWIG++ sample for Wt.

Additionally, for Wt uncertainties connected to variable shower radiation are estimated, as well as uncertainties on the interference between $t\bar{t}$ and Wt contributions by comparing two different overlap subtraction strategies.

6.9.3 Impact of Systematics Uncertainties

The impact of the systematics is measured for each sample in the single categories by how much the yield in each bin is changed by the $\pm 1\sigma$ variation or one-sided variation with relation to the nominal event yield. Table 6.20 shows the impacts of the maximal relative up or down variation in percent of the experimental systematics, Table 6.21 of the systemtatics related to the misidentified leptons estimates and Table 6.22 of the theoretical systematics. The columns in the tables indicate individual uncertainty parameters, the rows indicate categories (signal or control regions) and samples contributing to the estimates. The cell entries indicate the relative effects of the $\pm 1\sigma$ variations on the total yield. The cell colors encode the treatment of the corresponding parameter.

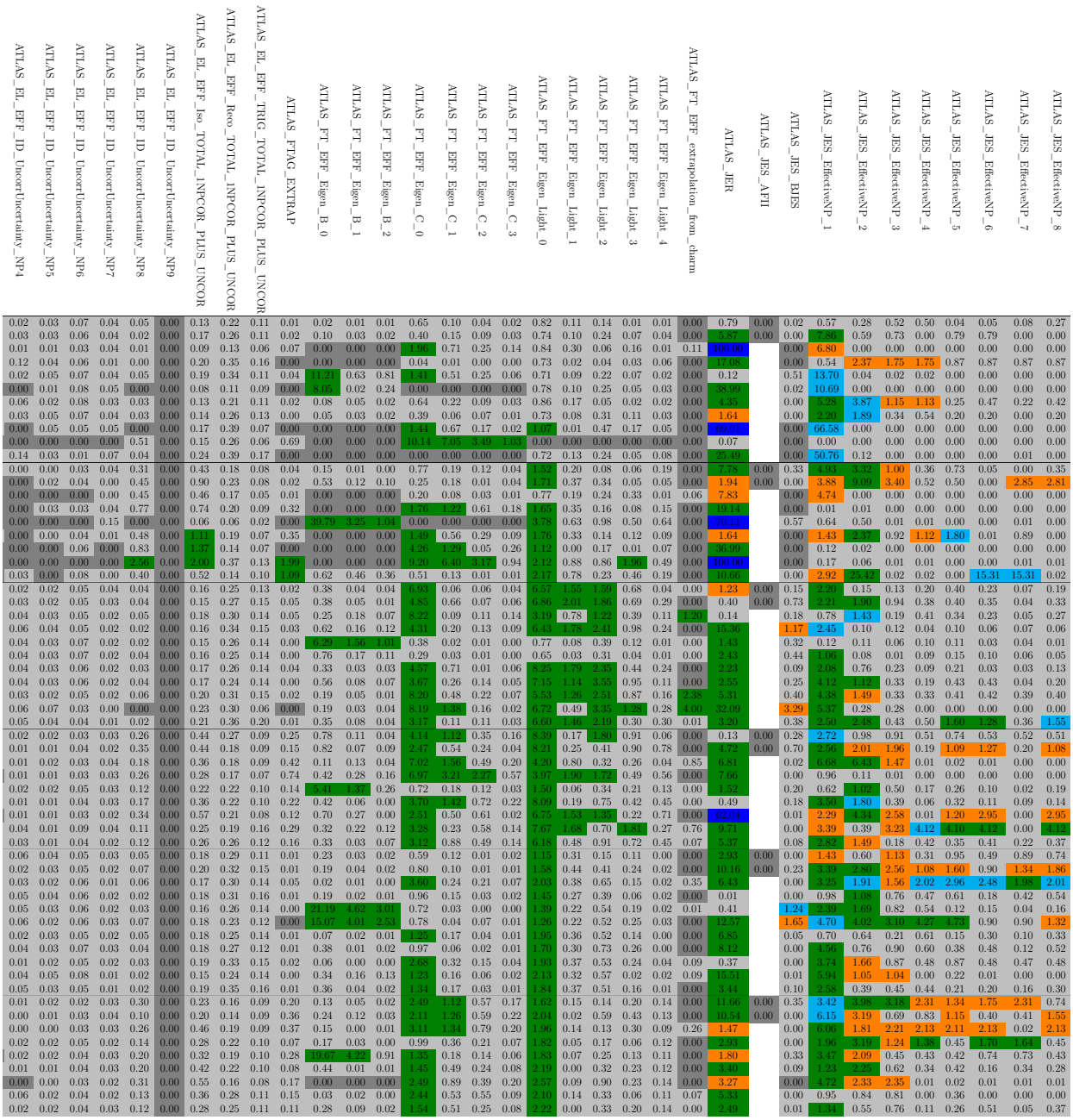
- Dark and light gray cells indicates uncertainties neglected because they have no effect on the yield or under 1% yield variation, which is compared to other systematics too small to have a noticeable impact on the likelihood fit.
- Cyan indicates cases where the templates have been symmetrized. This is only executed if both up and down variation have the same sign, i.e. both variations increase or decrease the event yield. Then the larger of the two variations is taken and used as a symmetric uncertainty.

- Blue indicates uncertainties capped at 100% because the statistical power of the variational template is so low that unrealistic high uncertainties emerge and thus the uncertainty can not conclusively determine the effect.
- P4 systematics affecting only the normalization are pruned if the variation in units of MC stat uncertainty differ from the nominal event yield by less than 0.2σ. These cells are colored in orange.
- Shape systematics are dropped if both relative up and down variation with respect to the nominal histogram are tested for being compatible with a flat ratio by performing a χ^2 fit to a polynomial of 0th order. The systematic is pruned unless the probability to encounter a χ^2 value lower than the one found in the fit of the relative variation exceeds 0.05. If both, up and down variation histograms are found, the systematic is kept, if one of the variations is deemed significant by the aforementioned criterion. These cases are color coded in orange, too.
- Green finally indicates that the systematic is used as-is.

This pruning helps significantly to stabilize the likelihood fit which would not be able to converge including systematics with no to little effect on the event yield, since a variation of the respective nuisance parameter would not have any or negligible effect to determine the minimum of the likelihood function. The construction of the likelihood function, the statistical treatment and the fit results are presented in the next sections.

Table 6.20.: Impact of experimental systematics on the event yields on the individual samples in all regions. The cell color indicates different treatment of the systematics in the fit and is described in the text.

		ATLAS_EG_RESOLUTION_ALL	ATLAS_EG_SCALE_ALLCORR	ATLAS_EG_SCALE_E4SCINTILLATOR	ATLAS_EG_SCALE_LARCALIB_EXTRA2015PRE	ATLAS_EG_SCALE_LARTEMPERATURE_EXTRA2015PRE	ATLAS_EG_SCALE_LARTEMPERATURE_EXTRA2016PRE	ATLAS_EL_EFF_ID_CorrUncertainty_NP0	ATLAS_EL_EFF_ID_CorrUncertainty_NP1	ATLAS_EL_EFF_ID_CorrUncertainty_NP10	ATLAS_EL_EFF_ID_CorrUncertainty_NP11	ATLAS_EL_EFF_ID_CorrUncertainty_NP12	ATLAS_EL_EFF_ID_CorrUncertainty_NP13	ATLAS_EL_EFF_ID_CorrUncertainty_NP14	ATLAS_EL_EFF_ID_CorrUncertainty_NP2	ATLAS_EL_EFF_ID_CorrUncertainty_NP3	ATLAS_EL_EFF_ID_CorrUncertainty_NP4	ATLAS_EL_EFF_ID_CorrUncertainty_NP5	ATLAS_EL_EFF_ID_CorrUncertainty_NP6	ATLAS_EL_EFF_ID_CorrUncertainty_NP7	ATLAS_EL_EFF_ID_CorrUncertainty_NP8	ATLAS_EL_EFF_ID_CorrUncertainty_NP9	ATLAS_EL_EFF_ID_UncorrUncertainty_NP0	ATLAS_EL_EFF_ID_UncorrUncertainty_NP1	ATLAS_EL_EFF_ID_UncorrUncertainty_NP10	ATLAS_EL_EFF_ID_UncorrUncertainty_NP11	ATLAS_EL_EFF_ID_UncorrUncertainty_NP12	ATLAS_EL_EFF_ID_UncorrUncertainty_NP13	ATLAS_EL_EFF_ID_UncorrUncertainty_NP14	ATLAS_EL_EFF_ID_UncorrUncertainty_NP15	ATLAS_EL_EFF_ID_UncorrUncertainty_NP2	ATLAS_EL_EFF_ID_UncorrUncertainty_NP3
low-p_T SR	VBF$_{SM}$	0.22	0.23	0.04	0.14	0.00	0.00	0.00	0.00	0.04	0.08	0.06	0.21	0.42	0.00	0.00	0.00	0.02	0.05	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.03
	ggF$_{SM}$	0.05	0.42	0.00	0.42	0.00	0.00	0.00	0.00	0.04	0.05	0.12	0.44	0.35	0.00	0.00	0.00	0.04	0.15	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.07
	VH	0.20	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.08	0.08	0.11	0.18	0.00	0.00	0.00	0.00	0.01	0.02	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.01
	$H \to \tau\tau$	0.67	0.25	0.00	0.25	0.00	0.00	0.00	0.00	0.07	0.03	0.22	0.23	0.90	0.00	0.00	0.00	0.02	0.11	0.03	0.03	0.13	0.00	0.00	0.00	0.00	0.02	0.03	0.01	0.00	0.00	0.05
	$t\bar{t}$	0.32	0.06	0.00	0.06	0.00	0.00	0.00	0.00	0.04	0.04	0.11	0.50	0.42	0.00	0.00	0.00	0.04	0.13	0.02	0.00	0.03	0.00	0.00	0.00	0.00	0.03	0.01	0.02	0.02	0.00	0.08
	Wt	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.11	0.00	0.11	0.28	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00
	$qq \to WW$	2.49	0.45	0.00	0.24	0.00	0.00	0.00	0.00	0.04	0.09	0.03	0.10	0.43	0.00	0.00	0.00	0.00	0.03	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.01
	$gg \to WW$	3.25	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.07	0.09	0.30	0.44	0.00	0.00	0.00	0.02	0.09	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.06
	$WZ/ZZ/W\gamma^{*}$	0.29	1.68	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.21	0.73	0.48	0.00	0.00	0.01	0.10	0.36	0.07	0.04	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.14
	$V\gamma$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.18	0.05	0.03	0.91	0.00	0.00	0.00	0.00	0.00	0.05	0.08	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Z/DY	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.01	0.29	0.39	1.02	0.00	0.00	0.00	0.09	0.33	0.03	0.03	0.03	0.00	0.00	0.00	0.00	0.00	0.05	0.03	0.03	0.00	0.12
high-p_T SR	VBF$_{SM}$	1.49	1.69	0.00	0.38	0.00	0.00	0.00	0.00	0.05	0.06	0.02	0.12	0.26	0.00	0.00	0.00	0.00	0.02	0.02	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	ggF$_{SM}$	8.80	5.77	0.00	2.41	0.00	0.00	0.00	0.00	0.08	0.04	0.03	0.17	0.13	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	VH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.09	0.00	0.02	0.55	0.00	0.00	0.00	0.00	0.00	0.03	0.05	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$H \to \tau\tau$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.10	0.00	0.05	0.20	0.00	0.00	0.00	0.00	0.00	0.04	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	top	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.13	0.01	0.05	0.21	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$qq \to WW$	2.19	0.65	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.09	0.16	0.00	0.00	0.00	0.00	0.02	0.02	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	$gg \to WW$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.02	0.01	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.07	0.02	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$WZ/ZZ/W\gamma^{(*)}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.14	0.04	0.01	0.10	0.00	0.00	0.00	0.00	0.00	0.18	0.10	0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Z/DY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.09	0.03	0.10	0.01	0.00	0.00	0.00	0.00	0.01	0.03	0.05	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
low-p_T Top CR	VBF$_{SM}$	0.23	0.11	0.00	0.06	0.00	0.00	0.00	0.00	0.04	0.06	0.07	0.29	0.48	0.01	0.00	0.00	0.02	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.01	0.03	0.00	0.04
	ggF$_{SM}$	0.44	0.70	0.00	0.27	0.00	0.00	0.00	0.00	0.05	0.06	0.09	0.27	0.45	0.00	0.00	0.01	0.02	0.07	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.02	0.00	0.03
	VH	0.07	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.06	0.10	0.26	0.54	0.00	0.00	0.00	0.02	0.08	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.02	0.01	0.04	0.00	0.03
	$H \to \tau\tau$	0.78	1.31	0.00	0.29	0.00	0.00	0.00	0.00	0.05	0.03	0.09	0.23	0.65	0.01	0.00	0.02	0.03	0.06	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.03	0.02	0.01	0.04	0.00	0.04
	$t\bar{t}$	0.05	0.06	0.01	0.06	0.00	0.00	0.00	0.00	0.05	0.07	0.09	0.27	0.53	0.00	0.00	0.00	0.02	0.08	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.00	0.04
	Wt	0.02	0.02	0.00	0.07	0.00	0.00	0.00	0.00	0.04	0.07	0.08	0.24	0.49	0.00	0.00	0.00	0.02	0.07	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.00	0.03
	$qq \to WW$	0.16	0.18	0.00	0.06	0.00	0.00	0.00	0.00	0.04	0.07	0.09	0.27	0.52	0.00	0.00	0.00	0.02	0.06	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.02	0.02	0.04	0.00	0.03
	$gg \to WW$	0.29	0.17	0.04	0.18	0.00	0.00	0.00	0.00	0.04	0.07	0.09	0.25	0.51	0.00	0.00	0.00	0.02	0.07	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.00	0.04
	$WZ/ZZ/W\gamma^{*}$	1.53	0.40	0.00	0.06	0.00	0.00	0.00	0.00	0.07	0.07	0.14	0.39	0.49	0.00	0.00	0.00	0.04	0.13	0.00	0.01	0.03	0.00	0.00	0.00	0.00	0.02	0.01	0.01	0.04	0.00	0.06
	$V\gamma$	3.10	3.10	0.00	2.96	0.00	0.00	0.00	0.00	0.08	0.02	0.32	1.15	0.41	0.00	0.00	0.01	0.06	0.25	0.03	0.00	0.05	0.00	0.00	0.00	0.00	0.05	0.00	0.02	0.04	0.00	0.23
	Z/DY	0.35	0.27	0.00	0.14	0.00	0.00	0.00	0.00	0.05	0.05	0.19	0.44	0.74	0.00	0.00	0.00	0.04	0.15	0.00	0.01	0.04	0.00	0.00	0.00	0.00	0.04	0.03	0.03	0.04	0.00	0.06
high-p_T Top CR	VBF$_{SM}$	0.03	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06	0.06	0.19	0.39	0.00	0.00	0.00	0.02	0.05	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.03
	ggF$_{SM}$	3.45	1.04	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.08	0.01	0.04	0.36	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	VH	0.10	0.13	0.00	0.03	0.00	0.00	0.00	0.00	0.02	0.09	0.01	0.13	0.38	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.00	0.00
	$H \to \tau\tau$	0.56	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.04	0.09	0.01	0.02	0.35	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	top	0.09	0.13	0.03	0.08	0.00	0.00	0.00	0.00	0.03	0.08	0.05	0.18	0.41	0.00	0.00	0.00	0.01	0.05	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.02	0.00	0.03
	$qq \to WW$	0.05	0.07	0.00	0.12	0.00	0.00	0.00	0.00	0.02	0.08	0.04	0.19	0.38	0.00	0.00	0.00	0.01	0.04	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.00	0.02
	$gg \to WW$	1.53	1.53	0.00	1.53	0.00	0.00	0.00	0.00	0.04	0.07	0.06	0.10	0.34	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00
	$WZ/ZZ/W\gamma^{(*)}$	0.36	0.03	0.00	0.03	0.00	0.00	0.00	0.00	0.06	0.10	0.09	0.04	0.58	0.00	0.00	0.00	0.00	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.01
	Z/DY	1.29	0.21	0.00	0.07	0.00	0.00	0.00	0.00	0.04	0.05	0.10	0.28	0.47	0.00	0.00	0.01	0.02	0.06	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.01	0.02	0.00	0.04
low-p_T $Z \to \tau\tau$ CR	VBF$_{SM}$	0.72	0.64	0.00	0.45	0.00	0.00	0.00	0.00	0.04	0.07	0.07	0.21	0.53	0.00	0.00	0.00	0.02	0.09	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.04
	ggF$_{SM}$	0.75	0.63	0.01	0.23	0.00	0.00	0.00	0.00	0.03	0.04	0.10	0.36	0.41	0.00	0.00	0.00	0.04	0.14	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.07
	VH	0.34	0.09	0.00	0.08	0.00	0.00	0.00	0.00	0.06	0.04	0.12	0.35	0.58	0.02	0.00	0.01	0.03	0.08	0.00	0.02	0.03	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.00	0.06
	$H \to \tau\tau$	0.49	0.29	0.00	0.24	0.00	0.00	0.00	0.00	0.05	0.05	0.14	0.36	0.62	0.00	0.00	0.01	0.03	0.11	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.03	0.02	0.02	0.03	0.00	0.05
	$t\bar{t}$	0.39	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.07	0.10	0.23	0.58	0.00	0.00	0.00	0.02	0.06	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.01	0.04	0.00	0.03
	Wt	1.40	1.75	0.00	0.03	0.00	0.00	0.00	0.00	0.04	0.11	0.04	0.08	0.63	0.00	0.00	0.00	0.00	0.04	0.02	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.02	0.00	0.01
	$qq \to WW$	0.38	0.20	0.00	0.26	0.00	0.00	0.00	0.00	0.03	0.08	0.07	0.26	0.52	0.00	0.00	0.00	0.02	0.05	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.02	0.04	0.00	0.03
	$gg \to WW$	0.57	0.31	0.00	0.13	0.00	0.00	0.00	0.00	0.05	0.05	0.08	0.24	0.51	0.01	0.00	0.01	0.02	0.07	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.01	0.01	0.02	0.00	0.04
	$WZ/ZZ/W\gamma^{*}$	0.19	1.22	0.00	0.44	0.00	0.00	0.00	0.00	0.02	0.02	0.10	0.54	0.57	0.00	0.00	0.01	0.04	0.11	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.04	0.04	0.02	0.04	0.00	0.08
	$V\gamma$	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.12	0.06	0.16	0.63	0.00	0.00	0.00	0.01	0.02	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.03	0.05	0.02	0.05	0.00	0.01
	Z/DY	0.09	0.14	0.04	0.14	0.00	0.00	0.00	0.00	0.04	0.04	0.14	0.40	0.65	0.00	0.00	0.00	0.04	0.12	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.03	0.03	0.02	0.04	0.00	0.07
high-p_T $Z \to \tau\tau$ CR	VBF$_{SM}$	0.15	0.57	0.00	0.57	0.00	0.00	0.00	0.00	0.02	0.12	0.03	0.13	0.44	0.00	0.00	0.00	0.01	0.02	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.00	0.01
	ggF$_{SM}$	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.11	0.03	0.05	0.44	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.06	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.00
	VH	0.06	3.54	0.00	0.07	0.00	0.00	0.00	0.00	0.01	0.06	0.02	0.12	0.31	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.02
	$H \to \tau\tau$	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.06	0.08	0.26	0.37	0.00	0.00	0.00	0.02	0.07	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.00	0.04
	top	1.47	0.62	0.00	0.82	0.00	0.00	0.00	0.00	0.05	0.11	0.01	0.11	0.41	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.00	0.00
	$qq \to WW$	1.11	1.11	0.09	0.08	0.00	0.00	0.00	0.00	0.04	0.06	0.04	0.20	0.36	0.00	0.00	0.00	0.01	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.02
	$gg \to WW$	0.05	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.06	0.00	0.05	0.25	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
	$WZ/ZZ/W\gamma^{(*)}$	0.09	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.07	0.14	0.20	0.64	0.00	0.00	0.00	0.00	0.04	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.01	0.00	0.02	0.00	0.00
	Z/DY	0.13	0.10	0.01	0.10	0.00	0.00	0.00	0.00	0.03	0.06	0.07	0.28	0.40	0.00	0.00	0.00	0.02	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.01	0.03	0.00	0.03


