Studies of effective theories beyond the Standard Model

Stella Riad

Theoretical Particle Physics, Department of Theoretical Physics,
School of Engineering Sciences
Royal Institute of Technology, SE-106 91 Stockholm, Sweden

Stockholm, Sweden 2014
Abstract

The vast majority of all experimental results in particle physics can be described by the Standard Model (SM) of particle physics. However, neither the existence of neutrino masses nor the mixing in the leptonic sector, which have been observed, can be described within this model. In fact, the model only describes a fraction of the known energy in the Universe. Thus, we know there must exist a theory beyond the SM. There is a plethora of possible candidates for such a model, such as supersymmetry, extra dimensional theories, and string theory. So far, there are no evidence in favor of these models.

These theories often reside at high energies, and will therefore be manifest as effective theories at the low energies experienced here on Earth. A first example is extra-dimensional theories. From our four-dimensional point of view, particles which propagate through the extra dimensions will effectively be perceived as towers of heavy particles. In this thesis we consider an extra-dimensional model with universal extra dimensions, where all SM particles are allowed to propagate through the extra dimensions. Especially, we place a bound on the range of validity for this model. We study the renormalization group running of the leptonic parameters as well as the Higgs self-coupling in this model with the neutrino masses generated by a Weinberg operator.

Grand unified theories, where the gauge couplings of the SM are unified into a single one at some high energy scale, are motivated by the electroweak unification. The unification must necessarily take place at energies many orders of magnitude greater than those that ever can be achieved on Earth. In order to make sense of the theory, which is given at the grand unified scale, at the electroweak scale, the symmetry at the grand unified scale is broken down to the SM symmetry. Within these models the SM is considered as an effective field theory. We study renormalization group running of the leptonic parameters in a non-supersymmetric SO(10) model which is broken in two steps via the Pati-Salam group.

Finally, the discovery of the new boson at the LHC provides a new opportunity to search for physics beyond the SM. We consider an effective model where the magnitudes of the couplings in the Higgs sector are scaled by so-called coupling scale factors. We perform Bayesian parameter inference based on the LHC data. Furthermore, we perform Bayesian model comparison, comparing models where one or several of the Higgs couplings are allowed, to the SM, where couplings are fixed.

**Key words:** Effective field theories, neutrino physics, extra dimensions, universal extra dimensions, Higgs physics, renormalization group running, Bayesian statistics, coupling scale factor, grand unified theories.
Sammanfattning


Nyckelord: Effektiva fältteorier, Higgsfysik, storförena teorier, Bayesisk statistik, neutrinomassor, renorneringsgruppslöpande, universella extra dimensioner.
Preface

This thesis is the result of my research carried out at the Department of Theoretical Physics at KTH Royal Institute of Technology from June 2012 to September 2014. The thesis is divided into two parts. The first part is an introduction to the subjects relevant for the scientific papers. These subjects include a set of different theories, Higgs physics, renormalization group running, and Bayesian statistics. The second part consists of the scientific papers that was the result of the research. These are listed below.

Overview of the thesis

In Ch. 1, I give an introduction to science, the subject of physics in general, and particle physics in particular. In Ch. 2, I give a short introduction to quantum field theory in general, and to the SM, which is the quantum field theory used in particle physics. Furthermore, I discuss the limitations and problems of the SM and hints for physics beyond. In Ch. 3, I discuss the concept of effective field theories and give an introduction to the theories that are studied in this thesis. I discuss the concept of an extra-dimensional theory known as Universal Extra Dimensions. Furthermore, I discuss the concept of grand unified theories, and especially a non-supersymmetric model based on the $SO(10)$ gauge group. Finally, I investigate the possibility to use the Higgs sector as a probe for new physics using effective scale couplings. In Ch. 4, I discuss the concepts of regularization, renormalization, and renormalization group running. Finally, in Ch. 5, I discuss the notion of probability and the two fundamental interpretations which leads to two schools of statistics, Bayesian and frequentist statistics.

List of papers included in this thesis

[1] T. Ohlsson, and S. Riad
Running of Neutrino Parameters and the Higgs Self-Coupling in a Six-Dimensional UED Model
arXiv:1208.6297
The thesis author’s contribution to the papers

[1] I performed the analytical and numerical calculations and I produced the figures. I did part of the writing.

[2] I performed all of the numerical calculations. I produced parts of the figures and did part of the writing.

[3] I performed the numerical calculations for the model comparison in Sec. 4. I made some of the plots. I did part of the writing.

Notations and conventions

The metric tensor on Minkowski space used is

\[(g_{\mu\nu}) = \text{diag}(1, -1, -1, -1).\]  

(1)

Unless otherwise stated we use Einstein’s summation convention, \textit{i.e.} we implicitly sum over repeated spacetime indices, one covariant and one contravariant.

For the sake of clarity we use natural units, \(\hbar = c = 1\).
Acknowledgements

First, I would like to thank my supervisor Tommy Ohlsson for giving me the opportunity to do my PhD studies in theoretical particle physics. Also thanks for the collaboration which led to two papers in this thesis and for proof-reading of the thesis.

Special thanks to Davide Meloni for interesting discussions and especially for the collaboration that lead to Paper II. Also special thanks to Johannes Bergström, for providing great company in the office, the collaboration which resulted in Paper III, help with many small things, and the careful proof-reading of chapter 5 of this thesis. Furthermore, I would like to thank my co-supervisor Mattias Blennow for good advice, for mediating the joy of physics, and for the proof-reading of the thesis.

During these past years, I have had the privilege to share the office with a number of eminent people. I would like to thank Henrik for the supervision during my time as a master student, and for the many discussions by the coffee machine; Marcus and Björn for great discussions and plenty of laughs; and Shun for being of help whenever needed. Thank you Jessica for the chocolate, the careful proof-reading, and for being a true friend. I am glad you have only moved upstairs. Furthermore thanks to all other members of the group, who have contributed to a good working atmosphere. Also great thanks to the other PhD students at the department, both past and present, for providing good company at both lunch and fika.

Outside of physics, first of all I would like to thank my family. My mother Cecilia, my father Tomas, and my sister Magdalena for their encouragement, love, and support throughout the years. Thank you Maria, Johanna, Victoria, and the rest of my extended family for support and interest even though nobody actually understood what I was doing. Finally, I would like to thank all of my friends, who make my world a better place. You know who you are.
Contents

Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iii
Sammanfattning . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iii
Preface v
Acknowledgements vii
Contents ix

I Introduction and background material 1

1 Introduction 3

2 The Standard Model and slightly beyond 9
  2.1 Quantum Field Theory . . . . . . . . . . . . . . . . . . . . . . . . 9
  2.2 The Standard Model . . . . . . . . . . . . . . . . . . . . . . . . . 10
    2.2.1 The SM gauge group . . . . . . . . . . . . . . . . . . . . . . . 10
    2.2.2 The Standard Model Lagrangian . . . . . . . . . . . . . . . 11
    2.2.3 Gauge bosons . . . . . . . . . . . . . . . . . . . . . . . . . . 11
    2.2.4 Fermions . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
    2.2.5 The Higgs sector and electroweak symmetry breaking . . . 13
    2.2.6 Neutrinos: mixings and masses . . . . . . . . . . . . . . . . 16
    2.2.7 Problems with the SM . . . . . . . . . . . . . . . . . . . . . 18

3 Models beyond the SM 19
  3.1 Effective field theories . . . . . . . . . . . . . . . . . . . . . . . 19
  3.2 Universal Extra Dimensions . . . . . . . . . . . . . . . . . . . . 21
    3.2.1 Kaluza-Klein decomposition . . . . . . . . . . . . . . . . . 22
    3.2.2 Six-dimensional UED model . . . . . . . . . . . . . . . . . 23
  3.3 A non-supersymmetric SO(10) grand unified model . . . . . . 26
    3.3.1 A short introduction to Grand Unified Theories . . . . . . 27
    3.3.2 A non-supersymmetric SO(10) model . . . . . . . . . . . . 27
  3.4 New physics in the Higgs Sector . . . . . . . . . . . . . . . . . . 30
3.4.1 Coupling scale factors ........................................... 32

4 Renormalization group running ..................................... 35
  4.1 Regularization ...................................................... 35
  4.1.1 Dimensional regularization .................................... 36
  4.2 Renormalization ..................................................... 38
    4.2.1 Minimal subtraction ........................................... 39
  4.3 Renormalization group equations ................................. 39

5 Statistics .................................................................. 41
  5.1 Probability ............................................................ 42
    5.1.1 Axioms of probability .......................................... 42
    5.1.2 Interpretation of probability ................................... 42
  5.2 Bayesian statistics .................................................. 43
    5.2.1 Parameter inference ............................................. 44
    5.2.2 Priors .............................................................. 44
    5.2.3 Model comparison ............................................... 46
  5.3 Frequentist statistics ............................................... 48
    5.3.1 Parameter estimation ........................................... 48
    5.3.2 Hypothesis testing .............................................. 49

6 Summary and conclusions ............................................. 51

Bibliography .................................................................. 52

II Scientific papers ....................................................... 63
Part I

Introduction and background material
Chapter 1

Introduction

Curiosity is a most characteristic feature of human nature. The ubiquitous ques-
tions of “how” and “why” are asked by humans from their childhood. This curiosity
is the driving force for increasing our knowledge of Nature and the world surround-
ing us. It has taken us to the Moon, allowed us to explore our genes, and provided
us with knowledge about the largest scales of our Universe.

A natural part of human societies has for many centuries been the building and
organization of the systematic studies of the material world. Stemming from the
latin word for knowledge, “scientia”, is the concept of science. Or, perhaps better
put in the facetious words of the famous American scientist Richard Feynman, “Sci-
ence is the belief in the ignorance of experts” [4]. The scientific branch known as
physics broadly describes our knowledge of Nature. It describes the natural world,
the Universe, and our knowledge of energy, matter, time, space, and the connec-
tions between them. Hence, physics encompasses the study of the smallest known
building stones of Nature such as elementary particles, atoms, and molecules, as
well as phenomena at the very largest scales of the Universe and beyond. Conse-
quently, the time scale of physical experiments ranges from the age of the Universe
to small fractions of a second which is the decay times of heavy, unstable elementary
particles.

Physics can in turn be divided into two distinct branches, theoretical and exper-
imental physics. Theoretical physics uses mathematics and abstraction of physical
objects as a means for describing and predicting phenomena in Nature. Experi-
mental physics is concerned with the observations of physical phenomena and ex-
periments. The two branches are strongly interconnected and one cannot survive
without the other. Theoretical predictions motivate the design of experiments
and from experimental results, new theoretical ideas are born. When physics first
emerged as a discipline there was no distinction between a theorist and an exper-
imentalist. It was feasible as long as the experiments could be contained in a small
room and the theoretical field was limited enough to be possible to grasp for a single
person. However, as the science has rapidly developed, the option of being both
is naturally no longer viable for any physicist. The experiments have developed
from table-top designs constructed by a single scientist to huge apparatuses built
and governed by collaborations of hundreds of physicists contained in enormous
cavities. Experiments are not confined to the Earth but sent into space, to the
Moon, and further beyond. In theoretical physics the changes have perhaps not
been as obvious as in experimental physics, but nonetheless a significant evolution
has taken place concurrently with the growth of the field. Especially, the desktop
computers have allowed for a tremendous evolution of the discipline. It should be
emphasized that the aim and scope of physics never is to explain Nature, only to
describe it. Once more, in the words of Feynman “While I am describing to you
how Nature works, you won’t understand why Nature works that way. But you see,
nobody understands that” [5].

The task of constructing a new physical theory is far from a simple one. When
trying to grapple it, the principle of Occam’s razor is of great importance. Stated as
“entities should not be multiplied unnecessarily” by the friar William of Ockham
in the 14th century it advocates that, among a number of hypotheses, the one
making the fewest assumptions should always be favoured. Hence, models should be
constructed as simple as experimental results allow them. On the other hand beauty
is often an elevated motive in the quest for a theory of Nature. It is as important
to construct the theory neither more complicated nor simpler than necessary. We
must make the choice that is right and not the one that is easy. Simplicity and
beauty, should always be second to truth, once again expressed by Feynman “it
doesn’t make a difference how smart you are who made the guess, or what his
name is. If it disagrees with experiment it is wrong”.

Any good physical theory makes predictions that could, and should, be investi-
gated experimentally. This should not only be an abstract notion but also carried
out in reality. The notion of falsifiability, or refutability, presents an opportunity
to judge theories. A physical theory is falsifiable if a counterexample can be found
which proves the theory wrong, given that the theory is false. Hence the theoretical
predictions of the theory should be possible to investigate in an experiment, not
only possible to construct in theory, but also in reality in a not-too-distant future.

The smallest known building stones of our Universe are so-called elementary
particles, particles without any known substructures. The scientific field which
comprises the study of such particles is known as elementary particle physics. Since
the study of Nature’s smallest building stones requires the highest energies, the field
is often denominated high-energy physics (HEP). Studies of these particles can be
carried out at large facilities called particle accelerators. In these facilities, particles
are accelerated to speeds close to that of light and then collided head on. In order to
produce a new particle the collisions need to take place at energies higher than the
rest mass of the particle. Hence, the heavier the particle, the higher the accelerator
energy. To discover new interactions and new particles, energy colliders of even
higher energies have been constructed. At the time of the writing of this thesis,
the most powerful particle accelerator in the world is the Large Hadron Collider
(LHC) at CERN which began operating in 2009. In the final stage of operation,
the LHC will collide particles at energies of 14 TeV. The part of theoretical particle physics bordering to experiment is known as phenomenology, which encompasses the interpretation of the results of HEP. It acts as a bridge between theory and experiment and deals with applications of theory to experiments, especially the experimental signals which can be expected from the theories.

At present the best theory for describing these particles and the interactions among them is called the Standard Model (SM) of particle physics. It describes all known elementary particles: the fermions, which are the particles that make up all matter, and the bosons, which consists of the force carrying particles and the Higgs boson. Three of the four fundamental interactions are included in the SM: electromagnetism, the strong, and the weak force. The fourth fundamental interaction, gravity, is missing. This is perhaps the force which is most obvious to us in our everyday life, it keeps us standing firmly on the ground and the Earth in orbit around the Sun. In particle physics, however, gravity is absolutely negligible. Electromagnetism concerns the interactions of electrically charged objects and is the force responsible for all interactions above nuclear level which cannot be accredited to gravity; for instance all forces experienced when pushing and dragging physical objects. The weak force is responsible for both radioactive decay and nuclear fusion. It is a short range interaction since the corresponding gauge bosons are massive. The strong force has a very short range but over the range which it does interact it triumphs over the other interactions and it is the force responsible for keeping nuclei and nucleons together. The electromagnetic and weak force is unified into a single force, the electroweak force.

