Precision Physics in Extensions of the Standard Model

PHILIP ADAM WAITE

Precision Physics in Extensions of the Standard Model

WAITE, PHILIP,ADAM

Precision Physics in Extensions of the Standard Model

Philip Adam Waite

A thesis submitted for the degree of
Doctor of Philosophy

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

August 2019

Dedicated to

Hannah

Precision Physics in Extensions of the Standard Model

Philip Adam Waite

August 2019

Abstract: The Standard Model of particle physics has been remarkably successful at explaining the behaviour of nature at very high energies. It has been thoroughly tested by experiments at the Large Hadron Collider and almost all of its predictions have agreed closely with observations. Despite this, there are many phenomena that it cannot explain, such as the origins of neutrino masses. Therefore, the Standard Model alone cannot provide a complete explanation of reality, and so there must exist physics beyond it. However, finding it is particularly difficult due to the aforementioned success of the Standard Model.

In this thesis, we continue the search for this new physics by carrying out five separate studies that could contribute towards the overall goal. We calculate new limits on specific extensions of the Standard Model by using lepton-flavour-violating decays of τ leptons. We also use electroweak precision measurements for the first time to constrain the trilinear self-coupling of the Higgs boson, and perform the most precise calculation to date of the production of a heavy neutrino via gluon fusion. Considering machine-learning techniques, we improve the robustness of an autoencoder used for unsupervised searches for new physics, and we develop a new approach to using neural networks for solving differential equations, which we apply to the calculation of cosmological phase transitions. A general theme across all of these investigations is the importance of a high precision, both in terms of the theoretical calculations, and the experiments through which they are tested.

Declaration

The work in this thesis is based on research carried out at the Institute for Particle Physics Phenomenology, Department of Physics, Durham University. No part of this thesis has been submitted elsewhere for any other degree or qualification. It is based on my own work, and on work undertaken in collaboration with the other co-authors of the following publications:

- Chapter 2 contains results which have been published in: C. Hays, M. Mitra, M. Spannowsky and **P. Waite**, *Prospects for new physics in $\tau \to l\mu\mu$ at current and future colliders*, *JHEP* **05** (2017) 014 [1701.00870].

- Chapter 3 contains results which have been published in: G. D. Kribs, A. Maier, H. Rzehak, M. Spannowsky and **P. Waite**, *Electroweak oblique parameters as a probe of the trilinear Higgs boson self-interaction*, *Phys. Rev.* **D95** (2017) 093004 [1702.07678].

- Chapter 4 contains results which have been published in: R. Ruiz, M. Spannowsky and **P. Waite**, *Heavy neutrinos from gluon fusion*, *Phys. Rev.* **D96** (2017) 055042 [1706.02298].

- Chapter 5 contains results which have been published in: A. Blance, M. Spannowsky and **P. Waite**, *Adversarially-trained autoencoders for robust unsupervised new physics searches*, *JHEP* **10** (2019) 047 [1905.10384].

- Chapter 6 contains results which have been published in: M. L. Piscopo, M. Spannowsky and **P. Waite**, *Solving differential equations with neural networks: Applications to the calculation of cosmological phase transitions*, *Phys. Rev.* **D100** (2019) 016002 [1902.05563].

During my doctoral studies, I also co-authored the following publications, but their results are not presented in this thesis:

- V. V. Khoze, J. Reiness, M. Spannowsky and **P. Waite**, *Precision measurements for the Higgsploding Standard Model*, *J. Phys.* **G46** (2019) 065004 [1709.08655].
- M. Chala, V. V. Khoze, M. Spannowsky and **P. Waite**, *Mapping the shape of the scalar potential with gravitational waves*, *Int. J. Mod. Phys.* **A34** (2019) 1950223 [1905.00911].
- P. Schichtel, M. Spannowsky and **P. Waite**, *Constraining strongly coupled new physics from cosmic rays with machine learning techniques*, *EPL* **127** (2019) 61002 [1906.09064].

Acknowledgements

I would like to acknowledge all those who have played a part in helping me along the four-year journey towards the completion of this PhD. First and foremost, I would like to thank my supervisor, Michael Spannowsky, for his many words of wisdom and interesting discussions about topics in physics and beyond, and for allowing me to work on such a wide variety of research projects. On these projects I worked with many co-authors, whose support and contributions have been invaluable. Of particular note was Richard Ruiz, who went above and beyond to teach me the intricate details of the physics we were working on together. This thesis was examined by Jeppe Andersen and Malcolm Fairbairn, and I am grateful for their friendliness during the viva examination and for passing the thesis without any corrections.

Outside of my academic work I met many amazing people, without which the last four years would have been far harder, and I would not wish to list them all here for the fear of missing someone. However, I would particularly like to mention my housemates at Park House Road, Oliver Wales, Jack Richings, Jacob Kegerreis and Kristian Moffat, whose companionship and friendship over the last four years has made this journey much more enjoyable.

Of course, none of this would have been possible at all without the support of my parents, Carolyn and Eddy, my brother, Matthew, and my girlfriend, Hannah, who have been there with me the whole way. I thank you for all that you have done.

Finally, I am grateful for the financial support of the Science and Technology Facilities Council in allowing me to study for this PhD.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model of particle physics is an extremely well-tested theory and is the best approximation to the behaviour of nature at length scales of $\mathcal{O}(10^{-18}$ m)—the smallest scales that can currently be probed experimentally to a high precision—that humanity has ever achieved. In fact, it can be argued that it is the most successful theory from any branch of physics. Although it has its flaws—there are many unexplained phenomena, as we shall see later—and thus cannot be the complete story of our universe, the model has been so successful that almost every attempt to construct an even better description of the physics at such scales begins by taking the Standard Model and extending it in some way. In any case, there is no denying the fact that even if the absolute description of nature is radically different, then this behaviour would mainly be seen at even smaller lengths (higher energies), and must reduce to something that resembles the Standard Model in the lower-energy limit.

It is therefore no surprise that attempts to find what lies beyond the Standard Model often prove futile, and verifying such models experimentally is very difficult since the Standard Model already agrees very well with many observations. However, we are guided in the search both by the unexplained phenomena, and areas of the Standard Model which are currently experimentally more loosely constrained than

others. The unexplained phenomena motivate extensions of the Standard Model which can explain such effects, while loosely constrained parts of the model motivate more precise studies to see whether a discrepancy with data can be found.

Thus, in this thesis we explore various avenues where it may be possible to shed light on what lies beyond the Standard Model. In Chapter 2, we consider three extensions of the Standard Model that address unexplained phenomena, and examine the constraints that can be placed on their parameter spaces due to their prediction of lepton-flavour-violating decays of the τ lepton. We then perform a calculation in Chapter 3 of the effects of an extension to the Higgs sector on the trilinear self-coupling of the Higgs boson via its imprint on the electroweak oblique parameters, using perturbative methods from quantum field theory. In Chapter 4, we make use of techniques from quantum chromodynamics to carry out a threshold-resummed calculation of the cross section for the production of heavy neutrinos from gluon fusion.

In the final part of the thesis, we focus on some applications of machine learning to helping with the search for physics beyond the Standard Model. In Chapter 5, we consider the application of autoencoders to performing model-independent searches for new physics, and improve their robustness by combining them with adversarial neural networks. Additionally, in Chapter 6 we develop a new approach for solving differential equations with neural networks, and apply this to the calculation of tunnelling rates which can be used to study the scalar sector in the Standard Model and beyond. Finally, we summarise the work in the thesis in Chapter 7.

Before we delve into any of these studies, we first discuss the basic ideas on which they are based, with more specific details included later along with each particular investigation. We begin by introducing the Standard Model in Sec. 1.1, and discuss some of its further interesting features which are important in Sec. 1.2. We then give a brief overview of the shortcomings of the Standard Model and the ways in which it can be extended in Sec. 1.3, before concluding the chapter in Sec. 1.4 by introducing some of the machine-learning techniques and ideas that are used in particle physics.

1.1 The Standard Model of Particle Physics

We will begin with a discussion of the Standard Model of particle physics. In order to do so, we must first introduce the symmetries and particle content from which the model is constructed. We will then describe the concept of spontaneous symmetry breaking, which is necessary for the generations of masses for particles that would otherwise be massless. Furthermore, we will introduce the topic of flavour physics which arises as a result of the mechanism from which the masses are generated.

1.1.1 Overview and Basic Ideas

The Standard Model [1–7] is a quantum field theory which describes the interactions between fermions (particles with half-integer spin) and bosons (particles with integer spin), and is based upon the notion of symmetry, which lies at the heart of all of particle physics. All the matter content in the theory is associated to fermions, while all the fundamental interactions are mediated by bosons. A key concept is the statement that the theory is invariant under a gauge transformation[1] [8], which is a local transformation of the form,

$$\psi(x) \to e^{-i\alpha^a(x)t^a}\psi(x) \ . \tag{1.1.1}$$

Here, $\psi(x)$ is a matter field which depends on the spacetime coordinate x, t^a are the generators of the corresponding Lie group and $\alpha^a(x)$ parameterise the transformation. The relevant Lie groups for the Standard Model are the abelian U(1) group and the non-abelian groups SU(n), with $n \geq 2$. The generators for SU(n) satisfy the commutation relation,

$$[t^a, t^b] = if^{abc}t^c \ , \tag{1.1.2}$$

where f^{abc} are the structure constants of the group. The generators t^a are traceless Hermitian matrices with a dimension that depends on their representation. In the

[1]The theory is also invariant under the Poincaré group—the combination of both Lorentz transformations and translations in spacetime.

fundamental representation they are $n \times n$ matrices, while in the adjoint representation they are $(n^2-1) \times (n^2-1)$ matrices, where n^2-1 is the dimension of the group and corresponds to the total number of generators. Thus the indices a, b and c run from 1 to n^2-1.

The Standard Model is constructed from the gauge group,

$$G_{\text{SM}} \equiv \text{SU}(3)_c \times \text{SU}(2)_L \times \text{U}(1)_Y \ . \tag{1.1.3}$$

The $\text{SU}(3)_c$ gauge group corresponds to the strong interactions of quantum chromodynamics [3, 4], which is mediated by gluons that interact with colour charges. Similarly, the $\text{SU}(2)_L \times \text{U}(1)_Y$ gauge group describes the electroweak interactions [1,2] mediated by vector bosons that couple to either the left-handed weak isospin or the weak hypercharge. After spontaneous symmetry breaking [5–7], this electroweak gauge group is broken down to the $\text{U}(1)_{\text{EM}}$ group that describes the interactions of quantum electrodynamics [9–11] between electrically charged particles and mediated by photons. This is due to the addition of a complex scalar field in the Lagrangian, which takes a non-zero vacuum expectation value (VEV) and is able to generate mass terms for both the fermions and gauge bosons without breaking the gauge invariance of the theory.

The Lagrangian of the theory consists of the terms allowed by gauge invariance, and can be expressed as a sum of several distinct parts[2] [12]:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{Fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} \ . \tag{1.1.4}$$

The first piece contains the kinetic terms for the gauge fields, with the fermion part describing the kinetic terms for the fermion fields and their interactions with the gauge fields. Furthermore, the Higgs terms contain the potential for the additional complex scalar field and its interactions with the gauge bosons. Finally, the Yukawa terms describe the interactions between the fermions and the complex scalar field. Here, we will describe each part separately.

[2]There are also gauge-fixing and ghost terms, which will be defined later.

Gauge Terms

Associated to each gauge group are sets of gauge fields, which arise from the requirement for the Lagrangian to be invariant under such a gauge transformation. For the Standard Model, these gauge fields and their representations under the gauge groups are,

$$G^a_\mu \in (\mathbf{8}, \mathbf{1}, 0) \ , \qquad W^i_\mu \in (\mathbf{1}, \mathbf{3}, 0) \ , \qquad B_\mu \in (\mathbf{1}, \mathbf{1}, 0) \ , \tag{1.1.5}$$

corresponding to the $\mathrm{SU}(3)_c$, $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ gauge groups, respectively. The kinetic terms for the gauge fields in the Lagrangian are then,

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu} - \frac{1}{4} W^i_{\mu\nu} W^{i\,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \ , \tag{1.1.6}$$

where the field-strength tensors are defined in terms of the gauge fields as,

$$\begin{aligned} G^a_{\mu\nu} &= \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu \ , \\ W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g \epsilon^{ijk} W^j_\mu W^k_\nu \ , \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \ . \end{aligned} \tag{1.1.7}$$

The structure constants for $\mathrm{SU}(3)_c$ are denoted by f^{abc}, while the structure constants for $\mathrm{SU}(2)_L$ are the Levi-Civita tensor ϵ_{ijk}. The parameters g_s and g are the coupling strengths of their respective gauge groups. In addition, the gauge group $\mathrm{U}(1)_Y$ also has a coupling strength, which we denote by g'. Note that the additional term in the non-abelian field-strength tensors gives rise to trilinear and quartic self-interactions of the associated gauge boson.

Under a gauge transformation parameterised by α, the abelian field B_μ transforms as,

$$B_\mu \to B_\mu - \frac{1}{g'} \partial_\mu \alpha \ , \tag{1.1.8}$$

which means that a term proportional to $B_\mu B^\mu$ in the Lagrangian is not invariant under such a transformation. A more complicated transformation holds for the non-abelian fields, but the equivalent quadratic terms $G^a_\mu G^{a\,\mu}$ and $W^i_\mu W^{i\,\mu}$ are also

not invariant. Therefore, these mass terms for the gauge fields are forbidden from the Lagrangian, with the consequence that in the unbroken theory, the gauge bosons are massless.

Fermion Terms

As we mentioned earlier, all matter in the theory is described by fermions. These are particles with half-integer spin, which determines their transformation under the Lorentz group. The fermions can be further classified according to their representations under the gauge groups. In the Standard Model, these classes are,

$$\begin{aligned} Q_L &= \begin{pmatrix} u_L \\ d_L \end{pmatrix} \in (\mathbf{3}, \mathbf{2}, 1/3) \ , \qquad u_R \in (\mathbf{3}, \mathbf{1}, 4/3) \ , \qquad d_R \in (\mathbf{3}, \mathbf{1}, -2/3) \ , \\ L_L &= \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \in (\mathbf{1}, \mathbf{2}, -1) \ , \qquad e_R \in (\mathbf{1}, \mathbf{1}, -2) \ . \end{aligned} \tag{1.1.9}$$

The first line are the quarks, which are in the triplet representation of $\mathrm{SU}(3)_c$ and thus interact via the strong force. The second line are the leptons, which are in the trivial representation of $\mathrm{SU}(3)_c$ and are thus uncharged under colour. Furthermore, u and d are up-type quarks and down-type quarks while ν and e are neutrinos and charged leptons, respectively. For both quarks and leptons there exist left-handed doublets under the gauge group $\mathrm{SU}(2)_L$. Specifically, this means that the gauge group $\mathrm{SU}(2)_L$ only interacts with the left-handed components of the fields, and right-handed fields are thus singlets under $\mathrm{SU}(2)_L$. For a field ψ, the left-handed component is projected out by,

$$\psi_L \equiv P_L \psi = \frac{1}{2}(1 - \gamma^5)\psi \ , \tag{1.1.10}$$

where γ^5 is the fifth Dirac matrix $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. Similarly, there is a right-handed projection operator $P_R = (1 + \gamma^5)/2$. Finally, the third number describing the representations is the weak hypercharge Y, which is the constant of proportionality of the coupling strength under the abelian group $\mathrm{U}(1)_Y$. There is a copy of these

representations for each of three 'generations' of fermions, which are identical except for their masses and mixings.

In the doublet representations of $\mathrm{SU}(2)_L$, the upper field has a third component of weak isospin $T_3 = 1/2$. Similarly, the lower field has $T_3 = -1/2$. The electromagnetic charge Q is related to the weak isospin and weak hypercharge by the Gell-Mann-Nishijima formula [13, 14],

$$Q = T_3 + \frac{Y}{2} \ . \tag{1.1.11}$$

In the broken theory after spontaneous symmetry breaking, the charge Q is the constant of proportionality of the coupling strength under the remaining symmetry $\mathrm{U}(1)_{\mathrm{EM}}$.

The kinetic terms for the fermions in the Lagrangian can then be constructed from these representations. They take the form,

$$\begin{aligned}\mathcal{L}_{\mathrm{Fermion}} = \sum_{i=1}^{3} \Big[& \overline{Q}_{Li} i\gamma^\mu D_\mu Q_{Li} + \overline{u}_{Ri} i\gamma^\mu D_\mu u_{Ri} + \overline{d}_{Ri} i\gamma^\mu D_\mu d_{Ri} \\ & + \overline{L}_{Li} i\gamma^\mu D_\mu L_{Li} + \overline{e}_{Ri} i\gamma^\mu D_\mu e_{Ri} \Big] \ ,\end{aligned} \tag{1.1.12}$$

where the sum is over the fermion generations. There is also an implicit sum over the Dirac spin indices, as well as the indices corresponding to colour and weak isospin. For a fermion field ψ charged under colour and weak isospin with a weak hypercharge Y, the covariant derivative D_μ is defined by,

$$D_\mu \psi = \partial_\mu \psi + i g_s \frac{\lambda^a}{2} G^a_\mu \psi + i g \frac{\sigma^i}{2} W^i_\mu \psi + i g' \frac{Y}{2} B_\mu \psi \ . \tag{1.1.13}$$

Here, λ^a and σ^i are the Gell-Mann and Pauli matrices, which are the generators for $\mathrm{SU}(3)_c$ and $\mathrm{SU}(2)_L$, respectively[3]. The covariant derivatives in Eq. (1.1.12) only contain the gauge fields corresponding to the gauge groups under which the relevant fermion field is charged.

Similar to the gauge bosons, it might be tempting to write down a fermion mass term proportional to $\overline{\psi}_L \psi_R + \overline{\psi}_R \psi_L$. However, the left-handed and right-handed

[3]Strictly speaking, the generators for $\mathrm{SU}(3)_c$ and $\mathrm{SU}(2)_L$ are $\lambda^a/2$ and $\sigma^i/2$, respectively.

fields have different quantum numbers under the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ gauge groups. This means that they both transform differently under a gauge transformation, so the transformation of the left-handed field cannot cancel the transformation of the right-handed field, and vice versa. Therefore, a fermion mass term would not be gauge invariant and in the unbroken theory fermions are also massless.

Higgs and Yukawa Terms

Observations have shown that there exist gauge bosons and fermions with masses [15]. This is in contrast with the prediction of the unbroken theory of the Standard Model that mass terms for these are forbidden. To solve such a problem, it was proposed [5–7] to include in the theory a complex scalar Φ, which, when it takes a non-zero VEV, can break the gauge invariance spontaneously and generate mass terms for the gauge bosons and fermions. While the full Lagrangian is invariant under the gauge symmetry, the vacuum of the scalar field is not[4]. The complex scalar Φ, also called the 'Higgs doublet', has the following representation under the gauge symmetries:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \in (\mathbf{1}, \mathbf{2}, 1) \ . \tag{1.1.14}$$

The kinetic and potential terms for the Higgs doublet, which are required to satisfy the conditions of gauge invariance and renormalisability, are then,

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \ . \tag{1.1.15}$$

Furthermore, its interactions with the fermions are described by the Yukawa terms,

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j} \left[(Y_u)_{ij} \overline{Q}_{Li} \tilde{\Phi} u_{Rj} + (Y_d)_{ij} \overline{Q}_{Li} \Phi d_{Rj} + (Y_e)_{ij} \overline{L}_{Li} \Phi e_{Rj} \right] + \text{h.c.} \ , \tag{1.1.16}$$

where the sum runs over the generation indices, and the conjugate doublet $\tilde{\Phi} \equiv i\sigma^2 \Phi^*$ is required to construct a term for up-type quarks which is invariant under $\mathrm{SU}(2)_L$.

[4] It is crucial that the field here is a scalar. If a field such as a fermion or a vector boson—which have non-trivial representations under the Lorentz group—has a non-zero VEV, then it would violate Lorentz invariance.

1.1.2 Higgs Sector and Spontaneous Symmetry Breaking

We will now discuss the complex scalar Φ and its interactions in more detail. The Higgs potential from Eq. (1.1.15) is,

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \ . \tag{1.1.17}$$

It is useful to parameterise the Higgs doublet in terms of four real scalar fields as,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \ . \tag{1.1.18}$$

Of the four degrees of freedom, two arise because the doublet is charged under $\mathrm{SU}(2)_L$, and two arise because the scalars are complex. The potential can then be written as,

$$V(\Phi) = \frac{\mu^2}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2)^2 \ . \tag{1.1.19}$$

In the scenario where both μ^2 and λ are positive, this quartic potential has a minimum at the origin describing a set of scalars with masses μ but is otherwise relatively uninteresting. However, when $\mu^2 < 0$ and $\lambda > 0$, the vacuum of the potential is shifted away from the origin. This minimum is located at,

$$|\langle \Phi \rangle|^2 = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2} \ . \tag{1.1.20}$$

In principal, this VEV could be chosen to lie in any of the four real fields. However, in order to ensure that electromagnetism remains a symmetry, the VEV must be chosen in either ϕ_3 or ϕ_4 which are electrically neutral, or any complex rotation between them. For convenience, ϕ_3 is chosen to have the non-zero VEV, $\langle \phi_3 \rangle = v$, with the other fields having a VEV of zero.

It is then possible to expand the Lagrangian around the VEV of the Higgs doublet. Denoting the physical Higgs field as h, the Higgs doublet becomes,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + h + i\phi_4 \end{pmatrix} \ . \tag{1.1.21}$$

The $\phi_{1,2,4}$ are Goldstone bosons, which are unphysical degrees of freedom and can be removed through a gauge transformation—this is known as the unitary gauge. Since the Higgs field h is a physical degree of freedom, no gauge transformation can remove this field. The Higgs potential after expanding around the VEV contains the terms,

$$V(\Phi) \supset \lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4 \ . \tag{1.1.22}$$

Therefore, the mass of the Higgs boson in terms of the Lagrangian parameters is $m_h = \sqrt{2\lambda v^2}$, and it has trilinear and quartic self-interactions with coupling strengths proportional to λv and $\lambda/4$, respectively. The mass of the Higgs boson has been measured experimentally to be $m_h \sim 125$ GeV [16], and the VEV—which sets the scale for electroweak physics—is measured from the muon lifetime to be $v \sim 246$ GeV [17]. These two quantities are sufficient to completely describe the Higgs potential of the Standard Model. This is to be expected, since there were only two independent parameters—μ^2 and λ—to begin with.

We can now expand the rest of the terms in the Lagrangian which contain the Higgs field around the VEV. In order to do so, it is useful to redefine the gauge fields in such a way that mass terms that arise in the Lagrangian can be expressed solely in terms of these new gauge fields—these are the mass eigenstates. For the W^1_μ and W^2_μ gauge fields, they can be linearly combined into the charged bosons,

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp i W^2_\mu) \ . \tag{1.1.23}$$

Furthermore, the gauge fields W^3_μ and B_μ, which have the same quantum numbers under the gauge symmetries, mix together into the physical gauge fields A_μ and Z_μ,

$$\begin{aligned} A_\mu &= \cos\theta_{\mathrm{W}} B_\mu + \sin\theta_{\mathrm{W}} W^3_\mu \ , \\ Z_\mu &= -\sin\theta_{\mathrm{W}} B_\mu + \cos\theta_{\mathrm{W}} W^3_\mu \ , \end{aligned} \tag{1.1.24}$$

where θ_{W} is the weak mixing angle defined by $g \sin\theta_{\mathrm{W}} = g' \cos\theta_{\mathrm{W}} = e$. Here, e is the coupling constant of the residual $\mathrm{U}(1)_{\mathrm{EM}}$ symmetry. With these definitions, after expanding Eq. (1.1.15) around the VEV of the Higgs doublet the Lagrangian

contains the terms,

$$\mathcal{L}_{\text{Higgs}} \supset \frac{1}{4}g^2v^2W_\mu^+W^{-\mu} + \frac{1}{8}(g^2+g'^2)v^2Z_\mu Z^\mu \ , \tag{1.1.25}$$

which generates the following masses for the $W^\pm$ and Z bosons:

$$m_W = \frac{gv}{2} \ , \qquad m_Z = \frac{\sqrt{g^2+g'^2}v}{2} \ . \tag{1.1.26}$$

No mass term is generated for the A boson, which means that this can be identified with the photon γ which mediates the unbroken $\mathrm{U}(1)_{\text{EM}}$ symmetry. The relation between the masses and the coupling strengths g and g' allows the weak mixing angle to be expressed in terms of the gauge boson masses as,

$$\sin^2\theta_{\text{W}} = 1 - \frac{m_W^2}{m_Z^2} \ . \tag{1.1.27}$$

Furthermore, the Higgs VEV can be related to the other parameters by,

$$v = \frac{2m_W \sin\theta_{\text{W}}}{e} \ . \tag{1.1.28}$$

Counting the degrees of freedom allows us to see what has happened during spontaneous symmetry breaking. The original Higgs doublet contained four degrees of freedom, but we found after spontaneous symmetry breaking that only one of these was physical and the others were carried by the Goldstone bosons—these represent the ability to make gauge transformations, and thus can be removed in the unitary gauge. Furthermore, when the $W^\pm$ and Z bosons became massive, they each obtained an additional longitudinal polarisation. This means that the three degrees of freedom have effectively been 'eaten up' by the three gauge bosons to generate their masses, preserving the total number of degrees of freedom in the theory. This is the Higgs mechanism [5–7]—for every broken continuous local symmetry, the gauge boson associated to the generator of the broken symmetry acquires a mass, and the gauge boson associated to any unbroken symmetry (here a $\mathrm{U}(1)_{\text{EM}}$ symmetry) remains massless. It is also important to note that since the Higgs doublet is not charged under $\mathrm{SU}(3)_c$, this remains a symmetry of the theory and the gluon fields

G^a_μ do not acquire a mass.

The relation between the masses of the $W^\pm$ and Z bosons is actually rather interesting. Rearranging Eq. (1.1.27) further, we get,

$$\frac{m_W^2}{m_Z^2} = \cos^2\theta_\mathrm{W} \ . \tag{1.1.29}$$

This means that the ratio between the masses of the gauge bosons is determined by the amount of mixing between the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ gauge groups. If there was no mixing at all, corresponding to $\cos\theta_\mathrm{W} = 1$, then the $W^\pm$ and Z boson masses would be equal, with the Z boson being made up entirely from the W^3_μ gauge eigenstate, which can be seen from Eq. (1.1.24). We can then define the ρ parameter,

$$\rho \equiv \frac{m_W^2}{m_Z^2 \cos^2\theta_\mathrm{W}} \ . \tag{1.1.30}$$

Clearly at tree level, we get the result that $\rho = 1$. However, remarkably this is also very stable when higher-order radiative corrections are included [18]. This is due to custodial symmetry, which arises because the Higgs doublet obeys an approximate $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$ global symmetry[5], which after spontaneous symmetry breaking is broken to the approximate global symmetry $\mathrm{SU}(2)_V$. This custodial symmetry relates the strengths of the interactions of the $W^\pm$ and Z bosons and protects the relation between their masses from higher-order corrections. However, it is not respected by the Yukawa terms—the higher-order corrections are proportional to terms that break the custodial symmetry, so the largest deviations from $\rho = 1$ are due to the top quark—but these deviations can be absorbed into the definitions of the parameters in Eq. (1.1.30) such that $\rho = 1$ in the Standard Model.

1.1.3 Flavour Physics

In the previous section we showed how the addition of a complex scalar doublet to the Lagrangian resulted in a physical scalar particle—the Higgs boson h—and generated

[5]This global symmetry is exact in the vanishing weak hypercharge and Yukawa coupling limits, $g' \to 0$ and $Y_{u,d,e} \to 0$, respectively.

masses for the gauge bosons when expanding the Lagrangian around the VEV. We will now look at what happens when the same expansion is carried out in the Yukawa terms in Eq. (1.1.16). Performing such an expansion and only considering terms proportional to the VEV, we get,

$$\mathcal{L}_{\text{Yukawa}} \supset -\sum_{i,j}\Bigg[(Y_u)_{ij}\overline{u}_{Li}\frac{v}{\sqrt{2}}u_{Rj} + (Y_d)_{ij}\overline{d}_{Li}\frac{v}{\sqrt{2}}d_{Rj} + (Y_e)_{ij}\overline{e}_{Li}\frac{v}{\sqrt{2}}e_{Rj}\Bigg] + \text{h.c.}\ . \tag{1.1.31}$$

These are exactly the fermion mass terms which were forbidden by gauge invariance from being explicitly included in the Lagrangian. However, there is an issue. Since the original fermion states in the Lagrangian were eigenstates of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry which has since been broken, there is no guarantee that these should coincide with the mass eigenstates which arise after spontaneous symmetry breaking. Specifically, this means that the Yukawa matrices $(Y_{u,d,e})_{ij}$ are not necessarily diagonal in the gauge basis. This is the same behaviour that we saw for the gauge fields—the fields W^i_μ and B_μ associated to the gauge symmetries are not aligned with the physical gauge eigenstates $W^\pm_\mu$, Z_μ and A_μ.

Therefore, it is necessary to rotate the Yukawa matrices such that in the basis of mass eigenstates, they are diagonal. This is done using two unitary matrices $V_L^{u,d,e}$ and $V_R^{u,d,e}$ for each of the three Yukawa matrices. The Yukawa matrices are then diagonalised by,

$$\widetilde{Y}_u = V_L^u Y_u V_R^{u\dagger}\ , \qquad \widetilde{Y}_d = V_L^d Y_d V_R^{d\dagger}\ , \qquad \widetilde{Y}_e = V_L^e Y_e V_R^{e\dagger}\ , \tag{1.1.32}$$

where $\widetilde{Y}_{u,d,e}$ are diagonal. These unitary matrices can also be used to rotate the left-handed and right-handed quark and lepton states into the mass basis, where the Yukawa terms are diagonal in flavour, and each fermion thus has a mass,

$$m_f = \frac{y_f v}{\sqrt{2}}\ , \tag{1.1.33}$$

where y_f is a diagonal element from one of the three Yukawa matrices $\widetilde{Y}_{u,d,e}$. Having

rotated the Yukawa matrices into the mass basis, it is important to consider how this affects the rest of the Lagrangian. For the charged leptons, since there is only one Yukawa matrix, the gauge eigenstates can be defined such that they coincide with the mass eigenstates and so the interactions between the leptons and gauge bosons remains diagonal. However, since there are two Yukawa matrices for the quarks, it is not possible to simultaneously keep the gauge interactions both diagonal and aligned with the mass eigenstates. In the interactions with the photon and Z boson (neutral-current interactions) the effects of the rotations cancel, but in the interactions with the $W^{\pm}$ bosons (charged-current interactions) they take the form,

$$\mathcal{L}_{\text{Fermion}} \supset \frac{e}{\sqrt{2}\sin\theta_{\text{W}}} \overline{u}_L \gamma^{\mu} W_{\mu}^{+} V_L^{u} V_L^{d\dagger} d_L + \text{h.c.} \,. \tag{1.1.34}$$

The non-diagonal matrix $V_{\text{CKM}} \equiv V_L^{u} V_L^{d\dagger}$ is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [19], and means that the charged-current interactions between quarks in their gauge eigenstate can change the flavour of their mass eigenstates. By convention, the gauge basis for the up-type quarks is chosen to align with the mass basis, and the rotation from V_{CKM} is applied exclusively to the down-type quarks.

The CKM matrix is a 3×3 unitary matrix and contains four independent real parameters after absorbing five parameters into the quark fields. These consist of three mixing angles and one CP-violating phase. In theory there is nothing to prevent this matrix from being equal to the identity matrix. However, experiments have shown that mixings between the quark generations do indeed take place [15], and that they have a pronounced hierarchy with mixings between adjacent generations far stronger than those between the first and third. Furthermore, the non-zero CP-violating phase leads to CP-violation in the weak interactions which was first discovered in 1964 by Cronin and Fitch [20]. Importantly, the discovery of CP-violation showed that more than two generations of fermions must exist, since a 2×2 unitary matrix would not contain a CP-violating phase [21]. This is because of the four real parameters required to describe a 2×2 unitary matrix, three can be

absorbed into the quark fields, leaving only a single independent mixing angle.

1.2 Further Aspects of the Standard Model

So far, we have discussed the classical Lagrangian of the Standard Model which is built from the gauge group $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ describing the interactions between fermions and gauge bosons, and which has an additional complex scalar field that generates masses for the fermions and gauge bosons. In addition to the topics we have described above, there are some further aspects of the Standard Model which are both interesting and important, which we will now discuss.

1.2.1 Gauge Fixing and Ghosts

Since the Lagrangian is invariant under the gauge symmetries, a gauge transformation of the fields will still describe the same physical state. However, this introduces unphysical degrees of freedom, such as those arising from the Goldstone bosons we described earlier. During the quantisation of the Lagrangian, it becomes necessary to fix a particular gauge in order to be able to carry out calculations. This is achieved by adding to the Lagrangian a set of gauge-fixing terms,

$$\mathcal{L}_{\mathrm{GF}} = -\frac{1}{2\xi}F_G^2 - \frac{1}{2\xi}F_A^2 - \frac{1}{2\xi}F_Z^2 - \frac{1}{\xi}F_+F_- \ , \tag{1.2.1}$$

where we have the definitions,

$$F_G^a = \partial^\mu G_\mu^a \ , \qquad F_A = \partial^\mu A_\mu \ , \qquad F_Z = \partial^\mu Z_\mu - \xi m_Z \phi^0 \ ,$$
$$F_\pm = \partial^\mu W_\mu^\pm \mp i\xi m_W \phi^\pm \ . \tag{1.2.2}$$

These are known as the covariant, or R_ξ, gauges and the specific gauge is determined by the single parameter ξ. The Goldstone bosons are denoted by $\phi^\pm \equiv (\phi_1 \pm i\phi_2)/\sqrt{2}$ and $\phi^0 \equiv \phi_4$, with $\phi_{1,2,4}$ defined in Eq. (1.1.21).

The choice of the gauge parameter affects the propagators of the gauge bosons. In the R_ξ gauges, the propagators of the gauge bosons take the form,

$$\Pi^{\mu\nu}_{VV}(p) = \frac{-i}{p^2 - m_V^2 + i\varepsilon}\left[g^{\mu\nu} + (\xi - 1)\frac{p^\mu p^\nu}{p^2 - \xi m_V^2}\right] , \tag{1.2.3}$$

where m_V is the mass of the gauge boson V. Some particular gauge choices are:

- Unitary gauge: $\xi \to \infty$,
- Feynman gauge: $\xi \to 1$,
- Landau gauge: $\xi \to 0$.

Note that the unitary gauge only makes sense for massive gauge bosons where $m_V > 0$. The massive gauge bosons have corresponding Goldstone bosons with masses $\sqrt{\xi} m_V$, and so in the unitary gauge these degrees of freedom do not propagate. Although the particular choice of gauge parameter is arbitrary, in any meaningful calculation it is necessary for the dependence on ξ to cancel at the end. Therefore, the calculation can be performed in any gauge, and it is usually simplest to use the Feynman gauge, where only the first term in Eq. (1.2.3) proportional to the metric tensor $g^{\mu\nu}$ contributes. However, this means that in the Feynman gauge the Goldstone bosons must be included as intermediate unphysical particles, although they cannot appear in the final states. Furthermore, highly relativistic gauge bosons are dominated by their longitudinal polarisations, which are precisely the corresponding Goldstone bosons. This means that at very high energies, gauge boson scattering can be approximated by Goldstone bosons—this is the Goldstone boson equivalence theorem [22].

A further subtlety arises when performing the gauge fixing for non-abelian gauge theories. In the calculation of Green's functions, the determinant for abelian theories does not contribute, but for non-abelian theories it causes the appearance of unphysical Faddeev-Popov ghost particles [23]. These occur as a solution to the problem of maintaining the consistency between the Feynman diagram approach

and the path integral formulation of quantum field theory, and to preserve unitarity. In the Lagrangian, the ghost fields are described by the terms,

$$\mathcal{L}_{\text{Ghost}} = \sum_{a,b} \overline{c}^a \frac{\delta F_G^a}{\delta \theta^b} c^b + \sum_{\alpha,\beta} \overline{c}^\alpha \frac{\delta F_\alpha}{\delta \theta^\beta} c^\beta \ . \tag{1.2.4}$$

Here, the first sum runs over the gluon indices from 1 to 8 and the second sum runs over the electroweak gauge boson indices A, Z, $+$ and $-$. The variational derivatives represent the infinitesimal change of the gauge-fixing functions F with respect to their corresponding gauge-transformation parameter θ. The c are the ghost particles, which are anti-commuting scalars and must be included in Feynman diagrams along with the Goldstone bosons unless the calculation is performed in the unitary gauge.

1.2.2 Loop Integrals and Divergences

When Standard Model calculations are performed beyond tree level, the Feynman diagrams which describe the perturbative expansion contain loops. These are parts of the diagrams that contain unconstrained momenta which are not specified by the external kinematics of the process, and so these momenta must be integrated over all possible values. They can arise when calculating higher-order corrections to observables that appear at tree level, such as the Drell-Yan production of lepton pairs, or they appear at the lowest order in processes that are forbidden at tree level—these are loop-induced processes, such as a Higgs boson decaying to two photons.

A generic N-point one-loop integral contains various tensor structures and denominators involving external momenta and masses, and has the form [24],

$$T_N^{\mu_1 \dots \mu_P}(p_1, \dots, p_{N-1}, m_0, \dots, m_{N-1}) = \frac{(2\pi\mu)^{4-d}}{i\pi^2} \int d^d k \frac{k^{\mu_1} \dots k^{\mu_P}}{D_0 \dots D_{N-1}} \ , \tag{1.2.5}$$

where the denominator contains the propagator factors,

$$D_j = (k + r_j)^2 - m_j^2 + i\varepsilon \ , \qquad r_j = \sum_{i=1}^{j} p_i \ , \qquad r_0 = 0 \ . \tag{1.2.6}$$

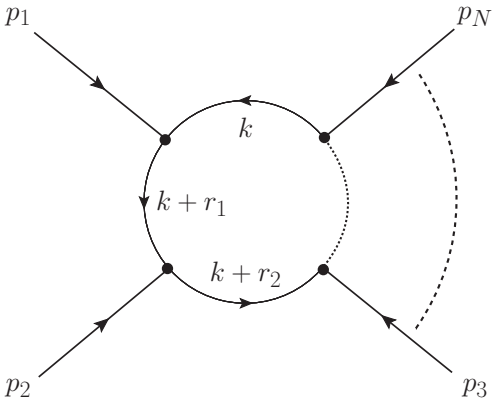


Figure 1.1: General N-point function showing the configuration of external and internal momenta flowing through the loop. All momenta are defined to be incoming, and the arrows show the direction of the momenta.

The integral corresponds to the N-point function shown in Fig. 1.1. The p_i are the external momenta of the N-point integrals, and the r_i are convenient definitions to simplify the notation. The number of spacetime dimensions in these integrals has been written as $d \equiv 4 - 2\epsilon$, as a precursor to performing dimensional regularisation. Furthermore, μ is an unphysical mass scale which is required to maintain the correct mass dimension for the integral.

The general tensor integrals in Eq. (1.2.5) can be reduced using the Passarino-Veltman reduction procedure [25] to a set of four independent scalar basis integrals containing no powers of momenta in the numerators. The one-point, two-point and three-point scalar basis integrals are:

$$A_0(m_0^2) = \frac{(2\pi\mu)^{2\epsilon}}{i\pi^2} \int d^d k \frac{1}{k^2 - m_0^2 + i\varepsilon} \ , \tag{1.2.7}$$

$$B_0(p_1^2, m_0^2, m_1^2) = \frac{(2\pi\mu)^{2\epsilon}}{i\pi^2} \int d^d k \prod_{i=0}^{1} \frac{1}{(k + r_i)^2 - m_i^2 + i\varepsilon} \ , \tag{1.2.8}$$

$$C_0(p_1^2, p_2^2, r_2^2, m_0^2, m_1^2, m_2^2) = \frac{(2\pi\mu)^{2\epsilon}}{i\pi^2} \int d^d k \prod_{i=0}^{2} \frac{1}{(k + r_i)^2 - m_i^2 + i\varepsilon} \ . \tag{1.2.9}$$

These integrals can be evaluated in dimensional regularisation by Feynman parameterisation, Wick rotating into Euclidean space, and then carrying out the d-dimensional spherical integrals. The results are:

$$A_0(m_0^2) = m_0^2 \left(\frac{1}{\epsilon} - \gamma_{\mathrm{E}} + \ln 4\pi - \ln \frac{m_0^2}{\mu^2} + 1 \right) \ , \tag{1.2.10}$$

$$B_0(p_1^2, m_0^2, m_1^2) = \frac{1}{\epsilon} - \gamma_{\mathrm{E}} + \ln 4\pi - \int_0^1 dx_1 \ln \frac{M_B^2}{\mu^2} \ , \tag{1.2.11}$$

$$C_0(p_1^2, p_2^2, r_2^2, m_0^2, m_1^2, m_2^2) = -\Gamma(3) \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \frac{1}{2M_C^2} \ . \tag{1.2.12}$$

Here, γ_{E} is Euler's constant, and the dummy variables x_1 and x_2 result from Feynman parameterisation, where we have defined,

$$M_B^2 = (1 - x_1)m_1^2 + x_1 m_2^2 + x_1(x_1 - 1)p_1^2 \ , \tag{1.2.13}$$

for the B_0 integral, and for the C_0 integral we use the definition,

$$\begin{aligned} M_C^2 = (1 - x_1 - x_2)m_1^2 + x_1 m_2^2 + x_2 m_3^2 + x_1(x_1 + x_2 - 1)p_1^2 \\ - x_1 x_2 p_2^2 + x_2(x_1 + x_2 - 1)r_2^2 \ . \end{aligned} \tag{1.2.14}$$

The A_0 tadpole integral and the B_0 self-energy integral are ultraviolet (UV) divergent, which is manifest by the $1/\epsilon$ poles arising from dimensional regularisation, where the limit $\epsilon \to 0$ is to be taken. However, this is not just an artefact of the regularisation procedure—another way of regularising the integrals is by cutting off the integrals above the value of a cutoff scale Λ. The divergences would then show up as quadratic and logarithmic terms in Λ, with the limit $\Lambda \to \infty$ to be taken. In practice, the regularisation procedure is just a convenient way of parameterising the divergences so that calculations can be carried out. These UV-divergent contributions must vanish in any physical observable, otherwise the calculations would be not be meaningful. This is achieved through a process called 'renormalisation', which we will shortly discuss.

Another source of divergences that arise in loop integrals are infrared (IR) divergences. These occur when massless particles, such as photons and gluons, appear in the loops. When these particles become very soft—which means that their momenta become zero—their propagators diverge. However, these IR divergences cancel in the complete calculation of any observable when considering the emission of massless particles into the final state. Divergences from these arise when their energy becomes

soft or they become collinear[6] to a massless particle that emits them, but they cancel with the loop divergences in any complete calculation due to the Kinoshita-Lee-Nauenberg (KLN) theorem [26, 27].

1.2.3 Renormalisation

The parameters which appear in the classical Lagrangian of the Standard Model, such as the coupling constants and the masses of the gauge bosons, satisfy very simple relations. At tree level, these parameters can be identified with physical observables. Unfortunately, this identification with physical observables is destroyed when considering higher-order corrections, which involve UV-divergent loop integrals. Although it is still possible to relate physical observables to other physical observables in such a way as to cancel the divergences that arise, this has the consequence of losing an intuitive meaning for the Lagrangian parameters themselves.

A systematic approach can be taken to absorb these infinite corrections into redefinitions of the parameters and fields in the Lagrangian, through the process of renormalisation. In this process, the 'bare' parameters and fields that appear in the Lagrangian are split into two parts. These consist of 'renormalised' parameters and fields, along with corresponding 'counterterms'. The renormalised parameters and fields are finite, while the counterterms contain the UV-divergent contributions. This allows the Lagrangian to be expressed as,

$$\mathcal{L}_0 = \mathcal{L} + \delta\mathcal{L} \ , \tag{1.2.15}$$

where $\mathcal{L}_0$ is the original bare Lagrangian, while $\mathcal{L}$ contains renormalised parameters and fields and $\delta\mathcal{L}$ is the counterterm Lagrangian. For example, the bare Higgs boson mass can be written as $m^2_{h,0} = m^2_h + \delta m^2_h$, while the bare Higgs boson field may

[6] Actually, the complete cancellation of collinear divergences also requires the consideration of jets of particles in the initial state, and they can be absorbed into the momentum-scale dependence of the parton distribution functions. The KLN theorem only ensures the cancellation of collinear divergences in the final state.

be expressed as $h_0 = \sqrt{Z_h}h = (1 + \delta Z_h/2)h$. This means that calculations can be performed in terms of Feynman diagrams involving the renormalised parameters and fields, with the counterterms being simply added through additional diagrams, thus cancelling all the divergences that arise order-by-order in perturbation theory.

While the minimal requirement for the counterterms is to cancel the UV-divergent quantities that appear, the exact values of the finite parts are not fixed by this condition. The finite parts are instead fixed by choosing a set of renormalisation conditions, which enable the calculation of the counterterms to be carried out for a particular order in perturbation theory. A common way of defining the counterterms is through the on-shell renormalisation scheme [28]. In this scheme, the renormalised parameters, such as the gauge boson masses, become equal to their experimentally measurable equivalents, since the propagators are defined such that their poles lie at the values of the physical masses. For the Higgs boson, this results in the following expressions for the counterterms:

$$\delta t = -T\,, \qquad \delta m_h^2 = \mathrm{Re}\left[\Sigma_{hh}(m_h^2)\right]\,, \qquad \delta Z_h = -\mathrm{Re}\left[\left.\frac{\partial \Sigma_{hh}(p^2)}{\partial p^2}\right|_{p^2=m_h^2}\right]\,, \quad (1.2.16)$$

where δt, δm_h^2 and δZ_h are the tadpole, mass and field counterterms respectively. These are calculated from the Higgs tadpole diagrams T, and the Higgs self-energy diagrams Σ_{hh}. Other counterterms for the gauge bosons, fermions, CKM mixing matrix and electric charge in the on-shell scheme can be found in Ref. [28].

Another way of defining the counterterms is through the modified minimal subtraction ($\overline{\mathrm{MS}}$) scheme, where the counterterms simply contain the divergent $1/\epsilon$ poles along with the factors γ_E and $\ln 4\pi$ which appear in Eqs. (1.2.10) and (1.2.11). While losing the equivalence between the Lagrangian parameters and the experimentally measured quantities, this scheme can make calculations simpler since the finite parts of the counterterms do not require computation, and is also usually used to renormalise quark masses for which pole masses are difficult to measure. Furthermore, it causes the renormalised parameters to be dependent on the unphysical scale μ at which the renormalisation is performed.

1.2.4 Quantum Chromodynamics

We have previously discussed how the gauge group $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ is broken down to the group $\mathrm{U}(1)_{\mathrm{EM}}$ by spontaneous symmetry breaking. However, since the Higgs doublet Φ is not charged under the gauge group $\mathrm{SU}(3)_c$, this remains an unbroken symmetry of the theory. Therefore, the eight gluon fields G^a_μ remain massless and mediate the strong interactions between the quarks, which carry colour charges. Specifically, the quarks can have one of three different colour charges since they exist in the triplet representation of $\mathrm{SU}(3)_c$. The strong interactions are described by quantum chromodynamics, which we will now briefly discuss.

One of the features of quantum chromodynamics which distinguishes it from the other unbroken interaction of the Standard Model, $\mathrm{U}(1)_{\mathrm{EM}}$, is the concept of asymptotic freedom [29, 30]. This is the phenomenon that at higher energy scales, the coupling strength becomes weaker until at infinite energy in the asymptotic limit, it vanishes. On the other hand, the electromagnetic coupling constant $\alpha_e \equiv e^2/(4\pi)$ increases with energy until it reaches a Landau pole, which signifies the breakdown of the perturbative behaviour of the theory. The implication of this for the strong coupling constant $\alpha_s \equiv g_s^2/(4\pi)$ is that the increase in coupling strength occurs at the opposite end of the energy scale—at energies of $\mathcal{O}(200\ \mathrm{MeV})$ and below. A similar breakdown of the perturbative applicability of the theory happens, and non-perturbative effects become important at low energies. This leads to the colour-confinement of quarks and gluons—they can cannot be observed individually, but are instead bound into colourless objects such as baryons and mesons.

The dependence of the coupling constant on the energy scale arises as a result of the renormalisation of the parameters in the Lagrangian, which we described earlier. Specifically, the requirement that the bare coupling does not depend on the unphysical scale μ leads to a renormalisation-group (RG) evolution equation for the renormalised coupling [31]:

$$\frac{d\alpha_s(\mu)}{d\ln\mu} = \beta(\alpha_s)\ . \tag{1.2.17}$$

The function $\beta(\alpha_s)$ is known as the β-function for quantum chromodynamics. It can be computed by calculating the higher-order perturbative corrections that contribute to the vertices or propagators and their resulting counterterms, and has an expansion of the form,

$$\beta(\alpha_s) = -2\alpha_s(\mu) \sum_{n\geq 0} \beta_n \left(\frac{\alpha_s(\mu)}{4\pi}\right)^{n+1} . \tag{1.2.18}$$

The first term contains the one-loop coefficient β_0, which in the $\overline{\text{MS}}$ scheme has the value[7],

$$\beta_0 = \frac{11}{3} C_A - \frac{4}{3} T_F n_f , \tag{1.2.19}$$

where we have used the colour-algebra definitions,

$$C_F = \frac{N_c^2 - 1}{2N_c} , \qquad C_A = N_c , \qquad T_F = \frac{1}{2} . \tag{1.2.20}$$

The number of light quark flavours is denoted by n_f, while the number of quark colours is denoted by N_c. Clearly for a small number of light quark flavours, the β-function in Eq. (1.2.18) at one loop is negative which leads to the asymptotic freedom property, and this result holds to all orders. The solution to the RG-evolution equation in Eq. (1.2.17) at one loop is then given in terms of the coupling strength at another scale μ_0 by,

$$\alpha_s(\mu) = \frac{\alpha_s(\mu_0)}{1 + \beta_0 \frac{\alpha_s(\mu_0)}{2\pi} \ln \frac{\mu}{\mu_0}} = \alpha_s(\mu_0) \sum_{n\geq 0} \left(-\beta_0 \frac{\alpha_s(\mu_0)}{2\pi} \ln \frac{\mu}{\mu_0}\right)^n . \tag{1.2.21}$$

When the scales μ and μ_0 differ by a large amount, the logarithms that arise can be large and make the series formally divergent. However, the running of the strong coupling resums these large logarithms to all orders, thus giving a sensible result. If we define $\Lambda \equiv \mu_0 e^{-2\pi/(\beta_0 \alpha_s(\mu_0))}$, the strong coupling constant can alternatively be written as,

$$\alpha_s(\mu) = \frac{2\pi}{\beta_0 \ln \frac{\mu}{\Lambda}} , \tag{1.2.22}$$

where Λ is the $\mathcal{O}(200$ MeV) scale that denotes the perturbative breakdown of quantum chromodynamics.

[7]The first two β-function coefficients β_0 and β_1 are independent of the renormalisation scheme used, while the higher-order coefficients are scheme dependent.

1.3 Beyond the Standard Model

One glaring omission from the Standard Model, which we have yet to mention, is that it contains no right-handed fields for the neutrinos. The Standard Model generates masses for nine fermions—six quarks and three charged leptons—through the three mass matrices which arise from the Yukawa terms after spontaneous symmetry breaking. With no right-handed neutrinos, a Yukawa term cannot be constructed for the neutrinos and thus they are massless. However, observations have shown that neutrinos oscillate [32], which necessarily requires them to have masses, albeit very small ones, which is in conflict with the Standard Model prediction. Thus, there must exist a mass basis for the three generations of neutrinos, and they are related to the gauge eigenstates through the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix U_{P} [33] by,

$$\nu_\ell = \sum_i (U_{\mathrm{P}})_{\ell i}\, \nu_i \ . \tag{1.3.1}$$

Here, ν_ℓ are the neutrinos in the gauge basis and ν_i are the neutrinos in the mass basis. Note that by convention, the charged leptons are defined to have aligned gauge and mass eigenstates, so the PMNS matrix is analogous to the CKM matrix for quarks.

The most obvious way to extend the Standard Model to include masses for the neutrinos is to simply add three right-handed neutrino fields, which we will denote by N_{Ri}. This allows for the construction of a Yukawa term, which after spontaneous symmetry breaking generates masses for the neutrinos—which are known as 'Dirac masses'—given by $M_{Di} = y_{\nu i} v/\sqrt{2}$. However, for the current best-fit values for the neutrino masses, this would require Yukawa couplings of $y_\nu \sim \mathcal{O}(10^{-12})$. This is not satisfying, since it would raise the further problem of why this Yukawa coupling is so small compared with all the other Yukawa couplings, and relates to the problem of naturalness [34].

It is possible to extend this even further. Since the lepton doublet L_L and the conjugate Higgs doublet $\tilde{\Phi}$ have the same weak hypercharge, a right-handed neutrino

must have zero weak hypercharge in order for this Yukawa term to be gauge invariant. Therefore, these right-handed neutrinos must be singlets under all gauge groups of the Standard Model. In addition to explaining why it might be very difficult to detect such a neutrino[8], this also means that it is possible for these neutrinos to be their own anti-particle, $N_R^c = N_R$. A further mass term—known as a 'Majorana mass' term [35]—can then be constructed and added to the Lagrangian:

$$\mathcal{L}_{\text{Majorana}} = \sum_{i,j} \left[\frac{1}{2} \overline{N_{Ri}^c} i\gamma^\mu \partial_\mu N_{Ri} - \frac{1}{2} (M_R)_{ij} \overline{N_{Ri}^c} N_{Rj} \right] , \tag{1.3.2}$$

where the first term is the kinetic term for the Majorana neutrino fields, and $(M_R)_{ij}$ is the Majorana mass matrix.