ATLAS_EL_EFF_ID_UncorrUncertainty_NP4
ATLAS_EL_EFF_ID_UncorrUncertainty_NP5
ATLAS_EL_EFF_ID_UncorrUncertainty_NP6
ATLAS_EL_EFF_ID_UncorrUncertainty_NP7
ATLAS_EL_EFF_ID_UncorrUncertainty_NP8
ATLAS_EL_EFF_ID_UncorrUncertainty_NP9
ATLAS_EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR
ATLAS_EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR
ATLAS_EL_EFF_TRIG_TOTAL_1NPCOR_PLUS_UNCOR
ATLAS_FTAG_EXTRAP
ATLAS_FT_EFF_Eigen_B_0
ATLAS_FT_EFF_Eigen_B_1
ATLAS_FT_EFF_Eigen_B_2
ATLAS_FT_EFF_Eigen_C_0
ATLAS_FT_EFF_Eigen_C_1
ATLAS_FT_EFF_Eigen_C_2
ATLAS_FT_EFF_Eigen_C_3
ATLAS_FT_EFF_Eigen_Light_0
ATLAS_FT_EFF_Eigen_Light_1
ATLAS_FT_EFF_Eigen_Light_2
ATLAS_FT_EFF_Eigen_Light_3
ATLAS_FT_EFF_Eigen_Light_4
ATLAS_FT_EFF_extrapolation_from_charm
ATLAS_JER
ATLAS_JES_AFII
ATLAS_JES_BJES
ATLAS_JES_EffectiveNP_1
ATLAS_JES_EffectiveNP_2
ATLAS_JES_EffectiveNP_3
ATLAS_JES_EffectiveNP_4
ATLAS_JES_EffectiveNP_5
ATLAS_JES_EffectiveNP_6
ATLAS_JES_EffectiveNP_7
ATLAS_JES_EffectiveNP_8

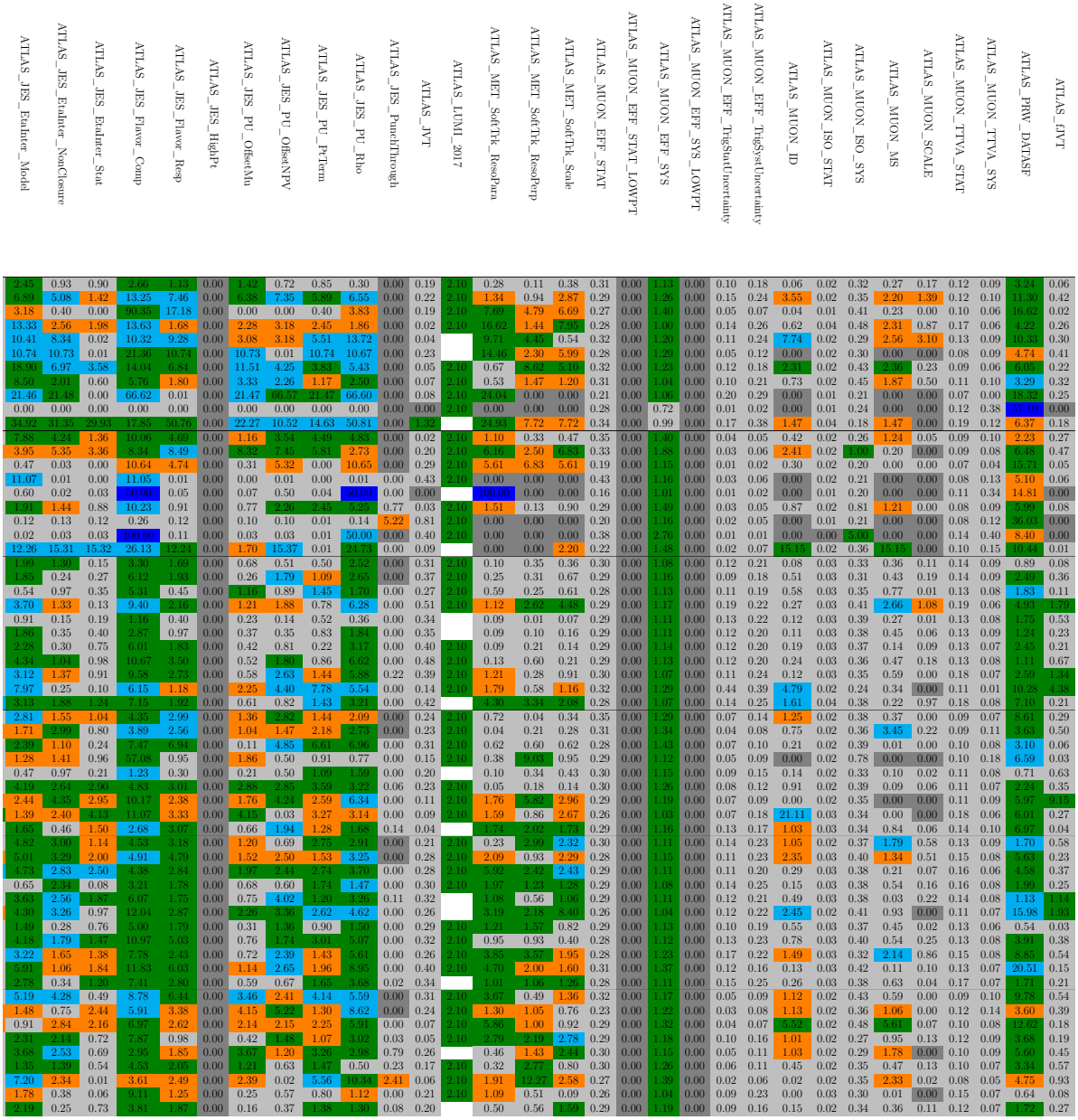

ATLAS_JES_EtaInter_Model
ATLAS_JES_EtaInter_NonClosure
ATLAS_JES_EtaInter_Stat
ATLAS_JES_Flavor_Comp
ATLAS_JES_Flavor_Resp
ATLAS_JES_HighPt
ATLAS_JES_PU_OffsetMu
ATLAS_JES_PU_OffsetNPV
ATLAS_JES_PU_PtTerm
ATLAS_JES_PU_Rho
ATLAS_JES_PunchThrough
ATLAS_JVT
ATLAS_LUMI_2017
ATLAS_MET_SoftTrk_ResoPara
ATLAS_MET_SoftTrk_ResoPerp
ATLAS_MET_SoftTrk_Scale
ATLAS_MUON_EFF_STAT
ATLAS_MUON_EFF_STAT_LOWPT
ATLAS_MUON_EFF_SYS
ATLAS_MUON_EFF_SYS_LOWPT
ATLAS_MUON_EFF_TrigStatUncertainty
ATLAS_MUON_EFF_TrigSystUncertainty
ATLAS_MUON_ID
ATLAS_MUON_ISO_STAT
ATLAS_MUON_ISO_SYS
ATLAS_MUON_MS
ATLAS_MUON_SCALE
ATLAS_MUON_TTVA_STAT
ATLAS_MUON_TTVA_SYS
ATLAS_PRW_DATASF
ATLAS_fJVT

Table 6.21.: Impact of systematics on the event yields of the misidentified leptons estimation in all regions. The cell color indicates different treatment of the systematics in the fit and is described in the text.

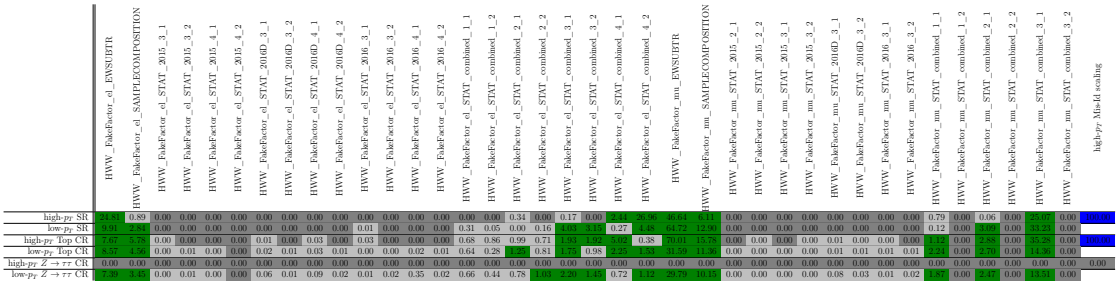


Table 6.22.: Impact of theoretical systematics on the event yields on the individual samples in all regions. The cell color indicates different treatment of the systematics in the fit and is described in the text.

		VBF PDF	VBF PS/UE	VBF QCD scale	VBF generator	ggF PDF	ggF PS/UE	ggF QCD scale	ggF generator	$t\bar{t}$ PDF	$t\bar{t}$ PS radiation	$t\bar{t}$ PS/UE	$t\bar{t}$ generator	Wt DS/DR scheme	Wt PS radiation	Wt PS/UE	Wt generator	WW CKKW	WW PDF	WW QCD scale	WW QSF	WW generator	WW recoil scheme
high-p_T SR	VBF$_{SM}$	2.91	6.40	2.49	10.71																		
	ggF$_{SM}$					4.21	20.25	14.89	29.16														
	top									68.60	12.76	100.00	100.00										
	$qq \to WW$																	51.59	3.88	29.11	2.36	64.74	27.99
low-p_T SR	VBF$_{SM}$	1.35	9.01	3.76	0.95																		
	ggF$_{SM}$					4.80	6.58	21.01	29.21														
	$t\bar{t}$									57.70	46.88	36.70	5.15										
	Wt													13.85	32.24	73.01	100.00						
	$qq \to WW$																	15.81	2.62	23.45	15.10	19.18	12.63
high-p_T Top CR	VBF$_{SM}$	1.92	10.87	4.65	28.47																		
	ggF$_{SM}$					3.94	8.96	34.52	43.45														
	top									67.25	7.86	14.83	8.94	7.60	1.96	2.12	8.55						
	$qq \to WW$																	6.13	2.50	27.33	3.41	29.65	5.14
low-p_T Top CR	VBF$_{SM}$	1.53	6.98	2.17	6.08																		
	ggF$_{SM}$					4.51	9.96	25.78	18.41														
	$t\bar{t}$									67.73	1.14	2.32	0.10										
	Wt													0.93	7.77	13.03	15.94						
	$qq \to WW$																	4.20	2.04	22.94	3.73	18.25	5.45
high-p_T $Z\tau\tau$ CR	VBF$_{SM}$	2.92	8.80	2.94	8.15																		
	ggF$_{SM}$					4.40	27.81	38.00	44.86														
	top									68.65	24.53	15.12	37.23	30.37	23.38	29.95	33.77						
	$qq \to WW$																	3.57	2.32	25.71	5.27	28.04	2.47
low-p_T $Z\tau\tau$ CR	VBF$_{SM}$	1.05	3.82	1.40	3.96																		
	ggF$_{SM}$					5.08	0.07	24.79	42.64														
	$t\bar{t}$									67.99	0.33	5.31	15.86										
	Wt													7.26	2.31	7.59	12.39						
	$qq \to WW$																	2.56	2.12	23.05	3.82	5.54	2.01

6.10 Statistical Treatment

The statistical evaluation of the analysis is carried out by the maximum likelihood method described in Section 5.1. First a likelihood is constructed as function of

- the parameters of interest, which are the EFT coupling parameters $\boldsymbol{\kappa}$, and
- all nuisance parameter describing the effects of systematic variations $\boldsymbol{\theta}$, normalisations by control regions $\boldsymbol{\beta}$ or finite MC statistics $\boldsymbol{\gamma}$.

6.10.1 Likelihood Construction

The event selection described in Section 6.8 defines two disjoint categories, one for low- and one for high-p_{T} leading jets ($p_{\mathrm{T}}^{j_1} \lessgtr 200\,\mathrm{GeV}$). In each category an individual signal region is selected. The events passing the low-p_{T} signal region selection are further subdivided into three bins in the final discriminant variable, the signed azimuthal angle of the leading and subleading jet $\Delta\phi_{jj,\mathrm{sign}}$.[1] The distribution of events in $\Delta\phi_{jj,\mathrm{sign}}$ including all data-driven corrections, but preceding the likelihood minimisation, is shown in Figure 6.36. The modelling of these regions in the likelihood is realised by a Poisson distribution, which describes an event counting experiment. Thus, each bin enters the likelihood with the factor

$$P\left(N \,\middle|\, s+b\right) = \frac{(s+b)^N}{N!} e^{-(s+b)}, \tag{6.10.1}$$

where N is the data event yield, and s and b the number of expected simulated events for signal and background.

Two control regions per category are taken into account by the fit, the top and $Z \to \tau\tau$ control regions, where in each only the total event yield is considered. Corresponding normalisation factors are derived for each category by minimising the likelihood. Similar to the signal regions, these regions are also modelled with a Poisson distribution. The normalisation factors are included by splitting the background contributions into individual process categories equally in both signal and background regions. In this way the normalisation factors for top $\beta^{\mathrm{top}}_{\mathrm{high/low}}$ and $Z{\to}\tau\tau$ $\beta^{Z\tau\tau}_{\mathrm{high/low}}$ are incorporated e.g. for a control region cr with

$$P\left(N_{cr} \,\middle|\, s_{cr} + \beta^{\mathrm{top}}_{\mathrm{high/low}} b^{\mathrm{top}}_{cr} + \beta^{Z\tau\tau}_{\mathrm{high/low}} b^{Z\tau\tau}_{cr} + \sum_{\phi \in \mathrm{bkgs}}^{\notin \mathrm{top}, Z\tau\tau} b^{\phi}_{cr}\right). \tag{6.10.2}$$

[1] Each bin will be seen as individual region to simplify the description of the statistical treatment, such that the term *region* and *bin* can be used interchangeably. In the used statistical analysis framework ROOSTATS different bins in a region are processed as histograms and not as individual regions. Nevertheless, this does not affect the results.

All systematic uncertainties are integrated as subsidiary calibration measurement modelled as constraint term $C(\theta)$ with one nuisance parameter θ corresponding to one systematic each. The systematic constraints are simplified with a normalised Gaussian term

$$C(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}. \tag{6.10.3}$$

They are treated equally to the physics measurement by multiplying every constraint likewise to the likelihood. The effects of systematic uncertainties are measured in the event yields of the different samples involved per systematic and independently in each region or bin. A $\pm 1\sigma$ variation of a systematic, e.g. different energies of the jets due to the variations of the jet energy scale systematic, might result in different event rates in the final regions of the analysis. Since only the nominal and the up and down event rates are determined, it is necessary to interpolate between these three values (similar to the *vertical morphing* explained in Section 5.2.3) and extrapolate to lower or higher values of systematic variations. This is done by polynomial interpolation and exponential extrapolation as follows exemplary for a background sample

$$b(\boldsymbol{\theta}) = b^0 \cdot \prod_{\theta \in \boldsymbol{\theta}} \begin{cases} \left(\frac{b_\theta^+}{b^0}\right)^\theta & \theta \geq 1 \\ 1 + \sum_{i=1}^{6} a_i \theta^i & |\theta| < 1 \\ \left(\frac{b_\theta^-}{b^0}\right)^{-\theta} & \theta < -1. \end{cases} \tag{6.10.4}$$

The event yield of a background sample b in a region depends on the set of nuisance parameters $\boldsymbol{\theta}$ affecting this background. The nominal event yield b^0 is determined by setting all nuisance parameters to $\boldsymbol{\theta} = \mathbf{0}$. Between the up and down event yields per systematic $b_\theta^\pm$ an interpolating polynomial is evaluated, where the polynomial coefficients $a_i,\ i = 1, \ldots, 6$ are fixed satisfying the boundary conditions that the function and its first and second derivative matches at $\theta = \pm 1$. Thus, no kink in the interpolating function is present, which might cause some difficulties in numerical minimisation packages such as MINUIT used in this evaluation.

The finite MC statistics introduce a statistical uncertainty which needs to be included into the likelihood. Barlow and Beeston outlined a treatment in which for each region a nuisance parameter is given to each sample for the true rate, which is then fit using both the data measurement and the MC estimate [133]. This approach would lead to as many parameters as bins times MC samples that are present in the analysis. To simplify the fit procedure and not trigger fitting problems due to an exceeding numbers of fit parameters, only one nuisance parameter for the total MC background per bin is introduced. In one individual region the contribution

to the full statistical model is the Poisson term

$$P(M|\gamma\tau)\,, \tag{6.10.5}$$

where the factor γ is the nuisance parameter reflecting that the true rate may differ from the MC estimate $b^{\mathrm{MC}} = \sum_{\phi \in \mathrm{bkgs}}^{\notin \mathrm{MisId}} b^{\phi}$ by some amount. If the total statistical uncertainty is δb^{MC}, then the relative statistical uncertainty is given by $\frac{\delta b^{\mathrm{MC}}}{b^{\mathrm{MC}}}$. This corresponds to a total MC sample in that region of size $M = \left(\frac{\delta b^{\mathrm{MC}}}{b^{\mathrm{MC}}}\right)^2$. Treating the uncertainty of the MC estimate as an auxiliary measurement, the Poisson constraint term in Equation 6.10.5 is included to the likelihood, where M would fluctuate about $\gamma\tau$ as if a new MC sample is generated. The nuisance parameter γ has been scaled to be a factor at about 1 leading to $\tau = \left(\frac{b^{\mathrm{MC}}}{\delta b^{\mathrm{MC}}}\right)^2$, which is treated as a fixed constant. In each region in the physics Poisson terms, the number of background events estimated using MC techniques b^{MC} are multiplied by the respective γ factor reflecting that the true rate may differ from the MC estimate.

Besides the total MC background, the signal samples in each signal region receive an individual nuisance parameter γ_s to model the MC statistical uncertainty over the phase space of the EFT parameters. The MC statistical uncertainties of the signal samples are usually small, because they have been created with generator level requirements suited to the selection of the specific analysis to retain high statistics in the final regions. Still, especially if the input coupling parameters, at which the signal samples have been generated, are not well chosen, the uncertainties at different EFT parameters can be orders of magnitude larger than the uncertainties of the input samples themselves. More insights in these morphing uncertainties and effects on the likelihood are given in Section 6.11.2.

Finally, the parameters of interest, the EFT coupling parameters $\boldsymbol{\kappa}$, are modelled using the Effective Lagrangian Morphing. The event yields for VBF and ggF are calculated according to Equation 5.3.13 with

$$s(\boldsymbol{\kappa}, \boldsymbol{\theta}, \boldsymbol{\gamma}) = \sum_{s \in S} \gamma_s w_s(\boldsymbol{\kappa})\, s_s(\boldsymbol{\theta})\,. \tag{6.10.6}$$

The sum comprises a linear combination of the morphing weights $w_s(\boldsymbol{\kappa})$, which are functions of the EFT parameters of interest, and the input samples $s_s(\theta)$. They themselves depend on the set of nuisance parameters affecting the respective input signal sample. The EFT coupling parameters at which the input samples have been generated do not change the likelihood function. Still, they are used to calculate the formulas for the morphing weights and influence the magnitude of each weight. Bad chosen input parameters might lead to large cancellations of the weight factor and

cause large statistical uncertainties. Therefore, as mentioned before, MC statistical uncertainty nuisance parameters γ_s are assigned to each input signal sample.

Summary

Putting all ingredients together, the full likelihood formula reads as follows

$$\begin{aligned}\mathcal{L}(\boldsymbol{\kappa},\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma}) = &\prod_{sr\in\text{SRs}} P\left(N_{sr}\middle|\nu_{sr}(\boldsymbol{\kappa},\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma})\right)\cdot \\ &\prod_{cr\in\text{CRs}} P\left(N_{cr}\middle|\nu_{cr}(\boldsymbol{\kappa},\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\gamma})\right)\cdot \\ &\prod_{r\in\text{SRs,CRs}} P\left(M_r\middle|\gamma_r\tau_r\right) \prod_{sr\in\text{SRs}}\prod_{\sigma\in\text{sigs}_{sr}}\prod_{s\in\sigma} P\left(M_s\middle|\gamma_s\tau_s\right)\cdot \\ &\prod_{\theta\in\boldsymbol{\theta}} C(\theta)\end{aligned} \tag{6.10.7}$$

with SRs = {low-p_T SR bin1, low-p_T SR bin2, low-p_T SR bin3, high-p_T SR}
CRs = {low-p_T Top CR, low-p_T $Z\tau\tau$ CR, high-p_T Top CR, high-p_T $Z\tau\tau$ CR}.