The SM was finalized in the 1970s as a joint effort of many prominent physicists after a struggle lasting well over a decade. The first step was taken when Sheldon Glashow in 1961 discovered a way of combining the electromagnetic and weak forces into one [6]. By the incorporation of the Brout-Englert-Higgs (BEH) mechanism [7–10] by Steven Weinberg and Abdus Salam in 1968, the electroweak theory took its modern form [11, 12]. The theory was commonly accepted after the experimental discovery of the weak neutral current, mediated by the \( Z \) boson, by the Gargamelle experiment at CERN in 1973 [13, 14]. At the time that the electroweak theory was formulated the strong interaction was still not properly understood. Instead all there was, was a zoo of hadrons. However, in 1964, Murray Gell-Mann and George Zweig independently proposed that hadrons were not elementary particles but rather constituted by smaller entities [15–17], which were named quarks after a word coined by James Joyce in the novel Finnegans wake [18]. Following on this proposition, which was experimentally confirmed in a deep inelastic scattering experiment in 1968 at Stanford Linear Accelerator Center (SLAC) [19, 20], more quarks were suggested until all six flavors had been discovered [21–25]. Through the introduction of color charge and the formulation of the color group as a gauge group a quantum field theory known as quantum chromodynamics (QCD) could be formulated. The discovery of asymptotic freedom, i.e. that the bonds between the color charged particles become weaker as the energy becomes higher, in the beginning of the 1970s allowed for calculations of high-energy experiments in QCD.
using perturbation theory. Finally, all pieces needed for the SM were there.

The SM has ever since been an excellent example of a good physical theory. At the time not all of the particles present in the theory had been discovered, it was a theory with a lot of blanks. However, all of the predicted but missing particles were discovered in form of the bottom quark, the W and Z boson, the top quark, the tau neutrino, and the Higgs boson in 1977, 1983, 1995, 2000, and 2012, respectively \[21, 24–30\]. The search for the Higgs boson was carried out in increasingly large particle accelerators, and finally it was discovered by the ATLAS and CMS experiments at the LHC. The discovery cements our knowledge on the generation of particle masses and marks the completion of the SM and the end of the quest for a particle which has been going on ever since it was postulated in the 1960s as a rest product from electroweak symmetry breaking (EWSB) \[7, 9, 10\]. Furthermore the discovery of the Higgs boson does provide a new possibility for new studies of the SM and beyond.

The SM is the by far most successful theory in particle physics, since it describes the vast majority of all experimental particle physics results. We do, however, know that it cannot be the final theory. In fact the theory only describes about 5 % of the energy in the Universe since it does not accommodate neither dark matter nor dark energy. Furthermore, it does not describe the asymmetry of particles and anti-particles in the Universe. As previously mentioned, the perhaps most obvious shortcoming of the theory is that gravity is not contained within the SM framework. Instead gravity is well described by Einstein’s theory of general relativity \[31\]. It has however proven extremely difficult, if not impossible, to unify gravity and quantum mechanics, and thus construct a theory of quantum gravity. In addition, there is no place for neutrino masses in the SM.

The neutrino is a SM particle, a fermion that only interacts via the weak force. Its existence was suggested by Wolfgang Pauli in 1930 in a letter to the radioactive ladies and gentlemen meeting in Thübingen, as a solution to the problem of missing mass and momentum in beta decay. The neutrinos come in three different flavors: the electron, muon, and tau neutrino, one corresponding to each charged lepton in the SM. Since there are no right-handed neutrinos in the SM, they must be massless within the SM framework. However, experiments show that the neutrinos oscillate, i.e. change flavor, as they travel through space and time. These oscillations are only possible if neutrinos have non-zero mass.

In order to describe this, new physics extensions need to be made to the SM. Examples of simple extensions are supersymmetry, where each SM particle is endowed with a superpartner, or extra dimensions, where one or several extra spatial dimensions are added to the already existing three. Supersymmetry has been fervently studied in different theoretical contexts during the past decades and there are ongoing searches for supersymmetry at the LHC. So far, there has not been any indications of its existence and the limits on the simplest supersymmetric models are by now rather severe, see e.g. Refs. \[32, 33\]. The introduction of one extra dimension was first suggested by Theodor Kaluza in 1921, when he extended general relativity to five dimensions as a method to unify electromagnetism and gravity \[34\].
This idea was developed further by Oskar Klein in 1926, who suggested that the extra dimensions should be compact, i.e. small and finite [35]. Extra-dimensional theories constructed in the same spirit are nowadays referred to as Kaluza-Klein theories. The sizes of the extra dimensions are rather constrained, since we neither perceive them in everyday life nor detect them at experiments such as the LHC.

The subject of theories beyond the SM is flourishing. Guided by the unification of the electromagnetic and weak interactions, there is hope for a theory where the gauge couplings of the three fundamental interactions are unified into a single one. This happens at an energy far beyond what we can, and will be able to, achieve in any experiment here on Earth. Such theories are known as Grand Unified Theories (GUTs), and the energy scale of the unification is called the GUT energy. There is a plethora of these types of theories, based on different gauge groups and symmetry considerations. So far none have been successful but hopefully, one of them will bear fruit one day. Ultimately, the goal is set even higher, on the unification of all four fundamental forces in Nature, and a Theory for Everything.
Chapter 2

The Standard Model and slightly beyond

The SM is the current framework, which describes physics at the smallest scales, or equivalently, the highest energies. It describes the known elementary particles and the interactions of the electromagnetic, weak, and strong force and it is an example of a quantum field theory (QFT).

In this chapter we give an introduction to QFTs in general, and the SM in particular with emphasis on the parts of the theory which are of particular interest for the subjects studied in the following chapters. Especially we discuss EWSB and the Brout-Englert-Higgs (BEH) mechanism, as well as neutrino masses. Finally, we discuss some of the conceptual issues of the model.

2.1 Quantum Field Theory

QFT is a framework which is used in both particle physics and condensed matter physics in order to describe systems with an infinite number of degrees of freedom, such as fields and many-body systems. In this framework quantum mechanics is applied to systems of fields. Fields, which are physical quantities with specified values in each point in space and time in a classical theory, are quantized and instead specified in terms of operators. Particles are associated with excitations of quantum fields, and particle interactions are described through interactions of the underlying fields.

Normally, a QFT is formulated in terms of a Lagrangian density, or Lagrangian for short, \( \mathcal{L}(\phi_i(x), \partial_\mu \phi_i(x)) \), which is a function of the fields and the first order
derivatives of the fields. The dynamics of the system is described by the action, which is dimensionless in natural units and given by

\[ S[\phi] = \int d^4x \mathcal{L}(\phi_i(x), \partial_\mu \phi_i(x)). \] (2.1)

From the Lagrangian the equations of motion are obtained via a generalization of the Euler-Lagrange equations from classical mechanics, which are given by

\[ \frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} = 0, \] (2.2)

where \( \phi_i \) is the studied field.

A QFT is, in principle, determined by the Lagrangian. However, the observables, and hence the interesting parts of the theory are not the values of the fields themselves but rather quantities such as cross sections and decay rates. These can be determined from the S-matrix, relating the ingoing and outgoing states involved in a scattering process.

### 2.2 The Standard Model

The particle content of the SM can be categorized by spin, masses, and other quantum numbers. Based on spin there are three main categories: spin-1/2 fermions, spin-1 gauge bosons, and the spin-0 Higgs boson, which is the only scalar particle of the SM. Corresponding to each particle in the SM there is an antiparticle with the same mass but opposite quantum numbers.

In physics various kinds of symmetries are of great importance. The SM is a local gauge theory. A gauge theory is a field theory where the Lagrangian is invariant under a continuous Lie group of local transformations. A local theory is dependent on the spatial coordinates in each point, and thus transforms depending on the spacetime point. Furthermore, the SM is a relativistic QFT, which implies that it is invariant under transformations of the Lorentz group.

#### 2.2.1 The SM gauge group

The gauge group of the SM is the 12-dimensional non-Abelian symmetry group

\[ G = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \] (2.3)

The gauge group \( SU(2)_L \otimes U(1)_Y \) corresponds to the electroweak sector and \( SU(3)_C \) to the strong sector. The electroweak sector is described by the Glashow-Weinberg-Salam theory [6, 11, 36]. Corresponding to the four-dimensional electroweak group there are four group generators. Three of them come from \( SU(2)_L \), which is generated by \( T^a = \frac{ie^a}{2} \), where \( \tau^a, \ a \in \{1,2,3\} \), are the three Pauli matrices. The subscript of \( SU(2)_L \) stands for left implying that the SM is a chiral theory, i.e. only
2.2. The Standard Model

particles with left-handed chirality acts under $SU(2)_L$. The fourth group generator is the one of $U(1)_Y$, $T^a = Y$, where $Y$ is the hypercharge, related to $Q$, electric charge, and $I_3$, weak isospin, through the Gell-Mann-Nishijima formula. We use the convention $Y = Q - I_3$ [37, 38]. The subscript $Y$ distinguishes it from the gauge group $U(1)_Q$ of electrical charge, which is the symmetry remaining after EWSB. The strong interactions are described in QCD with an eight-dimensional gauge group. Hence, the group has eight generators, $T^a = \frac{i\lambda^a}{2}$, where $\lambda^a$, $a \in \{1, \ldots, 8\}$, are the Gell-Mann matrices. The subscript, C, stands for color charge, the quantum number of the group, which comes in three variants: red, green, and blue.

2.2.2 The Standard Model Lagrangian

The Lagrangian describes the dynamics of a system. It can contain three different types of terms: kinetic terms, which are quadratic in the derivative of a single field, mass terms, which are quadratic in a single field, without any derivatives, and interaction terms which contain more than two fields of one or several types, with or without derivatives. The SM Lagrangian can be written as a sum of four different parts

$$L_{SM} = L_{\text{gauge}} + L_{\text{fermion}} + L_{\text{Higgs}} + L_{\text{Yukawa}},$$

where $L_{\text{gauge}}$ contains the kinetic terms and self-interactions of the gauge bosons, $L_{\text{fermion}}$ contains all kinetic terms and gauge interaction terms of the fermions, $L_{\text{Higgs}}$ is the scalar sector of the Lagrangian with the kinetic terms, gauge boson interactions, and self-interaction of the Higgs boson, and finally $L_{\text{Yukawa}}$ describes all the interactions of the Higgs boson with leptons and quarks. The different parts of the Lagrangian are discussed in the subsequent sections.

2.2.3 Gauge bosons

The gauge bosons are the spin-1 particles mediating the fundamental interactions. In total there are twelve gauge bosons, which correspond to each of the twelve group generators previously discussed. The gauge bosons corresponding to the electroweak sector are the B boson, from $U(1)_Y$, and a triplet, $W^1$, $W^2$, and $W^3$, from the three-dimensional $SU(2)_L$. After EWSB the B boson mix with the three $W^i$, and will thus, upon rediagonalization, form the physical bosons, i.e. the massless photon, $\gamma$, which is the gauge boson of $U(1)_Q$ and the massive $Z$ and $W^\pm$ bosons, which mediate the neutral and charged weak current interactions. The strong force is mediated by eight gluons $G^a$, corresponding to the eight-dimensional gauge group $SU(3)_C$.

The SM part of the Lagrangian that describes the kinetic terms and the self-interactions of the gauge fields is given by

$$L_{\text{gauge}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W^{i,\mu\nu} - \frac{1}{4} G_{\mu\nu}^a G^{a,\mu\nu},$$

(2.5)
where \( i \in \{1, 2, 3\} \) and \( a \in \{1, 2, \ldots, 8\} \). The field strength tensors are given by

\[
B_{\mu \nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \\
W^i_{\mu \nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g \epsilon^{ijk} W^j_\mu W^k_\nu, \\
G^a_{\mu \nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_3 f^{abc} G^b_\mu G^c_\nu,
\]

where \( g \) and \( g_3 \) are the gauge couplings of \( SU(2)_L \) and \( SU(3)_C \) respectively, and \( \epsilon^{ijk} \) and \( f^{abc} \) are the corresponding structure constants. The self-interactions of \( W^i \) and \( G^a \) are due to the fact that the gauge groups are non-Abelian. It is not possible to construct gauge invariant mass terms for the gauge bosons directly, therefore EWSB, which is discussed in Sec. 2.2.5, is required in order to obtain massive gauge bosons.

### 2.2.4 Fermions

There are in total twelve spin-1/2 fermions, which are further divided into two main categories, quarks and leptons, with six particles in each category. The quarks and leptons are then divided into three fermion generations, which couple identically to gauge bosons. The generations are copies of each other and thus their properties are the same except for the mass, which increases with the generation index. However, it turns out that matter here on Earth is made up of fermions only from the first generation. Why there experimentally seems to be exactly three fermion generations remains an unresolved theoretical problem.

Leptons only interact via the electroweak interaction and are thus not charged under \( SU(3)_C \). The six leptons are the electron (\( e^- \)), muon (\( \mu^- \)), and tau (\( \tau^- \)), each with a corresponding neutrino (\( \nu_e \), \( \nu_\mu \), \( \nu_\tau \)). The electron, muon, and tau all have electric charge \( Q = -|e| \), whereas the neutrinos are electrically neutral. The three lepton generations are organized as

\[
\left( \begin{array}{c} \nu_e \\ e^- \\ \nu_\mu \\ \mu^- \\ \nu_\tau \\ \tau^- \end{array} \right).
\]

The quarks carry color charge, and thus come in three different varieties: red, green, and blue. The six quarks are the up (u), down (d), charm (c), strange (s), top (t), and bottom (b). The three generations of quarks are

\[
\left( \begin{array}{c} u \\ d \\ c \\ s \\ t \\ b \end{array} \right).
\]

The upper-level members of the quark doublets carry electric charge \( Q = +\frac{2}{3}|e| \), whereas the lower-level members carry charge \( Q = -\frac{1}{3}|e| \). The third generation is the heaviest and the top quark is the heaviest particle in the SM. The mass eigenstates and the weak eigenstates of the quarks are not the same. The quark mixing is given by the \( CKM matrix \), which normally is parametrized by three mixing
angles $\theta_{12}^q$, $\theta_{13}^q$, and $\theta_{23}^q$, and one CP-violating phase $\delta_{\text{CKM}}$. The mixing in the quark sector is small and the CKM matrix can be viewed as a small perturbation of the identity matrix.

The leptons can couple to the gauge fields in a gauge invariant manner under transformations of $SU(2)_L \otimes U(1)_Y$ and the part of the SM Lagrangian that contains the kinetic terms of the fermions and their interactions with the gauge bosons is given by

$$L_{\text{fermion}} = i \bar{\psi} \gamma^\mu D_\mu \psi.$$ (2.11)

In order to preserve gauge invariance, the ordinary derivative $\partial_\mu$ is replaced by a covariant derivative in an ordinary kinetic term for $\psi$. The covariant derivative is defined as

$$D_\mu = \partial_\mu + igA^a_\mu T^a,$$ (2.12)

where $A^a_\mu$ are gauge fields and $T^a_\mu$ are the generators of the representation of the gauge group that the fermion field belongs to. This change of derivatives will introduce gauge interaction terms of the form

$$L_{\text{fermion-gauge}} = -g \bar{\psi} \gamma^\mu A^a_\mu T^a \psi.$$ (2.13)

Thus, once the gauge structure and fermion representations are chosen, the gauge interactions of the fermions are completely determined.

