Durham E-Theses

Constraining New Physics at the LHC

PARISA THERESA GREGG

How to cite:

GREGG, PARISA THERESA (2022) Constraining New Physics at the LHC. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/14434/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

Constraining New Physics at the LHC

Parisa Gregg

A Thesis presented for the degree of
Doctor of Philosophy

Durham University

Institute for Particle Physics Phenomenology
Department of Physics
Durham University
United Kingdom

April 2022

Constraining New Physics at the LHC

Parisa Gregg

Submitted for the degree of Doctor of Philosophy

April 2022

Abstract: We analyse constraints that can be placed on CP violating operators in the dimension-six Standard Model Effective Field Theory. We consider the associated production of a Higgs boson and a photon in weak boson fusion. We construct an asymmetry in a CP-sensitive observable which we use to analyse the sensitivity of this channel to possible CP violation. Additionally, we constrain CP violating charged and neutral anomalous triple gauge couplings with diboson and VBF Vjj production at the LHC. We analyse charged triple gauge couplings in the dimension-six Standard Model Effective Field Theory using asymmetries to place constraints of CP-violating Wilson coefficients. For neutral triple gauge couplings we present bounds in terms of a general anomalous coupling framework using high energy transverse momentum bins. Finally, we consider next-to-leading order electroweak virtual corrections to $W+$ jets production at the LHC. We investigate the ability of a neural network to learn electroweak k-factors for partonic channels with two, three and four final partons.

Contents

Abstract **3**

Contents **5**

Declaration **9**

Acknowledgements **11**

1 Introduction **13**

1.1 The Standard Model 14

1.1.1 The Gauge Lagrangian 16

1.1.2 The Fermion Lagrangian 17

1.1.3 The Higgs Lagrangian 18

1.1.4 The Yukawa Lagrangian 19

1.1.5 Beyond the Standard Model 20

1.2 Collider Phenomenology 21

1.2.1 Collider parameters 22

1.2.2 Measurements and observables 23

1.2.3 Resonance searches 24

1.2.4 Higgs phenomenology 25

1.2.5 Monte Carlo event generators 26

1.2.6 Statistical analyses 28

I Constraining CP violating operators beyond the Standard Model 31

2 Effective Field Theory 33

2.1 EFTs in particle physics 33

2.2 Fermi Theory 35

2.3 Standard Model Effective Field Theory 36

2.3.1 Phenomenology of SMEFT operators 37

2.3.2 Constraining Wilson coefficients 40

3 Constraining SMEFT operators with associated hγ production in WBF 43

3.1 Introduction 43

3.2 Signal and Backgrounds in the Standard Model 45

3.2.1 Process simulation 45

3.2.2 Extracting the signal 46

3.3 Determination of the b-Yukawa coupling 49

3.4 EFT analysis 51

3.4.1 Selection of operators 52

3.4.2 CP structure of the EFT and observable consequences . . . 55

3.5 Conclusions and outlook 58

4 Constraining CP violating operators in Triple Gauge Couplings 61

4.1 Introduction 61

4.2 Charged aTGCs 62

4.2.1 Process simulation 65

4.2.2 Combined constraints 67

4.3 Neutral aTGCs 69

4.3.1 ZZ production 70

4.3.2 Zγ production 72

4.4 Conclusions and outlook 74

II Machine Learning 77

5 Machine Learning in High Energy Physics 79

5.1 Neural Networks 80

5.2 Decision Trees 82

5.2.1 Averaging methods 84

5.2.2 Boosting methods 84

6 Learning virtual corrections 85

6.1 Introduction 85

6.2 Sherpa set up 86

6.3 Neural network 88

6.4 Results 89

6.5 Conclusions and outlook 95

7 Conclusions 97

A Additional information for "Learning virtual corrections" 101

B Industrial placement: Monitoring aircraft turnaround with AI 105

B.1 Introduction and motivation 105

B.2 Convolutional neural networks and computer vision 107

B.2.1 Convolutional neural networks 107

B.2.2 Object detection 109

B.3 Monitoring the turnaround 110

B.3.1 Object detection model 110

B.3.2 Turning data into knowledge 112

B.3.3 Aircraft sensors 114

B.4 Conclusions 115

Bibliography 117

Declaration

The work in this thesis is based on research carried out in the Institute for Particle Physics Phenomenology at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification. This thesis is partly based on the joint research below.

- Chapter 2 is based on the article *Constraining SMEFT operators with associated $h\gamma$ production in Weak Boson Fusion* published in Physics Letters B [1].
- Chapter 3 is based on the article *Constraining CP violating operators in charged and neutral triple gauge couplings* published in Physics Letters B [2].
- Chapter 5 is based on research undertaken in collaboration with Frank Krauss and Marek Schönherr.

Copyright © 2022 Parisa Gregg.

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Acknowledgements

First and foremost, I would like to thank my supervisor Frank Krauss for his expertise, patience and guidance, without which this work would not have been possible. I would also like to thank Anke Biekötter, Marek Schönherr and Raquel Gomez-Ambrosio, with whom I have had the privilege of collaborating and whose expertise and insights have taught me a great deal.

The IPPP has been a wonderful, welcoming and inspiring place to work over the course of my PhD. I feel extremely fortunate to have studied here and to have met a great number of incredible people. In particular, thank you to the members of OC321 (lets have some fun) for being the best office pals anyone could ask for; wine time, book club and ballroom dancing have been some of the highlights of the past few years. Special mention must also be given to Jack, Lois, Dorian, Oscar, Mia, Elliott, Daniel, Joe, Henry, Ryan and Andrew for generously volunteering their time to proofread parts of this thesis.

I would like to thank my housemates past and present: Danny, Vincent, Elliott and Andrew, for their friendship and for making the various lockdowns bearable.

To Andrew, thank you for your love and kindness and for your ceaseless support and encouragement throughout.

Finally, thank you to Mum, Dad and Maziar for your endless love and support, without which I would not be where I am today.

Chapter 1

Introduction

In 2012 a scalar resonance was discovered at the Large Hadron Collider (LHC) igniting intense efforts to measure its properties. To date these have proven consistent with those of a Standard Model Higgs boson, consolidating the Standard Model (SM) as our best current theory of the universe. In chapter 1 we introduce the Standard Model and its phenomenology at collider experiments.

Although no new resonances have been discovered since the Higgs boson, we know from other contradictory phenomena that the Standard Model is incomplete, motivating the search for new physics beyond it. With the possibility that signals of new physics models may occur at scales beyond the reach of the LHC, as well as the accumulation of a vast LHC dataset over the past nine years, focus has shifted towards placing model independent constraints on experimentally allowed deviations from the Standard Model. In chapter 2 we introduce the Standard Model Effective Field Theory (SMEFT). This theory parametrises the effect of new physics at high energies on measurements at the lower energy scale of the LHC. With this framework we can look for hints of Beyond the Standard Model (BSM) phenomena in current and future LHC data.

The effects we are searching for can be extremely subtle, necessitating the combination of as many measurements as possible across the multitude of particle interaction processes occurring at the LHC. In chapter 3 we look at the weak boson fucion

(WBF) Higgs production channel with an associated photon. The constraints obtained on charge and parity (CP) violating parameters of the SMEFT with this channel are comparable with those from the WBF channel alone, motivating its inclusion in global fits.

The study of BSM CP-odd interactions is important as they provide additional sources of CP-violation which are unaccounted for in the SM yet necessary to explain the baryon asymmetry observed in the universe. In chapter 4 we study diboson and vector boson fusion Vjj production at the LHC in order to probe couplings between three gauge bosons. We investigate the constraints that can be placed on CP-odd SMEFT operators in these couplings.

As we gather increasing amounts of data from the LHC and enter the high luminosity era, we are able to probe progressively smaller deviations from SM predictions. It is therefore vital that our theoretical predictions for signal and background processes are accurate. In some cases this necessitates the inclusion of next-to-leading order (NLO) calculations. However, these can come with a large computational overhead. In chapters 5 and 6 we introduce common machine learning techniques used in high energy physics and present an initial exploration of the use of machine learning in calculating NLO electroweak virtual corrections to $W+$jet processes at the LHC.

In a wider context machine learning has a vast number of applications in the modern world. In appendix B we give details of a project undertaken at Jeppesen, which demonstrates a use case of convolutional neural networks for object detection at airports.

1.1 The Standard Model

In 1897 J.J.Thomson discovered the first fundamental particle, the electron, through the deflection of cathode rays in a magnetic field. Over the next century a theoretical framework was built up to describe electrons and the other fundamental particles

u	c	t	γ	h
d	s	b	g	
ν_e	ν_μ	ν_τ	W	
e	μ	τ	Z	

Table 1.1: The Standard Model particle content. The three generations of quarks and leptons are shown in top-left and bottom-left respectively. The centre column contains the gauge bosons and the scalar Higgs is shown in the top right.

as well as their interactions, leading to the Standard Model (SM) of particle physics. The SM has proven to be extremely successful in explaining and predicting almost all observed experimental phenomena, its latest triumph being the discovery of a new particle consistent with the theorised Higgs boson at CERN's LHC in 2012 [3,4].

The Standard Model is a quantum field theory (QFT) constructed to describe three of the four known fundamental forces of nature: strong, weak and electromagnetic. The fourth, the gravitational force, is not considered in the SM as its effect on physics at the particle scale is negligible. The model describes the fundamental particles of matter as a set of fermions that interact under these forces via the exchange of vector bosons. These fermions are further split into two types: quarks, which participate in strong, weak and electromagnetic interactions; and leptons, which are not acted upon by the strong force. The particle content of the SM is summarised in table 1.1.

In QFT particles are described as excitations of quantum fields. The dynamics and interactions of these fields are determined by the Lagrangian of the theory. We can decompose the SM Lagrangian into the sum,

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{1.1.1}$$

where $\mathcal{L}_{\text{gauge}}$ describes the SM gauge fields, $\mathcal{L}_{\text{fermion}}$ describes the dynamics of the fermions, $\mathcal{L}_{\text{Higgs}}$ contains the kinetic and potential Higgs terms which determine the properties of the Higgs field and $\mathcal{L}_{\text{Yukawa}}$ describes the interaction between the Higgs field and fermions. We will briefly examine each of these terms in the following sections. A more detailed formulation can be found in [5–7].

The SM is a gauge theory and eq. (1.1.1) is invariant under transformations of the local symmetry group,

$$\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y. \tag{1.1.2}$$

Here we show the colour, c, left, L, and hypercharge, Y, charge subscripts which we generally drop in the remainder of this thesis. The number of bosons that mediate interactions is determined by the number of generators of the group. In general an SU(N) group has $N^2 - 1$ generators. Thus, the SU(3) group has eight generators which give the eight gluons. The SU(2) group has three generators and the U(1) group has one generator. The generators of the SU(2) and U(1) groups give the $W^\pm$ and Z bosons and the photon after electroweak symmetry breaking (EWSB).

1.1.1 The Gauge Lagrangian

The gauge Lagrangian contains the kinetic and self interaction terms of the SM gauge fields,

$$\begin{aligned} \mathcal{L}_{\text{gauge}} &= -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \\ &= -\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a - \frac{1}{4} W^i_{\mu\nu} W^{\mu\nu}_i - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \end{aligned} \tag{1.1.3}$$

with field strength tensors for each gauge group given by

$$\begin{aligned} G^a_{\mu\nu} &= \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_3 f^{abc} G^b_\mu G^c_\nu \\ W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g_2 \epsilon^{ijk} W^j_\mu W^k_\nu \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{1.1.4}$$

where the indices, a and i are the SU(3) colour and SU(2) weak indices. The coefficients g_3 and g_2 are coupling constants, f^{abc} and ϵ^{ijk} are structure constants for the SU(3) and SU(2) groups and G^a_μ, W^i_μ and B_μ are the gauge fields of SU(3), SU(2) and U(1) respectively.

1.1.2 The Fermion Lagrangian

The second term in eq. (1.1.1) describes the dynamics of fermions and their interactions with the gauge fields. For one generation of fermions this can be written as

$$\mathcal{L}_{\text{fermion}} = i\overline{Q_L}\not{D}Q_L + i\overline{u_R}\not{D}u_R + i\overline{d_R}\not{D}d_R + i\overline{L_L}\not{D}L_L + i\overline{e_R}\not{D}e_R, \qquad (1.1.5)$$

where $\not{D} = \gamma^\mu D_\mu$ and D_μ is the covariant derivative given by

$$D_\mu = \partial_\mu - ig_3 G^a_\mu t^a - ig_2 W^i_\mu \tau^i - ig_Y B_\mu Y. \qquad (1.1.6)$$

Here g_3, g_2 and g_Y are coupling constants and t^a, τ^i and Y are the generators of the SU(3), SU(2) and U(1) groups. The τ^i can be written in terms of the Pauli matrices as $\sigma^i/2$. The left chiral quark and lepton fields form SU(2) doublets

$$L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \qquad Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \qquad (1.1.7)$$

whereas the right chiral quark and charged lepton fields, u_R, d_R and e_R, are singlets under SU(2) transformations. Furthermore, the quark fields are triplets under SU(3),

$$u_L = \begin{pmatrix} u^r_L \\ u^b_L \\ u^g_L \end{pmatrix}, \qquad u_R = \begin{pmatrix} u^r_R \\ u^b_R \\ u^g_R \end{pmatrix} \qquad (1.1.8)$$

where the r, g, b superscripts represent the QCD colours: red, blue, green.

Any fermion mass terms in the Lagrangian must involve a coupling between left chiral and right chiral states. However, as these states transform differently under SU(2), direct coupling terms are not gauge invariant. Instead the fermion masses are generated via the Higgs Mechanism.

1.1.3 The Higgs Lagrangian

The Higgs field is an SU(2) doublet of two complex scalar fields with U(1) hypercharge $Y = 1/2$:

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.1.9}$$

Its properties are described by

$$\mathcal{L}_{\text{Higgs}} = \left| D_\mu H \right|^2 - \mu H^\dagger H - \lambda (H^\dagger H)^2, \tag{1.1.10}$$

where the last two terms describe the Higgs potential $V(H)$, and the covariant derivative of the Higgs field is given by,

$$D_\mu H = (\partial_\mu - i g_2 W^i_\mu \tau^i - i g_Y B_\mu Y) H \tag{1.1.11}$$

When $\mu^2 < 0$ the minimum of the potential occurs for a non-zero value of the Higgs field

$$\langle H \rangle = |H| = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}, \tag{1.1.12}$$

where $|H| = \sqrt{H^\dagger H}$ and v is the vacuum expectation value (VEV) which has been experimentally measured to be $v \sim 246$ GeV. This results in the spontaneous breaking of the electroweak symmetry as this non-zero vacuum state of the Higgs field is not invariant under the SU(2) × U(1) group.

In the unitary gauge the Higgs doublet can be written as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \tag{1.1.13}$$

where h is the perturbation about the minimum corresponding to the Higgs boson. Substituting this definition of the Higgs field into the potential defined in eq. (1.1.10) we obtain the Higgs boson mass

$$m_h^2 = 2\lambda v^2. \tag{1.1.14}$$

The gauge boson mass terms come from evaluating the kinetic term of eq. (1.1.10) after electroweak symmetry breaking,

$$\left|D_\mu\langle H\rangle\right|^2 = \left|\frac{1}{2}\begin{pmatrix} g_Y B_\mu + g_2 W^3_\mu & g_2\left(W^1_\mu - iW^2_\mu\right) \\ g_2\left(W^1_\mu + iW^2_\mu\right) & g_Y B_\mu - g_2 W^3_\mu \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}\right|^2 \tag{1.1.15}$$

$$= \frac{v^2 g_2^2}{4} W^+_\mu W^{-\mu} + \frac{v^2}{4}\left(g_2^2 + g_Y^2\right) Z_\mu Z^\mu. \tag{1.1.16}$$

In the last line the new massive fields are defined as linear combinations of the old fields by

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right), \quad Z_\mu = \frac{g_Y B_\mu - g_2 W^3_\mu}{\sqrt{g_2^2 + g_Y^2}} = c_\theta W^3_\mu - s_\theta B_\mu, \tag{1.1.17}$$

where $c_\theta = \cos\theta_W$, $s_\theta = \sin\theta_W$ and θ_W is the rotation angle of the Z boson with respect to W^3 and B. The remaining orthogonal combination of fields, $A_\mu = (g_Y B_\mu + g_2 W^3_\mu)/\sqrt{g_2^2 + g_Y^2} = s_\theta W^3_\mu + c_\theta B_\mu$ does not couple to the Higgs field. It is therefore left massless and is identified as the photon.

1.1.4 The Yukawa Lagrangian

The final part of the SM Lagrangian describes the interaction of the Higgs field with fermions

$$\mathcal{L}_{\text{Yukawa}} = -\left(\overline{Q_L}\tilde{H}\right) Y_u u_R - \left(\overline{Q_L} H\right) Y_d d_R - \left(\overline{L_L} H\right) Y_e e_R + h.c. \tag{1.1.18}$$

where Y_u, Y_d and Y_e are the Yukawa matrices and $\tilde{H} = i\sigma^2 H^*$. Only couplings between an SU(2) singlet and SU(2) doublet are included to preserve gauge invariance. As the Y_u and Y_d Yukawa matrices are not diagonal in the flavour basis they lead to flavour and CP violating effects through the Cabbibo-Kobayashi-Maskawa (CKM) matrix.

The fermion masses can be obtained from the Yukawa Lagrangian by substituting the Higgs doublet from eq. (1.1.13). Only including terms linear in the vacuum expectation value and h gives

$$\mathcal{L}_{\text{Yukawa}} = -\frac{v}{\sqrt{2}}\,(v+h)\,Y_u^{\alpha\beta}\overline{u_l^{\alpha}}u_R^{\alpha} - \frac{v}{\sqrt{2}}\,(v+h)\,Y_d^{\alpha\beta}\overline{d_l^{\alpha}}d_R^{\alpha} - \frac{v}{\sqrt{2}}\,(v+h)\,Y_e^{\alpha\beta}\overline{e_l^{\alpha}}e_R^{\alpha} + h.c., \tag{1.1.19}$$

were the indices, α and β indicate the fermion flavours. After diagonalising the Yukawa matrices we find the fermion masses $m_\alpha = y_\alpha v/\sqrt{2}$, where y_α are the matrix eigenvalues. We can see eq. (1.1.19) also contains terms which couple the Higgs to the fermions with a coupling strength proportional to the fermion masses.

1.1.5 Beyond the Standard Model

The Standard Model has been hugely successful in describing the fundamental particles and their interactions via the strong, weak and electromagnetic forces and its predictions have been experimentally verified with discoveries such as the top quark at Fermilab [8, 9] and the Higgs Boson at CERN [3, 4]. However, despite its usefulness and outstanding track record, the SM is known to be incomplete. Unsolved problems include, but are by no means limited to:

Neutrino masses - In the SM, there are no right handed neutrino states and therefore no neutrino mass terms, as these terms require a coupling between left and right chiral fields. However, the observation of neutrino oscillations [10] indicates that neutrinos do have mass.

Dark matter and dark energy - There is an inconsistency between the matter we can see, and measurements of galaxy rotation velocities [11] and gravitational lensing [12]. This implies the existence of additional matter that we have not yet detected, referred to as dark matter. Cosmological models predict that dark matter makes up around 85% of the matter in the universe. Additionally, explanations of the observed accelerating expansion of the universe require a new form of energy, referred to as dark energy [13], which is not described in the SM.

Baryon asymmetry - There is an asymmetry between the amount of matter and antimatter observed in our universe, necessitating mechanisms of CP violation [14].

However, the amount of observed CP violation from SM processes is insufficient to explain the observed baryon asymmetry of the universe. This motivates the exploration of further sources of CP violation beyond the SM. In chapters 4 and 5 we will look at constraints on CP violating effects from higher dimensional operators.

Flavour problem - The masses of the fermions range from the electron at 0.511 MeV to the top quark at 173 GeV. They enter the SM in a rather ad-hoc way via the Yukawa couplings to the Higgs boson and there is no clear theoretical motivation for them to span over multiple orders of magnitude.

1.2 Collider Phenomenology

High-energy particle colliders have been key tools in testing predictions of the Standard Model and in the search for new physics. The Large Electron-Positron (LEP) collider at CERN led to the observation of the W and Z bosons in 1983 [15, 16] and determined the number of lepton generations through the measurement of the Z boson decay width [17]. In 1994, the Tevatron proton-proton collider at Fermilab observed the top quark [18], and in 2012 the Large Hadron Collider (LHC) at CERN discovered a resonance in proton-proton collisions consistent with the Higgs Boson [3]. The discoveries of these particles have all been made through the observation of resonances, and the properties of the colliders determine whether the resonance will be kinematically allowed.

Complementary to the physical experiments at colliders, particle physics phenomenology provides a vital role in determining the observable consequences of theories such as the Standard Model. A significant part of this relies on the simulation of collisions through Monte Carlo event generators.

In this section we look at the properties of high energy colliders and what information about the initial and final states we can gain with particle detectors. We then review the principles behind resonance searches and conclude with an introduction to the computational and statistical techniques used to compare theory with experiment. Much of this section is based on [19–21].

1.2.1 Collider parameters

The beam energy at a collider determines which physical processes are kinematically allowed. The total centre of momentum (c.m.) energy in a collision between two ultra-relativistic ($p \gg m$) particles can be approximated as

$$E_{\mathrm{CM}} = \sqrt{s} \approx 2E_1 \approx 2E_2 \tag{1.2.1}$$

where s is the Mandelstam invariant $s = (p_1 + p_2)^2$. Here p_1 and p_2 denote the particle 4-momenta $p = (E, \vec{p})$ where E is the energy and $\vec{p}$ is the 3-momentum.

Colliders should ideally maximise the c.m. energy of the colliding particles, as this results in reaching a higher new physics threshold. However, this energy is reduced through synchrotron radiation. For a circular collider with radius R

$$\Delta E \propto \frac{1}{R}\left(\frac{E}{m}\right)^4 . \tag{1.2.2}$$

Therefore, accelerators with a larger radius using heavier particles will experience less energy loss for a given beam energy. A circular hadron collider can thus reach much higher energies than a lepton collider for a given radius.

It is also necessary to consider the structure of our targets when designing a collider. In electron-positron collisions all the energy is converted into the final state and this energy is well known. However, for a proton-proton collider only a fraction of the c.m. energy is carried by the partons (quarks and gluons). It is therefore necessary to deduce the partons involved in the collision and their energies from the resulting decay products. On the other hand, as the initial state inherently varies with each collision, hadron colliders do not need to operate over an energy range and different physical processes can occur. This can be both a blessing and a curse, as interesting interactions can take place but at the cost of higher background events. As leptons only interact under the electroweak force, their collisions are much cleaner with fewer background events.

Another important parameter is the luminosity. Instantaneous luminosity is the ratio of the number of particles passing each other per unit time to the cross section. A useful related measure is the integrated luminosity over time. This tells us how much data the experiment will collect in a certain run time of the collider. For example Run 1 of the LHC collected 20 fb^{-1}. Run 2 then collected a further 150 fb^{-1} and Run 3 which will commence in 2022 is expected to accumulate 300 fb^{-1}. The planned High Luminosity LHC (HL-LHC) is estimated to further increase the integrated luminosity by a factor of 10 to 3000 fb^{-1}.

The total number of scattering events is given by

$$N = \mathcal{L}\sigma(s), \tag{1.2.3}$$

where $\mathcal{L}$ is the integrated luminosity and $\sigma(s)$ is the total scattering cross section. With larger luminosities, the increased number of events reduces the statistical uncertainty and we are able to gain sensitivity to more subtle new physics effects. However, the uncertainty may then become dominated by experimental systematic errors and theoretical uncertainties. It is important that these are also reduced in order to gain the maximum amount of information from the data.

1.2.2 Measurements and observables

In a hadronic collision the parton and hadron (lab) centre-of-mass frames do not coincide. It is therefore desirable that any observable measured is invariant between these frames. As the particles are initially travelling in the z direction, taken to be along the beam axis, we look for observables that are invariant under longitudinal Lorentz boosts. Variables that only have transverse components satisfy this condition. Some useful observables are the transverse momentum, $p_T = \sqrt{p_x^2 + p_y^2}$ and the azimuthal angle about the z axis, ϕ.

Another common observable is the rapidity, given by

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \tag{1.2.4}$$

For ultra relativistic particles in the massless limit, $E \approx |\vec{p}|$, and the rapidity coincides with the definition of the pseudo-rapidity variable, η, defined by

$$y \to \frac{1}{2} \ln \left(\frac{1 + \cos\theta}{1 - \cos\theta} \right) = - \ln \left(\tan \left(\frac{\theta}{2} \right) \right) \equiv \eta, \tag{1.2.5}$$

where θ is the polar angle between $\vec{p}$ and the beam axis. As the rapidity and pseudo-rapidity are additive under longitudinal boosts, the difference between two rapidities is invariant.

We can use the pseudo-rapidity and azimuthal angle to define the separation between two objects

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{1.2.6}$$

This is particularly useful in jet definitions where ΔR gives the cone size; hadrons within this cone are considered to constitute the jet.

These are a sample of common observables used in experimental analyses. We shall see in chapter 3 how we can combine observables using multi-variate techniques to enhance sensitivity to a signal process and reduce background events. Furthermore, we can construct different observables to be sensitive to certain effects such as CP violation. This will be examined in chapters 3 and 4.

1.2.3 Resonance searches

The most common way of discovering high-mass particles is through a resonance signal. In QFT the propagator contribution of an unstable particle of mass M and total decay width Γ is given by

$$R(s) = \frac{1}{(s - M^2)^2 + \Gamma^2 M^2}, \tag{1.2.7}$$

which is known as the Breit-Wigner function. For two particles a and b, decaying to n final state particles via an unstable particle V^*, this function has a peak at the invariant mass given by

$$s = (p_a + p_b)^2 = (\sum_i^n p_i)^2 \approx M_{V^*}^2, \tag{1.2.8}$$

where p_a, p_b and p_i are the 4-momenta of a, b and the final state particles respectively, and M_{V*} is the mass of the unstable particle V^*. The distribution of the number of events with the invariant mass will therefore exhibit a peak at the mass of the unstable particle. The LHC was built to search for a Higgs resonance in the 115-140 GeV mass range, which was previously constrained by Tevatron and LEP. In 2012, the LHC succeeded in its goal, identifying a resonance consistent with the Higgs at mass $m_h = 125$ GeV.

However, if the centre of mass energy of the collider is below the invariant mass of a new particle then we will be unable to detect a resonance. This highlights the importance of looking at and optimising other observables that can be used to indirectly seach for new physics.

1.2.4 Higgs phenomenology

The dominant Higgs production mechanisms at the LHC are gluon fusion (ggF) and vector boson fusion (VBF). Higgs production in association with vector bosons (Vh) and top quarks ($t\bar{t}h$) are also important. This is due to the fact the Higgs predominantly couples to heavy particles. The Feynman diagrams for the production mechanisms are shown in fig. 1.1. Gluon fusion and associated top quark production modes provide sensitivity to the top quark Yukawa coupling, whilst the Vh and VBF modes allow us to measure the coupling strength of the Higgs to gauge bosons. The initial Higgs discovery was made through the analysis of events from $H \to \gamma\gamma$, $H \to ZZ^* \to 4l$ and $H \to WW^* \to l\nu l\nu$ decays. The low branching ratios of these decay modes are mitigated by the fact they have low backgrounds and easily

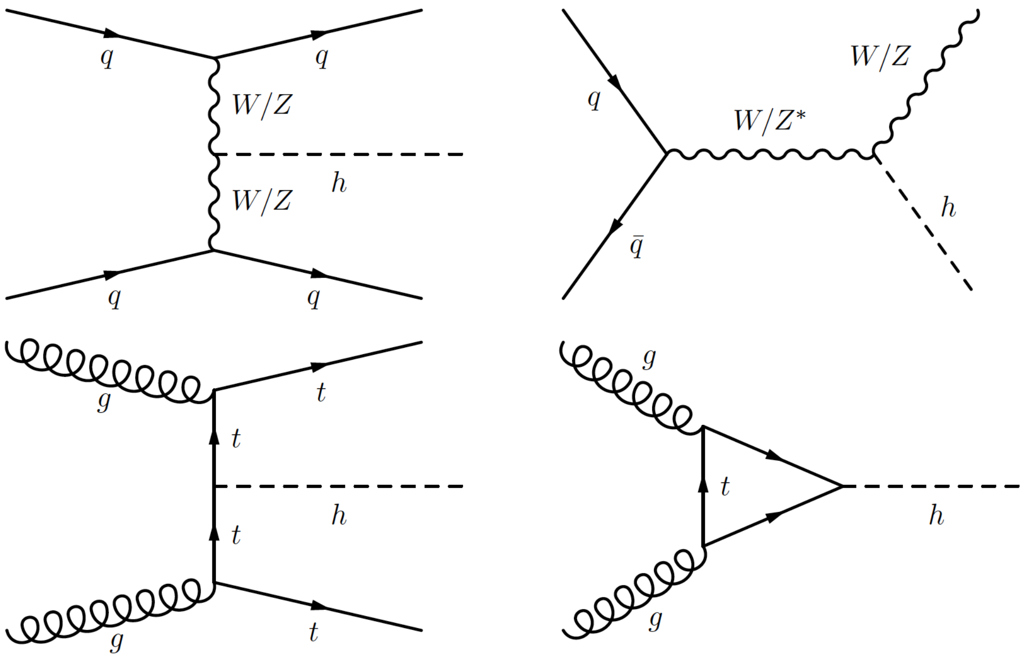


Figure 1.1: Feynman diagrams of the main Higgs boson production mechanisms; vector boson fusion (VBF) (top left), production in association with a vector boson (Vh) (top right), $t\bar{t}$ production (bottom left) and gluon fusion (ggF) (bottom right).

identifiable topologies. The largest Higgs branching ratio in the SM is $H \to b\bar{b}$. However, analysis of this process is complicated by the presence of a large QCD background. In chapter 3 we look at the associated production of a Higgs boson with a photon in weak boson fusion with the Higgs decaying to b quarks. This process is of interest as the additional photon in this channel suppresses the dominant QCD background [22, 23].