The first row is a product of the Poisson terms of the signal regions SRs and the second row over the control regions CRs. These Poisson terms return the probability of obtaining $N_{sr/cr}$ recorded data events when $\nu_{sr/cr}$ are expected. The expected events are calculated as follows in the signal and control regions

$$\begin{aligned}\nu_{sr} = &\sum_{\sigma\in\text{sigs}_{sr}}\sum_{s\in\sigma}\gamma_s w_s(\boldsymbol{\kappa})\, s^s_{sr}(\boldsymbol{\theta}) + \\ &\gamma_{sr}\left(\beta^{\text{top}}_{\text{high/low}} b^{\text{top}}_{sr}(\boldsymbol{\theta}) + \beta^{Z\tau\tau}_{\text{high/low}} b^{Z\tau\tau}_{sr}(\boldsymbol{\theta}) + \sum_{\phi\in\text{bkgs}}^{\notin\text{top},Z\tau\tau,\text{MisId}} b^{\phi}_{sr}(\boldsymbol{\theta})\right) + b^{\text{MisId}}_{sr}(\boldsymbol{\theta})\end{aligned} \tag{6.10.8}$$

$$\begin{aligned}\nu_{cr} = &\sum_{\sigma\in\text{sigs}_{cr}}\sum_{s\in\sigma} w_s(\boldsymbol{\kappa})\, s^s_{cr}(\boldsymbol{\theta}) + \\ &\gamma_{cr}\left(\beta^{\text{top}}_{\text{high/low}} b^{\text{top}}_{cr}(\boldsymbol{\theta}) + \beta^{Z\tau\tau}_{\text{high/low}} b^{Z\tau\tau}_{cr}(\boldsymbol{\theta}) + \sum_{\phi\in\text{bkgs}}^{\notin\text{top},Z\tau\tau,\text{MisId}} b^{\phi}_{cr}(\boldsymbol{\theta})\right) + b^{\text{MisId}}_{cr}(\boldsymbol{\theta})\,.\end{aligned} \tag{6.10.9}$$

In the first row of both equations the number of signal events is computed, where the first sum runs over the simulated EFT signal processes $\text{sigs}_r = \{\text{VBF}_r, \text{ggF}_r\}$

for each region r. The second sum iterates over the respective input samples constructing the morphing function in Equation 6.10.6. Additionally, the signal regions are provided with the MC statistical uncertainty γ_s factors. In the control regions, the influence of the signal event yield to the total yield is small enough to neglect the MC statistical uncertainty modelling and reduce the number of free parameters in the likelihood.

The second row of Equations 6.10.8 and 6.10.9 calculate the background event yield. Only the background samples estimated by MC and not with data-driven methods are included in the MC statistical uncertainty modelling by the $\gamma_{sr/cr}$ factors. The $Z\to\tau\tau$ and top background have control regions with high statistics in the respective processes, such that potential mismodelling of the overall normalisation in these processes are obviated with the $\beta^{\text{top}/Z\to\tau\tau}_{\text{high/low}}$ normalisation factors individually in the high- and low-p_{T} categories. The dependence of the signal and background event yields $s(\boldsymbol{\theta})$ and $b(\boldsymbol{\theta})$ follow the inter- and extrapolation of the nuisance parameters given by Equation 6.10.4. Each sample may only depend on only a subset of systematic uncertainties due to pruning described in Section 6.9.3, which is not explicitly noted in the equations for simplicity.

Further, the third row in Equation 6.10.7 contains the Poisson terms to model the MC statistical uncertainty of the full MC background samples combined for each region, and for all signal morphing input samples in each signal region for both VBF and ggF. In total 8 γ_r nuisance parameters emerge for the backgrounds. Depending on the set of $\boldsymbol{\kappa}$ parameters, the number of morphing input samples vary as seen in Equation 5.3.22. For example in the case of κ_{SM} and κ_{AWW} (5 input samples in VBF and 3 in ggF) the likelihood obtain 32 γ_s parameters.

Finally, the fourth row of Equation 6.10.8 contains the product over all systematic constraint terms modelled as normalised Gaussians shown in Equation 6.10.3.

6.10.2 Statistical Evaluation

The goal of the analysis is to find the confidence interval for the EFT coupling parameter space, which is compatible with the recorded data. With the log-likelihood ratio test statistic[2]

$$t(\boldsymbol{\kappa}) = -2\ln\left(\frac{\mathcal{L}\left(\boldsymbol{\kappa},\widehat{\widehat{\boldsymbol{\theta}}}\right)}{\mathcal{L}\left(\boldsymbol{\kappa},\widehat{\boldsymbol{\theta}}\right)}\right), \tag{6.10.10}$$

first defined in Equation 5.1.4, one or more parameters of interest can be profiled and evaluated at different values. The best fit values $\widehat{\boldsymbol{\kappa}}$ at the minimum of $t(\boldsymbol{\kappa})$ gives the most probable value for the EFT couplings given the data.

[2] All nuisance parameters $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are joined into $\boldsymbol{\theta}$ for convenience.

Since the raw number of events in each signal region is sufficiently large the large-sample limit can be applied and hence the likelihood approaches a Gaussian. In this limit $t(\boldsymbol{\kappa})$ follows a χ^2 distribution with n degrees of freedom according to Wilks' theorem [134], where n is the number of parameters in $\boldsymbol{\kappa}$.

One can use the quantiles $\chi^2_{1-\alpha}$ of the χ^2 distribution to evaluate $1-\alpha$ confidence regions shown in Figure 6.37 for a Gaussian distribution. These quantiles define the rise in $t(\boldsymbol{\kappa})$ corresponding to the points of $\boldsymbol{\kappa}$ on the border of the confidence region. The value of the quantile is obtained from $F^{-1}_{\chi^2}(1-\alpha, n)$, the inverse of the cumulative function $F^{-1}_{\chi^2}$ of a χ^2 distribution with n degrees of freedom. For example, in the case of one κ parameter, the 1σ interval (i.e. $1-\alpha = 68.27\%$) is obtained as

$$\begin{aligned} t(\kappa_{\text{low}}) = t(\kappa_{\text{high}}) &= F^{-1}_{\chi^2}(0.6827, 1) = 1 \\ \text{with} \quad \kappa_{\text{low}} = \widehat{\kappa} - \Delta\widehat{\kappa}_- \quad &\text{and} \quad \kappa_{\text{high}} = \widehat{\kappa} - \Delta\widehat{\kappa}_+, \end{aligned} \tag{6.10.11}$$

illustrated in Figure 6.38. Similarly, $s \cdot \sigma$ confidence intervals can be obtained from the $t\big(\kappa_{\text{low/high}}\big) = s^2$ contour.

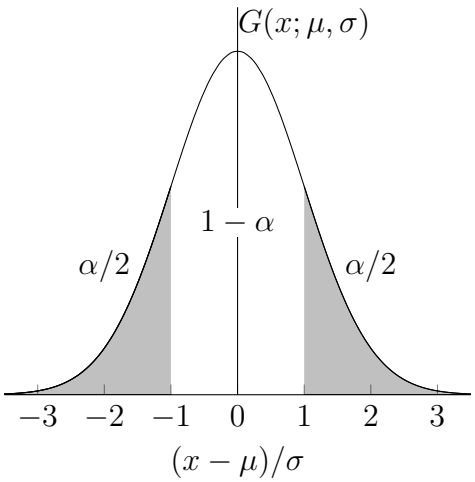


Figure 6.37.: Illustration of a 1σ or 68.27% confidence interval (unshaded) for a Gaussian-distributed measurement of a single quantity.

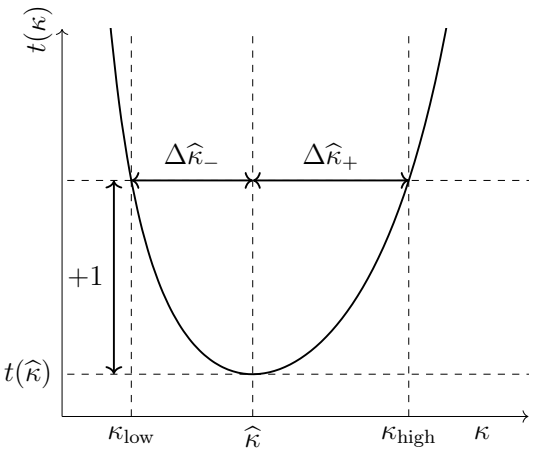


Figure 6.38.: Exemplary $t(\kappa)$ function with the estimated negative and positive uncertainties $\Delta\widehat{\kappa}_\pm$ defining a 1σ or 68.27% confidence level interval estimated at $t(\widehat{\kappa}) + 1$ (cf. Equation 6.10.11).

In the case of two parameters, the contours of κ_1 and κ_2 are enclosing a two dimensional confidence region at a given confidence level by finding the correspond-

ing values in the $t(\kappa_1, \kappa_2)$ function. The 68.27% contour is obtained by finding the set of values for κ_1 and κ_2 which satisfy

$$t(\kappa_1, \kappa_2) = F^{-1}_{\chi^2}(0.6827, 2) = 2.30. \tag{6.10.12}$$

Likewise, for 95.45% (2σ) the contour is given at $t(\kappa_1, \kappa_2) = 6.18$.

6.10.3 Model Diagnostics

The method and machinery for making a measurement should be tested to ensure that it yields on average the input values and appropriate uncertainties. Therefore, a set of tools are utilised.

Before looking at the observed data from the actual experiment a representing artificial data set, the so-called *Asimov data set*, is used to determine the expected confidence intervals of the coupling parameters. The interesting case is given by a data set at the SM parameters which can be generated with the MC prediction that have already been used as input for the statistical model. Not only the expected sensitivity on the estimate of the coupling parameters is obtained by an unconditional fit, but also an important check is made by comparing the best fit values of the coupling parameters to match to the SM and are therefore un-biased.

A similar validation is done for every nuisance parameter by studying the so-called *pull*, which is directly obtained from the unconditional fit with the best fit value $\widehat{\theta}$ and the expected constrain value θ_0 as

$$p_\theta = \frac{\widehat{\theta} - \theta_0}{\theta_0}. \tag{6.10.13}$$

For systematics modelled as a normalised Gaussian, θ_0 is 0. Using the Asimov data set no bias in the pull is expected, because the best fit values have to be by construction at the estimators values. With recorded physics data, small deviations are expected, since the exact systematic values are not known and the data may prefer different estimates. As long as the bias is within the uncertainty of the systematic, the modelling of the likelihood indicates no problem. A larger bias may point to some shortcoming. One reason could be that the probability functions used in the likelihood are not a perfect description of the MC simulations due to e.g. some modelling simplifications.

The constraint on the estimates of a nuisance parameter is calculated using the optimisation software MINOS [135], which does a profiled scan over the nuisance parameter and computes the 1σ uncertainties using Equation 6.10.11. As a result, the constraint might be asymmetrical. The uncertainty is normalised to the respective expected uncertainty, which is 1 for nuisance parameters associated

with a unit Gaussian subsidiary measurement, such that a constraint of ± 1 is expected. A deviation might hint to a constraint of the parameter from the physics measurement, which in some circumstances may limit θ better than the subsidiary measurement leading to an underestimation of the parameter. For example, if the estimators of θ is correlated due to similar response in the physics measurement, the estimated uncertainty on θ is not precisely 1, as one would expect from unit Gaussian subsidiary measurement. This is most likely caused by an insufficiently detailed model of the physics measurement, but may also have a physics reason due to sensitive response of the physics measurement on the systematic.

Larger constraints might result from correlations with other parameters. The correlation matrix, as yet another important diagnostic, shows all correlation between two parameters. High (anti-)correlated parameters might induce problems to the fit, such that it might not converge. Therefore, one needs to understand, if possible correlations have a physics reason. For example high correlations between electron and muon identification systematics would not have a physics reason and can hint to problems in the model.

Further, the relevance of the individual nuisance parameter or groups of nuisance parameters to the total uncertainty of the parameters of interest is substantial, such that one can focus on modelling problems that matter. Over-constraint nuisance parameters that have no impact on the parameter of interest may be a nonissue, because changes to the model might not adversely affect the result and only complicate the model. The uncertainty of each parameter of interest is broken down to groups of individual nuisance parameters by performing a partially conditional fit, where the selected nuisance parameter(s) are fixed to the best fit value of the unconditional fit and in that way the uncertainties of the parameter of interest is computed. Finally, the magnitude of the impact on the parameter of interest is estimated with $\Delta\kappa = \sqrt{\sigma^2_{\text{uncond.}} - \sigma^2_{\text{cond.}}}$ for the upper and lower uncertainties separately.

6.11 Results

This section describes the results of the statistical evaluations and the conclusions that can be drawn. The procedure outlined in Sections 5.1 and 6.10 is performed with the ROOT Data Analysis Framework [136] and the additional ROOFIT [137] package. The probability density functions are incorporated with the tools provided by the S-FRAMEWORK, which is a custom build analysis package linked to Q-FRAMEWORK [138] originally developed for ATLAS analyses in the $H \to WW$ channel. To obtain the estimators of the profiled likelihood ratio $t(\boldsymbol{\kappa})$, the analysis framework is interfaced with the minimisation package MINUIT [139].

Two different sets of EFT coupling parameters are tested separately using the MC samples listed in Table 6.7 as inputs for the morphing functions in the respective likelihoods. The first set consists of the SM and the CP-odd AVV coupling, and the second set includes the SM, the HVV and the derivative H∂V couplings.

6.11.1 Measurement of SM and AVV

Starting with a closer look on the systematic uncertainties and auxiliary measurements, the impact of groups of nuisance parameters and the leading individual nuisance parameters on the coupling parameters κ_{SM} and κ_{AVV} are listed in Figure 6.39 for the SM Asimov dataset and the observed data. While for the κ_{SM} parameter the statistical uncertainty from data is balanced with the full systematic uncertainty, the κ_{AVV} coupling is clearly limited by the statistics of the recorded data. The uncertainty of the jet energy scale has the largest impact on both EFT couplings, followed by the theoretical uncertainties from the parton showering of top processes especially for κ_{SM}. With more and more recorded data in the future, the generation of top-related MC samples and the event selection have to address these systematic uncertainties and minimise their influence on the coupling parameters.

In addition, the pull and uncertainty of the individual nuisance parameters are shown for the Asimov and observed data. The pulls for the Asimov dataset are as expected precisely at 0, or in case of the γ parameters at 1 and in case of the background normalisation factors at the values listed in Table 6.23 at which the Asimov data set has been generated. Only small deviations within the uncertainty of the pull are observed for the experimental dataset. Likewise, the uncertainties have only small deviations from $\pm 1\sigma$. Thus, no mismodelling of the likelihood or unexpected influence of the physics measurement to the auxiliary measurement is observed.

Figure 6.40 shows correlations between pairs of parameters above 20% for both the SM Asimov and observed data. A large anti-correlation between the normalisation factor β^{top} and the PDF top uncertainty is observed in both low- and high-p_{T} regions. Both have only a small impact on the uncertainty of the coupling parameters. Comparing the data sets the correlations are the same within small fluctuations except between the coupling parameters, which are moderately correlated for the Asimov dataset and anti-correlated for the observed data. These correlation reflect the shape of the two dimensional likelihood contours shown in Figure 6.42 and discussed in more detail later. The best fit value of the Asimov data set is located exactly at the SM parameters for which the C-shaped contour is only very lightly rotated anti-clockwise leading to a small positive correlation. However, the best fit value for the observed data prefers a positive κ_{AVV} value and is located therefore in the right part of the C-shape contour, which indicates a negative correlation between both coupling parameters.

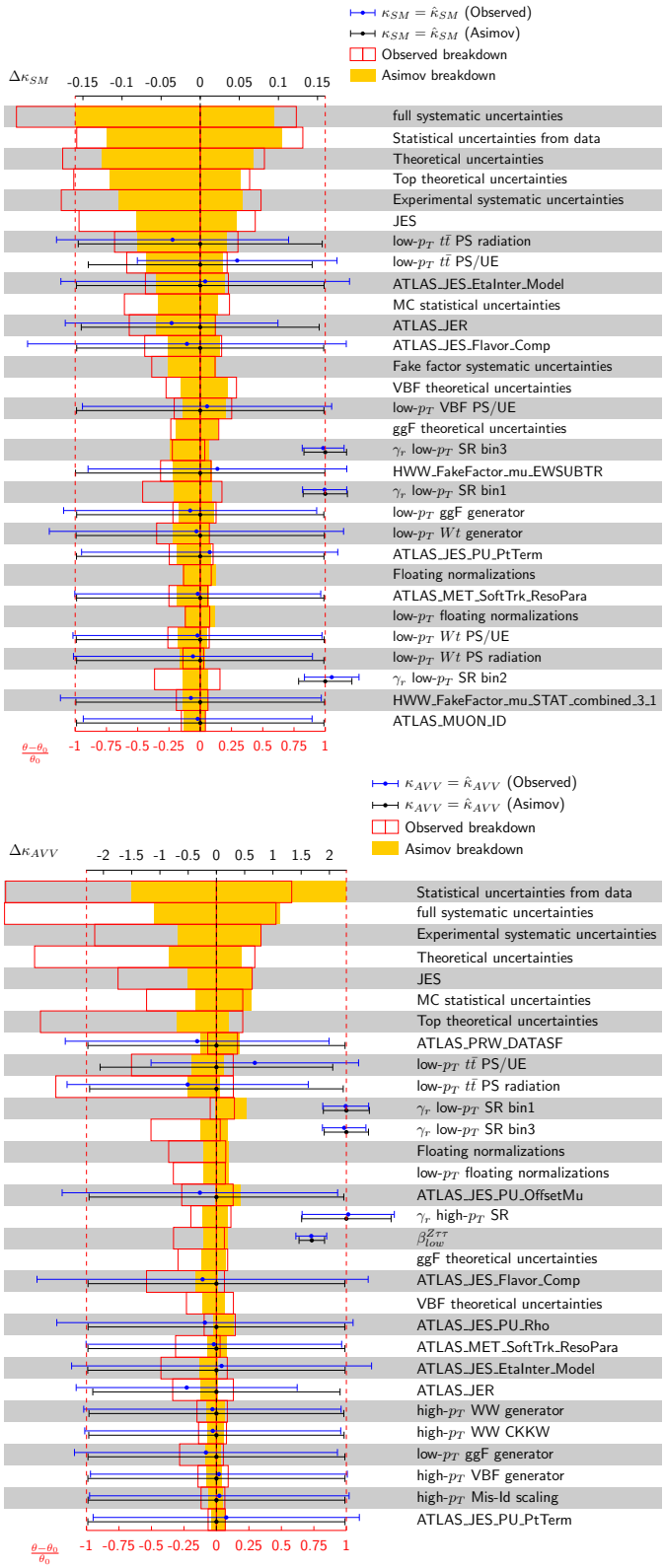


Figure 6.39.: Breakdown of the total uncertainty on κ_{SM} (top) and κ_{AVV} (bottom) into the individual components and nuisance parameters, shown for the fit to observed data (yellow) and the Asimov data set (red). Only the leading sources of uncertainty are shown. The post-fit values and uncertainties of the individual parameters are also shown in either case, normalised to their respective pre-fit constraint where applicable.

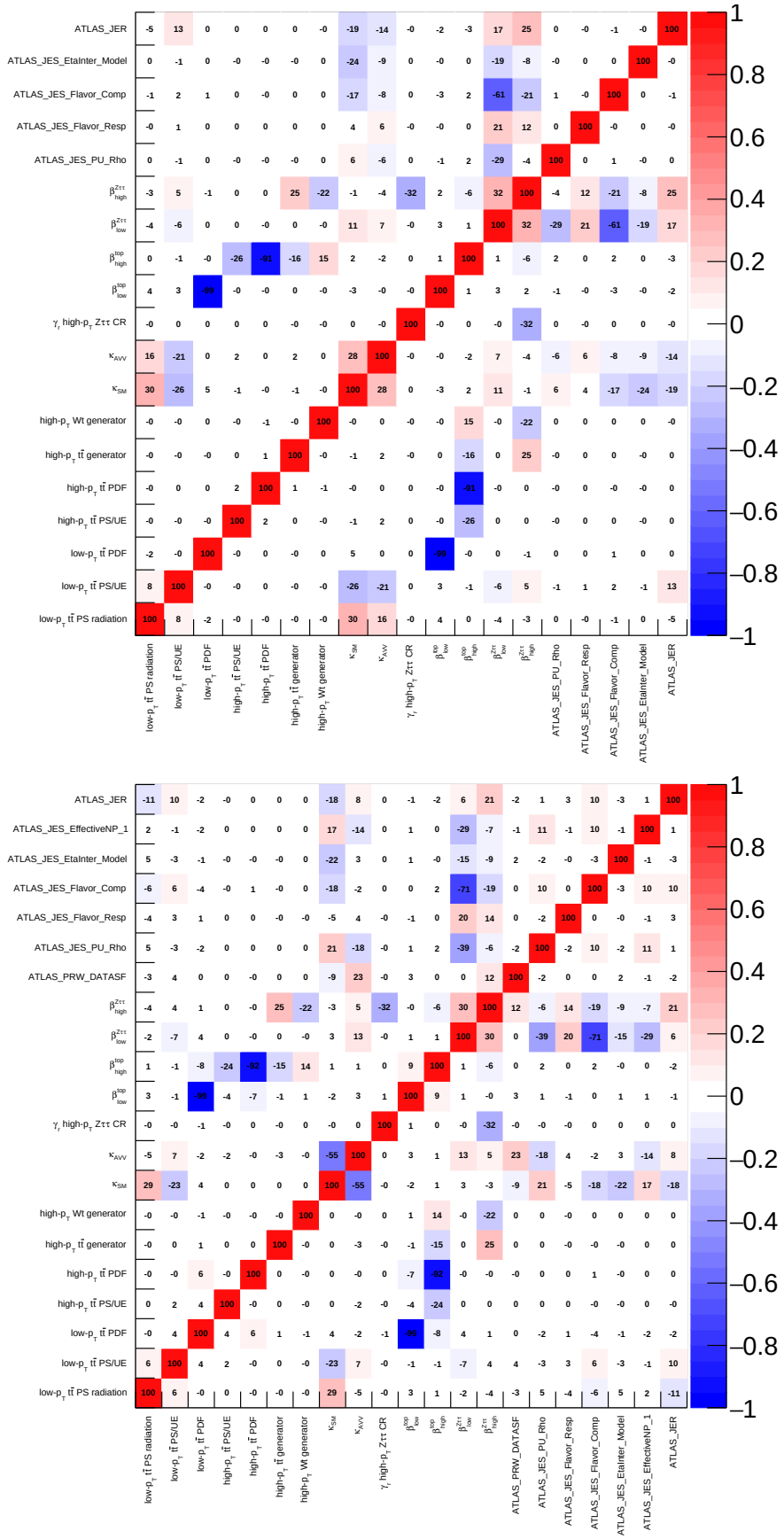


Figure 6.40.: Correlations between nuisance parameters observed in the likelihood fit for the Asimov data set (top) and the experimental data (bottom). Only parameters with correlations above 20% are shown.