Fermions are chiral, they can be either left- or right-handed, and the particles interact differently depending on their chirality. The fields are assigned to different representations of $SU(2)_L$, the left-handed fields form $SU(2)_L$ doublets, whereas the right-handed fields instead form a pair of $SU(2)_L$ singlets, which do not couple to the $SU(2)_L$ vector fields. Therefore, only left-handed fields interact via the weak interaction and the SM is a chiral theory. The entire particle content of one fermion generation is thus

$$L_L = \left( \begin{array}{c} \nu_i \\ l_i^- \end{array} \right)_L, \quad l_i^- R, \quad Q_L = \left( \begin{array}{c} u_i \\ d_i \end{array} \right)_L, \quad u_i R, \quad d_i R,$$ (2.14)

where $i \in \{1, 2, 3\}$ is the generation index. Note however that there are no right-handed neutrinos in the SM, which implies that neutrinos are massless. This will be discussed further in Sec. 2.2.6.

### 2.2.5 The Higgs sector and electroweak symmetry breaking

The Higgs field, $\Phi$, in the SM is a doublet under $SU(2)_L$, and transforms as a particle with hypercharge $Y = \frac{1}{2}$ under $U(1)_Y$. The Higgs sector in the SM is given by

$$L_{\text{Higgs}} = |D_\mu \Phi|^2 - V(\Phi),$$ (2.15)

with the covariant derivative defined in Eq. (2.12). The Higgs potential, $V(\Phi)$ is given by

$$V(\Phi) = -\mu^2 |\Phi|^2 + \frac{\lambda}{2} |\Phi|^4.$$ (2.16)
This is the most general renormalizable gauge invariant potential which can be written in the SM. Every term allowed by the symmetries of the model should be added to the model, and hence this full potential is the one that should be used.

Finally, the Higgs field’s interactions with the fermion fields are given by the Yukawa sector

\[
L_{\text{Yukawa}} = -Y_{ij}^u \bar{Q}_i L \Phi u_R - Y_{ij}^d \bar{Q}_i L \Phi d_R - Y_{ij}^e \bar{Q}_i L \Phi e_R + \text{h.c.,}
\]  

where \( Y_{ij} \) are the Yukawa matrices, \( \tilde{\Phi} = i\tau_2 \Phi^* \), \( \tau_2 \) is the second Pauli matrix, and \( i, j \in \{1, 2, 3\} \) are again the generation indices. The scalar field in the SM forms a two-dimensional, complex representation of two of the gauge groups.

The Brout-Englert-Higgs mechanism

Our world is not symmetric. Still, we do wish for the Lagrangian that describes our non-symmetrical world, to be symmetric. A central theme of particle physics is thus the study of how the symmetry of the Lagrangian can be broken in order to generate the world we see around us. Symmetries can always be broken "by hand", simply by the introduction of terms in the Lagrangian that do not respect the symmetry, which initially was there. This is however not very interesting from a theoretical point of view since it is in principle equivalent to having an asymmetry in the Lagrangian from the very beginning. Instead, a system which breaks the symmetry itself via a process called spontaneous symmetry breaking is of interest.

The mechanism through which the electroweak symmetry of the SM is broken i.e.

\[
SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \to SU(3)_C \otimes U(1)_Q
\]  

and the masses of the \( Z \) and \( W^\pm \) gauge boson masses are generated is known as the Brout-Englert-Higgs (BEH) mechanism. The Higgs potential has its minimum at

\[
|\Phi|_0 = \sqrt{\frac{\mu^2}{\lambda}} \equiv v \simeq 246 \text{ GeV.}
\]  

The Higgs field can thus be expanded around this minimum, which is known as the vacuum expectation value (vev), since the physical states are excitations around this point. Thus we can express the Higgs field as

\[
\Phi = \left( \frac{1}{\sqrt{2}} (h + i\phi + v) \right),
\]  

where \( h \) and \( \phi \) are real fields. We construct the real part, i.e. \( h + v \), in such a way that \( h \) is treated as an exitation around \( v \). The ground state of the field does not respect the \( SU(2)_L \) symmetry, since the value of the vev is non-zero. The symmetry has been spontaneously broken. When the Higgs field acquires a vev the
2.2. The Standard Model

Gauge bosons will obtain masses, since the gauge interacting part of Eq. (2.15) can be rewritten as

\[ L_{\text{mass, gauge}} = \frac{v^2}{8} \left[ g^2 (W_1^1)^2 + g^2 (W_2^2)^2 + (-gW_3^3 + g'B_\mu)^2 \right] , \]  
\tag{2.21}

where \( g' \) and \( g \) are the \( U(1)_Y \) and \( SU(2)_L \) gauge coupling constants respectively. The two components \( W^1 \) and \( W^2 \) obtain equal masses \( m_W = gv/2 \). The linear combination

\[ Z_\mu \equiv \frac{1}{\sqrt{g'^2 + g^2}} (-gW_3^3 + g'B_\mu) , \]  
\tag{2.22}

will obtain the mass \( m_Z = \sqrt{g'^2 + g^2}v/2 \). The orthogonal field combination

\[ A_\mu \equiv \frac{1}{\sqrt{g'^2 + g^2}} (g'W_3^3 + gB_\mu) , \]  
\tag{2.23}

corresponds to the unbroken direction of the gauge group, is massless and identified as the photon, which is the gauge boson of the gauge group \( U(1)_Q \).

Since three out of the four generator directions of \( SU(2)_L \otimes U(1)_Y \) are broken, one would normally expect that three of the four components of the Higgs field to become Goldstone bosons \[36, 39\]. However, as was shown in 1964, when a gauge symmetry is broken, the extra degrees of freedom instead can become longitudinal polarizations of the gauge bosons. We say that the components are eaten by the now massive gauge boson fields. Thus only one component of the Higgs field remains as a physical field, corresponding to the new scalar particle known as the Higgs boson. It is worth noting that the total number of degrees of freedom is left unchanged by the BEH mechanism, since a massive gauge boson has one more degree of freedom than a massless one.

The matrix which describes the mixing of the \( W^3 \) and \( B \) fields into the \( Z \) and \( A \) fields is given by

\[ \left( \begin{array}{cc} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{array} \right) , \]

where the \( Weinberg\ angle \), or weak mixing angle, is defined as \( \tan \theta_W = g'/g \). It also gives the relationship between the masses as \( m_W = \frac{m_Z^2 \cos \theta_W}{\cos \theta_W} \). The fields \( W^1 \) and \( W^2 \) are rotated into the charged, massive fields \( W^\pm \), defined as

\[ W_\mu^\pm = \frac{1}{\sqrt{2}} (W_1^\mu \mp iW_2^\mu) . \]  
\tag{2.24}

Rather by coincidence the fermions also acquire their masses from the BEH mechanism. Similarly to the gauge boson case, mass terms of the kind \(-m\bar{\psi}_L\psi_R\) are not allowed directly in the Lagrangian as they violate gauge symmetry. However at EWSB the terms in the Yukawa sector will generically take the following form

\[ L_{\text{mass, fermion}} = m^{ij}\bar{\psi}_L^i\psi^j_R , \]  
\tag{2.25}

where the fermion mass matrix now is given by \( m^{ij} = \frac{1}{\sqrt{2}} Y^{ij}v \). Note that the mass matrix is not diagonal in general.
Chapter 2. The Standard Model and slightly beyond

2.2.6 Neutrinos: mixings and masses

There are no right-handed neutrinos in the SM. Therefore the neutrinos cannot form Yukawa couplings, and thus not obtain their mass through the BEH mechanism like the other fermions. Local gauge invariance of the SM Lagrangian forbids any other type of mass terms and the neutrinos must therefore be massless in the SM framework. However experimental evidence points to the contrary, since the neutrinos oscillate, i.e. change flavor as they travel through space, a phenomenon which only can occur if the neutrinos are massive. Intuitively one can think of this in the following way. Had the neutrinos been massless they would have travelled at the speed of light. When travelling at the speed of light, different eigenstates will not propagate at different velocities, which will change the interference pattern.

Neutrino mass models

The charged leptons in the SM, which obtain their mass from the BEH mechanism, have so-called Dirac masses, which are on the form

\[ \mathcal{L}_{\text{Dirac}} = m_D (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L), \]  

(2.26)

where \( m_D \) is the Dirac mass. This obviously involves both left- and right-handed components of the fermion field.

There is however another option for the neutrino. Since it is an electrically neutral particle, it is in theory possible that it is its own anti-particle. Such spin-1/2 fermions which are their own anti-particle are known as Majorana fermions [40]. Mathematically this can be written as \( \psi_R = (\psi_L)^c \), where \( c \) denotes charge conjugation. Through this relation we can conclude that the left- and right-handed components of the fields are related and there is only one chiral component for the Majorana particle, whereas the Dirac fermions have two. The Dirac or Majorana nature of the neutrinos is one of the most discussed issues in neutrino physics since it is of great importance for the mechanism for neutrino mass generation. The Majorana mass term is given by

\[ \mathcal{L}_{\text{Majorana}} = -\frac{1}{2} (\bar{\psi}^c M \psi + \text{h.c.}), \]  

(2.27)

where \( M \) is the mass matrix. This term changes a particle to its antiparticle, violating the \( B - L \) symmetry, where \( B \) is baryon number and \( L \) lepton number. This symmetry is accidentally conserved in the SM.

The mechanism for generating neutrino masses must necessarily lie beyond the SM. The simplest way of giving the neutrino mass is by a dimension-5 Weinberg operator [41]

\[ -\mathcal{L}^{d=5}_{\text{Weinberg}} = \frac{(\bar{L}_L \phi) \kappa (\phi^T L_L^c)}{2} + \text{h.c.}, \]  

(2.28)

where \( \kappa \) is a complex, symmetric \( 3 \times 3 \) matrix, inversely proportional to the energy scale of new physics. This operator is in fact the only operator of dimension five
that is allowed by SM symmetries and neither the SM nor dimension-six operators can do the job of generating neutrino masses [42]. With the addition of this higher-order operator, the SM will no longer be renormalizable. The neutrino masses are generated at EWSB and are given by

\[
-\mathcal{L}_{\text{Majorana}} = \frac{1}{2} \nu_L \nu_L^c + \text{h.c.},
\]

where the neutrino mass matrix is given by

\[
m_{\nu} = \kappa v^2.
\]

**Seesaw mechanism**

The Majorana mass matrix in Eq. (2.30) can be generated in a group of models called *seesaw models*, where heavy degrees of freedom are introduced at a high energy scale and then integrated out. The requirement on \( B - L \) symmetry is renounced. There are three main types of seesaw models, and in addition several models which are combinations thereof. In the type I seesaw model a number of fermionic SM singlets, typically right-handed neutrinos, are added to the SM [43–46]. In the type II seesaw, the scalar sector is enlarged by a \( SU(2)_L \) triplet field [47–49]. Lepton number will then be violated by the triplet field’s interaction with both the lepton and Higgs doublets. Finally, in a type III seesaw model a fermionic \( SU(2)_L \) triplet is added, which violates lepton number through its Majorana mass term [50]. We shall discuss the type I seesaw model in some more detail.

In the type I seesaw the SM particle content is extended by three heavy right-handed neutrino fields, \( N_R \), with masses well above the electroweak scale. A Yukawa term as well as a Majorana mass term for the right-handed fields are added and the leptonic sector of the Lagrangian for one generation is then given by

\[
-\mathcal{L}_{\text{Lepton}} = \bar{L}_L Y_\ell \Phi_R + \bar{L}_R Y_\nu \tilde{\Phi} N_R + \frac{1}{2} N_R^C M_R N_R + \text{h.c.},
\]

where \( M_R \) is the symmetric Majorana mass matrix. After the heavy neutrinos have been integrated out and spontaneous symmetry breaking a Majorana mass term is obtained for the light neutrinos, with the Majorana mass matrix given as

\[
m_{\nu} = -\frac{1}{2} Y_\nu \frac{v^2}{M_R} Y_\nu^T.
\]

The smallness of the SM neutrinos can now be easily attributed to the large \( M_R \).

**Neutrino oscillations**

The eigenstates of the neutrinos in the mass basis and in the flavor basis are not the same. Neutrinos propagating through space change flavor. In vacuum however,
the mass eigenstate cannot change. Thus, the two states cannot be the same. The flavor eigenstates can thus be expressed in terms of the mass eigenstate as
\[
|\nu_\alpha\rangle = U^*_{\alpha i} |\nu_i\rangle, \quad (2.33)
\]
\[
|\nu_i\rangle = U_{\alpha i} |\nu_\alpha\rangle, \quad (2.34)
\]
where \( \alpha \in \{e, \mu, \tau\} \) and \( i \in \{1, 2, 3\} \). The mixing is described by the unitary matrix \( U_{\text{PMNS}} \), called the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [51, 52]. The PMNS matrix can be parametrized, in similarity with the CKM matrix, in terms of three mixing angles, \( \theta_{12}, \theta_{13}, \) and \( \theta_{23} \); a CP-violating phase, \( \delta \); and two Majorana phases \( \rho \), and \( \sigma \), that are only non-zero if the neutrinos are Majorana fermions. Thus the matrix is given by
\[
U_{\text{PMNS}} = \begin{pmatrix}
c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\
-s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & -c_{23}s_{13}e^{i\delta} \\
s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}s_{13}
\end{pmatrix}
\begin{pmatrix}
e^{i\rho} & 0 & 0 \\
0 & e^{i\sigma} & 0 \\
0 & 0 & 1
\end{pmatrix},
\]
where \( s_{ij} = \sin(\theta_{ij}) \) and \( c_{ij} = \cos(\theta_{ij}) \), \( ij = \{12, 13, 23\} \).

### 2.2.7 Problems with the SM

Apart from the experimental problems with the SM discussed in the introduction, there are problems of a more conceptual nature. The conventional wisdom of the past 30 years in the HEP community is that there is a hierarchy or naturalness problem in the SM. It boils down to the question of why the ratio of the W and Z boson mass and the Planck mass is so small, \( M_Z/M_{\text{Planck}} \approx 10^{-16} \). A problem, which has to do with the value of the non-zero Higgs field, that determines the boson masses. This problem is not one of consistency but of naturalness and fine-tuning, which needs explaining. Many theories, such as supersymmetry, technicolor, GUT theories, and string theory, try to address this problem. A common denominator of these theories is the need for new physics at the TeV scale, which in turn provides a strong motivation for the LHC. So far, however, there are no indications of this new physics in the data.

The SM is not fully determined from theory but contains 18 free parameters (plus another nine if neutrinos with non-zero mass is included), which have to be determined experimentally and put into the theory by hand. There is no way of knowing their values directly. This obvious shortcoming of the theory does heighten physicists beliefs that there is a more general theory where every parameter is unambiguously determined as a prediction of the theory. So far, however, none of the proposed extensions to the SM has provided a satisfying solution to this issue. On the contrary, supersymmetry, one of the most popular SM extensions, include another hundred or so parameters.
Chapter 3

Models beyond the SM

There are plenty of possible models going beyond the SM. Here we shall discuss three of them. The first model is an extra-dimensional theory with Universal Extra Dimensions (UEDs), which are effectively manifest in terms of heavy particles in our four-dimensional world. This model is discussed in Ref. [1]. Then, we shall discuss a non-supersymmetric SO(10) GUT model, which is studied in Ref. [2]. Within such a model the SM is considered an EFT. Finally, we consider a generic effective model, studied in Ref. [3], where the magnitude of the couplings in the SM Higgs sector is rescaled. In Refs. [1, 2] we investigate neutrino masses generated by the Weinberg operator.