In principle there could be any number of right-handed neutrinos, but for every massive left-handed neutrino there needs to be at least a corresponding right-handed neutrino to generate its mass. Including both the Dirac and Majorana masses together, the diagonalisation of the full mass matrix results in three light mass states with masses $m_{\text{Light}} \simeq M_D^2/M_R$ and a heavy mass state for each additional right-handed neutrino with masses $m_{\text{Heavy}} \simeq M_R$. In fact, the presence of the Majorana mass term means that the left-handed and right-handed neutrinos do not pair up to create a Dirac fermion—instead, two separate Majorana fermions are created for each pair. This is the Type-I Seesaw Model [36], and provides a solution to the problem of small neutrino masses since the masses are suppressed by the large scale M_R, while still allowing for $\mathcal{O}(1)$ Yukawa couplings. If the Yukawa couplings are as small as that of the electron, $\mathcal{O}(10^{-5})$, then this would allow for a heavy neutrino to exist close to the electroweak scale while still providing the required suppression of the masses of the light neutrinos [37]. We will discuss seesaw models and neutrino mixing further in later chapters.

There are many other issues with the Standard Model. It suffers from a hierarchy problem when considering energy scales much higher than the electroweak scale [38],

[8] Although the right-handed neutrinos are singlets and thus their gauge eigenstates do not couple to the Standard Model gauge bosons, their mass eigenstates do couple via their mixing with the light neutrinos.

and does not include an explanation of gravity [39]. The rotation curves of galaxies imply the existence of dark matter [40], for which no viable candidate is provided by the Standard Model. Furthermore, the amount of CP-violation in the Standard Model is not sufficient to explain why the Universe contains far more matter than anti-matter [41]. Many high-energy models have been proposed which can explain such issues, normally by including some extra degrees of freedom at a high scale, such as a heavy right-handed neutrino which we have just discussed. This is referred to as a 'top-down' approach, since we start with the physics at a high scale and derive the consequences at lower energies.

A more general way to extend the Standard Model is to take a 'bottom-up' approach, by starting with the Standard Model Lagrangian and adding further terms constructed entirely out of Standard Model fields. Since the Standard Model Lagrangian contains all renormalisable and gauge-invariant terms up to a mass dimension of four, these terms must have a higher mass dimension. This effective field theory [42] can be written as,

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_{n,i} \frac{1}{\Lambda^{n-4}} c_n^{(i)} \mathcal{O}_n^{(i)} \ , \tag{1.3.3}$$

where Λ is the scale of new physics, which is required to ensure the full Lagrangian still has a mass dimension of four, and describes the cutoff scale where the effective field theory is valid up to. The operators $\mathcal{O}_n^{(i)}$ have mass dimension n and are constructed from Standard Model fields, and are multiplied by the dimensionless Wilson coefficients $c_n^{(i)}$ which represent the coupling strength of each operator.

The Wilson coefficients describe the behaviour of the new physics at high energies, and the operators represent their effects at low energies after the new physics is integrated out. This is analogous to the Fermi theory of weak interactions [43], where the four-fermion operators that describe weak decays appear after the higher-energy $W^{\pm}$ boson is integrated out, and allows one to calculate that the interaction strength G_{F} is proportional to m_W^{-2}. However, in the case where the physics at the higher scale is not known, the Wilson coefficients cannot be calculated directly by matching

onto a lower-energy theory and must be left as parameters to be determined, such as from their contribution to experimental observables. In later chapters, we will consider both top-down and bottom-up approaches to physics beyond the Standard Model.

1.4 Machine Learning in Particle Physics

The data that is generated by the experiments at colliders such as the Large Hadron Collider (LHC) [44]—as well as pseudo-data from Monte Carlo simulations—contain many possible kinematic features which have highly non-linear correlations due to the multitude of perturbative and non-perturbative processes involved [45]. In a particle collision these processes consist of the initial-state partons arising from the incoming protons, which then interact through a hard process that can be described perturbatively. The particles produced can then decay, before undergoing non-perturbative parton showering and hadronisation processes. Finally, the final-state objects are measured by the various detectors in the experiment, which have a finite resolution and produce an imperfect reconstruction of the true final state. In even a single particle collision, these processes can result in there being hundreds of final-state objects, each of which have a four-momentum in addition to extra information that can be deduced such as the charge and type of the object.

Clearly with a large number final-state objects there are many possible observables that can be extracted, so it is important to develop appropriate analysis strategies that make the most effective use of them. Traditionally, the data-analysis techniques used in particle physics have involved sequences of Boolean decisions followed by statistical analyses of a smaller number of individual observables that have been selected based on the physics being studied [46]. Importantly, there is no guarantee that these observables are optimal and crucial information from other kinematic features may be ignored.

In recent years, however, machine-learning algorithms have become increasingly popular for extracting correlations in high-dimensional parameter spaces and classifying the structure within the data. Such techniques are also often referred to as the field of 'multivariate analysis', and include algorithms such as logistic regression, support vector machines and boosted decision trees [47]. Even more recently, the increasing power of computers has prompted a renewed interest in the study of large artificial neural networks, which is commonly called 'deep learning' [48]. There are a wide variety of structures of neural network, and they each have many applications such as in image classification, natural language processing, and also in particle physics [49].

Machine-learning algorithms generally fall into two main categories [47]. The first category is 'supervised learning', where the algorithm has the objective of predicting a label associated to each data point. This could be a discrete label for classification tasks, or a continuous value for regression tasks. To train the algorithm, it must be supplied with a set of training data, which consists of pairs of inputs $\vec{x}^i$ and target labels $\vec{y}^i$. On this training data, a loss function $\mathcal{L}(\vec{y}, f_\theta(\vec{x}))$ is calculated and minimised in an attempt to obtain the optimal values of the parameters $\vec{\theta}$ of the algorithm $f_\theta(\vec{x})$. Supervised-learning techniques are used in particle physics for tasks such as the tagging of jets based on whether they were initiated by a quark, gluon or $W^\pm$ boson [50, 51], or for the classification of entire events into whether they were induced by new physics or the Standard Model [52]. An application of a regression task that a supervised-learning approach can be used for is pileup mitigation [53].

The other main category is 'unsupervised learning'. Here, the training data does not have labels attached, so the purpose of the algorithm is to identify structure within the data that is a priori unknown. Furthermore, this means that a loss function based on the true label of each data point cannot be constructed, and so a different optimisation objective is required. Typically, unsupervised learning is used for clustering with algorithms such as k-means clustering, or for anomaly detection or dimensionality reduction [47]. In terms of particle physics, one of its most promising applications is for model-independent searches for new physics [54–56]. In this

scenario, the algorithm only needs to be trained on background events induced by processes in the Standard Model, with events arising from new physics then flagged up as anomalies.

In order for a trained machine-learning algorithm to be generalisable, it is important to avoid overfitting to the training data. This can be achieved through the use of a cross-validation dataset to monitor its performance so that training can be stopped when the loss on the cross-validation set beings to increase. Furthermore, this allows the hyperparameters[9] to be selected such that the algorithm has the best performance on the cross-validation set. Additional ways to avoid overfitting include regularisation methods such as dropout [57].

As we mentioned above, artificial neural networks are becoming increasingly important in the analysis of high-dimensional datasets that arise in particle physics and elsewhere. In essence, these are functions that map a vector of inputs to a vector of outputs through a series of matrix multiplications combined with non-linear activation functions. In each layer of the network there are a number of 'nodes' (also called 'units') where the activation function is applied, and which are connected to nodes in the adjacent layers through weights and biases. Neural networks are trained using backpropagation algorithms [58] on batches of training points for a given number of epochs, with the loss function determined by the particular task of the network.

When a neural network consists of multiple hidden layers, it is described as a 'deep neural network'. These can be very computationally expensive to train since they contain an enormous number of parameters and require very large training sets, but are able to perform remarkably well at their desired tasks. Fully-connected neural networks involve connections between all nodes in adjacent layers, but convolutional neural networks are another type of neural network that can be used for image classification, such as jets images [59], due to their translationally-invariant properties. Another form of neural network are recurrent neural networks, which can handle

[9]Hyperparameters are adjustable parameters of the algorithm which are not optimised during the minimisation of the loss function and must be chosen before the algorithm is trained.

inputs of varying length with a sequential ordering, and thus can be applied to natural language processing. They are also ideally suited for dealing with events or individual jets in particle physics, where there is no fixed total number of objects, and they have recently been applied to such cases [60, 61]. We will study neural networks further in Chapters 5 and 6.

Chapter 2

Lepton-Flavour-Violating τ Decays as Probes of New Physics

In this chapter, we perform our first investigation into physics beyond the Standard Model by considering the capability of current and future collider experiments at discovering new physics in lepton-flavour-violating decays of the τ leptons. These decays cannot occur in the Standard Model, and so their detection would be conclusive evidence for the existence of physics beyond the Standard Model. However, in the current absence of their detection, estimates of their maximum allowed branching ratios can be used to exclude regions of the parameter spaces of models of new physics which give rise to such decays.

2.1 Motivation

The Standard Model exhibits an accidental global $\mathrm{U}(3)^5$ flavour symmetry group in the absence of Yukawa interactions. This symmetry group can be expressed as,

$$G_{\mathrm{F}} \equiv \mathrm{SU}(3)_Q \times \mathrm{SU}(3)_u \times \mathrm{SU}(3)_d \times \mathrm{SU}(3)_L \times \mathrm{SU}(3)_e \times \mathrm{U}(1)^5 \ , \tag{2.1.1}$$

where each separate SU(3) group corresponds to the representations of the fermions: Q_L, u_R, d_R, L_L and e_R. The Yukawa couplings violate this global symmetry and it is subsequently broken to a smaller symmetry [62]:

$$G_{\mathrm{F}} \xrightarrow{\text{Yukawa}} \mathrm{U}(1)_B \times \mathrm{U}(1)_{L_e} \times \mathrm{U}(1)_{L_\mu} \times \mathrm{U}(1)_{L_\tau} \ . \tag{2.1.2}$$

This implies that in the Standard Model, baryon number (B) is conserved, along with the lepton number of each individual family ($L_{e,\mu,\tau}$), which we refer to as lepton flavour conservation. Therefore, to all orders in perturbation theory, no process can exist in the Standard Model which violates this[1].

However, as we mentioned in the previous chapter, the Standard Model is not sufficient to explain all the experimental observations that have been made of nature. For example, the neutrinos have been observed to oscillate which implies that they have small non-zero masses, and it does not provide a satisfactory solution to the hierarchy problem. Many models which are able to address these issues also introduce interactions which can violate the lepton flavour symmetry. The amount of lepton flavour violation (LFV) can be very large depending on the parameters of the model, and arise either at tree level or at higher orders in perturbation theory.

So far, however, no observations of such LFV processes have been made. In fact, the branching ratios (BR) of the transitions between the first two generations of leptons are quite constrained experimentally: $\mathrm{BR}(\mu \to e\gamma) \leq 5.7 \times 10^{-13}$ at 90% confidence level (CL) [63, 64] and $\mathrm{BR}(\mu^\mp \to e^\pm e^\mp e^\mp) \leq 10^{-12}$ at 90% CL [64, 65]. This already restricts large parts of the parameter spaces of such models that allow for LFV decays. On the other hand, the outlook for transitions from the third generation is quite different—for the τ lepton, the limit on the LFV decays is much larger at $\mathrm{BR}(\tau \to \ell\ell\ell) \lesssim 10^{-8}$ at 90% CL [64]. Consequently, large LFV couplings in models are still allowed provided that the stringent limits between the first two generations are obeyed.

[1]There is actually a small amount of B- and L-violation possible on the quantum level from sphalerons in the Standard Model, leaving $\mathrm{U}(1)_{B-L}$ as the conserved symmetry. However, the amount is tiny and sphalerons have not currently been observed.

There exist many models of physics beyond the Standard Model which allow for LFV decays of the τ lepton. There are several seesaw models which are used to introduce non-zero masses to the neutrinos, and account for their smallness through a suppression by a high scale (hence the term seesaw). In the Type-I [36] and Type-III [66] seesaw models, fermionic SU(2) singlets and triplets, respectively, are added to the lepton sector, and in the Type-II Seesaw Model [67,68], a Higgs triplet extends the scalar sector. In the Left-Right Symmetric Model (LRSM) [69–72], the gauge group of the Standard Model is extended such that there is an equivalent right-handed symmetry analogous to the $\mathrm{SU}(2)_L$ symmetry of the Standard Model, and contains phenomenological aspects of both the Type-I and Type-II seesaw models. In these models, the heavy neutrinos or additional scalar particles interact with the charged leptons to give rise to the LFV decays. A further model allowing for LFV τ decays is the Minimal Supersymmetric Standard Model (MSSM) [73–75]. Non-diagonal terms in the slepton mass matrices allow for interactions at the one-loop level which violate the lepton flavour symmetry.

In this study, we will focus on three of the aforementioned models: the Type-II Seesaw Model, the Left-Right Symmetric Model and the Minimal Supersymmetric Standard Model. We will use the constraints on the three decay modes $\tau^{\mp} \to \mu^{\pm}\mu^{\mp}\mu^{\mp}$, $\tau^{\mp} \to e^{\pm}\mu^{\mp}\mu^{\mp}$ and $\tau^{\mp} \to e^{\mp}\mu^{\mp}\mu^{\pm}$ from current and future lepton and hadron colliders to derive limits on the parameter spaces of the models. As well as constraining the models in the case of current limits, this will also show which areas of the parameter space the future experiments will be sensitive to and thus could be ruled out in the event of no discovery.

The remainder of this chapter is structured as follows: in Sec. 2.2 we will discuss the limits on the branching ratios that are derivable from current and future collider experiments. Then, in Secs. 2.3-2.5 we will describe each model in turn and the corresponding limits on their parameter spaces obtainable from the limits on the branching ratios. Finally, in Sec. 2.6 we conclude the chapter.

2.2 Limits from Collider Experiments

The decays $\tau^{\mp} \to \mu^{\pm}\mu^{\mp}\mu^{\mp}$, $\tau^{\mp} \to e^{\pm}\mu^{\mp}\mu^{\mp}$ and $\tau^{\mp} \to e^{\mp}\mu^{\mp}\mu^{\pm}$ have been constrained by current collider experiments. This has been possible at e^+e^- colliders for all three channels, and more recently at the LHC for the $\tau^{\mp} \to \mu^{\pm}\mu^{\mp}\mu^{\mp}$ decay, although the B-factories at the e^+e^- colliders are currently able to provide the best limits due to the cleaner nature of their collisions. The decays involving electrons have not currently been constrained at the LHC due to the larger fake rate of electrons.

We consider the limits on all three decays from the Belle [76] and BaBar [77] experiments at the KEK and PEP-II e^+e^- colliders, respectively, as well as limits on the $\tau^{\mp} \to \mu^{\pm}\mu^{\mp}\mu^{\mp}$ decay from LHCb [78] and ATLAS [79] at the LHC. The limits from each experiment represent the maximum possible branching ratios that each decay mode could have while still remaining undetected amongst the expected background events for the respective process. This depends on the total integrated luminosity of the collisions and the production cross section of the τ leptons, which together determine the total number of τ leptons produced. Furthermore, the limits are also affected by the selection efficiency of the events—stronger cuts are required in the cases where a larger background rejection is needed, such as at the LHC.

Upgrades to the KEK and LHC accelerators, as well as an upgraded Belle detector (Belle-II) and improvements to the LHCb and ATLAS detectors mean that with future data, much stronger limits will be obtainable. Projections can be made for these limits by scaling the integrated luminosities and production cross sections to their expected values, and assuming that no discovery is made. Conservative estimates for these can be made by assuming that the background scales in the same way as the potential signal events, but more optimistic limits can be calculated by assuming the background remains similar to current levels through better detector performance and more sophisticated analysis techniques. Further ahead in time, a future circular hadron (FCC-hh) or lepton (FCC-ee) collider [80] could provide even more stringent limits on these processes.

Experiment	Current	Projected
Belle	2.1×10^{-8}	$(4.7 - 10) \times 10^{-10}$
BaBar	3.3×10^{-8}	–
FCC-ee	–	$(5 - 10) \times 10^{-12}$
LHCb	4.6×10^{-8}	$(1.5 - 11) \times 10^{-9}$
ATLAS	3.8×10^{-7}	$(1.8 - 8.1) \times 10^{-9}$
FCC-hh	–	$(3 - 30) \times 10^{-10}$

Table 2.1: Current and projected 90% CL limits on the $\tau^{\mp} \to \mu^{\pm}\mu^{\mp}\mu^{\mp}$ branching ratio. The current limits from the LHC experiments utilise only the 8 TeV data, while the projected limits are based on the complete 13 TeV data sets of 3 ab^{-1} for ATLAS and 50 fb^{-1} for the LHCb from the high-luminosity run of the LHC.

	$\tau^{\mp} \to e^{\pm}\mu^{\mp}\mu^{\mp}$		$\tau^{\mp} \to e^{\mp}\mu^{\mp}\mu^{\pm}$	
Experiment	Current	Projected	Current	Projected
Belle	1.7×10^{-8}	$(3.4 - 5.1) \times 10^{-10}$	2.7×10^{-8}	$(5.9 - 12) \times 10^{-10}$
BaBar	2.6×10^{-8}	–	3.2×10^{-8}	–
FCC-ee	–	$(5 - 10) \times 10^{-12}$	–	$(5 - 10) \times 10^{-12}$

Table 2.2: Current and projected 90% CL limits on the $\tau^{\mp} \to e^{\pm}\mu^{\mp}\mu^{\mp}$ and $\tau^{\mp} \to e^{\mp}\mu^{\mp}\mu^{\pm}$ branching ratios.

In Table 2.1, we summarise the current and projected limits on the branching ratio of $\tau^{\mp} \to \mu^{\pm}\mu^{\mp}\mu^{\mp}$, while in Table 2.2 the current and projected limits for $\tau^{\mp} \to e^{\pm}\mu^{\mp}\mu^{\mp}$ and $\tau^{\mp} \to e^{\mp}\mu^{\mp}\mu^{\pm}$ are shown. We now proceed to use these limits to explore the parameter spaces of the models.

2.3 Type-II Seesaw Model

2.3.1 Overview of the Model

The Type-II Seesaw Model extends the scalar sector of the Standard Model and is used to generate masses for the light neutrinos. In addition to the Standard Model

Higgs doublet Φ, there is a further Higgs scalar triplet[2] Δ with a weak hypercharge $Y = 2$,

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} , \qquad \Delta = \begin{pmatrix} \frac{\Delta^+}{\sqrt{2}} & \Delta^{++} \\ \Delta^0 & -\frac{\Delta^+}{\sqrt{2}} \end{pmatrix} . \tag{2.3.1}$$

The new terms in the Lagrangian then consist of an additional kinetic term for the triplet Δ, a potential involving the scalars Δ and Φ, and Yukawa interactions between Δ and the leptons [81]. The kinetic term can be written as,

$$\mathcal{L}_{\text{Kinetic}}(\Delta) = \text{Tr}[(D_\mu \Delta)^\dagger (D^\mu \Delta)] , \tag{2.3.2}$$

where the covariant derivative $D_\mu \Delta$ is,

$$D_\mu \Delta = \partial_\mu \Delta + i\frac{g}{2}[\sigma^i W_\mu^i, \Delta] + ig' B_\mu \Delta , \tag{2.3.3}$$

with W_μ^i and B_μ the $\text{SU}(2)_L$ and $\text{U}(1)_Y$ gauge fields, with couplings g and g', respectively. The scalar potential has the following additional terms:

$$\begin{aligned} V(\Phi, \Delta) = {} & M_\Delta^2 \text{Tr}(\Delta^\dagger \Delta) + (\mu_\Delta \Phi^{\text{T}} i\sigma_2 \Delta^\dagger \Phi + \text{h.c.}) + \lambda_1 \Phi^\dagger \Phi \text{Tr}(\Delta^\dagger \Delta) \\ & + \lambda_2 [\text{Tr}(\Delta^\dagger \Delta)]^2 + \lambda_3 \text{Tr}[(\Delta^\dagger \Delta)^2] + \lambda_4 \Phi^\dagger \Delta \Delta^\dagger \Phi . \end{aligned} \tag{2.3.4}$$

Here, λ_{1-4} are dimensionless quartic couplings, the trace is over the 2×2 matrices, M_Δ is the mass scale of the Higgs triplet and μ_Δ is a lepton-number-violating parameter. Finally, there are the Yukawa interactions with the Standard Model lepton doublet[3],

$$\mathcal{L}_{\text{Yukawa}} = \sum_{i,j} (Y_\Delta)_{ij} \overline{L_{Li}^c} i\sigma_2 \Delta L_{Lj} + \text{h.c.} , \tag{2.3.5}$$

where Y_Δ is a 3×3 complex symmetric Yukawa matrix in flavour space in the gauge-interaction basis, and the sum runs over the flavour indices i and j.

The neutral components of the two Higgs multiplets, Φ^0 and Δ^0, each have a non-zero VEV, denoted as v_Φ and v_Δ, respectively. In the case where the mass scale of the Higgs triplet is much larger than that of the electroweak scale, the VEV takes

[2]Note that terms doublet and triplet here refer to their representation under the $\text{SU}(2)_L$ gauge group of the Standard Model.

[3]An equivalent term with quark doublets is forbidden by $\text{U}(1)_Y$ symmetry.

the form,

$$v_\Delta = \frac{\mu_\Delta v_\Phi^2}{\sqrt{2} M_\Delta^2} \ . \tag{2.3.6}$$

Under spontaneous symmetry breaking, the triplet VEV generates Majorana masses for the light neutrinos, $M_\nu = \sqrt{2} Y_\Delta v_\Delta$. This can be diagonalised by the PMNS mixing matrix $U_{\rm P}$, defined in Eq. (1.3.1), to the diagonal mass matrix $M_d = \text{diag}(m_1, m_2, m_3)$,

$$M_d = U_{\rm P}^* M_\nu U_{\rm P}^\dagger \ . \tag{2.3.7}$$

$U_{\rm P}$ takes the form,

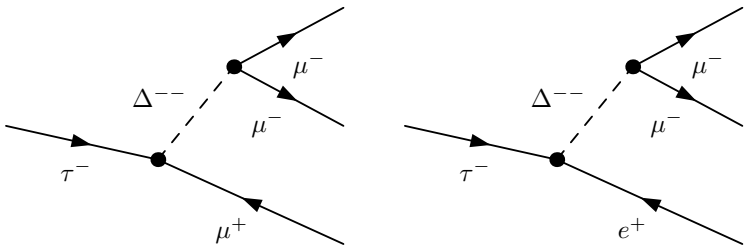
$$U_{\rm P} = \begin{pmatrix} c_{12}\, c_{13} & s_{12}\, c_{13} & s_{13}\, e^{-i\delta_{\rm CP}} \\ -c_{23}\, s_{12} - s_{23}\, s_{13}\, c_{12}\, e^{i\delta_{\rm CP}} & c_{23}\, c_{12} - s_{23}\, s_{13}\, s_{12}\, e^{i\delta_{\rm CP}} & s_{23}\, c_{13} \\ s_{23}\, s_{12} - c_{23}\, s_{13}\, c_{12}\, e^{i\delta_{\rm CP}} & -s_{23}\, c_{12} - c_{23}\, s_{13}\, s_{12}\, e^{i\delta_{\rm CP}} & c_{23}\, c_{13} \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha_1} & 0 \\ 0 & 0 & e^{i\alpha_2} \end{pmatrix} , \tag{2.3.8}$$

where $s_{ij} \equiv \sin\theta_{ij}$ and $c_{ij} \equiv \cos\theta_{ij}$ with θ_{ij} the neutrino oscillation parameters, $\delta_{\rm CP}$ is the CP-violating phase and $\alpha_{1,2}$ are the Majorana phases. This allows the Yukawa couplings in Eq. (2.3.5) to be expressed in terms of the neutrino masses and mixing parameters, for which best-fit values can be extracted from neutrino oscillation data [82, 83].

Furthermore, interactions between the doubly-charged Higgs field Δ^{++} and the charged leptons also arise. The terms which are relevant to us are,

$$\mathcal{L}_{\rm Yukawa}(\Delta^{++}) = (Y_\Delta)_{e\tau} \overline{e^c} \tau \Delta^{++} + (Y_\Delta)_{\mu\tau} \overline{\mu^c} \tau \Delta^{++} + (Y_\Delta)_{\mu\mu} \overline{\mu^c} \mu \Delta^{++} + \text{h.c.} \ . \tag{2.3.9}$$

It is the off-diagonal elements in the matrix Y_Δ which violate lepton flavour, and give rise to the LFV decay processes which we are interested in. These are shown by the Feynman diagrams in Fig 2.1 for the $\tau^\mp \to \mu^\pm \mu^\mp \mu^\mp$ and $\tau^\mp \to e^\pm \mu^\mp \mu^\mp$ decay modes. Note that the decay mode $\tau^\mp \to e^\mp \mu^\mp \mu^\pm$ involves two separate LFV Yukawa couplings, so it is generally suppressed by several orders of magnitude compared



(a) **(b)**

Figure 2.1: Feynman diagrams for the decays (a) $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and (b) $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ in the Type-II Seesaw Model.

with the other two modes, depending on the particular parameters. For the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, the partial width at tree level is [84],

$$\Gamma(\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp) = \frac{m_\tau^5}{192\pi^3}\left|\frac{(Y_\Delta)_{\tau\mu}(Y_\Delta)_{\mu\mu}}{m^2_{\Delta^{\pm\pm}}}\right|^2 . \tag{2.3.10}$$

Expressing the Yukawa couplings in terms of the neutrino mass matrix, this can be written as,

$$\Gamma(\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp) = \frac{m_\tau^5}{192\pi^3}\left|\frac{(M_\nu)_{\mu\tau}(M_\nu)_{\mu\mu}}{2v_\Delta^2 m^2_{\Delta^{\pm\pm}}}\right|^2 , \tag{2.3.11}$$

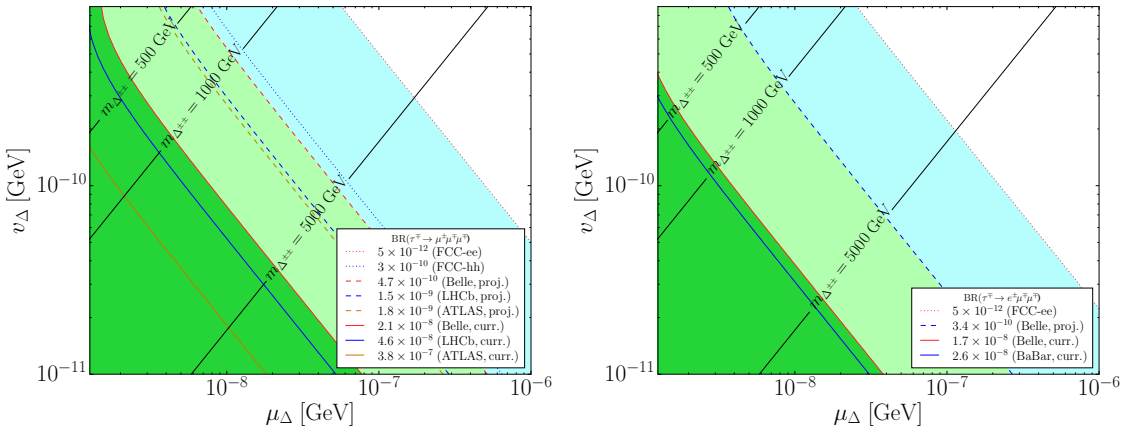
where the doubly-charged Higgs mass $m_{\Delta^{\pm\pm}}$ is given by,

$$m^2_{\Delta^{\pm\pm}} = M^2_\Delta - v^2_\Delta\lambda_3 - \frac{\lambda_4}{2}v^2_\Phi , \qquad M^2_\Delta = \frac{\mu_\Delta v^2_\Phi}{\sqrt{2}v_\Delta} . \tag{2.3.12}$$

Replacing $(M_\nu)_{\mu\tau}$ with $(M_\nu)_{e\tau}$ in Eq. (2.3.11) will give the rate for the $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ decay mode.

2.3.2 Limits on the Parameter Space

To calculate the branching ratios in the Type-II Seesaw Model, we use the program MadGraph5_aMC@NLO [85] with a model file generated by FeynRules [86]. These perform all the computation steps from the Lagrangian of the model to the matrix element calculation and integration over the phase space. This includes the PMNS mixing of the neutrino oscillation parameters into the values that enter the triplet Yukawa couplings. The output from MadGraph5_aMC@NLO is then the partial decay width, which we divide by the total measured decay width of the τ

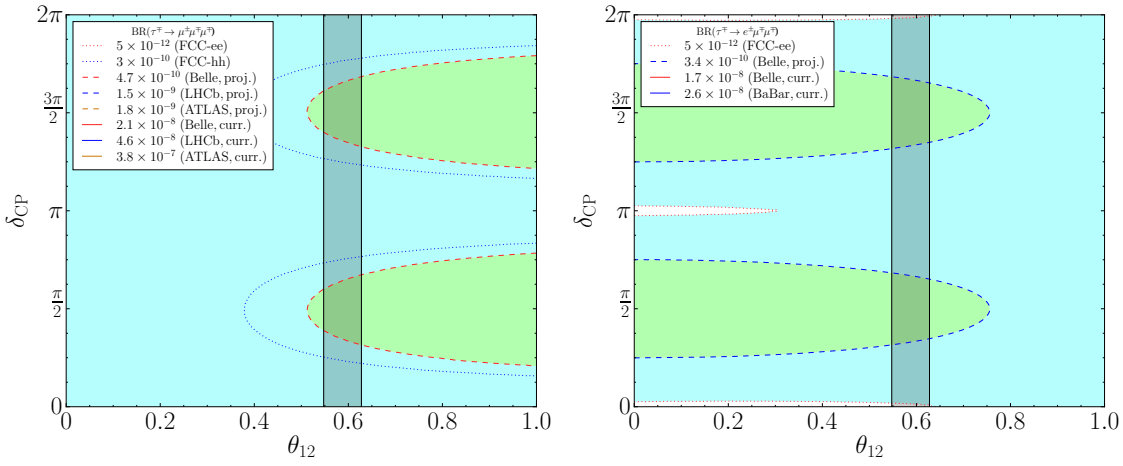


(a) **(b)**

Figure 2.2: Current and future branching-ratio limits in the parameter plane of μ_Δ and v_Δ for the Type-II Seesaw Model. (a) Shows the limits from the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, and (b) shows the limits from the decay $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$. The solid black lines represent constant values of the mass of the doubly-charged Higgs $\Delta^{\pm\pm}$.

lepton to obtain the branching ratio. For this we use the mean lifetime from the Particle Data Group [64], $(290.3 \pm 0.5) \times 10^{-15}$ s.

In Fig. 2.2 we show the current and future branching-ratio limits in the plane of the parameters μ_Δ and v_Δ. To calculate the branching ratios across the parameter plane, we perform the above calculation for each pair of μ_Δ and v_Δ values, keeping all other parameters fixed. The neutrino masses and oscillation parameters are held at their best-fit values [82,83], with the lightest neutrino mass at 0.1 eV and the CP-violating phase δ_{CP} set to zero. The solid black lines represent constant values of the doubly-charged Higgs mass across the parameter plane. The dark green regions show the parameter space restricted by the current limits, while the pale green regions show the exclusions that can be obtained by projections of current experiments. Furthermore, the pale blue regions show the restrictions from the future circular colliders FCC-hh and FCC-ee, while the white region is the part of the parameter space that will be allowed by the FCC-ee limit. For the projected limits we show the lower values of the limit ranges in Figs. 2.2a and 2.2b, corresponding to the best possible sensitivity for each experiment. All other parameter plots in this chapter will follow the same scheme for the region colours, and will use the lower values of



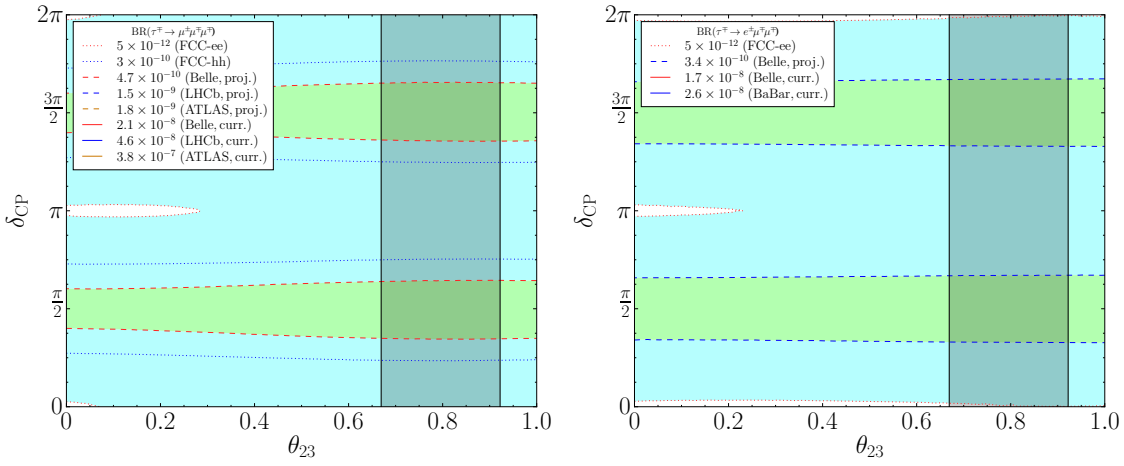
(a) **(b)**

Figure 2.3: Current and future branching-ratio limits in the parameter plane of the neutrino oscillation parameter θ_{12} and the CP-violating phase $\delta_{\rm CP}$ for the Type-II Seesaw Model. (a) Shows the limits from the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, and (b) shows the limits from the decay $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$. The dark shaded bands represent the allowed 3σ values of θ_{12}.

the limit ranges.

In Fig. 2.2, we show the variation across a range of small v_Δ values, $(10^{-11} - 10^{-9})$ GeV, that can naturally explain the small neutrino masses, $m_\nu \sim (0.01 - 0.1)$ eV, with $\mathcal{O}(1)$ coupling. For a moderate $v_\Delta = 10^{-10}$ GeV, and with the neutrino mass $m_\nu \sim 0.1$ eV, the present constraints on μ_Δ and the doubly-charged Higgs mass coming from Belle are $\mu_\Delta \geq 7.8 \times 10^{-9}$ GeV and $m_{\Delta^{\pm\pm}} \geq 1.8$ TeV, using the $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ decay. The future experiments Belle-II and FCC-ee could further constrain the doubly-charged Higgs mass up to $m_{\Delta^{\pm\pm}} \geq 4.6$ TeV and 14.5 TeV with $\mu_\Delta \geq 5.0 \times 10^{-8}$ GeV and 4.9×10^{-7} GeV, respectively.

In Fig. 2.3, we allow for a non-zero PMNS phase $\delta_{\rm CP}$ in the range $0 - 2\pi$, and investigate the effect of varying $\delta_{\rm CP}$ along with the neutrino oscillation parameter θ_{12} on the two decay processes, while fixing the other oscillation parameters to their best-fit values and the lightest neutrino mass to $m_1 = 0.1$ eV. The dark vertical shaded bands show the region of the parameter space allowed by the current 3σ limits on θ_{12}. For the $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ decay, we consider $\mu_\Delta = 1.5 \times 10^{-7}$ GeV and



(a) **(b)**

Figure 2.4: Current and future branching-ratio limits in the parameter plane of the neutrino oscillation parameter θ_{23} and the CP-violating phase $\delta_{\rm CP}$ for the Type-II Seesaw Model. (a) Shows the limits from the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, and (b) shows the limits from the decay $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$. The dark shaded bands represent the allowed 3σ values of θ_{23}.

$v_\Delta = 10^{-10}$ GeV, resulting in $m_{\Delta^{\pm\pm}} = 8.0$ TeV. In the case of $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$, we use an increased $\mu_\Delta = 2.5 \times 10^{-7}$ GeV and $v_\Delta = 10^{-10}$ GeV, giving $m_{\Delta^{\pm\pm}} = 10.3$ TeV. The Belle-II experiment could rule out $\delta_{\rm CP}$ in the ranges $1.1 - 2.0$ and $4.2 - 5.1$, while experiments at the FCC-ee could exclude all values of $\delta_{\rm CP}$ for these choices of μ_Δ and v_Δ. We find similar constraints when using the $\theta_{23} - \delta_{\rm CP}$ contours instead, and these are shown in Fig. 2.4.

We conclude this section by justifying our approach of only considering limits from the LFV τ lepton decays. The current bound on the branching ratio for $\mu^\mp \to e^\pm e^\mp e^\mp$ is $\text{BR}(\mu^\mp \to e^\pm e^\mp e^\mp) \leq 10^{-12}$ [65]. This tight bound imposes stronger limits in the plane of μ_Δ and v_Δ than those arising from the τ lepton decays, shown in Fig. 2.2. However, when varying the neutrino oscillation parameters and phases within experimental bounds, it is possible to suppress the branching ratio of $\mu^\mp \to e^\pm e^\mp e^\mp$ while leaving that of $\tau \to \ell\mu\mu$ essentially unchanged. We can consider the oscillation effects by defining the ratio,

$$\mathcal{R} = \frac{\text{BR}(\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp)}{\text{BR}(\mu^\mp \to \text{e}^\pm\text{e}^\mp\text{e}^\mp)} \propto \frac{|(M_\nu)_{\mu\tau}(M_\nu)_{\mu\mu}|^2}{|(M_\nu)_{\mu e}(M_\nu)_{ee}|^2} \,, \tag{2.3.13}$$

and varying all the oscillation parameters and phases within their allowed 3σ ranges. For quasi-degenerate neutrino masses with an inverted-hierarchy spectrum, and with $m_3 = 0.1$ eV, we find that $\mathcal{R}$ can be as large as 10^6, due to cancellations in the neutrino mass matrix M_ν, which is calculated via Eq. (2.3.7). Such regions of the parameter space suppress the branching ratio of $\mu^\mp \to e^\pm e^\mp e^\mp$ enough so that the strongest limits on μ_Δ and v_Δ arise from the LFV τ lepton decays, which can remain largely unaffected. This was discussed in detail in Ref. [87], for both hierarchical and quasi-degenerate neutrino masses, where branching ratios of as large as 10^{-8} for $\tau^\mp \to \mu^\pm \mu^\mp \mu^\mp$ were obtained, while still being consistent with the other bounds. Therefore, we focus on the bounds derived from the LFV τ lepton decays, independent of other constraints, and Fig. 2.2 qualitatively demonstrates the constraints that can be obtained in regions where the LFV τ lepton decays provide the dominant source of all LFV decays.

2.4 Left-Right Symmetric Model

2.4.1 Overview of the Model

We now consider a more phenomenologically extensive model. The Left-Right Symmetric Model is constructed by extending the gauge sector of the Standard Model to put the left and right symmetries on an equal footing—at least at the level of the Lagrangian before any spontaneous symmetry breaking. The simplest allowed model with left-right symmetry replaces the $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry of the Standard Model with an $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{SU}(2)_R \times \mathrm{U}(1)_{B-L}$ gauge symmetry[4]. The fermion content of the Standard Model is then extended to include right-handed neutrinos, and their representations under the gauge symmetry are,

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \in (\mathbf{3}, \mathbf{2}, \mathbf{1}, 1/3) \ , \qquad Q_R = \begin{pmatrix} u_R \\ d_R \end{pmatrix} \in (\mathbf{3}, \mathbf{1}, \mathbf{2}, 1/3) \ ,$$

[4]Note that here $\mathrm{U}(1)_{B-L}$ is a local symmetry of the model and not a global symmetry as it is in the Standard Model. The generation of neutrino masses then spontaneously breaks this symmetry.

$$L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \in (\mathbf{1}, \mathbf{2}, \mathbf{1}, -1) \ , \qquad L_R = \begin{pmatrix} N_R \\ e_R \end{pmatrix} \in (\mathbf{1}, \mathbf{1}, \mathbf{2}, -1) \ . \tag{2.4.1}$$

Thus there is a quark and lepton doublet corresponding to each of the SU(2) symmetries, and there is a copy of each of these doublets for each generation of fermions. The electromagnetic charge associated to each fermion field can be calculated from $Q = T_{3L} + T_{3R} + (B-L)/2$, where $T_{3L(R)}$ are the third components of the $\mathrm{SU}(2)_{L(R)}$ isospin. The scalar sector consists of a bidoublet Φ and two triplets Δ_L and Δ_R:

$$\Phi = \begin{pmatrix} \phi_1^0 & \phi_1^+ \\ \phi_2^- & \phi_2^0 \end{pmatrix} \in (\mathbf{1}, \mathbf{2}, \mathbf{2}, 0) \ ,$$

$$\Delta_L = \begin{pmatrix} \frac{\delta_L^+}{\sqrt{2}} & \delta_L^{++} \\ \delta_L^0 & -\frac{\delta_L^+}{\sqrt{2}} \end{pmatrix} \in (\mathbf{1}, \mathbf{3}, \mathbf{1}, 2) \ , \qquad \Delta_R = \begin{pmatrix} \frac{\delta_R^+}{\sqrt{2}} & \delta_R^{++} \\ \delta_R^0 & -\frac{\delta_R^+}{\sqrt{2}} \end{pmatrix} \in (\mathbf{1}, \mathbf{1}, \mathbf{3}, 2) \ . \tag{2.4.2}$$

The neutral components of each of these three Higgs multiplets take non-zero VEVs. These are v_L for Δ_L, v_R for Δ_R and $\mathrm{diag}(\kappa_1, \kappa_2)/\sqrt{2}$ for the bidoublet Φ. At a high scale, the right-handed triplet VEV v_R spontaneously breaks the $\mathrm{SU}(2)_R \times \mathrm{U}(1)_{B-L}$ symmetry down to the $\mathrm{U}(1)_Y$ symmetry of the Standard Model. This generates masses for the new gauge bosons $W_R^\pm$ and Z', in addition to masses for the right-handed neutrinos. The VEV of the bidoublet then breaks the remaining $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry down to $\mathrm{U}(1)_Q$, similar to the Higgs mechanism in the Standard Model. This generates masses for the gauge bosons $W_L^\pm$ and Z, and the fermions of the Standard Model. For phenomenological reasons, it is necessary for the VEV of the left-handed doublet to be very small [88], which leads to the hierarchy $v_L \ll \kappa_{1,2} \ll v_R$.

The complete Lagrangian involves many terms and is comprehensively described in Ref. [72]. Here, we will outline the parts which are important for the LFV decay processes.

Lepton Sector

In the lepton sector, the Yukawa Lagrangian takes the form,

$$\mathcal{L}_{\text{Yukawa}} = \sum_{i,j} \Big[h_{ij} \overline{L}_{Li} \Phi L_{Rj} + \tilde{h}_{ij} \overline{L}_{Li} \tilde{\Phi} L_{Rj} + (f_L)_{ij} \overline{L^c_{Li}} i\sigma_2 \Delta_L L_{Lj} + (f_R)_{ij} \overline{L^c_{Ri}} i\sigma_2 \Delta_R L_{Rj} \Big] + \text{h.c.} \,, \tag{2.4.3}$$

where the conjugate bidoublet is $\tilde{\Phi} = \sigma_2 \Phi^* \sigma_2$, and the sum runs over the flavour indices. The matrices $h, \tilde{h}, f_L$ and f_R are Yukawa couplings. After spontaneous breaking of the left-right and electroweak symmetries, the following 3×3 neutrino mass matrices are generated:

$$M_L = \sqrt{2} v_L f_L \,, \qquad M_R = \sqrt{2} v_R f_R \,, \qquad M_D = \frac{\kappa_1 h + \kappa_2 \tilde{h}}{\sqrt{2}} \,. \tag{2.4.4}$$

These correspond to left-handed Majorana, right-handed Majorana and Dirac masses, respectively. The full 6×6 neutrino mass matrix is then,

$$\mathcal{M}_\nu = \begin{pmatrix} M_L & M_D \\ M_D^{\mathrm{T}} & M_R \end{pmatrix} \,, \tag{2.4.5}$$

which can be diagonalised into the mass basis with the unitary matrix $\mathcal{V}$,

$$\mathcal{V}^{\mathrm{T}} \mathcal{M}_\nu \mathcal{V} = \begin{pmatrix} \widetilde{M}_\nu & \mathbf{0} \\ \mathbf{0} & \widetilde{M}_R \end{pmatrix} \,, \tag{2.4.6}$$

where $\widetilde{M}_\nu = \text{diag}(m_1, m_2, m_3)$ and $\widetilde{M}_R = \text{diag}(m_{N_4}, m_{N_5}, m_{N_6})$. The mixing matrix $\mathcal{V}$ can be expressed as,

$$\mathcal{V} = \begin{pmatrix} U & V \\ X & Y \end{pmatrix} \,. \tag{2.4.7}$$

The upper-left 3×3 matrix U is approximately equal[5] to the PMNS mixing matrix U_{P}, which only involves the left-handed neutrinos. The other entries in the matrix arise due to the additional right-handed neutrinos. Similarly, the matrix Y is approximately equal to the matrix Y_R which diagonalises the right-handed neutrino

[5]The differences between the mixing matrices U and U_{P} are $\mathcal{O}(M_D M_R^{-1})$ [89].

mass matrix.

As was the case in the Type-II Seesaw Model, the Yukawa terms also induce LFV interactions through the couplings with the doubly-charged Higgs. The relevant terms in the Lagrangian are,

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}}(\delta^{++}) = (f_L)_{e\tau}\overline{e_L^c}\delta_L^{++}\tau_L + (f_L)_{\mu\tau}\overline{\mu_L^c}\delta_L^{++}\tau_L + (f_L)_{\mu\mu}\overline{\mu_L^c}\delta_L^{++}\mu_L \\ + (f_R)_{e\tau}\overline{e_R^c}\delta_R^{++}\tau_R + (f_R)_{\mu\tau}\overline{\mu_R^c}\delta_R^{++}\tau_R + (f_R)_{\mu\mu}\overline{\mu_R^c}\delta_R^{++}\mu_R + \text{h.c.}\ .\end{aligned} \tag{2.4.8}$$

Note that if discrete parity or charge conjugation are imposed as symmetries, then $f_L = f_R$ or $f_L = f_R^*$ with a Hermitian or symmetric M_D, respectively.

Scalar Sector

As we described above, the scalar sector contains the Higgs bidoublet Φ and two triplets Δ_L and Δ_R. The full scalar potential has the form,

$$\begin{aligned}V(\Phi, \Delta_L, \Delta_R) = \\ &- \mu_1^2\text{Tr}(\Phi^\dagger\Phi) - \mu_2^2\text{Tr}(\Phi^\dagger\tilde{\Phi} + \tilde{\Phi}^\dagger\Phi) - \mu_3^2\text{Tr}(\Delta_L^\dagger\Delta_L + \Delta_R^\dagger\Delta_R) \\ &+ \lambda_1\left[\text{Tr}(\Phi^\dagger\Phi)\right]^2 + \lambda_2\left[\text{Tr}(\Phi^\dagger\tilde{\Phi})\right]^2 + \lambda_2\left[\text{Tr}(\tilde{\Phi}^\dagger\Phi)\right]^2 \\ &+ \lambda_3\text{Tr}(\Phi^\dagger\tilde{\Phi})\text{Tr}(\tilde{\Phi}^\dagger\Phi) + \lambda_4\text{Tr}(\Phi^\dagger\Phi)\text{Tr}(\Phi^\dagger\tilde{\Phi} + \tilde{\Phi}^\dagger\Phi) \\ &+ \rho_1\left[\text{Tr}(\Delta_L^\dagger\Delta_L)\right]^2 + \rho_1\left[\text{Tr}(\Delta_R^\dagger\Delta_R)\right]^2 + \rho_3\text{Tr}(\Delta_L^\dagger\Delta_L)\text{Tr}(\Delta_R^\dagger\Delta_R) \\ &+ \rho_2\text{Tr}(\Delta_L\Delta_L)\text{Tr}(\Delta_L^\dagger\Delta_L^\dagger) + \rho_2\text{Tr}(\Delta_R\Delta_R)\text{Tr}(\Delta_R^\dagger\Delta_R^\dagger) \\ &+ \rho_4\text{Tr}(\Delta_L\Delta_L)\text{Tr}(\Delta_R^\dagger\Delta_R^\dagger) + \rho_4\text{Tr}(\Delta_L^\dagger\Delta_L^\dagger)\text{Tr}(\Delta_R\Delta_R) \\ &+ \alpha_1\text{Tr}(\Phi^\dagger\Phi)\text{Tr}(\Delta_L^\dagger\Delta_L + \Delta_R^\dagger\Delta_R) + \alpha_3\text{Tr}(\Phi\Phi^\dagger\Delta_L\Delta_L^\dagger + \Phi^\dagger\Phi\Delta_R\Delta_R^\dagger) \\ &+ \left[\alpha_2 e^{i\delta_2}\text{Tr}(\Phi^\dagger\tilde{\Phi})\text{Tr}(\Delta_L^\dagger\Delta_L) + \alpha_2 e^{i\delta_2}\text{Tr}(\tilde{\Phi}^\dagger\Phi)\text{Tr}(\Delta_R^\dagger\Delta_R) + \text{h.c.}\right] \\ &+ \beta_1\text{Tr}(\Phi^\dagger\Delta_L^\dagger\Phi\Delta_R + \Delta_R^\dagger\Phi^\dagger\Delta_L\Phi) + \beta_2\text{Tr}(\Phi^\dagger\Delta_L^\dagger\tilde{\Phi}\Delta_R + \Delta_R^\dagger\tilde{\Phi}^\dagger\Delta_L\Phi) \\ &+ \beta_3\text{Tr}(\tilde{\Phi}^\dagger\Delta_L^\dagger\Phi\Delta_R + \Delta_R^\dagger\Phi^\dagger\Delta_L\tilde{\Phi})\ .\end{aligned} \tag{2.4.9}$$

The large number of terms arise due to the many possible gauge-invariant combinations of the three scalar multiplets. Before spontaneous symmetry breaking, there

are 20 degrees of freedom in these complex multiplets. Six of these degrees of freedom give masses to the $W_R^{\pm}$, Z', $W^{\pm}$ and Z gauge bosons. This leaves 14 physics Higgs states, denoted as h, $H^0_{1,2,3}$, $A^0_{1,2}$, $H_1^{\pm}$, $H_2^{\pm}$, $\delta_L^{\pm\pm}$, and $\delta_R^{\pm\pm}$. The masses of these are:

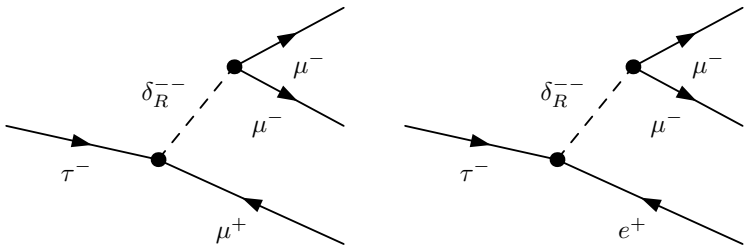
$$
\begin{aligned}
& m_h^2 \sim (125\ \text{GeV})^2 \simeq 2\kappa_+^2 \left(\lambda_1 + 4\frac{\kappa_1^2\kappa_2^2}{\kappa_+^4}(2\lambda_2+\lambda_3) + 4\lambda_4 \frac{\kappa_1\kappa_2}{\kappa_+^2} \right) , \\
& M^2_{H_1^0} = M^2_{A_1^0} \simeq \alpha_3 \frac{v_R^2}{2}\frac{\kappa_+^2}{\kappa_-^2} , \qquad M^2_{H_3^0} = M^2_{A_2^0} \simeq (\rho_3 - 2\rho_1)\frac{v_R^2}{2} , \quad M^2_{H_2^0} \simeq 2\rho_1 v_R^2 , \\
& M^2_{H_1^{\pm}} \simeq (\rho_3 - 2\rho_1)\frac{v_R^2}{2} + \alpha_3 \frac{\kappa_-^2}{4} , \quad M^2_{\delta_L^{\pm\pm}} \simeq (\rho_3 - 2\rho_1)\frac{v_R^2}{2} + \alpha_3 \frac{\kappa_-^2}{2} , \\
& M^2_{H_2^{\pm}} \simeq \alpha_3 \frac{v_R^2}{2}\frac{\kappa_+^2}{\kappa_-^2} + \alpha_3 \frac{\kappa_-^2}{4} , \qquad M^2_{\delta_R^{\pm\pm}} \simeq 2\rho_2 v_R^2 + \alpha_3 \frac{\kappa_-^2}{2} ,
\end{aligned}
\tag{2.4.10}
$$

where $\kappa_{\pm}^2 \equiv \kappa_1^2 \pm \kappa_2^2$ and $\kappa_+ \sim 246$ GeV.

Note that the scalar states H_1^0 and H_3^0 interact with both the up and down quark sectors and hence mediate the $\Delta F = 2$ flavour transitions in the neutral K and B mesons [90,91]. To avoid the flavour-changing neutral Higgs (FCNH) constraints, the neutral Higgs states H_1^0, H_3^0 and $A_{1,2}^0$ are required to be heavier than 20 TeV [90,91]. We also consider the other neutral Higgs state H_2^0 to be heavy in order to be in agreement with the heavy Higgs searches at the LHC. In the Higgs spectrum, we consider the case where the right-handed doubly-charged Higgs boson is somewhat lighter than the other additional Higgs states and hence significantly contributes to the LFV processes. We consider the following two benchmark scenarios—BP1 and BP2—with a lower and a higher symmetry-breaking scale v_R, respectively:

- BP1: $\alpha_3 = 18.88$, $v_R = 8.68$ TeV ,
- BP2: $\alpha_3 = 1.00$, $v_R = 30.00$ TeV .

For both of the benchmark scenarios, we consider the right-handed mixing matrix Y_R to be non-diagonal with unit entries everywhere. In order for v_R to be less than 10 TeV, the FCNH constraints on the neutral Higgs bosons necessarily require α_3 to be large ($\alpha_3 \sim 8$). Conversely, when α_3 is well within the perturbative limit, the FCNH constraints on the neutral Higgs bosons demand a large value of the



(a) **(b)**

Figure 2.5: Feynman diagrams for the decays (a) $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and (b) $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ in the LRSM.

symmetry-breaking scale v_R [91]. In our subsequent analysis we consider the two possibilities—both the large and the natural α_3—and show the restrictions that can be obtained on the heavy neutrino masses and the ρ_2 parameter.