1.2.5 Monte Carlo event generators

Monte Carlo event generators are vital tools in particle physics. They are useful for calculating observables predicted by a theory in order to match against observables measured by experiments. The simulation of a collision at the LHC is a complicated task. In event generators the process is broken down into several stages, which are depicted schematically in fig. 1.2. Generally, the simulation starts with the hard scattering. This defines the main collision process. The parton distribution functions

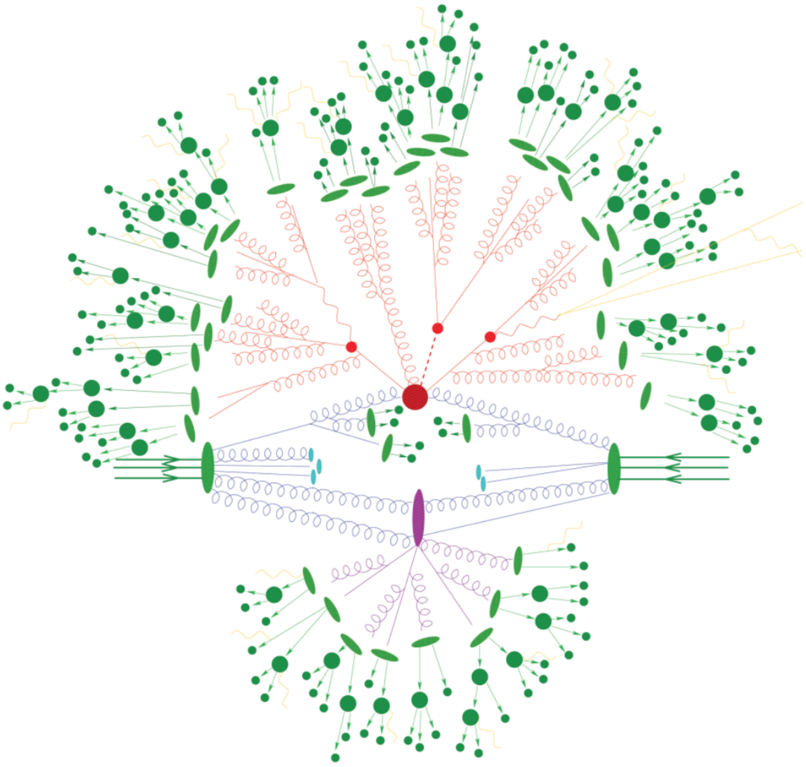


Figure 1.2: Schematic representation of the different stages of a particle collision from [24]. The hard scattering and subsequent parton shower are shown in red. The hadronisation stage and hadron decay are shown in green, whilst the purple oval indicates secondary interactions.

(PDFs) describe the initial partons in the process and matrix element generators give the probability distribution of the outgoing partons. Most event generators have inbuilt matrix element generators for tree level processes. However, for loop level processes an external tool is often interfaced.

The next stage is the parton shower. This describes the evolution of the hard scale down to the hadronization scale, the lower momentum scale at which the perturbation theory breaks down. The initial and final state particles of the hard process emit QED and QCD radiation. Gluons can trigger radiation themselves, producing an extended shower.

The final products of the collision must be colourless due to confinement. Hadronisation is the non-perturbative procedure used to cluster partons produced in the hard scattering and parton shower into hadrons. These hadrons may not be stable particles and so the final step is the modelling of hadronic decay.

In this thesis we make use of the `MadGraph` [25] and `SHERPA` [26] Monte Carlo event generators and the `OpenLoops` [27] matrix element generator for loop level processes.

1.2.6 Statistical analyses

In this section we outline the statistical procedure used in high energy particle physics to place limits on parameters of interest. We use this prescription in chapter 3, implemented using the `CheckMATE` [28] software.

The question of discovering or placing limits on a certain signal process is framed as a hypothesis test. The parameter of interest is the signal strength factor, μ, which is often expressed as $\mu = \sigma/\sigma_{\rm SM}$, where σ is the measured cross section of a certain process and $\sigma_{\rm SM}$ is the cross section of the same process predicted by the SM.

The number of events expected in the ith bin of the histogram of a certain variable of interest is given by,

$$E[N_i] = \mu S_i + B_i, \tag{1.2.9}$$

where S_i is the number of signal events and B_i is the number of background events. Therefore, we can have a null background-only hypothesis, H_0, with $\mu = 0$, and an alternate signal + background hypothesis, H_1, with a non-zero value of μ . For the SM signal hypothesis, $\mu = 1$.

For a given set of data the likelihood function is given by the product of the Poisson probabilities for all bins,

$$L(\mu, \theta) = \prod_{j=1}^{N} \frac{\left(\mu S_j + B_j\right)^{n_j}}{n_j!} e^{-\left(\mu S_j + B_j\right)}. \tag{1.2.10}$$

We can then construct the likelihood profile ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}_\mu)}{L(\hat{\mu}, \hat{\theta})} \tag{1.2.11}$$

where θ are nuisance parameters that affect the probability distribution of S_j and B_j but are not of interest to us. The variables $\hat{\mu}$ and $\hat{\theta}$ are values of μ and θ that maximize the likelihood function, whereas $\hat{\hat{\theta}}_\mu$ is the value of θ that maximises the likelihood for a given value of μ.

We can use the variable $\lambda(\mu)$ to construct the test statistic

$$q_\mu = -2\ln\left[\lambda(\mu)\right], \tag{1.2.12}$$

which is used to measure the discrepancy between a given dataset and a hypothesised value of μ. The discrepancy is quantified by the computation of the p-value

$$p_\mu = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu|\mu)dq_\mu \tag{1.2.13}$$

where $f(q_\mu|\mu)$ is the probability density function of q_μ under the assumption of signal strength μ.

In the CL$_s$ limit setting procedure, p_μ is calculated under the background only assumption, CL$_b$, and under the signal assumption for a certain value of μ, CL$_{s+b}$, with CL$_s$ = CL$_{s+b}$/(1-CL$_b$). The 95% CL$_s$ upper limit on μ is thus obtained by solving for CL$_s$ = 5%.

In experiments CL$_s$ is calculated with respect to the observed data thus the value of $q_{\mu,\mathrm{obs}}$ in the lower limit of the integral in eq. (1.2.13) is obtained from experimental data. In our case we are forecasting the expected CL$_s$ limits that could be obtained at a given luminosity. The value of $q_{\mu,\mathrm{obs}}$ is thus obtained using a simulated dataset with a standard model signal.

Part I

Constraining CP violating operators beyond the Standard Model

Chapter 2

Effective Field Theory

Effective field theories (EFTs) are founded on the principle of the separation of scales. This means that we can describe phenomena using only the relevant physics at the energy or length scale at which they occur. For example, you do not need to consider the topography of planets when analysing planetary motion. Likewise it is possible to calculate the hydrogen spectrum to some precision without considering the fact that the proton is made up of quarks and gluons. In this chapter we introduce the concept of EFTs with the example of Fermi Theory, before presenting the Standard Model Effective Field Theory (SMEFT) framework used in chapters 3 and 4. We discuss some key aspects of the phenomenology of SMEFT operators in the Warsaw basis as well as the methods and experimental processes used to constrain their Wilson coefficients. Parts of this chapter are based on [29–31].

2.1 EFTs in particle physics

In particle physics the assumption of a hierarchy of scales means that new physics could appear at a scale Λ which is much greater than the energy scale, E, at which an experiment is taking place. We can use effective field theories to describe the suppressed effects of the new physics at the experimentally available scale.

Heavy particles beyond the scale Λ are "integrated out" of the action by only performing the path integral over them. So if we have a Lagrangian describing light particles ψ_L and heavy particles ψ_H,

$$\mathcal{L}_{\text{full}}(\psi_L, \psi_H) = \mathcal{L}_H(\psi_H, \psi_L) + \mathcal{L}_L(\psi_L), \tag{2.1.1}$$

we can then integrate out the heavy states,

$$\int \mathcal{D}\psi_H e^{i\int \mathcal{L}_{\text{full}}(\psi_H, \psi_L)} = e^{\int i\mathcal{L}_{\text{eff}}(\psi_L)}. \tag{2.1.2}$$

where the effective Lagrangian is given by,

$$\mathcal{L}_{\text{eff}}(\psi_L) = \mathcal{L}_L(\psi_L) + \delta\mathcal{L}(\psi_L) \tag{2.1.3}$$

The $\delta\mathcal{L}(\psi_L)$ piece can be expanded into an infinite tower of terms of increasing dimension giving the EFT Lagrangian,

$$\mathcal{L}_{\text{EFT}}(\psi_L) = \mathcal{L}_L(\psi_L) + \sum_i c_i \frac{\mathcal{O}_i}{\Lambda^{\text{dim}(\mathcal{O}_i)-4}}, \tag{2.1.4}$$

Higher-order operators are suppressed by increasing powers of the scale Λ. Consequently, for lower energy processes, only a certain number of operators will realistically contribute. However, as we approach the Λ energy scale, the decrease in importance of terms with increase in expansion order no longer holds true. Thus, the EFT is only valid at energies below Λ which is sometimes referred to as the cut-off scale.

Effective field theories are extremely useful for studying both low and high energy effects. For example, if the theory is known to all scales, it may still be cumbersome to use the full theory in calculations at a particular scale. Therefore, EFTs are often used to simplify perturbative calculations. Conversely, if the full theory is unknown, then EFTs can be used to parametrise the unknown higher order interactions and constrain their effects and a lower attainable scale.

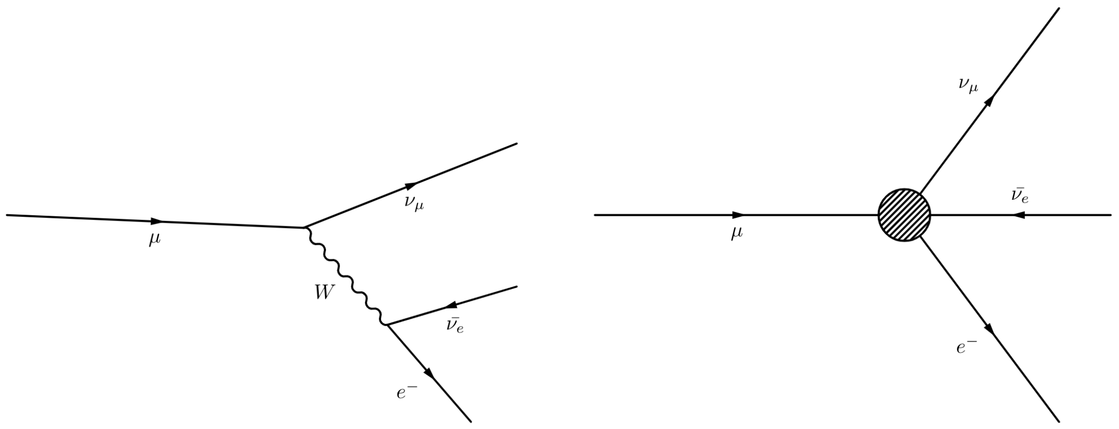


Figure 2.1: Feynman diagrams for muon decay in the full electroweak theory (left) and in the Fermi effective theory (right), where the shaded circle indicates the effective four point interaction vertex.

2.2 Fermi Theory

An example of an EFT in which we start with the full theory and integrate out the heavy states is Fermi Theory. This theory is an EFT of the SM below the weak scale in which the W and Z bosons are integrated out. It is therefore a good approximation of low energy phenomena, such as the decay of the muon $\mu \to \nu_\mu e \overline{\nu_e}$. In Fermi theory the amplitude for this process is given by the Feynman diagram in fig. 2.1 (right),

$$i\mathcal{M} = -i2\sqrt{2}G_F[\bar{u}(\nu_\mu)\gamma^\rho P_L u(\mu)][\bar{u}(e)\gamma^\sigma P_L v(\overline{\nu_e})], \tag{2.2.1}$$

where $P_L = 1/2(1-\gamma_5)$ is the left handed projection operator and G_F is the Fermi constant.

However, in the full SM theory this process is mediated by the W boson as show in the left Feynman diagram in fig. 2.1. Using the electroweak Feynman rules the SM amplitude for this process is given by,

$$i\mathcal{M} = \left(-i\frac{g}{\sqrt{2}}\right)^2 [\bar{u}(\nu_\mu)\gamma^\rho P_L u(\mu)]\frac{-i}{p^2 - M_W^2}\left(g^{\rho\sigma} - \frac{p_\rho p_\sigma}{M_W^2}\right)[\bar{u}(e)\gamma^\sigma P_L v(\overline{\nu_e})], \tag{2.2.2}$$

where M_W is the mass of the W boson and g is the weak coupling strength. At scales much lower than the mass of the W boson, $p^2 \ll M_W^2$, we can expand this amplitude in powers of $p^2/M_W^2 \ll 1$ giving,

$$i\mathcal{M} \approx -i\frac{g^2}{2M_W^2}[\bar{u}(\nu_\mu)\gamma^\rho P_L u(\mu)][\bar{u}(e)\gamma^\sigma P_L v(\overline{\nu_e})]\left(1 + \mathcal{O}\left(\frac{p^2}{M_W^2}\right)\right). \tag{2.2.3}$$

The first term in the expansion is equivalent to the four point interaction term in eq. (2.2.1). We can match the coefficients of these two terms to find the relation $G_F/\sqrt{2} = g^2/8M_W^2$. Thus, at energies, $E \ll M_W$, Fermi theory can be used as a valid approximation to the full electroweak theory.

2.3 Standard Model Effective Field Theory

As discussed in chapter 1, the Standard Model is an incredibly successful description of nature. However, we know that it is not complete. Given the absence of any significant direct or indirect deviation from its predictions at current colliders it is reasonable to assume that there could be new particle content which is much heavier than the weak scale. It is therefore possible to integrate out any potential heavy BSM particles and describe their effects in a model independent way using a Standard Model Effective Field Theory (SMEFT).

The Lagrangian of the SMEFT is built of the same fermion, gauge and higgs fields as the SM. It is constructed by considering all the gauge-invariant operators that can be built from these fields that are invariant under the SU(3)×SU(2)×U(1) gauge group. Lepton and baryon number conservation are also imposed. The terms in the Lagrangian can be organised by operator dimension. Here we will only consider operators up to dimension six giving,

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^{(5)}}{\Lambda}\mathcal{O}_i^{d=5} + \sum_i \frac{c_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{d=6}, \tag{2.3.1}$$

where the c_i are the Wilson coefficients which parametrise the importance of the effective interactions. It turns out that there is only one operator at dimension-five, the Weinberg Operator,

$$\mathcal{O}_5 = \left(H\overline{l_p^c}\right)\left(l_r H\right) + h.c \tag{2.3.2}$$

However, it violates lepton number conservation and so we only need to consider operators at dimension-six. A minimal, non-redundant set of operators is called a basis. For three generations of fermions, there are 2499 operators possible at dimension-six, thus constructing a basis is a highly non-trivial task. The first complete construction of a basis of dimension-six operators was derived in [32] and is referred to as the Warsaw basis. Ignoring flavour structure, the 59 [1] independent operators that are allowed are organised into eight classes: X, H^6, H^4D^2, X^2H^2, ψ^2H^3, ψ^2XH, ψ^2H^2D and ψ^4, where X is a field strength tensor, H is the higgs doublet and $\psi = q, u, d, l, e$ is a fermion field. The operators in each class are shown in table 2.1

The Warsaw basis is defined in terms of SU(3)×SU(1)×U(1) invariant operators before spontaneous symmetry breaking. However, to study their phenomenology it is more convenient to study interactions of the mass eigenstates in the broken phase.

2.3.1 Phenomenology of SMEFT operators

Dimension-six operators result in deviations from SM predictions through modifying the couplings of interactions that appear the SM Lagrangian, as well as though the addition of new vertices that are not present in the SM. Furthermore, the introduction of dimension-six operators results in shifts in the SM fields and parameters such as the boson masses, coupling constants and the Higgs vacuum expectation value.

Upon the addition of the Warsaw basis dimension-six SMEFT operators, the Higgs part of the Lagrangian becomes

$$\mathcal{L}_{\text{Higgs}} = \left|D_\mu H\right| - \mu H^\dagger H - \lambda\left(H^\dagger H\right)^2 + C_H\mathcal{O}_H + C_{H\Box}\mathcal{O}_{H\Box} + C_{HD}\mathcal{O}_{HD}, \tag{2.3.3}$$

---

[1] There are four additional four-fermion operators which violate the constraint of baryon number conservation.

where we redefine the Wilson coefficients $C_i = c_i/\Lambda^2$ and the operators, $\mathcal{O}_H, \mathcal{O}_{H\Box}, \mathcal{O}_{HD}$, are given in table 2.1. This results in a shift in the vacuum expectation value [33],

$$v \to v\left(1 + \frac{3v^2}{8\lambda}C_H\right). \tag{2.3.4}$$

The Higgs field is then expanded around the vacuum. However, the resulting form of the kinetic Higgs terms is now not canonical. It is therefore necessary to rescale the Higgs field in order to restore canonically normalised kinetic terms,

$$h \to h\left(1 + \frac{v^2}{4}C_{HD} - v^2 C_{H\Box}\right). \tag{2.3.5}$$

The phenomenological consequence of this field redefinition is that every interaction vertex involving the Higgs field will have a dependence on the C_{HD} and $C_{H\Box}$ Wilson coefficients. Furthermore, this also results in a shift in the Higgs mass with dependence on C_H, $C_{H\Box}$ and C_{HD} coefficients compared to the SM mass defined in eq. (1.1.14).

Similarly, for the gauge sector it is also necessary to rescale the W^i_μ, B_μ and G^a_μ fields, as well as the g_1, g_2 and g_3 couplings. These shifts depend on the C_{HW}, C_{HB} and C_{HG} Wilson coefficients respectively. Thus, after defining the physical $W^\pm$, Z and γ bosons in the mass basis, we get modifications to the W and Z boson masses dependent on C_{HW}, and C_{HW}, C_{HB}, C_{HWB} and C_{HD}, respectively. The electromagnetic and electroweak couplings are also modified with dependence on C_{HW}, C_{HB} and C_{HWB}. Therefore, the corrections to dimension-four vertices originating from parameter shifts in the gauge and Higgs sector can have dependence on these Wilson coefficients.

As mentioned previously, SM vertices can also receive direct contributions as a result of higher dimensional operators. For this the operator must have the same field content as the SM vertex after electroweak symmetry breaking. (Note that this is not a condition for the vertex to be modified by parameter shifts.) In order for this to be realised, the dimension-six SMEFT operator must have two or more powers of the Higgs field, or two or more derivatives compared to the dimension-four SM operator. In this manner, vertices can have additional contributions which can be roughly described as proportional to v^2/λ or p^2/λ respectively. When the additional contribution is proportional to v^2/λ, the SM coupling is rescaled. However, the p^2/λ proportionality means that the kinematics of the interaction change. In particular, the deviation from the SM will be greater at higher energies.

The final phenomenological effect of higher-dimensional operators is the addition of new vertices that are not present in the SM. For example, the four-point interaction between three gauge bosons and a Higgs boson. Both the $\mathcal{O}_{HWB}$ and $\mathcal{O}_{HW}$ operators, as well as their CP-odd counterparts contribute to this vertex. We will study this interaction in more detail in chapter 3.

In the gauge sector, the deviations from SM gauge couplings and new additional vertex structures are often parametrized in terms of anomalous couplings, κ_V, g_1^V, g_4^V, g_5^V, Λ_V, $\tilde{\Lambda}_V$ and $\tilde{\kappa}_V$, which can be written in terms of the dimension-six Wilson coefficients. For example, the triple gauge WWV vertex can be parameterised in terms of the effective Lagrangian,

$$\begin{aligned}
i\mathcal{L}_{\text{eff}}^{WWV} =& g_{WWV} \Big[g_1^V V^\mu \left(W^-_{\mu\nu} W^{+\nu} - W^+_{\mu\nu} W^{-\nu} \right) \\
&+ \kappa_V W^+_\mu W^-_\nu V^{\mu\nu} + \frac{\lambda_V}{M_W^2} V^{\mu\nu} W^{+\rho}_\nu W^-_{\rho\mu} \\
&+ i g_5^V \epsilon_{\mu\nu\rho\sigma} \left(\left(\partial^\rho W^{-\mu} \right) W^{+\nu} - W^{-\mu} \left(\partial^\rho W^{+\nu} \right) \right) V^\sigma \\
&+ i g_4^V W^-_\mu W^+_\nu \left(\partial^\mu V^\nu + \partial^\nu V^\mu \right) - \frac{\tilde{\kappa}_V}{2} W^-_\mu W^+_\nu \epsilon^{\mu\nu\rho\sigma} V_{\rho\sigma} \\
&- \frac{\tilde{\lambda}_V}{2M_W^2} W^-_{\rho\mu} W^{+\mu} \epsilon^{\nu\rho\alpha\beta} V_{\alpha\beta},
\end{aligned} \tag{2.3.6}$$

where $V \equiv Z, \gamma$. In the SM $g_1^V = \kappa_V = 1$ and all the other couplings are zero. The $\tilde{\kappa}_V$ and $\tilde{\lambda}_V$ couplings parameterise CP violating effects. It is also important to consider the possible heavy particles that can lead to these couplings. For example, the g_1^Z coupling can be generated by integrating out a new heavy vector boson. A model in which the SM is extended by a vector triplet and a vector singlet field is studied in [34]. This and other possible extensions to the SM which result in dimension-six operators in the low-energy regime, have been analysed in [35–38].

In this thesis we will mainly constrain SMEFT parameters in terms of Wilson coefficients. In section 4.3 we use a more general form of anomalous couplings to constrain CP violating effects in neutral triple gauge couplings.

2.3.2 Constraining Wilson coefficients

Constraints on SMEFT operators from experiment

In order to constrain the values of the SMEFT Wilson coefficients we look to experiments. The sources of data used to constrain the SMEFT operators relevant for Higgs physics include electroweak precision observables from the LEP experiment, as well as Higgs production and decay measurements at the LHC. Measurements of diboson production at LEP and the LHC also provide constraints. In particular, in the Warsaw basis, the $\mathcal{O}_{Hl}^{(3)}$, $\mathcal{O}_{Hl}^{(1)}$, $\mathcal{O}_{ll}$, $\mathcal{O}_{HD}$, $\mathcal{O}_{HWB}$, $\mathcal{O}_{He}$, $\mathcal{O}_{Hu}$, $\mathcal{O}_{Hd}$, $\mathcal{O}_{Hq}^{(3)}$, $\mathcal{O}_{Hq}^{(1)}$ and $\mathcal{O}_{W}$ operators are involved, and can be constrained, in diboson production and EW precision measurements. On the other hand, Higgs measurements are affected by the $\mathcal{O}_{eH}$, $\mathcal{O}_{dH}$, $\mathcal{O}_{uH}$, $\mathcal{O}_{G}$, $\mathcal{O}_{H\Box}$, $\mathcal{O}_{uG}$, $\mathcal{O}_{HW}$, $\mathcal{O}_{HB}$ and $\mathcal{O}_{HG}$ operators [39].

Higher dimensional operators not only affect the total production cross section of processes, they can also result in different vertex structures to the SM, as mentioned in the previous section. Therefore, differential distributions often provide greater sensitivity to the effects of SMEFT operators. Furthermore, additional derivatives in higher dimension operators mean they can be constrained through kinematic distributions, where their effect scales to be greater with larger p_T.

In a study of contributions from dimension-six operators to Higgs data [40] the VBF channel was found to be more sensitive to higher dimensional operators than gluon fusion and associated vector boson production modes. More recently, in a global fit [39] including Z-pole observables, WW production at LEP and LHC, and Higgs Run I and II data from the LHC, the largest impact on the constraint of the c_W Wilson coefficient came from WW production at LEP. Higgs Run II data provided the predominant constraint on c_{HW} and additionally constrained the c_{HWB}

parameter along with Z-pole measurements. Several other global fits combining these data sources (or subsets of) include [41–45].

Constraining CP-odd Wilson coefficients

Consider a scattering process with matrix element, $\mathcal{M}_{\mathrm{SM}} + \mathcal{M}_{\mathrm{NP}}$, where $\mathcal{M}_{\mathrm{SM}}$ is the SM matrix element, and $\mathcal{M}_{\mathrm{NP}}$ is the contribution from a particular dimension-six operator. This gives a squared matrix element of,

$$|\mathcal{M}|^2 = |\mathcal{M}_{\mathrm{SM}}|^2 + 2\mathcal{R}\left(\mathcal{M}^*_{\mathrm{SM}}\mathcal{M}_{\mathrm{NP}}\right) + \mathcal{O}(\Lambda^{-4}). \tag{2.3.7}$$

The resulting leading change in the cross section is thus proportional to the interference term in the equation above. However, for CP odd operators, the integration over this term vanishes in the calculation of the total cross section. Therefore, the CP-odd operators do not change the total event rate at the level of interference with the SM. Additionally, they do not contribute to CP-even observables such as invariant masses and transverse momenta. If we want to study CP-violating physics in the interference, it is necessary to look for a CP-odd observable. This is an observable whose expectation value vanishes if CP is conserved. For example, a commonly used CP-odd variable is the triple product of three momenta [46],

$$(\vec{p_1} \times \vec{p_2}) \cdot \vec{p_3}. \tag{2.3.8}$$

On the other hand, the quadratic $\mathcal{M}^2_{\mathrm{NP}}$ terms can contribute to the total cross section for CP-odd operators. This is due to the fact that these terms will always produce a CP-even effect. Thus, it is also possible to study CP-odd operators with momentum dependent observables. However these effects will generally be more subtle compared to the interference level effects in CP-odd observables as they are suppressed by higher powers of Λ.

1:X^3		2:H^6		3:H^4D^6	
$\mathcal{O}_G$	$f^{abc}G^{a\nu}_{\mu}G^{b\rho}_{\nu}G^{c\mu}_{\rho}$	$\mathcal{O}_H$	$\left(H^{\dagger}H\right)^3$	$\mathcal{O}_{H\Box}$	$\left(H^{\dagger}H\right)\Box\left(H^{\dagger}H\right)$
$\mathcal{O}_{\widetilde{G}}$	$f^{abc}\widetilde{G}^{a\nu}_{\mu}G^{b\rho}_{\nu}G^{c\mu}_{\rho}$			$\mathcal{O}_{HD}$	$\left(H^{\dagger}D^{\mu}H\right)^*\left(H^{\dagger}D^{\mu}H\right)$
$\mathcal{O}_W$	$\epsilon^{ijk}W^{i\nu}_{\mu}W^{j\rho}_{\nu}W^{k\mu}_{\rho}$				
$\mathcal{O}_{\widetilde{W}}$	$\epsilon^{ijk}\widetilde{W}^{i\nu}_{\mu}G^{j\rho}_{\nu}G^{k\mu}_{\rho}$				

4:X^2H^2		6:$\phi^2XH+h.c$		7:ϕ^2H^2D	
$\mathcal{O}_{HG}$	$H^{\dagger}H\,G^{A}_{\mu\nu}G^{A\mu\nu}$	$\mathcal{O}_{eW}$	$(\bar{l}_p\sigma^{\mu\nu}e_r)\tau^I HW^I_{\mu\nu}$	$\mathcal{O}^{(1)}_{Hl}$	$(H^{\dagger}i\overleftrightarrow{D}_{\mu}H)(\bar{l}_p\gamma^{\mu}l_r)$
$\mathcal{O}_{H\widetilde{G}}$	$H^{\dagger}H\,\widetilde{G}^{A}_{\mu\nu}G^{A\mu\nu}$	$\mathcal{O}_{eB}$	$(\bar{l}_p\sigma^{\mu\nu}e_r)HB_{\mu\nu}$	$\mathcal{O}^{(3)}_{Hl}$	$(H^{\dagger}i\overleftrightarrow{D}^I_{\mu}H)(\bar{l}_p\tau^I\gamma^{\mu}l_r)$
$\mathcal{O}_{HW}$	$H^{\dagger}H\,W^{I}_{\mu\nu}W^{I\mu\nu}$	$\mathcal{O}_{uG}$	$(\bar{q}_p\sigma^{\mu\nu}T^Au_r)\widetilde{H}\,G^A_{\mu\nu}$	$\mathcal{O}_{He}$	$(H^{\dagger}i\overleftrightarrow{D}_{\mu}H)(\bar{e}_p\gamma^{\mu}e_r)$
$\mathcal{O}_{H\widetilde{W}}$	$H^{\dagger}H\,\widetilde{W}^{I}_{\mu\nu}W^{I\mu\nu}$	$\mathcal{O}_{uW}$	$(\bar{q}_p\sigma^{\mu\nu}u_r)\tau^I\widetilde{H}\,W^I_{\mu\nu}$	$\mathcal{O}^{(1)}_{Hq}$	$(H^{\dagger}i\overleftrightarrow{D}_{\mu}H)(\bar{q}_p\gamma^{\mu}q_r)$
$\mathcal{O}_{HB}$	$H^{\dagger}H\,B_{\mu\nu}B^{\mu\nu}$	$\mathcal{O}_{uB}$	$(\bar{q}_p\sigma^{\mu\nu}u_r)\widetilde{H}\,B_{\mu\nu}$	$\mathcal{O}^{(3)}_{Hq}$	$(H^{\dagger}i\overleftrightarrow{D}^I_{\mu}H)(\bar{q}_p\tau^I\gamma^{\mu}q_r)$
$\mathcal{O}_{H\widetilde{B}}$	$H^{\dagger}H\,\widetilde{B}_{\mu\nu}B^{\mu\nu}$	$\mathcal{O}_{dG}$	$(\bar{q}_p\sigma^{\mu\nu}T^Ad_r)H\,G^A_{\mu\nu}$	$\mathcal{O}_{Hu}$	$(H^{\dagger}i\overleftrightarrow{D}_{\mu}H)(\bar{u}_p\gamma^{\mu}u_r)$
$\mathcal{O}_{HWB}$	$H^{\dagger}\tau^I H\,W^{I}_{\mu\nu}B^{\mu\nu}$	$\mathcal{O}_{dW}$	$(\bar{q}_p\sigma^{\mu\nu}d_r)\tau^I H\,W^I_{\mu\nu}$	$\mathcal{O}_{Hd}$	$(H^{\dagger}i\overleftrightarrow{D}_{\mu}H)(\bar{d}_p\gamma^{\mu}d_r)$
$\mathcal{O}_{H\widetilde{W}B}$	$H^{\dagger}\tau^I H\,\widetilde{W}^{I}_{\mu\nu}B^{\mu\nu}$	$\mathcal{O}_{dB}$	$(\bar{q}_p\sigma^{\mu\nu}d_r)H\,B_{\mu\nu}$	$\mathcal{O}_{Hud}$ + h.c.	$i(\widetilde{H}^{\dagger}D_{\mu}H)(\bar{u}_p\gamma^{\mu}d_r)$