3.1 Effective field theories

Our physical world seems to contain interesting physics across all energy scales, spanning from the age of the Universe, to the lifetime of a W boson. Typically, however, one is only interested in physics at a particular scale. Therefore, it is convenient to isolate physics relevant at one scale from physics relevant at other scales, using different descriptions of physics at different scales. Hence, instead of using the full underlying theory, which might not even be known, it can be useful to study an EFT. Finite effects of parameters at other scales are then included as perturbations to the parameters relevant at the given scale. The effective theory might be computationally convenient and more physically intuitive, even in the case when the full theory is known. In particle physics, where the SM is often considered to be a low-energy effective theory, an EFT framework allows us to study phenomena at energies which are relevant here on Earth, independently of physics at for instance the GUT scale or the scale of a Theory of Everything [53, 54].

In the “top-down” approach the full physical theory is known and heavy degrees of freedom are eliminated from the theory in a step-by-step procedure. This

\^1Here we shall only be considering effective quantum field theories.
will naturally simplify the theory, even though this procedure is non-trivial due to
the necessity of ultraviolet regularization, which requires that the limit when the
heavy particles are eliminated is treated carefully. These heavy degrees of freedom
will thus be manifest in the low-energy EFT in terms of effective operators of mass
dimension greater than 4. Such operators are non-renormalizable. Hence the in-
formation from high energies, or equivalently small distances, will be encoded in a
few parameters used to describe the dynamics at low energies, or equivalently large
distances [55]. An EFT is thus only valid up to a given cut-off scale, Λ, and the
effective theory is described by operators containing only the light fields, i.e. fields
with mass below Λ [54]. In principle the region of validity of the effective theory is
bounded from below as well as from above. In the energy domain below the mass
of the heaviest particle in the effective theory, it is rather convenient to replace the
previous EFT with a new one, where these heavy particles have once again been
removed. Thus, pushing this argument to the extreme, the mass of every particle
in the theory marks a boundary between two effective theories.

An EFT can be constructed even when the full physical theory is not known.
The physics at the present scale can still be described by the EFT. In this case,
however, the model must be built from scratch and the interesting physics must
be put in by hand. This is known as the “bottom-up” approach and the choice of
operators is motivated by criteria such as naturalness, symmetries of nature, and
interactions assumed to be of interest.

Non-renormalizability

The SM is a renormalizable quantum field theory, which often is argued by theorists
to be an important feature of the theory. The notion of renormalizability is that
all divergences that appear in a theory should be possible to cancel with a finite
number of counterterms. This can be shown to be equivalent to having coupling
constants with only non-negative mass dimensions. After fixing the free parameters
of a renormalizable QFT, any observable can be predicted at any scale to any order
of perturbation. Thus a renormalizable theory is independent of the energy scale
whereas the non-renormalizable is not. We shall discuss the concept of renormal-
ization in more depth in Ch. 4. However, the modern view on renormalizability
of a theory is that it is not necessary as long as the limitations of the theory are
properly taken into account. Therefore EFTs, which are non-renormalizable, are
considered equally valid as a fully renormalizable theory, as long as they are only
applied within their region of validity. At some energy the EFT breaks down, which
is referred to as the cutoff energy, Λ. Above this energy, the EFT is replaced by a
more fundamental theory, a UV completion.

Effective Lagrangian

Independent of whether the EFT is top-down or bottom-up the Lagrangian will
be given as an infinite sum of local operators. In general the low-energy effective
Lagrangian can be expanded in operator dimensions and thus it can be written as

\[ \mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \mathcal{L}_{d=5} + \mathcal{L}_{d=6} + \ldots \]  

(3.1)

Here \( \mathcal{L}_{\text{SM}} \) is the SM Lagrangian, \( \mathcal{L}_{d=5} \) contains the the one dimension-5 operator which is allowed by SM symmetries, the Weinberg operator, which gives rise to neutrino mass. The term \( \mathcal{L}_{d=6} \) includes some operators which modify the Higgs couplings. The number of terms in the Lagrangian is infinite, however, the theory still have an approximate predictive power. The couplings of the term \( \mathcal{L}_{d=di} \) are of the form \( c_i \cdot \Lambda^{4-d_i} \), where \( c_i \)'s are a dimensionless constants and \( \Lambda \) is the scale of new physics. These terms obviously have couplings with negative mass dimension and are therefore not renormalizable. At a given energy scale \( E \) the contribution from the term will be of the order \( (E/\Lambda)^{d_i-4} \). Therefore at \( E < \Lambda \) the contribution will be smaller in size the higher the dimensionality.

### 3.2 Universal Extra Dimensions

One of the simplest ways to extend the SM is by adding extra spatial dimensions. The first to explore the idea of such extra dimensions were Theodor Kaluza and Oskar Klein, in the beginning of the 1920s, as an effort to unify electromagnetism and gravity. The idea was to introduce one extra spatial dimension, which was both circular and finite in size, *i.e.* compactified. Nowadays, theories similar to this first theory bear their name and are normally referred to as Kaluza–Klein (KK) theories. There are several different types of extra dimensional models. The extra dimensional model chanced upon most often in popular science, *string theory*, is a theory of extended objects. Here we will only consider an essentially different type of extra-dimensional theory, namely extra-dimensional QFTs. In this category there is a plethora of popular models. The first example of such a model is large extra dimensions or the Arkani-Hamed–Dimopoulos–Dvali (ADD) model [56, 57]. In this model all SM fields are confined to a four-dimensional brane, whereas gravity is allowed to move in one or several additional extra dimensions, which possibly would explain the weakness of gravity compared to the other forces in Nature. The extra dimensions in this model are referred to as large, since they are large compared to the Planck scale. This model has been extensively studied by the experimental collaborations at the LHC, and there are at present rather severe limits on it [58, 59].

A second example is the Randall–Sundrum (RS) model [60]. In this model spacetime is extended with one extra dimension on \( S^1/Z_2 \). Space is extremely warped and contains two branes, the Planck brane, where gravity is relatively strong, and the TeV-brane, where our SM particles live. This model provides a solution to the hierarchy problem.

In the present work we will discuss a model with UEDs, introduced by Appelquist, Cheng, and Dobrescu in 2000 [61]. Here, all SM fields are propagating in the extra dimensions. However, macroscopic objects, such as ourselves, are not
Chapter 3. Models beyond the SM

able to probe the extra dimensions. Instead we will perceive the extra dimensions as towers of heavy particles corresponding to each SM one. Furthermore, extradimensional theories are necessarily non-renormalizable, independent of the the size and shape of the extra dimensions. The model will therefore only be valid in the region up to a cutoff, $\Lambda$, which will limit the number of modes in these particle towers that are allowed to contribute.

The simplest UED model has naturally only one extra dimension, compactified on a circle, corresponding to a $S^1/Z_2$ orbifold. This model has been extensively studied in the literature [62–66]. In this thesis we shall especially study a minimal six-dimensional UED model [67, 68]. In both the five- and six-dimensional models there is a stable dark matter candidate as well as a mechanism for electroweak symmetry breaking [69–71]. The six-dimensional UED model also requires that the number of fermion generations is a multiple of three in order to obtain anomaly cancellations, and proton stability is ensured [72, 73].

3.2.1 Kaluza-Klein decomposition

The obvious caveat regarding theories with extra dimensions is naturally that no extra spatial dimensions have been detected. Thus the extra dimensions cannot be on the same footing as the other ones. The most intuitive way is to make them small, and finite, i.e. compact. The extra dimensions could then only be probed by particles with small enough wavelength, or equivalently, high enough energy. To macroscopic objects they would then effectively be hidden. Instead, in our four-dimensional world the modes of the fields propagating through the extra dimensions will effectively be perceived as towers of KK modes, corresponding to each of the SM fields. Without any loss of generality we can make an expansion of the fields in the extra, compactified dimensions. A concept known as KK decomposition. Thus we are effectively treating the extra-dimensional theory as a four-dimensional theory where the extra dimensions instead are manifest in the structure of the particles and their interactions. This concept is relevant for all the extra-dimensional QFTs and not only in the context of UEDs.

In order to illustrate this concept more qualitatively we shall study a toy model with a complex scalar particle in five dimensions. The fifth dimension is compactified on a circle with a radius $R$. Since the fifth dimension is a circle we know that the scalar field is periodic in the parameter $x^4$. Thus the scalar field must fulfill

$$\phi(x^{\mu}, x^4) = \phi(x^{\mu}, x^4 + 2\pi R), \quad (3.2)$$

where $\mu \in \{0, 1, 2, 3\}$. Especially this implies that we can expand this field in a Fourier series, without any loss of generality, we have

$$\phi(x^{\mu}, x^4) = \frac{1}{\sqrt{2\pi R}} \sum_{n=-\infty}^{\infty} \phi^{(n)}(x^{\mu}) e^{in\pi x^4}. \quad (3.3)$$
The action of this model is given by
\[ S = \int d^5(\partial_M \phi^* \partial^M \phi - m^2 |\phi|^2) \] (3.4)

and explicitly inserting this expression into Eq. (3.3) we obtain the following expression for the action
\[ S = \int d^4(\partial_M \phi^{(n)*} \partial^M \phi^{(n)} - (m^2 + \frac{n^2}{R^2} |\phi^{(n)}|^2)), \] (3.5)

where \( M \in \{0, 1, \ldots, 4\} \). This now represents a four-dimensional theory with an infinite number of fields, a KK tower corresponding to the SM field. The number \( n \) is called the KK number and each of these modes has a KK mass given by
\[ m_n = \sqrt{m^2 + M_n^2} = \sqrt{m^2 + \frac{n^2}{R^2}}, \] (3.6)

Within each of these towers the particles only differ in mass, otherwise the physical properties are the same. The lowest order field, i.e. the field with \( n = 0 \) has no momentum along the extra dimension and corresponds to the SM particle. Thus the extra dimensions will effectively be manifest in terms of heavy particles.

As previously discussed these extra dimensions will only be accessible to particles of large enough energies, i.e. energies larger than \( R^{-1} \), which is the threshold energy. In principle this means that the higher the energy available the larger the number of modes which can be observed. This means that such extra dimensions should be possible to probe for instance in the LHC, if it reaches a high-enough energy.

So far we have only considered a toy model with a single scalar particle in a five-dimensional model. The generalization to other fields and internal spaces is in principle straightforward. In the next section we shall discuss a six-dimensional UED model with flat extra dimensions.

### 3.2.2 Six-dimensional UED model

As we discussed in Chap. 2, fermions have definite chirality in four-dimensional spacetime. We require the that the six-dimensional theory reproduce the observed four-dimensional chirality when the two extra dimensions are integrated out. Thus the compactification of the two extra dimensions must be made on an orbifold in accordance with this requirement. The simplest choice which matches the criterion of four-dimensional chirality is the chiral square [67]. This compactification is equivalent to the orbifold \( T^2/\mathbb{Z}_4 \), where \( T^2 \) is a torus and \( \mathbb{Z}_4 \) corresponds to a 90 degree rotational invariance. This topology corresponds to a sphere with conical singularities in \((0,0), (L,L)\), and the two identified points \((L,0) \sim (0,L)\), where there in principle could be localized kinetic terms, however these are not considered here. The compactification is thus made on the square which is illustrated in
Chapter 3. Models beyond the SM

Figure 3.1. The chiral square which is used for compactification of the extra dimensions. The sides with thick lines are identified with each other, and the same is valid for the sides with double lines. Figure taken from Ref. [67].

Fig. 3.1, where the sides are of length $L$ and the extra-dimensional coordinates take on the values $0 < x_4, x_5 < L$. The adjacent sides are identified according to

$$ (x^\mu, 0, y) = (x^\mu, y, 0), $$
$$ (x^\mu, L, y) = (x^\mu, y, L), \tag{3.7} $$

where $\mu = 0, 1, 2, 3$. Furthermore, we require that the physics at identified points is equal. This can be done by requiring that the Lagrangian in the identified points have the same values independent of the field configuration at these points, i.e. that it fulfills

$$ \mathcal{L}(x^\mu, 0, y) = \mathcal{L}(x^\mu, y, 0), $$
$$ \mathcal{L}(x^\mu, L, y) = \mathcal{L}(x^\mu, y, L). \tag{3.8} $$

In addition, there is a six-dimensional chirality which, however, is not the same as the four-dimensional one. In general, spinors in an even number of dimensions, $D = 2k$, are $2^k$-component objects [74]. Hence, the spinors in six dimensions have eight components. The six-dimensional chiral representations are referred to as having $+/-$ chirality, compared to the left- and right-handed chiralities in four dimensions. These $+/-$-chirality representations are four-component objects since they carry half of the number of degrees of freedom of the Dirac representations. The zero mode components of the six-dimensional fields, i.e. the fields with zero
momentum in the extra dimensions, are identified with the SM particles. From the definite chirality of the four-dimensional fermions we can make restrictions on the six-dimensional fields which do have a zero mode component. However, as will be discussed in the subsequent section, in this six-dimensional model not all KK towers have a zero mode component.

**Kaluza-Klein decomposition in the six-dimensional UED model**

The concept of KK decomposition was introduced in Sec. 3.2.1. We shall now proceed to discuss the concept in the context of the six-dimensional UED model. We shall here only consider a gauge field, but the decomposition is done in a similar way for fermionic and scalar field. Abelian gauge fields, $A^\alpha(x^\beta)$, where $\alpha, \beta \in 0, 1, \ldots, 5$, i.e. fields with six components, which are allowed to propagate in the full six dimensions of the theory. We can write the fields as

$$A^\alpha(x^\beta) = (A^\mu(x^\nu, x^4, x^5), A^4(x^\nu, x^4, x^5), A^5(x^\nu, x^4, x^5)),$$

where $x^\mu$ and $x^\nu$ are the Minkowski spacetime coordinates, with $\mu, \nu \in 0, 1, 2, 3$. The decomposition can now be made in a complete set of KK functions, $f_n$, which are defined in [68]

$$f^{(j,k)}_n(x^4, x^5) = \frac{1}{1 + \delta_{j,0} \delta_{k,0}} \left[ e^{-i n \pi/2} \cos \left( \frac{j x^4 + k x^5}{L} + \frac{n}{2} \right) \right]$$

where $j, k$, where $k \geq 0, j \geq 0$, are the integers which label the KK modes. Conservation of momentum in the extra dimensions will be translated to conservation of KK number at vertices at tree level in the effective theory. Only the functions with $n = 0$ have a zero mode. Especially, this implies that only functions which are decomposed in this function can have a zero mode, corresponding to the SM mode. The functions, $f_n$, fulfill the six-dimensional Klein-Gordon equation

$$(\partial_4^2 + \partial_5^2 + M_{j,k}^2) f_n = 0,$$

with the KK mass defined as

$$M_{j,k}^2 = \frac{j^2 + k^2}{R^2}.$$ 

Furthermore the functions $f_n$ are normalized over $x_4, x_5$. We introduce $A_\pm = A_4 \pm i A_5$ and can now make the decomposition of the six-dimensional gauge field
as

\[
A_\mu(x^\nu, x^4, x^5) = \frac{1}{L} \left( A_\mu^{(0,0)}(x^\nu) + \sum_{j \geq 1} \sum_{k \geq 0} f_0^{(j,k)}(x^4, x^5) A_\mu^{(j,k)}(x^\nu) \right), \tag{3.13}
\]

\[
A_+(x^\nu, x^4, x^5) = -\frac{1}{L} \sum_{j \geq 1} \sum_{k \geq 0} f_3^{(j,k)} A_+^{(j,k)}(x^\nu), \tag{3.14}
\]

\[
A_-(x^\nu, x^4, x^5) = \frac{1}{L} \sum_{j \geq 1} \sum_{k \geq 0} f_1^{(j,k)} A_-^{(j,k)}(x^\nu). \tag{3.15}
\]

The mode with \((j, k) = (0, 0)\), corresponds to the SM fields. All other values of \(j, k\) correspond to KK modes. Furthermore, note that there is no \((0, 1)\) mode and therefore the lowest mode is \((1, 0)\), with mass \(1/R\). In the four-dimensional theory the six-dimensional gauge field will be perceived as one KK tower of heavy spin-1 modes, and two towers of spin-0 modes. Note that the spin-0 tower lack a zero mode component. Thus it does not directly correspond to a SM field and will provide additional degrees of freedom at each KK level. We can redefine the spin-0 fields into two real scalars, \(A^{(j,k)}_H\) and \(A^{(j,k)}_G\), at each KK level as

\[
A_{\pm}^{(j,k)} = r_{j, \pm k} (A_H^{(j,k)} \mp i A_G^{(j,k)}), \tag{3.16}
\]

with complex phases defined as \(r_{j, \pm k} = (j \pm ik)/\sqrt{j^2 + k^2}\). At each KK level \(A_G\) can be eaten by the KK mode corresponding to the four-dimensional particle. There will however be an additional adjoint scalar \(A_H^{j,k}\) at each KK level, corresponding to the gauge bosons.