LFV Decay Amplitudes

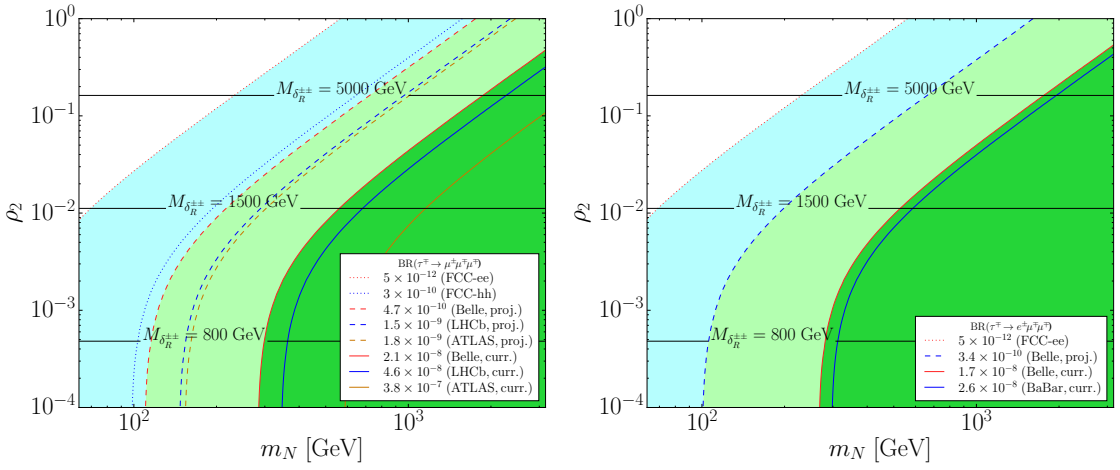
The two doubly charged Higgs states $\delta_L^{\pm\pm}$ and $\delta_R^{\pm\pm}$ mediate the τ LFV processes at tree level. The amplitude for the LFV process $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ is proportional to the coefficient $C_{\tau\mu\mu\mu}$, which is defined as,

$$C_{\tau\mu\mu\mu} = \frac{(f_L)_{\mu\tau}(f_L)_{\mu\mu}}{M^2_{\delta_L^{\pm\pm}}} + \frac{(f_R)_{\mu\tau}(f_R)_{\mu\mu}}{M^2_{\delta_R^{\pm\pm}}} \ . \tag{2.4.11}$$

Since in our case the chosen parameter $M_{\delta_L^{\pm\pm}}$ is much heavier than $M_{\delta_R^{\pm\pm}}$, the dominant contribution arises due to $\delta_R^{\pm\pm}$:

$$C_{\tau\mu\mu\mu} \simeq \frac{(f_R)_{\mu\tau}(f_R)_{\mu\mu}}{M^2_{\delta_R^{\pm\pm}}} = \frac{(M_R)_{\mu\tau}(M_R)_{\mu\mu}}{2v_R^2 M^2_{\delta_R^{\pm\pm}}} = \frac{(Y_R^*\widetilde{M}_R Y_R^\dagger)_{\mu\tau}(Y_R^*\widetilde{M}_R Y_R^\dagger)_{\mu\mu}}{2v_R^2(2\rho_2 v_R^2 + \alpha_3\frac{k^2}{2})} \ . \tag{2.4.12}$$

The amplitude for the LFV process $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ can then be obtained by replacing the $(M_R)_{\mu\tau}$ element in Eq. (2.4.12) with the $(M_R)_{e\tau}$ element. The Feynman diagrams for this process are shown in Fig. 2.5. As in the Type-II Seesaw Model, the decay process $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$ involves two LFV Yukawa couplings and thus is suppressed.



(a) **(b)**

Figure 2.6: Current and future branching-ratio limits in the parameter plane of the right-handed neutrino masses m_N and the parameter ρ_2 for the LRSM for the benchmark scenario BP1. (a) Shows the limits from the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, and (b) shows the limits from the decay $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$. The solid black lines represent constant values of the mass of the doubly-charged Higgs $\delta_R^{\pm\pm}$.

2.4.2 Limits on the Parameter Space

To calculate the branching ratios for the LFV decay processes in the LRSM, we again use the program MadGraph5_aMC@NLO with a model file generated by FeynRules. In Fig. 2.6, corresponding to BP1, we show the branching-ratio limits for the case where the three right-handed neutrino masses are all equal and denoted by m_N, and are varied along with the parameter ρ_2. In Fig. 2.7, we show the equivalent plots for BP2. For BP1, the current limit from Belle imposes the constraint on the right-handed neutrino masses $m_N \leq 290$ GeV for the doubly-charged Higgs mass $M_{\delta_R^{\pm\pm}} = 420$ GeV for the $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ decays. This $M_{\delta_R^{\pm\pm}}$ mass is the lower limit set by the 13 TeV ATLAS search for the right-handed triplet [92]. For BP2, with a higher value of the symmetry-breaking scale v_R, the mass limits are much higher: $m_N \lesssim 10$ TeV for the doubly-charged Higgs mass $M_{\delta_R^{\pm\pm}} = 8$ TeV. For both of the scenarios, a future circular collider will be able to probe much smaller values of m_N.

In Fig. 2.8, we consider the scenario of non-degenerate right-handed neutrino masses

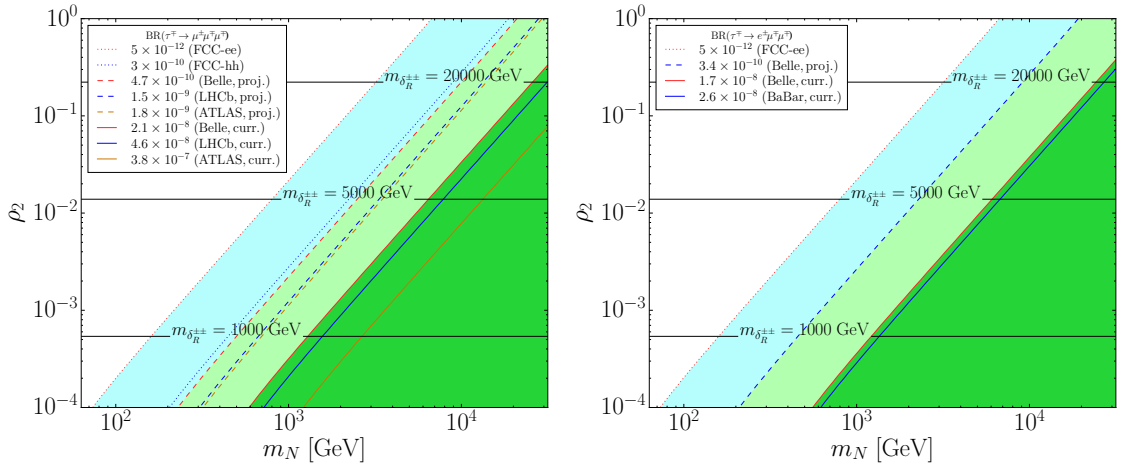
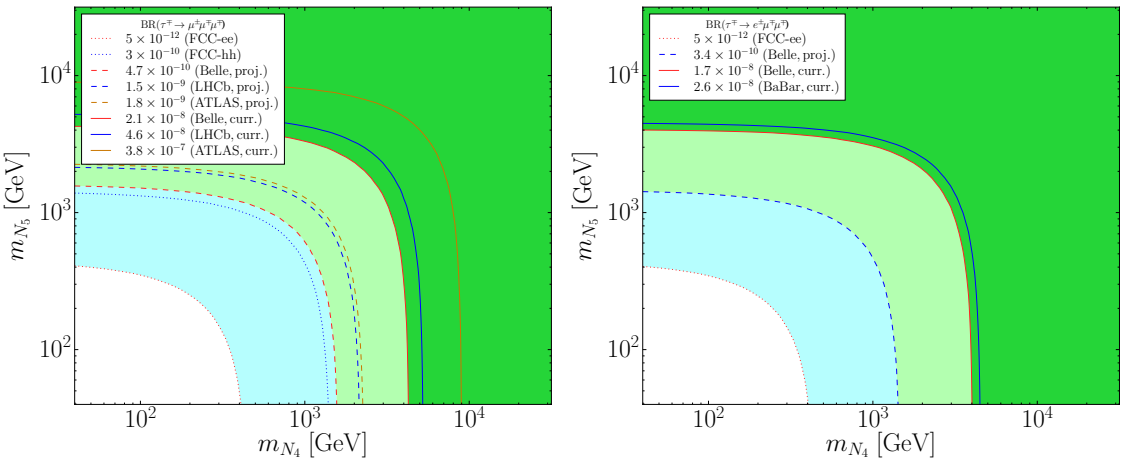


Figure 2.7: Current and future branching-ratio limits in the parameter plane of the right-handed neutrino masses m_N and the parameter ρ_2 for the LRSM for the benchmark scenario BP2. (a) Shows the limits from the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, and (b) shows the limits from the decay $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$. The solid black lines represent constant values of the mass of the doubly-charged Higgs $\delta_R^{\pm\pm}$.

$m_{N_{4,5,6}}$. We show the branching-ratio limits in the plane of the right-handed neutrino masses m_{N_4} and m_{N_5} for the case of BP1, while fixing $m_{N_6} = 100$ GeV and the doubly-charged Higgs mass to $M_{\delta_R^{\pm\pm}} = 4$ TeV. The present stringent limit from Belle constrains both of the m_{N_4} and m_{N_5} masses to be smaller than ~ 1 TeV, while the FCC-ee could probe these masses down to ~ 100 GeV.

In our analysis, we considered the possibilities of both a lower and a higher symmetry-breaking scale v_R. While a lower symmetry-breaking scale and a right-handed gauge boson with mass $m_{W_R} \lesssim (5-6)$ TeV is within the reach of the 13 TeV LHC, a higher symmetry-breaking scale, such as that in BP2, along with a much heavier W_R could be probed at a 100 TeV future circular collider [93, 94]. In Refs. [95] and [94], the impact of RG evolution of the quartic couplings on the discovery of W_R and the Higgs states has been discussed and bounds on the quartic couplings have been derived by analysing stability conditions. A lower symmetry-breaking scale with a W_R accessible at the 13 TeV LHC implies a larger ρ_2 (for a cut-off scale $10m_{W_R}$ with $m_{W_R} = 6$ TeV, then $\rho_2 \geq 0.35$ [94]) and hence a larger $M_{\delta_R^{\pm\pm}}$. This cannot be directly produced at the LHC, but instead can be tested through indirect detection. Conversely, for



(a) **(b)**

Figure 2.8: Current and future branching-ratio limits in the parameter plane of the right-handed neutrino masses m_{N_4} and m_{N_5} for the LRSM. (a) Shows the limits from the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, and (b) shows the limits from the decay $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$.

a larger symmetry-breaking scale with $m_{W_R} \sim (20 - 30)$ TeV the bounds on ρ_2 are relaxed. In our discussion, we do not specify any particular mass of the other Higgs states or the cut-off scale of the theory. Instead, we independently analyse the implication of the branching-ratio limits for the LFV processes $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ on the relevant model parameter ρ_2 and the doubly-charged Higgs mass $M_{\delta_R^{\pm\pm}}$.

2.5 Minimal Supersymmetric Standard Model

2.5.1 Overview of the Model

The final model that we consider is the Minimal Supersymmetric Standard Model, which is the simplest extension to the Standard Model that can include supersymmetry. This is a symmetry between bosons and fermions, which is the only non-trivial extension of the Poincaré group allowed by the Coleman-Mandula theorem [96], and was proposed as a solution to the hierarchy problem. The bosons and fermions exist

in supermultiplets, where they have a corresponding partner particle. If supersymmetry is unbroken, the supersymmetric partners necessarily have the same masses as their corresponding Standard Model particles. However, we do not observe such partners with the same mass, thus supersymmetry must be broken to ensure their masses are much larger. One way to break supersymmetry is to add terms to the Lagrangian by hand. This is called 'soft supersymmetry breaking', and the terms are required to be gauge invariant and protect the cancellation of quadratic singularities, which is a feature of supersymmetry. The allowed terms are scalar mass terms, gaugino mass terms and bilinear and trilinear couplings between the scalars.

In the MSSM, the soft supersymmetry breaking parameters in the slepton sector are a generic source of LFV. Without assuming a specific supersymmetry-breaking mechanism that ensures a suppression of the off-diagonal terms in the slepton mass matrix, their presence can induce a misalignment in flavour space between the lepton and slepton mass matrices, which cannot be rotated away. The non-diagonal Hermitian 6×6 slepton mass matrix receives contributions from D, F, A and M terms [75], where the latter two can induce mixing between different slepton generations. In the electroweak interaction basis $(\tilde{e}_L, \tilde{\mu}_L, \tilde{\tau}_L, \tilde{e}_R, \tilde{\mu}_R, \tilde{\tau}_R)$, the slepton mass matrix has the form[6] [97],

$$\mathcal{M}^2_{\tilde{\ell}} = \begin{pmatrix} M^2_{\tilde{\ell}\,LL} & M^2_{\tilde{\ell}\,LR} \\ M^{2\,\dagger}_{\tilde{\ell}\,LR} & M^2_{\tilde{\ell}\,RR} \end{pmatrix} , \tag{2.5.1}$$

where each of the blocks $M^2_{\tilde{\ell}\,LL}$, $M^2_{\tilde{\ell}\,RR}$, $M^2_{\tilde{\ell}\,LR}$ and $M^{2\,\dagger}_{\tilde{\ell}\,LR}$ are 3×3 matrices:

$$\begin{aligned} M^2_{\tilde{\ell}\,LL\,ij} &= m^2_{\tilde{L}\,ij} + \left(m^2_{\ell_i} + (-\frac{1}{2} + \sin^2\theta_{\mathrm{W}}) m^2_Z \cos 2\beta\right) \delta_{ij} \ , \\ M^2_{\tilde{\ell}\,RR\,ij} &= m^2_{\tilde{E}\,ij} + \left(m^2_{\ell_i} - \sin^2\theta_{\mathrm{W}} m^2_Z \cos 2\beta\right) \delta_{ij} \ , \\ M^2_{\tilde{\ell}\,LR\,ij} &= v_1 \mathcal{A}^{\ell}_{ij} - \delta_{ij} m_{\ell_i} \mu \tan\beta \ . \end{aligned} \tag{2.5.2}$$

In these equations the indices $i, j \in \{1, 2, 3\}$ denote the three generations, m_{ℓ_i} are

[6] L and R refer here to the slepton partners corresponding to the left- and right-handed leptonic degrees of freedom, respectively.

the lepton masses, $\theta_{\rm W}$ is the weak mixing angle, m_Z is the Z boson mass, μ is the Higgsino mass term, and $\tan\beta = v_2/v_1$ with $v_1 = \langle H_1 \rangle$ and $v_2 = \langle H_2 \rangle$ being the two VEVs of the corresponding SU(2) Higgs doublets. Here, δ_{ij} is the Kronecker delta symbol. Furthermore, $\mathcal{A}^{\ell}_{ij}$ are the slepton trilinear couplings. The flavour-violating terms in the LL and RR mixing matrices correspond to off-diagonal terms in the soft masses $m^2_{\tilde{L}\,ij}$ and $m^2_{\tilde{E}\,ij}$, respectively.

Within the MSSM, the sneutrino mass matrix has a one-block 3×3 form denoted as $\mathcal{M}^2_{\tilde{\nu}}$, where in the electroweak basis $(\tilde{\nu}_{eL}, \tilde{\nu}_{\mu L}, \tilde{\nu}_{\tau L})$,

$$\mathcal{M}^2_{\tilde{\nu}} = M^2_{\tilde{\nu}\,LL} \ , \qquad M^2_{\tilde{\nu}\,LL\,ij} = m^2_{\tilde{L}\,ij} + \left(\frac{1}{2} m_Z^2 \cos 2\beta\right) \delta_{ij} \ . \tag{2.5.3}$$

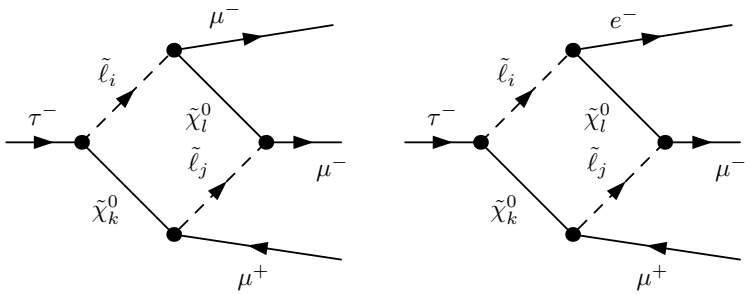
To parameterise the off-diagonal entries, we introduce the dimensionless real parameters δ^{AB}_{ij}. Written explicitly, the mixing matrices are then,

$$\begin{aligned}
m^2_{\tilde{L}} &= \begin{pmatrix} m^2_{\tilde{L}_1} & \delta^{LL}_{12} m_{\tilde{L}_1} m_{\tilde{L}_2} & \delta^{LL}_{13} m_{\tilde{L}_1} m_{\tilde{L}_3} \\ \delta^{LL}_{21} m_{\tilde{L}_2} m_{\tilde{L}_1} & m^2_{\tilde{L}_2} & \delta^{LL}_{23} m_{\tilde{L}_2} m_{\tilde{L}_3} \\ \delta^{LL}_{31} m_{\tilde{L}_3} m_{\tilde{L}_1} & \delta^{LL}_{32} m_{\tilde{L}_3} m_{\tilde{L}_2} & m^2_{\tilde{L}_3} \end{pmatrix} \ , \\
m^2_{\tilde{E}} &= \begin{pmatrix} m^2_{\tilde{E}_1} & \delta^{RR}_{12} m_{\tilde{E}_1} m_{\tilde{E}_2} & \delta^{RR}_{13} m_{\tilde{E}_1} m_{\tilde{E}_3} \\ \delta^{RR}_{21} m_{\tilde{E}_2} m_{\tilde{E}_1} & m^2_{\tilde{E}_2} & \delta^{RR}_{23} m_{\tilde{E}_2} m_{\tilde{E}_3} \\ \delta^{RR}_{31} m_{\tilde{E}_3} m_{\tilde{E}_1} & \delta^{RR}_{32} m_{\tilde{E}_3} m_{\tilde{E}_2} & m^2_{\tilde{E}_3} \end{pmatrix} \ , \\
v_1 \mathcal{A}^{\ell} &= \begin{pmatrix} m_e A_e & \delta^{LR}_{12} m_{\tilde{L}_1} m_{\tilde{E}_2} & \delta^{LR}_{13} m_{\tilde{L}_1} m_{\tilde{E}_3} \\ \delta^{LR}_{21} m_{\tilde{L}_2} m_{\tilde{E}_1} & m_\mu A_\mu & \delta^{LR}_{23} m_{\tilde{L}_2} m_{\tilde{E}_3} \\ \delta^{LR}_{31} m_{\tilde{L}_3} m_{\tilde{E}_1} & \delta^{LR}_{32} m_{\tilde{L}_3} m_{\tilde{E}_2} & m_\tau A_\tau \end{pmatrix} \ ,
\end{aligned} \tag{2.5.4}$$

where $m_{\tilde{L}_i}$ and $m_{\tilde{E}_i}$ are the soft mass scales. We further assume that $|\delta^{AB}_{ij}| \leq 1$, and the hermiticity of $\mathcal{M}^2_{\tilde{\ell}}$ implies $\delta^{AB}_{ij} = \delta^{BA}_{ji}$.

After rotating the sleptons and sneutrinos into their mass eigenstates,

$$\begin{aligned}
\mathrm{diag}\{m^2_{\tilde{\ell}_1}, m^2_{\tilde{\ell}_2}, m^2_{\tilde{\ell}_3}, m^2_{\tilde{\ell}_4}, m^2_{\tilde{\ell}_5}, m^2_{\tilde{\ell}_6}\} &= R^{\tilde{\ell}} \mathcal{M}^2_{\tilde{\ell}} R^{\tilde{\ell}\dagger} \ , \\
\mathrm{diag}\{m^2_{\tilde{\nu}_1}, m^2_{\tilde{\nu}_2}, m^2_{\tilde{\nu}_3}\} &= R^{\tilde{\nu}} \mathcal{M}^2_{\tilde{\nu}} R^{\tilde{\nu}\dagger} \ ,
\end{aligned} \tag{2.5.5}$$



(a) **(b)**

Figure 2.9: Feynman diagrams for the decays (a) $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and (b) $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$ in the MSSM.

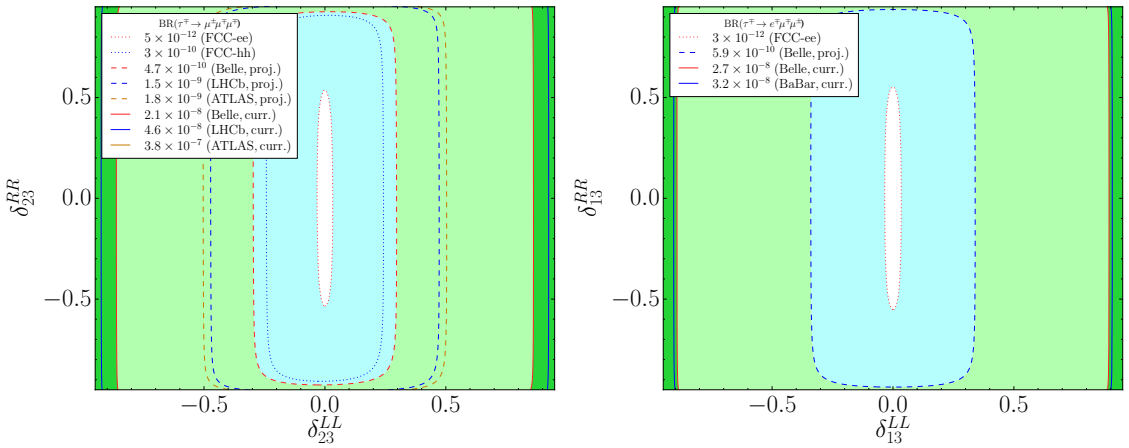
the soft breaking terms $m^2_{\tilde{L}\,ij}$, $m^2_{\tilde{E}\,ij}$ and $v_1\mathcal{A}^\ell_{ij}$ can induce flavour-changing neutral current interactions, such as that between a lepton, slepton and neutralino. This is shown by the Feynman diagrams in Fig. 2.9 for the LFV decay modes $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$, which arise at the one-loop level.

2.5.2 Limits on the Parameter Space

We consider the impact of the present and future LFV constraints on the flavour-violating parameters δ^{LL}_{ij} and δ^{RR}_{ij}. To numerically compute the branching ratios in the MSSM, we use the supersymmetry spectrum generator SPheno [98,99], with the source code for the flavour observables produced by SARAH [100]. We work with the following benchmark point for the MSSM parameters that provides a particle spectrum in agreement with the present collider limits:

$$\begin{aligned} \tan\beta = 10\ , \qquad & \mu = -100\ \text{GeV}\ , \\ M_A = 1000\ \text{GeV}\ , \qquad & M_1 = 250\ \text{GeV}\ , \\ M_2 = 500\ \text{GeV}\ , \qquad & M_3 = 2000\ \text{GeV}\ , \\ m_{\tilde{L}_i} = m_{\tilde{E}_j} = 1000\ \text{GeV}\ , \qquad & A_\tau = 200\ \text{GeV}\ . \end{aligned} \tag{2.5.6}$$

Here, M_A is the mass of the pseudoscalar Higgs boson, $M_{1,2,3}$ are the gaugino masses and A_τ is the slepton trilinear coupling for the third generation. We do not specify squark supersymmetry-breaking parameters, as their values are not relevant for the

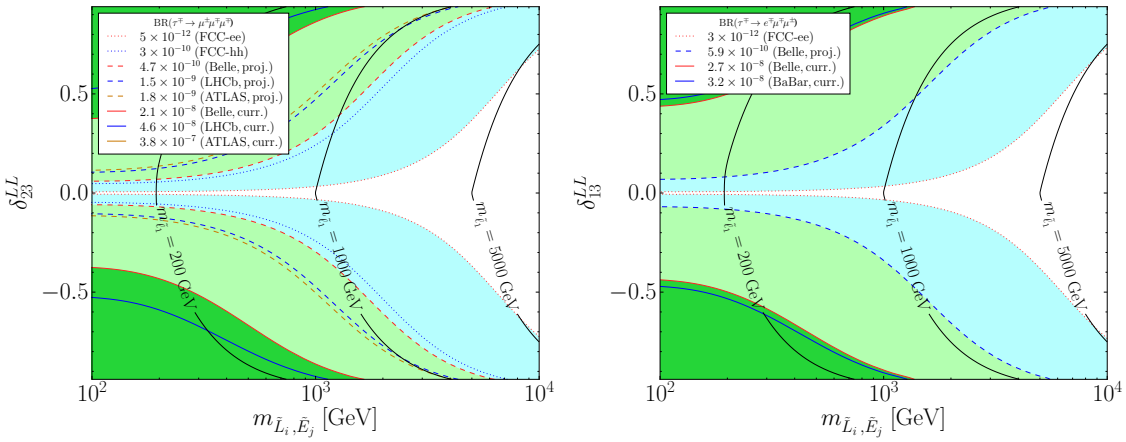


(a) **(b)**

Figure 2.10: Current and future branching-ratio limits in the parameter plane of (a) δ_{23}^{LL} and δ_{23}^{RR} for the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and (b) δ_{13}^{LL} and δ_{13}^{RR} for the decay $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$ in the MSSM.

processes we calculate. While searches for squarks and gluinos by ATLAS [101, 102] and CMS [103, 104] have pushed their respective mass limits to already rather large values, limits for slepton masses are still fairly weak [64]. Direct slepton pair production requires the exchange of electroweak gauge bosons and is thus strongly suppressed compared with squark or gluino pair production at hadron colliders. Hence, assuming LFV is realised in nature, much stronger limits on the slepton masses can be obtained indirectly by measuring rare LFV decays.

In Figs. 2.10a and 2.10b, we show present and future constraints on the pair $(\delta_{23}^{LL}, \delta_{23}^{RR})$ from the process $\tau^\mp \to \mu^\mp\mu^\mp\mu^\pm$, and the pair $(\delta_{13}^{LL}, \delta_{13}^{RR})$ from the process $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$, respectively. In analogy with the squark sector [105], we find that the δ_{13}^{RR} and δ_{23}^{RR} parameters are much less constrained than their LL counterparts. This is because the processes are mediated by flavour-violating neutralino interactions. In the gauge-interaction basis, the exchanged particles are the bino ($\tilde{B}$), wino ($\tilde{W}^0$) or Higgsino ($\tilde{H}_i$) particles. The $\tilde{H}_i - \ell_R - \tilde{\ell}_L$ interactions are proportional to the lepton Yukawa coupling y_l and are thus subleading, while $\tilde{B} - \ell_{R/L} - \tilde{\ell}_{R/L}$ and $\tilde{W}^0 - \ell_L - \tilde{\ell}_L$ interactions occur with the strength of their associated gauge couplings. Therefore, the branching ratios $\tau^\mp \to \mu^\mp\mu^\mp\mu^\pm$ and $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$ are amplified for a light



(a) **(b)**

Figure 2.11: Current and future branching-ratio limits in the parameter plane of (a) $m_{\tilde{L}_i,\tilde{E}_j}$ and δ^{LL}_{23} for the decay $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$ and (b) $m_{\tilde{L}_i,\tilde{E}_j}$ and δ^{LL}_{13} for the decay $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$ in the MSSM. The solid black lines represent constant values of the mass of the slepton $\tilde{\ell}_1$.

wino-type neutralino—small M_2, and large δ^{LL}_{ij}.

In Fig. 2.11, we show the LFV branching-ratio limits where the soft slepton mass scale is allowed to vary along with a single mixing parameter. We vary the slepton mass scale over a wide range. For slepton masses at the current lower bound from direct searches ($\sim$ 100 GeV), future experiments could place very strong constraints on LFV parameters. Since the slepton masses are large when the soft slepton mass scales $m_{\tilde{L}_i} = m_{\tilde{E}_j}$ are large, their contribution to LFV processes decouples and the sensitivity to the mixing parameters is reduced.

2.6 Conclusions

In this chapter, we have discussed three models of physics beyond the Standard Model and shown the restrictions on their parameter spaces that can be derived from the limits on LFV decays of the τ lepton. We focussed on the Type-II Seesaw Model, the LRSM and the MSSM and considered the decays $\tau^\mp \to \mu^\pm\mu^\mp\mu^\mp$, $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ and $\tau^\mp \to e^\mp\mu^\mp\mu^\pm$. The limits on the LFV were calculated from current and future lepton and hadron colliders.

For the Type-II Seesaw Model with a small triplet VEV v_Δ in the range $(10^{-11} - 10^{-9})$ GeV—which naturally explains the $(0.01 - 1)$ eV light neutrino masses with $\mathcal{O}(1)$ Yukawa coupling Y_Δ—the model parameter μ_Δ is presently constrained as $\mu_\Delta \geq (2\times10^{-9} - 7\times10^{-8})$ GeV. The future circular collider FCC-ee could provide improved constraints on μ_Δ by almost two orders of magnitude. Constraints on the CP-violating phase $\delta_{\rm CP}$ of the PMNS mixing matrix could be obtained by the Belle-II experiment in regions around $\pi/2$ and $3\pi/2$ for a quasi-degenerate neutrino spectrum with the oscillation angles equal to their best-fit values.

For the LRSM we considered two extreme regimes, with a lower and higher value of the symmetry-breaking scale v_R, respectively. For the first benchmark point BP1, we considered a lower $v_R = 8$ TeV and a large $\alpha_3 \sim \mathcal{O}(10)$. For BP2, we considered a larger $v_R = 30$ TeV with a smaller $\alpha_3 \sim \mathcal{O}(1)$, which is well within the perturbative regime. In BP1, and for a doubly-charged Higgs mass $M_{\delta_R^{\pm\pm}} = 800$ GeV, we find that the right-handed neutrino masses can be constrained to $m_N \leq 290$ GeV by the present stringent limit from Belle. The future limits from LHCb and Belle-II will further constrain the right-handed neutrino masses down to the $m_N \leq 100$ GeV mass range. Further improvements at the future circular colliders will allow for tighter constraints on the ρ_2 parameter and the doubly-charged Higgs mass $M_{\delta_R^{\pm\pm}}$ to be obtained.

Finally, for the MSSM, we explored the present and future constraints on the dimensionless LFV parameters δ_{13}^{LL}, δ_{23}^{LL} (and their RR equivalents) and the soft slepton masses from the $\tau^\mp \to \mu^\mp\mu^\mp\mu^\pm$ and $\tau^\mp \to e^\pm\mu^\mp\mu^\mp$ decays. We find that δ_{13}^{LL} and δ_{23}^{LL} are at present bounded by Belle to $|\delta_{13,23}^{LL}| \lesssim 0.9$ for the benchmark scenario we chose. The future constraints from existing colliders will improve the limits to ~ 0.2, while an FCC-ee collider could further constrain this parameter to as low as 0.03.

Chapter 3

An Electroweak Approach to the Trilinear Higgs Self-Interaction

In our second investigation into physics beyond the Standard Model, we consider potential modifications to the trilinear self-coupling of the Higgs boson due to new physics, and calculate how these affect the electroweak oblique Peskin-Takeuchi parameters S and T. The effect is fairly small since the electroweak oblique parameters arise at one-loop level at the lowest order, with the Higgs trilinear self-interaction only entering at the next-to-leading two-loop order. However, since di-Higgs production has not yet been observed—which would allow for a direct measurement of the Higgs trilinear self-coupling—the bounds on this coupling are currently very weak. Therefore, the constraints on the Peskin-Takeuchi parameters can allow competitive and complementary (due to the orthogonal approach) limits to be set on the trilinear self-coupling of the Higgs boson.

3.1 Motivation

The discovery of the Higgs boson [106, 107] has been a major success of the LHC and confirms the underlying mechanism of spontaneous symmetry breaking in the

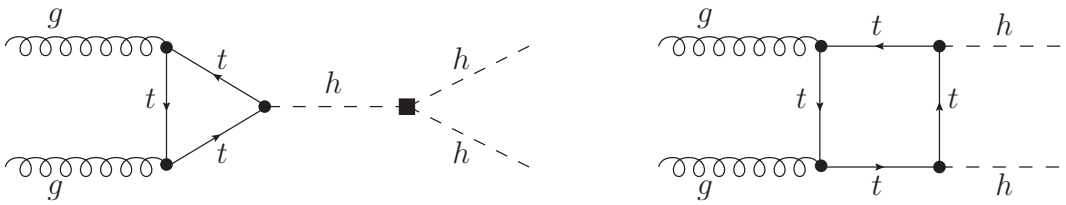


Figure 3.1: Feynman diagrams contributing to the gluon fusion di-Higgs production mode at leading order. The black square represents the trilinear self-interaction of the Higgs boson.

electroweak sector, which we described in Chapter 1. However, although the mass of the Higgs boson has been directly experimentally measured [16], its couplings to the gauge bosons and heavy fermions are much more loosely constrained—they lie approximately within 10% and 20% of their Standard Model predictions, respectively [108]. The situation for the self-couplings of the Higgs boson is far worse [109].

An observation of di-Higgs production would allow a direct measurement of the trilinear self-coupling of the Higgs boson to be made since it enters the process at leading order. This is represented by the black square in Fig. 3.1, which shows the Feynman diagrams contributing to the gluon fusion di-Higgs production mode, which is dominated by top quarks in the loop. However, because the cross section for di-Higgs production is very small, with a Standard Model expectation of $\mathcal{O}(10\text{ fb})$, this process has never been observed [110]. Analogously to the LFV τ lepton decays in the previous chapter, this means that the best one can do is obtain upper limits for the cross section of the process, which can be translated into limits on the trilinear self-coupling. The outlook for the quartic self-coupling is even bleaker [111]. Therefore, an important goal for current and future high-energy experiments is to probe the Higgs self-couplings with a much higher precision, from both direct measurements and indirect determinations. Such measurements would allow for a thorough testing of the structure and shape of the scalar potential, and uncover any deviations from the Standard Model.

We can parameterise a deviation from the Standard Model prediction for the Higgs trilinear self-coupling by multiplying it by κ_λ. A value of $\kappa_\lambda = 1$ would correspond to

the Standard Model result, and any deviations from this would indicate disagreement with the Standard Model. Direct searches for di-Higgs production constrain κ_λ to $-14.5 \leq \kappa_\lambda \leq 19.1$ [110, 112] and $-8.4 \leq \kappa_\lambda \leq 13.4$ [113, 114] using Run I and Run II data, respectively. In addition, Higgs coupling measurements performed in single Higgs production lead to the combined bound of $-9.4 \leq \kappa_\lambda \leq 17.0$ [115]. Clearly these bounds are very weak and allow a lot of room for improvement in future searches. Therefore, it is reasonable to suppose that alternative approaches might perform at least as well, and so we propose a different method for deriving constraints on the trilinear self-coupling of the Higgs boson.

In the case of observables involving leptons and gauge bosons, the experimental precision is currently competing with the theoretical precision. This state-of-the-art data allows electroweak precision tests of the Standard Model to be performed. In general, this procedure works by taking some input measurements, carrying out a theoretical calculation to predict what should be seen for another observable, and then comparing it to the experimental measurement.

For the theoretical prediction to have any chance of success, the calculation of the observables has to be performed beyond tree level to include the effects of loop diagrams. In the electroweak sector, the loops which affect gauge boson propagators directly have by far the dominant effect. These are known as oblique corrections, and include the effects of gauge bosons, ghosts, scalars and fermions in the loops. In the Standard Model, the dominant oblique corrections arise from loops involving the top and bottom quarks, and the Higgs boson.

This also provides an excellent place to look for the effects of physics beyond the Standard Model. By making some simple assumptions about the new physics model, the entire contribution to the oblique corrections can be parameterised by the three Peskin-Takeuchi parameters S, T and U [116, 117]. By calculating the values of these parameters in the new model, their value can be compared to experimental fits to provide constraints on the model parameters.

Therefore, we will calculate the effect of a modified trilinear self-coupling of the Higgs boson on the Peskin-Takeuchi parameters S and T. These arise at the two-loop level, and so will require the computation of two-loop self-energy diagrams of the gauge bosons. From the experimental fits for the Peskin-Takeuchi parameters, we will derive the corresponding constraints on the trilinear Higgs self-coupling. The rest of this chapter is structured as follows: in Sec. 3.2 we will describe the basics of electroweak precision measurements and introduce the Peskin-Takeuchi parameters. We will then modify the Higgs potential in Sec. 3.3, and calculate the effect of this on the Peskin-Takeuchi parameters in Sec. 3.4. The results of this calculation will be presented in Sec. 3.5, before the chapter is concluded in Sec. 3.6.

3.2 Testing the Electroweak Sector of the Standard Model

3.2.1 Electroweak Precision Basics

The electroweak sector of the Standard Model is constructed from the gauge group $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. At tree level, only three parameters in the Lagrangian are required to calculate all electroweak observables at energies where the quark and lepton masses can be ignored. These are the $\mathrm{SU}(2)_L$ gauge coupling g, the $\mathrm{U}(1)_Y$ gauge coupling g' and the Higgs VEV v. Equivalently, g and g' can be exchanged for the weak mixing angle $s \equiv \sin\theta_\mathrm{W}$ and the $\mathrm{U}(1)_\mathrm{EM}$ charge e through the relations $g = e/s$ and $g' = e/c$, where $c \equiv \cos\theta_\mathrm{W}$. It is then possible to express any electroweak parameters in terms of these Lagrangian parameters via simple tree-level relations. For example, the fine-structure constant is $\alpha_e \equiv e^2/(4\pi)$ and the $W^\pm$ and Z gauge boson masses are,

$$m_Z = \frac{ev}{2sc} \,, \qquad m_W = \frac{ev}{2s} \,, \tag{3.2.1}$$

while the Fermi constant G_{F} can be expressed solely in terms of the Higgs VEV as,

$$G_{\mathrm{F}} = \frac{1}{\sqrt{2}v^2} \ . \tag{3.2.2}$$

In order to perform electroweak precision tests of the Standard Model, a set of observables needs to be measured experimentally. These can be used to determine the parameters in the Lagrangian by performing a best fit to the data, from which a prediction can be made for another observable and compared to the experimental result. Crucially, the Standard Model is over-constrained, which means that with enough measurements, it is possible to determine every free parameter and thus meaningful predictions are possible. If it was under-constrained, then there would not be enough observables to determine every free parameter.

To begin with, three measurements must be used to determine e, s and v. For example, some precisely measured observables that could be used are the electron magnetic dipole moment g_e, the muon lifetime τ_μ or the Z boson pole mass m_Z^{Pole}. From the measurement of the electron magnetic dipole moment, the fine-structure constant at zero momentum has been computed to be $\hat{\alpha}_e(0) = (137.035999139 \pm 0.000000031)^{-1}$ [64]. Running this value to the $\overline{\mathrm{MS}}$ Z boson mass m_Z gives,

$$\hat{\alpha}_e(m_Z)^{-1} = 127.950 \pm 0.017 \ . \tag{3.2.3}$$

Following the notation used in Ref. [118], we denote further quantities derived from these measurements via simple tree-level relations with a hat. In addition, we can use the measured Z boson pole mass to define $\hat{m}_Z = m_Z^{\mathrm{Pole}}$, and using the tree-level relations in Eq. (3.2.2) the muon decay rate gives $\hat{v} = 246.48$ GeV. We can rearrange Eq. (3.2.1) for the Z boson mass to obtain the relation for the weak mixing angle,

$$\hat{s}^2(1-\hat{s}^2) = \frac{\pi\hat{\alpha}_e(m_Z)}{\sqrt{2}\hat{G}_F\hat{m}_Z^2} \ , \tag{3.2.4}$$

which evaluates to $\hat{s}^2 = 0.234289$. We note here that the input measurements chosen for this tree-level analysis are arbitrary, and choosing a different set of measurements for the input would result in different values for the derived parameters.

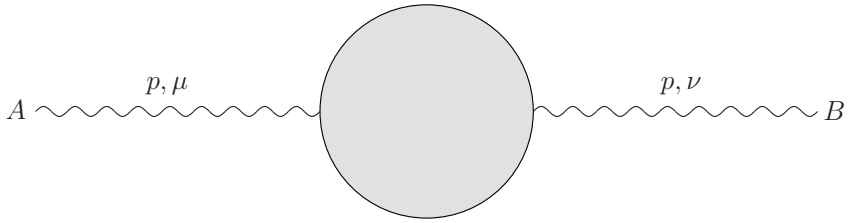


Figure 3.2: Oblique corrections to the gauge boson A propagating into the gauge boson B.

We can now perform a test of the Standard Model by making a prediction for a new observable. Using the relation for the $W^{\pm}$ boson mass in Eq. (3.2.1), one can predict that $\hat{m}_W = 79.794$ GeV. If we were to equate this to the $W^{\pm}$ boson pole mass, then we would find that it is far outside the experimental uncertainty on the measured value, $m_W^{\text{Pole}} = (80.385 \pm 0.015)$ GeV.

3.2.2 Oblique Corrections

Clearly, the tree-level analysis performed above is not sufficient to be able to make predictions at the level of the experimental precision, and so it is necessary to compute to higher orders in perturbation theory. Including all possible loop corrections would be a formidable task, but a simplification can be made by noting that since the electroweak observables come from the exchange of gauge bosons at tree level, the dominant corrections will be from loops inside the gauge boson propagators. These oblique corrections (also known as vacuum polarisation or self-energy corrections) can be probed experimentally using four-fermion scattering since they do not depend on the specific external fermions.

For a gauge boson A with incoming four-momentum p propagating via a loop into a gauge boson B, the oblique corrections have the form,

$$i\Pi_{AB}^{\mu\nu}(p) = i\Pi_{AB}(p^2)g^{\mu\nu} + i\Pi_{AB}^{pp}(p^2)p^{\mu}p^{\nu} \ . \tag{3.2.5}$$

This is the only form for the result allowed by Lorentz invariance and the Ward identity [119], and is represented diagrammatically in Fig. 3.2. The second term will always vanish when we consider observables with light external fermions whose masses can be neglected, since it will be contracted with the fermion current. If we sum up all the one-particle-irreducible (1PI) insertions into the propagator for a Z boson, then the pole mass can be written in terms of the $\overline{\text{MS}}$ mass from the Lagrangian as,

$$(m_Z^{\text{Pole}})^2 = m_Z^2 + \text{Re}\left[\Pi_{ZZ}(m_Z^2)\right] \ . \tag{3.2.6}$$

A similar relation holds for the $W^\pm$ boson pole mass. The imaginary parts of the vacuum polarisation functions arise when it is kinematically possible for all particles in a particular loop diagram contribution to the self-energy to go on-shell. By the optical theorem, the total imaginary part is proportional to the decay width of the gauge boson. Thus it is only the real parts of the vacuum polarisation functions which can contribute to a physical shift in the gauge boson masses, and we will always assume that the real part has been taken from now on.

The goal is to be able to use the measurements from some observables to make a prediction for a new observable which can be compared to data. Similar to the tree-level approach above, we will determine the Lagrangian parameters using the measured values and then substitute the results into the expressions for the new observables.

The electromagnetic charge $\hat{e}^2(m_Z)$ can be extracted from the measurement of the electron magnetic dipole moment and compared to the one-loop result for Coulomb scattering to give the relation,

$$e^2 = \hat{e}^2(m_Z)\left[1 - \frac{\Pi_{\gamma\gamma}(\hat{m}_Z^2)}{\hat{m}_Z^2}\right] \ , \tag{3.2.7}$$

where e is the $\overline{\text{MS}}$ parameter from the Lagrangian and $\hat{m}_Z$ is equal to m_Z^{Pole}. Inverting Eq. (3.2.6), we get the expression for the Z boson $\overline{\text{MS}}$ mass,

$$m_Z^2 = \hat{m}_Z^2\left[1 - \frac{\Pi_{ZZ}(\hat{m}_Z^2)}{\hat{m}_Z^2}\right] \ . \tag{3.2.8}$$

The limit of the $W^{\pm}$ boson propagator at low energies allows us to define $\hat{G}_F$ through,

$$\hat{G}_F = \sqrt{2}\frac{e^2}{8s^2c^2m_Z^2}\left(1 - \frac{\Pi_{WW}(0)}{m_W^2}\right) , \tag{3.2.9}$$

where all quantities on the right-hand side are $\overline{\text{MS}}$ Lagrangian parameters which satisfy the tree-level relations. Solving Eqs. (3.2.7)–(3.2.9) and keeping only the one-loop terms, we can obtain an expression for the weak mixing angle,

$$s^2 = \hat{s}^2\left(1 + \frac{\hat{c}^2}{\hat{c}^2 - \hat{s}^2}\Pi_R\right) , \tag{3.2.10}$$

where we define,

$$\Pi_R \equiv -\frac{\Pi_{\gamma\gamma}(\hat{m}_Z^2)}{\hat{m}_Z^2} + \frac{\Pi_{ZZ}(\hat{m}_Z^2)}{\hat{m}_Z^2} - \frac{\Pi_{WW}(0)}{\hat{m}_W^2} . \tag{3.2.11}$$

Finally, using the relation $m_W^2 = c^2 m_Z^2$ between the $\overline{\text{MS}}$ parameters, we can make a prediction for the pole mass of the $W^{\pm}$ boson,

$$(m_W^{\text{Pole}})^2 = \hat{c}^2\hat{m}_Z^2\left(1 - \frac{\hat{s}^2}{\hat{c}^2 - \hat{s}^2}\Pi_R - \frac{\Pi_{ZZ}(\hat{m}_Z^2)}{\hat{m}_Z^2} + \frac{\Pi_{WW}(\hat{m}_W^2)}{\hat{c}^2\hat{m}_Z^2}\right) . \tag{3.2.12}$$

Using only the three input measurements, an analytic calculation of the vacuum polarisations allows a prediction for the $W^{\pm}$ boson pole mass (and many more observables) to be made which will have far more reasonable agreement with the data.

The only remaining issue is to compute all the vacuum polarisations. Due to gauge invariance, the photon is massless and so $\Pi_{\gamma\gamma}(0)$ must be zero to all orders in perturbation theory. Although there are many loop insertions possible for each gauge boson propagator, the dominant Standard Model effects will come from the Higgs boson and the top (and bottom, for $\text{SU}(2)_L$ invariance) quarks since the result increases with the mass of the particle. We show the analytic calculation for the insertion of a fermion loop into the gauge boson propagators in Appendix A. It is worth noting that all divergences in the vacuum polarisation functions cancel when they are combined into an observable, as well as the unphysical mass scale μ. This is necessary for ensuring that the predictions for the measured values are finite.

3.2.3 Peskin-Takeuchi Parameters

We have seen that the oblique corrections in the electroweak sector are a useful testing ground for the effects of calculating to higher orders in perturbation theory, which facilitates precise tests of Standard Model predictions. In the same way, the oblique corrections allow for a remarkably simple method to constrain new physics models beyond the Standard Model, as long as a few assumptions about the model are made [120]:

1. The gauge group which describes the electroweak sector is still the Standard Model $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group, with the same gauge bosons.

2. The couplings of any new physics particles to gauge bosons dominate over their couplings to light fermions.

3. The scale Λ at which the new physics enters is much larger than the masses of the gauge bosons—that is, $\Lambda \gg m_Z$.

The first two assumptions mean that the dominant effect of new physics contributions will be through the vacuum polarisations of the gauge bosons—exactly the oblique corrections which we introduced above. Furthermore, by redefining the fields, it is possible to entirely constrain the effects of new physics solely by the oblique corrections.

Since the Standard Model itself describes electroweak precision data very well, the contributions from new physics must be small, so the Standard Model oblique corrections can be perturbed by adding a new physics contribution as,

$$\Pi_{AB}(p^2) = \Pi^{\mathrm{SM}}_{AB}(p^2) + \Delta\Pi_{AB}(p^2) \ . \tag{3.2.13}$$

The third assumption above means that the new physics contribution can be Taylor expanded in p^2 as,

$$\Delta\Pi_{AB}(p^2) = \Delta\Pi_{AB}(0) + p^2\Delta\Pi'_{AB}(0) + \mathcal{O}\left(m_Z^2/\Lambda^2\right) \ , \tag{3.2.14}$$

where the derivative of the vacuum polarisation function is defined by,

$$\Delta\Pi'_{AB}(p^2) \equiv \frac{d}{dp^2}\Delta\Pi_{AB}(p^2) \ . \tag{3.2.15}$$

Neglecting terms of $\mathcal{O}(m_Z^2/\Lambda^2)$, the derivative $\Delta\Pi'_{AB}(0)$ can be expressed as,

$$\Delta\Pi'_{AB}(0) = \frac{\Delta\Pi'_{AB}(m_Z^2) - \Delta\Pi'_{AB}(0)}{m_Z^2} \ . \tag{3.2.16}$$

Naively, there are eight independent parameters needed to describe the effects of new physics—two quantities for each of the four vacuum polarisation functions. However, $\Delta\Pi_{\gamma\gamma}(0)$ is zero by gauge invariance, and four more can be removed by the renormalisation conditions involving the three input parameters and $\Pi_{Z\gamma}(0)$. This leaves three independent quantities, and a convenient parameterisation for them is given by [121],

$$\begin{aligned} S = \frac{4c^2s^2}{\alpha_e m_Z^2}\Bigg(&\Delta\Pi_{ZZ}(m_Z^2) - \Delta\Pi_{ZZ}(0) - \Delta\Pi_{\gamma\gamma}(m_Z^2) \\ &- \frac{c^2-s^2}{cs}\left[\Delta\Pi_{Z\gamma}(m_Z^2) - \Delta\Pi_{Z\gamma}(0)\right]\Bigg) \ , \end{aligned} \tag{3.2.17}$$

$$T = \frac{1}{\alpha_e}\left(\frac{\Delta\Pi_{WW}(0)}{m_W^2} - \frac{c^2}{m_W^2}\left[\Delta\Pi_{ZZ}(0) + \frac{2s}{c}\Delta\Pi_{Z\gamma}(0)\right]\right) \ . \tag{3.2.18}$$

$$\begin{aligned} U = \frac{4s^2}{\alpha_e}\Bigg(&\frac{\Delta\Pi_{WW}(m_W^2) - \Delta\Pi_{WW}(0)}{m_W^2} - c^2\frac{\Delta\Pi_{ZZ}(m_Z^2) - \Delta\Pi_{ZZ}(0)}{m_Z^2} \\ &- 2cs\frac{\Delta\Pi_{Z\gamma}(m_Z^2) - \Delta\Pi_{Z\gamma}(0)}{m_Z^2} - s^2\frac{\Delta\Pi_{\gamma\gamma}(m_Z^2)}{m_Z^2}\Bigg) \ . \end{aligned} \tag{3.2.19}$$

The quantities S, T and U are known as the Peskin-Takeuchi parameters. S measures the effect of neutral-current processes at different energies, while T measures the difference between neutral- and charged-current processes at low energies and so is sensitive to a violation of weak isospin. In fact, T is related to the ρ parameter from Eq. (1.1.30) by $T = (\rho - 1)/\alpha_e$, so T parameterises the amount of custodial symmetry violation in a new physics model. U is usually insensitive to new physics since it is only constrained by the $W^\pm$ boson mass and width, and thus S and T are normally used for testing new physics—we will assume that $U = 0$ throughout.

By construction, the Peskin-Takeuchi parameters only include the effects of new physics and so if there is no new physics then all three parameters are zero. Equivalently, the parameters could be defined to include the complete vacuum polarisation functions with Standard Model effects. In this case, the Standard Model contribution must subsequently be subtracted, for given reference values of the masses of the top quark and Higgs boson—which we take to be $m_t = 173$ GeV and $m_h = 125$ GeV, respectively—leaving only the contributions from the new physics behind.

The Peskin-Takeuchi parameters can be measured experimentally by performing fits of electroweak observables. The constraints on S and T can then be obtained by fixing $U = 0$ in the fits. The Standard Model result ($S = T = 0$) is consistent with current experimental constraints on the Peskin-Takeuchi parameters. These are $S|_{U=0} = 0.06 \pm 0.09$ and $T|_{U=0} = 0.10 \pm 0.07$, as calculated by the Gfitter Group [122]. However, this does not rule out new physics—rather, it means that the new physics contributions must be small enough such that their predictions for S and T remain within the confidence limits. This allows one to calculate constraints on parameters of new physics models, by calculating the range of parameters that result in experimentally allowed values of S and T.

Although m_t and m_h have now been measured very precisely, historically the Peskin-Takeuchi parameters were used to obtain bounds on these masses. The dependence of S and T on m_t and m_h, valid when $m_t, m_h \gg m_Z$ and at one-loop level, can be written as [116],

$$S \simeq \frac{1}{12\pi} \ln\left(\frac{m_h^2}{(m_h^{\mathrm{Ref}})^2}\right) - \frac{1}{6\pi} \ln\left(\frac{m_t^2}{(m_t^{\mathrm{Ref}})^2}\right) , \tag{3.2.20}$$

$$T \simeq -\frac{3}{16\pi c^2} \ln\left(\frac{m_h^2}{(m_h^{\mathrm{Ref}})^2}\right) + \frac{3}{16\pi s^2 c^2}\left(\frac{m_t^2 - (m_t^{\mathrm{Ref}})^2}{m_Z^2}\right) , \tag{3.2.21}$$

where m_h^{Ref} and m_t^{Ref} are reference values of the Higgs boson and top quark masses, respectively.

