



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Posgrado en Ciencias Físicas

Instituto de Física

Low Energy Electroweak Precision Constraints

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS FÍSICAS

PRESENTA:

RODOLFO FERRO HERNÁNDEZ

Tutor Principal:

Dr. Paul Artur Jens Erler Weber
Instituto de Física

Miembros del Comité tutor:

Dr. Saúl Noé Ramos Sánchez
Instituto de Física

Dr. Luis Fernando Urrutia Ríos
Instituto de Ciencias Nucleares

Ciudad Universitaria, Cd. Mx. Diciembre 2020



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

To my parents

Acknowledgments

It is clear that the results and projects achieved during this PhD would not have been done without the help of many people. First of all, I want to thank my supervisor Jens Erler. His patience and incredible physics intuition made me look physics in a totally different way. I learned a lot during this process, and the vast majority of this learning is a result of the strong collaboration with him. Thank you Jens. Thanks to the members of my committee, Saúl Ramos and Luis Fernando Urrutia, they always supervised with detail my development in this program.

I would also like to thank to all my collaborators, but specially to Eduardo Peinado and Aurora Cuortoy, the discussions, talks and work done with them certainly enriched my knowledge and interest for different topics in particle physics. I learned a lot from much more people at the Instituto de Física, like Genaro Toledo, Saúl Ramos, Myriam Mondragón, Isaac Pérez and many others. They always had the door and mind open to hear me say and ask any question about physics. Thanks to all of them.

On the personal level I want to thank my parents, for their unconditional love and support. They were always there for me, I traveled this far because they taught me how to persist in the pursuit of what you love. Thank you Alicia and Rodolfo.

The work of this thesis had the support of institutions and governmental programs. CONACyT gave me a phd scholarship which was very important to be focused to get this work done. The research of this work was also supported by DGAPA-PAPIIT IA102418 and IA101720, the German–Mexican research collaboration grant SP 778/4–1 (DFG), and the CONACyT projects 252167, 278017 and CB2017-2018/A1-S-13051.

Resumen

Esta tesis cubre una gran variedad de temas, aunque de alguna manera todos ellos orbitan alrededor de física electrodébil de precisión. La estructura de la tesis sigue una lógica muy específica, trata de ir de lo general a lo particular. Primero, empiezo discutiendo cómo mejorar la combinación de resultados de varios experimentos y doy como ejemplo aplicado el tiempo de vida media del neutrón. La combinación de varios experimentos es estudiada usando un modelo Bayesiano jerárquico. Uno de los resultados principales es una justificación más formal del escalamiento que el Particle Data Group (PDG) aplica cuando combina errores para un número grande de experimentos, pero al mismo tiempo mejora la forma en la que se combinan los datos cuando tenemos pocos experimentos. Estos resultados fueron publicados en el año 2020. El modelo Bayesiano usado para mejorar el escalamiento del PDG también fue usado para estimar un error en una restricción a las funciones de distribución partónica, esto es parte de una publicación más amplia que fue publicada en los promedios de la comunidad PDFLattice. Después de analizar estos temas, discuto uno de los resultados más importantes de mi doctorado, que involucró cálculos teóricos dentro del Modelo Estándar. Este cálculo fue el corrimiento del ángulo de mezcla débil desde el polo del bosón Z hasta bajas energías, donde se consiguió una reducción en la incertidumbre por casi un factor de cuatro. Esta reducción es muy importante cuando se desea comparar experimentos y teoría. Los resultados fueron publicados en 2018 y la gráfica más representativa de este trabajo fue mostrada en la versión 2018 de la sección de física electrodébil del PDG. Al final de la tesis, discuto modelos específicos más allá del Modelo Estándar, en particular, se estudian dos modelos de masa de neutrinos usando simetrías discretas. Los modelos fueron publicados

en dos artículos independientes en 2019 y 2020, con resultados compatibles con los datos experimentales de oscilaciones de neutrinos.

Abstract

This thesis covers a wide variety of topics, but all of them orbit around electroweak precision physics. The structure follows a specific logic, first, I start discussing how to improve the combination of results of several experiments, and I give an applied example to the neutron lifetime. The combination of experiments is studied using a Bayesian hierarchical model. One of the main results is a formal justification of the Particle Data Group (PDG) scaling of errors, but at the same time, we get a way to improve the combination of measurements for a small number of experiments. These results were published in 2020. I also used a similar hierarchical model to estimate the error on a bound to parton distribution functions, which was part of broader work that was published in the averages of the PDFLattice community.

Then I move to theoretical calculations inside the Standard Model. I compute the running of the weak mixing angle to low energies, where I achieved a reduction of the uncertainty by almost a factor of four. This reduction is important when comparing future experiments and theory. The results were published in 2018 and the main plot of this work was shown in the 2018 version of the PDG.

At the end of the thesis I discuss specific models beyond the Standard Model, particularly neutrino mass models. Two models of neutrino masses were published from this work, and both involved discrete flavor symmetries. The resulting correlations between the oscillation parameters were consistent with the experimental data. These works turned into two publications one in 2019 and the other one in 2020.

In this thesis I will give a detailed explanation and background of three of the mentioned papers: the Bayesian average, the running of the weak mixing angle to low energies, and one

of the models of neutrino masses with discrete symmetries. Although a brief introduction of the content of the other two is also shown.

Table of Contents

	Page
1 Introduction	1
1.1 Particle physics: Historical perspective	1
1.1.1 Particle physics: The beginning	1
1.1.2 Particle Physics: The Standard Model	4
1.2 Modern Particle Physics: The search for new physics	4
2 Bayesian Inference	10
2.1 A simple introduction	10
2.2 Binomial model, the coin toss example.	12
2.3 Gaussian Case: Single parameter	16
2.3.1 Known Variance	16
2.3.2 Known central value	19
2.4 Unknown variance and central value	20
2.5 Gaussian Approximation	21
2.6 Model evaluation	22
2.6.1 PDG scaling	23
2.7 Hierarchical Modelling	25
2.7.1 Election of the prior	28
2.7.2 Equal errors case	28
2.7.3 Gaussian Approximation	32
2.7.4 Errors of different size.	33
2.8 Applied example: Neutron Lifetime	34
2.9 Comparison with other models	37
2.10 Other applications of the Hierarchical model	38
3 Standard Model	41
3.1 Definitions	42

3.2	Higgs Mechanism	44
3.3	Interaction with Fermions	46
3.4	Renormalization	49
3.4.1	Vacuum Polarization Function	50
3.4.2	Renormalization group equation in QED	52
3.4.3	Matching relations at one loop	56
3.4.4	Running of $\hat{\alpha}_s$	58
3.4.5	QCD contribution to the running of $\hat{\alpha}$	61
3.4.6	Breaking of perturbation theory, dispersion relations	63
3.5	The running of the weak mixing angle	67
4	The Weak Mixing Angle at Low Energies	75
4.1	Renormalization group evolution	77
4.1.1	Matching conditions	80
4.2	Implementation of experimental input	82
4.3	Singlet contribution	84
4.4	Flavor separation	88
4.4.1	Experimental data	89
4.4.2	Lattice data	91
4.4.3	Threshold masses	92
4.5	Theoretical uncertainties	96
4.6	Results	97
5	BSM physics and other topics	100
5.1	Neutrino physics	100
5.1.1	Origin of neutrino mass	100
5.1.2	Seesaw Formula	102
5.1.3	Neutrino Oscillations: Theory	104
5.1.4	Neutrino Oscillations: Experiments	108
5.1.5	The road map to construct neutrino models.	110
5.2	Previous works on the topic	111
5.2.1	The framework	112
5.2.2	Results	115
5.2.3	Phenomenology	119
5.2.4	Other textures: B_1 and B_2	120
5.2.5	Other works in the same direction.	121

6 Conclusions	122
--------------------------------	-----

Appendices

A Fisher's Information	126
---	-----

B Coefficients of the RGE	128
--	-----

B.1 Calculations of $\alpha(M_Z^2)$	130
---	-----

B.2 Calculation of $\Delta_{\text{disc}}\hat{\alpha}$	131
---	-----

List of Figures

1.1	Plot of the year vs the mass of the particle discovered.	2
1.2	Plot of the costs of super conducting magnets as the center of mass energy is increased. This plot was taken from [6].	3
2.1	Plot of the reduced χ^2 distribution as the number of degrees of freedom increase.	23
2.2	Ordinary averaging. We assume that the y_i are random outcomes of measurements of the <i>same</i> parameter θ	25
2.3	Hierarchical model. Each experimental parameter θ_i arises from a random draw from a parent distribution with hyper-parameters μ and τ , and each experimental central value y_i is then considered to be the result of a random draw from a Gaussian distribution with central value θ_i and error σ_i	26
2.4	Plot of the scale factor as a function of the reduced χ^2 for different values of N and α	29
2.5	Comparison of the exact result with the approximate formula for $\alpha = 0$	32
2.6	The measurement points with small error are shown in blue, the usual averages with the PDG scaling in red, and the hierarchical averages in black. The labels at the horizontal axis show by how many σ_M the blue points deviate from the gray point. The gray band represents the ordinary weighted averages of the bulk of measurements.	34

2.7	Neutron Lifetime. The blue points are beam experiments, while the green points are bottle experiments. The black points are the Bayesian averages, and the red points are the average with the PDG scaling. We did this averaging for the case when all experiments are included and also when only bottle experiments are taking into account.	35
3.1	Diagrams that contribute the the photon two point Green function.	50
3.2	Quark diagrams that contribute to the photon two-point Green's function.	61
3.3	Contour to obtain α from experimental results	64
3.4	$R(s)$ for different center of mass energy s	64
4.1	Examples of a connected (top) and a disconnected (bottom) Feynman diagram.	81
4.2	Scale dependence of the singlet contribution to $\Delta\alpha$ (solid line) and its step function approximation (dashed line).	87
4.3	Scale dependence of the weak mixing angle in the $\overline{\text{MS}}$ renormalization scheme. The dots indicate the scales where a particle is integrated out. The total uncertainty corresponds to the thickness of the line. The β -function of $SU(2)_L$ changes sign at $\mu = M_W$, where the fermionic screening effects of the effectively Abelian gauge theory are being overcompensated by the anti-screening effects of the full non-Abelian electroweak theory.	98
5.1	Fermions mass scales. The values of the neutrino masses are just shown for illustrative purpose taking into account the neutrino mass differences, the real value is still not known exactly.	101
5.2	Process that contributes to the neutrino-less double beta decay.	102
5.3	Loop diagram which generates a non-zero m_{ee} and consequently a small double beta decay rate.	115

List of Tables

2.1	Binomial simulation	14
3.1	Standard Model charge assignments.	46
4.1	RGE contributions of different particle types, where the minus sign is indicative for the asymptotic freedom in non-Abelian gauge theories.	78
4.2	Coefficients entering the higher order RGE for the weak mixing angle.	79
4.3	Channels associated with the strange quark external current (top) and possible further channels originating from it (bottom).	88
4.4	Theoretical uncertainties in the low energy mixing angle.	96
5.1	Texture zero matrices.	112
5.2	A_4 generators for the irreducible representations in the S diagonal basis, $\omega = e^{i2/3\pi}$ is the cubic root of 1.	113
5.3	Summary of the particle content and quantum numbers. $\alpha, \beta, \gamma, \delta$ and ϵ can be any of the singlet representations $\mathbf{1}, \mathbf{1}', \mathbf{1}''$	113

5.4	Particle transformation under A_4 that gives rise to texture zeros in the neutrino mass matrices. We get the same texture matrix if we exchange $\mathbf{1}' \leftrightarrow \mathbf{1}''$ in each line. In the same way we get the same texture if we exchange the representations of N_4 and N_5 . The last column gives the matrix type according to [173].	114
-----	--	-----

Chapter One

Introduction

This introduction gives a brief historical perspective of particle physics, the Standard Model, and the search for new physics. A more detailed introduction to concepts is given in each chapter.

1.1 Particle physics: Historical perspective

1.1.1 Particle physics: The beginning

There has been an incredible amount of progress in our understanding of nature during the last centuries. In particular, the comprehension of the behavior of things at small scales improved dramatically. The idea of a particle is a very old one (since the ancient Greeks). Nevertheless, a modern version of it was incorporated in the last centuries. The modern idea of fundamental particle has its roots in 1897 [1] when it was found that the cathode rays were particles instead of waves. The particles discovered were much lighter than any known atom or molecule, and they have a negative charge. This new particle was the electron ¹. The mass

¹The idea of the experiment is very simple, the first step is to separate the electrons from the atoms in a gas by applying an intense electric field with two charged plates. Once the ions are formed, the electrons will try to move to the positive charged plate and will pass through a small hole in it. After this step, an array of perpendicular electric and magnetic fields is placed in, and when the particles pass through them their path will change, so the idea is to tune the values of the electric and magnetic field in such a way that their effects cancel. Once the equations of motion are solved, the charge mass ratio can be written as a function of these fields.

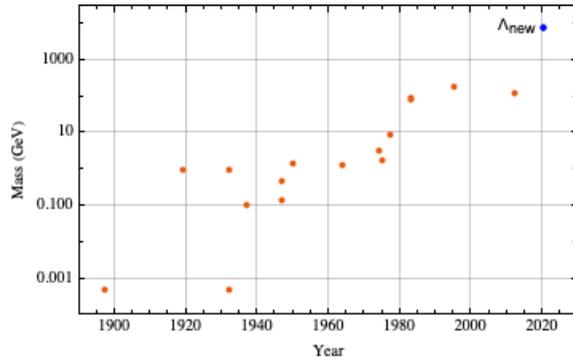


Figure 1.1: Plot of the year vs the mass of the particle discovered.

of the electron is very small: 0.51 MeV . A few years later, in 1919, E. Rutherford directed alpha particles into nitrogen [2]. The result of the collision was oxygen and some positive charged particles, with a much heavier mass than the electron. This was the discovery of the proton, which has a mass of 938.3 MeV . Later, in 1932, J. Chadwick studied the scattering of protons off from a paraffin wax target, when this was bombarded with an unknown source of radiation². He found that the unknown radiation was made of charge-less particles of nearly the same mass as the proton: 939.6 MeV . This new particle was the neutron [3]. In the same year the positron was discovered [4], which is the antiparticle of the electron and has exactly the same mass. The discovery of the positron is a very interesting one. This is because the positron was already proposed theoretically by Paul Dirac in 1928 [5], i.e., before the discovery. He was trying to construct an equation which incorporates quantum mechanics and special relativity, and in the process he found some negative energy states which he associate with antiparticles, or in other words the positrons.

Many more particles were found since then. In Fig. [1.1] each point in the plane represents a particle discovery. It is clear that we have been able to produce heavier and heavier particles as the technology increases. It is expected that new physics (as we call these new particles or interactions) should be somewhere in this plot. But two questions remain open: what kind of new physics are we going to find? and at what energy?.

²The unknown source of radiation was produced from the collision of alpha particles sent to a beryllium target.

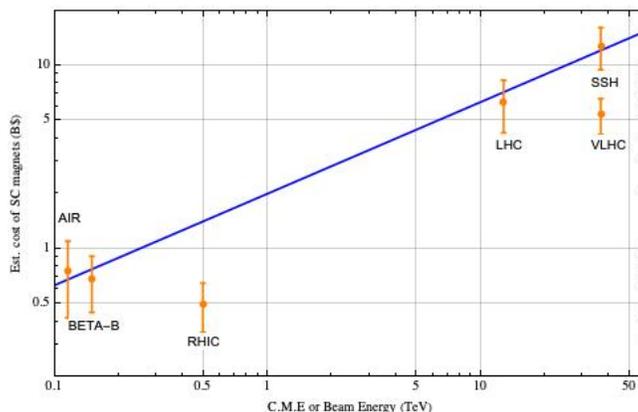


Figure 1.2: Plot of the costs of super conducting magnets as the center of mass energy is increased. This plot was taken from [6].

From the theoretical point of view there has also been amazing progress. In the early 1900s quantum mechanics was born and formalized, through works done by Einstein, Schroedinger, Pauli, Heisenberg [7–10] among many others. In the mid century, with the contribution of another generation of great physicists (Fermi, Feynman, Tomonaga, Schwinger, Gell-mann, Weinberg among others) a formal theory that included relativistic phenomena and quantum mechanics in the same framework was written, a framework known as Quantum Field Theory (QFT). A particular case of QFT is Quantum Electrodynamics (QED) which explains the interaction between light and matter. One of the earliest triumphs of QED was the computation of the correction to the magnetic moment of the electron. Dirac theory predicted it to be equal to two, but experimentally it was slightly different. Schwinger [11] computed the next order correction in perturbation theory of the anomalous magnetic moment in the QED framework. With this computation, theory and experiment agreed perfectly [12].

In parallel to the development of QED, Fermi [13] proposed a four fermion interaction for beta decay, which is the decay of a neutron into a proton, an electron, and an electron anti-neutrino. This model also had great success. This particular type of interaction (four fermion) is one of the simple examples of interactions that emerge as effective operators of a more complete theory. In this case the “complete theory” is the Standard Model (SM), and the four fermion interaction is the effective interaction of a W boson exchange between

a pair of fermions (nucleons and leptons).

1.1.2 Particle Physics: The Standard Model

The SM was constructed in the last century. It assumes that there is an $SU(3)_c \times SU(2)_L \times U(1)_Y$ symmetry, where the index correspond to color, left-handed fields and hyper-charge. The $SU(2)_L \times U(1)_Y$ symmetry was proposed by Weinberg and Salam [14, 15]. This part of the SM allowed to explain parity violating effects seen in parity violating experiments like the famous ^{60}Co Wu experiment [16], which did not have an explanation under the non-chiral model given by Fermi. It also incorporated the Higgs mechanism [17], which gives mass to the gauge bosons through a spontaneous breaking of the SM symmetry $SU(2)_L \times U(1)_Y \rightarrow U(1)_Q$ (where Q means the charge). The $SU(3)_C$ symmetry part, called Quantum Chromodynamics (QCD), concerns a symmetry that explains the interaction between quarks and gluons. Here spontaneous gauge symmetry breaking does not occur, which is why the 8 gluons are massless. The final pieces that entered the SM are the neutrino oscillations and masses. The idea of oscillations was suggested in 1967 by Pontecorvo [18]. The idea is that the states that interact with the gauge bosons (flavor eigenstates) are not the same as the states with definite mass. Therefore there is a non-zero superposition between the flavor eigenstates and the mass eigenstates. This, in turn, implies that neutrinos can oscillate. Even though the effect was predicted a long time ago it was not proven experimentally until the 90's [19].

1.2 Modern Particle Physics: The search for new physics

Even with all the theoretical comprehension that we have of the universe, there are still many open questions that need to be answered. One of them is the stability of the Higgs mass [20]. If there is any scale of new physics, then the quantum corrections to the Higgs mass would be proportional to this new physics scale. This implies that there should be a large tuning between the parameters and the scale of new physics to get just exactly the

Higgs mass. It is then more appealing that there exists an underlying mechanism that protects the mass in a natural way. Another open question is the nature of the dark matter. From astronomical observations it is clear that there should be some type of matter that interacts very weakly with the charged particles [21]. But we don't know exactly what is that particle(s) or the type of interactions that it has with the ordinary matter. Another question is the origin of the neutrino masses. There is a large difference between the scale of the neutrino masses and the scale of the rest of the SM particles. The question is whether there is an explanation to this atypical scale difference without using fine tuning. Another intriguing phenomena is the unification of the gauge couplings. It is interesting to see that as we run the three couplings of the SM up to high energies, they almost converge to the same value at some particular energy (around 10^{16} GeV). Other open questions are: are there more sources of charge-parity (CP) violation?, why θ_{QCD} is small?³, what is the nature of dark energy?.

There is a long list of models of new physics⁴ trying to explain some of the questions mentioned before, for example Supersymmetry [22, 23], String Theory [24], Grand Unification, discrete symmetries, WIMPS (Weakly interacting massive particle) or axions, among many others⁵. Each of them have their own well motivated theoretical or experimental support. It is also true that the community expected to find the new physics before (at not so high energies), but nature has been elusive.

A natural solution to the problem mentioned before is to create colliders which produce collisions at higher and higher energies. Nevertheless, as the energy is increased, experiments become more complicated and expensive as shown in Fig. [1.2] taken from [6]. It is here where a different and complementary approach can be taken, which is the precision approach: instead of looking for new interactions at high energy, the effects of new interactions

³ θ_{QCD} gives the size of boundary term that respects the QCD symmetry. This term is important because it contributes to physical quantities.

⁴These are generically called Beyond Standard Model models (BSM models).

⁵The references that I give here are not the original papers on these topics but reviews or books with a pedagogical perspective. More detailed references can be found in them.

are extracted from extremely precise low energy experiments. Thus we have two fronts to find new physics: the energy frontier (high energy colliders for example), and the precision frontier.

For example, a very well known experiment at the precision frontier is the measurement of the anomalous magnetic moment of the muon (denoted by a_μ), which (as in the case of the electron) measures the interaction of the muon with a magnetic field. Theoretically it is possible to predict very precisely a_μ . Nevertheless, the calculation is rather complicated, due to several reasons. The main one is that at low energies it is not possible to use a perturbative approach to compute physical quantities in QCD. This is because the QCD coupling $\hat{\alpha}_s$ increases at low energies, so perturbation theory breaks down. To overcome these complications different approaches need to be used, such as chiral perturbation theory, dispersion relations (which exploit the optical theorem) or Lattice QCD. The largest source of relative uncertainty in a comes from the so called Light-by-Light diagrams, while the largest total uncertainty comes from the hadronic vacuum polarization.

One of the things that will be discussed in this thesis are the radiative corrections and hadronic contributions to the weak mixing angle $\sin^2 \theta_W$ at low energies in the $\overline{\text{MS}}$ scheme, which is a combination of the gauge couplings of the $SU(2)_L \times U(1)_Y$ part of the SM. But, why is it important to get a precise value of the weak angle at low energies? Well, many experiments have and will measure parity violation at low energies. The physical quantities that these experiments measure can usually be related to the weak charge, and ultimately to the weak mixing angle. If there are some new interactions beyond the Standard Model that violate parity, it might possible to see them through these low energy experiments. In some cases, new physics models can be parametrized as four fermion operators. There the deviation of the experimental result from the Standard Model value can be mapped to this new physics scale [25, 26]. In other models such as a dark Z with kinetic mixing and extended mass matrix, there will be a strong shift of the weak mixing angle at low energies, which might not be seen at higher ones [27, 28].

In these low energy observables, the weak mixing angle is usually redefined as an effective running weak mixing angle. This weak mixing angle absorbs the contribution from the γZ mixing bubble. It is usually written as $\sin^2 \theta_W(Q^2) = \kappa(Q^2, M_Z^2) \sin^2 \hat{\theta}_W(M_Z^2)$. Where $\sin^2 \hat{\theta}_W(M_Z^2)$ is the weak mixing angle at the Z scale in the $\overline{\text{MS}}$ scheme. The "problem" with this formula is that $\kappa(Q^2, M_Z^2)$ will contain large logs when $Q^2 \approx 0$. If we want to reduce the uncertainty in $\sin^2 \theta_W(Q^2)$ we should properly sum these logs. To re-sum them we can use the relation $\sin^2 \theta_W(Q^2) = \kappa(Q^2, \mu^2) \sin^2 \hat{\theta}_W(\mu^2)$ where μ is the renormalization scale in the $\overline{\text{MS}}$ scheme⁶, and set $\mu^2 \approx Q^2$. Then we would have to compute $\sin^2 \hat{\theta}_W(0)$ through the renormalization group equations. This is what I do in this thesis.

In a different context, we should mention that a key point in the search for new physics is the combination of results from several experiments that measure a quantity that we will call θ . For example, the Particle Data Group (PDG) [29] uses results from different experiments and returns the combined value. Usually a simple weighted average is enough to get the combined value of the parameter of interest θ . Under the Gaussian approximation, with N experiments that have errors σ_i , the combined result will have an error (on θ) of

$$\sigma_{com} = \left(\sum_{i=1}^N \frac{1}{\sigma_i^2} \right)^{-\frac{1}{2}}, \quad (1.1)$$

(when $\sigma_i = \sigma$ for all i , the combined error is $\sigma_{com} = \frac{\sigma}{\sqrt{N}}$). Thus adding more and more experiments will increase our knowledge about the parameter θ . Nevertheless, if there is tension between the central values of the experiments (and we suspect that there might be an unknown systematic shift), this naive combination will underestimate the final errors. The PDG mitigates this problem multiplying the final error by a scale factor defined as

⁶The $\overline{\text{MS}}$ scheme is just a specific prescription to absorb the infinities in the parameters of the Lagrangian. We will discuss the proper definition later in the text.

$S_{PDG} \equiv \sqrt{\chi_{red}^2}$ for $\chi_{red}^2 > 1$ and non scaling for $\chi_{red}^2 < 1$, where

$$\chi_{red}^2 \equiv \frac{\chi_{min}^2}{N-1} = \frac{1}{N-1} \sum_{i=1}^N \frac{(y_i - \bar{y})^2}{\sigma_i^2} \quad (1.2)$$

is the minimum of the reduced χ^2 function. \bar{y} is the result of the average of the experimental values y_i . But even in this prescription there are some problems. For example, it shows an un-natural discontinuity in the derivative of the scale factor with respect to the reduced χ^2 . This creates an unnecessary dichotomy. The second problem is a bit more subtle. As shown in [30], the PDG scaling of the error is not a sufficient statistic to describe the entire probability distribution. As a consequence, the PDG scaling changes the final result depending on how the scale factor is computed. For example, suppose a subset of the experiments (points) is averaged and scaled with the PDG method. Now take this subset (now just a point) and combine it with the rest of the experiments, and apply the PDG scaling again. This process will give a different result compared to analyzing and scaling the whole set of points at the same time.

There have been several works trying to take into account the possibility of an underestimation of the error of the experiments.

In this thesis we present a method inspired by Bayesian meta-analysis studies. Meta-analysis is a statistical analysis that emerged in biological sciences, and tries to combine the results of multiple scientific studies. To understand this better take for example the study of the effects of some drug in human beings. There might be different experiments, in different parts of the world trying to estimate the effect of the same drug. The overall effect can in principle be obtained from the combination of all the experiments. But, we should accept that each lab might be measuring a different parameter, due to the differences in populations and laboratories. A way to take this into account is to assume that the parameters that each lab is estimating came from a random sample of a common hyper-population. This is the heart of the hierarchical model.

In the physics context, we used the same kind of model to take into account systematic shifts that could have not been taken into account by the experimental collaborations. We obtained the PDG scaling as a particular limit of this model and solved the problem of the averaging order.

Finally in this thesis I also discuss two models to generate neutrino mass. These models try to solve two of the problems that I mentioned before, the neutrino mass values and the origin of dark matter. The idea is to use a discrete flavor symmetry which is spontaneously broken. This symmetry will leave some patterns in the neutrino mass matrix (texture zeros), which can be used to find correlations between the oscillation parameters. At the same time the remaining symmetry protects some new particles to decay, making them a possible source of dark matter.

The structure of the thesis will go from the general knowledge to specific cases. In chapter 2, I discuss Bayesian statistics. I re-derived, within the Bayesian framework, some widely used formulas by the physics community. Then I explain the Bayesian Hierarchical Model, and show its results and consequences. In chapter 3, I review some basic concepts about the Standard Model and renormalization. Chapter 4 is about the weak mixing angle and its running to low energies. Chapter 5 contains Beyond Standard Model (BSM) models for neutrino oscillations and mass hierarchy. Chapter 6 contains my conclusions and outlook.

Chapter Two

Bayesian Inference

2.1 A simple introduction

The Bayesian¹ approach was probably born long ago in 1763, when the paper “*An Essay towards solving a Problem in the Doctrine of Chances*” written by Thomas Bayes was published posthumously [32]. There Bayes gave the first detailed description of the well known basic theorem of probability that has his name: “Bayes’ Theorem”. In its discrete form Bayes’ Theorem reads:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_i P(A|B_i)P(B_i)}. \quad (2.1)$$

Explained in words, this equation tells that the probability that the event B_i occurs *given* A is proportional to the probability that A occurs given B_i multiplied by the “prior” probability of B_i . This simple and easy to understand formula has had a tremendous impact² in the scientific community [33–35] and in the technology sector [36, 37]. A simple frequentist proof can be given as follows. Suppose you have two types of objects, spheres and cubes, $\{s, c\}$. These objects might be red or blue, and the total number of objects is $n_{red} + n_{blue} = n$ where n_{red} is the number of red objects and n_{blue} is the number of blue objects. Similarly,

¹The origin of the word “Bayesian” used as way to identify a complete branch of statistics has a long history, for details see [31].

²The references given here are just examples of applications and general books on the subject, the bibliography is vast and the reader is invited to look up those references to find more applications and detailed explanations.

the number of spheres (cubes) with a specific color can be labeled as $n_{red}^{s(c)}$. If an object is chosen at random, the probability of choosing a sphere will be $P(s) = \frac{n^s}{n}$. Now, *given* that a sphere is chosen, the probability of choosing a red sphere is $P(r|s) = \frac{n_{red}^s}{n^s}$. On the other hand, if we do not care about the shape of the object, then the probability of choosing a red object is $P(r) = \frac{n_{red}}{n}$. Finally, the probability of choosing a red object that also is a sphere must be $P(r, s) = \frac{n_{red}^s}{n}$. Thus we have

$$P(r, s) = \frac{n_{red}^s}{n} = \frac{n_{red}^s n^s}{n^s n} = P(r|s)P(s), \quad (2.2)$$

this argument can be done again but now changing $r \leftrightarrow s$. Comparing this results, we easily get

$$P(r|s) = \frac{P(s|r)P(r)}{P(s)} = \frac{P(s|r)P(r)}{P(s|r)P(r) + P(s|b)P(b)}, \quad (2.3)$$

which is Bayes' theorem. This theorem can be generalized to continuous variables, where sums get replaced by integrals,

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}. \quad (2.4)$$

This is a good point to start with our notation. In the following, y will represent a measurement i.e., the result of a coin flip, the value of a card, the neutron lifetime measured by a experimental collaboration etc. θ is a parameter that we want to infer, for example the probability that a flip coin will return head or tail, or the true underlying value of the neutron lifetime. $p(y|\theta)$ is known as the *likelihood*, $p(\theta)$ the *prior* and $p(\theta|y)$ is called the *posterior distribution*. The factor $\int p(y|\theta)p(\theta)d\theta$ is a normalization factor that does not depend on θ , and is usually omitted, so we write

$$p(\theta|y) \propto p(y|\theta)p(\theta), \quad (2.5)$$

keeping in mind that to find the normalized distribution an integral over θ must be done. In the next section I will present some typical examples on parameter inference. The first one is one of the oldest applications of Bayesian inference. It consists on making a prediction of the probability that a coin is biased given some measurements. Then I will move to the Gaussian case, which is one of the most important cases in scientific research. This Gaussian case will be analyzed in the single parameter (known variance) model and also when both variance and central value are unknown. Later I will introduce hierarchical modeling and the results that we obtained compared to the PDG. Many of the basic concepts and notation are borrowed from [35]. Since my intention here is to be as brief and clear as possible, I might omit some proofs and profound discussions that can be found in the literature. If the reader is interested, further references and details can be found in [35].

2.2 Binomial model, the coin toss example.

This is one of the oldest applications of Bayesian inference [32]. Suppose we flip a coin n times and save the number of head and tails that appear. Let us save this in a variable y which is equal to the number of tails that appear. If $n = 1$ then the probability of having a tail given θ is

$$p(\text{tail}|\theta, n = 1) = \theta, \tag{2.6}$$

if we flip it two times, then the probability of having one tail will be

$$p(\text{tail}|\theta, n = 2) = \theta(1 - \theta) + (1 - \theta)\theta = 2(1 - \theta)\theta, \tag{2.7}$$

the factor of two appears because the tail could have appeared in the first flip of the coin or in the second one. On the other hand, the probability of having two tails in two coin flips is $p(\text{tail}, \text{tail}|\theta, n = 2) = \theta^2$. This can be easily generalized to an arbitrary n number of coin

flips and tail results y . The probability distribution is then

$$p(y|\theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y}, \quad (2.8)$$

which is the binomial distribution. In this derivation we have used an important concept of probability theory, the notion of independent measurements. That is to say that the occurrence of one experiment does not affect the probability of occurrence of the other. This means that the *true* value of θ does not change after one experiment is done³. Using Bayes theorem the posterior distribution of θ given the data is

$$p(\theta|y) \propto \theta^y (1 - \theta)^{n-y} p(\theta). \quad (2.9)$$

If the prior is assumed to be uniform $p(\theta) = 1$, the normalization factor can be obtained analytically. Thus we get

$$p(\theta|y) = (n + 1) \binom{n}{y} \theta^y (1 - \theta)^{n-y}. \quad (2.10)$$

Suppose that we want to know if our coin is biased. I will define a coin as biased if the probability of θ being larger or smaller than 0.5, is more than 95%. To make this example more real let us simulate some random flips with known θ . Then we will pretend that we do not know its value and try to infer it from those measurements (flips). We summarize the results in Tab. [2.1]. The first three rows of the table show the simulations from $\theta_{true} = 0.8$. We see that as the number of experiments n becomes larger, the posterior 95% confidence interval for θ shrinks around 0.8. Already at $n=20$ we can say that the coin is biased. In a similar way the posterior shrinks around $\theta = 0.5$, for $\theta_{true} = 0.5$, and we can not conclude

³It is interesting though that *our knowledge* about the parameter θ does change after one experiment is done, this is the heart of Bayesian statistics.

θ_{true}	n	y	$\max(\{P(\theta > 0.5 y), P(\theta < 0.5 y)\})$	θ_{mean}	95% confidence
0.8	5	3	0.656	0.571	[0.243,0.899]
0.8	10	7	0.887	0.667	[0.417,0.917]
0.8	20	16	0.996	0.772	[0.607,0.939]
0.8	100	82	0.999	0.814	[0.739,0.889]
0.5	5	3	0.656	0.571	[0.243,0.899]
0.5	10	3	0.887	0.333	[0.083,0.583]
0.5	20	13	0.905	0.636	[0.441,0.831]
0.5	100	52	0.655	0.52	[0.422,0.617]

Table 2.1: Binomial simulation

that the coin is biased.

An interesting and common question that frequently emerges is: what happens if we change our prior? what prior should we use? Well, it depends on our knowledge about the problem. Frequently a *conjugate prior* is used. This is defined as a prior such that, when combined with the likelihood, will give a posterior with the same form as the likelihood itself. In this case, the conjugate prior is

$$p(\theta|\alpha, \beta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad (2.11)$$

note that the flat prior is just a special case where $\alpha = \beta = 1$. The posterior distribution of θ using this prior will be

$$p(\theta|y, n, \alpha, \beta) \propto \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1}. \quad (2.12)$$

Changing the value of α and β modifies our prior believe about the true distribution of θ . Nevertheless, it is important to mention that if enough data is included i.e., if $y \gg \alpha$ and $n \gg \beta$, then data will dominate over the prior used beforehand. Although this is an obvious result from this example, it is a general statement for other distributions⁴ [38, 39]. When the number of experiments is small, the prior will have more and more weight. It is here

⁴This is closely related to the central limit theorem, for a pedagogical derivation see Appendix B of [35].

where a smart election of prior can give powerful inferences. Sometimes, we do not have any previous knowledge of the problem, so we want to be as agnostic as possible about the prior that we choose, in order to avoid an unwanted bias in the posterior distribution. One might argue that a flat distribution always gives a non-informative prior distribution. Nevertheless, there is an ambiguity here, because under a change of variables a flat prior distribution will not be flat. To overcome this issue, a formal definition of a non informative prior can be constructed. Jeffrey [40, 41] defined a prior in terms of Fisher's information $\mathcal{I}(\theta)$, which in turn is defined as

$$\mathcal{I}(\theta) \equiv \int p(y|\theta) \left(\frac{d \ln p(y|\theta)}{d\theta} \right)^2 dy. \quad (2.13)$$

Jeffrey's prior is then $p(\theta) \equiv \sqrt{\mathcal{I}(\theta)}$. It is straightforward to show that Jeffrey's prior is invariant under a change of variables,

$$\begin{aligned} p(\phi) &= p(\theta) \frac{d\theta}{d\phi} \\ &= \sqrt{\int p(y|\theta) \left(\frac{d \ln p(y|\theta)}{d\theta} \right)^2 dy} \frac{d\theta}{d\phi} \\ &= \sqrt{\int p(y|\phi) \left(\frac{d \ln p(y|\theta)}{d\phi} \right)^2 dy} \\ &= \sqrt{\mathcal{I}(\phi)}. \end{aligned}$$

In the particular case of the binomial distribution, we can use an alternative form of Fisher's information to obtain Jeffrey's prior. This alternative form is⁵

$$\mathcal{I}(\theta) \equiv - \int p(y|\theta) \left(\frac{d^2 \ln p(y|\theta)}{d\theta^2} \right) dy. \quad (2.14)$$

⁵The proof is shown in appendix A

So we get

$$\mathcal{J}(\theta) = \sum_{y=0}^n \binom{n}{y} y \theta^{y-2} (1-\theta)^{n-y} dy + \sum_{y=0}^n \binom{n}{y} (n-y) \theta^y (1-\theta)^{n-y-2} dy. \quad (2.15)$$

With a few algebraic steps we arrive to

$$p(\theta) = \sqrt{\mathcal{J}} \propto \theta^{-\frac{1}{2}} (1-\theta)^{-\frac{1}{2}}, \quad (2.16)$$

which in our previous notation can be mapped to $\alpha = 1/2$ and $\beta = 1/2$. The analysis that we did with the flat prior can be done using this prior too. After doing so, one can easily see that as the number of data point increases, the prior become less and less relevant. For example, for the case of $\theta_{true} = 0.8$ and $n = 5$, $y = 3$ we get $[0.237, 0.929]$ as the 95% confidence interval, while for the mean we get 0.583. On the other hand, for $n = 100$ and $y = 82$ we get $[0.742, 0.892]$ as the 95% confidence interval, while for the mean we get 0.817. If we compare this results with Tab. 2.1 we see that the effect of the prior is smaller when the amount of data is larger.

2.3 Gaussian Case: Single parameter

2.3.1 Known Variance

We can now move to the Gaussian case. This is by far one of the most important distributions in science. One of the reasons for this is that it emerges naturally due to the Central limit theorem [42, 43]. The distribution is usually labeled as $\mathcal{N}(y|\theta, \sigma^2)$ where

$$\mathcal{N}(y|\theta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2} \frac{(y-\theta)^2}{\sigma^2}\right]. \quad (2.17)$$

In the most general case (which we will analyse in the next section) both σ^2 and θ are unknown. Here, we assume that σ is known, so we make an inference on θ . From the definitions and procedure mentioned in previous sections, we know that the posterior distribution for θ is

$$p(\theta|y) \propto \mathcal{N}(y|\theta, \sigma^2)p(\theta). \quad (2.18)$$

The conjugate prior for this posterior is $\mathcal{N}(\theta|\mu, \tau^2)$. With simple algebraic steps, and normalizing the posterior distribution, it can be shown that the posterior is $\mathcal{N}(\theta|\hat{\theta}, \hat{\sigma}^2)$ where

$$\frac{1}{\hat{\sigma}^2} = \frac{1}{\sigma^2} + \frac{1}{\tau^2}, \quad (2.19)$$

and

$$\hat{\theta} = \left(\frac{1}{\sigma^2} + \frac{1}{\tau^2} \right)^{-1} \left(\frac{y}{\sigma^2} + \frac{\mu}{\tau^2} \right). \quad (2.20)$$

Thus we see that the precision (defined as the inverse of the variance) increases after new data is included. Now suppose that we have n independent measurements $\{y_1, y_2, \dots, y_n\}$. Due to the independence of the experiments, the posterior distribution can be written as

$$p(\theta|y_i) \propto \left[\prod_{j=1}^n \mathcal{N}(y_j|\theta, \sigma^2) \right] \mathcal{N}(\theta|\mu, \tau^2), \quad (2.21)$$

and the posterior distribution in this case will also be a Gaussian $\mathcal{N}(\theta|\hat{\theta}, \hat{\sigma}^2)$ where

$$\frac{1}{\hat{\sigma}^2} = \sum_{j=1}^n \frac{1}{\sigma_i^2} + \frac{1}{\tau^2}, \quad (2.22)$$

and

$$\hat{\theta} = \left(\sum_{j=1}^n \frac{1}{\sigma_i^2} + \frac{1}{\tau^2} \right)^{-1} \left(\sum_{j=1}^n \frac{y_i}{\sigma_i^2} + \frac{\mu}{\tau^2} \right). \quad (2.23)$$

In this case we have used a conjugate prior to obtain the posterior distribution. It is intuitive that as we make τ^2 larger and larger, the prior would give less and less information⁶ to the system. In the limit when $\tau^2 \rightarrow \infty$ the prior distribution is completely flat, and is non-informative. Usually, in parameter inference this is implicitly assumed. This is why the combined precision of a set of experiments is just the sum of experimental precisions and the posterior central value is the precision weighted average. We can also show that the flat ($\tau^2 \rightarrow \infty$) is a non-informative prior through Jeffrey's method, because:

$$\frac{d^2 \ln p(y|\theta)}{d\theta^2} = -\frac{1}{\sigma^2}, \quad (2.24)$$

and

$$\mathcal{J}(\theta) = - \int p(y|\theta) \left(\frac{d^2 \ln p(y|\theta)}{d\theta^2} \right) dy = \frac{1}{\sigma^2} \int p(y|\theta) dy = \frac{1}{\sigma^2}, \quad (2.25)$$

so $p(\theta) = \sqrt{\mathcal{J}(\theta)} = \text{constant}$. An important property of the Gaussian distribution is that the likelihood can be written as

$$p(y_i|\theta) = \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \right] \exp \left[-\frac{1}{2} \chi^2(\theta) \right], \quad (2.26)$$

where

$$\chi^2(\theta) = \sum_{i=1}^n \frac{(\theta - y_i)^2}{\sigma_i^2} = \chi_{min}^2(\hat{\theta}) + \frac{(\theta - \hat{\theta})^2}{\hat{\sigma}^2} \quad (2.27)$$

is the χ^2 function. This quadratic form of the χ^2 shows the idea behind the minimization of the χ^2 in a typical problem of parameter inference. The central value $\hat{\theta}$ of the posterior distribution can be obtained minimizing the χ^2 function with respect to θ .