In order to ensure that the lightest KK particle is a viable dark matter candidate a KK parity, defined as \((-1)^{j+k}\), can be imposed. This \(\mathbb{Z}_2\) symmetry will then distinguish among the KK modes acting according to

\[
\Phi^{(j,k)}(x^\mu) \rightarrow (-1)^{j+k}\Phi^{(j,k)}(x^\mu), \tag{3.17}
\]

where \(\Phi\) is a field of any spin, and the indices \(j, k\) are integers labeling the KK levels [67]. An important effect is that conservation of momenta in the extra dimensions will be translated to conservation of KK number in the tree level vertices.

### 3.3 A non-supersymmetric \(SO(10)\) grand unified model

In Ref. [2] we study a non-supersymmetric GUT model based on the gauge group \(SO(10)\). In this model an intermediate scale is introduced, which is necessary in non-supersymmetric models in order to be able to break \(SO(10)\) down to the SM gauge group [75]. In this section we briefly discuss GUT models in general and the non-supersymmetric \(SO(10)\) model in particular.
3.3. A non-supersymmetric $SO(10)$ grand unified model

3.3.1 A short introduction to Grand Unified Theories

In a GUT all fundamental interactions in the SM are considered having a common origin, characterized by one larger gauge group with several force carriers but only one, single coupling constant. The idea that all the gauge interactions of the SM are unified into one single interaction an energy scale, the GUT scale, is inspired by the electroweak unification. The gauge group must necessarily contain the SM to be able to reproduce EW data. The smallest simple Lie group which can accommodate the SM is $SU(5)$. The first GUT model, proposed by Howard Georgi and Sheldon Glashow in 1974, was based on this group [76]. In $SU(5)$ the SM fields of one generation belong to two irreducible representations of the group, with $10 = [Q_L, u_R^c, e_R^c]$ and $\bar{5} = [d_R^c, L_L^c]$. In addition, a Higgs doublet is needed for electroweak symmetry breaking, which can sit either in a $5_H$ or a $\bar{5}_H$. The additional three states in the quintuples are color triplet Higgs scalars. The masses of these color triplets are bounded by the non-observation of nucleon decay.

The next group of choice is $SO(10)$, discovered by Howard Georgi in 1974, where all SM particles are in a 16-dimensional spinor representation, including the right-handed neutrino which will naturally give rise to a seesaw mechanism and small neutrino masses [77]. Due to the incorporation of all SM particles and the accommodation of neutrino masses $SO(10)$ has become the canonical choice for GUTs [78]. There are two inequivalent maximal breaking patterns of $SO(10)$: $SO(10) \rightarrow SU(5) \times U(1)_X$ and $SO(10) \rightarrow SU(4)_C \times SU(2)_L \times SU(2)_R$. The first gives rise either to the Georgi-Glashow $SU(5)$ or so-called flipped $SU(5)$, depending on whether or not $Q_{EM}$ is partly in $U(1)_X$ or not. The second breaking pattern has a Pati-Salam (PS) symmetry. The PS group was the GUT model, proposed by Abdus Salam and Jogesh Pati in 1974 as an alternative to the $SU(5)$ unification, where lepton flavor was embedded in the theory as a fourth color, lilac [79]. In addition to the models mentioned here there are an abundance of similar models as well as models based on other gauge groups.

3.3.2 A non-supersymmetric $SO(10)$ model

GUTs are often assumed to be supersymmetric, which is well motivated since supersymmetry provides a solution to the gauge hierarchy problem. However, in $SO(10)$, no supersymmetry is needed for gauge coupling unification, as long as there is an intermediate scale, $M_I$ present somewhere in between the electroweak scale, $M_Z$, and the GUT scale, $M_{GUT}$. In the present thesis, we study an $SO(10)$ GUT model which breaks down to the SM via the PS group. The breaking to this group occurs at an intermediate scale $M_I \sim 10^{11}$ GeV. In addition we will assume that the neutrinos obtain their mass through a type I seesaw, discussed in Sec. 2.2.6. For simplicity, the seesaw scale is assumed to coincide with the intermediate scale.
Higgs representations

Even though the SM fits nicely into the unifying gauge groups, the Higgs sector must necessarily be enlarged in order to break the enhanced gauge group down to the SM. The Higgs representations must be chosen in such a way that the desired breaking chain is obtained. Motivated by phenomenological arguments of tractability and predictivity the Higgs sector should also be chosen in a minimal way. This leads us to the following breaking chain \[75, 80, 81\]

\[
SO(10) \xrightarrow{M_{GUT}} 210 \xrightarrow{126} 4C_2L_2R \xrightarrow{M_{I}} 3C_2L_1Y \xrightarrow{M_{GUT}} 3C_1Y. \tag{3.18}
\]

Thus the gauge group at the GUT scale is broken by a \(210_H\), which gives suitable values of \(M_{GUT}\) and \(M_I\). The PS group is broken by a \(126_H\) and finally, EW symmetry breaking is made by a \(10_H\). In principle a \(45_H\) could be added in the second step in order to ensure the existence of an axion dark matter candidate, however this does not effect the behavior of the lepton sector and can therefore be excluded in the present discussion. Both the \(126_H\) and the \(10_H\) are necessary for the existence of fermion masses.

The decomposition of the Higgs representations under the PS group is given by

\[
\begin{align*}
10_H & = (1, 2, 2) \oplus (6, 1, 1), \\
16 & = (4, 2, 1) \oplus (4, 1, 2), \\
\overline{126}_H & = (6, 1, 1) \oplus (10, 1, 3) \oplus (\overline{10}, 3, 1) \oplus (15, 2, 2), \\
210_H & = (1, 1, 1) \oplus (15, 1, 3) \oplus (15, 1, 1) \oplus (15, 3, 1) \oplus (\overline{10}, 2, 2) \oplus 6, 2, 2, \\
\end{align*}
\]

and we shall introduce a shorthand notation of the vevs

\[
\begin{align*}
k_{u,d} & = \langle (1, 2, 2)_{u,d} \rangle_{10}, \\
v_R & = \langle (10, 1, 3) \rangle_{126}, \\
v_{u,d} & = \langle (15, 2, 2)_{u,d} \rangle_{126}, \tag{3.20}
\end{align*}
\]

where \(v_R\) is needed to generate a right-handed neutrino mass matrix and is at \(M_I\) whereas \(v_{u,d}\) contribute to fermion mass matrices and must be at the EW scale.

Gauge coupling evolution

The gauge couplings unify at the GUT scale, per definition, in these types of models. Thus the values of the GUT scale, \(M_{GUT}\), and the intermediate scale, \(M_I\), are given from the condition of unification. The evolution of the gauge couplings between two scales \(M_1\) and \(M_2\) is given by the formula

\[
\alpha_i^{-1}(M_2) = \alpha_i^{-1}(M_1) - \frac{a_i}{2\pi} \log \frac{M_2}{M_1}, \tag{3.21}
\]

where the coefficients \(a_i\) are group theoretical quantities which, for a generic \(SU(N)\), can be written as

\[
a_i = \frac{4}{3} n_g - \frac{11}{3} N + \frac{1}{3} \eta S_2(R_p), \tag{3.22}
\]
3.3. A non-supersymmetric $SO(10)$ grand unified model

Figure 3.2. The running of the gauge couplings in a non-supersymmetric $SO(10)$-model. Figure taken from Ref. [75].

where $n_g$ is the number of fermion generations, $N$ is the dimension of the group $SU(N)$, and $\eta = 1, 1/2$ for complex and real scalar fields, respectively. Given a scalar field $S$, which transforms according to the representation $R = R_1 \otimes \ldots \otimes R_{N'}$, where $R_p$ is an irreducible representation of the group $G_p$ of dimension $d(R_p)$, the factor $S_2(R_p)$ is defined in e.g. Ref. [82] as

$$S_2(R_p) \equiv T(R_p) \frac{d(R)}{d(R_p)}, \quad (3.23)$$

where $T(R_p)$ is the Dynkin index of $R_p$ [83].

Together with the evolution equations are matching conditions needed at the intermediate scale. These are given in Ref. [84]. The running of the gauge couplings in this model is presented in Fig. 3.2.

Yukawa sector

The evolution is performed from the GUT scale down to the EW scale. We are therefore interested in the $SO(10)$ Yukawa sector at the GUT scale. The masses of fermions, which belong to the spinorial 16 representation of $SO(10)$, arise from couplings of Higgs fields which belong to

$$\overline{16} \otimes \overline{16} = 10 \oplus 120 \oplus \overline{126}, \quad (3.24)$$

where 10 and $\overline{126}$ are symmetric and 120 anti-symmetric. Thus at renormalizable level there are only three possible choices for the Higgs representation, i.e. $10_H$, $120_H$, and $\overline{126}_H$. The running of the gauge couplings in this model is presented in Fig. 3.2.
120_H, and \( \overline{126}_H \) and the most general Lagrangian for these couplings are given on the form

\[
16_F(Y_{10}10_H + Y_{120}120_H + Y_{126}\overline{126}_H)16_F,
\]

(3.25)

where \( 16_F \) is the matter field, \( Y_{10} \) and \( Y_{126} \) are complex symmetric matrices, and \( Y_{120} \) is a complex antisymmetric matrix. As previously discussed, the most economical choice for the Yukawa sector, which do accommodate all known low-energy data, contains the Higgs representations 126_H and 10_H. The Yukawa sector at the GUT scale is then

\[
\mathcal{L}_Y = 16_F(h10_H + f\overline{126}_H)16_F + \text{h.c.},
\]

(3.26)

where \( h \) and \( f \) are complex and symmetric coupling matrices.

The representation 10_H cannot be real, since this would imply that the ratio \( m_t/m_b \) would be of order one. Thus, as in the case of supersymmetry, it is necessary to complexify 10_H. However, taking the 10_H complex the real and imaginary parts can couple differently to the 16_F, which in turn will lead to two independent Yukawa matrices. This can be dealt with through the introduction of a Peccei-Quinn symmetry, \( U(1)_{PQ} \) given by [75, 81]

\[
16_F \rightarrow e^{i\alpha}16_F, \quad \overline{126}_H \rightarrow e^{-2i\alpha}\overline{126}_H, \quad 10_H \rightarrow e^{-2i\alpha}10_H.
\]

(3.27)

In general the matching condition at the GUT scale for the Yukawa matrices is in the minimal case given by

\[
Y_u = r(H + sF),
\]

(3.28)

\[
Y_d = H + F,
\]

(3.29)

\[
Y_d = r(H - 3sF),
\]

(3.30)

\[
Y_e = H - 3F,
\]

(3.31)

\[
M_R = r_R^{-1}F,
\]

(3.32)

where \( H \) and \( F \) correspond to \( Y_{10} \) and \( Y_{126} \), respectively, i.e. \( H \) and \( F \) are symmetric. The parameter \( s \) is complex whereas \( r \) and \( r_R \) can be chosen to be real without any loss of generality. These parameters can be expressed in terms of the vevs introduced in Eq. (3.20) as \( H = h k_d \), \( F = f v_d \), \( r = k_u/k_d \), \( s = v_u/(r v_d) \), and \( r_R = v_d/v_R \). The evolution of the Yukawa matrices, as well as the matching conditions at the Pati-Salam scale, is discussed in more detail in Ref. [2].

### 3.4 New physics in the Higgs Sector

The discovery of the Higgs boson in 2012 has opened new doors in the search for BSM physics. The analysis of the data which has been made subsequently to the discovery indicates that the new particle is in fact a CP-even scalar boson, which forms a \( SU(2)_L \) doublet together with the longitudinal polarizations of the W and Z bosons such that the electroweak symmetry is linearly realized at higher scales as
predicted in the SM [85–87]. Over the coming years the SM Higgs properties, i.e. its mass, spin, parity, and couplings to other particles will be probed to even higher precision. There is however an upper limit to the precision which can be achieved in hadron colliders in general, and in the LHC in particular, and therefore a linear collider is necessary in order to achieve the greatest precision [88, 89]. Now, the question is whether the detected particle really is the Higgs boson, i.e. the scalar particle of the SM. Studies of the Higgs sector beyond the SM can be done in different ways.

The obvious first possibility is studying the Higgs sector in the context of complete, renormalizable physical models, such as supersymmetry [90–92], technicolor [93, 94], or two Higgs doublet models [95–97]. All of these have extended Higgs sectors with several Higgs bosons or enhanced properties of the existing Higgs boson. These models address the conceptual problems of the SM, such as naturalness, and solutions to the hierarchy problem. These, and many other models, have been studied and limited individually, by the LHC [98–100].

Another possibility is to disregard the underlying theory and make a model-independent study in the context of effective operators added to the SM Lagrangian. As previously discussed, this can only occur if we waive our requirements of renormalizability. Furthermore, there is only one five-dimensional operator allowed by the symmetries and that is the Weinberg operator. This operator, however, gives rise to neutrino masses and does not affect the Higgs sector. At the next level, i.e. operators of \( d = 6 \), there are in total 58 non-renormalizable operators for each generation, which are allowed by symmetries of the SM. Any subset of these can be studied individually, which allow for a large number of possible models. Ideally, all six-dimensional operators should be studied simultaneously. At present this is however not feasible. A comprehensive overview of these operators is for example given in Ref. [42]. Naturally, operators of even higher dimensions can contribute to the Higgs couplings. However, at the current precision these contributions can be ignored. In principle one should take all of the six-dimensional effective operators into account in the analysis, which however is not feasible at the present stage. These operators will effectively change the Higgs couplings, their magnitudes and possibly also their tensor structure [101–103].