As a simple example of a new physics model, consider an extension to the number of generations of fermions, such that there are four generations of quarks and four

generations of leptons. Since they would obtain their masses via Yukawa couplings to the Higgs field, they should be accessible by experiments at the electroweak scale. Including the fourth generation fermions $(Q_{L4}, u_{R4}, d_{R4}, L_{L4}, e_{R4}, \nu_{R4})$, the Yukawa terms become [123],

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j}\Big[(Y_u)_{ij}\overline{Q}_{Li}\tilde{\Phi}u_{Rj} + (Y_d)_{ij}\overline{Q}_{Li}\Phi d_{Rj} + (Y_e)_{ij}\overline{L}_{Li}\Phi e_{Rj} + (Y_\nu)_{ij}\overline{L}_{Li}\tilde{\Phi}\nu_{Rj}\Big] + \text{h.c.}\,, \tag{3.2.22}$$

where the sum runs over the four generations and we only consider a Dirac mass for the fourth generation of neutrinos. For simplicity, we consider diagonal CKM and PMNS matrices, so that the only new parameters of the model are the masses m_{u4}, m_{d4}, $m_{\nu 4}$ and m_{e4}. Current bounds from CDF exclude m_{u4} below 358 GeV [124] and m_{d4} below 372 GeV [125]. We require $m_{\nu 4} > m_Z/2$ for the neutrino not to contribute to the invisible width of the Z boson, and both of the fourth generation leptons have a lower limit of 101 GeV from LEP [126].

If we consider a single fermion doublet with masses (m_1, m_2), where $(m_1, m_2) \equiv (m_{u4}, m_{d4})$ for quarks and $(m_1, m_2) \equiv (m_{\nu 4}, m_{e4})$ for leptons, then the one-loop contributions to the Peskin-Takeuchi parameters S and T, in the limit that $m_1, m_2 \gg m_Z$, are [116],

$$S = \frac{N_c}{6\pi}\left[1 - Y\ln\left(\frac{m_1^2}{m_2^2}\right)\right]\,, \tag{3.2.23}$$

$$T = \frac{1}{16\pi s^2 c^2 m_Z^2}\left[m_1^2 + m_2^2 - \frac{2m_1^2 m_2^2}{m_1^2 - m_2^2}\ln\left(\frac{m_1^2}{m_2^2}\right)\right]\,, \tag{3.2.24}$$

where $N_c = 3(1)$ and $Y = \frac{1}{3}(-1)$ for quarks (leptons). For $m_1 = m_2$, $T = 0$ for both quarks and leptons since there is no weak isospin violation, and T grows as the mass splitting increases. S grows when m_1 increases with respect to m_2 for leptons, with the opposite behaviour for quarks due to the differing signs of the weak hypercharge Y. A calculation of S and T can then be used to obtain complementary constraints on the allowed values of the fourth generation masses.

3.3 A Modified Higgs Potential

In Chapter 1, we introduced the Higgs mechanism for spontaneous symmetry breaking in the electroweak sector. To make clear the distinction between Standard Model and new physics effects, we now write the Standard Model Higgs potential in terms of the Higgs doublet Φ as,

$$V_{\mathrm{SM}}(\Phi) = \mu_{\mathrm{SM}}^2 \Phi^\dagger \Phi + \lambda_{\mathrm{SM}} (\Phi^\dagger \Phi)^2 \ . \tag{3.3.1}$$

After spontaneous symmetry breaking and expanding around the VEV of the neutral component of the Higgs doublet, this can be expressed in terms of the physical Higgs boson h as,

$$V_{\mathrm{SM}}(h) \supset \frac{m_h^2}{2} h^2 + \lambda_{\mathrm{SM}} v h^3 + \frac{\lambda_{\mathrm{SM}}}{4} h^4 \ , \tag{3.3.2}$$

where $m_h^2 = -2\mu_{\mathrm{SM}}^2 = 2\lambda_{\mathrm{SM}} v^2$. The Higgs mass has been directly experimentally measured, with the Higgs VEV inferred from the muon lifetime. Since there are only two independent parameters in the Higgs potential in the Standard Model, this constrains all the other self-couplings of the Higgs boson, and its couplings with the Goldstone bosons. However, because the self-couplings have not been directly measured, it is reasonable to suppose that the Standard Model might not be the full description—new physics effects could introduce extra parameters which allow the self-couplings to vary independently of the Higgs mass and VEV. Thus, we parameterise modifications to the Higgs self-couplings as,

$$V_{\mathrm{Mod}}(h) \supset \frac{m_h^2}{2} h^2 + \kappa_\lambda \lambda_{\mathrm{SM}} v h^3 + \kappa_4 \frac{\lambda_{\mathrm{SM}}}{4} h^4 \ . \tag{3.3.3}$$

In general, new physics effects that modify the Higgs potential would also cause modifications to other couplings in the Standard Model. However, for the purposes of this study, we consider only modifications to the Higgs trilinear and quartic self-couplings in isolation from the other Standard Model couplings. This is reasonable if we can formulate these modifications in a gauge-invariant way. The modifications in Eq. (3.3.3) make sense in the unitary gauge, but in the Feynman gauge—which

contains interactions with Goldstone bosons—a more formal approach is required. We implement the modified couplings using higher-dimensional operators that only affect the Higgs potential [127, 128],

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} - \sum_{n\geq 3} \frac{\bar{c}_{2n}\lambda_{\text{SM}}}{v^{2n-4}}(\Phi^\dagger\Phi)^n \ , \tag{3.3.4}$$

where we have normalised the couplings with a factor of $\lambda_{\text{SM}} \equiv m_h^2/(2v^2)$. The modified Higgs scalar potential becomes,

$$V(H) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 + \sum_{n\geq 3} \frac{\bar{c}_{2n}\lambda_{\text{SM}}}{v^{2n-4}}(\Phi^\dagger\Phi)^n \ , \tag{3.3.5}$$

where now μ^2 and λ are in general different from their corresponding Standard Model values.

For now, consider extending the Standard Model with just the additional dimension-6 operator $\mathcal{O}_6$. The minimisation conditions are shifted, and so μ^2 and λ develop different relations in terms of the physical Higgs boson mass m_h and VEV v, which remain fixed to their experimental values. These relations are,

$$\mu^2 = -\lambda_{\text{SM}}v^2\left(1 - \frac{3}{4}\bar{c}_6\right) \ , \qquad \lambda = \lambda_{\text{SM}}\left(1 - \frac{3}{2}\bar{c}_6\right) \ , \tag{3.3.6}$$

where the Standard Model results are recovered in the limit $\bar{c}_6 \to 0$. Expanding the potential around the VEV once again, the Higgs potential becomes Eq. (3.3.3) with the identifications,

$$\kappa_\lambda = 1 + \bar{c}_6 \ , \qquad \kappa_4 = 1 + 6\bar{c}_6 \ . \tag{3.3.7}$$

At this stage, we have a gauge-invariant correlated modification of the trilinear and quartic Higgs self-couplings. This also introduces modifications of the couplings between the Higgs and Goldstone bosons—the $hh\phi^0\phi^0$ coupling is modified by a factor of $1 + 3\bar{c}_6$ and the $hh\phi^+\phi^-$ coupling is also modified by the same factor. All the other three-point and four-point interactions involving the Higgs and Goldstone bosons remain exactly as they are in the Standard Model.

This can be generalised to two separate uncorrelated modifications by also including the dimension-8 operator from Eq. (3.3.4) with coefficient $\bar{c}_8$. The modified trilinear

and quartic Higgs self-couplings become,

$$\kappa_\lambda = 1 + \bar{c}_6 + 2\bar{c}_8 \ , \qquad \kappa_4 - 1 = 1 + 6\bar{c}_6 + 16\bar{c}_8 \ . \tag{3.3.8}$$

Furthermore, the $hh\phi^0\phi^0$ and $hh\phi^+\phi^-$ couplings are both modified by factors of $1 + 3\bar{c}_6 + 6\bar{c}_8$. That is, the corrections to the couplings between the Higgs and Goldstone bosons are exactly a factor of three larger than those for the trilinear Higgs self-coupling, and so they remain correlated. Only the Higgs quartic self-coupling is modified independently. If we include even higher-dimensional operators $(\Phi^\dagger\Phi)^n$ with $n \leq n_{\max}$, we again find two different linear combinations,

$$\kappa_\lambda = 1 + \sum_{n=3}^{n_{\max}} a_{2n}\bar{c}_{2n} \ , \qquad \kappa_4 = 1 + \sum_{n=3}^{n_{\max}} b_{2n}\bar{c}_{2n} \ . \tag{3.3.9}$$

The coefficients a_{2n} and b_{2n}, where in general $a_{2n} \neq b_{2n}$, have to be evaluated for the chosen $n_{\max}$. We will see that it is not necessary to include operators beyond the additional dimension-6 operator $\mathcal{O}_6$ since the quartic self-coupling, and hence κ_4, will be shown to not contribute to S and T at two loops. Furthermore, the modified couplings between the Higgs and Goldstone bosons remain correlated to the trilinear Higgs self-coupling. The calculation will therefore be performed in terms of the single additional dimension-6 operator $\mathcal{O}_6$, with the result expressed in terms of $\bar{c}_6$. This will then allow a direct translation of the result into a bound on the κ_λ trilinear self-coupling modification. The higher-dimensional operators in Eq. (3.3.4) also generate even higher order Higgs boson interactions $\mathcal{O}(h^n)$ with $n \geq 5$, but since they do not contribute to the observables at the order to which we calculate, we do not need to consider them further.

3.4 Effects on the Electroweak Oblique Parameters

We will now calculate the effect of the modified Higgs potential on the Peskin-Takeuchi parameters S and T. Contributions to S and T involving the dimension-6

operator $\mathcal{O}_6$ first appear at the two-loop level. At this order in perturbation theory, self-energy diagrams containing both trilinear and quartic Higgs self-interactions appear. Due to their modifications from $\bar{c}_6$ outlined above, these are manifest as non-zero corrections to S and T. However, as we will see shortly, contributions from the quartic Higgs self-interaction exactly cancel in these observables. It is also important to note that at this order in perturbation theory, there are no vertex or box diagrams that depend on $\bar{c}_6$ which contribute to observables involving light external fermions—light enough that their Yukawa couplings can be neglected. Furthermore, there are also no two-loop corrections to vertex or box diagrams involving both $\bar{c}_6$ and heavy external fermions in the electroweak observables, so the relevant two-loop $\bar{c}_6$ contributions to the self-energies must be separately gauge invariant.

3.4.1 Self-energy diagrams

To evaluate the electroweak oblique parameters S and T, all two-loop self-energy diagrams involving corrections from $\bar{c}_6$ need to be calculated. From the definitions of S and T, all Standard Model contributions are subtracted and so only terms proportional to $\bar{c}_6$ and $\bar{c}_6^2$ can remain. Working in the Feynman gauge, and discarding all two-loop diagrams that do not contain a contribution from $\bar{c}_6$, there are 26 diagrams for $\Delta\Pi_{ZZ}$, 26 for $\Delta\Pi_{WW}$, 5 for $\Delta\Pi_{Z\gamma}$ and 5 for $\Delta\Pi_{\gamma\gamma}$. An example Feynman diagram for each of the self-energies is shown in Fig. 3.3. From Eqs. (3.2.17) and (3.2.18), the ZZ, $Z\gamma$ and $\gamma\gamma$ self-energies are required to be evaluated at both zero and non-zero external momenta, whereas the WW self-energies are only required with zero external momenta.

The two-loop self-energies can be reduced to linear combinations of a set of basis integrals using the reduction algorithm from Tarasov [129], based on integration-by-parts relations [130]. This is analogous to the Passarino-Veltman reduction procedure for one-loop integrals. The reduction procedure is implemented in the MATHEMATICA package TARCER [131], which is part of the program FEYNCALC

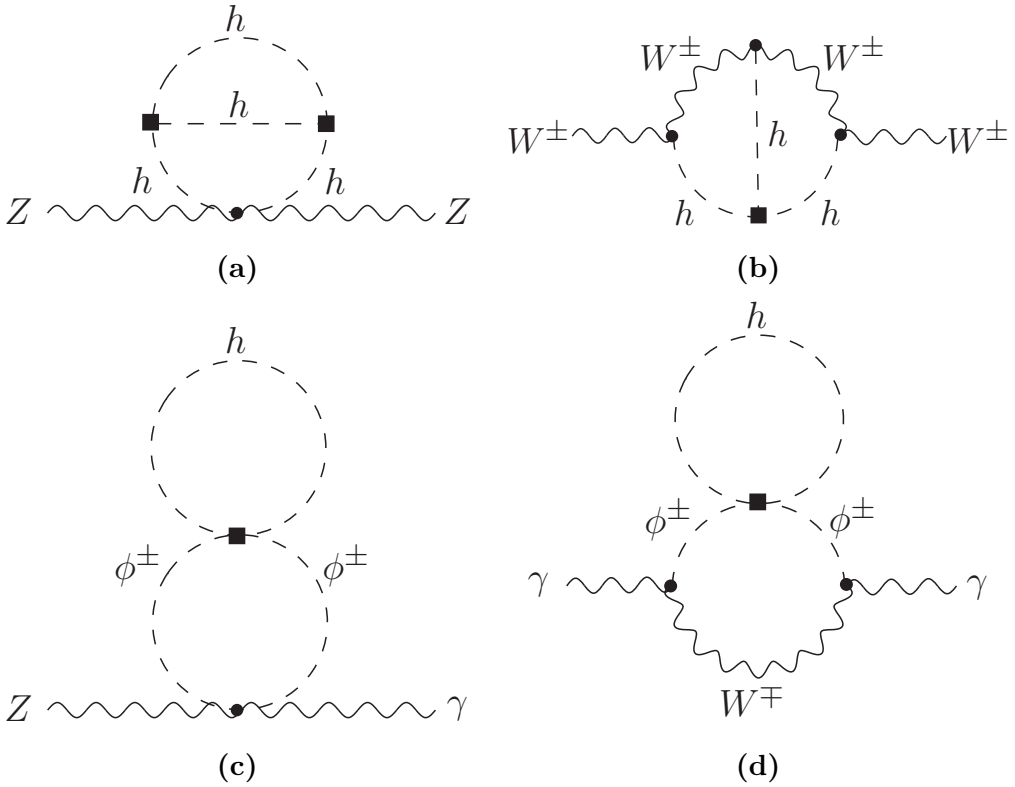


Figure 3.3: Example Feynman diagrams for the (a) ZZ, (b) WW, (c) $Z\gamma$ and (d) $\gamma\gamma$ two-loop self-energies. The black square represents a vertex where there is a contribution from the dimension-6 operator.

[132,133]. We generate the amplitudes for the self-energy diagrams using a model file in FeynArts [134], before using TARCER for the integral reduction. The reduction algorithm allows for the calculation of self-energies with non-zero external momenta and requires a total of eight basis integrals, but this reduces to a simplified set of two basis integrals when the external momenta are zero. A numerical implementation for the evaluation of all the basis integrals is given by the TSIL package [135]. Following the notation from TSIL, the eight basis integrals are,

$$\begin{aligned}
\mathbf{A}(x) &= -C \int d^d k \frac{1}{[k^2 - x]} \,, \\
\mathbf{B}(p, x, y) &= C \int d^d k \frac{1}{[k^2 - x][(k-p)^2 - y]} \,, \\
\mathbf{S}(p, x, y, z) &= -C^2 \int d^d k \int d^d q \frac{1}{[k^2 - x][q^2 - y][(k+q-p)^2 - z]} \,, \\
\mathbf{I}(x, y, z) &= \mathbf{S}(p, x, y, z)|_{p^2=0} \,,
\end{aligned}$$

$$
\begin{aligned}
\mathbf{T}(p,x,y,z) &= -\frac{\partial}{\partial x}\mathbf{S}(p,x,y,z)\ , \\
\mathbf{U}(p,x,y,z,u) &= C^2\int d^dk\int d^dq\frac{1}{[k^2-x][(k-p)^2-y][q^2-z][(q+k-p)^2-u]}\ , \\
\mathbf{V}(p,x,y,z,u) &= -\frac{\partial}{\partial y}\mathbf{U}(p,x,y,z,u)\ , \\
\mathbf{M}(p,x,y,z,u,v) &= -C^2\int d^dk\int d^dq \\
&\quad\times\frac{1}{[k^2-x][q^2-y][(k-p)^2-z][(q-p)^2-u][(k-q)^2-v]}\ ,
\end{aligned}
\tag{3.4.1}
$$

where $C \equiv (2\pi\mu)^{2\epsilon}/\pi^2$. Note that the **A** and **B** integrals are equivalent to the tadpole and self-energy integrals in Eqs. (1.2.7) and (1.2.8), respectively, and the correspondence between the notations for the basis integrals in TARCER, TSIL and other approaches is given in the appendix of Ref. [135].

Upon evaluating the basis integrals, they can be split up into their divergent and finite parts. For the **A**, **B** and **S** integrals these are,

$$
\begin{aligned}
\mathbf{A}(x) &= -\frac{x}{\epsilon} + A(x) + \epsilon A_\epsilon(x) + \mathcal{O}(\epsilon^2) \\
\mathbf{B}(p,x,y) &= \frac{1}{\epsilon} + B(p,x,y) + \epsilon B_\epsilon(p,x,y) + \mathcal{O}(\epsilon^2) \\
\mathbf{S}(p,x,y,z) &= -\frac{x+y+z}{2\epsilon^2} + \frac{1}{\epsilon}\left[A(x)+A(y)+A(z) - \frac{x+y+z}{2} + \frac{p^2}{4}\right] \\
&\quad + S(p,x,y,z) + A_\epsilon(x) + A_\epsilon(y) + A_\epsilon(z) + \mathcal{O}(\epsilon)
\end{aligned}
\tag{3.4.2}
$$

where $A(x)$, $A_\epsilon(x)$, $B(p,x,y)$, $B_\epsilon(p,x,y)$ and $S(p,x,y,z)$ are finite functions defined in Ref. [135]. The **A** and **B** integrals are one-loop basis integrals and only contain $1/\epsilon$ poles. The other six basis integrals are two-loop basis integrals and can contain $1/\epsilon^2$ poles in addition to $1/\epsilon$ poles. For self-energy diagrams with non-zero external momenta, only **A** and **I** are required.

3.4.2 Renormalisation

The leading-order contribution to the electroweak oblique parameters from the Standard Model (and modifications to the renormalisable couplings) begins at the one-loop

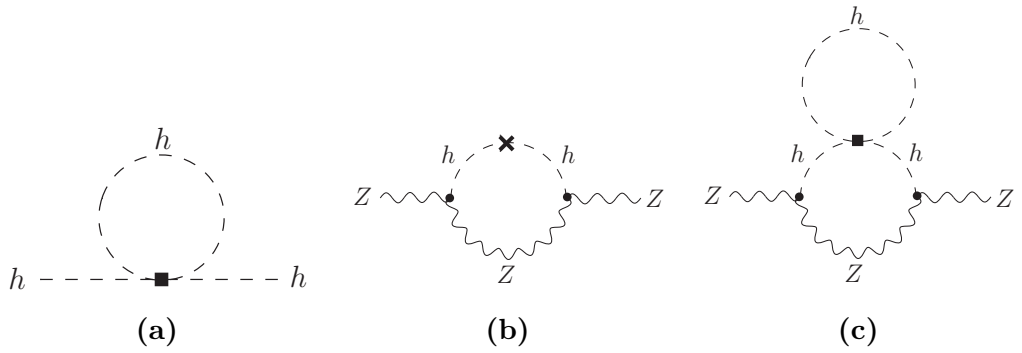


Figure 3.4: Feynman diagrams demonstrating the cancellation of the quartic Higgs self-coupling. (a) Shows the quartic contribution to the Higgs self-energy and (b) shows a counterterm insertion containing the quartic Higgs self-coupling which cancels with the contribution arising in (c). The black square represents a vertex where there is a contribution from the dimension-6 operator, and the black cross represents a counterterm insertion.

level[1]. This means, for the calculation of these parameters at next-to-leading (two-loop) order, no actual two-loop counterterms are needed. However, all the tree-level parts entering into the one-loop leading-order result, such as vertices and propagators, have a corresponding one-loop counterterm contribution in the next-to-leading-order calculation of the oblique parameters. Since contributions of the $\mathcal{O}_6$ operator and the corresponding $\bar{c}_6$ parameter only enter at the two-loop level, no renormalisation condition is needed for this parameter. All the other parameters are Standard Model parameters, and we perform the renormalisation procedure in the on-shell scheme which we defined in Chapter 1.

As already stated, we only take $\bar{c}_6$-dependent corrections into account. Since the one-loop results for S and T are independent of $\bar{c}_6$, in order to obtain a $\bar{c}_6$-dependent contribution at the two-loop level, the one-loop counterterm insertions must depend on $\bar{c}_6$. In the counterterm vertices, the only $\bar{c}_6$-dependent contributions originate from the field renormalisation of the Higgs boson, but these field renormalisation constants cancel together with the field renormalisation constants from the counterterm insertions in the Higgs boson propagator. The only contributing counterterms are

[1]We are assuming that the only higher-dimensional operators present are $\mathcal{O}_{2n}$, and in particular, the dimension-6 operators that give tree-level contributions to S and T are absent.

then the Higgs mass and tadpole counterterms inserted into the Higgs and Goldstone boson propagators.

It should be noted that the counterterm insertion into the Higgs boson propagator contains a part that is proportional to the quartic Higgs self-coupling. It originates from the on-shell Higgs mass counterterm, $\delta m_h^2 = \mathrm{Re}\,[\Sigma_{hh}(m_h^2)]$, and the corresponding contribution to the Higgs self-energy Σ_{hh} shown in Fig. 3.4a. The correction due to Feynman diagrams with a counterterm insertion into the Higgs propagator, an example of which is shown in Fig. 3.4b, cancels the corresponding quartic Higgs self-couplings arising in the two-loop self-energy diagrams, such as in Fig. 3.4c. Therefore, there is no contribution from the quartic Higgs self-coupling to the S and T parameters.

3.5 Results

We have performed the calculation of the contribution from the dimension-6 operator $\mathcal{O}_6$ to the electroweak oblique parameters S and T and we find that after renormalisation all UV divergences from the loop integrals cancel out, leaving non-zero and finite contributions to S and T. Specifically, this means that all $1/\epsilon^2$ and $1/\epsilon$ poles cancel out, and the dependence on the unphysical mass scale μ is also cancelled. Analytic expressions for the two-loop contributions to S and T from $\bar{c}_6$ are given in Appendix B, where the results are expressed in terms of the finite parts of the basis integrals in Eq. (3.4.1), for which we use TSIL for the numerical evaluation. For the numerical analysis, we take as input parameters [64]:

$$\begin{aligned} m_W &= 80.385 \text{ GeV} , & m_Z &= 91.1876 \text{ GeV} , \\ m_h &= 125 \text{ GeV} , & G_\mathrm{F} &= (1.16637870 \times 10^{-5}) \text{ GeV}^{-2} . \end{aligned} \tag{3.5.1}$$

The $W^\pm$ and Z boson masses are the pole masses, and the electroweak scheme is specified by the tree-level relations between the parameters [136]. We find that the

contribution of $\bar{c}_6$ to S and T is,

$$\begin{aligned} S &= -0.000138\ \bar{c}_6^2 + 0.000180\ \bar{c}_6\ , \\ T &= \quad 0.000206\ \bar{c}_6^2 - 0.000324\ \bar{c}_6\ . \end{aligned} \tag{3.5.2}$$

As there are no contributions from the quartic Higgs self-coupling, we can use the relation between $\bar{c}_6$ and κ_λ in Eq. (3.3.6) to write this result as,

$$\begin{aligned} S &= -0.000138\ (\kappa_\lambda^2 - 1) + 0.000456\ (\kappa_\lambda - 1)\ , \\ T &= \quad 0.000206\ (\kappa_\lambda^2 - 1) - 0.000736\ (\kappa_\lambda - 1)\ . \end{aligned} \tag{3.5.3}$$

The distinction between the contribution from two insertions of a modified Higgs self-coupling and a single insertion is made explicit here, since a term proportional to $(\kappa_\lambda^2 - 1)$ is exactly the contribution we get from two insertions.

The path of the κ_λ contribution in the S-T plane is shown in Fig. 3.5. The light blue ellipse shows the current 95% CL bound on the S and T parameters, as obtained by the Gfitter Group [122]. Also shown in the plot are possible future bounds on these parameters. The ellipses are constructed for $U = 0$ and are centred on $(0, 0)$. From the intersection points of the path of κ_λ in the S-T plane with the current ellipse, we estimate for the 95% CL a bound of:

$$-14.0 \leq \kappa_\lambda \leq 17.4\ . \tag{3.5.4}$$

The potential future limits are shown in Fig. 3.5. Similar bounds have also recently been derived using the observables m_W and $\sin\theta_{\mathrm{W}}$ instead of S and T [137]. The limits of Eq. (3.5.4) can be compared to the existing bounds from searches for di-Higgs final states and Higgs coupling measurements, which we stated in Sec. 3.1. While current limits from single Higgs production are stronger than bounds derived from electroweak precision measurements, they provide complementary information and thus can be used to extract a combined limit. It is interesting to note that the relative sensitivity of the bounds that we achieve is only possible because the path of the modifications to the trilinear self-coupling in the S-T plane is parallel to the

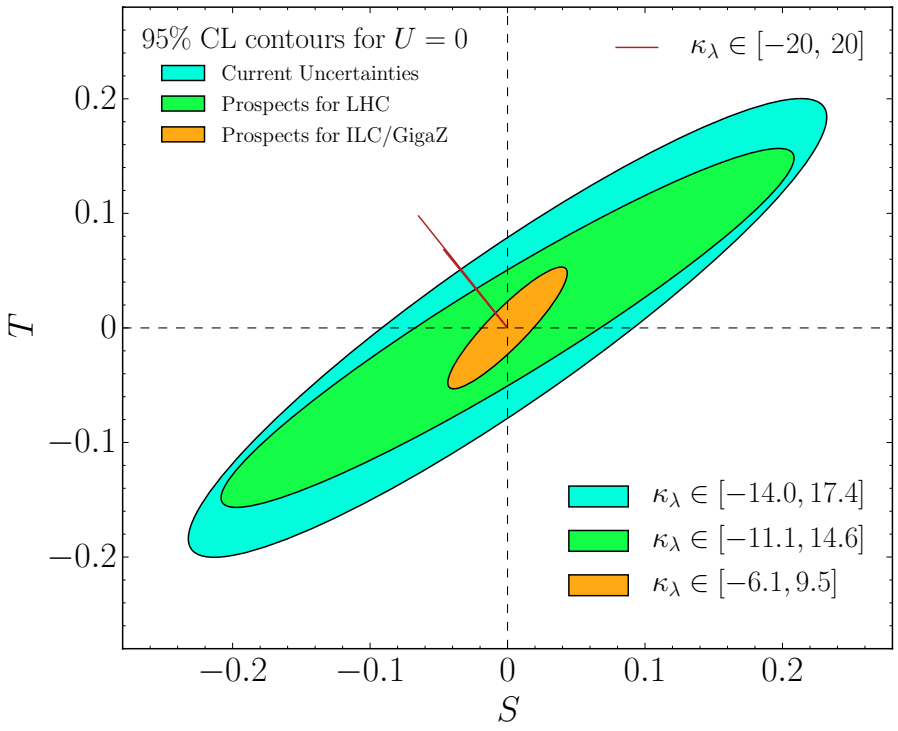


Figure 3.5: Current limits and projected sensitivities on κ_λ from the electroweak oblique parameters S and T. The light blue area in the S-T plane corresponds to the 95% CL region based on measurements at LEP and the LHC. The green and orange areas correspond to projected LHC and ILC/GigaZ sensitivities, respectively. The longer (shorter) thin red lines show the shift in S and T as κ_λ extends up to -20 ($+20$). The intersection of these lines with the current limits and projected sensitivities gives the ranges of κ_λ as shown in the figure.

axis of tightest experimental constraint. If this was not the case, then the obtainable limits would be orders of magnitude higher.

3.6 Conclusions

In this chapter, we derived limits on the trilinear self-coupling of the Higgs boson by calculating the effects of modifications to this coupling on the electroweak oblique Peskin-Takeuchi parameters S and T. We found that they are comparable to those limits obtained from direct searches and from indirect effects on single Higgs production, thus verifying the self-consistency of the bounds. Furthermore, they

can provide complementary information since they are derived from an independent approach with separate uncertainties.

To carry out the calculation in a gauge-invariant way, we parameterised the modifications using an effective field theory. Since we found that the contributions from the quartic Higgs self-coupling did not contribute at the order to which we calculated, we only had to include a single dimension-6 operator $\mathcal{O}_6$. This allowed the limits on the Wilson coefficient $\bar{c}_6$ to be directly translated into limits on the modified coupling κ_λ.

It is worth considering the implications of such a large modification to the Higgs trilinear self-coupling on the validity of the effective field theory approach. The effective field theory terms in Eq. (3.3.4) could instead be written as $-(\Phi^\dagger\Phi)^n/\Lambda^{2n-4}$, where Λ is the scale of new physics and an $\mathcal{O}(1)$ Wilson coefficient is assumed. The bounds on $\bar{c}_6$ could then be translated into bounds on the new physics scale via $\Lambda \gtrsim (v/\sqrt{2}) \times \sqrt{15.5/\bar{c}_6}$. This would then imply that Λ could be close to the electroweak scale, given the allowed values of $\bar{c}_6$ that we find. Furthermore, the large dimension-6 term would cause problems with the global properties of the scalar potential, resulting in it becoming unstable. However, if new physics was to enter at such a scale, it would no longer make sense to truncate an effective field theory to only dimension-6 terms, and one would have to include the higher dimensional operators regardless, which could address the issues with stability. It would also probably be necessary to include other effects besides a modification to the Higgs self-couplings, since it is likely that new physics that allows such a large modification to the Higgs trilinear self-coupling would not be restricted to this modification alone. We leave these investigations to future work, and in this study we only focused on the effects of performing this modification in isolation, independent of a specific model of new physics.

Chapter 4

Heavy Neutrinos from Threshold-Resummed Gluon Fusion

In this chapter, we perform our third investigation into physics beyond the Standard Model by considering the production of heavy neutrinos from the gluon fusion process. Heavy neutrinos, as we explained in Chapter 1, arise in extensions of the Standard Model to provide a mechanism for generating small sub-eV scale masses for the light neutrinos. These heavy neutrinos have not yet been observed, but could be produced at colliders such as the LHC since they couple directly to the electroweak bosons via their mixing. The conventional wisdom suggests that the charged-current Drell-Yan process involving the $W^{\pm}$ boson should be the dominant production mode. However, we study the impact of soft-gluon corrections to the production cross section from gluon fusion by resumming threshold logarithms.

4.1 Motivation

The observation that neutrinos oscillate [32] implies that they have non-zero masses and non-trivial mixings between their gauge and mass eigenstates [138, 139]. Many

models that are proposed to produce small non-zero masses for the neutrinos, such as the various seesaw mechanisms [140–143], involve the existence of right-handed neutrinos. This results in the generation of both heavy and light mass eigenstates for the neutrinos, with the heavy neutrinos expected to lie at the electroweak scale or beyond [144]. Through their mixing with the gauge eigenstates, these heavy mass eigenstates can be produced at a collider since they interact directly with the electroweak bosons [145].

However, no direct detection of a heavy neutrino at a collider has yet been made [146]. Clearly if such a scenario occurred, it would confirm the existence of physics beyond the Standard Model and answer the question of why the neutrinos in the Standard Model have non-zero masses which are far smaller than those of all the other fermions. To help towards this goal, it is important to study the processes that can produce such a heavy neutrino. The dominant production mode at the LHC is currently believed to be the charged-current Drell-Yan process, with the vector boson fusion process dominating at much higher neutrino masses. However, there has recently been interest in the study of gluon fusion production processes for heavy neutrinos [147]. Although gluon fusion is known to dominate at very high collider energies—namely those above ~ 40 TeV—it has been assumed that it is not so important for LHC phenomenology [148].

These statements rely on the assumption that the leading-order calculation for the cross section of a heavy neutrino via gluon fusion is accurate. While the higher-order corrections from quantum chromodynamics to the charged-current Drell-Yan and vector boson fusion processes are fairly small [148], it is known that the leading-order calculation for the gluon fusion production modes for the Higgs boson [149, 150], heavy scalars [151] and heavy pseudoscalars [152, 153] vastly underestimates the full cross section. Therefore, it is reasonable to expect that higher-order corrections to the heavy neutrino production mode, which is of a similar structure, could similarly be large.

Although a complete higher-order calculation would be computationally difficult since the gluon fusion process arises at leading order at the one-loop level, we carry out a resummation of large logarithms that arise when real gluon radiation that is emitted into the final state is forced to be soft. It is known that for the gluon fusion production of a Higgs boson, the resummation of these threshold logarithms captures the dominant contribution to the higher-order cross section [154]. Therefore, it is reasonable to suppose that such a result may hold for the production of a heavy neutrino via gluon fusion, and we will carry out this computation.

The remainder of this chapter is structured as follows: in Sec. 4.2, we will introduce the various production modes of a heavy neutrino and show explicitly the calculation of the partonic cross section for a heavy neutrino via the Drell-Yan process. We will further define the hadronic cross section and show how higher-order corrections to the Drell-Yan process are computed. In Sec. 4.3, we will describe the threshold resummation framework and define the computational setup that we use to calculate the resummed cross section for gluon fusion. In Sec. 4.4 we will present the results and a comparison of the various production modes, before concluding in Sec. 4.5.

4.2 Production of Heavy Neutrinos

4.2.1 Heavy Neutrino Model

We consider the existence of right-handed neutrinos, N_{Ri}, which are singlets under the Standard Model gauge groups. Following the formalism in Ref. [155], we write the gauge eigenstates of the neutrinos as[1] $(\nu_{L1}, \nu_{L2}, \nu_{L3}, N^c_{R1}, \ldots, N^c_{Rn})$, where there are n right-handed neutrinos. This basis can then be rotated into the mass basis by,

$$\begin{pmatrix} \nu_{Li} \\ N^c_{Rj} \end{pmatrix} = \begin{pmatrix} U_{3\times 3} & V_{3\times n} \\ X_{n\times 3} & Y_{n\times n} \end{pmatrix} \begin{pmatrix} \nu_m \\ N^c_{m'} \end{pmatrix} , \tag{4.2.1}$$

[1] Note that the neutrinos $\nu_{L1,2,3}$ are equivalent to the alternative notation $\nu_{e,\mu,\tau}$.

where the diagonalising matrix is equivalent to the matrix $\mathcal{V}$ from Eq. (2.4.7), which was defined for $n = 3$. This allows the gauge eigenstates ν_ℓ to be written explicitly in the mass basis as,

$$\nu_\ell = \sum_{m=1}^{3} U_{\ell m}\nu_m + \sum_{m'=1}^{n} V_{\ell m'}N^c_{m'} \ . \tag{4.2.2}$$

This means that the mass eigenstates of the heavy neutrinos, N^c_m, interact with the electroweak bosons through their mixing with the gauge eigenstates ν_ℓ, for which electroweak interaction terms in the Standard Model exist. Since N_{Ri} are singlets under the Standard Model gauge groups, these are the only couplings that exist between the Standard Model particles and the heavy neutrinos, and so the couplings are suppressed by the 'active-heavy' mixing matrix $V_{\ell m'}$. For simplicity we only consider the lightest heavy neutrino state, which we will now denote as N, with a mass of m_N. Current constraints on the mixing from electroweak precision data are $|V_{\ell N}| \lesssim 10^{-2} - 10^{-1}$ [144, 156], for a heavy neutrino at the electroweak scale.

We can now write the electroweak interaction Lagrangian for the neutrino sector as,

$$\begin{aligned}\mathcal{L}_{\text{Neutrino}} = &-\frac{g}{\sqrt{2}}W^+_\mu \sum_{\ell=e}^{\tau}\sum_{m=1}^{3} \ \overline{\nu}_m \ U^*_{\ell m} \ \gamma^\mu P_L \ell^- - \frac{g}{\sqrt{2}}W^+_\mu \sum_{\ell=e}^{\tau} \ \overline{N^c} \ V^*_{\ell N} \ \gamma^\mu P_L \ell^- \\ &- \frac{g}{2\cos\theta_{\text{W}}}Z_\mu \sum_{\ell=e}^{\tau}\sum_{m=1}^{3} \ \overline{\nu}_m \ U^*_{\ell m} \ \gamma^\mu P_L \nu_\ell - \frac{g}{2\cos\theta_{\text{W}}}Z_\mu \sum_{\ell=e}^{\tau} \ \overline{N^c} \ V^*_{\ell N} \ \gamma^\mu P_L \nu_\ell \\ &- \frac{g m_N}{2m_W} h \sum_{\ell=e}^{\tau} \ \overline{N^c} \ V^*_{\ell N} P_L \nu_\ell + \text{h.c.} \ . \end{aligned} \tag{4.2.3}$$

The final term involving the Higgs boson h results from the Yukawa interactions, which also generate the Dirac masses for the neutrinos. This interaction term only exists due to the presence of the right-handed neutrino states, and would not exist in the Standard Model alone. From this Lagrangian, the Feynman rules for the interactions involving N can be extracted, and they are all proportional to the active-heavy mixing $V_{\ell N}$.

4.2.2 Drell-Yan Production

We will now consider the production of a heavy neutrino. It can be produced at tree level via its interactions with the gauge bosons $W^\pm$ and Z. A resonant $W^{\pm *}$

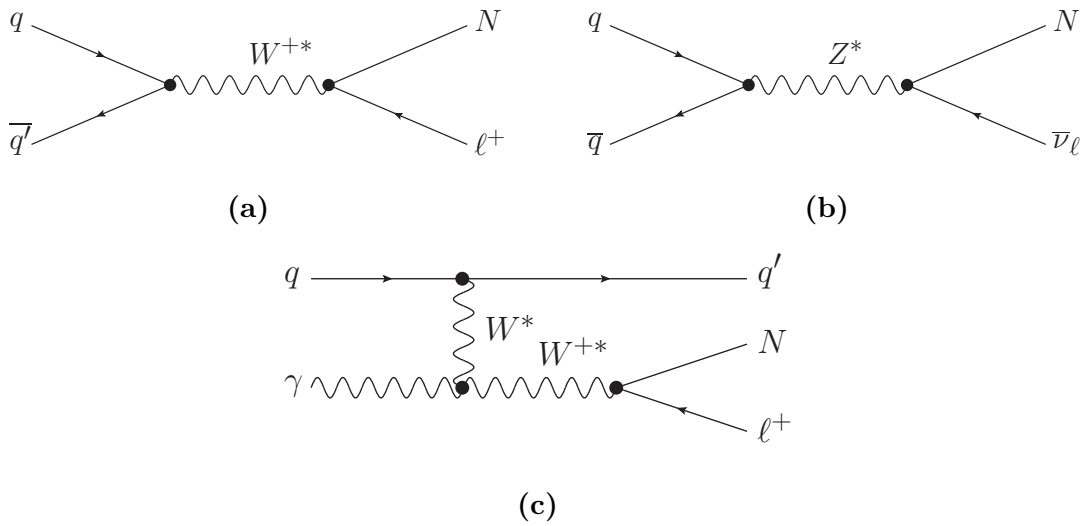


Figure 4.1: Feynman diagrams for the production of a heavy neutrino N via (a) CC DY, (b) NC DY and (c) VBF at leading order.

boson can be produced which decays to the heavy neutrino N and charged lepton $\ell^\pm$ through the charged-current (CC) Drell-Yan (DY) process [157]. Alternatively, the resonance can be a Z^* boson which decays to the heavy neutrino N and light neutrino ν_ℓ, and this is known as the neutral-current (NC) Drell-Yan process. These two processes can be expressed as,

$$\text{CC DY}: \qquad q\overline{q}' \to W^{\pm *} \to N\ell^\pm \ , \qquad q, q' \in \{u, c, d, s, b\} \ , \tag{4.2.4}$$

$$\text{NC DY}: \qquad q\overline{q} \to Z^* \to N\overset{(-)}{\nu}_\ell \ , \qquad q \in \{u, c, d, s, b\} \ . \tag{4.2.5}$$

A further process which can generate a heavy neutrino is the vector boson fusion (VBF) process [158] initiated by a photon,

$$\text{VBF}: \qquad q\gamma \overset{W^\pm\gamma \to N\ell^\pm}{\longrightarrow} N\ell^\pm \, q' \ , \qquad q, q' \in \{u, c, d, s, b\} \ . \tag{4.2.6}$$

All three partonic-level processes are shown by the Feynman diagrams in Fig. 4.1 at leading order (LO). There are also additional diagrams for the VBF process with different topologies. As seen in Fig. 4.1c, the VBF process is driven by the $W^\pm\gamma \to N\ell^\pm$ subprocess which receives longitudinal $W^\pm$ boson enhancements for $m_N \gg m_W$ [159].

Drell-Yan Production at Leading Order

For the CC DY process, we will now calculate the cross section. At the partonic level, the matrix element is constructed from the Feynman rules by,

$$\begin{aligned}\mathcal{M}(q\overline{q'} \to & W^{+*} \to N\ell^+) = \\ & \overline{v}(p_{q'}) \left(-i\frac{g}{\sqrt{2}}\gamma^\mu V^*_{qq'} P_L\right) u(p_q) \times \frac{-ig_{\mu\nu}}{\hat{s} - m_W^2 + im_W\Gamma_W} \\ & \times \overline{u}(p_N) \left(-i\frac{g}{\sqrt{2}}\gamma^\nu V^*_{\ell N} P_L\right) v(p_\ell) \ , \end{aligned} \tag{4.2.7}$$

where u and v are Dirac spinors, $V_{qq'}$ is the CKM mixing between the quarks q and q', and Γ_W is the decay width of the W^+ boson. Furthermore, p are the four-momenta of the incoming and outgoing particles, and $\hat{s} \equiv (p_q + p_{q'})^2$ is the total centre-of-mass energy squared. The propagator for the W^+ boson is written in the Feynman gauge, and includes the decay width since the W^+ is unstable—this leads to a Breit-Wigner resonance distribution in the cross section [160]. However, the choice of gauge is unimportant since the propagator is contracted with two incoming fermions which can be assumed to be massless, so the gauge-dependent $p^\mu p^\nu$ contribution from the propagator in Eq. (1.2.3) vanishes.

Carrying out the computation in the centre-of-mass frame, the four-momenta of the particles can be expressed as,

$$\begin{aligned} p_q &= \frac{\sqrt{\hat{s}}}{2}(1,\ 0,\ 0,\ 1) \ , \qquad p_{q'} = \frac{\sqrt{\hat{s}}}{2}(1,\ 0,\ 0,\ -1) \ , \\ p_\ell &= (E_\ell,\ |\vec{p}_\ell| \sin\theta_\ell \cos\phi_\ell,\ |\vec{p}_\ell| \sin\theta_\ell \sin\phi_\ell,\ |\vec{p}_\ell| \cos\theta_\ell) \ , \\ p_N &= (E_N,\ -|\vec{p}_\ell| \sin\theta_\ell \cos\phi_\ell,\ -|\vec{p}_\ell| \sin\theta_\ell \sin\phi_\ell,\ -|\vec{p}_\ell| \cos\theta_\ell) \ , \end{aligned} \tag{4.2.8}$$

where θ_ℓ and ϕ_ℓ are the polar and azimuthal angles, respectively, of the charged lepton $\ell^\pm$ with respect to incoming quark q. The kinematic setup for the incoming and outgoing particles is shown in Fig. 4.2. From the definition of $\hat{s}$, it is clear that $E_q = E_{q'} = \sqrt{\hat{s}}/2$. Assuming that $m_N \gg m_\ell$, we can derive the relations,

$$E_N = \frac{\sqrt{\hat{s}}}{2}(1 + r_N) \ , \qquad E_\ell = \frac{\sqrt{\hat{s}}}{2}(1 - r_N) \ , \qquad |\vec{p}_\ell| = \frac{\sqrt{\hat{s}}}{2}(1 - r_N) \ , \tag{4.2.9}$$

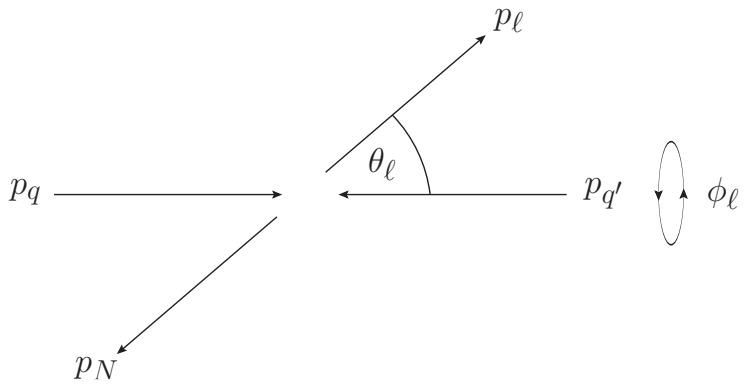


Figure 4.2: Kinematic setup for the incoming and outgoing particles in the CC DY process.

where $r_N \equiv m_N^2/\hat{s}$. The calculation can then proceed by making use of states of definite helicity [161]. These are eigenstates of the helicity operator, which projects the spin of a particle along its direction of propagation. A right-handed helicity state has its spin parallel to its direction of propagation, while for a left-handed helicity state they are anti-parallel. The W^+ boson only couples to left-handed chiral spinors and left-handed chiral anti-spinors. These correspond to left-handed (helicity) particles and right-handed anti-particles, respectively. The q, $\overline{q'}$ and ℓ^+ are assumed to be massless, so this forces their helicity states to be q_L, $\overline{q'}_R$ and ℓ_R^+, respectively, where the subscript denotes the helicity. The N is massive so its helicity and chiral states do not coincide with each other, and thus we need to consider both the N_L and N_R helicity states[2]. This leads to two possible matrix elements: $\mathcal{M}(q_L\overline{q'}_R \to W^{+*} \to N_L\ell_R^+)$ and $\mathcal{M}(q_L\overline{q'}_R \to W^{+*} \to N_R\ell_R^+)$.

The four-component Dirac spinors can then be written in terms of two-component helicity eigenstates as,

$$u_L(p_q) = \sqrt{2E_q}\begin{pmatrix}\chi_L(\hat{z})\\ 0\end{pmatrix}, \qquad v_R(p_{q'}) = -\sqrt{2E_{q'}}\begin{pmatrix}\chi_L(-\hat{z})\\ 0\end{pmatrix},$$

[2]Note that since we are considering the mass eigenstates of the heavy neutrino, these can have both left-handed and right-handed chiral states. On the other hand, the gauge eigenstates only exist in right-handed chiral states, by construction.

$$v_R(p_\ell) = -\sqrt{2E_\ell}\begin{pmatrix}\chi_L(\hat{p}_\ell)\\ 0\end{pmatrix}\,, \tag{4.2.10}$$

$$u_R(p_N) = \begin{pmatrix}\sqrt{E_N - |\vec{p}_\ell|}\;\chi_R(-\hat{p}_\ell)\\ \sqrt{E_N + |\vec{p}_\ell|}\;\chi_R(-\hat{p}_\ell)\end{pmatrix}\,, \qquad u_L(p_N) = \begin{pmatrix}\sqrt{E_N + |\vec{p}_\ell|}\;\chi_L(-\hat{p}_\ell)\\ \sqrt{E_N - |\vec{p}_\ell|}\;\chi_L(-\hat{p}_\ell)\end{pmatrix}\,,$$

where $\hat{z}$ is the axis of the incoming quark q and $\hat{p}_\ell$ is the axis of the charged lepton. The relevant helicity eigenstates are,

$$\chi_L(\hat{z}) = \begin{pmatrix}0\\1\end{pmatrix}\,, \qquad \chi_L(-\hat{z}) = \begin{pmatrix}-1\\0\end{pmatrix}\,, \qquad \chi_L(\hat{p}_\ell) = \begin{pmatrix}-\sin\frac{\theta_\ell}{2}\\ \cos\frac{\theta_\ell}{2}\end{pmatrix}\,,$$

$$\chi_L(-\hat{p}_\ell) = \begin{pmatrix}\cos\frac{\theta_\ell}{2}\\ \sin\frac{\theta_\ell}{2}\end{pmatrix}\,, \qquad \chi_R(-\hat{p}_\ell) = \begin{pmatrix}\sin\frac{\theta_\ell}{2}\\ -\cos\frac{\theta_\ell}{2}\end{pmatrix}\,, \tag{4.2.11}$$

and we have used the azimuthal symmetry to set $\phi_\ell = 0$. The Dirac spinors can then be inserted into Eq. (4.2.7) for the two helicity combinations and contracted with the Dirac matrices to get the amplitudes,

$$\begin{aligned}\mathcal{M}(q_L\overline{q}'_R \to W^{+*} \to N_L\ell^+_R) &= -i\frac{g^2}{2}V^*_{qq'}V^*_{\ell N}\;(1-\cos\theta_\ell)\;\frac{\hat{s}\sqrt{1-r_N}}{\hat{s}-m_W^2+im_W\Gamma_W}\,,\\ \mathcal{M}(q_L\overline{q}'_R \to W^{+*} \to N_R\ell^+_R) &= -i\frac{g^2}{2}V^*_{qq'}V^*_{\ell N}\;\sin\theta_\ell\;\frac{\hat{s}\sqrt{(1-r_N)r_N}}{\hat{s}-m_W^2+im_W\Gamma_W}\,.\end{aligned} \tag{4.2.12}$$

The dependence on θ_ℓ can be understood from the conservation of spin along the axis $\hat{z}$. When $\theta_\ell = 0$, both processes are forbidden as this does not conserve spin. The $LR \to LR$ process is maximal when $\theta_\ell = \pi$ and the $LR \to RR$ process is maximal when $\theta_\ell = \pm\pi/2$. Furthermore, in the massless limit $r_N \to 0$, then only the $LR \to LR$ process contributes—this is expected because the interaction with the W^+ boson forces N into a left-handed chiral state, which means that it is also forced into a left-handed helicity state if it is massless.

The full partonic cross section is calculated by squaring the matrix elements, averaging over the initial-state degrees of freedom, summing over the final-state degrees of freedom, and integrating over the phase space. In general, for the scattering of

massless particles A and B into n final-state particles, this is given by [119],

$$\hat{\sigma}(A+B \to 1+\ldots+n) = \frac{1}{2\hat{s}} \frac{1}{(2s_A+1)(2s_B+1)N_c^A N_c^B} \sum_{\text{d.o.f.}} \int |\mathcal{M}|^2 \, d\text{PS}_n \; . \quad (4.2.13)$$

Here, $2s_A+1$ and N_c^A are the number of spin states and colour states, respectively, of particle A, with the quantities for particle B defined analogously. The sum over the degrees of freedom simply consists of adding the squares of the helicity matrix elements. The n-body differential phase space is,

$$d\text{PS}_n = (2\pi)^4 \delta^{(4)}(P - p_1 - \ldots - p_n) \prod_{k=1}^{n} \frac{d^3 p_k}{(2\pi)^3 \; 2E_k} \; , \quad (4.2.14)$$

where P is the total outgoing four-momentum. The phase space accounts for all the possible kinematically-allowed unique configurations that can be produced by the system. If there are more allowed configurations for a system to have, then it has a larger phase-space volume and a correspondingly higher probability of occurring.

For the CC DY process, there are two outgoing particles, and the incoming quarks each have two spin states and three colour states. With one massive particle in the final state, the two-body differential phase space takes the form,

$$d\text{PS}_2 = \frac{1}{32\pi^2}(1 - r_N) \; d\cos\theta_\ell \; d\phi_\ell \; . \quad (4.2.15)$$

Inserting this into Eq. (4.2.13) and performing the sum over the helicity matrix elements we get,

$$\hat{\sigma}(q\overline{q'} \to W^{+*} \to N\ell^+) = G_\text{F}^2 \; \frac{m_W^4 |V_{qq'}|^2 |V_{\ell N}|^2}{36\pi} \; \frac{\hat{s}(1-r_N)^2(2+r_N)}{(\hat{s}-m_W^2)^2 + (\Gamma_W m_W)^2} \; . \quad (4.2.16)$$

For the NC DY process, the calculation proceeds very similarly, with the main difference arising because the Z boson also couples to right-handed chiral states. Since the light (anti-)neutrino ν_ℓ ($\overline{\nu}_\ell$) can only exist in the left-handed (right-handed) helicity state, there are four possible helicity matrix elements to consider. After calculating these, the partonic cross section is,

$$\hat{\sigma}(q\overline{q} \to Z^* \to N\overline{\nu}_\ell) = G_\text{F}^2 \; \frac{m_Z^4 |V_{\ell N}|^2 (c_V^{q2} + c_A^{q2})}{72\pi} \; \frac{\hat{s}(1-r_N)^2(2+r_N)}{(\hat{s}-m_Z^2)^2 + (\Gamma_Z m_Z)^2} \; , \quad (4.2.17)$$

where the vector and axial-vector couplings of the Z boson are,

$$c_V^q = T_3^q - 2Q^q \sin^2\theta_{\mathrm{W}} \ , \qquad c_A^q = T_3^q \ . \tag{4.2.18}$$

Here, Q^q is the charge and T_3^q is the third component of weak isospin of the quark q.

Hadronic Cross Section

At a hadron collider such as the LHC, the incoming particles are protons, which consist of strongly-interacting quarks and gluons. The quarks that contribute to the quantum numbers of the proton are the 'valence quarks'—for a proton these are two up quarks and one down quark—along with a 'sea' of pairs of quarks and anti-quarks, and gluons that carry the strong force. In a collision between two protons, a hard interaction takes place between a parton from each of the protons. These partons carry fractions ξ_1 and ξ_2 of the total longitudinal momentum of their parent protons. The probability density for each of these partons to carry such fractions of momentum is given by the parton distribution functions (PDFs) $f_{i/p}(\xi_1, \mu_f)$ and $f_{j/p}(\xi_2, \mu_f)$ where parton 1 is type i and parton 2 is type j. These are process independent and universal—this is important, because it allows independent predictions to be made for hadronic cross sections. Furthermore, μ_f is the factorisation scale, which can be understood as representing the boundary between the high-scale perturbative interactions and the low-scale non-perturbative physics. The PDFs are evolved to this scale through DGLAP evolution [162–164] from a scale at which they have been extracted through fits to experimental data.