The post-fit background strength parameters $\boldsymbol{\beta}$ in Table 6.23 resemble the normalisation factors determined from the control regions that are evaluated simultaneously in the fit. They are compatible within the uncertainties with the values shown in Table 6.10 derived from the independent normalisation factor calculation with the Matrix method. The differences between the single normalisation factors can be attributed to the indirect influence of and correlations with constrained nuisance parameters.

Table 6.23.: Background normalisation factors obtained from the likelihood fit to the observed data, including the total uncertainty propagated from all sources.

Control regions	Top	$Z \to \tau\tau$
low-p_{T}	$1.046^{+0.004}_{-0.016}$	$0.730^{+0.012}_{-0.007}$
high-p_{T}	$0.859^{+0.004}_{-0.007}$	$0.750^{+0.005}_{-0.017}$

Two floating coupling parameters enable one and two dimensional likelihood scans. In case of one dimensional scans the second κ parameter is profiled. Figure 6.41 shows the likelihood scans for $\kappa_{\mathrm{SM}}\cos(\alpha)$ and $\kappa_{\mathrm{AVV}}\sin(\alpha)$, respectively, where the CP angle parameter α is fixed at $\cos(\alpha) = \sin(\alpha) = 1/\sqrt{2}$. The SM Asimov data set yields the expected graph and the physics data the observed one. In the case of the $\kappa_{\mathrm{SM}}\cos(\alpha)$ scan the graph flattens fast to values less than 1 and increases much faster for values larger than 1. This is due to the fact that for small κ_{SM} values the event yields decreases and reaches 0 for $\kappa_{\mathrm{SM}} = 0$ at which point only an increased κ_{AVV} coupling component is able to fulfil the physics data set event yield expectation. Going to large κ_{SM} values the event yield of the signal processes grows quickly even when κ_{AVV} has no contribution. Thus the expected event yield from the input data set is less compatible with high κ_{SM} values resulting in a higher likelihood.

The SM expected likelihood scan over $\kappa_{\mathrm{AVV}}\sin(\alpha)$ is nearly symmetrical as anticipated. Here, equal positive and negative κ_{AVV} values lead to the same event yield due to the cancellation of the CP-even SM and CP-odd AVV interference terms in the event yield. Only the $\Delta\phi_{jj,\mathrm{sign}}$ distribution is sensitive to the interference between both couplings, which can be noted in the observed graph, where the experimental data prefers a small positive κ_{AVV} value. In this case without any sensitivity to the interference, the same negative value should have been equally likely and a symmetrical curve would emerge. Nevertheless, with the low statistics the influence of the $\Delta\phi_{jj,\mathrm{sign}}$ distribution in the low-p_{T} SR is not very large, but will be more and more relevant for future measurements with larger statistics.

In the two dimensional likelihood scan in Figure 6.42 the sign ambiguity of the κ_{AVV} couplings is visible, too. For smaller κ_{SM} values larger absolute κ_{AVV} couplings are preferred to compensate the decreasing event yield from the SM coupling. Overall a good agreement of the observed data to the SM prediction is

apparent. Only the 68% confidence contour is shown here, because the likelihood saturates to higher coupling values and remains flat. The reason is the increasing statistical uncertainty propagated through the morphing functions. At these high coupling values the parameters for the input morphing samples are not chosen well anymore. The next section will address this effect in more detail.

From the likelihood scans the measured coupling values can be extracted to

$$\begin{aligned}\kappa_{\mathrm{SM}}\cos(\alpha) &= 0.948\,^{+0.127}_{-0.200}(\mathrm{tot.}) \\ &= 0.948\,^{+0.103}_{-0.111}(\mathrm{stat.})\ \ ^{+0.055}_{-0.126}(\mathrm{exp.sys.})\ ^{+0.058}_{-0.124}(\mathrm{theo.}) \\ &\qquad\quad ^{+0.010}_{-0.015}(\mathrm{norm.})\ ^{+0.013}_{-0.044}(\mathrm{fake.})\ \ ^{+0.026}_{-0.069}(\mathrm{MC\ stat.}),\end{aligned}$$

and

$$\begin{aligned}\kappa_{\mathrm{AVV}}\sin(\alpha) &= 1.652\,^{+1.202}_{-3.741}(\mathrm{tot.}) \\ &= 1.652\,^{+0.943}_{-2.638}(\mathrm{stat.})\ \ ^{+0.558}_{-1.523}(\mathrm{exp.sys.})\ ^{+0.481}_{-2.276}(\mathrm{theo.}) \\ &\qquad\quad ^{+0.117}_{-0.598}(\mathrm{norm.})\ ^{+0.072}_{-0.414}(\mathrm{fake.})\ \ ^{+0.330}_{-0.874}(\mathrm{MC\ stat.}),\end{aligned}$$

where the individual contributions to the uncertainties have been disentangled into the component stemming from statistical uncertainties on the data (stat.), from experimental systematic uncertainties (exp. sys.), from systematic uncertainties from theory (theo.), from the floating background normalisation factors (norm.), from the fake estimate (fake), and from statistical uncertainties arising from the finite size of the MC samples (MC stat.). Both coupling values are compatible with the SM and no evidence of EFT contribution in this measurement is seen.

The post-fit distributions are obtained by applying the minimised floating fit parameters to the event rates of the respective MC samples. Compared to the pre-fit distribution in Figure 6.36, the post-fit distributions in Figure 6.43 show no data excess anymore and a good agreement in both signal regions.

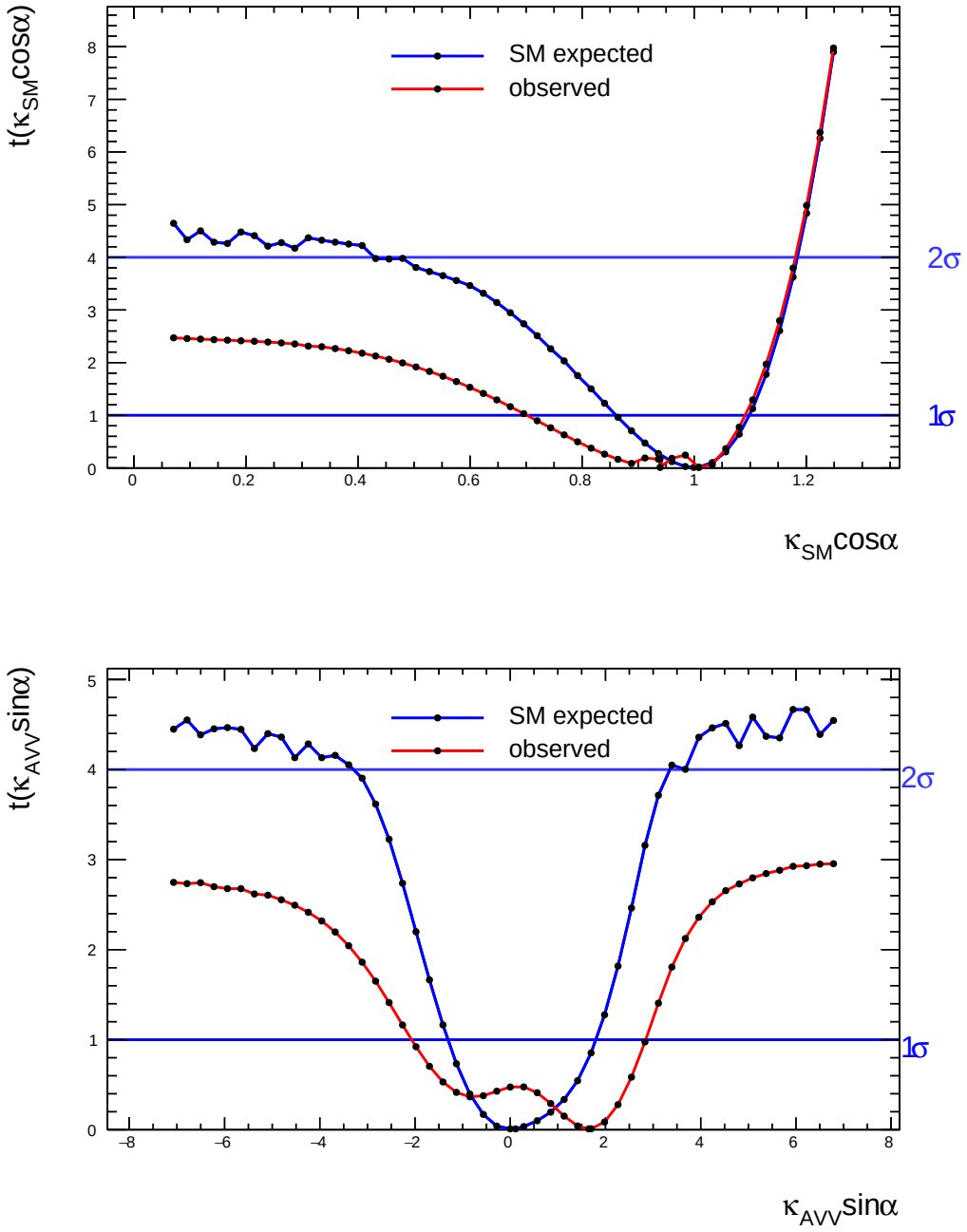


Figure 6.41.: The likelihood scans of $\kappa_{\mathrm{SM}} \cos(\alpha)$ (top) and $\kappa_{\mathrm{AVV}} \sin(\alpha)$ (bottom) using the SM Asimov data set (SM expected) and the experimental data (observed). The respective other κ parameter and all nuisance parameters are profiled. The CP angle α is fixed at $\frac{\pi}{4}$.

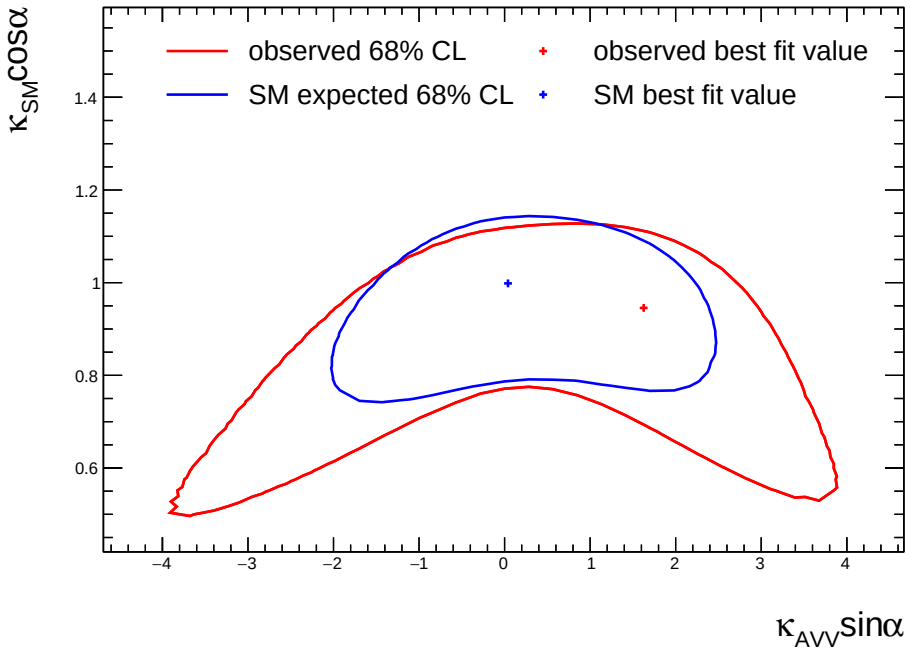


Figure 6.42.: The simultaneous likelihood scan of $\kappa_{\mathrm{SM}}\cos(\alpha)$ and $\kappa_{\mathrm{AVV}}\sin(\alpha)$ showing the 68% confidence level contours and best fit values of the SM Asimov data set (SM expected) and the experimental data (observed). All nuisance parameters are profiled. The CP angle α is fixed at $\frac{\pi}{4}$.

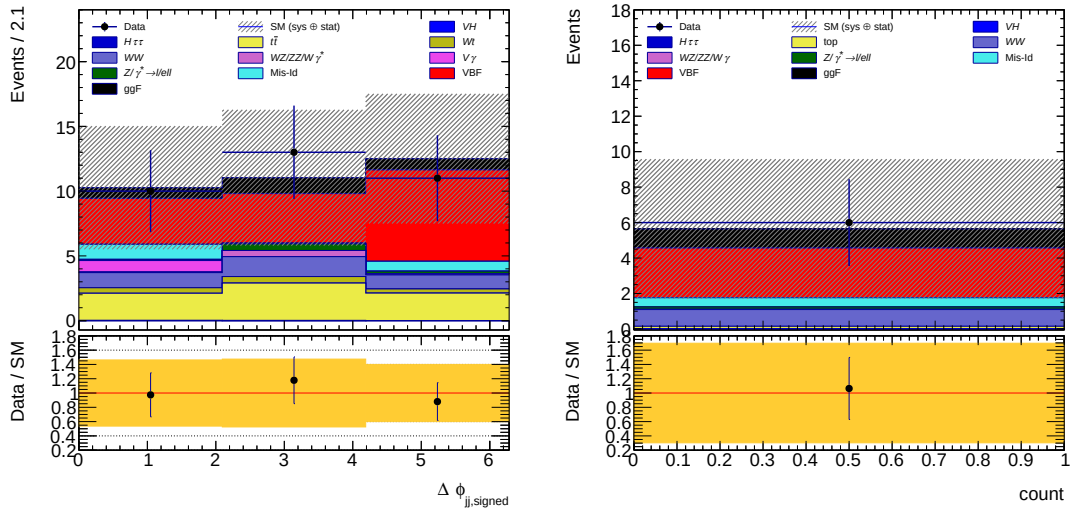


Figure 6.43.: Post-fit signal region distributions $\Delta\phi_{jj,\mathrm{sign}}$ in low-p_{T} SR (left) and event yield count in high-p_{T} SR (right).

6.11.2 Measurement of SM, HVV and H∂V

The simultaneous measurement of three coupling parameters is a lot more challenging due to the higher dimensionality as previously explained in Section 5.5.1. And indeed huge cancellation of the weights in each morphing function renders the unconditional fit to the Asimov and to the observed data set impossible. To visualise the emerging problems a minimal version of the likelihood is constructed, where only the total event yield in the low-p_{T} SR with the total background including the SM ggF process in this region is considered and no systematic uncertainty is incorporated. Thus only one morphing function and the effects of its high statistical MC uncertainties can be analysed. The likelihood simplifies to

$$\begin{aligned}\mathcal{L}(\boldsymbol{\kappa},\boldsymbol{\gamma}) =& P\left(N_{\mathrm{SRlow}}\middle|\nu_{\mathrm{VBF_{SRlow}}}(\boldsymbol{\kappa},\boldsymbol{\gamma}_s)+\gamma_{\mathrm{SRlow}}\sum_{\phi\in\mathrm{bkgs,ggF_{SM}}} b^{\phi,0}_{\mathrm{SRlow}}\right)\cdot \\ & P\left(M_{\mathrm{SRlow}}\middle|\gamma_{\mathrm{SRlow}}\tau_{\mathrm{SRlow}}\right)\prod_{s\in\mathrm{VBF_{SRlow}}} P\left(M_s\middle|\gamma_s\tau_s\right),\end{aligned} \tag{6.11.1}$$

with the expected VBF event yield in the signal region given by the morphing function

$$\nu_{\mathrm{VBF_{SRlow}}}(\boldsymbol{\kappa},\boldsymbol{\gamma}_s) = \sum_{s\in\mathrm{VBF_{SRlow}}} \gamma_s w_s(\boldsymbol{\kappa})\, s^s_{\mathrm{SRlow}}. \tag{6.11.2}$$

First, the relative statistical uncertainty $\frac{\sigma_N}{|N|}$ of the morphing function over the parameter space is examined. The uncertainty is thereby determined using toys. Each input sample is randomly varied by a normal distribution with the actual nominal expected event yield s^s_{SRlow} as mean value and the statistical MC uncertainty on the yield as standard deviation. In total 100 000 variations are generated for each parameter point and a grid scan is performed over two coupling parameters while the third is fixed to a particular value.

Figure 6.44 shows for three fixed κ_{SM} values the relative uncertainty in the κ_{HVV}-$\kappa_{\mathrm{H\partial V}}$ coupling space, as well as the positions at which the input samples have been generated. At these input parameters the uncertainties match exactly the MC statistical uncertainties of the corresponding generated input sample. Except for a small area around the SM values the uncertainty increases further away from the input parameter points. The most striking characteristics are distinct surfaces of very high relative uncertainties seen as yellow bands in the figures. These areas appear when the expected event yield drops to zero and the relative uncertainty grows to infinity. Inside these surfaces the event rate prediction of the morphing function, although still compatible with a positive yield within the uncertainty, becomes negative.

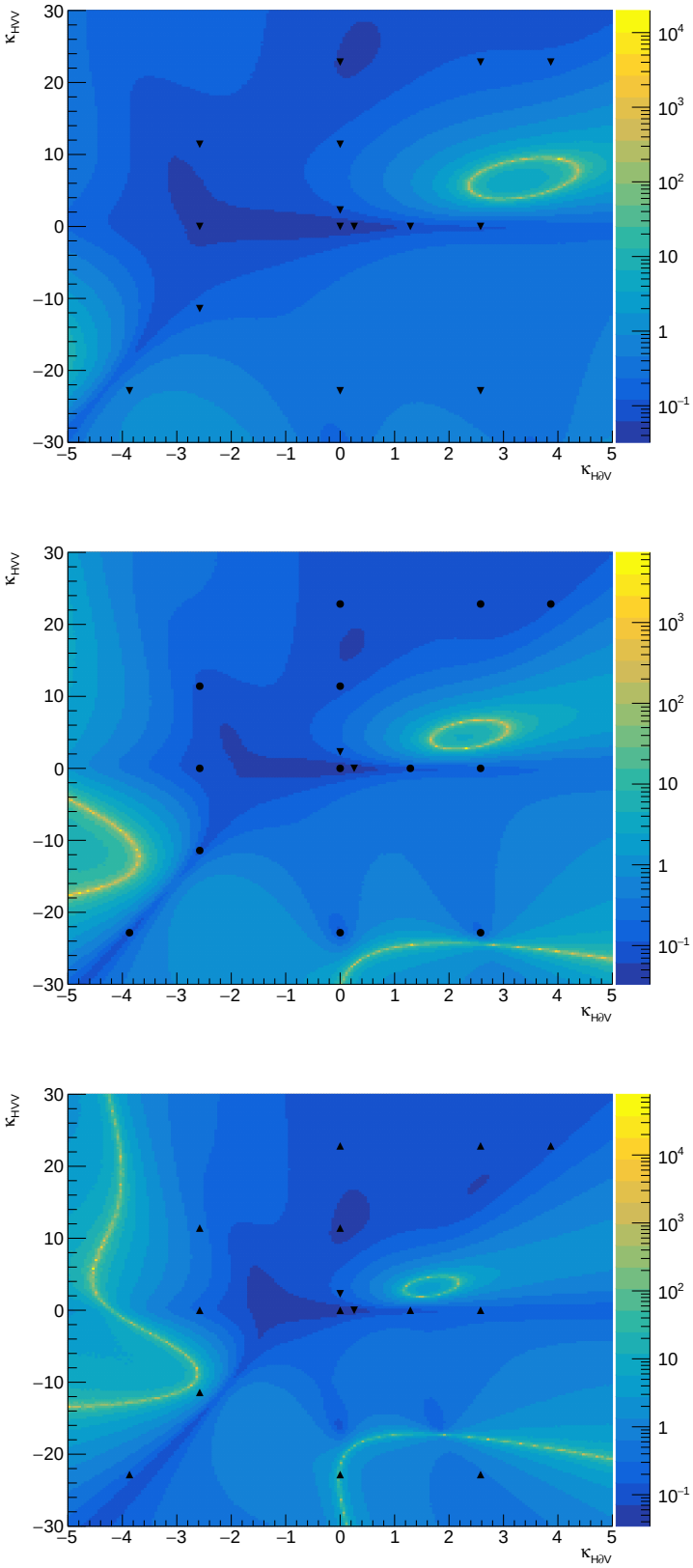


Figure 6.44.: The relative uncertainty of the VBF event yield expectation calculated by the morphing function in Equation 6.11.2 over the κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ coupling phase space for fixed values for κ_{SM} at 2 (top) $\sqrt{2}$ (middle) and 1 (bottom). The black markers show the positions at which the morphing input samples have been generated, where ▲/●/▼ indicates that the κ_{SM} value of the input sample is above/at/below the illustrated plain.