The simplification can be taken one step further by directly studying the Higgs boson couplings disregarding the underlying mechanism, whether it is a proper physical model or the effective operators. In particular, one can study an extended Higgs sector where the tensor structure of the couplings is assumed to be the same as in the SM and only the magnitudes of the couplings in the Lagrangian are rescaled, which has been done in e.g. Refs. [104–106]. In Ref. [3] we perform a Bayesian model comparison on the scale factors using the interim framework introduced by the Higgs cross section working group in Ref. [107]. In this framework the rescaling of the magnitudes is made by coupling scale factors which are defined in terms of the Higgs production cross sections and decay rates.


3.4.1 Coupling scale factors

In this interim framework, it is assumed that there is a single underlying Higgs boson state of a mass of about 125 GeV. The question is however if it is the SM Higgs boson, or if it is only a part of a bigger picture. No further assumptions are made on other states which might affect the phenomenology of the state which has been observed at 125 GeV, such as additional Higgs bosons or other scalars which do not acquire a vev. The purpose of using the framework with coupling scale factors is to quantitatively investigate whether the properties of the discovered boson agree with the SM Higgs boson or not. This framework does not provide any information on the physical cause of the deviation, therefore subsequent analyses of explicit physical models are necessary.

Furthermore, a narrow-width approximation is assumed, which implies that the signal cross section of a given process $ii \rightarrow H \rightarrow ff$ can be written as

$$(\sigma \cdot BR)(ii \rightarrow H \rightarrow ff) = \sigma(ii \rightarrow H) \cdot \frac{\Gamma_H}{\Gamma_H},$$  \hspace{1cm} (3.33)

where $i$ denotes the initial state, $f$ is the final state, $\Gamma_{ff}$ is the partial decay width of the particular process and $\Gamma_H$ is the total decay width of the Higgs.

At the LHC there are four production modes which is of interest for a light Higgs boson. The predominant process is gluon fusion, $gg \rightarrow H$, which is loop-induced with primarily top and bottom quarks contributions. The other three processes are vector boson fusion, $qq' \rightarrow qq'H$, associated production with a vector boson $q\bar{q} \rightarrow WH/ZH$, and the associated production with a top-quark pair $q\bar{q}/gg \rightarrow t\bar{t}H$.

Five of the Higgs decay modes have been detected, namely the bosonic decays to $\gamma\gamma$, $WW^{(*)}$, and $ZZ^{(*)}$ and the fermionic decays to $b\bar{b}$ and $\tau^+\tau^-$. The $H \rightarrow \gamma\gamma$ is loop induced, whereas the others are tree level processes. Furthermore, lose upper limits have been placed on the decays to $\mu^+\mu^-[108]$.

The coupling scale factors, $\kappa_j$, for the process $ii \rightarrow H \rightarrow ff$, can thus be written as

$$(\sigma \cdot BR)(ii \rightarrow H \rightarrow ff) = (\sigma^{SM} \cdot BR^{SM})(ii \rightarrow H \rightarrow ff) \cdot \frac{\kappa_{ii} \cdot \kappa_{ff}}{\kappa_H},$$  \hspace{1cm} (3.34)

where $\kappa_{ii}$ and $\kappa_{ff}$ are the coupling scale factors for the production and decay processes respectively, and $\kappa_H$ is the scale factor for the total Higgs decay width. The SM is recovered in the limit when $\kappa_i = 1$. In general all of the scale factors depend on the mass of the Higgs boson, however not stated explicitly here.

The $\kappa$s need to be defined both for the SM tree level processes as well as the SM loop induced ones. For the tree level decays they are given by

$$\kappa_j^2 = \frac{\Gamma_{jj}}{\Gamma_{jj}^{SM}},$$  \hspace{1cm} (3.35)

where $\Gamma_{jj}$ is the partial decay width in some model and $\Gamma_{jj}^{SM}$ is the partial decay width in the SM.
3.4. New physics in the Higgs Sector

In addition to the tree level processes we need to define the scaling of vector boson fusion, both the loop induced processes, and the total width of the Higgs. All of these can be determined in terms of the other scale factors. In addition, the coupling scale factors both for the loop induced processes, i.e. \( \kappa_g \) and \( \kappa_\gamma \), and for the total Higgs decay width \( \kappa_H \) can be defined either in terms of the other scale factors or treated as free parameters of the fit.

**Scaling of vector boson fusion**

The scaling factor for the vector boson fusion process is given in terms of the scale factors \( \kappa_W \) and \( \kappa_Z \)

\[
\kappa_{VBF}^2 = \frac{\sigma_{WF} \cdot \kappa_W^2 + \sigma_{ZF} \cdot \kappa_Z^2}{\sigma_{WF} + \sigma_{ZF}},
\]

where \( \sigma_{ZF} \) and \( \sigma_{WF} \) are the W and Z fusion cross sections, respectively.

**Scaling of gluon fusion and \( H \rightarrow gg \)**

The scaling of \( \kappa_g \) is done either via the gluon fusion process or via the Higgs decay to a pair of gluons. In either case \( \kappa_g \) is dependent on \( \kappa_b \) and \( \kappa_t \). Thus \( \kappa_g(\kappa_b, \kappa_t) \) can be defined either in terms of the partial production cross section for gluon fusion or in terms of partial decay rates, for \( H \rightarrow gg \). Thus in terms of cross sections, we have

\[
\kappa_g^2(\kappa_b, \kappa_t) = \frac{\kappa_t^2 \cdot \sigma_{ggH}^{tt} + \kappa_b^2 \cdot \sigma_{ggH}^{bb} + \kappa_t \kappa_b \cdot \sigma_{ggH}^{tb}}{\sigma_{ggH}^{tt} + \sigma_{ggH}^{bb} + \sigma_{ggH}^{tb}},
\]

where \( \sigma_{ggH}^{tt}, \sigma_{ggH}^{bb}, \) and \( \sigma_{ggH}^{tb} \) are the square of the top quark contribution, the square of the bottom quark contribution, and the interference term, respectively. For a light Higgs, the interference term will be negative.

In terms of the partial decay widths \( \kappa_g(\kappa_b, \kappa_t) \) is given by

\[
\kappa_g^2(\kappa_b, \kappa_t) = \frac{\kappa_t^2 \cdot \Gamma_{ggH}^{tt} + \kappa_b^2 \cdot \Gamma_{ggH}^{bb} + \kappa_t \kappa_b \cdot \Gamma_{ggH}^{tb}}{\Gamma_{ggH}^{tt} + \Gamma_{ggH}^{bb} + \Gamma_{ggH}^{tb}},
\]

where \( \Gamma_{ggH}^{tt}, \Gamma_{ggH}^{bb}, \) and \( \Gamma_{ggH}^{tb} \) are the partial decays for \( t \) and \( b \), and the interference term respectively. The term \( \Gamma_{ggH}^{ii} \) is determined for \( \kappa_i = 1 \) and \( \kappa_j \), for \( j \neq i \) and the cross term is obtained from setting \( \kappa_t = \kappa_b \) and subtracting \( \Gamma_{ggH}^{bb} \) and \( \Gamma_{ggH}^{tt} \).

Furthermore, if new physics is allowed to propagate in the loop, the coupling scale factor should be treated as a free parameter. It is then defined as \( \sigma_{ggH}/\sigma_{ggH}^{SM} = \kappa_g^2 \) in the gluon fusion case and \( \Gamma_{gg}/\Gamma_{gg}^{SM} = \kappa_g^2 \) in form of decay widths. In this case the scale factor is assumed to be the same in both the production and the decay process.
Scaling of $H \to \gamma\gamma$

In terms of the other $\kappa$s, $\kappa_\gamma$ is given by

$$\kappa_\gamma(\kappa_b, \kappa_t, \kappa_\tau, \kappa_W) = \frac{\sum_{i,j} \kappa_i \kappa_j \cdot \Gamma_{\gamma\gamma}^{ij}}{\sum_{i,j} \Gamma_{\gamma\gamma}^{ij}},$$

(3.39)

where $\Gamma_{\gamma\gamma}^{ij}$ are the partial decay widths and the pair $(i, j)$ are given by $bb, tt, \tau\tau, WW, bt, b\tau, bW, t\tau, tW, \tau W$ [107]. The partial decay widths are determined in the same way as for the $\kappa_g(\kappa_b, \kappa_t)$.

If one consider the possibility of a contribution of new physics to the loop process, the coupling scale factor can also be treated as a free parameter of the fit. It is then defined as $\kappa_\gamma^2 = \Gamma_{\gamma\gamma}/\Gamma_{\gamma\gamma}^{SM}$. 
Chapter 4

Renormalization group running

Physical quantities, such as cross sections, decay rates, and masses, are divergent when calculated to loop-level Feynman diagrams in a QFT, i.e. to higher orders in perturbation theory. These divergences can either be ultraviolet (UV), corresponding to high energies, or infrared (IR), corresponding to low energies. The latter type of divergences only occur in theories with massless particles and here we shall only consider theories where divergences of the first kind are the relevant ones. In order to make physical sense of QFTs these infinities have to be treated in a consistent way. This is done with a technique, divided into two steps: regularization and renormalization. In the first step, the divergence is parametrized by a small parameter, a regulator such as $\epsilon$, so that the original theory is recovered in the limit $\epsilon \to 0$. Then the terms in the Lagrangian are modified by finite corrections, so-called counterterms, which are allowed to diverge in the limit $\epsilon \to 0$. By choosing the counterterms wisely the theory can stay finite and well defined even in this limit. The Lagrangian which appears in the path integral, is called the bare Lagrangian, and the parameters and fields in the bare Lagrangian are referred to as bare quantities. After the renormalization procedure, the renormalized quantities are left.

4.1 Regularization

A theory is regularized by introduction of a regulator, in an intermediate step, as a method for handling infinities and divergences. Different regularization methods, with different regulators, naturally have different advantages and disadvantages. It should however be emphasized that the fundamental, physical, result does not depend on the regularization method.
Chapter 4. Renormalization group running

A first approach is *cutoff regularization*, where simply an upper limit, $\Lambda$, is introduced in the integral so that the integration no longer extends to infinity, *i.e.* one makes the replacement

$$\int_{0}^{\infty} d^{4}k \to \int_{0}^{\Lambda} d^{4}k. \quad (4.1)$$

In practice this type of brute force method is however rather awkward since it has the undesired property of destroying translational invariance and furthermore it is difficult to maintain gauge invariance [109].

Another method is *Pauli-Villars regularization* [110], where the photon propagator is modified according to

$$\frac{1}{k^{2} + i\epsilon} \to \frac{1}{k^{2} + i\epsilon} - \frac{1}{k^{2} - \Lambda^{2} + i\epsilon}, \quad (4.2)$$

with $\Lambda$ interpreted as the mass of a fictitious photon. Note the opposite sign in front of the second propagator. This sign difference implies that the theory cannot be completely physical. Finite values of $\Lambda$ allow the integral to converge, whereas in the limit $\Lambda \to \infty$ we retrieve the original expression of the propagator. The drawback of this method is that these modified propagators do not describe ordinary particles and hence this method is only physically valid in models where there is a limit on the total energy. Furthermore, this regularization method does not conserve gauge invariance in massive theories.

There are plenty of other regularization methods, in the present thesis we use *dimensional regularization*, which is discussed in more detail in the following section.

### 4.1.1 Dimensional regularization

Dimensional regularization was invented by Martinus Veltman an Gerardus ‘t Hooft in 1972 [111]. It is the most commonly used method of regularization, since it preserves both Lorentz and gauge invariance. In dimensional regularization integrals are formally evaluated in $d < 4$ dimensions, where $d$ need not be an integer, instead of the ordinary $d = 4$ dimensions. Integrals diverging in $d = 4$ dimensions can be shown to converge in $d < 4$ dimensions. Hence, the following change in the integral is made

$$\int \frac{d^{4}k}{(2\pi)^{4}} \to \int \frac{d^{d}k}{(2\pi)^{d}}. \quad (4.3)$$

The $d$ dimensions can be defined as $d = 4 - \epsilon$, where $\epsilon$ is a complex number and the regulator of the theory. Analytical continuation is invoked in order to ensure that the original integral is retrieved when $\epsilon \to 0$. Upon this extension we have not violated any physical law and the properties of the regularized theories are conserved, with the only difference that spacetime is no longer four-dimensional. Note that the use of dimension here only refers to a complex number and not to
4.1. Regularization

actual spacetime dimensions otherwise implied in this work. Thus such a non-
integer complex dimension cannot be interpreted physically and it has no function
outside of perturbation theory.

Upon the continuation to $d$ dimensions the dimensionalities of the fields will
change. The action

$$S = \int d^d x \mathcal{L}$$ (4.4)

is required to be dimensionless. Hence the Lagrangian has to have the mass di-
mension $d$, which allow us to determine the dimensionality of the fields from the
kinetic terms in the Lagrangian since they contain no further parameters and we
know that the partial derivative has the dimension $[\partial_\mu] = 1$. As an example we can
determine the dimensionality of the Higgs field to be

$$[\phi] = \frac{d - 2}{2}.$$ (4.5)

In a similar way the dimensionality of the fermionic and gauge fields are

$$[\psi] = \frac{d - 1}{2},$$ (4.6)
$$[A] = \frac{d - 2}{2}.$$ (4.7)

If all expansion parameters have mass-dimension greater than or equal to zero, the
theory is power counting renormalizable. This is an important feature of the SM
and the reason that the highest order terms in a four-dimensional theory that we
can accept are quartic interactions among scalars and gauge fields (in the SM there
is only quartic gluon interactions). Furthermore, in order to be able to actually
perform calculations in $d$ dimensions we need to define a Clifford algebra, which is
non-trivial. This ought to be done in a consistent way so that the original integrals
are obtained in the limit $d \to 4$. A possible definition is using a similar one to the
four-dimensional case, i.e.

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu},$$ (4.8)

where $\eta^{\mu\nu}$ is the ordinary Minkowski metric. However, the traces need to change
and thus we have

$$\eta_{\mu\nu}\eta^{\mu\nu} = d,$$ (4.9)
$$\gamma_\mu\gamma^\mu = d.$$ (4.10)

As a consequence, the possible contractions and traces of the gamma matrices will
differ from the four-dimensional ones. A full list of these contractions is given in
Ref. [109]. In order to keep the mass dimension intact the couplings need to be redefined in the SM and similar theories. The following redefinitions are made

\[ \lambda \rightarrow \lambda_0 = \mu^\epsilon \lambda, \]  
(4.11)

\[ g_i \rightarrow g_{i,0} = \mu^{\frac{\epsilon}{2}} g_i, \]  
(4.12)

\[ Y_f \rightarrow Y_{f,0} = \mu^{\frac{\epsilon}{2}} Y_f, \]  
(4.13)

where \( \mu \) is the renormalization scale with mass dimension, \( i \in \{ 1, 2, 3 \} \), and \( f \in \{ e, \nu, u, d \} \). Contrary to the case in Pauli-Villars and cut-off regularizations, \( \mu \) is not a large mass scale. Hence the original integrals are retrieved in the case \( d \rightarrow 4 \) and not \( \mu \rightarrow \infty \). Quantities with subscript 0 are the bare quantities, which have mass dimensions, whereas quantities without subscripts are the renormalized, physical ones. The renormalization scale is an arbitrary scale which is not taken to infinity, but will eventually drop out of physical calculations. After the renormalization procedure the physical parameters will depend both on the renormalization scale and the bare parameters. These relations are valid to lowest order and are thus modified by loop corrections.