⁶This statement can actually be formalized in terms of information theory. We do not discuss this any further, but the interested reader is invited to see [44].

2.3.2 Known central value

Previously we were discussing the case of several experiments measuring the same parameter θ . Now we will compute the case of known central value but unknown variance. For this case, we will restrict to a single experiment. In this hypothetical experiment, the central value θ is known but not σ^2 . Now suppose that we have n observations from the same experiment and that these observations follow a Gaussian distribution. Thus, given both σ and θ , the likelihood is

$$p(y|\sigma^2) \propto \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \theta)^2 \right] = (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{n}{2\sigma^2} v \right] \quad (2.28)$$

where

$$v \equiv \frac{1}{n} \sum_{i=1}^n (y_i - \theta)^2 \quad (2.29)$$

A conjugate prior on σ^2 , will be⁷

$$p(\sigma^2) \propto \left(\frac{\sigma_0^2}{\sigma^2} \right)^{\frac{\nu_0}{2} + 1} e^{-\frac{\nu_0 \sigma_0^2}{2\sigma^2}}, \quad (2.30)$$

Using this prior, the posterior distribution for σ^2 reads

$$p(\sigma^2|y) \propto (\sigma^2)^{-\frac{(n+\nu_0)}{2} + 1} \exp \left[-\frac{1}{2\sigma^2} (\nu_0 \sigma_0^2 + nv) \right], \quad (2.31)$$

this is just an Inv- χ^2 distribution with $\nu_0 + n$ degrees of freedom and scale given by $\frac{\nu_0 \sigma_0^2 + nv}{\nu_0 + n}$, namely

$$p(\sigma^2|y) = \text{Inv-}\chi^2 \left(\nu_0 + n, \frac{\nu_0 \sigma_0^2 + nv}{\nu_0 + n} \right). \quad (2.32)$$

⁷A non informative prior has $\nu_0 = 0$.

2.4 Unknown variance and central value

Now let us jump to the next case: a single experiment where both σ^2 and θ are unknown. This is an example of a multi-parameter problem. We use the likelihood for a set of n Gaussian i.i.d points, a flat prior for θ , and a prior with $\nu_0 = 0$ for σ^2 . After some straightforward algebraic steps, it can be shown that the likelihood has the form

$$p(\theta, \sigma^2 | y) \propto \sigma^{-n-2} \exp \left[-\frac{1}{2\sigma^2} \left((n-1)s^2 + n(\bar{y} - \theta) \right) \right], \quad (2.33)$$

where

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2, \quad (2.34)$$

is the sample variance, and \bar{y} is the sample mean. Marginalizing over σ^2 a posterior distribution for μ given the data can be obtained. The result is:

$$p(\mu | y) = \int p(\mu, \sigma^2 | y) d\sigma^2 \propto \left[1 + \frac{n(\mu - \bar{y})^2}{(n-1)s^2} \right]^{-\frac{n}{2}}, \quad (2.35)$$

this is just the $t_{n-1}(\bar{y}, s^2/n)$ density. In the very same way a posterior distribution for the parameter σ^2 can be found, the result is

$$p(\sigma^2 | y) \propto (\sigma^2)^{-\frac{(n+1)}{2}+1} \exp \left[-\frac{(n-1)s^2}{2\sigma^2} \right], \quad (2.36)$$

which is scaled inverse- χ^2 density. By definition this distribution is the distribution of the inverse of a variable that has a χ^2 distribution. Then, $\frac{(n-1)s^2}{\sigma^2}$ has a χ_{n-1}^2 , which is a χ^2 distribution with $n-1$ degrees of freedom.

2.5 Gaussian Approximation

In the last sections we have seen that it is easy to manipulate Gaussian distributions. In particular, it is an incredible useful property that when new experiments are included, the final precision of the combination of this experiment with another set of experiments can be obtained adding the precisions in quadrature. As I mentioned before, Gaussian distributions emerge naturally from the central limit theorem. Nevertheless, sometimes the distributions are not exactly a Gaussian. In this cases the formal methodology is to perform the integrals over the posterior distributions to find confidence intervals for the parameters. This formal process can be computationally demanding, so if the distributions have a shape similar to a Gaussian (like for example a t-student for a large number of degrees of freedom) it is better to approximate the distribution to a Gaussian. In this section we show how this is done. Suppose we have the probability distribution of a random variable x . Then a Gaussian approximation can be obtained by expanding the logarithm of the distribution to second order around the maximum of the likelihood. Explicitly

$$\ln p(x) = \ln p(x_0) + \frac{d \ln p(x)}{dx} (x - x_0) + \frac{1}{2!} \frac{d^2 \ln p(x)}{dx^2} (x - x_0)^2 + \frac{1}{3!} \frac{d^3 \ln p(x)}{dx^3} (x - x_0)^3 + \dots \quad (2.37)$$

The first term is just a constant, the first derivative vanishes when we expand around the peak, so the quadratic term is the one that gives the Gaussian distribution. Comparing this to the logarithm of a Gaussian distribution we immediately deduce that $p(x)$ can be approximated as $\mathcal{N}\left(x|x_0, -\left[\frac{d^2 \ln p(x)}{dx^2}\right]^{-1}\right)$.

As an example, let us take the Student-t distribution for n degrees of freedom. The logarithm of this distribution is

$$\ln p(x|x_0, \sigma^2) = C - \frac{n+1}{2} \ln \left[1 + \frac{(x-x_0)^2}{\sigma^2} \right], \quad (2.38)$$

where C is the normalization constant. The maximum of this distribution is easily obtained

and corresponds to $x = x_0$. The second derivative at this point is then

$$\frac{d^2 \ln p(x|x_0, \sigma^2)}{dx^2} = -\frac{(1+n)}{n} \frac{1}{\sigma^2}. \quad (2.39)$$

Then we can approximate $t_n(x|x_0, \sigma^2) \sim \mathcal{N}(x|x_0, \frac{n}{n+1}\sigma^2)$. The error on this approximation can be obtained computing the next term in the series. The third derivative vanishes, so the next term would be the fourth derivative which goes as $1/n$ so it is clearly suppressed for large n . In general, the power expansion of this distribution has vanishing odd derivatives. So only even derivatives contribute, and go as $n^{1-l/2}$ where l is the order of the derivative. Then we conclude that for large n we get closer and closer to a Gaussian distribution.

2.6 Model evaluation

There is still an important part that is missing: the evaluation of the model. After computing the parameters, it is a good practice to estimate how likely it is that the sample distribution came from the proposed parent distribution. To do so, the next step is to define a test statistic which is a scalar function of the measured values y_i . Then the p-value is computed, which is defined as the probability that the test statistic is larger than the measured one. For the Gaussian case, a typical test statistic is the χ^2 function. Using Cochran's theorem, it can easily be shown that the χ^2 function for n experiments with known central value follow a χ^2 distribution with n degrees of freedom. On the other hand, the χ^2 function for n measurements with central value approximated by the weighted average follows a χ^2 distribution function with $n-1$ degrees of freedom. The change of n to $n-1$ is called Bessel's correction and it comes from the use of a function of the measurements to infer the central value. This effect also shows up in the bayesian case, compare for example the posterior distribution of σ^2 for the case of known and unknown central values.

It is sometimes convenient to use the distribution of the variable $\chi_n^2 \equiv \chi^2/(n-1)$ instead

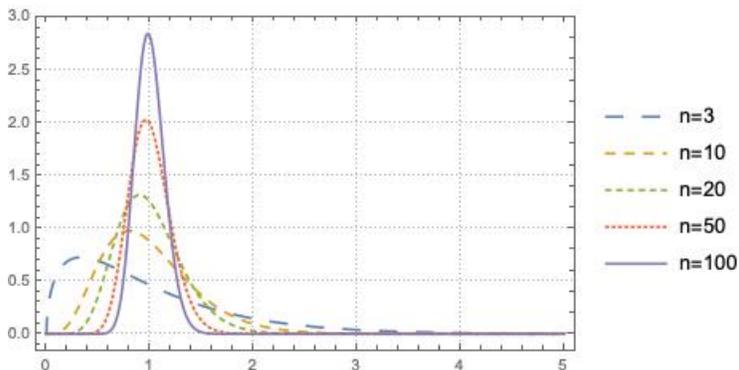


Figure 2.1: Plot of the reduced χ^2 distribution as the number of degrees of freedom increase.

of χ^2 . The variable χ_n^2 is called the reduced χ^2 . In Fig. 2.1 the χ_n^2 distribution is shown for different values of n . Note that as the number of degrees of freedom increases, the distribution becomes more and more peaked around one. Thus, if a set of measurements returns a reduced χ^2 that is greater than one for a large number of experiments, we might suspect that data can not be described by the model. The PDG knows about this issue and proposes a solution: scale the errors of all experiments by the same amount, in such a way that the reduced χ^2 is one. In the next subsection we describe the PDG methodology.

2.6.1 PDG scaling

The PDG [29] methodology to compute the average of several experiments is the following:

"...To average data, we use a standard weighted least-squares procedure and in some cases, discussed below, increase the errors with a "scale factor." We begin by assuming that measurements of a given quantity are uncorrelated, and calculate a weighted average and error as

$$\bar{x} + \delta\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \pm \left(\sum_i \right)^{-\frac{1}{2}} \quad (2.40)$$

where

$$w_i = \frac{1}{(\delta x_i)^2} \quad (2.41)$$

Here x_i and δx_i are the value and error reported by the i -th experiment, and the sums run

over the N experiments. We then calculate $\chi^2 = \sum_{i=1} w_i (\bar{x} - x_i)$ and compare it with $N-1$, which is the expectation value of χ^2 if the measurements are from a Gaussian distribution. If $\chi^2/(N-1)$ is less than or equal to 1, and there are no known problems with the data, we accept the results. If $\chi^2/(N-1)$ is very large, we may choose not to use the average at all. Alternatively, we may quote the calculated average, but then make an educated guess of the error, a conservative estimate designed to take into account known problems with the data. Finally, if $\chi^2/(N-1)$ is greater than 1, but not greatly so, we still average the data, but then also do the following:

We increase our quoted error, x in Eq. (1), by a scale factor S defined as $S = \sqrt{\chi^2/(N-1)}$

Our reasoning is as follows. The large value of the χ^2 is likely to be due to underestimation of errors in at least one of the experiments. Not knowing which of the errors are underestimated, we assume they are all underestimated by the same factor S . If we scale up all the input errors by this factor, the χ^2 becomes $N-1$, and of course the output error δx scales up by the same factor. When combining data with widely varying errors, we modify this procedure slightly. We evaluate S using only the experiments with smaller errors. Our cutoff or ceiling on δx_i is arbitrarily chosen to be $\delta_0 = 3N^{\frac{1}{2}}\delta\bar{x}$, where $\delta\bar{x}$ is the unscaled error of the mean of all the experiments. Our reasoning is that although the low-precision experiments have little influence on the values \bar{x} and $\delta\bar{x}$, they can make significant contributions to the χ^2 , and the contribution of the high-precision experiments thus tends to be obscured. Note that if each experiment has the same error δx_i , then $\delta\bar{x}$ is $\delta x_i/N^{\frac{1}{2}}$, so each δx_i is well below the cutoff. (More often, however, we simply exclude measurements with relatively large errors from averages and fits: new, precise data chase out old, imprecise data.) Our scaling procedure has the property that if there are two values with comparable errors separated by much more than their stated errors (with or without a number of other values of lower accuracy), the scaled-up error $\delta\bar{x}$ is approximately half the interval between the two discrepant values. We emphasize that our scaling procedure for errors in no way affects central values. And if you wish to recover the unscaled error x , simply divide the quoted error by $S...$ "

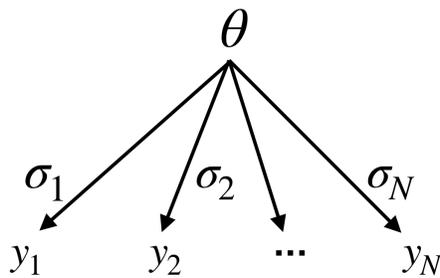


Figure 2.2: Ordinary averaging. We assume that the y_i are random outcomes of measurements of the *same* parameter θ .

Although this method provides a way to include cases where the χ^2 is large, it has several drawbacks. First of all it creates an unnecessary dichotomy around $\sqrt{\chi^2/(N-1)} \approx 1$, i.e., the derivative of the scale factor is discontinuous around this value. Another issue, as is pointed out by [30], is that the final central value and final errors depend on how the experiments are combined. For example, imagine that there is a sub-set of experiments that are combined with the PDG methodology and converted to a single point(experiment). If we now combine this result with the rest of experiments (under the PDG methodology) we would get a different result than if we combine the full set of experiments with the PDG methodology from the beginning.

In the next section I will explain the results of our publication. I will put particular attention to those steps that are omitted in the original publication, while things that are clearly explained and computed there will not be shown here. For further details, the publication is included at the end of this chapter.

2.7 Hierarchical Modelling

The following sections of this chapter are a modified version of the paper that we published in [45]. We proposed a Bayesian Hierarchical Model to solve the issues that the PDG methodology has. Bayesian Hierarchical Models had been used previously in other areas, particularly in biological sciences [46–48]. The idea is to assume that the the experimental

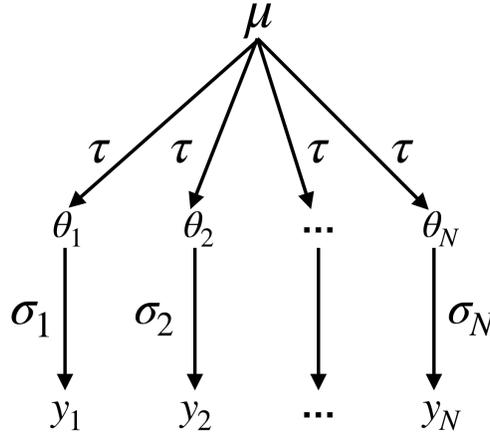


Figure 2.3: Hierarchical model. Each experimental parameter θ_i arises from a random draw from a parent distribution with hyper-parameters μ and τ , and each experimental central value y_i is then considered to be the result of a random draw from a Gaussian distribution with central value θ_i and error σ_i .

values y_i are not random samples from the same parameter θ as shown in Fig.2.2. But instead that each experiment is measuring a different parameter θ_i , which in turn comes from a random sample of a common hyper-distribution among all the experiments as can be seen in Fig. 2.3. Mathematically, this is equivalent to using as a prior

$$p(\theta_1, \theta_2, \dots) = \int [p(\theta_1|\mu, \tau^2)p(\theta_2|\mu, \tau^2)\dots] p(\mu, \tau)d\mu d\tau^2, \quad (2.42)$$

the posterior distribution is

$$p(\vec{\theta}|\vec{y}, \vec{\sigma}) \propto \left[\prod_{i=1}^N p(y_i|\theta_i) \right] p(\theta_1, \theta_2, \dots). \quad (2.43)$$

Basic probability theory⁸ allows us to write,

$$p(\mu, \tau^2, \vec{\theta}|\vec{y}, \vec{\sigma}) \propto \left[\prod_{i=1}^N p(y_i|\theta_i)p(\theta_i|\mu, \tau^2\theta_i) \right] p(\mu, \tau^2). \quad (2.44)$$

⁸Here we just used $p(a) = \sum_b p(a|b)p(b)$ and $p(a, b) = p(a|b)p(b)$.

For the Gaussian case we have (assuming a flat prior for μ)

$$p(\mu, \tau^2, \vec{\theta} | \vec{y}, \vec{\sigma}) \propto \left[\prod_{i=1}^N \mathcal{N}(y_i | \theta_i) \mathcal{N}(\theta_i | \mu, \tau^2) \right] p(\tau^2). \quad (2.45)$$

We take μ as the overall common effect of the experiments, and it is indeed our quantity of interest. i.e., the real parameter that we want to infer. The idea is that there is a possibility that the experimental collaborations did not take into account possible systematic shifts in their measurements, not because they were not careful, but because there is always a possibility of an unaccounted systematic shift. These shifts have the net effect of making each experiment to measure a different parameter θ_i . On the other hand, we know that they are indeed trying to measure a common quantity μ . It is important to point out that the spread of μ , has two contributions, one which comes from the experimental errors σ_i , and another one which measures the difference among the experimental collaborations given by τ^2 . In other words, if we increase the value of σ_i for each experiment, then the error on μ will also increase. Similarly, if we increase the discrepancy between the experiments leaving σ_i constant, the error on μ will also increase. To see this, look to the marginalized posterior distribution of μ and τ^2 after integrating over each θ_i , this is

$$p(\mu, \tau^2 | \vec{y}) \propto \prod_{i=1}^N \mathcal{N}(\mu | y_i, \sigma_i^2 + \tau^2) p(\tau^2). \quad (2.46)$$

So we have a sum of two terms, the usual σ_i^2 contribution and the τ^2 that is common to all experiments. The posterior distribution of τ is highly controlled by the discrepancy of the experiments, so a higher discrepancy between experiments will imply a probability distribution that allows large values of τ , and in consequence a larger error on μ .

Integrating over τ^2 , would give the posterior distribution of μ given the data. It is clear that if $p(\tau^2)$ is sharply peaked around zero, then the posterior distribution for the non

hierarchical case is recovered,

$$p(\mu|\vec{y}) \propto \int \prod_{i=1}^N \mathcal{N}(\mu|y_i, \sigma_i^2 + \tau^2) \delta(\tau^2) d\tau^2 = \prod_{i=1}^N \mathcal{N}(\mu|y_i, \sigma_i^2). \quad (2.47)$$

Thus, the distribution of τ^2 encodes the χ^2 value of the usual non-hierarchical fit, and turns it back as a source of uncertainty. To compare our results with the PDG prescription we define the scale factor (S_b) as the ratio of the sizes of the 68% highest confidence intervals of the hierarchical and non-hierarchical models.

2.7.1 Election of the prior

The next step is to choose a prior that is convenient for our problem. We could use Jeffrey's method to get a non informative prior, but in this case we would like to have something that is weakly informative and also simplifies the computations. Another interesting property would be to have a prior that can naturally be peaked around zero, due to the arguments of the previous paragraphs. We found that a plausible option is

$$p(\tau^2) \propto \prod_{i=1}^N \frac{1}{(\sigma_i^2 + \tau^2)^{\frac{\alpha}{2N}}}. \quad (2.48)$$

If we want the prior to have a defined variance, (i.e., the variance to be finite) then $\alpha > 6$, which is a reasonable supposition. We computed the posteriors using different values of α . We will analyse its effects in the next subsections.

2.7.2 Equal errors case

Suppose that all the experiments have similar errors $\sigma_i \sim \sigma$. Then the posterior distribution using the prior mentioned before reads

$$p(\mu, \tau^2|\vec{y}) \propto \frac{1}{(\sigma^2 + \tau^2)^{\frac{N+\alpha}{2}}} \exp \left[-\frac{1}{2} \sum_{i=1}^N \frac{(\mu - y_i)^2}{\sigma^2 + \tau^2} \right], \quad (2.49)$$

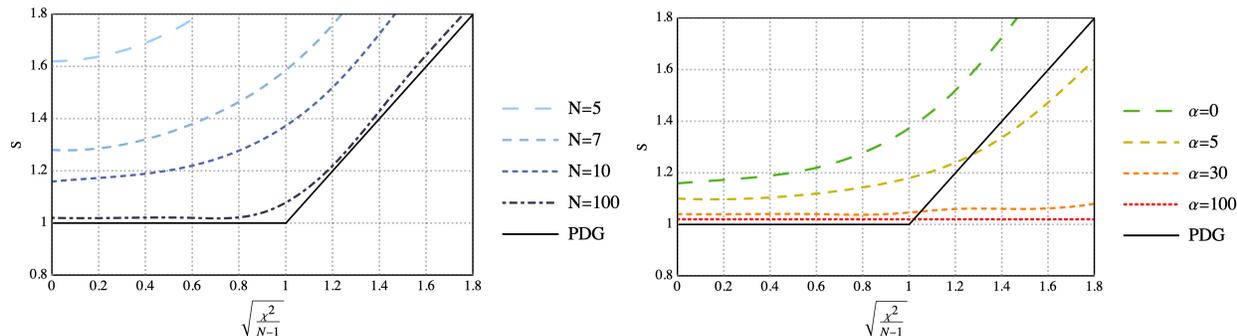


Figure 2.4: Plot of the scale factor as a function of the reduced χ^2 for different values of N and α .

this can easily be rewritten as

$$p(\mu, \tau^2 | \vec{y}) \propto \frac{1}{(\sigma^2 + \tau^2)^{\frac{N+\alpha}{2}}} \exp \left[-\frac{1}{2} \frac{\sigma^2}{\sigma^2 + \tau^2} \chi^2(\mu) \right], \quad (2.50)$$

where we have used Eq. 2.27. The form of this equation has strong implications. First of all, let us compute the peak of the posterior distribution of μ given the data, $p(\mu | \vec{y})$, this is obtained by solving

$$\frac{dp(\mu | \vec{y})}{d\mu} = \frac{dp(\mu | \vec{y})}{d\chi^2(\mu)} \frac{d\chi^2(\mu)}{d\mu} = 0. \quad (2.51)$$

Remember that the value of θ that minimizes the χ^2 is just the weighted average of the experimental results. Then this equation makes explicit the fact that the central value of the posterior distribution in the Hierarchical model is the same as in the non-hierarchical model when the errors of all experiments are of the same size. This result obviously breaks down for the case of different errors, because then it is not possible to rewrite the distribution as a function of the χ^2 only. Thus we get a similar result than what the PDG assumes: the central value after scaling, does not get modified when the experimental errors are of similar size. In Fig. 2.4 I show the Bayesian scale factor as a function of χ_{min}^2 for different values of N and α . The PDG scale factor is also shown there as a solid black line. It is under one's nose that the Bayesian scaling approximates to the PDG prescription as the number of degrees of freedom increases. Let us now prove this observation.

Asymptotic limit

Let us make the following change of variables

$$r = \frac{\sigma^2}{\tau^2 + \sigma^2}, \quad (2.52)$$

in the posterior distribution of μ given the data. The result is

$$p(\mu|\vec{y}) \propto \int_0^1 \exp \left[-\frac{\nu-2}{2} (r\chi_{\nu-1}^2(\mu) - \ln r) \right] dr, \quad (2.53)$$

where $\nu \equiv N + \alpha - 2$. Now, in the limit of large ν this integral can be approximated using Laplace method [49]. Since ν is very large, the dominant contribution to the integral will come from the minimum of the exponential. Thus is $r_0 = [\chi_{\nu-1}^2(\mu)]^{-1}$. So the position of the minimum depends on the value of $\chi_{\nu-1}^2(\mu)$. There are two cases, $\chi_{\nu-1}^2(\mu) > 1$ and $\chi_{\nu-1}^2(\mu) < 1$.

Case 1: $\chi_{\nu-1}^2(\mu) > 1$ In this case then the minimum is inside the integration limits. So we can approximate

$$r\chi_{\nu-1}^2(\mu) - \ln r \approx 1 + \ln [\chi_{\nu-1}^2(\mu)] + \frac{1}{2} [\chi_{\nu-1}^2(\mu)]^2 (r - r_0)^2 + \mathcal{O} [(r - r_0)^3], \quad (2.54)$$

this implies

$$p(\mu|\vec{y}) \propto e^{-\frac{\nu-2}{2} \ln[\chi_{\nu-1}^2(\mu)]} \int_{-\infty}^{\infty} \exp \left[-\frac{\nu-2}{4} [\chi_{\nu-1}^2(\mu)]^2 (r - r_0)^2 \right] dr, \quad (2.55)$$

this is just a Gaussian integral which gives something proportional to $[\chi_{\nu-1}^2(\mu)]^{-1}$, then

$$p(\mu|\vec{y}) \propto [\chi_{\nu-1}^2(\mu)]^{-\frac{\nu}{2}}, \quad (2.56)$$

expanding the χ^2 around its minimum (Eq. 2.27) allows to write this as

$$p(\mu|\vec{y}) \propto \left[\chi_{\nu-1}^2(\hat{\mu}) + \frac{1}{\nu-2} \frac{(\mu - \hat{\mu})^2}{\hat{\sigma}^2} \right]^{-\frac{\nu}{2}}, \quad (2.57)$$

this is

$$p(\mu|\vec{y}) \propto \left[1 + \frac{2}{\nu-2} \frac{(\mu - \hat{\mu})^2}{2\chi_{\nu-1}^2(\hat{\mu})\hat{\sigma}^2} \right]^{-\frac{\nu}{2}}, \quad (2.58)$$

which is proportional to a Student-t distribution for $\nu - 1$ degrees of freedom. We can now use the well known relation

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n \quad (2.59)$$

to get

$$p(\mu|\vec{y}) \propto \exp \left[-\frac{(\mu - \hat{\mu})^2}{2\chi_{\nu-1}^2(\hat{\mu})\hat{\sigma}^2} \right], \quad (2.60)$$

and the final error becomes $\hat{\sigma}_{bay} = \sqrt{\chi_{\nu-1}^2(\hat{\mu})}\hat{\sigma}$ this is just the prescription that the PDG suggest to scale the errors.

Case 2: $\chi_{\nu-1}^2(\mu) < 1$ In this case the minimum is outside of the integration limits. Then we must approximate around $r = 1$. This implies

$$p(\mu|\vec{y}) \propto \exp \left[-\frac{\nu-2}{2} (\chi_{\nu-1}^2(\mu)) \right], \quad (2.61)$$

if we replace $\chi_{\nu-1}^2(\mu)$ by its expansion around the minimum we get

$$p(\mu|\vec{y}) \propto \exp \left[-\frac{(\mu - \hat{\mu})^2}{2\hat{\sigma}^2} \right], \quad (2.62)$$

then there is no scaling, in accordance with the PDG prescription.

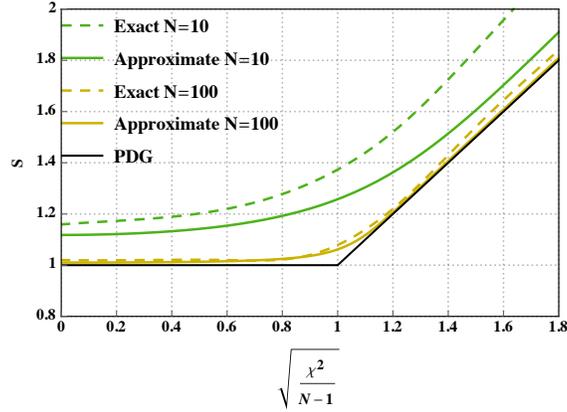


Figure 2.5: Comparison of the exact result with the approximate formula for $\alpha = 0$.

2.7.3 Gaussian Approximation

As we have shown before, the second derivative of $\ln p$ can be compared to the corresponding term of the expansion of a Gaussian distribution, which gives

$$\frac{1}{\sigma_{\text{Bayes}}^2} \approx - \left. \frac{d^2 \ln p}{d\mu^2} \right|_{\mu_0} = - \frac{2N}{\sigma^2} \left. \frac{d \ln p}{d\chi^2} \right|_{\chi_0^2}. \quad (2.63)$$

Using Eq. (2.53) we have,

$$-2 \left. \frac{d \ln p}{d\chi^2} \right|_{\chi_0^2} = \frac{\nu}{\chi^2} - \frac{(\chi^2/2)^{(\frac{\nu}{2}-1)} e^{-\chi^2/2}}{\gamma(\nu/2, \chi^2/2)}, \quad (2.64)$$

where γ is the incomplete Gamma function, defined by

$$\gamma(s, x) \equiv \int_0^x t^{s-1} e^{-t} dt. \quad (2.65)$$

As we mentioned before, the scale factor S_{Bayes} is defined as the ratio of the sizes of the 68% highest confidence intervals of the hierarchical and non-hierarchical models. In the Gaussian

approximation we find,

$$S_{\text{Bayes}} \approx \sqrt{N} \frac{\sigma_{\text{Bayes}}}{\sigma} \approx \sqrt{\frac{\chi^2}{\nu}} \left[1 + \frac{1}{\sum_{k=1}^{\infty} \frac{(\chi^2)^k \nu!!}{(\nu+2k)!!}} \right]^{\frac{1}{2}} \quad (2.66)$$

where we have used the power series expansion of the incomplete Gamma function,

$$\gamma(s, x) = x^s \Gamma(s) e^{-x} \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(s+k+1)}. \quad (2.67)$$

In Fig. 2.5 we compare the approximate formula with the exact result. As expected, the approximation improves for larger values of ν .

2.7.4 Errors of different size.

We are now ready to discuss the general case of unequal errors, $\sigma_i \neq \sigma_j$.

To understand this case, we fix the value of τ in Eq. (2.46). The distribution of μ is then Gaussian, with total error,

$$\frac{1}{\sigma_t^2} = \sum_{i=1}^N \frac{1}{\sigma_i^2 + \tau^2}, \quad (2.68)$$

and central value,

$$\mu_0 = \left(\sum_{i=1}^N \frac{1}{\sigma_i^2 + \tau^2} \right)^{-1} \sum_{i=1}^N \frac{y_i}{\sigma_i^2 + \tau^2}. \quad (2.69)$$

Thus, experiments with smaller errors are more sensitive to τ than less precise ones. Suppose that M of the experiments have an error σ_M , and that σ_M is much smaller than the error σ of the rest of the experiments. Then, for $\sigma_M \simeq \tau \ll \sigma$ the scaling will mainly affect the experiments with small errors. This resembles the PDG prescription reviewed before, but we avoid reference to a hard and *ad hoc* cutoff. Since we were unable to find an analytical formula for the peak or mean of τ , we proceed with a numerical analysis.

As a first example, we randomly generated eleven fictitious measurement points from a Gaussian with standard deviation $\sigma = 1$ centered at the value of 10. The last point is from

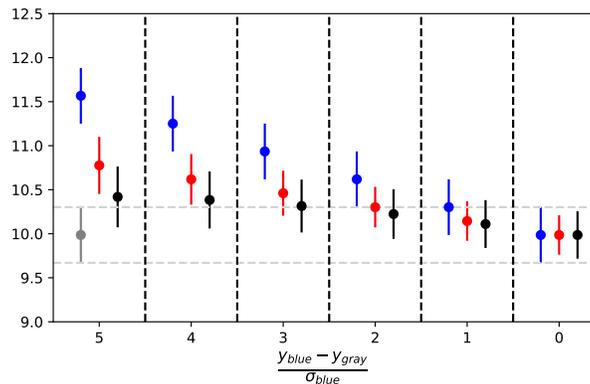


Figure 2.6: The measurement points with small error are shown in blue, the usual averages with the PDG scaling in red, and the hierarchical averages in black. The labels at the horizontal axis show by how many σ_M the blue points deviate from the gray point. The gray band represents the ordinary weighted averages of the bulk of measurements.

a Gaussian centered at $10 + 5/\sqrt{10}$ with $\sigma_M = 1/\sqrt{10}$, which is chosen so that its precision is the same as the combined precision of the other ten.

In Fig. 2.6 we show how the two kind of averages change when we move the central value of the 11th measurement (in blue) while leaving the other 10 unchanged. The red point denotes the ordinary weighted average with PDG scaling applied. The black point, on the other hand, is the average obtained as the result of our Bayesian hierarchical model ($\alpha = 10$). Just for orientation, the gray band represents the ordinary average (non-hierarchical) of the bulk of measurements with the same error. Clearly, as we approach the bulk the combined error shrinks.

2.8 Applied example: Neutron Lifetime

As I mentioned in the previous section, when the errors are not of the same size the full distribution needs to be used and integrated numerically. We applied this method to a case of particular interest in the physics community: the lifetime of the neutron. The process is

$$n \rightarrow p + e^- + \bar{\nu} \quad (2.70)$$

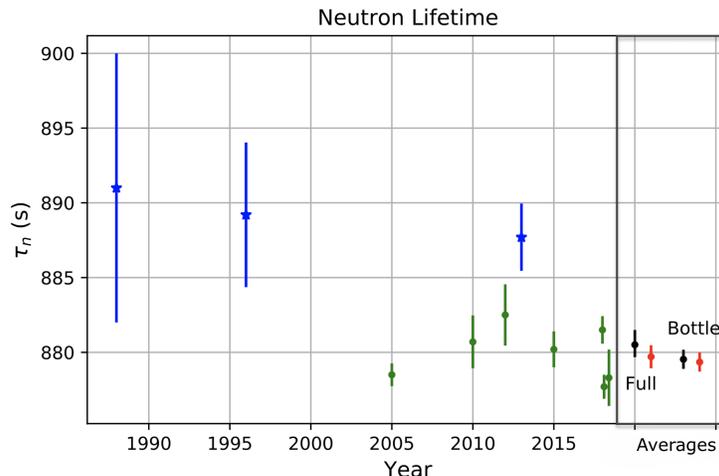


Figure 2.7: Neutron Lifetime. The blue points are beam experiments, while the green points are bottle experiments. The black points are the Bayesian averages, and the red points are the average with the PDG scaling. We did this averaging for the case when all experiments are included and also when only bottle experiments are taking into account.

with the amplitude given by

$$\mathcal{M} = [G_V \bar{p} \gamma_\mu n - G_A \bar{p} \gamma_5 \gamma_\mu n] [\bar{e} \gamma_\mu (1 + \gamma_5) \nu]. \quad (2.71)$$

Using first order perturbation theory and integrating over the electron and neutrino energies, the neutron lifetime can be calculated from this amplitude. The result is

$$\tau_n = \frac{2\pi^3 \hbar^7}{(G_V^2 + 3G_A^2) m_e^5 c^4 f_R}, \quad (2.72)$$

where m_e is the electron mass, and f_R is the result of the integral over the energy spectrum. The neutron lifetime plus the angular dependence of the decay rate can be used to estimate the two coupling constants G_V and G_A . Furthermore, the determination of G_V allows to calculate the first element V_{ud} of the Cabibbo-Kobayashi-Maskawa (CKM) matrix [50, 51] through the relation⁹ $G_V = V_{ud} G_F$, where G_F is the Fermi constant which is measured in

⁹This relation assumes a conserved vector current hypothesis [52, 53], see [54] for a quick and pedagogical explanation.

muon decays. The CKM by construction is unitary, and a violation of this property would imply the presence of new physics, so this is an important test for the Standard Model. I would like to emphasize that the most precise measurements of G_V are super-allowed beta decays $0^+ \rightarrow 0^+$ [55–57]. Nevertheless, due to the nature of these processes, only G_V can be constrained. The neutron lifetime is then relevant for two reasons: to complement the measurements of G_V and to get G_V/G_A which can also put constraint on new physics [58, 59].

There are two types of experiments that measure the lifetime of the neutron: beam and bottle experiments. A state of art review of them can be found in Ref. [60].

In the so called beam experiments [61–63], a neutron beam is targeted to the decay volume. The decay volume is a physical space where the neutrons decay. The decay products (like p^+) produced in this region are detected through different methods. After leaving this region, the neutron beam passes through another detector, which measures the neutron rate. The neutron lifetime will be then proportional to the neutron rate and inversely proportional to the rate of products detected.

On the other hand, bottle experiments [64–69], consist on ultra cold neutrons (UCN) confined inside a bottle. First neutrons at a temperature lower than 1mK are placed in a container. After waiting some specific time, the number of neutrons inside the bottle are counted. From here, the neutron lifetime can be calculated. The neutrons are confined through several forces, like gravitational, magnetic and Fermi effective potential from coherent scattering of the neutron from atomic nuclei.

We now apply our method with $\alpha = 6$ to the results of these experiments. The averaging and experiments are shown in Fig. 2.7. PDG χ^2 scaling (shown in red) yields the lifetime $\tau_n = 879.71 \pm 0.78$ s, while the Bayesian method (black point to the left) gives $\tau_n^{\text{Bayes}} = 880.51_{-0.83}^{+0.98}$ s. We find that our Bayesian hierarchical method increases the central value when the beam experiments are included. We also observe that even when only bottle experiments are considered, our method still gives a slightly larger average value ($\tau_n^{\text{Bayes}} = 879.53_{-0.63}^{+0.64}$ s)

than the PDG method ($\tau_n = 879.35 \pm 0.64$ s). This is due to the bulk of the bottle experiments that prefer lifetimes longer than 880 s. It is important to recall that the tails of the Bayesian hierarchical model do not fall as fast as a Gaussian, so that there is still a non-negligible probability for τ_n to be lower.

2.9 Comparison with other models

There are a couple of papers in the same direction than our work. The first one [70] discusses the kaon mass in the context of a skeptical combination of experiments, scaling each experimental error independently but correlated. The second one [30] studies the discrepancy that arises when the PDG scaling is applied to sub-sets of experiments and then to the combination of the sets, *vs.* (for example) applying it to the whole data at the same time. The conclusion is that *the χ^2/ν prescription used to enlarge the standard deviation does not hold sufficiency*. This means that the scaling is not enough to properly describe the full probability distribution. Our model would have the same problem had we used the marginalized (over τ^2) distribution of μ . This is because the “correlations” that emerge through τ^2 would be absent. Nevertheless, it is clear from Eq. (2.46) that if we use the posterior distribution of μ and τ^2 of a subset of experiments as the prior for the remaining subset, then the updated posterior will be the same as combining the whole data set simultaneously.

Another interesting point discussed in Ref. [30] is the fact that the PDG scaling treats any value of N equally, while for fixed χ^2/N the p -value decreases with N . Our model with α chosen close to zero would aggravate this problem, as there is more scaling for low N as can be seen in the left plot of Fig. 2.4. However, we can use the freedom to choose a value of α to improve on this issue, by demanding that the variance of the τ distribution be finite, which corresponds to $\alpha = 6$. Indeed, the right plot of Fig. 2.4 shows how for N fixed, our scaling will intersect the PDG scaling for moderate α .

We have seen that in the case of equal errors, our method is very similar to the PDG.

On the other hand, for different errors, we have seen that in the Hierarchical Model the central value of the combined data will in general move towards the experiments with lower precision, this is because the shift induced by τ^2 affects more the high precision experiments, so gives more relative weight to the experiments with low precision. This is the previously mentioned effect that the bulk weights more than a single experiment.

We have shown that our methodology resembles the recommendation of the Particle Data Group whenever the number of degrees of freedom (data points) is large. Our approach connects smoothly to cases with fewer degrees of freedom, though. Another important advantage is that it makes the underlying assumptions in the averaging process transparent. *E.g.*, a large value of the parameter α appearing in our proposed form of the prior, implies a strong believe that the experiments do not have an unknown systematic error, while a small value corresponds to a more agnostic point of view. Our method can be extended to experiments with correlated errors, but we leave this generalization for the future.