8:$(\bar{L}L)(\bar{L}L)$		8:$(\bar{R}R)(\bar{R}R)$		8:$(\bar{L}L)(\bar{R}R)$	
$\mathcal{O}_{ll}$	$(\bar{l}_p\gamma_{\mu}l_r)(\bar{l}_s\gamma^{\mu}l_t)$	$\mathcal{O}_{ee}$	$(\bar{e}_p\gamma_{\mu}e_r)(\bar{e}_s\gamma^{\mu}e_t)$	$\mathcal{O}_{le}$	$(\bar{l}_p\gamma_{\mu}l_r)(\bar{e}_s\gamma^{\mu}e_t)$
$\mathcal{O}^{(1)}_{qq}$	$(\bar{q}_p\gamma_{\mu}q_r)(\bar{q}_s\gamma^{\mu}q_t)$	$\mathcal{O}_{uu}$	$(\bar{u}_p\gamma_{\mu}u_r)(\bar{u}_s\gamma^{\mu}u_t)$	$\mathcal{O}_{lu}$	$(\bar{l}_p\gamma_{\mu}l_r)(\bar{u}_s\gamma^{\mu}u_t)$
$\mathcal{O}^{(3)}_{qq}$	$(\bar{q}_p\gamma_{\mu}\tau^I q_r)(\bar{q}_s\gamma^{\mu}\tau^I q_t)$	$\mathcal{O}_{dd}$	$(\bar{d}_p\gamma_{\mu}d_r)(\bar{d}_s\gamma^{\mu}d_t)$	$\mathcal{O}_{ld}$	$(\bar{l}_p\gamma_{\mu}l_r)(\bar{d}_s\gamma^{\mu}d_t)$
$\mathcal{O}^{(1)}_{lq}$	$(\bar{l}_p\gamma_{\mu}l_r)(\bar{q}_s\gamma^{\mu}q_t)$	$\mathcal{O}_{eu}$	$(\bar{e}_p\gamma_{\mu}e_r)(\bar{u}_s\gamma^{\mu}u_t)$	$\mathcal{O}_{qe}$	$(\bar{q}_p\gamma_{\mu}q_r)(\bar{e}_s\gamma^{\mu}e_t)$
$\mathcal{O}^{(3)}_{lq}$	$(\bar{l}_p\gamma_{\mu}\tau^I l_r)(\bar{q}_s\gamma^{\mu}\tau^I q_t)$	$\mathcal{O}_{ed}$	$(\bar{e}_p\gamma_{\mu}e_r)(\bar{d}_s\gamma^{\mu}d_t)$	$\mathcal{O}^{(1)}_{qu}$	$(\bar{q}_p\gamma_{\mu}q_r)(\bar{u}_s\gamma^{\mu}u_t)$
		$\mathcal{O}^{(1)}_{ud}$	$(\bar{u}_p\gamma_{\mu}u_r)(\bar{d}_s\gamma^{\mu}d_t)$	$\mathcal{O}^{(1)}_{qd}$	$(\bar{q}_p\gamma_{\mu}q_r)(\bar{d}_s\gamma^{\mu}d_t)$
		$\mathcal{O}^{(8)}_{ud}$	$(\bar{u}_p\gamma_{\mu}T^Au_r)(\bar{d}_s\gamma^{\mu}T^Ad_t)$	$\mathcal{O}^{(8)}_{qd}$	$(\bar{q}_p\gamma_{\mu}T^Aq_r)(\bar{d}_s\gamma^{\mu}T^Ad_t)$
				$\mathcal{O}^{(1)}_{qd}$	$(\bar{q}_p\gamma_{\mu}q_r)(\bar{d}_s\gamma^{\mu}d_t)$
				$\mathcal{O}^{(8)}_{qd}$	$(\bar{q}_p\gamma_{\mu}T^Aq_r)(\bar{d}_s\gamma^{\mu}T^Ad_t)$

8:$(\bar{L}R)(\bar{R}L)+h.c$		8:$(\bar{L}R)(\bar{L}R)+h.c$		5:$\psi^2H^3+h.c$	
Q_{ledq}	$(\bar{l}^j_pe_r)(\bar{d}_sq_{tj})$	$Q^{(1)}_{quqd}$	$(\bar{q}^j_pu_r)\epsilon_{jk}(\bar{q}^k_sd_t)$	Q_{eH}	$(H^{\dagger}H)(\bar{l}_pe_rH)$
		$Q^{(8)}_{quqd}$	$(\bar{q}^j_pT^Au_r)\epsilon_{jk}(\bar{q}^k_sT^Ad_t)$	Q_{uH}	$(H^{\dagger}H)(\bar{q}_pu_r\widetilde{H})$
		$Q^{(1)}_{lequ}$	$(\bar{l}^j_pe_r)\epsilon_{jk}(\bar{q}^k_su_t)$	Q_{dH}	$(H^{\dagger}H)(\bar{q}_pd_rH)$
		$Q^{(3)}_{lequ}$	$(\bar{l}^j_p\sigma_{\mu\nu}e_r)\epsilon_{jk}(\bar{q}^k_s\sigma^{\mu\nu}u_t)$		

Table 2.1: The 59 independent dimension-six operators in the Warsaw basis following the notation in [47]. The p, r, s, t subscripts are flavour indices and σ^I are the Pauli matrices.

Chapter 3

Constraining SMEFT operators with associated hγ production in WBF

3.1 Introduction

Since the discovery of a new particle with a mass of 125 GeV by the ATLAS and CMS collaborations at the LHC, intense efforts have been made to measure its properties to confirm they are consistent with those predicted by the Standard Model. Precise measurements of the Higgs couplings to other particles are of great importance as they probe the underlying mechanism of electroweak spontaneous symmetry breaking (EWSB). These couplings determine the production and decay rates of the Higgs boson and therefore deviations in rates from those predicted in the SM are a key indication of new physics.

However, to date no significant evidence of deviation from the SM has been found. This has lead to the emergence of the Standard Model Effective Field Theory as one of the most useful tools to systematically study the data from the LHC and other experiments for hints of new physics in a model independent way. With the possibility of another, higher energy, collider many years away, it is vital to extract as much information as possible from processes accessible at the LHC. The analysis of data from multiple complementary channels is therefore very useful.

In this chapter we study the associated production of a Higgs boson with a photon in weak boson fusion (WBF), with the Higgs decaying to a b-quark pair. The dominant decay of the Higgs boson is to b-quarks. However, with the production mechanisms of gluon fusion and weak boson fusion, the resulting final state consists solely of jets. This suffers from a large background of non-resonant b-quark production. In [22, 23], WBF with an associated photon is proposed as a possibly interesting channel to observe the Higgs boson decay to b-quarks, as the additional photon suppresses the otherwise dominant QCD background. The ATLAS collaboration studied this channel in [48] with a boosted decision tree (BDT) at 30.6 fb^{-1} and found a significance of 1.4σ. In section 3.3 we analyse the potential of this channel for an independent measurement of the b-quark Yukawa coupling at higher luminosities.

In section 3.4 we investigate the effects of possible Beyond the Standard Model physics on WBF $h\gamma$ production using the SMEFT framework. The Wilson coefficients of operators relevant to Higgs physics have been constrained in [40–45, 49–68]. These works study a variety of channels including WBF. Here we propose the use of WBF $h\gamma$ production as an additional independent constraint. We particularly focus on CP-odd operators in the gauge-Higgs sector of the SMEFT. The limits on CP-odd Wilson coefficients are important as they provide constraints on additional sources of CP violation which are needed to explain electroweak baryogenesis. Furthermore, both the h and $h\gamma$ WBF final states have comparable sensitivity to CP-odd operators. The CP-odd operators relevant for our study have already been constrained in Higgs boson [69–73] and diboson production processes [74–77] . In this work we show that the WBF $h\gamma$ final state is also a relevant signature in placing constraints. We additionally propose the use of a CP-sensitive observable in order to constrain the $c_{H\tilde{W}}$ and $c_{H\tilde{W}B}$ Wilson coefficients in the Warsaw basis.

3.2 Signal and Backgrounds in the Standard Model

3.2.1 Process simulation

For our study we assume $\sqrt{s} = 13$ TeV throughout. The signal process of $h\gamma$ production in association with two jets at $\mathcal{O}(\alpha^4)$, is simulated with `MadGraph5 v2.6.6` [25] at leading order (LO) and with the default `NNPDF23_NLO` parton distribution function [78]. We use `PYTHIA 8.2` [79] to model secondary emissions through parton showering, perform hadronization and add the underlying event. It also decays the Higgs boson into the b-quarks. We select the WBF topology through the usual invariant mass cut on the tagging jets m_{jj}. All jets, at both parton and hadron level, are defined through the anti-kT algorithm [80] with $R = 0.4$. In the following, the indices j and b refer to the light and b-jets respectively. At generation level the following parton-level cuts are applied;

$$\begin{aligned} &p_{T,j} > 30\,\text{GeV}, \quad |\eta_j| < 5., \\ &p_{T,\gamma} > 20\,\text{GeV}, \quad |\eta_\gamma| < 2.5, \\ &\Delta R_{\gamma j} > 0.4, \quad m_{jj} > 1200\,\text{GeV}. \end{aligned} \tag{3.2.1}$$

The combination of these cuts ensures that non-WBF contributions (gluon fusion, tth and Vh) to the signal are negligible at the 10% level [48].

All irreducible background processes are simulated at LO using `Sherpa-2.2.7` [26] with the default `NNPDF30_NNLO` parton distribution function [81] from `LHAPDF 6.2.1` [82]; matrix elements are calculated with `COMIX` [83] and jets are parton showered with `CSSHOWER++` [84] [85]. For hadronisation we use the Sherpa default settings.

As background contributions to the signal final-state feature the direct production of b-jets in the simulation, it is necessary to implement additional generation-level cuts. We consider the following processes:

- Continuum production of a b-jet pair, two light jets and a photon, $b\bar{b}jj\gamma$. In particular, we consider $\mathcal{O}(\alpha_s^4\alpha)$ contributions which we denote QCD, as well as electroweak (EW) $Z\gamma jj$ production with the Z boson decaying to b-quarks. We impose the following additional cuts on these processes;

  $$\begin{aligned} &p_{T,b} > 20\,\text{GeV}, \quad m_{bb} \in [90,\ 200]\,\text{GeV}, \\ &\Delta R_{\gamma b} > 0.4, \quad \Delta R_{jb} > 0.4\,, \end{aligned} \tag{3.2.2}$$

  We have explicitly checked that the contributions from $\mathcal{O}(\alpha_s^2\alpha^3)$ are negligible at the 5% level and $\mathcal{O}(\alpha_s^3\alpha^2)$ as well as $\mathcal{O}(\alpha_s\alpha^4)$ contribute less than 1% each.

- $t\bar{t}\gamma$ production and single top production with an associated photon. For the $t\bar{t}\gamma$ and single top processes we force the decay of the $W^\pm$ boson to light quarks. We do not apply specific cuts on the decay products of the on-shell top quarks, but we require, again,

  $$\Delta R_{\gamma j} > 0.4 \tag{3.2.3}$$

  for the single-top processes.

3.2.2 Extracting the signal

In the initial analysis with `Rivet 2.7.0` [86] we apply the following baseline cuts to all signal and background processes:

1. We require an isolated photon with

   $$p_{T\gamma} > 20\,\text{GeV}, \quad |\eta_\gamma| < 2.5 \tag{3.2.4}$$

   and the isolation given by

   $$\sum_{i,\Delta R_{i\gamma}<0.4} p_\perp^i < 10\text{GeV}. \tag{3.2.5}$$

2. We require exactly two light jets and two b-jets,

$$N_{\text{jets}} = N_{b-\text{jets}} = 2, \tag{3.2.6}$$

where both are defined with the anti-kT algorithm with R=0.4 and

$$\begin{aligned} &p_{Tj} > 40\,\text{GeV}, \quad p_{Tb} > 30\,\text{GeV}, \\ &|\eta_j| < 4.5, \quad |\eta_b| < 2.5\,. \end{aligned} \tag{3.2.7}$$

We assume perfect b-tagging efficiency.

3. To select the WBF topology, we cut on the invariant light-jet mass and the pseudo-rapidity difference of the light jets

$$m_{jj} > 1500\,\text{GeV}, \quad \Delta\eta_{jj} > 4.5\,. \tag{3.2.8}$$

4. We require the invariant b-jet mass to be close to the Higgs mass

$$m_{bb} \in [100,\, 140]\,\text{GeV}\,. \tag{3.2.9}$$

This finalizes our baseline selection which we will use in the multivariate analysis in section 3.3.

5. To allow for a fair comparison between a cut-and-count approach and the multivariate analysis below, we apply the following additional cuts in our cut-and-count analysis

$$\begin{gathered} |\eta_{j_1}| > 1.5\,, \quad |\eta_{j_2}| > 2\,, \\ \eta^{\text{cen}}_{\gamma bb},\, \eta^{\text{cen}}_{\gamma}\,,\, \eta^{\text{cen}}_{bb} < 0.5\,, \\ m_{jj} > 2000\,\text{GeV}, \end{gathered} \tag{3.2.10}$$

where the centralities η^{cen}_x relative to the WBF tagging jets are defined as

$$\eta^{\text{cen}}_x = \left| \frac{\eta_x - \frac{\eta_{j_1}+\eta_{j_2}}{2}}{\eta_{j_1} - \eta_{j_2}} \right|. \tag{3.2.11}$$

The signal and background process cutflows are shown in Fig. 3.1. The baseline

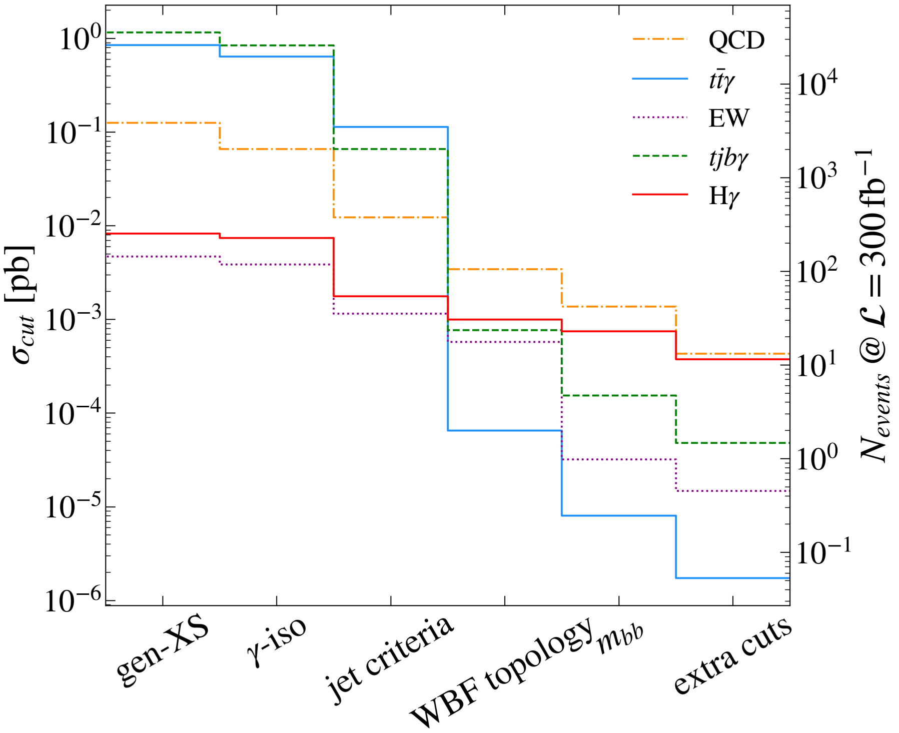


Figure 3.1: Cross section after different cuts in our cutflow, as given in eqs. (3.2.4) to (3.2.10). On the right axis, we display the number of events for an integrated luminosity of $300\,\mathrm{fb}^{-1}$. Be aware that we applied stronger cuts on the QCD and EW backgrounds at generator level which explains their lower generator-level cross section compared to the top backgrounds.

set of cuts, eqs. (3.2.4) to (3.2.9), reduces the contribution from $t\bar{t}\gamma$ and single top processes by six and four orders of magnitude respectively, whilst only loosing one order of magnitude in the signal. With the top-based backgrounds irrelevant after cuts, the dominant background contribution for associated $h\gamma$ production stems from the continuum QCD process.

After the final cuts in eq. (3.2.10), we reach a signal-over-background ratio of $S/B = 0.8$ in our cut-and-count analysis. We translate this into a CL$_s$ limit [87] on the signal strength

$$\mu = \frac{\sigma(\mathrm{pp} \to hjj\gamma)\ \mathrm{BR}(h \to b\bar{b})}{\sigma^{\mathrm{SM}}(\mathrm{pp} \to hjj\gamma)\ \mathrm{BR}^{\mathrm{SM}}(h \to b\bar{b})} \tag{3.2.12}$$

using the CL$_s$ limit setting implementation in `CheckMATE` [28]. The resulting limits are $\mu < 1.1$ for $\mathcal{L}_{\mathrm{int}} = 30.6\,\mathrm{fb}^{-1}$ at 95% CL ($\mu < 0.4$ for $\mathcal{L}_{\mathrm{int}} = 300\,\mathrm{fb}^{-1}$ and $\mu < 0.3$

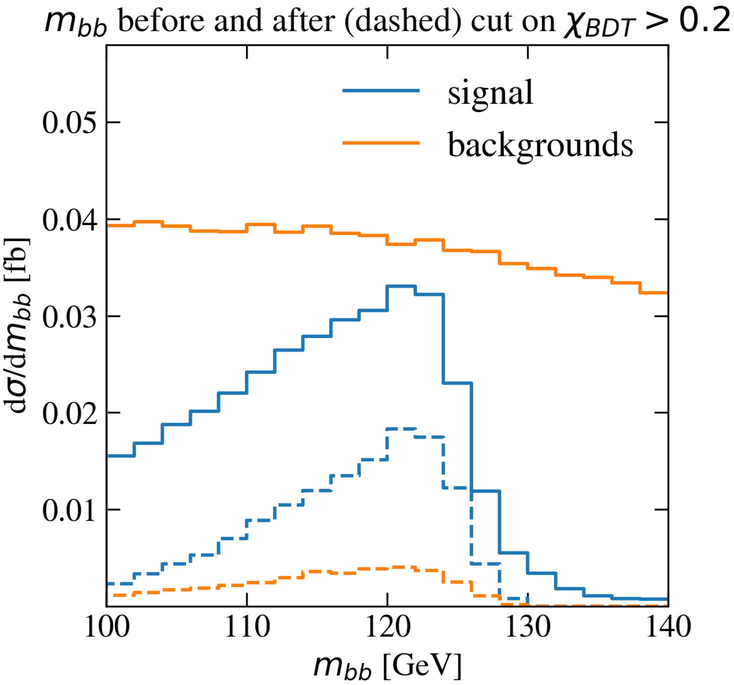


Figure 3.2: Distribution of the invariant mass of the b-jet pair m_{bb} before (solid lines) and after (dashed lines) a cut on the BDT classifier of $\chi_{\mathrm{BDT}} > 0.2$.

for $\mathcal{L}_{\mathrm{int}} = 3000\,\mathrm{fb}^{-1}$) assuming negligible systematic uncertainties.

3.3 Determination of the b-Yukawa coupling

The majority of studies of the decay of the Higgs to bottom quarks have been made with the Vh Higgs production mechanism. When the vector boson is required to decay leptonically, this channel has good sensitivity to the $H \to b\bar{b}$ decay. This decay has also been measured with VBF production in association with a photon in [48]. The measurement of the Higgs coupling to a b-quark pair in the VBF production mode is important as it provides an independent measurement complementary to the Vh mode. In this study we will look at the prospects of constraining the b-Yukawa coupling in the VBF topology at future luminosities.

In section 3.2.2 we applied a baseline set of cuts to events to extract the signal. As the final state is somewhat complicated, in order to enhance our sensitivity to the signal we perform a further multivariate analysis on events passing the cuts in

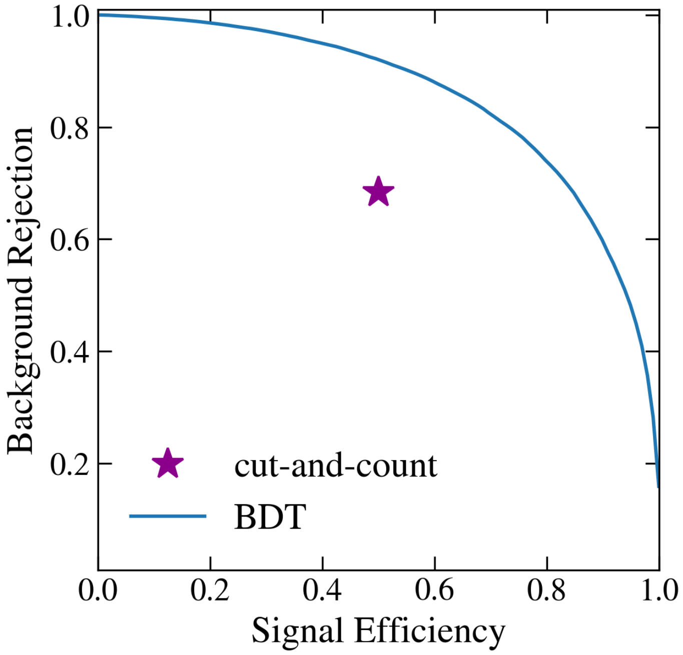


Figure 3.3: Receiver operating characteristic curve for the BDT analysis. The asterisk marks the signal and background efficiencies after the baseline and additional cuts in eqs. (3.2.4) to (3.2.10).

eqs. (3.2.4) to (3.2.9). We use a Boosted Decision Tree (BDT) with `TMVA` [88] in `Root 6.22` [89]. We generate N=200 trees with a maximum depth of 3 and set the minimum node size to 6%. As inputs we use the p_T and η of all the final state particles as well as the following variables,

$$\begin{aligned}
&m_{jj}, \quad \Delta\eta_{jj}, \quad \Delta\phi_{jj}, \quad \Delta R_{\gamma,j_1}, \quad \Delta R_{\gamma,j_2},\\
&m_{bb}, \quad \Delta\eta_{bb}, \quad \Delta\phi_{bb}, \quad \Delta R_{\gamma,b_1}, \quad \Delta R_{\gamma,b_2},\\
&p_{T,bb}, \quad \eta_{bb}, \\
&m_{bb\gamma}, \quad \Delta\eta_{\gamma,bb}, \quad \Delta\phi_{\gamma,bb}, \quad \Delta R_{\gamma,bb},\\
&\eta^{\text{cen}}_{\gamma bb}, \quad \eta^{\text{cen}}_{\gamma}, \quad \eta^{\text{cen}}_{bb}\,.
\end{aligned} \tag{3.3.1}$$

We find that the variable that is used most by the BDT is the invariant mass of the b-quarks. This is understandable as for the signal the m_{bb} distribution is peaked around the Higgs mass, whereas for the QCD background it is flat. The detector has a finite resolution and so we explicitly check that the BDT cut used for limit setting does not use an m_{bb} range finer than the detector resolution (see fig. 3.2).

$\mathcal{L}$ fb^{-1}	95% CL$_s$	χ_{BDT}
30.6	$\mu < 0.8$	0.1
300	$\mu < 0.25$	0.2
3000	$\mu < 0.1$	0.2

Table 3.1: Limits on the signal strength μ assuming negligible systematic uncertainties.

$\mathcal{L}$ fb^{-1}	95% CL$_s$	χ_{BDT}
30.6	$\mu < 0.9$	0.1
300	$\mu < 0.3$	0.2
3000	$\mu < 0.15$	0.2

Table 3.2: Limits on the signal strength μ assuming a systematic uncertainty of 50%.

When looking at the importance of other observables in the BDT classifier, we find that although they are less important individually than the invariant mass of the b-quark pair, their collective contribution is greater. We find that removing the m_{bb} variable reduces the efficiency of the signal classification by about 10%, for a fixed background efficiency of 10%.

We find our multivariate analysis with a BDT classifier clearly outperforms a cut and count analysis (including the cuts in eq. (3.2.10)). This can be seen in fig. 3.3, where we show the BDT receiver operating characteristic (ROC) curve in contrast with the cut-and-count efficiency.

The BDT classifier is chosen to minimise the CL$_s$ limit on the signal strength μ, for a given luminosity. Our results for integrated luminosities of $\mathcal{L} = 30.6$ fb^{-1}, 300 fb^{-1}, 3000 fb^{-1} are shown in table 3.1. Here we have neglected systematic uncertainties and assume statistical uncertainties to be dominant. These results show that an observation of the $h \to b\bar{b}$ decay channel will be possible at the HL-LHC.

However, at higher luminosities our assumption of negligible systematics no longer holds. If we instead assume a 50% systematic uncertainty on the background, the limits would be weakened to those shown in table 3.2.

3.4 EFT analysis

In this section we investigate the possible effects of Beyond the Standard Model physics in WBF $h\gamma$ production. We do this using the Standard Model Effective

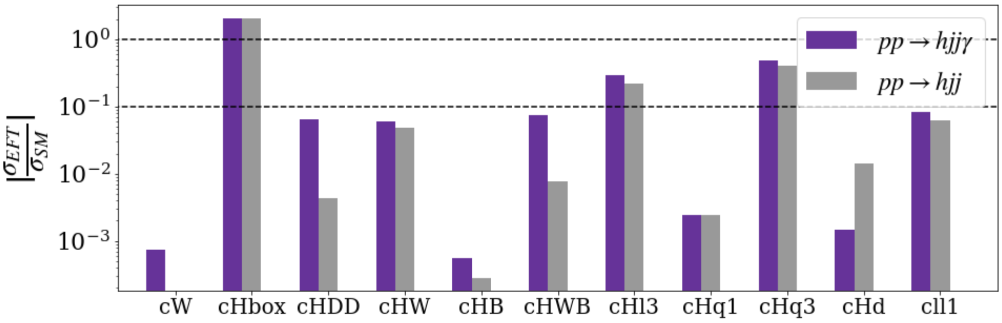


Figure 3.4: Relative cross section of relevant EFT operators with respect to the SM cross section for the WBF channel (grey) and for WBF with an additional photon (purple). The value of the Wilson coefficients is set to $c_i = \overline{c_i} + 2\sigma$ where $\overline{c_i}$ is the global fit central value from [39].

Field Theory framework introduced in section 2.3. We will look at which operators are relevant for our signal and in which areas of the phase space we have the most sensitivity to them. As the cross section of our signal process is relatively small, it is interesting to also look at CP violating operators. These do not contribute to the overall cross section at the level of interference with the SM, however we will study how they can be constrained using asymmetries.

3.4.1 Selection of operators

In principle there are many operators which contribute to our signal. The ratios of contributions from different operators to the total cross section with respect to the SM cross section are shown in fig. 3.4. The WBF $h\gamma$ channel has a relatively small cross section and as such, many of the Wilson coefficients of these operators will be constrained by other processes or by WBF without an additional photon, before our signal becomes sensitive.

In our study we will focus on operators which lead to interactions between three gauge bosons and the Higgs boson, as well as effective three-point interactions. In the SM, the signal is suppressed by three t-channel W propagators when the photon is radiated off the W boson and not an external quark. The Feynman diagram for

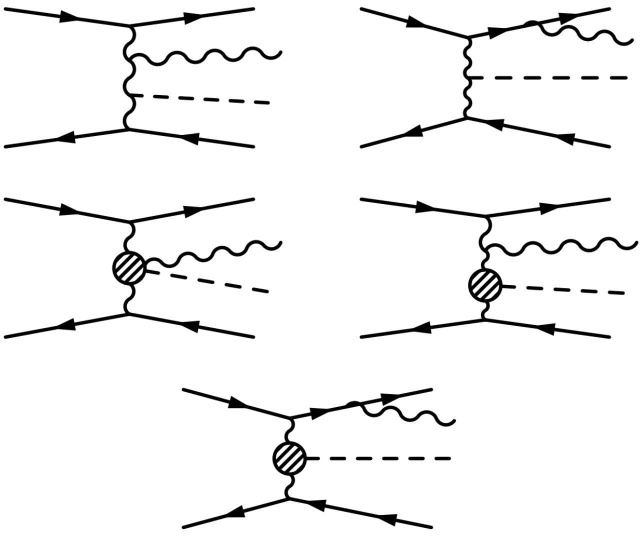


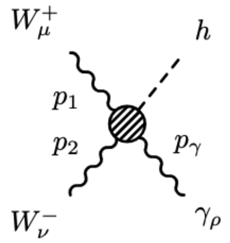
Figure 3.5: Example diagrams for WBF Higgs production in association with a photon in the SM (top row) and in the EFT (centre and bottom row).

this process is shown in the top left of fig. 3.5. However, the effective four-point interactions have the benefit of only being suppressed by two t-channel propagators (centre left of fig. 3.5). In order to enhance the relative importance of these diagrams, we impose the requirement of a large separation between the photon and the jets, $\Delta R_{\gamma j}$.

The following operators lead to interactions between three gauge bosons and a Higgs boson,

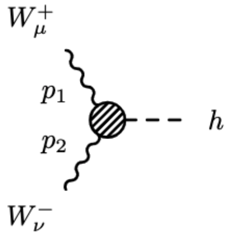
$$\begin{aligned} \mathcal{O}_{HW} &= H^\dagger H W^I_{\mu\nu} W^{I\mu\nu} & \mathcal{O}_{H\tilde{W}} &= H^\dagger H \tilde{W}^I_{\mu\nu} W^{I\mu\nu} \\ \mathcal{O}_{HWB} &= H^\dagger \tau^I H W^I_{\mu\nu} B^{\mu\nu} & \mathcal{O}_{H\tilde{W}B} &= H^\dagger \tau^I H \tilde{W}^I_{\mu\nu} B^{\mu\nu} \,. \end{aligned} \tag{3.4.1}$$

The four point $WWh\gamma$ interaction which arises from these operators has the vertex structure [30],



$$
\begin{aligned}
&+c_{HWB}\frac{2iev}{\Lambda^2}\frac{s_\theta}{c_\theta}\left(p_\gamma^\mu g_{\alpha\nu}-p_\gamma^\nu g_{\alpha\mu}\right)\\
&-c_{HW}\frac{4iev}{\Lambda^2}\left(g_{\alpha\mu}(p_\gamma-p_1)^\nu+g_{\alpha\nu}(p_2-p_\gamma)^\mu+g_{\mu\nu}(p_1-p_2)^\alpha\right)\\
&-c_{H\tilde{W}B}\frac{2iev}{\Lambda^2}\frac{s_\theta}{c_\theta}\,\epsilon_{\alpha\nu\mu\rho}\,p_\gamma^\rho\\
&+c_{H\tilde{W}}\frac{4iev}{\Lambda^2}\,\epsilon_{\alpha\nu\mu\rho}\left(p_\gamma+p_1+p_2\right)^\rho .
\end{aligned}
\tag{3.4.2}
$$

The contribution to this vertex arising from the $\mathcal{O}_{HW}$ operator is exactly the same as the structure of the SM $WW\gamma$ vertex. For this operator the SM and EFT diagrams differ only by the additional W t-channel propagator in the SM. The resulting three point VVh interaction from the $\mathcal{O}_{HW}$ operator has additional momentum enhancement due to derivatives in the $W_{\mu\nu}$ field strength tensors. The WWh vertex resulting from this operator and its CP-odd counterpart has the structure [30],



$$
\begin{aligned}
&+c_{HW}\frac{4iev}{\Lambda^2}\left(p_1^\nu p_2^\mu-g_{\mu\nu}\,p_1\cdot p_2\right)\\
&+c_{H\tilde{W}}\frac{4iev}{\Lambda^2}\,\epsilon_{\mu\nu\rho\delta}\,p_1^\rho p_2^\delta.
\end{aligned}
\tag{3.4.3}
$$

The $\mathcal{O}_{HWB}$ and $\mathcal{O}_{H\tilde{W}B}$ operators only contribute to $hZ\gamma$, $h\gamma\gamma$ and ZZh interactions. These vertices do not allow for the photon in our process to be radiated from the t-channel propagator. Their contribution will therefore be suppressed by the jet-photon separation cut.