Calculating the loop integrals which arise from the Feynman diagrams is in principle rather tedious. First, the integral is extended to \( d \) dimensions, then a Feynman parametrization of the integrand is performed [109]. This parametrization, as the name indicates, due to Richard Feynman, and provides a way to rewrite fractions in the denominator as products, which then more easily can be evaluated. The next step is a Wick rotation from Minkowski to Euclidean space. This is done by defining a Euclidean \( d \)-dimensional vector \( \mathbf{k} \) such that

\[ k^0 = i k_E, \quad \mathbf{k} = \mathbf{k}, \]  
(4.14)

with \( k^2 = k_1^2 + k_2^2 + \ldots + k_d^2 \), i.e. the ordinary dot product on Euclidean space. The integrals are then written in spherical coordinates in the complex plane, followed by the regularization step. Normally one is only interested in the divergent parts of the integral and in the case of dimensional regularization the Passarino-Veltman decomposition can be conveniently used. The divergent parts of the integrals are tabulated, see Ref. [112] and can be used directly without the tedious computations.

### 4.2 Renormalization

A relationship between the bare and renormalized quantities is needed in order to make any physical sense of our initial, bare Lagrangian. This is done through the introduction of renormalization constants, \( Z_i \) for the fields, couplings, and masses. The renormalization constants are on the form \( Z_i = 1 + \delta Z_i \), and in terms of the Lagrangian this amounts to the expression

\[ \mathcal{L}_0 = \mathcal{L} + \delta \mathcal{L}, \]  
(4.15)
where $\mathcal{L}_0$ is the bare Lagrangian, $\mathcal{L}$ is the renormalized one, and $\delta\mathcal{L}$ contains the necessary counterterms. In dimensional regularization, the counterterms replace the dependence on $\Lambda$ with a dependence on $\mu$.

After regularization, a choice of renormalization scheme has to be made. Once again, the choice of renormalization scheme should not affect any physical results. However, there are practical advantages and disadvantages which should be taken into account. The two most commonly used ones are on-shell renormalization and minimal subtraction (MS). The renormalization constants in on-shell renormalization schemes are defined such that the inverse renormalized propagators, their first derivatives and the renormalized vertices are equal to the corresponding Born values on the mass shell. Such schemes are often used in QED but never in QCD, due to problems regarding infrared divergences. The method used in this thesis is MS, or rather its close relative modified minimal subtraction (\overline{MS}), which will be discussed in the following section.

### 4.2.1 Minimal subtraction

Minimal subtraction only works together with dimensional regularization, contrary to other renormalization schemes, which work with any regularization scheme. Normally \overline{MS}, is convenient to use together with dimensional regularization. The procedure is minimal in the sense that only the pole in $1/\epsilon$ is absorbed into the counterterms, which therefore do not contain any finite parts. From the renormalization scale, $\mu$ of MS, is \overline{MS} obtained from the through the replacement

$$\mu^2 \to \overline{\mu}^2 = \frac{\mu^2}{4\pi} e^{\gamma_E},$$

(4.16)

where $\gamma_E$ is the Euler constant. The reason for this choice is the fact that in one-loop calculations $1/\epsilon$ always appears in the combination

$$\frac{1}{\epsilon} - \gamma_E + \ln 4\pi = \frac{1}{\epsilon},$$

(4.17)

Furthermore, since this scheme has a minimal structure the counterterms are independent of mass, $m$. The MS scheme was devised as a method for renormalization of electromagnetic and weak interactions independently by t’Hooft and Weinberg in the beginning of the 1970s [113, 114].

### 4.3 Renormalization group equations

An important consequence of renormalization is that the renormalized coupling constants and masses depend on the energy scale where they are measured. We say that the parameters run. However, the theory should still be invariant under a change of the renormalization prescription. Thus a change in $\mu$ will be cancelled by a change in the renormalized mass, $m(\mu)$, or coupling constant, $g(\mu)$. In other words
the bare quantities are independent of the renormalization scale. The behavior of the evolution of the renormalized quantities is determined by Renormalization Group Equations (RGEs). These are determined from the relation between the bare and renormalized Green’s functions. The bare Green’s function is given by

$$G_0^{(n)}(\{x_i\}; \mu, \lambda) \equiv \langle 0 | T \phi_0(p_1) \ldots \phi_0(p_n) | 0 \rangle,$$  (4.18)

where $T$ is time ordering. We now relate the bare Green’s function to the renormalized one through

$$G_0^{(n)}(\{x_i\}; \mu, \lambda) = Z^{-\frac{n}{2}} \phi G^{(n)}(\{x_i\}; \mu, \lambda),$$  (4.19)

where $\phi_0 = Z^{\frac{n}{2}}$. From the fact that the bare quantities are independent of the renormalization scale, the Green’s function must also be independent of $\mu$. This can be expressed as

$$\frac{dG_0}{d\mu} = 0.$$  (4.20)

Carrying out this differentiation explicitly we arrive at the Callan-Symanzik equation [115–117], which is given by

$$\left( \mu \frac{\partial}{\partial \mu} + \beta(\lambda) \frac{\partial}{\partial \lambda} + n \gamma(\mu) \right) G^{(n)}(\{x_i\}; \mu, \lambda) = 0.$$  (4.21)

This equation describes the evolution of the $n$-point correlation function when the energy is varied and the functions $\beta(\mu)$ and $\gamma(\mu)$ encode the change in the couplings and the masses which compensates for the change in $\mu$. They are given by

$$\beta = \mu \frac{d\lambda}{d\mu}, \quad \gamma = \frac{1}{2Z_\phi} \mu \frac{dZ_\phi}{d\mu}.$$  (4.22)

In the case of more than one coupling and field in the model, there are one $\beta$-function for each coupling and one $\gamma$-function for each field. Since these functions must be the same for any $n$, they must be independent of the coordinates $x_i$. Since the renormalized Green’s function is finite, $\beta$ and $\gamma$ will also be finite and in addition independent of the cutoff. The RGEs are naturally dependent on the model and thus the behavior of the RG running of the parameters could in principle be used to distinguish different models.

We have studied the RG running of the quark masses and mixings, the Higgs self-coupling, the charged lepton masses, and the neutrino masses and leptonic mixing angles, in the context of a six-dimensional UED in Ref. [1]. In the context of a non-supersymmetric $SO(10)$ model we study RG running of the fermion parameters, especially the leptonic mixing parameters and masses in Ref. [2]. A more detailed description of the equations used is given in the respective paper.
Chapter 5

Statistics

Access to sophisticated statistical tools has become very important in particle physics. Searches for new physics often involve searches for incredibly small effects in tremendously large data sets from experiments which are both intricate and difficult to perform. A few decades ago particles, such as $J/\psi$, was discovered as tracks in a bubble chamber which could be seen by eye, whereas the recent discovery of the Higgs boson required a sophisticated statistical analysis in order to find an effect which, if anything, was like finding the proverbial needle in a statistical haystack [118, 119]. Still, some argue that the use of statistics only points to a lack in the quality of the data. However, since we nowadays, in some cases, need to deal with the fact that there will be neither better experiments nor better data, at least not for a very long time, we do need to make the most of what we have. Statistics is the tool for quantifying our knowledge about a measurement and its uncertainties and the way experimental results are communicated to the rest of the community.

In the field of particle physics is statistics used in many contexts. Point estimation is used for estimating the value of a certain quantity such as a particle mass, confidence intervals to put bounds on for example cross sections and signal strengths, model comparison or hypothesis testing in order to test the hypotheses of the existence of for instance a Higgs boson, supersymmetry, or extra dimensions. The list can be made long. There are plenty of pitfalls in statistics and many are the phenomena with a statistical significance of $2 - 3\sigma$, which have historically caused a stir in the community, only to disappear a few months later upon the inclusion of more data. These statistical fluctuations are expected and should be treated accordingly.

There are two fundamentally different types of statistics, originating from two essentially different interpretations of probability, known as frequentist and Bayesian statistics. In the high-energy physics community in particular, the frequentist framework has been the prevalent one during the 20th century. The main exception to this rule has been cosmology, where the existence of one, single Universe,
has made the frequentist framework difficult, if not impossible, to use. However, for example the data analyses at particle accelerators do include elements of both frequentist and Bayesian approaches [120]. Especially, the nuisance parameters are treated efficiently within in the Bayesian framework. We shall here discuss the concepts of primarily Bayesian statistics which is applied to LHC and Tevatron Higgs data in Ref. [3], but also briefly discuss some important concepts in the context of frequentist statistics.

5.1 Probability

The concept of probability lies at the basis of the field of statistics. Any function which satisfies the axioms of probability, discussed in Sec. 5.1.1, can, by definition, be called a probability function. Still, one must specify how the elements of the sample space should be interpreted and also how the probability values should be assigned and interpreted [121]. We shall discuss the two most common interpretations, starting from the mathematical definition of probability.

5.1.1 Axioms of probability

The definition of probability were stated in 1933 by Kolmogorov [122]. Consider a set $S$, called the sample space, consisting of a number of elements. One then assigns a probability, $P(A)$, to each subset $A \in S$ which is defined by the following axioms

1. For every subset $A \in S$, $P(A) \geq 0$.

2. For any two disjoint subsets $A$ and $B$ is the probability assigned to the union of $A$ and $B$ equal to the sum of the two corresponding probabilities, $P(A \cup B) = P(A) + P(B)$.

3. The probability assigned to the whole sample space is $P(S) = 1$.

Further properties of the probability function can be derived from these axioms. Conditional probability $P(B|A)$, which is read as “the probability of $B$ given $A$”, can be reversed using Bayes’ theorem is defined as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$  \hfill (5.1)

As the name indicates Bayesian inference is based on Bayes’ theorem. The formula per se is however perfectly valid when using the frequentist framework as well.

5.1.2 Interpretation of probability

An interpretation is necessary in order to make any partial use of these axioms. In the frequentist framework the probability of an event is interpreted as a limiting relative frequency. More elaborately expressed as “the number of times the
event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions” [123]. Based on such a model probabilities can obviously never be determined with perfect precision and thus the task becomes to estimate them, which are assumed to have fixed but unknowable values. Here data is a repeatable random sample whereas parameters of a theory are considered fixed. The interpretation of probability as a frequency is problematic in the context of unique phenomena, for example regarding questions of our Universe, such as assigning a probability to the event that it rained the day Richard Feynman was born. The problem might be resolved by assuming that our Universe is just one out of many universes and that a certain event occur in some fraction of these, which does feel rather contrived.

In the Bayesian view is probability interpreted as a degree of belief in a proposition [123]. Elements of the sample space are then hypotheses or propositions which are either true or false. It is often referred to as subjective probability, since in this interpretation prior knowledge is included in determining the probability. This interpretation does however have a great advantage in giving an unproblematic and meaningful probability to a unique event. In this framework observed data are considered fixed and the parameters unknown and described probabilistically.

## 5.2 Bayesian statistics

The Bayesian method provides the best known decision rule and an elegant way to update previous beliefs with new information. However, it does not provide a way to summarize results which are independent of previous beliefs and knowledge. As the name indicates Bayes’ theorem is of central importance in Bayesian statistics. In terms of an hypothesis, $H$; the information given, $I$; and the known data, $d$, it can be written

$$P(H|d, I) = \frac{P(d|H, I)P(H|I)}{P(d|I)},$$

(5.2)

where $P(H|d, I)$ is the posterior probability representing our knowledge after the data have been taken into account. The data are given in the likelihood, $P(d|H, I)$, which is a function of the hypothesis, derived from fixed data. It can in the continuous case be written $\mathcal{L}(H) = P(d|H, I)$. The likelihood allows us to estimate unknown parameters based on known outcomes. Our knowledge prior to taking the data into account is summarized in $P(H|I)$, which is called the prior probability. The quantity in the denominator has several names, we shall refer to it as the evidence. Through marginalization over all possible hypotheses this is given as

$$P(d|I) = \sum_{H} P(d|H, I)P(H|I)P(d|I).$$

(5.3)

This quantity is of great importance for model selection but is irrelevant for parameter estimation.
5.2.1 Parameter inference

A typical problem in statistics is determining the best fit values of a set of parameters within a model\footnote{We shall only consider parametric inference, although non-parametric inference, which is not parametrized by a fixed number of parameters, also is in principle possible.}. The best fit value of the parameter is determined using point estimation and credible intervals for the parameters are determined by interval estimation. Both the point and interval estimates are determined from the posterior distribution, which is determined by Bayes’ theorem. By conditioning on an explicit choice of model, $M$, Bayes’ theorem can be reformulated. Then the model is given as a set of hypotheses in terms of a set of parameters $\theta$. Together with the data, $d$, Bayes’ theorem can thus be written

$$P(\theta|d,M) = \frac{P(d|\theta,M)P(\theta|M)}{P(d|M)},$$

(5.4)

where $P(d|M)$ is only a normalizing factor and therefore irrelevant in parameter estimation. Together with the model, we need to determine which prior to use, taking into consideration the caveats which previously have been discussed. In addition, the likelihood should be constructed, given the experiment under consideration. This is a non-trivial task and cannot always be done properly. In some cases the correct likelihood can be approximated by a Gaussian, in order to give a hint of the true character of the problem.

Depending on the what parts of the model is decided interesting, the parameters of the model, $\theta = (\phi, \psi)$, can be divided into relevant parameters $\phi$, and nuisance parameters, $\psi$. In order to only make inference on the relevant parameters $\phi$ it is possible to marginalize over the nuisance parameters. For the continuous case this is done as

$$P(\phi|d,M) = \int P(\phi, \psi|d,M) d\psi.$$

(5.5)

Once the posterior has been determined inference on $\phi$. The information can be summarized using a summary statistic of choice, such as mean, median, or mode of the posterior distribution together with standard deviations and correlation matrix \cite{124}.

Given the posterior distribution a credible interval can be constructed in such a way that it has a probability of 90\% (for example) that the true parameter is contained in the interval. The only requirement on the credible interval is that it should contain 90\% of the posterior (in this example) so it can be constructed in many different ways.

5.2.2 Priors

As the name indicates the prior contains the \textit{a priori} knowledge of the parameters which are to be estimated. The choice of prior distribution is a fundamental ingredient of Bayesian statistics. It can be argued that this introduces a subjective
component in the analysis, which however should be thought of as a feature of
the statistical model rather than a limitation. The gist of the prior is that it as
accurately as possible should reflect all previous knowledge of the problem at hand.
It is sensible to include prior knowledge of a problem as long as it is made in a
consistent and correct way.

The choice of prior is not trivial and since all physicists have different prior
knowledge of a problem they will make different choices of prior. One needs to take
care to include all knowledge from the theory and from previous experiments which
are not included in the current analysis. There are a few customary choices of prior
distributions which tend to make sense in many physical situations. The perhaps
most intuitive example is the uniform prior, where the prior has a constant value
over the interval of interest, and zero otherwise, \( P(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}. \) (5.6)

Note, however, that the use of a uniform prior is not a way to implement ignorance
of the parameters in a model. It should be used in the case when it is reasonable to
believe that the parameters lie in some given range and that the all intervals of equal
length in that range have equal probability. Another common example of a prior
is the logarithmic prior, which is uniform in the logarithm of the parameter. This
choice of prior is convenient to use when the order of magnitude is unknown, which
is rather common in physics. Since many phenomena in Nature are distributed
according to a normal distribution, a gaussian prior might be relevant for some
quantities. There is an ongoing debate regarding objective priors, as a method of
quantifying ignorance. Some of them are model independent and based on symmetry
arguments [124–127]. We shall however not delve any deeper into that subject here.