An inclusive scattering process consists of $pp \to V + X$, where p are the colliding protons, V is a colour-singlet final state and X are any other final-state products produced in the collision. To calculate the cross section for this process, we can make use of the factorisation theorem [165], which allows the cross section to be

written as a convolution between the PDFs and the partonic cross section as,

$$\sigma(pp \to V + X) = \frac{1}{1+\delta_{ij}} \sum_{i,j} \int_{\tau_0}^{1} d\xi_1 \int_{\frac{\tau_0}{\xi_1}}^{1} d\xi_2 \Big[f_{i/p}(\xi_1, \mu_f) f_{j/p}(\xi_2, \mu_f) + (i \leftrightarrow j) \Big] \hat{\sigma}(ij \to V) \ . \quad (4.2.19)$$

The sum is over all parton types i and j for which the partonic process $ij \to V$ is possible. We have used the definitions,

$$\hat{s} = \xi_1 \xi_2 s \equiv \tau s \ , \qquad \tau_0 = \frac{\hat{s}_{\min}}{s} \ , \quad (4.2.20)$$

where s is the hadronic centre-of-mass energy squared, $\hat{s}$ is the partonic centre-of-mass energy squared and $\hat{s}_{\min}$ is the kinematic threshold below which the process $ij \to V$ is kinematically forbidden. For the case of Drell-Yan production of a heavy neutrino, the threshold is $\hat{s}_{\min} = m_N^2$. The term $(i \leftrightarrow j)$ is added because the assignment of each parton type to a particular proton was arbitrary and the process where the parton assignments are reversed is equally likely. The Kronecker delta symbol δ_{ij} is used as a symmetry factor because the reverse assignment is not required when the parton types are identical. It can also sometimes be useful to use the definition $\tau \equiv \xi_1 \xi_2$ to replace one of the ξ integrals with an integral over τ, so that Eq. (4.2.19) becomes,

$$\sigma(pp \to V + X) = \frac{1}{1+\delta_{ij}} \sum_{i,j} \int_{\tau_0}^{1} d\tau \int_{\tau}^{1} \frac{d\xi}{\xi} \Big[f_{i/p}(\xi, \mu_f) f_{j/p}(\tau/\xi, \mu_f) + (i \leftrightarrow j) \Big] \hat{\sigma}(ij \to V) \ . \quad (4.2.21)$$

Since the PDFs are only known numerically, their convolution with the partonic cross sections must be carried out numerically as well. This can be performed using a Monte Carlo integration tool, such as the CUBA libraries [166]. Here, the integrals over ξ_1 and ξ_2 (or τ and ξ) are shifted and scaled such that they run from 0 to 1. This allows the integration to be carried out by randomly sampling a value for each integration variable, from which the corresponding integrand can be calculated, with appropriate Jacobian factors to account for the scaling. The average of this integrand over many samples of the integration variables is then the value of the

integral. Furthermore, this method allows cuts on the phase space of the final-state particles to be included by leaving the partonic cross sections unintegrated over the $\cos\theta_\ell$ variable—these correspond to the differential cross sections $d\hat{\sigma}/d\cos\theta_\ell$. The integration over $\cos\theta_\ell$ is then combined with the integration over the PDFs into a three-dimensional Monte Carlo integration. The partonic four-momenta corresponding to a particular phase-space point can then be boosted into the lab frame with a Lorentz factor $\beta = (\xi_1 - \xi_2)/(\xi_1 + \xi_2)$, and if it falls outside the region allowed by the cuts, the value for this integrand is set to zero.

Drell-Yan Production at Next-to-Leading Order

The previous calculation only included the leading-order contribution to the Drell-Yan process. Since the initial state at the partonic level includes quarks which interact via the strong force, higher-order corrections from quantum chromodynamics will modify the total cross section. These arise from virtual loops involving gluons, real emissions of gluons into the final state and gluon-induced processes with quarks emitted into the final state. Importantly, since only the initial state contains strongly-interacting particles—with the new physics arising solely in the final state—the next-to-leading-order (NLO) corrections will essentially factor out and the calculation will be equivalent to that for the Standard Model Drell-Yan production of lepton pairs.

The squared matrix element and phase space for the NLO corrections to an n-body process can be written schematically as,

$$\begin{aligned}\left|\mathcal{M}^{\mathrm{NLO}}\right|^2 d\mathrm{PS} &= \left|\mathcal{M}^{\mathrm{LO}} + \mathcal{M}^{\mathrm{V}}\right|^2 d\mathrm{PS}_n + \left|\mathcal{M}^{\mathrm{RE}}\right|^2 d\mathrm{PS}_{n+1} \\ &= \left|\mathcal{M}^{\mathrm{LO}}\right|^2 d\mathrm{PS}_n + 2\mathrm{Re}\left[\mathcal{M}^{\mathrm{LO}\,*}\mathcal{M}^{\mathrm{V}}\right] d\mathrm{PS}_n \\ &\quad + \left|\mathcal{M}^{\mathrm{RE}}\right|^2 d\mathrm{PS}_{n+1} + \mathcal{O}(\alpha_s^2)\ . \end{aligned} \tag{4.2.22}$$

Note that the real emission of an extra particle into the final state involves an $(n+1)$-body phase-space calculation. The virtual and real emission contributions

are also both IR divergent. However, due to the KLN theorem which we mentioned in Chapter 1, these divergences necessarily cancel, as we shall shortly see.

To compute the NLO corrections and allow a systematic treatment of the IR divergences, we can use the phase-space slicing method [167, 168]. In this method, the $(n+1)$-body phase space for the real emission contribution is sliced into soft and collinear regions with the small dimensionless parameters $\delta_{\rm S}$ and $\delta_{\rm C}$. A region of the phase space is soft if the partonic real emission with energy E_j has,

$$E_j < \frac{\sqrt{\hat{s}}}{2}\delta_{\rm S} \ . \tag{4.2.23}$$

Similarly, a region of the phase space is collinear if,

$$\hat{s}_{ik}, |\hat{t}_{ik}| < \delta_{\rm C}\hat{s} \ , \tag{4.2.24}$$

where we have used the definitions,

$$\hat{s}_{ik} \equiv (p_i + p_k)^2 \ , \qquad \hat{t}_{ik} \equiv (p_i - p_k)^2 \ , \tag{4.2.25}$$

for any partonic four-momenta p_i or p_k in the initial or final states. The dependence on the dimensionless parameters $\delta_{\rm S}$ and $\delta_{\rm C}$ cancels, provided $\delta_{\rm C} \ll \delta_{\rm S} \ll 1$. This slicing then defines four regions of the $(n+1)$-body phase space: soft non-collinear (S), soft-collinear (SC), hard-collinear (HC) and hard non-collinear (HNC) regions.

The total hadronic cross section for the CC DY heavy neutrino production can then be expressed as a sum from each of these four regions, as well as the leading-order and virtual contributions:

$$\sigma^{\rm NLO}(pp \to N\ell^{\pm} + X) = \sigma^{\rm LO} + \sigma^{\rm V} + \sigma^{\rm S} + \sigma^{\rm SC} + \sigma^{\rm HC} + \sigma^{\rm HNC} \ . \tag{4.2.26}$$

Note that we have also included the conjugate process with the charged lepton ℓ^-, which is produced for the conjugate assignment of initial-state quarks, since both processes contribute to the CC DY rate. The partonic cross sections for these two processes are equal. It is implicit here that the leading-order and virtual corrections involve the integration of the partonic cross sections over a two-body phase space,

while the other terms a priori involve a three-body phase space. However, when the real emission is soft, the leading-order process remains unaffected by the extra emission and the phase space of its particles are unchanged due to factorisation [167]. Therefore, the real-emission matrix element in the soft limit can be written as a product of the leading-order matrix element and an extra factor, with an integration over a two-body phase space. Furthermore, since the IR divergences arise due to soft and collinear emissions, the hard non-collinear region is finite and its cross section can be evaluated numerically with Monte Carlo integration.

To address the IR divergences, the virtual, soft and soft-collinear corrections are computed in $d = 4 - 2\epsilon$ dimensions. The IR divergences then arise explicitly as $1/\epsilon^2$ and $1/\epsilon$ poles, similar to the behaviour for UV divergences we have seen in previous chapters. The virtual corrections to the partonic cross section take the form [168],

$$\hat{\sigma}^{\mathrm{V}} = \hat{\sigma}^{\mathrm{LO}} \frac{\alpha_s(\mu_r)}{2\pi} C_F C_\epsilon(\hat{s}) \left(\frac{A_2^{\mathrm{V}}}{\epsilon^2} + \frac{A_1^{\mathrm{V}}}{\epsilon} + A_0^{\mathrm{V}} \right) , \tag{4.2.27}$$

where,

$$\begin{aligned} A_2^{\mathrm{V}} &= -2 , \qquad A_1^{\mathrm{V}} = -3 , \qquad A_0^{\mathrm{V}} = -8 + \frac{2\pi^2}{3} , \\ C_F &= \frac{4}{3} , \qquad C_\epsilon(\hat{s}) = \left(\frac{4\pi\mu_r^2}{\hat{s}} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} . \end{aligned} \tag{4.2.28}$$

Here, μ_r is the renormalisation scale at which UV divergences are renormalised. Because the incoming fermions are massless, the $\mathcal{O}(\alpha_s)$ virtual corrections to the electroweak vertex factorise [169] such that they are proportional to the leading-order cross section.

Similarly, the soft corrections are,

$$\hat{\sigma}^{\mathrm{S}} = \hat{\sigma}^{\mathrm{LO}} \frac{\alpha_s(\mu_r)}{2\pi} C_F C_\epsilon(\hat{s}) \left(\frac{A_2^{\mathrm{S}}}{\epsilon^2} + \frac{A_1^{\mathrm{S}}}{\epsilon} + A_0^{\mathrm{S}} \right) , \tag{4.2.29}$$

where,

$$A_2^{\mathrm{S}} = 2 , \qquad A_1^{\mathrm{S}} = -4 \ln \delta_{\mathrm{S}} , \qquad A_0^{\mathrm{S}} = 4 \ln^2 \delta_{\mathrm{S}} , \tag{4.2.30}$$

and the soft-collinear corrections can be written as,

$$\hat{\sigma}^{\mathrm{SC}} = \hat{\sigma}^{\mathrm{LO}} \frac{\alpha_s(\mu_r)}{2\pi} C_F C_\epsilon(\hat{s}) \left(\frac{2A_1^{\mathrm{SC}}}{\epsilon} + 2A_0^{\mathrm{SC}} \right) , \tag{4.2.31}$$

where,

$$A_1^{\mathrm{SC}} = 2\ln\delta_{\mathrm{S}} + \frac{3}{2} , \qquad A_0^{\mathrm{SC}} = \left(2\ln\delta_{\mathrm{S}} + \frac{3}{2} \right) \ln\left(\frac{\hat{s}}{\mu_f^2} \right) . \tag{4.2.32}$$

Clearly, the sum of these three contributions is finite since the explicit poles cancel out due to the relations,

$$A_2^{\mathrm{V}} + A_2^{\mathrm{S}} = 0 , \qquad A_1^{\mathrm{V}} + A_1^{\mathrm{S}} + A_1^{\mathrm{SC}} = 0 . \tag{4.2.33}$$

The treatment of the hard-collinear corrections is more subtle. Since the PDFs already contain a resummation of hard-collinear splittings, the $\mathcal{O}(\alpha_s)$ contribution to these should be subtracted. This involves a redefinition of the PDFs, and the explicit expressions for the redefinitions can be found in the appendix of Ref. [168]. The calculation of the hard-collinear hadronic cross section is then simply a convolution of the redefined PDFs with the leading-order partonic cross section.

The complete expression for the NLO hadronic cross section for CC DY heavy neutrino production is therefore,

$$\begin{aligned} \sigma^{\mathrm{NLO}}(pp \to N\ell^\pm + X) &= \sigma^{\mathrm{HC}} + \sigma^{\mathrm{HNC}} \\ &+ \frac{1}{1+\delta_{ij}} \sum_{i,j} \int_{\tau_0}^1 d\xi_1 \int_{\frac{\tau_0}{\xi_1}}^1 d\xi_2 \Big[f_{i/p}(\xi_1,\mu_f) f_{j/p}(\xi_2,\mu_f) \hat{\sigma}^{(2)}(ij \to N\ell^\pm) + (i \leftrightarrow j) \Big] , \end{aligned} \tag{4.2.34}$$

where the finite two-body contributions are,

$$\begin{aligned} \hat{\sigma}^{(2)} &= \hat{\sigma}^{\mathrm{LO}} + \hat{\sigma}^{\mathrm{V}} + \hat{\sigma}^{\mathrm{S}} + \hat{\sigma}^{\mathrm{SC}} \\ &= \hat{\sigma}^{\mathrm{LO}} \left[1 + \frac{\alpha_s(\mu_r)}{2\pi} C_F \left(A_0^{\mathrm{V}} + A_0^{\mathrm{S}} + 2A_0^{\mathrm{SC}} \right) \right] , \end{aligned} \tag{4.2.35}$$

and the limit $\epsilon \to 0$ has been taken, which results in $C_\epsilon(\hat{s}) \to 1$. The calculation for the NC DY process can be performed analogously, where the processes with both ν_ℓ and $\overline{\nu}_\ell$ in the final state are included. The results of these computations will be shown later, when we compare with the rates for gluon fusion. We additionally

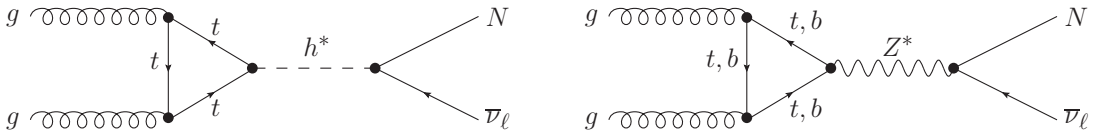


Figure 4.3: Feynman diagrams for the production of a heavy neutrino N from GF via an intermediate (a) h^* and (b) Z^* at leading order.

define the K-factor for the NLO corrections as,

$$K^{\mathrm{NLO}} \equiv \frac{\sigma^{\mathrm{NLO}}}{\sigma^{\mathrm{LO}}} \ . \tag{4.2.36}$$

4.2.3 Gluon Fusion Production

The CC DY, NC DY and VBF processes allow for the production of a heavy neutrino at tree level. However, it is also possible for a heavy neutrino to be produced by the gluon fusion (GF) process involving an intermediate Z or Higgs boson [147, 170, 171]:

$$gg \to Z^*/h^* \to N\overset{(-)}{\nu}_\ell \ . \tag{4.2.37}$$

This process is shown by the Feynman diagrams in Fig. 4.3. It arises at the one-loop level from a quark loop, which couples to the intermediate Z^*/h^* particle. The two contributions add incoherently due to the (anti-)symmetric nature of the $(Z)h$ coupling [148, 170], which means that their partonic cross sections contribute independently, and they are given by [147, 170],

$$\hat{\sigma}^Z(Q) = G_{\mathrm{F}}^2 \frac{\alpha_s^2(\mu_r)|V_{\ell N}|^2}{16(4\pi)^3} m_N^2 (1 - r_N)^2 |F_Z(Q^2)|^2 \ , \tag{4.2.38}$$

$$\hat{\sigma}^h(Q) = G_{\mathrm{F}}^2 \frac{\alpha_s^2(\mu_r)|V_{\ell N}|^2}{16(4\pi)^3} \frac{m_N^2 Q^4 (1 - r_N)^2}{(Q^2 - m_h^2)^2} |F_h(Q^2)|^2 \ , \tag{4.2.39}$$

where Q is the hard process scale—the total centre-of-mass energy of the outgoing final state. At leading order $Q = \hat{s}$, but we will shortly see that this can be changed by soft-gluon corrections. For quarks with third component of weak isospin T_3^q, the

Z/h one-loop form factors are,

$$F_Z(Q^2) = \sum_q 2T_3^q \left[1 - 2r_q f(r_q)\right] \ , \tag{4.2.40}$$

$$F_h(Q^2) = \sum_q 2r_q \left[2 + (1 - 4r_q) f(r_q)\right] \ , \tag{4.2.41}$$

where,

$$f(r) = \begin{cases} 2\left(\sin^{-1}\frac{1}{2\sqrt{r}}\right)^2 , & r > \frac{1}{4} \ , \\ -\frac{1}{2}\left[\ln\left(\frac{1+\sqrt{1-4r}}{1-\sqrt{1-4r}}\right) - i\pi\right]^2 , & r \leq \frac{1}{4} \ , \end{cases} \tag{4.2.42}$$

and we have used the definitions $r_N \equiv m_N^2/Q^2$ and $r_q \equiv m_q^2/Q^2$ in these equations.

When the Z^* form factor is summed over the full constituents of each SU$((2)_L$ quark doublet, such as the top and bottom quarks, the first term cancels and thus the total contribution is proportional to r_q. Specifically, the coupling becomes a pseudoscalar-like coupling proportional to the Yukawa coupling, in accordance with the Goldstone boson equivalence theorem, and in the high-energy (large-Q) limit, the two partonic cross sections become identical. Furthermore, there is no pole in the Z^* process because the Z boson cannot be produced on-shell in the collision between two massless spin-1 vector bosons due to the Landau-Yang theorem [172, 173]. On the other hand, it is possible for a Higgs boson to be produced on-shell—this is exactly the GF production mode which is the dominant contribution to Higgs boson production at the LHC [174].

Indeed, the correspondence between the axial-vector and pseudoscalar couplings is more general. For a massive, colourless vector boson V with four-momentum q and mass m_V produced in the loop process $gg \to V^*$, the contraction between the current and propagator can be written in the most general form as,

$$\Gamma^\mu \Pi_{\mu\nu} \simeq \left(g_V \gamma^\mu + g_A \gamma^\mu \gamma^5\right)\left(g_{\mu\nu} - \frac{q_\mu q_\nu}{m_V^2}\right) \ , \tag{4.2.43}$$

where the propagator has been written in the unitary gauge. The contribution of the vector current $g_V \gamma^\mu$ vanishes due to Furry's theorem [118], which states that matrix elements involving an odd number of vector currents vanishes. Furthermore,

the contribution of the transverse polarisations $g_{\mu\nu}$ vanishes due to the Landau-Yang theorem. Therefore, the result reduces to,

$$\Gamma^{\mu}\Pi_{\mu\nu} \simeq g_A \gamma^{\mu}\gamma^5 \frac{q_\mu q_\nu}{m_V^2} \ . \tag{4.2.44}$$

The quark propagators in the triangle loop can be decomposed by using spinor-completeness relations, and from the Dirac equation we get,

$$\Gamma^{\mu}\Pi_{\mu\nu} \simeq \gamma^5 2m_q \frac{q_\nu}{m_V^2} \ . \tag{4.2.45}$$

Therefore, the contribution reduces to a pseudoscalar coupling proportional to the quark mass m_q.

The hadronic cross sections for the GF process can be computed at leading-order analogously to the Drell-Yan process using Eq. (4.2.19), and the Z^*/h^* contributions should be added together. We will now extend these cross sections to include additional contributions from soft-gluon radiation.

4.3 Soft-gluon Corrections

Computing higher-order corrections to the GF process is difficult since at leading order it is already a one-loop process and involves two powers of the strong-coupling constant α_s. However, it is possible to use the framework of threshold resummation to resum large logarithms which arise when gluon radiation becomes soft. Although not a complete higher-order calculation, the resummation of the threshold logarithms up to next-to-next-to-next-to-leading-logarithmic ($\mathrm{N^3LL}$) accuracy captures the dominant contributions to the inclusive cross section of Higgs boson production via GF up to next-to-next-to-leading-order ($\mathrm{N^2LO}$) accuracy [154]. Therefore, it is reasonable to assume that such a result would still hold for the production of a heavy neutrino, and hence it is still a useful calculation for estimating the full effect of perturbative corrections to the cross section.

4.3.1 Threshold Logarithms

To describe how threshold logarithms arise, we first clarify the notation that we use for the scales involved in the process $pp \to V + X$. The PDFs give the probability densities that the two hard partons carry longitudinal momentum fractions ξ_1 and ξ_2 of their parent proton momenta, respectively. We define the hard scattering scale $Q = \sqrt{p_V^2}$ to be the total centre-of-mass energy of the outgoing state V. The partonic scale is defined to be $\sqrt{\hat{s}} = \sqrt{\xi_1 \xi_2 s}$, where $s = (P_1 + P_2)^2$ is the square of the hadronic centre-of-mass energy. The kinematic threshold for the process to be kinematically allowed is $\tau_0 = Q^2_{\text{min}}/s$. Crucially for the discussion of threshold logarithms is the relation $Q^2 = z\hat{s}$ where $z \leq 1$, and therefore $\tau = Q^2/s = \xi_1 \xi_2 z$ is the hard threshold.

The reason why the partonic scale and hard scale are not equal, which they were when we considered the kinematics of the leading-order Drell-Yan process, is because when the initial state or intermediate partons radiate gluons, these carry a fraction of the momentum away. The fraction of momentum of the partons that the gluons can carry is $(1 - z)$, which leaves a fraction z of the momentum left to go into the hard process. Specifically, this means that the energy of the radiated gluons is $E_g \simeq \sqrt{\hat{s}}(1-z)$, which generates logarithms of the form $\ln(1-z)$ in the perturbative expansion of the cross section. Clearly, in the limit $z \to 1$, which corresponds to the region of phase space where the gluons are forced to be soft, these logarithms can become large which would suggest that that they should spoil the perturbative expansion of the cross section if this region is phenomenologically realised.

One region of the phase space where the gluons are forced to be soft is in the limit $\tau \to 1$, where the hadronic centre-of-mass energy almost entirely goes into the final state V. However, this requires the parton distribution functions to both give momentum fractions close to $\xi_1, \xi_2 \to 1$. In this limit, the PDFs take very small numerical values and this region of phase space has a negligible contribution[3]. On

[3]This is the reason why final states with invariant masses of $\mathcal{O}(10$ TeV) have such negligible cross sections in LHC collisions at 14 TeV.

the other hand, this sharply-falling behaviour of the PDFs dynamically enhances the contribution at the partonic threshold where $Q^2/\hat{s} \to 1$. This corresponds to the soft-gluon limit where $z \to 1$. Therefore, this 'dynamical threshold enhancement' effect [175] can cause the threshold logarithms to contribute significantly. Hence it is necessary to resum the logarithms to prevent the perturbative behaviour of the cross section from being destroyed, and this is possible since the threshold logarithms factorise and thus can be resummed to all orders in $\alpha_s \ln(1-z)$ via exponentiation [176–178].

4.3.2 Threshold Resummation Formalism

We will now describe the formalism that we employ to resum threshold logarithms for the GF production of heavy neutrinos. An important result that we make use of is that the emissions of soft gluons off the fermions in the quark loop leave the structure of the loop unchanged due to soft factorisation. Hence the soft-gluon corrections to $gg \to Z^*$, where the Z^* coupling reduces to that of a pseudoscalar coupling via Eq. (4.2.45), can be approximated by the soft-gluon corrections to the production of a pseudoscalar. This allows us to perform the threshold resummation for the $gg \to Z^*$ subprocess in an approach that is consistent with the $gg \to h^*$ subprocess, and was not observed in previous approaches to the resummation of $gg \to Z^*$.

To account for the soft-gluon radiation, we modify Eq. (4.2.21) to give the fixed-order (FO) hadronic cross section,

$$\begin{aligned}\sigma^{\mathrm{FO}}(pp \to V + X) = \\ \frac{1}{1+\delta_{ij}} \sum_{i,j,\beta} \int_{\tau_0}^1 d\tau \int_\tau^1 \frac{d\xi}{\xi} \int_{\tau/\xi}^1 \frac{dz}{z} \Big[f_{i/p}(\xi,\mu_f) f_{j/p}(\tau/(z\xi),\mu_f) + (i \leftrightarrow j) \Big] \\ \times \Delta_{ij}^{\beta\,\mathrm{FO}}(z)\ \hat{\sigma}^\beta(ij \to V)\ ,\end{aligned} \tag{4.3.1}$$

where $\beta \in \{Z, h\}$. The soft-coefficient function Δ accounts for the soft-gluon radiation, and due to soft factorisation, the entire process involves a convolution with

the leading-order partonic cross section $\hat{\sigma}$. The partonic cross sections are given in Eqs. (4.2.39) and (4.2.38). The soft-coefficient function explicitly has the form,

$$\Delta_{ij}^{\beta\,\mathrm{FO}}(z) = \delta(1-z) + \mathcal{O}(\alpha_s) \ . \tag{4.3.2}$$

At leading order—where there is no soft-gluon radiation—the delta function enforces $z = 1$ and the result reduces to the leading-order hadronic cross section from Eq. (4.2.21). The higher-order terms in the soft-coefficient function contain threshold logarithms, $\ln(1-z)$, and should be resummed to all orders to maintain perturbative convergence. The most divergent logarithmic terms at each order of the perturbative expansion are denoted as the leading-logarithmic (LL) terms, with the second-most divergent referred to as the next-to-leading-logarithmic terms (NLL). An all-orders resummation at $\mathrm{N}^k\mathrm{LL}$ accuracy resums the $k+1$ most divergent logarithmic terms at each order of the perturbative expansion.

We perform this resummation by using the Soft-Collinear Effective Theory (SCET) framework [179–181]. This allows Eq. (4.3.1) to be factorised directly in momentum space [182, 183] by segmenting and regularising divergent regions of phase space with hard and soft scales, μ_h and μ_s. The requirement for physical observables to be independent of the choice of scales implies that factorised components can be independently RG evolved and matched, using exponentiation [176, 177]. Thus, numerically large quantities are replaced with perturbative ones regulated by μ_h and μ_s and with RG-evolution coefficients that run μ_h and μ_s to μ_f and Q.

The resummation procedure that we employ simply reduces to replacing the fixed-order soft-coefficient function in Eq. (4.3.1) with a resummed version,

$$\sigma^{\mathrm{FO}} \to \sigma^{\mathrm{res}} \ : \quad \Delta_{ij}^{\beta\,\mathrm{FO}}(z) \to \Delta_{ij}^{\beta\,\mathrm{res}}(z) \ , \tag{4.3.3}$$

which can then be calculated using the SCET framework. In the case of GF production of a heavy neutrino, the resummed version of the soft-coefficient function takes

the form [175, 184],

$$\Delta_{gg}^{\beta\,\mathrm{res}}(z) = |C_\beta(Q^2,\mu_h^2)|^2\; U(Q^2,\mu_\alpha^2,\mu_h^2,\mu_s^2,\mu_f^2) \\ \times \frac{\sqrt{z}\; z^{-\eta}}{(1-z)^{1-2\eta}}\; \tilde{s}_{\mathrm{Higgs}}\left(\ln\frac{Q^2(1-z)^2}{\mu_s^2 z} + \partial_\eta, \mu_s\right) \frac{e^{-2\gamma_{\mathrm{E}}\eta}}{\Gamma(2\eta)}\;. \tag{4.3.4}$$

This result contains several distinct pieces, which we shall now discuss.

The hard function C_β is process dependent and accounts for virtual corrections to the hard process. This is the main part of the soft-coefficient function which depends on the process, and thus the main differences between the scalar ($\beta = h$) and pseudoscalar ($\beta = Z$) mediators is accounted for by it. For $\beta = h$, the function is given by the two-step SCET matching coefficients C_t and C_S of Ref. [184], with

$$C_h(Q^2,\mu_h^2) \equiv C_t(m_t^2,\mu_t^2) C_S(-Q^2,\mu_h^2)\;, \tag{4.3.5}$$

where,

$$C_X(Q^2,\mu^2) = \sum_{n=0}^{\infty} C_X^{(n)}(Q^2,\mu^2)\;\left(\frac{\alpha_s(\mu)}{4\pi}\right)^n\;, \tag{4.3.6}$$

and $X \in \{t, S\}$. The product of C_t and C_S—which can be expanded individually as power series in $(\alpha_s/4\pi)$—is equivalent to a one-step SCET matching procedure when setting $\mu_t = \mu_h$ [185]. For $\beta = Z$, the one-step matching hard function can also be expanded as a power series. In the notation of Refs. [186] and [187], this is,

$$C_Z(Q^2,\mu_h^2) \equiv C_g^{A\,\mathrm{eff}}(Q^2,\mu_h^2) \\ = \sum_{n=0}^{\infty} C_{g,n}^{A\,\mathrm{eff}}(Q^2,\mu_h^2)\;\left(\frac{\alpha_s(\mu_h)}{4\pi}\right)^n\;. \tag{4.3.7}$$

It is important to note that the $\ln(\mu_h^2/m_t^2)$ term that appears in $C_{g,2}^{A\,\mathrm{eff}}$ of Ref. [186] should be replaced with $\ln(Q^2/m_t^2)$ in order to preserve the scale independence of the total cross section, which is a physical observable [187]. With this modification, both C_h and C_Z satisfy the evolution equation,

$$\frac{d}{d\ln\mu} C_\beta(Q^2,\mu^2) = \left[\Gamma_{\mathrm{Cusp}}^A \ln\left(\frac{Q^2}{\mu^2}\right) + \gamma^S + \gamma^t\right] C_\beta(Q^2,\mu^2)\;, \tag{4.3.8}$$

for anomalous dimensions Γ_{Cusp}^A, γ^S and γ^t as given in Refs. [175] and [184].

The evolution function U accounts for the RG running between the various scales, and is given by,

$$U(Q^2,\mu_\alpha^2,\mu_h^2,\mu_s^2,\mu_f^2) = \frac{\alpha_s^2(\mu_s)}{\alpha_s^2(\mu_f)}\left[\frac{\beta(\alpha_s(\mu_s))/\alpha_s^2(\mu_s)}{\beta(\alpha_s(\mu_\alpha))/\alpha_s^2(\mu_\alpha)}\right]^2\left(\frac{Q^2}{\mu_h^2}\right)^{-2a_\Gamma(\mu_h^2,\mu_s^2)} \times\left|e^{4\mathcal{S}(\mu_h^2,\mu_s^2)-2a_{\gamma^S}(\mu_h^2,\mu_s^2)+4a_{\gamma^B}(\mu_s^2,\mu_f^2)}\right| , \tag{4.3.9}$$

where $\mu_\alpha = \mu_t$ for h^* mediators, and $\mu_\alpha = \mu_h$ for Z^* mediators. This differing scale choice is the only other difference between the soft-coefficient function for scalar and pseudoscalar mediators aside from the hard function. We have also used the definitions,

$$\mathcal{S}(\mu_1^2,\mu_2^2) = -\int_{\alpha_s(\mu_1)}^{\alpha_s(\mu_2)} d\alpha\frac{\Gamma^A_{\rm Cusp}(\alpha)}{\beta(\alpha)}\int_{\alpha_s(\mu_1)}^{\alpha}\frac{d\alpha'}{\beta(\alpha')} ,$$

$$a_\Gamma(\mu_1^2,\mu_2^2) = -\int_{\alpha_s(\mu_1)}^{\alpha_s(\mu_2)} d\alpha\frac{\Gamma^A_{\rm Cusp}(\alpha)}{\beta(\alpha)} , \tag{4.3.10}$$

with similar definitions for a_{γ^S} and a_{γ^B} .

The soft-scalar function $\tilde{s}_{\rm Higgs}$ describes soft radiation off incoming gluons and hence is universal for both scalars and pseudoscalars. The derivatives in $\tilde{s}_{\rm Higgs}$ are regular partial derivatives that act to the right, before $\eta \equiv 2a_\Gamma(\mu_s^2,\mu_f^2)$ is evaluated numerically. This is valid for the region $\eta > 0$, but can be extended to negative values using analytic continuation, which involves a series of subtraction terms, and is described in Appendix C. Furthermore, we include an additional factor of $\sqrt{z}$ in Eq. (4.3.4) with respect to Refs. [175] and [184], which accounts for power corrections that are manifest in the alternative Mellin-space resummation formalisation [175, 184].

Definitions and explicit expressions of the quantities in Eqs. (4.3.4)–(4.3.9) up to $\mathcal{O}(\alpha_s^2)$ are presented in Refs. [175] and [184]. The mappings between N^{k}LL accuracy and the required ingredients can be found in Refs. [175] and [188], where the power counting is such that at N^{k}LL accuracy terms up to $\mathcal{O}(\alpha_s^{k-1})$ accuracy are included, with the large logarithm $\ln(1-z)$ counted as $\mathcal{O}(1/\alpha_s)$. At N^{3}LL, one needs C_β at two loops for both pseudoscalars [186] and scalars [184, 189], as well as $\tilde{s}_{\rm Higgs}$ at two

loops [175, 184, 190]. Furthermore, the anomalous dimension Γ^A_{Cusp} and β-function coefficients are required at four loops [175], while the anomalous dimensions γ^S and γ^B are needed at three loops [184]. Note that while the results of Refs. [184, 186, 189] are derived in the heavy top limit, $\mathcal{O}(Q^2/m_t^2)$ corrections to inclusive scalar and pseudoscalar cross sections are known to be $\mathcal{O}(1-10\%)$ [151, 191], even for $Q^2 \gg m_t^2$, justifying their use in our calculation.

4.3.3 Computational Setup

Using the methodology presented in the previous sections, we can now compute the NLO hadronic cross sections for the CC DY and NC DY processes, as well as the threshold-resummed hadronic cross section up to N^{3}LL for the GF process. We implement these using a custom code written in the Fortran [192] programming language, and perform Monte Carlo integration using the CUBA [166] libraries. The PDF sets and extraction of $\alpha_s(\mu)$ are evaluated using the LHAPDF6 libraries [193]. We use the following Standard Model inputs [15]:

$$m_Z = 91.1876 \text{ GeV} , \quad \alpha_e(m_Z)^{-1} = 127.940 , \quad G_{\text{F}} = (1.17456 \times 10^{-5}) \text{ GeV}^{-2} ,$$
$$m_b = 0 \text{ GeV} , \quad m_t = 173.2 \text{ GeV} , \quad m_h = 125.7 \text{ GeV} . \tag{4.3.11}$$

For the Drell-Yan processes, we use the NNPDF 3.0 QED NLO PDF set [194, 195] with no phase-space cuts applied. We also cross-check these calculations against an implementation in MadGraph5_aMC@NLO [85] with a model file generated by FeynRules [86], following a similar methodology to Ref. [148]. Furthermore, we also use MadGraph5_aMC@NLO to calculate the NLO hadronic cross section for the VBF process, using the same PDF set, and with the scale choices and regulating VBF phase-space cuts identical to those in Ref. [148].

For the GF process, we use the NNPDF 3.0 NNLO+NNLL PDF set [196] to best match the logarithmic accuracy. Although the uncertainties in the PDF set are

sizeable, the use of a fixed-order PDF set would double-count the initial-state gluons. The calculation of the leading-order GF cross section is checked against Refs. [147] and [148], while we check the soft-coefficient function against Refs. [187, 188, 197].

To minimise the numerical impact of missing perturbative corrections, we follow Refs. [175] and [184] and choose the scale scheme,

$$\mu_r, \mu_f, \mu_t, \mu_h = Q \ , \qquad \mu_s = \frac{Q(1-\tau)}{1+7\tau} \ , \tag{4.3.12}$$

for both the leading-order and resummed GF calculations. For GF, we report the scale dependence arising from simultaneously varying μ_f, μ_r, and μ_s up and down by factors of two. For the CC DY, NC DY and VBF processes, we quantify scale dependence by simultaneously varying μ_f and μ_r up and down by factors of two. While the μ_s dependence itself is numerically small, we vary it jointly with μ_f for the GF process to ensure that the subtraction terms required for numerical evaluation lead sufficiently to numerical convergence [175, 184]. Missing fixed-order terms that would otherwise stabilise μ_f represents the largest source of uncertainty. Indeed, we find other scale uncertainties to be relatively small due to our high logarithmic accuracy.

4.4 Results

We present the cross sections for heavy neutrino production by factoring out the active-heavy mixing matrix $V_{\ell N}$, which always appears in cross sections as $|V_{\ell N}|^2$. This result holds to all orders in α_s [148,168] since the higher-order quantum chromodynamics corrections do not modify the vertex where the heavy neutrino is produced. This means that the cross sections we present are model independent across all low-scale seesaw models that result in the production of a heavy neutrino, with all model-dependent information encoded in $V_{\ell N}$.

In Fig. 4.4 we show the cross sections for the production of a heavy neutrino as a function of the heavy neutrino mass m_N at a collider centre-of-mass energy of

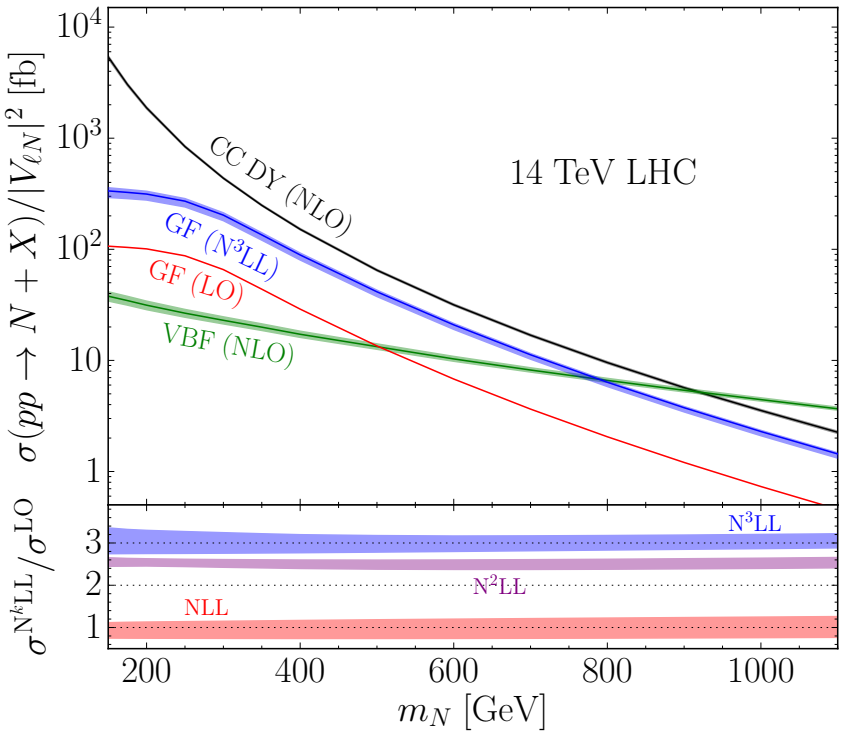


Figure 4.4: Cross sections for the production of a heavy neutrino N as a function of the heavy neutrino mass m_N, divided by the active-heavy mixing $|V_{\ell N}|^2$. The CC DY at NLO, VBF at NLO, and GF at LO and N^{3}LL production modes are shown. The lower panel shows the K-factors for the resummed GF production modes. The thickness of each curve shows the uncertainty from scale dependence.

$\sqrt{s} = 14$ TeV, divided by the active-heavy mixing $|V_{\ell N}|^2$. The processes shown are CC DY at NLO and VBF at NLO, and GF at LO and N^{3}LL. Each curve has a thickness which represents the uncertainty due to the scale dependence, except for GF at LO where we omit the scale uncertainty. The lower panel shows the K-factors for the resummed GF cross section, which we define as,

$$K^{\mathrm{N}^k\mathrm{LL}} \equiv \frac{\sigma^{\mathrm{N}^k\mathrm{LL}}}{\sigma^{\mathrm{LO}}} \ , \tag{4.4.1}$$

where k represents the order of the logarithmic accuracy. For $m_N = 150 - 1000$ GeV, the cross sections cover the ranges:

$$\begin{aligned} \text{CC DY (NLO)} \ : &\quad 3.5 - 5400 \text{ fb} \ , \\ \text{GF (LO)} \ : &\quad 0.73 - 110 \text{ fb} \ , \\ \text{GF (N}^3\text{LL)} \ : &\quad 1.9 - 280 \text{ fb} \ , \end{aligned}$$

$$\text{VBF (NLO)} : \quad 4.4 - 37 \text{ fb} . \tag{4.4.2}$$

For the resummed GF production mode, the K-factors and associated scale uncertainties cover the ranges:

$$\begin{aligned} \text{GF (NLL)} &: \quad K = 1.00 - 1.06 , \quad \delta\sigma/\sigma = \pm 25 - 27\% , \\ \text{GF (N}^2\text{LL)} &: \quad K = 2.59 - 2.66 , \quad \delta\sigma/\sigma = \pm 6 - 9\% , \\ \text{GF (N}^3\text{LL)} &: \quad K = 3.07 - 3.14 , \quad \delta\sigma/\sigma = \pm 8 - 13\% . \end{aligned} \tag{4.4.3}$$

These K-factors can be compared with the DY (VBF) K-factors of $K^{\text{NLO}}_{\text{DY (VBF)}} = 1.15-1.25\ (0.98-1.06)$ and uncertainties of $(\delta\sigma/\sigma)^{\text{DY (VBF)}} = \pm 1-5\ (5-11)\%$ [148].

It is clear that over the entire range of heavy neutrino masses, the GF N^{2}LL and N^{3}LL cross sections are several factors larger than the corresponding LO rate. Furthermore, they are approximately flat across the entire range of m_N. At NLL, we find that $C_\beta,\ \tilde{s}_{\text{Higgs}} \sim 1$ with the main contribution arising from the evolution of $\alpha_s(\mu)$ and its associated scale uncertainty, which results in the K-factor being close to 1. We further find that $\sigma^{\text{N}^3\text{LL}}/\sigma^{\text{N}^2\text{LL}} \sim 1.1 - 1.2$, which indicates the convergence of the perturbative series. It is important to note that choosing a different set of scales in Eq. (4.3.12) could reduce the size of the K-factors, but also increase them as well. However, our choice of scales was driven by a desire to minimise missing FO corrections.

The large residual scale uncertainty at N^{3}LL arises due to missing FO contributions which are not contained in the resummation of threshold logarithms. This includes hard initial-state radiation which is not accounted for by the DGLAP-evolution of the PDFs, and is likely to be positive definite [154] and thus increase the overall rate. Including a matching to FO contributions would stabilise the dependence on μ_f and μ_r. However, the sizes of our corrections are consistent with those for the production of heavy scalars and pseudoscalars [149–153, 187, 198, 199].

We find for $m_N \gtrsim 300$ GeV that the heavy neutrino production GF rate is now comparable to the CC DY and NC DY (not shown for clarity) rates. When ba-

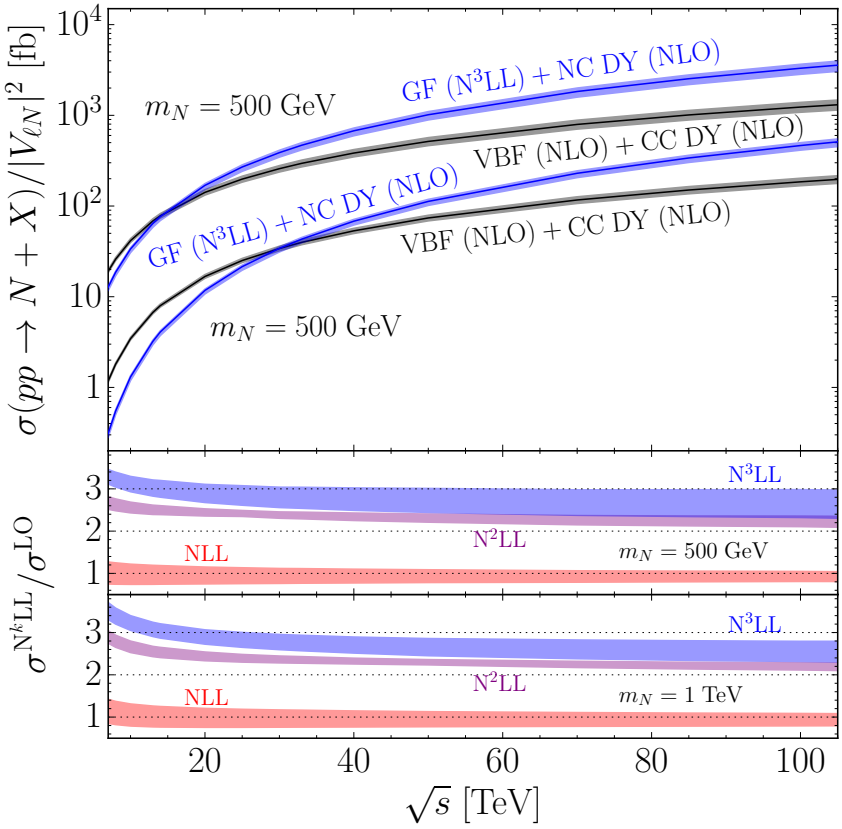


Figure 4.5: Cross sections for the production of a heavy neutrino N as a function of the collider centre-of-mass energy $\sqrt{s}$ for heavy neutrino masses $m_N = 500$ GeV and $m_N = 1$ TeV, divided by the active-heavy mixing $|V_{\ell N}|^2$. The combined CC DY at NLO and VBF at NLO, and NC DY at NLO and GF at N^{3}LL production modes are shown. The lower panels show the K-factors for the resummed GF production modes for both heavy neutrino masses. The thickness of each curve shows the uncertainty from scale dependence.

sic fiducial cuts are applied on the charged lepton in the CC processes the combined GF + NC DY rate is slightly larger than the combined VBF + CC DY channel. These combinations make sense because GF is a non-interfering $\mathcal{O}(\alpha_s^2)$ correction to the NC DY process, while VBF is a non-interfering $\mathcal{O}(\alpha_e)$ correction to the CC DY process. For $m_N \lesssim 600$ GeV, which correspond to masses that are most relevant for LHC phenomenology due to mixing suppression [155, 159], the GF channel is several factors larger than the VBF mechanism, indicating its potential importance at the LHC and its upgrades and successors.

The applicability to the upgrades and successors of the LHC is shown by Fig. 4.5

$\sqrt{s}$	14 TeV				100 TeV			
m_N	300 GeV		500 GeV		300 GeV		1 TeV	
$\sigma/\lvert V_{\ell N}\rvert^2$ [fb]	σ [fb]	K	σ [fb]	K	σ [fb]	K	σ [fb]	K
GF LO	65.4	–	13.5	–	2.84×10^3	–	154	–
GF NLL	$65.9\ ^{+14\%}_{-26\%}$	1.01	$13.7\ ^{+17\%}_{-27\%}$	1.01	$2.83\ ^{+2\%}_{-18\%}\times10^3$	1.00	$154\ ^{+8\%}_{-21\%}$	1.00
GF N^{2}LL	$170\ ^{<1\%}_{-7\%}$	2.61	$34.9\ ^{<1\%}_{-8\%}$	2.59	$6.83\ ^{+2\%}_{-13\%}\times10^3$	2.40	$351\ ^{<1\%}_{-8\%}$	2.28
GF N^{3}LL	$202\ ^{+5\%}_{-11\%}$	3.09	$41.3\ ^{+3\%}_{-9\%}$	3.06	$7.88\ ^{+13\%}_{-16\%}\times10^3$	2.77	$401\ ^{+8\%}_{-11\%}$	2.60

Table 4.1: Cross sections for the production of a heavy neutrino N via GF for representative values of the heavy neutrino mass m_N and collider energy $\sqrt{s}$, divided by the active-heavy mixing $|V_{\ell N}|^2$. Also shown are the scale dependence (%) and K-factors.

where we plot the combined CC DY at NLO and VBF at NLO, and NC DY at NLO and GF at N^{3}LL production modes as a function of the centre-of-mass energy $\sqrt{s}$ of the collider, for heavy neutrino masses $m_N = 500$ GeV and $m_N = 1$ TeV. The relative scale uncertainties for the combined processes are added in quadrature. For $\sqrt{s} = 7 - 100$ TeV, the resummed K-factors cover the ranges:

$$\begin{aligned}
\text{GF (NLL)} &: \quad K = 1.0 - 1.2\ , \quad \delta\sigma/\sigma = \pm 19 - 29\%\ , \\
\text{GF (N}^2\text{LL)} &: \quad K = 2.3 - 3.0\ , \quad \delta\sigma/\sigma = \pm 6 - 11\%\ , \\
\text{GF (N}^3\text{LL)} &: \quad K = 1.0 - 1.2\ , \quad \delta\sigma/\sigma = \pm 8 - 14\%\ .
\end{aligned} \tag{4.4.4}$$

The drop in the K-factors with increasing $\sqrt{s}$ is likely due to an increasing importance of hard initial-state radiation as the collider energy becomes larger. We observe that for $m_N = 500 - 1000$ GeV, the total inclusive production rate of $N\overset{(-)}{\nu}_\ell$ overtakes the inclusive $N\ell^\pm$ production rate at $\sqrt{s} \gtrsim 15 - 30$ TeV. For $\sqrt{s} = 33$ (100) TeV, this difference is a factor of $1 - 1.6$ $(2.5 - 2.7)$ and is driven by the GF rate, for which the luminosity grows much faster than the $q\overline{q'}$ (DY) and qq (VBF) luminosities with increasing collider energies. While not shown, we find for $m_N = 500 - 1000$ GeV that the GF rate individually exceeds the CC DY rate for $\sqrt{s} \gtrsim 20 - 25$ TeV. Our results for the resummed and LO GF cross sections are summarised numerically in Table 4.1 for representative values of the heavy neutrino mass and collider energy.

4.5 Conclusions

In this chapter, we calculated the inclusive cross section for the threshold-resummed GF production of a heavy neutrino up to N^{3}LL accuracy. By treating the $gg \to Z^*$ subprocess as a pseudoscalar, we have computed the soft-gluon corrections to the GF production mode for the first time. We found that the K-factors can be large, with enhancements to the GF production cross section of approximately a factor of three, and that GF dominates over the CC DY and VBF production modes at collider energies of $\sqrt{s} \gtrsim 20 - 25$ TeV. Therefore, at a future collider the GF process would be the most important channel. Including additional FO corrections would likely enhance the rate even further. Although the actual cross section is heavily dependent on the value of the active-heavy mixing $V_{\ell N}$, this affects all channels equally so the conclusions about the relative importance of the GF channel remain valid.

These results are important because the discovery of heavy neutrinos would solve one of the unexplained phenomena of the Standard Model—specifically, the mechanism that allows the neutrinos of the Standard Model to have small non-zero masses. Many low-scale seesaw models predict the existence of a heavy neutrino beyond the electroweak scale as a solution to this problem, and through the mixing between their gauge and mass eigenstates, these heavy neutrinos couple directly to the electroweak bosons such that they could be produced at a collider. Therefore, studying the various production mechanisms and making predictions for their cross sections provides useful information for the ongoing and future searches for heavy neutrinos at such colliders.

Chapter 5

Adversarial Autoencoders for Robust New Physics Searches

We now consider an application of machine learning to the search for physics beyond the Standard Model. Autoencoders are a type of neural network that can be used for model-independent searches for new physics because after training them only on background events, they can detect new signal events by flagging them up as anomalies. Therefore they could be trained directly on experimental data, which would avoid sensitivity to theoretical uncertainties. This would then leave the experimental uncertainties as the final source of uncertainties affecting the classification. We propose a method of combining the autoencoder with an adversarial neural network to remove its sensitivity to the smearing of final-state objects which are a major source of experimental uncertainty. Although we apply it to $t\bar{t}$ events generated with a Monte Carlo event generator as a proof-of-concept, the same method could be used directly on data.

5.1 Motivation

The use of machine-learning techniques to aid with the search for new physics is becoming increasingly popular [49, 200–202]. These techniques and models are able

to analyse large quantities of data to disentangle the slight differences between signal and background events which can be hidden amongst the high dimensionality of the data. Machine-learning algorithms are able to analyse multiple observables simultaneously to find a region in this multi-dimensional parameter space that shows a relative enhancement of signal over background.

Using a supervised-learning approach, it is necessary to use training data where each event has been appropriately labelled as either signal or background. This requires the generation of pseudo-data for both the signal and background using Monte Carlo event generators [203–205], where it is already known which physics model was used to generate any particular event. However, these samples are plagued by theoretical uncertainties such as the dependence on the factorisation and renormalisation scales, and so the classification algorithm will be affected by these same uncertainties. Recently, it has been proposed that adversarial networks can be used to reduced the sensitivity of classification methods to some of these theoretical uncertainties [206]. In any case, there is no guarantee that features that the algorithm learns would be replicated in an actual experiment, and it may simply learn artefacts of the particular choice of event generator.

One approach to overcome the deficiencies arising from training on pseudo-data is to train directly on experimental data. While this is not affected by theoretical uncertainties, it is necessary to ensure that the signal and background samples are sufficiently pure in order to use a supervised-learning technique. However, obtaining such training samples is impossible because if one was to already know whether a particular event came from the signal or background, then there would be no need for the machine-learning algorithm in the first place. The way to overcome this bottleneck is not to train on signal events at all, but to identify the kinematic features of background samples and design a method that flags up events that do not possess the same features—these would then be classified as signal events.

Autoencoders are a type of neural network that can be used for this purpose [55, 56, 207, 208]. These are fully-connected neural networks that contain a hidden layer with

a fewer number of nodes than there are inputs to the network. The network maps the set of inputs to this latent-compressed layer, and then decodes this compressed representation back into a set of outputs of the same size as the input. The loss function measures the squared difference between the inputs and outputs, and thus the optimisation objective of the network is to reconstruct the inputs as closely as possible. In order to do so, the autoencoder must learn an efficient mapping of the input to the latent-compressed layer which performs well across all training samples.

After training an autoencoder on background samples, it is expected that applying the autoencoder to signal samples will result in a higher value of the loss function. This is because the efficient mapping that the autoencoder learns has only been optimised to work well for background events, and it is likely that signal events will not be reconstructed as effectively since their kinematic features will be different. Therefore, by applying a threshold on the loss function, it is possible to use the autoencoder as an unsupervised classifier, with events that have a value of the loss above the chosen threshold predicted to be signal events.