We generate events for the contributions of these operators to the signal with the `SMEFTsim` implementation [47] of the Warsaw basis with `MadGraph` [25], neglecting dimension six squared terms. After applying the same cuts as for the SM signal we parametrise the WBF $h\gamma$ cross section as

$$
\begin{aligned}
\left.\frac{\sigma^{(\mathrm{LO})}_{\mathrm{SM+EFT}}}{\sigma^{(\mathrm{LO})}_{\mathrm{SM}}}\right|_{\mathrm{cuts}}-1 =&10^{-3}\cdot\left\{\left(\frac{1\,\mathrm{TeV}}{\Lambda}\right)^2\left[-44\,c_{HW}-240\,c_{HWB}\right]\right.\\
&\left.+\left(\frac{1\,\mathrm{TeV}}{\Lambda}\right)^4\left[83\,c_{HW}^2+23\,c_{HWB}^2+80\,c_{H\tilde{W}}^2+8\,c_{H\tilde{W}B}^2\right]\right\}.
\end{aligned}
$$

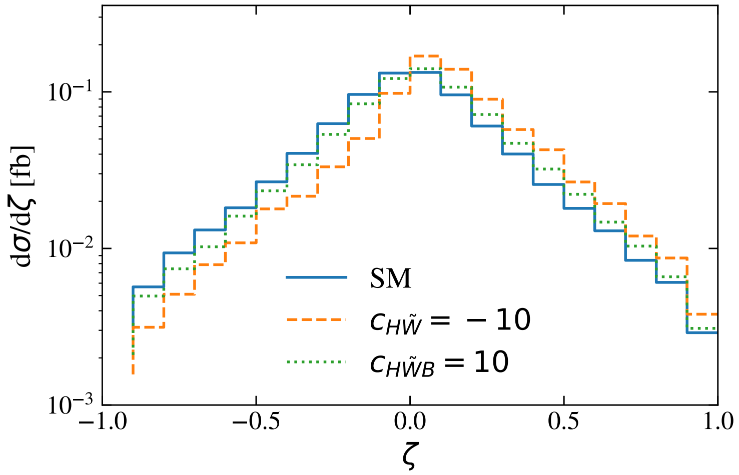


Figure 3.6: Distribution of the CP angle ζ in the SM and including the interference with the EFT.

Here we have shown the dimension squared terms although they are generally neglected elsewhere in our analysis. The CP-odd contributions to the cross section cancel out at the level of interference with the SM. However, their effects can still be observed through angular distributions of final state particles.

3.4.2 CP structure of the EFT and observable consequences

CP violation can only be detected without ambiguity through CP-odd observables. These can be constructed from scalar and triple products of the momenta of final state particles. As we have four particles in our final state, we scan over possible combinations and find the best sensitivity with the combination,

$$\zeta = \frac{\vec{p}_\gamma \cdot (\vec{p}_{j_2} \times \vec{p}_{bb})}{|\vec{p}_\gamma||\vec{p}_{j_2}||\vec{p}_{bb}|}\,, \tag{3.4.4}$$

where $\vec{p}_\gamma$ is the momentum of the photon, $\vec{p}_{j_2}$ is the momentum of the second p_T-ordered tagging jet and $\vec{p}_{bb}$ is the momentum of the Higgs boson reconstructed from the two b-jets.

For CP-odd operators this quantity should exhibit an asymmetry between the number of positive sign events, N_{ζ^+} and the number of negative sign events, N_{ζ^-}. This is shown in fig. 3.6 where we compare the distribution of ζ for the CP-even SM and the CP-odd EFT. Here we have chosen rather large values of the Wilson coefficients for illustrative purposes. We can see the SM distribution is symmetric, whereas the EFT clearly introduces an asymmetry.

There is no contribution to the total cross section from the interference of the CP-odd operators with the SM. It is therefore convenient to study a normalised asymmetry between negative and positive events, rather than looking at N_{ζ^+} and N_{ζ^-} directly. We define the asymmetry as,

$$A_\zeta = \frac{N_\zeta^+ - N_\zeta^-}{N_\zeta^+ + N_\zeta^-}, \qquad A_\zeta^{\mathrm{SM}} = 0. \tag{3.4.5}$$

After the baseline cuts in eqs. (3.2.4) to (3.2.9) this asymmetry can be parametrised in terms of the Wilson coefficients as,

$$A_\zeta = 10^{-3} \cdot \left(\frac{1\mathrm{TeV}}{\Lambda}\right)^2 \cdot \Big[-39\, c_{H\tilde{W}} + 12\, c_{H\tilde{W}B} \Big]. \tag{3.4.6}$$

In this parametrisation we neglect the dimension six squared terms for the CP-odd operators. These would modify the above numbers by less than 10% for $\mathcal{O}(1)$ Wilson coefficients.

Considering only statistical uncertainties, we are able to constrain the Wilson coefficients $c_{H\tilde{W}}$ and $c_{H\tilde{W}B}$ to,

$$\frac{c_{H\tilde{W}}}{\Lambda^2} < \frac{1.1}{\mathrm{TeV}^2} \qquad \frac{c_{H\tilde{W}B}}{\Lambda^2} < \frac{3.6}{\mathrm{TeV}^2} \quad \text{at 95\% CL.} \tag{3.4.7}$$

The magnitude of the asymmetry, A_ζ, depends on the kinematic region. This is due to the fact that the relative contributions of particular diagrams are enhanced in certain regions. Therefore, the cuts that we impose on the final state can have an effect on the asymmetry. In fig. 3.7 we show how A_ζ varies with the cut on the invariant mass of the Higgs-photon pair, $m_{bb\gamma}$. The more drastic the cut, the greater the asymmetry. However, from the lower panel, we can see the drop in the cross section as the cut on $m_{bb\gamma}$ is increased. There is therefore a trade off between

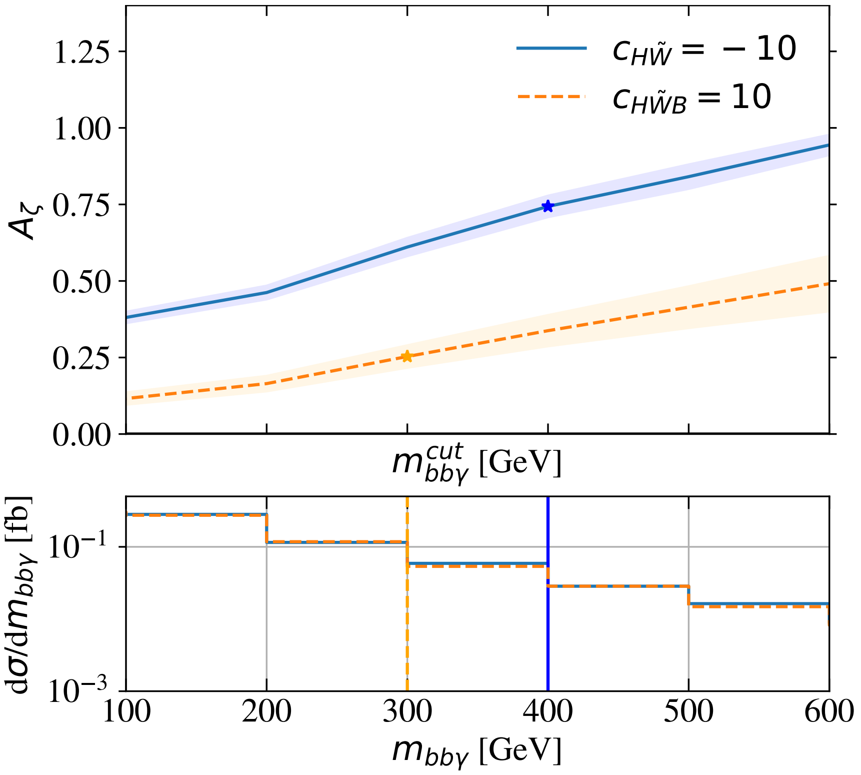


Figure 3.7: Dependence of the asymmetry A_ζ on a cut on the invariant mass of the Higgs-photon pair $m_{bb\gamma} > m^{\text{cut}}_{bb\gamma}$. The shaded band represents the statistical uncertainty on the asymmetry assuming an integrated luminosity of 3 ab^{-1}. The asterisk highlights the optimal cut on the invariant mass for the given Wilson coefficient. In the lower panel we show the distribution of the cross section as a function of the $m_{bb\gamma}$ invariant mass.

selecting a signal region with a large asymmetry whilst still remaining inclusive enough to reduce statistical uncertainties.

If we assume an optimal cut on the invariant mass of the Higgs-photon pair we can improve the limits in eq. (3.4.7) to ,

$$\frac{c_{H\tilde{W}}}{\Lambda^2} < \frac{1.1}{\text{TeV}^2} \qquad \frac{c_{H\tilde{W}B}}{\Lambda^2} < \frac{3.1}{\text{TeV}^2} \qquad \text{at 95\% CL.} \tag{3.4.8}$$

The current best constraints on the Wilson coefficients $c_{H\tilde{W}}$ and $c_{H\tilde{W}B}$ come from observables in WBF and the Higgs decay $h \to ZZ \to 4l$ respectively [71]. These are quoted as $|\frac{c_{H\tilde{W}}}{\Lambda^2}| < \frac{1.2}{\text{TeV}^2}$ and $|\frac{c_{H\tilde{W}B}}{\Lambda^2}| < \frac{1.5}{\text{TeV}^2}$ for an integrated luminosity of 3 ab^{-1}. In comparing these constraints we note that our analysis did not include systematic errors, which we assume are larger for our process than for WBF without the additional photon. Nevertheless, this shows that a combination of the WBF $h\gamma$ signal with other processes probing the same dimension-six Wilson coefficients is useful in testing the underlying paradigms of the EFT and may resolve degeneracies in global fits.

3.5 Conclusions and outlook

In this chapter we studied WBF Higgs production with an associated photon. This channel is interesting as it probes the b-yukawa coupling in weak boson fusion, with the additional photon supressing the otherwise dominant QCD background. Using a cut-and-count analysis we find the 95% CL limit on the signal strength, $\mu < 1.1$ for $\mathcal{L}_{\text{int}} = 30.6$ fb^{-1}. The five-particle final state of this signature means that this process is a prime candidate for the use of multivariate techniques. Using a BDT we find an improved limit on the signal strength of $\mu < 0.8$ at 95% CL for $\mathcal{L}_{\text{int}} = 30.6$ fb^{-1}. At the HL-LHC we find the limit $\mu < 0.1$ at 95% CL. In the calculation of these limits we have assumed negligible systematic uncertainties. However, at high luminosities these may well become relevant. Assuming a systematic uncertainty of 50% on the background, we find the limit at the HL-LHC full luminosity of 3 ab^{-1} worsens to $\mu < 0.15$ at 95% CL.

We additionally investigated the sensitivity of the WBF $h\gamma$ channel to Beyond the Standard Model couplings in the dimension-six SMEFT framework. In particular, we focused on the CP-odd operators $\mathcal{O}_{H\tilde{W}}$ and $\mathcal{O}_{H\tilde{W}B}$ in the Warsaw basis. To constrain the Wilson coefficients of these operators we constructed a CP-sensitive observable from scalar and triple products of momenta of final state particles. We defined an asymmetry in this variable which we used to extract the limits, $\frac{c_{H\tilde{W}}}{\Lambda^2} < \frac{1.1}{\text{TeV}^2}$ and $\frac{c_{H\tilde{W}B}}{\Lambda^2} < \frac{3.6}{\text{TeV}^2}$ at 95% CL for the luminosity of the full HL-LHC dataset of 3 ab^{-1}, where we have assumed dominant statistical uncertainty. We find these limits to be comparable to those placed in [71] in a global fit of the Higgs sector including WBF without the additional photon. Therefore, including the WBF $h\gamma$ signature in global fits provides an independent test of the paradigms underlying the EFT.

Chapter 4

Constraining CP violating operators in Triple Gauge Couplings

4.1 Introduction

The non-Abelian structure of the SU$(2)_L$ symmetry group leads to the triple and quartic self-couplings of the gauge bosons. These can be probed at colliders through diboson production. In this chapter we study the CP-odd charged anomalous triple gauge couplings (aTGCs) that contribute to the WWZ and $WW\gamma$ vertices. Again we will use the framework of dimension-six operators in the SMEFT [90–94]. In particular, we look at constraints on the CP-odd Wilson coefficients coming from the measurement of, $WW \to ll'\nu\nu$ [95], $WZ \to l^+l^-l^\pm\nu$ [96], $W\gamma \to l^\pm\nu\gamma$ [97], $Zjj \to l^+l^-jj$ [98] and $Wjj \to l\nu jj$ [99] production. Constraints on these operators have been previously placed using Higgs [1, 69–73] and diboson [74–77, 100] production, as well as vector boson scattering [101]. The same set of processes as listed above is used in [102]. Our analysis differs from theirs in the use of asymmetries instead of full differential distributions. This has the benefit of reducing theoretical and experimental systematic errors, which cancel in the normalised asymmetry. Our

SM and BSM events are simulated with `SHERPA` and therefore provide independent confirmation of the results found in [102]. They also serve to test the validity of the `SHERPA` implementation of this sector.

Additionally, we study the interactions of neutral gauge bosons, $ZZ\gamma$, ZZZ and $Z\gamma\gamma$ in the processes, $ZZ \to 4l, 2l2\nu$ [103,104] and $Z\gamma \to 2l\gamma, 2\nu\gamma$ [105,106]. These couplings do not exist in the SM and are only introduced at dimension-8 in the SMEFT [107]. Furthermore, the squared neutral aTGCs dominate over the linear due to the polarization suppression of the SM interference [75,108,109]. This complicates the interpretation of neutral aTGCs in the dimension-8 SMEFT. Therefore, we instead use the most general parametrization of neutral aTGC's given in [110] to investigate CP-odd BSM effects. As the squared couplings are the most important, rather than constructing an asymmetry, we constrain these effects in the high-pT regime.

4.2 Charged aTGCs

Couplings between three gauge bosons exist in the Standard Model. However, higher dimensional operators of the SMEFT produce deviations from the couplings predicted by the SM. In diboson production the $\mathcal{O}_{\tilde{W}}$ and $\mathcal{O}_{H\tilde{W}B}$ operators introduce the possibility for CP violation through their modification of the $WW\gamma$ and WWZ interactions. Their effects can be described by the effective Lagrangian,

$$\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}} + \frac{c_{\tilde{W}}}{\Lambda^2}\mathcal{O}_{\tilde{W}} + \frac{c_{H\tilde{W}B}}{\Lambda^2}\mathcal{O}_{H\tilde{W}B}. \tag{4.2.1}$$

In this study we look at WW, WZ, Wγ and VBF Zjj and Wjj production channels at the LHC, to constrain the $c_{\tilde{W}}$ and $c_{H\tilde{W}B}$ Wilson coefficients. For each channel, i, we define an angle, ζ_i, from triple products of final state momenta. For the CP violating operators $\mathcal{O}_{\tilde{W}}$ and $\mathcal{O}_{H\tilde{W}B}$, the differential distributions of the angle exhibit modulations. An example is shown in fig. 4.1 for the $\Delta\phi_{jj}$ distribution in Zjj production. We can see from this plot that the modulation over the distribution

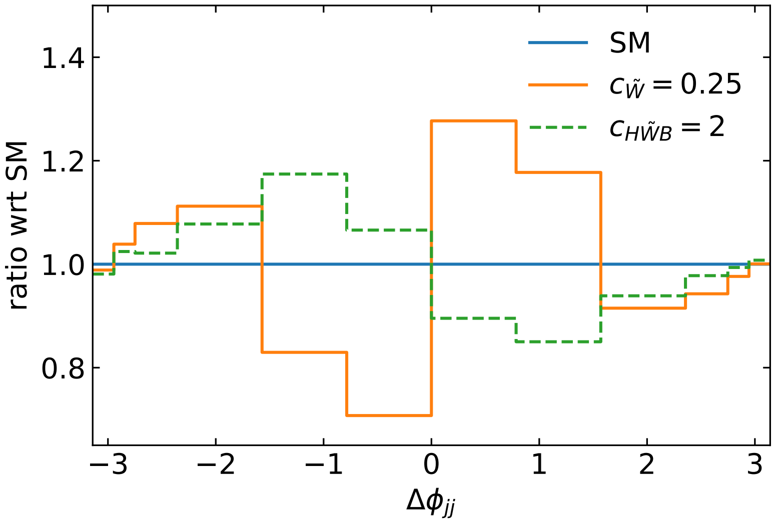


Figure 4.1: $\Delta\phi_{jj}$ distribution for the Zjj analysis.

varies above and below the SM for positive and negative values of $\Delta\phi_{jj}$. Therefore, by simply counting the difference between positive and negative sign events, as we do with our asymmetry definition in chapter 3, it is possible we could loose some sensitivity. Instead, we divide each of the $\Delta\phi$ distributions into pairs of bins with bin boundaries $\pm[0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, 1] \cdot \pi$. From these we construct the asymmetries, A_{ij}, which compare the difference in positive and negative sign events between each bin pair. For example, we compare the number of events in bin $[-\frac{\pi}{4}, 0]$ with those in bin $[0, \frac{\pi}{4}]$. The asymmetry for a certain bin, j, in a given channel,i, is given by,

$$\begin{gathered} A_{ij} = \frac{N_{i,-j} - N_{i,+j}}{N_{i,-j} + N_{i,+j}}, \\ i = WW,\, WZ,\, W\gamma,\, Zjj,\, Wjj, \quad j = 1, \ldots, 6\,, \end{gathered} \tag{4.2.2}$$

and is defined such that it vanishes for the CP-even SM and SM background.

In fig. 4.2 we show the asymmetry for the most relevant bins of the $\Delta\phi$ distribution for Zjj and $W\gamma$ production. We can see that for the Zjj distribution we get large asymmetries with larger uncertainties, compared to $W\gamma$, in which the asymmetries and their uncertainties are smaller. In using bin-wise asymmetries, the statistical uncertainty is increased compared to the global asymmetry (as shown in fig. 4.2). However, the magnitude of the asymmetry also increases, allowing us to trade off better statistics for larger generated asymmetries in a constrained phase-space region.

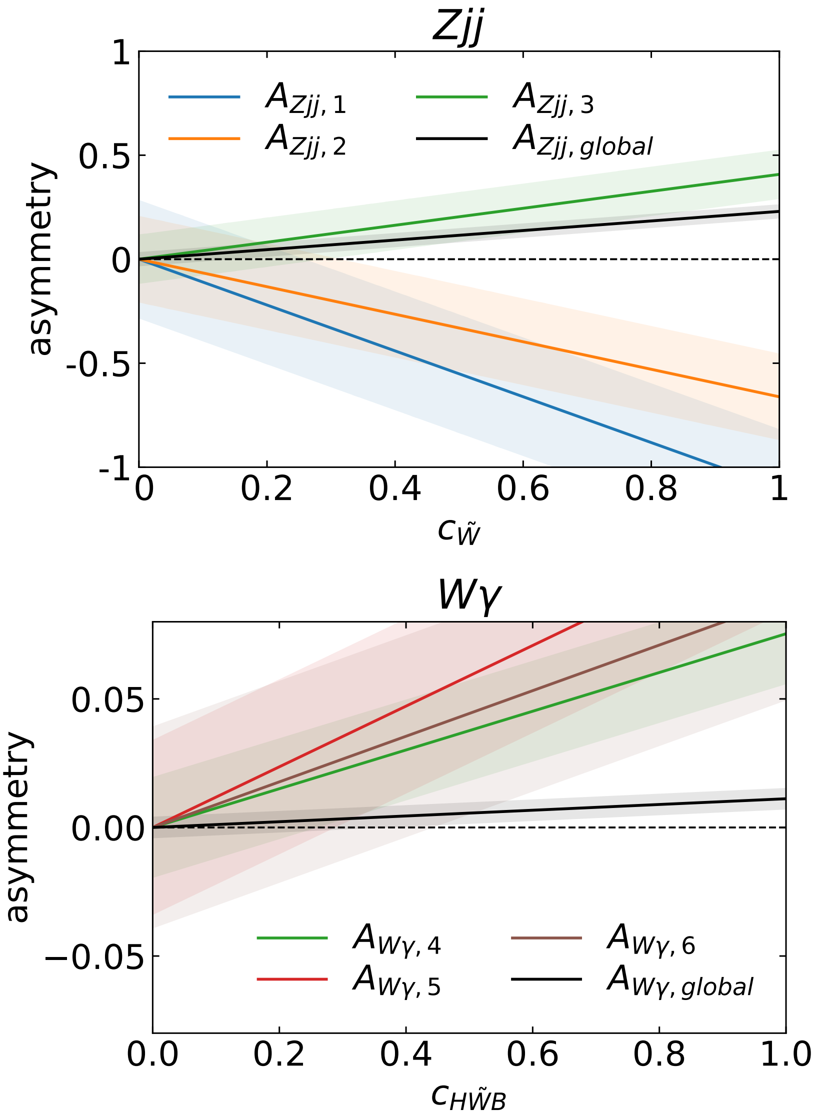


Figure 4.2: Dominant bin-wise asymmetries as defined in Eq. (4.2.2) for the Zjj (top) and the $W\gamma$ (bottom) channels. The shaded bands show the 2σ uncertainty on the asymmetry from statistics and background subtraction. The channels shown result in the strongest bounds on $c_{\tilde{W}}$ and $c_{H\tilde{W}B}$ respectively. The global asymmetry is defined using the full $\Delta\phi$ phase space.

An advantage of asymmetries is that by taking a ratio of events, the systematic uncertainties largely cancel. Therefore, the limits are determined by statistical uncertainties and we do not consider systematic uncertainties. The only exception from this is the error from background subtraction, which can be relevant for processes with a small signal to background ratio. For channels with $S/B < 1$ we take the background subtraction uncertainty to be, $\sigma^{\text{bkg}}_{A_{ij}} = \sqrt{N^{\text{bkg}}_{ij}}/N^{\text{sig}}_{ij}$. For the WZ, $W\gamma$

and WW processes, $S/B > 1$ and the inclusion of σ^{bkg} affects the limits by less than 3%, 12% and 20% respectively.

4.2.1 Process simulation

Events are generated at leading order with `SHERPA-2.2.10` [26] with the default `NNPDF30_nnlo_as_0118` parton distribution function [81] from `LHAPDF` 6.2.1 [82]; matrix elements are calculated with `COMIX` [83] and parton showered with `CSS` [84]. QED corrections are effected through a YFS soft-photon resummation [111, 112]. For multi-parton interactions, hadronisation, and subsequent hadron decays we use the `SHERPA` default settings. EFT contributions are generated using the `SMEFTsim` model [47] in `SHERPA` through its UFO [113] interface [114]. We consider the interference of the SM with the dimension-six operator only and neglect contributions from the dimension-six squared terms.

For each of these channels we normalise the cross section for the SM to the cross section observed in experiment and assume the same normalisation factor for the EFT contribution. In order to take into account detector effects, for each channel we multiply by a flat efficiency calculated from the ratio of the number of predicted events to the predicted cross section, $\epsilon_{\text{det}} = N_{\text{events}}/(\sigma_{\text{pred}}\mathcal{L}_{\text{int}})$. In the following we give the specific details for each process.

WW production

For WW production, we consider an asymmetry in the sine of the difference of the azimuthal angles ϕ of the two final state leptons ordered by their pseudorapidity, $\zeta_{WW} = \Delta\phi_{\ell\ell}$. We make use of the existing `Rivet` [86] analysis to reproduce the experimental cuts and normalize the `Sherpa` cross section to the measured value of $\sigma_{\text{fid},EW} = 379.1 \pm 27.1$ fb [95]. The detector efficiency is deduced from the difference between the predicted cross section and the predicted number of events

$\epsilon_{\text{det}} = 0.61$. Since the signal-over-background ratio $S/B > 1$, we can safely neglect the uncertainty from background subtraction.

WZ production

For WZ production, the CP-sensitive observable considered is $\zeta_{WZ} = \Delta\phi_{Z\ell'}$, where ℓ' denotes the lepton from the decay of the W boson and Z denotes the reconstructed Z boson from the same-flavor-opposite-sign lepton pair. We normalize the `SHERPA` cross section to the measured value of $\sigma_{\text{fid},EW} = 254.7 \pm 11.5$ fb [96] and assume a detector efficiency of $\epsilon_{\text{det}} = 0.52$. Since the signal-over-background ratio $S/B > 1$, we can safely neglect the background contributions.

$W\gamma$ production

For $W\gamma$ production in the $\ell\nu\gamma$ final state we define the CP-sensitive observable $\zeta_{W\gamma} = \Delta\phi_{\gamma\ell}$, where ℓ and γ denote the lepton from the W boson decay and of the photon, respectively. CMS has performed an analysis for $W\gamma$ production at 13 TeV for an integrated luminosity of $\mathcal{L} = 127.1$ fb^{-1} [97]. Including the decay of the W boson, the analysis has measured a cross section of $\sigma_{\text{fid}} = (3.32 \pm 0.16)$ pb. We implemented the experimental cuts in `Rivet` and normalized the cross section after cuts to this value. From the expected number of signal events and the expected cross section, we deduce a detector efficiency of $\epsilon_{\text{det}} = 0.59$. For $W\gamma$ production, the signal-over-background ratio after cuts is $S/B \sim 0.54$ and we therefore explicitly take the uncertainty from background subtraction into consideration. Since the dominant background contributions arise from experimental effects such as nonprompt leptons and photons and e-induced photons, it is difficult to estimate their shape. For this reason, we assume the background shape to closely follow the signal shape in $\Delta\phi_{\gamma\ell}$.

Zjj production

In vector boson fusion Zjj production, CP violation in the ZWW and γWW couplings causes modulations in the $\Delta\phi_{jj}$ distribution of the η-ordered jets, see Fig. 4.1. We normalize the `SHERPA` cross section to the measured value of $\sigma_{\text{fid},EW} = 37.4 \pm 6.5$ fb [98] and take into account a factor of $\epsilon_{\text{det}} = 0.85$ for detector effects. Since $S/B \sim 0.59$, we consider the uncertainty from background subtraction using the $\Delta\phi_{jj}$ distribution for the background as given in the experimental reference.

Wjj production

For VBF Wjj production, we again base our analysis on the $\Delta\phi_{jj}$ distribution of the η-ordered jets. On top of the baseline selection used in [99], we apply a stricter cut on the invariant mass of the tagging jets $m_{jj} > 1100\,\text{GeV}$, resulting in a signal-over-background ratio of $S/B \sim 0.13$. For the background, we have generated the dominant QCD Wjj contribution with `SHERPA` to obtain the shape. We normalize the event numbers to match the predicted number of total signal and background events in [99] rescaled by the luminosity.

4.2.2 Combined constraints

We combine the constraints on the Wilson coefficients from measurements of the WW, WZ, $W\gamma$, Zjj and Wjj channels in a χ^2 analysis. Since systematic uncertainties cancel out in our observables, we do not need to consider correlations between uncertainties of the different channels and directly calculate the χ^2 from the differential asymmetries A_{ij} via

$$\chi^2 = \sum_{i,j} \frac{(A_{ij} - 0.)^2}{\sigma^2_{A_{ij}}}\,, \qquad i = WW,\, WZ,\, W\gamma,\, Zjj,\, Wjj, \quad j = 1, \ldots, 6\,, \tag{4.2.3}$$

where $\sigma_{A_{ij}}$ denotes the combined statistical uncertainty from signal and background on the asymmetry in bin j of channel i. The number of expected events is greater than 20 for all bins so we can assume they follow a Gaussian distribution.

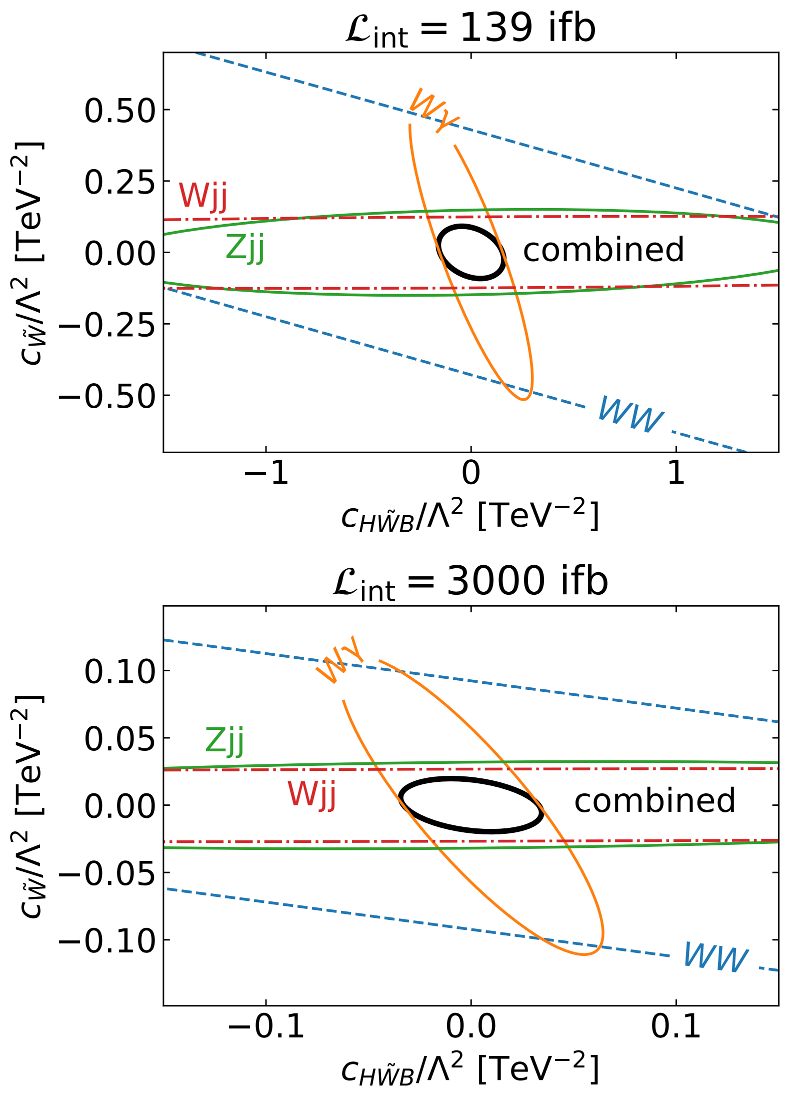


Figure 4.3: Constraints on Wilson coefficients at the 95% CL from WW, Zjj, Wjj and $W\gamma$ production for $\mathcal{L}_{\text{int}} = 139$ fb^{-1} (top) and $\mathcal{L}_{\text{int}} = 3000$ fb^{-1} (bottom). The limits from WZ production are too weak to be shown in these plots.