2.10 Other applications of the Hierarchical model

Here I briefly explain how the hierarchical model can be used in other contexts. For example, in [71] we used a Hierarchical approach to estimate the error on a bound to parton distribution functions. Let us ignore the physics for the moment, and suppose that some quantity must lie between a certain region. In other words, given a value of "x" in a (x,y) plane, the value of y must be between two functions, $f_1(x)$ and $f_2(x)$. Now, assume that we have experiments that measure y for several values of x. Since experiments have a finite precision, there is not full certainty if the true value is inside or outside the mentioned region. The idea is then to compute the "common" probability that the true value can be outside the region. This is where the Hierarchical Model enters. First let us introduce θ_i , which will be understood as the probability that true value of one experiment is inside the region. Suppose that the true value A_i is known, then it is obvious that $\theta_i = 1$ if A_i is inside the region, while $\theta_i = 0$ if it

is not. Mathematically this can be written as:

$$p(\theta_j|A_j) = \begin{cases} \delta(\theta_j - 1) & \text{for } A_j \in \text{region} \\ \delta(\theta_j) & \text{for } A_j \notin \text{region} \end{cases} \quad (2.73)$$

We can give a prior weight to each of these two cases, through a prior distribution of the form

$$p(\theta_j|t) = (1 - t)\delta(\theta_j - 1) + t\delta(\theta_j) \quad , \quad (2.74)$$

The idea is then to use the data to obtain the values of the hyperparameter t , which measures the weight between both cases. This is just a Hierarchical Model. The next step is to write $p(\theta_i|\text{data})$. This can easily be obtained through the well known probability rule

$$p(\theta_j|\text{data}) = \int p(\theta_j|A_j)p(A_j|\text{data})dA_j \quad , \quad (2.75)$$

here, $p(A_j|\text{data})$ is just a Gaussian distribution given by the experimental collaborations for each experiment. Then through Bayes theorem we have

$$p(t|\text{data}) \propto p(t) \sum_j \int p(\theta_j|t)p(\theta_j|\text{data})d\theta_j. \quad (2.76)$$

This probability distribution can be understood as the distribution of the weight for the two options: $\theta_j = 0$ and $\theta_j = 1$. As τ becomes smaller, there is a larger probability that the true value is inside the region. In the paper we get

$$\bar{t} = \int_0^1 t F(t) dt = 0.049 \pm 0.040 \quad , \quad (2.77)$$

which is actually compatible with zero¹⁰.

If the reader is interested and wants more detail on how this result was used I recommend

¹⁰The error here is defined as the symmetric error around the mean that contains the 68% probability mass.

to check [71]. But now let us jump to a different topic: The Standard Model.

Chapter Three

Standard Model

In this chapter I will introduce the Standard Model (SM), which unifies the weak, strong and electric forces into a single model. This introduction will be useful for the next two chapters, where some of the concepts studied here will be used in detail. A key ingredient in the Standard Model is the concept of gauge symmetries. Gauge symmetries allow us to understand forces as the interchange of a particle called the "mediator". Examples of mediators are the photon, the Z boson, the W boson, the gluon and hypothetically the graviton. The concept of non Abelian gauge symmetries was introduced in 1954 by C. N. Yang and R. L. Mills [72] in the context of isotopic spin gauge invariance. The final part of the paper had a discussion about the mass of the mediator: the conclusion was that they were not able to explain its origin or size. A few years later in 1961, Sheldon Glashow combined the weak and electromagnetic interactions under the same gauge theory [73]. But the mass of the gauge bosons was still a problem of the theory¹. The main problem is that an explicit mass term for the gauge boson would imply breaking of the gauge symmetry. This was solved later in 1964 via the Higgs mechanism [74, 75], where the mass of the mediator is acquired when a scalar field (the Higgs boson) acquires a non zero vacuum expectation value (vev). This mechanism was incorporated later in the Electroweak Standard Model introduced by Weinberg and Salam [14, 15] in 1967 .

¹He introduced the term partial symmetry to avoid this issue.

3.1 Definitions

The full Lagrangian of the Standard Model is:

$$\mathcal{L}_{SM} = \mathcal{L}_{Gauge} + \mathcal{L}_{Fermion} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} + \mathcal{L}_{GF} + \mathcal{L}_{ghost}. \quad (3.1)$$

The first term (\mathcal{L}_{Gauge}) contains the kinetic terms and self interactions of the Gauge Bosons namely,

$$\mathcal{L}_{Gauge} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}, \quad (3.2)$$

here $G_{\mu\nu}^a$ is the gluon field strength tensor given by

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A^b A^c \quad (3.3)$$

and f^{abc} are the structure constants of $SU(3)$ which are defined through the Lie algebra:

$$[T^a, T^b] = i f^{abc} T^c, \quad (3.4)$$

where T^a are the $SU(3)$ generators of the Lie algebra and g_s is the QCD coupling constant. $W_{\mu\nu}^a$ is defined in an analogous way to $G_{\mu\nu}^a$, except that we change the symmetry group to $SU(2)$ and also $g_s \rightarrow g$. This implies that the structure constants are $f^{abc} = \epsilon^{abc}$ where ϵ is the completely antisymmetric Levi-Civita tensor. After spontaneous symmetry breaking, we will see that the bosons W^1 and W^2 are linear combinations of the physical charged W^\pm bosons, and W^3 is a linear combination of the photon and the Z boson. Finally, $B_{\mu\nu}$ is also defined in the same way as $G_{\mu\nu}^a$ but using $U(1)$ as the symmetry group. This implies that the structure constants are equal to zero, so the non-Abelian terms vanish and self interactions for the gauge bosons disappear as in QED. The B boson can be split as a linear combination of the photon and the Z boson. This linear combination is orthogonal to the one of the W^3 . The second term on the right hand side, $\mathcal{L}_{Fermion}$, contains the kinetic terms

of the fermions. But when the partial derivatives are replaced by the covariant derivatives $\partial_\mu \rightarrow \partial_\mu - igW_\mu^a T^a - ig'Y B_\mu$, this term will also contain the interactions of the gauge bosons with the the fermion currents.

The third term, \mathcal{L}_{Higgs} , contains the kinetic term and potential of the Higgs field. In a similar way to the kinetic term of the fermions, the partial derivative is replaced by the covariant derivative, giving rise to the interactions between the Higgs boson (and Goldstone bosons) and the gauge bosons. After the Higgs acquires a vacuum expectation value, the gauge fields acquire masses through the Higgs mechanism.

The fourth term, \mathcal{L}_{Yukawa} , contains the interactions of the Higgs field with the fermions. After spontaneous symmetry breaking the fermions acquire a mass. The wide difference in the mass of the neutrinos is an open question. This is because if they all acquire their mass from the Higgs vev, then the difference between the yukawa couplings among the fermions must be huge to explain the measured masses for the neutrinos. A possible solution to this issue is to include a dimension 5 operator, the so called Weinberg operator. This will naturally suppress the mass of the neutrinos with the new physics scale. There are specific UV completions that are very appealing, like the see-saw mechanism, where right-handed neutrinos are introduced, with large Majorana masses. The mass of this right-handed neutrinos will then suppress the mass of the light left-handed neutrinos. We will talk more about this in chapter 5.

The fifth and sixth terms are, as far as I understand, a consequence of our lack of a better mathematical tool (Quantum Field Theory) to describe relativistic and quantum phenomena together. \mathcal{L}_{GF} is a term that fixes the gauge in which we are working. The gauge freedom emerges as a consequence of embedding our mass-less particle with two degrees of freedom in a polarization vector ϵ_μ with four d.o.f. The result is that we have to fix the gauge in our Lagrangian to eliminate the redundancy originated from this embedding. In the case of non-Abelian gauge theories, this gauge fixing induces a change in the measure of the path integral. After some algebraic tricks, this change of measure can be rewritten as an

extra term in the Lagrangian: \mathcal{L}_{ghost} . This term contains non physical fields called ghosts, which have one degree of freedom and obey fermion statistics. These ghosts are key in the cancellation of the gauge dependence in physical quantities. We will not discuss in detail how these cancellations occur but if the reader is interested I recommend to revise [76–78] for further information.

3.2 Higgs Mechanism

The Higgs Mechanism is closely connected to spontaneous symmetry breaking. Spontaneous symmetry breaking states the following: "If a continuous global symmetry of the Lagrangian is broken by the ground state of a scalar field, $Q_i |\Omega\rangle \neq 0$, then there are as many massless Goldstone bosons as generators were broken". When this continuous global symmetry is upgraded to a gauge symmetry, the Higgs Mechanism says: "The degrees of freedom of the Goldstone Bosons are absorbed by the Gauge bosons, giving them a mass". To see this explicitly, the unitary gauge [79] is frequently used, in this gauge the Goldstone boson fields are completely eliminated from the Lagrangian, and only physical degrees of freedom propagate. Let us now study the Higgs mechanism in the Standard Model. The first thing that is required is a scalar field that will acquire a vev. Thus, let us introduce scalar field (the Higgs) H that transforms as a doublet under $SU(2)$, and with hypercharge $Y = \frac{1}{2}$ under the $U(1)$ symmetry of the B bosons. Just as usual, to get the interaction of this field with the gauge bosons, the partial derivative is promoted to a covariant derivative. The covariant derivative is then

$$D_\mu H = \partial_\mu H - igW_\mu^a \tau^a H - \frac{1}{2}ig' B_\mu H, \quad (3.5)$$

and the Lagrangian is

$$\mathcal{L}_{Higgs} = (D_\mu H)^\dagger (D_\mu H) + \mu^2 H^\dagger H - \lambda (H^\dagger H)^2. \quad (3.6)$$

Now, we expand the Higgs field around the minimum. To do so we use the Kibble parametrization,

$$H(x) = e^{-2\frac{i\tau^a\pi^a(x)}{v}} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} + \frac{h(x)}{\sqrt{2}} \end{pmatrix}, \quad (3.7)$$

where the unitary gauge is defined as the gauge such that the π^a fields are not present in the Lagrangian. Given the form of the breaking, the unbroken generator will be

$$\hat{Q} = \hat{T}^3 + \hat{Y} \quad (3.8)$$

which we should identify with the electric charge, since we know that this is the remaining symmetry after the breaking. The relation between these three generators is known as the Gell-Mann–Nishijima formula [80]. It is obvious then, that the field h must be neutral, since $T^3 = -\frac{1}{2}$ and $Y = \frac{1}{2}$ for this field. Inserting this expression in the kinetic term of the Higgs Lagrangian, and leaving only the terms that are proportional to v^2 we get,

$$\mathcal{L}_{Higgs}^{mass} = \frac{1}{8} \left[g^2 W^{\mu a} W_\mu^a v^2 + 2gg' B^\mu W_\mu^b \begin{pmatrix} 0 & v \end{pmatrix} \tau^b \begin{pmatrix} 0 \\ v \end{pmatrix} + g'^2 B_\mu B^\mu v^2 \right]. \quad (3.9)$$

If we define

$$W_\mu^+ = \frac{1}{\sqrt{2}} [W_\mu^1 - iW_\mu^2] \quad W_\mu^- = \frac{1}{\sqrt{2}} [W_\mu^1 + iW_\mu^2], \quad (3.10)$$

then

$$\mathcal{L}_{Higgs}^{mass} = \frac{1}{2} \left[g^2 v^2 W^{\mu+} W_\mu^- + v^2 (g^2 + g'^2) \left(\frac{g}{\sqrt{g^2 + g'^2}} W^{\mu 3} - \frac{g'}{\sqrt{g^2 + g'^2}} B^\mu \right)^2 \right]. \quad (3.11)$$

We can now make another definition,

$$Z^\mu = \frac{g}{\sqrt{g^2 + g'^2}} W^{\mu 3} - \frac{g'}{\sqrt{g^2 + g'^2}} B^\mu, \quad (3.12)$$

	e_L^i	ν_L^i	e_R^i	ν_R^i	$q_{L,u}^i$	$q_{L,d}^i$	$q_{R,u}^i$	$q_{R,d}^i$	h
T_3	$-\frac{1}{2}$	$\frac{1}{2}$	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	$-\frac{1}{2}$
Q	-1	0	-1	0	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	0
Y	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{2}$

Table 3.1: Standard Model charge assignments.

which leaves the orthogonal linear combination

$$A^\mu = \frac{g}{\sqrt{g^2 + g'^2}} W^{\mu 3} + \frac{g'}{\sqrt{g^2 + g'^2}} B^\mu, \quad (3.13)$$

as a mass less particle which we identify as the photon. We define

$$\sin \theta_W \equiv \frac{g'}{\sqrt{g^2 + g'^2}}, \quad (3.14)$$

as the weak mixing angle or Weinberg angle. The study of the running of this quantity with energy is one of the main topics of this thesis and will be described in the next chapter. At tree level the weak mixing angle is related in a simple way to other parameters of the Standard Model. For example, we can identify

$$M_W^2 = \frac{g^2 v^2}{2}, \quad M_Z^2 = v^2 \frac{g^2 + g'^2}{2}, \quad (3.15)$$

so

$$\sin^2 \theta_W = 1 - \frac{M_W^2}{M_Z^2}, \quad (3.16)$$

We are now ready to review the interactions of the gauge bosons with fermions.

3.3 Interaction with Fermions

The hypercharge of the fermions can easily be obtained from the Gell-Mann–Nishijima formula. First we embed the left-handed charged leptons and left-handed neutrinos in a doublet

with quantum numbers $T_3^\nu = \frac{1}{2}$ and $T_3^l = -\frac{1}{2}$ under the $SU(2)$ mentioned previously. Then, the right-handed leptons are defined to transform as singlets under this symmetry. We also arrange the left-handed quarks as doublets, and the right-handed quarks as singlets. Because this symmetry only interacts with left-handed fields, we call this an $SU(2)_L$ symmetry. To compute the hypercharge we simply use

$$Y = (Q - T_3), \quad (3.17)$$

for example, for a left-handed charged fermion we have $Y = (-1 + \frac{1}{2}) = -\frac{1}{2}$. In the same way, for a right-handed charged lepton we would have $Y = -1$. Table 3.3 shows the SM particles with their $SU(2)_L \otimes U(1)_Y$ quantum numbers. With the hypercharge known, we can write the kinetic Lagrangian for the fermions,

$$\mathcal{L}_{fermion} = i\bar{\Psi}_L^i \gamma^\mu D_\mu \Psi_L^i + i\bar{e}_R^i \gamma^\mu D_\mu e_R^i + i\bar{Q}_L^i \gamma^\mu D_\mu Q_L^i + i\bar{q}_{u,R}^i \gamma^\mu D_\mu q_{u,R}^i + i\bar{q}_{d,R}^i \gamma^\mu D_\mu q_{d,R}^i \quad (3.18)$$

where i runs over all the families. Here Ψ_L and Q_L are the left-handed lepton and quark doublets respectively. e_R is a right-handed charged lepton and $q_{u,R}$ and $q_{d,R}$ are right-handed quarks. For a general field ψ the covariant derivative is

$$D_\mu \psi = \partial_\mu \psi - igW_\mu^a \tau^a \psi - ig'Y B_\mu \psi, \quad (3.19)$$

where the term proportional to the τ^a is zero for all the right-handed fields. To study the running of the weak mixing angle, we will need the coupling of the fermions to the Z and the photon. Thus we rewrite the covariant derivative in terms of those fields. After doing so, it is straightforward to find that the interaction between the Z and the photon with the

fermions is given by (ignoring the strong force),

$$\mathcal{L}_{int}^{fermions} = i \sum_j \bar{\psi}_j \gamma^\mu \left[A_\mu \frac{gg'}{\sqrt{g^2 + g'^2}} Q_j + Z_\mu \frac{gg'}{\sin \theta_W \cos \theta_W \sqrt{g^2 + g'^2}} (t_j^3 - \sin^2 \theta_W Q_j) \right] \psi_j, \quad (3.20)$$

where we are summing over all chiralities, families and fermion type. Due to the form of this equation, we can easily see that we must have

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}}, \quad (3.21)$$

and,

$$\mathcal{L}_{int}^{fermions} = i \sum_j \bar{\psi}_j \gamma^\mu \left[A_\mu e Q_j + Z_\mu \frac{e}{\sin \theta_W \cos \theta_W} (t_j^3 - \sin^2 \theta_W Q_j) \right] \psi_j. \quad (3.22)$$

Since the charge Q_j is the same for left and right-handed fields, the interaction with the photon is of vector type. On the other hand, for the interactions with the Z bosons, only the left ones will contribute to t_3 . Thus if we want to consider only the vector interaction, we should divide by two this contribution due to the form of the projector P_L . Thus the vector interaction between the fermions and the Z and photon is given by

$$\mathcal{L}_{int}^{fermions} = i \sum_i \bar{\psi}_i \gamma^\mu \left[A_\mu e Q_i + Z_\mu \frac{e}{2 \sin \theta_W \cos \theta_W} (t_i^3 - 2 \sin^2 \theta_W Q_i) \right] \psi_i, \quad (3.23)$$

where now the sum is over each Dirac field, taking t_3 to be the eigenvalue of the corresponding left-handed field. We define the vector coupling as $g_V^i = t_i^3 - 2Q_i \sin^2 \theta_W$.

3.4 Renormalization

One of the most interesting and intriguing topics in Quantum Field theory is renormalization. The idea behind it is that the "bare" parameters such as masses or couplings that appear in the Lagrangian are not necessarily the same as the ones measured in experiments. To refresh our minds about this concepts, let us start with the QED Lagrangian.

The QED Lagrangian (ψ is the field of the charged fermion) is

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}^{(0)}F^{(0)\mu\nu} + i\bar{\psi}^{(0)}\gamma^\mu \left(\partial_\mu + ie^0 A_\mu^0 \right) \psi^{(0)} + m^{(0)}\bar{\psi}^{(0)}\psi^{(0)}. \quad (3.24)$$

But interesting things happen when interactions are included. For example, the residue and pole of the propagators change. In a similar way, the effective values of the couplings also change. To take into account these effects, it is a useful and common practice to re-scale the fields and parameters. In this way we can actually force these parameters to be equal to the physical values (like masses in the on shell-scheme). First, let us assume that we are in a space of $4 - d$ dimensions. Now, let us rewrite the parameters and fields as

$$\psi^{(0)} = Z_2^{\frac{1}{2}}\psi \quad e^{(0)} = Z_e e \mu^{\frac{4-d}{2}} \quad A_\mu^{(0)} = Z_3^{\frac{1}{2}}A_\mu \quad m^{(0)} = Z_m m, \quad (3.25)$$

where the renormalization scale μ appears because we would like e to be dimensionless, and $e^{(0)}$ is only dimensionless in 4 dimensions. Thus we have,

$$\mathcal{L}_{QED} = -\frac{1}{4}Z_3 F_{\mu\nu} F^{\mu\nu} + iZ_2 \bar{\psi} \gamma^\mu \left(\partial_\mu + iZ_e Z_3^{\frac{1}{2}} e A_\mu \right) \psi + Z_m Z_2 m \bar{\psi} \psi. \quad (3.26)$$

After particle loops are included in physical processes, infinities will appear. An intermediate step in the renormalization procedure is regularization, which is the method that we use to handle these infinities. The regularization of our integrals (the way we extract the infinities in them) will be done in dimensional regularization, were we set the dimension



Figure 3.1: Diagrams that contribute the the photon two point Green function.

equal to $d = 4 - \epsilon$. The infinities in the loops will be absorbed in the parameter ϵ in such way that when $\epsilon \rightarrow 0$ the relations between bare and physical quantities become divergent, or in other words, the relation between the bare parameters and the renormalized ones will contain a power expansion of the form ϵ^{-n} where n is the number of loops that are to be computed. This implies that if we rewrite $Z_i = 1 + \delta Z_i$, then at one loop, we should have something of the form $\delta Z_i = \alpha \equiv \frac{e^2}{4\pi} (\frac{1}{c_1 \epsilon} + c_2)$, where c_1 is a constant to be computed, and c_2 depends on the renormalization scheme. The δZ_i are usually called counter terms.

3.4.1 Vacuum Polarization Function

Let us compute the quantum corrections to the two-point Green function for the photon in momentum space. This will lead us to obtain Z_3 . I will only sketch the steps, since this is done in any standard book of Quantum Field Theory [76, 77, 81]. Here I will follow the notation of [76]. At tree level in an class of gauges we have for the photon propagator

$$iD_{\mu\nu}^{tree}(p) = \frac{-i}{p^2 + i\epsilon} \left[g_{\mu\nu} - (1 - \xi) \frac{p_\mu p_\nu}{p^2} \right]. \quad (3.27)$$

To get this propagator, an extra term must be included in the Lagrangian, (which is the result of performing a path integral over the gauge configurations), which is

$$\mathcal{L}_{fix} = -\frac{1}{2\xi_{(0)}} (\partial^\mu A_\mu^{(0)})^2. \quad (3.28)$$

Note that we have also used bare quantities here. We rewrite $\xi_0 = Z_\xi \xi$ to have everything written in terms of the renormalized quantities. At one loop, the diagrams that contribute to the running are given by Fig. 3.1. The first element is just the bare propagator, the second one contains a fermion loop, and the third diagram corresponds to the counter terms. The counter terms have two contributions, one that comes from scaling the kinetic term of the photon field in the Lagrangian, and another one that comes from the gauge fixing term in the Lagrangian,

$$iD_{\mu\nu} = iD_{\mu\nu}^{tree} + iD_{\mu\alpha}^{tree} \left(\Pi^{\alpha\beta}(p) - i\delta_3(p^2 g^{\alpha\beta} - p^\alpha p^\beta) + i \frac{\delta_\xi - \delta_3}{\xi} p^\alpha p^\beta \right) iD_{\beta\nu}^{tree} + \dots \quad (3.29)$$

where $\Pi^{\alpha\beta}(p)$ comes from the fermion loop. Due to Ward Identity [82], this fermion loop must be proportional to $p^2 P^{\mu\nu} \equiv p^2 g^{\mu\nu} - p^\mu p^\nu$. Then we can write $\Pi^{\mu\nu}(p) = -ie^2 (p^2 g^{\mu\nu} - p^\mu p^\nu) \Pi(p^2)$ so

$$iD_{\mu\nu} = iD_{\mu\nu}^{tree} + iD_{\mu\alpha}^{tree} \left[-i(p^2 g^{\alpha\beta} - p^\alpha p^\beta) (e^2 \Pi(p) + \delta_3) + i \frac{\delta_\xi - \delta_3}{\xi} p^\alpha p^\beta \right] iD_{\beta\nu}^{tree} + \dots \quad (3.30)$$

the propagator contains two terms: a transverse part and a longitudinal one. This is clearly stated if we write

$$iD_{\mu\nu}^{tree}(p) = \frac{-i}{p^2 + i\epsilon} \left[P_{\mu\nu} + \xi \frac{p_\mu p_\nu}{p^2} \right]. \quad (3.31)$$

The projector $P_{\mu\nu}$ gives $P_{\mu\nu} p^\mu = 0$ and $P_{\mu\nu} P^{\nu\alpha} = P_\mu^\alpha$ so the power expansion of the propagator can be written as

$$iD_{\mu\nu}(p) = \frac{-i}{p^2 + i\epsilon} \left[P_{\mu\nu} (1 - e^2 \Pi(p) - \delta_3) + \xi \frac{p_\mu p_\nu}{p^2} \right] + iD_{\mu\alpha}^{tree} \left[i \frac{\delta_\xi - \delta_3}{\xi} p^\alpha p^\beta \right] iD_{\beta\nu}^{tree} + \dots \quad (3.32)$$

Since the computation of the vacuum polarization function is done in any standard book of quantum field theory, here I will just quote the one loop result, which is

$$\Pi(p^2, \mu^2) = \frac{1}{2\pi^2} \int_0^1 dx x(1-x) \left[\frac{2}{\epsilon} + \ln \left(\frac{\mu^2}{m^2 - p^2 x(1-x)} \right) \right], \quad (3.33)$$

It is clear that δ_3 is related to the transverse polarization and hence to the physical degrees of freedom. On the other hand δ_ϵ is related to the longitudinal part which is not physical. Let us focus then on the transverse part. If we force the residue of this part to be equal to one, then we must have $\delta_3 = -e^2 \Pi(0)$, where we may define e as the value of the charge at zero momentum. This is the on-shell renormalization scheme. On the other hand, the $\overline{\text{MS}}$ scheme is defined as that one where δ_3 absorbs only the divergent terms² contained in $\Pi(p)$.

Then we must have

$$\delta_3 = -\frac{e(\mu)^2}{2\pi^2} \int_0^1 dx x(1-x) \frac{2}{\epsilon} = -\frac{e^2(\mu)}{6\pi^2 \epsilon}. \quad (3.34)$$

This implies that the pole of the transverse part is still at zero, with residue $-i(1 - e^2(\mu)\bar{\Pi}(0, \mu))$ where $\bar{\Pi}$ is the fermion vacuum polarization function with the infinity part removed.

3.4.2 Renormalization group equation in QED

If we want the propagator (including charge at the external legs) to be a μ independent quantity, then the μ dependence that appears in $e(\mu)$ must necessarily cancel the μ dependence in $\bar{\Pi}$. Then we must have

$$\frac{d}{d \ln \mu^2} \left[e^2(\mu) \left(1 - e^2(\mu) \bar{\Pi}(0, \mu) \right) \right] = 0, \quad (3.35)$$

using $e^2 = 4\pi\alpha$ we get (to order α^2)

²To be more specific, it is the MS scheme the one that only absorbs the infinities. The $\overline{\text{MS}}$ scheme also absorbs some well defined finite terms.

$$\frac{d\alpha}{d \ln \mu^2} = 4\pi\alpha^2(\mu) \left[\frac{d\bar{\Pi}(0, \mu)}{d \ln \mu^2} \right] \quad (3.36)$$

so the one loop solution is given by

$$\frac{1}{\alpha(\mu)} - \frac{1}{\alpha(\mu_0)} = 4\pi \left[\bar{\Pi}(0, \mu) - \bar{\Pi}(0, \mu_0) \right] \quad (3.37)$$

or

$$\alpha(\mu) = \frac{\alpha(\mu_0)}{1 - \alpha(\mu_0) 4\pi \left[\bar{\Pi}(0, \mu) - \bar{\Pi}(0, \mu_0) \right]} \quad (3.38)$$

If we had explicitly computed the derivative of the vacuum polarization function we would have obtained,

$$\frac{d\alpha(\mu)}{d \ln \mu^2} = \frac{\alpha^2(\mu)}{3\pi}, \quad (3.39)$$

which is known as the RGE equation for the electric charge. In this derivation we have assumed that in a physical process, the μ dependence of the external legs in some scattering amplitude is cancelled, i.e., that the μ dependence of the fermion legs is cancelled by the μ dependence of the vertex. This cancellation turns out to happen in QED, and it is closely related to the Ward-Takashi identity [82, 83].

The RGE equation can easily be generalized to higher orders. Once more loops are included in the theory, the RGE will become a power series in $\hat{\alpha}$, so we can rewrite it as

$$\frac{d\hat{\alpha}(\mu)}{d \ln \mu^2} = \beta(\hat{\alpha}) = -\pi \left(\frac{\hat{\alpha}}{\pi} \right)^2 \sum_{n=0}^{L-1} \beta_n \left(\frac{\hat{\alpha}}{\pi} \right)^n, \quad (3.40)$$

Where L is the number of loops up to which the RGE has been computed. β is known as the beta function.

There is a more formal way to derive the RGE equation through the counter terms. The

idea is that the bare charge is related to the renormalized charge, by

$$e^{(0)} = Z_e e \mu^{\frac{4-d}{2}}. \quad (3.41)$$

On the other hand, the correction to the vertex interaction between the photon and a pair of fermions should be absorbed in $Z_1 \equiv Z_e Z_3^{\frac{1}{2}} Z_2$ as can be seen from Eq. (3.26). Z_1 is a counter term that absorbs the infinities that come from the vertex correction to the fermion photon interaction. The remaining counter terms are Z_2 , Z_3 and Z_e . We have just computed the value of Z_3 through the vacuum polarization function. Z_2 can be computed from the correction to the external leg of a fermion i.e., the emission and absorption of a photon on the same fermion leg. After computing Z_1 , Z_2 and Z_3 it is straightforward to solve for Z_e ,

$$Z_e = \frac{Z_1}{Z_3^{\frac{1}{2}} Z_2}. \quad (3.42)$$

Here is where the Ward identity enters. In its original form, the Ward identity tells us that $Z_1 = Z_2$. Then

$$e^{(0)} = Z_3^{-\frac{1}{2}} e(\mu) \mu^{\frac{\epsilon}{2}}, \quad (3.43)$$

taking the log

$$\ln e^{(0)} = -\frac{1}{2} \ln Z_3 + \ln e(\mu) + \frac{\epsilon}{2} \ln \mu, \quad (3.44)$$

multiplying by two

$$2 \ln e^{(0)} = -\ln Z_3 + 2 \ln e(\mu) + \frac{\epsilon}{2} 2 \ln \mu, \quad (3.45)$$

or

$$\ln e^{(0)2} = -\ln Z_3 + \ln e(\mu)^2 + \frac{\epsilon}{2} \ln \mu^2, \quad (3.46)$$

since the left hand side is independent of μ we must have

$$\frac{d}{d \ln \mu^2} \left[-\ln Z_3 + \ln e(\mu)^2 + \frac{\epsilon}{2} \ln \mu^2 \right] = 0. \quad (3.47)$$

On the other hand, since $Z_3 = 1 - \frac{2\alpha}{3\pi\epsilon}$ then

$$\ln Z_3 = -\frac{2\alpha}{3\pi\epsilon} \rightarrow \frac{d}{d \ln \mu^2} \ln Z_3 = \frac{d \ln Z_3}{d\alpha} \frac{d\alpha}{d \ln \mu^2} = -\frac{2}{3\pi\epsilon} \frac{d\alpha}{d \ln \mu^2}. \quad (3.48)$$

Plugging in this result back into Eq. (3.47) we must have (we have used $e^2 = 4\pi\alpha$)

$$\frac{2}{3\pi\epsilon} \frac{d\alpha}{d \ln \mu^2} + \frac{d \ln \alpha}{d \ln \mu^2} + \frac{\epsilon}{2} = 0, \quad (3.49)$$

or equivalently

$$\frac{2\alpha}{3\pi\epsilon} \frac{d\alpha}{d \ln \mu^2} + \frac{d\alpha}{d \ln \mu^2} + \frac{\epsilon\alpha}{2} = 0. \quad (3.50)$$

Now we can solve this equation perturbatively. First suppose that

$$\frac{d\alpha}{d \ln \mu^2} = A\alpha + B\alpha^2 + \dots, \quad (3.51)$$

then

$$\frac{2\alpha}{3\pi\epsilon} (A\alpha + B\alpha^2) + A\alpha + B\alpha^2 + \frac{\epsilon\alpha}{2} = 0, \quad (3.52)$$

grouping by powers of α

$$\left(\frac{2}{3\pi\epsilon}A + B\right)\alpha^2 + \left(A + \frac{\epsilon}{2}\right)\alpha + \frac{2\alpha^3}{3\pi\epsilon}B = 0, \quad (3.53)$$

we arrive to

$$A = -\frac{\epsilon}{2} \quad B = -\frac{2}{3\pi\epsilon}A = \frac{1}{3\pi}, \quad (3.54)$$

then, Eq. (3.51) turns out to be

$$\frac{d\alpha}{d \ln \mu^2} = -\frac{\epsilon\alpha}{2} + \frac{\alpha^2}{3\pi} + \dots, \quad (3.55)$$

which, in the limit when $\epsilon \rightarrow 0$ gives

$$\frac{d\alpha}{d \ln \mu^2} = \frac{\alpha^2}{3\pi} + \dots \quad (3.56)$$

the fact that the divergences originated by ϵ vanished is a consequence of the renormalizability of the theory. This result is exactly the same result that we got before, and it is the RGE equation for α in the $\overline{\text{MS}}$ scheme. To be consistent with the notation of our paper, I will denote any quantity in this scheme by a caret. For example I will replace $\alpha \rightarrow \hat{\alpha}$.

To conclude this subsection, I would like to talk about the running of the charge in the on-shell scheme. In this case the renormalized charge is the same as the charge measured at $p^2 = 0$. Then the Z_3 term absorbs all the μ dependence. Technically speaking there is no running of the charge in terms of μ . Nevertheless, one can define an effective charge where the "running" is given explicitly by the physical momenta p^2 . The dependence of this effective charge on the physical momenta can easily be obtained performing a geometric sum of an arbitrary number of insertions of one-particle irreducible diagrams. To one loop this re summation is given by

$$\alpha(p) = \frac{\alpha_0}{1 + 4\pi\alpha_0 (\Pi(p, \mu) - \Pi(0, \mu))}. \quad (3.57)$$

3.4.3 Matching relations at one loop

In this thesis, we will use an effective field theory framework. This is a wide topic that has had a tremendous impact and applications in recent years. A list of lectures on the topic can be found in [84, 85]. Here I will base the analysis of this section on the lectures given by Grozin [85].

Suppose that we have QED with two types of particles, electrons and muons. If the energy of the process is small, the effects of the muons should be suppressed by a power expansion

of the form $\frac{E}{m_\mu}$. The full Lagrangian includes both particles: electrons and muons,

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}^{(0)}F^{(0)\mu\nu} + i\sum_j\bar{\psi}_j^{(0)}\gamma^\mu(\partial_\mu + ie^0A_\mu^0)\psi_j^{(0)} + \sum_j m_j^{(0)}\bar{\psi}_j^{(0)}\psi_j^{(0)}, \quad (3.58)$$

where $m_2 = M$ is the mass of the muon. The effective Lagrangian must respect the $U(1)$ symmetry and it is a power expansion in $\frac{1}{m_\mu}$ with no muon fields present in the theory. It is clear then, that we must have something of the form

$$\mathcal{L}'_{QED} = -\frac{1}{4}F_{\mu\nu}'^{(0)}F'^{(0)\mu\nu} + i\bar{\psi}'^{(0)}_e\gamma^\mu D'_\mu\psi_e'^{(0)} + m_e'^{(0)}\bar{\psi}'^{(0)}_e\psi_e'^{(0)} + \frac{c}{m_\mu}\bar{\psi}'_e F'^{(0)\mu\nu}\sigma_{\mu\nu}\psi'_e + \mathcal{O}\left(\frac{1}{m_\mu^2}\right), \quad (3.59)$$

The effective theory should give the same physical results as the full theory. To ensure this, matching conditions between fields and couplings must be computed. Here we will only worry about the matching condition of the photon field³. Thus we define

$$A = (\zeta_A)^{\frac{1}{2}}A'. \quad (3.60)$$

The factor ζ_A will scale the pole of the propagator. The next step in the matching procedure is to require the photon propagator to have the same residue for the pole at $p^2 = 0$. The $\overline{\text{MS}}$ scheme is defined as the one where the counter terms absorb only the UV infinities. Then the residue of the propagator for the full theory is given by

$$\text{R}^{full} = 1 - e^2\bar{\Pi}(0, \mu^2, m_e^2) - e^2\bar{\Pi}(0, \mu^2, m_\mu^2) \quad (3.61)$$

while the residue for the effective theory is given only by the electron contribution

$$\text{R}^{eff} = 1 - e^2\bar{\Pi}(0, \mu^2, m_e^2) \quad (3.62)$$

³This is enough to compute the matching of α due to the Ward-Takashi identity.

If we want both theories to have the same residue we must have

$$\zeta_A = 1 - e^2 \bar{\Pi}(0, \mu^2, m_\mu^2) \quad (3.63)$$

up to high order corrections. From the Ward identity, it can be shown that the decoupling constant for the electric charge is $\zeta_e = \zeta_A^{-\frac{1}{2}}$, which implies that the decoupling constant for α is equal to $\zeta_\alpha = \zeta_A^{-1}$. Then we must have

$$\alpha(\mu) = \alpha'(\mu) \left[1 - \frac{2\alpha(\mu)}{3\pi} \ln \left(\frac{\mu}{m_\mu} \right) \right]. \quad (3.64)$$

From here it is clear that if we match at m_μ there is no discontinuity in the fine structure constant α at the one loop level.

Now, what is the next order correction? Well, for leptons it would be the two-loop QED diagrams, like the one with a photon inside the one loop bubble. But for quarks we can have a gluon inside the quark loop. This gives mixed QED-QCD corrections. This is the topic of the next subsections.

3.4.4 Running of $\hat{\alpha}_s$

Before taking into account the effect of QCD in the running of $\hat{\alpha}$ we must understand the running of the strong coupling constant. As we have mentioned before, QCD is the quantum field theory that emerges from the $SU(3)_c$ gauge symmetry among quarks. As consequence of this non Abelian gauge theory there will be interactions among the gauge bosons. This is different to the Abelian case (like QED) where the photon does not have an electric charge and cannot couple to itself. This difference has an important implication on the running of the gauge coupling g_s . Due to the different statistics between bosons and fermions, the vacuum bubbles originated from bosons and fermions will have opposite sign, so if there are not enough fermions in the theory, the boson bubbles will dominate and the sign of the beta

function will be opposite to the one in the Abelian case ($\beta > 0$). A positive β function implies that as we increase the scale, the coupling also increases. On the other hand, if the beta function is negative the coupling constant increases as we decrease the energy scale⁴. At some point the coupling is so large that the theory becomes strongly coupled and non perturbative. The one loop RGE for the strong coupling constant is derived in any standard QFT book. The result is

$$\frac{d\hat{\alpha}_s(\mu)}{d\ln\mu^2} = \frac{\hat{\alpha}_s^2}{\pi} \left(\frac{n_f}{6} - \frac{11}{4} \right) \quad (3.65)$$

where we defined $\hat{\alpha}_s \equiv \frac{g_s^2}{4\pi}$. If $n_f < 33/2$ the beta function is negative and $\hat{\alpha}_s$ becomes large at low scales. The solution to this equation is

$$\hat{\alpha}_s(\mu) = \frac{\hat{\alpha}_s(\mu_0)}{1 - \frac{\hat{\alpha}_s(\mu_0)}{\pi} \left(\frac{n_f}{6} - \frac{11}{4} \right) \ln \frac{\mu^2}{\mu_0^2}}. \quad (3.66)$$

Due to this non perturbative behavior it is necessary to compute many loops to get accurate results. The five loop QCD running constant has already been computed in 2016 [87]:

$$\mu^2 \frac{d}{d\mu^2} \hat{\alpha}_s(\mu) = -\pi \left(\frac{\hat{\alpha}_s}{\pi} \right)^2 \sum_{i=0} \beta_i^{QCD} \left(\frac{\hat{\alpha}_s}{\pi} \right)^i, \quad (3.67)$$

where

$$\beta_0^{QCD} = \frac{1}{4} \left\{ 11 - \frac{2}{3} n_f \right\}, \quad \beta_1 = \frac{1}{42} \left\{ 102 - \frac{38}{3} n_f \right\}, \quad (3.68)$$

$$\beta_2^{QCD} = \frac{1}{43} \left\{ \frac{2857}{2} - \frac{5033}{18} n_f + \frac{325}{54} n_f^2 \right\}, \quad (3.69)$$

$$\begin{aligned} \beta_3^{QCD} = \frac{1}{44} \left\{ \frac{149753}{6} + 3564\zeta_3 - \left[\frac{1078361}{162} + \frac{6508}{27}\zeta_3 \right] n_f \right. \\ \left. + \left[\frac{50065}{162} + \frac{6472}{81}\zeta_3 \right] n_f^2 + \frac{1093}{729} n_f^3 \right\}, \end{aligned} \quad (3.70)$$

⁴This is the idea behind asymptotic freedom [86]: non Abelian gauge theories become weakly coupled as the energy becomes larger.

(3.71)

$$\beta_4^{QCD} = \frac{1}{4^5} \left\{ \frac{8157455}{16} + \frac{621885}{2} \zeta_3 - \frac{88209}{2} \zeta_4 - 288090 \zeta_5 \right.$$

$$+ n_f \left[-\frac{336460813}{1944} - \frac{4811164}{81} \zeta_3 \right.$$

$$\left. \left. + \frac{33935}{6} \zeta_4 + \frac{1358995}{27} \zeta_5 \right] \right.$$

(3.72)

$$+ n_f^2 \left[\frac{25960913}{1944} + \frac{698531}{81} \zeta_3 - \frac{10526}{9} \zeta_4 - \frac{381760}{81} \zeta_5 \right]$$

$$+ n_f^3 \left[-\frac{630559}{5832} - \frac{48722}{243} \zeta_3 + \frac{1618}{27} \zeta_4 + \frac{460}{9} \zeta_5 \right]$$

$$+ n_f^4 \left[\frac{1205}{2916} - \frac{152}{81} \zeta_3 \right] \left. \right\},$$

(3.73)

and n_f denotes the number of active quark flavors.