The autoencoder can then be used to train directly on experimental data, acting as an anomaly—or novelty—detector. Since this avoids all theoretical uncertainties, the final remaining source of uncertainties affecting the classifier are then experimental uncertainties from the reconstruction of the final-state objects. Analogously to how an adversarial neural network can reduce the sensitivity of supervised classification methods to the theoretical uncertainties, we propose the combination of such a network with an unsupervised autoencoder to reduce its sensitivity to experimental uncertainties. As a proof-of-concept, we generate the data using a Monte Carlo event generator, and simulate experimental uncertainties by smearing the jets and missing energy in the events.

The rest of this chapter is structured as follows: in Sec. 5.2 we will describe how the events are generated and analysed, and the smearing procedure that we apply to simulate the experimental uncertainties. In Sec. 5.3, we will first use a supervised classifier and combine this with an adversarial neural network to remove the

sensitivity to the smearing. We will then extend this approach to an unsupervised autoencoder in Sec. 5.4, and consider its application to other new physics models and the effects of an impure training sample. Finally, we conclude the chapter in Sec. 5.5.

5.2 Analysis Setup and Smearing Procedure

We use MadGraph5_aMC@NLO [85] to generate the events for the study, followed by Pythia 8.2 [205] for the parton showering and hadronisation of the events. The background events consist of $pp \to t\bar{t}$ at a centre-of-mass energy of 14 TeV, with one top quark forced to decay leptonically—the other is forced to decay hadronically. The signal events are generated from a heavy Z' boson [209] via $pp \to Z' \to t\bar{t}$, also with semileptonic decays of the top quarks. As a benchmark for this study, we select the Z' mass to be 2 TeV with a width of 89.6 GeV. A transverse momentum cut of $p_T > 500$ GeV is applied directly to the top quarks at the generator level, for both signal and background events.

Following the concept of reconstructing highly boosted top quarks with fat jets [210, 211], the hadrons and non-isolated leptons from the event are initially clustered into jets using the Cambridge-Aachen algorithm [212] with a radius of $R = 1$. The constituents of the two hardest fat jets[1] are then reclustered into jets using the k_T algorithm with $R = 0.2$, implemented in FastJet [213]. Jets are required to have $p_T > 30$ GeV and are b-tagged through their association to a final-state B meson. Isolated leptons are required to have $p_T > 10$ GeV. Events are selected which have a scalar-summed visible (everything except neutrinos) transverse momentum of $H_T > 1$ TeV, and which have at least one b jet inside one fat jet, and at least one b jet and two light jets inside the other fat jet. Furthermore, we require at least one isolated lepton in the events.

[1] These are the fat jets with the highest p_T, and are chosen to capture the two top quarks from the hard process.

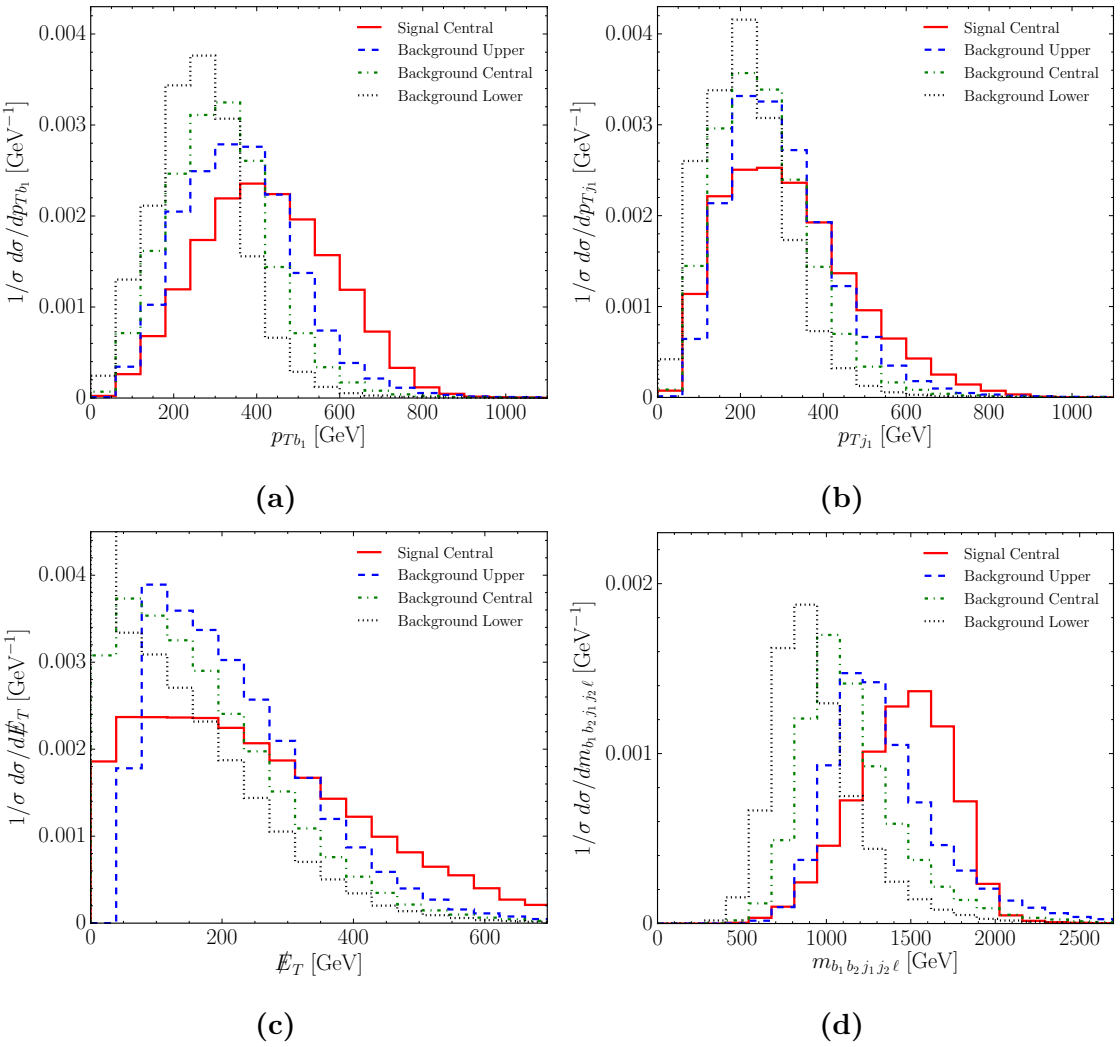


Figure 5.1: Effect of smearing on the distributions of (a) the p_T of the hardest b jet, (b) the p_T of the hardest light jet, (c) the missing energy and (d) the invariant mass of the jets and lepton, compared with the unsmeared background and the signal samples.

The observables that we consider for the analysis are the four-momenta of each of the two b jets, two light jets and isolated lepton, as well as the total missing energy $\not{E}_T$ in the event (21 observables in total). To represent possible systematic uncertainties that can arise in detectors from jet energy scales, we apply a smearing procedure to both the jets and the missing energy in the events. For the jets and leptons, we use a smearing based on Refs. [214] and [215] where the three-momenta of each object is smeared with a Gaussian, and the mass held fixed. In our case, we take the extremities of this Gaussian so that the smearing is either applied upwards or

downwards for all objects, with the relative width of the smearing envelope being larger for smaller p_T values. Similarly, we apply a shift to the missing energy based on Ref. [216], where the width of the shift is proportional to $\sqrt{H_T}$, and we also use the two extremities of the envelope. We fix the direction of the missing-energy smearing to always be the same as that of the jets and leptons—if they are smeared upwards then we also smear the missing energy upwards, and vice versa. For the purposes of this study, we increase the size of the Gaussian smearing envelope by a further factor of three, to be conservative on the systematic uncertainties and highlight the ability of our setup to correct for it.

We apply the smearing to the background sample such that two extra datasets are created for smearing in the upwards and downwards directions, in addition to the unsmeared central background sample. No smearing is applied to the signal sample. The three background samples are each generated from statistically independent generator samples, and after all cuts we select 100,000 events from each of the four samples, with 20% of these retained for testing. This gives a total of 320,000 training events, and 80,000 events for testing.

In Fig. 5.1 we show the effect of smearing on the p_T of the hardest b jet and light jet, the missing energy in the event, and the invariant mass of the jets and lepton, compared with the equivalent distributions for the signal events. Clearly the smearing of the background has the potential to make it either easier or harder for a classifier to discriminate between signal and background, depending on which direction the smearing shifts the background distribution. This can be seen from the amount of overlap that exists between the distributions of the background samples and the signal sample—if there is a larger overlap in their distributions, then a classifier would be more likely to recognise a signal event as having resulted from the background than it would if there was a smaller overlap. In the later sections, we will attempt to reduce the sensitivity of both supervised and unsupervised classifiers to this smearing.

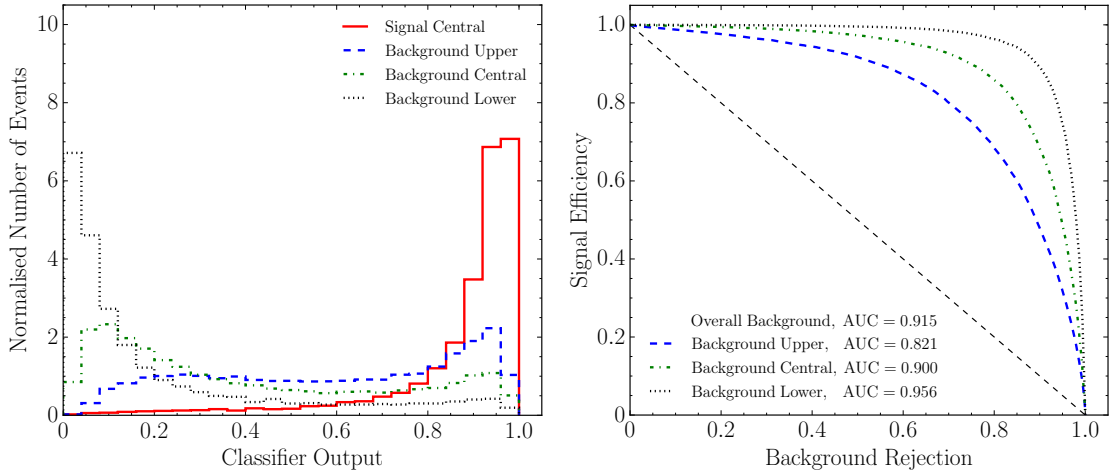
5.3 Decorrelated Smearing with Supervised Adversarial Classifier

To set a benchmark for the signal-to-background separation, we first train a simple neural network as a classifier to discriminate signal events from background events. We expect this supervised-learning approach to perform better than the unsupervised approach which we perform later.

The network consists of two hidden layers each with 20 nodes, with ReLu activations[2], and a final layer with a single sigmoid output. We use a binary cross-entropy loss function since there are two possible classes—signal or background. A class weighting in the loss function is used to account for the higher frequency of background events in the training data—the contribution to the loss of the signal events is therefore weighted higher. The network is trained using the Adam optimiser [217] for 500 epochs with a learning rate of 0.01 and a batch size of 500. The network is implemented in Keras [218] with a TensorFlow [219] backend, and we use these throughout the rest of this study.

The results are shown by the distributions of the classifier outputs and the corresponding receiver-operating-characteristic (ROC) curves in Fig. 5.2. These are obtained by testing the network on each of the three background test sets separately, and performing a classification against the central signal sample for each one. Also shown are the corresponding area-under-curve (AUC) scores for each ROC curve as well as the AUC score for all the background test samples combined. The network performance is strongly dependent on the direction that the sample has been smeared in. This can be understood from the distributions of the observables in Fig. 5.1, where there is a larger overlap between the background which has been smeared upwards and the signal distribution—conversely, there is a smaller overlap between the downwards-smeared background and the signal distribution.

[2]The ReLu, or rectified linear unit, is defined by $g(x) = \max(0, x)$.



(a) **(b)**

Figure 5.2: (a) Supervised neural network classifier output and (b) ROC curves for a classifier trained to classify signal and background events. The three background distributions result from the three different directions of smearing.

We now extend this classifier with an adversarial neural network which is designed to discriminate the smearing class that the background sample came from, based upon the output of the classifier. The aim for such an extension to the classifier is to attempt to remove such a large dependence of its performance on the smearing of the background [206, 220]. The adversary and classifier are forced to take part in a zero-sum game—the classifier must learn to make its prediction without using any information derived from the smearing, in order to make it as hard as possible for the adversary to be able to discriminate the background samples. This is achieved by the two networks having opposite optimisation objectives, so that the classifier is penalised when the adversary performs better.

The adversarial neural network consists of two hidden layers, each with 20 nodes and ReLu activation functions, and takes the output of the classifier as an input. The output of the adversary has three nodes (one for each smearing class) with a softmax activation function and a categorical cross-entropy loss. The network is then trained as follows:

1. The classifier is trained for three epochs using the Adam optimiser with a

learning rate of 0.01 and a batch size of 500. A class weighting is applied to account for the higher frequency of background events in the training data.

2. The adversary is trained on background events for three epochs using mini-batch gradient descent with a learning rate of 0.01 and a batch size of 500.

3. The classifier is trained for one epoch with mini-batch gradient descent with a batch size of 500 and with a total loss function,

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_{\text{Classifier}} - \alpha \mathcal{L}_{\text{Adversary}} \ . \tag{5.3.1}$$

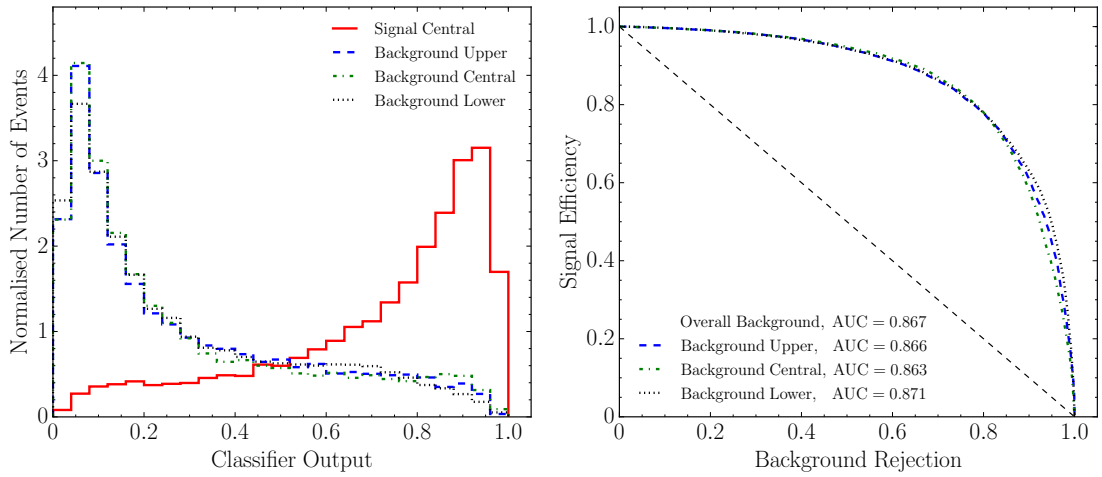
   Furthermore, two class weightings are applied: one to account for the higher frequency of background events that the classifier is trained on, and one to account for the fact that the signal events are unsmeared, resulting in a higher frequency of unsmeared events that the adversary is trained on.

4. The adversary is trained on background events for one epoch using mini-batch gradient descent with a batch size of 500.

5. Steps 3 and 4 are repeated until they have been performed a total of 1000 times, with the learning rate decaying every 100 epochs to a factor of 0.75 of its previous value, starting from an initial value of 0.01.

The weight factor α in Eq. (5.3.1) determines the relative importance of the two optimisation objectives. If it is set to zero, then the adversary has no effect on the training of the classifier. If it is too large, however, the performance of the classifier is severely affected. We find a value of 100 works well for our setup. Furthermore, there is another approach to training the adversarial network, where one updates the weights of both networks simultaneously. However, we find the approach of alternating the training—where the classifier is trained with the adversary weights frozen, and vice versa—to be more stable.

In Fig. 5.3, we show the performance of the adversarial classifier through the classifier output and ROC curves. The adversary has clearly had the effect of shaping the



(a) **(b)**

Figure 5.3: (a) Supervised neural network classifier output and (b) ROC curves for an adversarial classifier trained to classify signal and background events, with sensitivity to the smearing removed. The three background distributions result from the three different directions of smearing.

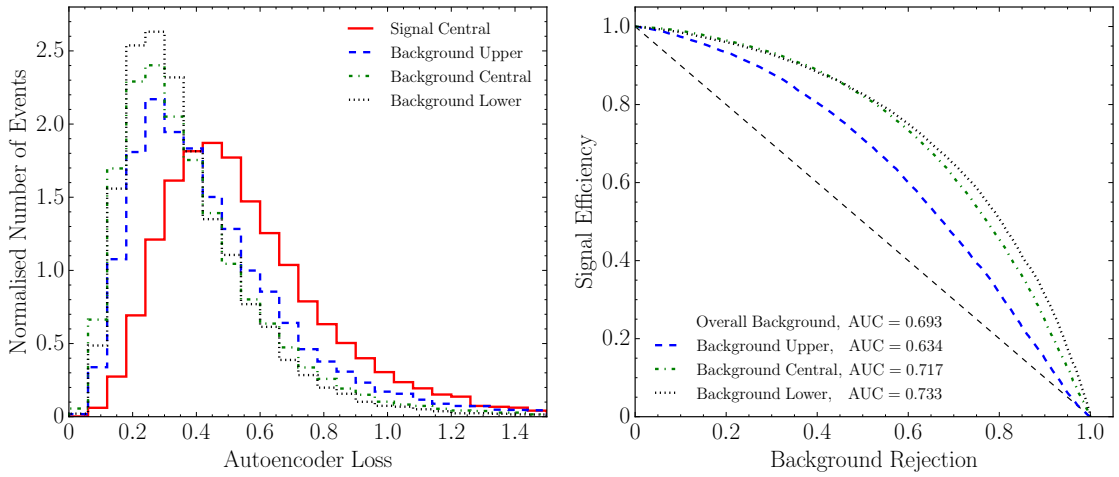
classifier outputs such that their dependence on the background smearing has been almost entirely removed. Thus, the ROC curves and AUC scores become very close together since the classification performance is now barely affected by the smearing.

5.4 Extension to Unsupervised Autoencoder

5.4.1 Adversarial Autoencoder

As we have already described, autoencoders are a type of neural network which can be used in an unsupervised manner as anomaly detectors to search for new physics since they only need to be trained on the background. To this aim, we consider an autoencoder constructed from three hidden layers with 10, 3 and 10 nodes respectively, each with sigmoid activation functions. After the hidden layers, there is a linear output layer[3] with the same dimension as the number of inputs,

[3]A linear output layer is equivalent to no activation function having been applied to the node. Specifically, the activation function is $g(x) = x$.



(a) **(b)**

Figure 5.4: (a) Autoencoder loss and (b) ROC curves for an autoencoder trained only on background events. The three background distributions result from the three different directions of smearing.

which correspond to the 21 observables. The loss is the mean squared error between the inputs and outputs—namely, the autoencoder has the goal of reconstructing the inputs as well as possible, having encoded the information into the latent-compressed layer.

We train the autoencoder on the three background samples using the ADAM optimiser with a learning rate of 0.01 for 500 epochs, and the results are shown in Fig. 5.4. Since the autoencoder is trained only on the background events, it learns how to reconstruct background events better than the signal events, and so the distribution of the losses for the signal events in Fig. 5.4a is at higher values. The ROC curves in Fig. 5.4b are obtained by performing a cut on the loss function and labelling all events above the cut as signal events, and all events below the cut as background events, and then scanning this threshold across all values. This is similar to how the ROC curves are calculated from the output of the supervised classifier, where the threshold is varied between 0 and 1 instead.

As we saw for the classifier, the smearing of the background has an effect on how well the autoencoder can be used to classify events, with the events which have been smeared upwards being mislabelled as signal events more often. It is important to

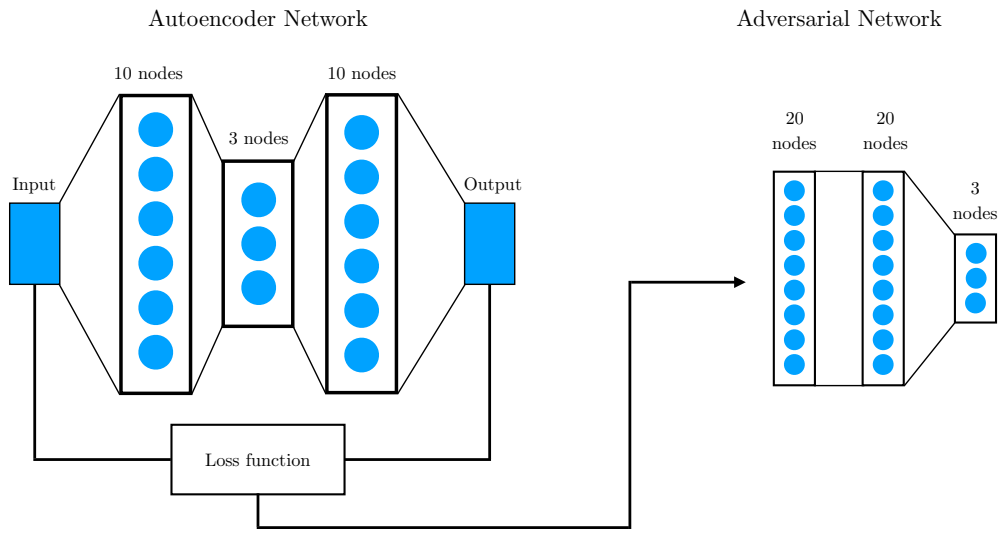


Figure 5.5: Architecture of the adversarial autoencoder. The loss function of the autoencoder is used as an input to the adversary for it to discriminate the smeared background samples.

note that the overall classification performance of the autoencoder is much worse than for the dedicated supervised classifier in Sec. 5.3. However, this is not surprising—the autoencoder is only ever trained on background events, and only sees the signal events during testing. Thus, to distinguish between signal and background it learns the intrinsic kinematic features of the background only. Furthermore, the optimisation objective of the classifier is for it to achieve a strong classification performance—that is not the case for the autoencoder, which simply has the objective of reconstructing the input events well.

We now combine the autoencoder with an adversarial network to improve the reliability and robustness of this unsupervised-learning approach. To achieve the aim of the autoencoder being able to make its predictions independent of the smearing of the background, we use the autoencoder loss as an input to the adversary. Since a threshold on the autoencoder loss is used to perform the classification between signal and background, it is completely analogous to the output of the dedicated classifier used above, on which a cut is placed to classify the events. This input is then followed by two hidden layers each with 20 nodes and ReLu activation functions, with three softmax output nodes and a categorical cross entropy loss. This

architecture is illustrated by the diagram in Fig. 5.5. The training proceeds similarly to the adversarial classifier, but with only background events in the training sample:

1. The autoencoder is trained for three epochs using the ADAM optimiser with a learning rate of 0.01 and a batch size of 500.

2. The adversary is trained for three epochs using mini-batch gradient descent with a learning rate of 0.01 and a batch size of 500.

3. The autoencoder is trained for one epoch with mini-batch gradient descent with a batch size of 500 and with a total loss function,

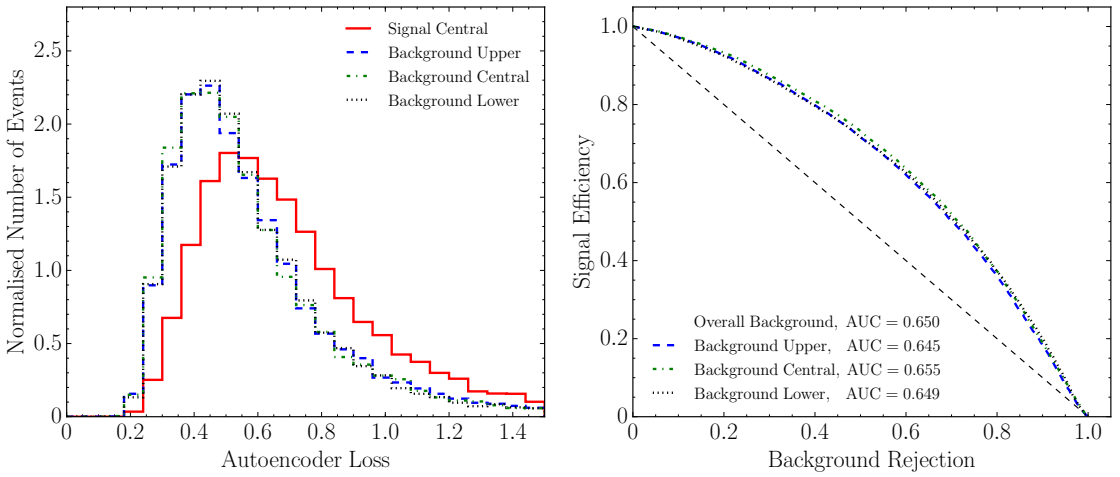
$$\mathcal{L}_{\text{tot}} = \mathcal{L}_{\text{Autoencoder}} - \alpha \mathcal{L}_{\text{Adversary}} \ . \tag{5.4.1}$$

4. The adversary is trained for one epoch using mini-batch gradient descent with a batch size of 500.

5. Steps 3 and 4 are repeated until they have been performed a total of 1500 times, with the learning rate decaying every 100 epochs to a factor of 0.75 of its previous value, starting from an initial value of 0.01.

We find this procedure to provide stable and numerically reliable results. Again, the relative weighting between the autoencoder and the adversary is set to $\alpha = 100$. The performance of the adversarially-trained autoencoder is shown in Fig. 5.6. The background distributions shown in Fig. 5.6a have been shaped such that they are independent of the direction of smearing, which results in the ROC curves in Fig. 5.6b becoming almost identical. This shows that the classification method has become independent of uncertainties inherent to the reconstruction of the final-state objects of LHC events.

In addition, we observe that our setup also has the ability to interpolate to smaller amounts of smearing—although we have trained using background data which has been systematically smeared by a very large amount, we find that if it is subsequently



(a) **(b)**

Figure 5.6: (a) Autoencoder loss and (b) ROC curves for an adversarial autoencoder trained only on background events, with sensitivity to the smearing removed. The three background distributions result from the three different directions of smearing.

tested on samples which have been smeared by a much smaller amount, then the output of the adversarially-trained autoencoder (and also for the classifier in the previous section) is still insensitive to the smearing.

We will now briefly recap what we have achieved by combining an autoencoder with an adversarial neural network. We started with three sets of background events—one which had been smeared upwards, one which had been smeared downwards, and one which had not been smeared at all. This smearing corresponded to the extremities of a Gaussian envelope, and was applied to jets, leptons and the missing energy in each event accordingly. Furthermore, we also had a set of signal events which had not been smeared. The smearing had the effect of shifting the kinematic features of the background such that the events which had been smeared upwards appeared more like signal events, and the ones which had been smeared downwards appeared less like signal events. This can be seen from the distributions in Fig. 5.1.

We then trained an autoencoder on all the background events for the purpose of using it to detect signal events, which have a higher expected reconstruction loss. In Fig. 5.4b, the ROC curves are the result of testing the classification performance of the autoencoder for the signal separately against each background, and as expected,

the autoencoder had a harder time discriminating the signal events against background events which had been smeared upwards. We then combined this with an adversarial neural network, which had the objective of recognising which direction each background sample had been smeared in based upon the loss of the autoencoder. The autoencoder and adversary were trained using a combined loss function, which penalised the autoencoder for outputting reconstruction losses from which the adversary could discriminate the samples. The result of this is that the autoencoder has learnt to reconstruct events without using any information derived from the smearing, which can be seen from the fact that the ROC curves in Fig. 5.6b have converged.

5.4.2 Further Applications

Thus far, the analysis has been carried out on training sets consisting of pure background events. Realistically, data may not actually look like this because if new physics exists, then it would also form part of that same data. However, since the Standard Model already explains many observations well, then the constraints on new physics are such that is reasonable to assume that most events would still be background events, with only a small fraction of the data actually consisting of signal events.

To begin accounting for this, we therefore need to include in the training data a small fraction of signal events. The smearing procedure that is applied to the background events would also then indiscriminately apply to the signal events in the same way. Thus, we inject into the three background sets appropriately smeared signal events. By training on these newly contaminated sets we can investigate how sensitive the performance of the adversarial autoencoder is to an increase in signal corruption in the training set. In Fig. 5.7, we show these results. The band represents the difference between the upper and lower AUC scores, which shows how well the adversary desensitises the autoencoder from the smearing, and the central

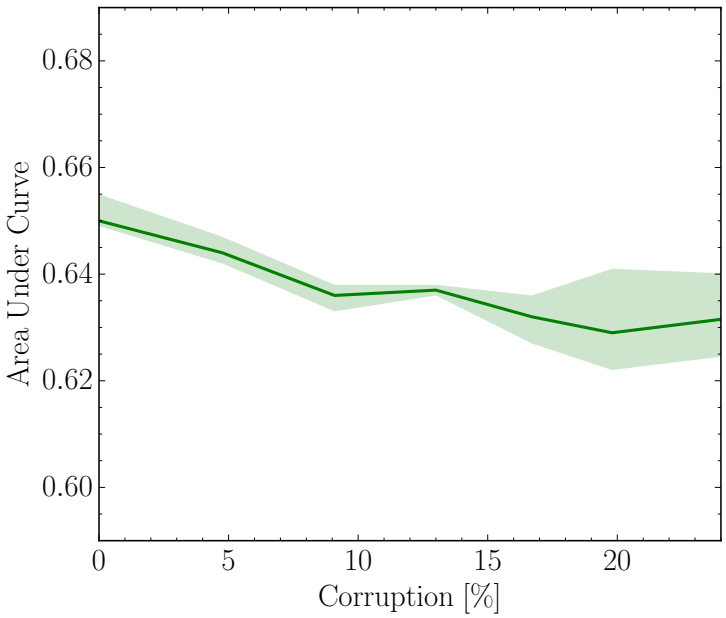


Figure 5.7: Effect of contaminating the training sample with an increasing fraction of signal events. The central line shows the overall AUC score, and the band represents the difference between the upper and lower AUC scores.

line is the overall AUC score. All model hyperparameters of the network setup are left unchanged during the training, with only the relative fraction of corruption changing, defined as a percentage of the total training set. From the plot it is clear that injecting signal events during training has little effect on the overall performance until the fraction of corruption becomes unrealistically large, showing the potential applicability of the method to real data.

Since the performance is not drastically affected by a corruption of the training data, we can proceed with a training sample consisting purely of background events. One of the advantages of the autoencoder only needing to be trained on background events is that it can then be tested for signal events arising from any model. Here, we test our adversarially-trained autoencoder on a variety of new physics models. The models used are:

- Two further Z' cases with widths of 10 GeV and 200 GeV. In both cases the masses are held at 2 TeV.

Signal	Overall AUC	Upper-Lower Difference	Cross Section Limit [pb]
$Z'_{\Gamma=10\ \text{GeV}}$	0.662	0.009	0.0101
$Z'_{\Gamma=89.6\ \text{GeV}}$	0.656	0.009	0.0098
$Z'_{\Gamma=200\ \text{GeV}}$	0.650	0.009	0.0105
Scalar	0.654	0.010	0.0104
Octet	0.659	0.010	0.0102

Table 5.1: Result of training the adversarial autoencoder only on background events and testing it on the original Z' case and four other signals. Shown are the overall AUC score, the different between the largest and smallest AUC scores, and the cross section limits.

- A scalar colour-octet [221], with a mass of 2 TeV and the scalar and axial parameters fixed to ensure the width is ~ 89.6 GeV.

- A scalar colour-singlet with a mass of 2 TeV and a width of 89.6 GeV.

Table 5.1 shows the results of testing the adversarially-trained autoencoder on the new signals. In each case the adversary is able to perform well, with the difference between the upper and lower AUC scores showing that the new signals do not hinder the ability of adversary to desensitise the autoencoder to the smearing. This behaviour is of course expected, since the same background samples are used to test against each new signal.

We also show estimates of the potential limits on the cross sections that can be obtained using the classification performance of the autoencoder. These are calculated by finding the points on the ROC curves that maximise $S/\sqrt{B}$, then comparing them to the background cross section and assuming an integrated luminosity of 100 fb^{-1}. We then require that $S/\sqrt{B} > 2$ to set a 95% confidence limit. The limits we find are insensitive to the nature of the resonance (with respect to the quantum numbers of the resonance), and they are comparable to the limits found by ATLAS in Ref. [222].

5.5 Conclusions

In this chapter, we proposed a new method for reducing the sensitivity of an unsupervised autoencoder—which can be used as an anomaly detector to search for new physics—to experimental uncertainties. This involved combining the autoencoder with an adversarial neural network which had the objective of discriminating the direction that a particular background event had been smeared in, so that the autoencoder learnt to reconstruct the background in a way which was unaffected by this smearing. In order to do this, we used the loss of the autoencoder as the input to the adversarial neural network. The reason for this is that since the loss of the autoencoder is used to discriminate signal from background, it is this particular quantity which needs to be made insensitive to the smearing.

We applied the procedure to resonance searches in semileptonic $t\bar{t}$ events, and showed that the performance is barely affected when the background is corrupted by a small amount of signal. Furthermore, we showed that it also performs equally as well on other signals, which is expected since the autoencoder only needs to learn the kinematic features of the background.

The ideal scenario for the usage of machine-learning techniques is when they can be applied directly on experimental data, without the requirement to train them on pseudo-data generated by Monte Carlo methods, which have inherent theoretical uncertainties. The experimental data alone would be sufficient to be to identify anomalous signal events, which could be isolated and studied further. In this ideal scenario, residual uncertainties due to the imperfect reconstruction of final-state objects remains. Although our method to reduce the sensitivity to these uncertainties was applied on pseudo-data, it may be possible to apply the same procedure on experimental data by also smearing the final-state objects upwards and downwards, and using an adversarially-trained autoencoder to detect anomalous events in the data in a robust way.

Chapter 6

Solving Differential Equations with Neural Networks

In our final investigation, we consider a more unconventional use for neural networks. Neural networks are uniquely suited to solving optimisation problems, and many physics problems can be cast as an optimisation task. Specifically, differential equations—which arise in many areas of physics—can be written in the style of a loss function, with the minimisation of this loss providing the solution to the differential equation. This allows a neural network to be used for solving differential equations, and in this chapter we develop a new approach for doing so. We then use this method to solve the differential equations that arise in the calculation of cosmological phase transitions, which can be used to study the scalar sector of the Standard Model and beyond.

6.1 Motivation

Neural networks are designed to perform an optimisation procedure, with the loss function providing a measure of the performance of the optimisation. Common loss functions used in machine-learning problems are the mean squared error for

regression tasks and categorical cross entropy for classification. Through algorithms based on backpropagation [58], the parameters in the network are updated so that the loss function on the training examples is minimised.

If a physics problem can be written in the form $\mathcal{F}(\vec{x}) = 0$, then an approximation to its solution can be calculated by minimising $\mathcal{F}^2$. Moreover, this encourages one to make the identification between $\mathcal{F}^2$ and the loss function of a neural network. In this scenario, the solution to the problem $\mathcal{F}(\vec{x}) = 0$ could then be approximated by a neural network, with appropriate inputs and outputs determined by the particular problem being solved. This utility of a neural network can be expected from the universal approximation theorem [223, 224], which states that an artificial neural network containing a single hidden layer can approximate any arbitrarily complex function, provided it has enough nodes.

A general type of problem that arises in many areas of physics and which can be cast in this form is the solution of differential equations. For large classes of differential equations, analytic solutions cannot be found, and so numerical or approximative methods are required to solve them. Standard methods to solve differential equations numerically include the Runge-Kutta method, linear multi-step methods, finite-element and finite-volume methods, and spectral methods [225].

One such case where differential equations arise is in the calculation of tunnelling rates for cosmological phase transitions. These are non-perturbative processes that occur in quantum field theory where a scalar field can transition from a metastable vacuum state into a global vacuum state via the propagation of a bubble. This is of importance to determining whether a model of physics beyond the Standard Model can allow for a strong first-order phase transition during the early Universe, which is one of the necessary conditions to explain the abundance of matter over anti-matter [226, 227]. Such a process could lead to a stochastic gravitational wave signal which could be measured at future gravitational wave experiments [228, 229]. Thus the reliable and precise calculations of tunnelling rates is important.

Therefore, we propose a new method for using neural networks to solve differential equations, which can then be applied to the calculation of tunnelling rates. In contrast to previous approaches where the neural network is part of a full trial solution which is fixed to satisfy the boundary conditions [230–235], our approach includes the boundary conditions as additional terms in the loss function instead. This means that the output of the neural network alone solves the differential equation, which improves the flexibility and applicability of the method.

The remainder of this chapter is structured as follows: in Sec. 6.2 we will describe the design of the neural network and its loss function, and the procedure that we use to optimise it. We will then apply this to various simple cases of differential equations in Sec. 6.3, before using it for the calculation of cosmological phase transitions in Sec. 6.4. Finally, the chapter is concluded in Sec. 6.5.

6.2 Design of the Network and Optimisation Procedure

We consider an artificial feedforward neural network (NN) with n inputs and m outputs. For a single hidden layer with k units, the outputs of the network, N_m, can be written compactly as,

$$N_m(\vec{x}, \{w, \vec{b}\}) = \sum_{k,n} w^f_{mk} g(w^h_{kn} x_n + b^h_k) + b^f_m \ , \tag{6.2.1}$$

where the activation function $g : \mathbb{R}^k \mapsto \mathbb{R}^k$ is applied element-wise to each unit, and h and f denote the hidden and final layers, respectively. In this work, we use a single neural network with m outputs to predict the solutions to m coupled differential equations—for the case of one differential equation, we use $m = 1$. Similarly, we use a single neural network with n inputs to predict the solution to a partial differential equation in n variables.

A set of m coupled jth order differential equations can be expressed in the general form,

$$\mathcal{F}_m(\vec{x}, \phi_m(\vec{x}), \nabla\phi_m(\vec{x}), \ldots, \nabla^j\phi_m(\vec{x})) = 0 \ , \tag{6.2.2}$$

with boundary or initial conditions imposed on the solutions $\phi_m(\vec{x})$. Writing the differential equations in such a way allows us to easily convert the problem of finding a solution into an optimisation one. An approximate solution $\hat{\phi}_m(\vec{x})$ is one which approximately minimises the square of the left-hand side of Eq. (6.2.2), and thus the analogy can be made to the loss function of a neural network. In previous approaches [230–233], $\hat{\phi}_m(\vec{x})$ is a trial solution composed of two parts: one which satisfies the boundary conditions, and one which is a function of the output of a neural network and vanishes at the boundaries. However, this requires one to choose a special form of the trial solution which is dependent on the boundary conditions. Furthermore, for some configurations of boundary conditions, finding such a trial solution is a very complex task, such as in the case of phase transitions. Instead, we identify the trial solution with the network output, $\hat{\phi}_m(\vec{x}) \equiv N_m(\vec{x}, \{w, \vec{b}\})$, and include the boundary conditions as extra terms in the loss function. If the domain is discretised into a finite number of training points $\vec{x}^i$, then approximations to the solutions, $\hat{\phi}_m(\vec{x})$, can be obtained by finding the set of weights and biases, $\{w, \vec{b}\}$, such that the neural network loss function is minimised on the training points. For $i_{\max}$ training examples, the full loss function that we use is,

$$\begin{aligned}\mathcal{L}(\{w, \vec{b}\}) = \frac{1}{i_{\max}} \sum_{i,m} \hat{\mathcal{F}}_m(\vec{x}^i, \hat{\phi}_m(\vec{x}^i), \ldots, \nabla^j\hat{\phi}_m(\vec{x}^i))^2 \\ + \sum_{\text{BC}} (\nabla^p\hat{\phi}_m(\vec{x}_b) - K(\vec{x}_b))^2 \ ,\end{aligned} \tag{6.2.3}$$

where the second term[1] represents the sum of the squares of the boundary conditions, defined at the boundaries $\vec{x}_b$. These can be Dirichlet or Neumann, or they can be initial conditions if defined only at the initial part of the domain.

[1]Here, p represents the order of derivative for which the boundary condition is defined, and K is a function on the boundary. For example, for the condition $\frac{d}{dx}\phi(0) = 1$ the second term would be $\left(\frac{d}{dx}\hat{\phi}(0) - 1\right)^2$.

The problem is then to minimise $\mathcal{L}(\{w, \vec{b}\})$ by optimising the weights and biases in the network, for a given choice of network setup. To calculate the loss, it is necessary to compute the derivatives of the network output with respect to its input. Since each part of the network, including the activation functions, are differentiable, the derivatives can be obtained analytically. Ref. [233] outlines how to calculate these derivatives. The optimisation can then proceed via backpropagation by further calculating the derivatives of the loss itself with respect to the network parameters. We use the KERAS framework [218] with a TENSORFLOW [219] backend to implement the network and perform the optimisation of the loss function. In addition, we also cross check these results with a custom implementation written using the NUMPY [236] and AUTOGRAD [237] packages.

As with any neural network, the choice of hyperparameters will have an effect on the performance. For our setup, the important hyperparameters are the number of hidden layers, the number of units in each hidden layer, the number of training points $\vec{x}^i$ (corresponding to the number of anchors in the discretisation of the domain of the differential equation), the activation function in each hidden layer, the optimisation algorithm, the learning rate, and the number of epochs the network is trained for. Furthermore, a choice must be made for the size of the domain that contains the points that the network is trained on, but this will usually be determined by the problem being solved.

In all the examples, we use the ADAM optimiser [217] with learning rate reduction on plateau (when the loss plateaus, the learning rate is reduced) and an initial learning rate of 0.01. We find that the network is not susceptible to overfitting—the training points are chosen exactly from the domain that one is trying to find the solution on, and are not subject to statistical fluctuations. Thus, finding a solution for which the loss at every training point is zero would not limit the generalisation of the solution to other points within the domain. Therefore, we use a large number of epochs such that the training loss becomes very small. For all examples we use a conservative number of 50,000 epochs. Furthermore, we use batches consisting of the entire set of

training points so that each update of the network parameters includes every point in addition to the boundary conditions. We also find that, in general, a single hidden layer with a small number of units—$\mathcal{O}(10)$—is sufficient to obtain very accurate solutions.

In order to assess and improve the stability and performance in certain cases, there are some additional technical methods which we employ beyond the basic setup. Firstly, the differentiability of the network solution allows us to calculate the differential contribution, $\hat{\mathcal{F}}$, to the loss across the entire training domain. This shows the degree of accuracy to which each part of the network solution satisfies the differential equation, and can be used for assessing the performance in cases where the analytic solution is not known. Secondly, for coupled differential equations with initial conditions, we find that the stability of the solution can be improved by iteratively training on increasing domain sizes—that is, a solution is found on a smaller domain, before keeping the trained network parameters and training again on a larger domain. Finally, for the calculation of phase transitions, we employ a two-step training where initially the boundaries are chosen to be the true and false vacua, before the precise boundary conditions are used in the second training. This prevents the network from finding the trivial solution where the field is always in the false vacuum.

6.3 Application to Simple Examples

6.3.1 Ordinary Differential Equation Examples

To show how well the method can solve ordinary differential equations (ODEs), we apply it to both a first and a second order ODE, which have known analytic solutions. The equations we study are,

$$\frac{d\phi}{dx} + \left(x + \frac{1+3x^2}{1+x+x^3} \right) \phi - x^3 - 2x - x^2 \frac{1+3x^2}{1+x+x^3} = 0 \ , \tag{6.3.1}$$

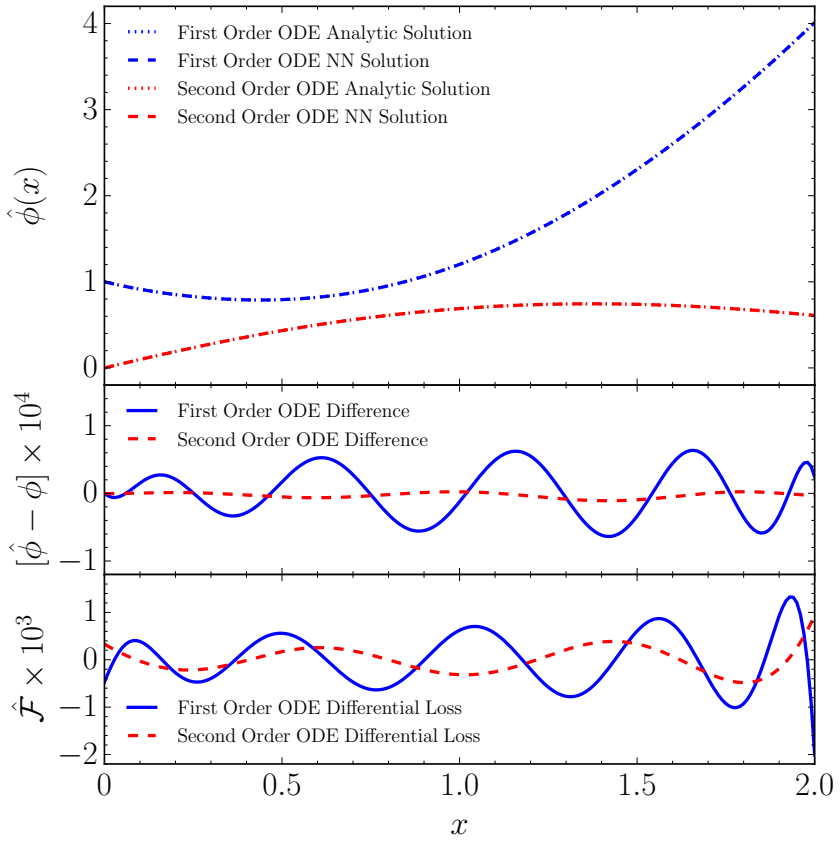


Figure 6.1: The upper panel shows the solutions to the first and second order ODEs of Eqs. (6.3.1) and (6.3.2), with boundary conditions as outlined in the text. The middle panel shows the numerical difference between the analytic solution and the NN predicted solution for both cases. The lower panel shows the differential contribution $\hat{\mathcal{F}}$ to the loss across the entire domain.

with the boundary condition $\phi(0) = 1$ in the domain $x \in [0, 2]$ and,

$$\frac{d^2\phi}{dx^2} + \frac{1}{5}\frac{d\phi}{dx} + \phi + \frac{1}{5}e^{-\frac{x}{5}}\cos x = 0 \ , \tag{6.3.2}$$

with boundary conditions $\phi(0) = 0$ and $\frac{d}{dx}\phi(0) = 1$ in the domain $x \in [0, 2]$.

As a simple neural network structure, we choose a single hidden layer of 10 units with sigmoid activation functions, and we discretise the domain into 100 training examples. It is then just a case of passing the differential equations and boundary conditions to the loss function, as described in Eq. (6.2.3), and proceeding with the optimisation.

Fig. 6.1 shows the results of the neural network output, compared to the analytic solutions of Eqs. (6.3.1) and (6.3.2). The middle panel of Fig. 6.1 shows the absolute

numerical difference between the numerical and analytic solutions. This difference can be reduced further by increasing the number of epochs, the number of points in the discretisation of the domain, or the number of units in the hidden layer. Thus, the neural network provides handles to consistently improve the numerical accuracy one aims to achieve. The lower panel of Fig. 6.1 shows the differential contribution to the loss function from each training example. As we will describe shortly, if the solution is not analytically known, this provides a measure to assess whether the found solution is the correct one or if a numerical instability led the network to settle in a local minimum for the loss.

6.3.2 Coupled Differential Equation Example

When discussing the calculation of cosmological phase transitions, we will study the solution of coupled non-linear differential equations, for which no closed analytic form is known. Here, we will first show that such solutions can be obtained with our approach, for a case where analytic solutions are already known. We consider,

$$\begin{aligned} \frac{d\phi_1}{dx} - \cos x - \phi_1^2 - \phi_2 + 1 + x^2 + \sin^2 x = 0 \ , \\ \frac{d\phi_2}{dx} - 2x + (1 + x^2)\sin x - \phi_1\phi_2 = 0 \ , \end{aligned} \tag{6.3.3}$$

with boundary conditions,

$$\phi_1(0) = 0 \ , \qquad \phi_2(0) = 1 \ . \tag{6.3.4}$$

If the boundary conditions are defined at one edge of the domain, such as at $x = 0$, it requires an increasingly elaborate network to maintain numerical stability for the solution over a large domain where $x \gg 1$. This is due to small numerical instabilities during backpropagation because of the complexity of the loss hypersurface. If such numerical instabilities lead the network to choose a path that is in close proximity to the true solution, the NN can settle on a local minimum with a small value for the loss function. To solve this problem, we propose to incrementally extend the domain

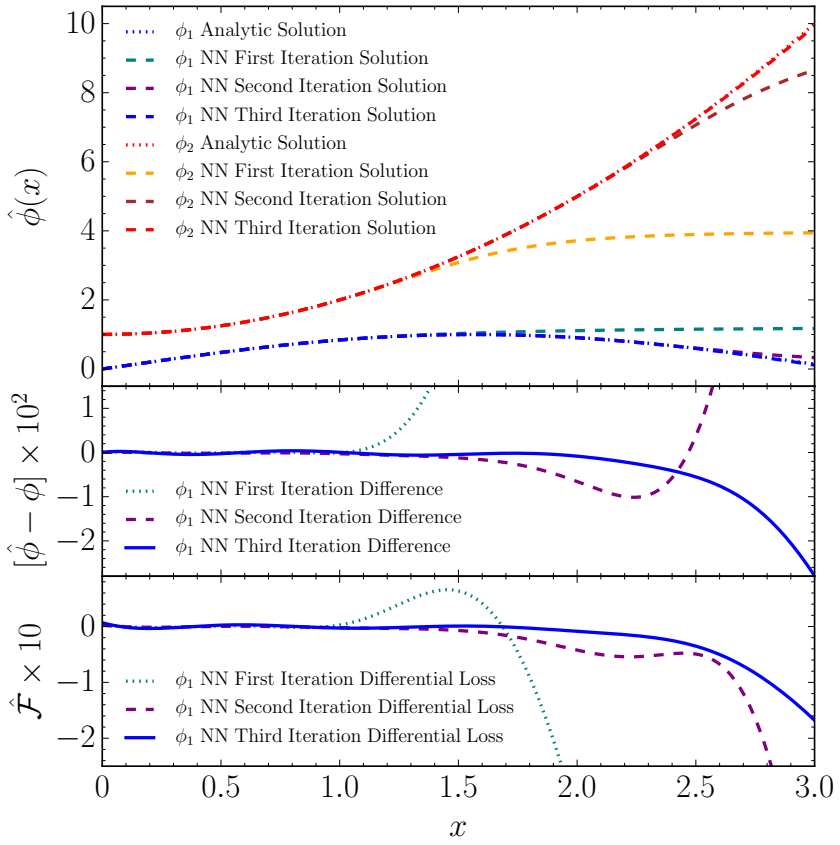


Figure 6.2: The upper panel shows the solutions for the functions ϕ_1 and ϕ_2 to the coupled differential equation of Eq. (6.3.3). The middle panel displays the numerical difference between the analytic solution and the NN predicted solution for ϕ_1. The lower panel shows the differential contribution $\hat{\mathcal{F}}$ to the loss across the entire domain, from the equation for ϕ_1. The three NN curves in each panel correspond to the first, second and third iteration steps in the training of the network.

on which a solution should be found, by partitioning the training examples and increasing the number of partitions the NN is trained on in each step. If the weights the NN has learned in the previous step are then retained before training for the next step—such that the network only has to learn the function on the part of the domain that was incrementally increased—we find that one can achieve numerical stability for an arbitrarily large domain.

We show this mechanism in Fig. 6.2, where we have partitioned the full domain containing 100 training examples into three regions each of size 1. The network structure again consists of a single hidden layer of 10 units with sigmoid activation functions, and with two units in the final layer, since there are two coupled equations.

The upper panel shows the solutions for ϕ_1 and ϕ_2 for each iterative step. While the first iteration only allows a solution to be found on a smaller domain—here from 0 to 1—subsequent steps, and in particular the third step, allow an accurate solution to be found over the entire domain. Again, the differential $\hat{\mathcal{F}}$ proves to be a good indicator of whether the calculated solution is satisfying the differential equation over the entire domain—this can be seen from the lower panel of Fig. 6.2.

6.3.3 Partial Differential Equation Example

While we do not study partial differential equations (PDEs) in the later physics examples of calculating phase transitions, we showcase here the flexibility of our NN method. With the same network architecture as used for the solution of the ordinary differential equations (except for an extra input unit for each additional variable), we can apply our approach to the solution of partial differential equations. The precise solution of such equations is a widespread problem in physics, such as in mechanics, thermodynamics and quantum field theory. As an example, we choose the second order partial differential equation,

$$\nabla^2\phi - e^{-x}(x - 2 + y^3 + 6y) = 0 \ , \tag{6.3.5}$$

with boundary conditions,

$$\begin{aligned} &\phi(0, y) = y^3 \ , && \phi(1, y) = (1 + y^3)e^{-1} \ , \\ &\phi(x, 0) = xe^{-x} \ , && \phi(x, 1) = e^{-x}(x + 1) \ , \end{aligned} \tag{6.3.6}$$

for which an exact analytic solution is known. In Fig. 6.3 we show the difference between the numerical solution as predicted by the NN and the analytic solution over the domain $(x, y) \in [0, 1] \times [0, 1]$. The 100 training examples were chosen from an evenly spaced 10×10 grid. As the value of $\phi(x, y)$ is of $\mathcal{O}(1)$ for most of the domain, the relative and absolute accuracies are similar, so we only show the absolute accuracy here. Across the entire domain, we find a numerical solution with

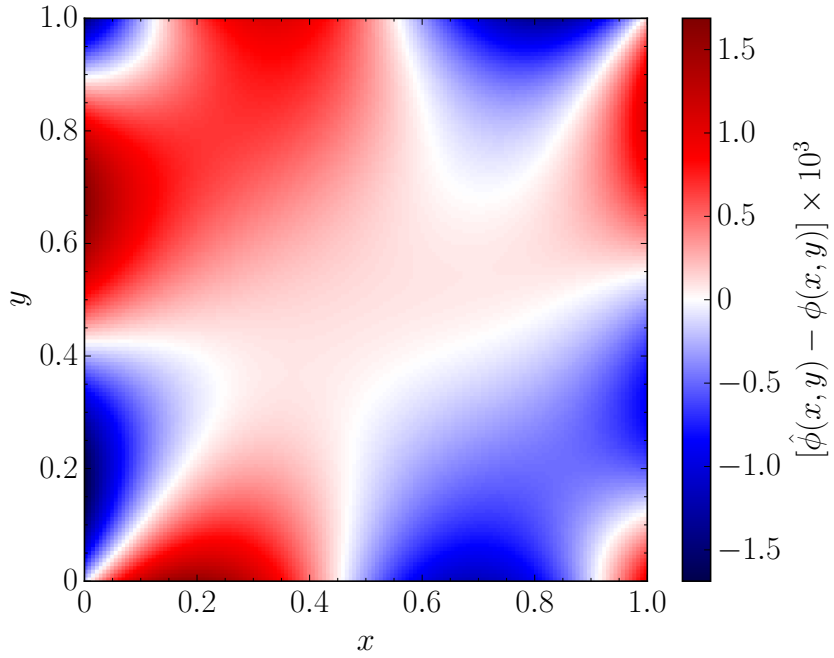


Figure 6.3: Numerical difference between the analytic solution and the NN predicted solution of Eq. (6.3.5), with boundary conditions as given in Eq. (6.3.6), over the domain $(x, y) \in [0, 1] \times [0, 1]$.

very good absolute and relative accuracy for this second order partial differential equation. However with a deeper NN, such as the addition of a second hidden layer with 10 tanh units, we find that the accuracy improves by an order of magnitude further. Deeper and wider networks result in even better accuracies.