In fig. 4.3 we present our results for the constraints on Wilson coefficients, $c_{H\tilde{W}B}$ and $c_{\tilde{W}}$, expected for the LHC Run II at $\mathcal{L}_{\text{int}} = 139$ fb^{-1} and for the HL-LHC at $\mathcal{L}_{\text{int}} = 3000$ fb^{-1}. We show the limits coming from individual processes as well as their combination. The strongest constraint on the $c_{\tilde{W}}$ operator comes from Wjj, whereas for $c_{\tilde{W}}$ the tightest limit is from $W\gamma$ production. Our results approximately agree with those in [102]. There are slight differences due to the inclusion of detector efficiencies. Note that we have the ability to use an existing 13 TeV analysis and therefore do not have to make any assumptions for the implemented cuts. This leads to a factor 10 decrease in the cross section in our analysis compared to the reference.

We find the expected limit on the CP-odd Wilson coefficient $c_{H\tilde{W}B}/\Lambda^2 < 0.04\,\text{TeV}^{-2}$ at $\mathcal{L} = 3000$ fb^{-1}. This improves upon the expected limits of $c_{H\tilde{W}B}/\Lambda^2 < 3.1\,\text{TeV}^{-2}$ we found in chapter 3 using Higgs production in WBF with an additional photon, as well as the limit $c_{H\tilde{W}B}/\Lambda^2 < 1.5\,\text{TeV}^{-2}$ from standard Higgs processes [71].

The Wilson coefficient $c_{\tilde{W}}$ is constrained in our analysis to $c_{\tilde{W}}/\Lambda^2 < 0.02\,\text{TeV}^{-1}$ at $\mathcal{L} = 3000$ fb^{-1}. Comparable expected limits at the HL-LHC can be found with diboson processes and vector boson scattering [101], where distributions at high-p_T are used rather than a genuine CP-odd observable.

4.3 Neutral aTGCs

Neutral triple-gauge couplings do not exist in the Standard Model. Thus, their observation in an experiment would be a clear hint of physics Beyond the SM [115]. Neutral TGCs are also absent in the SMEFT up to dimension-six and are only introduced at dimension-eight. Additionally, the linear interference of the EFT with the SM is suppressed due to the allowed polarizations of the gauge bosons and the squared terms dominate. This complicates the interpretation of neutral aTGCs in the SMEFT. In this analysis we instead rely on the most general parametrization of neutral TGCs in ZZ and $Z\gamma$ production given by [110],

$$\begin{aligned}
\mathcal{L} = \mathcal{L}_{\text{SM}} + \frac{e}{M_Z^2}\Bigg[& - [f_4^\gamma\,(\partial_\mu F^{\mu\beta}) + f_4^Z\,(\partial_\mu Z^{\mu\beta})] Z_\alpha (\partial^\alpha Z_\beta) \\
& + [f_5^\gamma\,(\partial^\sigma F_{\sigma\mu}) + f_5^Z\,(\partial^\sigma Z_{\sigma\mu})] \widetilde{Z}^{\mu\beta} Z_\beta \\
& - [h_1^\gamma\,(\partial^\sigma F_{\sigma\mu}) + h_1^Z\,(\partial^\sigma Z_{\sigma\mu})] Z_\beta F^{\mu\beta} \\
& - [h_3^\gamma\,(\partial_\sigma F^{\sigma\rho}) + h_3^Z\,(\partial_\sigma Z^{\sigma\rho})] Z^\alpha \widetilde{F}_{\rho\alpha} \\
& - \frac{1}{M_Z^2}\Big\{ h_2^\gamma\,[\partial_\alpha\partial_\beta\partial^\rho F_{\rho\mu}] + h_2^Z\,[\partial_\alpha\partial_\beta(\Box + M_Z^2) Z_\mu] \Big\} Z^\alpha F^{\mu\beta} \\
& + \frac{1}{2M_Z^2}\Big\{ h_4^\gamma\,[\Box\,\partial^\sigma F^{\rho\alpha}] + h_4^Z\,[(\Box + M_Z^2)\partial^\sigma Z^{\rho\alpha}] \Big\} Z_\sigma \widetilde{F}_{\rho\alpha} \Bigg] ,
\end{aligned} \tag{4.3.1}$$

where $\widetilde{Z}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\sigma\rho}Z^{\sigma\rho}$ with $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$. Here the f_4^V, h_1^V and h_2^V coupling parameters are CP-odd and f_5^V, h_3^V and h_4^V are CP-even and in the SM $h_i = f_i = 0$ at tree level.

As we can only constrain these anomalous couplings in a regime where the squared terms dominate, the bounds on these parameters come primarily from the cross section of high-pT bins rather than from CP asymmetries [75,108,109]. This also means that we expect similar limits for both CP conserving and CP violating parameters.

For our analysis of neutral triple gauge couplings, we generate events at leading order with the SM+AGC package in `SHERPA-2.2.1` as well as with a UFO model [114,116,117]. This simulation includes the supressed contribution from the linear interference with the SM as well as the quadratic.

4.3.1 ZZ production

We study the interaction of neutral gauge bosons in ZZ production with the leptonic decay to $4l$ [103] and $2l2\nu$ [104]. The measured experimental cross sections in the fiducial regions for these processes are $\sigma_{4l} = (46.2 \pm 2.4)$ fb [103] and $\sigma_{2l2\nu} = (25.4 \pm 1.7)$ fb [104]. With these processes we place constraints on neutral TGCs using the p_T^{ll} distribution. In order to be able to compare with published data, we use the same binning for this distribution as the experimental collaborations. For the $4l$ decay mode we use the two leptons coming from the leading reconstructed Z boson.

In our event generation we include both the gg and qq initial states at leading order. However, we only consider the effect on neutral TGCs for the qq initial state which accounts for 90% of the events. The NNLO QCD and NLO EW corrections for qq are accounted for with bin-wise k-factors which we assume to be the same for both SM and BSM. We calculate the k-factor from the ratio of leading order results to the most precise `SHERPA` prediction available. For the gg contribution we include a relative k-factor of 1.67. For the $4l$ final state, we account for detector inefficiencies with a bin-by-bin factor, ϵ_{det}, which is deduced from the ratio of the expected number of events to the expected cross section. This ranges from 0.57 to 0.69. For the $2l2\nu$ process we include a flat detector efficiency $\epsilon_{\text{det}} = 0.57$. The total number of events in the i$^{\text{th}}$ bin is thus given by, $N_i^{\text{SHERPA}} = N_i^{qq} + N_i^{gg} = \epsilon_{\text{det},i}\mathcal{L}_{\text{int}}(\sigma_i^{qq,\text{NLO}} + 1.67\sigma_i^{gg})$.

We perform a χ^2 analysis on each bin in the p_T^{ll} distribution for $4l$ and $2l2\nu$ final states,

$$\chi^2 = \sum_i = \frac{(N_i^{\text{data}} - N_i^{\text{pred}})^2}{N_i^{\text{data}} + (\sigma_i^{\text{syst}})^2}, \tag{4.3.2}$$

where N_i^{data} and N_i^{pred} are the number of observed and predicted events in each bin respectively, and σ_i^{syst} is the systematic error for each bin. For both channels the dominant constraints come from the highest p_T^{ll} bin, which is $p_T^{ll} \in [555, 3000]$ GeV for $4l$ and $p_T^{ll} \in [350, 1000]$ GeV for $2l2\nu$.

We validate our analysis by checking that we can reproduce the limits on the coupling, f_i^V, at $\mathcal{L}_{\text{int}} = 36.1$ fb^{-1}, found by the experimental collaborations [103, 104]. We agree with their constraints to within 15%. This discrepancy can be explained by the use of different Monte Carlo event generators and the fact that we only have access to the global detector efficiency for the $2l2\nu$ final state, rather than the bin-by-bin factors.

Combining the $4l$ and $2l2\nu$ final states of ZZ production we find limits on the CP-odd neutral aTGCs of,

$$|f_4^\gamma| < 7.2 \times 10^{-4}, \quad |f_4^Z| < 6.1 \times 10^{-4} \tag{4.3.3}$$

at the 95% CL for $\mathcal{L}_{\text{int}} = 3000$ fb^{-1}. Here we have constrained one parameter at a time. The predicted limits at various luminosities are given in table 4.1. At higher luminosities the statistical error will decrease and the error on the final bin will be dominated by systematics. To give an idea of how the limits will be affected by the reduction in systematic uncertainties between 36.1 fb^{-1} and 3000 fb^{-1}, in table 4.1 we show the limits for $\mathcal{L}_{\text{int}}$ assuming the systematic uncertainty were to reduce by a factor two.

lumi [fb^{-1}]	$\lvert f_4^\gamma \rvert \times 10^4$	$\lvert f_4^Z \rvert \times 10^4$
139	11.	9.1
300	9.1	7.7
3000	7.2	6.1
300 ($\sigma_{\text{syst}}/2$)	8.2	7.0
3000 ($\sigma_{\text{syst}}/2$)	5.3	4.5

Table 4.1: Expected limits on nTGCs for the combination of the $ZZ \to 4\ell$ and $ZZ \to 2\ell 2\nu$ analyses at different luminosities. The limits on the parameters f_5^V which lead to CP-conserving interactions are equivalent to those on their CP-violating counterparts. In the two bottom rows, we present the limits assuming that the relative systematic uncertainties in each bin have been halved with respect to the value quoted by the experimental collaborations at 36.1 fb^{-1}.

At $\mathcal{L}_{\text{int}} = 3000$ fb^{-1}, our combined limits approximately agree with those forecast by CMS for LHC-RunII in the $4l$ final state [118]. This analysis draws most of its sensitivity from the final overflow bin $m_{ZZ} > 1300$ GeV. If we were to include an overflow bin instead of using fixed binning, our limit would tighten by roughly 20%. However, the inclusion of an overflow bin could lead to potential issues when translating the limits to other frameworks such as EFTs. Furthermore, by constraining the final bin we ensure that our events lie in a kinematic regime in which the detector is well understood.

4.3.2 Zγ production

We study $Z\gamma$ production with leptonic $2l\gamma$ [105] and $2\nu\gamma$ [106] final states. By analysing the $E_{T,\gamma}$ distribution in these channels we can place constraints on the CP-odd couplings h_1^V and h_2^V. The inclusive cross section for the $2l\gamma$ final state is measured in [105] as $\sigma_{2l\gamma} = (1065.4 \pm 23.5)$ fb. The $2\nu\gamma$ analysis vetoes additional jets and measures the cross section as $\sigma_{2\nu\gamma} = (52.4 \pm 4.8)$ fb [106]. In our analysis we assume flat detector efficiencies for the $2l\gamma$ and $2\nu\gamma$ final states of $\epsilon_{\text{det}} = 0.54$ and $\epsilon_{\text{det}} = 0.89$ respectively. As for ZZ production we include a bin-by-bin k-factor for NNLO QCD and NLO EW corrections. This is obtained by rescaling to the

lumi [fb^{-1}]	$10^4\,|h_1^\gamma|$	$10^4\,|h_1^Z|$	$10^7\,|h_2^\gamma|$	$10^7\,|h_2^Z|$
139	3.6	3.2	8.1	8.1
300	3.2	2.9	7.3	7.2
3000	2.7	2.4	6.1	6.1
300 ($\sigma_{\text{syst}}/2$)	2.7	2.4	6.1	6.1
3000 ($\sigma_{\text{syst}}/2$)	2.0	1.8	4.5	4.4

Table 4.2: Expected limits on nTGCs for the combination of the $Z\gamma \to 2\ell\gamma$ and $Z\gamma \to 2\nu\gamma$ analyses at different luminosities. In the two bottom rows, we present the limits assuming that the relative systematic uncertainties in each bin have been halved with respect to the value quoted by the experimental collaborations.

predictions in [105, 106]. We use the same binning as the experimental references for each final state, the only difference being we use a constrained final bin as opposed to an overflow.

We again use a χ^2 analysis to obtain constraints on the anomalous couplings given by eq. (4.3.2). We find the final bins of $E_{T,\gamma} \in [500, 1200]$ GeV for $2l\gamma$ and $E_{T,\gamma} \in [600, 1100]$ GeV for $2\nu\gamma$, provide the best sensitivity to neutral TGCs. We explicitly check that we can reproduce the expected limits for the $2l\gamma$ final state at $\mathcal{L}_{\text{int}} = 36.1$ fb^{-1} [106] when we include the overflow bin.

Combining both final states we obtain one-parameter limits at 95% CL for $\mathcal{L}_{\text{int}} =$ 3000 fb^{-1} of,

$$\begin{aligned} |h_1^\gamma| &< 2.7 \times 10^{-4}, \quad & |h_1^Z| &< 2.4 \times 10^{-4}, \\ |h_2^\gamma| &< 6.1 \times 10^{-7}, \quad & |h_2^Z| &< 6.1 \times 10^{-7}, \end{aligned} \tag{4.3.4}$$

for the CP-odd aTGCs. Here we have assumed the same relative systematic errors as given in the experimental references at 36.1 fb^{-1} and 139 fb^{-1}. We also present the limits for different luminosities in table 4.2, where we again additionally display the $\mathcal{L}_{\text{int}}$ limit assuming the systematic uncertainty is halved. As the limits are expected to be roughly the same for the CP-even TGCs we do not present them explicitly here.

If we include an overflow bin, the limits on h_1^V reduce by roughly 20%. However, for the h_2^V coupling they are approximately halved. This implies that great care is needed when translating limits from analyses with overflow bins into the EFT framework.

4.4 Conclusions and outlook

In this chapter we have studied constraints on anomalous triple gauge couplings. We analysed limits on charged aTGCs coming from $WW \to ll'\nu\nu$, $WZ \to l^+l^-l^\pm\nu$, $W\gamma \to l^\pm\nu\gamma$, $Zjj \to l^+l^-jj$ and $Wjj \to l\nu jj$ production. This was done in the dimension-six SMEFT framework. We studied the CP violating $c_{H\tilde{W}B}$ and $c_{\tilde{W}}$ Wilson coefficients which modify the $WW\gamma$ and WWZ interactions in diboson production. Using CP-asymmetries and marginalising over the other coefficient, we find constraints of $|c_{H\tilde{W}B}|/\Lambda^2 < 0.04$ TeV^{-2} and $|c_{\tilde{W}}|/\Lambda^2 < 0.02$ TeV^{-2} at $\mathcal{L}_{\text{int}} = 3000$ fb^{-1} . We find the strongest limits come from $W\gamma$, Wjj and Zjj production processes. The limit presented here for $c_{H\tilde{W}B}$ improves upon that found in chapter 3 and from standard Higgs production processes. The inclusion of diboson data along with Higgs and vector boson scattering would therefore be advantageous in placing combined constraints on CP violating operators.

We analysed neutral anomalous triple gauge couplings, which are absent from the Standard Model. Their observation would thus be a clear sign of new physics. We constrained neutral aTGCs in ZZ and $Z\gamma$ production using the general parameterisation given in [110]. As the squared coupling terms are dominant, we use the differential p_T^{ll} and $E_{T,\gamma}$ distributions of $ZZ \to 2l2\nu, 4l$ and $Z\gamma \to 2l\gamma, 2\nu\gamma$ respectively, rather than a CP-sensitive observable. The dominant limits come from the final bins on of these distributions, which in our analysis have a defined upper limit.

Combining the two final states of ZZ production, we find one-parameter bounds on the f_4^γ and f_4^Z couplings of $f_4^\gamma < 7.2 \times 10^{-4}$ and $f_4^Z < 6.1 \times 10^{-4}$ for $\mathcal{L}_{\text{int}} = 3000$ fb^{-1}. For $Z\gamma$, combining the $2l\gamma$ and $2\nu\gamma$ channels, we find limits on the CP-odd couplings $|h_1^\gamma| < 2.7 \times 10^{-4}$, $|h_1^Z| < 2.4 \times 10^{-4}$, $|h_2^\gamma| < 6.1 \times 10^{-7}$, and $|h_2^Z| < 6.1 \times 10^{-7}$. When an overflow bin above 1 TeV was included, the limits on h_2^V were found to tighten significantly. This should be taken into account when translating these limits to an EFT framework.

To summarise, in this chapter we have examined the limits that can be placed on anomalous triple gauge couplings with diboson production at future runs of the LHC. This is important as it will allow us to constrain additional sources of CP-violation needed to explain the baryon asymmetry observed in our universe.

Part II

Machine Learning

Chapter 5

Machine Learning in High Energy Physics

Many different machine learning methods have a vast number of uses in high-energy particle physics. In this chapter we introduce the two methods used in this thesis: neural networks and decision trees. Machine learning tasks are generally split into regression and classification, although both algorithms can perform either of these tasks. Here we will focus on using neural networks (NN) for regression and decision trees for classification.

Generally in a machine learning task we have a set of data with some features and a target variable which we want to predict. In particle physics, features of the data could be the 4-momenta of particles in a process at the LHC, or a set of experimental observables. This data is usually then randomly shuffled and split into training, validation and test sets. The validation set is necessary to check the performance of different algorithms and to tune hyperparameters. Furthermore, the use of a validation set can help prevent overfitting. This is when the model does not generalise well from the training data to unseen data. By monitoring the performance of the validation and training sets simultaneously we can determine when overfitting is occurring. The validation data is independent of the test set, which is used to evaluate the final performance of an algorithm without feeding back into its optimisation. The training data is used to update the parameters of an algorithm in order to accurately reproduce the target variable. We shall see how this is done for neural networks and decision trees in the following sections.

5.1 Neural Networks

Artificial neural networks are a popular machine learning technique. They are particularly powerful due to their ability to represent non-linear functions. There are different types of neural network architectures and learning methods. In this section we will focus on supervised learning with fully connected feedforward networks as this will be most relevant for chapter 6.

The core element of a neural network is the neuron, also referred to as the perceptron. This receives a set of input variables, $x = (x_1, ..., x_N)$, and returns a function of their weighted sum,

$$f(x) = a\left(\sum_{i=0}^{N} w_i x_i\right), \tag{5.1.1}$$

where $a(z)$ is the activation function, included to break the linearity of the node, and $x_0 = 1$ such that $w_0 \cdot x_0$ gives a constant bias term.

Most neural networks consist of many perceptrons grouped together in layers, where each layer is a function of the layer before, as shown in fig. 5.1. The input x is fed in and information flows forwards through each layer, known as forward propagation.

We want to adjust the weights such that the output of the network is as close to the original data as possible. We therefore introduce a measure to be optimised called the loss function. For regression tasks this is commonly taken as the mean squared error,

$$L_{\mathrm{MSE}} = \frac{1}{N}\sum_{i=1}^{N} \left(y_i - f(x_i)\right)^2. \tag{5.1.2}$$

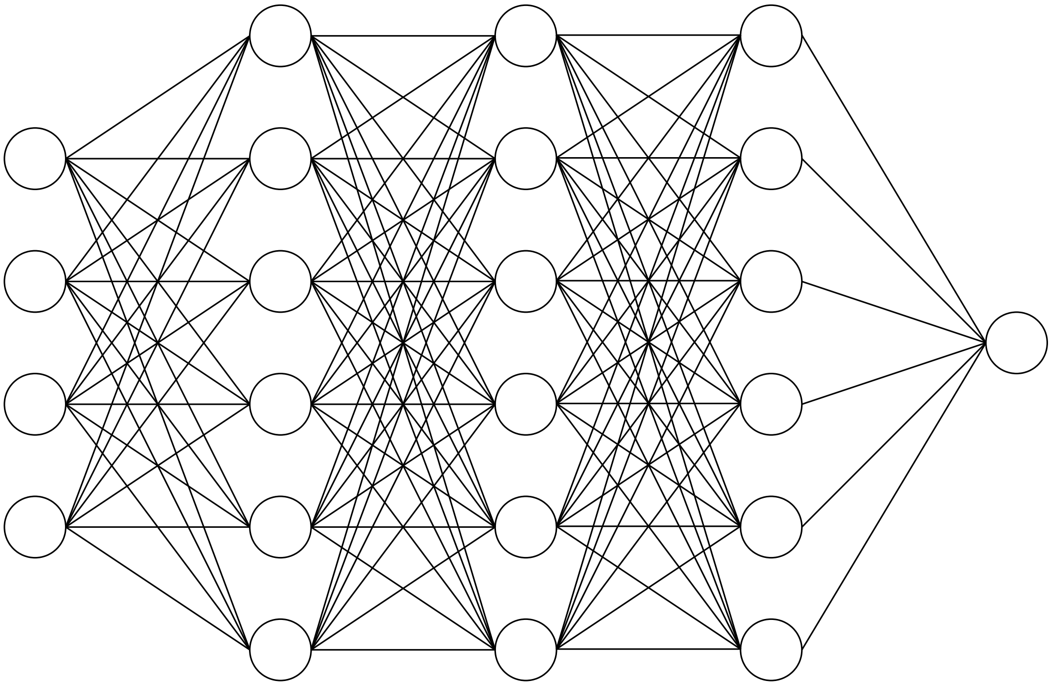


Figure 5.1: Schematic of the architecture of a densely connected neural network constructed of three hidden layers, four input nodes and one output.

In principle, the loss function is a function of all the parameters of the network. In order to evaluate its gradient with respect to the parameters, an algorithm called back propagation is used. The gradient is then used to update the weights,

$$w_{t+1} = w_t + \eta \nabla_w L(w_t), \tag{5.1.3}$$

where η is known as the learning rate.

In practice, computing the gradient exactly over the entire training set is very computationally expensive. Instead, it is common to split the training data into batches and evaluate the gradient over only one subset at a time in order to update the weights. This is done iteratively with each pass of the entire training set through the network referred to as an epoch.

In addition to the weights of the neural network, there are several other parameters of its architecture that can be tuned; the number of neurons per layer and the number of layers, the activation function a of the neurons, and the learning rate η. Additionally, the batch size and number of epochs for which to train the network can also be optimised. These are known as hyperparameters and typically need to be determined by experimentation or by iterating over combinations in a hyperparameter fit. The same activation function tends to be used across all neurons in the same layer and across the hidden layers. Although the learning rate can be tuned, having a constant learning rate over the training is disadvantageous. As the network converges to a global minima of the loss function, too large a learning-rate can cause it to overshoot. This is mitigated by the use of adaptive learning-rate algorithms that change the rate dynamically over the course of training.

In chapter 6 we investigate the ability of neural networks to learn next-to-leading order electroweak corrections in $W+$jet production. These are essential to include as they have a measurable effect in the same regions of phase space used to probe new physics effects, however they can be computationally expensive to calculate.

5.2 Decision Trees

The decision tree algorithm is a supervised learning algorithm used for classification and regression. It predicts the value of a target based on a set of learnt decision rules using the input features. One of their uses in particle physics is the separation of signal and background events.

Decision trees are made up of nodes and leaves. At each node the optimal cut on each data feature, that best splits the data into the target categories, is found. In order to quantify the purity of a data sample we can define the entropy as

$$E = -\sum_{i=1}^{N} p_i \mathrm{log}_2 p_i, \tag{5.2.1}$$

where N is the number of target categories and p_i is the probability of randomly selecting category i. The information gain is then given by the difference between the entropy of the initial node and the average entropy of the new data samples after the split. The feature which gives the maximum information gain is then used to divide the data into further nodes.

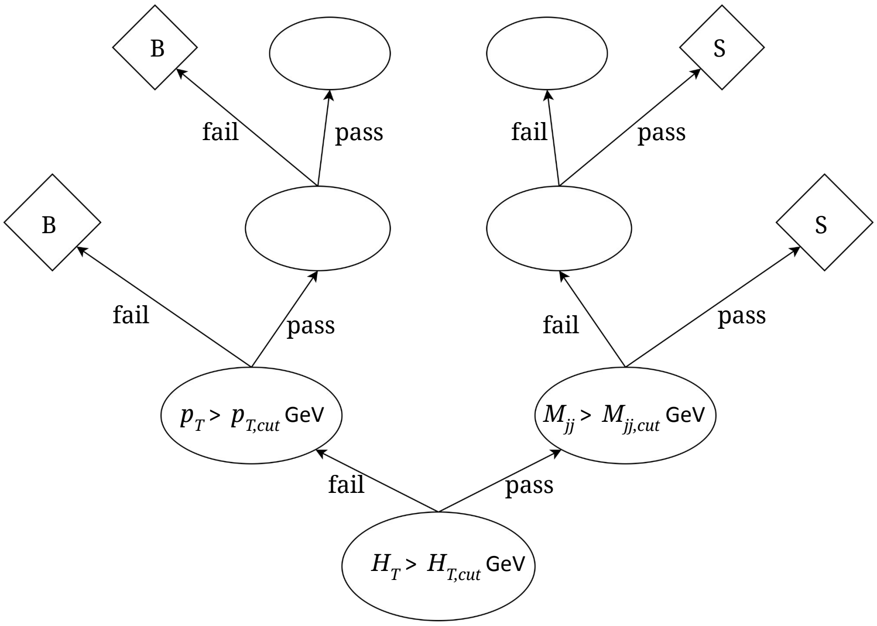


Figure 5.2: Schematic of a decision tree separating signal, S, and background, B, events using different observables as features. The ovals indicate nodes and the diamonds represent leaves.

This is done iteratively until the purity of the node exceeds a certain value and it becomes a leaf, meaning this node gives the final classification. Additionally, if the separation cannot be improved by a cut on any of the features then the node is turned into a leaf. A schematic depiction of a decision tree separating signal and background events with cuts on observables is shown in fig. 5.2.

One of the advantages of decision trees is that they are very easy to interpret. However, they are weak learners and are prone to overfitting to the training data. Their performance can be enhanced using ensemble methods. The aim of ensemble methods is to combine the predictions of several models to improve the generalizability and robustness of the prediction. There are two main groups of ensemble methods for decision trees: averaging and boosting.

5.2.1 Averaging methods

In averaging methods such as bagging and random forests [119], several decision trees are built independently and their predictions are then averaged. In bagging algorithms several decision trees are trained using randomly sampled subsets of the training data. Their individual predictions are then aggregated to form the final prediction. This reduces overtraining as no one model is trained on all of the data. There are various bagging methods that can be used which differ in the way they sample the data.

In random forests the training data used to train each tree is sampled with replacement. Additionally, another source of randomness is introduced, as at each node a random subset of features is used to determine the best split. This helps decorrelate the individual trees in the ensemble and further reduces the chance of the model overfitting.

5.2.2 Boosting methods

Boosting methods combine several trees sequentially. For each tree in the sequence the data points are weighted based on the performance of the tree before. If a subset of data points are poorly predicted by the previous model then they are given higher weighting when training the next tree. Therefore, each subsequent model that is trained is forced to concentrate on the data samples that are difficult to predict.

The boosting procedure can be optimised using gradient boosting algorithms. Similarly to neural networks, the objective is to minimise a loss function of the model, in this case by adding decision trees and using gradient descent. In this procedure the parameters of the added tree are modified so as to minimise the loss function.

For example, in chapter 4 we used a boosted decision tree classifier to separate signal and background events to enhance sensitivity to the Higgs boson coupling to b-quarks in weak boson fusion with an associated photon, and find it to clearly outperform a traditional cut-and-count analysis.

Chapter 6

Learning virtual corrections

6.1 Introduction

As we are able to probe processes at the LHC with increasing precision, it is essential for the accuracy of theoretical predictions to match experimental uncertainties. One of the key processes at the LHC is $W+$jets production. This is a background to many searches for BSM physics and therefore needs to be well constrained. Furthermore, due to its clean experimental signature and large cross section, $W+$ jets processes can be measured experimentally with a high precision. Therefore, the accuracy of theoretical simulations is of great importance.

Next-to-leading QCD corrections for $W+$ jets production have been calculated in [120–136] for up to five associated jets. Electroweak (EW) corrections at NLO for $W+1$ jet production have been calculated in [137–139]. Much work has been done to automate the evaluation of NLO EW corrections within event generators. This has been done within `Madgraph5_aMC@NLO` [140], with the loop generator `RECOLA` [141] and also with `OpenLoops` interfaced with the `SHERPA` Monte Carlo event generator for on- [142] an off-shell W bosons [143].

Electroweak corrections are particularly relevant at high energies, where they experience strong enhancement from so called Sudakov logarithms. In this region, NLO

EW corrections are the dominant uncertainty and can reach up to the order of 10%. An automated evaluation of Sudakov logarithms implemented in `SHERPA` is presented in [144], which is valid in the high energy regime.

Due to large corrections to the differential Born cross sections for large values of energy dependent observables, it is essential that NLO EW calculations are included in theoretical predictions for $W+$jets processes in order to increase the sensitivity to BSM effects. However, the computation time needed to calculate corrections can be impractical, especially as the number of final state particles increases.

In this chapter we investigate the potential of neural networks to approximate the ratio of the NLO EW correction to the Born process, referred to as the EW k-factor. The application of neural networks to speed up the evaluation of production cross sections has been shown in [145, 146]. Additionally in [147, 148] two different NN based models are presented to approximate computationally expensive matrix element calculations for the process $e^+e^- \to$ jets, for up to 5 jets. In [149] a NN model for the evaluation of matrix elements for diphoton production via gluon fusion is investigated and interfaced with the `SHERPA`.

We study W boson production with up to two associated jets. This includes the $u_i\bar{d}_i \to Wg$ partonic process with $W+1$jet and $u_i\bar{d}_i \to Wq\bar{q}$ and $u_i\bar{d}_i \to Wgg$ for W+2 jets, where $u_i = (u, c)$ and $d_i = (d, s)$. The other partonic channels are related to these through permutations of initial and final state particles.

This chapter is organised as follows. In section 6.2 we detail the event generation set up used, in section 6.3 we describe the neural network architecture and set-up specifics used to perform the inference. We analyse the performance of our model for each process in section 6.4 and summarise our conclusions in section 6.5.