In the same way as for QED one can compute the matching conditions for QCD. This has already been computed to four loops in [88, 89], and their result is

$$\frac{\hat{\alpha}'_s(\mu)}{\hat{\alpha}_s(\mu)} = 1 + \frac{\hat{\alpha}_s(\mu)}{\pi} \left(-\frac{1}{6} \ln \frac{\mu^2}{\hat{m}^2} \right) + \left(\frac{\hat{\alpha}_s(\mu)}{\pi} \right)^2 \left(\frac{11}{72} - \frac{11}{24} \ln \frac{\mu^2}{\hat{m}^2} + \frac{1}{36} \ln^2 \frac{\mu^2}{\hat{m}^2} \right)$$

$$+ \left(\frac{\hat{\alpha}_s(\mu)}{\pi} \right)^3 \left[\frac{564731}{124416} - \frac{82043}{27648} \zeta(3) - \frac{955}{576} \ln \frac{\mu^2}{\hat{m}^2} + \frac{53}{576} \ln^2 \frac{\mu^2}{\hat{m}^2} - \frac{1}{216} \ln^3 \frac{\mu^2}{\hat{m}^2} \right.$$

$$+ n_l \left(-\frac{2633}{31104} + \frac{67}{576} \ln \frac{\mu^2}{\hat{m}^2} - \frac{1}{36} \ln^2 \frac{\mu^2}{\hat{m}^2} \right) + \left(\frac{\hat{\alpha}_s(\mu)}{\pi} \right)^4 \left[\frac{291716893}{6123600} \right.$$

$$+ \frac{3031309}{1306368} \ln^4 2 - \frac{121}{4320} \ln^5 2 - \frac{3031309}{217728} \zeta(2) \ln^2 2 + \frac{121}{432} \zeta(2) \ln^3 2 - \frac{2362581983}{87091200} \zeta(3)$$

$$- \frac{76940219}{2177280} \zeta(4) + \frac{2057}{576} \zeta(4) \ln 2 + \frac{1389}{256} \zeta(5) + \frac{3031309}{54432} a_4 + \frac{121}{36} a_5 - \frac{151369}{2177280} X_0$$

$$+ \left(\frac{7391699}{746496} - \frac{2529743}{165888} \zeta(3) \right) \ln \frac{\mu^2}{\hat{m}_h^2} + \frac{2177}{3456} \ln^2 \frac{\mu^2}{\hat{m}^2} - \frac{1883}{10368} \ln^3 \frac{\mu^2}{\hat{m}^2} + \frac{1}{1296} \ln^4 \frac{\mu^2}{\hat{m}^2}$$

$$+ n_l \left(-\frac{4770941}{2239488} + \frac{685}{124416} \ln^4 2 - \frac{685}{20736} \zeta(2) \ln^2 2 + \frac{3645913}{995328} \zeta(3) \right.$$

$$- \frac{541549}{165888} \zeta(4) + \frac{115}{576} \zeta(5) + \frac{685}{5184} a_4 + \left(-\frac{110341}{373248} + \frac{110779}{82944} \zeta(3) \right) \ln \frac{\mu^2}{\hat{m}^2}$$

$$- \frac{1483}{10368} \ln^2 \frac{\mu^2}{\hat{m}^2} - \frac{127}{5184} \ln^3 \frac{\mu^2}{\hat{m}^2} \left. \right) + n_l^2 \left(-\frac{271883}{4478976} + \frac{167}{5184} \zeta(3) + \frac{6865}{186624} \ln \frac{\mu^2}{\hat{m}^2} \right.$$

$$\left. - \frac{77}{20736} \ln^2 \frac{\mu^2}{\hat{m}^2} + \frac{1}{324} \ln^3 \frac{\mu^2}{\hat{m}^2} \right) + \mathcal{O} \left(\left(\frac{\hat{\alpha}_s(\mu)}{\pi} \right)^5 \right),$$

(3.74)

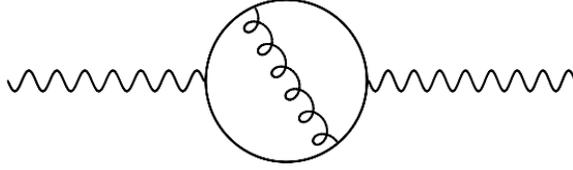


Figure 3.2: Quark diagrams that contribute to the photon two-point Green's function.

where $\zeta(n)$ is Riemann's zeta function, $a_n = \text{Li}_n(1/2) = \sum_{k=1}^{\infty} 1/(2^k k^n)$ and X_0 is a constant with value given in [88, 89]. With an input value at some scale μ_0 for $\hat{\alpha}_s(\mu_0)$ (typically $\mu_0 = M_z$) we can obtain the full running of the strong coupling constant using the RGE and the matching conditions. The running can be solved numerically by explicit integration or approximated, being careful of re-summing all the logarithms to the respective order. A compendium of methods and algorithms to re-sum the logs can be found in [90]. As we increase the number of loops we get more and more accurate approximations in the perturbative region of the theory. Unfortunately it gives us no useful results for the non-perturbative limit. To get useful results in that region, one usually has to rely on other methods, such as dispersion relations.

3.4.5 QCD contribution to the running of $\hat{\alpha}$

The type of diagrams that give a contribution to the running of α at two loops in QCD are shown in Fig. 3.2. These diagrams will contribute in two ways, in the RGE running and in the matching condition. The RGE running can be obtained in the usual way: computing Z_A and using it in the relation between the bare charge and the renormalized charge. Another way is to extract the logs from the high energy expansion of the vacuum polarization function, or from the logs of $\bar{\Pi}(0, \mu)$. The one-loop QED and two loop QCD RGE for the fine structure constant is [91]

$$\mu^2 \frac{d\hat{\alpha}}{d\mu^2} = \frac{\hat{\alpha}}{3\pi} \left[\sum_i K_i Q_i^2 \right], \quad (3.75)$$

where Q_i is the charge of the fermion in the loop (it appears in the vertices in the one loop bubble diagram), K_i is equal to one for charged leptons, and $K_i = N_i^C \left(1 + \frac{\hat{\alpha}_s}{\pi}\right)$ for quarks (where $N_i^C = 3$ is the number of colors). These contributions come from diagrams like the one shown in Fig. 3.2 where a gluon can be interchanged by the quarks in the loop.

For the RGE matching we must also include the constant terms in $\bar{\Pi}(0, \mu)$. These terms can be obtained from the low energy expansion⁵ of this vacuum polarization function in [92–95]. In the case of quarks, we have [93]

$$\bar{\Pi}(0, \mu, m) = \frac{Q_i^2}{4\pi^2} \left[\ln \frac{\mu^2}{m^2} + \frac{\hat{\alpha}_s(\mu)}{\pi} \left(\frac{15}{4} + \ln \frac{\mu^2}{m^2} \right) \right], \quad (3.76)$$

where m is the QCD pole mass and Q_i is the charge of the corresponding quark. Note that the coefficient of the log in the $\hat{\alpha}_s/\pi$ term is the same as the coefficient of the first order logarithm. This is equivalent to our statement that for quarks $K_i = N_i^C(1 + \hat{\alpha}_s/\pi)$.

In our work we use the $\overline{\text{MS}}$ mass, denoted as $\hat{m}(\mu)$. Thus we need to change m by \hat{m} . The relation between both masses is given in [93]. To order $\hat{\alpha}_s$ the relation is

$$\frac{\hat{m}(\mu)}{m} = 1 + \frac{\hat{\alpha}_s}{\pi} \left(-\frac{4}{3} - \ln \left(\frac{\mu^2}{m^2} \right) \right), \quad (3.77)$$

and we have

$$\bar{\Pi}(0, \mu, \hat{m}) = \frac{Q_i^2}{4\pi^2} \left[\ln \frac{\mu^2}{\hat{m}^2(\mu)} + \frac{\hat{\alpha}_s(\mu)}{\pi} \left(\frac{13}{12} - \ln \frac{\mu^2}{\hat{m}^2(\mu)} \right) \right]. \quad (3.78)$$

With this results and Eq. (3.63) we get that the decoupling relation of the fine structure for a quark of charge Q_i at one loop in QCD at $\mu = \hat{m}(\mu)$ is

$$\frac{\hat{\alpha}(\hat{m})}{\pi} = \frac{\hat{\alpha}'(\hat{m})}{\pi} \left[1 + \frac{\hat{\alpha}'(\hat{m})}{\pi} Q_i^2 \frac{\hat{\alpha}_s(\hat{m})}{\pi} \frac{13}{12} \right]. \quad (3.79)$$

In an ideal world, where QCD would be perturbative, we would have everything to

⁵[92] already computed the decoupling relation for the QED fine structure constant at order α_s^3

compute the running of the fine structure constant. The only inputs required would be the value of the fine structure constant at zero, the masses and charges of the particles involved in the loops, and the value of $\hat{\alpha}_s$ at some scale. Then, we can solve the RGE numerically or re-sum the logs with an approximate analytical formula [96]. Nevertheless, due to the non perturbative nature of QCD, one more ingredient needs to be added to this formula: the quark contributions at low energies. This is the topic of the next section.

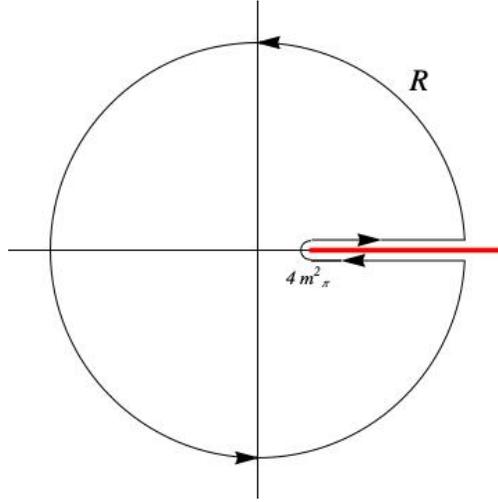
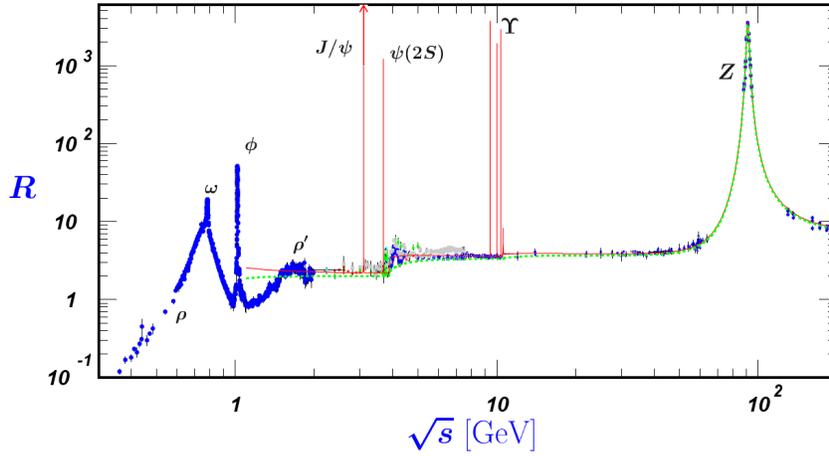
3.4.6 Breaking of perturbation theory, dispersion relations

As we have mentioned before, if QCD would be perturbative, this is all we would need here. Once we have the RGE formula we get the full running. Nevertheless, there is a problem that emerges once QCD is included in the theory: at low energies the theory becomes non perturbative. Then we must change our method to compute the vacuum polarization function and rely on experimental results. The vacuum polarization from quarks (which are the ones affected by QCD) is an analytic function, with a branch-cut along the real positive axis. This branch cut starts at $s = 4m_\pi^2$, which is the energy required to produce a pion pair. Using Cauchy theorem, we can find the value of the vacuum polarization function at any scale,

$$\bar{\Pi}(p, \mu) = \frac{1}{2\pi i} \oint \frac{\bar{\Pi}(p', \mu)}{p'^2 - p^2} dp'^2 \quad (3.80)$$

where we use the contour shown in Fig. 3.3. At low energies, we do not have an equation for this vacuum polarization function, due to the non perturbativity of the series. Nevertheless, let us continue with the computation. The path integral is (I will rewrite $\bar{\Pi}(p, \mu)$ as $\bar{\Pi}(p^2, \mu)$ since we know that it is really a function of p^2)

$$\bar{\Pi}(p^2, \mu) = \frac{1}{2\pi i} \int_R^{4m_\pi^2} \frac{\bar{\Pi}(t - i\epsilon, \mu)}{t - i\epsilon - p^2} dt + \frac{1}{2\pi i} \int_{4m_\pi^2}^R \frac{\bar{\Pi}(t + i\epsilon, \mu)}{t + i\epsilon - p^2} dt + \frac{1}{2\pi i} \int_C \frac{\bar{\Pi}(p'^2, \mu)}{p'^2 - p^2} dp'^2, \quad (3.81)$$


 Figure 3.3: Contour to obtain α from experimental results

 Figure 3.4: $R(s)$ for different center of mass energy s .

where C denotes an integral along the circle. Now we use the fact that for an analytical function, $f(z^*) = f(z)^*$, and with a bit of algebra we get,

$$\bar{\Pi}(p^2, \mu) = \frac{1}{\pi} \int_{4m_\pi^2}^R \frac{\text{Im}\bar{\Pi}(t, \mu)}{t - p^2 - i\epsilon} dt + \frac{1}{2\pi i} \int_C \frac{\bar{\Pi}(p', \mu)}{p'^2 - p^2} dp'^2. \quad (3.82)$$

The next step is to use the optical theorem, which relates the imaginary part of the vacuum polarization function to the square of the amplitude, or in other words to a cross section. In this case we are taking into account only the contributions of hadrons, since

for leptons we can compute everything in a perturbative way. Experimental collaborations usually quote $R(s)$ (shown in Fig. 3.4) which is defined as the ratio

$$R(s) \equiv \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)}. \quad (3.83)$$

It is straightforward to relate this quantity to $\text{Im}\bar{\Pi}(p^2, \mu)$ with the help of the optical theorem, with the result $R(s) = 12\pi\text{Im}\bar{\Pi}(s, \mu)$. Note that the left-hand side of this equation must not depend on the renormalization scale μ , so it might seem that there is something odd going on here, but if the computation is explicitly done, it turns out that the imaginary part of the vacuum polarization function is μ independent as it should be. Now, let us compute the contribution to the running of $\hat{\alpha}$ of the light quarks. To compute this running, we need to calculate $\bar{\Pi}(0, \mu/\mu_0)$. Consider the following trick [96]: set the radius R of the circle to be equal to μ^2 (the renormalization scale), and p^2 equal to 0. Thus we get

$$\bar{\Pi}(0, \mu) = \frac{1}{12\pi^2} \int_{4m_\pi^2}^{\mu^2} \frac{R(t)}{t - i\epsilon} dt + \frac{1}{2\pi i} \int_C \frac{\bar{\Pi}(p'^2, \mu)}{p'^2} dp'^2. \quad (3.84)$$

we see that if $\mu = 4m_\pi^2$ then the contribution from the first term in the right-hand side is equal to zero. The second term will be later understood as a matching condition from a theory without light quarks to a theory which contains them.

The idea of splitting $\bar{\Pi}(0, \mu)$ in this way is to choose a μ such that the circle integral can be performed perturbatively while the integral over the line is computed using the experimental data. It is a smart idea to choose a μ to reduce the error, taking the most advantage of both, the experimental and the theoretical estimations. A common value of such μ is 2 GeV [97]. If we write the RGE running as

$$\hat{\alpha}(\mu) = \frac{\alpha}{1 - \Delta_{\text{had}}\hat{\alpha}(\mu) - \Delta_{\text{lep}}\hat{\alpha}(\mu)}, \quad (3.85)$$

where α is the fine structure constant at zero momentum and compare with Eq. (3.38) we get

$$\Delta_{\text{had}}\hat{\alpha}^{(3)}(2.0 \text{ GeV}) = \frac{\alpha}{3\pi} \int_{4m_\pi^2}^{4.0 \text{ GeV}^2} \frac{R(t)}{t - i\epsilon} dt - 2i\alpha \int_C \frac{\bar{\Pi}(p'^2, 2.0 \text{ GeV})}{p'^2} dp'^2. \quad (3.86)$$

Here, the superscript (3) means that we are only considering the three light quarks. We can follow a similar approach to compute the hadronic contribution to the running of the effective fine structure constant. Except that in this case, the running is given by the "physical" momenta p^2 . Then in the on-shell scheme $\Delta\alpha(p^2) = -4\pi\alpha (\bar{\Pi}(p^2, \mu) - \bar{\Pi}(0, \mu))$ (see Eq.(3.57)). So we need to compute the difference $\bar{\Pi}(p^2, \mu) - \bar{\Pi}(0, \mu)$. To do so, we use Eq. 3.82 with $R \rightarrow \infty$. After a bit of algebra, it is easy to obtain

$$\bar{\Pi}(p^2, \mu) - \bar{\Pi}(0, \mu) = \frac{p^2}{12\pi^2} \int_{4m_\pi^2}^{\infty} \frac{R(s)}{s(s - p^2 - i\epsilon)} ds. \quad (3.87)$$

There is another consideration that we have to take into account here. When $p^2 < 0$ (t-channel), the vacuum polarization function is real. On the other hand, when $p^2 > 4m_{\pi^2}$, like in the s-channel, the vacuum polarization function can acquire an imaginary part. To get a real α in the s channel, we define $\Delta\alpha$ as the real part of this difference. Actually, we are formally not allowed to take explicitly the value of p^2 in the real axis above the threshold $4m_\pi^2$, due to branch cut on this axis. Nevertheless, we can take the limit when we include a small imaginary part for p^2 and taking the limit when this imaginary term goes to zero. With all this in mind we can take

$$\Delta\alpha(M_Z^2)_{\text{had}} = -\frac{\alpha M_Z^2}{3\pi} \int_{4m_\pi^2}^{\infty} \frac{R(s)}{s(s - M_Z^2 - i\epsilon)} ds. \quad (3.88)$$

The real part is picked up for this integral if it is computed as the Cauchy principal value [98].

3.5 The running of the weak mixing angle

In previous sections I have highlighted how the renormalization procedure takes place in QCD and QED. Now it is time to explain how this procedure takes place in the Standard Model. This is a wide and complex topic and it is not the intention of this thesis to explain it in full detail. In this section I follow and compress the results explained in [99–104]

First, let me recall that at tree level the following relations of the Standard Model hold,

$$\sin^2 \theta_W = 1 - \frac{M_W^2}{M_Z^2} = \frac{g'^2}{g^2 + g'^2} = \frac{\pi\alpha}{\sqrt{2}G_F M_Z^2 \cos^2 \theta_W}. \quad (3.89)$$

Nevertheless, when loops appear things start to get messy. Different collaborations use different definitions of the weak angle, depending on the quantities measured by the experimental collaborations. There has been discussions and papers that tackle these different definitions of the weak mixing angle [99, 101]. A common definition is the on-shell weak mixing angle, which imposes the relation

$$\sin^2 \theta_W = 1 - \frac{M_{poleW}^2}{M_{poleZ}^2}, \quad (3.90)$$

to all orders in perturbation theory. This definition is easy to understand, but it has the drawback that some observables have an enhanced dependence on m_t [105], leading to a weak mixing angle with larger uncertainty. On the other hand, we have the $\overline{\text{MS}}$ scheme where the weak mixing angle is defined as

$$\sin^2 \hat{\theta} = \frac{\hat{g}'^2}{\hat{g}'^2 + \hat{g}^2}, \quad (3.91)$$

which is a theoretician definition, and sometimes its physical meaning is somewhat more obscure. The advantage is that it is easy to implement, and it has a straightforward connection and applications to BSM models like Grand Unification. It also has the advantage that the inclusion of hadrons corrections can be implemented in a systematic way, which is the main topic of the next chapter. Now let us give a brief explanation of the renormalization of the

Standard Model. The scalings of the couplings and fields are the following⁶

$$\begin{aligned}
 W_\mu^a &\rightarrow (Z_3^W) W_\mu^a, & Q_L^i &\rightarrow (Z_{qL}^i) Q_L^i, \\
 B_\mu &\rightarrow (Z_3^B) B_\mu, & q_{u,R}^i &\rightarrow (Z_{qu,R}^i) q_{u,R}^i, \\
 \Psi_L^i &\rightarrow (Z_{\psi L}^i) \Psi_L^i, & H &\rightarrow (Z_H)^{\frac{1}{2}} H, \\
 \psi_R^i &\rightarrow (Z_{\psi R}^i) \psi_R^i, & g &\rightarrow Z_g^W (Z_3^W)^{-\frac{3}{2}} g, \\
 v &\rightarrow (Z_H)^{\frac{1}{2}} (v - \delta v), & g' &\rightarrow Z_{g'}^B (Z_3^B)^{-\frac{3}{2}} g', \\
 \mu^2 &\rightarrow (Z_H)^{-1} (\mu^2 - \delta\mu^2), & y_u^j &\rightarrow (Z_H)^{-\frac{1}{2}} Z_y^j y_u^j, \\
 \lambda &\rightarrow (Z_H)^{-2} Z_\lambda \lambda,
 \end{aligned} \tag{3.93}$$

To compute the running of the weak angle it is easier to go to the basis of W, Z and γ . The counter terms of the W_3 and A , are related to the ones of the W and B by ⁷

$$\begin{pmatrix} \delta Z_{e(3)}^\gamma \\ \delta Z_{g_z(3)}^Z \end{pmatrix} = \begin{pmatrix} \sin^2 \theta & \cos^2 \theta \\ \cos^2 \theta & \sin^2 \theta \end{pmatrix} \begin{pmatrix} \delta Z_{g(3)}^W \\ \delta Z_{g'(3)}^B \end{pmatrix}. \tag{3.94}$$

A nice and reassuring feature of the weak angle, is that we do not need to compute the value of all the counter terms in the $\overline{\text{MS}}$ scheme to get its running. Actually, the only thing that we need are the counter terms related to the masses of the gauge bosons, or the γZ mixing. The mass counter terms can be obtained from the pole of the Z and W propagators. Due to the symmetries and structure of the Lagrangian, these counter terms will be functions of the ones defined in Eq. (3.94), and in consequence functions of the original counter terms from Eq. (3.93). The first step to get the renormalization constants for the masses is to compute the one loop correction to the propagator of the Z gauge boson. In this section I will only

⁶One can choose the counter term for the Higgs vacuum in such a way that it cancels the linear term in the Higgs potential induced by tadpole diagrams, to keep the relation

$$v = \frac{2\mu}{\sqrt{\lambda}} \tag{3.92}$$

remains valid for the renormalized parameters.

⁷ $Z_i = 1 + \delta Z_i$

consider the contribution that comes from a diagram with a fermion running in an internal bubble just like the one in Fig. 3.1 but with external Z bosons. This one loop diagram can be split in a term proportional to $g^{\mu\nu}$ and another one proportional to $p^\mu p^\nu$. To compute the pole of the Green's function, we only need the part proportional to $g^{\mu\nu}$. Furthermore, the $p^\mu p^\nu$ term is suppressed in physical amplitudes by a factor of m_i^2/M_V^2 . Thus we get

$$iD_{\mu\nu}^Z(p) = \frac{-ig^{\mu\nu}}{p^2 - M_Z^2} \left(1 + (\Pi_{ZZ}(p) + \delta M_Z^2 - \delta Z_3^Z(p^2 - M_Z^2)) \frac{1}{p^2 - M_Z^2} \right) + p^\mu p^\nu \text{ terms}, \quad (3.95)$$

where $\Pi_{ZZ}(p)$ is the term proportional to $g^{\mu\nu}$ in the one loop diagram, δM_Z^2 is the counter term that relates the mass in the corresponding scheme to the bare mass and δZ_3^Z is the wave function renormalization counter term. The next step is to re-sum this expression including insertions of one particle irreducible diagrams, in a similar way as how we did it for QED.

The result is

$$iD_{\mu\nu}^Z(p) = \frac{-ig^{\mu\nu}}{p^2 - M_Z^2 - \delta M_Z^2 + \delta Z_3^Z(p^2 - M_Z^2) - \Pi_{ZZ}(p^2)} + p^\mu p^\nu \text{ terms}. \quad (3.96)$$

The mass is related to the real part of the pole of this propagator (the complex part of the pole is related to the decay width of the Z boson). Then, the pole mass is at

$$M_{Z(\text{pole})}^2 - M_Z^2 - \delta M_Z^2 + \delta Z_3^Z(M_{Z(\text{pole})}^2 - M_Z^2) - \Re \left[\Pi_{ZZ}(M_{Z(\text{pole})}^2) \right] = 0, \quad (3.97)$$

where the term $(M_{Z(\text{pole})}^2 - M_Z^2)$ is of one loop order. This implies that $\delta Z_3^Z(M_{Z(\text{pole})}^2 - M_Z^2)$ is higher order, and we can deduce

$$M_{Z(\text{pole})}^2 - M_Z^2 - \delta M_Z^2 - \Re \left[\Pi_{ZZ}(M_{Z(\text{pole})}^2) \right] = 0. \quad (3.98)$$

If we are to set $M_Z = M_{\text{pole}}$, or in other words if we make the parameter M_Z of the

Lagrangian equal to the pole mass, then we have

$$\delta M_Z^2 = -\Re \left[\Pi_{ZZ} \left(M_{Z(pole)}^2 \right) \right]. \quad (3.99)$$

If, on the other hand, we set M_Z^2 to be equal to the $\overline{\text{MS}}$ mass, then only the divergent terms are absorbed in δM_Z^2 . In a similar way we get for the W boson

$$\delta M_W^2 = -\Re \left[\Pi_{WW} \left(M_{W(pole)}^2 \right) \right]. \quad (3.100)$$

To obtain the running of $\sin^2 \hat{\theta}_W$ we can make use of a useful property in the $\overline{\text{MS}}$ scheme, which is the fact that in this scheme, the relation between the $\overline{\text{MS}}$ masses of the gauge bosons and the weak mixing angle is the same as the one at three level⁸. Then we have

$$\sin^2 \theta_W^0 = 1 - \frac{M_W^{(0)2}}{M_Z^{(0)2}} = 1 - \frac{\hat{M}_W^2 + \delta M_W^2}{\hat{M}_Z^2 + \delta M_Z^2} = \sin^2 \hat{\theta}_W + \cos^2 \hat{\theta}_W \left(\frac{\delta M_Z^2}{\hat{M}_Z^2} - \frac{\delta M_W^2}{\hat{M}_W^2} \right). \quad (3.102)$$

The next step is to calculate the bubble contributions to the gauge boson propagators. Here I will quote the results for the fermion bubbles. I do not compute these bubbles here, but the interested reader can look up chapter 31 of [76], or if a more complete treatment is desired [102]. For a fermion doublet we have⁹ with $\sin^2 \hat{\theta} \equiv \hat{s}^2$:

$$\Pi_{WW} \left(p^2, \mu^2 \right) = \frac{N^c}{\sin^2 \hat{\theta}_W} \frac{1}{2} \Pi_{LL} \left(p^2, \Delta_{ud}, \mu^2 \right), \quad (3.103)$$

⁸If instead of using the $\overline{\text{MS}}$ masses for the gauge bosons, one uses the on-shell definition for them, then the tree level relation gets modified to

$$M_W^2 = \hat{\rho} M_Z^2 (1 - \sin^2 \hat{\theta}_W). \quad (3.101)$$

⁹Here the index u and d refer not only to the up and down quarks, but the up and down elements of any fermion doublet.

$$\Pi_{ZZ}(p^2, \mu^2) = \frac{1}{\hat{s}^2 \hat{c}^2} \sum_{f=u,d} N_f^c \left((t_f^3)^2 \Pi_{LL}(p^2, \Delta_{ff}) - \hat{s}^2 Q_f (t_f^3 - \hat{s}^2 Q_f) \Pi_{VV}(p^2, m_f) \right), \quad (3.104)$$

$$\Pi_{\gamma Z}(p^2, \mu^2) = \frac{1}{\hat{s} \hat{c}} \sum_{f=u,d} N_f^c \left(t_f^3 Q_f \frac{1}{2} - \hat{s}^2 Q_f^2 \right) \Pi_{VV}(p^2, m_f), \quad (3.105)$$

where

$$\Pi_{LL}(p^2, \Delta_{ud}, \mu^2) = \frac{\hat{\alpha}}{\pi} \left\{ \frac{m_u^2 + m_d^2 - \frac{2}{3} p^2}{2\epsilon} - \frac{1}{2} \int_0^1 dx [x(1-x)p^2 \Delta_{ud}] \ln \frac{\mu^2}{\Delta_{ud}} \right\}, \quad (3.106)$$

$$\Pi_{VV}(p^2, m_f^2, \mu^2) = -\frac{2\alpha}{\pi} p^2 \left\{ \frac{1}{3\epsilon} + \int_0^1 dx [x(1-x)] \ln \frac{\mu^2}{m_f - p^2 x(1-x)} \right\}, \quad (3.107)$$

and

$$\Delta_{ij} = x m_i^2 + (1-x) m_j^2 - x(1-x) p^2, \quad (3.108)$$

and N_f^C is the number of colors. After some algebraic steps and the relation $Q_u - Q_d = 1$ one can show that the divergent part satisfies

$$\cos^2 \hat{\theta}_W \left(\frac{\delta M_Z^2}{M_Z^2} - \frac{\delta M_W^2}{M_W^2} \right) = - \sum_{f=u,d} N_f^c \left(t_f^3 Q_f \frac{1}{2} - \sin^2 \hat{\theta}_W Q_f^2 \right) \left\{ \frac{\alpha}{\pi} \frac{2}{3\epsilon} \right\}. \quad (3.109)$$

Notice that the coefficient is just proportional to the γ - Z mixing bubble, and in consequence, to the vector vacuum polarization bubble. Substituting this result into Eq. (3.102) we get

$$\sin^2 \theta_W^0 = \sin^2 \hat{\theta}_W - \sum_{f=u,d} N_f^c \left(t_f^3 Q_f \frac{1}{2} - \sin^2 \hat{\theta}_W Q_f^2 \right) \left\{ \frac{\alpha}{\pi} \frac{2}{3\epsilon} \right\}. \quad (3.110)$$

Taking the derivative with respect to $\ln \mu^2$, we obtain

$$\frac{d \sin^2 \hat{\theta}_W}{d \ln \mu^2} = \frac{d\alpha}{d \ln \mu^2} \frac{1}{\pi} \frac{2}{3\epsilon} \sum_{f=u,d} N_f^c \left(t_f^3 Q_f \frac{1}{2} - \sin^2 \hat{\theta}_W Q_f^2 \right), \quad (3.111)$$

but recalling that to zeroth order

$$\frac{d\hat{\alpha}}{d \ln \mu^2} = -\frac{\hat{\alpha}}{2}\epsilon, \quad (3.112)$$

one finds

$$\frac{d \sin^2 \hat{\theta}_W}{d \ln \mu^2} = -\frac{\hat{\alpha}}{\pi} \frac{1}{6} \sum_{f=u,d} N_f^c Q_f \left(t_f^3 - 2 \sin^2 \hat{\theta}_W Q_f \right). \quad (3.113)$$

This can easily be rewritten as

$$\frac{d \sin^2 \hat{\theta}_W}{d \ln \mu^2} = -\frac{\hat{\alpha}(\mu)}{\pi} \frac{1}{(2)24} \sum_i N_i^c \gamma_i Q_i \hat{v}_i(\mu), \quad (3.114)$$

where $\hat{v}_i \equiv T_i - Q_i \sin^2 \hat{\theta}$ is the vector coupling of the Z boson with the fermions, and γ_i is a numerical factor with value 4.

There is another way to get the running of the weak mixing angle in a more heuristic but intuitive way. In the $\overline{\text{MS}}$ Lagrangian, the Z boson and photon couple to fermions through the interaction given in Eq. (3.22). Let us now take the measurement of a physical quantity, like for example the Z-boson production asymmetry, defined as

$$A_e \equiv \frac{\sigma(e_L^- e_L^+ \rightarrow Z) - \sigma(e_R^- e_R^+ \rightarrow Z)}{\sigma(e_L^- e_L^+ \rightarrow Z) + \sigma(e_R^- e_R^+ \rightarrow Z)}. \quad (3.115)$$

The corresponding part of the Lagrangian given in Eq. (3.22) that contributes to this asymmetry is

$$\mathcal{L}_{\gamma Z} = -\frac{\hat{e}}{\hat{s}\hat{c}} Z_\mu \left[\left(\frac{1}{2} - \hat{s}^2 \right) \bar{e}_L \gamma^\mu e_L - \hat{s}^2 \bar{e}_R \gamma^\mu e_R \right] - \hat{e} A_\mu [\bar{e}_L \gamma^\mu e_L + \bar{e}_R \gamma^\mu e_R]. \quad (3.116)$$

At tree level, the asymmetry can then be written as

$$A_e = \frac{\left(\frac{1}{2} - s^2\right)^2 - s^4}{\left(\frac{1}{2} + s^2\right)^2 + s^4}, \quad (3.117)$$

where I have used s^2 instead of \hat{s}^2 because this is a tree level relation. At one loop level the largest contribution to this asymmetry comes from corrections affecting the gauge boson propagators¹⁰. In this case, the contributions come from corrections to the Z boson propagator and the γZ mixing bubble. The contribution of the Z boson propagator can be taken into account if the mass in the propagator is taken to be the pole Z boson mass. On the other hand, the contribution of the γZ mixing can be taken into account with an effective Z-boson coupling, and as a consequence an effective weak mixing angle defined as

$$s_{eff}^2 = \hat{s}^2 - \hat{s}\hat{c} \frac{\bar{\Pi}_{\gamma Z}(M_Z^2, \mu^2)}{M_Z^2}, \quad (3.118)$$

using this definition, the asymmetry has the same form as Eq. (3.117) but with the replacement $s \rightarrow s_{eff}$. Since A_e is a physical quantity, it should not depend on the renormalization scale. Furthermore, the asymmetry only depends on s_{eff}^2 , so s_{eff}^2 should not depend on the renormalization scale neither. This implies that the μ dependence of the weak mixing angle in the $\overline{\text{MS}}$ scheme should cancel the μ dependence in the γZ vacuum polarization function. This is how we can deduce that

$$\frac{d \sin^2 \hat{\theta}}{d \ln \mu^2} = \sum_f N_f^c \left(t_f^3 Q_f \frac{1}{2} - \hat{s}^2 Q_f^2 \right) \frac{1}{M_Z^2} \frac{d}{d \ln \mu^2} \Pi_{VV}(M_Z^2, m_f, \mu^2), \quad (3.119)$$

which can easily be written as Eq. (3.114). As it is pointed out at the end of Ref. [106], for

¹⁰Sometimes these are called oblique corrections. For more details at an introductory level see chapter 31 of [76].

processes at low energy we would have an equation of the form (similarly to Eq. (3.118))

$$\sin^2 \theta(q^2)_{eff} = \hat{\kappa}(q^2, \mu^2 = M_Z^2) \sin^2 \hat{\theta}(\mu = M_Z), \quad (3.120)$$

where the q^2 dependence comes from the γZ vacuum polarization function. To properly sum all the logs in this equation it is more convenient to set $\mu^2 = q^2$ in this equation, and move the large logs in $\hat{\kappa}$ from the bubble to the running of the weak mixing angle. To do so, we need to solve the RGE for $\sin^2 \hat{\theta}(\mu)$. This is done in Ref. [106] and improved in the work of this thesis. That is the topic of the next chapter.

Chapter Four

The Weak Mixing Angle at Low Energies

This chapter is a slightly modified version of the paper [107]. Due structure of the Standard Model, parity violation experiments can be usually mapped into a measurement of the weak mixing angle, as we have shown with one example in the previous chapter. There are several experiments that measure parity violation at very low energies. Due to their usually clean environments and different range of energies, these experiments are important and complementary to the high energy experiments. For example, the Qweak experiment [108] at Jefferson Laboratory (JLab) has measured the weak charge of the proton, $Q_W(p) \sim 1 - 4 \sin^2 \theta_W$, in polarized electron scattering from a fixed liquid hydrogen target at $Q^2 \approx 0.026 \text{ GeV}^2$. The same observable, but at an even lower $Q^2 \approx 0.0045 \text{ GeV}^2$, will also be targeted by the P2 experiment [109] at the MESA facility which is currently under construction at the University of Mainz in Germany. In a very similar setup, the MOLLER Collaboration [110] at JLab will build and improve on the completed E158 experiment [111] at SLAC (that occurred at almost the same Q^2 as Qweak) and measure the analogous weak charge of the electron, $Q_W(e)$, in polarized Møller scattering at $Q^2 \approx 0.0056 \text{ GeV}^2$. The PVDIS Collaboration [112] at the 6 GeV CEBAF complex at JLab scattered polarized electrons deep-inelastically from deuterium, and the SoLID Collaboration [113] will increase the PVDIS precision in the future by benefiting from the energy upgraded CEBAF and a correspondingly higher and broader Q^2 range. Other approaches include neutrino and anti-neutrino deep inelastic scat-

tering [114], $\bar{\nu}$ - e scattering near nuclear reactors [115], and parity violation in atoms [116] and ions [117]. For more details, see the recent reviews on low energy measurements of the weak mixing angle [26], on the weak neutral current [118], and on weak polarized electron scattering [119].

When physical quantities are computed at low momentum transfers, there will be large logarithms of the form $\ln(\frac{M_Z^2}{m_i^2})$ where m_i is the mass of a fermion coming from the vacuum polarization function. The logs will appear explicitly in $\hat{\kappa}(q^2, M_Z^2)$. These large logs have an important contribution to physical quantities like parity violating asymmetries at low energy. For example, it has been shown in Ref. [120, 121] that the one loop contributions from the γZ mixing reduce the Møller asymmetry by 40% with respect to its tree level value. This is why it is important to properly re-sum this logarithms. The re-summation can be done by setting $\mu^2 \approx q^2$ instead of $\mu^2 \approx M_Z^2$ in $\hat{\kappa}$, this, in turn, will move the logs to the weak mixing angle in the $\overline{\text{MS}}$ scheme, which is solved through the RGE and in consequence will re-sum them automatically.

As it is natural from this discussion and the results of the previous chapter, the weak mixing angle running is controlled by the vector part of the Z coupling to fermions. This is because the mixing bubble only gets contribution from this vector coupling. Ref. [106] exploited this idea and used it to compute the running of $\sin^2 \hat{\theta}$. In the work presented here, we follow the same approach, updating and improving the uncertainty of the weak mixing angle at low energies.

Since QCD at low energies does not allow for reliable perturbative calculations, the theoretical uncertainty of the RGE running from the Z-pole to low energies arises dominantly from the hadronic region. A phenomenological approach to address this region was developed in Ref. [106]. Working in the $\overline{\text{MS}}$ scheme¹, the main idea was to relate the case of the weak mixing angle to that of the electromagnetic coupling, $\hat{\alpha}$, as far as possible, and then to consider both maximal and minimal $SU(3)$ flavor symmetry breaking to constrain the flavor

¹Quantities defined in the $\overline{\text{MS}}$ scheme will be denoted by a caret.

separation of the three light quarks (u, d, s). In the present work, we extend the analysis to the next order in the strong coupling constant, $\hat{\alpha}_s$, and introduce a number of new elements. We employ the most recent values and uncertainties of the input parameters, such as $\hat{\alpha}_s$ and the heavy quark masses. The hadronic vacuum polarization contribution to the RGE running of $\hat{\alpha}$ is obtained dispersively from e^+e^- annihilation data for hadronic final states, which are supplemented by isospin rotated τ decay spectral functions corrected for isospin breaking effects [122, 123]. We tie experimental data [122] and lattice gauge theory calculations [124, 125] together to obtain the individual contributions of strange and first generation quarks. This flavor separation at the quark level to high accuracy is consistent with and almost an order of magnitude more precise than previous calculations [106, 121]. It is also necessary to constrain OZI-rule [126–128] violating effects, for which we utilize the recent lattice gauge theory calculation of disconnected contributions to the anomalous magnetic moment of the muon [129]. These refinements allow for significant reduction of the theoretical uncertainty of the RGE evolution. As a by-product, our method sheds light on the dual description of quarks and hadrons in the non-perturbative regime and may open new ways to extract the strange quark mass from the electro-production of hadrons.

4.1 Renormalization group evolution

In an approximation in which all fermions are either massless and active or infinitely heavy and decoupled, the RGE for the electromagnetic coupling in the $\overline{\text{MS}}$ scheme [96], $\hat{\alpha}$, can be written in the form [106],

$$\mu^2 \frac{d\hat{\alpha}}{d\mu^2} = \frac{\hat{\alpha}^2}{\pi} \left[\frac{1}{24} \sum_i K_i \gamma_i Q_i^2 + \sigma \left(\sum_q Q_q \right)^2 \right], \quad (4.1)$$

where the sum is over all active particles in the relevant energy range. This equation is the generalization of Eq. (3.75) to higher order in QCD. Just as before, the Q_i are the electric

boson	γ_i	fermion	γ_i
real scalar	1	chiral fermion	4
complex scalar	2	Majorana fermion	4
massless gauge boson	-22	Dirac fermion	8

Table 4.1: RGE contributions of different particle types, where the minus sign is indicative for the asymptotic freedom in non-Abelian gauge theories.

charges, while the γ_i are constants depending on the field type and shown in Table 4.1. The K_i and σ contain higher-order corrections and are given by [91],

$$\begin{aligned}
 K_i = & N_i^c \left\{ 1 + \frac{3}{4} Q_i^2 \frac{\hat{\alpha}}{\pi} + \frac{\hat{\alpha}_s}{\pi} + \frac{\hat{\alpha}_s^2}{\pi^2} \left[\frac{125}{48} - \frac{11}{72} n_q \right] \right. \\
 & + \frac{\hat{\alpha}_s^3}{\pi^3} \left[\frac{10487}{1728} + \frac{55}{18} \zeta_3 - n_q \left(\frac{707}{864} + \frac{55}{54} \zeta_3 \right) - \frac{77}{3888} n_q^2 \right] \\
 & + \frac{\hat{\alpha}_s^4}{4\pi^4} \left[\frac{2665349}{41472} + \frac{182335}{864} \zeta_3 - \frac{605}{16} \zeta_4 - \frac{31375}{288} \zeta_5 \right. \\
 & - n_q \left(\frac{11785}{648} + \frac{58625}{864} \zeta_3 - \frac{715}{48} \zeta_4 - \frac{13325}{432} \zeta_5 \right) \\
 & \left. - n_q^2 \left(\frac{4729}{31104} - \frac{3163}{1296} \zeta_3 + \frac{55}{72} \zeta_4 \right) + n_q^3 \left(\frac{107}{15552} + \frac{1}{108} \zeta_3 \right) \right\}, \quad (4.2)
 \end{aligned}$$

and,

$$\begin{aligned}
 \sigma = & \frac{\hat{\alpha}_s^3}{\pi^3} \left[\frac{55}{216} - \frac{5}{9} \zeta_3 \right] + \frac{\hat{\alpha}_s^4}{\pi^4} \left[\frac{11065}{3456} - \frac{34775}{3456} \zeta_3 + \frac{55}{32} \zeta_4 + \frac{3875}{864} \zeta_5 \right. \\
 & \left. - n_q \left(\frac{275}{1728} - \frac{205}{576} \zeta_3 + \frac{5}{48} \zeta_4 + \frac{25}{144} \zeta_5 \right) \right], \quad (4.3)
 \end{aligned}$$

with n_q the number of active quarks and $N_i^c = 3$ the color factor for quarks. For leptons one substitutes $N_i^c = 1$ and $\hat{\alpha}_s = 0$, while $K_i = 1$ for bosons.