6.4 Application to the Calculation of Phase Transitions

As we discussed in Chapter 1, the observed abundance of matter over antimatter in the Universe [227, 238] is an unsolved problem in physics. Furthermore, the need for a dynamical generation of this baryon asymmetry is dictated by inflation—the entropy production occurring in the reheating phase of inflation would have washed out any asymmetry already present at the beginning of the Universe's evolution [239]. A model of baryogenesis was proposed by Sakharov in 1967 [226], and must

be accommodated by any fundamental theory capable of addressing the baryon asymmetry problem. This is commonly translated into three necessary conditions: baryon number violation, C- and CP-violation, and loss of thermal equilibrium. While the first condition can be satisfied in the Standard Model by sphalerons, the second[2] and third conditions require it to be extended [240–242]. Departure from thermal equilibrium can be obtained during a strong first-order phase transition, which is usually accompanied by a sudden change of symmetry [243]. Within the Standard Model, this could have occurred during the spontaneous symmetry breaking of the electroweak gauge symmetry, when the Universe had the temperature $T \sim 100$ GeV [244, 245]. In order to assess whether this might have been the case and to further test its possibility in extensions of the Standard Model, it is important to discuss the conditions for scalar-field phase transitions at finite temperature.

Quantum fluctuations allow the transition between two vacua of the potential $V(\vec{\phi})$, where $\vec{\phi}(x)$ is an m-dimensional real scalar field. When these vacua are not degenerate, the false vacuum $\vec{\phi}_F$—which corresponds to a local minimum—becomes unstable under barrier penetration, and can decay into the true vacuum $\vec{\phi}_\mathrm{T}$ of the potential. The tunnelling process converts a homogeneous region of false vacuum into one of approximate true vacuum—a bubble. Far from this region the false vacuum persists undisturbed [246]. The four-dimensional Euclidean action for this process has the form,

$$\mathcal{S}_4(\vec{\phi}) = \int d\tau d^3x \left[\frac{1}{2}\left(\frac{d\vec{\phi}}{d\tau}\right)^2 + \frac{1}{2}(\nabla\vec{\phi})^2 + V(\vec{\phi})\right] . \tag{6.4.1}$$

The extension to the description of the tunnelling action at finite temperatures follows from the equivalence between the quantum statistics of bosons (fermions) at $T \neq 0$ and Euclidean quantum field theory, periodic (anti-periodic) in the Euclidean time τ with period T^{-1}. In the calculation of $\mathcal{S}_4(\vec{\phi})$, the integration over τ can be replaced by multiplication by T^{-1} [247], leaving the three-dimensional Euclidean

[2]There is an insufficient amount of CP-violation in the Standard Model to explain baryogenesis.

action,

$$\mathcal{S}_3(\vec{\phi}) = \int d^3x \left[\frac{1}{2} (\nabla \vec{\phi})^2 + V(\vec{\phi}, T) \right] , \tag{6.4.2}$$

with the relation $\mathcal{S}_4(\vec{\phi}) = T^{-1} \mathcal{S}_3(\vec{\phi})$. Suggested by the symmetry of the physical problem, we assume $\vec{\phi}(\vec{x})$ to be invariant under three-dimensional Euclidean rotations [248], and define $\rho \equiv \sqrt{\vec{x}^2}$. The bubble configuration $\vec{\phi}_b(\rho)$ is then the solution to the Euler-Lagrange equation of motion,

$$\frac{d^2 \vec{\phi}}{d\rho^2} + \frac{2}{\rho} \frac{d\vec{\phi}}{d\rho} = \nabla V , \tag{6.4.3}$$

where the gradient of the potential is with respect to the field $\vec{\phi}$. Furthermore, the boundary conditions are,

$$\frac{d}{d\rho} \vec{\phi}(0) = 0 , \qquad \lim_{\rho \to \infty} \vec{\phi}(\rho) = \vec{\phi}_{\mathrm{F}} . \tag{6.4.4}$$

The bubble configuration thus minimises the action in Eq. (6.4.2). The probability per unit time and unit volume for the metastable vacuum to decay is given by,

$$\frac{\Gamma}{V} = A e^{-B/T} . \tag{6.4.5}$$

This is maximised by the bounce,

$$B = \mathcal{S}_3(\vec{\phi}_b) - \mathcal{S}_3(\vec{\phi}_{\mathrm{F}}) , \tag{6.4.6}$$

where $\mathcal{S}_3(\vec{\phi}_{\mathrm{F}})$ is the action evaluated at the stationary configuration $\vec{\phi}_{\mathrm{F}}$. A complete expression for the factor A in Eq. (6.4.5) would require sophisticated computations of differential operator determinants [239]. Instead, an estimate can be obtained from dimensional analysis, which gives $A \simeq T^4$ [249].

Dedicated methods for calculating the nucleation rate, by finding a solution for the bubble profile $\vec{\phi}_b$ to the non-linear coupled differential equations of Eq. (6.4.3), exist and have been implemented in publicly available codes such as CosmoTransitions [250] and BubbleProfiler [251]. For the single-field case, both CosmoTransitions and BubbleProfiler use variants of the overshooting and undershooting method. In the multiple-field case, BubbleProfiler applies the Newton-

Kantorovich method [252], which is described in Ref. [253]. CosmoTransitions instead uses a method that splits the equation of motion into both parallel and perpendicular components along a test path through field space. The path is then varied until a configuration is found that simultaneously solves both directions of the equations of motion. A further code to calculate the tunnelling rates is given in Ref. [254]. An approach using neural networks to directly learn bounce actions from potentials was described in Ref. [255]. Recently, a novel approximative approach for single [256] and multiple fields [257] was proposed, and a new method based on exact analytic solutions of piecewise linear potentials is outlined in Ref. [258]. Older numerical approaches to calculating bubble profiles and tunnelling rates include Refs. [259–262].

6.4.1 Phase Transition with a Single Scalar Field

As a first application of our method to the computation of cosmological phase transitions, we consider the case of a single scalar field. In this scenario, Eq. (6.4.3) then has a straightforward classical analogy—it describes the motion of a particle with a coordinate $\phi(\rho)$ subject to the inverted potential $-V(\phi)$ and to a peculiar-looking damping force which decreases with time. The problem then reduces to finding the initial position ϕ_0, in the vicinity of ϕ_T, such that the particle stops at ϕ_F as $\rho \to \infty$.

The existence of a solution for this case was proven in Ref. [246]. Starting either too close or too far from ϕ_T would result in the solution missing the final configuration ϕ_F, due to overshooting and undershooting, respectively. The continuity of $\phi(\rho)$ thus implies that there must exist an intermediate initial position ϕ_0 which solves the boundary conditions in Eq. (6.4.4). The solution presents two limiting profiles, determined by the ratio of $\Delta \equiv V(\phi_\mathrm{F}) - V(\phi_\mathrm{T})$ to the height of the potential barrier $V(\phi_\mathrm{bar})$. If this ratio is $\gtrsim 1$, which corresponds to the thick-wall case, the particle will overshoot unless its initial energy is similar to $V(\phi_\mathrm{F})$. Conversely, if this

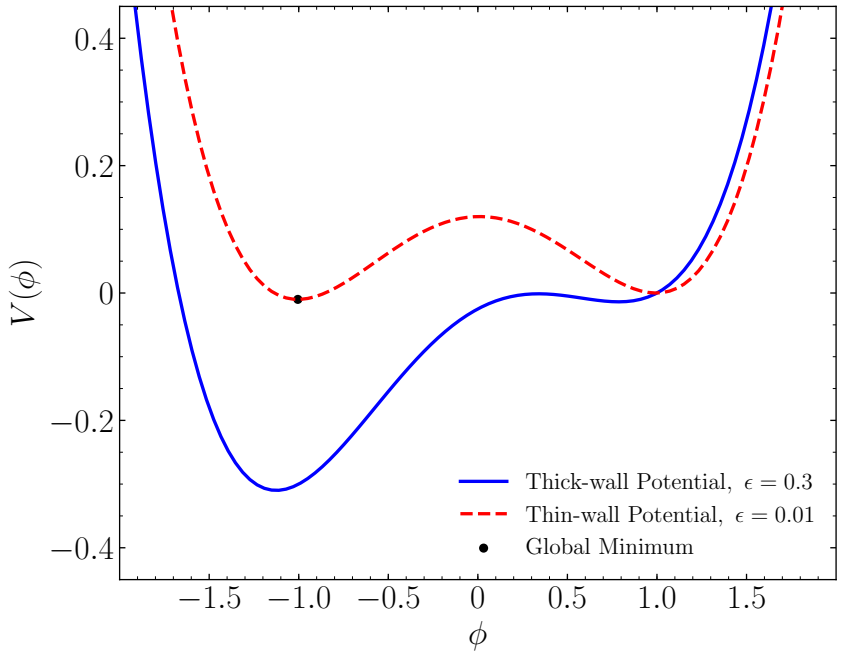


Figure 6.4: Plot of the potential in Eq. (6.4.9), with $\lambda = \alpha = 1$, for the thick-wall (blue solid) and the thin-wall (red dashed) cases. For the latter, the position of the global minimum is also marked by the black dot for clarity.

ratio is small, corresponding to the thin-wall case, in order to avoid undershooting the particle must wait close to $\phi_{\rm T}$ until the time $\rho \simeq R$, when the damping force has become negligible. The value of R can be determined exactly in the thin-wall limit [246] using,

$$R = \frac{2\sigma}{\Delta} \ , \tag{6.4.7}$$

where the surface tension σ is given by,

$$\sigma = \lim_{\Delta \to 0} \int_{\phi_{\rm F}}^{\phi_{\rm T}} d\phi \sqrt{2[V(\phi) - V(\phi_{\rm F})]} \ . \tag{6.4.8}$$

We test our method on the potential [251],

$$V(\phi) = \frac{\lambda}{8}(\phi^2 - a^2)^2 + \frac{\epsilon}{2a}(\phi - a) \ , \tag{6.4.9}$$

and set $\lambda = a = 1$. Two distinct and non-degenerate minima exist for $0 < \epsilon \lesssim 0.3$, with the upper bound representing the thick-wall limit and smaller values of ϵ representing progressively thinner cases. A plot of the potential is shown in Fig. 6.4, for the values $\epsilon = 0.01$ and $\epsilon = 0.3$ which we consider as our thin-wall and thick-wall

cases, respectively.

For the boundary conditions in Eq. (6.4.4), it is clearly not possible to implement an infinite domain for the training of a neural network, and the divergence in the second term of Eq. (6.4.3) prevents the equation from being evaluated at $\rho = 0$. Therefore, a training domain $\rho \in [\rho_{\min}, \rho_{\max}]$ must be chosen. Since the solution approaches the boundaries exponentially, it can be safely assumed that the numerical uncertainties induced by these choices can be neglected, provided that $\rho_{\min}$ is sufficiently small and $\rho_{\max}$ is sufficiently large. To help in choosing this domain, the identification of ϵ in Eq. (6.4.9) with Δ in Eq. (6.4.7) can be made, and the bubble radius R calculated. We then use $\rho_{\max} = 5R$ for the thick-wall case, and $\rho_{\max} = 2R$ for the thin-wall case (since Eq. (6.4.7) underestimates the true radius for thick-wall cases). Furthermore, we use $\rho_{\min} = 0.01$ for both cases. Although these choices may seem arbitrary, we find that the solution converges provided that the transition point is contained well inside the domain, and the result remains stable even if larger domains are used. The boundary conditions then have the form,

$$\frac{d}{d\rho}\phi(\rho_{\min}) = 0 \ , \qquad \phi(\rho_{\max}) = \phi_{\mathrm{F}} \ . \tag{6.4.10}$$

Our NN method can then be applied to find the bubble profile by solving the Euler-Lagrange equation in Eq. (6.4.3). In this context, the NN method corresponds to an approach where the neural network attempts to apply the minimum-action principle to the Euclidean action of Eq. (6.4.2). The test-field configuration, defined by the output layer of the neural network, is then adjusted using backpropagation until the classical trajectory is found. We discretise the domain into 500 training points and choose a network with a single hidden layer. For the thick-wall case, we use 10 hidden units with a sigmoid activation function, which we also used in the earlier simple examples. However, for the thin-wall case, we find that a single tanh unit is sufficient to achieve very good performance since the solution itself closely resembles a tanh function. To prevent the network from finding the trivial solution where the field remains in the false vacuum forever, we first train the network with the

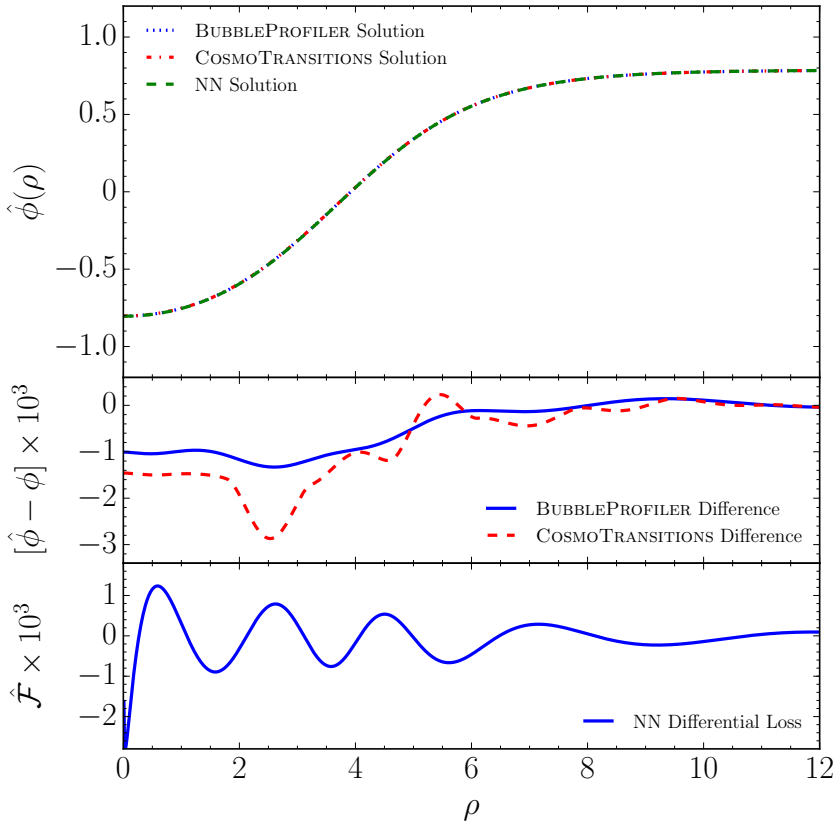


Figure 6.5: The upper panel shows the bubble profile for the thick-wall potential ($\epsilon = 0.3$) in Eq. (6.4.9) for one scalar field, as obtained by our NN method, BubbleProfiler and CosmoTransitions. The middle panel displays the numerical difference between the NN predicted solution and the solutions from the other two codes. The lower panel shows the differential contribution $\hat{\mathcal{F}}$ to the loss.

boundary condition at $\rho_{\min}$ modified to $\phi(\rho_{\min}) = \phi_{\mathrm{T}}$ so that the network finds a solution in the vicinity of the correct solution, since the starting point is close to the true vacuum, before training again with the correct boundary conditions. We use this two-step training for all phase transition calculations.

Our results for the thick-wall and thin-wall cases are shown in Figs. 6.5 and 6.6, respectively, together with the CosmoTransitions and BubbleProfiler solutions. While all three methods agree very well for the thick-wall case, there is a disagreement in CosmoTransitions compared with BubbleProfiler and the NN approach in the thin-wall case. The dotted vertical line indicates where the bubble radius should be according to Eq. (6.4.7). Both BubbleProfiler and NN find a solution that matches the analytic calculation for the bubble radius. Cosmo-

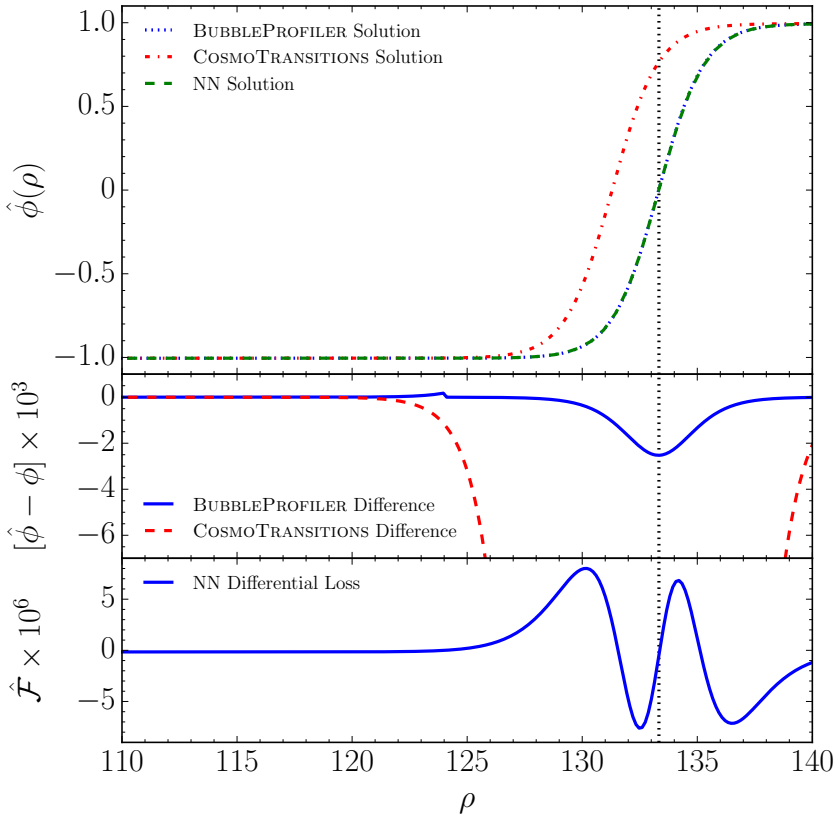


Figure 6.6: The upper panel shows the bubble profile for the thin-wall potential ($\epsilon = 0.01$) in Eq. (6.4.9) for one scalar field, as obtained by our NN method, BubbleProfiler and CosmoTransitions. The middle panel displays the numerical difference between the NN predicted solution and the solutions from the other two codes. The lower panel shows the differential contribution $\hat{\mathcal{F}}$ to the loss. The dotted vertical line shows the analytic location of the bubble radius.

Transitions instead finds a solution with a smaller bubble radius, and therefore a smaller action and a larger tunnelling rate.

For thin-wall cases, numerical stability is difficult to achieve. It is possible for an approximate solution to be found, which transitions at a much earlier ρ than it should, since a translated solution also approximately solves the differential equation [254]. For our method, $\hat{\mathcal{F}}$ can be monitored during the course of the training. During the early stages of the training where the solution does not yet have the correct transition point, $\hat{\mathcal{F}}$ will be sharply distributed in the region of the incorrect transition. As the training proceeds and the solution converges, the function will flatten out until an accurate solution is found across the entire domain.

We have shown that the NN achieved very good stability for the thin-wall case using a single tanh function. We also explored the idea of using an adaptive distribution of training examples, such that more examples are distributed close to the region where the transition of the NN solution happens, and this distribution is then modified over the course of the training. A larger contribution to the loss in this region will be amplified by having more training examples, which can speed up learning. We found that the results can be improved by using this procedure, and this is an idea which could be investigated further in future work.

6.4.2 Phase Transition with Two Scalar Fields

To investigate how well the NN approach can solve the differential equation of Eq. (6.4.3) for multiple fields, we consider a potential for two scalar fields [250],

$$V(\phi_1, \phi_2) = (\phi_1^2 + \phi_2^2)\left[\frac{9}{5}(\phi_1 - 1)^2 + \frac{1}{5}(\phi_2 - 1)^2 - \delta\right], \tag{6.4.11}$$

which has a local minimum at $\phi_1 = \phi_2 = 0$ and a global minimum near $\phi_1 \simeq \phi_2 \sim 1$. We focus again on the thick- and thin-wall cases, setting $\delta = 0.4$ for the former and $\delta = 0.02$ for the latter. For the thick-wall potential, we solve the coupled equations in Eq. (6.4.3) with the boundary conditions,

$$\begin{aligned} &\frac{d}{d\rho}\phi_1(\rho_{\min}) = 0\ , \qquad \phi_1(\rho_{\max}) = 0\ , \\ &\frac{d}{d\rho}\phi_2(\rho_{\min}) = 0\ , \qquad \phi_2(\rho_{\max}) = 0\ , \end{aligned} \tag{6.4.12}$$

in the training domain $\rho \in [0.01, 6]$ with 500 training points. Again, the NN is built with 10 units in a single hidden layer with a sigmoid activation function. Since there are two fields, the NN has two units in the final layer.

The two components ϕ_1 and ϕ_2 of the bubble solution, and the associated path through field space, are shown in Figs. 6.7 and 6.8, respectively. Once more, BubbleProfiler and the NN predictions agree very well, both for the one-dimensional profiles for ϕ_1 and ϕ_2, and for the path in the ϕ_1-ϕ_2 plane. CosmoTransitions

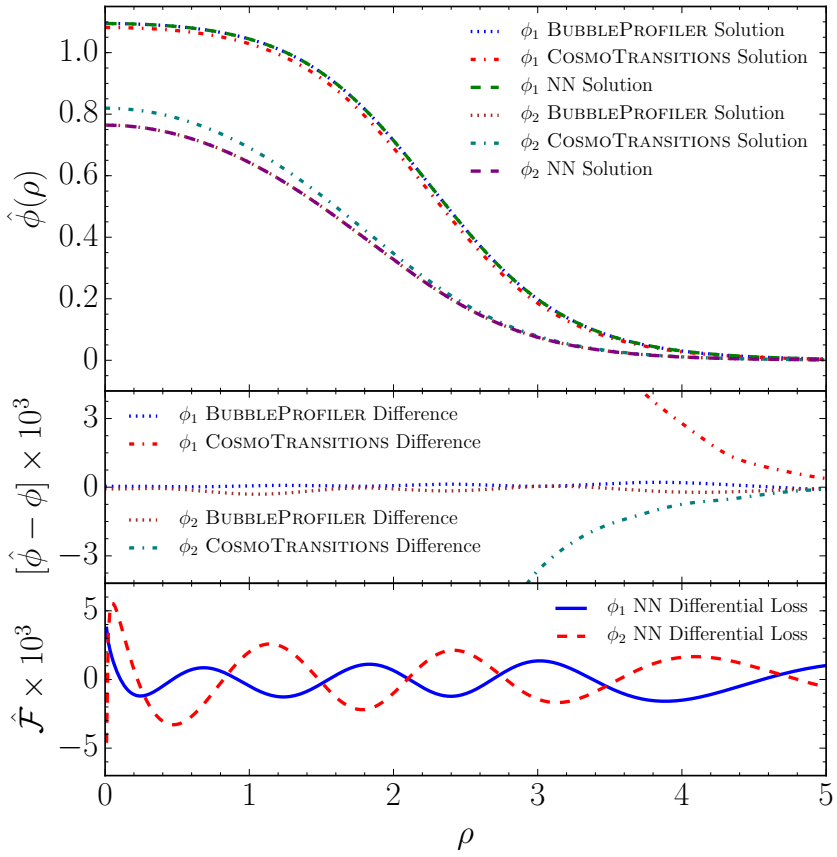


Figure 6.7: The upper panel shows the bubble profiles for the thick-wall potential in Eq. (6.4.11) for two scalar fields, as obtained by our NN method, BubbleProfiler and CosmoTransitions. The middle panel displays the numerical difference between the NN predicted solutions and the solutions from the other two codes. The lower panel shows the differential contribution $\hat{\mathcal{F}}$ to the loss from ϕ_1 and ϕ_2.

shows a slightly different shape for the solutions of $\phi_1(\rho)$ and even more so for $\phi_2(\rho)$, resulting in a slightly modified escape path in Fig. 6.8. The behaviour and small numerical value of the differential contribution $\hat{\mathcal{F}}$ to the loss suggests that the NN has converged to a correct solution for the profiles. Since it also agrees very closely with the result from BubbleProfiler, we conclude that in this case the BubbleProfiler result is correct. Note also that our NN solution has found initial positions for the fields which agree with those from BubbleProfiler. In thick-wall cases, these can differ significantly from the true vacuum ϕ_T—these initial positions have been independently found by the network during optimisation and have not been used as an input during training.

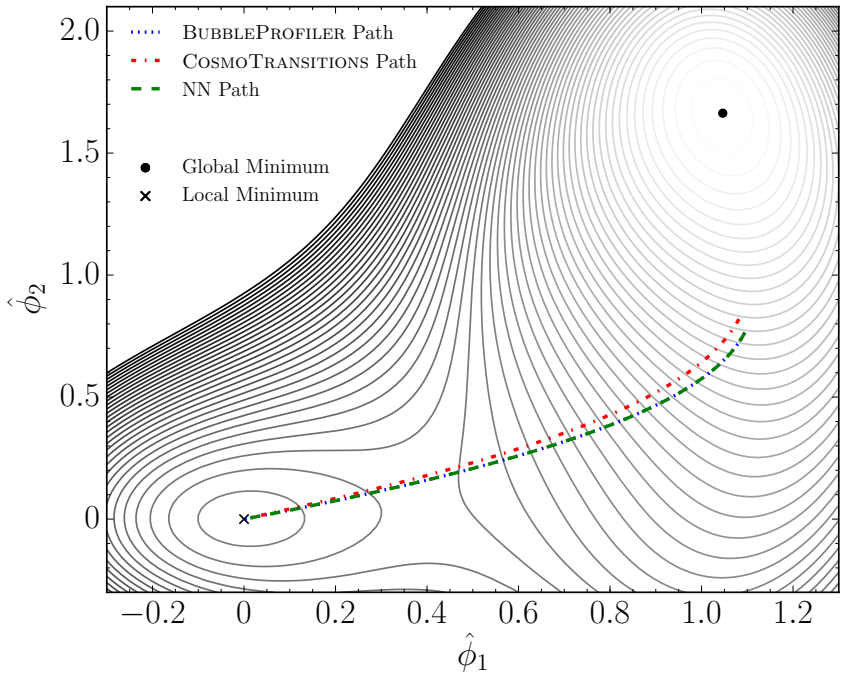


Figure 6.8: Calculated solutions for the tunnelling path obtained by our NN method, BubbleProfiler and CosmoTransitions. The paths range from the local minimum to the exit point of the tunnelling barrier. Also shown are the contours of the potential, where the global minimum is denoted by the black dot and the local minimum by the black cross.

For the thin-wall potential we find that the performance can be significantly improved if a deeper network is used. BubbleProfiler instead does not find a solution at all, while the NN agrees very well with the path found by CosmoTransitions. Since there is not a solution from all three codes, we do not show the plot here. Thus, we have shown examples where CosmoTransitions or BubbleProfiler fail to provide a correct result, while the NN approach can cope well with both the thick-wall and the thin-wall cases.

6.4.3 Scalar-Singlet-Extended Standard Model

As a final example, we study a scenario of phenomenological interest—the extension of the Standard Model Higgs sector by a single scalar field [263, 264]. Despite its simplicity, the scalar-singlet-extended Standard Model could potentially provide

solutions to puzzles such as the existence of dark matter [265, 266] and baryogenesis [267, 268], where a crucial requirement is a strong electroweak phase transition, as we discussed previously. The tree-level potential has the form,

$$V^{(0)}(h,s) = \frac{\mu^2}{2}h^2 + \frac{\lambda}{4}h^4 + \frac{\mu_s^2}{2}s^2 + \frac{\lambda_s}{4}s^4 + \frac{\lambda_m}{4}s^2h^2 \ , \tag{6.4.13}$$

where h denotes the Higgs field and s the additional $\mathbb{Z}_2$-symmetric[3] scalar field. Furthermore, we have the relation $\mu^2 = -\lambda v_{\rm EW}^2$, where $v_{\rm EW}$ is the electroweak VEV at zero temperature. It is possible to consider a scenario in which the potential barrier separating the symmetric and broken phases is generated already at tree level [269]. In this scenario, to study the evolution of the parameters with T, it is enough to include only the high-temperature expansion of the one-loop thermal potential, which results in thermal corrections to the mass parameters [260] of the form,

$$V^{(1)}(h,s,T) = \left(\frac{1}{2}c_h h^2 + \frac{1}{2}c_s s^2\right)T^2 \ , \tag{6.4.14}$$

where,

$$c_h = \frac{1}{48}\left[9g^2 + 3g'^2 + 2(6h_t^2 + 12\lambda + \lambda_m)\right] \ , \qquad c_s = \frac{1}{12}(2\lambda_m + 3\lambda_s) \ , \tag{6.4.15}$$

with g and g' being the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ gauge couplings, respectively, and h_t is the Yukawa coupling of the top quark. We then consider Eq. (6.4.3) with the potential,

$$V(h,s,T) = V^{(0)}(h,s) + V^{(1)}(h,s,T) \ . \tag{6.4.16}$$

At high temperatures, the thermal contribution in Eq. (6.4.14) dominates, and the global minimum is the electroweak- and $\mathbb{Z}_2$-symmetric configuration ($h = 0, s = 0$). The behaviour as T decreases is determined by the choice of parameters. These are constrained to the parameter region in which the potential develops a strong tree-level barrier at the critical temperature $T_{\rm C}$ [269]. In particular, at $T > T_{\rm C}$ after $\mathbb{Z}_2$-symmetry breaking, s acquires a non-zero VEV, $\langle s \rangle = w$, along the $\langle h \rangle = 0$

[3]This condition could also be relaxed, since in models with no $\mathbb{Z}_2$-symmetry the most general renormalisable potential would have three more parameters [263, 269].

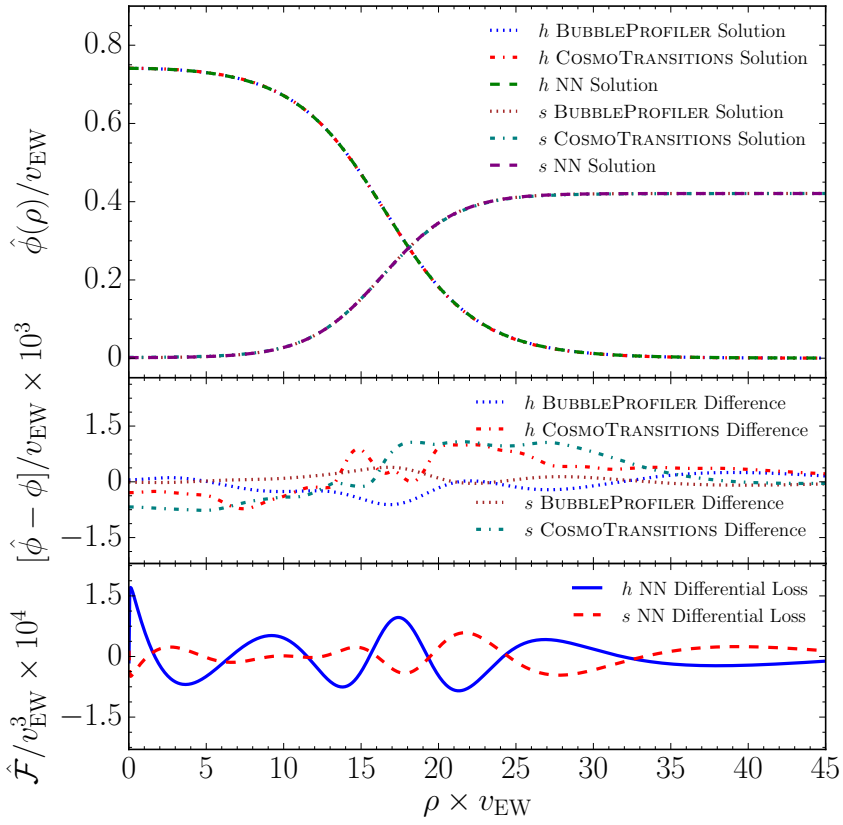


Figure 6.9: The upper panel shows the bubble profiles for the scalar-singlet-extended Standard Model potential in Eq. (6.4.16), as obtained by our NN method, BubbleProfiler and CosmoTransitions. The middle panel displays the numerical difference between the NN predicted solutions and the solutions from the other two codes. The lower panel shows the differential contribution $\hat{\mathcal{F}}$ to the loss from h and s.

direction. This configuration constitutes a global minimum for the potential. At $T = T_{\rm C}$ a second degenerate minimum appears at the symmetry-breaking VEV[4] $\langle h\rangle = v$ and at the restored $\mathbb{Z}_2$-symmetric VEV, $\langle s\rangle = 0$. Finally, at $T < T_{\rm C}$ the electroweak minimum $(v, 0)$ represents the only energetically favourable configuration. The nucleation temperature $T_{\rm N}$ at which the phase transition from $\vec{\phi}_{\rm F} = (0, w)$ to $\vec{\phi}_{\rm T} = (v, 0)$ occurs is found from the requirement $\mathcal{S}_3(T_{\rm N})/T_{\rm N} \sim 140$ [240, 270].

As an example parameter configuration, we consider $T_{\rm C} = 110$ GeV, $\lambda_m = 1.5$ and $\lambda_s = 0.65$, as used in Ref. [251], and a temperature of $T = 85$ GeV, which is the nucleation temperature that BubbleProfiler finds. We thus solve Eq. (6.4.3)

[4]Note that the VEVs w and v are themselves dependent on the temperature, with $v_{\rm EW}$ the zero-temperature limit of v.

with the boundary conditions,

$$\begin{aligned} \frac{d}{d\rho}h(\rho_{\min}) &= 0 \ , \qquad & h(\rho_{\max}) &= 0 \ , \\ \frac{d}{d\rho}s(\rho_{\min}) &= 0 \ , \qquad & s(\rho_{\max}) &= w \ . \end{aligned} \tag{6.4.17}$$

We use a neural network with 10 units in a single hidden layer with a sigmoid activation function, on a training domain of $\rho \in [0.01, 50]$ with 500 training points. To avoid large numerical values in the loss function, we scale all mass parameters in the potential by v_{EW}. Our result, along with the comparison to CosmoTransitions and BubbleProfiler, is shown in Fig. 6.9. We find very good agreement between all three methods to calculate the bubble profiles $h(\rho)$ and $s(\rho)$, and the small values of $\hat{\mathcal{F}}$ across the domain show that good convergence has been achieved.

6.5 Conclusions

In this chapter, we built upon the capabilities of neural networks for solving optimisation problems by using them to find solutions to differential equations. Our method extended upon existing approaches—where the neural network is included as part of a full trial solution—by providing the information about the boundary conditions as extra terms to the loss function, so that the output of the neural network alone is the solution to the differential equation. Furthermore, we showed that the differential contribution $\hat{\mathcal{F}}$ to the loss can be used as a test for the performance of the method. The calculation of this quantity is possible because the neural network simply consists of a series of matrix multiplications with activation functions applied, and so the solution is fully analytic and thus is differentiable.

We applied this approach to finding differentiable solutions to ordinary, coupled and partial differential equations, for which analytic solutions are known. Various network architectures have been studied, and even relatively small networks showed a very good performance. We further proposed that in regions of numerical stability, it

can be useful to iteratively extend the domain to find stable solutions over arbitrarily large domains. We also found that for solutions that vary quickly over a small part of the domain, it can be numerically beneficial to self-adaptively distribute more training points in such regions.

To show the applicability of this method to a task of direct phenomenological interest, we used it to calculate the tunnelling profiles arising in cosmological phase transitions and compared them to those obtained by CosmoTransitions and BubbleProfiler. We have presented explicit examples where the neural network finds correct solutions, while either CosmoTransitions or BubbleProfiler fail. We find an optimised neural network to be very flexible and reliable, and is able to converge to solutions for all the examples tested with an accuracy that is competitive with the dedicated programs for calculating the profiles. However, further work—such as in developing an approach to choosing the domain sizes for phase transitions in a more robust way—would be required to develop a fully automated tool using this approach.

Another idea which is made possible by our new approach to using neural networks for solving differential equations is that the boundary conditions could themselves be included as additional inputs to the network. The term in the loss function that involves the boundary conditions would then obtain its values from these inputs. This would allow a differential equation to be solved for multiple values of the boundary conditions simultaneously, with the solution for fixed values of the boundary conditions recovered by taking a slice along the full solution. We leave the exploration of this idea to future work. Furthermore, as the entire method could be extended beyond the calculation of differential equations, we envisage it to be applicable to a wide range of problems in perturbative and non-perturbative quantum field theory.

Chapter 7

Summary and Outlook

In this thesis, we have performed several investigations into physics beyond the Standard Model, with a general focus on improving the precision of searches for this new physics. Since the Standard Model has been so successful and explains many observations very well, it has proven to be extremely difficult to know exactly where to look for any new physics. However, with multiple unexplained phenomena, clearly the Standard Model is not the complete answer and there has to be something else that lies beyond it. There are a multitude of techniques, both theoretical and experimental, that can be applied to help with this search, and we explored some of them in this thesis.

In Chapter 2 we derived new limits on the parameter spaces of specific extensions of the Standard Model by considering lepton-flavour-violating decays of τ leptons, which are currently weakly constrained, but which can be constrained further at future colliders to obtain more precise limits. Then in Chapter 3 we used electroweak precision measurements to calculate new bounds on the trilinear self-coupling of the Higgs boson, which we found to be comparable and complementary to the existing limits. In Chapter 4 we improved the precision of the calculation of the cross section for the production of a heavy neutrino via gluon fusion by resumming threshold logarithms, and found that its rate is increased by approximately a factor of

three compared with the leading-order calculation. We considered machine-learning techniques in Chapter 5, where we improved the robustness of an autoencoder used for unsupervised searches for new physics by combining it with an adversarial neural network. Finally, in Chapter 6 we developed a new approach to using neural networks for solving differential equations, and we showed it could be applied to the case of cosmological phase transitions with a performance that is competitive with dedicated methods.

There is still a long way to go in the search for an absolute theory of nature that could explain everything we observe in the Universe, and this thesis makes a small contribution along the grand path towards doing so. With rapidly improving machine-learning techniques and advances in computational power, it may be that the best chance physicists have of making the next significant step is from applying these to the vast quantities of data that colliders such as the LHC produce. The adversarial autoencoder we developed in Chapter 5, which can perform model-independent searches for new physics without being sensitive to the experimental uncertainties, is an example of such an approach. If these techniques can reveal even a small hint of where the new physics lies, the advanced theoretical and experimental analysis methods built up over many decades could then be deployed to hone in on and find the true theory of physics beyond the Standard Model.

Appendix A

One-Loop Calculation of Gauge Boson Vacuum Polarisations

In this appendix we describe the calculation for the insertion of a fermion loop into the electroweak gauge boson propagators. The Feynman diagram for this process is shown in Fig. A.1.

The fermion loop inside the diagram in Fig. A.1 is the part we calculate, so we will remove the external gauge boson polarisations. This amplitude is then denoted by $i\Pi^{\mu\nu}_{AB}(p)$. It is convenient to break down the calculation by considering the fermion currents to be left- or right-handed. Since the electroweak gauge bosons couple to left- or right-handed fermions differently, this makes it easier to build up the

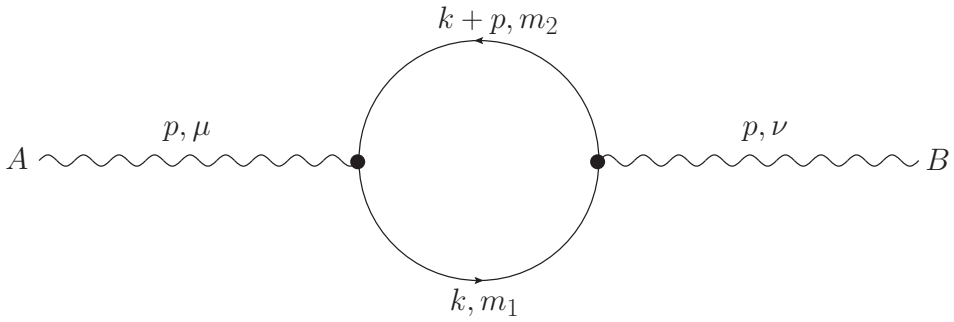


Figure A.1: Feynman diagram for the insertion of a fermion loop into the gauge boson propagators.

full vacuum polarisation functions. The relation between the LL, RR, LR and RL currents is,

$$i\Pi^{\mu\nu}_{LL}(p) = i\Pi^{\mu\nu}_{RR}(p) \ , \qquad i\Pi^{\mu\nu}_{LR}(p) = i\Pi^{\mu\nu}_{RL}(p) \ , \tag{A.1}$$

which means that there are only two independent loop diagrams to calculate. The first one is the LL contribution, and using the Feynman rules, this is given by,

$$i\Pi^{\mu\nu}_{LL}(p) = -e^2\mu^{4-d}\int \frac{d^dk}{(2\pi)^d}\mathrm{Tr}\left[(i\gamma^\mu)P_L\frac{i(\not{k}+m_1)}{k^2-m_1^2}(i\gamma^\nu)P_L\frac{i(\not{k}+\not{p}+m_2)}{(k+p)^2-m_2^2}\right] \ . \tag{A.2}$$

We can combine the denominators using the Feynman parameterisation,

$$\frac{1}{[k^2-m_1^2][(k+p)^2-m_2^2]} = \int_0^1 dx \frac{1}{(l^2-\Delta)^2} \ , \tag{A.3}$$

where,

$$l = k + xp \ , \qquad \Delta = xm_2^2 + (1-x)m_1^2 - x(1-x)p^2 \ . \tag{A.4}$$

Performing a Wick rotation and using the results for the standard scalar integrals in dimensional regularisation, we get,

$$i\Pi^{\mu\nu}_{LL}(p) = \frac{ie^2\mu^{4-d}}{(4\pi)^{d/2}}\int_0^1 dx \frac{\Gamma(2-\frac{d}{2})}{\Delta^{2-d/2}}\Big[\Big(2xm_2^2 + 2(1-x)m_1^2 - 4x(1-x)p^2\Big)g^{\mu\nu} + x(1-x)p^\mu p^\nu\Big] \ . \tag{A.5}$$

The term proportional to $p^\mu p^\nu$ is dropped because it vanishes when it is contracted with light external fermion currents. Writing the number of dimensions as $d \equiv 4-2\epsilon$, where $\epsilon \ll 1$, we can use the expansion of the gamma function to make the replacement,

$$\frac{\mu^{4-d}}{(4\pi)^{d/2}}\frac{\Gamma(2-\frac{d}{2})}{\Delta^{2-d/2}} \to \frac{1}{(4\pi)^2}\left[\frac{1}{\epsilon} - \gamma_\mathrm{E} + \ln 4\pi + \ln\frac{\mu^2}{\Delta}\right] \ , \tag{A.6}$$

where γ_E is Euler's constant. After stripping off the $ig^{\mu\nu}$ factor, the vacuum polarisation function then becomes,

$$\Pi_{LL}(p^2) = \frac{e^2}{(4\pi)^2}\int_0^1 dx \left[\frac{1}{\epsilon} - \gamma_\mathrm{E} + \ln 4\pi + \ln\frac{\mu^2}{\Delta}\right]\Big[2xm_2^2 + 2(1-x)m_1^2 - 4x(1-x)p^2\Big] \ . \tag{A.7}$$

If we separate the divergent part from the finite part, we get,

$$\Pi_{LL}(p^2) = \frac{e^2}{(4\pi)^2}\left[\frac{m_1^2 + m_2^2 - \frac{2}{3}p^2}{\epsilon} + \int_0^1 dx[2\Delta - 2x(1-x)p^2]\ln\frac{4\pi e^{-\gamma_{\rm E}}\mu^2}{\Delta}\right] . \tag{A.8}$$

For $i\Pi_{LR}^{\mu\nu}(p)$ the calculation is the same except that one of the P_L operators is replaced by P_R. The result is,

$$\Pi_{LR}(p^2) = -\frac{e^2}{(4\pi)^2}m_1 m_2\left[\frac{2}{\epsilon} + 2\int_0^1 dx\ln\frac{4\pi e^{-\gamma_{\rm E}}\mu^2}{\Delta}\right] . \tag{A.9}$$

For two vector currents, the result is the sum of all four combinations of the left- and right-handed currents,

$$\Pi_{VV}(p^2) = \Pi_{LL}(p^2) + \Pi_{RR}(p^2) + \Pi_{LR}(p^2) + \Pi_{RL}(p^2) . \tag{A.10}$$

Using Eq. (A.1) and plugging in the results from from Eqs. (A.8) and (A.9), we obtain,

$$\Pi_{VV}(p^2) = \frac{e^2}{4\pi^2}\left[\frac{m_1^2 + m_2^2 - 2m_1m_2 - \frac{2}{3}p^2}{2\epsilon} + \int_0^1 dx[xm_2^2 + (1-x)m_1^2 - m_1m_2 - 2x(1-x)p^2]\ln\frac{4\pi e^{-\gamma_{\rm E}}\mu^2}{\Delta}\right] . \tag{A.11}$$

We now have everything we need to assemble the analytic expressions of the vacuum polarisation functions of all the electroweak gauge bosons. They can be written entirely in terms of VV and LL currents, and for the insertion of a single doublet of top and bottom quarks they are [118],

$$\Pi_{\gamma\gamma}(p^2) = N_c\sum_{i=t,b}(Q^i)^2\Pi_{VV}(\Delta_{ii}) , \tag{A.12}$$

$$\Pi_{Z\gamma}(p^2) = \frac{1}{sc}N_c\sum_{i=t,b}\left[T_3^iQ^i\frac{1}{2}\Pi_{VV}(\Delta_{ii}) - s^2(Q^i)^2\Pi_{VV}(\Delta_{ii})\right] , \tag{A.13}$$

$$\Pi_{WW}(p^2) = |V_{tb}|^2\,\frac{1}{s^2}N_c\frac{1}{2}\Pi_{LL}(\Delta_{tb}) , \tag{A.14}$$

$$\Pi_{ZZ}(p^2) = \frac{1}{s^2c^2}N_c\sum_{i=t,b}\left[(T_3^i)^2\Pi_{LL}(\Delta_{ii}) - 2s^2(T_3^iQ^i\frac{1}{2}\Pi_{VV}(\Delta_{ii})\right.$$

$$+ s^4 (Q^i)^2 \Pi_{VV}(\Delta_{ii}) \Big] \,, \tag{A.15}$$

where Δ_{ij} represents Δ with $m_1 = m_i$ and $m_2 = m_j$.

Appendix B

Analytic Expressions for $\bar{c}_6$ Contributions to S and T

In this appendix, we present the analytic results for the $\bar{c}_6$ contributions to S and T in terms of the finite parts of the basis integrals, which then require numerical evaluation. The notation for the basis integrals follows that used by TSIL [135]. For the self-energy integrals B, S, T, U, and M, the first argument is the square of the external four-momentum. The results are:

$$
\begin{aligned}
S = \frac{\alpha_e \bar{c}_6}{1024\pi^2 s^2 m_W^2 m_Z^4 (m_h^2 - 4m_Z^2)(m_h^2 - m_Z^2)^2} \Big\{ \\
& 36(2+\bar{c}_6) m_h^2 (m_h^2 - m_Z^2) B(m_h^2, m_h^2, m_h^2) \Big(- m_Z^2 (m_h^6 - 3 m_h^4 m_Z^2 \\
& \quad + 4 m_h^2 m_Z^4 + 16 m_Z^6) + 2(m_h^2 - 2m_Z^2)^3 (m_h^2 - m_Z^2) B(m_Z^2, m_h^2, m_Z^2) \Big) \\
& + 8 m_h^2 A(m_Z^2) \Big(- 4(m_h^2 - 4m_Z^2)(m_h^2 - 2m_Z^2)(m_h^2 - m_Z^2)^2 B(m_Z^2, m_h^2, m_Z^2) \\
& \quad - (m_h^2 - m_Z^2)[(10 + 3\bar{c}_6) m_h^6 - 3(18 + 5\bar{c}_6) m_h^4 m_Z^2 \\
& \quad + 48(3 + \bar{c}_6) m_h^2 m_Z^4 - 4(34 + 9\bar{c}_6) m_Z^6] \\
& \quad - 9(2 + \bar{c}_6)(m_h^8 - 6 m_h^6 m_Z^2 + 14 m_h^4 m_Z^4 - 8 m_h^2 m_Z^6 + 8 m_Z^8) B(m_h^2, m_h^2, m_h^2) \Big) \\
& + 8 A(m_h^2) \Big(- 2 m_h^2 (2 m_h^6 - 9 m_h^4 m_Z^2 + 16 m_Z^6) A(m_Z^2) \\
& \quad - m_h^2 (m_h^2 - m_Z^2)[(14 + 3\bar{c}_6) m_h^6 - 6(10 + \bar{c}_6) m_h^4 m_Z^2 \\
& \quad + 12(7 + \bar{c}_6) m_h^2 m_Z^4 + 8(20 + 9\bar{c}_6) m_Z^6]
\end{aligned}
$$

$$
\begin{aligned}
&+ 2(m_h^2 - 4m_Z^2)(m_h^2 - m_Z^2)^2[(4+\bar{c}_6)m_h^4 - 4(3+\bar{c}_6)m_h^2 m_Z^2 \\
&\quad + 12(2+\bar{c}_6)m_Z^4]B(m_Z^2, m_h^2, m_Z^2) + 9(2+\bar{c}_6)m_h^2(m_h^8 - 7m_h^6 m_Z^2 \\
&+ 19m_h^4 m_Z^4 - 24m_h^2 m_Z^6 + 20m_Z^8)B(m_h^2, m_h^2, m_h^2)\Big) \\
&+ m_h^2\Big(- 8[(12+7\bar{c}_6)m_h^8 - 9(9+5\bar{c}_6)m_h^6 m_Z^2 + 99(2+\bar{c}_6)m_h^4 m_Z^4 \\
&\quad - 8(15+8\bar{c}_6)m_h^2 m_Z^6 + 12(6+7\bar{c}_6)m_Z^8]I(m_h^2, m_h^2, m_h^2) \\
&+ (m_h^2 - 4m_Z^2)\Big[8(4m_h^4 - 5m_h^2 m_Z^2 - 2m_Z^4)A(m_Z^2)^2 \\
&\quad + 24m_Z^2\Big((m_h^2 - 2m_Z^2)A(m_h^2)^2 + m_Z^2[20m_Z^2 - (20+9\bar{c}_6)m_h^2]I(m_h^2, m_h^2, m_Z^2)\Big) \\
&\quad - 8(m_h^2 + 2m_Z^2)(2m_h^4 - 9m_h^2 m_Z^2 + 16m_Z^4)I(m_h^2, m_Z^2, m_Z^2)\Big] \\
&+ (m_h^2 - m_Z^2)\Big\{128m_h^8 + 32\bar{c}_6 m_h^8 - 554m_h^6 m_Z^2 - 99\bar{c}_6 m_h^6 m_Z^2 + 986m_h^4 m_Z^4 \\
&\quad + 279\bar{c}_6 m_h^4 m_Z^4 + 432m_h^2 m_Z^6 + 268\bar{c}_6 m_h^2 m_Z^6 + 304m_Z^8 + 168\bar{c}_6 m_Z^8 \\
&\quad + 8(m_h^2 - m_Z^2)\Big[\\
&\quad [(22+9\bar{c}_6)m_h^4 - 12(8+3\bar{c}_6)m_h^2 m_Z^2 + 8(22+9\bar{c}_6)m_Z^4]S(m_Z^2, m_h^2, m_h^2, m_Z^2) \\
&\quad + 8(m_h^2 - m_Z^2)[(4+\bar{c}_6)(m_h^4 - 4m_h^2 m_Z^2) + 12(2+\bar{c}_6)m_Z^4]T(m_Z^2, m_h^2, m_h^2, m_Z^2) \\
&\quad + (m_h^2 - 4m_Z^2)\Big(2(m_h^6 - 12m_h^2 m_Z^4 + 24m_Z^6)M(m_Z^2, m_h^2, m_h^2, m_Z^2, m_Z^2, m_h^2) \\
&\quad\quad + [(2+\bar{c}_6)m_h^4 - 4(2+\bar{c}_6)m_h^2 m_Z^2 + 4(10+3\bar{c}_6)m_Z^4]B(m_Z^2, m_h^2, m_Z^2) \\
&\quad\quad - 4m_Z^2(m_h^2 - 2m_Z^2)[B(m_Z^2, m_h^2, m_Z^2)^2 + 2U(m_Z^2, m_h^2, m_Z^2, m_h^2, m_Z^2)]\Big) \\
&\quad - [(10+7\bar{c}_6)m_h^6 - 2(32+19\bar{c}_6)m_h^4 m_Z^2 + 4(36+13\bar{c}_6)m_h^2 m_Z^4 \\
&\quad\quad + 24(\bar{c}_6 - 2)m_Z^6]U(m_Z^2, m_Z^2, m_h^2, m_h^2, m_h^2)\Big]\Big\}\Big)\Big\}\,, \qquad \text{(B.1)}
\end{aligned}
$$

$$
\begin{aligned}
T = \frac{3\alpha_e \bar{c}_6 m_h^2}{512\pi^2 s^4 m_W^4 (m_h^2 - m_W^2)^2 (m_h^2 - m_Z^2)^2}\Big\{ \\
&A(m_h^2)\Big((22+9\bar{c}_6)m_h^2(m_h^2 - m_W^2)m_Z^2(m_h^2 - m_Z^2)s^2 \\
&\quad + 2(m_h^2 - 2m_W^2)(m_h^2 - m_Z^2)^2 A(m_W^2) - 2(m_h^2 - m_W^2)^2(m_h^2 - 2m_Z^2)A(m_Z^2) \\
&\quad - 9(2+\bar{c}_6)m_h^2 m_Z^2[m_h^2(m_W^2 + m_Z^2) - 2m_W^2 m_Z^2]s^2 B(m_h^2, m_h^2, m_h^2)\Big) \\
&- m_Z^2 s^2\Big([m_h^2(m_W^2 + m_Z^2) - 2m_W^2 m_Z^2]A(m_h^2)^2 \\
&\quad - 3m_h^2[2(2+\bar{c}_6)m_h^4 - 2(1+2\bar{c}_6)m_W^2 m_Z^2 + (\bar{c}_6 - 1)m_h^2(m_W^2 + m_Z^2)]I(m_h^2, m_h^2, m_h^2) \\
&\quad + (m_h^2 - m_W^2)(m_h^2 - m_Z^2)[(20+9\bar{c}_6)m_h^4 + 2m_W^2 m_Z^2
\end{aligned}
$$

$$
\begin{aligned}
&\quad - 2m_h^2(m_W^2+m_Z^2) - 9(2+\bar{c}_6)m_h^4 B(m_h^2,m_h^2,m_h^2)]\Big) \\
&+ (m_h^2-m_Z^2)^2\Big(- (m_h^2-2m_W^2)A(m_W^2)^2 \\
&\quad + m_W^2\Big[A(m_W^2)[4(m_h^2-m_W^2) - 9(2+\bar{c}_6)m_h^2 B(m_h^2,m_h^2,m_h^2)] \\
&\quad\quad + [(20+9\bar{c}_6)m_h^2 - 20m_W^2]I(m_h^2,m_h^2,m_W^2)\Big] \\
&\quad + (m_h^4 - 4m_h^2m_W^2 + 12m_W^4)I(m_h^2,m_W^2,m_W^2)\Big) \\
&+ (m_h^2-m_W^2)^2\Big((m_h^2-2m_Z^2)A(m_Z^2)^2 - (m_h^4 - 4m_h^2m_Z^2 + 12m_Z^4)I(m_h^2,m_Z^2,m_Z^2) \\
&\quad + m_Z^2\Big[A(m_Z^2)[\,-4m_h^2 + 4m_Z^2 + 9(2+\bar{c}_6)m_h^2 B(m_h^2,m_h^2,m_h^2)] \\
&\quad\quad - [(20+9\bar{c}_6)m_h^2 - 20m_Z^2]I(m_h^2,m_h^2,m_Z^2)\Big]\Big)\Big\}\,. \qquad \text{(B.2)}
\end{aligned}
$$

Appendix C

Analytic Continuation in SCET

In this appendix, we discuss the source of the subtraction terms required to analytically continue the integral over the soft-coefficient function $\Delta_{gg}^{\beta\,\mathrm{res}}(z)$ to negative values of $\eta \equiv 2a_\Gamma(\mu_s^2, \mu_f^2)$, which is needed for numerical evaluation. The soft coefficient function is given in Eq. (4.3.4), while the complete integral is given in Eq. (4.3.1) with the FO coefficient function replaced by the resummed version.