To have a closer look, Figure 6.45 displays three graphs that show the event yield with uncertainties as function of κ_{HVV} with a fixed κ_{SM} value at $\sqrt{2}$ and fixed values of $\kappa_{\mathrm{H}\partial\mathrm{V}}$ at 2.58, 0 and -5. The vertical lines indicate the positions of the input samples, at which the event yields and uncertainties are determined by the MC generator and the event selection of the low-p_{T} SR. They determine the inter- and extrapolation of the predicted event yield. Since the morphing function has a polynomial structure, small variations of the input event rate may lead to undesired effects visible in the upper two graphs. There, the negative extrapolation to $\kappa_{\mathrm{HVV}} \simeq -30$ as well as the interpolation of the upper graph at around $\kappa_{\mathrm{HVV}} \simeq 5$ leads to a negative event rates. In addition, the lower graph shows large uncertainties further away from the input parameters.

The effect of these large uncertainties and especially the negative event rate predictions affect the negative logarithm of the likelihood presented in Figure 6.46. Here, the SM Asimov data set is used as pseudo-data input. All nuisance parameters and κ_{SM} have been fixed to their nominal values such that only the effects of the event yield is apparent. The negative logarithm of the likelihood reaches always its minimal value, if the VBF plus total background yield is exactly equal to the SM Asimov event rate, which is by construction precisely at $\kappa_{\mathrm{HVV}} = \kappa_{\mathrm{H}\partial\mathrm{V}} = 0$. However, not only at the SM parameter point the minimal value is present, but likewise near locations, where the VBF event rate prediction drops to zero. At these positions the morphing function has to cross the SM VBF event rate. Therefore, the same structures emerge in the Figures 6.46 as have been seen in the Figures 6.44. Just before the very large relative event yield uncertainties (yellow contours), the negative logarithm of the likelihood becomes minimal. Further into the negative VBF event yield prediction the $-\log\mathcal{L}$ rises again until the VBF plus total background event rate becomes negative and $-\log\mathcal{L}$ has no solution seen in the white areas in the middle and lower plots.

By profiling the nuisance parameters $\boldsymbol{\gamma}_s$ associated to the MC statistical uncertainties of the VBF morphing input samples the negative logarithm of the likelihood keeps in general the same set of minima and even lowers its $-\log\mathcal{L}$ values at other parameter regions as can be seen in the top Figure 6.47. Small changes in the broad subsidiary Poisson terms have small increasing effects on the likelihood but large minimising effect on the signal region Poisson term. Additional profiling of the κ_{SM} parameter aggravates this effect leading to the major κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ phase space at the $-\log\mathcal{L}$ minimum (cf. bottom Figure 6.46).

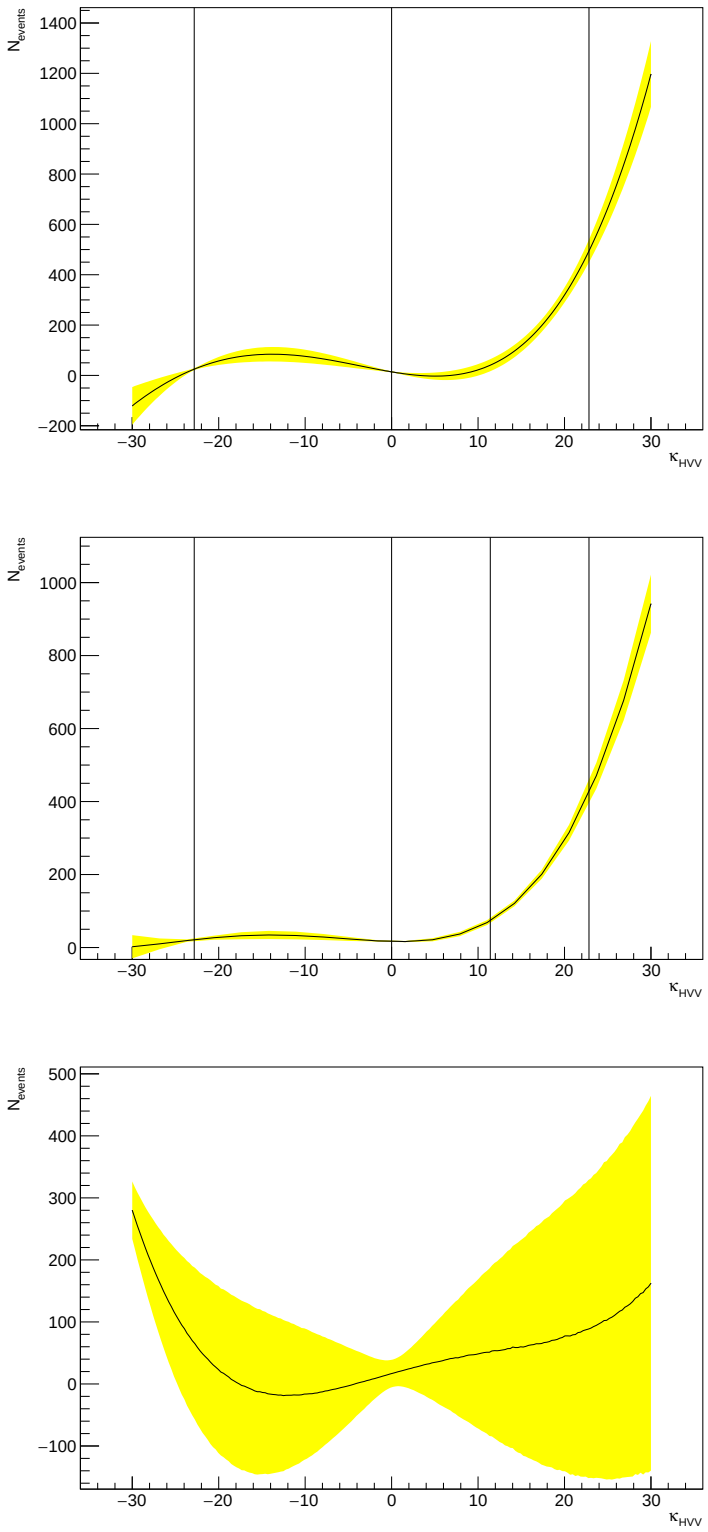


Figure 6.45.: Predicted event yield with uncertainties of the VBF morphing function as function of κ_{HVV} with fixed κ_{SM} value at $\sqrt{2}$ and fixed $\kappa_{\mathrm{H\partial V}}$ values at 2.58 (top), 0 (middle) and -5 (bottom). The vertical lines indicate the positions of the morphing input samples. The uncertainties are determined using toy variations of the MC statistical uncertainty of the signal samples.

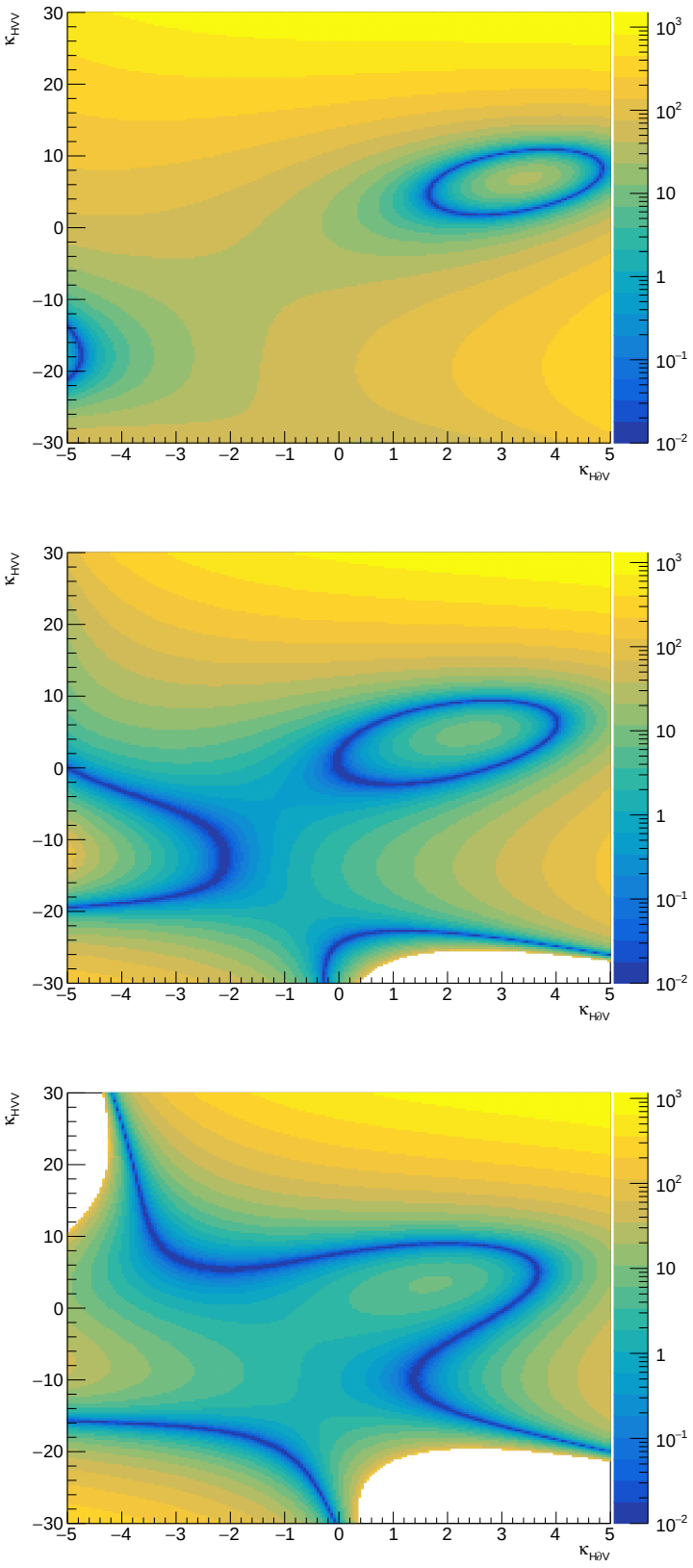


Figure 6.46.: Negative logarithm of the likelihood for the SM Asimov data set in the κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ coupling phase space with fixed κ_{SM} at $\sqrt{2}$ and fixed nuisance parameters at $\boldsymbol{\gamma} = \mathbf{1}$. For the logarithmic scale of $-\log\mathcal{L}$ the minimal value is set to 0.01.

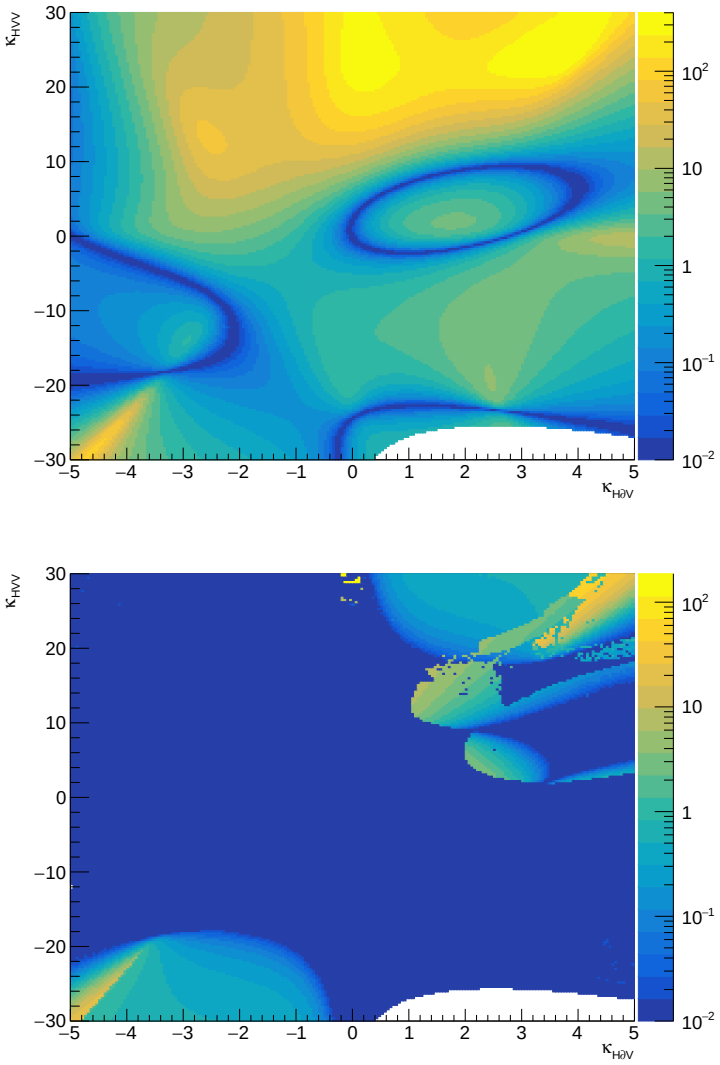


Figure 6.47.: Negative logarithm of the likelihood for the SM Asimov data set in the κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ coupling phase space with profiled nuisance parameters $\boldsymbol{\gamma}$ and fixed κ_{SM} at $\sqrt{2}$ (top) and additional profiled κ_{SM} (bottom). For the logarithmic scale of $-\log\mathcal{L}$ the minimal value is set to 0.01.

Evidently, a measurement is not feasible having these huge discrepancies of the event rate predictions to the ones a directly generated MC sample would predict, although they would be presumably compatible within the large uncertainties. Hence, despite the physically exact modelling of the coupling phase space within the morphing, the small MC statistical uncertainties of the input samples magnify in certain coupling phase space regions due to large cancellations of the morphing weights and lead to both very large uncertainties and bad event yield predictions. Especially domains with negative event yields render a measurement therefore impossible.

The most promising solution to yet enable such a measurement is an optimisation of the input coupling parameters at which the MC samples are generated in such a way, that the sum of the morphing weights times the respective cross sections of the input samples are as small as possible in the coupling parameter space of interest. The algorithm presented in Section 5.5.4 is designed especially for this purpose and may be used to repeat this or conduct similar future high-dimensional coupling measurements.

Still, many aspects remain subject to investigative research. For instance, given a set of coupling parameters it is not clear, that an optimal set of input parameters can generate a sufficient large area of small event rate uncertainty needed by the measurement. If not, other concepts or enhancements needs to be developed e.g. by dynamically adapting an optimal set of input samples as function of the coupling parameters or combinations of the morphing interpolation with other techniques.

7 Conclusions and Outlook

The discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations is surely one of the biggest scientific achievements of the recent years and has opened a new domain in physics that is sought to be explored. Thereby, not only the intrinsic properties of the new boson like mass and spin are of interest but as well coupling strengths to other elementary particles. In these couplings new potential insights may be found manifested in deviations to the SM predictions, which might hint to yet undiscovered particles even beyond the energy scales of the LHC.

7.1 Summary

Many experimental evidence, like cosmological observations exposing the nature of the dark matter, show the incompleteness of the SM and indicate that at higher energies the SM eventually will break down. In a bottom-up approach the SM seen as a low-energy effective theory can be systematically and model independently expanded with higher dimensional operations. Thereby, free coupling parameters emerge that are matter of experimental measurement.

7.1.1 Effective Lagrangian Morphing

The number of relevant free coupling parameters of such an effective field theoretical extension depend on the considered physics process and the number of operators included. Limiting factor in a measurement is thereby the computational power and hardware equipment to generate costly and time consuming MC samples for different parameter settings to continuously model observables in the parameter phase space. In this context, the novel technique of Effective Lagrangian Morphing was developed in the course of this thesis and has been included in the CERN Yellow report [7]. It 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.

7.1.2 Effective Coupling Measurement

For the first time a full analysis of the measurement of EFT coupling parameters in the $H \to WW$ channel has been performed using the method of Effective Lagrangian Morphing. The first $39.5\,\text{fb}^{-1}$ of pp collision data recorded at $\sqrt{s} = 13\,\text{TeV}$ by the ATLAS detector is used targeting a final state with leptonical W decays mainly focused at the VBF production process. Initially, the most sensitive coupling parameters constributing to the HVV vertex are identified as the CP-even κ_{HVV}, the CP-even derivative couplings $\kappa_{\text{H}\partial\text{V}}$, and the CP-odd couplings κ_{AVV}. The best choice for an observable to distinguish different EFT contributions is determined to be the azimuthal angle between the tagging jets, where the leading and subleading jets are sorted by pseudorapidity, $\Delta\phi_{jj,\text{sign}}$. A highly optimized event selection, including the utilization of boosted decision trees and sophisticated techniques of background estimation have been applied to seperate the background processes from the signal process without loosing the sensitivity to the effects of the EFT operators on the observable and event yields in the final signal regions.

The coupling parameters that scales the SM constribution $\kappa_{\text{SM}}\cos(\alpha)$ and the CP-odd coupling $\kappa_{\text{AVV}}\sin(\alpha)$ are probed using the maximum likelihood method both simultaneously and one at a time assuming that all other EFT couplings vanish and with fixed CP angle $\alpha = \frac{\pi}{4}$. The measured best-fit values with the 68% confidence level are

$$\kappa_{\text{SM}}\cos(\alpha) = 0.948^{+0.127}_{-0.200}$$
$$\kappa_{\text{AVV}}\sin(\alpha) = 1.652^{+1.202}_{-3.741},$$

showing no significant deviation from the SM prediction.

The simultaneous fit of κ_{SM}, κ_{HVV} and $\kappa_{\text{H}\partial\text{V}}$ demonstrates the challenges to the Effective Lagrangian method facing high dimensionalities. Large cancelations of the weights in each morphing function renders the unconditional fit to the Asimov and to the observed data set impossible.

7.2 Outlook

The full Run-2 data set consists of $147\,\text{fb}^{-1}$ recorded by the ATLAS detector and in the future about $300\,\text{fb}^{-1}$ are planned to be recorded in the upcoming Run-3 of the LHC until end of 2023. With this amount of physics data, precision measurements of coupling parameters from EFT expansions in the Higgs sector will become feasible. Statistically limited analyses like the one presented in this thesis will benefit greatly from the increased data set.

Future measurements may not only include well-chosen kinematic observables, which can only yield precise bounds along individual directions in the parameter space, and often lead to weak constraints in other directions. With an increasing amount of recorded physics data the sensitivity in measurements to multiple parameters can be substantially improved by using many such variables including their correlations or even fully differential cross sections of the processes.

With a larger data set, a measurement of one individual scattering process may also comprise all sensitive EFT operators, which require simultaneous inference over a possibly large multi-dimensional parameter space. A desired goal would be a combination of all measurements of processes sensitive to all effective couplings to gain the full constraining or discovery power of BSM effects to the EFT expansion of the SM.

In the future the limiting factor for such an appreciable measurement will potentially lie in the computing power and data storage capacity, since it becomes prohibitively expensive to scan for more than a few parameters using just simulated events with simple interpolation methods. The novel technique of Effective Lagrangian Morphing will, in connection with other state-of-the-art techniques, enable such desirable measurements and derivations of limits for coupling parameters in model-independent effective theories.

With a proof-of-concept analysis conducted in this thesis and additionally employing advancements of dimensionality reduction and uncertainty minimization presented here, this method can be applied to a wide variety of physics measurements in the future, deciphering the nature of what might lie beyond the Standard Model of particle physics.

A BDT Supplementary Distributions

In the measurement of the effective couplings in Chapter 6 BDTs are used to define the final signal regions (cf. Section 6.6). Here, supplementary distributions of the BDT inputs are shown to examine the data to MC agreement.

A.1 BDT Validation Regions

The lower range of the BDTs discriminant distributions, which are removed from the final signal regions, can be used as validation regions to examine the data to MC agreement. In the Figure A.1 the validation plots are shown for the BDT_{low} input variables and in Figure A.2 for the BDT_{high} input variables, respectively. The data to MC agreement is reasonable good. A small mismodelling in the distribution of the number of jets for BDT_{high} is seen. Presumably, the inclusion of the theory systematic uncertainties comprise most of the mismodelling.