We can relate the RGE of $\hat{\alpha}$ to that of $\sin^2 \hat{\theta}_W$ since both, the γZ mixing tensor $\bar{\Pi}_{\gamma Z}$ and the photon vacuum polarization function $\bar{\Pi}_{\gamma\gamma}$ are pure vector-current correlators. Including higher order corrections, the RGE for the Z boson vector coupling to fermion f , $\hat{v}_f =$

Energy range	λ_1	λ_2	λ_3	λ_4
$\bar{m}_t \leq \mu$	$\frac{9}{20}$	$\frac{289}{80}$	$\frac{14}{55}$	$\frac{9}{20}$
$M_W \leq \mu < \bar{m}_t$	$\frac{21}{44}$	$\frac{625}{176}$	$\frac{6}{11}$	$\frac{3}{22}$
$\bar{m}_b \leq \mu < M_W$	$\frac{21}{44}$	$\frac{15}{22}$	$\frac{51}{440}$	$\frac{3}{22}$
$m_\tau \leq \mu < \bar{m}_b$	$\frac{9}{20}$	$\frac{3}{5}$	$\frac{2}{19}$	$\frac{1}{5}$
$\bar{m}_c \leq \mu < m_\tau$	$\frac{9}{20}$	$\frac{2}{5}$	$\frac{7}{80}$	$\frac{1}{5}$
$\bar{m}_s \leq \mu < \bar{m}_c$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{5}{36}$	0
$\bar{m}_d \leq \mu < \bar{m}_s$	$\frac{9}{20}$	$\frac{2}{5}$	$\frac{13}{110}$	$\frac{1}{20}$
$\bar{m}_u \leq \mu < \bar{m}_d$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{40}$	0
$m_\mu \leq \mu < \bar{m}_u$	$\frac{1}{4}$	0	0	0
$m_e \leq \mu < m_\mu$	$\frac{1}{4}$	0	0	0

Table 4.2: Coefficients entering the higher order RGE for the weak mixing angle.

$T_f - 2Q_f \sin^2 \hat{\theta}_W$, where T_f is the third component of weak isospin of fermion f , is then

$$\mu^2 \frac{d\hat{v}_f}{d\mu^2} = \frac{\hat{\alpha} Q_f}{24\pi} \left[\sum_i K_i \gamma_i \hat{v}_i Q_i + 12\sigma \left(\sum_q Q_q \right) \left(\sum_q \hat{v}_q \right) \right]. \quad (4.4)$$

Eqs. (4.1) and (4.4) can be used [106] to obtain

$$\begin{aligned} \hat{s}^2(\mu) &= \hat{s}^2(\mu_0) \frac{\hat{\alpha}(\mu)}{\hat{\alpha}(\mu_0)} + \lambda_1 \left[1 - \frac{\hat{\alpha}(\mu)}{\hat{\alpha}(\mu_0)} \right] + \\ &\frac{\hat{\alpha}(\mu)}{\pi} \left[\frac{\lambda_2}{3} \ln \frac{\mu^2}{\mu_0^2} + \frac{3\lambda_3}{4} \ln \frac{\hat{\alpha}(\mu)}{\hat{\alpha}(\mu_0)} + \tilde{\sigma}(\mu_0) - \tilde{\sigma}(\mu) \right], \end{aligned} \quad (4.5)$$

where the λ_i are known [106] constants² given in Table 4.2 and the explicit K_i dependence has disappeared. The $\tilde{\sigma}$ terms,

$$\tilde{\sigma}(\mu) = \frac{\lambda_4}{33 - 2n_q} \frac{5}{36} \left[(11 - 24\zeta_3) \frac{\hat{\alpha}_s^2(\mu)}{\pi^2} + b \frac{\hat{\alpha}_s^3(\mu)}{\pi^3} \right], \quad (4.6)$$

²See Appendix B.1 for mote details.

with,

$$\begin{aligned}
 b \equiv & \frac{2213}{24} - \frac{6955}{24}\zeta_3 + \frac{99}{2}\zeta_4 + \frac{775}{6}\zeta_5 - n_q \left(\frac{55}{12} - \frac{41}{4}\zeta_3 + 3\zeta_4 + 5\zeta_5 \right) \\
 & - \frac{(153 - 19n_q)(11 - 24\zeta_3)}{99 - 6n_q}, \tag{4.7}
 \end{aligned}$$

represent the singlet contributions to the RGE evolution of the weak mixing angle at four and five loop order. These terms arise from quark-antiquark annihilation (disconnected) diagrams (see Figure 4.1) and are suppressed in perturbative QCD (PQCD). In the non-perturbative domain these give rise to so-called OZI-rule [126–128] violations.

Eq. (4.5) together with the solution of the four-loop QCD β -function [130, 131] represents a complete solution, as long as all matching scales μ at which an active particle decouples are known, because there the λ_i change their values. The matching scales of all bosons [132], charged leptons, and heavy (t , b , and c) quarks [92, 94, 95] can be calculated as what we call threshold masses \bar{m}_q , where the QCD corrections to the matching relations vanish by definition.

4.1.1 Matching conditions

At each particle threshold the RGE coefficients need to be modified to reflect the particle content of the associated effective field theory (EFT), and in the $\overline{\text{MS}}$ scheme it is also convenient to change the definitions of $\hat{\alpha}$ and \hat{s} to correspond to this same EFT. This is analogous to the usual treatment of $\hat{\alpha}_s$ and leads to very small matching discontinuities in the RGE running of the couplings.

Denoting the electromagnetic coupling with and without the fermion near the threshold by $\hat{\alpha}(m_f)^+$ and $\hat{\alpha}(m_f)^-$, respectively³, the matching condition for $\hat{\alpha}$ which was introduced

³We assume m_f is an $\overline{\text{MS}}$ mass with respect to QCD, but a pole mass for both leptons and quarks with respect to QED.

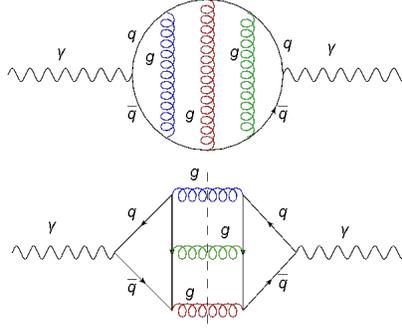


Figure 4.1: Examples of a connected (top) and a disconnected (bottom) Feynman diagram.

in the last chapter⁴, reads [92, 94, 95],

$$\begin{aligned}
 \frac{\pi}{\hat{\alpha}(m_f)^+} &= \frac{\pi}{\hat{\alpha}(m_f)^-} - \frac{15}{16} N_f^c \frac{\hat{\alpha}(m_f)}{\pi} Q_f^4 \\
 &- \frac{N_f^c - 1}{2} \left[\frac{13}{12} \frac{\hat{\alpha}_s^+}{\pi} + \left(\frac{655}{144} \zeta_3 - \frac{3847}{864} + \frac{361}{1296} n_q \right) \frac{\hat{\alpha}_s^{+2}}{\pi^2} \right. \\
 &+ \left. \left(-0.55739 - 0.92807 n_q + 0.01928 n_q^2 \right) \frac{\hat{\alpha}_s^{+3}}{\pi^3} \right] Q_f^2 \\
 &- \frac{N_f^c - 1}{2} \left[\frac{295}{1296} \frac{\hat{\alpha}_s^{+2}}{\pi^2} + (\mathcal{K}_1 + \mathcal{K}_2 n_q) \frac{\hat{\alpha}_s^{+3}}{\pi^3} \right] \sum_{\ell} Q_{\ell}^2 . \quad (4.8)
 \end{aligned}$$

The first three lines derive from heavy quark vector-current correlators. The last line involves a sum over all quarks ℓ with $m_{\ell} \ll m_q$, and arises from the decoupling of the heavy quark q propagating in inner loops of multi-bubble type diagrams in which the outer loop (the one coupled to the currents) is occupied by a light quark ℓ . The corresponding contribution at order $\hat{\alpha}_s^3$ is parametrized by the coefficients \mathcal{K}_i and is unknown at present. The known $\hat{\alpha}_s^2$ term for the charm and bottom quarks, and the $\hat{\alpha}_s^3$ terms from the charm and bottom quark vector-current correlators amount to about 9×10^{-6} and -9×10^{-6} , respectively. Taking these as conservative bounds on the unknown higher-order terms and combining them in quadrature results in an estimated truncation error of $\pm 1.3 \times 10^{-5}$ in $\hat{\alpha}$.

⁴See sections 3.4.3 and 3.4.5.

The matching conditions of \hat{s}^2 and $\hat{\alpha}$ can also be related [106],

$$\sin^2 \hat{\theta}_W(\hat{m}_f)^- = \frac{\hat{\alpha}(\hat{m}_f)^-}{\hat{\alpha}(\hat{m}_f)^+} \sin^2 \hat{\theta}_W(\hat{m}_f)^+ + \frac{Q_i T_i}{2Q_i^2} \left[1 - \frac{\hat{\alpha}(\hat{m}_f)^-}{\hat{\alpha}(\hat{m}_f)^+} \right]. \quad (4.9)$$

Applying the numerical analysis of the previous paragraph to Eq. (4.9), we find 2.4×10^{-6} and -1.4×10^{-6} , respectively, and we estimate a truncation error related to the matching of about $\pm 3 \times 10^{-6}$ in \hat{s}^2 .

For completeness we recall that integrating out the W^\pm bosons induces the one-loop matching condition [106, 132],

$$\frac{1}{\hat{\alpha}^+} = \frac{1}{\hat{\alpha}^-} + \frac{1}{6\pi}. \quad (4.10)$$

For \hat{s}^2 this implies

$$\sin^2 \hat{\theta}_W(M_W)^+ = 1 - \frac{\hat{\alpha}(M_W)^+}{\hat{\alpha}(M_W)^-} \cos^2 \hat{\theta}_W(M_W)^-. \quad (4.11)$$

4.2 Implementation of experimental input

The perturbative treatment of the previous section cannot be applied at hadronic energy scales and experimental input is required. This is usually taken from $R(s)$, i.e., the cross section $\sigma(e^+e^- \rightarrow \text{hadrons})$ normalized to $\sigma(e^+e^- \rightarrow \mu^+\mu^-)$. Additional information on $R(s)$ is encoded in hadronic τ decay spectral functions [133]. The traditional method to implement the $R(s)$ measurements is through a subtracted dispersion integral,

$$\Delta\alpha_{\text{had}}^{(5)}(M_Z^2) = \frac{\alpha}{3\pi} \int_{4m_\pi^2}^{\infty} ds \frac{R(s)M_Z^2}{s(M_Z^2 - s) - i\epsilon}, \quad (4.12)$$

which gives the hadronic contribution (with the top quark removed) to the Z scale value of the electromagnetic coupling in the on-shell scheme. One supplements the input data with the theoretical (perturbative) prediction for $R(s)$ at $s \geq s_0$, with s_0 large enough to be able

to trust QCD perturbation theory. A variant [134] of this approach evaluates Eq. (4.12) in the space-like region, $\Delta\alpha_{\text{had.}}^{(5)}(-M_Z^2)$, and obtains $\Delta\alpha_{\text{had.}}^{(5)}(M_Z^2)$ in a second step. More details about how different groups get the running of alpha are given in Appendix B.1.

In the $\overline{\text{MS}}$ scheme it is more natural to use the unsubtracted dispersion relation [96],

$$\Delta\hat{\alpha}^{(3)}(\mu_0) = \frac{\alpha}{3\pi} \int_{4m_\pi^2}^{\mu_0^2} ds \frac{R(s)}{s - i\epsilon} + 4\pi I^{(3)}, \quad (4.13)$$

this equation is the same as Eq. (3.86), where, as we have said before, the superscript indicates that we focus here on the currents produced by the three light quarks (bosons, leptons, charm and bottom quarks are included following Sec. 4.1). The upper integration limit can in principle be chosen as an arbitrary perturbative scale μ_0 , but in practice we take μ_0^2 to coincide with the cut-off value s_0 used in the traditional method, since this allows us to recycle results obtained there. Indeed [96],

$$\frac{\alpha}{3\pi} \int_{4m_\pi^2}^{\mu_0^2} ds \left[\frac{R(s)}{s - i\epsilon} - \frac{R(s)M_Z^2}{s(M_Z^2 - s) - i\epsilon} \right] < 10^{-6}, \quad (4.14)$$

for $\mu_0 \approx 2$ GeV. Using the results of Ref. [122] including inputs from τ decays which we correct for γ - ρ mixing [123], we obtain,

$$\frac{\alpha}{3\pi} \int_{4m_\pi^2}^{4 \text{ GeV}^2} ds \frac{R(s)M_Z^2}{s(M_Z^2 - s)} = (58.71 \pm 0.45) \times 10^{-4}. \quad (4.15)$$

We compute the second term in Eq. (4.13) at the scale $\mu = 2$ GeV perturbatively, extending the $\mathcal{O}(\hat{\alpha}_s^2)$ result of Ref. [96] to $\mathcal{O}(\hat{\alpha}_s^3)$,

$$\begin{aligned} 4\pi I^{(3)} &= 2\alpha \int_0^{2\pi} d\theta \hat{\Pi}^{(3)}(\mu^2 e^{i\theta}) \\ &= \frac{2\alpha}{3\pi} \left[\frac{5}{3} + \left(\frac{55}{12} - 4\zeta(3) + 2\frac{\hat{m}_s^2}{\mu^2} \right) \left(\frac{\hat{\alpha}_s}{\pi} + \frac{\hat{\alpha}}{4\pi} \right) \right] \end{aligned}$$

$$\begin{aligned}
 & + \left(\frac{34525}{864} - \frac{9}{4}\zeta(2) - \frac{715}{18}\zeta(3) + \frac{25}{3}\zeta(5) + \frac{125}{12} \frac{\hat{m}_s^2}{\mu^2} + F_2(\hat{m}_c, \hat{m}_b) \right) \frac{\hat{\alpha}_s^2}{\pi^2} \\
 & + \left(\frac{7012579}{13824} - \frac{961}{16}\zeta(2) - \frac{76681}{144}\zeta(3) + \frac{12515}{288}\zeta(5) \right. \\
 & \quad \left. - \frac{665}{36}\zeta(7) + \frac{81}{2}\zeta(2)\zeta(3) + \frac{155}{2}\zeta(3)^2 + F_3(\hat{m}_c, \hat{m}_b) \right) \frac{\hat{\alpha}_s^3}{\pi^3} \\
 & = (24.85 \pm 0.18 - 43 \Delta\hat{\alpha}_s) \times 10^{-4}, \tag{4.16}
 \end{aligned}$$

where the $F_i(\hat{m}_c, \hat{m}_b)$ are correction terms from the charm and bottom quarks, with values $F_2(\hat{m}_c, \hat{m}_b) \simeq -0.2348$ and $F_3(\hat{m}_c, \hat{m}_b) \simeq -0.390$. The numerical evaluation in the last line of Eq. (4.16) is for $\hat{\alpha}_s(M_Z) = 0.1182$, $\hat{\alpha}_s(2 \text{ GeV}) = 0.303$ and $\hat{m}_s(2 \text{ GeV}) = 98 \pm 6 \text{ MeV}$ [135]. The uncertainty is the size of the $\mathcal{O}(\hat{\alpha}_s^3)$ term, and we have defined

$$\Delta\hat{\alpha}_s \equiv \hat{\alpha}_s(M_Z) - 0.1182, \tag{4.17}$$

to display the dependence on $\hat{\alpha}_s$. Thus, from Eqs. (4.13)–(4.16) we obtain,

$$\Delta\hat{\alpha}^{(3)}(2 \text{ GeV}) = (83.56 \pm 0.45 \pm 0.18) \times 10^{-4}. \tag{4.18}$$

4.3 Singlet contribution

We recall that Eq. (4.6) exhibits an explicit dependence on α_s , which in the non-perturbative domain gives rise to the QCD induced OZI-rule [126–128] violations. These have to be known independently, since they affect $\hat{\alpha}$ and \hat{s}^2 differently. Thus, in addition to a quark flavor separation, one also needs a singlet piece separation, even though the singlet piece is expected to be small. To do so, we first relate $\Delta_{\text{disc}}\hat{\alpha}$, the disconnected part in $\Delta\hat{\alpha}^{(3)}(2 \text{ GeV})$, to the one entering the low energy weak mixing angle, $\Delta_{\text{disc}}\hat{s}^2$. Non-singlet and singlet contributions are separately gauge-invariant, and to gain information on $\Delta_{\text{disc}}\hat{\alpha}$, we will

adopt a lattice QCD calculation [129] of the disconnected quark line contributions to the anomalous magnetic moment of the muon, a_μ .

By construction, the $\tilde{\sigma}$ terms in Eq. (4.5) are related to the σ terms in Eq. (4.1),

$$\mu^2 \frac{d\tilde{\sigma}}{d\mu^2} = -\lambda_4 \sigma. \quad (4.19)$$

On the other hand, isolating the $\Delta_{\text{disc}} \hat{\alpha}$ term in Eq. (4.1) we obtain (working here in lowest order in α),

$$\mu^2 \frac{d\Delta_{\text{disc}} \hat{\alpha}}{d\mu^2} = \frac{\alpha}{\pi} \left(\sum_q Q_q \right)^2 \sigma, \quad (4.20)$$

so that,

$$\frac{d\tilde{\sigma}}{d\mu^2} = -\frac{\pi}{\alpha} \lambda_4 \left(\sum_q Q_q \right)^{-2} \frac{d\Delta_{\text{disc}} \hat{\alpha}}{d\mu^2} = -\lambda_1 \frac{\pi}{\alpha} \frac{d\Delta_{\text{disc}} \hat{\alpha}}{d\mu^2}, \quad (4.21)$$

where the last step applies for $\mu < \bar{m}_c$ (we are assuming approximate isospin symmetry which eliminates the interval $\bar{m}_u < \mu < \bar{m}_d$). Then,

$$\tilde{\sigma}(\mu) - \tilde{\sigma}(\mu_0) = -\lambda_1 \frac{\pi}{\alpha} [\Delta_{\text{disc}} \hat{\alpha}(\mu) - \Delta_{\text{disc}} \hat{\alpha}(\mu_0)]. \quad (4.22)$$

These relations are general, but there is a subtle point. In general, the singlet pieces effectively decouple at renormalization scales \bar{m}_q^{disc} that may differ from the scales \bar{m}_q at which the non-singlet pieces decouple. This would generate various energy intervals with generally different values for λ_1 . Implementing strong isospin symmetry in the form $\bar{m}_u = \bar{m}_d$ and $\bar{m}_u^{\text{disc}} = \bar{m}_d^{\text{disc}}$, as well as accepting the physical mass orderings $\bar{m}_s \geq \bar{m}_u$ and $\bar{m}_s^{\text{disc}} \geq \bar{m}_u^{\text{disc}}$, there remain a total of six different orderings.

As an example, consider the case,

$$\bar{m}_s^{\text{disc}} > \bar{m}_s > \bar{m}_u > \bar{m}_u^{\text{disc}}. \quad (4.23)$$

For scales $\mu > \bar{m}_s^{\text{disc}}$ there are three active quarks with $Q_u + Q_d + Q_s = 0$ and the singlet

contributions vanish. For scales in the range $\bar{m}_s^{\text{disc}} > \mu > \bar{m}_s$ we obtain the value $\lambda_1 = 1/2$. Similarly, for $\bar{m}_s > \mu > \bar{m}_u$ and for $\bar{m}_u > \bar{m}_u^{\text{disc}}$ we find $\lambda_1 = 9/20$ and $1/4$, respectively. Below \bar{m}_u^{disc} all singlet contributions vanish by definition. Inserting these results into Eq. (4.22) and summing the contributions from all intervals, we find the constraint,

$$-\frac{\Delta_{\text{disc}}\hat{\alpha}}{4} < \frac{\alpha}{\pi} \left[\tilde{\sigma}(\bar{m}_s^{\text{disc}}) - \tilde{\sigma}(\bar{m}_u^{\text{disc}}) \right] < -\frac{\Delta_{\text{disc}}\hat{\alpha}}{2}, \quad (4.24)$$

where we have anticipated that $\Delta_{\text{disc}}\hat{\alpha} < 0$ (see below).

The other five cases are dealt with in the same way, and one can check that the inequality (4.24) is never violated. For the three mass orderings satisfying $\bar{m}_u^{\text{disc}} \geq \bar{m}_u$, or generally if we can neglect the presumably small range $\bar{m}_u > \mu > \hat{m}_u^{\text{disc}}$, we find the much stronger constraint,

$$-\frac{9\Delta_{\text{disc}}\hat{\alpha}}{20} < \frac{\alpha}{\pi} \left[\tilde{\sigma}(\bar{m}_s^{\text{disc}}) - \tilde{\sigma}(\bar{m}_u^{\text{disc}}) \right] < -\frac{\Delta_{\text{disc}}\hat{\alpha}}{2}. \quad (4.25)$$

Since we do not expect the \bar{m}_q^{disc} to be numerically very different from the \bar{m}_q we choose our central value to correspond to $\bar{m}_q^{\text{disc}} = \bar{m}_q$, and we include twice the range in Eq. (4.25) as the uncertainty due to possible $\bar{m}_q^{\text{disc}} \neq \bar{m}_q$ effects. Thus,

$$\frac{\alpha}{\pi} \left[\tilde{\sigma}(\bar{m}_s^{\text{disc}}) - \tilde{\sigma}(\bar{m}_u^{\text{disc}}) \right] = - \left[\frac{9}{20} \pm \frac{1}{20} \right] \Delta_{\text{disc}}\hat{\alpha}, \quad (4.26)$$

which can be inserted into Eq. (4.5). Notice, however, that Eq. (4.5) also contains an implicit singlet contribution from each of the two terms in the first line. Taken together, the λ_1 term cancels exactly the central value in Eq. (4.26) and we finally arrive at

$$\Delta_{\text{disc}}\hat{s}^2 = \left[\hat{s}^2 \pm \frac{1}{20} \right] \Delta_{\text{disc}}\hat{\alpha}. \quad (4.27)$$

In Appendix B.2 we compute $\Delta_{\text{disc}}\alpha$ in the on-shell scheme by exploiting the lattice gauge

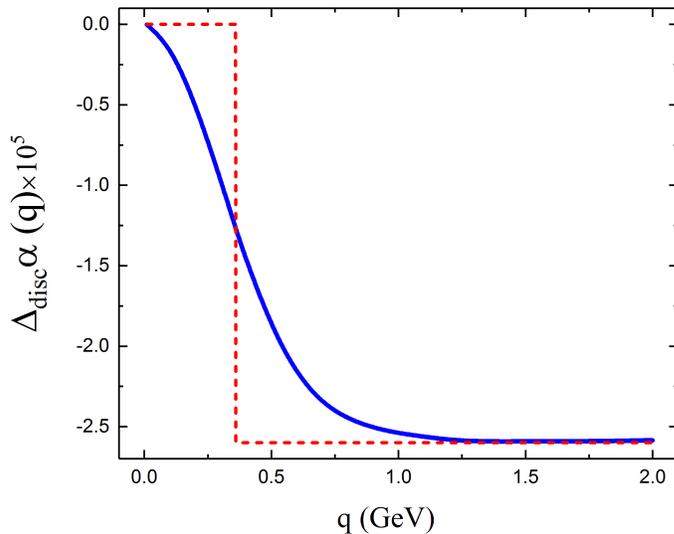


Figure 4.2: Scale dependence of the singlet contribution to $\Delta\alpha$ (solid line) and its step function approximation (dashed line).

theory calculation [129] of the corresponding contribution to a_μ with the result,

$$\Delta_{\text{disc}}\alpha(2.0 \text{ GeV}) = -2.6 \times 10^{-5}. \quad (4.28)$$

Note that because the sum of the charges of the three light quarks vanishes, and we enter the perturbative domain where the singlet piece is known to be tiny, we expect an asymptotically stable value at higher energies for $\Delta_{\text{disc}}\alpha(q)$. This is supported by Figure 4.2, showing that $\Delta_{\text{disc}}\alpha(q)$ is nearly q -independent for $q \sim 1.2 \text{ GeV}$. We also remark that the dominance of low scales notwithstanding, the sign in Eq. (4.28) coincides with that of the singlet piece in the perturbative regime. Also shown in Figure 4.2 is the step function approximation of $\Delta_{\text{disc}}\alpha(q)$, with the step defined as the value of q where it reaches half of its asymptotic value in Eq. (4.28). We interpret this as the value where the strange quark decouples from singlet diagrams, so that $\bar{m}_s^{\text{disc}} \sim 350 \text{ MeV}$. Our central value of \bar{m}_s to be derived in the next section, $\bar{m}_s = 342 \text{ MeV}$, is numerically very close to this providing evidence for $\bar{m}_s^{\text{disc}} \approx \bar{m}_s$.

Eq. (4.27) and Eq. (4.28) refer to quantities in the $\overline{\text{MS}}$ and on-shell schemes, respectively,

channel	$a_\mu \times 10^{10}$	$\Delta\alpha \times 10^4$
ϕ	38.43	5.13
$K\bar{K}\pi$	2.45	0.78
$\eta\phi$	0.36	0.13
PQCD [136] ($> 1.8 GeV$)	7.30	—
Total	48.54	6.04
$K\bar{K}$ (<i>non</i> - ϕ)	3.62	0.76
$K\bar{K}2\pi$	0.85	0.30
$K\bar{K}3\pi$	-0.03	-0.01
$K\bar{K}\eta$	0.01	0.00
$K\bar{K}\omega$	0.01	0.00
Total	4.46	1.05

Table 4.3: Channels associated with the strange quark external current (top) and possible further channels originating from it (bottom).

and in general these may differ. However, since we are working here in the three quark theory and the sum of the charges of three light quarks vanishes, the change of schemes is trivial. We can therefore use Eq. (4.28) in Eq. (4.27) and obtain,

$$\Delta_{\text{disc}}\hat{s}^2 = (-0.6 \pm 0.3) \times 10^{-5}, \quad (4.29)$$

where the uncertainty combines the errors from Eq. (4.27) and the one induced by the lattice calculation [129].

4.4 Flavor separation

In this section we perform a flavor separation of the contributions of up-type from down-type quarks, or — given that up and down quarks are linked by the approximate strong isospin symmetry — a separation of s from u and d quarks. Our strategy consists of first using exclusively the experimental electro-production data as tabulated in Ref. [122] to constrain the contribution $\Delta_s\alpha$ of the strange quark to $\Delta\alpha$. We then exploit the lattice gauge theory

results in Refs. [124, 125] to confirm and refine the purely data driven analysis. Then we introduce the *threshold mass* \bar{m}_q of a quark q as the value of the 't Hooft scale where the QCD contribution to the corresponding decoupling relation becomes trivial. \bar{m}_c and \bar{m}_b are treated in perturbation theory, while for u , d , and s quarks we derive bounds using phenomenological and theoretical constraints.

4.4.1 Experimental data

To obtain $\Delta_s\alpha$ we use Ref. [122] where the contribution of each hadronic channel to a_μ and $\Delta\alpha$ for energies up to 1.8 GeV is given. The main idea is to determine for each channel whether it was produced by an $\bar{s}s$ or a first generation quark current. For reasons that will become clear later, we consider both, $\Delta_s\alpha$, and the strange quark contribution to the anomalous magnetic moment, a_μ^s .

We begin by listing in the upper part of Table 4.3 the experimental channels [122] which we associate with an $\bar{s}s$ current. Up to OZI-rule violating ϕ - ω and ϕ - ρ mixing effects, the ϕ meson can be identified with strange quarks. We calculate its contribution using a Breit-Wigner shape with s -dependent total and partial widths, adopting the PDG values [135] for the ϕ meson branching ratios and applying a small correction for ϕ - ω mixing. As for the $\phi(1680)$, the main decay channel is $K\bar{K}^*$ with K^* mesons decaying almost entirely into $K\pi$. As can be seen from data [122], the $K\bar{K}\pi$ channel is indeed virtually saturated by $\phi(1680)$ decays. The η - ϕ channel also arises dominantly from the strange quark current since the contribution to this channel from light quarks is Zweig rule suppressed. Conversely, we expect channels involving an η meson accompanied by non-strange states to be mainly due to light quark currents. For a_μ^s we need to add the contribution from energies above 1.8 GeV . It can be computed within PQCD and taken as one sixth of the corresponding light quark contribution [136] of 43.8×10^{-10} . The lower part of Table 4.3 shows further channels involving strange quarks to which first generation quark currents could conceivably

contribute, and we conservatively assign $(50 \pm 50)\%$ of these to the $s\bar{s}$ current. The table also shows the corresponding contributions to a_μ . Adding the totals in this way we find,

$$a_\mu^s = (50.77 \pm 0.60 \pm 0.83 \pm 2.23) \times 10^{-10} = (50.77 \pm 2.45) \times 10^{-10}, \quad (4.30)$$

and,

$$\Delta_s \alpha(1.8 \text{ GeV}) = (6.56 \pm 0.11 \pm 0.19 \pm 0.53) \times 10^{-4} = (6.56 \pm 0.57) \times 10^{-4}. \quad (4.31)$$

The first errors are experimental [122] where we accounted for correlations. The second errors allow for differences in parametrizations when decay parameters are extracted from experimental data by different groups. The last errors are half of the totals in Table 4.3, but we expect the $s\bar{s}$ current to virtually saturate the kaon channels in Table 4.3 because the larger strange quark mass should suppress the probability amplitude to produce an $s\bar{s}$ sea quark pair relative to first generation quark pairs.

The uncertainty in Eq. (4.31) is already about three times smaller than in the past [106]. We can reduce it further by quantifying our expectation that the strange quark current actually saturates the kaon channels listed in the bottom part of Table 4.3. For this, we re-write Eqs. (4.30) and (4.31) in the form,

$$a_\mu^s = (53.00 - 4.46 \kappa \pm 0.60 \pm 0.83) \times 10^{-10}, \quad (4.32)$$

$$\Delta_s \alpha(1.8 \text{ GeV}) = (7.09 - 1.05 \kappa \pm 0.11 \pm 0.19) \times 10^{-4}, \quad (4.33)$$

with a parameter $0 \leq \kappa \leq 1$, where $\kappa = 0$ ($\kappa = 1$) corresponds to the case where all kaon contributions in Table 4.3 arise from the strange (first generation) quark current. In order to confirm that indeed $\kappa \approx 0$ and to compute an uncertainty for possible $\kappa \neq 0$ effects, we can use results on a_μ^s from lattice gauge theory, as we show next.

4.4.2 Lattice data

Two groups [124, 125] calculated the contribution of the strange quark to the vacuum polarization function within lattice gauge theory with a focus on a_μ^s . The two results agree and average to

$$a_\mu^s = (53.32 \pm 0.49) \times 10^{-10} \quad [\text{lattice}], \quad (4.34)$$

which is in perfect agreement with Eq. (4.32) and our expectation $\kappa \approx 0$. Since the analogous result for $\Delta_s \alpha(1.8 \text{ GeV})$ has not been provided by either of the groups, we follow a Bayesian procedure to quantify the parameter κ in Eq. (4.33), using as prior information the comparison of Eq. (4.32) with Eq. (4.34). The 68.3% highest probability interval of κ , namely $0 \leq \kappa \leq \kappa_{1\sigma}$, can be obtained from

$$N \int_0^{\kappa_{1\sigma}} \exp \left[-\frac{(53.00 - 4.46 \kappa - 53.32)^2}{2(0.60^2 + 0.83^2 + 0.49^2)} \right] d\kappa = 0.683, \quad (4.35)$$

where N is the normalization of the distribution. This yields $\kappa_{1\sigma} = 0.22$, and Eq. (4.33) now provides us with the desired result,

$$\Delta_s \alpha(1.8 \text{ GeV}) = (7.09 \pm 0.11 \pm 0.19 \pm 0.23) \times 10^{-4} = (7.09 \pm 0.32) \times 10^{-4}, \quad (4.36)$$

which is consistent with, but more precise than Eq. (4.31). We assigned the uncertainty from $\kappa_{1\sigma}$ symmetrically around $\kappa = 0$, which is both the physically favored and most probable value (the peak of the distribution). This rather conservative treatment effectively doubles the error from $\kappa_{1\sigma}$, and is meant to account for the fact that the kernels of $\Delta \alpha$ and a_μ differ.

The experimental values of a_μ and $\Delta \alpha$ are correlated, possibly impacting Eq. (4.36). However, we found that even assuming them to be fully correlated changes the central value only very slightly and reduces the uncertainty modestly. Thus, we keep Eq. (4.36) as our final result on $\Delta_s \alpha(1.8 \text{ GeV})$. As an additional cross-check we used the vacuum polarization

function of another lattice calculation [125] of a_μ^s (expressed as a Padé approximant which is the source of the largest uncertainty [125]) to first reproduce their results, and then we computed $\Delta_s\alpha$ which yields,

$$\Delta_s\alpha(1.8 \text{ GeV}) \approx (6.9 \pm 0.5) \times 10^{-4} \quad [\text{lattice}], \quad (4.37)$$

in excellent agreement with Eq. (4.36).

4.4.3 Threshold masses

Heavy quarks

We can compute \bar{m}_c and \bar{m}_b in perturbation theory by reincorporating the RGE summable logarithms of the form $\ln \hat{m}_q/\bar{m}_q$ into Eq. (4.8), and then solving for \bar{m}_q by setting the contribution from quark q equal to zero. Since $\bar{m}_q \rightarrow \hat{m}_q$ for $\hat{\alpha}_s \rightarrow 0$, these logarithms are at most of order $\hat{\alpha}_s$ and can be ignored in the $\hat{\alpha}_s^3$ coefficient. Thus, we can use a previous analysis [96] where the logarithms up to order $\hat{\alpha}_s^2$ are given. We find,

$$\begin{aligned} \bar{m} &= \hat{m} \left\{ 1 - \frac{13}{24} \frac{\hat{\alpha}_s}{\pi} + \left(\frac{10073}{3456} - \frac{655}{288} \zeta_3 - \frac{361}{2592} n_q \right) \frac{\hat{\alpha}_s^2}{\pi^2} \right. \\ &+ \left(1.61024 + 0.59599 n_q - 0.00964 n_q^2 \right) \frac{\hat{\alpha}_s^3}{\pi^3} \\ &+ \left. \left[-\frac{295}{2592} \frac{\hat{\alpha}_s^2}{\pi^2} + \left(\frac{5767}{62208} - \frac{\mathcal{K}_1 + \mathcal{K}_2 n_q}{2} \right) \frac{\hat{\alpha}_s^3}{\pi^3} \right] \frac{\sum Q_\ell^2}{Q_h^2} \right\}. \end{aligned} \quad (4.38)$$

Using the input values for the Z boson mass [135], $M_Z = 91.1876$ GeV, the charm quark mass [137], $\hat{m}_c(\hat{m}_c) = 1.272$ GeV, and the bottom quark mass [135], $\hat{m}_b(\hat{m}_b) = 4.18$ GeV, together with the 4-loop RGE [130] for $\hat{\alpha}_s$ with $n_q = 4$ and $n_q = 5$, respectively, we find

$$\bar{m}_c = 1.185 \text{ GeV}, \quad (4.39)$$

$$\bar{m}_b = 3.990 \text{ GeV}. \quad (4.40)$$

It will be useful for later to define quantities ξ_q [106] as ratios between the threshold mass of quark q and the $1S$ $\bar{q}q$ bound state mass,

$$\xi_q \equiv \frac{2\bar{m}_q}{M_{1S}}. \quad (4.41)$$

This definition implies that $\xi_q \rightarrow 1$ for $\bar{m}_q \rightarrow \infty$ and $\xi_q \rightarrow 0$ for $\bar{m}_q \rightarrow 0$. We expect ξ_q to be a monotonically increasing in the sense that $\xi_1 > \xi_2$ if $\bar{m}_1 > \bar{m}_2$. Using the PDG values for the bound state masses [135] we find $\xi_c = 0.766$ and $\xi_b = 0.844$, and thus $\xi_b > \xi_c$ as expected.

Light quarks

Next we constrain the individual contributions of the light quarks to $\Delta\hat{\alpha}$, evaluated at \bar{m}_c . Using the RGE and the starting value given in Eq. (4.18) we obtain,

$$\Delta\hat{\alpha}^{(3)}(\bar{m}_c) = (65.10 \pm 0.45 \pm 0.18) \times 10^{-4}. \quad (4.42)$$

From Eq. (4.36) we can also calculate $\Delta_s\hat{\alpha}$ at \bar{m}_c . To do so, we first invoke experimental data to obtain the shift,

$$\Delta_s\alpha(2 \text{ GeV}) = \Delta_s\alpha(1.8 \text{ GeV}) + (0.55 \pm 0.04) \times 10^{-4}, \quad (4.43)$$

given by one sixth of the continuum contribution [122] of $(3.31 \pm 0.26) \times 10^{-4}$ between the two scales. The uncertainty is the difference to using PQCD instead of data and accounts for quark-hadron duality violations. Changing to the $\overline{\text{MS}}$ scheme and employing again the RGE gives,

$$\Delta_s\hat{\alpha}(\bar{m}_c) = (8.71 \pm 0.32) \times 10^{-4}. \quad (4.44)$$

Since the threshold mass is the value of the 't Hooft scale corresponding to trivial matching conditions regarding the QCD contribution, we can write,

$$\Delta_s \hat{\alpha}(\bar{m}_c) = Q_s^2 \frac{\alpha}{\pi} K_{\text{QCD}}^s(\bar{m}_c) \ln \frac{\bar{m}_c^2}{\bar{m}_s^2}, \quad (4.45)$$

where we defined a scale dependent factor $K_{\text{QCD}}^q(\mu)$ as the average QCD correction to the β function between \bar{m}_q and the scale μ . Eq. (4.45) has two unknowns, $K_{\text{QCD}}^s(\bar{m}_c)$ and \bar{m}_s , and it shows that increasing $K_{\text{QCD}}^s(\bar{m}_c)$ forces the logarithm to decrease and in turn \bar{m}_s to increase. Thus, smaller (larger) values of $K_{\text{QCD}}^s(\bar{m}_c)$ correspond to a smaller (larger) values of \bar{m}_s . On the other hand, if we have two quarks with masses $\bar{m}_1 > \bar{m}_2$, we expect the average QCD contribution between \bar{m}_2 and μ to be larger than that between \bar{m}_1 and μ , since α_s is larger at lower scales. Thus,

$$\bar{m}_1 > \bar{m}_2 \quad K_{\text{QCD}}^1(\mu) < K_{\text{QCD}}^2(\mu), \quad (4.46)$$

and we must have,

$$K_{\text{QCD}}^c(\bar{m}_c) < K_{\text{QCD}}^s(\bar{m}_c). \quad (4.47)$$

$K_{\text{QCD}}^c(\bar{m}_c)$ can be computed from Eq. (4.2). Using $n_q = 3$ and $\alpha_s(\bar{m}_c) = 0.413$ yields $K_{\text{QCD}}^c(\bar{m}_c) = 1.178$, and implies the lower bound,

$$\bar{m}_s > \bar{m}_c \exp \left[-\frac{\pi \Delta_s \hat{\alpha}(\bar{m}_c)}{2Q_s^2 \alpha K_{\text{QCD}}^c(\bar{m}_c)} \right] = 289 \text{ MeV}, \quad (4.48)$$

where we used $\alpha = \alpha(\bar{m}_s) \approx 1/135$. We can also obtain an upper bound on \bar{m}_s ,

$$\frac{2\bar{m}_s}{M_\phi} = \xi_s < \xi_c = 0.766 \quad \bar{m}_s < 390 \text{ MeV}, \quad (4.49)$$

implying $K_{\text{QCD}}^s(\bar{m}_c) < 1.50$. We can summarize these results by writing,

$$K_{\text{QCD}}^s(\bar{m}_c) = 1.34 \pm 0.16, \quad \bar{m}_s = 342_{-53}^{+48} \text{ MeV}. \quad (4.50)$$

\bar{m}_u and \bar{m}_d can be obtained in a similar way. We have,

$$\Delta_{\text{conn}}\hat{\alpha}^{(3)}(\bar{m}_c) = \Delta_s\hat{\alpha}(\bar{m}_c) + \frac{2\alpha}{\pi} \left[(Q_u^2 + Q_d^2) K_{\text{QCD}}^{u,d} \ln \frac{\bar{m}_c}{\bar{m}_{u,d}} \right], \quad (4.51)$$

where the quark connected contribution to $\Delta\hat{\alpha}^{(3)}(\bar{m}_c)$ is given by,

$$\Delta_{\text{conn}}\hat{\alpha}^{(3)}(\bar{m}_c) \equiv \Delta\hat{\alpha}^{(3)}(\bar{m}_c) - \Delta_{\text{disc}}\hat{\alpha}^{(3)}(\bar{m}_c) = 65.36 \times 10^{-4}. \quad (4.52)$$

Following the same steps as for \bar{m}_s we find,

$$K_{\text{QCD}}^{u,d}(\bar{m}_c) = 1.38 \pm 0.20, \quad \bar{m}_{u,d} = 246_{-57}^{+54} \text{ MeV}, \quad (4.53)$$

where the errors in Eqs. (4.50) and (4.53) are strongly correlated. The light quark threshold masses are convenient for implementing the RGE and serve an illustrative purpose, but their precise values affect $\hat{s}(0)$ only at order $\mathcal{O}(\alpha^2)$ and beyond, as long as the central value in Eq. (4.44) remains fixed (the uncertainty there will give rise to the flavor separation error). Notice, that for the central values we have $\bar{m}_s - \bar{m}_u \approx 96$ MeV, which is of typical size for hadronic mass splittings within $SU(3)$ flavor multiplets.