6.2 Sherpa set up

We generate events for the $2 \to 2$ partonic channels; $u\bar{d} \to Wg$, $g\bar{d} \to W\bar{u}$, $g\bar{u} \to Wd$ and $u\bar{d} \to e\nu_e$. We also look at the $2 \to 3$ channels; $uu \to Wdu$, $u\bar{u} \to Wd\bar{u}$ and

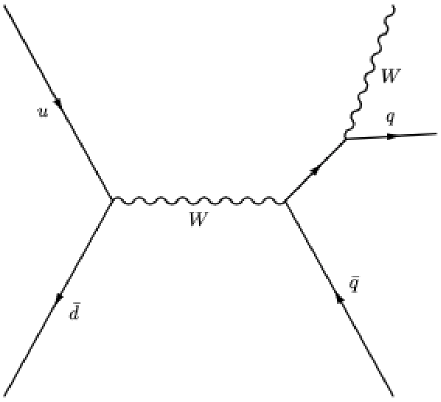


Figure 6.1: Example tree level diagram for the process $u\bar{d} \to Wq\bar{q}$ at $\mathcal{O}(e^3)$

$u\bar{d} \to Wgg$ as well as $u\bar{d} \to e\nu_e g$, $g\bar{d} \to e\nu_e\bar{u}$ and $g\bar{u} \to e\nu_e d$ and the $2 \to 4$ channels; $gg \to e\nu_e d\bar{u}$, $gu \to e\nu_e gd$, $uu \to e\nu_e du$, $u\bar{d} \to e\nu_e b\bar{b}$, $u\bar{d} \to e\nu_e u\bar{u}$, $u\bar{d} \to e\nu_e d\bar{d}$, $u\bar{d} \to e\nu_e s\bar{s}$ and $u\bar{d} \to e\nu_e gg$. We generate events for the LHC at 14 TeV with `SHERPA`. Next-to-leading order QCD and EW contributions are implemented with `OpenLoops`. For all processes we include a cut on the transverse mass of the final partons of $p_T > 20$ GeV and a cut on the separation between the final state partons of $\Delta R > 0.4$. We use the G_μ EW scheme and the m_z renormalisation scheme.

An additional cut is applied to processes which involve four quarks. An example tree level diagram for this process is shown in fig. 6.1. These diagrams contain an external quark pair coupled to a W or Z boson propagator. This leads to a resonance when the propagator momentum is equal to the square of the boson mass. Additionally, we can have a process where the boson is on-shell and has a zero width, this results in the invariant mass of the quark pair exhibiting a pole at $M^2_{W,Z}$. In order to avoid this divergence, for four quark processes we implement a cut on the invariant mass of the external quarks of $M_{qq} > 150$ GeV.

We generate 1M events for each process. This dataset is then randomly shuffled and split into a training, validation and test set with a ratio of 3:1:1. We explored the performance of the network with different dataset sizes; however, it was found that increasing the number of points beyond 1M gave little improvement compared to the large increase in generation and training time.

6.3 Neural network

Our neural network architecture consists of fully connected layers and is built using Keras with a TensorFlow backend. For each process we input the 4-momenta of the outgoing particles as well as the energy and z-component of the momenta of the initial partons, thus the number of nodes of the input layer is $4n_{\mathrm{FS}}+4$, where n_{FS} is the number of outgoing particles. The only exception is for the process $u\bar{d} \to e\nu_e$. The NLO EW correction to this process has a more straightforward dependence on the s and t Mandelstam invariants. Therefore we directly input $\log(s)$ and $\log(t)$ to the network with two input nodes. For each process we scale each input variable with standard scaling,

$$z = \frac{x - \bar{x}}{\sigma}, \tag{6.3.1}$$

where, x is a variable with mean, $\bar{x}$, and standard deviation, σ in the training set. The validation and test set variables are then scaled with the same mean and standard deviation.

The number of nodes of each hidden layer is determined by scanning over combinations in a hyperparameter fit and is tuned for each process. For all processes, the weights for each hidden layer are initialised with the Glorot uniform distribution. We use the ReLU activation function,

$$a(x) = \max(0, x) \tag{6.3.2}$$

for each hidden layer and a linear activation function for the output layer.

The network is optimised using the Adam optimisation algorithm and for our loss function we use the mean squared error (MSE),

$$L_{\mathrm{MSE}} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - f(x_i)\right)^2, \tag{6.3.3}$$

where N is the total number of training points, y_i is the target variable, x_i is the vector of input features and $f(x_i)$ is the neural network prediction.

We train for a maximum of 1000 epochs. However, this is shortened in most cases by using early stopping where we monitor the validation loss with a patience of 50 epochs. Early stopping also helps prevent the model from overfitting. We additionally reduce the learning rate by a factor of 0.5 when there is no improvement in the validation loss over 20 epochs. This helps reduce the training time of the model. Furthermore, reducing the learning rate when the improvement in the mean squared error stagnates enables the model to reach a new minimum that may have been unobtainable with a larger learning rate.

The neural network model is initialised with random weights. This means that the model will be slightly different and therefore give different predictions each time it is trained. Additionally, we train using mini-batches which are randomly sampled from the full training dataset. It is unlikely that each time the training is run the model will converge to the same minimum, we therefore run a ensemble of ten models and aggregate their predictions. Although this is more computationally expensive, with an ensemble of models we can gain a more robust evaluation of model performance and a measure of the model uncertainty in the prediction. We take the ensemble model target prediction to be the mean of the individual predictions and the uncertainty to be the standard error in the mean, i.e $\sigma_y = \sigma/\sqrt{N}$, where σ is the standard deviation and N is the number of ensembles.

6.4 Results

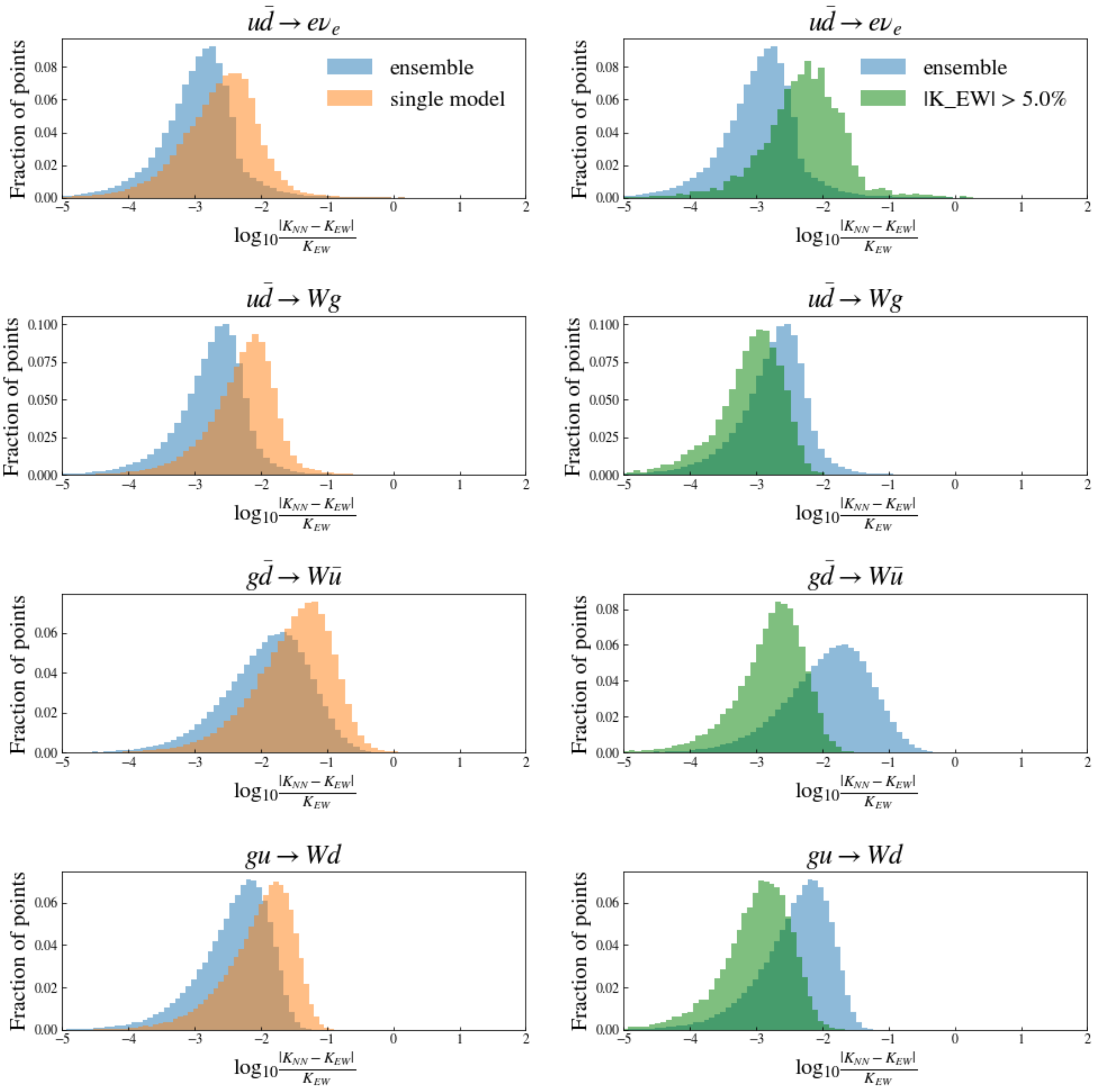
In figs. 6.2 to 6.5 we show the distribution of the absolute percentage error difference between the NN prediction and `OpenLoops` for 2 → 4, 2 → 3 and 2 → 4 processes. In fig. 6.2a we compare the performance between an ensemble of ten models and a single model from the ensemble for 2 → 2 processes. For each process the percentage error distribution is improved with an ensemble of models. In table 6.1 we show the

Process	**Single Model**		**Ensemble**		$\boldsymbol{\lvert K_{EW}\rvert > 5\%}$	
	MAPE	**95% limit**	**MAPE**	**95% limit**	**MAPE**	**95% limit**
$u\bar{d} \to e\nu_e$	0.0057	0.014	0.0028	0.0057	0.012	0.026
$u\bar{d} \to Wg$	0.014	0.025	0.0055	0.0093	0.0013	0.0037
$g\bar{d} \to W\bar{u}$	0.073	0.19	0.033	0.10	0.0028	0.0082
$gu \to Wd$	0.016	0.041	0.0069	0.02	0.0018	0.0054
$g\bar{d} \to e\nu_e\bar{u}$	0.74	1.3	0.57	0.89	0.04	0.12
$gu \to e\nu_e d$	2.2	2.5	1.60	2.1	0.045	0.15
$u\bar{d} \to Wgg$	0.041	0.081	0.025	0.049	0.007	0.021
$uu \to Wdu$	0.69	0.84	0.37	0.51	0.034	0.097
$u\bar{u} \to Wd\bar{u}$	0.33	0.48	0.14	0.22	0.023	0.071
$u\bar{d} \to e\nu_e b\bar{b}$	14	25	10	19	1.2	3.2
$u\bar{d} \to e\nu_e d\bar{d}$	3.1	3.5	2.5	2.5	0.22	0.74
$u\bar{d} \to e\nu_e s\bar{s}$	2.3	3.7	2.0	3.2	0.21	0.7
$u\bar{d} \to e\nu_e u\bar{u}$	6.1	4.2	2.9	3.1	0.26	0.79
$gg \to e\nu_e d\bar{u}$	1.9	3.2	1.6	2.7	0.1	0.3
$gu \to e\nu_e gd$	2.4	2.7	2.4	2.5	0.095	0.29
$u\bar{d} \to e\nu_e gg$	1.2	1.2	0.96	1.1	0.058	0.17
$uu \to e\nu_e du$	2.5	3.2	1.8	2.5	0.19	0.65

Table 6.1: Errors in the neural network prediction for each process. The mean absolute percentage error (MAPE) is given as well as the percentage error which 95% of the data points are equal to or below. Results are shown for an ensemble of 10 models as well as a single model from the ensemble. The final column shows the result for the region where $|K_{EW}| > 0.05$.

mean absolute percentage error as well as the absolute percentage error which 95% of the test set were equal to or below. This gives a measure of a more conservative error one can expect on the prediction. For the simplest process of $u\bar{d} \to e\nu_e$, 95% of the phase space points are predicted with a precision below the percent level with an ensemble of models. We achieve similar results for $u\bar{d} \to Wg$ and $gu \to Wd$. The results for $g\bar{u} \to W\bar{u}$ are slightly worse where 95% of points have a percentage difference of $\sim$ 10% or below.

In fig. 6.3a we show the results for processes with three outgoing particles. We can see the results are generally worse than for the 2 → 2 processes. Additionally, the network is worse at predicting the k-factor for the off-shell processes compared to the on-shell. As the performance is already poor, the improvement upon using an ensemble of model for these processes is minimal. Finally, in figs. 6.4a and 6.5a we show the results for 2 → 4 processes. We can see that the network has struggled to

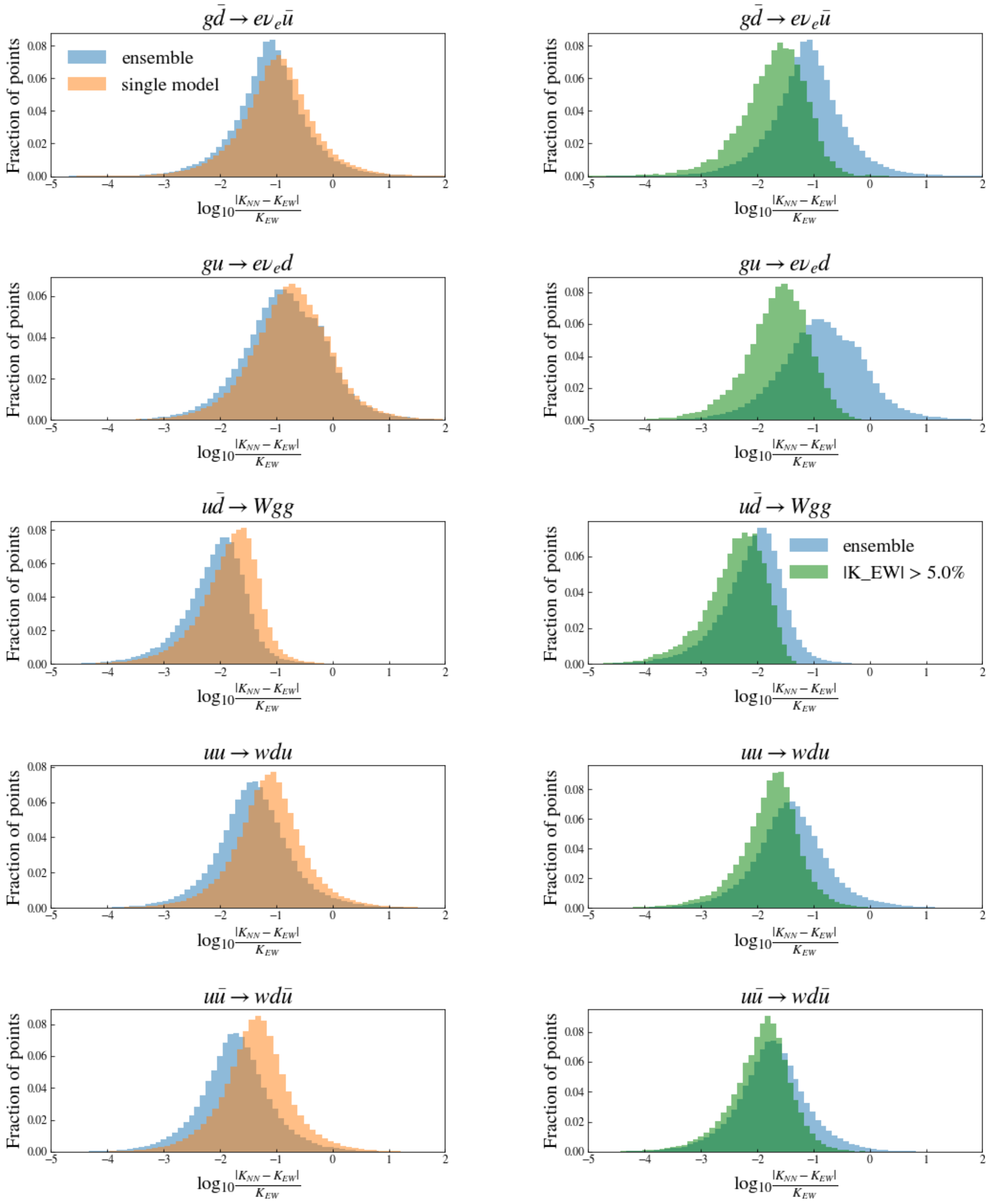


(a) (b)

Figure 6.2: Distributions of the absolute percentage difference in the NN prediction and the EW k-factor from `OpenLoops`.

learn the k-factor for these processes, especially for those involving four quarks.

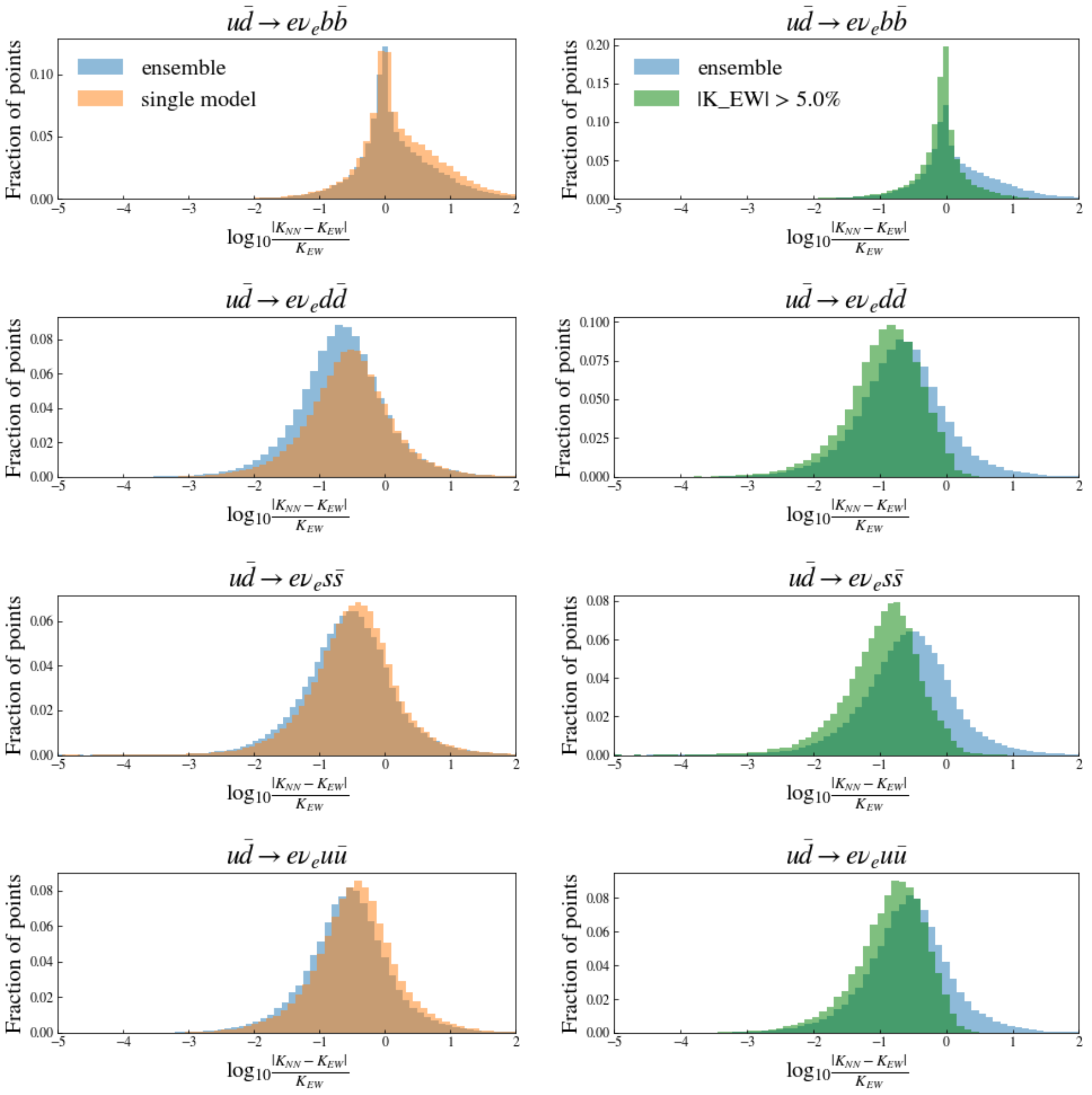
The absolute percentage error is generally a good metric to judge the performance of the neural network as it is straightforward and interpretable. However, our data contains predictions which are very close to zero, where this metric is undefined. This could result in very large percentage errors on very small corrections which skew the interpretation of the performance. (The variation of the percentage error with the k-factor is shown in appendix A.) In figs. 6.2b, 6.3b, 6.4b and 6.5b, we show



(a) (b)

Figure 6.3: Distributions of the absolute percentage difference in the NN prediction and the EW k-factor from `OpenLoops`.

the error distributions for points which have an absolute correction of greater than 5%. We can see the performance of the neural network is generally better in this region. For BSM searches at the TeV scale, the accuracy of theoretical predictions for W + jet production is particularly important at large transverse momentum. Therefore, as the virtual electroweak correction also grows with p_T, it is important

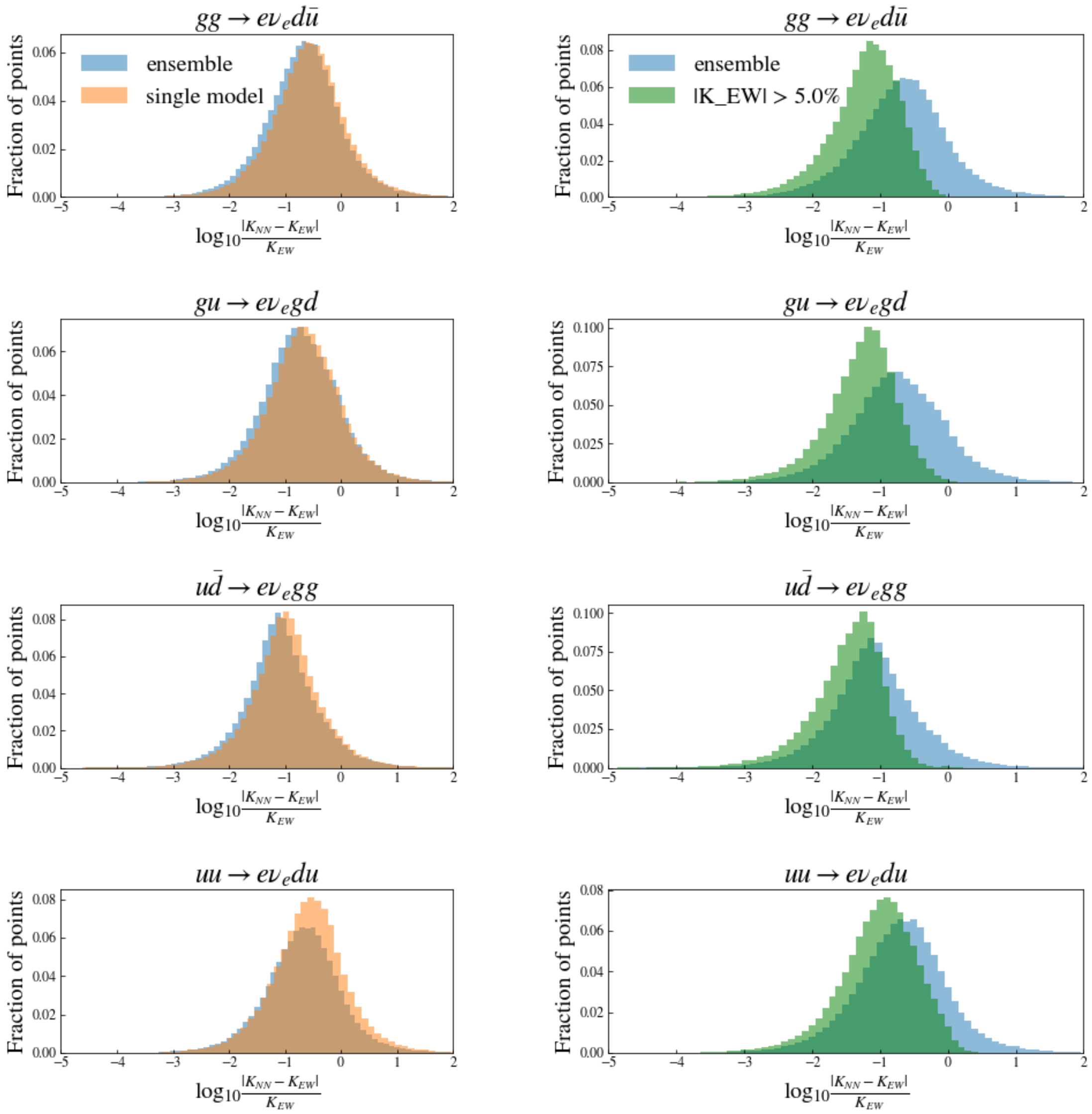


(a) (b)

Figure 6.4: Distributions of the absolute percentage difference in the NN prediction and the EW k-factor from `OpenLoops`.

to accurately predict large k-factors. In the final column of table 6.1 we show the absolute percentage error which 95% of points with $|K_{EW}| > 5\%$ were equal to or below.

In assessing the performance of our neural network it is also useful to evaluate the absolute error between the neural network prediction and the `OpenLoops` k-factor. This gives an idea of the magnitude of the errors in the prediction. In table 6.2 we show the fraction of points which have an error of 0.01 and 0.001 or less. For



(a) (b)

Figure 6.5: Distributions of the absolute percentage difference in the NN prediction and the EW k-factor from `OpenLoops`.

$2 \to 3$ processes we can predict almost all points to within $K_{EW} \pm 0.01$, and for $2 \to 2$ processes we can predict almost all points to within $K_{EW} \pm 0.001$. Therefore, when the percentage correction from virtual NLO electroweak contributions needs to be constrained to the 1% level, a neural network could provide an alternative to calculating the full correction for $2 \to 2$ and $2 \to 3$ processes. However, for cases which require greater precision or higher multiplicity, further work is needed.

Process	**% of pts with error < 0.01**	**% of pts with error < 0.001**
$u\bar{d} \to e\nu_e$	99.9	98.6
$u\bar{d} \to Wg$	100	100
$g\bar{d} \to W\bar{u}$	100	99.5
$gu \to Wd$	100	99.8
$g\bar{d} \to e\nu_e\bar{u}$	99.0	53.1
$gu \to e\nu_e d$	97.7	41.7
$u\bar{d} \to Wgg$	100	92.1
$uu \to Wdu$	97.9	59.6
$u\bar{u} \to Wd\bar{u}$	99.4	78.6
$u\bar{d} \to e\nu_e b\bar{b}$	11.0	1.1
$u\bar{d} \to e\nu_e d\bar{d}$	69.1	10.6
$u\bar{d} \to e\nu_e s\bar{s}$	62.1	8.0
$u\bar{d} \to e\nu_e u\bar{u}$	63.5	8.7
$gg \to e\nu_e d\bar{u}$	85.0	14.1
$gu \to e\nu_e gd$	87.8	17.2
$u\bar{d} \to e\nu_e gg$	97.3	39.9
$uu \to e\nu_e du$	63.1	89.2

Table 6.2: Percentage of points for which the absolute error in the prediction is less than 0.01 (middle column) and 0.001 (final column) for each process.

6.5 Conclusions and outlook

In this chapter we investigated the ability of a neural network to learn NLO EW virtual corrections in W boson production. In general we found aggregating the results from an ensemble of models improved the accuracy of the prediction. Furthermore, we found the model to perform better in regions where the k-factor is larger. The inclusion of EW NLO contributions is important in order to correctly interpret constraints on BSM models which are generally constrained in the high energy bins of distributions. The NLO correction is large and negative in this region due to the dominance of Sudakov logarithms which depend on the s and t invariants. Therefore, the imbalance in the performance of the network towards large k-factors is not necessarily a disadvantage. If we focus on the region where the k-factor is greater than 5%, for processes with two final partons we found 95% of the points have a percentage error of order 0.1% or below. For the processes with three final partons we found in general 95% of points have an error of order 10% or below. For the study of processes involving four final partons with four quarks more work is needed. Simple densely connected network architectures fail to learn these more complex processes, even with the removal of the resonance.

In the work presented here we have focused on the per point error in the output of the network. In further work it would be interesting to look at the performance in distributions of observables. Additionally, the interfacing of our model with a Monte Carlo event generator would enable an assessment of the computational time saved by using a neural network for inference.

Chapter 7

Conclusions

In this thesis we looked at constraining new physics at the LHC at current and future luminosities. We additionally presented an initial exploration of an application of machine learning to calculate NLO electroweak corrections necessary for the accurate calculation of background processes to BSM searches.

In chapter 1 we introduced the Standard Model of particle physics for which we are searching for an extension to. We also gave an overview of how new particles are measured at colliders and the techniques used in particle physics phenomenology. As large deviations from the SM may occur at energies beyond the kinematic reach of the LHC, it is useful to look for more subtle hints of BSM physics in LHC data in a model independent way.

In chapter 2 we introduced the Standard Model Effective Field Theory and discussed how the effects of physics at a higher energy scale can manifest at lower energies. We gave an overview of constraints on SMEFT operators placed by current and past experiments and how we can place constraints on operators which violate CP symmetry. This is particularly important as sources of CP violation, in addition to those introduced in the SM, are necessary to explain, for example, electroweak baryogenesis.

We examined constraints that can be placed on CP violating operators in WBF Higgs boson production with an associated photon in chapter 3. In particular we focused on $\mathcal{O}_{H\tilde{W}}$ and $\mathcal{O}_{H\tilde{W}B}$ operators in the Warsaw basis. We constructed a CP-sensitive observable from the scalar and triple products of the momentum of final state particles. By counting the normalised difference in the number of positive and negative events in the CP-sensitive variable, we calculated an asymmetry which was used to extract limits on the $c_{H\tilde{W}}$ and $c_{H\tilde{W}B}$ Wilson coefficients. We forecast limits that could be achieved at the HL-LHC full luminosity of 3 ab^{-1}. We found this limit to be comparable to that achieved with WBF without the additional photon. This motivates the inclusion of our signal channel in combination with WBF and other channels in global fits.