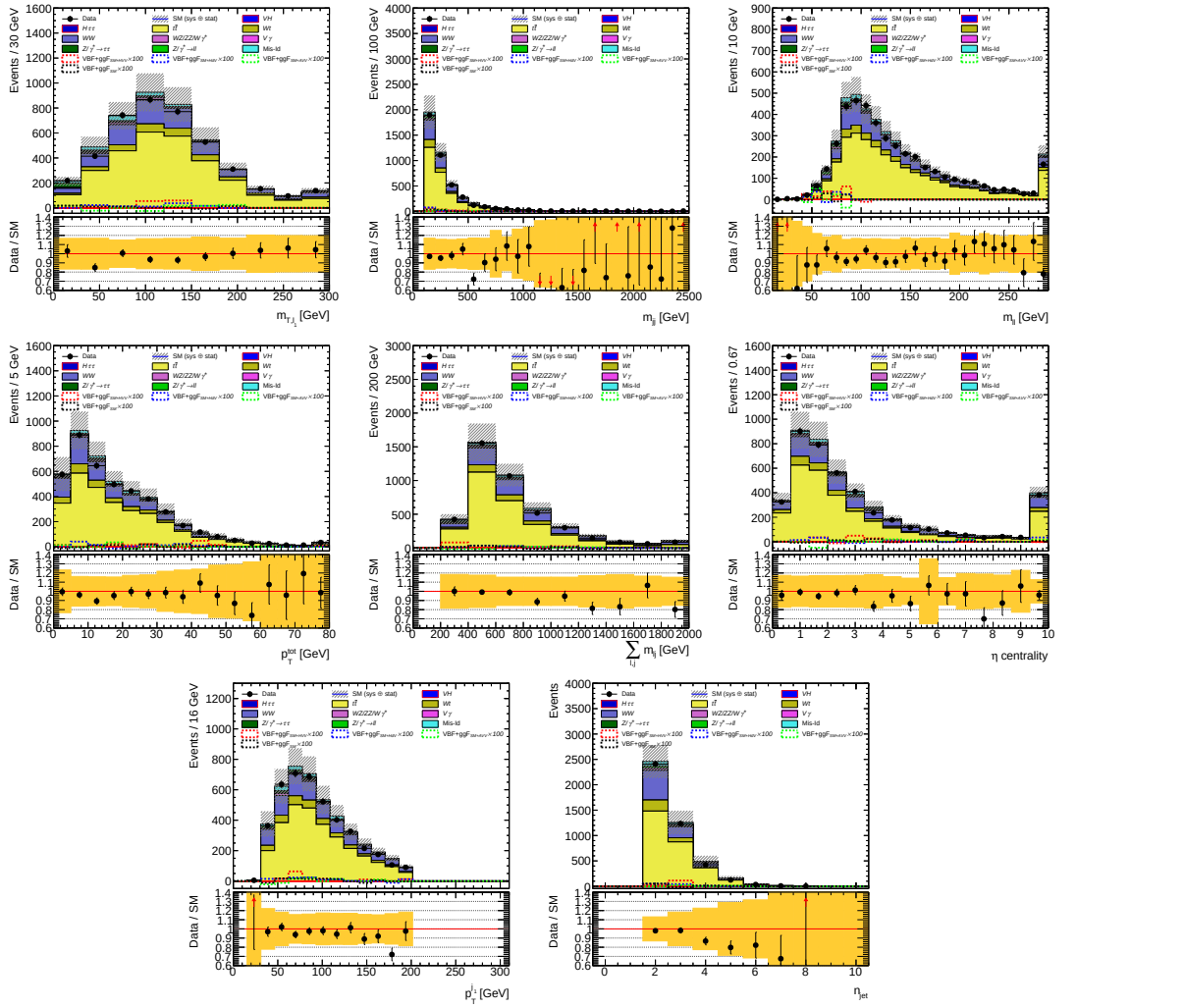


Figure A.1.: Distributions of the $\mathrm{BDT_{low}}$ input variables in the BDT validation region ($\mathrm{BDT_{low}} < -0.5$): $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\mathrm{tot}}$, $\sum_{\ell,j} m_{\ell j}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{j_1}$ and n_{jet}. The yellow band represents MC statistics and all detector systematics.

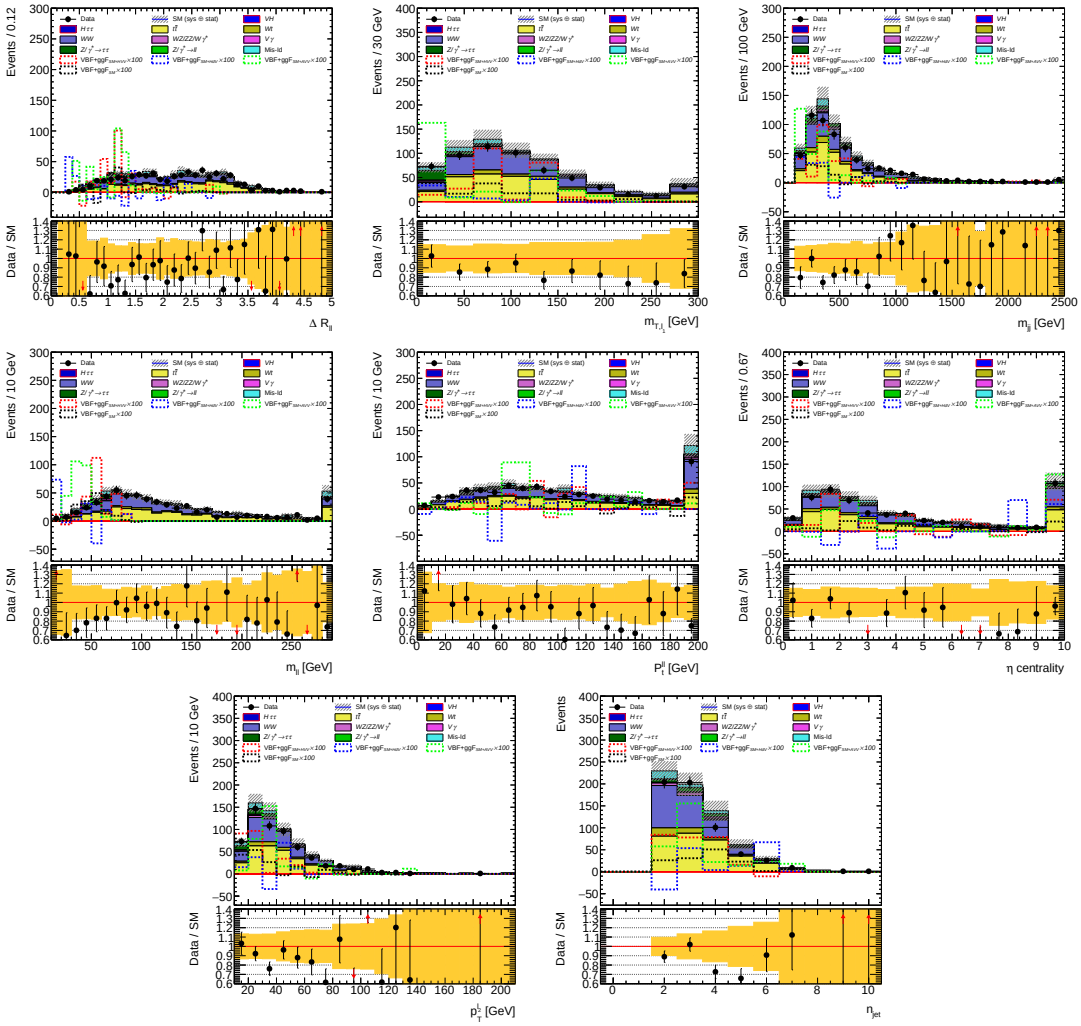


Figure A.2.: Distributions of the $\mathrm{BDT_{high}}$ input variables in the BDT validation region ($\mathrm{BDT_{high}} < -0.2$): $\Delta R_{\ell\ell}$, $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{\ell_2}$ and n_{jet}. The yellow band represents MC statistics and all detector systematics.

A.2 BDT Inputs in the Control Regions

Two control regions per category ($p_{\mathrm{T}}^{j_1} \leqslant 200\,\mathrm{GeV}$) are defined in Section 6.7.1, in which the same BDTs are applied as have been constructed for the signal regions.

A.2.1 Top Control Regions

The input variables for $\mathrm{BDT}_{\mathrm{low}}$ are shown in Figure A.3 and for $\mathrm{BDT}_{\mathrm{high}}$ in Figure A.4 for the top control region. A good data to MC agreement is observed for most distributions. As for the validation region in $\mathrm{BDT}_{\mathrm{high}}$ a small mismodelling is oberved for the n_{jet} distribution and additionally a skewness in $p_{\mathrm{T}}^{\ell\ell}$ for $\mathrm{BDT}_{\mathrm{high}}$ between 60 and 150 GeV. The incorporation of the theory systematic uncertainties, which are especially large for the top background, is expected to cover these mismodellings in the uncertainty.

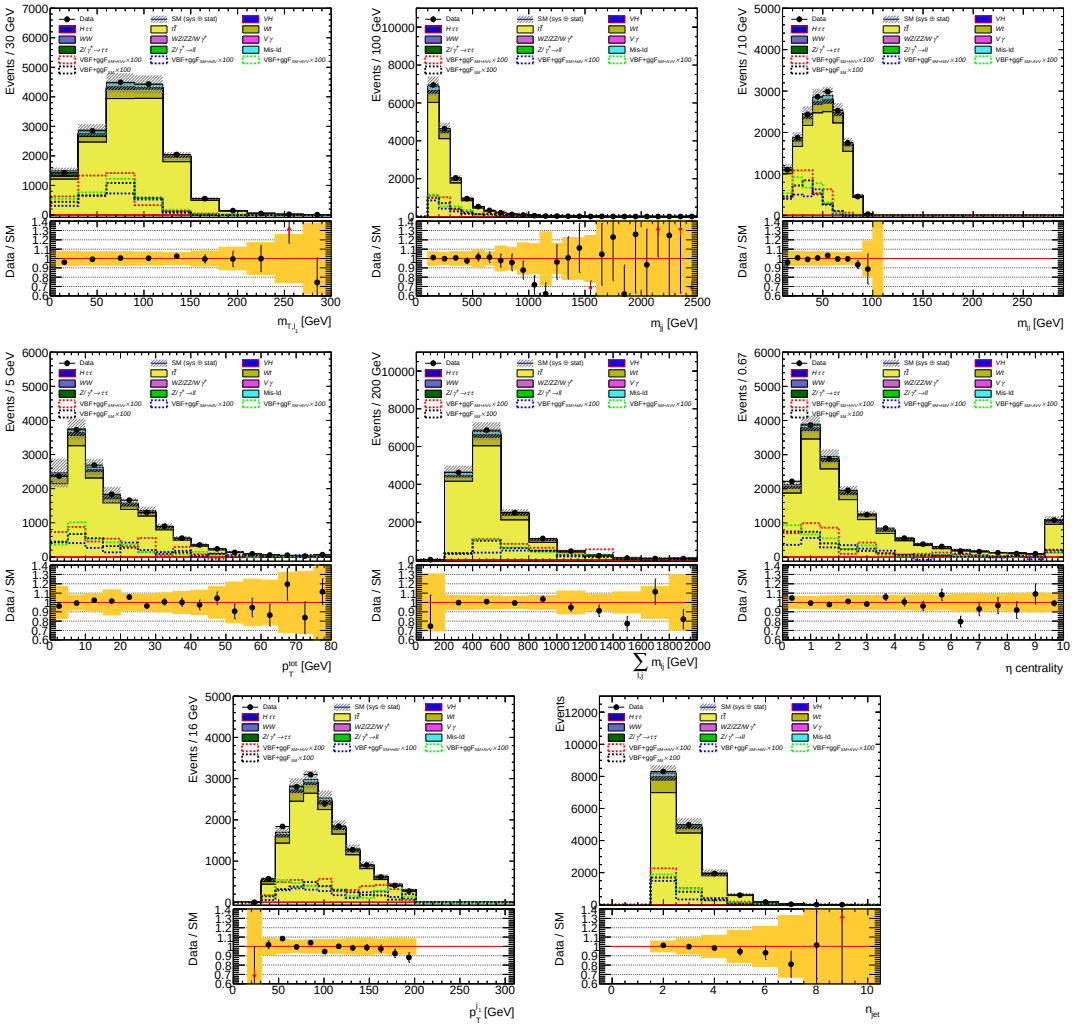


Figure A.3.: Distributions of the $\mathrm{BDT}_{\mathrm{low}}$ input variables in the low-p_{T} top control region: $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\mathrm{tot}}$, $\sum_{\ell,j} m_{\ell j}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{j_1}$ and n_{jet}. The yellow band represents MC statistics and all detector systematics.

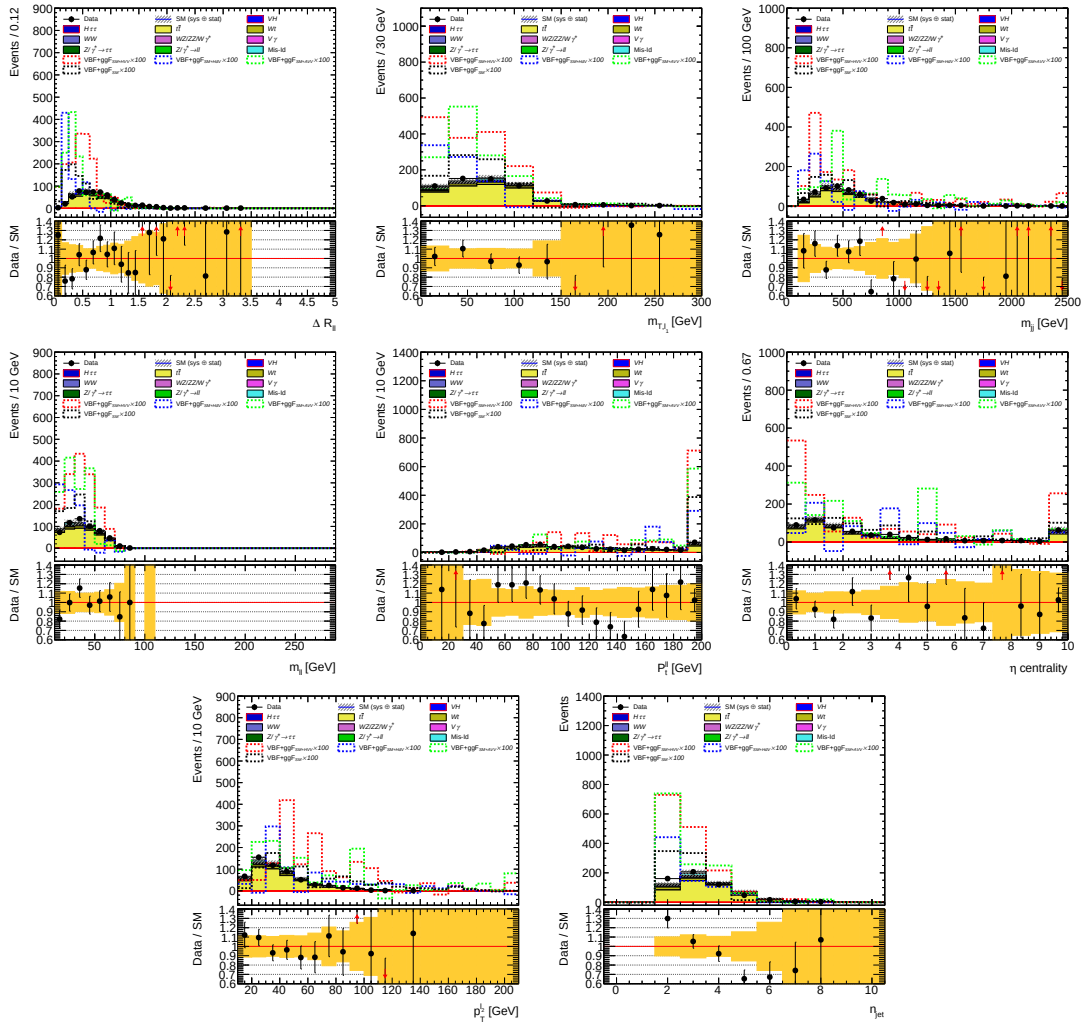


Figure A.4.: Distributions of the $\mathrm{BDT_{high}}$ input variables in the high-p_{T} top control region: $\Delta R_{\ell\ell}$, $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{\ell_2}$ and n_{jet}. The yellow band represents MC statistics and all detector systematics.

A.2.2 $Z \to \tau\tau$ Control Regions

The input obervables in the $Z \to \tau\tau$ control region are shown in Figure A.5 for BDT_{low} and A.6 for BDT_{high}. The data to MC agreement is reasonable good.

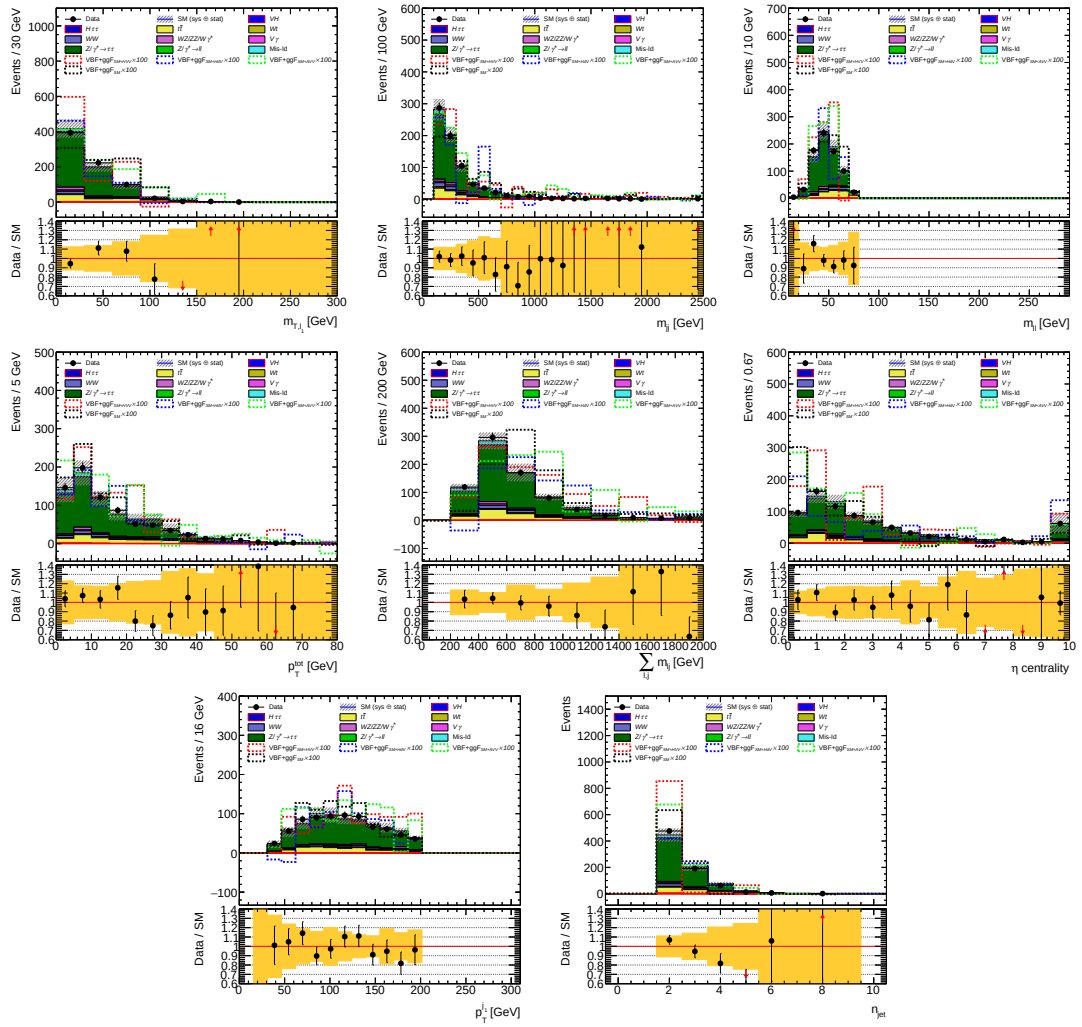


Figure A.5.: Distributions of the BDT_{low} input variables in the low-p_{T} $Z \to \tau\tau$ control region: $m_{\text{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\text{T}}^{\text{tot}}$, $\sum_{\ell,j} m_{\ell j}$, $\eta_{\text{lep}}^{\text{centrality}}$, $p_{\text{T}}^{j_1}$ and n_{jet}. The yellow band represents MC statistics and all detector systematics.