Finally, accounting for the squares of the electric charges we obtain the contributions from the first generation quarks at the scale \bar{m}_s ,

$$\Delta\hat{\alpha}^{(2)}(\bar{m}_s) = \Delta\hat{\alpha}^{(3)}(\bar{m}_c) - 6\Delta_s\hat{\alpha}(\bar{m}_c) = (12.9 \mp 1.9) \times 10^{-4}, \quad (4.54)$$

where we only quote the uncertainty from the flavor separation in Eq. (4.44).

source	$\delta \sin^2 \hat{\theta}_W(0) \times 10^5$
$\Delta \hat{\alpha}^{(3)}(2 \text{ GeV})$	1.2
flavor separation	1.0
isospin breaking	0.7
singlet contribution	0.3
PQCD	0.6
Total	1.8

Table 4.4: Theoretical uncertainties in the low energy mixing angle.

4.5 Theoretical uncertainties

In addition to parametric uncertainties, there are five sources of theoretical uncertainties for the weak mixing angle at low energies affecting our calculation. They are summarized in Table 4.4 and discussed in the following.

The first uncertainty is induced by the experimental error in the determination of $\Delta \hat{\alpha}^{(3)}(2.0 \text{ GeV})$. Eq. (4.5) propagates this uncertainty to the weak mixing angle [106],

$$\delta \hat{s}^2(0) = \left[\frac{1}{2} - \hat{s}^2 \right] \delta \Delta \hat{\alpha}^{(3)}(2 \text{ GeV}) = \mp 1.2 \times 10^{-5}, \quad (4.55)$$

where we have used $\delta \Delta \hat{\alpha}^{(3)}(2 \text{ GeV}) = \pm 0.45 \times 10^{-4}$ from Eq. (4.15).

The three light quarks enter with different electroweak weights into $\hat{s}^2(0)$ and $\Delta \alpha^{(3)}(\bar{m}_c)$. The flavor separation uncertainty is due to the imperfect knowledge of how much s quarks relative to u and d quarks contribute to $\Delta \alpha^{(3)}(\bar{m}_c)$. It is given by [106],

$$\delta \hat{s}^2(0) \simeq \frac{1}{20} \delta \Delta \hat{\alpha}^{(2)}(\bar{m}_c) = \pm 1.0 \times 10^{-5}, \quad (4.56)$$

where we used $\delta \Delta \hat{\alpha}^{(2)}(\bar{m}_s) = \pm 1.9 \times 10^{-4}$ from Eq. (4.54).

The flavor separation assumed isospin symmetry in the form $\bar{m}_u = \bar{m}_d$. To estimate the uncertainty associated with isospin breaking, we first consider the idealized case in which $SU(2)$ isospin violation was as large as $SU(3)$ breaking. This would occur for $\bar{m}_d = \bar{m}_s$, so

that from Eq. (4.54) the u quark current could at most contribute

$$\Delta\alpha^{(1)}(\bar{m}_d) < 14.8 \times 10^{-4}. \quad (4.57)$$

To propagate this uncertainty to $\hat{s}^2(0)$ we can use [106],

$$\delta\hat{s}^2(0) = -\frac{3}{40}\Delta\alpha^{(1)}(\bar{m}_d) > -1.1 \times 10^{-4}. \quad (4.58)$$

A measure of the breaking of $SU(2)$ relative to $SU(3)$ is given by the ratio,

$$\left| \frac{M_{K^{*\pm}}^2 - M_{K^{*0}}^2}{M_{K^{*\pm}}^2 - M_{\rho^0}^2} \right| \approx 0.06, \quad (4.59)$$

so that,

$$\delta\hat{s}^2(0) = {}_{-7}^{+0} \times 10^{-6}. \quad (4.60)$$

This error is asymmetric because we assume $\bar{m}_d \geq \bar{m}_u$, but it is convenient and conservative to treat it symmetrically in Table 4.4.

The uncertainty arising from the singlet contribution is given in Eq. (4.29). The last entry in Table 4.4 combines the truncation error from the perturbative matching conditions with the scheme conversion error shown as the second uncertainty in Eq. (4.18).

4.6 Results

Eq. (4.5) together with the Z pole value of the weak mixing angle from a global fit to the SM [135], $\sin^2 \hat{\theta}_W(M_Z) = 0.23129(5)$, can now be used to compute the weak mixing angle at zero momentum transfer,

$$\sin^2 \hat{\theta}_W(0) = 0.23868 \pm 0.00005 \pm 0.00002, \quad (4.61)$$

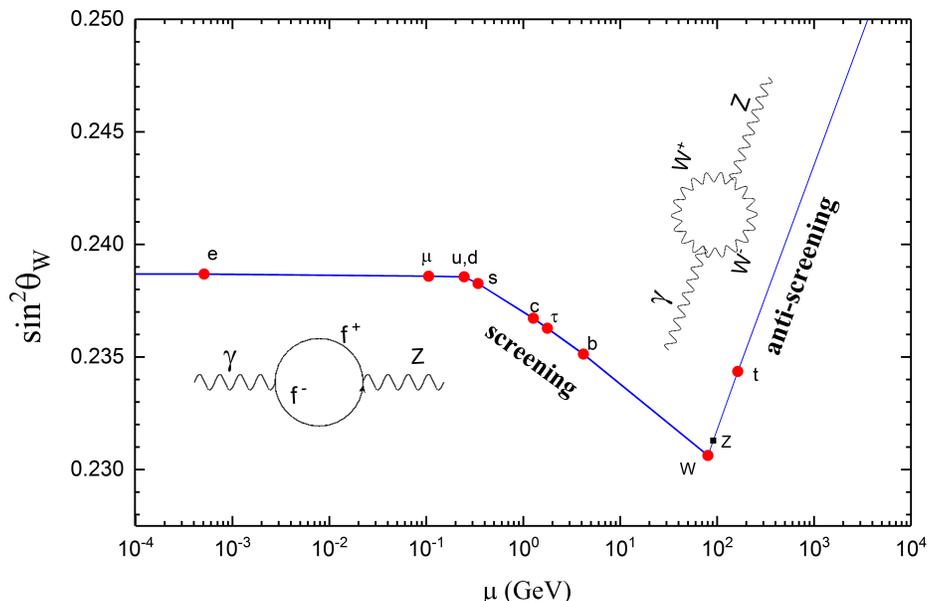


Figure 4.3: Scale dependence of the weak mixing angle in the $\overline{\text{MS}}$ renormalization scheme. The dots indicate the scales where a particle is integrated out. The total uncertainty corresponds to the thickness of the line. The β -function of $SU(2)_L$ changes sign at $\mu = M_W$, where the fermionic screening effects of the effectively Abelian gauge theory are being overcompensated by the anti-screening effects of the full non-Abelian electroweak theory.

where the second error is the total theoretical uncertainty from Table 4.4.

To facilitate the update of our results in the future, we also present a linearized formula of the form factor $\kappa(0)$,

$$\sin^2 \hat{\theta}_W(0) \equiv \hat{\kappa}(0) \sin^2 \hat{\theta}_W(M_Z), \quad (4.62)$$

in terms of variations of the input parameters, using $\Delta \hat{\alpha}_s(M_Z)$ in Eq. (4.17), as well as,

$$\tilde{\Delta} \alpha \equiv \Delta \alpha(2.0 \text{ GeV}) - 0.005871, \quad (4.63)$$

and,

$$\Delta \hat{m}_c \equiv \frac{\hat{m}_c(\hat{m}_c)}{1.272 \text{ GeV}} - 1, \quad \Delta \hat{m}_b \equiv \frac{\hat{m}_b(\hat{m}_b)}{4.180 \text{ GeV}} - 1. \quad (4.64)$$

We obtain,

$$\hat{\kappa}(0) = 1.03196 \pm 0.00006 + 1.14 \tilde{\Delta}\alpha + 0.025 \Delta\hat{\alpha}_s - 0.0016 \Delta\hat{m}_c - 0.0012 \Delta\hat{m}_b, \quad (4.65)$$

which shows that the current experimental uncertainties of $\pm 0.45 \times 10^{-4}$ in $\Delta\alpha(2 \text{ GeV})$ from Eq. (4.27) and of ± 0.0016 in $\hat{\alpha}_s(M_Z)$ induce errors of $\pm 5 \times 10^{-5}$ and $\pm 4 \times 10^{-5}$ in $\hat{\kappa}(0)$, respectively. Variations of $\pm 8 \text{ MeV}$ [137] in $\hat{m}_c(\hat{m}_c)$ and $\pm 30 \text{ MeV}$ in $\hat{m}_b(\hat{m}_b)$ both imply $\mp 2 \times 10^{-6}$ in $\hat{\kappa}(0)$ which is negligible. The resulting scale evolution of the weak mixing angle is illustrated in Figure 4.3.

When our result for the weak mixing angle in the Thomson limit or some other low momentum scale is used for the calculation of physical observables, there will generally be further process-dependent radiative corrections which need to be addressed. We expect this to be possible with theoretical uncertainties well below those in $\sin^2 \hat{\theta}_W(0)$ summarized in Table 4.4. Thus, we reduced the total theoretical uncertainty in the weak mixing angle at low energies from 7×10^{-5} [106] to less than 2×10^{-5} which can safely be neglected for any current or planned experiment.

In summary, we developed a new way of calculating the flavor separation which involved both $e^+e^- \rightarrow$ hadrons data and results from lattice gauge theory. We also better control now the uncertainty in the contribution of disconnected diagrams where we exploited results of Ref. [129] on the anomalous magnetic moment. Furthermore, we extended various formulas to the next order in perturbation theory, reducing the perturbative uncertainty. There has also been significant progress in the evaluations of $\Delta\alpha$ [97, 122] and $\Delta\hat{m}_c(\hat{m}_c)$ [137]. The theoretical uncertainty in $\sin^2 \hat{\theta}_W(0)$ is now at a negligible level.

Chapter Five

BSM physics and other topics

In the last chapters I have talked about physics within the Standard Model. Now I will talk about physics beyond the Standard Model. In general, this chapter contains research that I have done in collaborations with my colleagues at the Instituto de Física, of the Universidad Nacional Autónoma de Mexico.

5.1 Neutrino physics

5.1.1 Origin of neutrino mass

In chapter four I gave an introduction to the Standard Model, but there is a sector that I did not mention explicitly: the Yukawa sector. This sector is responsible of generating the masses of the SM fermions. Of particular interest is the mechanism that produces the masses of neutrinos since their masses are too small compared to the masses of the other particles, as can be seen in Figure 5.1. A simple possibility is that there are mass terms of the same form as the ones of the quarks, but with right handed neutrinos and Yukawa couplings that are very small by coincidence. Nevertheless, it is believed that physics beyond the SM is responsible of such small masses. In the SMEFT (Standard Model Effective Field Theory) framework, this small masses can be produced through a dimension 5 operator, the so called

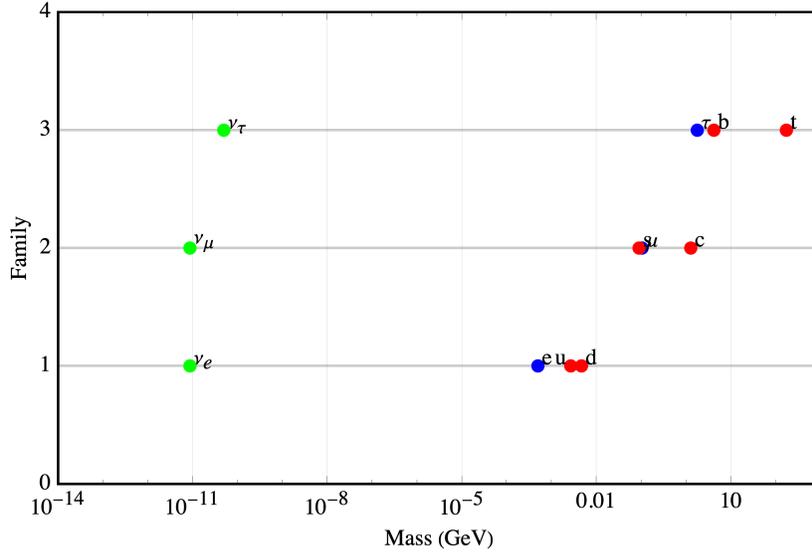


Figure 5.1: Fermions mass scales. The values of the neutrino masses are just shown for illustrative purpose taking into account the neutrino mass differences, the real value is still not known exactly.

Weinberg operator [138]. The idea is the following: Weinberg's operator has the form

$$\mathcal{O}_W^{(5)} = \frac{1}{2} \frac{C_{ij}}{\Lambda} (\overline{\Psi}_L^{ci} \widetilde{H}^*) (\widetilde{H}^\dagger \Psi_L^j) + \text{h.c.}, \quad (5.1)$$

where $\Psi_L = (\nu_L, \psi_L)^T$ is the left-handed lepton doublet of the SM, with ν the neutrino field, $H = (H^+, H^0)^T$ the Higgs doublet, $\widetilde{H} = i\sigma_2 H^*$, $C_{\alpha\beta}$ the dimensionless Wilson coefficient and Λ is the new physics scale. The indices i, j denote the families of the fermion fields. After spontaneous symmetry breaking, the dimension 5 operator will lead to terms of the form

$$\mathcal{O}_{Wmass}^{(5)} = \frac{1}{2} C_{ij} v \left(\frac{v}{\Lambda} \right) \overline{\nu}_L^{ci} \nu_L^j + \text{h.c.} \quad (5.2)$$

Using a Wilson coefficient of "natural" size $C_{\alpha\beta} \approx 0.1$ and $v = 246 \text{ GeV}$, we can explain neutrino masses of order $\sim 10^{-11} \text{ GeV}$ if we set $\Lambda \approx 10^{14} \text{ GeV}$. Then, this is the natural scale where we expect new physics in the neutrino sector to emerge.¹ After a bit of algebra

¹Nevertheless, some models can reduce this scale, for example if the dimension 5 Weinberg operator is produced through loops.

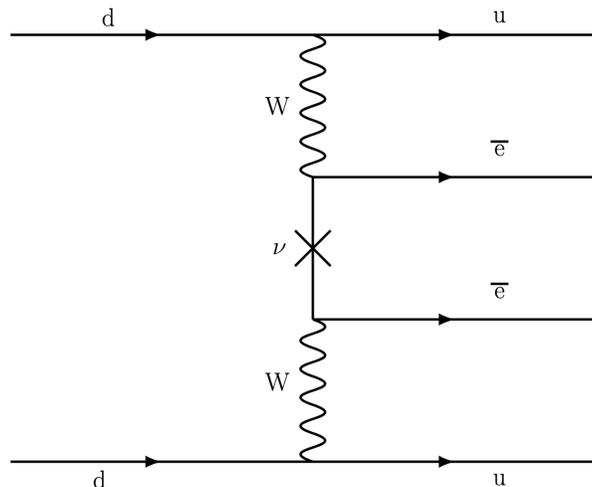


Figure 5.2: Process that contributes to the neutrino-less double beta decay.

and some rearrangements, it can be shown that the presence of a mass term with this form implies that the neutrinos are of Majorana nature, i.e that they are their own antiparticles, so we call this term in the Lagrangian a Majorana mass term. A Majorana mass term, (just as the Weinberg operator) violates lepton number by two units. This can be easily seen from the Lagrangian itself: the conjugated of the charge conjugated field has the same lepton number as the field itself, implying the non conservation of this global symmetry. There is a well known theorem known as the "Black-Box" theorem [139–142] which states that there is a one to one correspondence between having lepton number violating processes and a Majorana mass term in the Lagrangian. This implies that the measurement of a process that violates lepton number will elucidate the nature of the neutrino masses. The most common one is neutrino-less double beta decay, the Feynman diagram of this process is shown in Fig. 5.2. By the arguments mentioned before, neutrino-less double beta decay imposes constraints on the values of neutrino masses, see for example the PDG review [29].

5.1.2 Seesaw Formula

Now that we know that it is possible to understand neutrino masses as a consequence of new physics, we should ask ourselves: What kind of specific UV completions can induce

dimension 5 Weinberg operators? There are many possible models, for a detailed review see Ref. [143]. One of the simplest, earliest and most common set of models are the seesaw models [144–148]. These are generally divided in three types: type I, where heavy Majorana fermion singlets are included, type II, where a scalar triplet is included, and type III, with fermion triplets, although many models now include combinations of these mechanisms. Let us review the seesaw formula for the type I seesaw. Quite generally, the BSM Lagrangian of neutrino masses can be written as follows²

$$\mathcal{L}_M = -\frac{1}{2} \begin{pmatrix} \bar{\nu}_L & \bar{N}^c \end{pmatrix} \begin{pmatrix} m & m_D \\ m_D^T & M \end{pmatrix} \begin{pmatrix} \nu_L^c \\ N \end{pmatrix} + \text{h.c.} \quad (5.3)$$

This is actually a very compact notation, since here we define ν_L as the three left handed neutrino fields i.e $\nu_L = (\nu_L^e, \nu_L^\mu, \nu_L^\tau)^T$ and N as the set of right handed neutrino fields in the model. Thus, the matrix elements m , m_D and M are matrices themselves. To be specific, m is 3×3 matrix, m_D is a $3 \times n$ matrix (where n is the number of right handed fields) and M is a $n \times n$ matrix. Now let us integrate out the right handed heavy fields. The classical equations of motion are ³

$$\frac{\partial \mathcal{L}_M}{\partial N} = -\frac{1}{2} \bar{\nu}_L m_D - \frac{1}{2} \bar{N}^c M = 0, \quad (5.4)$$

$$\frac{\partial \mathcal{L}_M}{\partial N^c} = -\frac{1}{2} m_D^T \nu_L^c - \frac{1}{2} M N = 0, \quad (5.5)$$

using this equations, we can replace

$$\bar{N}^c = -\bar{\nu}_L m_D M^{-1}, \quad (5.6)$$

²We could have done the decoupling with the Higgs field included, but for our purposes it is just simpler to go to the broken phase from scratch.

³Since we are going to integrate out the heavy degrees of freedom, we neglect the kinetic term (this term is suppressed by factors of size $\frac{p}{M}$).

$$N = -M^{-1}m_D^T\nu_L^c, \quad (5.7)$$

in the Lagrangian, and we arrive to

$$\mathcal{L}_M = -\frac{1}{2} [\bar{\nu}_L m \nu_L^c - \bar{\nu}_L m_D M^{-1} m_D^T \nu_L^c] + \text{h.c.}, \quad (5.8)$$

then, the neutrino mass matrix for the light degrees of freedom can be written as

$$m_l = m - m_D M^{-1} m_D^T. \quad (5.9)$$

This is the seesaw formula for light neutrino masses. This formula can also be obtained via a diagonalization of the neutrino mass matrix.

5.1.3 Neutrino Oscillations: Theory

I guess that anyone who would read this thesis knows that an important consequence of a massive neutrino is the phenomena of flavor oscillations. The idea of neutrino oscillations was predicted many years ago by Bruno Pontecorvo [149]. Although Pontecorvo first proposed an oscillation of the type $\nu \rightarrow \bar{\nu}$, his pioneer work planted the idea of neutrino oscillations in the physics community. A few years later, Maki, Nakagawa, and Sakata [150] developed the formalism of neutrino flavor oscillation. This work was complemented in a second paper by Pontecorvo [18]. In this second paper, he mentioned possible ways to measure the phenomena, and estimated the fraction of electron neutrinos coming from the sun. But how do neutrino oscillations arise? This phenomena occurs because the flavor states are not the same as the mass eigenstates. In electroweak processes, flavor eigenstates are produced. This states will later travel through vacuum⁴. But in vacuum, the Hamiltonian eigenstates are the mass eigenstates and not the flavor ones. This implies that the flavor eigenstates will

⁴Not always the vacuum.

oscillate between each other. Let us review the formalism very quickly in a non formal way⁵.

A flavor eigenstate can be written as a linear combination of the mass eigenstates,

$$|\nu_\alpha\rangle = U_{\alpha i}^* |\nu_i\rangle, \quad (5.10)$$

where the latin indices indicate a mass state and the Greek indices a flavor state, the matrix $U_{\alpha i}$ is called the PMNS matrix⁶ and us a unitary 3×3 matrix. Now suppose a flavor neutrino is created with momentum p at $t = 0$. Then the wave function at time $t = 0$ is given by

$$|\psi(x, 0)\rangle = \int_{-\infty}^{\infty} d^3 \langle x | \tilde{p} \rangle U_{\alpha i}^* \langle \tilde{p} | \nu_i, p \rangle, \quad (5.11)$$

since the neutrinos have momentum p we get

$$|\psi(x, 0)\rangle = e^{ipx} U_{\alpha i}^* |\nu_i\rangle, \quad (5.12)$$

the time evolution of this wave function is given by the Hamiltonian, whose eigenvalues can be written as $E_i + \frac{m_i^2}{2E}$. Then after a time t , the wave function is given by

$$|\psi(x, t)\rangle = e^{i(px - E_i t)} U_{\alpha i}^* |\nu_i\rangle, \quad (5.13)$$

since the neutrinos almost move at the speed of light, we can approximate $t \approx x$ so the wave function at t is at position $x = t^7$, then

$$|\psi(t, t)\rangle = e^{-i \frac{m_i^2}{2E} x} U_{\alpha i}^* |\nu_i\rangle, \quad (5.14)$$

⁵There are many references that solve this (for me) non trivial problem. For example, [151] uses a wave packet to describe the oscillations. Using this method gives a proper definition of localization of the wave function, and as a consequence, a clear definition of the propagation of it. On the other hand, [152] uses the QFT formalism, and computes the full amplitude of a neutrino process (from production to detection).

⁶The name comes from Maki, Nakagawa, Sakata and Pontecorvo.

⁷This is much more clearer in the wavepacket approach, there, the wave packet is approximately located at $x = t$.

and the probability of finding a neutrino of flavor β at this point will be given by

$$P(\alpha \rightarrow \beta) = |\langle \nu_\beta | \psi(t, t) \rangle|^2 = \left| e^{-i \frac{m_i^2}{2E} x} U_{\alpha i}^* \langle \nu_\beta | \nu_i \rangle \right|^2 = \left| e^{-i \frac{m_i^2}{2E} x} U_{\alpha i}^* U_{\beta i} \right|^2. \quad (5.15)$$

We can extract an overall phase to the element inside the absolute value. Let us extract the phase that corresponds to m_3 . Then, we get

$$P(\alpha \rightarrow \beta) = \left| e^{-i \frac{\Delta m_{13}^2}{2E} x} U_{\alpha 1}^* U_{\beta 1} + e^{-i \frac{\Delta m_{23}^2}{2E} x} U_{\alpha 2}^* U_{\beta 2} + U_{\alpha 3}^* U_{\beta 3} \right|^2. \quad (5.16)$$

where $\Delta m_{ij} \equiv m_i^2 - m_j^2$. This implies that neutrino oscillation experiments can only measure differences of neutrino masses. The next step is to find out how many parameters do neutrino oscillation experiments should measure to get a full description of the unitary matrix U . First, let us do the counting of degrees of freedom of a 3×3 unitary matrix. An arbitrary complex squared matrix of dimension n has $2n^2$ free parameters. The orthogonality conditions of a unitary matrix will impose many constraints on this parameters. The orthogonality condition of the first column with the remaining $n - 1$ columns imposes $2(n - 1)$ constraints. The constraint of the second column with the remaining $n - 2$ columns will impose $2(n - 2)$ constraints, the third with the remaining ones gives $2(n - 3)$ and so on and so forth. Thus from orthogonality between columns we get $2 \sum_{i=1}^{n-1} (n - i) = n^2 - n$ constraints. Now, the normalization condition will add another n constraints (it is not $2n$ because these constraints are by construction always real). Then the number of free parameters for a unitary $n \times n$ matrix will be $2n^2 - (n^2 - n) - n = n^2$. This imply that for our 3×3 PMNS matrix we would have 9 parameters to determine from experiments. But this is not the whole story, since we can make re-definitions of our fields. First let us study the nature of this parameters. We can divide them in two types: angles and phases. To count the number of angles we should count the number of degrees of freedom in an orthogonal real matrix of dimension n . The counting is quite similar to the one of the unitary matrix and the result is that the number

of parameters is $n(n-1)/2$. For the particular case $n=3$, we have 3 angles. Then for a unitary matrix of dimension 3 we would have 9 parameters: 3 angles and 6 phases. Some of these phases might be removed by a redefinition of the fields.

To understand this properly, let us review how do the PMNS matrix shows up in the Lagrangian. To simplify the analysis, I will change my notation slightly. Instead of using Ψ_L^α to denote a Standard Model doublet of fermions of flavor α , I will write it as L_α . Furthermore, instead of using e_R^α to denote a right handed charged fermion, I will use l_α . Suppose that we are in a basis where the charged lepton mass matrix is diagonal and the neutrino mass matrix is not. To get a diagonal neutrino mass matrix, we must transform the states using Eq. (5.10). At the Lagrangian level, this shows up as a transformation on the fields itself⁸. This unitary transformation will cancel out in all the terms of the Lagrangian, except in two places, the charged current and in the mass term itself (since it sends it to its diagonal form). It does not cancel out in the charged current because the charge current has the form $\bar{l}_\alpha \gamma_\mu (1 - \gamma_5) \nu_\alpha$, so a transformation of the neutrino field would give something like $\bar{l}_\alpha \gamma_\mu (1 - \gamma_5) U_{\alpha i} \nu_i$.

Now, suppose that we have Dirac neutrinos. Then we can rotate each of the neutrino fields by a phase without changing the Lagrangian at all, except for the charged current term, this phase transformation can be assigned to the U matrix. The same can be done for the charged lepton fields. So we have three phases from the neutrinos and three from the charged fermions. That is 6 phases in total. But there is one redefinition that does nothing to the neutrino mass matrix: transforming each field by the same phase. In this case, the phase in \bar{l} will be cancelled by the phase in ν . Then the number of field phase redefinitions available for us is only 5. This would imply that the number of parameters in our PMNS matrix is $9-5=4$, three angles and one phase.

For Majorana neutrinos we cannot make a phase rotation to the neutrino fields since the Lagrangian is not invariant under this transformation (this is the same reason why lepton

⁸The transformation of the fields is the inverse of Eq. (5.10).

number is not conserved). Then we would only have the three phases of the charged leptons, and that is all. So in the case of Majorana neutrinos we would have $9-3=6$ free parameters, that is three angles and three phases. These two extra phases that emerge for Majorana neutrinos are called Majorana phases.

5.1.4 Neutrino Oscillations: Experiments

There are several ways to measure neutrino oscillations. This subsection is just a review of the literature about this measurements, and by no means has the intent to be a complete reference of experimental neutrino physics. If the reader is interested in understanding this topics in more detail, it is invited to look through the bibliography [29, 153, 154]. To start talking about measurements, let us settle down the notation. A 3-tuple of flavor neutrinos will be ordered as $(\nu_e, \nu_\mu, \nu_\tau)$, while for the PMNS matrix I will use the PDG parametrization

$$U = \begin{pmatrix} s_{12}c_{13} & c_{13}s_{12} & e^{-i\delta}s_{13} \\ -c_{23}s_{12} - e^{i\delta}c_{12}s_{13}s_{23} & c_{12}c_{23} - e^{i\delta}s_{12}s_{13}s_{23} & c_{13}s_{23} \\ s_{12}s_{23} - e^{i\delta}c_{12}c_{23}s_{13} & -e^{i\delta}c_{23}s_{12}s_{13} - c_{12}s_{23} & c_{13}c_{23} \end{pmatrix} \times \text{diag} \left(1, e^{i\frac{\alpha_{21}}{2}}, e^{i\frac{\alpha_{31}}{2}} \right), \quad (5.17)$$

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$. Let us review how each element of this matrix and the neutrino mass differences can be experimentally measured.

Measurements of θ_{12} and Δm_{21} . This are one of the first parameters ever measured. The pioneering work was done by Homestake [155] in 1968. This experiment measured the neutrino flux coming from the sun. In the sun, neutrinos are mainly produced through the pp chain, the result of these process can be summarized as $4p \rightarrow {}^4\text{He} + 2e^+ + 2\nu_e$. Homestake measured the amount of these neutrinos through the reaction $\nu_e + {}^{37}\text{Cl} \rightarrow e^- + {}^{37}\text{Ar}$, the ${}^{37}\text{Ar}$ was counted later observing their decay in a miniature gas-filled proportional counter. They found a deficiency of electron neutrinos, compared to what was expected from the Standard Solar Model [156, 157]. Several more radiochemical detectors were used to detect solar

neutrinos in different processes and energy ranges [158–160]. They all detected a deficiency of electron neutrinos with respect to the Solar Model. A different type of experiment started a few years later, Kamiokande [161]. It was set up to detect $\nu - e$ elastic scattering in the reaction $\nu_x + e^- \rightarrow \nu_x + e^-$. This process is particularly sensitive to electron neutrinos due to the higher cross section. They detected the Cherenkov radiation produced by the outgoing electron moving at high speeds in water. The difference with respect to the radiochemical experiments is that the detection was done in real time and that it was also possible to measure the directionality of the incoming and outgoing neutrinos. They also measured a deficiency of electron neutrinos. Later, another experiment (SNO)[162] used heavy water to measure the neutrino flux from the sun through two different reactions, the charged current one $\nu_e + d \rightarrow e^- + p + p$ and the neutral current $\nu_x + d \rightarrow \nu_x + p + n$. The charged current reaction is only sensitive to electron neutrinos, while the neutral current is sensitive to all neutrinos in the same way, i.e they all have the same cross section. This last process was crucial to solve the neutrino solar problem (the deficiency of electron neutrinos). This is because these experiment confirmed that the total flux of the three types of neutrinos was equal to the electron neutrino flux expected from the Solar Model. It is obvious that this experiment must be sensitive to $\theta_{12}, \theta_{13}, \Delta m_{12}$. It will turn out, that given the smallness of θ_{13} and the value of Δm_{23}^2 , solar neutrino experiments are mainly sensitive to θ_{12} and in less degree to Δm_{12}^2 .

A more precise measurement of Δm_{12}^2 can be obtained from long baseline neutrino experiments from reactors. This experiments measure electron antineutrinos emitted from nuclear reactors. One of such experiments is KamLand in Japan [163]. The detector is located in the Kamiokande cavern, and the average distance to the reactors around it is 180 Km until 2011. The detection process is $\bar{\nu}_e + p \rightarrow e^+ + n$.

Measurements of θ_{13} . The experiments that measure more precisely θ_{13} are short baseline reactor experiments. The detection reaction is the same as in Kamland, but the average distance to the detector is much smaller (around 1 Km). One of the first experiments of

this kind was CHOOZ [164] which showed that the parameter θ_{13} must be very small. After this experiments, many more came such as Double Chooz [165], Daya Bay [166] and RENO [167].

Measurements of θ_{23} , Δm_{23}^2 and δ_{CP} . The first two parameters can be obtained via atmospheric neutrinos. Atmospheric neutrinos are produced in a process that starts when the cosmic rays collide with the atmosphere. As a product of this collision, unstable particles such as pions are produced. Then, pions will decay through charged current interaction into muons and their corresponding antineutrino. The next step of this process is the decay of the muon into an electron, an electron antineutrino, and a muon neutrino. Thus the overall process will have the net result $\pi^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu + \bar{\nu}_\mu$. This imply that the expected ratio of muon type neutrinos over electron neutrinos must be roughly around two. One of the most important experiments constructed to measure atmospheric neutrinos is Super Kamiokande. To detect the neutrinos, they use the cherenkov radiation of the charged lepton produced when the neutrinos interact with water through a charged current interaction. To distinguish between an electron and a muon, the shape of the Cherenkov ring is observed, since muons will produce a ring that is less diffuse than the electron ring. What they observe is a disappearance of muon neutrinos with no excess of electron ones. This gave a way then to measure the remaining parameters θ_{23} and Δm_{23}^2 . The CP violation phase can be observed via accelerator long baseline electron neutrinos appearance. For detailed references see the PDG review [29].

5.1.5 The road map to construct neutrino models.

I have mentioned some of the pieces required to construct neutrino mass models. In particular, during my PhD I worked in extensions to the SM using discrete symmetries. Let us pull together the pieces to understand these models. The idea is to set a UV theory where the new and Standard Model fields transform in specific ways under this symmetry. This

symmetry will be broken through spontaneous symmetry breaking to a smaller symmetry group. After symmetry breaking, we can use the seesaw formula. But as a consequence of the original symmetry, this seesaw will give a specific shape to the neutrino mass matrix. Due to the specific form of this mass matrix, correlations between the oscillation parameters will emerge. These correlations must be compatible with observed neutrino oscillation data and neutrino-less double beta decay constraints. An interesting feature of our model is that the remaining symmetry protects some of the BSM particles from decaying into standard model particles, giving a possible source of dark matter. Now I present a slightly modified version of our paper [168].

5.2 Previous works on the topic

The pattern of neutrino masses and mixing, which is very different to the quarks pattern, has been extensively studied. Two main approaches for correlating neutrino oscillation observables have been used, one based on non-Abelian flavor symmetries, very useful to explain certain patterns in the mixing parameters such as the tri-bimaximal mixing [169], bi-maximal mixing [170] or the golden ratio mixing [171, 172], and one based on the assumption of zeros in the neutrino mass matrix and diagonal charged lepton mass matrix. The Glashow-Frampton-Marfatia classification for the two-zero texture Majorana neutrino mass matrices is given in Table 5.1 [173]. In this letter, we demonstrate how some of the two-zero textures in the neutrino mass matrix can be obtained in a framework of the non-Abelian ⁹ flavor symmetry A_4 . In this framework the dark matter (DM) stability is due to a residual Z_2 symmetry of A_4 . For a model based on A_4 where the Majorana neutrinos acquire masses through type I and type II see-saw mechanism giving rise to texture zeros see¹⁰ [177]. In a

⁹It is also possible to obtain texture zeros using Abelian symmetries like in [174], or [175], where several scalar fields are needed or also [176] where the symmetry group used is $Z_2 \times Z_8$ in a left-right symmetric model.

¹⁰In the model by Hirsch *et. al* [177], the texture zeros correspond to B_1 and B_2 in the classification of [173].

recent work [178], texture zeros were obtained corresponding to the B_3 and B_4 in Table 5.1. In this model, the texture zeros are related with the flavor symmetry breaking, the same that is responsible for the stability of DM. Following this approach, we obtain the two textures predicting a vanishing neutrino-less double beta decay at tree-level, namely A_1 and A_2 .

5.2.1 The framework

A_4 is the alternating group of four objects. It is formed by the even permutations of the larger permutation group S_4 . A_4 is generated by S and T where $S^2 = T^3 = (ST)^3 = I$. The dimensionality of A_4 is twelve and the number of conjugacy classes is four. Then we have four irreducible representations (irreps), three of them are one-dimensional $\mathbf{1}$, $\mathbf{1}'$, $\mathbf{1}''$ and one three-dimensional $\mathbf{3}$. Actually it is the smallest discrete group which contains a triplet irreducible representation, this has been extensively used by model builders because it is possible to accommodate the families in a triplet representation [179–181]. The generators S and T in the S -diagonal basis are given in Table 5.2.1.

Case	Texture zeros	((a,b),(c,d))
A_1	$(m_\nu)_{ee} = (m_\nu)_{e\mu} = 0$	$((1,1),(1,2))$
A_2	$(m_\nu)_{ee} = (m_\nu)_{e\tau} = 0$	$((1,1),(1,3))$
B_1	$(m_\nu)_{\mu\mu} = (m_\nu)_{e\tau} = 0$	$((2,2),(1,3))$
B_2	$(m_\nu)_{\tau\tau} = (m_\nu)_{e\mu} = 0$	$((3,3),(1,2))$
B_3	$(m_\nu)_{\mu\mu} = (m_\nu)_{e\mu} = 0$	$((2,2),(1,2))$
B_4	$(m_\nu)_{\tau\tau} = (m_\nu)_{e\tau} = 0$	$((3,3),(1,3))$
C	$(m_\nu)_{\mu\mu} = (m_\nu)_{\tau\tau} = 0$	$((2,2),(3,3))$

Table 5.1: Texture zero matrices.

To obtain singlets under the A_4 symmetry, it will be necessary to compute direct products of irreps. In particular, the product rule for two triplet representations [179] is

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{1}' \oplus \mathbf{1}'' \oplus \mathbf{3} \oplus \mathbf{3}, \quad (5.18)$$

Irrep	S	T
1	1	1
1'	1	ω
1''	1	ω^2
3	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

Table 5.2: A_4 generators for the irreducible representations in the S diagonal basis, $\omega = e^{i2/3\pi}$ is the cubic root of 1.

	L_e	L_μ	L_τ	l_e	l_μ	l_τ	N_T	N_4	N_5	H	η	ϕ
SU(2)	2	2	2	1	1	1	1	1	1	2	2	1
A_4	α	β	γ	α	β	γ	3	δ	ϵ	1	3	3

Table 5.3: Summary of the particle content and quantum numbers. $\alpha, \beta, \gamma, \delta$ and ϵ can be any of the singlet representations $\mathbf{1}, \mathbf{1}', \mathbf{1}''$.

where the representations $\mathbf{1}'$ and $\mathbf{1}''$ are complex conjugate to each other. The model we considered here contains an extended Higgs sector with the SM Higgs H transforming as a singlet $\mathbf{1}$, three copies of Higgses in a triplet representation of A_4 , $\eta = (\eta_1, \eta_2, \eta_3)^T$, and three scalar singlets of the SM also in a triplet of A_4 , $\phi = (\phi_1, \phi_2, \phi_3)^T$. We also considered five right-handed (RH) neutrinos, three of them in the triplet representation of A_4 , $N_T = (N_1, N_2, N_3)^T$ and two singlets N_4 and N_5 . The complete assignment of the matter fields to irreps of A_4 is shown in Table 5.3.