We continued to look at constraining CP violating operators in chapter 4, where we focused on constraints achieved through the study of triple gauge couplings in diboson production. We studied the $c_{H\tilde{W}B}$ and $c_{\tilde{W}}$ Wilson coefficients which modify $WW\gamma$ and WWZ charged triple gauge couplings. We used asymmetries to extract constraints on these coefficients in $W\gamma$, Zjj, Wjj and WZ production. We found the strongest constraints came from $W\gamma$, Wjj and Zjj channels. Neutral triple gauge couplings are absent from the SM, thus their observation would be a clear sign of new physics. nTGCs are introduced at dimension-eight in the SMEFT and moreover, the linear interference with the SM is suppressed due to allowed polarizations of the gauge bosons. We therefore used a more general parametrization of anomalous gauge couplings to investigate CP-odd BSM effects in ZZ and $Z\gamma$ production. As the squared aTGCs are dominant, we used the high energy bins of observables to place constraints on these couplings, rather than using asymmetries.

Machine learning techniques have a wide range of uses in particle physics. In chapter 5 we introduced neural networks and boosted decision trees which are commonly used in the discrimination of signal and background events.

Finally in chapter 6 we investigated using a neural network to predict virtual electroweak NLO k-factors in $W+$ jets production. This was motivated by the large computational expense of the full calculation, particularly for larger jet multiplicities. We found we could predict the k-factor for processes with two final partons within 1% error and k-factor for processes with three final state partons to around 10% error. However, our fairly straightforward network struggled to learn k-factors for processes with four final partons. Nevertheless, this chapter provides an initial exploration of the ability of neural networks to learn higher order corrections.

Appendix A

Additional information for "Learning virtual corrections"

This appendix contains supplementary tables and plots for chapter 6.

Process	**Layer Architecture**	**RMSE**
$u\bar{d} \to e\nu_e$	20-50-50-20	$1.80\ \times10^{-3}$
$u\bar{d} \to Wg$	200-100-50-100	7.00×10^{-5}
$g\bar{d} \to W\bar{u}$	50-200-100-50	2.81×10^{-4}
$gu \to Wd$	200-200-200-10	1.62×10^{-4}
$g\bar{d} \to e\nu_e\bar{u}$	200-200-100-100	3.39×10^{-3}
$gu \to e\nu_e d$	100-50-50-50	3.90×10^{-3}
$u\bar{d} \to Wgg$	200-200-50-200	5.77×10^{-4}
$uu \to Wdu$	200-200-200-100	7.16×10^{-3}
$u\bar{u} \to Wd\bar{u}$	200-200-200-200	2.75×10^{-3}
$u\bar{d} \to e\nu_e b\bar{b}$	50-50-50-50	7.80×10^{-1}
$u\bar{d} \to e\nu_e d\bar{d}$	200-200-200-100	2.11×10^{-2}
$u\bar{d} \to e\nu_e s\bar{s}$	200-200-200-200	3.24×10^{-2}
$u\bar{d} \to e\nu_e u\bar{u}$	200-200-200-100	3.20×10^{-2}
$gg \to e\nu_e d\bar{u}$	200-100-50-50	9.38×10^{-3}
$gu \to e\nu_e gd$	100-50-50-50	8.44×10^{-3}
$u\bar{d} \to e\nu_e gg$	100-200-100-100	4.77×10^{-3}
$uu \to e\nu_e du$	200-200-200-50	2.78×10^{-2}

Table A.1: Layer architecture used in the neural network for each process. We also show the root mean squared error (RMSE) achieved with this network.

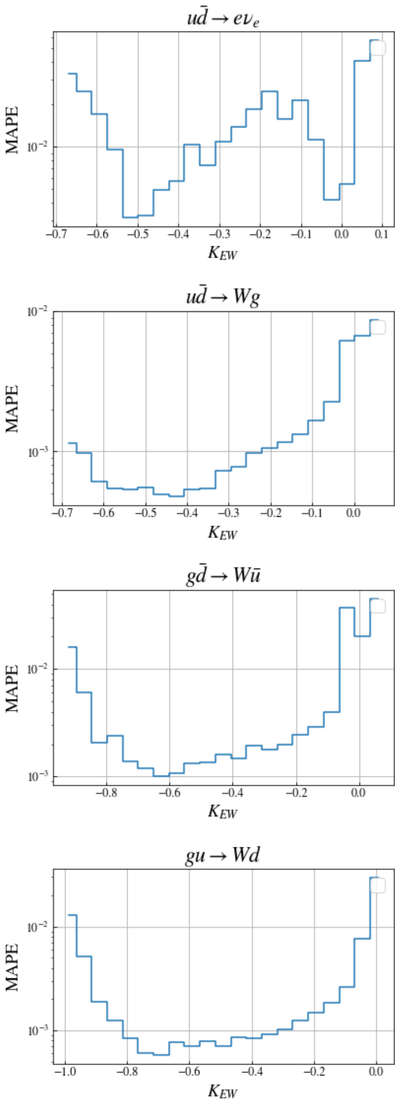


Figure A.1: Variation in the absolute percentage error (MAPE) with the true EW k-factor, K_{EW}, for 2→2 processes. Here we have binned the k-factor and show the MAPE calculated across the points in each bin.

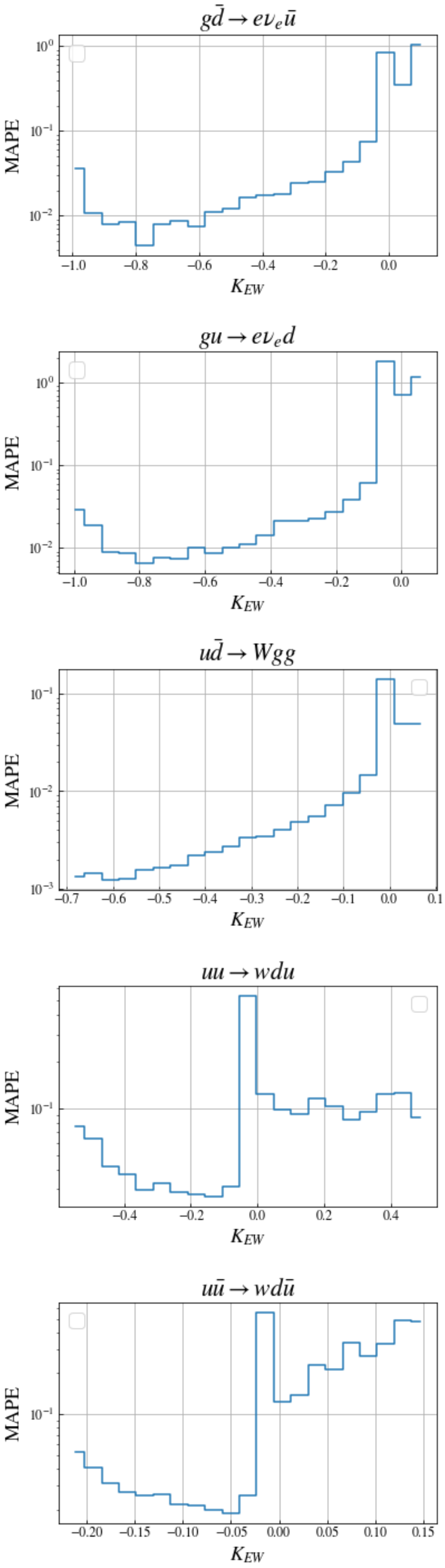


Figure A.2: Variation in the absolute percentage error (MAPE) with the true EW k-factor, K_{EW}, for 2→3 processes. Here we have binned the k-factor and show the MAPE calculated with the points in each bin.

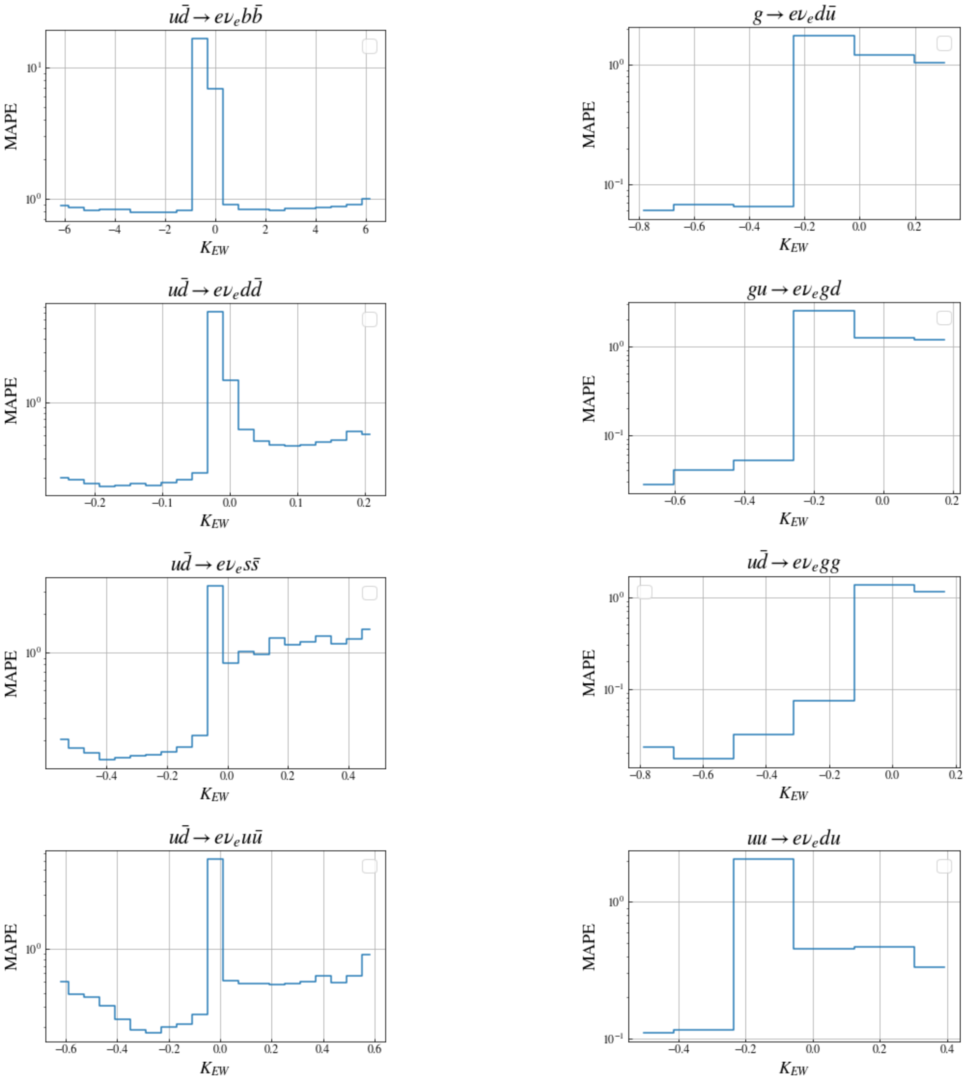


Figure A.3: Variation in the absolute percentage error (MAPE) with the true EW k-factor, K_{EW}, for 2→4 processes. Here we have binned the k-factor and show the MAPE calculated with the points in each bin.

Appendix B

Industrial placement: Monitoring aircraft turnaround with AI

In this appendix we present a report on an industrial placement undertaken with the Durham University Centre for Doctoral Training in Data-Intensive Science. The project was undertaken with Boeing at their Digital Solutions and Analytics lab in Frankfurt, Germany from 1^{st} April to 1^{st} July 2019.

B.1 Introduction and motivation

Computer vision has a wide range of uses from driverless vehicles to medical diagnostics. In this project we explored the use of computer vision in airports, specifically in the monitoring of aircraft turnarounds. The turnaround stage of a commercial aircraft is the time from when it lands at an airport to its next take-off. During this period the following procedures can take place:

- Park aircraft
- Connect/disconnect ground power
- Connect/disconnect bridge/stairs

- De-board/board passengers
- Unload/load cargo
- Refuel aircraft
- Load catering
- Pushback

These processes all happen to a tight schedule which is often specific to a certain aircraft or airport. If one operation is delayed this can have a knock-on effect thus delaying the take-off time. With a huge number of flights happening each day, it is essential to have an accurate tracking of the turnaround phase so delays can be flagged immediately and resolved or rearrangements put in place.

We investigate using computer vision to monitor the turnaround progress of an aircraft. With a camera positioned on the terminal or tail-fin of a plane we can view the activity around an aircraft. The various vehicles can then be identified with an object detection model, providing information as to which turnaround procedures are currently in progress.

Additionally we look at the possibility of combining visual information with sensor data directly from the aircraft and look at how this information is related to individual turnaround procedures.

The following report is structured as follows. In appendix B.2 we introduce the structures of neural networks used for object detection and classification. In appendix B.3 we show how visual and sensor data can help track the aircraft turnaround , and we give our conclusions in appendix B.4.

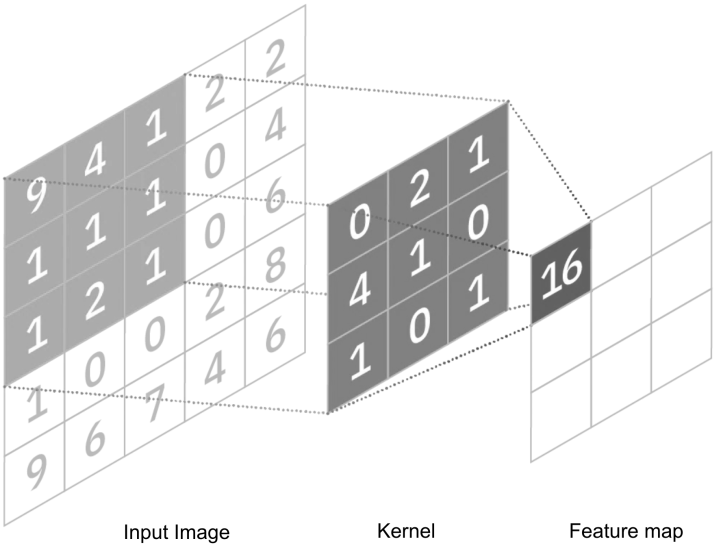


Figure B.1: Example of the convolution operation for a 3 × 3 kernel. This figure is taken from [151].

B.2 Convolutional neural networks and computer vision

B.2.1 Convolutional neural networks

Convolutional neural networks (CNNs) [150] are a specialised type of neural network that use convolution instead of general matrix multiplication in at least one of their layers [151]. Applied to an image, the convolution layer of a CNN uses an array kernel which passes over the image input tensor. At each location the sum of the element wise product of the kernel and the input tensor is calculated to produce a feature map. This is illustrated in fig. B.1.

In contrast to fully connected layers, which have a separate parameter for the interaction of every output with each input, convolutional layers have sparse connections when the kernel is smaller than the output size, meaning each output is only affected by a subset of inputs. This is illustrated in fig. B.2. This makes convolutional networks particularly useful for processing images. One image may contain thousands of pixels meaning a fully connected network could take an extremely long time to train. However, using the convolution operation fewer parameters need to be stored,

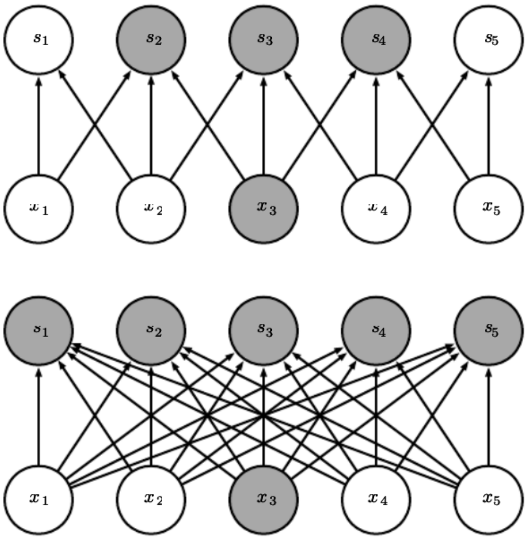


Figure B.2: Connections between input and output units for convolution with a kernel of width 3 (top) and matrix multiplication (bottom). This figure is taken from [151].

improving the efficiency of the model.

A convolutional "layer" can sometimes refer to three stages (or layers) which are often grouped together. In the first stage the convolution operation is performed. The outputs from this are then passed through a non-linear activation function, referred to as the detector stage. In the final stage an operation called pooling is carried out. Pooling takes an output of detector stage at a certain location and replaces it with a summary statistic calculated over neighbouring outputs. The purpose of this is to help make the network invariant to small translations in the input. Furthermore, by sampling over regions spaced more than one unit apart, pooling enables a decrease in the number of inputs to the next layer. This improves the computational efficiency of the network and is also useful for handling inputs of varying sizes. For example, if you wish to classify an image, the input to the classification layer must be of a fixed size. By varying the size of the offset between pooling regions it is possible to always have the same number of outputs from the pooling layer ready to be fed into the classification layer.

These stages are the building blocks of many variations of convolutional neural networks, such as those used for object detection.

B.2.2 Object detection

In object detection tasks a neural network outputs a bounding box region containing an object in the training set as well classification probabilities for each object label. Neural network architectures for object detection typically fall into two categories: one-stage and multi-stage detectors.

Multi-stage detectors

Multi-stage detectors use a region proposal algorithm to first generate regions of interest. These are then fed into a pipeline for object classification and bounding box regression. A popular multi-stage detector is R-CNN [152] (Regional- Convolutional Neural Network). R-CNN takes an input image and firstly extracts region proposals using the Selective Search [153] algorithm. This groups similar regions together over a range of scales to generate boxes that contain an entire object.

The next stage is feature extraction. In this stage each region of interest is passed through a CNN to produce a fixed length feature vector. These are then fed to a set of class specific Support Vector Machines (SVMs) which output the final classification probabilities for each class. Additionally, a bounding box regression stage is included to improve the localisation of the object region.

R-CNN has very good accuracy however it is too slow to use for real time inference. Fast R-CNN [154] is a variation of R-CNN which improves upon its speed as well as accuracy. It processes the entire image with a CNN. Then, for each region proposal, a pooling layer extracts a fixed length feature vector from the output of the CNN. By avoiding passing each region of interest through a separate CNN, this method is much quicker. Faster-RCNN [155] improves the speed and accuracy further through the use of a Region Proposal Network (RPN) instead of Selective Search.

These multi-stage detectors are typically have higher accuracy than one-stage detectors. However, even Faster-RCNN does not match their speed.

One-stage detectors

One-stage detectors use a single network to learn both the bounding boxes and class probabilities. A popular one-stage detector is Single Shot Multibox Detector (SSD) [156]. SSD doesn't use bounding box proposals. Instead, it evaluates a set of default boxes with different sizes at each location. For each box, the offset from the ground truth box and the classification probability for all classes is computed. It uses a loss which is a weighted sum between box localisation loss and a classification loss.

Another widely used one-stage detector is You Only Look Once (YOLO) [157]. YOLO also predicts bounding boxes and class probabilities simultaneously. It divides an image into an N × N grid. A number of bounding boxes is predicted for each cell of this grid along with a confidence score for how likely it is that the box contains an object and how accurate the box is around the object. Conditional class probabilities are also predicted for each cell. This is the probability that the cell contains an object of a certain class given the cell contains an object. The final scores are then given by the product of the conditional class probabilities with the individual box confidence predictions.

SSD and YOLO are much faster than multi-stage detectors as they eliminate the need for region proposals. SSD can evaluate images at $\sim$59 fps and YOLO at $\sim$49fps [156], thus these one-stage detectors are very useful for performing real time detection.

B.3 Monitoring the turnaround

B.3.1 Object detection model

We use pre-trained models from the TensorFlow Object Detection API [158] which have been trained on the COCO dataset [159]. This dataset contains the object

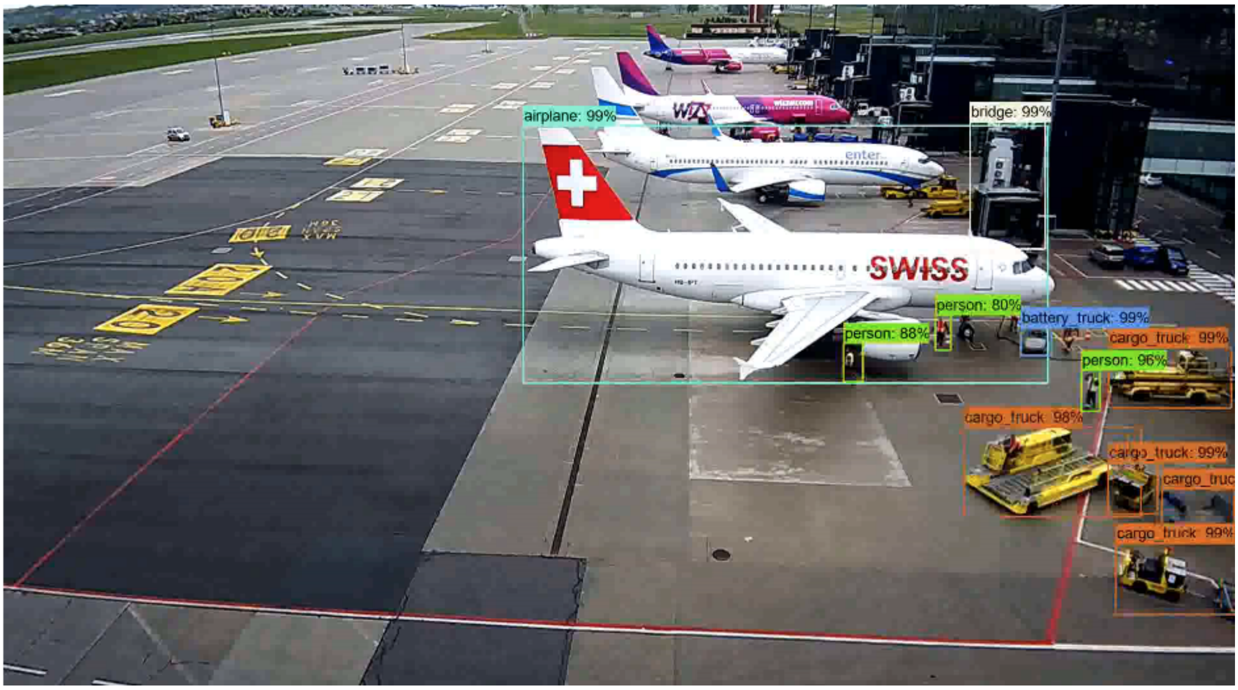

airplane: 99%
bridge: 99%
SWISS
enter
person: 80%
battery_truck: 99%
person: 88%
cargo_truck: 99%
person: 96%
cargo_truck: 98%
cargo_truck: 99%
cargo_truc
cargo_truck: 99%

Figure B.3: Example output of our object detection model. Bounding boxes are displayed on the image as well as the class probability.

classes of person, car, aeroplane bus and truck which are relevant for our use case. However, we also need to be able to differentiate between different types of trucks used at airports as well as other specialised objects such as the passenger bridge. We therefore further fine-tune the network with additional layers to classify specific objects relevant for our application. We use the following 16 classes: person, car, aeroplane, bus, bridge, cargo door, crew door, stair truck, catering truck, fuel truck, special assistance truck, cargo truck, pushback truck, battery truck, de-icing truck, other truck.

For training we use ∼2000 images as well as ∼25 hrs of video footage. We experimented with using Faster R-CNN, SSD and YOLO. We find that Faster R-CNN gives the best performance and is good at detecting smaller objects such as the cargo and crew doors. However, it is slow to train and infer. SSD is much faster to train and matches the performance of Faster R-CNN for larger objects but is worse at predicting smaller objects. YOLO is the fastest to train and infer however is less accurate that the other two models.

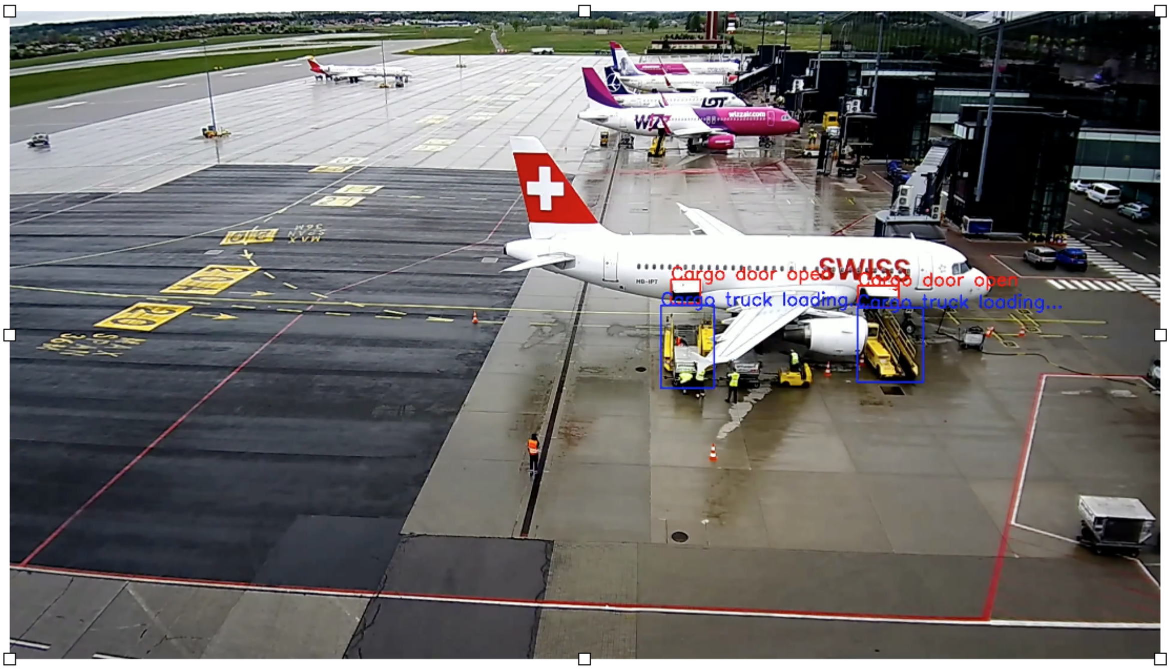


Figure B.4: Example output of our object detection model showing the cargo loading stage. The proximity of the cargo truck to an open cargo door flags it as loading.

In fig. B.3 we show an example output from the network. This shows its ability to classify the different types of truck in the image as well as the passenger bridge and individual people.

B.3.2 Turning data into knowledge

The object detection model provides information about which objects are present in a frame, however this does not necessarily on its own tell us much about which turnaround procedures are currently happening. For example, there are several cargo trucks detected in fig. B.3 although its clear the aircraft is not currently being loaded.

We therefore implement algorithms based on relative spatial information provided by the bounding box coordinates to help determine if a stage of the turnaround is occurring. Examples of this are shown in figs. B.4 and B.5 for cargo loading and pushback operations respectively. In the pushback operation a pushback truck connects to the aircraft and helps it reverse out of the bay. We can see that each

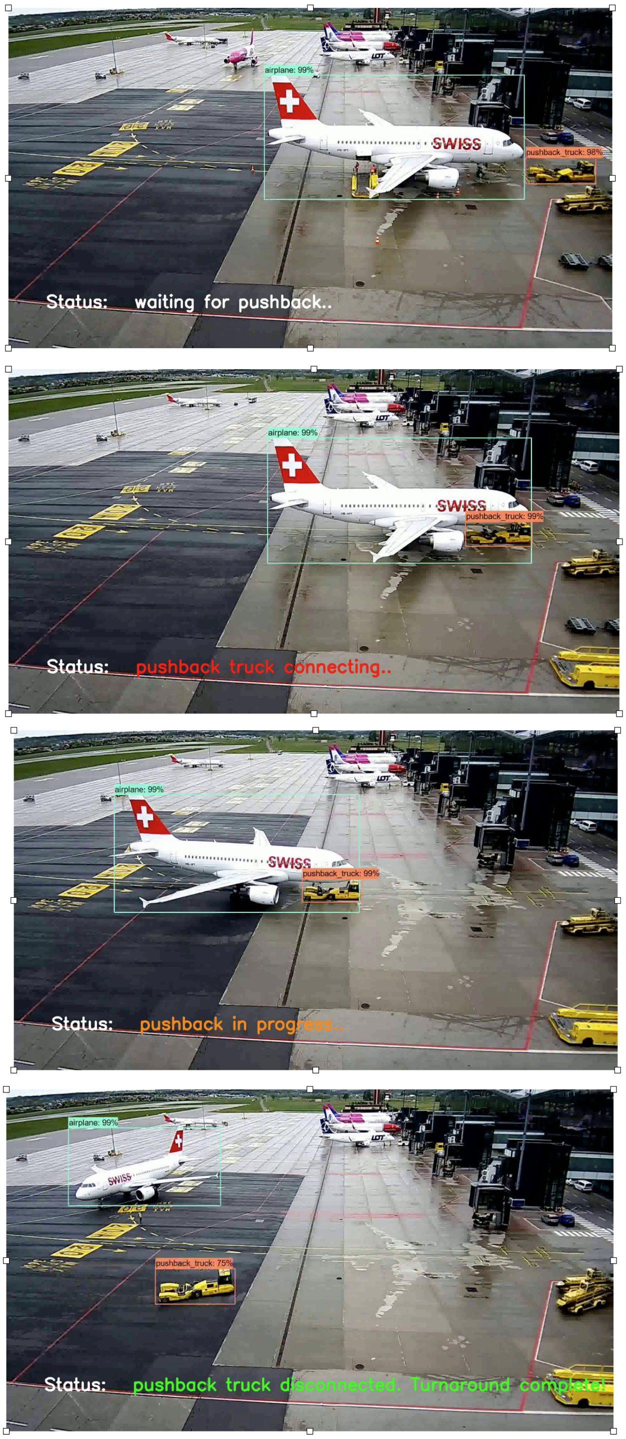
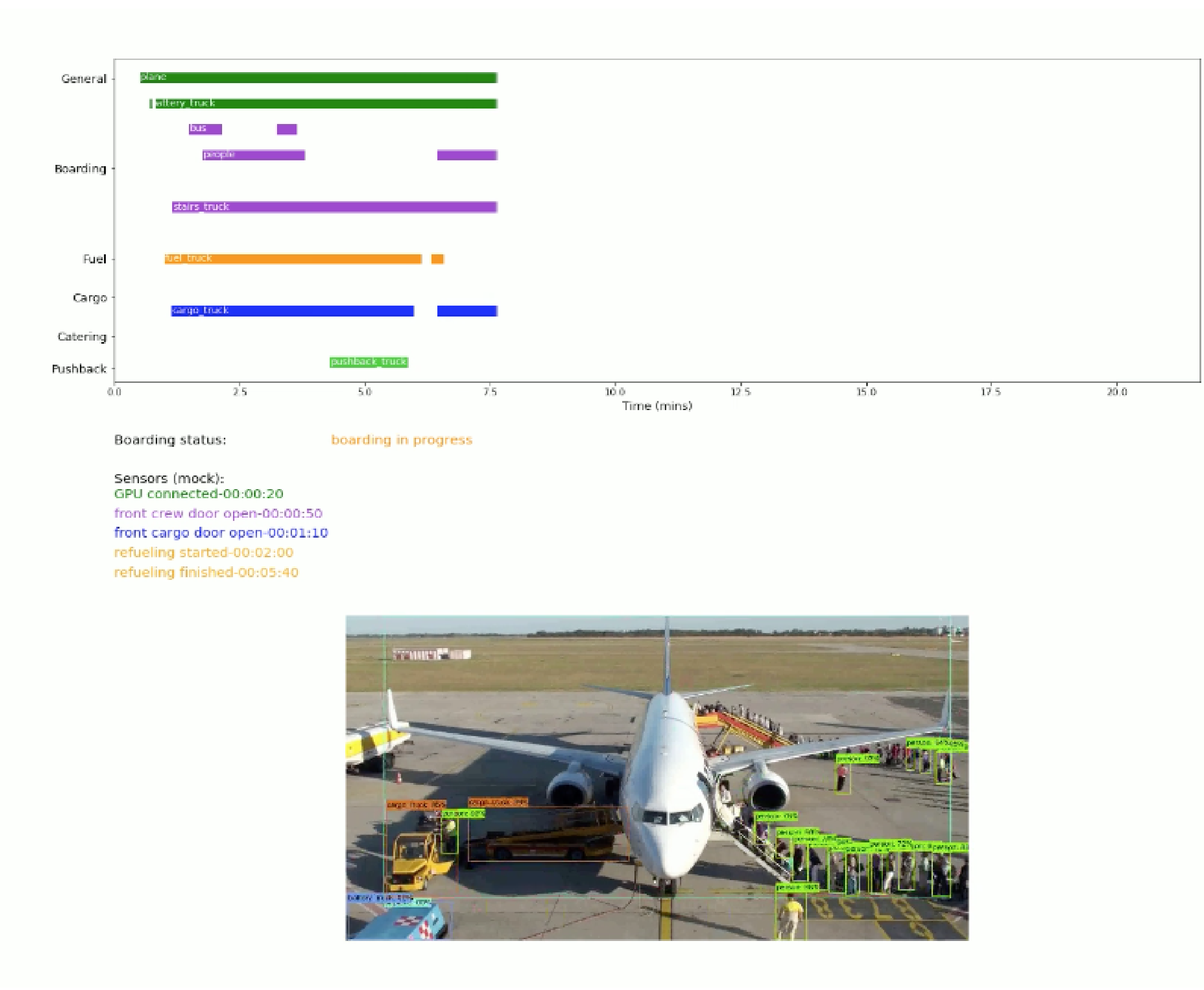

airplane: 99%
pushback_truck: 98%
Status: waiting for pushback..
airplane: 99%
pushback_truck: 99%
Status: pushback truck connecting..
airplane: 99%
pushback_truck: 99%
Status: pushback in progress..
airplane: 99%
pushback_truck: 75%
Status: pushback truck disconnected. Turnaround complete!