The integral—which involves a product of PDFs, resummed soft-coefficient function, and partonic cross sections—converges when integrated over z from τ/ξ to 1 even though the integrand itself diverges in the threshold region $z \to 1$. However, this convergence requires that $\eta > 0$, as we shall shortly show. The soft-scalar function $\tilde{s}_{\mathrm{Higgs}}$ involves powers of the operator $\ln[Q^2(1-z)^2/(\mu_s^2 z)]+\partial_\eta$ which act on a function of η. We will first consider the terms in the soft-coefficient function which involve the zeroth power of this operator. Up to NLL accuracy, the soft-scalar function is only required at tree level and is just the identity and so there are no derivatives in η at this order. Here, the integral over z of the PDFs, soft-coefficient function and partonic cross section can be written as,

$$\int_{\tau/\xi}^1 dz f_{g/p}(\tau/(z\xi), \mu_f)\hat{\sigma}(z)\frac{z^{-\eta-\frac{1}{2}}}{(1-z)^{1-2\eta}}\frac{e^{-2\gamma_{\mathrm{E}}\eta}}{\Gamma(2\eta)} \equiv \int_{\tau/\xi}^1 dz \frac{g(z)}{(1-z)^{1-2\eta}}\,, \tag{C.1}$$

where only the terms involving z or η have been included. The function $g(z)$ is finite as $z \to 1$. For simplicity, consider the change of variables $z = 1-\omega$, and write the

integral from 0 to 1 rather than τ/ξ to 1. This does not affect our argument, but the non-zero lower limit can be included again with an additional term. The integral that we wish to study is therefore,

$$I = \int_0^1 d\omega \frac{g(\omega)}{\omega^{1-2\eta}} \ . \tag{C.2}$$

The divergence in the integrand now occurs in the limit $\omega \to 0$. The integral itself is divergent when $\eta < 0$, which can be seen from setting $g(\omega) = 1$ and simply carrying out the integration.

For $\eta > 0$, we can write the integral as,

$$\begin{aligned} I &= \int_0^1 d\omega \frac{g(\omega) - g(0) + g(0)}{\omega^{1-2\eta}} \\ &= \int_0^1 d\omega \frac{g(\omega) - g(0)}{\omega^{1-2\eta}} + \int_0^1 d\omega \frac{g(0)}{\omega^{1-2\eta}} \ . \end{aligned} \tag{C.3}$$

We can evaluate the second integral, using the assumption that $\eta > 0$, to obtain,

$$I = \int_0^1 d\omega \frac{g(\omega) - g(0)}{\omega^{1-2\eta}} + \frac{g(0)}{2\eta} \ . \tag{C.4}$$

This is an equivalent expression to the integral that we started with, which was valid for $\eta > 0$. We now wish to extend the range of validity—or analytically continue the integral—down into negative values of η. To do so, we must determine for which values of η the new expression for I in Eq. (C.4) is valid. The second term is clearly valid for any value of η. It may seem that $\eta = 0$ would cause a problem, but in our case we have $g(0) \sim 1/\Gamma(2\eta) = 2\eta/\Gamma(1+2\eta)$ so the factors of 2η cancel and therefore the second term is valid $\forall\eta$. For the first term involving the integral, which we will denote as I_1, we can Taylor expand $g(\omega)$ around 0 to get,

$$\begin{aligned} I_1 &= \int_0^1 d\omega \frac{(g(0) + g'(0)\omega + \frac{g''(0)}{2!}\omega^2 + \ldots) - g(0)}{\omega^{1-2\eta}} \\ &= \int_0^1 d\omega \frac{g'(0)\omega + \frac{g''(0)}{2!}\omega^2 + \ldots}{\omega^{1-2\eta}} \\ &= \int_0^1 d\omega \left(g'(0)\omega^{2\eta} + \frac{g''(0)}{2!}\omega^{1+2\eta} + \ldots \right) \ . \end{aligned} \tag{C.5}$$

The first term in this integral now converges for $\eta > -0.5$, with the other terms having

larger validity ranges, and therefore the entire expression in Eq. (C.4) converges for $\eta > -0.5$. By including the addition and subtraction terms, we have been able to extend the range of validity of the integral.

Now, we would like to extend the range of validity even further. To do this, we must address the term that caused the expression to not be valid below $\eta = -0.5$—this was the term that was linear in ω in the Taylor expansion. We can apply the same method that we used previously to add and subtract this term from the integral, to get,

$$\begin{aligned} I &= \int_0^1 d\omega \frac{g(\omega) - g(0) - g'(0)\omega + g'(0)\omega}{\omega^{1-2\eta}} + \frac{g(0)}{2\eta} \\ &= \int_0^1 d\omega \frac{g(\omega) - g(0) - g'(0)\omega}{\omega^{1-2\eta}} + \int_0^1 d\omega \frac{g'(0)\omega}{\omega^{1-2\eta}} + \frac{g(0)}{2\eta} \\ &= \int_0^1 d\omega \frac{g(\omega) - g(0) - g'(0)\omega}{\omega^{1-2\eta}} + \int_0^1 d\omega g'(0)\omega^{2\eta} + \frac{g(0)}{2\eta} \\ &= \int_0^1 d\omega \frac{g(\omega) - g(0) - g'(0)\omega}{\omega^{1-2\eta}} + \frac{g'(0)}{1+2\eta} + \frac{g(0)}{2\eta} \ . \end{aligned} \tag{C.6}$$

We can ask the same question again: for what range of η is this expression valid? The last two terms are valid $\forall\eta$. The only issue would be at $\eta = -0.5$, but we can similarly regulate this by pulling an extra factor of $1 + 2\eta$ out of the Γ-function. For the integral term, which we denote as I_1 again, we can Taylor expand to get,

$$\begin{aligned} I_1 &= \int_0^1 d\omega \frac{(g(0) + g'(0)\omega + \frac{g''(0)}{2!}\omega^2 + \ldots) - g(0) - g'(0)\omega}{\omega^{1-2\eta}} \\ &= \int_0^1 d\omega \frac{\frac{g''(0)}{2!}\omega^2 + \ldots}{\omega^{1-2\eta}} \\ &= \int_0^1 d\omega \frac{g''(0)}{2!}\omega^{1+2\eta} + \ldots \ . \end{aligned} \tag{C.7}$$

This integral converges for $\eta > -1$, which means that the expression for I in Eq. (C.6) is also valid for $\eta > -1$. The process can then be repeated in the same manner: to extend the expression further into negative values of η by incremental amounts of 0.5, we must add and subtract in each step an additional term involving an extra power of ω from the Taylor expansion.

We will now convert back to the variable $z = 1 - \omega$ with integration limits from τ/ξ

to 1. We can write the integral in Eq. (C.1) as,

$$\begin{aligned}\int_{\frac{\tau}{\xi}}^1 dz \frac{g(z)}{(1-z)^{1-2\eta}} &= \int_{\frac{\tau}{\xi}}^1 dz \frac{g(z)-g(1)}{(1-z)^{1-2\eta}} + \frac{g(1)}{2\eta} - g(1)\int_0^{\frac{\tau}{\xi}} dz \frac{1}{(1-z)^{1-2\eta}} \\ &= \int_{\frac{\tau}{\xi}}^1 dz \frac{g(z)-g(1)}{(1-z)^{1-2\eta}} + \frac{g(1)}{2\eta} + g(1)\left(\frac{(1-\frac{\tau}{\xi})^{2\eta}}{2\eta} - \frac{1}{2\eta}\right) \\ &= \int_{\frac{\tau}{\xi}}^1 dz \frac{g(z)-g(1)}{(1-z)^{1-2\eta}} + \left(1-\frac{\tau}{\xi}\right)\frac{g(1)}{2\eta}\frac{1}{(1-\frac{\tau}{\xi})^{1-2\eta}} \,, \end{aligned} \tag{C.8}$$

which is valid for $\eta > -0.5$. These are the entire set of addition and subtraction terms required to analytically continue the NLL result down to $\eta = -0.5$.

Beyond NLL, the soft-scalar function $\tilde{s}_{\text{Higgs}}$ contains higher powers of the operator $\ln[Q^2(1-z)^2/(\mu_s^2 z)]+\partial_\eta$. In the region $\eta > 0$, these terms can simply be computed by calculating the action of multiple derivatives of ∂_η on the function $e^{-2\gamma_{\text{E}}\eta}/\Gamma(2\eta)$. The only complication arises when applying the aforementioned analytic continuation procedure to extend into negative values of η. The terms involving the action of the derivative ∂_η, and the logarithmic term $\ln[Q^2/(\mu_s^2 z)]$, can be absorbed into the function $g(z)$ since they are finite in the limit $z \to 1$. However, the logarithmic term $\ln(1-z)$ diverges in this limit and so we must consider integrals of the form,

$$\int_{\tau/\xi}^1 dz \frac{g(z)\ln^n(1-z)}{(1-z)^{1-2\eta}} \,. \tag{C.9}$$

Analogously to the earlier procedure, and in terms of an integral over ω from 0 to 1 for simplicity of notation, we can analytically continue the integral down to $\eta = -0.5$ using the result,

$$\int_0^1 d\omega \frac{g(\omega)\ln^n\omega}{\omega^{1-2\eta}} = \left[\int_0^1 d\omega \frac{(g(\omega)-g(0))\ln^n\omega}{\omega^{1-2\eta}}\right] + (-1)^n n! \frac{g(0)}{(2\eta)^{n+1}} \,. \tag{C.10}$$

Transforming this expression into an integral over z from τ/ξ to 1 then gives the exact form of the addition and subtraction terms required to analytically continue the full N^{k}LL result so that it is valid in the region $\eta > -0.5$. One can in principle extend this further by considering the additional terms in the Taylor expansion of $g(z)$, although this is difficult in practice, especially since the derivatives of $g(z)$ involve derivatives of the PDFs.

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579.

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

[3] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, *Phys. Rev. Lett.* **31** (1973) 494.

[4] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories - I*, *Phys. Rev.* **D8** (1973) 3633.

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

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

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

[8] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191.

[9] S. Tomonaga, *On a relativistically invariant formulation of the quantum theory of wave fields*, *Prog. Theor. Phys.* **1** (1946) 27.

[10] R. P. Feynman, *Space-time approach to nonrelativistic quantum mechanics*, *Rev. Mod. Phys.* **20** (1948) 367.

[11] J. S. Schwinger, *Quantum electrodynamics. I. A covariant formulation*, *Phys. Rev.* **74** (1948) 1439.

[12] S. F. Novaes, *Standard model: An Introduction*, in *Particles and fields. Proceedings, 10th Jorge Andre Swieca Summer School, Sao Paulo, Brazil, February 6-12, 1999*, pp. 5–102, 1999, hep-ph/0001283.

[13] K. Nishijima, *Charge Independence Theory of V Particles*, *Prog. Theor. Phys.* **13** (1955) 285.

[14] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, *Nuovo Cim.* **4** (1956) 848.

[15] Particle Data Group collaboration, *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 090001.

[16] ATLAS, CMS collaboration, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803 [1503.07589].

[17] T. van Ritbergen and R. G. Stuart, *Complete two loop quantum electrodynamic contributions to the muon lifetime in the Fermi model*, *Phys. Rev. Lett.* **82** (1999) 488 [hep-ph/9808283].

[18] S. Willenbrock, *Symmetries of the standard model*, in *Physics in $D \geq 4$. Proceedings, Theoretical Advanced Study Institute in elementary particle physics, TASI 2004, Boulder, USA, June 6-July 2, 2004*, pp. 3–38, 2004, hep-ph/0410370.

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

[20] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, *Phys. Rev. Lett.* **13** (1964) 138.

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

[22] L. Durand and K. Riesselmann, *The Goldstone boson equivalence theorem with fermions*, *Phys. Rev. D* **55** (1997) 1533 [hep-ph/9512224].

[23] L. D. Faddeev and V. N. Popov, *Feynman Diagrams for the Yang-Mills Field*, *Phys. Lett.* **B25** (1967) 29.

[24] A. Denner and S. Dittmaier, *Reduction schemes for one-loop tensor integrals*, *Nucl. Phys.* **B734** (2006) 62 [hep-ph/0509141].

[25] G. Passarino and M. J. G. Veltman, *One Loop Corrections for e^+e^- Annihilation Into $\mu^+\mu^-$ in the Weinberg Model*, *Nucl. Phys.* **B160** (1979) 151.

[26] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650.

[27] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (1964) B1549.

[28] A. Denner, *Techniques for calculation of electroweak radiative corrections at the one loop level and results for W physics at LEP-200*, *Fortsch. Phys.* **41** (1993) 307 [0709.1075].

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

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

[31] T. van Ritbergen, J. A. M. Vermaseren and S. A. Larin, *The Four loop beta function in quantum chromodynamics*, *Phys. Lett.* **B400** (1997) 379 [hep-ph/9701390].

[32] Super-Kamiokande collaboration, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562 [hep-ex/9807003].

[33] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870.

[34] J. Sayre and S. Wiesenfeldt, *Naturalness and the neutrino matrix*, *Phys. Rev.* **D77** (2008) 053005 [0711.1687].

[35] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, *Nuovo Cim.* **14** (1937) 171.

[36] P. Minkowski, $\mu \to e\gamma$ *at a Rate of One Out of* 10^9 *Muon Decays?*, *Phys. Lett.* **B67** (1977) 421.

[37] F. Borzumati and Y. Nomura, *Low scale seesaw mechanisms for light neutrinos*, *Phys. Rev.* **D64** (2001) 053005 [hep-ph/0007018].

[38] G. Burdman, *New solutions to the hierarchy problem*, *Braz. J. Phys.* **37** (2007) 506 [hep-ph/0703194].

[39] G. Amelino-Camelia, *Quantum theory's last challenge*, *Nature* **408** (2000) 661 [gr-qc/0012049].

[40] E. Corbelli and P. Salucci, *The Extended Rotation Curve and the Dark Matter Halo of M33*, *Mon. Not. Roy. Astron. Soc.* **311** (2000) 441 [astro-ph/9909252].

[41] M. Dine and A. Kusenko, *The Origin of the matter - antimatter asymmetry*, *Rev. Mod. Phys.* **76** (2003) 1 [hep-ph/0303065].

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

[43] F. L. Wilson, *Fermi's Theory of Beta Decay*, *Am. J. Phys.* **36** (1968) 1150.

[44] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[45] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1.

[46] R. Frühwirth and R. K. Bock, *Data analysis techniques for high-energy physics experiments*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **11** (2000) 1.

[47] G. James, D. Witten, T. Hastie and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014.

[48] Y. LeCun, Y. Bengio and G. Hinton, *Deep learning*, *Nature* **521** (2015) 436.

[49] D. Guest, K. Cranmer and D. Whiteson, *Deep Learning and its Application to LHC Physics*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161 [1806.11484].

[50] P. T. Komiske, E. M. Metodiev and M. D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination*, *JHEP* **01** (2017) 110 [1612.01551].

[51] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images—deep learning edition*, *JHEP* **07** (2016) 069 [1511.05190].

[52] P. Baldi, P. Sadowski and D. Whiteson, *Searching for Exotic Particles in High-Energy Physics with Deep Learning*, *Nature Commun.* **5** (2014) 4308 [1402.4735].

[53] P. T. Komiske, E. M. Metodiev, B. Nachman and M. D. Schwartz, *Pileup Mitigation with Machine Learning (PUMML)*, *JHEP* **12** (2017) 051 [1707.08600].

[54] J. H. Collins, K. Howe and B. Nachman, *Anomaly Detection for Resonant New Physics with Machine Learning*, *Phys. Rev. Lett.* **121** (2018) 241803 [1805.02664].

[55] J. Hajer, Y.-Y. Li, T. Liu and H. Wang, *Novelty Detection Meets Collider Physics*, 1807.10261.

[56] T. S. Roy and A. H. Vijay, *A robust anomaly finder based on autoencoder*, 1903.02032.

[57] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, *Improving neural networks by preventing co-adaptation of feature detectors*, 1207.0580.

[58] D. E. Rumelhart, G. E. Hinton and R. J. Williams, *Learning representations by back-propagating errors*, *Nature* **323** (1986) 533.

[59] J. Cogan, M. Kagan, E. Strauss and A. Schwarztman, *Jet-Images: Computer Vision Inspired Techniques for Jet Tagging*, *JHEP* **02** (2015) 118 [1407.5675].

[60] T. Cheng, *Recursive Neural Networks in Quark/Gluon Tagging*, *Comput. Softw. Big Sci.* **2** (2018) 3 [1711.02633].

[61] G. Louppe, K. Cho, C. Becot and K. Cranmer, *QCD-Aware Recursive Neural Networks for Jet Physics*, *JHEP* **01** (2019) 057 [1702.00748].

[62] J. Heeck, *Interpretation of Lepton Flavor Violation*, *Phys. Rev.* **D95** (2017) 015022 [1610.07623].

[63] MEG collaboration, *New constraint on the existence of the $\mu^+ \to e^+\gamma$ decay*, *Phys. Rev. Lett.* **110** (2013) 201801 [1303.0754].

[64] Particle Data Group collaboration, *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.

[65] SINDRUM collaboration, *Search for the Decay $\mu^+ \to e^+e^+e^-$*, *Nucl. Phys.* **B299** (1988) 1.

[66] R. Foot, H. Lew, X. G. He and G. C. Joshi, *Seesaw Neutrino Masses Induced by a Triplet of Leptons*, *Z. Phys.* **C44** (1989) 441.

[67] M. Magg and C. Wetterich, *Neutrino Mass Problem and Gauge Hierarchy*, *Phys. Lett.* **B94** (1980) 61.

[68] G. Lazarides, Q. Shafi and C. Wetterich, *Proton Lifetime and Fermion Masses in an SO(10) Model*, *Nucl. Phys.* **B181** (1981) 287.

[69] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, *Phys. Rev.* **D10** (1974) 275.

[70] R. N. Mohapatra and J. C. Pati, *A Natural Left-Right Symmetry*, *Phys. Rev.* **D11** (1975) 2558.

[71] G. Senjanovic and R. N. Mohapatra, *Exact Left-Right Symmetry and Spontaneous Violation of Parity*, *Phys. Rev.* **D12** (1975) 1502.

[72] P. Duka, J. Gluza and M. Zralek, *Quantization and renormalization of the manifest left-right symmetric model of electroweak interactions*, *Annals Phys.* **280** (2000) 336 [hep-ph/9910279].

[73] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, *Phys. Rept.* **110** (1984) 1.

[74] H. E. Haber and G. L. Kane, *The Search for Supersymmetry: Probing Physics Beyond the Standard Model*, *Phys. Rept.* **117** (1985) 75.

[75] S. P. Martin, *A Supersymmetry primer*, hep-ph/9709356.

[76] K. Hayasaka et al., *Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced Tau+Tau- Pairs*, *Phys. Lett.* **B687** (2010) 139 [1001.3221].

[77] BaBar collaboration, *Limits on tau Lepton-Flavor Violating Decays in three charged leptons*, *Phys. Rev.* **D81** (2010) 111101 [1002.4550].

[78] LHCb collaboration, *Search for the lepton flavour violating decay* $\tau^- \to \mu^-\mu^+\mu^-$, *JHEP* **02** (2015) 121 [1409.8548].

[79] ATLAS collaboration, *Probing lepton flavour violation via neutrinoless $\tau \longrightarrow 3\mu$ decays with the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) 232 [1601.03567].

[80] M. Benedikt and F. Zimmermann, *Future Circular Colliders*, *Proc. Int. Sch. Phys. Fermi* **194** (2016) 73.

[81] A. Arhrib, R. Benbrik, M. Chabab, G. Moultaka, M. C. Peyranere, L. Rahili et al., *The Higgs Potential in the Type II Seesaw Model*, *Phys. Rev.* **D84** (2011) 095005 [1105.1925].

[82] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler and T. Schwetz, *Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity*, *JHEP* **01** (2017) 087 [1611.01514].

[83] M. C. Gonzalez-Garcia, M. Maltoni and T. Schwetz, *Global Analyses of Neutrino Oscillation Experiments*, *Nucl. Phys.* **B908** (2016) 199 [1512.06856].

[84] A. Abada, C. Biggio, F. Bonnet, M. B. Gavela and T. Hambye, *Low energy effects of neutrino masses*, *JHEP* **12** (2007) 061 [0707.4058].

[85] 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].

[86] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [1310.1921].

[87] D. N. Dinh and S. T. Petcov, *Lepton Flavor Violating τ Decays in TeV Scale Type I See-Saw and Higgs Triplet Models*, *JHEP* **09** (2013) 086 [1308.4311].

[88] J. F. Gunion, J. Grifols, A. Mendez, B. Kayser and F. I. Olness, *Higgs Bosons in Left-Right Symmetric Models*, *Phys. Rev.* **D40** (1989) 1546.

[89] W. Grimus and L. Lavoura, *The Seesaw mechanism at arbitrary order: Disentangling the small scale from the large scale*, *JHEP* **11** (2000) 042 [hep-ph/0008179].

[90] Y. Zhang, H. An, X. Ji and R. N. Mohapatra, *General CP Violation in Minimal Left-Right Symmetric Model and Constraints on the Right-Handed Scale*, *Nucl. Phys.* **B802** (2008) 247 [0712.4218].

[91] A. Maiezza, M. Nemevšek and F. Nesti, *Perturbativity and mass scales in the minimal left-right symmetric model*, *Phys. Rev.* **D94** (2016) 035008 [1603.00360].

[92] ATLAS collaboration, *Search for doubly-charged Higgs bosons in same-charge electron pair final states using proton-proton collisions at* $\sqrt{s} = 13\,\text{TeV}$ *with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-051, CERN, Geneva, Aug, 2016.

[93] P. S. B. Dev, R. N. Mohapatra and Y. Zhang, *Probing the Higgs Sector of the Minimal Left-Right Symmetric Model at Future Hadron Colliders*, *JHEP* **05** (2016) 174 [1602.05947].

[94] A. Maiezza, G. Senjanović and J. C. Vasquez, *Higgs sector of the minimal left-right symmetric theory*, *Phys. Rev.* **D95** (2017) 095004 [1612.09146].

[95] J. Chakrabortty, J. Gluza, T. Jelinski and T. Srivastava, *Theoretical constraints on masses of heavy particles in Left-Right Symmetric Models*, *Phys. Lett.* **B759** (2016) 361 [1604.06987].

[96] S. R. Coleman and J. Mandula, *All Possible Symmetries of the S Matrix*, *Phys. Rev.* **159** (1967) 1251.

[97] M. Arana-Catania, S. Heinemeyer and M. J. Herrero, *New Constraints on General Slepton Flavor Mixing*, *Phys. Rev.* **D88** (2013) 015026 [1304.2783].

[98] W. Porod, *SPheno, a program for calculating supersymmetric spectra, SUSY particle decays and SUSY particle production at e^+e^- colliders*, *Comput. Phys. Commun.* **153** (2003) 275 [hep-ph/0301101].

[99] W. Porod and F. Staub, *SPheno 3.1: Extensions including flavour, CP-phases and models beyond the MSSM*, *Comput. Phys. Commun.* **183** (2012) 2458 [1104.1573].

[100] F. Staub, *SARAH 4 : A tool for (not only SUSY) model builders*, *Comput. Phys. Commun.* **185** (2014) 1773 [1309.7223].

[101] ATLAS collaboration, *Summary of the searches for squarks and gluinos using $\sqrt{s} = 8$ TeV pp collisions with the ATLAS experiment at the LHC*, *JHEP* **10** (2015) 054 [1507.05525].

[102] ATLAS collaboration, *ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider*, *Eur. Phys. J.* **C75** (2015) 510 [1506.08616].

[103] CMS collaboration, *Search for supersymmetry in events with one lepton and multiple jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev.* **D95** (2017) 012011 [1609.09386].

[104] CMS collaboration, *Inclusive search for supersymmetry using razor variables in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev.* **D95** (2017) 012003 [1609.07658].

[105] S. Dittmaier, G. Hiller, T. Plehn and M. Spannowsky, *Charged-Higgs Collider Signals with or without Flavor*, *Phys. Rev.* **D77** (2008) 115001 [0708.0940].

[106] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1 [1207.7214].

[107] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30 [1207.7235].

[108] C. Mariotti and G. Passarino, *Higgs boson couplings: measurements and theoretical interpretation*, *Int. J. Mod. Phys.* **A32** (2017) 1730003 [1612.00269].

[109] M. J. Dolan, C. Englert and M. Spannowsky, *Higgs self-coupling measurements at the LHC*, *JHEP* **10** (2012) 112 [1206.5001].

[110] ATLAS collaboration, *Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector*, *Phys. Rev.* **D92** (2015) 092004 [1509.04670].

[111] T. Plehn and M. Rauch, *The quartic higgs coupling at hadron colliders*, *Phys. Rev.* **D72** (2005) 053008 [hep-ph/0507321].

[112] M. Gorbahn and U. Haisch, *Indirect probes of the trilinear Higgs coupling: $gg \to h$ and $h \to \gamma\gamma$*, *JHEP* **10** (2016) 094 [1607.03773].

[113] ATLAS collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton−proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-049, CERN, Geneva, Aug, 2016.

[114] W. Bizon, M. Gorbahn, U. Haisch and G. Zanderighi, *Constraints on the trilinear Higgs coupling from vector boson fusion and associated Higgs production at the LHC*, *JHEP* **07** (2017) 083 [1610.05771].

[115] G. Degrassi, P. P. Giardino, F. Maltoni and D. Pagani, *Probing the Higgs self coupling via single Higgs production at the LHC*, *JHEP* **12** (2016) 080 [1607.04251].

[116] M. E. Peskin and T. Takeuchi, *A New constraint on a strongly interacting Higgs sector*, *Phys. Rev. Lett.* **65** (1990) 964.

[117] M. E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, *Phys. Rev.* **D46** (1992) 381.

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

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

[120] I. Maksymyk, C. P. Burgess and D. London, *Beyond S, T and U*, *Phys. Rev.* **D50** (1994) 529 [hep-ph/9306267].

[121] G. Degrassi, B. A. Kniehl and A. Sirlin, *Gauge invariant formulation of the S, T, and U parameters*, *Phys. Rev.* **D48** (1993) R3963.

[122] Gfitter Group collaboration, *The global electroweak fit at NNLO and prospects for the LHC and ILC*, *Eur. Phys. J.* **C74** (2014) 3046 [1407.3792].

[123] G. D. Kribs, T. Plehn, M. Spannowsky and T. M. P. Tait, *Four generations and Higgs physics*, *Phys. Rev.* **D76** (2007) 075016 [0706.3718].

[124] CDF collaboration, *Search for a Heavy Top-Like Quark in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. Lett.* **107** (2011) 261801 [1107.3875].

[125] CDF collaboration, *Search for heavy bottom-like quarks decaying to an electron or muon and jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. Lett.* **106** (2011) 141803 [1101.5728].

[126] L3 collaboration, *Search for heavy neutral and charged leptons in e^+e^- annihilation at LEP*, *Phys. Lett.* **B517** (2001) 75 [hep-ex/0107015].

[127] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *JHEP* **06** (2007) 045 [hep-ph/0703164].

[128] R. Contino, M. Ghezzi, C. Grojean, M. Mühlleitner and M. Spira, *Effective Lagrangian for a light Higgs-like scalar*, *JHEP* **07** (2013) 035 [1303.3876].

[129] O. V. Tarasov, *Generalized recurrence relations for two loop propagator integrals with arbitrary masses*, *Nucl. Phys.* **B502** (1997) 455 [hep-ph/9703319].

[130] K. G. Chetyrkin and F. V. Tkachov, *Integration by Parts: The Algorithm to Calculate beta Functions in 4 Loops*, *Nucl. Phys.* **B192** (1981) 159.

[131] R. Mertig and R. Scharf, *TARCER: A Mathematica program for the reduction of two loop propagator integrals*, *Comput. Phys. Commun.* **111** (1998) 265 [hep-ph/9801383].

[132] R. Mertig, M. Böhm and A. Denner, *FEYN CALC: Computer algebraic calculation of Feynman amplitudes*, *Comput. Phys. Commun.* **64** (1991) 345.

[133] V. Shtabovenko, R. Mertig and F. Orellana, *New Developments in FeynCalc 9.0*, *Comput. Phys. Commun.* **207** (2016) 432 [1601.01167].

[134] T. Hahn, *Generating Feynman diagrams and amplitudes with FeynArts 3*, *Comput. Phys. Commun.* **140** (2001) 418 [hep-ph/0012260].

[135] S. P. Martin and D. G. Robertson, *TSIL: A Program for the calculation of two-loop self-energy integrals*, *Comput. Phys. Commun.* **174** (2006) 133 [hep-ph/0501132].

[136] A. Sirlin, *Radiative Corrections in the $SU(2)_L \times U(1)$ Theory: A Simple Renormalization Framework*, *Phys. Rev.* **D22** (1980) 971.

[137] G. Degrassi, M. Fedele and P. P. Giardino, *Constraints on the trilinear Higgs self coupling from precision observables*, *JHEP* **04** (2017) 155 [1702.01737].

[138] SNO collaboration, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301 [nucl-ex/0204008].

[139] Super-Kamiokande collaboration, *A Measurement of atmospheric neutrino oscillation parameters by SUPER-KAMIOKANDE I*, *Phys. Rev.* **D71** (2005) 112005 [hep-ex/0501064].

[140] R. N. Mohapatra, *Mechanism for Understanding Small Neutrino Mass in Superstring Theories*, *Phys. Rev. Lett.* **56** (1986) 561.

[141] R. N. Mohapatra and J. W. F. Valle, *Neutrino Mass and Baryon Number Nonconservation in Superstring Models*, *Phys. Rev.* **D34** (1986) 1642.

[142] E. K. Akhmedov, M. Lindner, E. Schnapka and J. W. F. Valle, *Left-right symmetry breaking in NJL approach*, *Phys. Lett.* **B368** (1996) 270 [hep-ph/9507275].

[143] E. K. Akhmedov, M. Lindner, E. Schnapka and J. W. F. Valle, *Dynamical left-right symmetry breaking*, *Phys. Rev.* **D53** (1996) 2752 [hep-ph/9509255].

[144] S. Antusch and O. Fischer, *Non-unitarity of the leptonic mixing matrix: Present bounds and future sensitivities*, *JHEP* **10** (2014) 094 [1407.6607].

[145] A. Pilaftsis, *Radiatively induced neutrino masses and large Higgs neutrino couplings in the standard model with Majorana fields*, *Z. Phys.* **C55** (1992) 275 [hep-ph/9901206].

[146] S. Antusch, E. Cazzato and O. Fischer, *Sterile neutrino searches at future e^-e^+, pp, and e^-p colliders*, *Int. J. Mod. Phys.* **A32** (2017) 1750078 [1612.02728].

[147] A. G. Hessler, A. Ibarra, E. Molinaro and S. Vogl, *Impact of the Higgs boson on the production of exotic particles at the LHC*, *Phys. Rev.* **D91** (2015) 115004 [1408.0983].

[148] C. Degrande, O. Mattelaer, R. Ruiz and J. Turner, *Fully-Automated Precision Predictions for Heavy Neutrino Production Mechanisms at Hadron Colliders*, *Phys. Rev.* **D94** (2016) 053002 [1602.06957].

[149] S. Dawson, *Radiative corrections to Higgs boson production*, *Nucl. Phys.* **B359** (1991) 283.

[150] M. Spira, A. Djouadi, D. Graudenz and P. M. Zerwas, *Higgs boson production at the LHC*, *Nucl. Phys.* **B453** (1995) 17 [hep-ph/9504378].

[151] M. Spira, *QCD effects in Higgs physics*, *Fortsch. Phys.* **46** (1998) 203 [hep-ph/9705337].

[152] R. V. Harlander and W. B. Kilgore, *Production of a pseudoscalar Higgs boson at hadron colliders at next-to-next-to leading order*, *JHEP* **10** (2002) 017 [hep-ph/0208096].

[153] C. Anastasiou and K. Melnikov, *Pseudoscalar Higgs boson production at hadron colliders in NNLO QCD*, *Phys. Rev.* **D67** (2003) 037501 [hep-ph/0208115].

[154] M. Bonvini, S. Forte, G. Ridolfi and L. Rottoli, *Resummation prescriptions and ambiguities in SCET vs. direct QCD: Higgs production as a case study*, *JHEP* **01** (2015) 046 [1409.0864].

[155] A. Atre, T. Han, S. Pascoli and B. Zhang, *The Search for Heavy Majorana Neutrinos*, *JHEP* **05** (2009) 030 [0901.3589].

[156] E. Fernandez-Martinez, J. Hernandez-Garcia and J. Lopez-Pavon, *Global constraints on heavy neutrino mixing*, *JHEP* **08** (2016) 033 [1605.08774].

[157] W.-Y. Keung and G. Senjanovic, *Majorana Neutrinos and the Production of the Right-handed Charged Gauge Boson*, *Phys. Rev. Lett.* **50** (1983) 1427.

[158] A. Datta, M. Guchait and A. Pilaftsis, *Probing lepton number violation via majorana neutrinos at hadron supercolliders*, *Phys. Rev.* **D50** (1994) 3195 [hep-ph/9311257].

[159] D. Alva, T. Han and R. Ruiz, *Heavy Majorana neutrinos from $W\gamma$ fusion at hadron colliders*, *JHEP* **02** (2015) 072 [1411.7305].

[160] G. Breit and E. Wigner, *Capture of Slow Neutrons*, *Phys. Rev.* **49** (1936) 519.

[161] R. E. Ruiz, *Hadron Collider Tests of Neutrino Mass-Generating Mechanisms*, Ph.D. thesis, Pittsburgh U., 2015. 1509.06375.

[162] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[163] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641.

[164] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298.

[165] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization for Short Distance Hadron - Hadron Scattering*, *Nucl. Phys.* **B261** (1985) 104.

[166] T. Hahn, *CUBA: A Library for multidimensional numerical integration*, *Comput. Phys. Commun.* **168** (2005) 78 [hep-ph/0404043].

[167] B. W. Harris and J. F. Owens, *The Two cutoff phase space slicing method*, *Phys. Rev.* **D65** (2002) 094032 [hep-ph/0102128].

[168] R. Ruiz, *QCD Corrections to Pair Production of Type III Seesaw Leptons at Hadron Colliders*, *JHEP* **12** (2015) 165 [1509.05416].

[169] G. Altarelli, R. K. Ellis and G. Martinelli, *Large Perturbative Corrections to the Drell-Yan Process in QCD*, *Nucl. Phys.* **B157** (1979) 461.

[170] S. S. D. Willenbrock and D. A. Dicus, *Production of Heavy Leptons From Gluon Fusion*, *Phys. Lett.* **156B** (1985) 429.

[171] D. A. Dicus and P. Roy, *Supercollider signatures and correlations of heavy neutrinos*, *Phys. Rev.* **D44** (1991) 1593.

[172] L. D. Landau, *On the angular momentum of a system of two photons*, *Dokl. Akad. Nauk Ser. Fiz.* **60** (1948) 207.

[173] C.-N. Yang, *Selection Rules for the Dematerialization of a Particle Into Two Photons*, *Phys. Rev.* **77** (1950) 242.

[174] M. Y. Hussein, *Higgs Boson Production at the LHC*, 1703.03952.

[175] T. Becher, M. Neubert and G. Xu, *Dynamical Threshold Enhancement and Resummation in Drell-Yan Production*, *JHEP* **07** (2008) 030 [0710.0680].

[176] S. Catani and L. Trentadue, *Resummation of the QCD Perturbative Series for Hard Processes*, *Nucl. Phys.* **B327** (1989) 323.

[177] G. F. Sterman, *Summation of Large Corrections to Short Distance Hadronic Cross-Sections*, *Nucl. Phys.* **B281** (1987) 310.

[178] S. Catani and L. Trentadue, *Comment on QCD exponentiation at large x*, *Nucl. Phys.* **B353** (1991) 183.

[179] C. W. Bauer, S. Fleming, D. Pirjol and I. W. Stewart, *An Effective field theory for collinear and soft gluons: Heavy to light decays*, *Phys. Rev.* **D63** (2001) 114020 [hep-ph/0011336].

[180] C. W. Bauer, D. Pirjol and I. W. Stewart, *Soft collinear factorization in effective field theory*, *Phys. Rev.* **D65** (2002) 054022 [hep-ph/0109045].

[181] M. Beneke, A. P. Chapovsky, M. Diehl and T. Feldmann, *Soft collinear effective theory and heavy to light currents beyond leading power*, *Nucl. Phys.* **B643** (2002) 431 [hep-ph/0206152].

[182] T. Becher, M. Neubert and B. D. Pecjak, *Factorization and Momentum-Space Resummation in Deep-Inelastic Scattering*, *JHEP* **01** (2007) 076 [hep-ph/0607228].

[183] T. Becher and M. Neubert, *Threshold resummation in momentum space from effective field theory*, *Phys. Rev. Lett.* **97** (2006) 082001 [hep-ph/0605050].

[184] V. Ahrens, T. Becher, M. Neubert and L. L. Yang, *Renormalization-Group Improved Prediction for Higgs Production at Hadron Colliders*, *Eur. Phys. J.* **C62** (2009) 333 [0809.4283].

[185] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *Higgs Production with a Central Jet Veto at NNLL+NNLO*, *JHEP* **04** (2011) 092 [1012.4480].

[186] T. Ahmed, T. Gehrmann, P. Mathews, N. Rana and V. Ravindran, *Pseudo-scalar Form Factors at Three Loops in QCD*, *JHEP* **11** (2015) 169 [1510.01715].

[187] T. Ahmed, M. Bonvini, M. C. Kumar, P. Mathews, N. Rana, V. Ravindran et al., *Pseudo-scalar Higgs boson production at N^3LO_A+N^3LL'*, *Eur. Phys. J.* **C76** (2016) 663 [1606.00837].

[188] M. Bonvini and L. Rottoli, *Three loop soft function for N^3LL' gluon fusion Higgs production in soft-collinear effective theory*, *Phys. Rev.* **D91** (2015) 051301 [1412.3791].

[189] R. V. Harlander, *Virtual corrections to $gg \to H$ to two loops in the heavy top limit*, *Phys. Lett.* **B492** (2000) 74 [hep-ph/0007289].

[190] A. Idilbi, X.-d. Ji and F. Yuan, *Resummation of threshold logarithms in effective field theory for DIS, Drell-Yan and Higgs production*, *Nucl. Phys.* **B753** (2006) 42 [hep-ph/0605068].

[191] F. Caola and S. Marzani, *Finite fermion mass effects in pseudoscalar Higgs production via gluon-gluon fusion*, *Phys. Lett.* **B698** (2011) 275 [1101.3975].

[192] J. W. Backus and W. P. Heising, *Fortran*, *IEEE Transactions on Electronic Computers* **EC-13** (1964) 382.

[193] 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].

[194] NNPDF collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [1410.8849].

[195] NNPDF collaboration, *Parton distributions with QED corrections*, *Nucl. Phys.* **B877** (2013) 290 [1308.0598].

[196] M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, M. Ubiali, R. D. Ball et al., *Parton distributions with threshold resummation*, *JHEP* **09** (2015) 191 [1507.01006].

[197] M. Bonvini and S. Marzani, *Resummed Higgs cross section at N^3LL*, *JHEP* **09** (2014) 007 [1405.3654].

[198] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, *Phys. Rev. Lett.* **88** (2002) 201801 [hep-ph/0201206].

[199] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, *Phys. Rev. Lett.* **114** (2015) 212001 [1503.06056].

[200] A. J. Larkoski, I. Moult and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, 1709.04464.

[201] S. Carrazza, *Machine learning challenges in theoretical HEP*, *J. Phys. Conf. Ser.* **1085** (2018) 022003 [1711.10840].

[202] R. T. D'Agnolo and A. Wulzer, *Learning New Physics from a Machine*, *Phys. Rev.* **D99** (2019) 015014 [1806.02350].

[203] 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].

[204] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J.* **C76** (2016) 196 [1512.01178].

[205] 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 [1410.3012].

[206] C. Englert, P. Galler, P. Harris and M. Spannowsky, *Machine Learning Uncertainties with Adversarial Neural Networks*, *Eur. Phys. J.* **C79** (2019) 4 [1807.08763].

[207] M. Farina, Y. Nakai and D. Shih, *Searching for New Physics with Deep Autoencoders*, 1808.08992.

[208] T. Heimel, G. Kasieczka, T. Plehn and J. M. Thompson, *QCD or What?*, *SciPost Phys.* **6** (2019) 030 [1808.08979].

[209] G. Altarelli, B. Mele and M. Ruiz-Altaba, *Searching for New Heavy Vector Bosons in $p\bar{p}$ Colliders*, *Z. Phys.* **C45** (1989) 109.

[210] T. Plehn, M. Spannowsky and M. Takeuchi, *How to Improve Top Tagging*, *Phys. Rev.* **D85** (2012) 034029 [1111.5034].

[211] T. Plehn and M. Spannowsky, *Top Tagging*, *J. Phys.* **G39** (2012) 083001 [1112.4441].

[212] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323].

[213] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896 [1111.6097].

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

[215] ATLAS collaboration, *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at* $\sqrt{s} = 8\ TeV$, 2015.

[216] ATLAS collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, *Eur. Phys. J.* **C72** (2012) 1844 [1108.5602].

[217] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 1412.6980.

[218] F. Chollet et al., "Keras." https://github.com/fchollet/keras, 2015.

[219] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, 1603.04467.

[220] G. Louppe, M. Kagan and K. Cranmer, *Learning to Pivot with Adversarial Networks*, 1611.01046.

[221] R. Frederix and F. Maltoni, *Top pair invariant mass distribution: A Window on new physics*, *JHEP* **01** (2009) 047 [0712.2355].

[222] ATLAS collaboration, *Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at* $\sqrt{s} = 13\ TeV$ *with the ATLAS detector*, *Eur. Phys. J.* **C78** (2018) 565 [1804.10823].

[223] K. Hornik, M. Stinchcombe and H. White, *Multilayer feedforward networks are universal approximators*, *NeuralNetworks* **2** (1989) 359.

[224] K. Hornik, *Approximation capabilities of multilayer feedforward networks*, *NeuralNetworks* **4** (1991) 251.

[225] I. Fried, *Numerical Solution of Differential Equations*, Computer Science and Applied Mathematics Series. Acad. Press, 1979.

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

[227] V. A. Kuzmin, V. A. Rubakov and M. E. Shaposhnikov, *On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe*, *Phys. Lett.* **155B** (1985) 36.

[228] A. Kosowsky, M. S. Turner and R. Watkins, *Gravitational waves from first order cosmological phase transitions*, *Phys. Rev. Lett.* **69** (1992) 2026.

[229] C. Grojean and G. Servant, *Gravitational Waves from Phase Transitions at the Electroweak Scale and Beyond*, *Phys. Rev.* **D75** (2007) 043507 [hep-ph/0607107].

[230] H. Lee and I. S. Kang, *Neural algorithm for solving differential equations*, *Journal of Computational Physics* **91** (1990) 110.

[231] A. Meade and A. Fernandez, *The numerical solution of linear ordinary differential equations by feedforward neural networks*, *Mathematical and Computer Modelling* **19** (1994) 1.

[232] A. Meade and A. Fernandez, *Solution Of Nonlinear Ordinary Differential Equations By Feedforward Neural Networks*, *Mathematical and Computer Modelling* **20** (1994) 19.

[233] I. E. Lagaris, A. Likas and D. I. Fotiadis, *Artificial neural networks for solving ordinary and partial differential equations*, *IEEE Transactions on Neural Networks* **9** 987 [physics/9705023].

[234] S. Mall and S. Chakraverty, *Comparison of Artificial Neural Network Architecture in Solving Ordinary Differential Equations*, *Advances in Artificial Neural Systems* **2013** (2013) 1.

[235] A. Iftikhar and M. Bilal, *Numerical Solution of Blasius Equation through Neural Networks Algorithm*, *American Journal of Computational Mathematics* **04** (2014) 223.

[236] T. E. Oliphant, *A guide to NumPy*, vol. 1. Trelgol Publishing USA, 2006.

[237] D. Maclaurin et al., "Autograd." `https://github.com/HIPS/autograd`, 2015.

[238] M. E. Shaposhnikov, *Possible Appearance of the Baryon Asymmetry of the Universe in an Electroweak Theory*, *JETP Lett.* **44** (1986) 465.

[239] A. D. Linde, *Particle physics and inflationary cosmology*, *Contemp. Concepts Phys.* **5** (1990) 1 [hep-th/0503203].

[240] G. W. Anderson and L. J. Hall, *The Electroweak phase transition and baryogenesis*, *Phys. Rev.* **D45** (1992) 2685.

[241] M. Dine, R. G. Leigh, P. Huet, A. D. Linde and D. A. Linde, *Comments on the electroweak phase transition*, *Phys. Lett.* **B283** (1992) 319 [hep-ph/9203201].

[242] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, *Phys. Rev.* **D51** (1995) 379 [hep-ph/9404302].

[243] R. K. Pathria, *Statistical Mechanics*. Butterworth-Heinemann, 1996.

[244] V. A. Rubakov and M. E. Shaposhnikov, *Electroweak baryon number nonconservation in the early universe and in high-energy collisions*, *Usp. Fiz. Nauk* **166** (1996) 493 [hep-ph/9603208].

[245] D. E. Morrissey and M. J. Ramsey-Musolf, *Electroweak baryogenesis*, *New J. Phys.* **14** (2012) 125003 [1206.2942].

[246] S. R. Coleman, *The Fate of the False Vacuum. 1. Semiclassical Theory*, *Phys. Rev.* **D15** (1977) 2929.

[247] A. D. Linde, *Decay of the False Vacuum at Finite Temperature*, *Nucl. Phys.* **B216** (1983) 421.

[248] S. R. Coleman, V. Glaser and A. Martin, *Action Minima Among Solutions to a Class of Euclidean Scalar Field Equations*, *Commun. Math. Phys.* **58** (1978) 211.

[249] R. Apreda, M. Maggiore, A. Nicolis and A. Riotto, *Gravitational waves from electroweak phase transitions*, *Nucl. Phys.* **B631** (2002) 342 [gr-qc/0107033].

[250] C. L. Wainwright, *CosmoTransitions: Computing Cosmological Phase Transition Temperatures and Bubble Profiles with Multiple Fields*, *Comput. Phys. Commun.* **183** (2012) 2006 [1109.4189].

[251] P. Athron, C. Balázs, M. Bardsley, A. Fowlie, D. Harries and G. White, *BubbleProfiler: finding the field profile and action for cosmological phase transitions*, *Comput. Phys. Commun.* **244** (2019) 448 [1901.03714].

[252] L. V. Kantorovich, *Functional analysis and applied mathematics*, *Uspekhi Math. Nauk* **3** (1948) 89.

[253] S. Akula, C. Balázs and G. A. White, *Semi-analytic techniques for calculating bubble wall profiles*, *Eur. Phys. J.* **C76** (2016) 681 [1608.00008].

[254] A. Masoumi, K. D. Olum and B. Shlaer, *Efficient numerical solution to vacuum decay with many fields*, *JCAP* **1701** (2017) 051 [1610.06594].

[255] R. Jinno, *Machine learning for bounce calculation*, 1805.12153.

[256] J. R. Espinosa, *A Fresh Look at the Calculation of Tunneling Actions*, *JCAP* **1807** (2018) 036 [1805.03680].

[257] J. R. Espinosa and T. Konstandin, *A Fresh Look at the Calculation of Tunneling Actions in Multi-Field Potentials*, *JCAP* **1901** (2019) 051 [1811.09185].

[258] V. Guada, A. Maiezza and M. Nemevšek, *Multifield Polygonal Bounces*, *Phys. Rev.* **D99** (2019) 056020 [1803.02227].

[259] P. John, *Bubble wall profiles with more than one scalar field: A Numerical approach*, *Phys. Lett.* **B452** (1999) 221 [hep-ph/9810499].

[260] J. M. Cline, G. D. Moore and G. Servant, *Was the electroweak phase transition preceded by a color broken phase?*, *Phys. Rev.* **D60** (1999) 105035 [hep-ph/9902220].

[261] T. Konstandin and S. J. Huber, *Numerical approach to multi dimensional phase transitions*, *JCAP* **0606** (2006) 021 [hep-ph/0603081].

[262] J.-h. Park, *Constrained potential method for false vacuum decays*, *JCAP* **1102** (2011) 023 [1011.4936].

[263] S. Profumo, M. J. Ramsey-Musolf and G. Shaughnessy, *Singlet Higgs phenomenology and the electroweak phase transition*, *JHEP* **08** (2007) 010 [0705.2425].

[264] J. M. No and M. Spannowsky, *Signs of heavy Higgs bosons at CLIC: An e^+e^- road to the Electroweak Phase Transition*, *Eur. Phys. J.* **C79** (2019) 467 [1807.04284].

[265] A. Kusenko, *Sterile neutrinos, dark matter, and the pulsar velocities in models with a Higgs singlet*, *Phys. Rev. Lett.* **97** (2006) 241301 [hep-ph/0609081].

[266] C. P. Burgess, M. Pospelov and T. ter Veldhuis, *The Minimal model of nonbaryonic dark matter: A Singlet scalar*, *Nucl. Phys.* **B619** (2001) 709 [hep-ph/0011335].

[267] J. M. Cline and K. Kainulainen, *Electroweak baryogenesis and dark matter from a singlet Higgs*, *JCAP* **1301** (2013) 012 [1210.4196].

[268] M. Chala, G. Nardini and I. Sobolev, *Unified explanation for dark matter and electroweak baryogenesis with direct detection and gravitational wave signatures*, *Phys. Rev.* **D94** (2016) 055006 [1605.08663].

[269] J. R. Espinosa, T. Konstandin and F. Riva, *Strong Electroweak Phase Transitions in the Standard Model with a Singlet*, *Nucl. Phys.* **B854** (2012) 592 [1107.5441].

[270] J. M. Moreno, M. Quiros and M. Seco, *Bubbles in the supersymmetric standard model*, *Nucl. Phys.* **B526** (1998) 489 [hep-ph/9801272].