An important question is however, how to quantify ignorance in a problem. As
previously mentioned, it is often falsely assumed that the analysis can be made
objective by the choice of a uniform prior. However, this is not the case, since a
prior which is uniform in one parameter, \( \theta \), will not be uniform in another one, \( \phi \),
i.e.

\[
P(\theta) = \left| \frac{d\theta}{d\phi} \right| P(\phi). \quad (5.7)
\]

An example of a prior which is independent of the choice of parameter on which
the prior is placed is the Jeffreys prior [128]. However, this prior does depend on
all the data in an experiment which could have been observed and therefore cannot
really be considered to be objective.

When the information in the data is small, the prior is of great importance. However,
as soon as the data are good enough it does however turn out that in
most cases the choice of prior is, as long as it is made in a sensible way, rather
irrelevant for the resulting posterior. The evidence is however more sensitive to
choice of prior, and the prior range.
5.2.3 Model comparison

Apart from making parameter inference, which is performed within one model, it is of interest to be able to compare different models and quantify this comparison somehow. The Bayesian evidence for a model, \( Z = P(d|M) \), is the probability of the data given that the model is true and this quantity is of great importance in model selection, whereas it was irrelevant in the context of parameter estimation. Explicitly conditioning on the model under consideration and integrating over the parameters of the model, the evidence is given by

\[
P(d|M) = \int P(d|\theta, M)P(\theta|M)d\theta.
\]

The posterior of the model given the data, but independent of the parameters, can then be written in terms of the evidence and the prior of the model itself, \( P(M) \), as

\[
P(M|d) \propto P(d|M)P(M).
\]

The normalization constant \( P(d) \) is irrelevant and has therefore been dropped. Bayesian model selection can be made in different ways. The posteriors of the models are compared by taking the ratio of the posteriors:

\[
\frac{P(M_1|d)}{P(M_2|d)} = \frac{P(d|M_1)P(M_1)}{P(d|M_2)P(M_2)}.
\]

The last term on the right-hand side, \( i.e. \) the prior ratio, is normally assumed to be equal to one for fairness unless there is some particular reason to prefer one over the other. Thus this term can be dropped. The ratio

\[
B = \frac{P(d|M_1)}{P(d|M_2)}
\]

is called the \textit{Bayes factor} and is given by the ratio of the models’ evidences. If the Bayes factors is \( > 1 \ (\ < 1) \) the evidence is for (against) the model \( M_1 \) compared to \( M_2 \). Since all parameters of the models are integrated over, it does not depend on a single set of parameters. In addition, a natural consequence of this integral is an \textit{Occam effect}, which is built into the Bayes factor and serves as a protection against overfitting. A more complicated model will automatically be punished unless the fit to the data is significantly better. The strength of the evidence can for example be quantified using the \textit{Jeffreys scale} given in Table 5.1, as used \( e.g., \) Refs. [123, 129]. This scale has been used in a number of applications such as [123, 130, 131], historically a more aggressive scale has been used [132, 133].

Computing the Bayesian evidence

The numerical computation of the evidence requires multi-dimensional integration over all of parameter space, which is a computationally demanding task. The demand for computational power is one of the main reasons that frequentist statistics
5.2. Bayesian statistics

Table 5.1. The Jeffreys scale, which is used for interpretation of Bayes factors, odds, and model probabilities. Note that log here denotes natural logarithm.

| log(odds) | Odds   | \(P(H_1|D)\) | Interpretation   |
|----------|--------|-----------|-----------------|
| < 1.0    | \(\lesssim 3 : 1\) | \(\lesssim 0.75\) | Inconclusive   |
| 1.0      | \(\approx 3 : 1\) | \(\approx 0.75\) | Weak evidence  |
| 2.5      | \(\approx 12 : 1\) | \(\approx 0.92\) | Moderate evidence |
| 5.0      | \(\approx 150 : 1\) | \(\approx 0.993\) | Strong evidence |

has gained such an advantage over the Bayesian one. The Bayesian revival is driven both by improved hardware but also by problems in more dimensions. A combination of more computational power and more efficient algorithms has only recently made many evidence calculations viable. There is a number of algorithms for calculating the evidence and one which is particularly designed for it is nested sampling, constructed by John Skilling in 2004 [134, 135]. It reduces the many-dimensional integral to only one dimension, heavily reducing the computational cost. Starting with \(n\) points which are sampled according to the prior. The following points are sampled according to \(\mathcal{L} > \mathcal{L}^*\), where \(\mathcal{L}^*\) is the current smallest value of the likelihood. This can be equally expressed in terms of a parameter \(\zeta\), defined as

\[
\zeta(\lambda) = \int_{\mathcal{L}(x) > \lambda} \pi(x),
\]

i.e. the proportion of the prior, \(\pi\), with likelihood greater than a given value, \(\lambda\). Thus the points sampled according to the requirement on the likelihood are subject to \(\zeta < \zeta^*\). One iteration contains the following steps

- Determine \(\mathcal{L}_i\), the minimum of the current likelihoods of the points, and the corresponding \(\zeta_i\).
- Assuming the points are uniformly distributed in \((0, \zeta^*)\), the shrinkage ratio of the interval is \(t = \zeta_i / \zeta^*\), and assuming \(\log t = -\frac{1}{n}\) we obtain \(\zeta_i = \exp(-i/n)\)
- Determine the weight, \(w_i = \zeta_{i-1} - \zeta_i\).
- Update the evidence according to \(Z = Z + \mathcal{L}_i \cdot w_i\).
- Redefine \(\zeta^* = \zeta_i\) and update point with least likelihood.

The algorithm has been further developed so that it can be used in multimodal problems, which has been implemented in the software MultiNest [136–138]. This software has been used to determine the evidence in Ref. [3] in order to perform model comparison. When calculating the evidence the posterior distribution is determined as a by-product, which can be used for parameter estimation. This particular use of MultiNest has been made in Ref. [2]. The evidence can be estimated using other methods. One example is importance sampling with the prior as an importance sampling distribution, which will lead to the arithmetic mean
Chapter 5. Statistics

estimator, which however will be rather bad unless the sample size is large [139]. Another, better, method is thermodynamic integration, which is based on statistical thermodynamics. Then the evidence is considered as the partition function $Z(\beta)$ of a canonical ensemble where the loglikelihood is a fictitious energy function, $E(x) = -\log(P(d|M, x, I))$ and $\beta$ is a fictitious inverse temperature. The idea of thermodynamic integration then proceeds from

$$\frac{d \ln Z(\beta)}{d\beta} = -\langle E(\tilde{x}) \rangle_\beta.$$  \hspace{1cm} (5.13)

Thus the evidence is determined from the integral

$$\ln[Z(1)] = -\int_0^1 \langle E(\tilde{x}) \rangle_\beta d\beta.$$  \hspace{1cm} (5.14)

$\langle E(\tilde{x}) \rangle_\beta$ can be computed by an MCMC [140].

5.3 Frequentist statistics

Frequentist inference is based on the idea that probability is a limiting frequency. Data are a repeatable random sample, whereas the model, i.e. the parameters in parametric inference, remain fixed. Our data are noisy instances of the true values and the more data we can sample, the closer to the truth we will be. The frequentist approach to inference is the completely dominant one in particle physics, and the one mainly used in the analysis of LHC data.

5.3.1 Parameter estimation

Within the frequentist framework parameter estimation is performed in a diametrically different way to the Bayesian. In this case, the true value of the parameter is unknown and in fact unknowable, and thus can only be estimated from the data set.

Parameter estimation is made using different estimation methods which all will give a point estimate of the parameter. Examples of such methods are the maximum likelihood method, least squares method, and method of moments. The most commonly used and important method is the maximum likelihood method, where the best estimates of the parameters are, as the name indicates, given by maximizing the value of the likelihood.

For parameters confidence intervals can also be constructed. The true value of the parameter, $\theta_{\text{true}}$, is fixed and one wants to construct an interval such that

$$P(\theta_{\text{true}} \in I) = 1 - \alpha,$$  \hspace{1cm} (5.15)

where $\beta$ is a some pre-specified probability. The interval $(\theta_a, \theta_b)$ is the confidence interval and this interval, not $\theta_{\text{true}}$, is the random variable in this equation. The
5.3. Frequentist statistics

The confidence interval is fundamentally different from the Bayesian credible intervals. The confidence interval is constructed such that the interval covers $\theta_{\text{true}}$ in $1 - \alpha$ of the cases, whereas there is a probability, or degree of belief, $\beta$, that the Bayesian credible interval contains the true parameter. The proportion of the time the interval covers the true value is called coverage and is an important notion in the frequentist statistics. This difference does for example allow the entire confidence interval to end up in an unphysical part of the parameter space, the outcome just happens to be in the portion $\alpha$ which does not cover the true value. The construction of confidence intervals is non-trivial and can be done by a method due to Neyman-Pearson [141] but is most oftenly done by $\chi^2$ or likelihood minimization.

5.3.2 Hypothesis testing

Hypothesis testing, comparing two hypotheses, in the frequentist framework normally consists of the following steps.

- First, define a null hypothesis $H_0$ and an alternate hypothesis $H_1$. Typically the null hypothesis would be that the observations are pure chance (“there is no Higgs”), whereas the alternate hypothesis would require some new physics (“there is a Higgs”).

- In the second step a test statistic is defined in such a way that it can quantify, within the given distribution, the difference between the null and alternate hypotheses. Examples of possible test statistics is the log-likelihood ratio, the $\chi^2$-distributions, and the Student-t distribution.

- Then the p-value is determined. This concept is often mentioned in the literature and alarmingly often defined and used in the wrong way. The correct definition of the p-value is

  “the probability, assuming the null hypothesis $H_0$, to obtain a result as or more extreme than what was observed”.

- The final step is to determine if the p-value fulfills $p < \alpha$ for some pre-specified significance level, $1 - \alpha$, the null hypothesis will be rejected in favor of the alternate hypothesis. In high-energy physics, there is a rather arbitrary definition, that a phenomenon observed with a statistical significance of $5\sigma$ deviation from a Gaussian distribution, which is equivalent to $p < 3 \cdot 10^{-7}$, is said to be a discovery.

In the context of hypothesis testing two kinds of errors, type-I and type-II errors, are often discussed. A type-I error occurs when the null hypothesis is rejected even though it is true. The probability of such an error is $\alpha$. A type-II error occurs when the null hypothesis is retained even though it is false.
Chapter 6

Summary and conclusions

In Part I of this thesis, the models and techniques used in Part II were introduced. First, in Ch. 1, we gave an introduction to physics in general, and particle physics in particular. In Ch. 2, we gave an introduction to the SM with emphasis on the Higgs and neutrino sectors of the model. We also briefly discussed the problems of the SM. Then, in Ch. 3, we discussed the properties of effective field theories and their use in the search for a model beyond the SM. Since these theories contain operators of higher mass dimension than 4, they are non-renormalizable, which is not a problem if the theory is only employed within its energy range of validity. We continued by discussing three specific models. First, we discussed a six-dimensional model with UEDs. In this model the extra dimensions are manifest in terms of heavy particles. Then, we discussed a non-supersymmetric $SO(10)$ GUT model, with an intermediate scale and the Pati-Salam group as an intermediate group. Within this model the SM was regarded an effective theory and neutrino masses were generated through a type-I seesaw. Finally, we discussed effective models in the context of the Higgs sector. Especially we considered models where simply the magnitude of the tree-level Higgs couplings are rescaled with so-called effective scale couplings. A further extension was made by introducing effective scale couplings for the loop-induced processes, such as gluon fusion and Higgs boson decay to two photons. In Ch. 4, we discussed the concepts of regularization, renormalization, and renormalization group running in some detail. Finally, in Ch. 5, we gave an introduction to the concept of probability and to two of the schools of interpretation of probability, the frequentist and the Bayesian. Emphasis was placed on the Bayesian interpretation of probability and inference methods, as well as the subject of Bayesian model comparison.

Part II of this thesis consists of three scientific papers, which investigate the models discussed in Part I, using the methods described.

In Paper I, we discussed RG running of the neutrino parameters, masses and mixing angles, in the context of a six-dimensional UED model. We found that the KK modes from the extra dimensions enhance the running so that it shows a
power-law behavior in comparison to the logarithmic behavior, which is present in the SM. The most sizeable running was observed in the parameter $\theta_{12}$. Furthermore, we studied the running of the Higgs self-coupling in the model and could, from the requirement that it should be non-negative, put a limit on the cut-off scale, which was $\Lambda \approx 7$ TeV. This only allows for contributions from five KK modes.

In Paper II, we investigated the RG running of the fermion parameters, *i.e.* lepton masses and mixing angles as well as quark masses and mixing parameters, within a non-supersymmetric $SO(10)$ GUT model, with a minimal content in terms of Higgs representations. In such a non-supersymmetric model an intermediate scale is necessary in order to be able to incorporate all fermion masses. We compared the running in this model with a SM-like model, with the addition of three right-handed neutrinos necessary for a type-I seesaw mechanism. We found that the impact of the intermediate scale is significant for the outcome of the running. All parameters exhibited significant running, except for the quark mixing angles and the leptonic mixing angle $\theta_{13}^{\ell}$.

In Paper III, we discussed an effective extension of the Higgs sector. In this case, the magnitude of the Higgs boson couplings rescales by coupling scale factors, while the tensor structure of the couplings remains the same. We performed Bayesian inference on these parameters, with the result that all of the SM values were within the $1\sigma$ bound. Furthermore, we performed Bayesian model comparison between models with one or more parameters allowed to vary and the SM, where all parameters are fixed. The evidence in favor of the SM increases with the number of free parameters. The analysis was made in both the case where only the tree level couplings were allowed to vary, and the case where both the tree-level and loop-induced couplings were varied.
Bibliography


[33] ATLAS Collaboration, G. Aad et al., Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with $b$-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector, (2014), 1406.5375.


[58] ATLAS Collaboration, G. Aad et al., Search for dark matter candidates and large extra dimensions in events with a jet and missing transverse momentum with the ATLAS detector, JHEP 1304, 075 (2013), 1210.4491.

[59] CMS Collaboration, S. Chatrchyan et al., Search for dark matter and large extra dimensions in monojet events in pp collisions at $\sqrt{s} = 7$ TeV, JHEP 1209, 094 (2012), 1206.5663.


BIBLIOGRAPHY


[107] LHC Higgs Cross Section Working Group, A. David et al., LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle, (2012), 1209.0040.


[129] M. Hobson *et al.* (Eds.), *Bayesian methods in cosmology* (Cambridge University Press, 2010).


[131] F. Feroz *et al.*, *Bayesian Selection of sign(mu) within mSUGRA in Global Fits Including WMAP5 Results*, JHEP **0810**, 064 (2008), 0807.4512.