The most general Lagrangian consistent with the symmetries of our theory is

$$\begin{aligned}
 \mathcal{L}_Y &= y_e \bar{L}_e l_e H + y_\mu \bar{L}_\mu l_\mu H + y_\tau \bar{L}_\tau l_\tau H \\
 &+ y_1^\nu \bar{L}_e [N_T \eta]_\alpha + y_2^\nu \bar{L}_\mu [N_T \eta]_\beta + y_3^\nu \bar{L}_\tau [N_T \eta]_\gamma \\
 &+ y_4^{\nu 1} \delta_{\alpha\delta} \bar{L}_e N_4 \tilde{H} + y_4^{\nu 2} \delta_{\beta\delta} \bar{L}_\mu N_4 \tilde{H} + y_4^{\nu 3} \delta_{\gamma\delta} \bar{L}_\tau N_4 \tilde{H} \\
 &+ y_5^{\nu 1} \delta_{\alpha\epsilon} \bar{L}_e N_5 \tilde{H} + y_5^{\nu 2} \delta_{\beta\epsilon} \bar{L}_\mu N_5 \tilde{H} + y_5^{\nu 3} \delta_{\gamma\epsilon} \bar{L}_\tau N_5 \tilde{H} \\
 &+ M \bar{N}_T^c N_T + M_4 \delta_{\delta 1} \bar{N}_4^c N_4 + M_5 \delta_{\epsilon 1} \bar{N}_5^c N_5 \\
 &+ y_2^{N_1} \delta_{1\delta} [\bar{N}_T^c \phi]_1 N_4 + y_2^{N_{1''}} \delta_{1''\delta} [\bar{N}_T^c \phi]_{1'} N_4
 \end{aligned} \tag{5.19}$$

$$\begin{aligned}
 &+ y_2^{N_{1'}} \delta_{1'\delta} [\overline{N_T^c \phi}]_{1''} N_4 + y_3^{N_1} \delta_{1\epsilon} [\overline{N_T^c \phi}]_1 N_5 \\
 &+ y_3^{N_{1''}} \delta_{1''\epsilon} [\overline{N_T^c \phi}]_{1'} N_5 + y_3^{N_{1''}} \delta_{1'\epsilon} [\overline{N_T^c \phi}]_{1''} N_5 \\
 &+ y_1^N [\overline{N_T^c \phi}]_{3_1} N_T + (y_1^N)' [\overline{N_T^c \phi}]_{3_2} N_T \\
 &+ M_{45} \delta_{\delta\epsilon^*} \overline{N_4^c} N_5 + h.c. \quad ,
 \end{aligned}$$

where $\alpha, \beta, \gamma, \delta$ and ϵ can be any of the three singlet representations of A_4 . Notice that charged leptons are diagonal since $\mathbf{1}^* = \mathbf{1}''$ and $\mathbf{1}' \otimes \mathbf{1}'' = \mathbf{1}$. The scalar fields η, H and ϕ get a vacuum expectation value (vev). There are two possible configurations of the vev to get a minimum of the potential, these are $\langle \phi \rangle_{Z_2} = (v_\phi, 0, 0)^T$ (S invariant) and $\langle \phi \rangle_{Z_3} = (v_\phi, v_\phi, v_\phi)^T$ (T invariant). The first one leaves a Z_2 symmetry after the breaking and the other one will leave a Z_3 symmetry, which can be easily seen from the generators in Table 5.2.1. The same arguments apply for the vev of η .

L_e	L_μ	L_τ	N_4	N_5	Neutrino Matrix	Type
$\mathbf{1}$	$\mathbf{1}''$	$\mathbf{1}'$	$\mathbf{1}$	$\mathbf{1}'$	$\begin{pmatrix} X & 0 & X \\ 0 & 0 & X \\ X & X & X \end{pmatrix}$	B_3
$\mathbf{1}$	$\mathbf{1}''$	$\mathbf{1}'$	$\mathbf{1}$	$\mathbf{1}''$	$\begin{pmatrix} X & X & 0 \\ X & X & X \\ 0 & X & 0 \end{pmatrix}$	B_4
$\mathbf{1}''$	$\mathbf{1}$	$\mathbf{1}'$	$\mathbf{1}$	$\mathbf{1}'$	$\begin{pmatrix} 0 & 0 & X \\ 0 & X & X \\ X & X & X \end{pmatrix}$	A_1
$\mathbf{1}''$	$\mathbf{1}'$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}'$	$\begin{pmatrix} 0 & X & 0 \\ X & X & X \\ 0 & X & X \end{pmatrix}$	A_2

Table 5.4: Particle transformation under A_4 that gives rise to texture zeros in the neutrino mass matrices. We get the same texture matrix if we exchange $\mathbf{1}' \leftrightarrow \mathbf{1}''$ in each line. In the same way we get the same texture if we exchange the representations of N_4 and N_5 . The last column gives the matrix type according to [173].

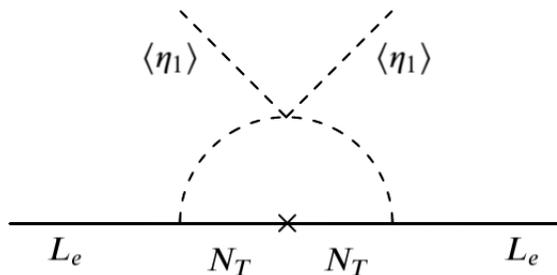


Figure 5.3: Loop diagram which generates a non-zero m_{ee} and consequently a small double beta decay rate.

5.2.2 Results

With the matter content of our model we were able to get the type A_1 , A_2 , B_3 and B_4 texture zeros. This textures are only obtained when the flavor symmetry is broken into the Z_2 subgroup. The residual Z_2 symmetry (corresponding to the S generator) is

$$\begin{aligned}
 \eta_1 &\rightarrow \eta_1, & \eta_2 &\rightarrow -\eta_2, & \eta_3 &\rightarrow -\eta_3, \\
 \phi_1 &\rightarrow \phi_1, & \phi_2 &\rightarrow -\phi_2, & \phi_3 &\rightarrow -\phi_3, \\
 N_1 &\rightarrow N_1, & N_2 &\rightarrow -N_2, & N_3 &\rightarrow -N_3.
 \end{aligned}
 \tag{5.20}$$

The lightest and Z_2 odd (and neutral) particle would play the role of DM since it will be stable [178, 182]. In the active sector besides the SM fields, we have two scalars (η_1 , ϕ_1) and three right handed neutrinos (N_1 , N_4 , N_5). While in the dark sector we have all the Z_2 odd fields in Eq. (5.20).

The models resulting from fixing the irreducible representations for the left handed fields and the (RH) neutrinos in Eq. (5.19) are summarized in Tab. 5.4. Textures B_3 and B_4 were already reported in [178]. The type A_1 and A_2 matrices have a zero tree level contribution to the neutrinoless double beta decay effective mass parameter, m_{ee} ,¹¹ and also zero $m_{e\mu}$ and $m_{e\tau}$ components respectively. These zeros are extremely powerful and predictive. To

¹¹A direct consequence of a vanishing m_{ee} is that these textures are only compatible with normal hierarchy.

see this, let us remark that, in general, neutrino oscillation experiments can only access to two neutrino mass differences, three mixing angles and one CP violating phase *i.e.* to six parameters of the theory. On the other hand, the theory has three masses, three mixing angles and three CP violating phases. This is where the zeros play an important role, they give four constraints (because the elements are in general complex) which give correlations between the six parameters to which the experiments have access. We will see explicitly the Lagrangians for models A_1 and A_2 by fixing the field irreps in Eq. (5.19).

Model for A_1

In this case, the irreps for the lepton fields are as follows: L_e , L_μ and L_τ transform as $\mathbf{1}''$, $\mathbf{1}$ and $\mathbf{1}'$ respectively, while N_4 and N_5 are in $\mathbf{1}$ and $\mathbf{1}'$, or in the notation of Table 5.3 $\alpha = \mathbf{1}''$, $\beta = \mathbf{1}$, $\gamma = \mathbf{1}'$, $\delta = \mathbf{1}$ and $\epsilon = \mathbf{1}'$. The Lagrangian in eq. (5.19) is reduced to

$$\begin{aligned}
 \mathcal{L}_Y &= y_e \bar{L}_e l_e H + y_\mu \bar{L}_\mu l_\mu H + y_\tau \bar{L}_\tau l_\tau H \\
 &+ y_1' \bar{L}_e [N_T \eta]_{1''} + y_2' \bar{L}_\mu [N_T \eta]_1 + y_3' \bar{L}_\tau [N_T \eta]_{1'} \\
 &+ y_4^{\nu^2} \bar{L}_\mu N_4 \tilde{H} + y_5^{\nu^3} \bar{L}_\tau N_5 \tilde{H} + M \bar{N}_T^c N_T \\
 &+ M_4 \bar{N}_4^c N_4 + y_2^{N_1} [\bar{N}_T^c \phi]_1 N_4 + y_3^{N_{1'}} [\bar{N}_T^c \phi]_{1''} N_5 \\
 &+ y_1^N [\bar{N}_T^c \phi]_3 N_T + (y_1^N)' [\bar{N}_T^c \phi]_{3_2} N_T + h.c.
 \end{aligned} \tag{5.21}$$

Once the flavor symmetry is broken by the scalar field ϕ , the mass matrix for the RH neutrino fields takes the form¹² [178]

$$M_R = \begin{pmatrix} M & 0 & 0 & v_\phi y_2^{N_1} & v_\phi y_3^{N_{1'}} \\ 0 & M & M_\phi & 0 & 0 \\ 0 & M_\phi & M & 0 & 0 \\ v_\phi y_2^{N_1} & 0 & 0 & M_4 & 0 \\ v_\phi y_3^{N_{1'}} & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{5.22}$$

¹²Note that the RH neutrino mass matrix is the same for case A_2 .

where $M_\phi = v_\phi(y_1^N + (y_1^N)')$, while the Dirac neutrino mass matrix is given by

$$m_D = \begin{pmatrix} v_\eta y_1^\nu & 0 & 0 & 0 & 0 \\ v_\eta y_2^\nu & 0 & 0 & v_H y_4^{\nu 2} & 0 \\ v_\eta y_3^\nu & 0 & 0 & 0 & v_H y_5^{\nu 3} \end{pmatrix}. \quad (5.23)$$

Using the type I seesaw formula for light neutrino masses $m_\nu = -m_D M_R^{-1} m_D^T$, the light left-handed neutrino mass matrix takes the form

$$m_\nu = \begin{pmatrix} 0 & 0 & w \\ 0 & x & y \\ w & y & z \end{pmatrix}, \quad (5.24)$$

which corresponds to the A_1 texture, where

$$\begin{aligned} w &= \frac{-v_H v_\eta y_1^\nu y_5^{\nu 3}}{(v_\phi y_3^{N_{1'}})} & x &= \frac{-v_H^2 (y_4^{\nu 2})^2}{M_4}, \\ y &= \frac{y_3^{N_{1'}} v_H (v_H v_\phi y_4^{\nu 2} y_2^{N_1} - v_\eta M_4 y_2^\nu)}{M_4 v_\phi y_3^{N_{1'}}}, \\ z &= \frac{y_3^{N_{1'}} v_H (M M_4 y_3^{N_{1'}} v_H - 2 M_4 y_3^\nu y_3^{N_{1'}} v_\phi v_\eta - v_H v_\phi^2 (y_2^{N_1})^2 y_3^{N_{1'}})}{M_4 v_\phi^2 (y_3^{N_{1'}})^2}. \end{aligned} \quad (5.25)$$

Model for A_2

In this case, the irreps for the lepton fields are as follows: L_e , L_μ and L_τ transform as $\mathbf{1}''$, $\mathbf{1}'$ and $\mathbf{1}$ respectively, while N_4 and N_5 are in $\mathbf{1}$ and $\mathbf{1}'$, or in the notation of Table 5.3 $\alpha = \mathbf{1}''$, $\beta = \mathbf{1}'$, $\gamma = \mathbf{1}$, $\delta = \mathbf{1}$ and $\epsilon = \mathbf{1}'$. The resulting Lagrangian is

$$\begin{aligned} \mathcal{L}_Y &= y_e \bar{L}_e l_e H + y_\mu \bar{L}_\mu l_\mu H + y_\tau \bar{L}_\tau l_\tau H \\ &+ y_1^\nu \bar{L}_e [N_T \eta]_{1''} + y_2^\nu \bar{L}_\mu [N_T \eta]_{1'} + y_3^\nu \bar{L}_\tau [N_T \eta]_1 \end{aligned} \quad (5.26)$$

$$\begin{aligned}
 & + y_4^{\nu^3} \bar{L}_\tau N_4 \tilde{H} + y_5^{\nu^2} \bar{L}_\mu N_5 \tilde{H} + M \bar{N}_T^c N_T \\
 & + M_4 \bar{N}_4^c N_4 + y_2^{N_1} [\bar{N}_T^c \phi]_1 N_4 + y_3^{N_{1''}} [\bar{N}_T^c \phi]_{1''} N_5 \\
 & + y_1^N [\bar{N}_T^c \phi]_{3_1} N_T + (y_1^N)' [\bar{N}_T^c \phi]_{3_2} N_T + h.c.
 \end{aligned}$$

After the breaking of the flavor and electroweak symmetries the Dirac neutrino mass matrix is

$$m_D = \begin{pmatrix} v_\eta y_1^\nu & 0 & 0 & 0 & 0 \\ v_\eta y_2^\nu & 0 & 0 & 0 & v_H y_5^{\nu^2} \\ v_\eta y_3^\nu & 0 & 0 & v_H y_4^{\nu^3} & 0 \end{pmatrix}, \quad (5.27)$$

and the light left-handed neutrino mass matrix takes the form

$$m_\nu = \begin{pmatrix} 0 & w' & 0 \\ w' & x' & y' \\ 0 & y' & z' \end{pmatrix}, \quad (5.28)$$

which corresponds to the A_2 texture, where

$$\begin{aligned}
 w' & = \frac{-v_H v_\eta y_5^{\nu^2} y_1^\nu}{v_\phi y_3^{N_{1'}}}, \\
 x' & = \frac{v_H y_5^{\nu^2} (-2M_4 v_\eta v_\phi y_2^\nu y_3^{N_{1'}} + M M_4 v_H y_5^{\nu^2} - v_H v_\phi^2 (y_2^{N_1})^2 y_5^{\nu^2})}{M_4 v_\phi^2 (y_3^{N_{1'}})^2}, \\
 y' & = \frac{v_H (-M_4 v_\eta y_3^\nu + v_H v_\phi y_2^{N_1} y_4^{\nu^3}) y_5^{\nu^2}}{M_4 v_\phi y_3^{N_{1'}}}, \\
 z' & = \frac{-v_h^2 (y_4^{\nu^3})^2}{M_4}.
 \end{aligned} \quad (5.29)$$

5.2.3 Phenomenology

These textures have been extensively studied [173, 183–198]. Here for completeness we present some results and give an estimate of the radiative correction for the neutrinoless double beta decay effective mass parameter. In order to do this, the first step is to notice that the neutrino mass in the flavor basis is related to the mass matrix in the mass basis by a unitary transformation U , *i.e.* $m_\nu = UDU^T$, where $D = \text{diag}(m_1, m_2, m_3)$ is the diagonal neutrino mass matrix. We use the PDG parametrization given in Eq. 5.17 and define the masses including the corresponding Majorana phases as $\mu_1 \equiv m_1$, $\mu_2 \equiv e^{i\alpha_{21}}m_2$, $\mu_3 \equiv e^{i\alpha_{31}}m_3$. Using the texture zeros we solve for the masses μ_2 and μ_3 in terms of $(\mu_1, \sin^2 \theta_{12}, \sin^2 \theta_{23}, \sin^2 \theta_{13}, \delta)$, see for instance [183, 184]. Then we compute the differences Δm_{12}^2 and Δm_{23}^2 , both will be proportional to μ_1 . Taking the ratio will give us something that is independent of μ_1 and is only a function of the parameters measured by experiments. The next step is to draw random points for $(\sin^2 \theta_{12}, \sin^2 \theta_{23}, \sin^2 \theta_{13})$ and predict the CP oscillation phase at each draw. Once we get the CP violating phase we can substitute its value in the functions Δm_{12}^2 or Δm_{23}^2 and solve for the mass μ_1 . Finally we check if our predictions are consistent with the neutrino allowed regions given by global neutrino oscillation fits [199, 200]. For the type A_1 and A_2 , the lightest neutrino mass we get is between $m_1 = [0.004, 0.008]eV$ and $m_1 = [0.003, 0.008]eV$ respectively. At leading order in θ_{13} one can easily extract the Majorana phases

$$\mu_2 \simeq -\mu_1 \cot^2 \theta_{12} \quad \mu_3 \simeq \pm \mu_1 e^{i\delta \frac{\cot \theta_{12} \cot \theta_{23}}{\sin \theta_{13}}} \quad , \quad (5.30)$$

where the $+$ is for A_1 and the $-$ for A_2 , then the Majorana phases are simply $\alpha_{21} = \pi$ and $\alpha_{31} = \delta$ ($\alpha_{31} = \delta + \pi$). An interesting property of the A_1 and A_2 matrices obtained here is that the neutrinoless double beta decay effective mass parameter is predicted to be zero at tree level. In any case, it should be pointed out that loop corrections like the one shown in Fig. 5.3 give a non-zero (but small) contribution to m_{ee} . The one loop contribution is not

computed exactly, instead of it, we estimate the bound with a few assumptions. We assume that the Yukawa couplings in the theory are of the same order of magnitude. The one loop diagram that contributes to m_{ee} is just the one shown in Fig. 5.3 (which depends on the coupling y_1'). Other possible diagrams such as one with external Higgs singlet of A_4 are forbidden by the flavor symmetry. Thus to estimate it, we use the neutrino matrix element¹³ which has vevs and Yukawas similar to the ones that appear in the one loop contribution in Fig. 5.3. We have performed the scan in our parameter space, and taken the maximal value for the neutrino matrix element in the three sigma scan and this value is divided by the usual one loop suppression factor $(4\pi)^2$. For A_1 we have $m_{e\tau} < 0.012eV$, thus we get $m_{ee}(0.012 \times (\frac{1}{4\pi})^2 = 8 \times 10^{-5})$ eV, which is well below current experimental limits [201]. For A_2 we get exactly the same constraint.

5.2.4 Other textures: B_1 and B_2

Finally, it is important to mention that if we break A_4 to Z_3 , in many cases the RH neutrino mass matrices are non invertible while some of them give only two non-vanishing light neutrino mass elements and therefore the phenomenology for masses and mixings are trivial and non compatible with the observations. Nevertheless, if we include another scalar field Δ which is triplet under $SU(2)_L$ and a singlet under A_4 , just as in [177] we are able to obtain two more texture zeros, namely B_1 and B_2 . For example, B_2 can be obtained if we use the irrep assignment $L_e = \mathbf{1}$, $L_\mu = \mathbf{1}''$, $L_\tau = \mathbf{1}'$, $N_4 = \mathbf{1}'$, $N_5 = \mathbf{1}''$ and $\Delta = \mathbf{1}'$. Thus a simple extension of our model allows to obtain six of the seven texture zeros. In these scenarios the DM stability is lost since the charged leptons transform also non-trivially under the residual Z_3 symmetry.

¹³ $m_{e\tau}$ ($m_{e\mu}$) for A_1 (A_2) texture.

5.2.5 Other works in the same direction.

I also worked with a left right symmetric model [202]. I will not show the details of the model or the theory behind it in this text since I want to keep it as compact as possible, and my contribution was a bit less than for the other three papers. I will explain the idea, but the interested reader is invited to look through the references. In left right symmetric models, the Standard Model is extended from $SU(2)_L \times U(1)_Y$ to $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$. In our model, we extended this, and included two discrete symmetries, D_4 and Z_2 . After successive spontaneous symmetry breaking, neutrino masses will emerge. At this level, an important difference with the respect to the A_4 model that I have just presented in this chapter, is the fact that in this case the charged leptons mass matrix has to be diagonalized to. When this is done, the measured PMNS matrix is not only the matrix that diagonalizes the neutrino mass matrix but a multiplication of both, the one that diagonalizes the charged leptons and the one that diagonalizes the neutrinos. The orthogonal matrix that diagonalizes the lepton mass matrix will not be unique in this model. This freedom leads to four solutions for the PMNS matrix. We found that two of these solutions return phenomenological available values for the neutrino oscillation parameters, the neutrino-less double beta decay and the process $\mu \rightarrow e + \gamma$.

Chapter Six

Conclusions

It is clear that even though the electroweak sector of the standard model was formulated many years ago, it is still an exciting and incredibly interesting subject to do research. The study of this subject made the work of this thesis to cover many topics, which (at first sight) might look disconnected. From Bayesian statistics, to the weak mixing angle at low energies, to models beyond the Standard Model. A total of five papers were published from this work [45, 71, 107, 168, 202].

In [45] we proposed a different way to compute the PDG scaling. The methodology is based on methods used in meta analysis. Essentially, a Bayesian Hierarchical method is used. In this way, we can be agnostic about each of the measurements. After that, a posterior distribution for the hyper-parameters of the model is obtained. The asymptotic behavior of this distribution is also studied. We have shown that in the large N limit, the Bayesian hierarchical model replicates the PDG scaling, while for low N more scaling is obtained. This method does not suffer from other problems of the PDG methodology such as the effects of the order of averaging in the final result. We also applied the methodology to the neutron lifetime, where there is a well known difference between the bottle and beam experiments. Due to the number of experiments we get a similar value to the weighted average suggested by the PDG. The model can be easily extended to experiments that are correlated. Another possible future line of research is to use a hierarchical model to impose

bounds on the Standard Model Effective Field theory. This should be possible, given that some families of Wilson coefficients might be similar and could have common origins. This was actually how we started this project, but at some point we realized out the possibility to use it for the PDG averages.

In [107] we computed the uncertainty in the running of the weak mixing angle in the $\overline{\text{MS}}$ scheme from the Z pole to low energies. In this work, an experimental separation of the strange quark contribution was used, instead of the the rougher theoretical estimate used in previous works. We also included results from lattice QCD for both, the strange quark contribution and the disconnected diagrams that appear at three loop order. The formulas for the running and matching were updated to the next loop order. We also updated the formula that gives the relationship between the experimental cross section for the process $e^+ + e^- \rightarrow \text{hadrons}$ and the fine structure constant in the $\overline{\text{MS}}$ scheme. The combination of these improvements reduced the uncertainty of the weak mixing angle at low energies by almost a factor of four. This reduction will allow a better comparison between theory and experiment, which can help in the search of physics beyond the Standard Model. This will be particularly important for several new experiments such as MOLLER, since the predicted sensitivity of $\delta(\sin^2 \theta_W) = \pm 0.00028$ will be close to LEP and LHC, and with our work, the theoretical error is an order of magnitude below the experimental one. We also made significant progress in the computation of the running of the fine structure constant α in the $\overline{\text{MS}}$ scheme. The plot of the running of the weak mixing angle was published in the 2018 version of the PDG.

In [168] we show how to generate neutrino mass texture zeroes from a discrete symmetry, A_1 and A_2 . We computed the experimental constraints that these matrices induce for the neutrino oscillation parameters, i.e the masses, the mixings, and the CP phases (Dirac and Majorana). The model, due to the residual Z_2 , gives a possible way to relate neutrino phenomenology with dark matter stability. The dark matter phenomenological constraints can be seen in one of my collaborators previous work [178]. In [202], we used a left right model

with discrete symmetries on the top of it. After going to the mass basis for all particles, we get four different solutions for the correlations and values of the neutrino oscillation parameters. Two of them require no tuning to be compatible with neutrino oscillations constraints. We also obtain non zero neutrino-less double beta decay which is compatible with experimental constraints. Our results are also compatible with the constraints from the process $\mu \rightarrow e + \gamma$.

There are many possible future lines of research that can follow the work presented here. For example, it might be interesting to study with detail all the relations and differences between the different schemes used for $\hat{\alpha}$. The same applies for the weak mixing angle. Another future line of research is to apply the statistical methods used here to perform global analysis of electroweak precision physics.

With respect to neutrino physics, there is an ongoing project where we have an extra U(1) with kinetic mixing and discrete symmetries. The idea is to perform an analysis that includes the SMEFT formalism to put constraints on the model in a global way.

To conclude, I would like to say electroweak physics turned out to be a delightful topic. I certainly learned a lot, but I also learned that there is still much more to be done.

APPENDICES

Appendix A

Fisher's Information

The proof is fairly easy and involves only basic calculus. First we use

$$\int p(y|\theta)dy = 1, \tag{A.1}$$

and we derive this equation by θ , thus

$$\int \frac{dp(y|\theta)}{d\theta} dy = 0, \tag{A.2}$$

with a bit of rearrangement we get

$$\int \frac{dp(y|\theta)/p(y|\theta)}{d\theta} p(y|\theta) dy = 0, \tag{A.3}$$

or

$$\int \frac{d \ln(p(y|\theta))}{d\theta} p(y|\theta) dy = 0. \tag{A.4}$$

Taking the derivative of this equation

$$\int \frac{d^2 \ln(p(y|\theta))}{d\theta^2} p(y|\theta) dy + \int \frac{d \ln(p(y|\theta))}{d\theta} \frac{dp(y|\theta)}{d\theta} dy = 0. \tag{A.5}$$

or

$$\int \frac{d^2 \ln(p(y|\theta))}{d\theta^2} p(y|\theta) dy + \int \left(\frac{d \ln(p(y|\theta))}{d\theta} \right)^2 p(y|\theta) dy = 0. \tag{A.6}$$

The second term on the left hand side is the Fisher's information. So we have proved that

$$\mathcal{I}(\theta) = - \int \frac{d^2 \ln(p(y|\theta))}{d\theta^2} p(y|\theta) dy. \quad (\text{A.7})$$

Appendix B

Coefficients of the RGE

The running of the weak mixing angle is

$$\mu^2 \frac{d\hat{s}^2}{d\mu^2} = \frac{\hat{\alpha}}{\pi} \left[\frac{1}{24} \sum_i K_i \gamma_i (Q_i^2 \hat{s}^2 - T_i Q_i) + \sigma \hat{s}^2 \left(\sum_q Q_q \right)^2 - \frac{\sigma}{2} \left(\sum_q T_q \right) \left(\sum_q Q_q \right) \right], \quad (\text{B.1})$$

on the other hand the running of the fine structure constant is

$$\mu^2 \frac{d\hat{\alpha}^2}{d\mu^2} = \frac{\hat{\alpha}^2}{\pi} \left[\frac{1}{24} \sum_i K_i \gamma_i Q_i^2 + \sigma \left(\sum_q Q_q \right)^2 \right], \quad (\text{B.2})$$

combining both equations we obtain

$$\mu^2 \frac{d}{d\mu^2} \left(\frac{\hat{s}^2}{\hat{\alpha}} \right) = -\frac{1}{24\pi} \sum_i K_i \gamma_i T_i Q_i - \frac{\sigma}{2\pi} \left(\sum_q T_q \right) \left(\sum_q Q_q \right), \quad (\text{B.3})$$

this is an equation that measures the difference between the running of the weak angle and the fine structure constant. This equation can be integrated, if one proposes the following solution,

$$\mu^2 \frac{d}{d\mu^2} \left(\frac{\hat{s}^2}{\hat{\alpha}} \right) = \frac{\lambda_2}{3\pi} + \lambda_1 \mu^2 \frac{d}{d\mu^2} \left(\frac{1}{\hat{\alpha}} \right) + \frac{3\lambda_3}{4\pi} \mu^2 \frac{d \ln \hat{\alpha}}{d\mu^2} + \mu^2 \frac{d}{d\mu^2} \left(\frac{\tilde{\sigma}}{\pi} \right), \quad (\text{B.4})$$

for unknown constants λ_1 , λ_2 and λ_3 that depend on the number of particles in the theory.

We can use again the RGE for $\hat{\alpha}$ to get values of these constants, then we must have

$$-\frac{1}{24\pi} \sum_i K_i \gamma_i T_i Q_i - \frac{\sigma}{2\pi} \left(\sum_q T_q \right) \left(\sum_q Q_q \right) = \frac{\lambda_2}{3\pi} - \frac{\lambda_1}{\hat{\alpha}^2} \mu^2 \frac{d\hat{\alpha}}{d\mu^2} + \frac{3\lambda_3}{4\pi} \frac{1}{\hat{\alpha}} \mu^2 \frac{d\hat{\alpha}}{d\mu^2} + \mu^2 \frac{d}{d\mu^2} \left(\frac{\tilde{\sigma}}{\pi} \right), \quad (\text{B.5})$$

inserting the RGE of $\hat{\alpha}$ we get

$$\begin{aligned} -\frac{1}{24\pi} \sum_i K_i \gamma_i T_i Q_i - \frac{\sigma}{2\pi} \left(\sum_q T_q \right) \left(\sum_q Q_q \right) &= \frac{\lambda_2}{3\pi} - \frac{\lambda_1}{\pi} \left[\frac{1}{24} \sum_i K_i \gamma_i Q_i^2 + \sigma \left(\sum_q Q_q \right)^2 \right] \\ &+ \frac{3\lambda_3}{4\pi^2} \hat{\alpha} \left[\frac{1}{24} \sum_i K_i \gamma_i Q_i^2 + \sigma \left(\sum_q Q_q \right)^2 \right] + \mu^2 \frac{d}{d\mu^2} \left(\frac{\tilde{\sigma}}{\pi} \right), \end{aligned} \quad (\text{B.6})$$

we fix λ_1 in such a way that it absorbs all the $\hat{\alpha}_s$ connected dependence in the RGE. Thus

$$-\frac{1}{24\pi} \sum_i T_i Q_q \frac{1}{2} = -\frac{\lambda_1}{\pi} \left[\frac{1}{24} \sum_q Q_q^2 \right], \quad (\text{B.7})$$

this reduces to

$$\lambda_1 = \frac{\sum_i T_i Q_q}{2 \sum_q Q_q^2}, \quad (\text{B.8})$$

there will remain terms that are constants, other ones that depend on $\hat{\alpha}$ and some others proportional to $\hat{\alpha}_s^3$ and come from the singlet. The constant terms can be easily obtained,

$$\lambda_2 = \frac{1}{8} \sum_{i \neq q} \gamma_i \left(\lambda_1 Q_i^2 - T_i Q_i \right), \quad (\text{B.9})$$

for the terms proportional to $\hat{\alpha}$ we would have

$$-\frac{1}{32\pi} \frac{\hat{\alpha}}{\pi} \sum_{i \neq W} N_i^c \gamma_i T_i Q_i^3 = -\frac{\lambda_1}{\pi} \left[\frac{1}{24} \frac{3}{4} \frac{\hat{\alpha}}{\pi} \sum_{i \neq W} N_i^c \gamma_i Q_i^4 \right] + \frac{3\lambda_3}{4\pi} \frac{\hat{\alpha}}{\pi} \left[\frac{1}{24} \sum_i N_i^c \gamma_i Q_i^2 \right], \quad (\text{B.10})$$

$$\lambda_3 = \frac{\sum_{i \neq W} N_i^c \gamma_i Q_i^2 [\lambda_1 Q_i^2 - T_i Q_i]}{\sum_i N_i^c \gamma_i Q_i^2}, \quad (\text{B.11})$$

B.1 Calculations of $\alpha(M_Z^2)$

Three independent groups presented recent evaluations of the hadronic contribution to the scale dependence of α . In this appendix we briefly compare their approaches and results.

In the Adler function approach [97, 134], one uses the relations,

$$\frac{D(Q^2)}{Q^2} \equiv 12\pi^2 \frac{d\Pi(q^2)}{dq^2} = -\frac{3\pi}{\alpha} \frac{d}{dq^2} \Delta_{\text{had}}(q^2) = \int_{4m^2}^{\infty} \frac{R(s)}{(s+Q^2)^2} ds, \quad (\text{B.12})$$

where $Q^2 = -q^2$, and where the dispersion integral in the latter expression can be used to implement experimental data up to some cut-off M_0 . One can then write,

$$\begin{aligned} \Delta_{\text{had}}\alpha^{(5)}(M_Z^2) &= \Delta_{\text{had}}\alpha^{(5)}(-M_0^2)_{\text{data}} + [\Delta_{\text{had}}\alpha^{(5)}(-M_Z^2) - \Delta_{\text{had}}\alpha^{(5)}(-M_0^2)]_{\text{PQCD}} \\ &+ [\Delta_{\text{had}}\alpha^{(5)}(M_Z^2) - \Delta_{\text{had}}\alpha^{(5)}(-M_Z^2)]_{\text{PQCD}}, \end{aligned} \quad (\text{B.13})$$

where the last two terms are computed using the operator product expansion (OPE) of $R(s)$, i.e., including the leading non-perturbative condensate corrections. Demanding consistency with the OPE of the Adler function itself suggests that a value of M_0 as low as 2 GeV appears to be a safe choice. Using this approach implies [97] for the on-shell definition,

$$\alpha(M_Z^2)^{-1} = 128.958 \pm 0.016. \quad (\text{B.14})$$

The approach of Ref. [203] is mostly data driven. Experimental data were used up to 11.09 GeV (except for the interval between 2.6 GeV and 3.73 GeV) and PQCD beyond that. The dispersion relation (4.12) then implied,

$$\alpha(M_Z^2)^{-1} = 128.944 \pm 0.019. \quad (\text{B.15})$$

Similarly, Ref. [122] uses data up to only 5 GeV (except for the interval between 1.8 GeV

and 3.7 GeV), with the result,

$$\alpha(M_Z^2)^{-1} = 128.947 \pm 0.012 . \quad (\text{B.16})$$

Here, we rely on the data handling of this work as it includes much more recent data than Ref. [203]. Moreover, the breakdown of individual channels and energy ranges is more explicit compared to Ref. [97].

Finally, changing our own result, with $\hat{\alpha}(M_Z^2)^{-1} = 127.959 \pm 0.010$, based on the direct application of the renormalization group and matching equations and including τ decay data, from the $\overline{\text{MS}}$ scheme to the on-shell scheme including the top quark contribution, we find,

$$\alpha(M_Z^2)^{-1} = 128.949 \pm 0.010 . \quad (\text{B.17})$$

The numerical difference of our result to Ref. [122] arises mostly from the different¹ value of α_s and our treatment of the charm quark contribution [137]. Thus, in view of the rather different approaches and differences in data sets, all numerical results are in good agreement with each other.

B.2 Calculation of $\Delta_{\text{disc}}\hat{\alpha}$

In the on-shell scheme one has [204],

$$\Delta_{\text{disc}}\alpha(q) = 4\pi\alpha \text{Re} \left[\Pi(q^2) - \Pi(0) \right]_{\text{disc}} , \quad (\text{B.18})$$

where,

$$\left[\Pi(q^2) - \Pi(0) \right]_{\text{disc}} = \sum_{t=0}^T \left[\frac{\cos(qt) - 1}{q^2} + \frac{t^2}{2} \right] C(t). \quad (\text{B.19})$$

¹The three groups use slightly different values for α_s , but this amounts to difference below the level of 0.004 in $\alpha(M_Z)^{-1}$.

$C(t)$ has been computed [129] in units set by the lattice cut-off scale $a^{-1} = 1.73$ GeV. To obtain Eq. (4.28), we plotted $\Delta_{\text{disc}}\alpha(q)$ as a function of T and observe a plateau near $T = 20$, which closely mirrors the result for the case of a_μ . The value of the plateau is interpreted as the physical value [129]. As an independent check we compute the ratio $\rho(a_\mu)$ of the disconnected contribution to the anomalous magnetic moment [129], $a_\mu^{\text{disc}} = -9.6 \times 10^{-10}$, to the total hadronic contribution [122] for energies up to 1.8 GeV, obtaining $\rho(a_\mu) = -0.015$. The integration kernel of a_μ enhances contributions from low q^2 momenta, and recalling that $Q_u + Q_d + Q_s = 0$, the disconnected piece also predominantly arises from such momenta. On the other hand, the integration kernel for $\Delta\alpha$ has greater support at higher scales compared to a_μ , so that $\rho(a_\mu)$ should imply an upper bound on the disconnected contribution to $\Delta\alpha$. Numerically,

$$|\Delta_{\text{disc}}\alpha(1.8 \text{ GeV})| < |\rho(a_\mu) \times \Delta_{\text{had}}\alpha(1.8 \text{ GeV})| = 8.3 \times 10^{-5}, \quad (\text{B.20})$$

where $\Delta_{\text{had}}\alpha(1.8 \text{ GeV}) = 55.26 \times 10^{-4}$. This confirms the finding in Eq. (4.28) that $\Delta_{\text{disc}}\alpha$ is very small.

Bibliography

- [1] J. J. Thomson. “XL. Cathode Rays”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 44.269 (1897), pp. 293–316.
- [2] E. Rutherford. “Collision of α particles with light atoms, an anomalous effect in nitrogen”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 37.222 (1919), pp. 581–587.
- [3] J. Chadwick. “Possible Existence of a Neutron”. In: *Nature* 129 (1932), p. 312.
- [4] C. D Anderson. “The apparent existence of easily deflectable positives”. In: *Science (New York, N.Y.)* 76 (Oct. 1932), pp. 238–9.
- [5] P. Dirac and R. H. Fowler. “The quantum theory of the electron”. In: *Proc. R. Soc. Lond A*.117 (1928).
- [6] V. Shiltsev. “A Phenomenological Cost Model for High Energy Particle Accelerators”. In: *JINST* 9 (2014), T07002.
- [7] A. Einstein. “On a Heuristic Viewpoint Concerning the Production and Transformation of Light”. In: *Annalen der Physik* 7 (1905), pp. 132–148.
- [8] W. Heisenberg. “Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen”. In: *Z. Physik* 33 (1925), p. 879.
- [9] W. Pauli. “Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren”. In: *Z. Physik* 31 (1925), p. 765.

-
- [10] E. Schrödinger. “An Undulatory Theory of the Mechanics of Atoms and Molecules”. In: *Phys. Rev.* 28 (6 Dec. 1926), pp. 1049–1070.
- [11] J. Schwinger. “On Quantum-Electrodynamics and the Magnetic Moment of the Electron”. In: *Phys. Rev.* 73 (4 Feb. 1948), pp. 416–417.
- [12] P. Kusch and H. M. Foley. “Precision Measurement of the Ratio of the Atomic ‘ g Values’ in the ${}^2P_{\frac{3}{2}}$ and ${}^2P_{\frac{1}{2}}$ States of Gallium”. In: *Phys. Rev.* 72 (12 Dec. 1947), pp. 1256–1257.
- [13] E. Fermi. “Tentativo di una teoria dell’emissione dei raggi beta”. In: *Ric. Sci.* 4 (1933), pp. 491–495.
- [14] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266.
- [15] A. Salam and J.C. Ward. “Electromagnetic and weak interactions”. In: *Physics Letters* 13.2 (1964), pp. 168–171. ISSN: 0031-9163.
- [16] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415.
- [17] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509.
- [18] B. Pontecorvo. “Neutrino Experiments and the Problem of Conservation of Leptonic Charge”. In: *Sov. Phys. JETP* 26 (1968). [Zh. Eksp. Teor. Fiz.53,1717(1967)], pp. 984–988.
- [19] Y. Fukuda et al. “Evidence for oscillation of atmospheric neutrinos”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003 [hep-ex].

-
- [20] G. 't Hooft. “Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking”. In: *NATO Sci. Ser. B* 59 (1980), pp. 135–157. DOI: 10.1007/978-1-4684-7571-5_9.
- [21] L. Baudis. “The Search for Dark Matter”. In: *European Review* 26.1 (2018), pp. 70–81.
- [22] S. P. Martin. “A Supersymmetry primer”. In: (1997). [Adv. Ser. Direct. High Energy Phys.18,1(1998)], pp. 1–98. arXiv: hep-ph/9709356 [hep-ph].
- [23] S. Weinberg. *The quantum theory of fields. Vol. 3: Supersymmetry*. Cambridge, UK: Univ. Pr. (2000) 419 p. 2000.
- [24] T. Mohaupt. “Introduction to string theory”. In: *Lect. Notes Phys.* 631 (2003). [173(2002)], pp. 173–251. arXiv: hep-th/0207249 [hep-th].
- [25] M.J. Ramsey-Musolf. “Low-energy parity violation and new physics”. In: *Phys. Rev. C* 60 (1999), p. 015501. DOI: 10.1103/PhysRevC.60.015501. arXiv: hep-ph/9903264.
- [26] K.S. Kumar et al. “Low Energy Measurements of the Weak Mixing Angle”. In: *Ann. Rev. Nucl. Part. Sci.* 63 (2013), pp. 237–267. DOI: 10.1146/annurev-nucl-102212-170556. arXiv: 1302.6263 [hep-ex].
- [27] H. Davoudiasl, H. S. Lee, and W. J. Marciano. “Muon Anomaly and Dark Parity Violation”. In: *Phys. Rev. Lett.* 109 (2012), p. 031802. DOI: 10.1103/PhysRevLett.109.031802. arXiv: 1205.2709 [hep-ph].
- [28] H. Davoudiasl, H.S. Lee, and W. J. Marciano. “‘Dark’ Z implications for Parity Violation, Rare Meson Decays, and Higgs Physics”. In: *Phys. Rev. D* 85 (2012), p. 115019. DOI: 10.1103/PhysRevD.85.115019. arXiv: 1203.2947 [hep-ph].
- [29] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001.