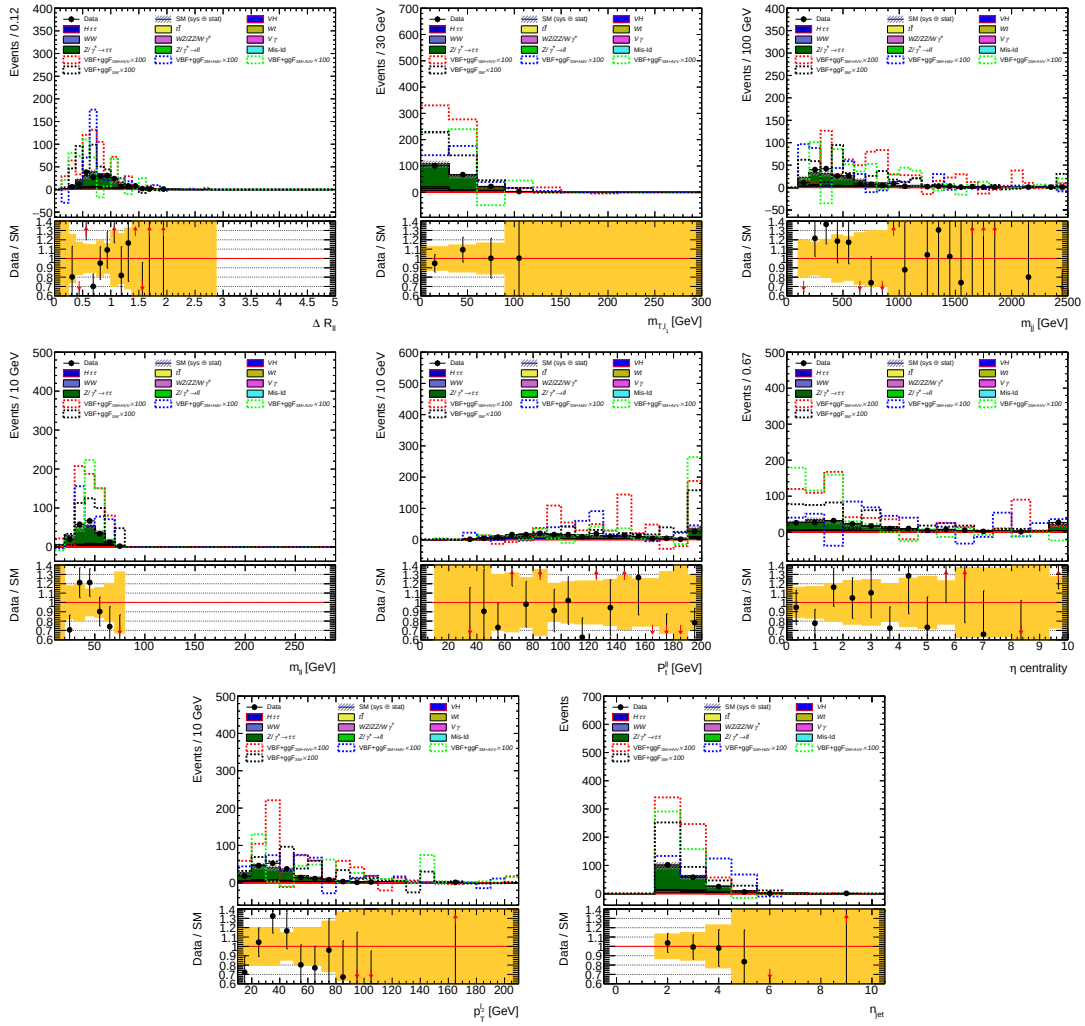


Figure A.6.: Distributions of the $\mathrm{BDT_{high}}$ input variables in the high-p_{T} $Z \to \tau\tau$ control region: $\Delta R_{\ell\ell}$, $m_{\mathrm{T}}^{\ell_1}$, m_{jj}, $m_{\ell\ell}$, $p_{\mathrm{T}}^{\ell\ell}$, $\eta_{\mathrm{lep}}^{\mathrm{centrality}}$, $p_{\mathrm{T}}^{\ell_2}$ and n_{jet}. The yellow band represents MC statistics and all detector systematics.

List of Figures

2.1. Hadron multiplets. 7
2.2. Higgs potential. 10
2.3. Example Feynman diagrams for Higgs boson production mechanisms. 13
2.4. The SM Higgs boson production cross sections as a function of the LHC center of mass energy. 14
2.5. SM predictions for the leading decay branching fractions of a Higgs boson. 16
2.6. Results of the cross section times branching ratio for each specific Higgs channel for the combined ATLAS measurements. 17
2.7. Top-down vs. bottom-up picture of effective field theories. 18
2.8. Feynman diagrams for β decay in the effective Fermi interaction and in the Standard Model. 20
2.9. Feynman diagram for a general $2 \to n$ particles process. 26
2.10. Typical pp collision event. 27
2.11. Comparison of different PDF sets. 28
2.12. Illustration of the factorization theorem. 29

3.1. The CERN accelerator complex. 33
3.2. LHC overview schematics. 34
3.3. Cut-away view of the ATLAS detector. 35
3.4. Coordinate systems of the ATLAS detector. 36
3.5. ATLAS Inner Detector. 37
3.6. Track in the Inner Detector. 38
3.7. Schematic view of the ATLAS calorimeter system. 40
3.8. Schematic layout of the muon spectrometer. 42
3.9. The ATLAS trigger and data acquisition system in Run-2. 44

4.1. Track reconstruction efficiency. 46
4.2. Average number of interactions per bunch crossing for ATLAS. . . . 47
4.3. Vertex reconstruction efficiency 48
4.4. Number of vertices vs. μ 48
4.5. Schematic view of the electron reconstruction and identification . . 49
4.6. Electron reconstruction and identification efficiency. 50

4.7. Electron identification efficiency. 51
4.8. Jet-shape comparison for different jet reconstruction algorithms. . . 55
4.9. JVT likelihood. 57
4.10. TST $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution. 60

5.1. Discrete and continuous representation of a two-dimensional likelihood. 64
5.2. Comparison of $\Delta\phi_{ll}$ and $p_{T,\mathrm{tot}}$ reweighted distributions to validation distributions. 67
5.3. Distribution of the difference between the weights predicted by the two Matrix Element reweighting strategies. 68
5.4. Example of integral morphing. 70
5.5. Example of moment morphing. 72
5.6. Feynman diagram of an exemplary process where two coupling parameters appear in one vertex. 76
5.7. Schematic depiction of a simple Effective Lagrangian Morphing. . . 78
5.8. Feynman diagram of an exemplary process where two coupling parameters appear in two vertices. 79
5.9. Number of input samples required for Effective Lagrangian Morphing. 81
5.10. Impact on the uncertainty for different morphing bases. 83
5.11. Distribution of $p_{\mathrm{T}}^{\mathrm{tot}}$ for the input and validation samples. 87
5.12. Morphing validation for the ggF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. . . . 88
5.13. Expected number of events as function of the free parameters κ_{SM} and κ_{AVV} and the respective relative uncertainty on the cross section. 89
5.14. Morphing basis and parameters for the validation samples for the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. 90
5.15. Input and validation distributions for the $\Delta\phi_{jj,\mathrm{sign}}$ observable for the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. 91
5.16. Morphing validation for the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. . . 92
5.17. Expected number of events as function of the free parameters in the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. 93
5.18. Relative uncertainty on the cross section as function of the free parameters in the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. 94
5.19. Relative uncertainty on the cross section as function of the free parameters κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$ in the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study after exchanging input sample `s14` with validation sample `v0`. . . . 94
5.20. Definitions of the angular observables sensitive to the spin and parity of the resonance in the $H\to ZZ^{*}\to\ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ decay 96
5.21. Effective Lagrangian Morphing validation using Φ and $\mathrm{TO}_2(\widetilde{\kappa}_{AZZ},\alpha)$ distributions of ggF $H\to ZZ^{*}\to\ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ events. 97
5.22. Example of a morphed distribution with large uncertainties. 98
5.23. Morphing basis for Figure 5.22. 99

5.24. Cumulative probability distributions of the relative uncertainties on the Effective Lagrangian prediction evaluated at equally distributed benchmark points for randomly chosen morphing bases. 101
5.25. VBF $H \to \mu\mu$ distributions for $\Delta\phi_{jj}$, m_{jj}, $p_{\mathrm{T}}^{j_1}$, and $\Delta\eta_{jj}$ with mixing of SM and BSM CP-even HWW couplings. 105
5.26. VBF $H \to \mu\mu$ distribution of $\mathrm{var}_{4\mathrm{d}}$, which is a combination of $\Delta\eta_{jj}$, dphijj, m_{jj} and $p_{\mathrm{T}}^{j_1}$ into one discriminating variable. 106
5.27. Correlation matrix of the coupling parameters after fitting to SM pseudo-data with 8% cross section uncertainty. 108
5.28. Illustration of the re-parametrization into the Eigen basis of the covariance matrix. 109
5.29. The profiled likelihood plots after a fit is performed on the reduced model for $\kappa_{\mathrm{H}\gamma\gamma}$ vs. $\kappa_{\mathrm{A}\gamma\gamma}$, $\kappa_{\mathrm{H}\gamma\gamma}$ vs. κ_{SM} and $\kappa_{\mathrm{A}\gamma\gamma}$ vs. κ_{SM}, and the corresponding correlation matrix. 110
5.30. The likelihood plots after a fit on the diagonalized reduced model for EV_0 vs. EV_1, EV_0 vs. EV_2 and EV_1 vs. EV_2, and the corresponding correlation matrix after diagonalization. 111
5.31. Prediction of the morphing of $\Delta\phi_{jj}$ and m_{jj} with and without interference. 114
5.32. Illustration of the interference effects in the observable $\Delta\phi_{jj,\mathrm{sign}}$ for the process VBF $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ between the CP-even coupling κ_{SM} and the CP-odd coupling κ_{AVV}. 115
5.33. Depiction of the morphing basis optimization algorithm. 118
5.34. Benchmark and validation parameter configuration shown in the parameter space used in the morphing basis optimization algorithm. 120
5.35. Development of the morphing bases during the interaction of the morphing basis optimization algorithm. 122
5.35. Development of the morphing bases during the interaction of the morphing basis optimization algorithm. (cont.) 123
5.35. Development of the morphing bases during the interaction of the morphing basis optimization algorithm. (cont.) 124
5.36. $\Delta\phi_{jj}$ distribution of the validation samples `v0-v3` for the single iterations in the morphing basis optimization algorithm. 125
5.36. $\Delta\phi_{jj}$ distribution of the validation samples `v0-v3` for the single iterations in the morphing basis optimization algorithm. (cont). . . 126
5.36. $\Delta\phi_{jj}$ distribution of the validation samples `v0-v3` for the single iterations in the morphing basis optimization algorithm. (cont). . . 127
5.37. Optimization score as a function of the parameter κ_{HVV} for one of the samples. 128

6.1. Leading order Feynman diagram of a Higgs boson produced via VBF. 130

6.2. Spin correlation in the $H \to W^{\pm}W^{\mp *} \to \ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ decay. 131
6.3. Distribution of p_{T}^{H} and $\Delta\phi_{jj}$ on generator level with acceptance cuts for the jets and additional VBF selection for different BSM scenarios. 133
6.4. $p_{T,jj}$ and $\Delta\phi_{jj,\mathrm{sign}}$ distributions for benchmarks showing characteristic features of EFT contributions to the $H \to VV$ vertex. 135
6.5. The $p_{T,jj}$ and $\Delta\phi_{jj,\mathrm{sign}}$ distributions for the benchmarks showing low impact of the EFT contributions to the $H \to VV$ vertex for the ggF production. 136
6.6. Trigger efficiencies for the single lepton triggers and single lepton triggers combined with the dilepton trigger. 138
6.7. Parameter configurations for VBF and ggF used to generate samples as inputs for the Effective Lagrangian Morphing to scan over the parameter space for κ_{SM}, κ_{HVV} and $\kappa_{\mathrm{H}\partial\mathrm{V}}$, and over κ_{SM} and κ_{AVV}. . . 139
6.8. Leading order Feynman diagram top-antitop pair production and single-top production. 143
6.9. Leading order Feynman diagram of the WW background production. 144
6.10. Leading order Feynman diagram of the $V\gamma$ background production. 145
6.11. Leading order Feynman diagram of the Drell-Yan background production with and without associated jets. 146
6.12. Leading order Feynman diagrams for the production of a W-boson. 147
6.13. The angular separation between the leading and subleading lepton $\Delta R_{\ell\ell}$ for events with a transverse momentum of the leading jet larger and smaller than 200 GeV. 150
6.14. The leading and subleading lepton p_{T} and η distributions after applying the lepton requirements. 154
6.15. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions after applying the lepton requirements. . . . 155
6.16. The $m_{\ell\ell}$ and n_{jet} distributions before applying respective requirements. 156
6.17. The leading and subleading jet p_{T} and η distributions after the $n_{\mathrm{jet}} > 2$ requirement. 157
6.18. The m_{jj} and $p_{\mathrm{T}}^{j_1}$ distributions before applying respective requirements. 158
6.19. The b-jet and $m_{\tau\tau}$ distributions before applying the respective b- and $Z \to \tau\tau$ vetos for the low- and high-p_{T} signal regions. 159
6.20. Relative contributions of each background and the SM signals at each cutstage of the signal region selection. 160
6.21. Schematic view of a decision tree. 162
6.22. Ranking of the input variables for $\mathrm{BDT}_{\mathrm{low}}$ and $\mathrm{BDT}_{\mathrm{high}}$. 165
6.23. Comparison of the BDT response between the `even` and `odd` subset of events taken as training and testing samples for both signal region. 167
6.24. Comparison of the ROC curves. 168
6.25. The distributions of $\mathrm{BDT}_{\mathrm{low}}$ and $\mathrm{BDT}_{\mathrm{high}}$ in the signal regions. . . . 169

6.26. Distributions of the $\mathrm{BDT_{low}}$ input variables in the low-p_{T} signal region. 170
6.27. Distributions of the $\mathrm{BDT_{high}}$ input variables in the high-p_{T} signal region. . . . 171
6.28. Correlation plots of BDT training variables for the BDT validation regions. . . . 172
6.29. The $m_{\tau\tau}$ distributions before applying respective requirements in the top control region. . . . 174
6.30. The distributions of $\mathrm{BDT_{low}}$ and $\mathrm{BDT_{high}}$ in the top control regions. 175
6.31. The $m_{\tau\tau}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\ell\ell}$ distributions before applying respective requirements in the $Z{\to}\tau\tau$ control region. . . . 176
6.32. The distributions of $\mathrm{BDT_{low}}$ and $\mathrm{BDT_{high}}$ in the $Z \to \tau\tau$ control regions. 177
6.33. Flavor compositions in the W+jets and Z+jets samples for identified and anti-identified electrons. . . . 185
6.34. Results of the fake factor calculations including statistical uncertainties. 187
6.35. Relative contributions of each background and the SM signals in the final regions. . . . 188
6.36. The $\Delta\phi_{jj,\mathrm{sign}}$ distributions for the low-pT signal region. . . . 189
6.37. Illustration of a 1σ confidence interval for a Gaussian-distributed measurement of a single quantity. . . . 207
6.38. Exemplary $t(\kappa)$ function with the estimated uncertainties. . . . 207
6.39. Breakdown of the total uncertainty on κ_{SM} and κ_{AVV} and the post-fit values and uncertainties of the nuisance parameters. . . . 211
6.40. Correlations between nuisance parameters observed in the likelihood fit κ_{SM} and κ_{AVV} measurement. . . . 212
6.41. The individual profiled likelihood scans of $\kappa_{\mathrm{SM}}\cos(\alpha)$ and $\kappa_{\mathrm{AVV}}\sin(\alpha)$. 215
6.42. The simultaneous profiled likelihood scans of $\kappa_{\mathrm{SM}}\cos(\alpha)$ and $\kappa_{\mathrm{AVV}}\sin(\alpha)$. 216
6.43. Post-fit signal region distributions in the measurement of κ_{SM} and κ_{AVV}. . . . 216
6.44. The relative uncertainty of the VBF event yield expectation calculated by the morphing function in the κ_{HVV}-$\kappa_{\mathrm{H}partial\mathrm{V}}$ coupling phase space. 218
6.45. Predicted event yield with uncertainties of the VBF morphing function. 220
6.46. Un-profiled negative logarithm of the likelihood for the SM Asimov data set in the κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ coupling phase space. . . . 221
6.47. Profiled negative logarithm of the likelihood for the SM Asimov data set in the κ_{HVV}-$\kappa_{\mathrm{H}\partial\mathrm{V}}$ coupling phase space. . . . 222

A.1. Distributions of the $\mathrm{BDT_{low}}$ input variables in the BDT validation region. . . . 230
A.2. Distributions of the $\mathrm{BDT_{high}}$ input variables in the BDT validation region. . . . 231

A.3. Distributions of the BDT_{low} input variables in the low-p_{T} top control region. 233
A.4. Distributions of the BDT_{high} input variables in the high-p_{T} top control region. 234
A.5. Distributions of the BDT_{low} input variables in the low-p_{T} $Z \to \tau\tau$ control region. 235
A.6. Distributions of the BDT_{high} input variables in the high-p_{T} $Z \to \tau\tau$ control region. 236

List of Tables

2.1. Standard Model particle content 5
2.2. Free parameters of the Standard Model. 6

4.1. Electron isolation operation points. 51
4.2. Muon identification efficiency. 52
4.3. Muon isolation operation points. 52
4.4. b-tag efficiencies and rejection rates for MV2c10. 58

5.1. Parameter settings for samples used for Matrix Element reweighting. 67
5.2. Set of coupling parameter points for a morphing example. 77
5.3. Morphing bases used to show the impact on the uncertainty. 84
5.4. Morphing basis and parameters for validation samples for the ggF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. 87
5.5. Correlation matrices for the fit of the morphing function to the validation distributions for the VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ study. . 91
5.6. Morphing basis and parameters for validation samples for the ggF $H\to ZZ^{*}\to\ell^{+}\ell^{-}\ell'^{+}\ell'^{-}$ study. 95
5.7. Calculated cross sections in the process VBF $H\to W^{\pm}W^{\mp *}\to\ell^{-}\bar{\nu}_{\ell}\ell'^{+}\nu_{\ell'}$ by MadGraph for different parameters settings. 100
5.8. The listed values of the coupling parameters reproduce the SM cross section at $\sqrt{s}=13\,\text{TeV}$. 104
5.9. Minimum and maximum values of the VBF production jet variables $\Delta\eta_{jj}$, m_{jj}, $p_{\text{T}}^{j_1}$ and $\Delta\phi_{jj}$. 106
5.10. Values of the coupling parameters and their respective uncertainties after fitting to SM pseudo-data with 8% cross section uncertainty. . 107
5.11. Morphing bases to visualize the absent interference effects between a CP-even (κ_{SM}) and a CP-odd (κ_{AVV}) coupling. 113

6.1. Benchmark scenarios used to illustrate the impact of different EFT parameters on typical VBF observables. 133
6.2. Parameters used for the benchmarks to show the signal characteristics. 134
6.3. The minimum p_{T} requirements used at the different levels of the trigger. 137

6.4. Parameter configuration of the benchmark EFT VBF samples. 140
6.5. Predicted signal process cross sections. 141
6.6. MC generators used to model the background processes, and corresponding cross sections. 148
6.7. Summary of the lepton selection. 151
6.8. MC statistics available for BDT trainings. 163
6.9. BDT parameter used for the training and hyper-parameter grid scan values used. 165
6.10. Background normalization factors from the control regions. 178
6.11. Expected event yields in the top control regions. 179
6.12. Expected event yields in the $Z \to \tau\tau$ control regions. 179
6.13. The requirements for fully identified and anti-identified electrons (left) and muons (right). 181
6.14. Corrections factors and corresponding systematic uncertainties. . . 185
6.15. Event selection criteria used to define the signal and control regions. 188
6.16. Expected event yields in the signal regions. 190
6.17. Event yields in the various signal and control regions. 190
6.18. Event yields in the signal regions of the VBF EFT benchmark samples. 190
6.19. Experimental systematic uncertainties. 191
6.20. Impact of experimental systematics. 198
6.21. Impact of misidentified leptons systematics. 201
6.22. Impact of theoretical systematics. 201
6.23. Post-fit Background normalisation factors in the likelihood fit κ_{SM} and κ_{AVV} measurement. 213

Bibliography

[1] D. Griffiths, *Introduction to Elementary Particles*. Wiley-VCH, 2004. ISBN 978-3-527-40601-2.

[2] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview, Boulder, CO, 1995. CDS:257493.

[3] M. J. Herrero, "The Standard Model", arXiv:9812242 [hep-ph].

[4] B. Gripaios, "Lectures on Physics Beyond the Standard Model", arXiv:1503.02636 [hep-ph].

[5] K. A. Olive and others for the Particle Data Group, "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001.

[6] M. Gell-Mann, "A Schematic Model of Baryons and Mesons", *Physics Letters* **8** no. 3, (1964) 214–215.