Figure B.5: Example output of the model showing the pushback stage. The proximity of the pushback truck to the aeroplane, as well as the change in bounding box position per frame monitors the pushback operation status.

stage of this operation can be identified in fig. B.5.


General
Boarding
Fuel
Cargo
Catering
Pushback
plane
battery_truck
bus
people
stairs_truck
fuel_truck
cargo_truck
pushback_truck
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Time (mins)

Boarding status: boarding in progress

Sensors (mock):
GPU connected-00:00:20
front crew door open-00:00:50
front cargo door open-00:01:10
refueling started-00:02:00
refueling finished-00:05:40

Figure B.6: Demonstration of how information from the object detection model and sensor data can feed into a display for turnaround monitoring.

B.3.3 Aircraft sensors

In some cases a more robust monitoring of the turnaround is extremely difficult or even impossible with visual information. For example in the refuelling of the aircraft. A fuel truck may be positioned next to the aircraft but it is very hard to detect the position of the fuel nozzle which could be obscured from view, thus it is uncertain as to whether the refuelling operation has started. We therefore propose to supplement visual information with data from aircraft sensors. There is a sensor in the fuel tank which shows how full it is. By monitoring this it is possible to definitively state that refuelling is in progress. There are similar sensors for the connection of ground power, and cargo and crew doors. fig. B.6 shows an example of how object detection and sensor information can feed into a live update of turnaround status. This is a mock up for demonstration purposes as we don't have access to video footage and sensor information at the same time.

B.4 Conclusions

In this project we demonstrated the ability of an object detection model to identify specific vehicles used in airports. Using this raw information, along with algorithms which make use of the relative spatial information, it is possible to determine when certain stages of the aircraft turnaround are in progress. Furthermore, we propose the combination of visual information with sensor data from the aircraft to help build robust monitoring of the turnaround. This can help with automatically flagging delays and provide accurate real time information to airport operation control. Data collected from monitoring the turnaround could further be used for the optimisation of this procedure in the future.

Bibliography

[1] A. Biekötter, R. Gomez-Ambrosio, P. Gregg, F. Krauss and M. Schönherr, *Constraining SMEFT operators with associated $h\gamma$ production in Weak Boson Fusion*, *Physics Letters B* **814** (2021) 136079, [`2003.06379`].

[2] A. Biekötter, P. Gregg, F. Krauss and M. Schönherr, *Constraining CP violating operators in charged and neutral triple gauge couplings*, *Phys. Lett. B* **817** (2021) 136311, [`2102.01115`].

[3] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1–29, [`1207.7214`].

[4] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [`1207.7235`].

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

[6] M. D. Schwartz, *Quantum Field Theory and the Standard Model.* Cambridge University Press, 3, 2014.

[7] H. E. Logan, *Tasi 2013 lectures on higgs physics within and beyond the standard model*, 2017.

[8] F. Abe, H. Akimoto, A. Akopian, M. G. Albrow, S. R. Amendolia, D. Amidei et al., *Observation of top quark production inpcollisions with the collider detector at fermilab*, *Physical Review Letters* **74** (Apr, 1995) 2626–2631.

[9] S. Abachi, B. Abbott, M. Abolins, B. S. Acharya, I. Adam, D. L. Adams et al., *Observation of the top quark*, *Physical Review Letters* **74** (Apr, 1995) 2632–2637.

[10] Y. Fukuda, T. Hayakawa, E. Ichihara, K. Inoue, K. Ishihara, H. Ishino et al., *Evidence for oscillation of atmospheric neutrinos*, *Physical Review Letters* **81** (Aug, 1998) 1562–1567.

[11] V. C. Rubin, J. Ford, W. K. and N. Thonnard, *Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa -> Sc.*, **225** (Nov., 1978) L107–L111.

[12] R. Massey, T. Kitching and J. Richard, *The dark matter of gravitational lensing*, *Reports on Progress in Physics* **73** (Jul, 2010) 086901.

[13] P. J. E. Peebles and B. Ratra, *The cosmological constant and dark energy*, *Reviews of Modern Physics* **75** (Apr, 2003) 559–606.

[14] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35.

[15] UA1 collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, *Phys. Lett. B* **122** (1983) 103–116.

[16] UA2 collaboration, P. Bagnaia et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\bar{p}p$ *Collider*, *Phys. Lett. B* **129** (1983) 130–140.

[17] *Physics Reports* **427** (May, 2006) 257–454.

[18] CDF collaboration, F. Abe et al., *Observation of top quark production in* $\bar{p}p$ *collisions*, *Phys. Rev. Lett.* **74** (1995) 2626–2631, [`hep-ex/9503002`].

[19] T. Han, *Collider phenomenology: Basic knowledge and techniques*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in $D \geqq 4$*, pp. 407–454, 8, 2005, `hep-ph/0508097`, DOI.

[20] K. Cranmer, *Practical statistics for the lhc*, 2015.

[21] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *The European Physical Journal C* **71** (Feb, 2011) .

[22] E. Gabrielli, F. Maltoni, B. Mele, M. Moretti, F. Piccinini and R. Pittau, *Higgs boson production in association with a photon in vector boson fusion at the lhc*, *Nuclear Physics B* **781** (Oct, 2007) 64–84.

[23] D. M. Asner, M. Cunningham, S. Dejong, K. Randrianarivony, C. Santamarina and M. Schram, *Prospects for observing the standard model higgs boson decaying intobbinal states produced in weak boson fusion with an associated photon at the lhc*, *Physical Review D* **82** (Nov, 2010) .

[24] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics: A Primer for the LHC Era.* Oxford University Press, 12, 2017.

[25] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [`1405.0301`].

[26] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [`0811.4622`].

[27] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang et al., *Openloops 2*, *The European Physical Journal C* **79** (Oct, 2019) .

[28] D. Dercks, N. Desai, J. S. Kim, K. Rolbiecki, J. Tattersall and T. Weber, *CheckMATE 2: From the model to the limit*, *Comput. Phys. Commun.* **221** (2017) 383–418, [`1611.09856`].

[29] W. Skiba, *Tasi lectures on effective field theory and precision electroweak measurements*, 2010.

[30] A. Dedes, W. Materkowska, M. Paraskevas, J. Rosiek and K. Suxho, *Feynman rules for the standard model effective field theory in r gauges*, *Journal of High Energy Physics* **2017** (Jun, 2017) .

[31] A. Falkowski, *Lectures on effective field theories*, September, 2020.

[32] B. Grzadkowski, M. Iskrzyński, M. Misiak and J. Rosiek, *Dimension-six terms in the standard model lagrangian*, *Journal of High Energy Physics* **2010** (Oct, 2010) .

[33] 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*, *Journal of High Energy Physics* **2014** (Apr, 2014) .

[34] A. Falkowski, M. González-Alonso, A. Greljo, D. Marzocca and M. Son, *Anomalous triple gauge couplings in the effective field theory approach at the LHC*, *Journal of High Energy Physics* **2017** (feb, 2017) .

[35] J. Brehmer, A. Freitas, D. Ló pez-Val and T. Plehn, *Pushing higgs effective theory to its limits*, *Physical Review D* **93** (apr, 2016) .

[36] I. Brivio, S. Bruggisser, E. Geoffray, W. Killian, M. Krämer, M. Luchmann et al., *From models to SMEFT and back?*, *SciPost Physics* **12** (jan, 2022) .

[37] A. Biekötter, A. Knochel, M. Krämer, D. Liu and F. Riva, *Vices and virtues of higgs effective field theories at large energy*, *Physical Review D* **91** (mar, 2015) .

[38] D. Pappadopulo, A. Thamm, R. Torre and A. Wulzer, *Heavy vector triplets: bridging theory and data*, *Journal of High Energy Physics* **2014** (sep, 2014) .

[39] J. Ellis, C. W. Murphy, V. Sanz and T. You, *Updated global smeft fit to higgs, diboson and electroweak data*, *Journal of High Energy Physics* **2018** (Jun, 2018) .

[40] S. Banerjee, S. Mukhopadhyay and B. Mukhopadhyaya, *Higher dimensional operators and the LHC Higgs data: The role of modified kinematics*, *Phys. Rev.* **D89** (2014) 053010, [`1308.4860`].

[41] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *The Global Electroweak and Higgs Fits in the LHC era*, *PoS* **EPS-HEP2017** (2017) 467, [`1710.05402`].

[42] A. Biekötter, T. Corbett and T. Plehn, *The Gauge-Higgs Legacy of the LHC Run II*, *SciPost Phys.* **6** (2019) 064, [`1812.07587`].

[43] E. da Silva Almeida, A. Alves, N. Rosa Agostinho, O. J. P. Éboli and M. C. Gonzalez–Garcia, *Electroweak Sector Under Scrutiny: A Combined Analysis of LHC and Electroweak Precision Data*, *Phys. Rev.* **D99** (2019) 033001, [`1812.01009`].

[44] T. Corbett, O. J. P. Eboli, D. Gonçalves, J. Gonzalez-Fraile, T. Plehn and M. Rauch, *The Higgs Legacy of the LHC Run I*, *JHEP* **08** (2015) 156, [`1505.05516`].

[45] A. Butter, O. J. P. Éboli, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, T. Plehn and M. Rauch, *The Gauge-Higgs Legacy of the LHC Run I*, *JHEP* **07** (2016) 152, [`1604.03105`].

[46] T. Han and Y. Li, *Genuine cp-odd observables at the lhc*, *Physics Letters B* **683** (Jan, 2010) 278–281.

[47] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070, [`1709.06492`].

[48] ATLAS collaboration, M. Aaboud et al., *Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *Phys. Rev.* **D98** (2018) 052003, [`1807.08639`].

[49] E. Massó and V. Sanz, *Limits on anomalous couplings of the Higgs boson to electroweak gauge bosons from LEP and the LHC*, *Phys. Rev.* **D87** (2013) 033001, [`1211.1320`].

[50] A. Falkowski, M. Gonzalez-Alonso, A. Greljo and D. Marzocca, *Global constraints on anomalous triple gauge couplings in effective field theory approach*, *Phys. Rev. Lett.* **116** (2016) 011801, [`1508.00581`].

[51] G. Brooijmans et al., *Les Houches 2013: Physics at TeV Colliders: New Physics Working Group Report*, `1405.1617`.

[52] M. Trott, *On the consistent use of Constructed Observables*, *JHEP* **02** (2015) 046, [`1409.7605`].

[53] A. Falkowski and F. Riva, *Model-independent precision constraints on dimension-6 operators*, *JHEP* **02** (2015) 039, [`1411.0669`].

[54] G. Belanger, B. Dumont, U. Ellwanger, J. F. Gunion and S. Kraml, *Global fit to Higgs signal strengths and couplings and implications for extended Higgs sectors*, *Phys. Rev.* **D88** (2013) 075008, [`1306.2941`].

[55] P. P. Giardino, K. Kannike, I. Masina, M. Raidal and A. Strumia, *The universal Higgs fit*, *JHEP* **05** (2014) 046, [`1303.3570`].

[56] B. Dumont, S. Fichet and G. von Gersdorff, *A Bayesian view of the Higgs sector with higher dimensional operators*, *JHEP* **07** (2013) 065, [`1304.3369`].

[57] P. Bechtle, S. Heinemeyer, O. Stål, T. Stefaniak and G. Weiglein, *Probing the Standard Model with Higgs signal rates from the Tevatron, the LHC and a future ILC*, *JHEP* **11** (2014) 039, [`1403.1582`].

[58] K. Cheung, J. S. Lee and P.-Y. Tseng, *Higgs precision analysis updates 2014*, *Phys. Rev.* **D90** (2014) 095009, [`1407.8236`].

[59] J. Ellis, V. Sanz and T. You, *Complete Higgs Sector Constraints on Dimension-6 Operators*, *JHEP* **07** (2014) 036, [`1404.3667`].

[60] C. Englert, R. Kogler, H. Schulz and M. Spannowsky, *Higgs coupling measurements at the LHC*, *Eur. Phys. J.* **C76** (2016) 393, [`1511.05170`].

[61] J.-B. Flament, *Higgs Couplings and BSM Physics: Run I Legacy Constraints*, `1504.07919`.

[62] L. Bian, J. Shu and Y. Zhang, *Prospects for Triple Gauge Coupling Measurements at Future Lepton Colliders and the 14 TeV LHC*, *JHEP* **09** (2015) 206, [`1507.02238`].

[63] G. Buchalla, O. Cata, A. Celis and C. Krause, *Fitting Higgs Data with Nonlinear Effective Theory*, *Eur. Phys. J.* **C76** (2016) 233, [`1511.00988`].

[64] S. Fichet and G. Moreau, *Anatomy of the Higgs fits: a first guide to statistical treatments of the theoretical uncertainties*, *Nucl. Phys.* **B905** (2016) 391–446, [`1509.00472`].

[65] L. Reina, J. de Blas, M. Ciuchini, E. Franco, D. Ghosh, S. Mishima et al., *Precision constraints on non-standard Higgs-boson couplings with HEPfit*, *PoS* **EPS-HEP2015** (2015) 187.

[66] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future*, *JHEP* **12** (2016) 135, [`1608.01509`].

[67] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *Electroweak precision constraints at present and future colliders*, *PoS* **ICHEP2016** (2017) 690, [`1611.05354`].

[68] J. Ellis, C. W. Murphy, V. Sanz and T. You, *Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data*, *JHEP* **06** (2018) 146, [`1803.03252`].

[69] F. Ferreira, B. Fuks, V. Sanz and D. Sengupta, *Probing CP-violating Higgs and gauge-boson couplings in the Standard Model effective field theory*, *Eur. Phys. J.* **C77** (2017) 675, [`1612.01808`].

[70] J. Brehmer, F. Kling, T. Plehn and T. M. P. Tait, *Better Higgs-CP Tests Through Information Geometry*, *Phys. Rev.* **D97** (2018) 095017, [`1712.02350`].

[71] F. U. Bernlochner, C. Englert, C. Hays, K. Lohwasser, H. Mildner, A. Pilkington et al., *Angles on CP-violation in Higgs boson interactions*, *Phys. Lett.* **B790** (2019) 372–379, [`1808.06577`].

[72] C. Englert, P. Galler, A. Pilkington and M. Spannowsky, *Approaching robust EFT limits for CP-violation in the Higgs sector*, *Phys. Rev.* **D99** (2019) 095007, [`1901.05982`].

[73] V. Cirigliano, A. Crivellin, W. Dekens, J. de Vries, M. Hoferichter and E. Mereghetti, *CP Violation in Higgs-Gauge Interactions: From Tabletop Experiments to the LHC*, *Phys. Rev. Lett.* **123** (2019) 051801, [`1903.03625`].

[74] J. Kumar, A. Rajaraman and J. D. Wells, *Probing CP-violation at colliders through interference effects in diboson production and decay*, *Phys. Rev.* **D78** (2008) 035014, [`0801.2891`].

[75] S. Dawson, S. K. Gupta and G. Valencia, *CP violating anomalous couplings in $W\gamma$ and $Z\gamma$ production at the LHC*, *Phys. Rev.* **D88** (2013) 035008, [`1304.3514`].

[76] M. B. Gavela, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, L. Merlo, S. Rigolin and J. Yepes, *CP violation with a dynamical Higgs*, *JHEP* **10** (2014) 044, [`1406.6367`].

[77] A. Azatov, D. Barducci and E. Venturini, *Precision diboson measurements at hadron colliders*, *JHEP* **04** (2019) 075, [`1901.04821`].

[78] NNPDF collaboration, R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti et al., *Parton distributions with QED corrections*, *Nucl. Phys.* **B877** (2013) 290–320, [`1308.0598`].

[79] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [`1410.3012`].

[80] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [`0802.1189`].

[81] NNPDF collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [`1410.8849`].

[82] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132, [`1412.7420`].

[83] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, [`0808.3674`].

[84] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038, [`0709.1027`].

[85] Z. Nagy and D. E. Soper, *Matching parton showers to NLO computations*, *JHEP* **10** (2005) 024, [`hep-ph/0503053`].

[86] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803–2819, [`1003.0694`].

[87] A. L. Read, *Presentation of search results: The CL(s) technique*, *J. Phys.* **G28** (2002) 2693–2704.

[88] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, `physics/0703039`.

[89] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.* **A389** (1997) 81–86.

[90] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys.* **B268** (1986) 621–653.

[91] H. Georgi, *Effective field theory*, *Ann. Rev. Nucl. Part. Sci.* **43** (1993) 209–252.

[92] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [`1008.4884`].

[93] 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*, *JHEP* **04** (2014) 159, [`1312.2014`].

[94] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1–98, [`1706.08945`].

[95] ATLAS collaboration, M. Aaboud et al., *Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C79** (2019) 884, [`1905.04242`].

[96] ATLAS collaboration, M. Aaboud et al., *Search for resonant WZ production in the fully leptonic final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett.* **B787** (2018) 68–88, [`1806.01532`].

[97] CMS collaboration, *Measurement of the inclusive $W\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on effective field theory coefficients*, .

[98] ATLAS collaboration, G. Aad et al., *Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS*, *Eur. Phys. J. C* **81** (2021) 163, [`2006.15458`].

[99] CMS collaboration, A. M. Sirunyan et al., *Measurement of electroweak production of a* W *boson in association with two jets in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **80** (2020) 43, [`1903.04040`].

[100] R. Rahaman and R. K. Singh, *Unravelling the anomalous gauge boson couplings in $ZW^{\pm}$ production at the LHC and the role of spin-1 polarizations*, *JHEP* **04** (2020) 075, [`1911.03111`].

[101] J. J. Ethier, R. Gomez-Ambrosio, G. Magni and J. Rojo, *SMEFT analysis of vector boson scattering and diboson data from the LHC Run II*, `2101.03180`.

[102] S. Das Bakshi, J. Chakrabortty, C. Englert, M. Spannowsky and P. Stylianou, *CP violation at ATLAS in effective field theory*, *Phys. Rev. D* **103** (2021) 055008, [`2009.13394`].

[103] ATLAS collaboration, M. Aaboud et al., *$ZZ \to \ell^+\ell^-\ell'^+\ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 032005, [`1709.07703`].

[104] ATLAS collaboration, M. Aaboud et al., *Measurement of ZZ production in the $\ell\ell\nu\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **10** (2019) 127, [`1905.07163`].

[105] ATLAS collaboration, G. Aad et al., *Measurement of the $Z(\to \ell^+\ell^-)\gamma$ production cross-section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **03** (2020) 054, [`1911.04813`].

[106] ATLAS collaboration, M. Aaboud et al., *Measurement of the $Z\gamma \to \nu\bar{\nu}\gamma$ production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector and limits on anomalous triple gauge-boson couplings*, *JHEP* **12** (2018) 010, [`1810.04995`].

[107] C. Degrande, *A basis of dimension-eight operators for anomalous neutral triple gauge boson interactions*, *JHEP* **02** (2014) 101, [`1308.6323`].

[108] U. Baur and E. L. Berger, *Probing the weak boson sector in $Z\gamma$ production at hadron colliders*, *Phys. Rev. D* **47** (1993) 4889–4904.

[109] R. Rahaman and R. K. Singh, *Anomalous triple gauge boson couplings in ZZ production at the LHC and the role of Z boson polarizations*, *Nucl. Phys. B* **948** (2019) 114754, [`1810.11657`].

[110] G. J. Gounaris, J. Layssac and F. M. Renard, *Signatures of the anomalous Z_γ and ZZ production at the lepton and hadron colliders*, *Phys. Rev. D* **61** (2000) 073013, [`hep-ph/9910395`].

[111] D. R. Yennie, S. C. Frautschi and H. Suura, *The infrared divergence phenomena and high-energy processes*, *Annals Phys.* **13** (1961) 379–452.

[112] M. Schönherr and F. Krauss, *Soft Photon Radiation in Particle Decays in SHERPA*, *JHEP* **12** (2008) 018, [`0810.5071`].

[113] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO - The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201–1214, [`1108.2040`].

[114] S. Höche, S. Kuttimalai, S. Schumann and F. Siegert, *Beyond Standard Model calculations with Sherpa*, *Eur. Phys. J. C* **75** (2015) 135, [`1412.6478`].

[115] T. Corbett, M. J. Dolan, C. Englert and K. Nordström, *Anomalous neutral gauge boson interactions and simplified models*, *Phys. Rev. D* **97** (2018) 115040, [`1710.07530`].

[116] Sherpa collaboration, E. Bothmann et al., *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034, [`1905.09127`].

[117] S. Banerjee, *private communication*, .

[118] CMS collaboration, A. M. Sirunyan et al., *Measurements of* pp → ZZ *production cross sections and constraints on anomalous triple gauge couplings at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **81** (2021) 200, [`2009.01186`].

[119] L. Breiman, *Random Forests*, *Machine Learning* **45** (2001) 5–32.

[120] R. K. Ellis and S. c. v. Veseli, *Strong radiative corrections to* $\mathrm{Wb}\bar{b}$ *production in* $p\overline{p}$ *collisions*, *Phys. Rev. D* **60** (Jun, 1999) 011501.

[121] P. B. Arnold, R. K. Ellis and M. H. Reno, *High p(t) W and Z Production at the Tevatron*, *Phys. Rev. D* **40** (1989) 912.

[122] P. B. Arnold and M. Hall Reno, *The complete computation of high-pt w and z production in second-order qcd*, *Nuclear Physics B* **319** (1989) 37–71.

[123] J. Campbell and R. K. Ellis, *Next-to-leading order corrections tow+2jet andz+2jet production at hadron colliders*, *Physical Review D* **65** (Jun, 2002) .

[124] F. Febres Cordero, L. Reina and D. Wackeroth, *NLO QCD corrections to W boson production with a massive b-quark jet pair at the Tevatron p anti-p collider*, *Phys. Rev. D* **74** (2006) 034007, [`hep-ph/0606102`].

[125] J. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, *Production of awboson and two jets with oneb-quark tag*, *Physical Review D* **75** (Mar, 2007) .

[126] F. F. Cordero, L. Reina and D. Wackeroth, *W- andz-boson production with a massive bottom-quark pair at the large hadron collider*, *Physical Review D* **80** (Aug, 2009) .

[127] S. Badger, J. M. Campbell and R. K. Ellis, *Qcd corrections to the hadronic production of a heavy quark pair and a w-boson including decay correlations*, *Journal of High Energy Physics* **2011** (Mar, 2011) .

[128] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau and P. Torrielli, *W and z/ boson production in association with a bottom-antibottom pair*, *Journal of High Energy Physics* **2011** (Sep, 2011) .

[129] J. M. Campbell, F. Caola, F. Febres Cordero, L. Reina and D. Wackeroth, *Next-to-leading order qcd predictions forw+1jet andw+2jet production with at least onebjet at the 7ev lhc*, *Physical Review D* **86** (Aug, 2012) .

[130] R. K. Ellis, Z. Kunszt, K. Melnikov and G. Zanderighi, *One-loop calculations in quantum field theory: From feynman diagrams to unitarity cuts*, *Physics Reports* **518** (Sep, 2012) 141–250.

[131] C. F. Berger, Z. Bern, L. J. Dixon, F. F. Cordero, D. Forde, T. Gleisberg et al., *Precise predictions forw+3jet production at hadron colliders*, *Physical Review Letters* **102** (Jun, 2009) .

[132] R. K. Ellis, K. Melnikov and G. Zanderighi, *Generalized unitarity at work: first nlo qcd results for hadronicw+3 jet production*, *Journal of High Energy Physics* **2009** (Apr, 2009) 077–077.

[133] R. K. Ellis, K. Melnikov and G. Zanderighi, *W+3jet production at the tevatron*, *Physical Review D* **80** (Nov, 2009) .

[134] C. F. Berger, Z. Bern, L. J. Dixon, F. F. Cordero, D. Forde, T. Gleisberg et al., *Next-to-leading order qcd predictions forw+3-jet distributions at hadron colliders*, *Physical Review D* **80** (Oct, 2009) .

[135] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg et al., *Precise predictions forw+4-jet production at the large hadron collider*, *Physical Review Letters* **106** (Mar, 2011) .

[136] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Höche, H. Ita, D. A. Kosower et al., *Next-to-leading orderw+5-jet production at the lhc*, *Physical Review D* **88** (Jul, 2013) .

[137] J. Kühn, A. Kulesza, S. Pozzorini and M. Schulze, *Electroweak corrections to large transverse momentum production of w bosons at the lhc*, *Physics Letters B* **651** (04, 2007) 160–165.

[138] J. H. Kühn, A. Kulesza, S. Pozzorini and M. Schulze, *Electroweak corrections to hadronic production of w bosons at large transverse momenta*, *Nuclear Physics B* **797** (Jul, 2008) 27–77.

[139] A. Denner, S. Dittmaier, T. Kasprzik and A. Mück, *Electroweak corrections to w + jet hadroproduction including leptonic w-boson decays*, *Journal of High Energy Physics* **2009** (Aug, 2009) 075–075.

[140] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *Journal of High Energy Physics* **2014** (Jul, 2014) .

[141] S. Actis, A. Denner, L. Hofer, A. Scharf and S. Uccirati, *EW and QCD One-Loop Amplitudes with RECOLA*, *PoS* **RADCOR2013** (2013) 034, [`1311.6662`].

[142] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini and M. Schönherr, *Nlo electroweak automation and precise predictions for w+multijet production at the lhc*, 2015.

[143] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini and M. Schönherr, *Nlo qcd+ew predictions for v + jets including off-shell vector-boson decays and multijet merging*, *Journal of High Energy Physics* **2016** (Apr, 2016) 1–51.

[144] E. Bothmann and D. Napoletano, *Automated evaluation of electroweak sudakov logarithms in sherpa*, *The European Physical Journal C* **80** (Nov, 2020) .

[145] S. Otten, K. Rolbiecki, S. Caron, J.-S. Kim, R. R. de Austri and J. Tattersall, *Deepxs: Fast approximation of mssm electroweak cross sections at nlo*, 2019.

[146] A. Buckley, A. Kvellestad, A. Raklev, P. Scott, J. V. Sparre, J. Van den Abeele et al., *Xsec: the cross-section evaluation code*, *The European Physical Journal C* **80** (Dec, 2020) .

[147] D. Maître and H. Truong, *A factorisation-aware matrix element emulator*, 2021.

[148] S. Badger and J. Bullock, *Using neural networks for efficient evaluation of high multiplicity scattering amplitudes*, *Journal of High Energy Physics* **2020** (Jun, 2020) .

[149] J. Aylett-Bullock, S. Badger and R. Moodie, *Optimising simulations for diphoton production at hadron colliders using amplitude neural networks*, *Journal of High Energy Physics* **2021** (Aug, 2021) .

[150] L. Bottou and Y. Le Cun, *Sn: A simulator for connectionist models*, in *Proceedings of NeuroNimes 88*, (Nimes, France), pp. 371–382, 1988, http://leon.bottou.org/papers/bottou-lecun-88.

[151] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning.* MIT Press, 2016.

[152] R. Girshick, J. Donahue, T. Darrell and J. Malik, *Rich feature hierarchies for accurate object detection and semantic segmentation*, 2014.

[153] J. Uijlings, K. Sande, T. Gevers and A. Smeulders, *Selective search for object recognition*, *International Journal of Computer Vision* **104** (09, 2013) 154–171.

[154] R. Girshick, *Fast r-cnn*, 2015.

[155] S. Ren, K. He, R. Girshick and J. Sun, *Faster r-cnn: Towards real-time object detection with region proposal networks*, 2016.

[156] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu et al., *Ssd: Single shot multibox detector*, *Lecture Notes in Computer Science* (2016) 21–37.

[157] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, *You only look once: Unified, real-time object detection*, 2016.

[158] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi et al., *Speed/accuracy trade-offs for modern convolutional object detectors*, *CoRR* **abs/1611.10012** (2016) , [`1611.10012`].

[159] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays et al., *Microsoft COCO: common objects in context*, *CoRR* **abs/1405.0312** (2014) , [`1405.0312`].