-
- [30] G. D’Agostini. “On a curious bias arising when the $\sqrt{\chi^2/\nu}$ scaling prescription is first applied to a sub-sample of the individual results”. In: (2020). arXiv: 2001.07562 [physics.data-an].
- [31] S. E. Fienberg. “When did Bayesian inference become "Bayesian"?” In: *Bayesian Anal.* 1.1 (2006), pp. 1–40.
- [32] T. Bayes and N. Price. “LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S”. In: *Philosophical Transactions of the Royal Society of London* 53 (1763), pp. 370–418. DOI: 10.1098/rstl.1763.0053.
- [33] K. Chaloner and I. Verdinelli. “Bayesian Experimental Design: A Review”. In: *Statist. Sci.* 10.3 (Aug. 1995), pp. 273–304.
- [34] G. D’Agostini. *Bayesian Reasoning in Data Analysis: A Critical Introduction*. Singapore: World Scientific, 2003. DOI: 10.1142/5262. URL: <https://cds.cern.ch/record/642515>.
- [35] A. Gelman et al. *Bayesian Data Analysis*. 2nd ed. Chapman and Hall/CRC, 2004.
- [36] B. Moghaddam, T. Jebara, and A. Pentland. “Bayesian face recognition”. In: *Pattern Recognition* 33.11 (2000), pp. 1771–1782.
- [37] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. 1st ed. Springer, 2007. ISBN: 0387310738.
- [38] L. Le Cam. *On some asymptotic properties of maximum likelihood estimates and related Bayes estimates*. Vol. 1. University of California Publications in Statistics, 1953, pp. 277–330.
- [39] L. Le Cam and G. L. Yang. *Asymptotics in Statistics: Some Basic Concepts*. Springer, 1990.

-
- [40] H. Jeffreys. “An invariant form for the prior probability in estimation problems”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186.1007 (1946), pp. 453–461.
- [41] H. Jeffreys. *Theory of Probability*. third edition. Oxford University Press, J, 1961.
- [42] W. Feller. “The fundamental limit theorems in probability”. In: *Bull. Amer. Math. Soc.* 51.11 (Nov. 1945), pp. 800–832.
- [43] M. R Spiegel. *Theory And Problems Of Probability And Statistics (Schaum S Outline Series)*. McGraw-Hill Education, 2003, pp. 112–113.
- [44] R. Liu et al. “On Divergence Measures Leading to Jeffreys and Other Reference Priors”. In: *Bayesian Analysis* 9 (June 2014), pp. 331–370.
- [45] J. Erler and R. Ferro-Hernández. “Alternative to the application of PDG scale factors”. In: *Eur. Phys. J. C* 80.6 (2020), p. 541. DOI: 10.1140/epjc/s10052-020-8115-3. arXiv: 2004.01219 [physics.data-an].
- [46] R. E. Tarone. “The Use of Historical Control Information in Testing for a Trend in Proportions”. In: *Biometrics* 38.1 (1982), pp. 215–220. ISSN: 0006341X, 15410420. URL: <http://www.jstor.org/stable/2530304>.
- [47] A. P. Dempster, R. S. Murray, and B. J. Weeks. “Combining Historical and Randomized Controls for Assessing Trends in Proportions”. In: *Journal of the American Statistical Association* 78.382 (1983), pp. 221–227.
- [48] L. Hedges and I. Olkin. “Statistical Methods in Meta-Analysis”. In: vol. 20. Jan. 1985. DOI: 10.2307/1164953.
- [49] P.S Laplace. “Memoir on the Probability of the Causes of Events”. In: *Statistical Science* 1.3 (1986), pp. 364–378. ISSN: 08834237. URL: <http://www.jstor.org/stable/2245476>.

-
- [50] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533.
- [51] M. Kobayashi and T. Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X.
- [52] R. P. Feynman and M. Gell-Mann. “Theory of the Fermi Interaction”. In: *Phys. Rev.* 109 (1 Jan. 1958), pp. 193–198. DOI: 10.1103/PhysRev.109.193. URL: <https://link.aps.org/doi/10.1103/PhysRev.109.193>.
- [53] S. S. Gerstein and Y. B. Zeldovich. “Meson Corrections in the Theory of Beta Decay”. In: *Sov. Phys. JETP* 2 (1956).
- [54] B. Krause. *Conserved Vector Current (CVC)*. 1996. URL: http://pibeta.phys.virginia.edu/docs/publications/ketevi_diss (visited on 02/23/2020).
- [55] J.C. Hardy et al. “Superallowed $0^+ \rightarrow 0^+$ nuclear -decays: A critical survey with tests of CVC and the standard model”. In: *Nuclear Physics A* 509.3 (1990), pp. 429–460. ISSN: 0375-9474.
- [56] J. C. Hardy and I. S. Towner. “Superallowed $0^+ \rightarrow 0^+$ nuclear decays: 2014 critical survey, with precise results for V_{ud} and CKM unitarity”. In: *Phys. Rev.* C91.2 (2015), p. 025501. DOI: 10.1103/PhysRevC.91.025501. arXiv: 1411.5987 [nucl-ex].
- [57] M. R. Dunlop et al. “High-Precision Half-Life Measurements for the Superallowed β^+ Emitter ^{10}C : Implications for Weak Scalar Currents”. In: *Phys. Rev. Lett.* 116 (17 Apr. 2016), p. 172501.
- [58] N. Severijns, M. Beck, and O. Naviliat-Cuncic. “Tests of the standard electroweak model in nuclear beta decay”. In: *Rev. Mod. Phys.* 78 (3 Sept. 2006), pp. 991–1040.

- [59] V. Cirigliano, S. Gardner, and B.R. Holstein. “Beta decays and non-standard interactions in the LHC era”. In: *Progress in Particle and Nuclear Physics* 71 (2013). Fundamental Symmetries in the Era of the LHC, pp. 93–118. ISSN: 0146-6410.
- [60] F. E. Wietfeldt. “Measurements of the Neutron Lifetime”. In: *Atoms* 6.4 (2018). ISSN: 2218-2004.
- [61] P. E. Spivak. “Neutron Lifetime from Atomic-Energy-Institute experiment”. In: *Sov. Phys. JETP* 67 (1988), pp. 1735–1740.
- [62] J. Byrne et al. “A revised value for the neutron lifetime measured using a Penning trap”. In: *Europhysics Letters (EPL)* 33.3 (Jan. 1996), pp. 187–192.
- [63] A. T. Yue et al. “Improved Determination of the Neutron Lifetime”. In: *Phys. Rev. Lett.* 111 (22 Nov. 2013), p. 222501.
- [64] A. Serebrov et al. “Measurement of the neutron lifetime using a gravitational trap and a low-temperature Fomblin coating”. In: *Phys. Lett. B* 605 (2005), pp. 72–78.
- [65] A. Pichlmaier et al. “Neutron lifetime measurement with the UCN trap-in-trap MAMBO II”. In: *Phys. Lett. B* 693 (2010), pp. 221–226.
- [66] A. Steyerl et al. “Quasielastic scattering in the interaction of ultracold neutrons with a liquid wall and application in a reanalysis of the Mambo I neutron-lifetime experiment”. In: *Phys. Rev. C* 85 (2012), p. 065503.
- [67] A. P. Serebrov et al. “Neutron lifetime measurements with a large gravitational trap for ultracold neutrons”. In: *Phys. Rev. C* 97.5 (2018), p. 055503.
- [68] V. F. Ezhov et al. “Measurement of the neutron lifetime with ultra-cold neutrons stored in a magneto-gravitational trap”. In: *JETP Lett.* 107.11 (2018), pp. 671–675.
- [69] R. W. Pattie Jr. et al. “Measurement of the neutron lifetime using a magneto-gravitational trap and in situ detection”. In: *Science* 360.6389 (2018), pp. 627–632.

-
- [70] G. D’Agostini. “Skeptical combination of experimental results using JAGS/rjags with application to the K^\pm mass determination”. In: (2020). arXiv: 2001.03466.
- [71] J. Benel Mogrovejo, A. Courtoy, and R. Ferro-Hernandez. “A constrained fit of the valence transversity distributions from dihadron production”. In: *The European Physical Journal C* 80 (May 2020). DOI: 10.1140/epjc/s10052-020-8039-y.
- [72] C. N. Yang and R. L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. In: *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191. URL: <https://link.aps.org/doi/10.1103/PhysRev.96.191>.
- [73] S.L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582.
- [74] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [75] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [76] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Mar. 2014. ISBN: 978-1-107-03473-0, 978-1-107-03473-0.
- [77] S. Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, June 2005. ISBN: 978-0-521-67053-1, 978-0-511-25204-4.
- [78] S. Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, Aug. 2013. ISBN: 978-1-139-63247-8, 978-0-521-67054-8, 978-0-521-55002-4.

-
- [79] S. Weinberg. “Physical Processes in a Convergent Theory of the Weak and Electromagnetic Interactions”. In: *Phys. Rev. Lett.* 27 (24 Dec. 1971), pp. 1688–1691. DOI: 10.1103/PhysRevLett.27.1688. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.27.1688>.
- [80] T. Nakano and K. Nishijima. “Charge Independence for V-particles*[†]”. In: *Progress of Theoretical Physics* 10.5 (Nov. 1953), pp. 581–582. ISSN: 0033-068X. DOI: 10.1143/PTP.10.581. eprint: <https://academic.oup.com/ptp/article-pdf/10/5/581/5364926/10-5-581.pdf>. URL: <https://doi.org/10.1143/PTP.10.581>.
- [81] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.
- [82] J. C. Ward. “An Identity in Quantum Electrodynamics”. In: *Phys. Rev.* 78 (2 Apr. 1950), pp. 182–182. DOI: 10.1103/PhysRev.78.182. URL: <https://link.aps.org/doi/10.1103/PhysRev.78.182>.
- [83] Y. Takahashi. “On the generalized Ward identity”. In: *Nuovo Cim.* 6 (1957), p. 371. DOI: 10.1007/BF02832514.
- [84] A. V. Manohar. “Introduction to Effective Field Theories”. In: *Les Houches summer school: EFT in Particle Physics and Cosmology*. Apr. 2018. arXiv: 1804.05863 [hep-ph].
- [85] A. Grozin. “Lectures on QED and QCD”. In: *3rd Dubna International Advanced School of Theoretical Physics*. Aug. 2005, pp. 1–156. arXiv: hep-ph/0508242.
- [86] D. J. Gross and F. Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.

-
- [87] P.A. Baikov, K.G. Chetyrkin, and J.H. Kühn. “Five-Loop Running of the QCD coupling constant”. In: *Phys. Rev. Lett.* 118.8 (2017), p. 082002. DOI: 10.1103/PhysRevLett.118.082002. arXiv: 1606.08659 [hep-ph].
- [88] Y. Schroder and M. Steinhauser. “Four-loop decoupling relations for the strong coupling”. In: *JHEP* 01 (2006), p. 051. DOI: 10.1088/1126-6708/2006/01/051. arXiv: hep-ph/0512058.
- [89] K.G. Chetyrkin, Johann H. Kuhn, and Christian Sturm. “QCD decoupling at four loops”. In: *Nucl. Phys. B* 744 (2006), pp. 121–135. DOI: 10.1016/j.nuclphysb.2006.03.020. arXiv: hep-ph/0512060.
- [90] L.R. Trujillo-Cruz. “Incertidumbres en la evolucion de la constante de acoplamiento de QCD”. MA thesis. UNAM: Instituto de Fisica (UNAM), Feb. 2014.
- [91] P.A. Baikov et al. “Vector Correlator in Massless QCD at Order $O(\alpha_s^4)$ and the QED beta-function at Five Loop”. In: *JHEP* 07 (2012), p. 017. DOI: 10.1007/JHEP07(2012)017. arXiv: 1206.1284 [hep-ph].
- [92] K.G. Chetyrkin, Bernd A. Kniehl, and M. Steinhauser. “Decoupling relations to (α_s^3) and their connection to low-energy theorems”. In: *Nucl. Phys. B* 510 (1998), pp. 61–87. DOI: 10.1016/S0550-3213(97)00649-4. arXiv: hep-ph/9708255.
- [93] K.G. Chetyrkin, Johann H. Kuhn, and M. Steinhauser. “Three loop polarization function and $O(\alpha_s^2)$ corrections to the production of heavy quarks”. In: *Nucl. Phys. B* 482 (1996), pp. 213–240. DOI: 10.1016/S0550-3213(96)00534-2. arXiv: hep-ph/9606230.
- [94] K.G. Chetyrkin, J. H. Kuhn, and C. Sturm. “Four-loop moments of the heavy quark vacuum polarization function in perturbative QCD”. In: *Eur. Phys. J. C* 48 (2006), pp. 107–110. DOI: 10.1140/epjc/s2006-02610-y. arXiv: hep-ph/0604234.

-
- [95] B. A. Kniehl and A. V. Kotikov. “Heavy-quark QCD vacuum polarisation function: Analytical results at four loops”. In: *Phys. Lett. B* 642 (2006), pp. 68–71. DOI: 10.1016/j.physletb.2006.09.008. arXiv: hep-ph/0607201.
- [96] J. Erler. “Calculation of the QED coupling $\alpha(M(Z))$ in the modified minimal subtraction scheme”. In: *Phys. Rev. D* 59 (1999), p. 054008. DOI: 10.1103/PhysRevD.59.054008. arXiv: hep-ph/9803453 [hep-ph].
- [97] F. Jegerlehner. “Variations on Photon Vacuum Polarization”. In: *EPJ Web Conf.* 218 (2019). Ed. by A. Denig and C.-F. Redmer, p. 01003. DOI: 10.1051/epjconf/201921801003. arXiv: 1711.06089 [hep-ph].
- [98] F. Jegerlehner. “The Running fine structure constant $\alpha(E)$ via the Adler function”. In: *Nucl. Phys. B Proc. Suppl.* 181-182 (2008), pp. 135–140. DOI: 10.1016/j.nuclphysbps.2008.09.010. arXiv: 0807.4206 [hep-ph].
- [99] Paolo Gambino and Alberto Sirlin. “Relation between $\sin^2 \Theta_w(m(z))$ and $\sin^2 \Theta_w^{\text{effective}}(\text{leptonic})$ ”. In: *Phys. Rev. D* 49 (1994), pp. 1160–1162. DOI: 10.1103/PhysRevD.49.R1160. arXiv: hep-ph/9309326.
- [100] A. Sirlin. “Radiative corrections in the $SU(2)_L \times U(1)$ theory: A simple renormalization framework”. In: *Phys. Rev. D* 22 (4 Aug. 1980), pp. 971–981. DOI: 10.1103/PhysRevD.22.971. URL: <https://link.aps.org/doi/10.1103/PhysRevD.22.971>.
- [101] G. Passarino and M. Veltman. “On the definition of the weak mixing angle”. In: *Physics Letters B* 237.3 (1990), pp. 537–544. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(90\)91221-V](https://doi.org/10.1016/0370-2693(90)91221-V).
- [102] Wolfgang Hollik. “Renormalization of the Standard Model”. In: vol. 14. 1995, pp. 37–116. DOI: 10.1142/9789814503662_0003.
- [103] P. Langacker. *Precision Tests of the Standard Electroweak Model*. World Scientific, 1995.

-
- [104] F. Jegerlehner. *Renormalizing the Standard Model*. 2011.
- [105] P. Langacker. *The standard model and beyond*. Series in High Energy Physics, Cosmology and Gravitation. Boca Raton, FL: Taylor and Francis, 2010. URL: <https://cds.cern.ch/record/1226768>.
- [106] J. Erler and M. J. Ramsey-Musolf. “The Weak mixing angle at low energies”. In: *Phys. Rev. D* 72 (2005), p. 073003. DOI: 10.1103/PhysRevD.72.073003. arXiv: hep-ph/0409169.
- [107] J. Erler and R. Ferro-Hernández. “Weak Mixing Angle in the Thomson Limit”. In: *JHEP* 03 (2018), p. 196. DOI: 10.1007/JHEP03(2018)196. arXiv: 1712.09146 [hep-ph].
- [108] D. Androic et al. “First Determination of the Weak Charge of the Proton”. In: *Phys. Rev. Lett.* 111.14 (2013), p. 141803. DOI: 10.1103/PhysRevLett.111.141803. arXiv: 1307.5275 [nucl-ex].
- [109] D. Becker et al. “The P2 experiment”. In: (Feb. 2018). DOI: 10.1140/epja/i2018-12611-6. arXiv: 1802.04759 [nucl-ex].
- [110] J. Benesch et al. “The MOLLER Experiment: An Ultra-Precise Measurement of the Weak Mixing Angle Using Møller Scattering”. In: (Nov. 2014). arXiv: 1411.4088 [nucl-ex].
- [111] P.L. Anthony et al. “Precision measurement of the weak mixing angle in Moller scattering”. In: *Phys. Rev. Lett.* 95 (2005), p. 081601. DOI: 10.1103/PhysRevLett.95.081601. arXiv: hep-ex/0504049.
- [112] D. Wang et al. “Measurement of parity violation in electron–quark scattering”. In: *Nature* 506.7486 (2014), pp. 67–70. DOI: 10.1038/nature12964.

-
- [113] P.A. Souder. “Parity Violation in Deep Inelastic Scattering with the SoLID Spectrometer at JLab”. In: *Int. J. Mod. Phys. Conf. Ser.* 40 (2016). Ed. by Haiyan Gao and Bo-Qiang Ma, p. 1660077. DOI: 10.1142/S2010194516600776.
- [114] G.P. Zeller et al. “A Precise Determination of Electroweak Parameters in Neutrino Nucleon Scattering”. In: *Phys. Rev. Lett.* 88 (2002). [Erratum: *Phys.Rev.Lett.* 90, 239902 (2003)], p. 091802. DOI: 10.1103/PhysRevLett.88.091802. arXiv: hep-ex/0110059.
- [115] B.C. Canas et al. “The weak mixing angle from low energy neutrino measurements: a global update”. In: *Phys. Lett. B* 761 (2016), pp. 450–455. DOI: 10.1016/j.physletb.2016.08.047. arXiv: 1608.02671 [hep-ph].
- [116] C.S. Wood et al. “Measurement of parity nonconservation and an anapole moment in cesium”. In: *Science* 275 (1997), pp. 1759–1763. DOI: 10.1126/science.275.5307.1759.
- [117] L. Willmann et al. “Trapped radioactive isotopes for fundamental symmetry investigations. The TRI μ P Facility”. In: *Hyperfine Interact.* 211 (2012), pp. 39–43. DOI: 10.1007/s10751-012-0603-x.
- [118] J. Erler and S. Su. “The Weak Neutral Current”. In: *Prog. Part. Nucl. Phys.* 71 (2013), pp. 119–149. DOI: 10.1016/j.ppnp.2013.03.004. arXiv: 1303.5522 [hep-ph].
- [119] J. Erler et al. “Weak Polarized Electron Scattering”. In: *Ann. Rev. Nucl. Part. Sci.* 64 (2014), pp. 269–298. DOI: 10.1146/annurev-nucl-102313-025520. arXiv: 1401.6199 [hep-ph].
- [120] A. Czarnecki and W. J. Marciano. “Electroweak radiative corrections to polarized Moller scattering asymmetries”. In: *Phys. Rev. D* 53 (1996), pp. 1066–1072. DOI: 10.1103/PhysRevD.53.1066. arXiv: hep-ph/9507420.

-
- [121] A. Czarnecki and W. J. Marciano. “Polarized Moller scattering asymmetries”. In: *Int. J. Mod. Phys. A* 15 (2000). Ed. by C.A. Heusch, pp. 2365–2376. DOI: 10.1016/S0217-751X(00)00243-0. arXiv: hep-ph/0003049.
- [122] M. Davier et al. “Reevaluation of the hadronic vacuum polarisation contributions to the Standard Model predictions of the muon $g - 2$ and $\alpha(m_Z^2)$ using newest hadronic cross-section data”. In: *Eur. Phys. J. C* 77.12 (2017), p. 827. DOI: 10.1140/epjc/s10052-017-5161-6. arXiv: 1706.09436 [hep-ph].
- [123] F. Jegerlehner and R. Szafron. “ $\rho^0 - \gamma$ mixing in the neutral channel pion form factor F_π^e and its role in comparing e^+e^- with τ spectral functions”. In: *Eur. Phys. J. C* 71 (2011), p. 1632. DOI: 10.1140/epjc/s10052-011-1632-3. arXiv: 1101.2872 [hep-ph].
- [124] T. Blum et al. “Lattice calculation of the leading strange quark-connected contribution to the muon $g2$ ”. In: *JHEP* 04 (2016). [Erratum: *JHEP* 05, 034 (2017)], p. 063. DOI: 10.1007/JHEP04(2016)063. arXiv: 1602.01767 [hep-lat].
- [125] Bipasha Chakraborty et al. “Strange and charm quark contributions to the anomalous magnetic moment of the muon”. In: *Phys. Rev. D* 89.11 (2014), p. 114501. DOI: 10.1103/PhysRevD.89.114501. arXiv: 1403.1778 [hep-lat].
- [126] S. Okubo. “Phi meson and unitary symmetry model”. In: *Phys. Lett.* 5 (1963), pp. 165–168. DOI: 10.1016/S0375-9601(63)92548-9.
- [127] G. Zweig. “An SU(3) model for strong interaction symmetry and its breaking. Version 2”. In: *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*. Ed. by D.B. Lichtenberg and Simon Peter Rosen. Feb. 1964, pp. 22–101.
- [128] J. Iizuka. “Systematics and phenomenology of meson family”. In: *Prog. Theor. Phys. Suppl.* 37 (1966), pp. 21–34. DOI: 10.1143/PTPS.37.21.

-
- [129] T. Blum et al. “Calculation of the hadronic vacuum polarization disconnected contribution to the muon anomalous magnetic moment”. In: *Phys. Rev. Lett.* 116.23 (2016), p. 232002. DOI: 10.1103/PhysRevLett.116.232002. arXiv: 1512.09054 [hep-lat].
- [130] T. van Ritbergen, J.A.M. Vermaseren, and S.A. Larin. “The Four loop beta function in quantum chromodynamics”. In: *Phys. Lett. B* 400 (1997), pp. 379–384. DOI: 10.1016/S0370-2693(97)00370-5. arXiv: hep-ph/9701390.
- [131] M. Czakon. “The Four-loop QCD beta-function and anomalous dimensions”. In: *Nucl. Phys. B* 710 (2005), pp. 485–498. DOI: 10.1016/j.nuclphysb.2005.01.012. arXiv: hep-ph/0411261.
- [132] Lawrence J. Hall. “Grand Unification of Effective Gauge Theories”. In: *Nucl. Phys. B* 178 (1981), pp. 75–124. DOI: 10.1016/0550-3213(81)90498-3.
- [133] R. Alemany, M. Davier, and A. Hocker. “Improved determination of the hadronic contribution to the muon ($g-2$) and to alpha ($M(z)$) using new data from hadronic tau decays”. In: *Eur. Phys. J. C* 2 (1998), pp. 123–135. DOI: 10.1007/s100520050127. arXiv: hep-ph/9703220.
- [134] S. Eidelman et al. “Testing nonperturbative strong interaction effects via the Adler function”. In: *Phys. Lett. B* 454 (1999), pp. 369–380. DOI: 10.1016/S0370-2693(99)00389-5. arXiv: hep-ph/9812521.
- [135] C. Patrignani et al. “Review of Particle Physics”. In: *Chin. Phys. C* 40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.
- [136] J. Erler and M. Luo. “Hadronic loop corrections to the muon anomalous magnetic moment”. In: *Phys. Rev. Lett.* 87 (2001), p. 071804. DOI: 10.1103/PhysRevLett.87.071804. arXiv: hep-ph/0101010.

-
- [137] J. Erler, P. Masjuan, and H. Spiesberger. “Charm Quark Mass with Calibrated Uncertainty”. In: *Eur. Phys. J. C* 77.2 (2017), p. 99. DOI: 10.1140/epjc/s10052-017-4667-2. arXiv: 1610.08531 [hep-ph].
- [138] S. Weinberg. “Baryon and Lepton Nonconserving Processes”. In: *Phys. Rev. Lett.* 43 (1979), pp. 1566–1570. DOI: 10.1103/PhysRevLett.43.1566.
- [139] J. Schechter and J.W.F. Valle. “Neutrinoless Double beta Decay in SU(2) x U(1) Theories”. In: *Phys. Rev. D* 25 (1982), p. 2951. DOI: 10.1103/PhysRevD.25.2951.
- [140] J. F. Nieves. “Dirac and Pseudodirac Neutrinos and Neutrinoless Double Beta Decay”. In: *Phys. Lett. B* 147 (1984), pp. 375–379. DOI: 10.1016/0370-2693(84)90136-9.
- [141] E. Takasugi. “Can the Neutrinoless Double Beta Decay Take Place in the Case of Dirac Neutrinos?” In: *Phys. Lett. B* 149 (1984), pp. 372–376. DOI: 10.1016/0370-2693(84)90426-X.
- [142] Martin Hirsch, Sergey Kovalenko, and Ivan Schmidt. “Extended black box theorem for lepton number and flavor violating processes”. In: *Phys. Lett. B* 642 (2006), pp. 106–110. DOI: 10.1016/j.physletb.2006.09.012. arXiv: hep-ph/0608207.
- [143] S. F. King. “Models of Neutrino Mass, Mixing and CP Violation”. In: *J. Phys. G* 42 (2015), p. 123001. DOI: 10.1088/0954-3899/42/12/123001. arXiv: 1510.02091 [hep-ph].
- [144] R. N. Mohapatra and G. Senjanovic. “Neutrino Mass and Spontaneous Parity Nonconservation”. In: *Phys. Rev. Lett.* 44 (1980), p. 912. DOI: 10.1103/PhysRevLett.44.912.
- [145] J. Schechter and J.W.F. Valle. “Neutrino Masses in SU(2) x U(1) Theories”. In: *Phys. Rev. D* 22 (1980), p. 2227. DOI: 10.1103/PhysRevD.22.2227.
- [146] R. N. Mohapatra and G. Senjanovic. “Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation”. In: *Phys. Rev. D* 23 (1981), p. 165. DOI: 10.1103/PhysRevD.23.165.

-
- [147] E. Ma. “Pathways to naturally small neutrino masses”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1171–1174. DOI: 10.1103/PhysRevLett.81.1171. arXiv: hep-ph/9805219.
- [148] E. Ma and D.P. Roy. “Heavy triplet leptons and new gauge boson”. In: *Nucl. Phys. B* 644 (2002), pp. 290–302. DOI: 10.1016/S0550-3213(02)00815-5. arXiv: hep-ph/0206150.
- [149] B. Pontecorvo. “Mesonium and anti-mesonium”. In: *Sov. Phys. JETP* 6 (1957), p. 429.
- [150] Z. Maki, M. Nakagawa, and S. Sakata. “Remarks on the unified model of elementary particles”. In: *Prog. Theor. Phys.* 28 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870.
- [151] B. Kayser. “On the quantum mechanics of neutrino oscillation”. In: *Phys. Rev. D* 24 (1 July 1981), pp. 110–116. DOI: 10.1103/PhysRevD.24.110. URL: <https://link.aps.org/doi/10.1103/PhysRevD.24.110>.
- [152] W. Grimus. “Revisiting the quantum field theory of neutrino oscillations in vacuum”. In: *J. Phys. G* 47.8 (2020), p. 085004. DOI: 10.1088/1361-6471/ab716f. arXiv: 1910.13776 [hep-ph].
- [153] R.N. Mohapatra and P.B. Pal. *Massive neutrinos in physics and astrophysics*. Vol. 41. 1991.
- [154] M.C. Gonzalez-Garcia, Michele Maltoni, and Thomas Schwetz. “Global analyses of neutrino oscillation experiments”. In: *Nuclear Physics B* 908 (2016). Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015, pp. 199–217. ISSN: 0550-3213. DOI: <https://doi.org/10.1016/j.nuclphysb.2016.02.033>. URL: <http://www.sciencedirect.com/science/article/pii/S0550321316000778>.
- [155] Jr. Davis R., D. S. Harmer, and K. C. Hoffman. “Search for neutrinos from the sun”. In: *Phys. Rev. Lett.* 20 (1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205.

-
- [156] J. N. Bahcall. “Standard solar models”. In: *Nucl. Phys. B Proc. Suppl.* 77 (1999). Ed. by Y. Suzuki and Y. Totsuka, pp. 64–72. DOI: 10.1016/S0920-5632(99)00399-0. arXiv: astro-ph/9808162.
- [157] C. Pena-Garay and A. Serenelli. “Solar neutrinos and the solar composition problem”. In: (Nov. 2008). arXiv: 0811.2424 [astro-ph].
- [158] W. Hampel et al. “GALLEX solar neutrino observations: Results for GALLEX IV”. In: *Phys. Lett. B* 447 (1999), pp. 127–133. DOI: 10.1016/S0370-2693(98)01579-2.
- [159] M. Altmann et al. “Complete results for five years of GNO solar neutrino observations”. In: *Phys. Lett. B* 616 (2005), pp. 174–190. DOI: 10.1016/j.physletb.2005.04.068. arXiv: hep-ex/0504037.
- [160] J.N. Abdurashitov et al. “Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period”. In: *Phys. Rev. C* 80 (2009), p. 015807. DOI: 10.1103/PhysRevC.80.015807. arXiv: 0901.2200 [nucl-ex].
- [161] K.S. Hirata et al. “Observation of B-8 Solar Neutrinos in the Kamiokande-II Detector”. In: *Phys. Rev. Lett.* 63 (1989), p. 16. DOI: 10.1103/PhysRevLett.63.16.
- [162] B. Aharmim et al. “Electron energy spectra, fluxes, and day-night asymmetries of B-8 solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory”. In: *Phys. Rev. C* 72 (2005), p. 055502. DOI: 10.1103/PhysRevC.72.055502. arXiv: nucl-ex/0502021.
- [163] S. Abe et al. “Measurement of the 8B Solar Neutrino Flux with the KamLAND Liquid Scintillator Detector”. In: *Phys. Rev. C* 84 (2011), p. 035804. DOI: 10.1103/PhysRevC.84.035804. arXiv: 1106.0861 [hep-ex].
- [164] M. Apollonio et al. “Search for neutrino oscillations on a long baseline at the CHOOZ nuclear power station”. In: *Eur. Phys. J. C* 27 (2003), pp. 331–374. DOI: 10.1140/epjc/s2002-01127-9. arXiv: hep-ex/0301017.

-
- [165] F. Ardellier et al. “Double Chooz: A Search for the neutrino mixing angle θ_{13} ”. In: (June 2006). arXiv: hep-ex/0606025.
- [166] X. Guo et al. “A Precision measurement of the neutrino mixing angle θ_{13} using reactor antineutrinos at Daya-Bay”. In: (Jan. 2007). arXiv: hep-ex/0701029.
- [167] J.K. Ahn et al. “RENO: An Experiment for Neutrino Oscillation Parameter θ_{13} Using Reactor Neutrinos at Yonggwang”. In: (Mar. 2010). arXiv: 1003.1391 [hep-ex].
- [168] Leon M.G. De La Vega, R. Ferro-Hernandez, and E. Peinado. “Simple A_4 models for dark matter stability with texture zeros”. In: *Phys. Rev. D* 99.5 (2019), p. 055044. DOI: 10.1103/PhysRevD.99.055044. arXiv: 1811.10619 [hep-ph].
- [169] P.F. Harrison, D.H. Perkins, and W.G. Scott. “Tri-bimaximal mixing and the neutrino oscillation data”. In: *Phys. Lett. B* 530 (2002), p. 167. DOI: 10.1016/S0370-2693(02)01336-9. arXiv: hep-ph/0202074.
- [170] Vernon D. Barger et al. “Bimaximal mixing of three neutrinos”. In: *Phys. Lett. B* 437 (1998), pp. 107–116. DOI: 10.1016/S0370-2693(98)00880-6. arXiv: hep-ph/9806387.
- [171] A. Datta, F. Ling, and P. Ramond. “Correlated hierarchy, Dirac masses and large mixing angles”. In: *Nucl. Phys. B* 671 (2003), pp. 383–400. DOI: 10.1016/j.nuclphysb.2003.08.026. arXiv: hep-ph/0306002.
- [172] Y. Kajiyama, M. Raidal, and A. Strumia. “The Golden ratio prediction for the solar neutrino mixing”. In: *Phys. Rev. D* 76 (2007), p. 117301. DOI: 10.1103/PhysRevD.76.117301. arXiv: 0705.4559 [hep-ph].
- [173] P. H. Frampton, S. L. Glashow, and D. Marfatia. “Zeroes of the neutrino mass matrix”. In: *Phys. Lett. B* 536 (2002), pp. 79–82. DOI: 10.1016/S0370-2693(02)01817-8. arXiv: hep-ph/0201008.

-
- [174] L.J. Flores, Newton Nath, and Eduardo Peinado. “Non-standard neutrino interactions in U(1)’ model after COHERENT data”. In: *JHEP* 06 (2020), p. 045. DOI: 10.1007/JHEP06(2020)045. arXiv: 2002.12342 [hep-ph].
- [175] Walter Grimus et al. “Symmetry realization of texture zeros”. In: *Eur. Phys. J. C* 36 (2004), pp. 227–232. DOI: 10.1140/epjc/s2004-01896-y. arXiv: hep-ph/0405016.
- [176] H. Borgohain and Mrinal K. Das. “Phenomenology of two texture zero neutrino mass in left-right symmetric model with $Z_8 \times Z_2$ ”. In: *JHEP* 02 (2019), p. 129. DOI: 10.1007/JHEP02(2019)129. arXiv: 1810.04889 [hep-ph].
- [177] M. Hirsch et al. “Predictive flavour symmetries of the neutrino mass matrix”. In: *Phys. Rev. Lett.* 99 (2007), p. 151802. DOI: 10.1103/PhysRevLett.99.151802. arXiv: hep-ph/0703046.
- [178] J.M. Lamprea and E. Peinado. “Seesaw scale discrete dark matter and two-zero texture Majorana neutrino mass matrices”. In: *Phys. Rev. D* 94.5 (2016), p. 055007. DOI: 10.1103/PhysRevD.94.055007. arXiv: 1603.02190 [hep-ph].
- [179] E. Ma and G. Rajasekaran. “Softly broken A(4) symmetry for nearly degenerate neutrino masses”. In: *Phys. Rev. D* 64 (2001), p. 113012. DOI: 10.1103/PhysRevD.64.113012. arXiv: hep-ph/0106291.
- [180] K.S. Babu, E. Ma, and J.W.F. Valle. “Underlying A(4) symmetry for the neutrino mass matrix and the quark mixing matrix”. In: *Phys. Lett. B* 552 (2003), pp. 207–213. DOI: 10.1016/S0370-2693(02)03153-2. arXiv: hep-ph/0206292.
- [181] G. Altarelli and F. Feruglio. “Tri-bimaximal neutrino mixing from discrete symmetry in extra dimensions”. In: *Nucl. Phys. B* 720 (2005), pp. 64–88. DOI: 10.1016/j.nuclphysb.2005.05.005. arXiv: hep-ph/0504165.
- [182] M. Hirsch et al. “Discrete dark matter”. In: *Phys. Rev. D* 82 (2010), p. 116003. DOI: 10.1103/PhysRevD.82.116003. arXiv: 1007.0871 [hep-ph].

-
- [183] P.O. Ludl, S. Morisi, and E. Peinado. “The Reactor mixing angle and CP violation with two texture zeros in the light of T2K”. In: *Nucl. Phys. B* 857 (2012), pp. 411–423. DOI: 10.1016/j.nuclphysb.2011.12.017. arXiv: 1109.3393 [hep-ph].
- [184] D. Meloni, A. Meroni, and E. Peinado. “Two-zero Majorana textures in the light of the Planck results”. In: *Phys. Rev. D* 89.5 (2014), p. 053009. DOI: 10.1103/PhysRevD.89.053009. arXiv: 1401.3207 [hep-ph].
- [185] Z. Xing. “Texture zeros and Majorana phases of the neutrino mass matrix”. In: *Phys. Lett. B* 530 (2002), pp. 159–166. DOI: 10.1016/S0370-2693(02)01354-0. arXiv: hep-ph/0201151.
- [186] P. H. Frampton, Myoung C. Oh, and T. Yoshikawa. “Majorana mass zeroes from triplet VEV without majoron problem”. In: *Phys. Rev. D* 66 (2002), p. 033007. DOI: 10.1103/PhysRevD.66.033007. arXiv: hep-ph/0204273.
- [187] A. Kageyama et al. “Seesaw realization of the texture zeros in the neutrino mass matrix”. In: *Phys. Lett. B* 538 (2002), pp. 96–106. DOI: 10.1016/S0370-2693(02)01964-0. arXiv: hep-ph/0204291.
- [188] A. Merle and W. Rodejohann. “The Elements of the neutrino mass matrix: Allowed ranges and implications of texture zeros”. In: *Phys. Rev. D* 73 (2006), p. 073012. DOI: 10.1103/PhysRevD.73.073012. arXiv: hep-ph/0603111.
- [189] H. Fritzsch, Z. Xing, and S. Zhou. “Two-zero Textures of the Majorana Neutrino Mass Matrix and Current Experimental Tests”. In: *JHEP* 09 (2011), p. 083. DOI: 10.1007/JHEP09(2011)083. arXiv: 1108.4534 [hep-ph].
- [190] S. Zhou. “Update on two-zero textures of the Majorana neutrino mass matrix in light of recent T2K, Super-Kamiokande and NO ν A results”. In: *Chin. Phys. C* 40.3 (2016), p. 033102. DOI: 10.1088/1674-1137/40/3/033102. arXiv: 1509.05300 [hep-ph].

-
- [191] T. Kitabayashi and M. Yasuè. “Formulas for flavor neutrino masses and their application to texture two zeros”. In: *Phys. Rev. D* 93.5 (2016), p. 053012. DOI: 10.1103/PhysRevD.93.053012. arXiv: 1512.00913 [hep-ph].
- [192] S. Dev et al. “Near Maximal Atmospheric Neutrino Mixing in Neutrino Mass Models with Two Texture Zeros”. In: *Phys. Rev. D* 90.1 (2014), p. 013021. DOI: 10.1103/PhysRevD.90.013021. arXiv: 1405.0566 [hep-ph].
- [193] J. Alcaide, J. Salvado, and A. Santamaria. “Fitting flavour symmetries: the case of two-zero neutrino mass textures”. In: *JHEP* 07 (2018), p. 164. DOI: 10.1007/JHEP07(2018)164. arXiv: 1806.06785 [hep-ph].
- [194] J. Liao, D. Marfatia, and K. Whisnant. “Texture and Cofactor Zeros of the Neutrino Mass Matrix”. In: *JHEP* 09 (2014), p. 013. DOI: 10.1007/JHEP09(2014)013. arXiv: 1311.2639 [hep-ph].
- [195] P. O. Ludl and W. Grimus. “A complete survey of texture zeros in the lepton mass matrices”. In: *JHEP* 07 (2014). [Erratum: *JHEP* 10, 126 (2014)], p. 090. DOI: 10.1007/JHEP07(2014)090. arXiv: 1406.3546 [hep-ph].
- [196] S. Dev, L. Singh, and D. Raj. “Neutrino Mass Matrices with Two Vanishing Elements/Cofactors”. In: *Eur. Phys. J. C* 75.8 (2015), p. 394. DOI: 10.1140/epjc/s10052-015-3569-4. arXiv: 1506.04951 [hep-ph].
- [197] R. Sinha, R. Samanta, and A. Ghosal. “Maximal Zero Textures in Linear and Inverse Seesaw”. In: *Phys. Lett. B* 759 (2016), pp. 206–213. DOI: 10.1016/j.physletb.2016.05.080. arXiv: 1508.05227 [hep-ph].
- [198] Radha Raman Gautam and Sanjeev Kumar. “Zeros in the magic neutrino mass matrix”. In: *Phys. Rev. D* 94.3 (2016). [Erratum: *Phys.Rev.D* 100, 039902 (2019)], p. 036004. DOI: 10.1103/PhysRevD.94.036004. arXiv: 1607.08328 [hep-ph].

-
- [199] I. Esteban et al. “Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity”. In: *JHEP* 01 (2017), p. 087. DOI: 10.1007/JHEP01(2017)087. arXiv: 1611.01514 [hep-ph].
- [200] P.F. de Salas et al. “Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity”. In: *Phys. Lett. B* 782 (2018), pp. 633–640. DOI: 10.1016/j.physletb.2018.06.019. arXiv: 1708.01186 [hep-ph].
- [201] A. Gando et al. “Search for Majorana Neutrinos near the Inverted Mass Hierarchy Region with KamLAND-Zen”. In: *Phys. Rev. Lett.* 117.8 (2016). [Addendum: *Phys.Rev.Lett.* 117, 109903 (2016)], p. 082503. DOI: 10.1103/PhysRevLett.117.082503. arXiv: 1605.02889 [hep-ex].
- [202] C. Bonilla et al. “Neutrino phenomenology in a left-right D_4 symmetric model”. In: *Phys. Rev. D* 102.3 (2020), p. 036006. DOI: 10.1103/PhysRevD.102.036006. arXiv: 2003.06444 [hep-ph].
- [203] K. Hagiwara et al. “ $(g-2)_\mu$ and $\alpha(M_Z^2)$ re-evaluated using new precise data”. In: *J. Phys. G* 38 (2011), p. 085003. DOI: 10.1088/0954-3899/38/8/085003. arXiv: 1105.3149 [hep-ph].
- [204] D. Bernecker and H. B. Meyer. “Vector Correlators in Lattice QCD: Methods and applications”. In: *Eur. Phys. J. A* 47 (2011), p. 148. DOI: 10.1140/epja/i2011-11148-6. arXiv: 1107.4388 [hep-lat].