[7] **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].

[8] 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", arXiv:1602.00695 [hep-ph].

[9] P. Bolzoni, M. Zaro, F. Maltoni, and S.-O. Moch, "Higgs Production at NNLO in QCD: the VBF Channel", arXiv:1006.2323 [hep-ph].

[10] **The ATLAS Collaboration**, "Evidence for the $H \to b\bar{b}$ Decay with the ATLAS Detector", arXiv:1708.03299 [hep-ph].

[11] **The ATLAS Collaboration**, "Measurement of Fiducial Differential Cross Sections of Gluon-Fusion Production of Higgs Bosons Decaying to $WW^* \to e\nu\mu\nu$ with the ATLAS Detector at $\sqrt{s} = 8$ TeV", arXiv:1604.02997 [hep-ph].

[12] **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.

[13] **The ATLAS Collaboration**, "Evidence for the Higgs-Boson Yukawa Coupling to Tau Leptons with the ATLAS Detector", arXiv:1501.04943 [hep-ph].

[14] **The CMS Collaboration**, "Evidence for the 125 GeV Higgs Boson Decaying to a Pair of Tau Leptons", arXiv:1401.5041 [hep-ph].

[15] **The ATLAS Collaboration**, "Measurements of the Total and Differential Higgs Boson Production Cross Sections Combining the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ Decay Channels at $\sqrt{s} = 8$ TeV with the ATLAS Detector", arXiv:1504.05833 [hep-ph].

[16] **The CMS Collaboration**, "Measurement of Differential and Integrated Fiducial Cross Sections for Higgs Boson Production in the Four-Lepton Decay Channel in pp Collisions at $\sqrt{s} = 7$ and 8 TeV", arXiv:1512.08377 [hep-ph].

[17] **The CMS Collaboration**, "Measurement of Differential Cross Sections for Higgs Boson Production in the Diphoton Decay Channel in pp collisions at $\sqrt{s} = 8$ TeV", arXiv:1508.07819 [hep-ph].

[18] **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.

[19] **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.

[20] **The CMS Collaboration**, "Search for a Standard Model-like Higgs Boson in the $\mu^+ \mu^-$ and $e^+ e^-$ Decay Channels at the LHC", arXiv:1410.6679 [hep-ph].

[21] **The CMS Collaboration**, "Search for a Higgs Boson Decaying into a Z and a Photon in pp Collisions at $\sqrt{s} = 7$ and 8 TeV", arXiv:1307.5515 [hep-ph].

[22] **The ATLAS Collaboration**, "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS detector at the LHC", arXiv:1207.7214 [hep-ph].

[23] **The CMS Collaboration**, "Observation of a New Boson at a Mass of 125 GeV with the CMS experiment at the LHC", arXiv:1207.7235 [hep-ph].

[24] **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", arXiv:1503.07589 [hep-ph].

[25] J. R. Espinosa, "Implications of the Top (and Higgs) Mass for Vacuum Stability", arXiv:1512.01222 [hep-ph].

[26] **The ATLAS Collaboration**, "Study of the Spin and Parity of the Higgs Boson in Diboson Decays with the ATLAS Detector", arXiv:1506.05669 [hep-ph].

[27] **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", arXiv:1606.02266 [hep-ph].

[28] **The ATLAS Collaboration**, "Combined Measurements of Higgs Boson Production and Decay using up to 80 fb^{-1} of pp Collision Data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment", ATLAS-CONF-2019-005, CERN, Geneva, Mar., 2019. CDS:2668375.

[29] W. Buchmüller and D. Wyler, "Effective Lagrangian Analysis of new Interactions and Flavour Conservation", *Nuclear Physics B* **268** (May, 1986) 621–653.

[30] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, "Dimension-Six Terms in the Standard Model Lagrangian", arXiv:1008.4884 [hep-ph].

[31] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, and M. Spira, "Effective Lagrangian for a Light Higgs-like Scalar", arXiv:1303.3876 [hep-ph].

[32] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, "The Strongly-Interacting Light Higgs", arXiv:0703164 [hep-ph].

[33] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, "Low Energy Effects of New Interactions in the Electroweak Boson Sector", *Physical Review D* **48** (Sept., 1993) 2182–2203.

[34] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, "Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology", arXiv:1312.2014 [hep-ph].

[35] B. Henning, X. Lu, T. Melia, and H. Murayama, "2, 84, 30, 993, 560, 15456, 11962, 261485, ...: Higher Dimension Operators in the SM EFT", arXiv:1512.03433 [hep-ph].

[36] A. Alloul, B. Fuks, and V. Sanz, "Phenomenology of the Higgs Effective Lagrangian via FeynRules", arXiv:1310.5150 [hep-ph].

[37] L. Lavele Carter and E. Darrell Cashwell, "Particle-Transport Simulation with the Monte Carlo Method", *Los Alamos Scientific Lab.* (Jan., 1975) .

[38] M. Veltman, *Diagrammatica: The Path to Feynman Diagrams*. Cambridge Lecture Notes in Physics. Cambridge University Press, 1994. ISBN 9780521456920.

[39] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, "FeynRules 2.0 - A Complete Toolbox for Tree-Level Phenomenology", arXiv:1310.1921 [hep-ph].

[40] F. Krauss, *Representation of a pp Collision*, Apr., 2018. https://www.opensciencegrid.org/using-the-osg-to-test-theories-of-nature-at-the-lhc/.

[41] **The NNPDF Collaboration**, *Comparison between NNPDF 3.1 and the Sets entering the PDF4LHC Combination at NNLO*, Mar., 2018. http://pcteserver.mi.infn.it/~nnpdf/nnpdf31-gallery/specs0_each_report_report.html.

[42] **The NNPDF Collaboration**, "Parton Distributions from High-Precision Collider Data", arXiv:1706.00428 [hep-ph].

[43] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton Distributions in the LHC era: MMHT 2014 PDFs", arXiv:1412.3989 [hep-ph].

[44] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, "New Parton Distributions for Collider Physics", arXiv:1007.2241 [hep-ph].

[45] A. Buckley, J. Ferrando, S. Lloyd, K. Nordstrom, B. Page, M. Ruefenacht, M. Schoenherr, and G. Watt, "LHAPDF6: Parton Density Access in the LHC Precision Era", arXiv:1412.7420 [hep-ph].

[46] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001.

[47] S. Myers, *The LEP Collider, from Design to Approval and Commissioning*. John Adams' Lecture. CERN, Geneva, 1991. https://cds.cern.ch/record/226776.

[48] The CERN Low Power Converter Section, *The CERN Accelerator Complex*, May, 2018. http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/pagesources/Cern-Accelerator-Complex.jpg.

[49] The LHC Machine Outreach, *LHC Overview Schematic*, May, 2018. https://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/lhc-schematic.jpg.

[50] **The ATLAS Collaboration**, "The ATLAS Experiment at the CERN Large Hadron Collider", *Journal of Instrumentation* **3** no. 08, (2008) S08003. http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[51] T. Heck, *Search for $t\bar{t}$ Resonances in the Lepton + Jets Channel in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*. PhD thesis, Johannes Gutenberg Universität Mainz, 2016.

[52] K. Potamianos, "The Upgraded Pixel Detector and the Commissioning of the Inner Detector Tracking of the ATLAS Experiment for Run-2 at the Large Hadron Collider", arXiv:1608.07850 [hep-ph].

[53] ATLAS Collaboration, "Particle Identification Performance of the ATLAS Transition Radiation Tracker", ATLAS-CONF-2011-128, CERN, Geneva, Sept., 2011. CDS:1383793.

[54] ATLAS Collaboration, "The Run-2 ATLAS Trigger System", ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb., 2016. CDS:2133909.

[55] A. Salzburger, "Optimisation of the ATLAS Track Reconstruction Software for Run-2", *Journal of Physics: Conference Series* **664** no. 7, (2015) 072042.

[56] **The ATLAS Collaboration**, "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run-2", arXiv:1704.07983 [hep-ph].

[57] **The ATLAS Collaboration**, "Charged-particle Distributions in $\sqrt{s} = 13$ TeV pp Interactions Measured with the ATLAS Detector at the LHC", arXiv:1602.01633 [hep-ph].

[58] **The ATLAS Collaboration**, "Performance of Primary Vertex Reconstruction in pp Collisions at $\sqrt{s}$ =7 TeV in the ATLAS Experiment", ATLAS-CONF-2010-069, CERN, Geneva, July, 2010. CDS:1281344.

[59] **The ATLAS Collaboration**, "An Imaging Algorithm for Vertex Reconstruction for ATLAS Run-2", ATL-PHYS-PUB-2015-008, CERN, Geneva, Apr., 2015. CDS:2008700.

[60] **The ATLAS Collaboration**, *Luminosity Public Results Run-2*, July, 2015. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[61] **The ATLAS Collaboration**, "Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV", ATL-PHYS-PUB-2015-026, CERN, Geneva, July, 2015. CDS:2037717.

[62] **The ATLAS Collaboration**, "Reconstruction of Primary Vertices at the ATLAS experiment in Run-1 pp Collisions at the LHC", arXiv:1611.10235 [hep-ph].

[63] **The ATLAS Collaboration**, "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics", arXiv:0901.0512 [hep-ph].

[64] **The ATLAS Collaboration**, "Electron Efficiency Measurements with the ATLAS Detector using the 2015 LHC pp Collision Data", ATLAS-CONF-2016-024, CERN, Geneva, June, 2016. CDS:2157687.

[65] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, "Calorimeter Clustering Algorithms: Description and Performance", ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr., 2008. CDS:1099735.

[66] **The ATLAS Collaboration**, *Electron Reconstruction and Identification Efficiency Measurements in 2016 Data*, Mar., 2017. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2017-003/index.html.

[67] H. Herde, "Muon Reconstruction Performance in ATLAS at Run-2", ATL-PHYS-PROC-2015-121, CERN, Geneva, Oct., 2015. CDS:2059849.

[68] **The ATLAS Collaboration**, "Muon Reconstruction Performance of the ATLAS Detector in pp Collision Data at $\sqrt{s}$=13 TeV", arXiv:1603.05598 [hep-ph].

[69] **The ATLAS Collaboration**, "Topological Cell Clustering in the ATLAS Calorimeters and its Performance in LHC Run-1", arXiv:1603.02934 [hep-ph].

[70] **The ATLAS Collaboration**, "Jet Energy Scale Measurements and their Systematic Uncertainties in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector", arXiv:1703.09665 [hep-ph].

[71] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k_t Jet Clustering Algorithm", arXiv:0802.1189 [hep-ph].

[72] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", arXiv:1111.6097 [hep-ph].

[73] **The ATLAS Collaboration**, "Tagging and Suppression of Pile-up Jets with the ATLAS Detector", ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. CDS:1700870.

[74] **The ATLAS Collaboration**, "Forward Jet Vertex Tagging: A New Technique for the Identification and Rejection of Forward Pile-up Jets", ATL-PHYS-PUB-2015-034, CERN, Geneva, Aug., 2015. CDS:2042098.

[75] **The ATLAS Collaboration**, "Performance of b-Jet Identification in the ATLAS Experiment", arXiv:1512.01094 [hep-ph].

[76] **The ATLAS Collaboration**, "Optimisation of the ATLAS b-Tagging Performance for the 2016 LHC Run", ATL-PHYS-PUB-2016-012, CERN, Geneva, June, 2016. CDS:2160731.

[77] **The ATLAS Collaboration**, "Expected Performance of Missing Transverse Momentum Reconstruction for the ATLAS Detector at $\sqrt{s} = 13$ TeV", ATL-PHYS-PUB-2015-023, CERN, Geneva, July, 2015. CDS:2037700.

[78] **The ATLAS Collaboration**, *Missing Transverse Momentum Distribution and Performance in 2016 Data*, July, 2016. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-008/.

[79] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic Formulae for Likelihood-based Tests of New Physics", arXiv:1007.1727 [hep-ph].

[80] K. Cranmer, "Practical Statistics for the LHC", arXiv:1503.07622 [hep-ph].

[81] R. Jansky, "The ATLAS Fast Monte Carlo Production Chain Project", *Journal of Physics: Conference Series* **664** no. 7, (2015) 072024. http://stacks.iop.org/1742-6596/664/i=7/a=072024.

[82] S. Frixione and B. R. Webber, "Matching NLO QCD Computations and Parton Shower Simulations", arXiv:0204244 [hep-ph].

[83] E. Bothmann, M. Schönherr, and S. Schumann, "Reweighting QCD Matrix Element and Parton Shower Calculations", arXiv:1606.08753 [hep-ph].

[84] 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", arXiv:1405.0301 [hep-ph].

[85] A. L. Read, "Linear Interpolation of Histograms", *Nucl. Instrum. Meth.* **A425** (1999) 357–360.

[86] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, "Interpolation between Multi-dimensional Histograms using a New Non-Linear Moment Morphing Method", arXiv:1410.7388 [hep-ph].

[87] **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.

[88] *BOOST C++ Libraries*. http://www.boost.org.

[89] **The NNPDF Collaboration**, "Parton Distributions for the LHC Run-2", arXiv:1410.8849 [hep-ph].

[90] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2", arXiv:1410.3012 [hep-ph].

[91] **The ATLAS Collaboration**, "ATLAS Run-1 Pythia8 Tunes", ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov., 2014. CDS:1966419.

[92] S. Agostinelli *et al.*, "GEANT4: A Simulation Toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[93] **The ATLAS Collaboration**, "The ATLAS Simulation Infrastructure", arXiv:1005.4568 [hep-ph].

[94] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO Higgs Boson Production via Gluon Fusion Matched with Shower in POWHEG", arXiv:0812.0578 [hep-ph].

[95] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, and N. V. Tran, "Spin Determination of Single-produced Resonances at Hadron Colliders", arXiv:1001.3396 [hep-ph].

[96] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, N. V. Tran, and A. Whitbeck, "On the Spin and Parity of a Single-produced Resonance at the LHC", arXiv:1208.4018 [hep-ph].

[97] T. L. H. C. S. W. Group, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", arXiv:1307.1347 [hep-ph].

[98] **The ATLAS Collaboration**, "Evidence for the Spin-0 Nature of the Higgs Boson using ATLAS Data", arXiv:1307.1432 [hep-ph].

[99] F. Maltoni, K. Mawatari, and M. Zaro, "Higgs Characterisation via Vector Boson Fusion and Associated Production: NLO and Parton Shower Effects", arXiv:1311.1829 [hep-ph].

[100] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic Spin-entangled Decays of Heavy Resonances in Monte Carlo Simulations", arXiv:1212.3460 [hep-ph].

[101] J. Adelman, M. Baak, N. Boelaert, M. D'Onofrio, J. A. Frost, C. Guyot, M. Hauschild, A. Hoecker, K. J. C. Leney, E. Lytken, M. Martinez-Perez, J. Masik, A. M. Nairz, P. U. E. Onyisi, S. Roe, S. Schaetzel, and M. G. Wilson, "ATLAS Offline Data Quality Monitoring", *Journal of Physics: Conference Series* **219** no. 4, (2010) 042018. http://stacks.iop.org/1742-6596/219/i=4/a=042018.

[102] **The ATLAS Collaboration**, "Improved Luminosity Determination in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS Detector at the LHC", arXiv:1302.4393 [hep-ph].

[103] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. D. Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, "Parton Distributions with LHC Data", arXiv:1207.1303 [hep-ph].

[104] O. Brein, A. Djouadi, and R. Harlander, "NNLO QCD Corrections to the Higgs-strahlung Processes at Hadron Colliders", arXiv:0307206 [hep-ph].

[105] M. L. Ciccolini, S. Dittmaier, and M. Krämer, "Electroweak Radiative Corrections to Associated WH and ZH Production at Hadron Colliders", arXiv:0306234 [hep-ph].

[106] E. Re, "Reaching NNLOPS accuracy with POWHEG and MiNLO", arXiv:1401.2944 [hep-ph].

[107] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A. D. Roeck, J. Feltesse, S. Forte, J. Gao, S. Glazov, J. Huston, Z. Kassabov, R. McNulty, A. Morsch, P. Nadolsky, V. Radescu, J. Rojo, and R. Thorne, "PDF4LHC Recommendations for LHC Run-2", arXiv:1510.03865 [hep-ph].

[108] 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].

[109] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual", arXiv:0603175 [hep-ph].

[110] **The ATLAS Collaboration**, "Measurement of the Z/γ^* Boson Transverse Momentum Distribution in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector", arXiv:1406.3660 [hep-ph].

[111] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, and J. Winter, "Event Generation with SHERPA 1.1", arXiv:0811.4622 [hep-ph].

[112] S. Schumann and F. Krauss, "A Parton Shower Algorithm based on Catani-Seymour Dipole Factorisation", arXiv:0709.1027 [hep-ph].

[113] M. Chrzaszcz, T. Przedzinski, Z. Was, and J. Zaremba, "TAUOLA of Tau Lepton Decays - Framework for Hadronic Currents, Matrix Elements and Anomalous Decays", arXiv:1609.04617 [hep-ph].

[114] M. Czakon, P. Fiedler, and A. Mitov, "The Total Top Quark Pair Production Cross Section at Hadron Colliders through $\mathcal{O}(\alpha_S^4)$", arXiv:1303.6254 [hep-ph].

[115] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes", arXiv:1005.3457 [hep-ph].

[116] D. J. Lange, "The EvtGen Particle Decay Simulation Package", *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[117] 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 NNLO QCD", arXiv:1408.5243 [hep-ph].

[118] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, "QCD Corrections to W^+W^- Production through Gluon Fusion", arXiv:1511.08617 [hep-ph].

[119] K. Melnikov and F. Petriello, "Electroweak Gauge Boson Production at Hadron Colliders through $\mathcal{O}(\alpha_S^2)$", arXiv:0609070 [hep-ph].

[120] M. Bahr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Platzer, P. Richardson, M. H. Seymour, A. Sherstnev, J. Tully, and B. R. Webber, "HERWIG++ Physics and Manual", arXiv:0803.0883 [hep-ph].

[121] T. Plehn, D. Rainwater, and D. Zeppenfeld, "A Method for Identifying $H \to \tau\tau \to e\mu\, p_{\mathrm{T}}^{\mathrm{miss}}$ at the CERN LHC", arXiv:9911385 [hep-ph].

[122] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

[123] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", *Journal of Computer and System Sciences* **55** no. 1, (1997) 119 – 139. http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[124] J. H. Friedman, "Stochastic Gradient Boosting", *Comput. Stat. Data Anal.* **38** no. 4, (2002) 367–378.

[125] F. J. Massey, "The Kolmogorov-Smirnov Test for Goodness of Fit", *Journal of the American Statistical Association* **46** no. 253, (1951) 68–78.

[126] G. Cowan, *Discovery Sensitivity for a Counting Experiment with Background Uncertainty*, July, 2012. https://www.pp.rhul.ac.uk/~cowan/stat/medsig/medsigNote.pdf.

[127] 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.

[128] **The ATLAS Collaboration**, "Electron Identification Measurements in ATLAS using $\sqrt{s} = 13$ TeV Data with 50 ns Bunch Spacing", ATL-PHYS-PUB-2015-041, CERN, Geneva, Sept., 2015. CDS:2048202.

[129] **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.

[130] M. Goblirsch-Kolb, for the ATLAS Collaboration, "Muon Reconstruction Efficiency, Momentum Scale and Resolution in *pp* Collisions at 8 TeV with ATLAS", arXiv:1408.7086 [hep-ph].

[131] **The ATLAS Collaboration**, "Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV", ATL-PHYS-PUB-2015-015, CERN, Geneva, July, 2015. CDS:2037613.

[132] L. Lonnblad and S. Prestel, "Matching Tree-Level Matrix Elements with Interleaved Showers", arXiv:1109.4829 [hep-ph].

[133] R. J. Barlow and C. Beeston, "Fitting using Finite Monte Carlo Samples", *Comput. Phys. Commun.* **77** (1993) 219–228.

[134] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *The Annals of Mathematical Statistics* **9** no. 1, (Mar., 1938) 60–62. https://doi.org/10.1214/aoms/1177732360.

[135] B. A. Murtagh and M. A. Saunders, *MINOS 5.51 User's Guide*, Sept., 2013. https://web.stanford.edu/group/SOL/guides/minos551.pdf.

[136] R. Brun and F. Rademakers, "ROOT — An Object Oriented Data Analysis Framework", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **389** no. 1, (1997) 81 – 86. http:

//www.sciencedirect.com/science/article/pii/S016890029700048X. New Computing Techniques in Physics Research V.

[137] W. Verkerke and D. Kirkby, "The RooFit Toolkit for Data Modeling", arXiv:0306116 [physics].

[138] C. Burgard *et al.*, *Common Analysis Framework*, Oct., 2017. https://atlas-caf.web.cern.ch/.

[139] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput. Phys. Commun.* **10** (1975) 343–367.